frontiers | Research Topics

Genomics of immunoregulation and inflammatory responses in the tumor microenvironment

Edited by
Lan Zhao, Yunpeng Cai and William C. Cho

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 978-2-83251-748-2
DOI 10.3389/978-2-83251-748-2

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

Genomics of immunoregulation and inflammatory responses in the tumor microenvironment

Topic editors

Lan Zhao — Stanford University, United States

Yunpeng Cai — Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences (CAS), China

William C. Cho — QEH, Kowloon, SAR China

Citation

Zhao, L., Cai, Y., Cho, W. C., eds. (2023). *Genomics of immunoregulation and inflammatory responses in the tumor microenvironment*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-748-2

Table of contents

06 Editorial: Genomics of immunoregulation and inflammatory responses in the tumor microenvironment
Lan Zhao, Yunpeng Cai and William C. Cho

10 The Expression Pattern of Ferroptosis-Related Genes in Colon Adenocarcinoma: Highly Correlated to Tumor Microenvironment Characteristics
Jie Liu, Hui Li, Shen Zhao, Rongbo Lin, Jiaqing Yu and Nanfeng Fan

22 Establishment and Application of a Prognostic Risk Score Model Based on Characteristics of Different Immunophenotypes for Lung Adenocarcinoma
Hong Gao, Yanhong Liu, Yue Hu, Meiling Ge, Jie Ding and Qing Ye

35 Apoptosis-Related Signature Predicts Prognosis and Immune Microenvironment Infiltration in Lung Adenocarcinoma
Xiaoli Zou, Rong He, Zhenzhen Zhang and Yulan Yan

50 Prognostic Characteristics and Immune Effects of N^6-Methyladenosine and 5-Methylcytosine-Related Regulatory Factors in Clear Cell Renal Cell Carcinoma
Lei Li, Zijia Tao, Yiqiao Zhao, Mingyang Li, Jianyi Zheng, Zeyu Li and Xiaonan Chen

66 TGFβ2 is a Prognostic Biomarker for Gastric Cancer and is Associated With Methylation and Immunotherapy Responses
Bangling Han, Tianyi Fang, Yimin Wang, Yongle Zhang and Yingwei Xue

82 Amino Acid Metabolism-Related lncRNA Signature Predicts the Prognosis of Breast Cancer
Yin-wei Dai, Zhi-kai Wen, Zhi-xuan Wu, Hao-dong Wu, Lin-xi Lv, Cong-zhi Yan, Cong-hui Liu, Zi-qiong Wang and Chen Zheng

99 Comprehensive Analysis of the Immune Cell Infiltration Landscape and Immune-Related Methylation in Retinoblastoma
Peiyao Mao, Yinchen Shen, Xun Xu and Jiawei Zhong

109 A Systematic Analysis Reveals the Prognostic and Immunological Role of Reptin/RUVBL2 in Human Tumors
Xiaoru Su, Gaoming Zheng, Zhifang Gui, Xiao Yang, Lahong Zhang and Feng Pan

122 Unfolded Protein Response–Related Signature Associates With the Immune Microenvironment and Prognostic Prediction in Osteosarcoma
Zhao Zhang, Xincheng Liu, Debin Cheng, Jingyi Dang, Zhenzhou Mi, Yubo Shi, Lei Wang and Hongbin Fan

135 The Crosstalk Between Immune Infiltration, Circulating Tumor Cells, and Metastasis in Pancreatic Cancer: Identification of HMGB3 From a Multiple Omics Analysis
Hao-dong Tang, Yang Wang, Peng Xie, Si-yuan Tan, Hai-feng Li, Hao Shen, Zheng Zhang, Zheng-qing Lei and Jia-hua Zhou

148 The Immune Cell Infiltration Patterns and Characterization Score in Bladder Cancer to Identify Prognosis
Yongsheng Zhang, Yunlong Wang, Jichuang Wang and Kaixiang Zhang

160 Analysis of Pyroptosis-Related Immune Signatures and Identification of Pyroptosis-Related LncRNA Prognostic Signature in Clear Cell Renal Cell Carcinoma
Ming Zhong, Xiaohua Wang, Enyi Zhu, Lian Gong, Lingyan Fei, Liang Zhao, Keping Wu, Chun Tang, Lizhen Zhang, Zhongli Wang and Zhihua Zheng

178 Identification of a Prognostic Model Based on Fatty Acid Metabolism-Related Genes of Head and Neck Squamous Cell Carcinoma
Peiyu Du, Yue Chai, Shimin Zong, Jianxin Yue and Hongjun Xiao

193 CCR5 as a Prognostic Factor in Lower-Grade Glioma is Involved in the Remodeling of the Tumor Microenvironment
Fang Wang, Zhennan Tao, Zhen Tian, Jiaqi Jin, Jiawei Dong, Yuxiang Dai, Wanli Yu, Bin Tang and Shaoshan Hu

206 The molecular mechanisms of CTHRC1 in gastric cancer by integrating TCGA, GEO and GSA datasets
Lulu Zhao, Wanqing Wang, Penghui Niu, Xiaoyi Luan, Dongbing Zhao and Yingtai Chen

216 m7G-Associated subtypes, tumor microenvironment, and validation of prognostic signature in lung adenocarcinoma
Guangyao Wang, Mei Zhao, Jiao Li, Guosheng Li, Fukui Zheng, Guanglan Xu and Xiaohua Hong

233 Integrative analysis of a necroptosis-related gene signature of clinical value and heterogeneity in diffuse large B cell lymphoma
Yu-Biao Pan, Wei Wang, Hong-Kai Cai, Jia Zhang, Ya Teng, Jiji Xue, Min Zhu and Wen-Da Luo

250 Identification of novel tumor microenvironment-associated genes in gastric cancer based on single-cell RNA-sequencing datasets
Xujin Wei, Jie Liu, Zhijun Hong, Xin Chen, Kang Wang and Jianchun Cai

267 Integrated pan-cancer analysis of CSMD2 as a potential prognostic, diagnostic, and immune biomarker
Huiyun Zhang, Taobi Huang, Xiangqing Ren, Xidong Fang, Xia Chen, Hui Wei, Weiming Sun and Yuping Wang

284 A novel prognostic and therapeutic target biomarker based on necroptosis-related gene signature and immune microenvironment infiltration in gastric cancer
Dao Xin, Yuxin Man, Yalan Yang and Feng Wang

305 Establishment of a prognostic ferroptosis- and immune-related long noncoding RNAs profile in kidney renal clear cell carcinoma
Zhijun Han, Hao Wang, Yafei Liu and Xiao-Liang Xing

317 INHBB is a novel prognostic biomarker and correlated with immune infiltrates in gastric cancer
Weifeng Yu, Guihua He, Wang Zhang, Zhenhao Ye, Zishao Zhong and Suiping Huang

334 Pan-cancer analysis of CREB3L1 as biomarker in the prediction of prognosis and immunotherapeutic efficacy
Zhengjun Lin, Yanlin Wu, XunGang Xiao, Xianghong Zhang, Jia Wan, Tao Zheng, Hongxuan Chen, Tang Liu and Xianzhe Tang

353 PANoptosis-based molecular clustering and prognostic signature predicts patient survival and immune landscape in colon cancer
Xu Wang, Rui Sun, Shixin Chan, Lei Meng, Yuanmin Xu, Xiaomin Zuo, Zhenglin Wang, Xianyu Hu, Qijun Han, Longfei Dai, Tao Bai, Zhen Yu, Ming Wang, Wenqi Yang, Huabing Zhang and Wei Chen

368 Identification and validation of ferroptosis-related lncRNA signatures as a novel prognostic model for glioma
Liang Huang, Juan Zhang, Fanghua Gong, Yuhua Han, Xing Huang, Wanxiang Luo, Huaan Cai and Fan Zhang

387 TLR3 serves as a novel diagnostic and prognostic biomarker and is closely correlated with immune microenvironment in three types of cancer
Xiong Zou, Yi Guo and Zengnan Mo

404 Identification of a ferroptosis-related long noncoding RNA signature with a prognostic value in adrenocortical carcinoma
Weixi Wang, Guilin Chang, Ran Zhuo and Cong Ye

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 10 February 2023
DOI 10.3389/fgene.2023.1136342

Check for updates

OPEN ACCESS

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

*CORRESPONDENCE
Lan Zhao,
lanzhao20140101@gmail.com

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

RECEIVED 02 January 2023
ACCEPTED 03 February 2023
PUBLISHED 10 February 2023

CITATION
Zhao L, Cai Y and Cho WC (2023), Editorial: Genomics of immunoregulation and inflammatory responses in the tumor microenvironment.
Front. Genet. 14:1136342.
doi: 10.3389/fgene.2023.1136342

COPYRIGHT
© 2023 Zhao, Cai and Cho. 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: Genomics of immunoregulation and inflammatory responses in the tumor microenvironment

Lan Zhao[1]*, Yunpeng Cai[2] and William C. Cho[3]

[1]Department of Medicine, Stanford University School of Medicine, Stanford, CA, United States, [2]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, [3]Department of Clinical Oncology, Queen Elizabeth Hospital, Hong Kong, China

KEYWORDS

genomics, tumor microenvironment, integrative analysis, tumor-infiltrating cells, microbes, inflammation

Editorial on the Research Topic
Genomics of immunoregulation and inflammatory responses in the tumor microenvironment

The tumor microenvironment (TME) plays key roles in cancer, and targeting the TME has received significant attention in recent years. TME is characterized by infiltration of immune cells with both tumor-suppressing and tumor-promoting properties. We have edited the Research Topic (RT), entitled Genomics of immunoregulation and inflammatory responses in the tumor microenvironment with the aim to enhance our understanding of the mechanisms involving the TME as potential therapeutic targets in cancer. A total of 78 original research manuscripts and reviews were submitted between 02 November 2021 and 24 June 2022, among which 27 manuscripts were accepted for publication after peer review. Multi-omics approaches such as transcriptomics, (epi)genomics, proteomics studies have been applied for biomarker discovery and characterize the TME. The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), and other public databases were used to conduct meta-analysis to identify promising mRNAs, non-coding RNAs (ncRNAs), N6-methyladenosine (m6A) RNAs, and DNA methylation-based prognostic biomarkers in patients with different types of cancer. TME components such as tumor, stromal, infiltrating immune, and vascular cells, and their positive and negative roles in cancer progression and inflammation are defined.

Gastric cancer (GC) is among the most types of cancer investigated. Both bulk and single-cell RNA-seq (scRNA-seq) methods were used to determine cancer-associated mRNA biomarkers. For example, TGFβ2 is a highly expressed gene identified from a GC bulk RNA-seq data. ScRNA-seq analysis by Wei et al. identified four high-risk genes, namely *TMPRSS15*, *VIM*, *APOA1*, and *RNASE1*, associated with poor clinical outcomes in GC. Previous study in mice has shown that *Cthrc1* expression is limited to injured vascular tissue and its involvement in arterial remodeling has been defined (LeClair et al., 2007). Combining bulk and scRNA-seq data, Zhao and colleagues found that *CTHRC1* overexpression was significantly correlated with angiogenesis, macrophage infiltration and poor prognosis in GC. Validation by immunofluorescence staining showed that *CTHRC1* was present in the vascular tissue which is considered to be involved in angiogenesis. In addition,

immunohistochemistry and tissue microarray by Yu et al. indicated *INHBB* is overexpressed and correlated with poor prognosis in GC.

Besides GC, at least 10 other cancer types, and several pan-cancer studies were conducted to find new biomarkers related to prognosis, immune infiltration, and other signatures of interest. Zhang and colleagues performed an integrated bioinformatics analysis of a membrane protein, CSMD2, in pan-cancer using the TCGA, GEO, and International Cancer Genome Consortium (ICGC) databases. They found that elevated expression of *CSMD2* indicates poor prognosis and high immune infiltration levels in cancer-associated fibroblasts in multiple cancers. Another three pan-cancer studies reveal promising prognostic value and immune characteristics of CREB3L1, RUVBL2, and TLR3 respectively. Chemokines and their receptors regulate immune cell trafficking into the TME (Nagarsheth et al., 2017). CCR5 is a chemokine receptor abundantly expressed in various tumors, and has been suggested as the therapeutic targets in breast and colon cancer (Jiao et al., 2019). Through bioinformatics and experimental validations, Wang and colleagues showed that *CCR5* is upregulated at both mRNA and protein levels, and associated with poor prognosis in patients with lower-grade glioma (LGG). Five m6A and 5-methylcytosine (m5C) prognostic regulators were identified in clear cell renal cell carcinoma (ccRCC). Another integrated analysis of bulk and scRNA-seq identified HMGB3 as a hub gene associated with epithelial-mesenchymal transition (EMT) in circulating tumor cells from pancreatic ductal adenocarcinoma (PDAC).

Cancer displays heterogeneity in phenotypic properties, clinical presentations and underlying genetic causes. Molecular subtyping of cancer using data-driven and model-based methodologies offer new hope for personalized medicine (Zhao et al., 2018b; Zhao and Yan, 2019). We previously conducted an innovative molecular subtyping study in nasopharyngeal carcinoma (NPC) (Zhao et al., 2018a). We built a classifier based on a 10-microRNA (miRNA) signature to stratify NPC patients and NPC cell lines into immunogenic, classical and mesenchymal subtypes. Of note, gene set enrichment analysis (GSEA) and survival analysis indicated that the mesenchymal subtype has enriched with EMT signaling pathways and the worst clinical outcomes. A panel of 4 miRNAs were subsequently identified and employed to establish a prognostic model, by differential and uni-/multi-variate Cox regression analysis between mesenchymal and non-mesenchymal subtypes. Last, the classifier and the Cox prognostic model were assessed in both the training and validation datasets. Furthermore, heterogeneity has been observed not only in the tumor but also its microenvironment (Zhao et al., 2018c). In our current RT, there were three studies applying unsupervised hierarchical clustering of tumor-infiltrating immune cells (TIICs) to identify 2-3 immunophenotypes in lung adenocarcinoma (LUAD), bladder cancer (BLCA), and retinoblastoma, respectively.

A number of articles in the RT employed similar data processing techniques as described in our NPC study (Zhao et al., 2018a), although without proper citations. For example, Zou and colleagues selected 136 apoptosis-related genes in the KEGG apoptosis pathway (map04210), and identified 64-gene that were differentially expressed between tumor and non-tumor tissues in LUAD from the TCGA database. Unsupervised consensus clustering was applied to classify the patients into two prognostic subtypes based on these 64-gene signatures. Differential and Cox regression analysis between the two subtypes helped to identify and build a 11 apoptosis-gene signature model for survival risk prediction. The model was further validated in the patient and cancer cell line datasets. Another study in LUAD identified three 7-methylguanosine (m7G)-associated prognostic signatures.

Similar approaches were conducted in gastric and colorectal cancers. The authors focused on necroptosis- and panoptosis-related genes, respectively, and their contributions to the TME in their respective cancers. The two studies also performed comprehensive bioinformatics and quantitative real-time polymerase chain reaction (qRT-PCR) experimental validations, which may represent improvements in this field but still lack some novelties. Another set of unique necroptosis-related gene signatures were identified in diffuse large B cell lymphoma although without significant advances in terms of data processing and analysis.

Ferroptosis- and pyroptosis-related prognostic gene signatures were identified from another four (Han et al., Huang et al., Wang et al., and Liu et al.) and Zhong et al.'s studies, respectively, using similar or identical data processing steps. The gene signatures in these four studies (Han et al., Huang et al., Wang et al., and Zhong et al.) were long non-coding RNAs (lncRNAs), and no overlaps were found between any of the studies. RNAs longer than 200 nucleotides in length that do not encode proteins are broadly categorized as lncRNAs. In contrast to mRNAs, lncRNAs are less abundant, less evolutionarily conserved, more tissue-specific, and are localized predominantly in the nucleus (Statello et al., 2021). The discovery of lncRNAs and their diverse functions in various cellular processes, including oncogenic signaling, have provided new perspectives in cancer prognosis (Evans et al., 2016; Bridges et al., 2021). In the current RT, the above four studies noted the important immune-related roles of lncRNAs in the TME, and there is another study in breast cancer which emphasized the contribution of lncRNAs in amino acids metabolism. Altered metabolism is a hallmark of cancer. The role of amino acids and lipids in cancer metabolism is very much appreciated recently (Lieu et al., 2020; Snaebjornsson et al., 2020). A total of two studies in this RT identified metabolism-related gene signatures for predicting the prognosis of patients with breast and head and neck cancer, respectively. Although interesting, the two pure bioinformatics studies lacked experimental validations.

Considering there are at least 9 papers in the RT mentioning apoptosis, necrosis, ferroptosis, and other forms of cell death, it is necessary to add a paragraph or two to talk about the subject in terms of comparing their characteristics and roles in cancer. Apoptosis is a highly regulated and conserved form of normal programmed cell death (PCD) in multicellular organisms (Kerr et al., 1972). The apoptotic pathway is mainly activated by the intrinsic mitochondrial and extrinsic death receptor signals (Wong, 2011). In addition, under endoplasmic reticulum (ER) stress, the unfolded protein response (UPR) pathway is central for apoptotic cell death (Hetz et al., 2020). In this RT, there is a study carried out by Zhang and colleagues to identify UPR-related prognostic signatures in Osteosarcoma. Necrosis is an inflammatory-inducing cell death, which has been characterized as an accidental and passive process that results from external

causes of injury (Syntichaki and Tavernarakis, 2002). The dysregulation of apoptosis and the harsh necrotic microenvironment contribute to malignant transformation and progression. For example, it is well-known that apoptosis evasion is a hallmark of cancer. Cancer cells evade apoptosis either by overexpressing anti-apoptotic genes such as *BCL2* or silencing the expression of pro-apoptotic transcription factor *TP53* (Fernald and Kurokawa, 2013). Hypoxia and metabolic stress induce necrosis and provoke immune responses, which eventually lead to cancer formation and TME modification (Karsch-Bluman et al., 2019). Moreover, a necroptosis-associated cytokine microenvironment regulates liver cancer progression (Seehawer et al., 2018). Ferroptosis is an non-apoptotic, iron-dependent, and oxidative form of cell death (Dixon et al., 2012). Iron has been found to play a role in the TME and in metastasis (Torti and Torti, 2013). Thus, iron and ferroptosis are promising therapeutic targets in cancer. In our current RT, Huang and colleagues identified and validated a 3-ferroptosis-related lncRNA prognostic signature in glioma. They performed functional and biochemical analysis of one of the signatures, LINC01426, which can be useful for future references.

A number of pathogens such as enteric *Salmonella* and *Shigella* species can induce a novel form of PCD, namely pyroptosis, in infected host cells (Boise and Collins, 2001; Fink and Cookson, 2005; Jorgensen and Miao, 2015). Although pyroptosis is inflammatory, it differs from necrosis which is caspase dependent and genetically controlled (Boise and Collins, 2001; Fink and Cookson, 2005). Detailed and comprehensive discussion of apoptosis, necrosis, ferroptosis, pyroptosis and their roles in cancer can be found in reviews (Boise and Collins, 2001; Fink and Cookson, 2005; Yu et al., 2017). The four forms of cell death may act as defense systems against microbial infection (Fink and Cookson, 2005; Amaral et al., 2019). Crosstalk between microbes and hosts, in terms of microbe-induced host cell death and cancer prevention and survival outcomes, can be further expanded in future studies.

Taken together, the TME is characterized by infiltrated immune cells that can influence cancer progression and patient prognosis. The connection between chronic inflammation in TME and cancer is bidirectional and complex. Large amounts of genomics data deposited in public repositories advance the fields of oncology and immunology. Uncovering novel risk-associated biomarkers in cancer through systematic review and meta-analysis are necessary and proven to be useful.

Author contributions

LZ wrote the draft. LZ, WC, and YC critically reviewed and revised 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

Amaral, E. P., Costa, D. L., Namasivayam, S., Riteau, N., Kamenyeva, O., Mittereder, L., et al. (2019). A major role for ferroptosis in Mycobacterium tuberculosis–induced cell death and tissue necrosis. *J. Exp. Med.* 216, 556–570. doi:10.1084/jem.20181776

Boise, L. H., and Collins, C. M. (2001). Salmonella-induced cell death: Apoptosis, necrosis or programmed cell death? *Trends Microbiol.* 9, 64–67. doi:10.1016/s0966-842x(00)01937-5

Bridges, M. C., Daulagala, A. C., and Kourtidis, A. (2021). LNCcation: lncRNA localization and function. *J. Cell Biol.* 220, e202009045. doi:10.1083/jcb.202009045

Dixon, S. J., Lemberg, K. M., Lamprecht, M. R., Skouta, R., Zaitsev, E. M., Gleason, C. E., et al. (2012). Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* 149, 1060–1072. doi:10.1016/j.cell.2012.03.042

Evans, J. R., Feng, F. Y., and Chinnaiyan, A. M. (2016). The bright side of dark matter: lncRNAs in cancer. *J. Clin. Invest.* 126, 2775–2782. doi:10.1172/JCI84421

Fernald, K., and Kurokawa, M. (2013). Evading apoptosis in cancer. *Trends Cell Biol.* 23, 620–633. doi:10.1016/j.tcb.2013.07.006

Fink, S. L., and Cookson, B. T. (2005). Apoptosis, pyroptosis, and necrosis: Mechanistic description of dead and dying eukaryotic cells. *Infect. Immun.* 73, 1907–1916. doi:10.1128/IAI.73.4.1907-1916.2005

Hetz, C., Zhang, K., and Kaufman, R. J. (2020). Mechanisms, regulation and functions of the unfolded protein response. *Nat. Rev. Mol. Cell Biol.* 21, 421–438. doi:10.1038/s41580-020-0250-z

Jiao, X., Nawab, O., Patel, T., Kossenkov, A. V., Halama, N., Jaeger, D., et al. (2019). Recent advances targeting CCR5 for cancer and its role in immuno-oncology. *Cancer Res.* 79, 4801–4807. doi:10.1158/0008-5472.CAN-19-1167

Jorgensen, I., and Miao, E. A. (2015). Pyroptotic cell death defends against intracellular pathogens. *Immunol. Rev.* 265, 130–142. doi:10.1111/imr.12287

Karsch-Bluman, A., Feiglin, A., Arbib, E., Stern, T., Shoval, H., Schwob, O., et al. (2019). Tissue necrosis and its role in cancer progression. *Oncogene* 38, 1920–1935. doi:10.1038/s41388-018-0555-y

Kerr, J. F., Wyllie, A. H., and Currie, A. R. (1972). Apoptosis: A basic biological phenomenon with wide-ranging implications in tissue kinetics. *Br. J. Cancer* 26, 239–257. doi:10.1038/bjc.1972.33

LeClair, R. J., Durmus, T., Wang, Q., Pyagay, P., Terzic, A., and Lindner, V. (2007). Cthrc1 is a novel inhibitor of transforming growth factor-β signaling and neointimal lesion formation. *Circ. Res.* 100, 826–833. doi:10.1161/01.RES.0000260806.99307.72

Lieu, E. L., Nguyen, T., Rhyne, S., and Kim, J. (2020). Amino acids in cancer. *Exp. Mol. Med.* 52, 15–30. doi:10.1038/s12276-020-0375-3

Nagarsheth, N., Wicha, M. S., and Zou, W. (2017). Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. *Nat. Rev. Immunol.* 17, 559–572. doi:10.1038/nri.2017.49

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

Snaebjornsson, M. T., Janaki-Raman, S., and Schulze, A. (2020). Greasing the wheels of the cancer machine: The role of lipid metabolism in cancer. *Cell Metab.* 31, 62–76. doi:10.1016/j.cmet.2019.11.010

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, 96–118. doi:10.1038/s41580-020-00315-9

Syntichaki, P., and Tavernarakis, N. (2002). Death by necrosis. *EMBO Rep.* 3, 604–609. doi:10.1093/embo-reports/kvf138

Torti, S. V., and Torti, F. M. (2013). Iron and cancer: More ore to be mined. *Nat. Rev. Cancer* 13, 342–355. doi:10.1038/nrc3495

Wong, R. S. Y. (2011). Apoptosis in cancer: From pathogenesis to treatment. *J. Exp. Clin. Cancer Res.* 30, 87. doi:10.1186/1756-9966-30-87

Yu, H., Guo, P., Xie, X., Wang, Y., and Chen, G. (2017). Ferroptosis, a new form of cell death, and its relationships with tumourous diseases. *J. Cell. Mol. Med.* 21, 648–657. doi:10.1111/jcmm.13008

Zhao, L., Fong, A. H. W., Liu, N., and Cho, W. C. S. (2018a). Molecular subtyping of nasopharyngeal carcinoma (NPC) and a microRNA-based prognostic model for distant metastasis. *J. Biomed. Sci.* 25, 16. doi:10.1186/s12929-018-0417-5

Zhao, L., Lee, V. H. F., Ng, M. K., Yan, H., and Bijlsma, M. F. (2018b). Molecular subtyping of cancer: Current status and moving toward clinical applications. *Brief. Bioinform* 20, 572–584. doi:10.1093/bib/bby026

Zhao, L., and Yan, H. (2019). Mcnf: A novel method for cancer subtyping by integrating multi-omics and clinical data. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 17, 1682–1690. doi:10.1109/TCBB.2019.2910515

Zhao, L., Zhao, H., and Yan, H. (2018c). Gene expression profiling of 1200 pancreatic ductal adenocarcinoma reveals novel subtypes. *BMC Cancer* 18, 603. doi:10.1186/s12885-018-4546-8

frontiers in Genetics

ORIGINAL RESEARCH
published: 08 March 2022
doi: 10.3389/fgene.2022.837941

The Expression Pattern of Ferroptosis-Related Genes in Colon Adenocarcinoma: Highly Correlated to Tumor Microenvironment Characteristics

*Jie Liu†, Hui Li, Shen Zhao, Rongbo Lin, Jiaqing Yu and Nanfeng Fan**

Department of Gastrointestinal Medical Oncology, Fujian Cancer Hospital, Fujian Medical University Cancer Hospital, Fuzhou, China

OPEN ACCESS

Edited by:
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology, (CAS), China

Reviewed by:
Laura Mondragón Martínez,
Josep Carreras Leukaemia Research Institute (IJC), Spain
Emil Bulatov,
Kazan Federal University, Russia

***Correspondence:**
Nanfeng Fan
fnf_fjzlyy@hotmail.com

†These authors share first authorship

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

Received: *17 December 2021*
Accepted: *21 February 2022*
Published: *08 March 2022*

Citation:
Liu J, Li H, Zhao S, Lin R, Yu J and Fan N (2022) The Expression Pattern of Ferroptosis-Related Genes in Colon Adenocarcinoma: Highly Correlated to Tumor Microenvironment Characteristics. Front. Genet. 13:837941. doi: 10.3389/fgene.2022.837941

In the latest literatures, ferroptosis caused by T cells in cancerous cells provided new insights of improving curative effect of the PD-1/PD-L1 antibody. The microenvironment on which tumor cells develop and survive was also emphasized as its crucial role in tumor occurrence, development, metastasis and immune escape. Thus, the interaction of ferroptosis related genes and tumor microenvironment (TME) was urgently be detected in a comprehensive perspective. We comprehensively evaluated the transcriptional feature of ferroptosis related genes in colon adenocarcinoma (COAD), and systematically associated these ferroptosis subtypes with DNA damage repair (DDR) and TME characteristics. We found two unique patterns of ferroptosis characterized by distinct biological pathways activation. We also demonstrated that FRG score constructed based on ferroptosis subtypes has a significant correlation with prognosis of colon cancer and could act as an independent prognostic biomarker for predicting patients' survival. The higher immune infiltrating level, immune functional pathways activation was observed in the high FRG score group. Furthermore, these results were verified by an independent external GEO cohort. This work revealed ferroptosis was highly associated with TME complexity and diversity. A novel ferroptosis subtypes related gene scoring system can be used for prognostic prediction in COAD. Targeting ferroptosis may be a therapeutic alternative for COAD.

Keywords: prognosis, immune status, tumor microenvironment, ferroptosis, colon cancer

INTRODUCTION

Ferroptosis is a programmed cell death characterized by an iron-dependent oxidative alteration of phospholipid membranes (Conrad et al., 2016). A preliminary analysis of this route established that depletion of cysteine, which results in the exhaustion of the intracellular glutathione (reduced) (GSH) pool, initiates this type of cell death (Wang, et al., 2019) explicitly. The GSH requirement for ferroptosis protection is connected to the optimum functioning of the enzyme glutathione peroxidase 4 (GPX4). A selenoprotein is necessary to effectively reduce peroxidized

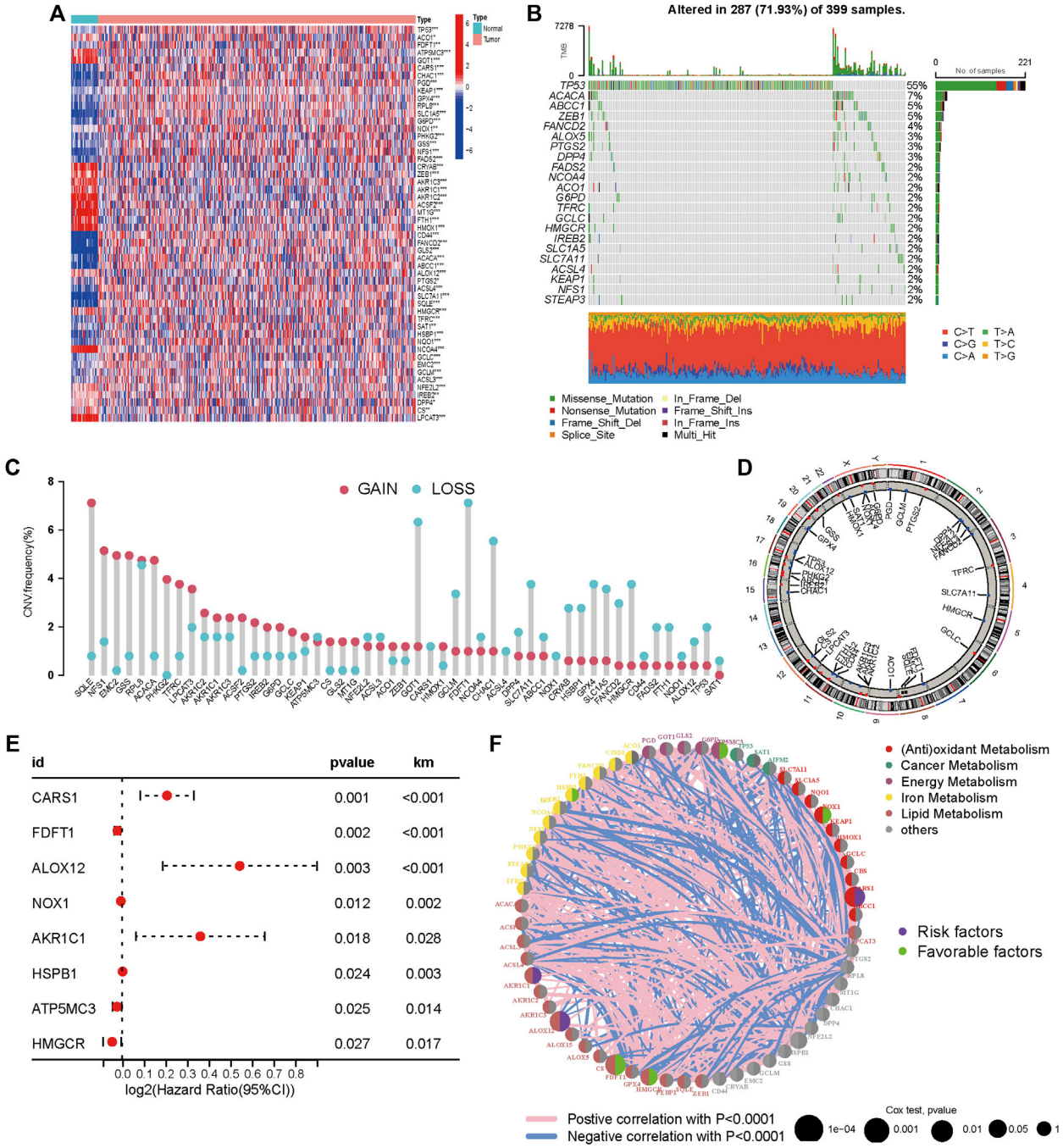


FIGURE 1 | Expression and genetic variation landscapes of m6A regulators with COAD. **(A)** The expression level of ferroptosis related genes in tumor and pericarcinomatous tissue in TCGA-COAD cohort. Blue and red represent low and high expression, respectively. The asterisks are indicative of statistical *p*-value (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). **(B)** Mutation frequencies of 22 FRGs (mutation frequency >2%) in 399 patients with COAD from the TCGA cohort. **(C)** Frequencies of CNV gain, loss, and non-CNV among FRGs. **(D)** Locations of CNV alterations in FRGs on 23 chromosomes. **(E)** Univariate Cox regression analysis conducted to estimate the significance of FRGs in prognosis prediction in TCGA-COAD cohort. The horizontal length stands for 95% CI of each gene. **(F)** Association among FRGs within CM. Circle size stands for the impact of every regulator on prognosis prediction. $p < 0.0001$, $p < 0.001$, $p < 0.01$, $p < 0.05$ and $p < 1$, respectively, were obtained from log-rank test. Green and purple dots within the circle indicate protective and risk factors for prognosis, respectively. The lines linking regulators indicate the mutual interactions, with the thickness showing the strength of association between them. Pink and blue colors indicate positive and negative correlation, respectively.

phospholipids and suppress arachidonic acid (AA)-metabolizing enzyme activity. This may contribute to the phospholipid peroxidation process (Friedmann Angeli et al., 2014); (Yang et al., 2014); (Ingold et al., 2018); (Kagan et al., 2017). Since then, it has become clear that a complex interaction between lipid, iron, and cysteine metabolism is critical for this cell death process. Ferroptosis has been implicated in the development and progression of tumors by activating various regulatory sites in signaling pathways, promoting tumor cell death (Friedmann Angeli et al., 2019) (Tang R. et al., 2020). Thus, elucidating the ferroptosis process and the associated mechanisms that regulate tumor formation may yield novel therapeutic strategies for malignancies (Wu et al., 2020); (Hassannia et al., 2019).

Immunotherapy, represented by the immune checkpoint blockade (ICB, PD-1/L1, and CTLA-4), has shown significant clinical success in a limited proportion of patients with long-lasting responses. Unfortunately, most patients show slight or no clinical improvement, far from meeting a clinical requirement (Topalian et al., 2012). Historically, tumor growth was thought to be a multistep process involving primarily epigenetic and genetic changes inside tumor cells. However, many reported studies have demonstrated that the microenvironment on which tumor cells develop and survive also plays a critical role in the progression of the tumor. In cancer, there is a complex microenvironment made up of macrophages and resident fibroblasts (CAF; cancer-associated fibroblasts) as well as immune cells infiltrating the tumor (myeloid cells and lymphocytes) cells derived from the bone marrow (BMDCs), and secreted factors like cytokines. TAMs, TANs, dendritic cells, myeloid-derived suppressor cells (MDSCs), and Tie2-expressing monocytes were all part of the tumor-associated myeloid cell (TAM) population (TAMCs) (Pitt et al., 2016). With other TME components interactions (directly or indirectly), cancer cells induced several biological behavior changes. Including proliferation and angiogenesis, inhibition of apoptosis, avoidance of hypoxia, and induction of immunological tolerance. As our understanding of the tumor microenvironment's diversity and complexity increased. Accumulating data demonstrates its crucial involvement in tumor development, immune escape, and immunotherapy response. To improve the efficacy of existing ICBs and to develop novel immunotherapeutic tactics, predicting ICB response based on TME cell infiltration characterization is a critical process (Quail and Joyce, 2013), (Ali et al., 2016). Therefore, the different tumor immune phenotypes could be uncovered through systematically analyzing the TME landscape's variability and complexity. Additionally, the capability to guide and anticipate immunotherapeutic response would improve.

Several studies have recently demonstrated a unique link between TME-infiltrating immune cells, particularly $CD8^+$ T cells, and ferroptosis. Wang et al. found that the ferroptosis-specific lipid peroxidation in tumor cells was induced by immunotherapy-activated $CD8^+$ T cells. Ferroptosis contributes to immunotherapy's anticancer efficiency via interferon-gamma (IFN) released by $CD8^+$ T cells (Wang, et al., 2019). The OTUD1 increases transferrin receptor protein-1 (TFRC)-mediated iron transport by deubiquitinating and stabilizing IREB2. Resulting in enhanced ROS production, ferroptosis and potentiates host antitumor immunity (Song et al., 2021). Previous colon adenocarcinoma (COAD) research has, however, been limited to one to several ferroptosis regulators. At the same time, the promising anticancer effect is characterized by many tumor suppressor factors associated in a well-coordinated manner. Widespread recognition of TME cell infiltration characterizations mediated by numerous ferroptosis-related genes will help us better understand Cancer-immune cycle controlled by TME. We integrated the genetic information from multiple colon cancer samples in this study and then identified the ferroptosis subtypes with thoroughly different TME cell-infiltrating features. We also discovered differences in the landscape of DNA damage repair (DDR) and somatic mutations between them. To assess ferroptosis in individual patients, we created a scoring system based on ferroptosis subtype-related genes. Following that, we established a consistent relationship between ferroptosis and immune infiltrating cells.

MATERIAL AND METHODS

Acquisition and Preprocessing of Data

We used the Gene-Expression Omnibus (GEO) and Cancer Genome Atlas (TCGA) databanks to obtain clinical annotation and public gene expression data. Colonic cancer cohorts (GSE17536, GSE39582, and TCGA-COAD) were collected for this study. From the University of California Santa Cruz Xena browser, Genomic Data Commons (https://xenabrowser.net/datapages and https://gdc.xenahubs.net accessed online 7 July 2020), the RNA sequencing data (FPKM values) for gene expression were acquired. The normalized matrix files from GEO's datasets were obtained directly. The FPKM values were converted into TPM (transcripts per kilobase million) values in the next step. The "sva" package's "ComBat" technique was used to correct batch effects caused by non-biological technical biases. To plot the copy number variation landscape of genes involved in ferroptosis in human DNA, the R package "Rcircos" was used. The somatic mutation data were obtained from the TCGA database, and the data were analyzed through R Bioconductor and R (v.3.6.1) packages.

Genetic Abnormalities and Highly Significant Tumor Mutational Patterns

The MutSigCV method was used to recognize genes that have undergone substantial mutations (SMGs) (Lawrence et al., 2013). The MutSigCV, in particular, evaluates the significant improvement of nonsilent somatic mutations in a gene through addressing mutational context-specific background mutation rates. The R "maftools" package displayed the mutational landscape of FRGs and SMGs in the TCGA-COAD cohort (Mayakonda et al., 2018). The RCircos R program visualized the copy number difference landscape of 51 differentially expressed FRGs across 23 chromosomal pairs.

Unsupervised Clustering for Ferroptosis Related Genes With Prognostic Ability

Survival analysis using the software "survminer" and univariate COX regression study was used to discover FRGs with prognostic significance in COAD. Differential analysis between tumor and paired tumor-adjacent tissue, different ferroptosis subtypes were conducted by "limma" package. Fifty-one differentially expressed ferroptosis-related genes were extracted for identifying different ferroptosis subtypes mediated by FRGs. An unsupervised clustering technique was applied to categorize patients into separate ferroptosis subtypes based on the expression of these 51 FRGs. The number of clusters and their stability were established using the consensus clustering technique. We performed the steps mentioned above using the ConsensuClusterPlus program and repeated them 1,000 times to ensure classification stability (Yu et al., 2012).

Estimation of Immune Infiltrating Cells and Immune Function

The levels of infiltration of various tumor infiltrating immune cells in melanoma were assessed through the R package "CIBERSORT" and the LM22 signature with 1,000 permutations (Yoshihara et al., 2013). ESTIMATE evaluated each melanoma sample's immune and stromal contents (immune and stromal score) (Newman et al., 2015). The single-sample gene set enrichment analysis (ssGSEA) was used for measuring the relative abundance of 28 immune cell types in the tumor microenvironment (Charoentong et al., 2017). A new study compiled unique feature gene panels for identifying each immune cell type. The relative abundance of each immune cell type was represented by an enrichment score on the MCP counter in ssGSEA analysis and standardized to a unity distribution between 0 and one simultaneously (Becht et al., 2016) and TIMER (Li et al., 2017), CIBERSORT-ABS (Yoshihara et al., 2013), QUANTISEQ (Finotello et al., 2019), Xcell (Aran et al., 2017) and EPIC (Racle et al., 2017). Algorithms were tested to determine the immune responses or cellular components in groups with a high low or a high FRG score. Using a Heatmap, we were able to deduce the variations in immunological response between different algorithms.

Analysis of the Gene Set Variations

We used GSVA enrichment analysis with the "GSVA" R packages to examine the biological processes that differentiate ferroptosis subtypes. GSVA is a nonparametric and unsupervised technique, frequently used to estimate the variance in pathways and biological process activity in samples from an expression dataset (Hänzelmann et al., 2013). A p-value of less than 0.05 was considered statistically significant.

Differentially Expressed Genes Identification Between Diverse Ferroptosis Subtypes

The prior consensus clustering approach categorized patients into two unique ferroptosis subtypes, and the R package "limma" was used to identify ferroptosis-related differentially expressed genes (DEGs). Data for gene expression was normalized using voom and then passed to the lmFit and eBayes algorithm to generate differential expression statistics. The DEGs were filtered for significance using a log fold change greater than 0.585 and an adjusted p value less than 0.05.

Construction of Principal Component Analysis Ferroptosis Related Genes Related Score

For the DEG prediction analysis, we employed a univariate Cox regression model. The genes with a good prognosis were identified for further research. Following that, we created a ferroptosis-relevant gene signature using principal component analysis (PCA). The signature scores were chosen from the central components 1 and 2. This technique benefits from concentrating the score on the set with the group's largest block of well-linked (or anticorrelated) genes, whereas down-weighting scores from genes that poorly track with other set members. The FRG score is then determined using a GGI-like mechanism (Sotiriou et al., 2006)1:

$$\text{FRG Score} = \sum (\text{PC1i} + \text{PC2i})$$

Where "I" represents the FRG subtypes-related gene expression.

Nomogram Construction

In multivariate analysis ($p < 0.05$), the "rms" of the R package was applied to produce the nomogram, which incorporates factors having predictive importance. Predicted and actual survival results were compared using calibration curves. ROC curves with time dependence were also used to assess the nomogram's accuracy as a predictor.

Statistical Methods

In this work, R 4.1.0 was used to do statistical analysis. Distance correlation and Spearman analyses were used for obtaining the correlation coefficients between two variables in this study. Statistical significance was determined through Wilcoxon rank-sum test and Student's t-tests, respectively, for analyzing quantitative data. One-way analysis of variance and Kruskal-Wallis tests were employed for nonparametric and parametric comparisons of more than two groups, respectively (Hazra and Gogtay, 2016). The R package "Survminer" utilized the Cox proportional hazards (v.0.4.6), surv-cutpoint from the "survival" package, and Kaplan-Meier survival analysis to divide samples into law and high FRG score groups. The FRG score model's prognostic classification performance was calculated through the ROC (receiver operating characteristic) curve, and the "timeROC" program was used to determine the AUC (area under the curve). The hazard ratios (HR) for ferroptosis-related genes and the FRG score were calculated using a univariate Cox regression model. In the multivariate regression model, the patients with comprehensive clinical data were included, and confusing factors were removed. For every p-value that fell within the range of 0.05–0.10, it was considered statistically significant. R 4.1.0 software was used for all of the data analysis and visualization.

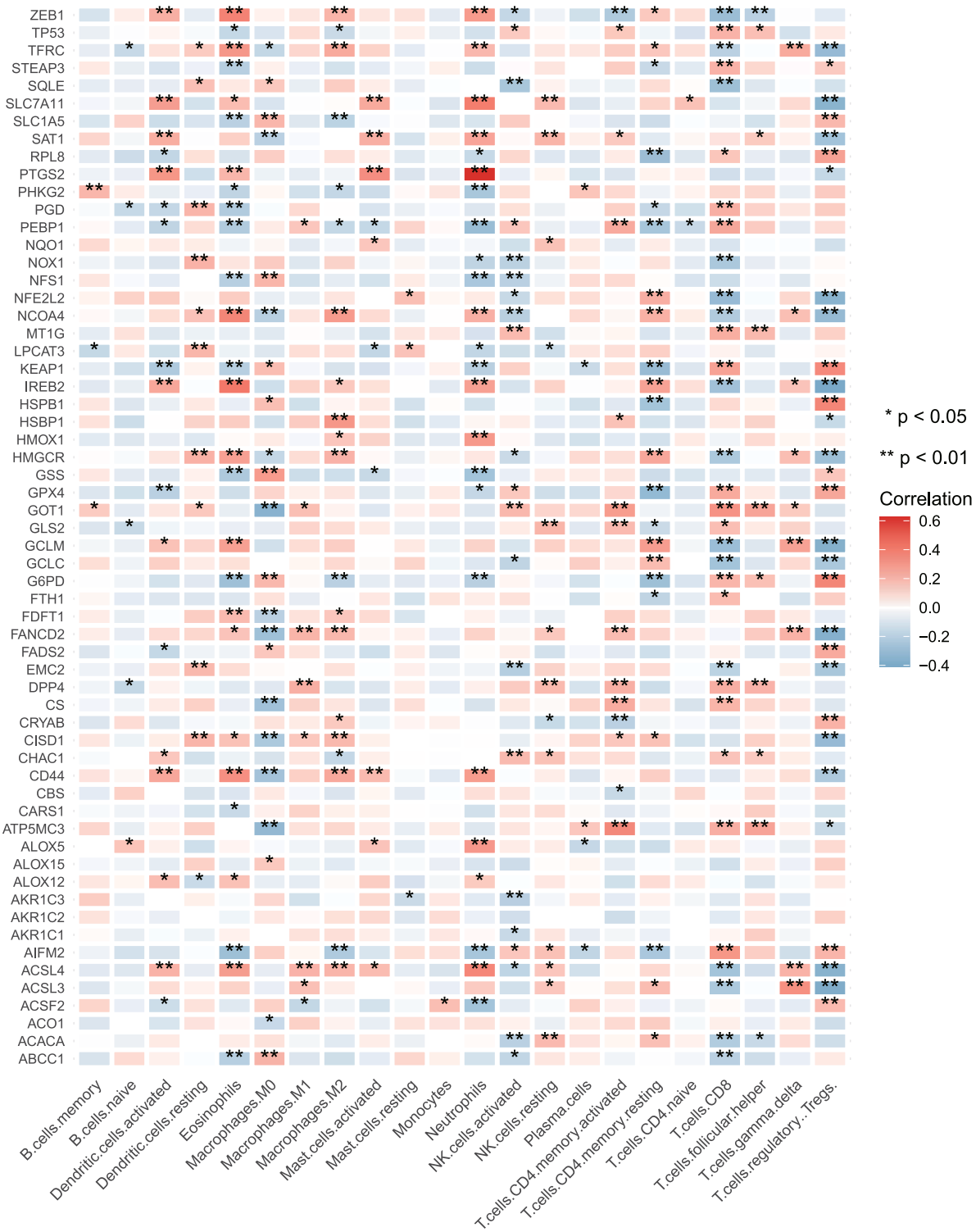


FIGURE 2 | Correlations between the tumor infiltrating immune cells and the expression of FRGs. Blue and red represent low and high expression, respectively. The asterisks are indicative of statistical *p*-value (*$p < 0.05$; **$p < 0.01$).

RESULTS

The Transcriptional and Genetic Alterations of Ferroptosis Related Genes in CRC

Genes that are involved in ferroptosis were summarized in this study by literature review. Tumors and surrounding nontumorous tissues were found to have significantly different levels of expression for 51/60 (85%) of the ferroptosis-related genes ($p < 0.05$, **Figure 1A**). Somatic mutations in 22 ferroptosis related genes (FRGs) with more than a 2% mutation rate in the TCGA-COAD cohort were summarized as shown in **Figure 1B**. FRG mutations were found in 287 of the 399 COAD samples (74.44%) (**Figure 1A**). TP53, a promising tumor treatment target (Mirgayazova et al., 2019) and ACACA were the two FRGs with the highest mutation rates (55% each), while the mutation rates of the remaining FRGs were all below 5%. We examined the prevalence of CNV alterations in 51 FRGs to see if genetic changes influenced the expression as mentioned earlier variances. The results showed that 51 regulators have a pervasive CNV change. The frequency of CNV amplification was higher than deletion in SQLE, NFS1, EMC2, GSS, ACACA, PHKG2, and TFRC. Furthermore, higher CNV frequency deletions were found in GOT1, GCLM, FDFT1 CHAC1, SLC7A11, SLC7A11, CRYAB, HSBP1, GPX4, SLC1A5, FANCD2, and HMGCR, which indicated that CNV could change gene expression in ferroptosis-related genes.

CNV gain-associated FRGs, such as FANCD2, showed decreased mRNA expression. Although CNV can elucidate several assessed variations in FRG expression, CNV is not the only mechanism regulating mRNA expression (Sebestyen et al., 2016). Other variables, such as transcription factors and DNA methylation, can influence gene expression (Koch et al., 2018) (Lambert et al., 2018). In **Figure 1D**, the locations of CNV alterations in ferroptosis-related genes on chromosomes were shown. In univariate COX analysis, eight prognostic ferroptosis-related DEGs were detected (all FDR< 0.05, **Figure 1E**). **Figure 1F** illustrates the network of FRG interactions, prognostic value, and regulator connections in COAD patients. Our findings showed a substantial variation in the genomic characteristics and levels of expression of FRGs among COAD and control samples, suggesting a dormant role for FRGs in COAD oncogenesis. Studies have revealed that ferroptosis can influence tumor immunity by regulating the adaptive immune response (Friedmann Angeli et al., 2019) (Wang, et al., 2019). As a result, we investigated the relationship between tumor-infiltrating immune cells and FRG expression. FRGs have a strong relationship with TIICs, particularly Macrophages M0, Eosinophils, Neutrophils, $CD8^+$ T cells, and Tregs (**Figure 2**).

Identification of Ferroptosis Subtypes in Colon Adenocarcinoma

To completely understand the expression patterns of FRGs implicated in tumor progression. The patients from the TCGA-COAD cohort were divided into two clusters through a consensus clustering method based on the expression profiles of the 51 differentially expressed FRGs. Our findings indicated that k = 2 seemed to be an ideal selection for categorising the complete population into cluster 1 (C1, n = 216) and cluster 2 (C2, n = 230) subtypes (**Figure 3A**). The patients classified as C1 had a significantly longer OS than those classified as C2 (log-rank test, p =0 .006; **Figure 3B**) as identified from Kaplan–Meier curves analysis. C1 was strongly enriched in DNA-damaging repair pathways, including nucleotide excision repair, mismatch repair, and non-homologous end joining, as shown by GSVA enrichment analysis.

Meanwhile, we discovered that C1 had activated protein-regulated pathways (KEGG Ubiquitin-mediated proteolysis and KEGG cell cycle) (**Figures 3C–E**). The GSVA data showed a latent link between DNA damage repair and the expression pattern of FRGs. A total of six signs for DNA damage repair (NHEJ, NER, BER, HR, FA, and MMR) were investigated. Surprisingly, in ferroptosis C1, these pathways were substantially more abundant than in C2 (**Figure 3D**). Our investigation of the TCGA-COAD cohort's mutation data revealed a lower TMB in C2 than C1 (**Figure 3F**). Then, in the TCGA-COAD cohort, we looked at how the distribution of somatic mutations differed across two ferroptosis subtypes. RYR2, ZFHX4, FAT4, MUC16, PIK3CA, SYNE1, KRAS, TTN, TP53, and APC were the top ten mutant genes in both two ferroptosis subtypes (**Figures 3G,H**). Patients in cluster one showed significantly greater rates of APC, TTN, OBSCN, PCLO and NMAH5 mutations than patients in cluster 2, but the exact opposite was discovered for TP53 and KRAS mutant levels. We also looked at a group of genes linked to Mariathasan et al. identified biological processes (Mariathasan et al., 2018). In cluster 2, EMT markers expression was higher such as EMT3, EMT2, and EMT1, pan-fibroblast TGFβ response characteristics (Pan-F TBRS), angiogenesis characteristics, and FGFR3 associated genes. Cluster one had a considerably higher expression of the CD8 effector (**Figure 3I**). Previous research found that tumors with an immune excluding phenotype have an abundance of immune cells. On the other hand, instead of invading their parenchyma, these immune cells remained in the stroma surrounding tumor cell nests. In TME, stroma activation was thought to inhibit T-cells (Chen and Mellman, 2017). Stromal activation-related pathways, such as EMT and Pan-F TBRS, enriched in cluster 2, suggesting an immune excluded phenotype. The results above also proposed that cluster one was significantly associated with DNA damage control, which usually associated with higher TMB, and implied a more favorable immunotherapy outcome.

Construction of Ferroptosis Related Genes Score and Validation of the Prognostic Ability

Given the unique complexity and heterogeneity of ferroptosis, a scoring system based on these phenotype-related genes was developed to calculate the expression pattern of FRGs in individual COAD patients, which we called the FRG score. First, we used the limma package to identify 586 ferroptosis

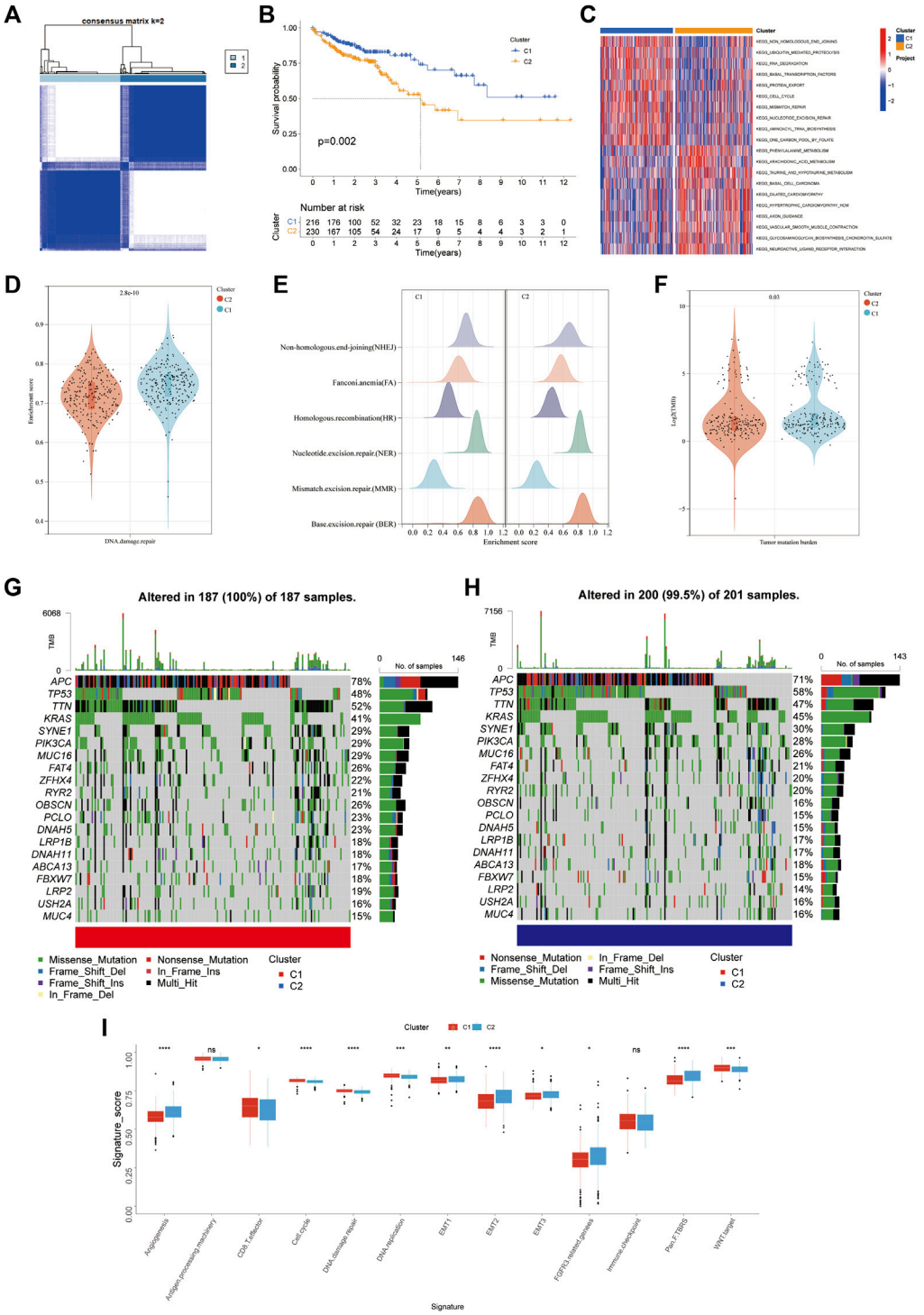


FIGURE 3 | FRGs subtypes, genomic and biological characteristics of two distinct subtypes of samples divided by consistent clustering. **(A)** Consensus matrix heatmap defining two clusters (k = 2) and their correlation area. **(B)** Kaplan-Meier curves for the OS of patients in the ferroptosis cluster 1 and 2. **(C)** GSVA enrichment analysis showing the activation states of biological pathways in distinct ferroptosis subtypes. The heatmap was used to visualize these biological processes, and red represents activated pathways and blue represents inhibited pathways. **(D)** Violin plots for enrichment score of DNA damage repair pathway in cluster 1 and 2. **(E)** Differences in enrichment score of DNA damage repair related pathways including NHEJ, FA, HR, NER, MMR, BER. **(F)** Violin plots for Tumor mutation burden in ferroptosis cluster 1 and 2. g-h. The waterfall plot of tumor somatic mutation established by those with high FRG score **(G)** and low FRG score **(H)** Each column represents individual patients. The upper barplot shows TMB, The number on the right indicates the mutation frequency in each gene. The right barplot shows the proportion of each variant type. **(I)** Differences in stroma-activated and immune-activated pathways among two distinct ferroptosis subtypes (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

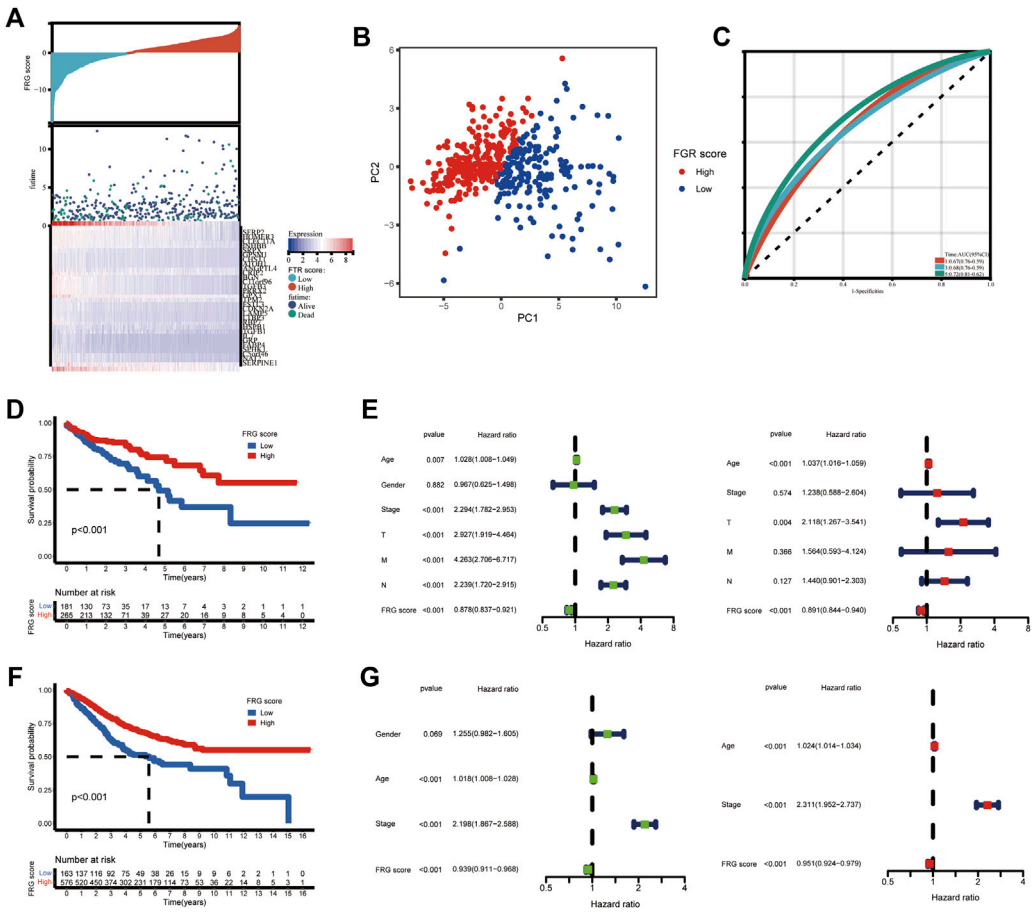


FIGURE 4 | Construction of the FGR score and validation of the prognostic ability. **(A)** Ranked dot and scatter plots showing the FRG score distribution and patient survival status. **(B)** PCA analysis based on the FRG coring system. The high and low risk patients are represented by red and blue dots, respectively. **(C)** ROC curves to predicting the sensitivity and specificity of 1,3, and 5 years survival, according to the FRG score. **(D)** Kaplan-Meier analysis of the OS of patients in TCGA cohort between the two groups. **(E)** Univariate and multivariate Cox regression results revealing the relationship of gender, age, TMN stage and FRG score with prognosis in TCGA cohort. **(F)** Kaplan-Meier analysis of the OS of patients in GEO cohort between the two groups. **(G)** Univariate and multivariate Cox regression results revealing the relationship of age, stage and the FRG score with prognosis in GEO cohort.

phenotype-related DEGs (log FC = 0.585, FDR = 0.05). "We next used univariate cox analysis to identify the DEGs that were connected to survival outcome." 31 of the 586 phenotype-related DEGs show predictive significance ($p < 0.05$). We then used principal component analysis (PCA) to create an FRG grading system based on the 31 DEGs. The FRG score distribution plot demonstrated that as FRG scores grew, so did survival times and rates (**Figure 4A**). The distinct dimensions between the high and low-FRG score groups were observed through PCA analysis (**Figure 4B**). The 1, 3, and 5 years FRG score survival rates were represented by AUC values of 0.67, 0.68, and 0.72, respectively (**Figure 4C**). Patients with high scores had significantly better survival rates than patients with low scores (log-rank test, $p < 0.001$) as per Kaplan–Meier survival curves (**Figure 4E**).

According to the Univariate findings, a lower FRG score was substantially related to an unsatisfactory OS [hazard ratio (HR): 0.878; 95% confidence interval (CI): 0.837–0.921; $p < 0.001$]. Higher age, advanced stage, deeper invasion, positive lymph nodes, and distant metastases are further clinicopathologic factors linked to poor survival (**Figure 4E**). In a multivariate analysis, FRG, T stage, and age remained independently linked with overall survival, with an HR of 0.891 (CI: 0.844–0.940, $p < 0.001$) as shown in **Figure 4F**. We confirmed these findings in GEO cohorts (GSE17536, GSE39582). The survival analysis revealed that the low-risk group had a much better prognosis than the high-risk group (**Figure 4G**). FRG score remains strongly linked with COAD patient survival time (**Figures 4H,I**).

Evaluation of the Difference Among High and Low FRG Score Groups in Tumor Microenvironment

Figure 5A illustrates a heatmap of tumor-infiltrating immune cells through EPIC algorithms, Xcell, MCP counter, TIMER, QUANTISEQ, CIBERSORT, and CIBERSORT-ABS. Based on ssGSEA of TCGA-COAD data, analysis of immune cell

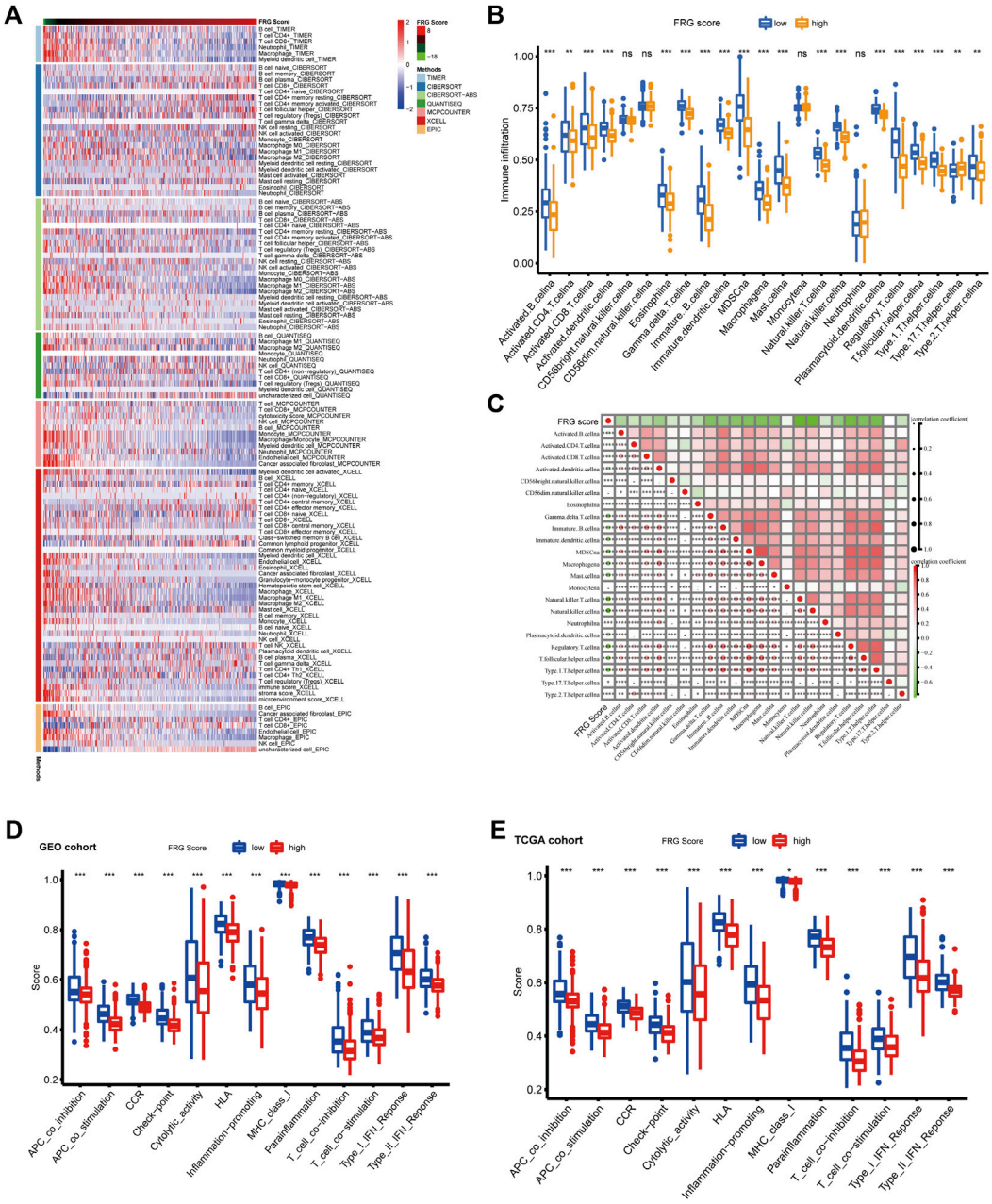


FIGURE 5 | Immune landscape in high and low FRG score groups. **(A)** Heatmap for immune infiltrating cells based on TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, MCP counter, Xcell and EPIC algorithms among high and low risk group. Blue and red represent low and high infiltrating level, respectively. **(B)** The abundance of each TME infiltrating cell in high and low FRG score groups. The upper and lower ends of the boxes represent interquartile range of values. The lines in the boxes represent median value, and black dots show outliers (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). **(C)** Correlation heatmap showing correlations between the FRG score and immune infiltrating cells evaluated by ssGSEA methods. Red and green represent low and high Pearson correlation coefficient, respectively (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). d-e. The enrichment score of immune function pathways in high and low FRG score groups. **(D)** GEO cohort **(E)** TCGA cohort. The upper and lower ends of the boxes represent interquartile range of values. The lines in the boxes represent median value, and black dots represent outliers (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

subpopulations revealed higher immune infiltration levels in the low-FRG score group (**Figure 5B**). The correlation heatmap showed a strong correlation between FRG scores with most types of immune infiltrating cells. Differential analysis of immune functions pathways using ssGSEA revealed immune stimulation pathways (INF response, CCR, co-stimulation, inflammation regulation, cytolytic activity, HLA and MHC) and immune inhibition pathways including the checkpoint, co-inhibition was significantly different between the high and low-risk groups. While patients with low FRG scores were prominently related to high level of immune infiltration and immune activation

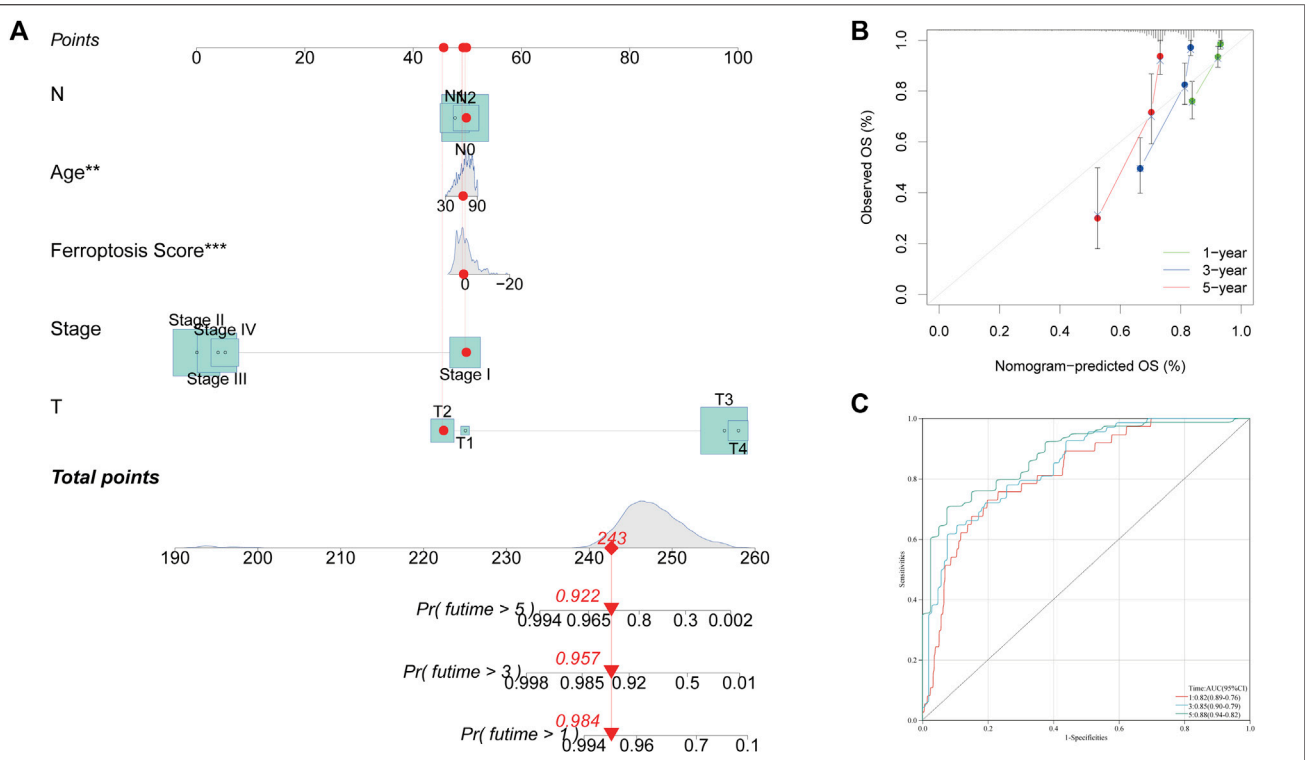


FIGURE 6 | Construction and validation of a nomogram. **(A)** Nomogram for predicting the 1, 3 and 5 years OS of COAD patients. **(B)** Calibration curves of the nomogram for predicting 1, 3 and 5 years OS of COAD patients. **(C)** ROC curves for predicting 1, 3 and 5 years OS of COAD patients.

pathways. However, patients with this low FRG score did not show a matching survival advantage. Previous studied suggested tumors with immune-excluded phenotype also showed the presence of abundant immune cells, while these immune cells were retained in the stroma surrounding tumor cell nests rather than penetrate their parenchyma. Thus, we speculated that our FGR score could serve as a predictor of immune-excluded subtypes.

Development of a Nomogram to Predict Survival

For predicting the 1, 3, and 5 years OS rates, a nomogram was developed using the FRG score and clinicopathological characteristics (**Figure 6A**). Age, FRG score, and patient stage were all used as predictors. The AUC trials of our study revealed good precision on the nomogram model for OS at 1, 3, and 5 (0.82, 0.85, and 0.88, respectively) in the TCGA-COAD cohort, indicating that the nomogram possessed a strong prediction ability (**Figure 6B**). Calibration plots revealed that the proposed nomogram performed similarly to an ideal model (**Figure 6C**).

DISCUSSION

Ferroptosis can overcome malignant cells' resistance to chemotherapy and assist in removing damaged cells. As a result, it could be a novel strategy for treating tumors (Mou et al., 2019) (Ghoochani et al., 2021) (Ye et al., 2020). According to two recent studies, ferroptosis caused by T cells in cancerous cells provided new insights of improving curative effect of the PD-1/PD-L1 antibody, and only a moderate influence of the PD-L1 antibody has been identified in ferroptosis-insensitive tumor cells (Wang, et al., 2019) (Lang et al., 2019). A lack of thorough understanding of the role of several FRGs in TME infiltration is because most research focuses on single TME infiltrating cells or single FRGs. The understanding of the role of ferroptosis in shaping TME, thereby effecting therapy efficiency and prognosis, can be improved by identifying the role of unique transcriptional ferroptosis subtypes in ferroptosis and TME antitumor immunity.

We found two unique forms of ferroptosis based on 51 FRGs that were differentially expressed. They were characterized by distinct biological processes. One of the characteristics of Cluster one was the activation of the DDR pathway. Angiogenesis features, FGFR3-related genes, and pan fibroblast TGF-β response characteristics (Pan-F TBRS) were detected in cluster two based on the TCGA data. According to previous research findings, immune cells were prevalent in tumors with an immune-exclusion pattern. Rather than entering the parenchyma of tumor cells, these immune cells remained in the stroma around them. T-cell suppression was thought to be inhibited by stroma activation in TME (Chen and Mellman, 2017). Stromal activation-related pathways, such as EMT pathways and Pan-F TBRS, enriched in cluster 2, suggested it has an immune excluded phenotype (Tang B. et al., 2020). Thus,

we defiend these two subtypes as DNA damage repair and immune-excluded phenotype. The genome's integrity is dependent on the DNA damage repair system (O'Connor, 2015). DDR pathway-related genes were strongly associated with a higher tumor mutation burden and predicted optimal anti-PD-1/PDL1 immunotherapy efficacy (Teo et al., 2018; Arora et al., 2019). Earlier research has found that gene alterations are linked to immune treatment response (Dong et al., 2017) (George et al., 2017) (Burr et al., 2017). Therefore, we evaluated the variation of TMB and somatic mutations between different ferroptosis subtypes. Consistent with previous studies, there was a significant difference between ferroptosis clusters one and two in COAD. TMB-high tumors are more likely to develop new immunogenic neo-antigens, increasing their immunotherapy susceptibility (Riaz et al., 2016). our study showed that ferroptosis was linked to the stromal and immunological TME as well as the tumor mutation burden. To above reasons, FRGs may play a significant role in the immune response to tumors.

A comprehensive assessment of the ferroptosis subtypes will enhance our understanding of TME cell-infiltrating characterization. Considering the individual heterogeneity of FRG expression, it was urgently demanded to quantify the mRNA transcriptional pattern related to ferroptosis of individual tumor. For that, we established a set of scoring system to evaluate the ferroptosis pattern of individual patients with colon cancer the FRG score. Significant links were found between the FRG score and TME features, and immunological functioning. A lower FRG score was associated with a higher level of infiltration as well as immune system activation and inhibitory pathways. (Liu et al., 2020). A comprehensive investigation of ferroptosis's aberrations and cancer-related functional consequences was conducted by Liu et al., 2020. The high-FPI group had a greater abundance of IL-6/JAK/STAT3 (immune) signaling pathways and epithelial-mesenchymal transition (stromal). Additional studies showed a link between higher ferroptosis level and poorer survival in adenocarcinoma of the lung, liver hepatitis cell carcinoma (LIHC), kidney renal papillary cell carcinoma (KIRP), kidney renal clear cell carcinoma (KIRC), and GBM, suggesting the dual role of ferroptosis in tumor progression. Therefore, regulating the ferroptosis level in TME by targeting multiple potential targets may help patients improve their prognosis.

In this study, we showed the transcriptional pattern of ferroptosis related genes played a nonnegligible role in shaping different stromal and immune TME landscape, implying ferroptosis could further determine the immune phenotypes of tumors and guide the more effective clinical practice. We also demonstrated that FRG score has a significant correlation with prognosis of colon cancer and could act as an independent prognostic biomarker for predicting patients' survival. This result has been verified in the large-scale TCGA and GEO cohorts.

However, there are a few limitations that need to be pointed out. The majority of our conclusions are based on information that was collected in the past. Prospective clinical validation in a larger COAD cohort is now required. The significance of ferroptosis in shaping TME characteristics in COAD was unknown, and it should be examined experimentally in future studies.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The data that support the findings of this study are available in TCGA database at (https://tcga-data.nci.nih.gov/tcga/) and GEO database at (https://www.ncbi.nlm.nih.gov/).

AUTHOR CONTRIBUTIONS

JL and HL organized data, performed calculations, and analyzed the data. Fan N co-designed the study, contributed to the methodological approach, analyzed data, discussed results, reviewed the manuscript, and co-supervised the project. SZ, RL, and JY envisioned the project, devised the methodological strategy, developed code, performed calculations, analyzed data, discuss results, drafted the manuscript, and co-supervised the project. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

We would like to thanks American Editor Company for language editing and proof-reading of this manuscript.

REFERENCES

Ali, H. R., Chlon, L., Pharoah, P. D. P., Markowetz, F., and Caldas, C. (2016). Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *Plos Med.* 13 (12), e1002194. doi:10.1371/journal.pmed.1002194

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

Arora, S., Velichinskii, R., Lesh, R. W., Ali, U., Kubiak, M., Bansal, P., et al. (2019). Existing and Emerging Biomarkers for Immune Checkpoint Immunotherapy in Solid Tumors. *Adv. Ther.* 36 (10), 2638–2678. doi:10.1007/s12325-019-01051-z

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 (1), 218. doi:10.1186/s13059-016-1070-5

Burr, M. L., Sparbier, C. E., Chan, Y.-C., Williamson, J. C., Woods, K., Beavis, P. A., et al. (2017). CMTM6 Maintains the Expression of PD-L1 and Regulates Anti-tumour Immunity. *Nature* 549 (7670), 101–105. doi:10.1038/nature23643

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. *Cel Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019

Chen, D. S., and Mellman, I. (2017). Elements of Cancer Immunity and the Cancer-Immune Set point. *Nature* 541 (7637), 321–330. doi:10.1038/nature21349

Conrad, M., Angeli, J. P. F., Vandenabeele, P., and Stockwell, B. R. (2016). Regulated Necrosis: Disease Relevance and Therapeutic Opportunities. *Nat. Rev. Drug Discov.* 15 (5), 348–366. doi:10.1038/nrd.2015.6

Dong, Z.-Y., Zhong, W.-Z., Zhang, X.-C., Su, J., Xie, Z., Liu, S.-Y., et al. (2017). Potential Predictive Value of TP53 and KRAS Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. *Clin. Cancer Res.* 23 (12), 3012–3024. doi:10.1158/1078-0432.Ccr-16-2554

Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., et al. (2019). Molecular and Pharmacological Modulators of the Tumor Immune Contexture Revealed by Deconvolution of RNA-Seq Data. *Genome Med.* 11 (1), 34. doi:10.1186/s13073-019-0638-6

Friedmann Angeli, J. P., Krysko, D. V., and Conrad, M. (2019). Ferroptosis at the Crossroads of Cancer-Acquired Drug Resistance and Immune Evasion. *Nat. Rev. Cancer* 19 (7), 405–414. doi:10.1038/s41568-019-0149-1

Friedmann Angeli, J. P., Schneider, M., Proneth, B., Tyurina, Y. Y., Tyurin, V. A., Hammond, V. J., et al. (2014). Inactivation of the Ferroptosis Regulator Gpx4 Triggers Acute Renal Failure in Mice. *Nat. Cel Biol* 16 (12), 1180–1191. doi:10.1038/ncb3064

George, S., Miao, D., Demetri, G. D., Adeegbe, D., Rodig, S. J., Shukla, S., et al. (2017). Loss of PTEN Is Associated with Resistance to Anti-PD-1 Checkpoint Blockade Therapy in Metastatic Uterine Leiomyosarcoma. *Immunity* 46 (2), 197–204. doi:10.1016/j.immuni.2017.02.001

Ghoochani, A., Hsu, E.-C., Aslan, M., Rice, M. A., Nguyen, H. M., Brooks, J. D., et al. (2021). Ferroptosis Inducers Are a Novel Therapeutic Approach for Advanced Prostate Cancer. *Cancer Res.* 81 (6), 1583–1594. doi:10.1158/0008-5472.CAN-20-3477

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics* 14, 7. doi:10.1186/1471-2105-14-7

Hassannia, B., Vandenabeele, P., and Vanden Berghe, T. (2019). Targeting Ferroptosis to Iron Out Cancer. *Cancer Cell* 35 (6), 830–849. doi:10.1016/j.ccell.2019.04.002

Hazra, A., and Gogtay, N. (2016). Biostatistics Series Module 3: Comparing Groups: Numerical Variables. *Indian J. Dermatol.* 61 (3), 251–260. doi:10.4103/0019-5154.182416

Ingold, I., Berndt, C., Schmitt, S., Doll, S., Poschmann, G., Buday, K., et al. (2018). Selenium Utilization by GPX4 Is Required to Prevent Hydroperoxide-Induced Ferroptosis. *Cell* 172 (3), 409–422. doi:10.1016/j.cell.2017.11.048

Kagan, V. E., Mao, G., Qu, F., Angeli, J. P. F., Doll, S., Croix, C. S., et al. (2017). Oxidized Arachidonic and Adrenic PEs Navigate Cells to Ferroptosis. *Nat. Chem. Biol.* 13 (1), 81–90. doi:10.1038/nchembio.2238

Koch, A., Joosten, S. C., Feng, Z., de Ruijter, T. C., Draht, M. X., Melotte, V., et al. (2018). Analysis of DNA Methylation in Cancer: Location Revisited. *Nat. Rev. Clin. Oncol.* 15 (7), 459–466. doi:10.1038/s41571-018-0004-4

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., et al. (2018). The Human Transcription Factors. *Cell* 172 (4), 650–665. doi:10.1016/j.cell.2018.01.029

Lang, X., Green, M. D., Wang, W., Yu, J., Choi, J. E., Jiang, L., et al. (2019). Radiotherapy and Immunotherapy Promote Tumoral Lipid Oxidation and Ferroptosis via Synergistic Repression of SLC7A11. *Cancer Discov.* 9 (12), 1673–1685. doi:10.1158/2159-8290.Cd-19-0338

Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., et al. (2013). Mutational Heterogeneity in Cancer and the Search for New Cancer-Associated Genes. *Nature* 499 (7457), 214–218. doi:10.1038/nature12213

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

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

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

Mirgayazova, R., Khadiullina, R., Mingaleeva, R., Chasov, V., Gomzikova, M., Garanina, E., et al. (2019). Novel Isatin-Based Activator of P53 Transcriptional Functions in Tumor Cells. *Mol. Biol. Res. Commun.* 8 (3), 119–128. doi:10.22099/mbrc.2019.34179.1419

Mou, Y., Wang, J., Wu, J., He, D., Zhang, C., Duan, C., et al. (2019). Ferroptosis, a New Form of Cell Death: Opportunities and Challenges in Cancer. *J. Hematol. Oncol.* 12 (1), 34. doi:10.1186/s13045-019-0720-y

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

O'Connor, M. J. (2015). Targeting the DNA Damage Response in Cancer. *Mol. Cel* 60 (4), 547–560. doi:10.1016/j.molcel.2015.10.040

Pitt, J. M., Marabelle, A., Eggermont, A., Soria, J.-C., Kroemer, G., and Zitvogel, L. (2016). Targeting the Tumor Microenvironment: Removing Obstruction to Anticancer Immune Responses and Immunotherapy. *Ann. Oncol.* 27 (8), 1482–1492. doi:10.1093/annonc/mdw168

Quail, D. F., and Joyce, J. A. (2013). Microenvironmental Regulation of Tumor Progression and Metastasis. *Nat. Med.* 19 (11), 1423–1437. doi:10.1038/nm.3394

Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E., and Gfeller, D. (2017). Simultaneous Enumeration of Cancer and Immune Cell Types from Bulk Tumor Gene Expression Data. *Elife* 6. doi:10.7554/eLife.26476

Riaz, N., Morris, L., Havel, J. J., Makarov, V., Desrichard, A., and Chan, T. A. (2016). The Role of Neoantigens in Response to Immune Checkpoint Blockade. *Intimm* 28 (8), 411–419. doi:10.1093/intimm/dxw019

Sebestyén, E., Singh, B., Miñana, B., Pagès, A., Mateo, F., Pujana, M. A., et al. (2016). Large-scale Analysis of Genome and Transcriptome Alterations in Multiple Tumors Unveils Novel Cancer-Relevant Splicing Networks. *Genome Res.* 26 (6), 732–744. doi:10.1101/gr.199935.115

Song, J., Liu, T., Yin, Y., Zhao, W., Lin, Z., Yin, Y., et al. (2021). The Deubiquitinase OTUD1 Enhances Iron Transport and Potentiates Host Antitumor Immunity. *EMBO Rep.* 22 (2), e51162. doi:10.15252/embr.202051162

Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., et al. (2006). Gene Expression Profiling in Breast Cancer: Understanding the Molecular Basis of Histologic Grade to Improve Prognosis. *J. Natl. Cancer Inst.* 98 (4), 262–272. doi:10.1093/jnci/djj052

Tang, B., Zhu, J., Li, J., Fan, K., Gao, Y., Cheng, S., et al. (2020a). The Ferroptosis and Iron-Metabolism Signature Robustly Predicts Clinical Diagnosis, Prognosis and Immune Microenvironment for Hepatocellular Carcinoma. *Cell Commun Signal* 18 (1), 174. doi:10.1186/s12964-020-00663-1

Tang, R., Xu, J., Zhang, B., Liu, J., Liang, C., Hua, J., et al. (2020b). Ferroptosis, Necroptosis, and Pyroptosis in Anticancer Immunity. *J. Hematol. Oncol.* 13 (1), 110. doi:10.1186/s13045-020-00946-7

Teo, M. Y., Seier, K., Ostrovnaya, I., Regazzi, A. M., Kania, B. E., Moran, M. M., et al. (2018). Alterations in DNA Damage Response and Repair Genes as Potential Marker of Clinical Benefit from PD-1/pd-L1 Blockade in Advanced Urothelial Cancers. *Jco* 36 (17), 1685–1694. doi:10.1200/jco.2017.75.7740

Topalian, S. L., Hodi, F. S., Brahmer, J. R., Gettinger, S. N., Smith, D. C., McDermott, D. F., et al. (2012). Safety, Activity, and Immune Correlates of Anti-PD-1 Antibody in Cancer. *N. Engl. J. Med.* 366 (26), 2443–2454. doi:10.1056/NEJMoa1200690

Wang, W., Green, M., Choi, J. E., Gijón, M., Kennedy, P. D., Johnson, J. K., et al. (2019). CD8+ T Cells Regulate Tumour Ferroptosis during Cancer Immunotherapy. *Nature* 569 (7755), 270–274. doi:10.1038/s41586-019-1170-y

Wu, Y., Yu, C., Luo, M., Cen, C., Qiu, J., Zhang, S., et al. (2020). Ferroptosis in Cancer Treatment: Another Way to Rome. *Front. Oncol.* 10, 571127. doi:10.3389/fonc.2020.571127

Yang, W. S., SriRamaratnam, R., Welsch, M. E., Shimada, K., Skouta, R., Viswanathan, V. S., et al. (2014). Regulation of Ferroptotic Cancer Cell Death by GPX4. *Cell* 156 (1-2), 317–331. doi:10.1016/j.cell.2013.12.010

Ye, Z., Liu, W., Zhuo, Q., Hu, Q., Liu, M., Sun, Q., et al. (2020). Ferroptosis: Final Destination for Cancer? *Cell Prolif* 53 (3), e12761. doi:10.1111/cpr.12761

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 (5), 284–287. doi:10.1089/omi.2011.0118

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, Li, Zhao, Lin, Yu and Fan. 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: 25 April 2022
doi: 10.3389/fgene.2022.850101

Establishment and Application of a Prognostic Risk Score Model Based on Characteristics of Different Immunophenotypes for Lung Adenocarcinoma

*Hong Gao[1], Yanhong Liu[1], Yue Hu[1], Meiling Ge[1], Jie Ding[1] and Qing Ye[2,3]**

[1]Biobank of Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing, China, [2]Department of Pathology, The First Affiliated Hospital of USTC, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China, [3]Intelligent Pathology Institute, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China

Objective: Lung adenocarcinoma (LUAD) is a highly heterogeneous tumor. Tumor mutations and the immune microenvironment play important roles in LUAD development and progression. This study was aimed at elucidating the characteristics of patients with different tumor immune microenvironment and establishing a prediction model of prognoses and immunotherapy benefits for patients with LUAD.

Materials and Methods: We conducted a bioinformatics analysis on data from The Cancer Genome Atlas and Gene Expression Omnibus (training and test sets, respectively). Patients in the training set were clustered into different immunophenotypes based on tumor-infiltrating immune cells (TIICs). The immunophenotypic differentially expressed genes (IDEGs) were used to develop a prognostic risk score (PRS) model. Then, the model was validated in the test set and applied to evaluate 42 surgery patients with early LUAD.

Results: Patients in the training set were clustered into high (Immunity_H), medium (Immunity_M), and low (Immunity_L) immunophenotype groups. Immunity_H patients had the best survival and more TIICs than Immunity_L patients. Immunity_M patients had the worst survival, characterized by most $CD8^+$ T and Treg cells and highest expression of PD-1 and PD-L1. The PRS model, which consisted of 14 IDEGs, showed good potential for predicting the prognoses of patients in both training and test sets. In the training set, the low-risk patients had more TIICs, higher immunophenoscores (IPSs) and lower mutation rates of driver genes. The high-risk patients had more mutations of DNA mismatch repair deficiency and APOBEC (apolipoprotein B mRNA editing enzyme catalytic polypeptide-like). The model was also a good indicator of the curative effect for immunotherapy-treated patients. Furthermore, the low-risk group out of 42 patients, which was evaluated by the PRS model, had more TIICs, higher IPSs and better progression-free survival. Additionally, IPSs and PRSs of these patients were correlated with EGFR mutations.

OPEN ACCESS

Edited by:
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Ming Yi,
Huazhong University of Science and Technology, China
Jelena Stojsic,
University of Belgrade, Serbia

****Correspondence:***
Qing Ye
qingye1998@126.com

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

Received: *07 January 2022*
Accepted: *17 March 2022*
Published: *25 April 2022*

Citation:
Gao H, Liu Y, Hu Y, Ge M, Ding J and Ye Q (2022) Establishment and Application of a Prognostic Risk Score Model Based on Characteristics of Different Immunophenotypes for Lung Adenocarcinoma. Front. Genet. 13:850101. doi: 10.3389/fgene.2022.850101

Conclusion: The PRS model has good potential for predicting the prognoses and immunotherapy benefits of LUAD patients. It may facilitate the diagnosis, risk stratification, and treatment decision-making for LUAD patients.

Keywords: lung adenocarcinoma, tumor-infiltrating immune cells, immunophenotypes, prognostic model, immunotherapy

INTRODUCTION

In the past few decades, the morbidity and mortality of lung cancer have increased year after year. According to the latest WHO data, lung cancer, with morbidity and mortality rates of 11.4 and 18.0% respectively, is the leading cause of cancer-related death (Bray et al., 2018). It also remains the most common cancer and leading cause of cancer-related death in China (Wu et al., 2019). Lung adenocarcinoma (LUAD) is the most common histologic subtype of non-small cell lung cancer (NSCLC), accounting for 40% of lung cancer incidence (Chen et al., 2014). For a long time, LUAD has been considered a non-immunogenic tumor with high heterogeneity. However, increasing evidence indicates the occurrence and development of LUAD depend on tumor mutations and are closely related to the tumor immune microenvironment (TIME).

The TIME is a complex assembly of the tumor, immune, stromal, and extracellular components (Schurch et al., 2020). The organization of these components at the cellular and tissue levels plays a crucial role in tumor progression (Binnewies et al., 2018; Junttila and de Sauvage, 2013). Tumor development and the immune system, with several innate and adaptive immune cell subpopulations, some of which show phenotypic plasticity and possess memory, are closely linked. The interactions and balance between them two directly influence immunotherapy response (Charoentong et al., 2017). Tumor-infiltrating immune cells (TIICs) play an important part in the TIME of LUAD (Bussard et al., 2016); however, the specific mechanisms remain controversial. With the development of detection techniques, researchers have found that the activation of TIICs and immune escape occur before lung cancer invasion, and TIICs are significantly associated with the survival rate (Mascaux et al., 2019). Furthermore, with the application of immune checkpoint inhibitors (ICIs) attracting widespread attention, the indispensable role of TIICs in immunotherapy has also become a research focus. The analysis of immunogenomic data by using bioinformatics tools can provide information on the composition, function, and localization of TIICs; predict tumor mutation burden (TMB) and tumor neoantigen; and indicate immunotherapy response (Schumacher and Schreiber, 2015).

Therefore, we conducted immunotyping of patients based on TIICs and constructed a prognostic risk model based on differentially expressed genes of each phenotype to evaluate the prognosis and immunotherapeutic benefits. We hoped to determine the characteristics of patients with different TIME; screen immune-related differentially expressed genes; establish an effective model to predict the benefits of immunotherapy and the prognoses of patients with LUAD.

MATERIALS AND METHODS

Downloading and Preprocessing of Data on mRNA Sequencing and Somatic Mutations

Data on mRNA sequencing (Fragments Per kilobase of exon model per Million mapped fragments, FPKM) and clinical data of LUAD were downloaded from TCGA as the training set for the next-step analysis. The mRNA sequencing (FPKM) and clinical data of GSE101929, GSE50081, GSE41271, and GSE42127 were downloaded from the Gene Expression Omnibus (GEO) platform. The batch effects between GEO datasets were corrected with the R package SVAR (Irizarry et al., 2003). The processed data were used as the test set for the subsequent analysis. The mRNA sequencing and clinical data of GSE13522 and GSE126044 were also downloaded to evaluate the predictive power of the PRS model for an immunotherapeutic response. The somatic mutation data for the training set were downloaded and analyzed using the R package maftools (Mayakonda et al., 2018). The TMBs and mutation rates of LUAD-related driver genes were calculated. The list of driver genes was derived from Integrative Onco Genomics (https://www.intogen.org/search).

Patient Recruitment and Sample Inclusion

A total of 42 patients (referred as NJDT patients) with stage I or II LUAD, who underwent surgeries in Nanjing Drum Tower Hospital from January 2017 to January 2018 were randomly selected. Paraffin-embedded samples of tumor and normal tissues were collected. Sections of the paraffin-embedded tissues were stained using hematoxylin–eosin and examined by two pathologists. The samples were graded and classified according to the Eighth Edition of TNM Classification for Lung Cancer proposed by IASLC (Goldstraw et al., 2016). mRNA high-throughput sequencing was performed on tumor and matching normal samples, and the FPKM data was used for follow-up analysis.

Consensus Clustering of TIICs

The Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression Data (ESTIMATE) algorithm was used to evaluate the stromal and immune components of samples in the training set and the stromal score, tumor purity, and immune score were calculated (Yoshihara et al., 2013). Based on signal sample Gene Set Enrichment Analysis (ssGSEA) using the R packages of gsva (Hanzelmann et al., 2013) and GSEABase (Reimand et al., 2019), 24 types of TIICs were classified (Bindea et al., 2013): innate immunity (natural killer cells [NKs], NK $CD56^{dim}$ cells, NK $CD56^{bright}$ cells, dendritic cells [DCs], activated DCs [aDCs], immature DCs [iDCs], plasmacytoid dendritic cells [pDCs], neutrophils, eosinophils, mast cells, and macrophages) and adaptive immunity (B, T, T helper 1 [Th1],

Th2, T gamma delta [Tgd], $CD4^+$ T, $CD8^+$ T, T central memory [Tcm], T effector memory [Tem], T follicular helper [Tfh], Th17, regulatory T [Treg], and cytotoxic cells). The training set was clustered hierarchically into high (Immunity_H), medium (Immunity_M), and low (Immunity_L) immunophenotype groups. Then, the CIBERSORT algorithm was used to calculate the relative content of each immune cell subset among 22 types of leukocyte subsets (LM22 signature) with 1,000 permutations (Newman et al., 2015). When the p value of the output for each subset was <0.05, the relative contents were considered accurate and suitable for further analysis.

Identification and Enrichment of IDEGs

For genes with multiple probes, the average of the probes was used as the gene expression. The R package limma (Ritchie et al., 2015) was used to identify DEGs between normal and tumor samples (DEGs_NT) in the training set. DEGs between Immunity_H and Immunity_M (DEGs_HM) and DEGs between Immunity_H and Immunity_L (DEGs_HL) groups were also screened in the same method. DEGs were defined by the false discovery rate (FDR) < 0.05 and Log2|FoldChange| > 1. The intersection of DEGs_NT, DEGs_HM, and DEGs_HL was used to determine IDEGs. Differential pathways were enriched using Gene Set Enrichment Analysis (GSEA). With the | normalized enrichment score (NES)| >1, nominal p value < 0.05, and FDR <25%, the enrichment was considered significant.

Establishment and Validation of the Prognostic Risk Score Model

Univariate Cox regression was used to analyze the correlation between IDEGs and overall survival (OS); genes with $p < 0.05$ were screened. Then, the above genes were analyzed by LASSO regression (Gao et al., 2010) and lambda (λ) values were calculated. Based on the λ value, which corresponded to the minimum mean standard error in the cross-validation, variables were obtained and regression coefficients were calculated. The regression coefficients multiplied by the mRNA levels of 14 genes were used to construct the formula. The median risk score in the training set was used as the grouping cut-off value. Patients with a risk score greater than the cut-off value were classified into the high-risk group; the rest were classified into the low-risk group. Meanwhile, the test set was divided into high- and low-risk groups by using the same cut-off value. The OS curves of the patients in the two sets were plotted, and Log-rank test was used to analyze the differences. The receiver operating characteristic curves (ROCs) of OS in the two sets were plotted, and the areas under curves were calculated to evaluate the predictive performance of the model. Multivariate Cox regression analysis was performed to construct nomograms in both sets.

Clustering Analysis of *de Novo* Somatic Mutation Signatures in the Training Set

The R package SomaticSignatures (Gehring et al., 2015) was used to identify and cluster *de novo* mutation signatures. The number of these signatures was determined by explained variance and residual sum of squares (RSS). The best number of *de novo* signatures was chosen for clustering. De novo signatures were then compared to 30 curated signatures in the Cancer Gene Census (COSMIC) by using cosine similarity (Cui et al., 2020), Cochran-Armitage trend test was used to examine the mutation signature contribution among groups.

Immunophenoscores

Immunophenoscores (IPSs) were calculated according to the recently published reports (Charoentong et al., 2017; Hakimi et al., 2016). In brief, consensus determinants including 20 single factors and 6 cell types were divided into four categories: effector cells, suppressive cells, MHC-related molecules, and checkpoints or immunomodulators. The Z scores of the determinants included in the particular category were positively weighted with one and negatively weighted with one. The weighted averaged Z score was then calculated by averaging the Z scores within the respective category leading to four values. The IPSs were calculated on an arbitrary scale of 0–10 based on the sum of the weighted average Z scores of the four categories.

Workflow

The workflow of this study is shown in **Figure 1**.

Statistical Analysis

All statistical analyses were conducted with the R software (version 4.0.2). The Wilcoxon test was used to compare continuous variables in two groups. The Kaplan-Meier plotter was employed to generate survival curves for the subgroups in each dataset. The Log-rank test was used to evaluate significant differences in survival. The Chi-square test or Fisher's exact test were used to analyze the clinicopathological categorical variables between the different PRS subgroups. Spearman correlation analysis was used to compute the correlation coefficient between indicators. The multiple hypothesis test with the Benjamini–Hochberg method was used to control FDR. All statistical tests were two-sided, and p values less than 0.05 were considered statistically significant ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

RESULTS

TIICs Evaluation and Immunotyping

We analyzed the contents of 24 types of TIICs in both sets and evaluated the results by principal component analysis (PCA). There were significant differences between tumor and normal samples. The differences could be used to distinguish normal and tumor tissues (**Supplementary Figures S1A–C**). The contents of adaptive immune cells increased in tumor tissues, while those of innate immune cells decreased (Wilcoxon test, $p < 0.05$) (**Supplementary Figures S1B–D**).

Furthermore, the ESTIMATE algorithm was used to evaluate mRNA profiles of tumor samples in the training set. The OS of the patients in the high score (greater than the median value) group based on the immune scores was higher than those of the patients in the low score group, and the intergroup difference was

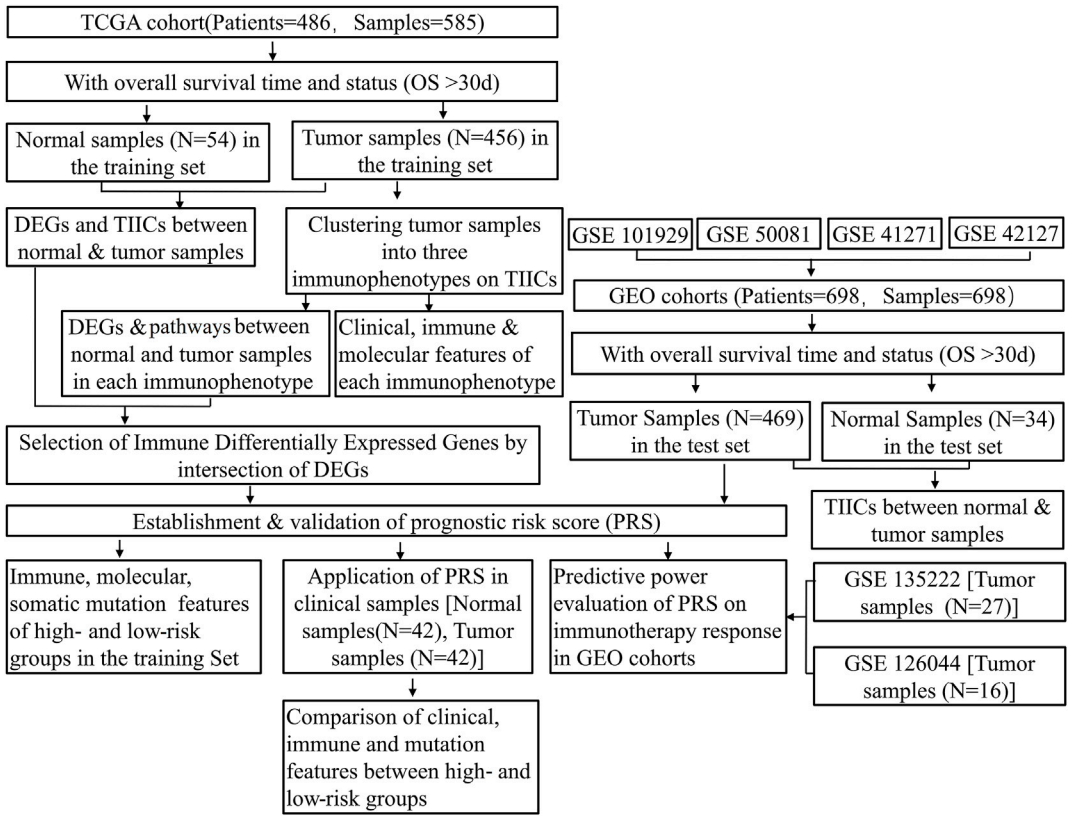


FIGURE 1 | Flowchart of the study protocol. TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; OS, overall survival; ESTIMATE, Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data; TIICs, tumor-infiltrating immune cells; DEGs, differentially expressed genes; PRS, prognostic risk score.

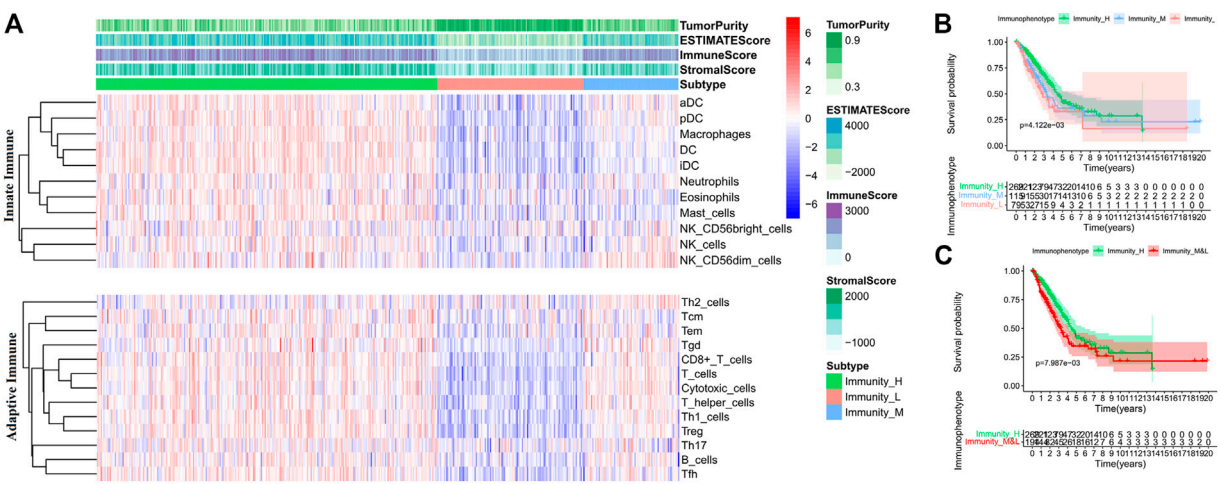


FIGURE 2 | Comparison of TIICs and overall survival (OS) of patients in each immunophenotype in the training set. **(A)** TIIC contents of patients in each immunophenotype. **(B)** Comparisons of OS among the three immunophenotypes. **(C)** Comparisons of OS between Immunity_H and other patients.

significant (Log-rank test, $p < 0.05$) (**Supplementary Figure S2A**). It indicates that the prognoses of patients with high immune scores are better than those of the patients with low immune scores. Therefore, hierarchical cluster analysis was performed on the TIICs in tumor samples (**Supplementary Figure S2B**). According to the immune scores, three clusters

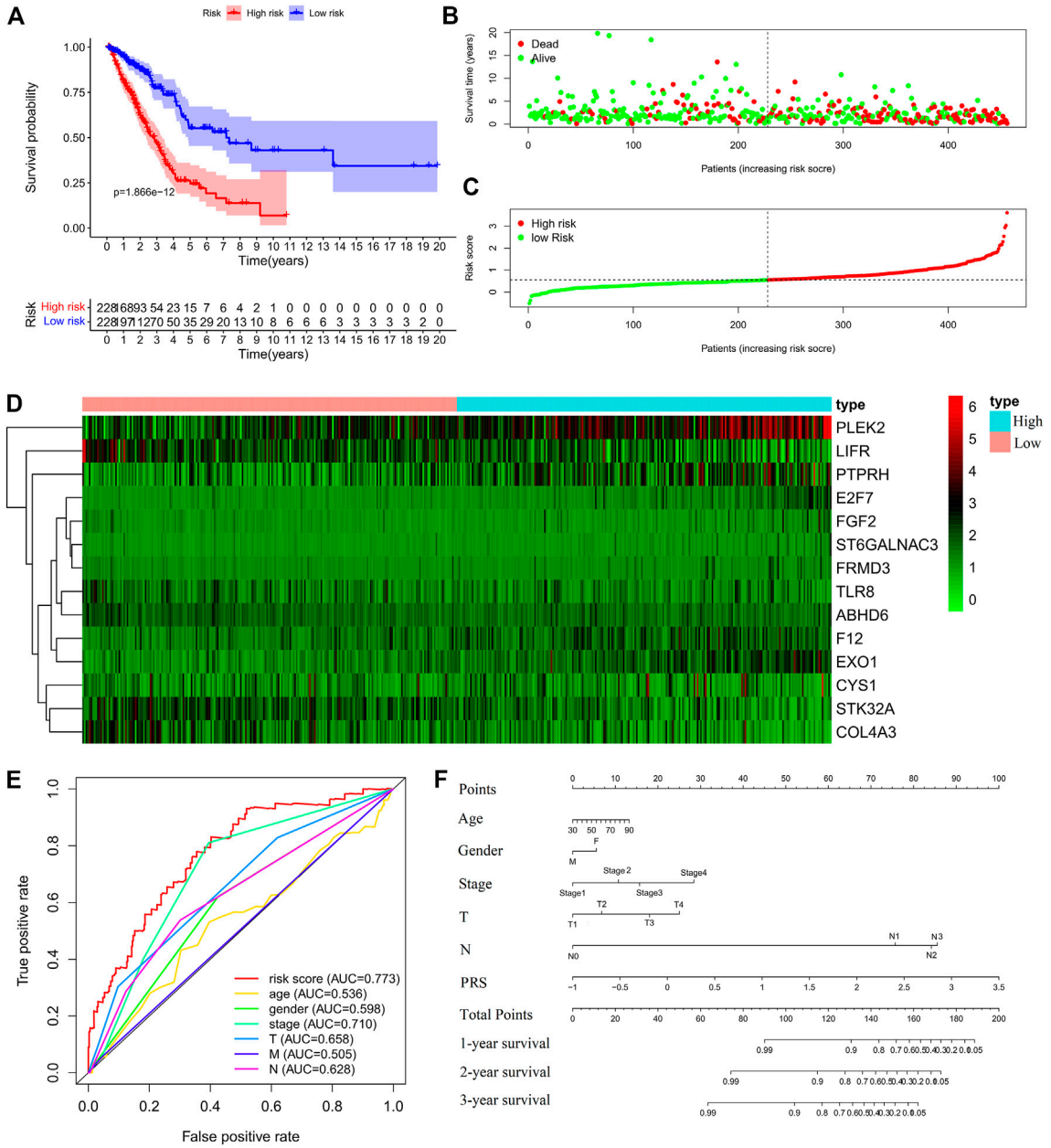


FIGURE 3 | Construction of the prognostic risk score (PRS) model in the training set. **(A)** Comparisons of OS between the high- and low-risk groups. **(B)**. Distribution of survival time of patients with different outcomes; **(C)** Distribution of increasing risk scores in high- and low-risk groups; **(D)** Heatmap of the fourteen-gene expression profiles in high- and low-risk groups; **(E)** ROCs of models with gender, stage, tumor size (T), lymph node metastasis (N), distant metastasis (M), and PRS as variables, respectively; **(F)** Nomogram to predict the 1-year, 2-years and 3-years survival rates of the patients in the training set by using gender, stage, tumor size (T), lymph node metastasis (N), distant metastasis (M), and risk score as variables, respectively.

were defined as high (Immunity_H), medium (Immunity_M), and low immunophenotypes (Immunity_L) (**Figure 2A**). The OS of the three immunophenotype groups was statistically different (Log-rank test, $p < 0.05$) (**Figure 2B**). The patients in the Immunity_H group had the better OS than others (Log-rank test, $p < 0.05$) (**Figure 2C**). The TIICs in each immunophenotype were further compared. The levels of mature immune cells in the Immunity_L group were the lowest. Almost all innate immune cells in the Immunity_H group were more than those in the Immunity_M group, except Tfh, $CD8^+$ T, and Treg cells. These three kinds of cell increased in the Immunity_M group (Wilcoxon test, $p < 0.05$) (**Supplementary Figures S3A,B**). We also used the CIBERSORT method to quantitate TIICs in each immunophenotype group. Twenty-two types of immune cells were quantified; however, the number of $CD4^+$ T naive cells was 0 in all samples. Hence, only 21 types of immune cells were

finally analyzed. The contents of most innate immune cells in the Immunity_H group were higher than those in the Immunity_M group (Wilcoxon test, $p < 0.05$) (**Supplementary Figure S4A**); however, the numbers of $CD8^+$ T and Tfh cells in the Immunity_H group were lower than those in the Immunity_M group (Wilcoxon test, $p < 0.05$) (**Supplementary Figure S4B**).

Feature Analysis of Different Immunophenotypes in the Training Set

We analyzed the clinical features of patients in the three immunophenotype groups (**Figure 3A**, **Supplementary Table S1**). The proportion of female patients in the Immunity_H group was the highest. The patients in the Immunity_H group had the lowest TMB and the highest IPSs (Wilcoxon test, $p < 0.05$) (**Supplementary Figure S5A**). We also compared HLA expressions and checkpoints in the three immunophenotype groups. The levels of PD-L1, PD-1, FASL, CTLA4, and CD244 in the Immunity_M group were higher than those in the Immunity_H group (**Supplementary Figure S4C**). In the Immunity_L group, the expression levels of HLA and checkpoints were lower than those in the Immunity_H group (Wilcoxon test, $p < 0.05$) (**Supplementary Figure S4D**).

GSEA analysis of the differential enrichment pathways of Kyoto Encyclopedia of Genes and Genomes (KEGG) showed that, the pathways of nucleotide sugar metabolism, DNA stability and autophagy regulation were significantly upregulated (**Supplementary Figure S5B**), but immune-related pathways and cell adhesion were significantly downregulated of tumor samples in all three immunophenotypes (**Supplementary Figure S5C**). The Immunity_H group had the most obvious upregulation of metabolism pathways, such as glucose, lipid and water salt metabolism and lysosome pathways. The Immunity_M group was more strongly related to homologous recombination, DNA replication and repair and gene transcription. The Immunity_L group was specifically associated with lowered immune-related pathways, including B and T cell receptor signaling pathways, NK cell mediated cytotoxicity, cytokine receptor interaction and complement-related pathways (**Supplementary Figures S5D,E**). Thereafter, we intersected DEG_HL, DEG_HM and DEG_NT and obtained 421 IDEGs for subsequent screening (**Supplementary Figure S5F**). The molecular function and biological processes of these genes covered immune response, regulation of gene silencing, glucolipid metabolism, cell adhesion and blood coagulation (**Supplementary Figure S5G**).

Construction and Validation of the PRS Model

Cox regression analysis was performed for the candidate genes among the IDEGs that were specifically associated with OS (Log-rank, $p < 0.05$), followed by LASSO logistic analysis. The most suitable tuning parameters (λ) and coefficients were calculated by cross-validation (**Supplementary Figure S6A**). Finally, 14 IDEGs were selected to construct the PRS model. The formula was as follows: Prognostic Risk Score = (−0.0461 × *TLR8* mRNA level) + (0.0992 × *FGF2* mRNA level) + (0.0467 × *F12* mRNA level) + (0.3515 × *ST6GALNAC3* mRNA level) + (0.0198 × *PTPRH* mRNA level) + (0.0368 × *EXO1* mRNA level) + (0.0182 × *FRMD3* mRNA level) + (0.1891 × *E2F7* mRNA level) + (−0.1644 × *ABHD6* mRNA level) + (-0.0423 × *STK32A* mRNA level) + (−0.0203 × *COL4A3* mRNA level) + (0.0178 × *PLEK2* mRNA level) + (−0.0222 * *LIFR* mRNA level) + (0.04453 × CYS1 mRNA level). The median score in the training set was considered as the cut-off value, and the patients were divided into high-risk (228 cases) and low-risk (228 cases) groups (**Supplementary Table S2**). Survival analysis showed that OS (**Figures 3A–C**), disease-free survival (DFS), progression-free survival (PFS), and disease-specific survival (DSS) (**Supplementary Figure S6B**) in the high-risk group were significantly worse than those of the low-risk group (Log-rank test, $p < 0.001$). The expression profiles of 14 genes were visualized as a heatmap (**Figure 3D**). The area under the ROC was 0.773 (**Figure 3E**). Then, the variables of age, stage, T, M, N, and PRS were analyzed to establish a nomogram for predicting the 3-years survival rate (**Figure 3F**).

Further, the PRS model was invalidated in the test set. The above PRS formula, cut-off value, and grouping method were used to divide patients into high-risk (248 cases) and low-risk (221 cases) groups (**Supplementary Table S3**). The OS of the high-risk group was significantly worse than that of the low-risk group (Log-rank, $p < 0.001$) (**Figures 4A–C**). The expression profiles of the PRS model genes were visualized as heatmaps (**Figure 4D**). The area under the ROC was 0.707 (**Figure 4E**). Then, the variables of age, stage, and risk score were analyzed to establish a nomogram for predicting the 3-years survival rate (**Figure 4F**).

Molecular, Immune, and Mutation Features of PRS Subgroups in the Training Set

After obtaining the reliable PRS model, we analyzed clinical and molecular features of PRS subgroups in the training set. The PRS subgroups showed significant differences in sex, T and N classifications, and stage. The proportion of female, T1, N0, and Stage I patients in the low-risk group was significantly higher than those in the high-risk group (Chi-square test, $p < 0.05$) (**Figure 5A**, **Supplementary Table S4**). GSEA analysis on the enrichment pathways of KEGG between the two PRS subgroups showed that pathways of cell cycle, DNA replication, homologous recombination, mismatch repair, p53 signal pathway, which were associated with gene mutation and chromosome instability, were significantly upregulated in the high-risk patients. In contrast, ABC transporters, B cell receptor signaling pathway, cell adhesion molecules (CAMs), histidine metabolism, and mTOR signaling pathway and other immune-related pathways were significantly upregulated in the low-risk patients (**Figure 5B**). In addition, PRSs of all patients were positively correlated with TMBs (Spearman correlation, $p < 0.001$), and negatively correlated with IPSs (Spearman correlation, $p < 0.05$) (**Figures 5C,D**) Meanwhile, in combination with literature data, we compared two prediction indicators of neoantigens (**Figure 5E**): the counts of mutations

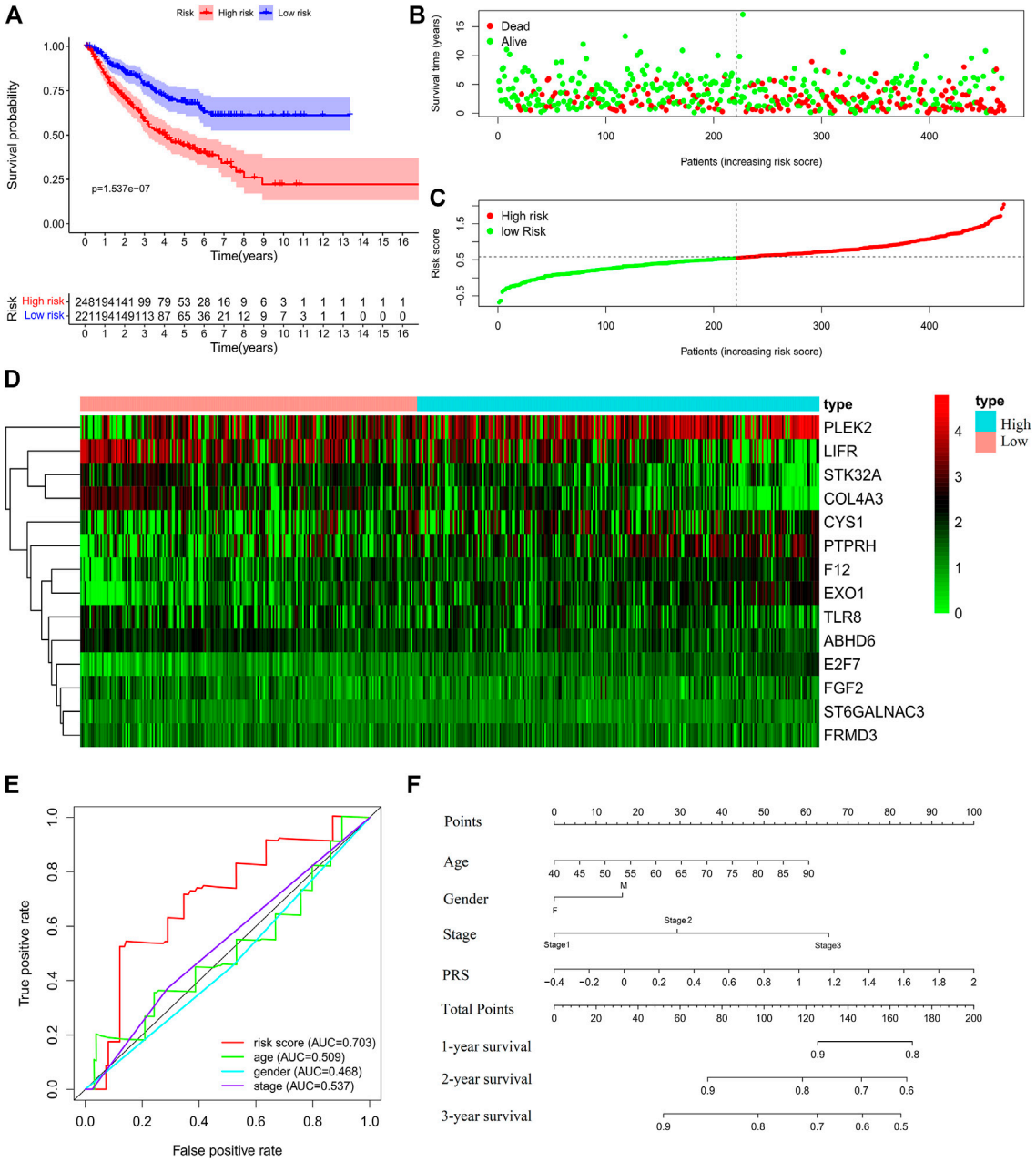


FIGURE 4 | Validation of the prognostic risk score (PRS) model in the test set. **(A)** Comparisons of overall survival (OS) between the high- and low-risk groups. **(B)** Distribution of survival time among patients with different outcomes. **(C)** Distribution of increasing risk scores in high- and low-risk groups. **(D)** Heatmap of the 14-gene expression profiles in high- and low-risk groups. **(E)** ROCs of models with sex, stage, and PRS as variables in the test set s, respectively. **(F)** Nomogram to predict the 1-, 2-, and 3-years survival rates of patients in the test set by using sex, stage, and risk score as variables, respectively.

predicted to yield HLA-binding neopeptides (Predicted NeoAgs) and the ratios of observed versus expected binders per non-silent mutation (Observed/Expected NeoAgs) (Charoentong et al., 2017; Hakimi et al., 2016). The Observed/Expected NeoAgs of the low-risk patients was higher than that of the high-risk ones (Wilcoxon test, $p < 0.05$). Patients in the low-risk group may have more effective neoantigens to promote immunity against tumor and obtain more benefits from immunotherapy.

Then, we also analyzed mutation rates of genes between PRS subgroups in the training set. The integral mutation rate of the high- and low risk group was 92.61 and 78.51%, respectively. Then, mutation frequencies of 42 driver genes associated with LUAD were calculated. Mutation ratios of the driver genes *TP53*, *LRP1B*, *CLIP1*, *EZH2*, *LRIG3*, *PIK3CA*, *RBM10,* and *KRAS* in the high-risk group were higher than those in the low-risk group (Chi-square test, $p < 0.05$) (**Supplementary**

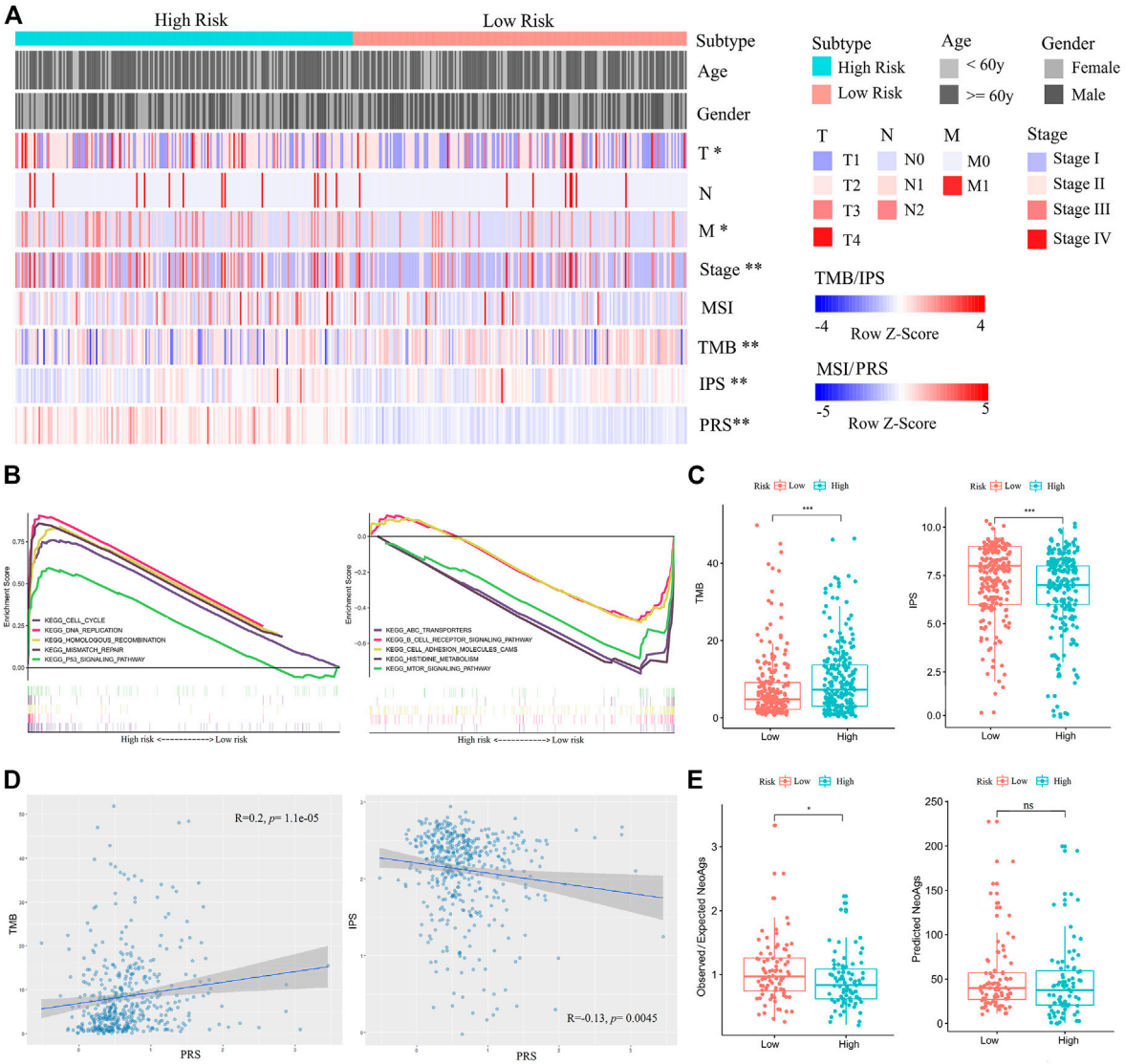


FIGURE 5 | Comparisons of clinical, immune, and molecular features between high- and low-risk patients in the training set. **(A)** Associations of three immunophenotypes with 10 variables. Fisher's exact test was used for categorical variables: age, sex, pathological stage, tumor size (T), lymph node metastasis (N), and distant metastasis (M). Wilcoxon test was used for continuous variables: MSI, TMB, IPS, and PRS. **(B)** GSEA (C2: curated gene sets, CP: KEGG) showed that the five top pathways upregulated in the high-risk group were cell cycle, DNA replication, homologous recombination, mismatch repair, and p53 signal pathway (left). The five top pathways upregulated in the low-risk group were ABC transporters, B cell receptor signaling pathway, cell adhesion molecules (CAMs), histidine metabolism, and mTOR signaling pathway (right). **(C)** Comparisons of TMBs and IPSs between high- and low-risk patients. **(D)** The positive correlation between TMB and PRS is shown on the left. The negative correlation between IPS and PRS is shown on the right. **(E)** Comparisons of Observed/Expected NeoAgs and Predicted NeoAgs between high- and low-risk patients. The symbol "ns" represents there is no significant difference between the two groups.

Table S5). These results provide new insights into mutation biomarkers and immunotherapeutic targets. To understand the effects of these mutations on LUAD development, we conducted a cluster analysis and identified 11 signatures of *de novo* mutation (S1-S11) (**Supplementary Figures S7A,B**). Among them, S2, S3, S4, and S5 were similar to the curated signatures in COSMIC (**Supplementary Figure S7C**; **Supplementary Table S6**). S2 was related to DNA mismatch repair deficiency (dMMR), S3 was related to APOBEC (apolipoprotein B mRNA editing enzyme catalytic polypeptide-like), S4 was related to age, and S5 was related to tobacco mutagens. The contributions of S2, S3, S4, S5, S10, and S11 in the high-risk group were higher than those in the low-risk group (Cochran-Armitage trend test, $p < 0.05$) (**Supplementary Figure S7D**).

The Predictive Potential of the PRS Model for Immunotherapy Benefits

In the subsequent analysis, we examined the ability of the PRS model to predict the response to immunotherapy in Asian

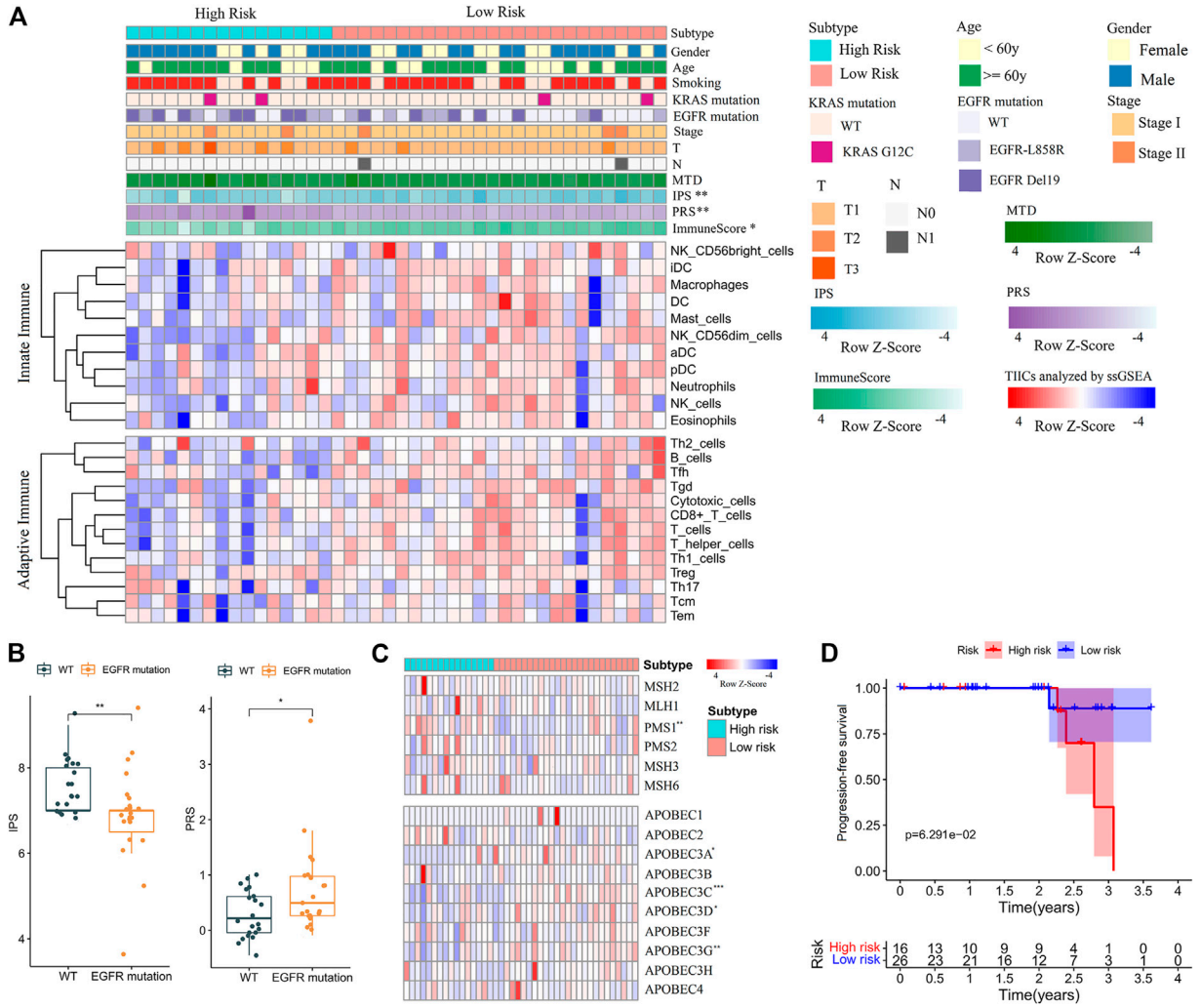


FIGURE 6 | Feature analysis and PFS comparisons of NJDT patients in different PRS subgroups. **(A)** Comparison of the variables between high- and low-risk patients. Fisher's exact test was used for categorical variables: sex, age, smoking, KRAS mutation, EGFR mutations, tumor size (T), lymph node metastasis (N). Wilcoxon test was used for continuous variables: maximum tumor diameter (MTD), IPS, PRS, Immune Score, and TIICs. **(B)** Comparisons of IPSs and PRSs between EGFR mutation and wild-type patients. **(C)** Heatmaps of MMR and APOBEC genes in high- and low-risk patients. **(D)** Comparisons of PFS between high- and low-risk patients.

patients. A total of 27 NSCLC patients (GSE135222) who received anti-PD-1/PD-L1 immunotherapy were selected. On the basis of the cut-off values in this study, the patients were divided into high-risk and low-risk groups. Survival analysis showed that low-risk patients had better OS than high-risk ones (Log-rank test, $p < 0.05$, **Supplementary Figure S8A**). Another dataset of NSCLC patients (GSE126044) who were also treated by anti-PD-1/PD-L1 immunotherapy showed that the PRSs of those who responded to immunotherapy were significantly lower than those who did not respond (Wilcoxon test, $p < 0.05$, **Supplementary Figure S8B**). The results of these two datasets indicated that the PRS model was also a good predictor for the efficacy of immunotherapy.

The Prognostic Potential of the PRS Model for NJDT Patients

Finally, the clinical (**Supplementary Table S7**) and mRNA sequencing data (**Supplementary Table S8**) of NJDT patients were analyzed by the PRS model and IPS algorithm (**Supplementary Material S1**). The PRSs of tumor samples were significantly higher than those of normal samples (Wilcoxon test, $p < 0.01$) (**Supplementary Figure S9A**); Among the tumor samples, 16 and 26 cases were categorized into the high- and low-risk groups, respectively. When the clinicopathological features were compared, the immune scores (calculated by ESTIMATE) and IPSs of the high-risk group were lower (**Figure 6A**), and *EGFR* mutations of high-risk patients were more frequent. In addition, patients with *EGFR* mutations

had higher PRSs and lower IPSs than wild-type (WT) patients (Wilcoxon test, $p < 0.05$) (**Figure 6B**). But *KRAS* mutations did not showed the similar phenomenon. These results may suggest that the immunity state against tumor of WT patients was superior to that of mutant patients. Considering the mutation results in the training set, we compared the expression of MMR and APOBEC proteins. *PMS1* was upregulated in the high-risk group, while *APOBEC3A*, *C*, *D* and *G* were upregulated in the low-risk group (Wilcoxon test, $p < 0.05$) (**Figure 6C**, **Supplementary Table S9**). Most checkpoints of patients in the low-risk group were higher than those in the high-risk group (**Supplementary Figure S9B**), and the immune-related pathways in the low-risk group were also upregulated (**Supplementary Figure S9C**). Furthermore, although the difference between groups was not statistically significant, the low-risk patients had better PFS than high-risk patients (**Figure 6D**).

DISCUSSION

TIME has been frequently reported to play an important role in the occurrence and development of LUAD. The evaluation of the dynamic changes in TIME to determine mechanisms underlying tumorigenesis and potential therapeutic targets is of great significance. In our research, the infiltration levels of TIICs in TIME were strongly correlated with patient outcomes. The patients in the Immunity_H group, who had more TIICs and stronger anti-tumor immunoactivity, also had better prognoses. The patients in the Immunity_L group, of whom nearly all mature TIICs were at low levels, did not trigger adequate immune response against LUAD and had worse prognoses. However, it was interesting that the patients in the Immunity_M group, with the worst OS, showed a deficiency of innate immune cells (DC, macrophages and NK cells), but a high level of CD8+T cells. Numerous studies have indicated that NK cells have a definite antitumor effect in the lung cancer (Bhome et al., 2015) and dendritic cells (DCs), as powerful antigen-presenting cells, play important roles in inducing the immune response of CD8+T cells (Hegde et al., 2020; Maier et al., 2020). However, studies on the role of CD8+T cells in TIME have yielded different conclusions. Some reports suggest that the number of these cells is positively correlated with the treatment response and survival of NSCLC patients (Donnem et al., 2015; Rashed et al., 2017). In contrast, recent studies have shown that only approximately 10% of tumor-infiltrating T cells in the TIME of NSCLC patients can recognize surrounding tumor cells, while the rest are "bystander T cells", which lack response to tumor antigens and are involved in tumor immune escape and progression (Scheper et al., 2019). In addition, Immunity_M showed more Treg cells and higher expressions of CTLA4 and PD-1. The CTLA4 expressed on Treg cells can mediate the downregulation of costimulatory molecules of DCs, reduce DC activation, and enhance the immunosuppressive activity of Treg cells (Chen et al., 2017). The dysfunction of T cells is positively correlated with a high expression of PD-1 (Thommen et al., 2015). These two factors of abnormal T cells and checkpoints might jointly contribute to the worst prognoses of patients in the Immunity_M group, whereas due to the increased CTLA4 and PD-1, they are likely to get more benefits from immunotherapy.

An interesting phenomenon was revealed by analysis of mutation characteristics between high- and low-risk patients in the training set. Somatic mutations in tumor may produce targeted neoantigens recognized by major histocompatibility complex (MHC) (Schumacher and Schreiber, 2015). TMB, as an indicator of somatic mutation in cancer, was lower in the low-risk group, but the predictive amount and proportion of neoantigens were higher. It suggested that although high-risk patients showed more mutations, they did not produce more neoantigens to induce adequate immune response against tumor. The phenomenon may be related to the unsatisfactory infiltration of TIICs and indicate less benefits from immunotherapy. Subsequent analyses on other East Asian patients who received anti-PD-1/PD-L1 immunotherapy repeated the consequence of less benefits from treatment in their high-risk groups. On the other hand, when the PRS model was used to evaluate early-stage LUAD patients, it could not only predict better TIME, but also demonstrate potentials for early LUAD diagnose. Furthermore, *EGFR* mutations were more frequent in high-risk patients, and it suggests that *EGFR* mutation may be associated with immunosuppression in NSCLC (Dong et al., 2017; Gainor et al., 2016). Although we did not detect more details of somatic mutations in these patients, abnormal MMR and APOBEC expressions suggest that there are more mutation differences between high- and low-risk patients, and these differences are also expected to be biomarkers for early diagnosis and prognosis prediction of LUAD.

In previous studies, several immune-related prognostic models of NSCLC based on TCGA datasets have been reported (Chen et al., 2021; Liu et al., 2020; Luo et al., 2020; Song et al., 2020). Some researchers divided TCGA data into the training and test sets, and obtained a prognostic model based on immune genes. The areas under the curves (AUCs) of the model were 0.74 for 3-years OS and 0.70 for 5-years OS in the training set. In the test set, they were 0.676 and 0.523, respectively (Yi et al., 2021a). In our research, AUCs of the PRS model were 0.706 for 3-years OS and 0.710 for 5-years OS in the training set. In the test set, they were 0.636 and 0.631 (**Supplementary Figures S6C,D**), respectively. The PRS model were performed by the external validation of GEO datasets and had more extensive and stable accuracy and sensitivity in prognostic prediction for LUAD patients. The 14 IDEGs of PRS model are involved in immune cell receptors, inflammatory factors, biological enzymes, gene transcription and blood coagulation. Some of these genes are deeply related to immune environment and immunotherapy. *F12* (coagulation factor XII) regulates a range of innate immune cells (Barbasz and Kozik, 2009; Vorlova et al., 2017), and promotes the differentiation of naive Th cells into TH17 cells (Gobel et al., 2016). *LIFR* (leukemia inhibitory factor receptor subunit alpha) mediates interleukin-6 signaling and is involved in immune regulation (Wang et al., 2020). *TLR8* (Toll-like receptor 8) initiates juvenile T cells, promotes the secretion of various cytokines by DCs and is involved in the regulation of tumor immune microenvironment (Tran et al., 2019), while *FGF2* (fibroblast growth factor 2) is

involved in the Wnt/β-catenin, TGF-β and PI3K/Akt pathways to affect the development of LUAD (Dai et al., 2019). TGF-β is an important signaling in promoting cancer metastasis, impairing the functions of immune cells and facilitating immune evasion (Batlle and Massague, 2019). Recently, the anti- TGF-β/PD-L1 bispecific antibody YM101 has reported to effectively overcome treatment resistance and exhibit a superior antitumor activity of non-inflamed tumors (Yi et al., 2021b). The antibody can promote the formation of "hot tumor" in increasing adaptive TIICs and DCs, regulating the ratio of M1/M2, and enhancing cytokine production in T cells. (Yi et al., 2021c). Our research also illustrated the balance of innate and adaptive immune cells and the recognization of T cells by surrounding tumor cells are the keys to improving prognosis and immunotherapy of LUAD. The PRS model may be applied to the predict suitability and efficacy of antibody YM101.

In conclusion, our study provided the risk model, which showed the good predictive ability for the prognosis and therapeutic benefits of LUAD. The exploration based on immunotyping revealed more immune characteristics and molecular mechanisms related to prognosis, and laid a foundation for further research on diagnosis, immunotherapy and drug development. Nevertheless, this study had many limitations, we will further improve the applicability of the PRS model for domestic patients, and conduct more biological experiments to verify the functions and pathways of IDEGs. We hope that our research will facilitate the diagnosis, risk stratification, prognostication, and treatment decision-making for LUAD patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: CNGBdb, CNP0002665.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Nanjing Drum Tower Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HG takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation, drafting the article. YL were responsible for evaluating paraffin-embedded samples. HY, MG, and JD collected samples and data. QY take responsibility for full text evaluation and guidance. All authors contributed to the article and approved the submitted version.

FUNDING

Our research was funded by Medical Science and Technology Development Foundation, Nanjing Municipality Health Bureau (YKK21105), Projects of Modern Hospital Management and Development Institute, Nanjing University (NDYG2021007, NDYG2021012), Open Projects of Jiangsu Biobank of Clinical Resources (SBK202006002,SBK202006003) and Jiangsu Biobank of Clinical Resources (BM2015004).

ACKNOWLEDGMENTS

We thank the Department of Pathology, Nanjing Drum Tower Hospital, for evaluating tumor samples from NJDT patients.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.850101/full#supplementary-material

Supplementary Figure S1 | Twenty-four types of TIICs between normal and tumor samples were analyzed in the training and test set. Principal component analysis (PCA) of 24 TIICs between normal and tumor samples in the training set **(A)** and test set **(B)**. The level comparison of 24 TIICs between normal and tumor samples in the training set **(C)** and test set **(D)**

Supplementary Figure S2 | OS comparison and hierarchical cluster analysis of TIICs of different groups in the training set. **(A)** OS comparisons of patients in the low- and high-level groups for immune score (left), stromal score (middle), and tumor purity (right). **(B)** Hierarchical cluster analysis of patients in the training set based on the TIICs. Immunity_H group in green, Immnuty_M group in red and Immnuty_L group in blue were shown.

Supplementary Figure S3 | Comparison of the levels of 24 TIICs of tumor samples in each immunophenotype. **(A)** The levels of the innate immune cells were compared among three different immunophenotypes. **(B)** The levels of the adaptive immune cells were compared among three different immunophenotypes. The symbol "ns" represents there is no significant difference between the two groups.

Supplementary Figure S4 | Comparison of the levels of 21 immune cells and expressions of checkpoints and HLA-related genes in each immunophenotype. **(A)** The levels of 21 immune cells analyzed by CIBERSORT were compared between Immunity_H and Immunity_M. **(B)** The levels of 21 immune cells analyzed by CIBERSORT were compared between Immunity_H and Immunity_L. **(C)** The expressions of checkpoints and HLA-related genes were compared between Immunity_H and Immunity_M. (B) The expressions of checkpoints and HLA-related genes were compared between Immunity_H and Immunity_L. The symbol "ns" represents there is no significant difference between the two groups.

Supplementary Figure S5 | Comparisons of clinical and molecular features and immunophenotypic differentially expressed genes (IDEGs) among the three immunophenotypes in the training set. **(A)** Association of the three immunophenotypes with 10 variables. Fisher's exact test was used for categorical variables: age, gender, pathological stage, tumor size (T), lymph node metastasis (N) and distant metastasis (M); Kruskal-Wallis test was used for continuous variables: MSI, TMB, IPS and Risk Score. **(B)** GSEA (C2: curated gene sets, CP: KEGG) revealed the pathways that were significantly upregulated in various immunophenotypes, compared with normal samples. **(C)** GSEA (C2: curated gene sets, CP: KEGG) revealed the pathways that were significantly downregulated in various immunophenotypes, compared with normal samples. **(D)** Comparison of significant enrichment pathways in tumor samples between Immunity_H and Immunity_M **(E)** Comparison of significant enrichment pathways in tumor samples between Immunity_H and Immunity_L **(F)** 421 immunophenotypic differentially expressed genes (IDEGs) were obtained by intersection of DEGs_NT, DEGs_HM and DEGs_HL. **(G)** Gene Ontology (GO) analysis of IDEGs

Supplementary Figure S6 | Construction of prognostic risk score (PRS) model analyzed by LASSO regression and comparison of DFS, DSS, and PFS between the high- and low-risk patients in the training set. **(A)** Tuning parameter (λ) and deviance in the LASSO regression (left). The partial likelihood deviance was plotted versus log (λ). The dotted vertical lines were drawn at the optimal values by using the minimum and 1-SE criteria. Fourteen features with the smallest binomial deviance were selected. LASSO coefficient profiles of texture features (right). Each line represented a variable with the regression coefficient on the vertical axis and the logarithm of λ on the abscissa. A 10-fold cross-validation was used in the log (λ) sequence to select 14 variables with non-zero coefficients. **(B)** Comparisons of disease-free survival (DFS) (left), disease-specific survival (DSS) (middle), and progression-free survival (PFS) (right) of high- and low-risk patients in the training set. **(C)** Time-dependent receiver operating characteristic curves of 3- and 5-year survival in the training set (left) and test set (right).

Supplementary Figure S7 | Identification and comparison of de novo mutational signatures in the training set. **(A)** Upper image, residual sum of square (RSS) of the signature number selection. Lower image, percentage of variance explained in the signature number selection. **(B)** Cosine similarity between 30 cosmic signatures (horizontal axis) and 11 de novo signatures (vertical axis) in the training set. **(C)** Contributions of point mutations of de novo mutation signatures (S1–S11) in the training set. **(D)** Comparison of the contributions of de novo mutation signatures (S1–S11) between the high- and low-risk groups

Supplementary Figure S8 | Prediction of the prognostic risk score (PRS) model for immunotherapy benefits of patients. **(A)** Comparisons of overall survival (OS) between the high- and low-risk patients in GSE135222. **(B)** Comparisons of PRSs of the responders and non-responders treated by immunotherapy in GSE126044.

Supplementary Figure S9 | Analysis of PRSs, checkpoints, HLA-related genes, and enrichment pathways of NJDT patients. **(A)** Comparison of PRSs between normal and tumor samples of NJDT patients. **(B)** The expressions of checkpoints and HLA-related genes were compared between high- and low-risk patients. **(C)** GSEA (C2: curated gene sets, CP: KEGG) showed that the five top pathways (antigen processing and presentation, cell adhesion molecules [CAMs], chemokine signaling pathway, cytokine receptor interaction, and natural killer cell-mediated cytotoxicity) were unregulated in the low-risk group. The symbol "ns" represents there is no significant difference between the two groups.

REFERENCES

Barbasz, A., and Kozik, A. (2009). The Assembly and Activation of Kinin-Forming Systems on the Surface of Human U-937 Macrophage-like Cells. *Biol. Chem.* 390 (3), 269–275. doi:10.1515/BC.2009.032

Batlle, E., and Massagué, J. (2019). Transforming Growth Factor-β Signaling in Immunity and Cancer. *Immunity* 50 (4), 924–940. doi:10.1016/j.immuni.2019.03.024

Bhome, R., Bullock, M. D., Al Saihati, H. A., Goh, R. W., Primrose, J. N., Sayan, A. E., et al. (2015). A Top-Down View of the Tumor Microenvironment: Structure, Cells and Signaling. *Front. Cell Dev. Biol.* 3, 33. doi:10.3389/fcell.2015.00033

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

Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., et al. (2018). Understanding the Tumor Immune Microenvironment (TIME) for Effective Therapy. *Nat. Med.* 24 (5), 541–550. doi:10.1038/s41591-018-0014-x

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. Clinicians* 68 (6), 394–424. doi:10.3322/caac.21492

Bussard, K. M., Mutkus, L., Stumpf, K., Gomez-Manzano, C., and Marini, F. C. (2016). Tumor-associated Stromal Cells as Key Contributors to the Tumor Microenvironment. *Breast Cancer Res.* 18 (1), 84. doi:10.1186/s13058-016-0740-2

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, X., Du, Y., Hu, Q., and Huang, Z. (2017). Tumor-derived CD4+CD25+regulatory T Cells Inhibit Dendritic Cells Function by CTLA-4. *Pathol. - Res. Pract.* 213 (3), 245–249. doi:10.1016/j.prp.2016.12.008

Chen, Z., Fillmore, C. M., Hammerman, P. S., Kim, C. F., and Wong, K.-K. (2014). Non-small-cell Lung Cancers: A Heterogeneous Set of Diseases. *Nat. Rev. Cancer* 14 (8), 535–546. doi:10.1038/nrc3775

Chen, Z., Huang, Y., Hu, Z., Zhao, M., Li, M., Bi, G., et al. (2021). Landscape and Dynamics of Single Tumor and Immune Cells in Early and Advanced-stage Lung Adenocarcinoma. *Clin. Translational Med.* 11 (3), e350. doi:10.1002/ctm2.350

Cui, Y., Chen, H., Xi, R., Cui, H., Zhao, Y., Xu, E., et al. (2020). Whole-genome Sequencing of 508 Patients Identifies Key Molecular Features Associated with Poor Prognosis in Esophageal Squamous Cell Carcinoma. *Cell Res* 30 (10), 902–913. doi:10.1038/s41422-020-0333-6

Dai, S., Zhou, Z., Chen, Z., Xu, G., and Chen, Y. (2019). Fibroblast Growth Factor Receptors (FGFRs): Structures and Small Molecule Inhibitors. *Cells* 8 (6), 614. doi:10.3390/cells8060614

Dong, Z.-Y., Zhang, J.-T., Liu, S.-Y., Su, J., Zhang, C., Xie, Z., et al. (2017). EGFR Mutation Correlates with Uninflamed Phenotype and Weak Immunogenicity, Causing Impaired Response to PD-1 Blockade in Non-small Cell Lung Cancer. *Oncoimmunology* 6 (11), e1356145. doi:10.1080/2162402X.2017.1356145

Donnem, T., Hald, S. M., Paulsen, E.-E., Richardsen, E., Al-Saad, S., Kilvaer, T. K., et al. (2015). Stromal CD8+ T-Cell Density-A Promising Supplement to TNM Staging in Non-small Cell Lung Cancer. *Clin. Cancer Res.* 21 (11), 2635–2643. doi:10.1158/1078-0432.CCR-14-1905

Gainor, J. F., Shaw, A. T., Sequist, L. V., Fu, X., Azzoli, C. G., Piotrowska, Z., et al. (2016). EGFR Mutations and ALK Rearrangements Are Associated with Low Response Rates to PD-1 Pathway Blockade in Non-small Cell Lung Cancer: A Retrospective Analysis. *Clin. Cancer Res.* 22 (18), 4585–4593. doi:10.1158/1078-0432.CCR-15-3101

Gao, J., Kwan, P. W., and Shi, D. (2010). Sparse Kernel Learning with LASSO and Bayesian Inference Algorithm. *Neural Networks* 23 (2), 257–264. doi:10.1016/j.neunet.2009.07.001

Gehring, J. S., Fischer, B., Lawrence, M., and Huber, W. (2015). SomaticSignatures: Inferring Mutational Signatures from Single-Nucleotide Variants: Fig. 1. *Bioinformatics* 31 (22), btv408–3675. doi:10.1093/bioinformatics/btv408

Göbel, K., Pankratz, S., Asaridou, C.-M., Herrmann, A. M., Bittner, S., Merker, M., et al. (2016). Blood Coagulation Factor XII Drives Adaptive Immunity during Neuroinflammation via CD87-Mediated Modulation of Dendritic Cells. *Nat. Commun.* 7, 11626. doi:10.1038/ncomms11626

Goldstraw, P., Chansky, K., Crowley, J., Rami-Porta, R., Asamura, H., Eberhardt, W. E., et al. (2016). The IASLC Lung Cancer Staging Project: Proposals for Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition of the TNM Classification for Lung Cancer. *J. Thorac. Oncol.* 11 (1), 39–51. doi:10.1016/j.jtho.2015.09.009

Hakimi, A. A., Reznik, E., Lee, C.-H., Creighton, C. J., Brannon, A. R., Luna, A., et al. (2016). An Integrated Metabolic Atlas of clear Cell Renal Cell Carcinoma. *Cancer Cell* 29 (1), 104–116. doi:10.1016/j.ccell.2015.12.004

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics* 14, 7. doi:10.1186/1471-2105-14-7

Hegde, S., Krisnawan, V. E., Herzog, B. H., Zuo, C., Breden, M. A., Knolhoff, B. L., et al. (2020). Dendritic Cell Paucity Leads to Dysfunctional Immune Surveillance in Pancreatic Cancer. *Cancer Cell* 37 (3), 289–307. doi:10.1016/j.ccell.2020.02.008

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., et al. (2003). Exploration, Normalization, and Summaries of High Density Oligonucleotide Array Probe Level Data. *Biostatistics* 4 (2), 249–264. doi:10.1093/biostatistics/4.2.249

Junttila, M. R., and de Sauvage, F. J. (2013). Influence of Tumour Microenvironment Heterogeneity on Therapeutic Response. *Nature* 501 (7467), 346–354. doi:10.1038/nature12626

Liu, Z., Wan, Y., Qiu, Y., Qi, X., Yang, M., Huang, J., et al. (2020). Development and Validation of a Novel Immune-Related Prognostic Model in Lung Squamous Cell Carcinoma. *Int. J. Med. Sci.* 17 (10), 1393–1405. doi:10.7150/ijms.47301

Luo, C., Lei, M., Zhang, Y., Zhang, Q., Li, L., Lian, J., et al. (2020). Systematic Construction and Validation of an Immune Prognostic Model for Lung Adenocarcinoma. *J. Cel. Mol. Med.* 24 (2), 1233–1244. doi:10.1111/jcmm.14719

Maier, B., Leader, A. M., Chen, S. T., Tung, N., Chang, C., LeBerichel, J., et al. (2020). A Conserved Dendritic-Cell Regulatory Program Limits Antitumour Immunity. *Nature* 580 (7802), 257–262. doi:10.1038/s41586-020-2134-y

Manthey, H. D., Cochain, C., Busch, M., Chaudhari, S. M., Stegner, D., Yepes, M., et al. (2017). Coagulation Factor XII Induces Pro-inflammatory Cytokine Responses in Macrophages and Promotes Atherosclerosis in Mice. *Thromb. Haemost.* 117 (1), 176–187. doi:10.1160/TH16-06-0466

Mascaux, C., Angelova, M., Vasaturo, A., Beane, J., Hijazi, K., Anthoine, G., et al. (2019). Immune Evasion before Tumour Invasion in Early Lung Squamous Carcinogenesis. *Nature* 571 (7766), 570–575. doi:10.1038/s41586-019-1330-0

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

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

Rashed, H. E., Abdelrahman, A. E., Abdelgawad, M., Balata, S., and Shabrawy, M. E. (2017). Prognostic Significance of Programmed Cell Death Ligand 1 (PD-L1), CD8+ Tumor-Infiltrating Lymphocytes and P53 in Non-small Cell Lung Cancer: An Immunohistochemical Study. *Tjpath* 1 (1), 211–222. doi:10.5146/tjpath.2017.01398

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., et al. (2019). Pathway Enrichment Analysis and Visualization of Omics Data Using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat. Protoc.* 14 (2), 482–517. doi:10.1038/s41596-018-0103-9

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

Scheper, W., Kelderman, S., Fanchi, L. F., Linnemann, C., Bendle, G., de Rooij, M. A. J., et al. (2019). Low and Variable Tumor Reactivity of the Intratumoral TCR Repertoire in Human Cancers. *Nat. Med.* 25 (1), 89–94. doi:10.1038/s41591-018-0266-5

Schumacher, T. N., and Schreiber, R. D. (2015). Neoantigens in Cancer Immunotherapy. *Science* 348 (6230), 69–74. doi:10.1126/science.aaa4971

Schürch, C. M., Bhate, S. S., Barlow, G. L., Phillips, D. J., Noti, L., Zlobec, I., et al. (2020). Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front. *Cell* 182 (5), 1341–1359. doi:10.1016/j.cell.2020.07.005

Song, C., Guo, Z., Yu, D., Wang, Y., Wang, Q., Dong, Z., et al. (2020). A Prognostic Nomogram Combining Immune-Related Gene Signature and Clinical Factors Predicts Survival in Patients with Lung Adenocarcinoma. *Front. Oncol.* 10, 1300. doi:10.3389/fonc.2020.01300

Thommen, D. S., Schreiner, J., Müller, P., Herzig, P., Roller, A., Belousov, A., et al. (2015). Progression of Lung Cancer Is Associated with Increased Dysfunction of T Cells Defined by Coexpression of Multiple Inhibitory Receptors. *Cancer Immunol. Res.* 3 (12), 1344–1355. doi:10.1158/2326-6066.CIR-15-0097

Tran, T. H., Tran, T. T. P., Truong, D. H., Nguyen, H. T., Pham, T. T., Yong, C. S., et al. (2019). Toll-like Receptor-Targeted Particles: A Paradigm to Manipulate the Tumor Microenvironment for Cancer Immunotherapy. *Acta Biomater.* 94, 82–96. doi:10.1016/j.actbio.2019.05.043

Wang, Q., Wu, J., Huang, H., Jiang, Y., Huang, Y., Fang, H., et al. (2020). LncRNA LIFR-AS1 Suppresses Invasion and Metastasis of Non-small Cell Lung Cancer via the miR-942-5p/ZNF471 axis. *Cancer Cell Int* 20, 180. doi:10.1186/s12935-020-01228-5

Wu, C., Li, M., Meng, H., Liu, Y., Niu, W., Zhou, Y., et al. (2019). Analysis of Status and Countermeasures of Cancer Incidence and Mortality in China. *Sci. China Life Sci.* 62 (5), 640–647. doi:10.1007/s11427-018-9461-5

Yi, M., Li, A., Zhou, L., Chu, Q., Luo, S., and Wu, K. (2021a). Immune Signature-Based Risk Stratification and Prediction of Immune Checkpoint Inhibitor's Efficacy for Lung Adenocarcinoma. *Cancer Immunol. Immunother.* 70 (6), 1705–1719. doi:10.1007/s00262-020-02817-z

Yi, M., Niu, M., Zhang, J., Li, S., Zhu, S., Yan, Y., et al. (2021b). Combine and Conquer: Manganese Synergizing Anti-TGF-β/pd-L1 Bispecific Antibody YM101 to Overcome Immunotherapy Resistance in Non-inflamed Cancers. *J. Hematol. Oncol.* 14 (1), 146. doi:10.1186/s13045-021-01155-6

Yi, M., Zhang, J., Li, A., Niu, M., Yan, Y., Jiao, Y., et al. (2021c). The Construction, Expression, and Enhanced Anti-tumor Activity of YM101: a Bispecific Antibody Simultaneously Targeting TGF-β and PD-L1. *J. Hematol. Oncol.* 14 (1), 27. doi:10.1186/s13045-021-01045-x

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

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 Gao, Liu, Hu, Ge, Ding 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.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 27 April 2022
doi: 10.3389/fgene.2022.818403

Apoptosis-Related Signature Predicts Prognosis and Immune Microenvironment Infiltration in Lung Adenocarcinoma

Xiaoli Zou[1†], *Rong He*[2†], *Zhenzhen Zhang*[1] *and Yulan Yan*[1*]

[1]Departments of Respiratory Medicine, Affiliated People's Hospital of Jiangsu University, Zhenjiang, China, [2]Departments of Cancer Institute, Affiliated People's Hospital of Jiangsu University, Zhenjiang, China

Lung adenocarcinoma (LUAD), a malignancy with high incidence and mortality rates worldwide, contains multiple genomic and epigenomic abnormalities. And the useful tumor markers associated with these abnormalities need further investigation. Whereas apoptosis is a form of programmed cell death, the expression of apoptosis-related genes in LUAD and its relationship with prognosis is unclear. In the present study, we identified 64 differentially expressed apoptosis-related genes (DEARGs) that were differentially expressed between LUAD tissue and normal lung tissue. Based on these DEARGs, all LUAD cases were classified into two subtypes using The Cancer Genome Atlas (TCGA) cohort to assess the prognostic value of apoptosis-related genes for survival. An 11-gene signature was established by applying the Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression method to construct a multigene prediction model and classify all LUAD patients in the TCGA cohort into high or low AS-score groups. Patients in the low AS-score group had significantly higher survival and prognosis than those in the high AS-score group. Taking the median risk score of the AS-score, LUAD patients in the GSE68465 cohort were divided into two risk groups, low and high. The overall survival (OS) time was longer in the low AS-score group. Combined with clinical characteristics, the AS-score was an independent predictor of LUAD patients. Gene ontology (GO) and Kyoto Encylopedia of Genes and Genomes (KEGG) analyses showed that the differential genes between the two groups were mainly enriched in cellular immunity. Further analysis revealed higher immune checkpoint protein expression and higher tumor mutational burden (TMB) in the high AS-score group, suggesting better efficacy of immunotherapy in the high AS-score group than the low AS-score group. And the high AS-score group was better in chemotherapy and targeted therapy efficiency. In conclusion, the AS-score constructed based on apoptosis-related genes can predict the

OPEN ACCESS

Edited by:
William C. Cho,
QEH, Hong Kong SAR, China

Reviewed by:
Xinwei Han,
Zhengzhou University, China
Liu Le Ping,
Central South University, China

***Correspondence:**
Yulan Yan
ylyan2005@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: *19 November 2021*
Accepted: *28 March 2022*
Published: *27 April 2022*

Citation:
Zou X, He R, Zhang Z and Yan Y (2022) Apoptosis-Related Signature Predicts Prognosis and Immune Microenvironment Infiltration in Lung Adenocarcinoma. Front. Genet. 13:818403. doi: 10.3389/fgene.2022.818403

Abbreviations: DEARGs, differentially expressed apoptosis-related gene; DEGs, differentially expressed genes; GDSC, genomics of drug sensitivity in cancer; GO, gene ontology; IC50, half-maximal inhibitory concentration; KEGG, kyoto encyclopedia of genes and genomes; LASSO, least absolute shrinkage and selection operator; LC, lung cancer; LUAD, lung adenocarcinoma; NSCLC, non-small cell lung cancer; PCA, principal component analysis; ROC, receiver operating characteristic curve; SCLC, small cell lung cancer; ss GSEA, single sample gene set enrichment analysis; STRING, search tool for the retrieval of interacting genes; TCGA, the cancer genome atlas; TMB, tumor mutational burden.

prognosis of LUAD patients and provide some guidance for the antitumor treatment of LUAD patients.

Keywords: gene signature, lung adenocarcinoma, apoptosis, tumormicroenvironment, immunotherapy

INTRODUCTION

The new Global Cancer Statistics 2020 showed that cancer incidence in China is the highest globally and that lung cancer (LC) remains the second most prevalent malignancy with high mortality (Sung et al., 2021). According to the pathological type of LC, it can be divided into small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC), of which non-small cell lung cancer includes LUAD and lung squamous carcinoma (LUSC). Meanwhile, LUAD accounts for 85% of non-small cell lung cancer and 40% of all types and is the most common type of LC (Denisenko et al., 2018; Liu L.-P. et al., 2021). With no symptoms in the early stages of LC, it is usually detected at an advanced stage that is not amenable to surgical treatment and has a poor prognosis. The current therapeutic modalities for LUAD include chemotherapy, radiotherapy, targeted drug therapy, immunotherapy, and surgery. However, the sensitivity and specificity of treatment are low due to the heterogeneity of the tumor and the complex immune microenvironment of cancer. Although an increasing number of studies have focused on the analysis of the characteristic death features of tumor cells to predict their prognosis (Liu L.-P. et al., 2021). However, the analysis of molecular features of tumor cell apoptosis to predict lung adenocarcinoma prognosis has not been demonstrated. In this study, we analyzed the molecular features related to tumor apoptosis. The study was carried out to compare the survival differences between the two groups and the efficacy of antitumor drugs by constructing an apoptosis-related prediction model for staging lung adenocarcinoma patients. These suggest that it is crucial to improve treatment specificity and establish a specific prognostic model (Hirsch et al., 2017; Greenawalt et al., 2019).

It is known that there are two main types of cell death, one is programmed death that is finely regulated by the cell, and the other is uncontrolled cell necrosis (Xu et al., 2019). In contrast, apoptosis, as a programming mechanism of cell death, is characterized by specific changes in cell structure and the biochemical processes of all enzyme-catalyzed reactions, mainly removing some damaged and potentially harmful cells from the body (Yaacoub et al., 2016; Carneiro and El-Deiry, 2020). The expression of phosphatidylserine in the outer layer of the cell membrane leads to early recognition and phagocytosis of dead cells by macrophages during apoptosis, without releasing pro-inflammatory cellular components (Hengartner, 2001). Apoptosis is characterized by several morphological features, including membrane vesicles, changes in organelle ultrastructure, and loss of membrane integrity, followed by the emergence of phagocytes that consume the apoptotic cells (Kroemer et al., 2005). The BCL-2 family of proteins is the main apoptosis regulator that directly controls the permeability of membranes (Singh et al., 2019). Cytochrome C is released from mitochondria to form apoptotic vesicles. At the same time, the caspases (cysteine, aspartate-specific proteases) family of proteases plays a crucial executive role in apoptosis (Li and Yuan, 2008), activating the executioner caspase 3 to initiate apoptosis (Ledgerwood and Morison, 2009; Xie et al., 2020). Necrosis, on the other hand, is not regulated and is primarily due to external factors that cause collapse and necrosis, releasing large amounts of harmful substances and causing severe damage to the cellular environment (Degterev et al., 2008).

There is increasing evidence that dysregulation of apoptosis signals the development and progression of malignant tumors, which can become resistant to therapeutic agents due to resistance to apoptosis while evading the immune system (Hassan et al., 2014). The majority of drugs currently used in clinical practice achieve their antitumor effects by affecting the apoptotic signaling pathway (Plati et al., 2008; Giménez-Bonafé et al., 2009).

Current evidence suggests that apoptosis not only plays a role in tumorigenesis, cancer metastasis, cancer immunity, and cancer subtypes, but that senescent or lost apoptotic cells are recognized and phagocytosed by macrophages, leading to the release of cytokines that participate in the complex tumor immune microenvironment, which also influences apoptosis (Carneiro and El-Deiry, 2020).

Apoptosis plays an important role in tumor development and antitumor therapy. However, less study of its specific functions and studies in LUAD drug resistance and tumor immune microenvironment are studied. Therefore, we conducted a systematic analysis to explore the prognostic value of apoptosis-related genes in LUAD and investigate the relevance of apoptosis in LUAD to antitumor drugs and the immune microenvironment.

MATERIALS AND METHODS

LUAD Data Sets and Preprocessing

Firstly, we draw a simple schematic diagram of the proposed apoptotic process based on the apoptotic pathway (Created with BioRender.com) (**Figure 1A**). The Cancer Genome Atlas downloaded and opened the LUAD gene expression dataset with complete clinical information, somatic mutation data, and FPKM transcriptome data (TCGA https://portal.gdc.cancer. gov/), excluding samples without survival information, including 494 cases of LUAD and 59 normal tissues. Detailed information on these LUAD patients is presented (**Supplementary Table S1**). In the KEGG pathway database (https://www.genome.jp/kegg/pathway.html), we found 136 apoptosis-related genes according to the apoptosis pathway (map04210). As no complete set of apoptosis-related genes has been reported before, we searched the literature related to apoptosis, performed a comprehensive analysis, and found that

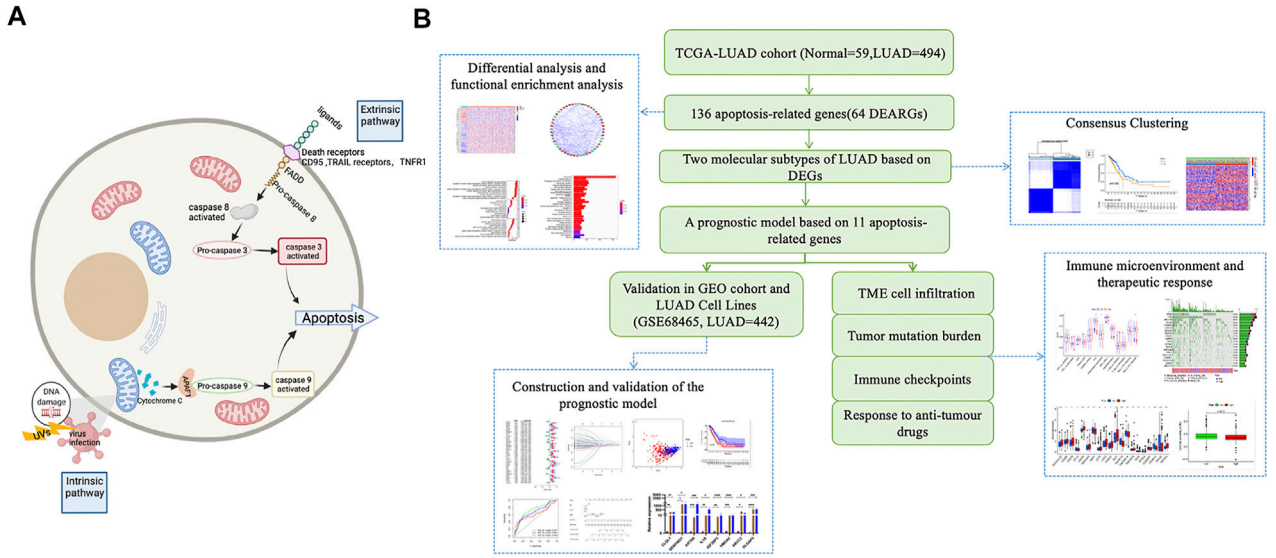


FIGURE 1 | Schematic diagram of the apoptosis pathway and a sketch of the research process. **(A)** Schematic diagram of the apoptosis pathway. **(B)** A sketch of this research process.

all of them could be found in the set of apoptosis-related genes in the kegg database (**Supplementary Table S2**). GSE68465 (N = 442) was downloaded from the gene expression omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) database and was used as the validation set. All data are publicly available online. This study did not require an author to perform experiments on humans or animals. A working diagram showing the overall research process (**Figure 1B**).

Data Processing of Differentially Expressed Genes and Functional Enrichment Analyses

The "limma" package was used to identify apoptosis-related genes that were differentially expressed between LUAD and normal tissues in the TCGA database. The screening criteria were A false discovery rate (FDR) < 0.05, |logFC| > 0.5 (Wu et al., 2021). GO and KEGG analysis was performed using the "cluster profile" package based on these differentially expressed genes. The Search Tool for the Retrieval of Interacting Genes (STRING) database (https://st ring-db. org/) was used to input differential genes for PPI information analysis, and Cytoscape software was used to visualize the PPI network.

Consensus Clustering

Consistent clustering identifies apoptosis-related patterns associated with the expression of apoptosis regulators by the k-means method. The number of clusters and their stability is determined by a consensus clustering algorithm using the "ConsensuClusterPlus" package, repeated 1,000 times to ensure classification stability. The prompt function was used to perform principal component analysis. Heat maps and Kaplan-Meyer (KM) curves were plotted using the R packages "heatmap," "survminer," and "survival."

Construction and Validation of Apoptosis-Related Gene Signature

The consensus clustering algorithm classifies lung cancer patients into two subtypes, and we next use the R package "limma" to identify differentially expressed genes between subtypes (| logFC | > 2 and FDR <0.001). Univariate Cox regression analysis was used to screen for prognosis-related DEGs, and LASSO—Cox analysis was used to narrow down candidate genes further, resulting in a validated predictive model (Liu et al., 2021b). AS-score = Σ (βi × expi) = 1 (where βi is the coefficient index and gene expression, respectively). The median cut-off value was determined using the "survminer" package. The Kaplan-Meier survival curves were used to identify the time to overall survival (OS) that distinguishes different subtypes of patients. The time-dependent subject operating characteristic curves (ROC) were used to assess the validity and accuracy of the model. The GSE68465 cohort was used as an external validation set to assess the value of the predictive model further.

TME Cell Infiltration and the Response to Anti-Tumour Drugs

The ssGSEA was performed by the "gsva" software package to calculate the infiltration score of 16 immune cells and the activity of 13 immune-related pathways. The data of the Genomics Of Drug Sensitivity in Cancer (GDSC) platform were used to predict the sensitivity of LUAD patients to common chemotherapeutic drugs and targeted therapeutic drugs (such as cisplatin, paclitaxel, gefitinib, and erlotinib). The "pRRophetic" R package was used to estimate the half-maximal inhibitory concentration (IC50) (Liu et al., 2021c).

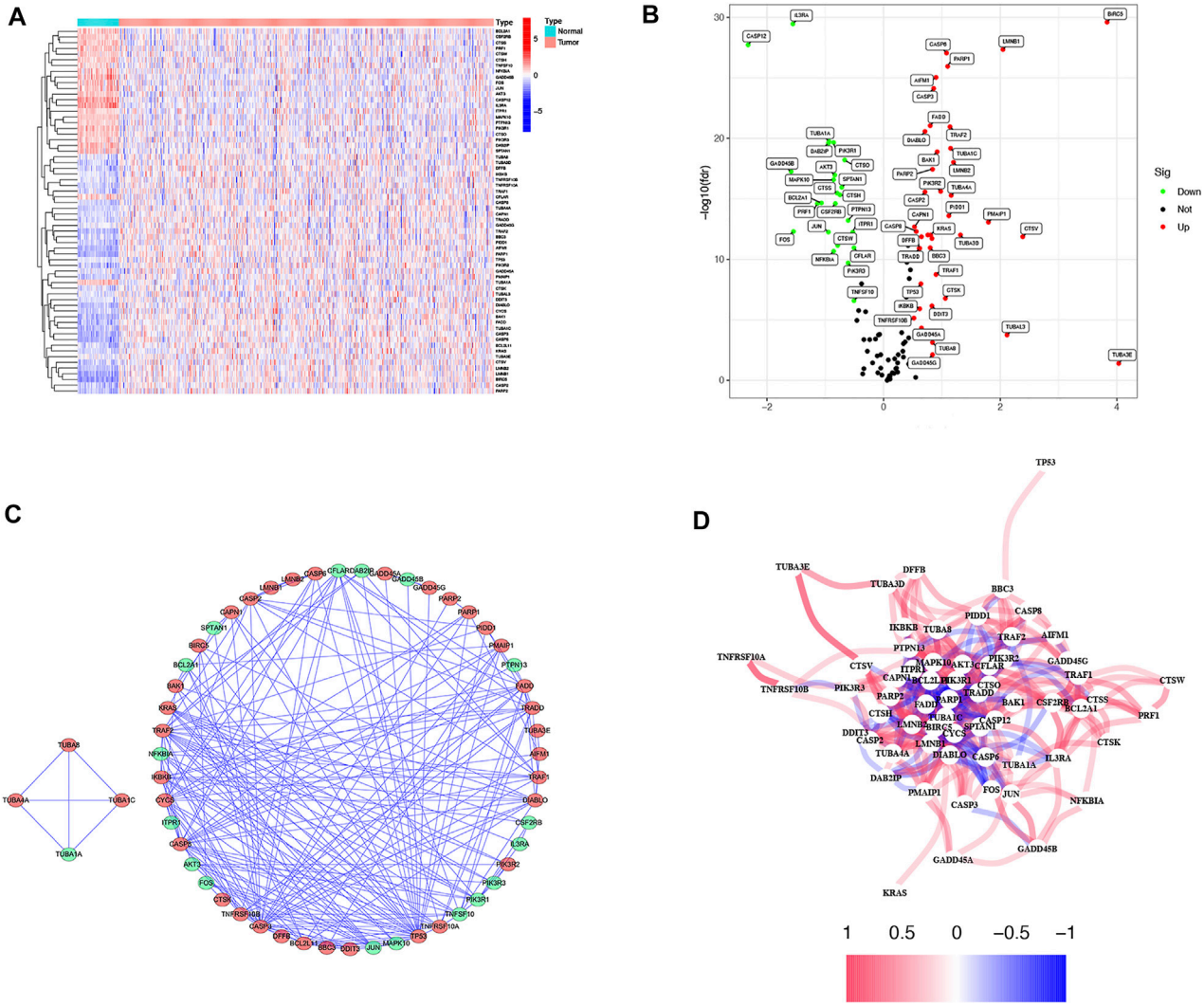


FIGURE 2 | Expressions of the 64 apoptosis-related genes and the interactions among them. **(A)** Heatmap (blue: low expression level; red: high expression level) of the apoptosis-related genes between the normal (N, brilliant blue) and the tumor tissues (T, red). **(B)** The volcano plot shows differential expression of apoptosis genes in LUAD and normal tissue in the TCGA cohort (green: low expression in LUAD; red: high expression in LUAD). **(C)** PPI network shows the apoptosis-related gene interactions (interaction score = 0.7) (green: low expression in LUAD; red: high expression in LUAD). **(D)** The correlation network of the apoptosis-related genes (red line: positive correlation; blue line: negative correlation. The depth of the colors reflects the strength of the relevance).

Cell Culture

The human LUAD cell lines (A549 and PC9) and the normal human lung epithelial cell line (BEAS-2B) used in this study were provided by the Institute of Cell Research, Chinese Academy of Sciences (Shanghai, China). The medium for A549 cells was DMEM medium with 10% fetal bovine serum and 1% double antibodies; PC9 and BEAS-2B cells were 1,640 medium with 10% fetal bovine serum and 1% dual antibodies. The cells were placed in a constant temperature incubator with a CO2 concentration of 5% and a temperature of 37°C.

RNA Extraction and Real-Time PCR

Total RNA was extracted by Trizol reagent (Invitrogen, Carlsbad, CA, United States) according to the instructions. The concentration of the extracted RNA was controlled to be 500 ng/ml with a purity between 1.80 and 2.00. Subsequently, extracted RNA was transcribed using the PrimeScript RT reagent Kit with gDNA Eraser (Takara, Japan). SYBR Green-based real-time PCR was used for analysis. PCR primers were designed and synthesized by Shanghai Bioengineering Co. (Shanghai.China). Primers used for real-time PCR assays are listed in **Supplementary Table S3**.

Statistical Analysis

Differences between groups were analyzed using the Wilcoxon test. Independent prognostic analysis was performed using univariate and multivariable cox regression analysis. Correlation tests were performed using Spearman analysis.

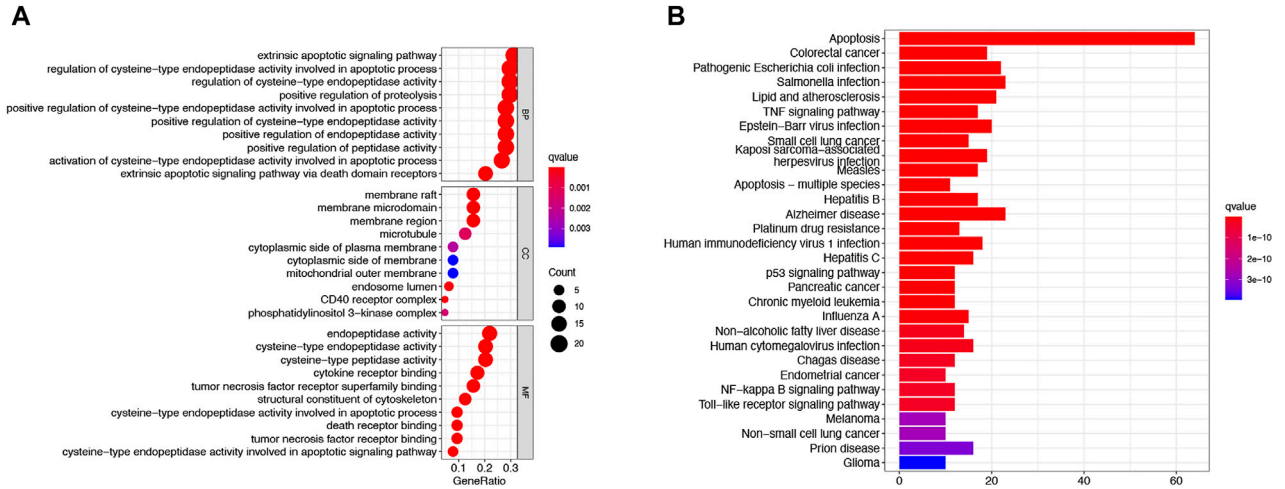


FIGURE 3 | Based on functional analysis of DEARGs between the normal and LUAD groups in the TCGA cohort.**(A)** Bubble shows KEGG pathway analysis predicted DEARGs. The circle size represents the number of genes enriched in the entry, and the color indicates the significance of the *p*-value. **(B)** Barplot shows GO enrichment analysis predicted DEARGs. The color indicates the significance of the *p*-value.

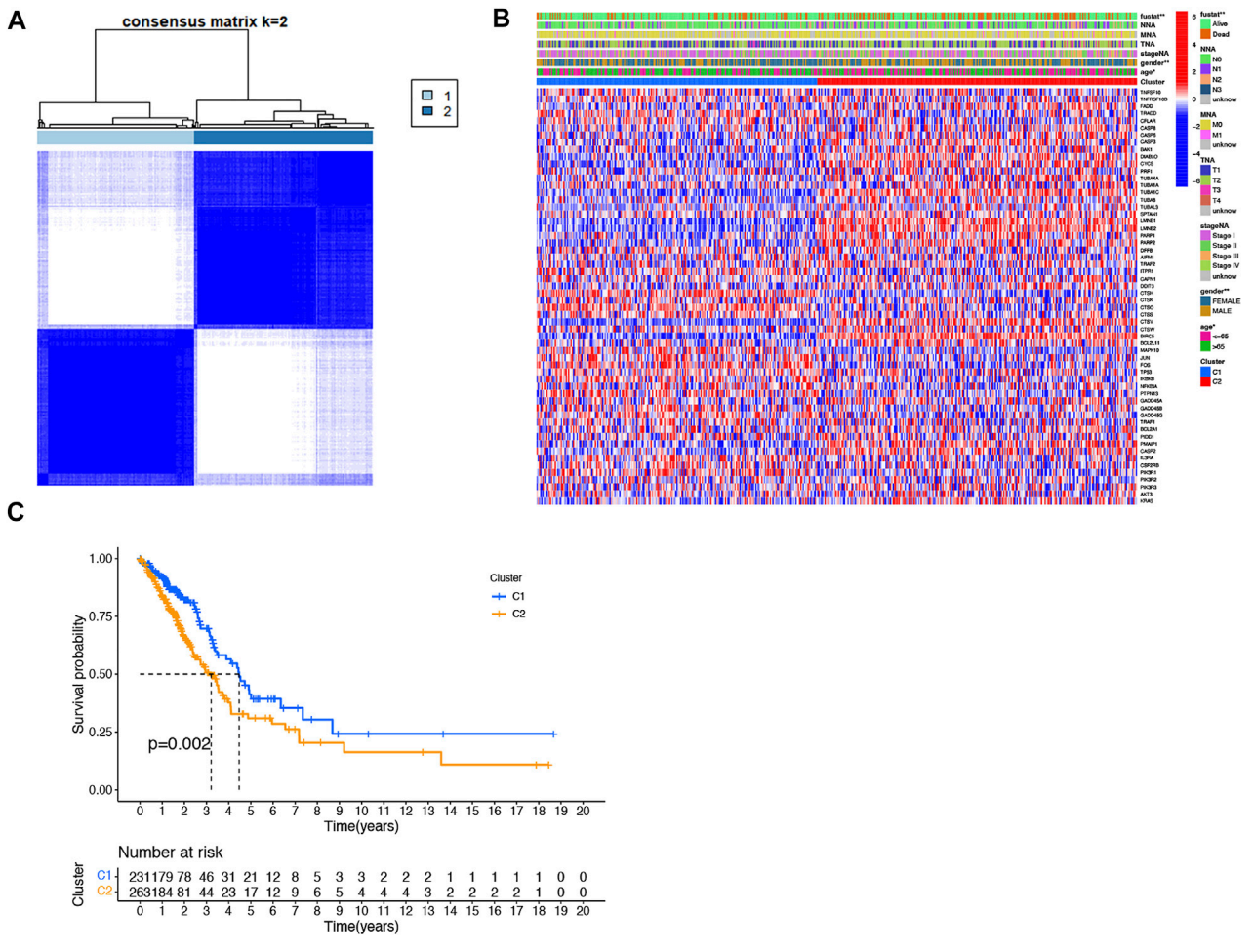


FIGURE 4 | Tumor classification based on the DEARGs. **(A)** 494 LUAD patients were grouped into two clusters according to the consensus clustering matrix (k = 2). **(B)** The heat map shows the clinicopathological characteristics of the differentially expressed genes between the two clusters. **(C)** Kaplan–Meier OS curves for the two clusters.

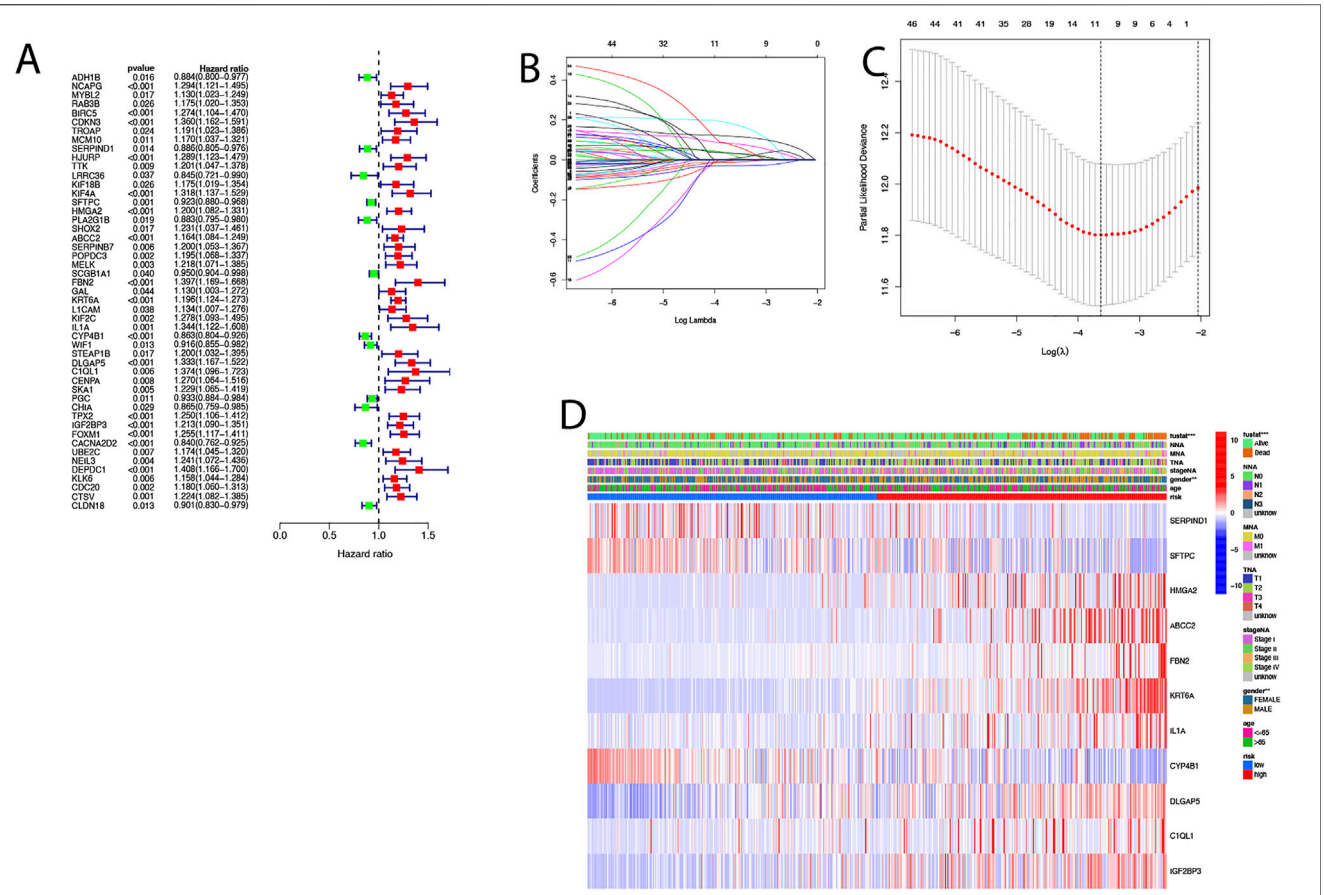


FIGURE 5 | Construction of risk signature in the TCGA cohort. **(A)** Univariate cox regression analysis of LUAD for each apoptosis-related gene, $p < 0.05$. **(B)** In the LASSO-Cox model of the TCGA cohort, the minimum standard was adopted to obtain the value of the super parameter λ by 10-fold cross-validation. **(C)** Cross-validation for tuning the parameter selection in the LASSO regression. **(D)** Heat map showing the clinical characteristics of 11 model genes.

Survival curves were plotted using log-rank and Kaplan-Meier tests. Mutations between groups were analyzed using the "maftools" R package. $p < 0.05$ was considered statistically significant. Data were processed using R 4.0.5 software.

RESULTS

LUAD Dataset Sourcing and Pre-Processing

The TCGA-LUAD database was downloaded from The Cancer Genome Atlas TCGA public database for a total of 594 samples, and patients with no survival information were excluded from further analysis. Comparing 59 normal tissues with 494 LUAD tissues for DEARGs, a total of 64 differentially expressed genes associated with apoptosis were identified, and the heatmap demonstrated the expression of each differential gene in each sample (**Figure 2A**). Twenty-one of these 64 DEARGs were genetically down-regulated and 38 differentially genetically up-regulated (**Figure 2B**; **Supplementary Table S4**). To further explore the interactions between these apoptosis-related genes, we constructed a protein-protein interaction network (PPI) with a minimum required score of 0.7 for PPI analysis, which was used to explore the interconnections between the genes (**Figure 2C**). Simultaneous drawing of a correlation network containing all apoptosis-related genes (red represents positive correlations and blue represents negative correlations) (**Figure 2D**).

Functional Analysis Based on DEARGs

To further explore the differences in the biological, behavioral functions of these DEARGs.We performed GO and KEGG enrichment analyses. The results showed that these 64 DEARGs are mainly involved in apoptosis, *salmonella* infection, and apoptotic pathways and are associated with cysteine-type endopeptidase activity, membrane rafts, and peptide chain endonuclease activity involved in the apoptotic process (**Figures 3A,B**).

Molecular Subtypes of LUAD Based on DEGs

To explore the association between DEARGs and LUAD, we performed a consensus clustering analysis on all 494 patients with survival information for LUAD in the TCGA cohort. Increasing the clustering variable k from 2 to 9 found that the highest intra-

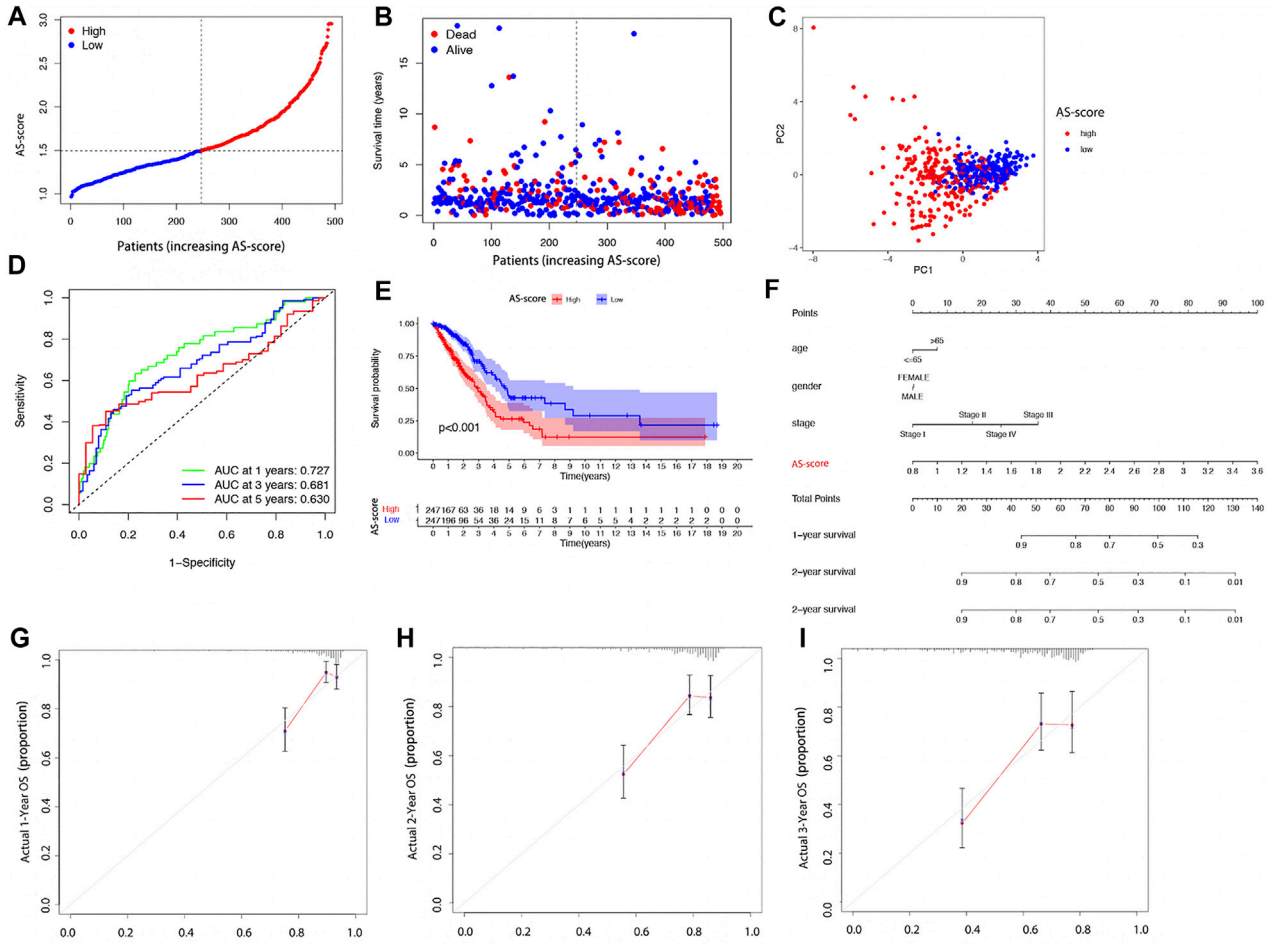


FIGURE 6 | Overall performance of the 11 gene signatures in all cohorts **(A)** Patient distribution based on risk scores. **(B)** Survival status of each patient (low-risk group: left side of the dashed line; high-risk group: right side of dashed line). **(C)** PCA plots of LUAD patients based on risk scores. **(D)** The ROC curve shows the predictive efficiency of the risk score. **(E)** Kaplan-Meier curves for LUAD patients in high and low-risk groups. **(F)** Nomogram for predicting 1, 2, and 3 years overall survival for LUAD patients in TCGA cohort. **(G–I)** Calibration plots of predicted recurrence after 1, 2, and 3 years. The x-axis represents the predicted probability of recurrence in the nomogram, and the y-axis represents the actual probability of recurrence.

group correlations and lower inter-group correlations when k = 2, indicating that the 494 LUAD patients could be well classified into two clusters based on 64 differentially expressed genes (DEGs) (**Figure 4A**). The heatmap showed the differential gene expression profile regarding clinical characteristics, including tumor grade, age (≤60 or >60 years), gender, TMN stage, and survival status (alive or dead). The results showed that the C1 group had better clinical performance than the C2 group (**Figure 4B**). We also compared the overall survival time (OS) of the two groups. There was a significant difference between the two groups (p = 0.002). The results suggested that the C1 group has a higher survival rate than the C1 group (**Figure 4C**).

Prognostic Gene Modeling in the TCGA Cohort

In order to develop a model that could quantify the ideal prognosis for each patient, we took a sample of 494 LUAD cases with complete survival information for the study. We firstly identified 153 differentially expressed genes between the two clusters and then used univariate Cox regression to analyze these 153 differentially expressed genes for initial screening (**Figure 5A**; **Supplementary Table S5**). The LASSO-Cox regression model was applied to include 11 of the differential genes with a minimum value of λ(**Figures 5B,C**). An apoptosis-related signature score was established which we named "AS-score"; a genetic risk score was constructed based on the optimal λ-value and calculated as: AS-score = (−0.003*SERPIND1 exp.) + (−0.118*SFTPC exp.) + (0.033*HMGA2 exp.) + (−0.081*ABCC2 exp.) + (0.178*FBN2 exp.) + (0.117*KRT6A exp.) + (0.022*IL1A exp.) + (−0.030*CYP4B1 exp.) + (0.007*DLGAP5 exp.) + (0.071*C1QL1 exp.) + (0.033*IGF2BP3exp.). The heatmap showed the relationship between 11 model genes and clinical characteristics (including stage TMN staging, gender, survival status, etc.) (**Figure 5D**).

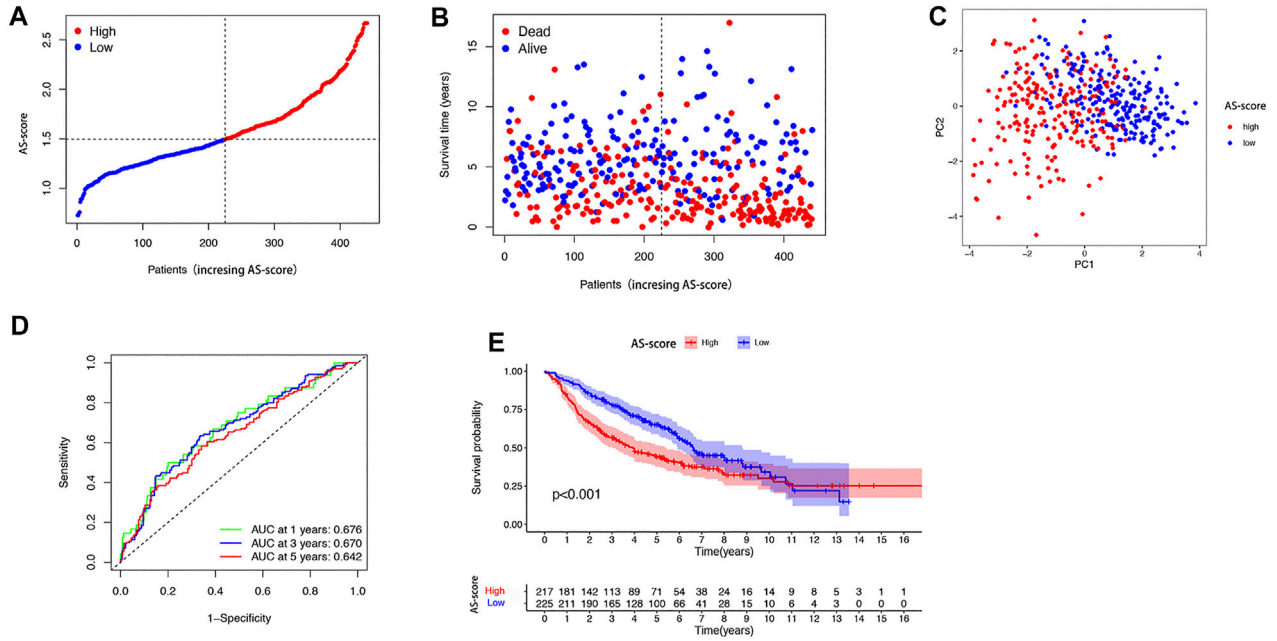


FIGURE 7 | Validation of risk prediction models in the GEO cohort **(A)** Distribution of patients based on risk score in the GEO cohort. **(B)** Survival status of each patient in the GEO cohort based on the risk prediction model (low-risk group: left side of the dashed line; high-risk group: right side of the dashed line). **(C)** PCA mapping of LUAD patients in the GEO cohort based on risk scores. **(D)** ROC curve showing the predictive efficiency of risk scores in the GEO cohort. **(E)** Kaplan-Meier curves for LUAD patients in high and low-risk groups.

The 494 LUAD patients were equally divided into high AS-score and low AS-score groups based on the median score calculated by the risk score calculation formula (**Figures 6A,B**). Principal component analysis (PCA) showed that patients with different risks could be divided into two groups (**Figure 6C**). The sensitivity and specificity of the prognostic model were assessed using time-dependent receiver operating characteristic curve (ROC) analysis. We found that the area under the ROC curve (AUC) was 0.727 at 1 year, 0.681 at 2 years, and 0.630 at 3 years (**Figure 6D**). Kaplan-Meier analysis suggested a significant difference in OS between the high AS-score group and the low AS-score group ($p < 0.001$), with the high AS-score group having a lower survival time than the low AS-score group (**Figure 6E**). To create a quantitative tool that can predict the clinical application of OS in LUAD patients, we developed a nomogram for predicting 1-, 2- and 3-years s overall survival for LUAD patients in the TCGA cohort (**Figure 6F**). Calibration plots showed that the nomogram agrees with the ideal model in the TCGA cohort (**Figures 6G–I**).

External Validation of the Risk Signature

The GSE68465 cohort of 443 LUAD patients with survival information was used as the validation set. According to the AS-score median risk score, 225 patients in the GSE68465 cohort were divided into low AS-score group and 217 into high AS-score group (**Figures 7A,B**). Principal component analysis (PCA) showed that patients could be well classified into high and low groups based on the AS-score (**Figure 7C**). ROC curve analysis of the GEO cohort showed that the constructed model was a good predictor (1-year AUC = 0.676, 2-years AUC = 0.670, 3-years survival 0.642) (**Figure 7D**). Kaplan-Meier analysis also showed a significant difference in survival between the low AS-score and high AS-score groups ($p < 0.001$). The high AS-score group had a significantly lower survival time than the low AS-score group (**Figure 7E**).

Independent Prognostic Value and Functional Analysis of Risk Signature

Univariate and multivariate Cox regression analyses were used to assess whether the risk score from the genetic trait model could be used as an independent prognostic factor. Univariate Cox regression analysis showed that period, T-stage, lymph node metastasis (N), and risk score were all factors associated with prognosis in the TCGA cohort and in the GSE68465 cohort (HR = 3.500, 95% CI: 2.536–4.829 and HR: 2.163, 95% CI: 1.600–2.926) (**Figures 8A,B**). Multivariate Cox regression analysis suggested that risk score is an independent prognostic factor after adjusting for other confounders (HR = 2.964, 95% CI: 2.120–4.145 and HR: 1.868, 95% CI: 1.372–2.544) (**Figures 8C,D**).

Comparison of Immunoreactivity

To further explore differences in gene function and pathways between risk model classifications, we identified differential genes between low AS-scores and high AS-scores in the TCGA cohort. GO enrichment analysis based on these DEGs showed that the

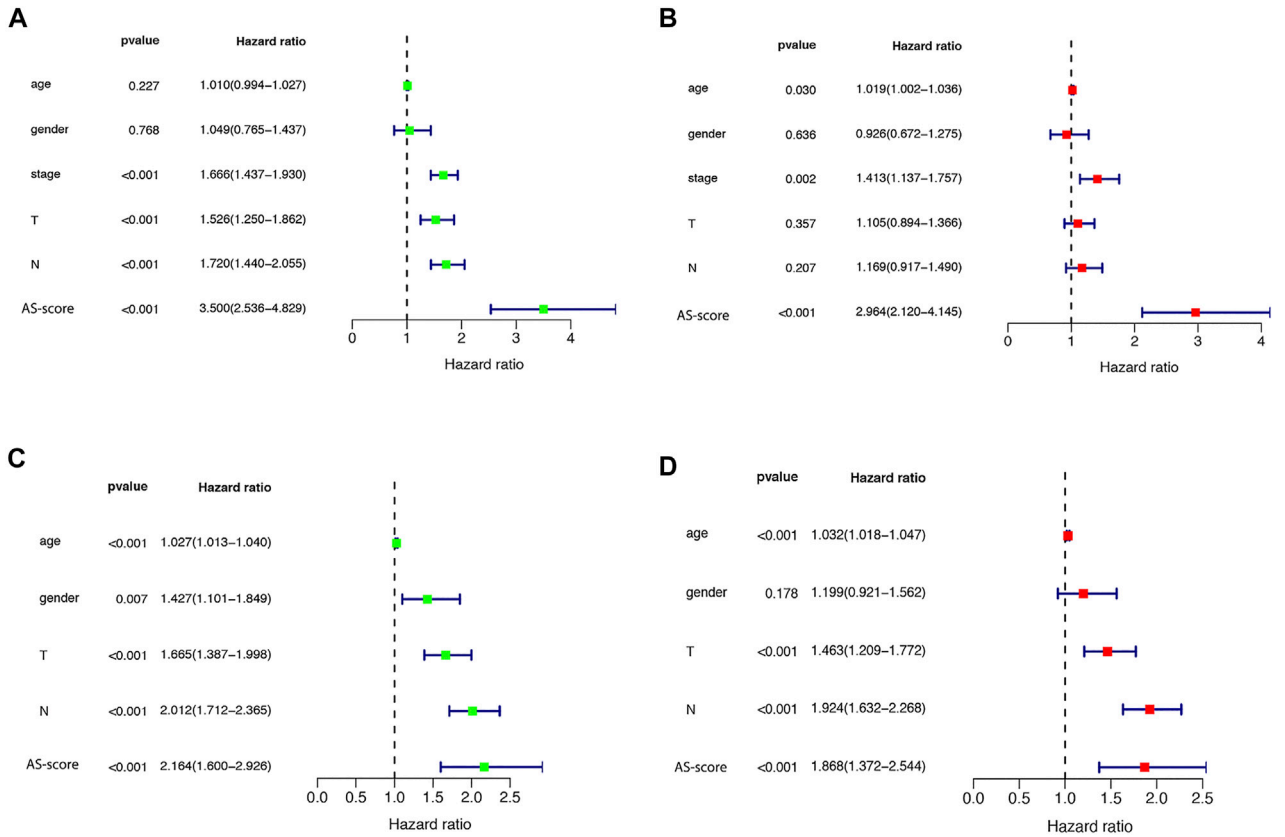


FIGURE 8 | Univariate and multivariate Cox regression analysis of risk scores **(A)** Univariate cox regression analysis in the TCGA cohort. **(B)** Multivariate Cox regression analysis in the TCGA cohort. **(C)** Univariate cox regression analysis in the GEO (GSE68465) cohort. **(D)** Multivariate Cox regression analysis in the GEO (GSE68465) cohort.

DEGs were mainly associated with signaling pathways for the immune response, epidermal development (**Figures 9A,B**).

Increasing studies have shown that tumor metastasis and invasion are inseparable from the tumor microenvironment. Based on the functional analysis, we further explored the role of the constructed AS-score in the immune microenvironment and immunotherapy. Single sample gene set enrichment analysis (ss GSEA) was used to compare the enrichment scores of 16 immune cell types and the activity of 13 immune-related pathways in the TCGA and GSE6846 cohorts in low AS-score versus high AS-score populations. In the TCGA cohort, the high AS-score group had lower levels of immune cell infiltration relative to the low AS-score group, except for p DCs Th1-cells, especially a DCs and I DCs cells. In contrast, in the immune pathway, the high AS-score was mainly enriched in HLA, MHC_class_1, Type_II_IFN_response (**Figures 10A,B**). Similar conclusions were reached in the study of immune infiltration in the GSE6846 cohort (**Figures 10C,D**).

Large studies have revealed that immunotherapy is emerging as a new hope for cancer treatment, and immune checkpoints play an important role in the immune response (Marin-Acevedo et al., 2018). To further explore the impact of high and low AS-score groups on immunotherapy, we compared the differences between immune checkpoints between the two groups. As we found, the expression of PDCDLG2, CD274, TNFSF15, CD40LG, HHLA2 was significantly upregulated, whereas CD276 and TNFSF4 were downregulated considerably in the high AS-score group, suggesting that apoptosis-related characteristics scored higher in patients who may have a better chance to immunotherapy (**Figure 11A**). Current studies suggested that patients with high TMB show more significant benefits from PD-1/PD-L1 inhibition than patients with low TMB (Hodges et al., 2017; Rizvi et al., 2018).

To further compare the mutations profile between the two AS-score groups, we examined mutation information for high and low AS-scores separately. In the TCGA cohort, 218 (91.6%) of the 238 samples in the high AS-score group had mutations, with the highest mutation frequency being in TTN (**Figure 11B**).In contrast, 211 of 246 samples (85.77%) in the low AS-score group showed mutations, with the most frequent mutation being TP53 (**Figure 11C**). At the genetic level, the TMB was higher in the high AS-score group compared to the low AS-score group ($p = 0.0097$) (**Figure 11D**). Furthermore, the risk index was positively correlated with the TMB, with an increased AS-score (p = 1.9e-05) (**Figure 11E**). This study indirectly suggested that AS- score plays a key role in mediating the clinical responses to checkpoint immunotherapy.

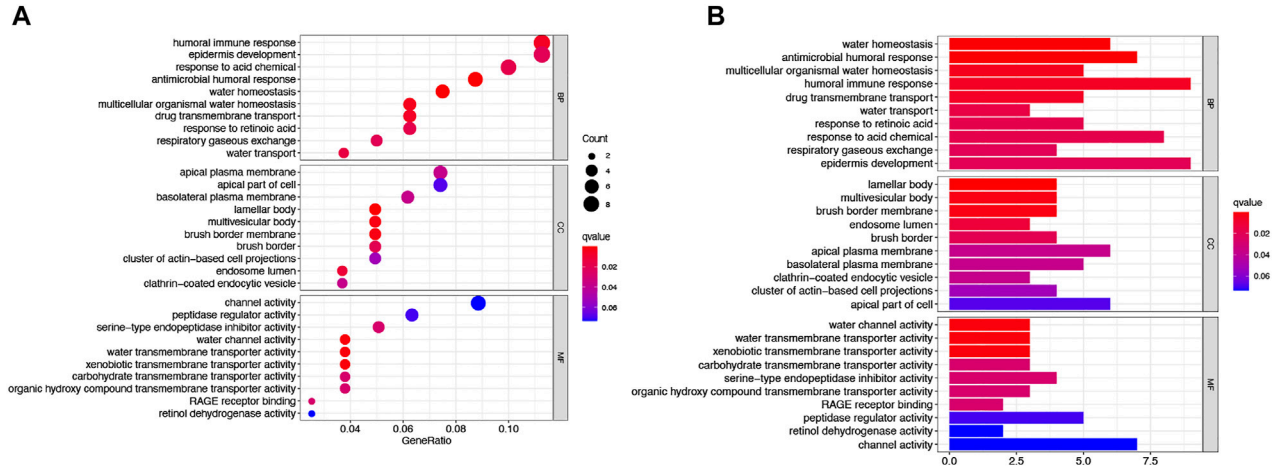


FIGURE 9 | Functional analysis of the DEGs between two risk groups in the TCGA cohort. **(A)** Bubble plot graph showing GO enrichment of DEGs between the two groups in the TCGA cohort (longer bars indicate more gene enrichment, and darker red indicates more pronounced differences q-value: the adjusted *p*-value). **(B)** Barplot showing the KEGG pathway of DEGs between the two groups in the TCGA cohort (longer bars indicate greater gene enrichment, and darker red indicates more pronounced differences).

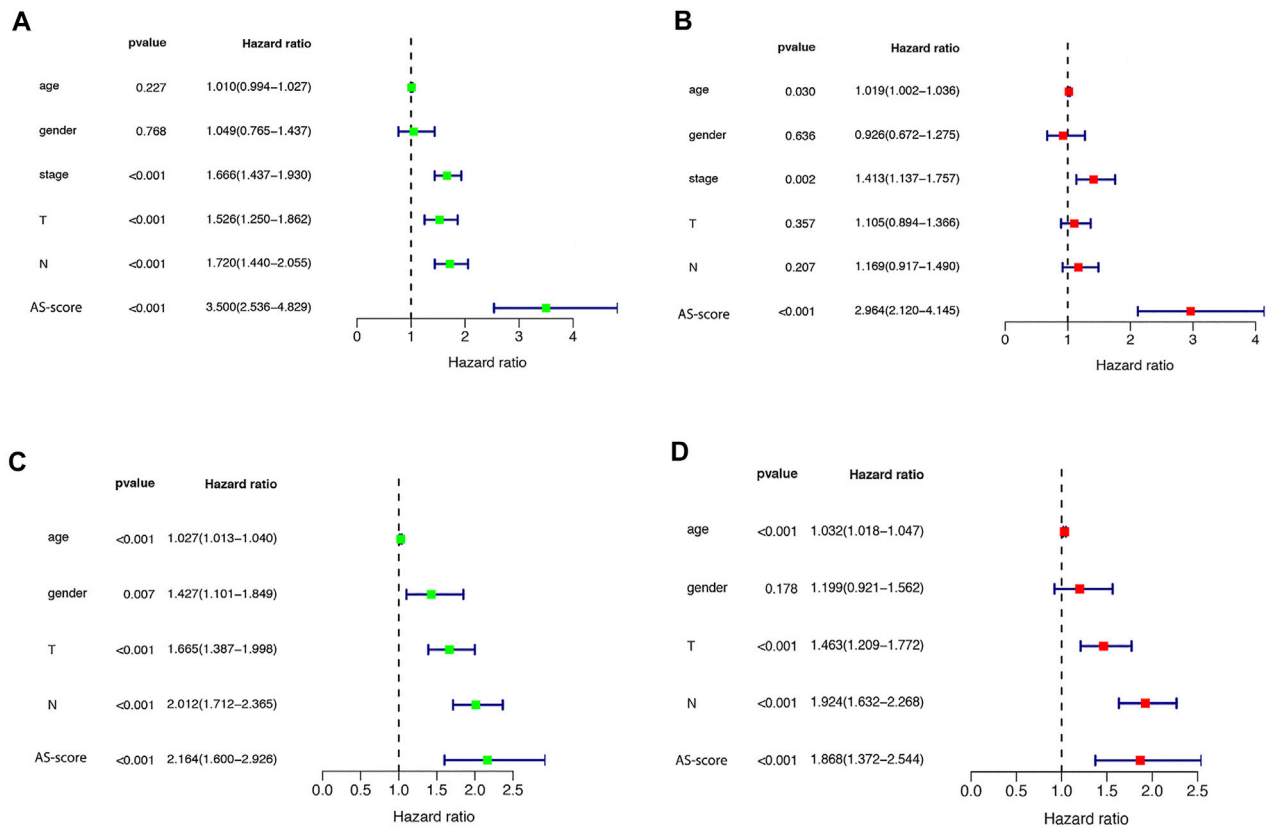


FIGURE 10 | Comparison of ssGSEA scores between the two risk groups **(A)** Immune cell infiltration between different risk groups in the TCGA cohort. **(B)** Immune-related functions between the two risk groups in the TCGA cohort. **(C)** Immune cell infiltration between different risk groups in the GEO (GSE68465) cohort. **(D)** Immune-related functions between the two risk groups in the GEO (GSE68465) cohort.

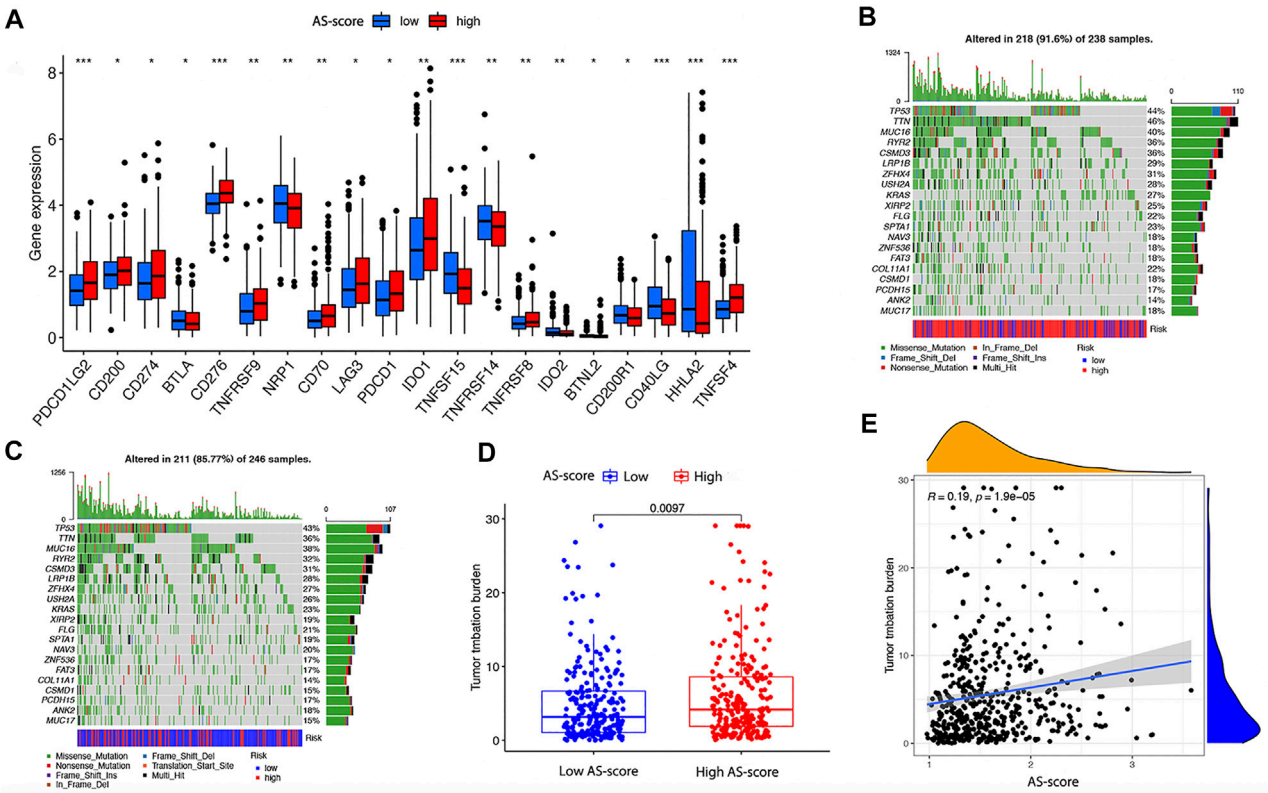


FIGURE 11 | Differences in immune checkpoint and TMB between the two risk groups in the TCGA cohort **(A)** Differences in immune checkpoints between the two risk groups in the TCGA cohort. **(B)** TMB in the high AS-score group in the TCGA cohort. **(C)** TMB in the low AS-score group in the TCGA cohort. **(D)** Differences of TMB between high and low AS-score groups in the TCGA cohort. **(E)** Association between AS-score and TMB in the TCGA cohort.

The Role of the AS-Score in Antineoplastic Drug Therapy

It has been shown that TMB predicts the efficacy of immunotherapy and influences the response to chemotherapy and targeted therapies. Therefore, we investigated the association between this AS-score and the effectiveness of chemotherapy and targeted therapies in LUAD patients (Choucair et al., 2020). To compare the efficacy of the high and low AS-score groups to commonly used chemotherapeutic and targeted drugs. The results showed that the high AS-score group had a lower half-maximal inhibitory concentration (IC50) of cisplatin (p = 2.2e-09) and paclitaxel (p < 2.22e-16) compared to the low AS-score group (**Figures 12A,B**), suggesting a higher sensitivity to treatment. Similarly, the high AS-score group had lower IC50 for erlotinib (p = 9.2e-08) and gefitinib (p = 0.0013) (**Figures 12C,D**), suggesting a better sensitivity to treatment with targeted drugs as well. These results suggested that the AS-score can predict the effect of treatment with chemotherapeutic drugs and targeted drugs.

Validation in LUAD Cell Lines

To better analyze the AS-score signatures of these 11 gene constructs, we predicted their expression in the TCGA cohort. The results showed that CYP4B1 and SFTPC were lowly expressed in LUAD, SERPIND1, HMGA2, ABCC2, KRT6A, IL1A, DLGAP5, C1QL1, IGF2BP3 were highly expressed, while FBN2 was not differentially expressed between LUAD and normal tissues (**Figure 13A**). Subsequently, we validated the expression of the 11 genes incorporated into the model constructs in the cell lines following the steps described above and found general agreement with the predicted results by comparing the differential expression of each gene in normal bronchial epithelial cells (BEAS-2B) versus A549 and PC9 (**Figures 13B,C**).

DISCUSSION

As the most common type of LC, LUAD is currently treated by surgery, radiotherapy, targeted therapy, and immunotherapy (Siegel et al., 2021).

Apoptosis, as a form of programmed cell death (PCD), is mediated through multiple signaling pathways (classified mainly as intrinsic and extrinsic pathways) (Pistritto et al., 2016). The intrinsic apoptotic pathway (mitochondrial-dependent) is mediated by intracellular signals that converge at the mitochondrial level in response to different stress conditions, with internal stimuli such as irreparable genetic damage, hypoxia, extremely high cytoplasmic Ca + concentrations, and severe

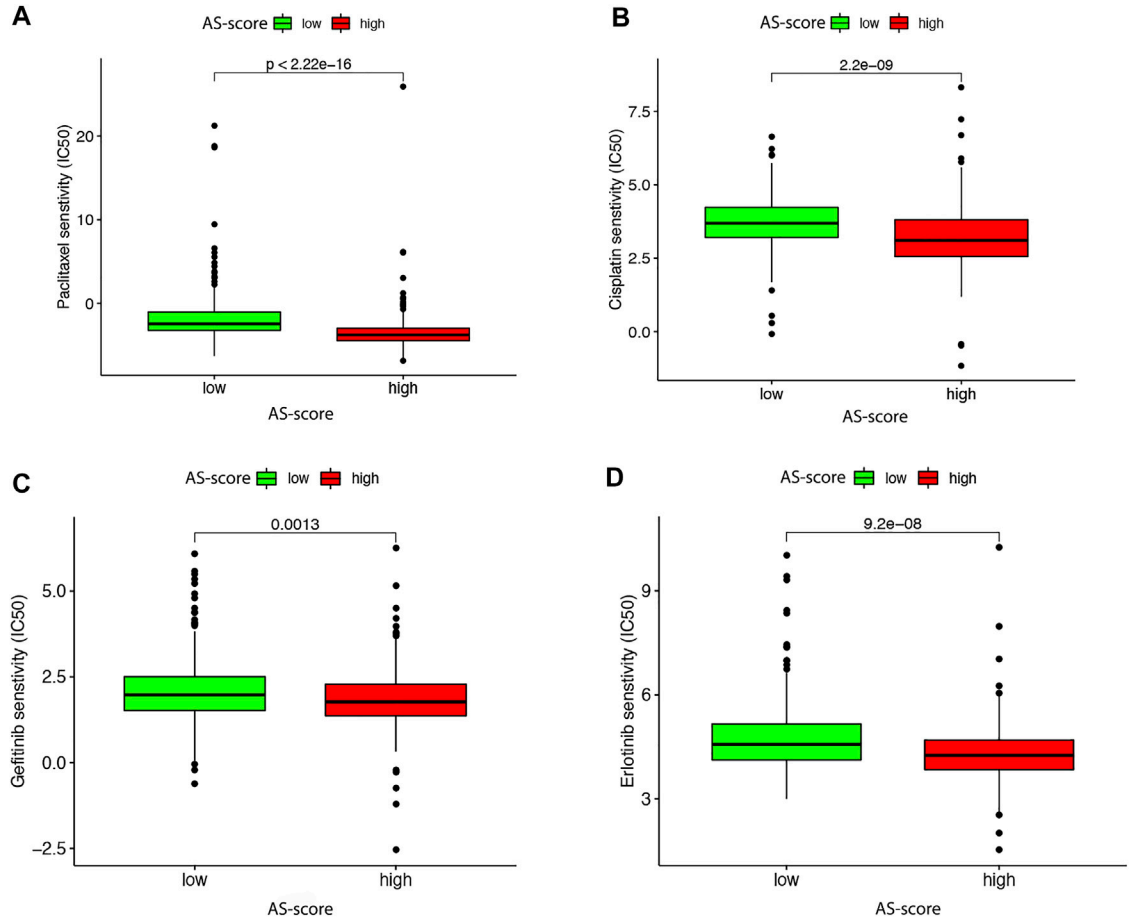


FIGURE 12 | Differences in antitumor drug therapeutic efficacy between high and low AS-scores in the TCGA cohort **(A)** Differences in paclitaxel therapeutic efficacy between high and low AS-score in the TCGA cohort. **(B)** Differences in the efficacy of cisplatin therapeutics between high and low AS-score in the TCGA cohort. **(C)** Differences in gefitinib therapeutic efficacy between high and low AS-score in the TCGA cohort. **(D)** Differences in the therapeutic efficacy of erlotinib between high and low AS-score in the TCGA cohort.

oxidative stress being important factors in initiating the intrinsic mitochondrial pathway (Green and Kroemer, 2004; Riley et al., 2018). In contrast, the extrinsic apoptotic pathway (death receptor-dependent) is initiated by the interaction of death receptors of the tumor necrosis factor receptor (TNFR) superfamily exposed on the cell surface with the respective protein TNF family ligands (Guicciardi and Gores, 2009). The apoptotic process sequentially and efficiently removes cells that cause damage (e.g., DNA damage or cells generated during development), thereby maintaining cell renewal, embryonic development, and immune system activity. Interaction of apoptotic pathways with other signaling mechanisms also affects cell death (Bauer and Helfand, 2006). Dysregulation of apoptotic cell death mechanisms is a feature of cancer, as shown by a growing body of literature (Lauber and Herrmann, 2015). The altered apoptosis is not only associated with tumor development and progression but also with resistance to antitumor drugs, and therapeutic strategies targeting apoptosis-resistant molecules are an effective way to restore the sensitivity of cancer cells to apoptosis and enhance antitumor effect (Kim, 2005; Mohammad et al., 2015).

However, it is unclear how apoptosis-associated genes interact in LUAD and whether these genes are associated with clinical characteristics of patients, prediction of antitumor drug efficacy, and infiltration in the immunological microenvironment. Classification of samples based on predefined gene expression profiles is a proven method (Cristescu et al., 2015). In this study, we used this method to analyze the data information of TCGA-LUAD. The 136 genes associated with apoptosis were investigated between LUAD and normal tissues, and the results revealed differences in the expression of most of these genes. Protein-protein interaction networks were mapped to demonstrate the interrelationships. According to the DEARGs associated with apoptosis, two distinct subtypes existed by consensus clustering analysis. There were significant differences in survival and prognosis between the two subtypes. To further assess the prognostic value of these associated regulators, we initially performed a preliminary screening of these differential genes using univariate analysis and then constructed an 11 (SERPIND1, SFTPC, HMGAA2, ABCC2, FBN2, KRT6A, IL-1A, CYP4B1, DLGAP5 C1QL1, IGF2BP3) genetic risk score model, named

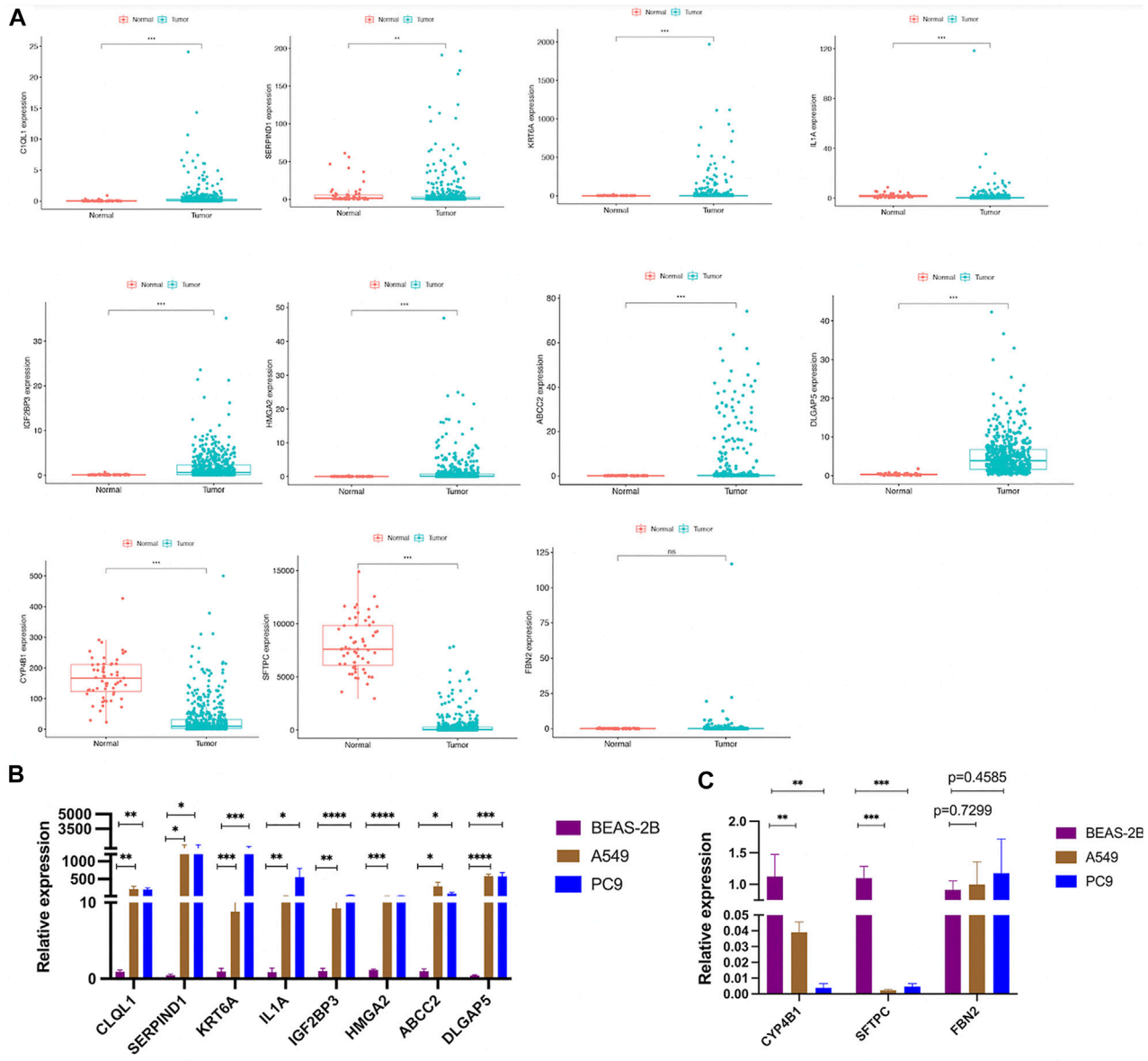


FIGURE 13 | Validation of the differential expression of the 11 model genes **(A)** Expression of 11 model genes from the TCGA cohort in LUAD tissue. **(B,C)** Validation of differential expression of the 11 model genes in LUAD cell lines.

AS-score. LUAD patients in the TCGA cohort were then divided into high and low AS-score groups based on the median AS-score score. The performance from this risk model score was validated in the GSE6846 cohort. Both the TCGA and GSE6846 cohorts showed that the low AS-score group had better prognostic survival and prognostic performance than the high AS-score group. On the basis that this AS-score is a good predictor of patient prognosis and an independent prognostic risk factor for LUAD, we validated its expression in LUAD cell lines versus normal bronchial epithelial cells. We found it to be consistent with database predictions.

It has been shown that the tumor microenvironment is significantly related to the clinical features, genomic expression, and biological characteristics of tumor patients (Zhang et al., 2020). A comprehensive analysis of the role of the tumor microenvironment in LUAD will help clarify the tumor immunophenotype of LUAD, explore independent prognostic indicators and new therapeutic targets, thus improving patient prognosis and predicting the effectiveness of immunotherapy (Liotta and Kohn, 2001; Fang and Declerck, 2013). Previous studies have also shown that the tumor microenvironment plays a key role in tumor carcinogenesis and revealed significant epigenetic regulators, opening new avenues for precision and personalized medicine (Gadiyar et al., 2020). According to our findings, differentially expressed genes between the low AS-score group and the high AS-score group were associated with immune-related pathways. Comparison of immune cell infiltration

and activation pathways between the low AS-score and high AS-score groups showed differences in multiple cellular and immune pathways.

Current immunotherapy for LUAD is mainly directed at immune checkpoint inhibitors (Marin-Acevedo et al., 2018; Wang et al., 2020). The most common and rapidly developing of these are PD-L1 and PD-1 inhibitors, which exert antitumor effects by blocking the binding of PD-L1 to PD-1, thereby reducing the inhibition of T-cell activation, suggesting that high PD-L1 expression is more effective for immunosuppressive therapy (Zhu et al., 2018; Wang et al., 2021).

In order to further explore the differential impacts on immunity between the high and low AS-score, we compared the TMB. We found that the high AS-score group had more mutational load than the low AS-score group and that AS-score was positively correlated with TMB. Rizvi et al. (2018) suggested that TMB and PD-L1 expression, although not correlated, independently predicted the efficacy of immunosuppressive therapy and combined TMB and PD-L1 expression into a multivariate prediction model should yield greater predictive capability. We also showed that the high AS-score group had higher PD-L1expression but TMB than the low AS-score group, and there may be a consistent synergistic predictive effect between the two. The high AS-score group had better immunotherapy efficacy than the low AS-score group (Samstein et al., 2019).

The prediction model based on DEARGs showed that the high AS-score group achieved better efficacy in immunotherapy. Similarly, the high AS-score group achieved better efficacy with chemotherapy and targeted drug therapy. At the same time, Ludovic Fournel (Fournel et al., 2019) found that cisplatin-based induction chemotherapy increased PD-L1 expression in tumor cells, suggesting that chemotherapy combined with immunotherapy could improve the overall prognosis of patients with LUAD. And the combination of cisplatin and anti-PD - L1 therapy improved the response to tumor treatment, which was consistent with the predictions of the AS-score model. The results of this study showed that although the low AS-score group had better survival than the high AS-score group. But fortunately, the high AS-score group was more sensitive to chemotherapy and targeted therapy. Meanwhile, because cisplatin-based therapy could increase the expression of PD-L1 in tumor cells, suggesting that the high AS-score group is more sensitive to chemotherapy combined with targeted therapy. Therefore, it is reasonable to assume that this model may indicate antitumor therapy.

In conclusion, our study showed that this AS-score plays a role in clinical prognosis and sensitivity to antitumor drug therapy. Our study provides a new genetic marker for predicting the prognosis of patients with LUAD and provides an important basis for further research into the relationship and new apoptosis-related antitumor drug treatments.

DATA AVAILABILITY STATEMENT

The data could be downloaded at (https://portal.gdc.cancer.gov/, https://xenabrowser.net/, and https://www.ncbi.nlm.nih.gov/geo/; GSE68465) and the code used during the current study are available from the corresponding author on reasonable request.

AUTHOR CONTRIBUTIONS

XZ analyzed data and wrote the manuscript. RH performed study concept and design. ZZ Analyzed data, interpreted results, and helped to write the manuscript. YY revised the manuscript and made final approval of the version. All authors have read and agreed to the published version of the manuscript.

FUNDING

Natural Science Foundation of Jiangsu Province (BK20151,333); Zhenjiang Key Research and Development project (SH2018059).

ACKNOWLEDGMENTS

We want to acknowledge the TCGA and the GEO (GSE68465) network for providing data. We want to express our sincere appreciation to the People's Hospital of Jiangsu University for their support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.818403/full#supplementary-material

REFERENCES

Bauer, J. H., and Helfand, S. L. (2006). New Tricks of an Old Molecule: Lifespan Regulation by P53. *Aging cell* 5 (5), 437–440. doi:10.1111/j.1474-9726.2006.00228.x

Carneiro, B. A., and El-Deiry, W. S. (2020). Targeting Apoptosis in Cancer Therapy. *Nat. Rev. Clin. Oncol.* 17 (7), 395–417. doi:10.1038/s41571-020-0341-y

Choucair, K., Morand, S., Stanbery, L., Edelman, G., Dworkin, L., and Nemunaitis, J. (2020). TMB: a Promising Immune-Response Biomarker, and Potential Spearhead in Advancing Targeted Therapy Trials. *Cancer Gene Ther.* 27 (12), 841–853. doi:10.1038/s41417-020-0174-y

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

Degterev, A., Hitomi, J., Germscheid, M., Ch'en, I. L., Korkina, O., Teng, X., et al. (2008). Identification of RIP1 Kinase as a Specific Cellular Target of Necrostatins. *Nat. Chem. Biol.* 4 (5), 313–321. doi:10.1038/nchembio.83

Denisenko, T. V., Budkevich, I. N., and Zhivotovsky, B. (2018). Cell Death-Based Treatment of Lung Adenocarcinoma. *Cell Death Dis* 9 (2), 117. doi:10.1038/s41419-017-0063-y

Fang, H., and Declerck, Y. A. (2013). Targeting the Tumor Microenvironment: from Understanding Pathways to Effective Clinical Trials. *Cancer Res.* 73 (16), 4965–4977. doi:10.1158/0008-5472.CAN-13-0661

Fournel, L., Wu, Z., Stadler, N., Damotte, D., Lococo, F., Boulle, G., et al. (2019). Cisplatin Increases PD-L1 Expression and Optimizes Immune Check-point

Blockade in Non-small Cell Lung Cancer. *Cancer Lett.* 464, 5–14. doi:10.1016/j.canlet.2019.08.005

Gadiyar, V., Lahey, K. C., Calianese, D., Devoe, C., Mehta, D., Bono, K., et al. (2020). Cell Death in the Tumor Microenvironment: Implications for Cancer Immunotherapy. *Cells* 9 (10), 2207. doi:10.3390/cells9102207

Giménez-Bonafé, P., Tortosa, A., and Pérez-Tomás, R. (2009). Overcoming Drug Resistance by Enhancing Apoptosis of Tumor Cells. *Ccdt* 9 (3), 320–340. doi:10.2174/156800909788166600

Green, D. R., and Kroemer, G. (2004). The Pathophysiology of Mitochondrial Cell Death. *Science* 305 (5684), 626–629. doi:10.1126/science.1099320

Greenawalt, E. J., Edmonds, M. D., Jain, N., Adams, C. M., Mitra, R., and Eischen, C. M. (2019). Targeting of SGK1 by miR-576-3p Inhibits Lung Adenocarcinoma Migration and Invasion. *Mol. Cancer Res.* 17 (1), 289–298. doi:10.1158/1541-7786.MCR-18-0364

Guicciardi, M. E., and Gores, G. J. (2009). Life and Death by Death Receptors. *FASEB j.* 23 (6), 1625–1637. doi:10.1096/fj.08-111005

Hassan, M., Watari, H., AbuAlmaaty, A., Ohba, Y., and Sakuragi, N. (2014). Apoptosis and Molecular Targeting Therapy in Cancer. *Biomed. Research International* 2014, 1–23. doi:10.1155/2014/150845

Hengartner, M. O. (2001). Apoptosis. *Cell* 104 (3), 325–328. doi:10.1016/s0092-8674(01)00219-7

Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Wu, Y.-L., et al. (2017). Lung Cancer: Current Therapies and New Targeted Treatments. *The Lancet* 389 (10066), 299–311. doi:10.1016/s0140-6736(16)30958-8

Hodges, T. R., Ott, M., Xiu, J., Gatalica, Z., Swensen, J., Zhou, S., et al. (2017). Mutational burden, Immune Checkpoint Expression, and Mismatch Repair in Glioma: Implications for Immune Checkpoint Immunotherapy. *Neuro-oncology* 19 (8), 1047–1057. doi:10.1093/neuonc/nox026

Kim, R. (2005). Recent Advances in Understanding the Cell Death Pathways Activated by Anticancer Therapy. *Cancer* 103 (8), 1551–1560. doi:10.1002/cncr.20947

Kroemer, G., El-Deiry, W. S., Golstein, P., Peter, M. E., Vaux, D., Vandenabeele, P., et al. (2005). Classification of Cell Death: Recommendations of the Nomenclature Committee on Cell Death. *Cell Death Differ* 12 Suppl 2 (Suppl. 2), 1463–1467. doi:10.1038/sj.cdd.4401724

Lauber, K., and Herrmann, M. (2015). Tumor Biology: with a Little Help from My Dying Friends. *Curr. Biol.* 25 (5), R198–R201. doi:10.1016/j.cub.2015.01.040

Ledgerwood, E. C., and Morison, I. M. (2009). Targeting the Apoptosome for Cancer Therapy. *Clin. Cancer Res.* 15 (2), 420–424. doi:10.1158/1078-0432.Ccr-08-1172

Li, J., and Yuan, J. (2008). Caspases in Apoptosis and beyond. *Oncogene* 27 (48), 6194–6206. doi:10.1038/onc.2008.297

Liotta, L. A., and Kohn, E. C. (2001). The Microenvironment of the Tumour-Host Interface. *Nature* 411 (6835), 375–379. doi:10.1038/35077241

Liu, L.-P., Lu, L., Zhao, Q.-Q., Kou, Q.-J., Jiang, Z.-Z., Gui, R., et al. (2021a). Identification and Validation of the Pyroptosis-Related Molecular Subtypes of Lung Adenocarcinoma by Bioinformatics and Machine Learning. *Front. Cel Dev. Biol.* 9, 756340. doi:10.3389/fcell.2021.756340

Liu, Z., Lu, T., Li, J., Wang, L., Xu, K., Dang, Q., et al. (2021b). Clinical Significance and Inflammatory Landscape of aNovel Recurrence-Associated Immune Signature in Stage II/III Colorectal Cancer. *Front. Immunol.* 12, 702594. doi:10.3389/fimmu.2021.702594

Liu, Z., Wang, L., Liu, L., Lu, T., Jiao, D., Sun, Y., et al. (2021c). The Identification and Validation of Two Heterogenous Subtypes and a Risk Signature Based on Ferroptosis in Hepatocellular Carcinoma. *Front. Oncol.* 11, 619242. doi:10.3389/fonc.2021.619242

Marin-Acevedo, J. A., Dholaria, B., Soyano, A. E., Knutson, K. L., Chumsri, S., and Lou, Y. (2018). Next Generation of Immune Checkpoint Therapy in Cancer: New Developments and Challenges. *J. Hematol. Oncol.* 11 (1), 39. doi:10.1186/s13045-018-0582-8

Mohammad, R. M., Muqbil, I., Lowe, L., Yedjou, C., Hsu, H.-Y., Lin, L.-T., et al. (2015). Broad Targeting of Resistance to Apoptosis in Cancer. *Semin. Cancer Biol.* 35 (Suppl. l), S78–S103. doi:10.1016/j.semcancer.2015.03.001

Pistritto, G., Trisciuoglio, D., Ceci, C., Garufi, A., and D'Orazi, G. (2016). Apoptosis as Anticancer Mechanism: Function and Dysfunction of its Modulators and Targeted Therapeutic Strategies. *Aging* 8 (4), 603–619. doi:10.18632/aging.100934

Plati, J., Bucur, O., and Khosravi-Far, R. (2008). Dysregulation of Apoptotic Signaling in Cancer: Molecular Mechanisms and Therapeutic Opportunities. *J. Cel. Biochem.* 104 (4), 1124–1149. doi:10.1002/jcb.21707

Riley, J. S., Quarato, G., Cloix, C., Lopez, J., O'Prey, J., Pearson, M., et al. (2018). Mitochondrial Inner Membrane Permeabilisation Enables Mt DNA Release during Apoptosis. *Embo J.* 37 (17). doi:10.15252/embj.201899238

Rizvi, H., Sanchez-Vega, F., La, K., Chatila, W., Jonsson, P., Halpenny, D., et al. (2018). Molecular Determinants of Response to Anti-programmed Cell Death (PD)-1 and Anti-programmed Death-Ligand 1 (PD-L1) Blockade in Patients with Non-small-cell Lung Cancer Profiled with Targeted Next-Generation Sequencing. *Jco* 36 (7), 633–641. doi:10.1200/JCO.2017.75.3384

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 (2), 202–206. doi:10.1038/s41588-018-0312-8

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

Singh, R., Letai, A., and Sarosiek, K. (2019). Regulation of Apoptosis in Health and Disease: the Balancing Act of BCL-2 Family Proteins. *Nat. Rev. Mol. Cel Biol* 20 (3), 175–193. doi:10.1038/s41580-018-0089-8

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 A. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Wang, L., Zhang, S., Li, H., Xu, Y., Wu, Q., Shen, J., et al. (2021). Quantification of m6A RNA Methylation Modulators Pattern Was a Potential Biomarker for Prognosis and Associated with Tumor Immune Microenvironment of Pancreatic Adenocarcinoma. *BMC Cancer* 21 (1), 876. doi:10.1186/s12885-021-08550-9

Wang, Q., Li, M., Yang, M., Yang, Y., Song, F., Zhang, W., et al. (2020). Analysis of Immune-Related Signatures of Lung Adenocarcinoma Identified Two Distinct Subtypes: Implications for Immune Checkpoint Blockade Therapy. *Aging* 12 (4), 3312–3339. doi:10.18632/aging.102814

Wu, X., Lan, T., Li, M., Liu, J., Wu, X., Shen, S., et al. (2021). Six Metabolism Related mRNAs Predict the Prognosis of Patients with Hepatocellular Carcinoma. *Front. Mol. Biosci.* 8, 621232. doi:10.3389/fmolb.2021.621232

Xie, Q., Liu, Y., and Li, X. (2020). The Interaction Mechanism between Autophagy and Apoptosis in colon Cancer. *Translational Oncol.* 13 (12), 100871. doi:10.1016/j.tranon.2020.100871

Xu, X., Lai, Y., and Hua, Z.-C. (2019). Apoptosis and Apoptotic Body: Disease Message and Therapeutic Target Potentials. *Biosci. Rep.* 39 (1). doi:10.1042/BSR20180992

Yaacoub, K., Pedeux, R., Tarte, K., and Guillaudeux, T. (2016). Role of the Tumor Microenvironment in Regulating Apoptosis and Cancer Progression. *Cancer Lett.* 378 (2), 150–159. doi:10.1016/j.canlet.2016.05.012

Zhang, B., Wu, Q., Li, B., Wang, D., Wang, L., and Zhou, Y. L. (2020). m6A Regulator-Mediated Methylation Modification Patterns and Tumor Microenvironment Infiltration Characterization in Gastric Cancer. *Mol. Cancer* 19 (1), 53. doi:10.1186/s12943-020-01170-0

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

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 Zou, He, Zhang and Yan. 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: 27 April 2022
doi: 10.3389/fgene.2022.864383

Prognostic Characteristics and Immune Effects of N^6-Methyladenosine and 5-Methylcytosine-Related Regulatory Factors in Clear Cell Renal Cell Carcinoma

*Lei Li, Zijia Tao, Yiqiao Zhao, Mingyang Li, Jianyi Zheng, Zeyu Li and Xiaonan Chen**

Department of Urology, Shengjing Hospital of China Medical University, Shenyang, China

In recent years, methylation modification regulators have been found to have essential roles in various tumor mechanisms. However, the relationships between N^6-methyladenosine (m6A) and 5-methylcytosine (m5C) regulators and clear cell renal cell carcinoma (ccRCC) remain unknown. This study investigated these relationships using the data from The Cancer Genome Atlas database. We calculated risk scores using a Lasso regression analysis and divided the patient samples into two risk groups (tumor vs. normal tissues). Furthermore, we used univariate and multivariate Cox analyses to determine independent prognostic indicators and explore correlations between the regulatory factors and immune infiltrating cell characteristics. Finally, quantitative reverse transcriptase–polymerase chain reaction (PCR) and The Human Protein Atlas were used to verify signature-related gene expression in clinical samples. We identified expression differences in 35 regulatory factors between the tumor and normal tissue groups. Next, we constructed a five-gene risk score signature (NOP2 nucleolar protein [*NOP2*], methyltransferase 14, N6-adenosine-methyltransferase subunit [*METTL14*], NOP2/Sun RNA methyltransferase 5 [*NSUN5*], heterogeneous nuclear ribonucleoprotein A2/B1 [*HNRNPA2B1*], and zinc finger CCCH-type containing 13 [*ZC3H13*]) using the screening criteria ($p < 0.01$), and then divided the cases into high- and low-risk groups based on their median risk score. We also screened for independent prognostic factors related to age, tumor grade, and risk score. Furthermore, we constructed a Norman diagram prognostic model by combining two clinicopathological characteristics, which demonstrated good prediction efficiency with prognostic markers. Then, we used a single-sample gene set enrichment analysis and the cell-type identification by estimating relative subsets of RNA transcripts (CIBERSORT) method to evaluate the tumor microenvironment of the regulatory factor prognostic characteristics. Moreover, we evaluated five risk subgroups with different genetic signatures for personalized prognoses. Finally, we analyzed the immunotherapy and immune infiltration response and demonstrated that the high-risk group was more sensitive to immunotherapy than the low-risk group. The PCR results showed that *NSUN5* and *HNRNPA2B1* expression was higher in tumor tissues than in normal tissues. In conclusion, we identified five m6A and m5C regulatory factors that might be promising biomarkers for future research.

OPEN ACCESS

Edited by:
William C. Cho, QEH, Hong Kong SAR, China

Reviewed by:
Ziheng Wang, Affiliated Hospital of Nantong University, China
Dan Bai, Northwestern Polytechnical University, China

***Correspondence:**
Xiaonan Chen
chenxn@cmu.edu.cn

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

Received: *28 January 2022*
Accepted: *14 March 2022*
Published: *27 April 2022*

Citation:
Li L, Tao Z, Zhao Y, Li M, Zheng J, Li Z and Chen X (2022) Prognostic Characteristics and Immune Effects of N^6-Methyladenosine and 5-Methylcytosine-Related Regulatory Factors in Clear Cell Renal Cell Carcinoma. Front. Genet. 13:864383. doi: 10.3389/fgene.2022.864383

Keywords: renal clear cell cancer, N6-methyladenose, prognosis, The Cancer Genome Atlas, 5-methylcytosine, tumor microenvironment

INTRODUCTION

Renal cell carcinoma (RCC) is the most common renal malignancy, resulting in more than 14,000 deaths annually in the United States (Siegel et al., 2018). Clear cell renal cell carcinoma (ccRCC) is the most common subtype (~70% of cases) with the worst degree of malignancy and prognosis (Jonasch et al., 2021). There has been considerable progress regarding ccRCC treatment. However, overall survival (OS) and relapse-free survival still require improvement (Greef and Eisen, 2016). Therefore, identifying new targets and prognostic biomarkers for ccRCC treatment is crucial.

N-Methyladenosine (m6A) and 5-methylcytosine (m5C) RNA modifications are newly discovered gene expression regulation mechanisms (Chen et al., 2019). These modifications affect the fate of modified RNA and play critical roles in biological processes such as tumor development (Huang et al., 2020). m6A modifications are the most abundant type of RNA modification, occurring in messenger RNA (mRNA), microRNA, and long noncoding RNAs. Furthermore, studies have demonstrated that m6A RNA modifications affect RNA processing, translation, and metabolism. Methyltransferases (writers), demethylases (erasers), and binding-protein (readers) primarily mediate the effects of m6A. The "writers" are responsible for RNA methylation and include methyltransferase-like (METTL) 33, METTL14, KIAA1429, Wilms' tumor 1-associating protein (WTAP), RNA binding motif protein 15 (RBM15), and zinc finger CCCH-type containing 13 (ZC3H13). The "erasers" specifically target m6A RNA and mainly include AlkB homolog 5, RNA demethylase (ALKBH5), and FTO alpha-ketoglutarate-dependent dioxygenase (FTO). The "readers" connect m6A sites and play a role in special regulatory RNA modifications, including YTH domain-containing (YTHDC) 1, YTHDC2, YTH domain-containing family protein (YTHDF) 1, YTHDF2, insulin-like growth factor binding protein (IGFBP) 1, IGFBP2, IGF2BP3, RNA binding motif protein X-Linked (RBMX), and heterogeneous nuclear ribonucleoprotein C (HNRNPC). m6A downregulation leads to reduced proliferation, self-renewal, survival, and differentiation. Furthermore, m6A methylation regulates all aspects of cellular RNA metabolism, including abundance, alternative splicing, stability, nuclear output, decay, and transformation (Chen et al., 2019; He et al., 2019; Lan et al., 2019).

m5C modifications are another prevalent RNA modification type occurring in mRNA, transfer RNA, ribosomal RNA, and some noncoding RNAs (Squires et al., 2012; Bohnsack et al., 2019; Trixl and Lusser, 2019). Studies have demonstrated that m5C is involved in gene expression related to RNA output, translation, and stabilization processes (Bohnsack et al., 2019; Huang et al., 2019). Writers, readers, and erasers also mediate the effects of m5C. NOP2 nucleolar protein (NSUN) 1–7, DNA methyltransferase (DNMT) 1, 2, 3A, and 3B are writers, regulating the process of RNA methylation modifications. Tet methylcytosine dioxygenase (TET) 2, an eraser, has m5C demethylation activity, removing the m5C modification, and Aly/REF export factor (ALYREF), a reader, recognizes and binds the m5C site on target mRNAs (Huang et al., 2019; Nombela et al., 2021).

Increasing evidence suggests that m6A and m5C regulators have essential roles in tumorigenesis and tumor progression (Chen et al., 2019; Chellamuthu and Gray, 2020). For example, tumorigenesis and proliferation, differentiation, invasion, and migration are related to methylation modifications (Lin et al., 2016; Ma et al., 2017; Liu et al., 2018a). Furthermore, m6A and m5C regulators have been reported as prognostic biomarkers. For instance, in hepatocellular carcinoma, METTL3 is associated with a poor prognosis and inhibits the suppressor of cytokine signaling 2 (i.e., SOCS2) expression through the miR-145/m 6 A/YTHDF2-dependent axis (Yang et al., 2017; Chen et al., 2018). In addition, METTL14 promotes cancer progression by regulating MYB proto-oncogene–transcription factor (i.e., MYB)/MYC proto-oncogene–bHLH transcription factor (i.e., MYC) in acute myeloid leukemia (Weng et al., 2018). m5C research has not become mainstream. However, studies have confirmed increased *NSUN2* expression in breast cancer (Frye and Watt, 2006), and *NOP2* is a non–small cell lung cancer prognostic biomarker (Sato et al., 1999). Furthermore, *NSUN5* is highly expressed in rectal cancer and promotes cancer progression through cell cycle regulation (Jiang et al., 2020).

RNA modifications do not drive tumor progression. However, abnormal expression of modification regulators can lead to changes in the biological behavior of tumors (Nombela et al., 2021). For example, *METTL3* is upregulated in breast cancer, which increases hepatitis B X-interacting protein (i.e., *HBXIP*) mRNA methylation and stability, inducing tumor cell proliferation and survival by inhibiting the tumor suppressor, let −7 g (Cai et al., 2018). METTL3 also regulates integrin subunit beta 1 (ITGB1) expression, thereby affecting the binding of ITGB1 to collagen I. This disruption affects tumor cell migration and promotes bone metastasis in prostate cancer (Li et al., 2020). In cervical cancer, FTO activates the β-catenin pathway, increasing ERCC excision repair 1–endonuclease noncatalytic subunit (i.e., *ERCC1*) expression, which is associated with worse prognosis (Zhou et al., 2018; Zou et al., 2019). In addition, FTO is overexpressed in lung cancer, promoting cell proliferation and invasion and inhibiting apoptosis by regulating myeloid zinc finger 1 (i.e., *MZF1*) expression, resulting in a poor prognosis (Liu et al., 2018b). This evidence demonstrates that RNA methylation significantly influences the biological behavior of tumors.

Many studies have explored the regulatory mechanisms among m6A- and m5C-related regulatory factors and various tumors. However, relationships between the clinicopathological characteristics of ccRCC and combined m6A–m5C regulatory factors remain unclear. Therefore, this study combined the gene signatures of m6A with m5C to explore these correlations. We downloaded kidney renal clear cell carcinoma (KIRC) transcriptome and clinical data from The Cancer Genome Atlas (TCGA) to analyze the differentially expressed regulatory factors in ccRCC. Next, we constructed a prognostic risk model

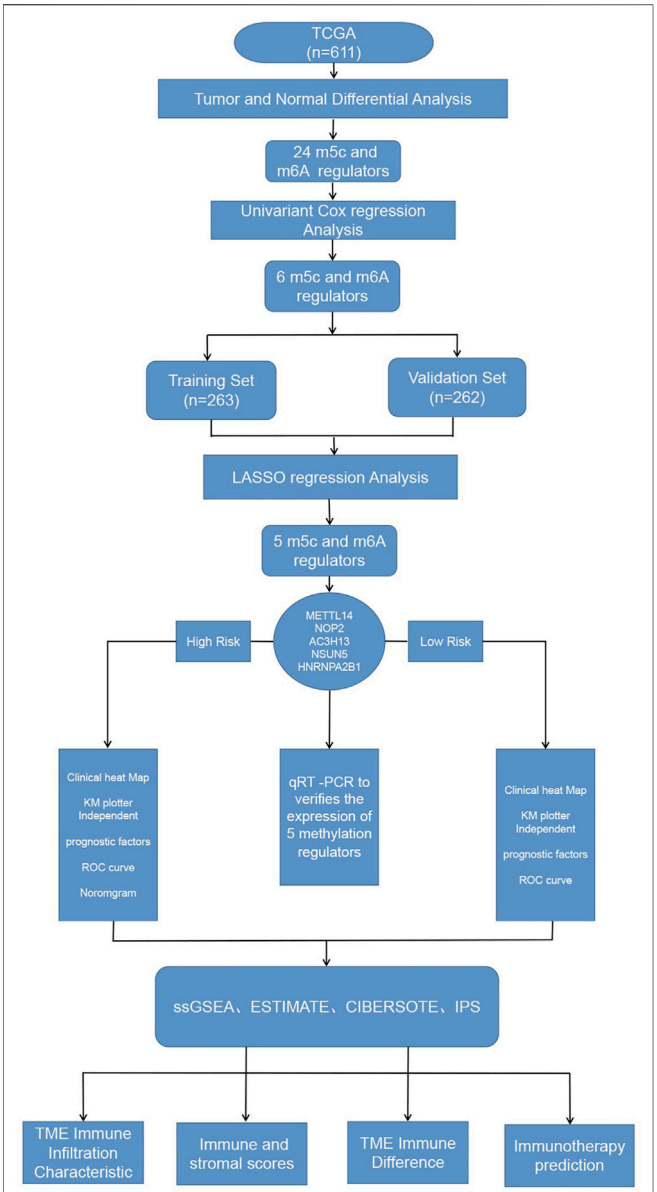


FIGURE 1 | Workflow of this study.

using Lasso regression and Cox analyses. Finally, we screened five prognostic regulatory factors as a model signature and combined the independent prognostic factors to construct a nomogram diagram.

Immune cell infiltration plays a decisive role in tumorigenesis and development. Furthermore, cancer cells shape their microenvironment by secreting various cytokines, chemokines, and other factors, leading to the reprogramming of surrounding cells. Therefore, they play a decisive role in tumor survival and progression (Hinshaw and Shevde, 2019). This study also aimed to identify the potential characteristics of m6A- and m5C-related regulatory factors to improve prognostic ccRCC evaluations.

Overall, we combined m6A and m5C to explore the influence of regulatory factors on ccRCC prognoses to provide new ccRCC biomarkers and construct a reliable prognostic model suitable for use in the clinic.

TABLE 1 | Classification of m6A and m5C combined regulatory factors.

Name	Regulators	Type
m6A	METTL3	writer
	METTL14	writer
	METTL16	writer
	YTHDF1	reader
	YTHDF2	reader
	YTHDF3	reader
	YTHDC1	reader
	YTHDC2	reader
	RBM15	writer
	RBM15 B	writer
	RBMX	reader
	IGFBP1	reader
	IGFBP2	reader
	IGFBP3	reader
	KIAA1429	writer
	FMR1	reader
	LRPPRC	reader
	HNRNPA2B1	reader
	HNRNPC	reader
	ZC3H13	writer
	FTO	eraser
	ALKBH5	eraser
	WTAP	writer
m5C	TET1	eraser
	TET2	eraser
	TET3	writer
	NOP2	writer
	NSUN2	writer
	NSUN3	writer
	NSUN4	writer
	NSUN5	writer
	NSUN6	writer
	ALYREF	reader
	TRDMT1	writer
	YBX1	reader

METHODS

Data Collection and Processing

We downloaded the ccRCC transcriptome (HTseq-FPKM) and clinical data from the TCGA-KIRC database (https://portal.gdc.cancer.gov/). We included 611 TCGA samples; 539 were tumor tissue (ccRCC) samples, and 72 were normal tissues samples. We removed all samples with missing data.

Expression Differences in N^6-methyladenosine and 5-methylcytosine-Related Regulatory Factors

We included 23 m6A and 12 m5C regulatory factors based on the literature (Bohnsack et al., 2019; He et al., 2019; Zhang et al., 2021) (m6A: KIAA1429, WTAP, RBM15, RBM15B, METTL16, METTL3, METTL14, ZC3H13, ALKBH5, FTO, FMRP translational regulator 1 [FMR1], heterogeneous nuclear ribonucleoprotein A2/B1

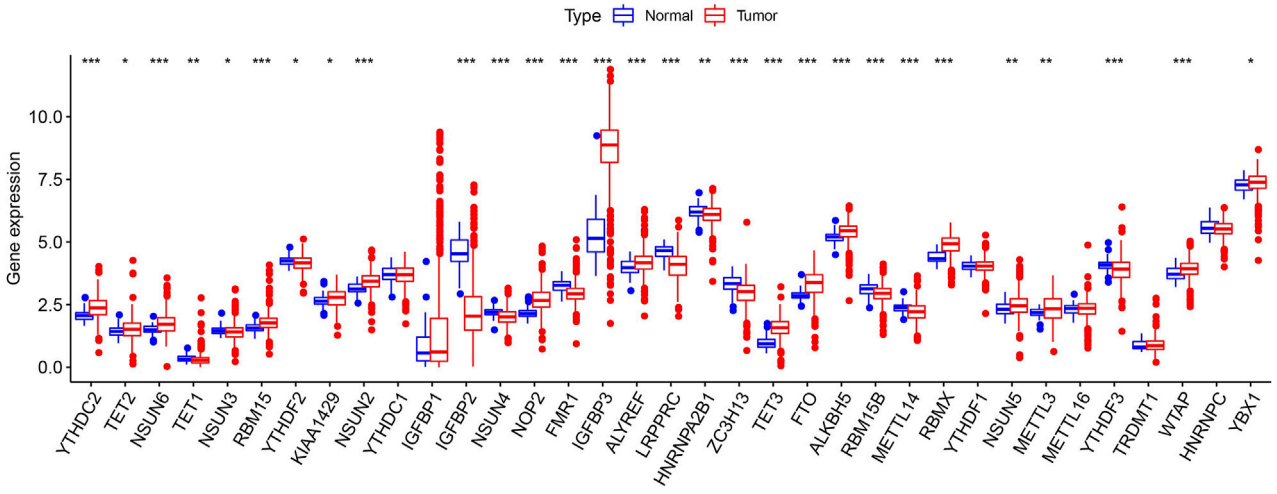


FIGURE 2 | Expression pattern of m6A and m5C regulators in TCGA ccRCC cohort. Difference expression of 35 m6A and m5C methylation regulators. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

[HNRNPA2B1], HNRNPC, IGFBP1, IGFBP2, IGFBP3, leucine-rich pentatricopeptide repeat-containing [LRPPRC], RBMX, YTHDC1, YTHDC2, YTHDF1, YTHDF2, and YTHDF3; m5C: TET1, TET2, TET3, NSUN2, NSUN3, NSUN4, NSUN5, NSUN6, NOP2, ALYREF, tRNA aspartic acid methyltransferase 1 [TRDMT1], and YBX1). We compared the gene expressions of these regulatory factors between the tumor and normal tissue groups. A p-value of <0.05 and an absolute log2 fold change value ≥ 1 were considered statistically significant.

Prognostic Characteristics of N^6-methyladenosine- and 5-methylcytosine-Related Regulatory Factors

We randomly divided the 525 ccRCC samples into training ($n = 263$) and validation ($n = 262$) groups using a 1:1 ratio and screened the prognostic adjustment factors. We calculated the risk score as follows:

$$Riskscore = \left(Exp_{gene1} \times Coef_{gene1}\right) + \left(Exp_{gene2} \times Coef_{gene2}\right) + \cdots + \left(Exp_{gene(n)} \times Coef_{gene(n)}\right),$$

where Exp: individual gene expression; Coef: correlation coefficient.

The samples were classified into high- and low-risk groups based on the median risk score. Next, we used univariate and multivariate Cox regression analyses to determine which pathological features were independent prognostic risk factors. Finally, we prepared receiver operating characteristic (ROC) curves. This process was performed in the training and validation groups.

Nomogram Prognostic Model Construction

We selected clinical indicators with a p-value of ≤0.001 in the multivariate Cox regression analysis and the risk score to construct the prognostic nomogram model. The calibration curve and C-index were used to predict the model's performance. We also created 1-, 3-, and 5-year ROC curves to verify the model's predictive ability and performed a Kaplan–Meier analysis to evaluate the OS of the low- and high-risk groups (statistical significance was set at $p < 0.05$).

Cell Infiltration in the Tumor Microenvironment

We used the single-sample gene set enrichment analysis (ssGSEA) algorithm to obtain the relative abundance of TME-infiltrated cells per sample. Next, we used the CIBERSORT algorithm to analyze differences in immune cell infiltration between the high- and low-risk groups. The CIBERSORT software deconvolves the matrix of immune cell subtypes according to linear support vector regression rules (Lee et al., 2013). We downloaded the ccRCC immune scores from the MD Anderson database (http://bioninformatics.mdanderson.org/estimate/) to assess correlations between regulatory factors and immune and matrix scores.

Cell Culture, Quantitative Real-Time Polymerase Chain Reaction, and Signature Gene Expression Analyses

We purchased human ccRCC (769-P) and immortalized proximal tubule epithelial (HK2) cell lines from the Cell Bank of the Chinese Academy of Sciences. 769-P and HK2 cells were cultured in Roswell Park Memorial Institute (i.e., RPMI)-1,640 medium (KeyGEN Biotech, Inc., Nanjing, China) and Dulbecco's Modified Eagle Medium (i.e., DMEM) (KeyGEN Biotech, Inc.,), respectively, containing 10% fetal bovine serum (Biological Industries, Shanghai, China) at 37°C and 5% carbon dioxide.

We verified the expression levels of the prognostic genes by qRT-PCR analysis. Total RNA was extracted using Trizol reagent (TaKaRa Bio Inc. Shiga, Japan), and complementary DNA was synthesized using PrimeScript RT reagent Kit (TaKaRa, Shiga, Japan). qRT-PCR was performed on 7,500 Real-Time PCR Systems (Applied Biosystems; Thermo Fisher

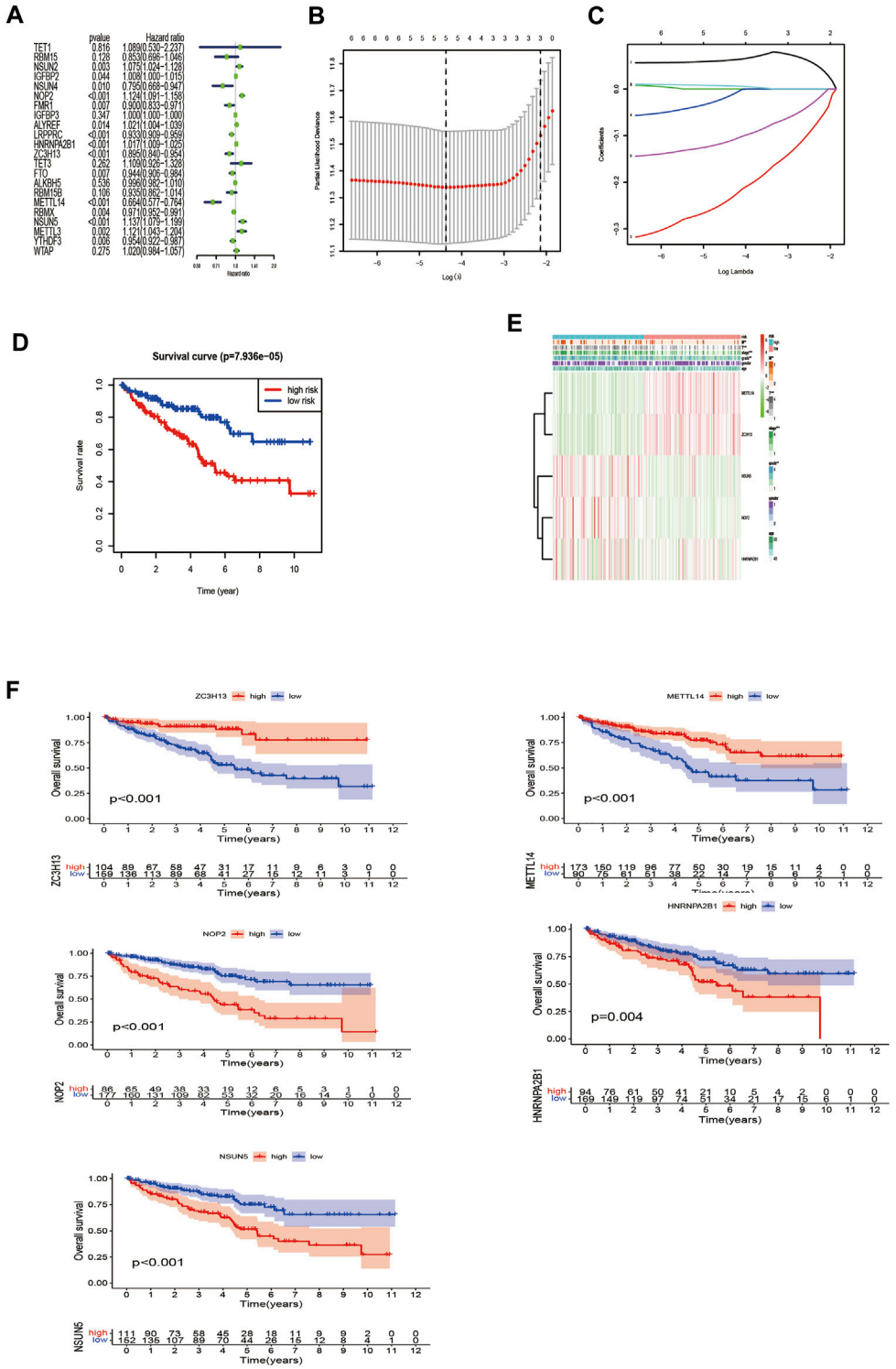


FIGURE 3 | Construction of the risk signature according to the m6A and m5C RNA methylation regulators. **(A)** Forest plot of the univariate Cox regression analysis for the 22 regulators. Identification of six significant regulators (*p < 0.05, **p < 0.01, and ***p < 0.001). **(B,C)** LASSO coefficient profiles of the six regulators. Cross-validation for tuning parameter selection in the LASSO model. **(D)** The K–M analysis showed that patients in the low-risk group presented better OS than those in the high-risk group. This analysis was based on the survival information of samples in the training set. The red line represents the high-risk cluster, whereas the blue line indicates the low-risk cluster. **(E)** The training set of the heat map of the relationship between the gene expression of the corresponding five regulatory factors and clinical features. *p < 0.05, **p < 0.01, and ***p < 0.001. **(F)** Kaplan–Meier survival curves for OS of five regulatory factors.

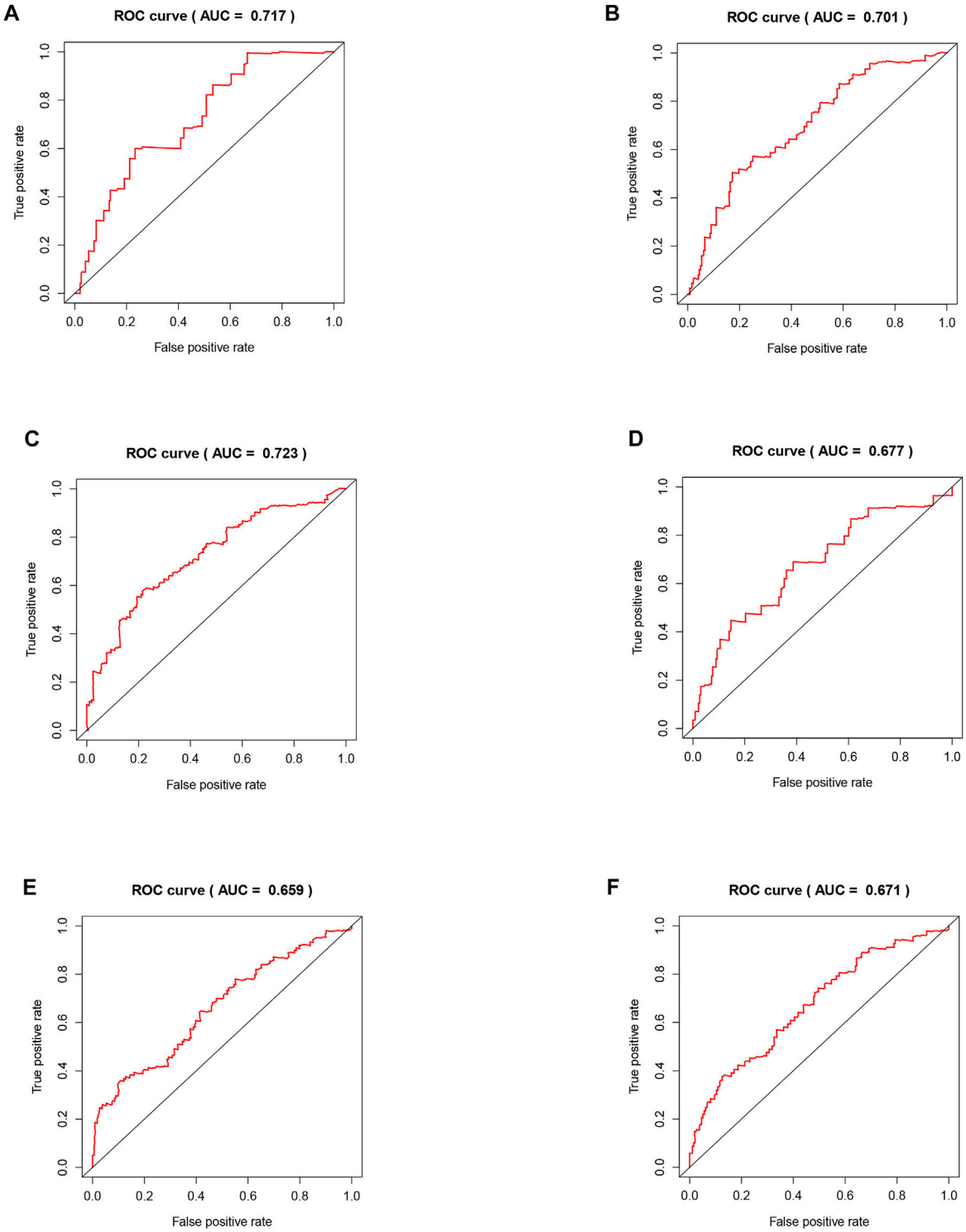


FIGURE 4 | Validation of the risk model related to five regulators. **(A–C)** The training set of the ROC curve for evaluating the prediction efficiency of the prognostic signature. **(D–F)** The validation set of the ROC curve for evaluating the prediction efficiency of the prognostic signature.

Scientific, Waltham, MA, United States) using SYBR GreenER Supermix (TaKaRa, Shiga, Japan). The PCR conditions comprised an initial melting step at 95°C for 1 min, followed by 35 cycles of 95°C for 90 s, 60°C for 30 s, 72°C for 30 s, and then 72°C for 10 min. We used the $2^{-\Delta\Delta Ct}$ method to analyze the relative expression of the prognostic genes based on the normalized relative expression of the β-actin gene. The primers were as follows: NSUN5: Forward,

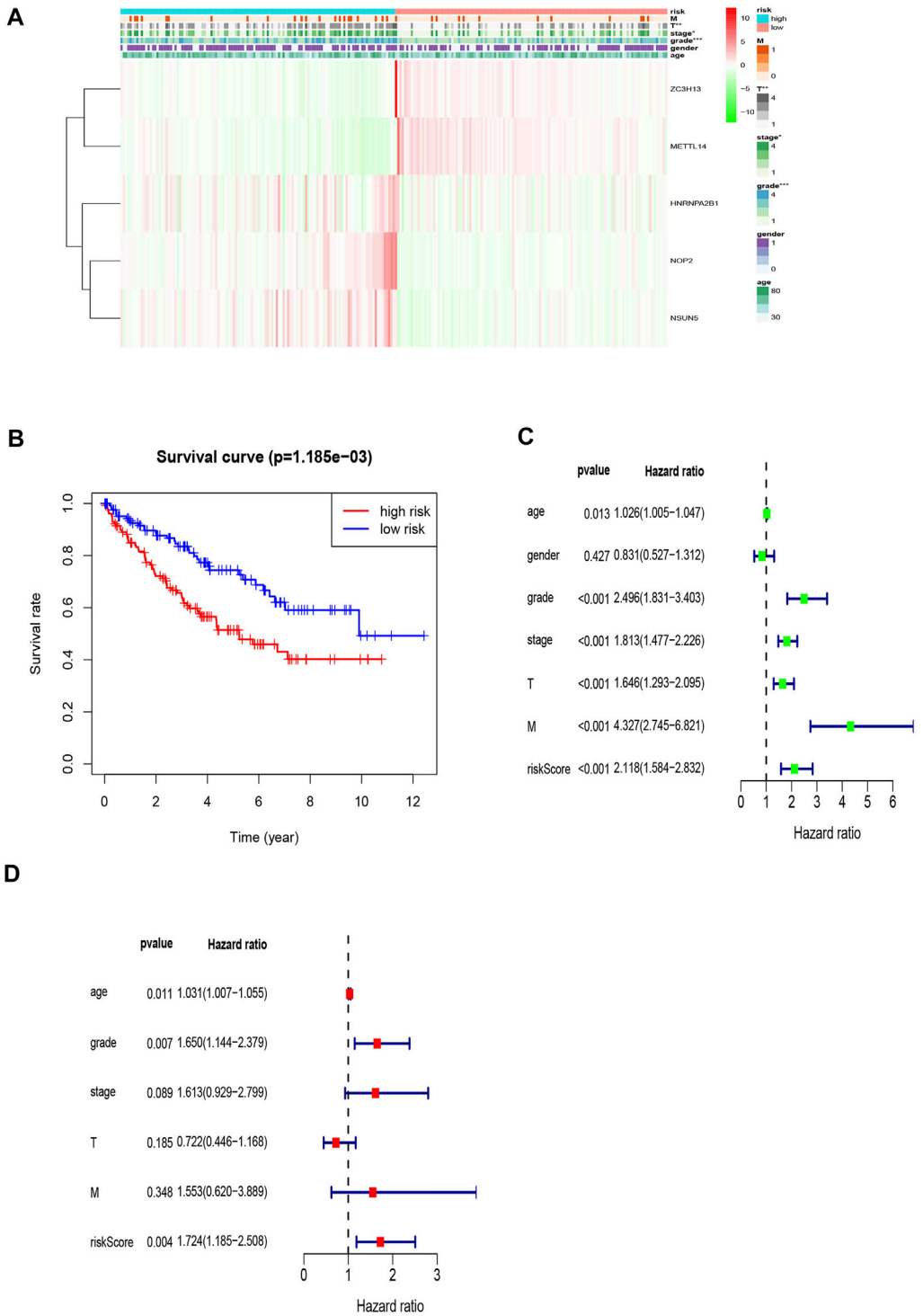


FIGURE 5 | Prognostic signature of the five related regulators in the validation set and the training set of independent prognostic factors for KIRC OS. **(A)** The validation set of the heat map of the relationship between the gene expression of the corresponding five regulatory factors and clinical features. **(B)** Kaplan–Meier survival curves for OS in the two groups of low and high risk. This analysis was based on the survival information of samples in the validation set. **(C,D)** Forrest plot of the independent prognostic factors in KIRC.

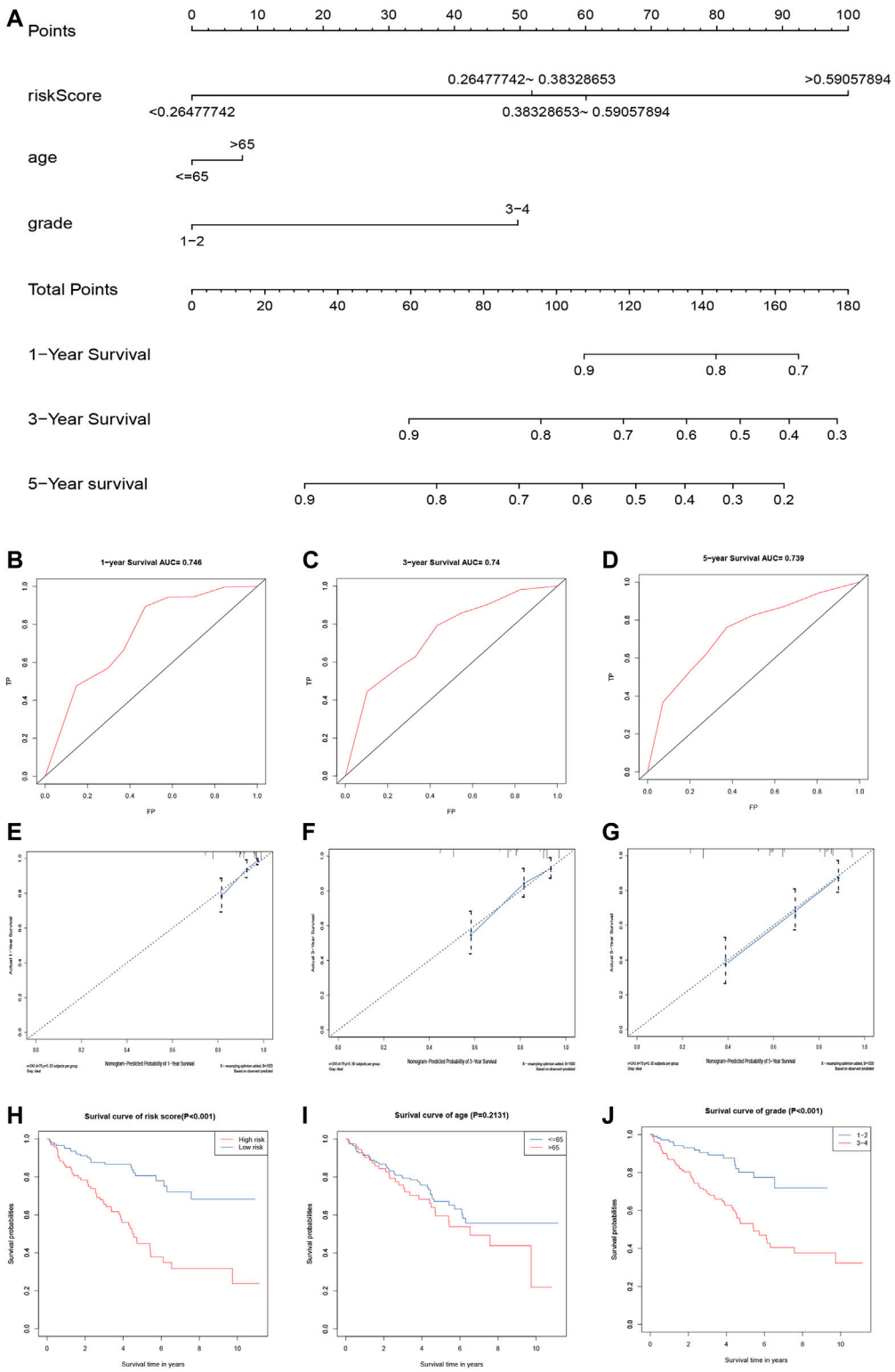


FIGURE 6 | Validation of the prognostic signature of the five related regulators. **(A)** The nomogram of the risk model for predicting the OS probability of ccRCC patients. The whole points projected on the bottom scales indicate the likelihood of 1-, 3-, and 5-year OS. **(B–D)** AUC of the ROC analysis showed the predicted efficacy of the risk model in the training set. **(E–G)** The calibration plot for the nomogram predicting 1-, 3-, and 5-year OS. The *y*-axis indicates the actual survival, as measured by the K–M analysis, while the *x*-axis shows the nomogram-predicted survival. **(H–J)** Kaplan–Meier survival curves stratified according to clinicopathological and risk scores.

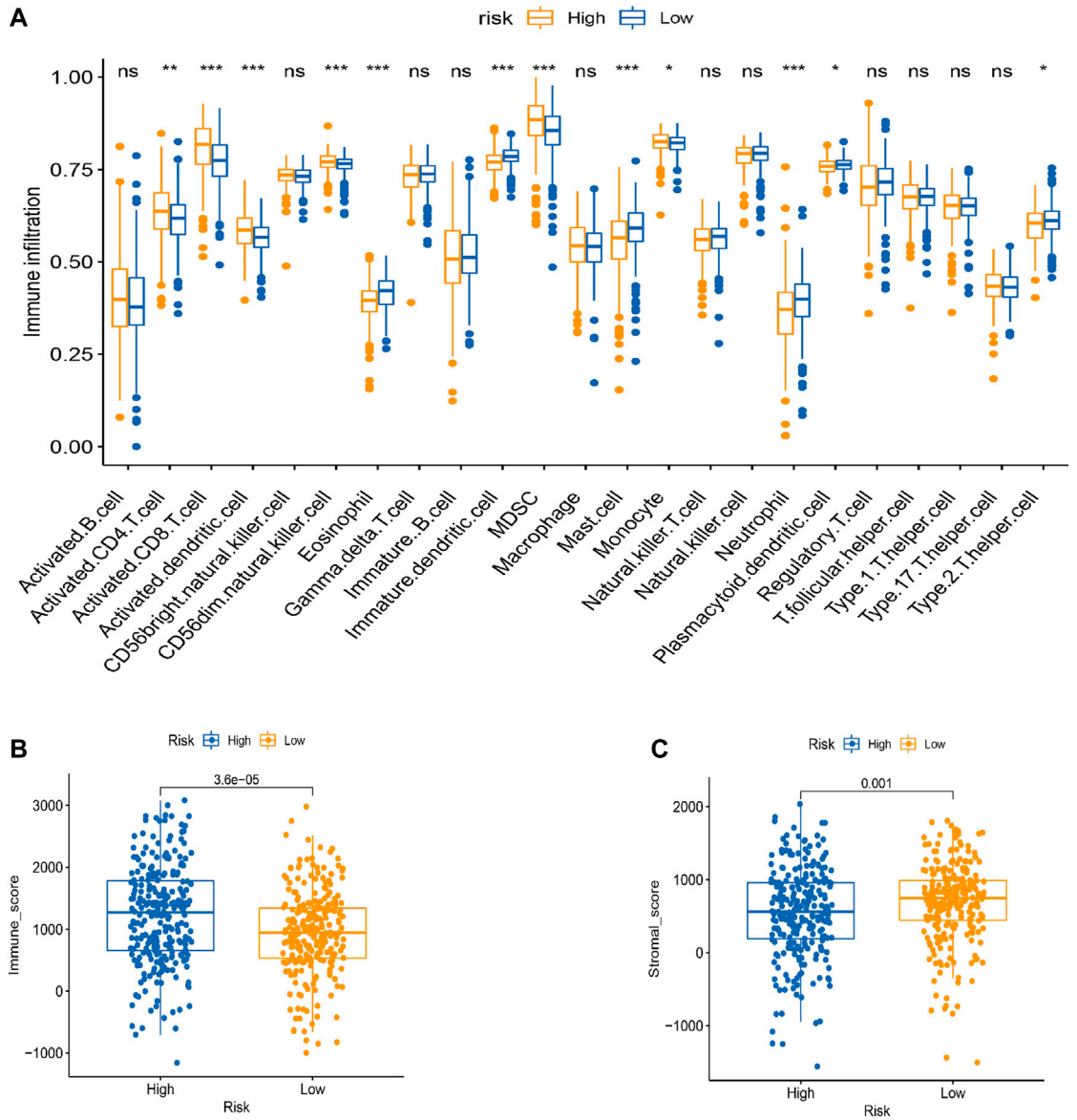


FIGURE 7 | TME cell infiltration characteristics in distinct modification patterns. **(A)** Box plot for the TME cells in distinct risk groups derived from KIRC patients based on the ssGSEA. The asterisks represented the statistical *p* value (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). **(B–C)** Immune and stromal scores within the low- and high-risk groups. **(C)** Effect of LRPPRC expression level on the expression of different immunomodulators. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

TGCCTCGATTTGTGCGTGTG, Reverse, GACAGCTGG CCCTGTCCT; GAPDH: Forward, TGACTTCAACAGCGA CACCCA, Reverse, CACCC-TGTTGCTGTAGCCAAA; ZC3H13: Forward, TGGTGCTGGAGAAGGATACGA, Reverse, CTATCACATCTAAGGGATCTGGCA; HNRNPA2B1: Forward, GCTTTGGGGATTCACGTGGT, Reverse, CCACTGCCATATCCATCAGATCC. We used The Human Protein Atlas database to analyze NSUN5 and HNRNPA2B1 protein expression in clinical specimens.

Statistical Analyses

All statistical analyses were performed using R (version 3.6.1; R Foundation for Statistical Computing, Vienna, Austria). We created box plots to visualize differential gene expression using the "reshape2" and "ggpubr" packages. We screened the prognostic adjustment factors using a Lasso regression analysis and the "glmnet" package and generated ROC curves using the "survivalROC" package. The survival curve was obtained using the "survminer" package, and we used the "rms" package to calculate the C-index and generate

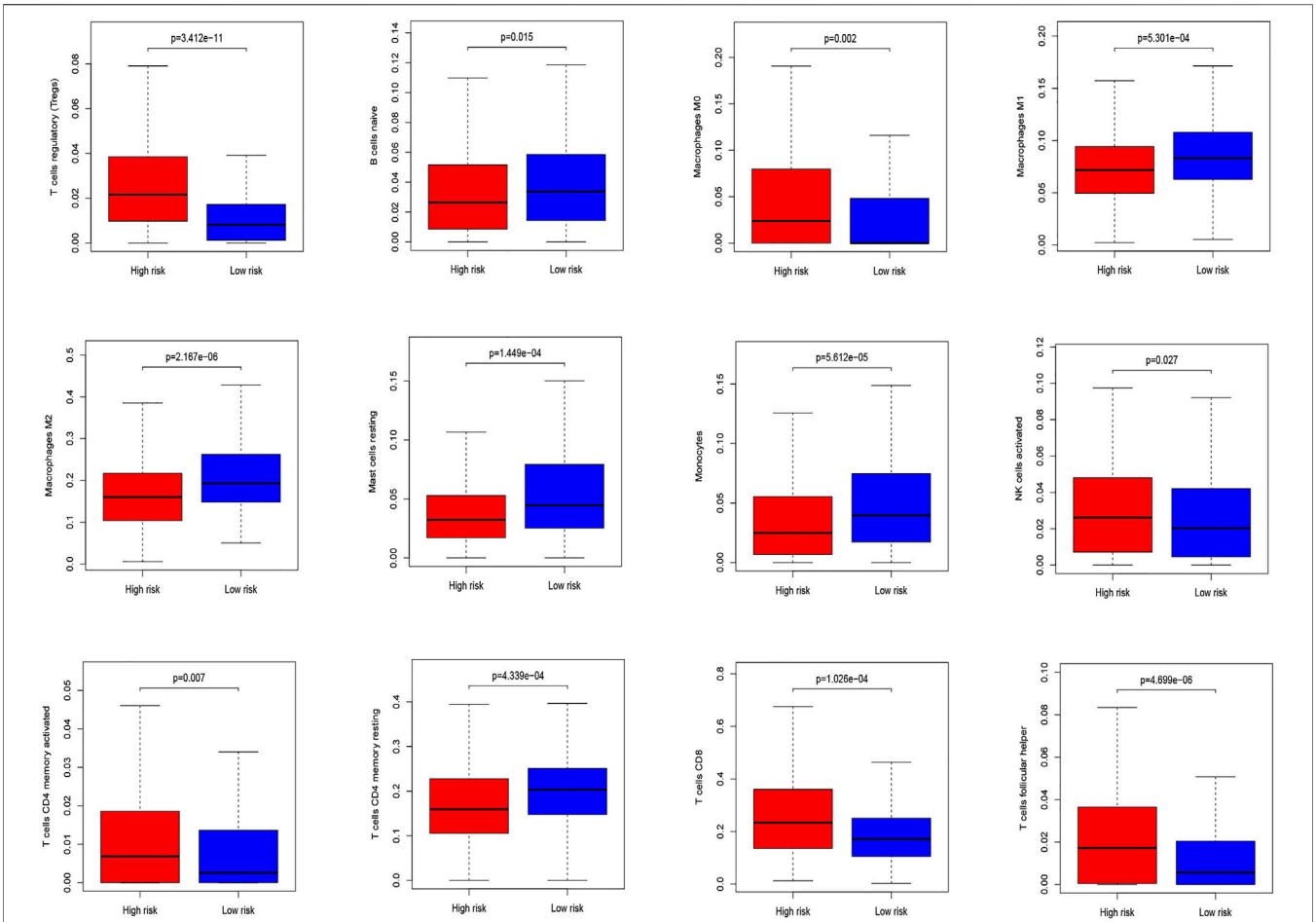


FIGURE 8 | Differential analysis of immune cells in two low-risk groups.

calibration curves. All *p*-values of <0.05 were considered statistically significant. We indicated the various levels of statistical significance as follows: $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$.

RESULTS

Expression Patterns of the N^6-methyladenosine and 5-methylcytosine Regulatory Factors

Figure 1presents a workflow chart, and **Table 1** lists the combined (m6A and m5C) regulatory factors. We identified 24 genes that differed between the tumor and normal tissue groups ($p < 0.01$; **Figure 2**). The expressions of *NSUN4*, *FMR1*, *LRPPRC*, *HNRNPA2B1*, *ZC3H13*, *RBM15B*, *METTL14*, *YTHDF3*, and *IGFBP2* were significantly higher in the normal tissue samples than in the tumor tissue samples. The expressions of *YTHDC2*, *NSUN6*, *RBM15*, *NSUN2*, *NOP2*, *IGFBP3*, *ALYREF*, *TET3*, *FTO*, *ALKBH5*, *RBMX*, *NSUN5*, *METTL3*, and *WTAP* were significantly higher in the tumor tissue samples than in the normal tissue samples. *TET1* expression did not differ among the groups (**Supplementary Table S1**).

Kidney Renal Clear Cell Carcinoma Prognostic Model Combined With Regulatory Factors

We further screened the methylation regulatory factors to explore their prognostic value. We selected 22 regulatory factors as the analysis object (**Figure 3A**). Univariate Cox regression analysis identified six regulatory factors related to OS (**Supplementary Table S2**), and the Lasso regression analysis identified five relevant prognostic factors, namely, *NOP2*, *METTL14*, *NSUN5*, *HNRNPA2B1*, and *ZC3H13* (**Figures 3B,C**; **Supplementary Table S3**). *NSUN5*, *NOP2*, and *HNRNPA2B1* were associated with risk (hazard ratio [HR] >1), and *METTL14* and *ZC3H13* were associated with a protective effect (HR <1). We used these five factors to generate the prognostic risk model.

Next, we compared the survival rates of the high- and low-risk groups based on the median risk. The survival rate of the low-risk group was distinctly better than that of the high-risk group (**Figure 3D**). Furthermore, the heat map illustrated relationships

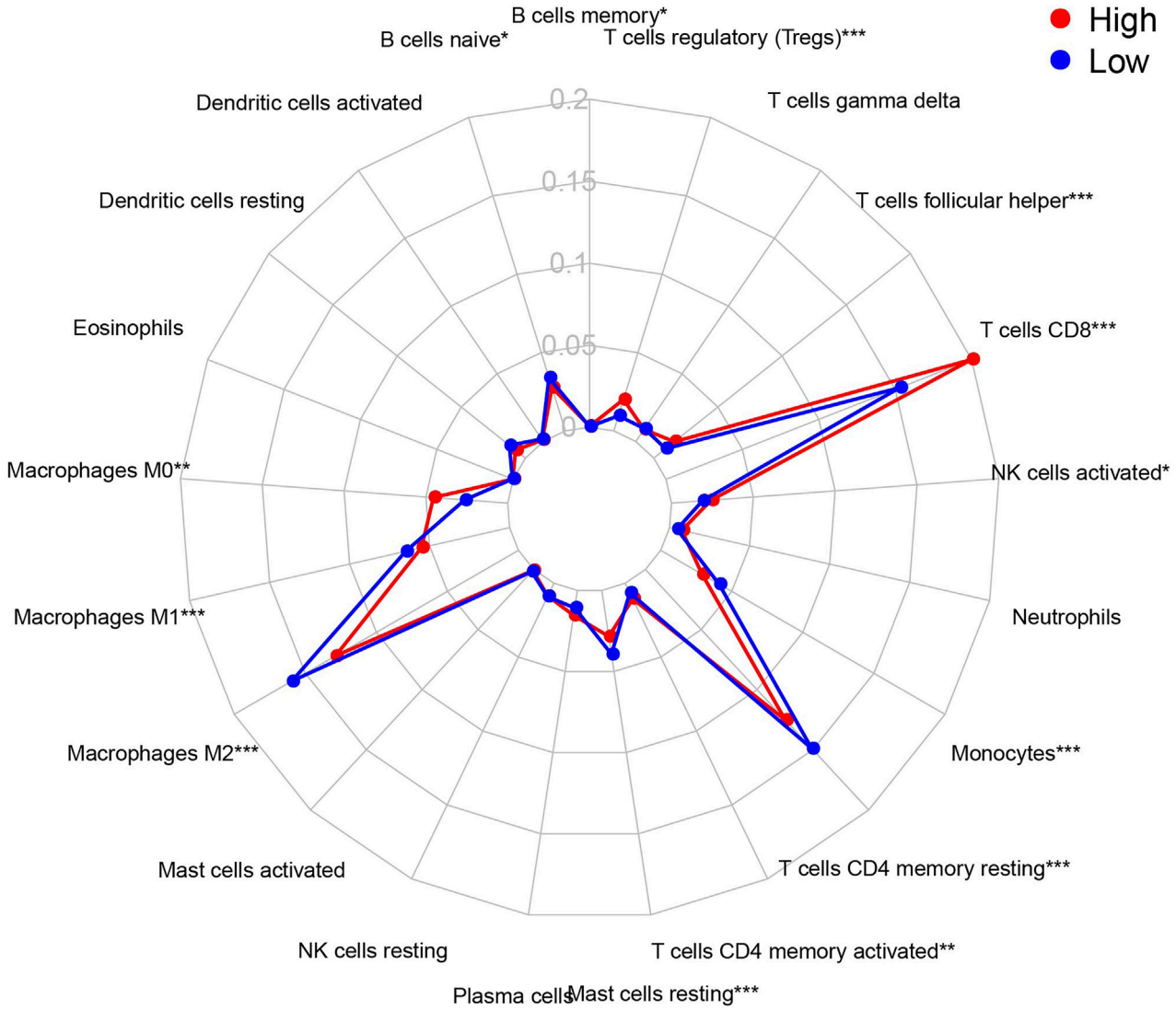


FIGURE 9 | Summary of the 21 immune cells' abundance for different risk groups of training cohort.

among the gene expressions of the five prognostic factors from the training set and various clinical traits and risk scores. The risk score correlated with tumor (T), metastasis (M), stage, and grade. In addition, the highly expressed genes *NOP2*, *HNRNPA2B1*, and *NSUN5* were highly expressed in the high-risk group, and the low-expressed genes *METTL14* and *ZC3H13* were highly expressed in the low-risk group (**Figure 3E**). High expression of *NOP2*, *HNRNPA2B1*, and *NSUN5* correlated with a poor prognosis, and high *METTL14* and *ZC3H13* expression correlated with a good prognosis (**Figure 3F**). These results support the results of our predicted risk genes and protective genes.

Validation of the Risk Model Related to Five Regulators

The areas under the curve (AUCs) for the 1-, 3-, and 5-year ROC curves were 0.717, 0.701, and 0.723, respectively, using the training set data and the prognostic risk model (**Figures 4A–C**). We performed the same analysis using the validation set data, finding that the AUCs for the 1-, 3-, and 5-year ROC curves were 0.677, 0.671, and 0.659, respectively (**Figures 4D–F**). Furthermore, the expression patterns of prognostic regulatory factors in the high- and low-risk groups were almost the same as those in the training set (**Figure 5A**), and the Kaplan–Meier analysis of the validation set was consistent with the training set results. Patient survival was better in the low-risk group than in the high-risk group (**Figure 5B**), indicating that these five regulatory factors positively affect the KIRC prognosis prediction. Next, the univariate analysis identified that prognosis was related to the age, grade, stage, T, M, and risk score. The multivariate analysis showed that the age, grade, and risk score were related to the OS ($p < 0.05$). Therefore, age, grade, and risk score are independent prognostic factors for KIRC OS (**Figures 5C,D**).

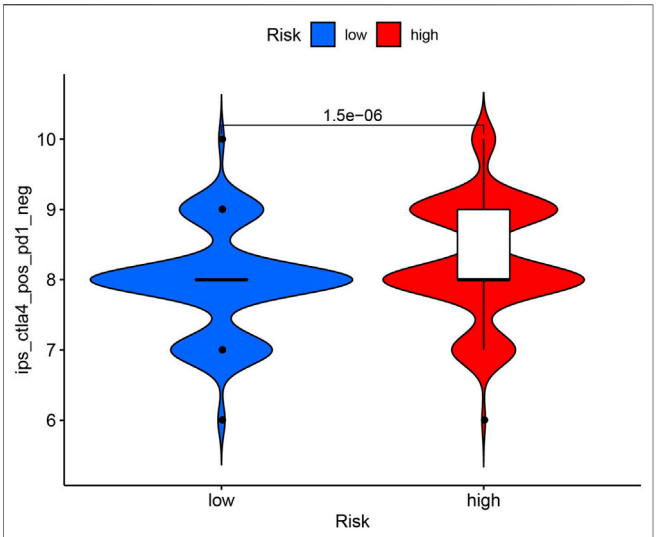


FIGURE 10 | Low-risk group was more likely to respond to anti-CTLA4 immunotherapy. ICPs: immune checkpoints; ICBs: immune checkpoint inhibitors.

Norman Graph Model Construction and Verification

We used the independent prognostic factors (i.e., age, grade, and risk score) to construct a Norman diagram model (**Figure 6A**). Next, we constructed a nomogram based on two independent prognostic factors, and then predicted the 1-, 3-, and 5-year OS using the Norman diagram. The C-index was 0.737, and the calibration curve was more consistent with the observed results. The AUCs for 1-, 3-, and 5-year OS ROC curves were 0.746, 0.740, and 0.739, respectively (**Figures 6B–D**). The calibration graphs displaying the curves illustrate that the nomogram model has a better predictive ability and accuracy (**Figures 6E–G**). The grade, age, and risk score survival curves indicate that the survival rate of patients in the early stage was much better than those in the late stage ($p < 0.001$). Furthermore, OS was better in the low-risk group than in the high-risk group ($p < 0.001$; **Figures 6H–J**).

Effects of N^6-methyladenosine Regulatory Factor Modifications on Immune Cell Infiltration

Tumor infiltration is crucial for tumor development. Therefore, we performed a ssGSEA analysis to evaluate correlation patterns between immune infiltrating cells and the risk score (**Figure 7A**). There was significant infiltration of eosinophils, immature dendritic cells, mast cells, and neutrophils in the low-risk group and abundant infiltration of the activated cluster of differentiation (CD) 4 +, $CD8^+$ T cells, activated dendritic cells, CD56 dim natural killer (NK) cells, and myeloid-derived suppressor cells (MDSCs) in the high-risk group.

Next, we used the estimation of stromal and immune cells in malignant tumor tissues using the expression data (ESTIMATE) algorithm to identify associations between the risk groups and the immune and stromal scores. The immune score was higher in the high-risk group than in the low-risk group, and the stromal score was higher in the low-risk group than in the high-risk group (**Figures 7B,C**).

The CIBERSORT algorithm identified the types of immune cells in ccRCC (**Figure 8**), and we found significant compositional differences between the high- and low-risk groups. M0 macrophages, CD4 memory-activated T cells, CD8 T cells, follicular helper T cells, activated NK cells, and regulatory T cells (Tregs) were significantly enriched in both groups. CD4 resting memory T cells, resting mast cells, M2 and M1 macrophages, monocytes, and naïve B cells were more abundant in the low-risk group than in the high-risk group. Memory B cells did not differ between the groups (**Figure 9**). These results emphasize that immune cell types vary between high- and low-risk groups. Therefore, exploring immune cell infiltration in ccRCC may help elucidate the mechanisms and improve prognosis predictions.

Immune Checkpoints Related to the Regulatory Factors

Immunotherapy is an emerging ccRCC therapy, and the current first-line treatment involves immune checkpoint inhibitors. Cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) and programmed cell death protein 1 (i.e., PD-1) are the two most important immune checkpoint molecules (ICMs) regarding improved OS (Dunn et al., 2002; Ladányi, 2015; Davis et al., 2019). Other immune checkpoints, such as T-cell immunoglobulin and mucin domain–containing protein 3 (i.e., TIM-3), lymphocyte-activation gene 3 (i.e., LAG-3), and T-cell immunoreceptor with Ig and ITIM domains (i.e., TIGIT), suppress the antitumor immune response (Anderson et al., 2016). Therefore, clarifying the expression correlations of ICMs among risk groups may improve the clinical application of immune checkpoint inhibitors in KIRC. In our study, the immunotherapy score was higher in the high-risk group than in the low-risk group, suggesting that the high-risk group may respond to anti-CTLA-4 treatment ($p < 0.001$; **Figure 10**).

NSUN5 and *HNRNPA2B1* Expression in Kidney Renal Clear Cell Carcinoma

We verified the expression of *NSUN5*, *ZC3H13*, *METL14*, *NOP2*, and *HNRNPA2B1* in ccRCC using the HK2 (epithelial; control) and 769-P (ccRCC) cell lines. *METL14* is under-expressed in ccRCC tissues (Wang et al., 2021), and *NOP2* expression is higher in tumor tissues than in normal tissues (Li et al., 2021). Furthermore, *ZC3H13* expression is low in tumor tissues (Guo et al., 2021). Based on this, we compared the NSUN5 and HNRNPA2B1 protein levels of ccRCC and normal kidney tissues to determine the *NSUN5* and *HNRNPA2B1* expression status. *NSUN5* and *HNRNPA2B1* expressions were higher in tumor tissues than in normal tissues (**Figure 11A**), consistent with our other findings. We also compared NSUN5 and HNRNPA2B1 protein levels between RCC and normal kidney tissues, finding more NSUN5 and HNRNPA2B1 staining in RCC tissues than in normal kidney tissues (**Figure 11B**).

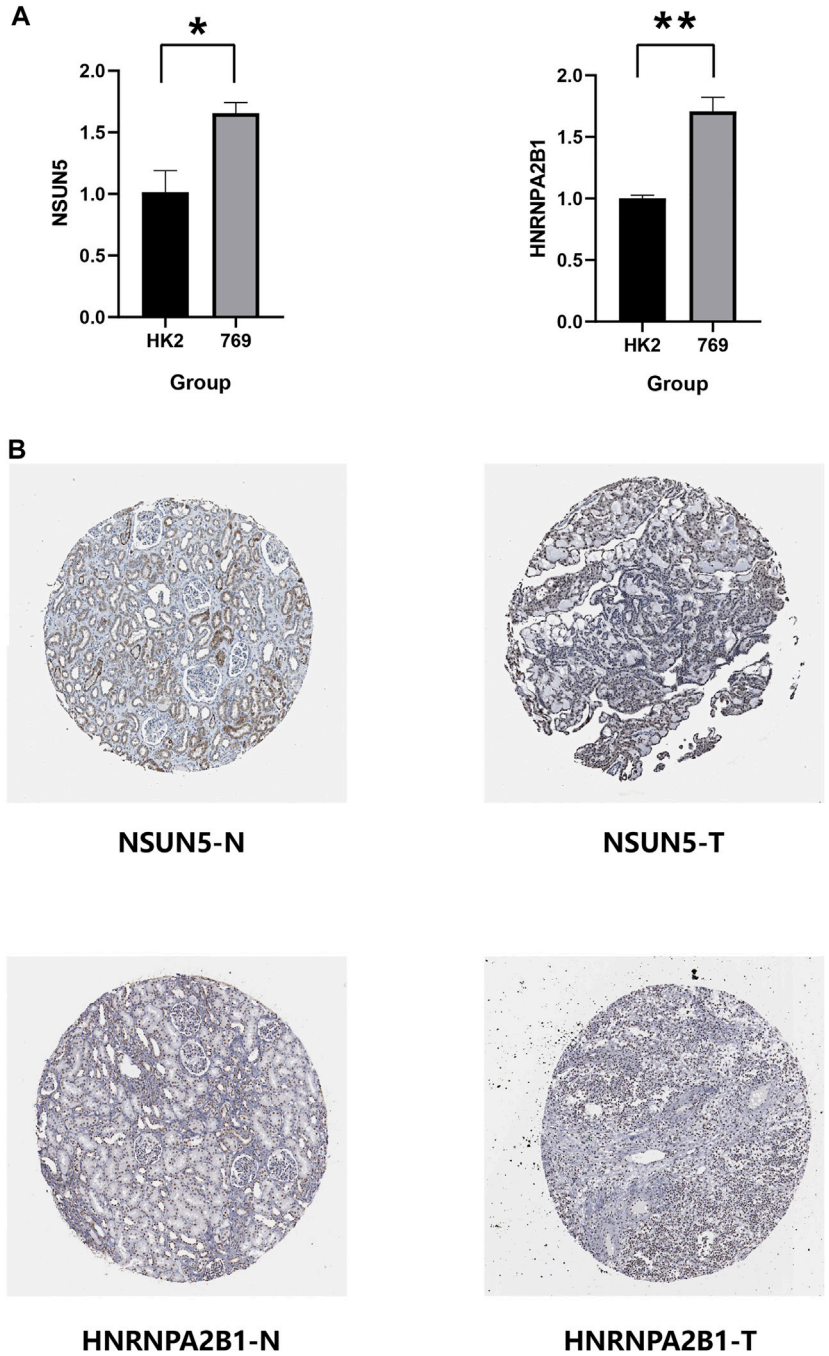


FIGURE 11 | Expression level of NSUN5 and HNRNPA2B1 in renal clear cell carcinoma (KIRC)**. (A)** The expression of NSUN5 and HNRNPA2B1 in ccRCC cell lines (HK2,769) was detected by the qRT-PCR assay. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$. **(B)** Protein expression of NSUN5 and HNRNPA2B1 in KIRC. KIRC, kidney renal clear cell carcinoma.

DISCUSSION

ccRCC is the third most common renal cancer, accounting for 3% of all adult tumors, and the most common sporadic RCC subtype (Padala and Kallam, 2022). Studies have demonstrated that inhibiting m6A and m5C regulatory factors may have therapeutic benefits for tumors, providing a new direction for tumor treatment. For example, R-2-hydroxyglutarate (R-2HG) sensitivity of leukemia increases with increasing the m6A levels (Su et al., 2018). Furthermore, aurora kinase B (i.e., AURKB) regulates NSUN2 at the protein level, phosphorylated by Ser-139 (Sakita-Suto et al., 2007). To date, the relevant ccRCC markers are insufficient for clinical diagnosis and prognosis

because ccRCC is often regulated by multiple genes. Thus, a single prognostic factor cannot accurately predict the clinical prognosis. This study explored the key regulators in ccRCC to provide new avenues for cancer treatment by identifying relevant regulatory factors and developing a new gene signature for more accurate prognosis predictions.

Our main goal was to clarify the relationship between m6A and m5C regulatory factors and determine how these combined factors affect KIRC prognosis and immune infiltrating cells. We used data from the TCGA database, identifying 35 relevant m6A and m5C regulatory factors, most of which were abnormally expressed in KIRC. Using these factors, we constructed a new prognostic model for more accurate prediction of OS in patients with KIRC. We randomly divided the TCGA dataset into two groups (training and validation) to verify the effectiveness of the risk model and introduced clinicopathological factors to improve the model's reliability.

Immune infiltration analysis found more M0 macrophages, CD4 memory-activated T cells, CD8 T cells, follicular helper T cells, activated NK cells, and Tregs in the high-risk group. The low-risk group had more resting CD4 memory T cells, resting mast cells, M1 and M2 macrophages, monocytes, and naïve B cells. Increased eosinophils may be related to a good prognosis, similar to gastric cancer (Iwasaki et al., 1986). However, it could also indicate a poor prognosis, similar to bladder cancer (Popov et al., 2018). Increased number of mast cells are associated with a poor prognosis in lung, colorectal, gastric, and cervical cancers, and melanoma, but it is associated with a good prognosis in breast and prostate cancers (Rajput et al., 2008; Fleischmann et al., 2009; Kormelink et al., 2009). In addition, $CD4^+$ T cells promote renal cancer cell proliferation by regulating YBX1 (Wang et al., 2018), and MDSCs accumulate in various tumors, promoting vascular survival and improving tumor immunity (De Cicco et al., 2020). Increased follicular helper T cells and Treg have been shown to promote cancer progression, which relates to a poor prognosis (Finotello and Trajanoski, 2017; Long et al., 2019). The role of mast and dendritic cells and their relationship with renal angiogenesis in KIRC remains unclear (Tamma et al., 2019).

Several studies have suggested that M2 and a small subset of M1 macrophages cannot phagocytose tumor cells and help tumor cells escape death and spread to other tissues and organs. These cells are called tumor-associated macrophages (TAMs) (Zhou et al., 2020), and they promote cancer progression and metastasis in human renal cell carcinoma, stimulating tumor inflammation (Hutterer et al., 2013). TAMs have also been shown to promote tumor metastasis, occurrence, and vascular lymphangiogenesis. During the initial stages of tumor development, macrophages either directly promote antitumor responses by killing tumor cells or indirectly recruit and activate other immune cells (Lopez-Yrigoyen et al., 2021). In our study, there were more M0 macrophages in the high-risk group, consistent with other clinical reports (Yi et al., 2021). Previous studies have shown that low-risk groups have more M1 and M2 macrophages and fewer M0 macrophages than high-risk groups. These descriptions are also consistent with our study's results. Furthermore, the prognostic outcome of the low-risk group is consistent with the previous survival advantage. The specific role of NK cells remains controversial and largely depends on the cancer type (Hinshaw and Shevde, 2019).

Finally, we used the CIBERSORT and ESTIMATE algorithms to generate immune and stromal scores per risk group. The immune score was higher in the high-risk group than in the low-risk group, and the stromal score was higher in the low-risk group than in the high-risk group. Furthermore, the high-risk group was sensitive to anti-CTLA-4. Liu et al. reported that CTLA-4 as an oncogene accelerates ccRCC development with high prognostic value (Liu et al., 2020).

However, our research has limitations. First, we only have internal verification and lack external platform verification. Therefore, the predictive power of the risk model may be limited. Also, additional basic experimental research is needed to determine the detailed mechanisms of these five regulators.

CONCLUSION

Our study evaluated the combined role of m6A and m5C regulatory factors in regulating the KIRC TME. First, we constructed a new gene signature with five relevant regulatory factors. Then, using this signature, we created a novel prognostic model to stratify KIRC and normal tissue samples based on risk. Finally, we explored the link between the new gene signature and immune infiltration and obtained new potential immune checkpoints. These tools may help clinicians make more personalized and accurate prognosis predictions.

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

LL: writing—original draft. ZT: conceptualization, methodology, and software. ML: validation, formal analysis. YZ: visualization. JZ: resources, data curation. ZL: writing—review and editing. XC: funding acquisition, project administration, and supervision.

FUNDING

This work was supported by the joint plan of key research and development program of Liaoning Province under Grant No. 2020JH 2/10300137 and 345 Talent Project under Grant No. M0716. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We thank XC, who serves for the Department of Urology, Shengjing Hospital of China Medical University, for his strong guidance and support for this review. He has put forward valuable suggestions for the improvement of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.864383/full#supplementary-material

Supplementary Table S1 | Different expression of 35 RNA methylation regulators.

Supplementary Table S2 | Multivariate Cox regression on the 24 selected regulatory factors.

Supplementary Table S3 | Five prognostic factors.

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, 989–1004. doi:10.1016/j.immuni.2016.05.001

Bohnsack, K. E., Höbartner, C., and Bohnsack, M. T. (2019). Eukaryotic 5-methylcytosine (m^5C) RNA Methyltransferases: Mechanisms, Cellular Functions, and Links to Disease. *Genes (Basel)* 10, 102. doi:10.3390/genes10020102

Cai, X., Wang, X., Cao, C., Gao, Y., Zhang, S., Yang, Z., et al. (2018). HBXIP-elevated Methyltransferase METTL3 Promotes the Progression of Breast Cancer *via* Inhibiting Tumor Suppressor Let-7g. *Cancer Lett.* 415, 11–19. doi:10.1016/j.canlet.2017.11.018

Chellamuthu, A., and Gray, S. G. (2020). The RNA Methyltransferase NSUN2 and its Potential Roles in Cancer. *Cells 9* 9, 1758. doi:10.3390/cells9081758

Chen, M., Wei, L., Law, C.-T., Tsang, F. H.-C., Shen, J., Cheng, C. L.-H., et al. (2018). RNA N6-Methyladenosine Methyltransferase-like 3 Promotes Liver Cancer Progression through YTHDF2-dependent Posttranscriptional Silencing of SOCS2. *Hepatology* 67, 2254–2270. doi:10.1002/hep.29683

Chen, X.-Y., Zhang, J., and Zhu, J.-S. (2019). The Role of m6A RNA Methylation in Human Cancer. *Mol. Cancer* 18, 103. doi:10.1186/s12943-019-1033-z

Davis, E. J., Perez, M. C., Ayoubi, N., Zhao, S., Ye, F., Wang, D. Y., et al. (2019) Clinical Correlates of Response to Anti-PD-1-based Therapy in Patients with Metastatic Melanoma, *J. Immunother.* 42, 221–227.doi:10.1097/cji.0000000000000258

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

Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J., and Schreiber, R. D. (2002). Cancer Immunoediting: from Immunosurveillance to Tumor Escape. *Nat. Immunol.* 3, 991–998. doi:10.1038/ni1102-991

Finotello, F., and Trajanoski, Z. (2017). New Strategies for Cancer Immunotherapy: Targeting Regulatory T Cells. *Genome Med.* 9, 10. doi:10.1186/s13073-017-0402-8

Fleischmann, A., Schlomm, T., Köllermann, J., Sekulic, N., Huland, H., Mirlacher, M., et al. (2009). Immunological Microenvironment in Prostate Cancer: High Mast Cell Densities Are Associated with Favorable Tumor Characteristics and Good Prognosis. *Prostate* 69, 976–981. doi:10.1002/pros.20948

Frye, M., and Watt, F. M. (2006). The RNA Methyltransferase Misu (NSun2) Mediates Myc-Induced Proliferation and Is Upregulated in Tumors. *Curr. Biol.* 16, 971–981. doi:10.1016/j.cub.2006.04.027

Greef, B., and Eisen, T. (2016). Medical Treatment of Renal Cancer: New Horizons. *Br. J. Cancer* 115, 505–516. doi:10.1038/bjc.2016.230

Guo, T., Duan, H., Chen, J., Liu, J., Othmane, B., Hu, J., et al. (2021). N6-Methyladenosine Writer Gene ZC3H13 Predicts Immune Phenotype and Therapeutic Opportunities in Kidney Renal Clear Cell Carcinoma. *Front. Oncol.* 11, 718644. doi:10.3389/fonc.2021.718644

He, L., Li, H., Wu, A., Peng, Y., Shu, G., and Yin, G. (2019). Functions of N6-Methyladenosine and its Role in Cancer. *Mol. Cancer* 18, 176. doi:10.1186/s12943-019-1109-9

Hinshaw, D. C., and Shevde, L. A. (2019). The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer Res.* 79, 4557–4566. doi:10.1158/0008-5472.can-18-3962

Huang, H., Weng, H., and Chen, J. (2020). m6A Modification in Coding and Non-coding RNAs: Roles and Therapeutic Implications in Cancer. *A Modification Coding Non-coding RNAs: Roles Ther. Implications CancerCancer Cel.* 37, 270–288. doi:10.1016/j.ccell.2020.02.004

Huang, T., Chen, W., Liu, J., Gu, N., and Zhang, R. (2019). Genome-wide Identification of mRNA 5-methylcytosine in Mammals. *Nat. Struct. Mol. Biol.* 26, 380–388. doi:10.1038/s41594-019-0218-x

Hutterer, G. C., Pichler, M., Chromecki, T. F., Strini, K. A., Klatte, T., Pummer, K., et al. (2013). Tumour-associated Macrophages Might Represent a Favourable Prognostic Indicator in Patients with Papillary Renal Cell Carcinoma. *Histopathology* 63, 309–315. doi:10.1111/his.12163

Iwasaki, K., Torisu, M., and Fujimura, T. (1986). Malignant Tumor and Eosinophils: I. Prognostic Significance in Gastric Cancer. *Cancer* 58, 1321–1327. doi:10.1002/1097-0142(19860915)58:6<1321::aid-cncr2820580623>3.0.co;2-o

Jiang, Z., Li, S., Han, M. J., Hu, G. M., and Cheng, P. (2020). Erratum: High Expression of NSUN5 Promotes Cell Proliferation *via* Cell Cycle Regulation in Colorectal Cancer. *Am. J. Transl Res.* 12, 6976–3870.

Jonasch, E., Walker, C. L., and Rathmell, W. K. (2021). Clear Cell Renal Cell Carcinoma Ontogeny and Mechanisms of Lethality. *Nat. Rev. Nephrol.* 17, 245–261. doi:10.1038/s41581-020-00359-2

Kormelink, T., Abudukelimu, A., and Redegeld, F. (2009). Mast Cells as Target in Cancer Therapy. *Cpd* 15, 1868–1878. doi:10.2174/138161209788453284

Ladányi, A. (2015). Prognostic and Predictive Significance of Immune Cells Infiltrating Cutaneous Melanoma. *Pigment Cel. Melanoma Res.* 28, 490–500. doi:10.1111/pcmr.12371

Lan, Q., Liu, P. Y., Haase, J., Bell, J. L., Hüttelmaier, S., and Liu, T. (2019). The Critical Role of RNA m6A Methylation in Cancer. *Cancer Res.* 79, 1285–1292. doi:10.1158/0008-5472.can-18-2965

Lee, Y.-Y., Kim, T.-J., Kim, J.-Y., Choi, C. H., Do, I.-G., Song, S. Y., et al. (2013). Genetic Profiling to Predict Recurrence of Early Cervical Cancer. *Gynecol. Oncol.* 131, 650–654. doi:10.1016/j.ygyno.2013.10.003

Li, E., Wei, B., Wang, X., and Kang, R. (2020). METTL3 Enhances Cell Adhesion through Stabilizing Integrin β1 mRNA *via* an m6A-HuR-dependent Mechanism in Prostatic Carcinoma. *Am. J. Cancer Res.* 10, 1012–1025.

Li, H., Jiang, H., Huang, Z., Chen, Z., and Chen, N. (2021). Prognostic Value of an m5C RNA Methylation Regulator-Related Signature for Clear Cell Renal Cell Carcinoma. *Cmar* Vol. 13, 6673–6687. doi:10.2147/cmar.s323072

Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R. I. (2016). The M 6 A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol. Cel.* 62, 335–345. doi:10.1016/j.molcel.2016.03.021

Liu, J., Eckert, M. A., Harada, B. T., Liu, S.-M., Lu, Z., Yu, K., et al. (2018a). m6A mRNA Methylation Regulates AKT Activity to Promote the Proliferation and

Tumorigenicity of Endometrial Cancer. *Nat. Cel Biol* 20, 1074–1083. doi:10.1038/s41556-018-0174-4

Liu, J., Ren, D., Du, Z., Wang, H., Zhang, H., and Jin, Y. (2018b). m 6 A Demethylase FTO Facilitates Tumor Progression in Lung Squamous Cell Carcinoma by Regulating MZF1 Expression. *Biochem. biophysical Res. Commun.* 502, 456–464. doi:10.1016/j.bbrc.2018.05.175

Liu, S., Wang, F., Tan, W., Zhang, L., Dai, F., Wang, Y., et al. (2020). CTLA4 Has a Profound Impact on the Landscape of Tumor-Infiltrating Lymphocytes with a High Prognosis Value in clear Cell Renal Cell Carcinoma (ccRCC). *Cancer Cel Int* 20, 519. doi:10.1186/s12935-020-01603-2

Long, J., Wang, A., Bai, Y., Lin, J., Yang, X., Wang, D., et al. (2019). Development and Validation of a TP53-Associated Immune Prognostic Model for Hepatocellular Carcinoma. *EBioMedicine* 42, 363–374. doi:10.1016/j.ebiom.2019.03.022

Lopez-Yrigoyen, M., Cassetta, L., and Pollard, J. W. (2021). Macrophage Targeting in Cancer. *Ann. N. Y Acad. Sci.* 1499, 18–41. doi:10.1111/nyas.14377

Ma, J. Z., Yang, F., Zhou, C. C., Liu, F., Yuan, J. H., Wang, F., et al. (2017). METTL14 Suppresses the Metastatic Potential of Hepatocellular Carcinoma by Modulating N 6 -methyladenosine-dependent Primary MicroRNA Processing. *Hepatology* 65, 529–543. doi:10.1002/hep.28885

Nombela, P., Miguel-López, B., and Blanco, S. (2021). The Role of m6A, m5C and Ψ RNA Modifications in Cancer: Novel Therapeutic Opportunities. *Mol. Cancer* 20, 18. doi:10.1186/s12943-020-01263-w

Padala, S. A., and Kallam, A. (2022) "Clear Cell Renal Carcinoma," in *StatPearls* (StatPearls Publishing; Treasure Island (FL)). Copyright © 2022

Popov, H., Donev, I. S., and Ghenev, P. (2018). Quantitative Analysis of Tumor-Associated Tissue Eosinophilia in Recurring Bladder Cancer. *Cureus* 10, e3279. doi:10.7759/cureus.3279

Rajput, A. B., Turbin, D. A., Cheang, M. C., Voduc, D. K., Leung, S., Gelmon, K. A., et al. (2008). Stromal Mast Cells in Invasive Breast Cancer Are a Marker of Favourable Prognosis: a Study of 4,444 Cases. *Breast Cancer Res. Treat.* 107, 249–257. doi:10.1007/s10549-007-9546-3

Sakita-Suto, S., Kanda, A., Suzuki, F., Sato, S., Takata, T., and Tatsuka, M. (2007). Aurora-B Regulates RNA Methyltransferase NSUN2. *MBoC* 18, 1107–1117. doi:10.1091/mbc.e06-11-1021

Sato, G., Saijo, Y., Uchiyama, B., Kumano, N., Sugawara, S.-i., Fujimura, S., et al. (1999). Prognostic Value of Nucleolar Protein P120 in Patients with Resected Lung Adenocarcinoma. *Jco* 17, 2721. doi:10.1200/jco.1999.17.9.2721

Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer Statistics, 2018. *Cancer J. clinicians* 68, 7–30. doi:10.3322/caac.21442

Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., et al. (2012). Widespread Occurrence of 5-methylcytosine in Human Coding and Non-coding RNA. *Nucleic Acids Res.* 40, 5023–5033. doi:10.1093/nar/gks144

Su, R., Dong, L., Li, C., Nachtergaele, S., Wunderlich, M., Qing, Y., et al. (2018). R-2HG Exhibits Anti-tumor Activity by Targeting FTO/m6A/MYC/CEBPA Signaling. *Cell* 172, 90–105. e123. doi:10.1016/j.cell.2017.11.031

Tamma, R., Rutigliano, M., Lucarelli, G., Annese, T., Ruggieri, S., Cascardi, E., et al. (2019) Microvascular Density, Macrophages, and Mast Cells in Human clear Cell Renal Carcinoma with and without Bevacizumab Treatment, *Urol. Oncol.* 37, 355.e11-355.e19.doi:10.1016/j.urolonc.2019.01.025

Trixl, L., and Lusser, A. (2019). The Dynamic RNA Modification 5-methylcytosine and its Emerging Role as an Epitranscriptomic Mark. *Wiley Interdiscip. Rev. RNA*, 10, e1510. doi:10.1002/wrna.1510

Wang, Y., Cong, R., Liu, S., Zhu, B., Wang, X., and Xing, Q. (2021). Decreased Expression of METTL14 Predicts Poor Prognosis and Construction of a Prognostic Signature for clear Cell Renal Cell Carcinoma. *Cancer Cel Int* 21, 46. doi:10.1186/s12935-020-01738-2

Wang, Y., Wang, Y., Xu, L., Lu, X., Fu, D., Su, J., et al. (2018). CD4 + T Cells Promote Renal Cell Carcinoma Proliferation *via* Modulating YBX1. *Exp. Cel. Res.* 363, 95–101. doi:10.1016/j.yexcr.2017.12.026

Weng, H., Huang, H., Wu, H., Qin, X., Zhao, B. S., Dong, L., et al. (2018) METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis *via* mRNA m6A Modification, *Cell stem cell* 22, 191–205.e199.doi:10.1016/j.stem.2017.11.016

Yang, Z., Li, J., Feng, G., Gao, S., Wang, Y., Zhang, S., et al. (2017). MicroRNA-145 Modulates N6-Methyladenosine Levels by Targeting the 3′-Untranslated mRNA Region of the N6-Methyladenosine Binding YTH Domain Family 2 Protein. *J. Biol. Chem.* 292, 3614–3623. doi:10.1074/jbc.m116.749689

Yi, M., Li, A., Zhou, L., Chu, Q., Luo, S., and Wu, K. (2021). Immune Signature-Based Risk Stratification and Prediction of Immune Checkpoint Inhibitor's Efficacy for Lung Adenocarcinoma. *Cancer Immunol. Immunother.* 70, 1705–1719. doi:10.1007/s00262-020-02817-z

Zhang, Q., Liu, F., Chen, W., Miao, H., Liang, H., Liao, Z., et al. (2021). The Role of RNA m5C Modification in Cancer Metastasis. *Int. J. Biol. Sci.* 17, 3369–3380. doi:10.7150/ijbs.61439

Zhou, J., Tang, Z., Gao, S., Li, C., Feng, Y., and Zhou, X. (2020). Tumor-Associated Macrophages: Recent Insights and Therapies. *Front. Oncol.* 10, 188. doi:10.3389/fonc.2020.00188

Zhou, S., Bai, Z.-L., Xia, D., Zhao, Z.-J., Zhao, R., Wang, Y.-Y., et al. (2018). FTO Regulates the Chemo-Radiotherapy Resistance of Cervical Squamous Cell Carcinoma (CSCC) by Targeting β-catenin through mRNA Demethylation. *Mol. carcinogenesis* 57, 590–597. doi:10.1002/mc.22782

Zou, D., Dong, L., Li, C., Yin, Z., Rao, S., and Zhou, Q. (2019). The m6A Eraser FTO Facilitates Proliferation and Migration of Human Cervical Cancer Cells. *Cancer Cel Int* 19, 321. doi:10.1186/s12935-019-1045-1

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 Li, Tao, Zhao, Li, Zheng, Li 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.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 10 May 2022
doi: 10.3389/fgene.2022.808041

TGFβ2 is a Prognostic Biomarker for Gastric Cancer and is Associated With Methylation and Immunotherapy Responses

*Bangling Han, Tianyi Fang, Yimin Wang, Yongle Zhang and Yingwei Xue**

Department of Gastrointestinal Surgery, Harbin Medical University Cancer Hospital, Harbin, China

TGFβ signaling plays a key role in cancer progression and by shaping tumor architecture and inhibiting the anti-tumor activity of immune cells. It was reported that high expression of TGFβ can promote the invasion and metastasis of cancer cells in a variety of tumors. However, there are few studies on TGFβ2 and its methylation in gastric cancer. We analyzed the Harbin Medical University Cancer Hospital (HMUCH) sequencing data and used public data to explore the potential function and prognostic value of TGFβ2 and its methylation in gastric cancer. In this study, we used the ssGSEA algorithm to quantify 23 methylation sites related to TGFβ2. Survival analysis showed that high expression of TGFβ2 and hypomethylation levels of TGFβ2 were negative factors in the prognosis of gastric cancer. Functional enrichment analysis of methylation revealed that methylation of different TGFβ2 methylation scores was mainly involved in energy metabolism, extracellular matrix formation and cell cycle regulation. In the gastric cancer microenvironment TGFβ2 was associated with high levels of multiple immune cell infiltration and cytokine expression, and high TGFβ2 expression was significantly and positively correlated with stemness markers, stromalscore and EMT. Gene set enrichment analysis also revealed an important role of TGFβ2 in promoting EMT. In addition, we discussed the relationship between TGFβ2 and immunotherapy. The expression of PD-1, PD-L1 and CTLA-4 was elevated in the TGFβ2 high expression group. Also when TGFβ2 was highly expressed, the responsiveness of immune checkpoint blockade (ICB) was significantly enhanced. This indicates that TGFβ2 may become an indicator for predicting the efficacy of immunosuppressive agents and a potential target for immunotherapy.

Keywords: gastric cancer, TGFβ2, NDA methylation, epithelial-mesenchymal transition, tumor microenvironment, progression

OPEN ACCESS

Edited by:
William C. Cho,
QEH, Hong Kong SAR, China

Reviewed by:
Varun Sasidharan Nair,
Helmholtz Center for Infection Research, Germany
Raheleh Roudi,
Stanford University, United States

****Correspondence:***
Yingwei Xue
xueyingwei@hrbmu.edu.cn

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

Received: *02 November 2021*
Accepted: *25 April 2022*
Published: *10 May 2022*

Citation:
Han B, Fang T, Wang Y, Zhang Y and Xue Y (2022) TGFβ2 is a Prognostic Biomarker for Gastric Cancer and is Associated With Methylation and Immunotherapy Responses. Front. Genet. 13:808041. doi: 10.3389/fgene.2022.808041

INTRODUCTION

Gastric cancer is a malignant tumor originating from the epithelium of the stomach, with a high incidence rate and mortality rate (Bray et al., 2018). At present, the conventional treatment of gastric cancer is difficult to remove the tumor cells completely by surgery and chemotherapy, and cancer recurrence often occurs. With the rapid development of medical biotechnology, cancer immunotherapy with strong targeting and low side effects is a rapidly developing research direction in oncology (Cho et al., 2020). Immunotherapy is mainly aimed at immune cells,

which can activate the immune system by inhibiting negative immune regulatory factors, enhance the recognition and killing of tumor by immune cells, so as to achieve the purpose of tumor clearance (Cho et al., 2012; Garon, 2017).

In the complex tumor microenvironment, TGFβ is a pleiotropic cytokine involved in the regulation of cancer cell proliferation, apoptosis and metastasis and other cellular processes (David and Massagué, 2018). TGFβ defines three subtypes (TGFβ1, TGFβ2 and TGFβ3), among which TGFβ2 is highly expressed in many cancers, especially those tumors that show high transmission potential (Massagué, 1998). In addition, the increased expression of TGFβ2 in a variety of cancers is often positively correlated with epithelial-mesenchymal transition (EMT) and coordinated with the expression of genes related to driving EMT (Vagenas et al., 2007; Yang et al., 2020). TGFβ signaling in the tumor microenvironment inhibits the anti-tumor function of a variety of immune cell populations, including T cells and natural killer cells, and the resulting immunosuppression severely limits the efficacy of immune checkpoint inhibitors and other immunotherapy approaches. Inhibitors of TGFβ signaling have been evaluated in a number of clinical trials as a major pathway to improve the immune effect of cancer, and combining TGFβ related signals can enhance the effect of other immunosuppressants (Holmgaard et al., 2018; Tauriello et al., 2018). Trabedersen (AP-12009) was an antisense molecule complementary to the mRNA expressed by human TGFβ2 gene, it had been applied II/Phase III clinical cases and has achieved encouraging results (Schlingensiepen et al., 2006).

However, the potential functions and mechanisms of TGFβ2 and TGFβ2 methylation involved in gastric cancer progression are unclear. In this study, we comprehensively analyzed the relationship between TGFβ2 and TGFβ2 methylation levels and the microenvironmental characteristics of gastric cancer based on sequencing data from the Harbin Medical University Cancer Hospital (HMUCH) and public databases. The expression levels of TGFβ2 were also validated in gastric cancer tissues. The results showed that TGFβ2 was highly expressed in gastric cancer and was a poor prognostic factor. TGFβ2 was closely related to the gastric cancer microenvironment, and functional enrichment analysis of TGFβ2 and TGFβ2-associated methylation was performed to explore the possible mechanisms of TGFβ2 action in gastric cancer.

MATERIALS AND METHODS

Patient and Clinical Databases

In this study, frozen tissues (cancer and normal paracancerous tissue more than 5 cm from the tumor cut edge) were collected from 231 gastric cancer surgery patients and subjected to high-throughput sequencing of the transcriptome. All gastric cancer tissues were certified by independent examination by two pathologists to confirm the histological type. Patients were not treated preoperatively with adjuvant therapy such as radiotherapy and chemotherapy. We uploaded and stored the sequencing data into the GEO Datasets (GSE184336). All patients signed an informed consent form. This study complied with the requirements of the Research Ethics Committee of the Harbin Medical University Cancer Hospital (2019-164-R).

We downloaded gene expression data for pan-cancer from the TCGA public database, and the cancer abbreviations and full names for pan-cancer are shown in **Supplementary Table S1**. The STAD download data also included mutation information, pathology data, and survival status. The GSE84437, GSE63089, GSE62254, GSE34942, GSE29272, GSE26253 and GSE15459 gastric cancer datasets were downloaded from the Gene Expression Omnibus (GEO) database for further analysis (Li et al., 2014; Chia et al., 2015; Cristescu et al., 2015; Zhang et al., 2015; Oh et al., 2018; Subhash et al., 2018; Yoon et al., 2020).

Western Blot Analysis and Immunohistochemistry

Gastric cancer paraffin blocks were serially sectioned with a section thickness of 4 μm. Immunohistochemical staining was performed as described previously (Kalantari et al., 2022). Total proteins of gastric cancer and normal tissues adjacent to the cancer were extracted and their concentrations were determined. PVDF membranes (Merck Millipore) were blocked with 5% skim milk powder and incubated with TGFβ2 antibody dilution (Proteintech, 19999-1-AP, 1:800) overnight at 4°C environment.

TGFβ2 Related Methylation

First, we assigned DNA methylation values for TGFβ2 with the average beta value of the probes mapped to the promoter region, including TSS200 (region from -200 bp upstream to the transcription start site (TSS) itself), TSS1500 (from -200 to -1,500 bp upstream of TSS), 1stExon (the first exon) and 5′UTR. Genome annotation of the CpG sites was based on GRCh38. methylation levels of the CpG sites were estimated as beta values (Ding et al., 2020).

Then we used the 23 methylation sites of TGFβ2 as a joint feature of TGFβ2 methylation and used these 23 methylation sites as a dataset for TGFβ2 methylation scoring. The single sample gene set enrichment analysis (ssGSEA) algorithm in the R package GSVA was used to calculate the TGFβ2 methylation score in each sample (Hänzelmann et al., 2013). Patients with STAD were grouped according to the median TGFβ2 methylation score and subjected to methylation differential analysis, and R package methylGSA was used for functional enrichment analysis of methylation (Geeleher et al., 2013).

Estimate, EMT and mRNAi

The R package ESTIMATE was used to calculate stromalscore, immunescore and ESTIMATEScore to assess the tumor microenvironment, where tumor purity = cos (0.6049872018 + 0.0001467884 * estimate score) (Yoshihara et al., 2013). We also used the ssGSEA algorithm to assess the EMT score in gastric cancer patients with the EMT gene set as described previously (Cristescu et al., 2015). We used the mRNAsi index to assess the tissue stemness characteristics

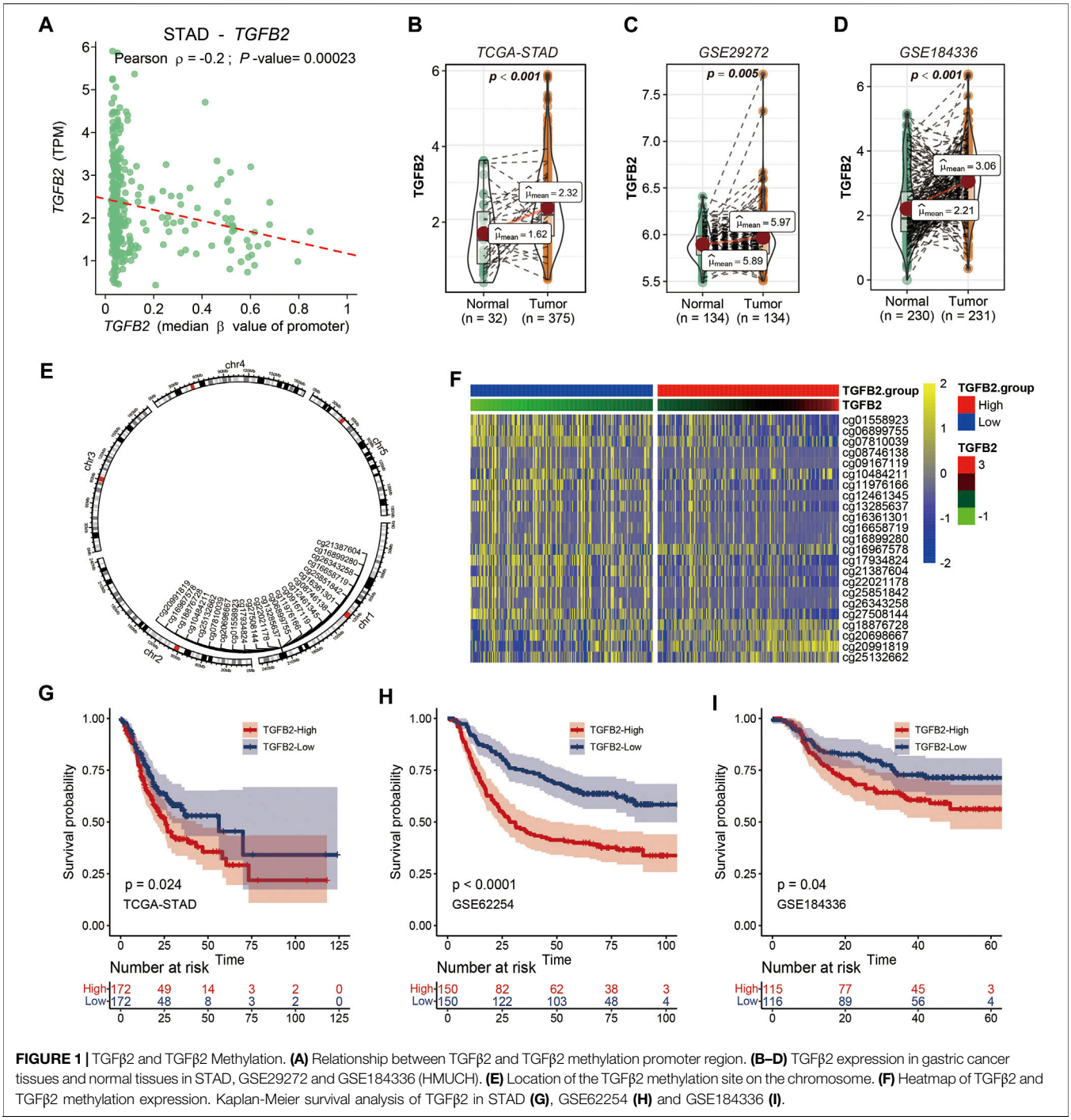


FIGURE 1 | TGFβ2 and TGFβ2 Methylation. **(A)** Relationship between TGFβ2 and TGFβ2 methylation promoter region. **(B–D)** TGFβ2 expression in gastric cancer tissues and normal tissues in STAD, GSE29272 and GSE184336 (HMUCH). **(E)** Location of the TGFβ2 methylation site on the chromosome. **(F)** Heatmap of TGFβ2 and TGFβ2 methylation expression. Kaplan-Meier survival analysis of TGFβ2 in STAD **(G)**, GSE62254 **(H)** and GSE184336 **(I)**.

of gastric cancer, and miRNA values were referenced to previous reports (Thorsson et al., 2018).

Immune Cell Infiltration and Cytokines

We quantified the level of immune cell infiltration in STAD using multiple methods, including CIBERSORT, MCPcounter, TIMER, ssGSEA and quanTIseq algorithms, and immune cells for each STAD patient were calculated as described previously (Becht et al., 2016; Charoentong et al., 2017; Li et al., 2017; Chen et al., 2018; Finotello et al., 2019). In addition, we analyzed the relationship between TGFβ2 and cytokines (receptor, chemokine, immunoinhibitor, immunostimulator and MHC) in the gastric cancer microenvironment (Ru et al., 2019).

Gene Set Enrichment Analysis

GSEA was presented between the high and low TGFβ2 groups. Pathways with nominal $p < 0.05$ and false

discovery rate (FDR) < 0.05 were considered signifificantly enriched (Subramanian et al., 2005). The "h.all.v7.1.entrez.gmt" was chosen as the reference.

Tumor Mutations

The R package maftools was used to analyze gene mutation information in STAD patients (Mayakonda et al., 2018). The Gistic 2.0 algorithm was used to identify copy number variation (CNV) and to display the frequency of CNV changes between different TGFβ2 groupings (Mermel et al., 2011). Tumor mutation burden (TMB) and neoantigens were calculated as previously described. The SCNA module in the TIMER database (http://timer.cistrome.org/) was used to compare the relationship between different somatic copy number alterations of TGFβ2 and immune cell infiltration.

Immunotherapy and Chemotherapy

In this study, we evaluated the relationship between TGFβ2 expression levels and immunotherapy. First, we used the submap algorithm to compare the similarity between the STAD expression data and the skin cancer immunotherapy dataset, a feature that can be reflected in the response of STAD to immunotherapy (Hoshida et al., 2007; Roh et al., 2017). We also used the ImmuCellAI database to predict immune checkpoint blockade (ICB) response (anti-PD-1 or anti-CTLA-4 treatment) in STAD patients (Miao et al., 2020). We estimated the sensitivity of STAD patients to chemotherapeutic agents using the genomics of drug sensitivity in cancer (GDSC) database (Yang et al., 2013). The half-maximal inhibitory concentration (IC50) was quantified and analyzed by the R package RpRophetic (Geeleher et al., 2014).

Nomogram and Calibration

Independent risk factors were identified by cox multi-factor regression analysis, and used R package rms to construct a nomogram to predict the probability of overall survival (Shariat et al., 2008). The calibration chart was used to evaluate the performance of the nomogram, and the 45° diagonal line represented the best predicted value. The index of concordance (C-index) was used to assess the agreement between the actual results and the probabilities predicted by the model.

Statistical Analysis

The chi-square test was used to analyze the association between different TGFβ2 subgroups and clinicopathological parameters. Kaplan-Meier survival curves were used to compare survival analyses between different subgroups followed by a log-rank test. The R package DESeq2 was used for differential analysis of count data of STAD patients, and the limma package was used for differential

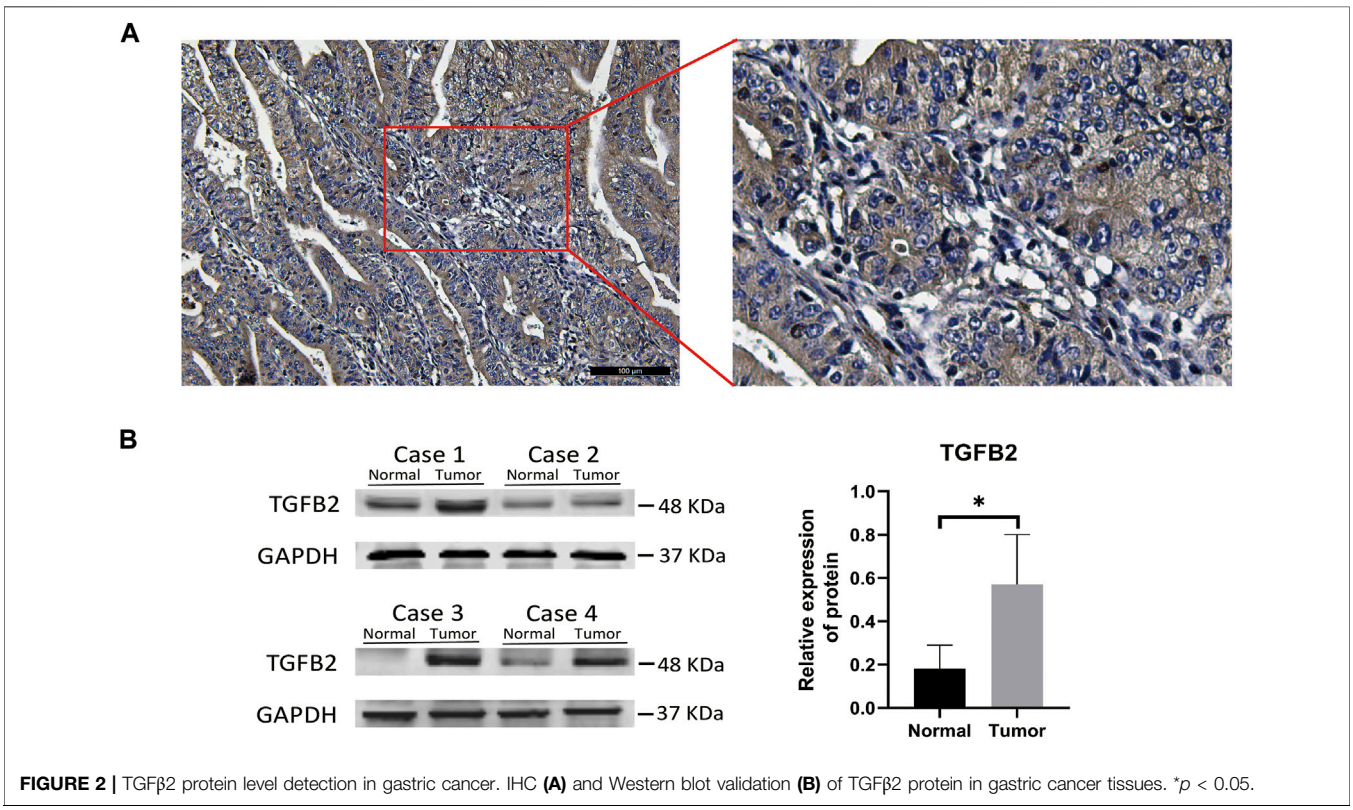


FIGURE 2 | TGFβ2 protein level detection in gastric cancer. IHC **(A)** and Western blot validation **(B)** of TGFβ2 protein in gastric cancer tissues. *$p < 0.05$.

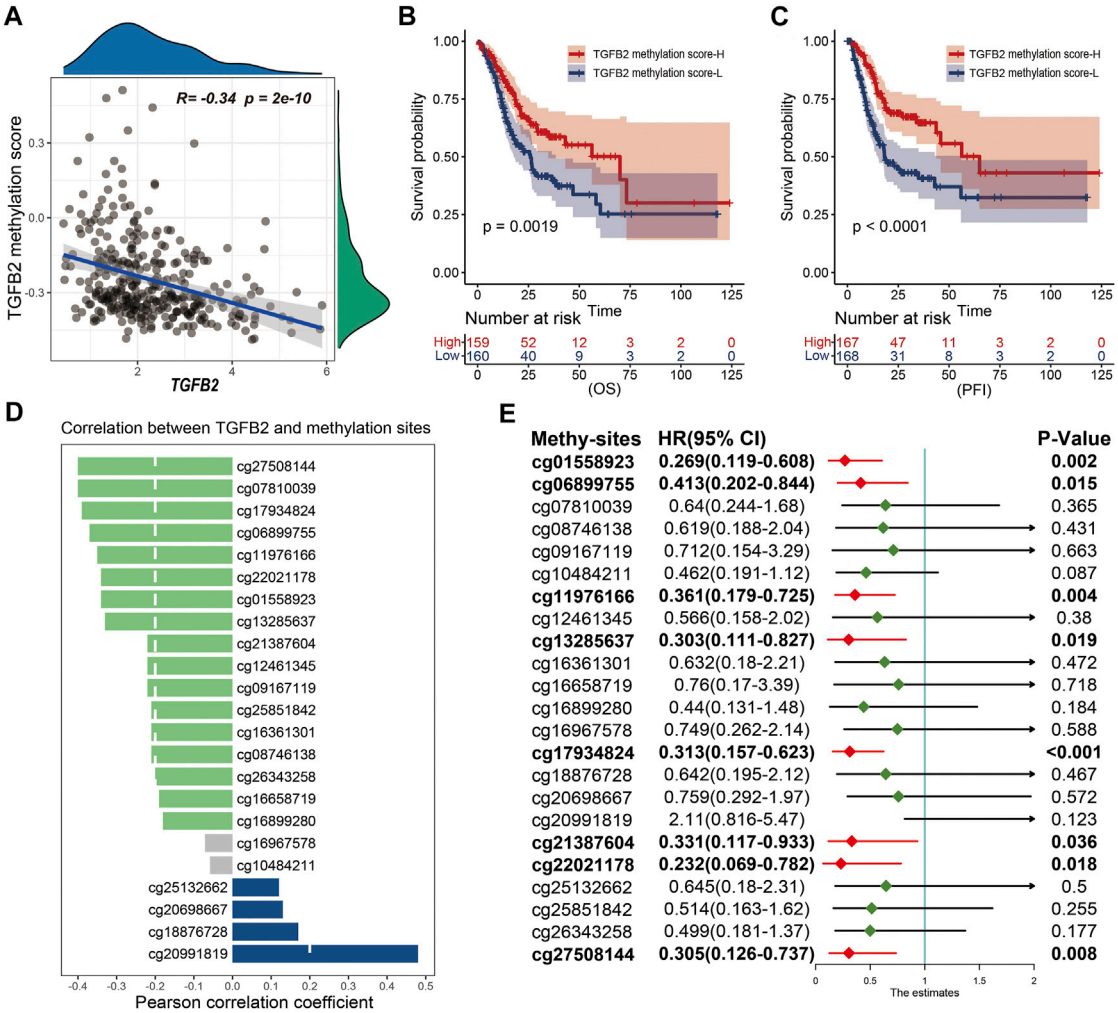


FIGURE 3 | TGFβ2 methylation score. **(A)** Correlation between TGFβ2 methylation score and TGFβ2 expression. Kaplan-Meier survival analysis of TGFβ2 methylation score in OS **(B)** and PFI **(C)**. **(D)** Correlation of TGFβ2 with methylation sites, and the column color represents the correlation of the difference, no statistical significance in gray. **(E)** Prognostic analysis of TGFβ2 methylation sites. HR < 1.0 (Red) indicates that the methylation site is a favorable prognostic marker.

analysis of methylation between different subgroups. The area under curve (AUC) value of ROC curve was calculated by the survivalROC R package. All statistical analyses were performed in R software (version 3.6.1). *p*-values less than 0.05 were considered statistically significant differences.

RESULT

TGFβ2 and TGFβ2 Methylation

The correlation between TGFβ2 and promoter region methylation in STAD was significantly negative, while TGFβ1 and TGFβ3 gene expression was not associated with promoter region methylation (**Figure 1A** and **Supplementary Figures S1A,B**). We compared the expression level of TGFβ2 in the STAD, GSE29272 and GSE184336 (HMUCH) data sets, and TGFβ2 was expressed at higher levels in cancer tissues than in the adjacent normal tissues ($p < 0.01$) (**Figures 1B–D**). We also showed the location distribution and detailed information of 23 TGFβ2 methylation sites in the chromosome (**Figure 1E** and **Supplementary Table S2**). Heatmap demonstrated TGFβ2 expression and 23 methylation site levels (**Figure 1F**). In the survival analysis of the STAD, GSE62254 and GSE184336 datasets, patients with high TGFβ2 expression all had shorter survival times ($p < 0.001$) (**Figures 1G–I**). We verified the protein expression level of TGFβ2 in gastric cancer tissues, and the results of IHC experiments showed that TGFβ2 was mainly distributed in cancer cells, with a small amount of distribution in the mesenchyme (**Figure 2A**). The results of Western blot assay showed that the protein expression level of TGFβ2 was higher in gastric cancer tissues than normal tissues, which was consistent with the results of transcriptome sequencing level (GSE184336) (**Figure 2B**).

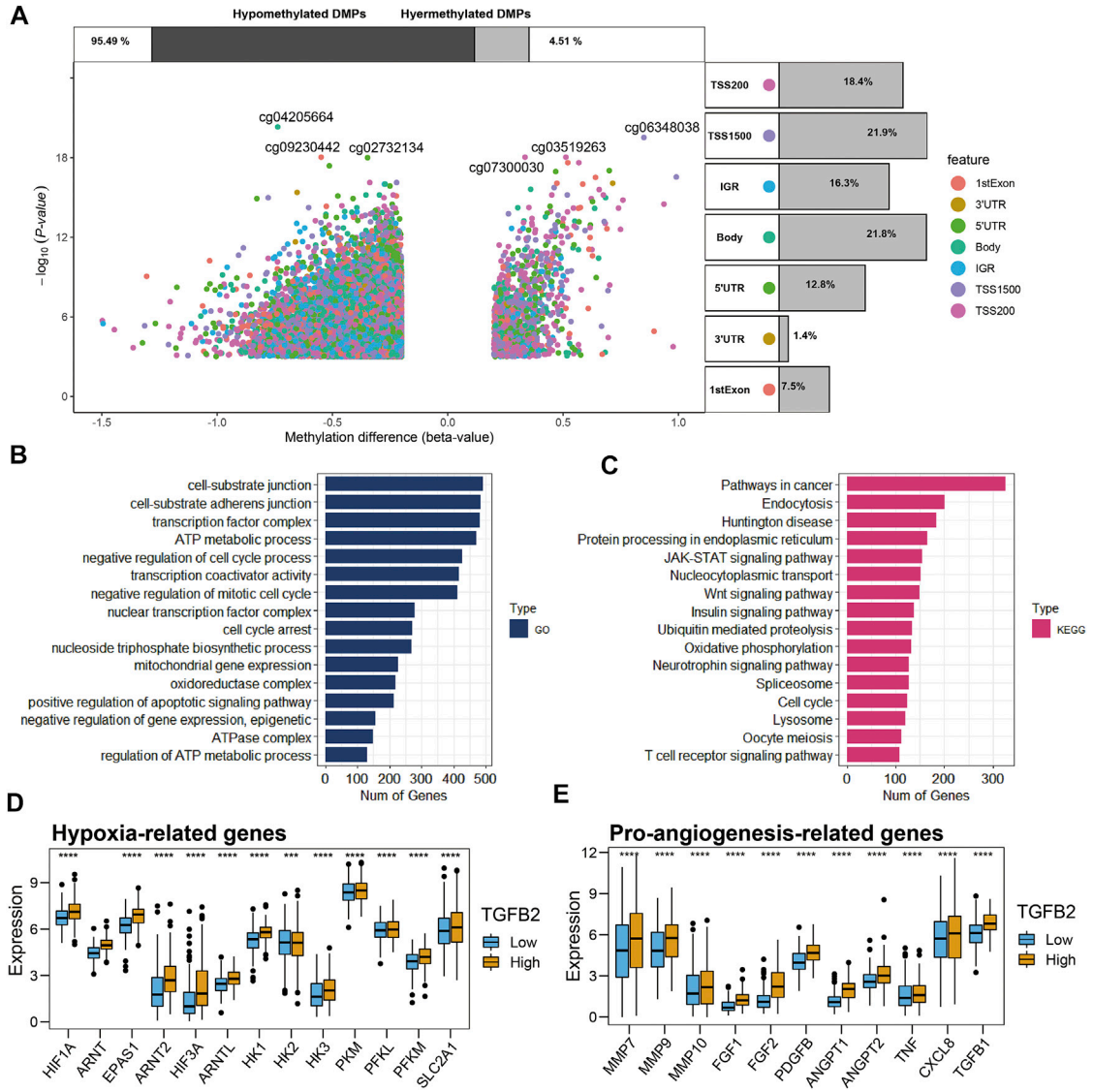


FIGURE 4 | Differential analysis of methylation and functional enrichment analysis. **(A)** Methylation volcano plots for different TGFβ2 methylation score groupings, with the location information of the methylation sites annotated on the right side of the figure notes and the expression differences of methylation on the top. **(B,C)** Methylation functional enrichment analysis, GO and KEGG results. Comparison of hypoxia-related genes **(D)** and pro-angiogenesis-related genes **(E)** in different TGFβ2 subgroups. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$.

We performed cox univariate analysis of the methylation sites and the results showed that cg01558923, cg06899755, cg11976166, cg13285637, cg17934824, cg21387604, cg22021178 and cg27508144 were statistically significant in STAD (**Figure 3E**). In addition, we also performed Kaplan-Meier survival analysis for 23 methylation sites, and most of them (cg01558923, cg06899755, cg08746138, cg10484211, cg11976166, cg12461345, cg13285637, cg17934824, cg21387604, cg22021178, cg26343258 and cg27508144) high expression was beneficial to prolong patient survival (**Supplementary Figure S2**).

TGFβ2 Methylation Score

To quantify the overall expression level of TGFβ2 methylation, we evaluated TGFβ2 methylation expression in STAD patients using the ssGSEA algorithm with 23 methylation sites as the reference set. Pearson correlation analysis showed a significant negative correlation between TGFβ2 and TGFβ2 methylation score (**Figure 3A**). Patients with high TGFβ2 methylation scores in STAD had longer survival times at OS (overall survival), PFI (progression-free interval), DPI (disease-free interval) and DSS (disease-specific survival) levels (**Figures 3B,C** and **Supplementary Figures S1C,D**). In addition, we also analyzed the correlation between 23 TGFβ2 methylation sites and TGFβ2 expression separately, and the results showed that 18 methylation sites were expressed at opposite levels to TGFβ2 expression. Although the levels of the four methylation sites (cg18876728, cg20698667, cg25132662 and cg20991819) followed the same trend as TGFβ2 expression, the cg18876728, cg20698667

and cg25132662 site was in the body region and the cg20991819 was in the 3′-UTR region at the end of the coding region, neither of which was within the promoter region that affects TGFβ2 expression (**Figure 3D** and **Supplementary Figure S3**).

Differential Analysis of Methylation and Functional Enrichment Analysis

We grouped STAD patients according to the median TGFβ2 methylation score, 169 cases in the TGFβ2-methy.score-Low group and 168 cases in the TGFβ2-methy.score-High group. After removing some undetected methylation probes, a total of 337 samples (pathological type of gastric adenocarcinoma) and 375,361 DNA methylation sites were obtained from TCGA. We identified 25,053 differentially methylated positions (DMPs) in the analysis of differences between different TGFβ2 methylation score subgroups with screening criteria of $p < 0.05$ and $|\text{delta}| > 0.2$ (**Figure 4A**).

Then, we performed functional enrichment analysis of methylation among different TGFβ2 methylation score groupings based on differential methylation analysis. The results of GO (Gene Onotology) analysis showed that methylation between different subgroups is mainly involved in cell cycle regulation (cell cycle arrest, nuclear transcription factor complex and negative regulation of mitotic cell cycle), extracellular matrix formation (cell-substrate adherens junction and cell-substrate junction) and energy metabolism regulation (regulation of ATP metabolic process, ATPase complex, oxidoreductase complex, mitochondrial gene expression, nucleoside triphosphate biosynthetic process and ATP metabolic process) (**Figure 4B**). The results of KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis showed that methylation among different subgroups may be involved in

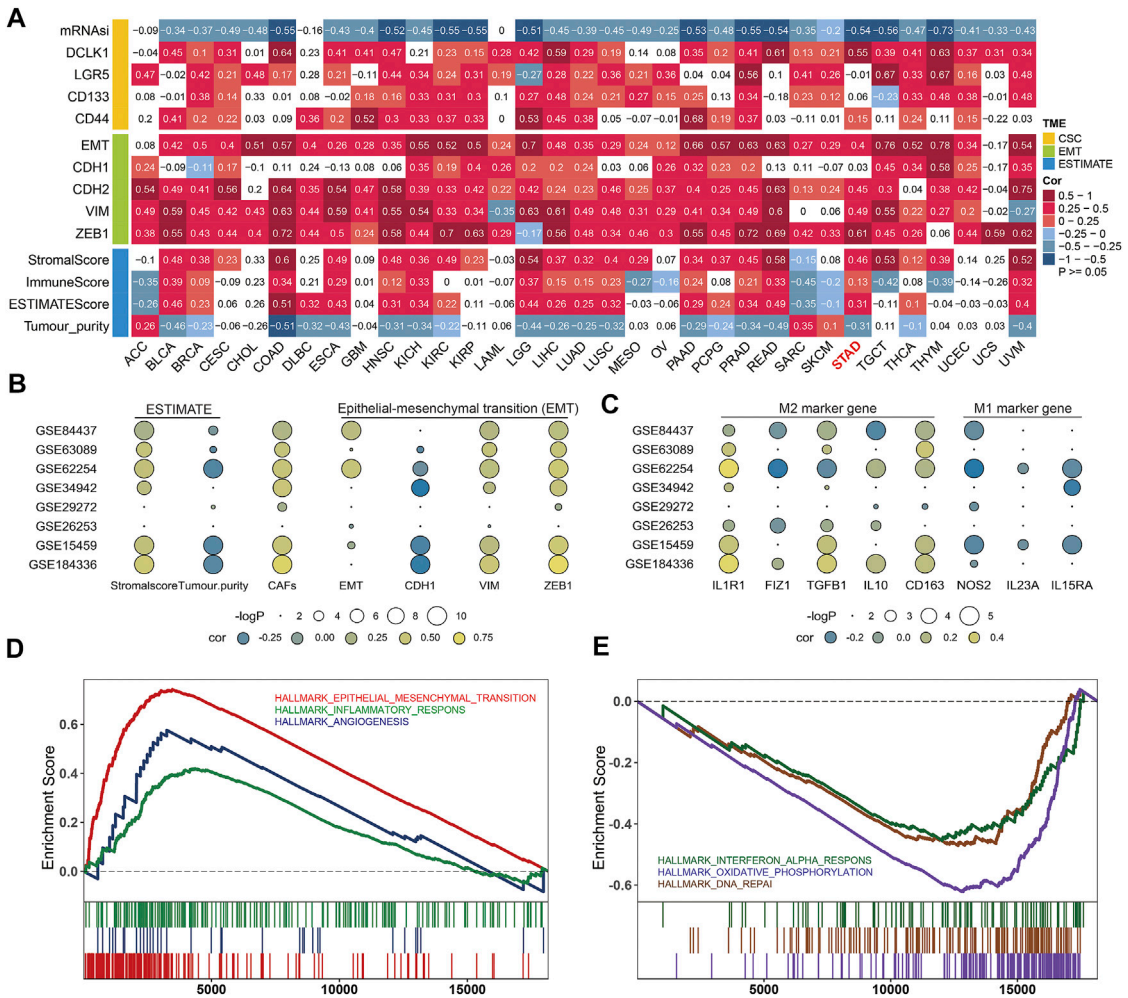


FIGURE 5 | TGFβ2 and gastric cancer microenvironment. **(A)** Correlation analysis of TGFβ2 with CSC, EMT and ESTIMATE in The Cancer Genome Atlas (TCGA). **(B)** Bubble plots of the correlation between TGFβ2 and ESTIMATE (Stromalscore and Tumor purity), CAFs and EMT (EMT, CDH1, VIM, ZEB1) in multiple GEO datasets. **(C)** Bubble plots of TGFβ2 correlation with M1 (IL1R1, FIZ1, TGFβ1, IL10 and CD163) and M2 (NOS2, IL23A and IL15RA) marker genes in multiple GEO datasets. Different colors represent different correlations. **(D,E)** Gene Set Enrichment Analysis (GSEA) was used to analyze the differences in gene function among different TGFβ2 subgroups.

cancer progression by regulating signaling pathways Wnt signaling pathway, JAK-STAT signaling pathway and Pathways in cancer (**Figure 4C**).

TGFβ2 and Hypoxia and Pro-Angiogenesis Genes

Given the important function of methylation on metabolic regulation among different TGFβ2 methylation score subgroups and that oxygen content is one of the important factors affecting energy metabolism, we compared hypoxia and proangiogenesis-related genes among patients with different TGFβ2 subgroups. The results showed that hypoxia-related genes (HIF1A family, HK1, HK3, PKM, PFKL, PFKM and SLC2A1) were higher in the high TGFβ2-expressing subgroup of patients, while pro-angiogenic genes (MMP7, MMP9, MMP10, FGF1, FGF2, PDGFB, ANGPT, ANGPT2, TNF, CXCL8 and TGFβ1) expression was also higher (**Figures 4D,E**).

TGFβ2 and Gastric Cancer Microenvironment

In the tumor microenvironment we analyzed the relationship between TGFβ2 and tumor stemness, EMT and ESTIMATE scores, respectively, with similar results

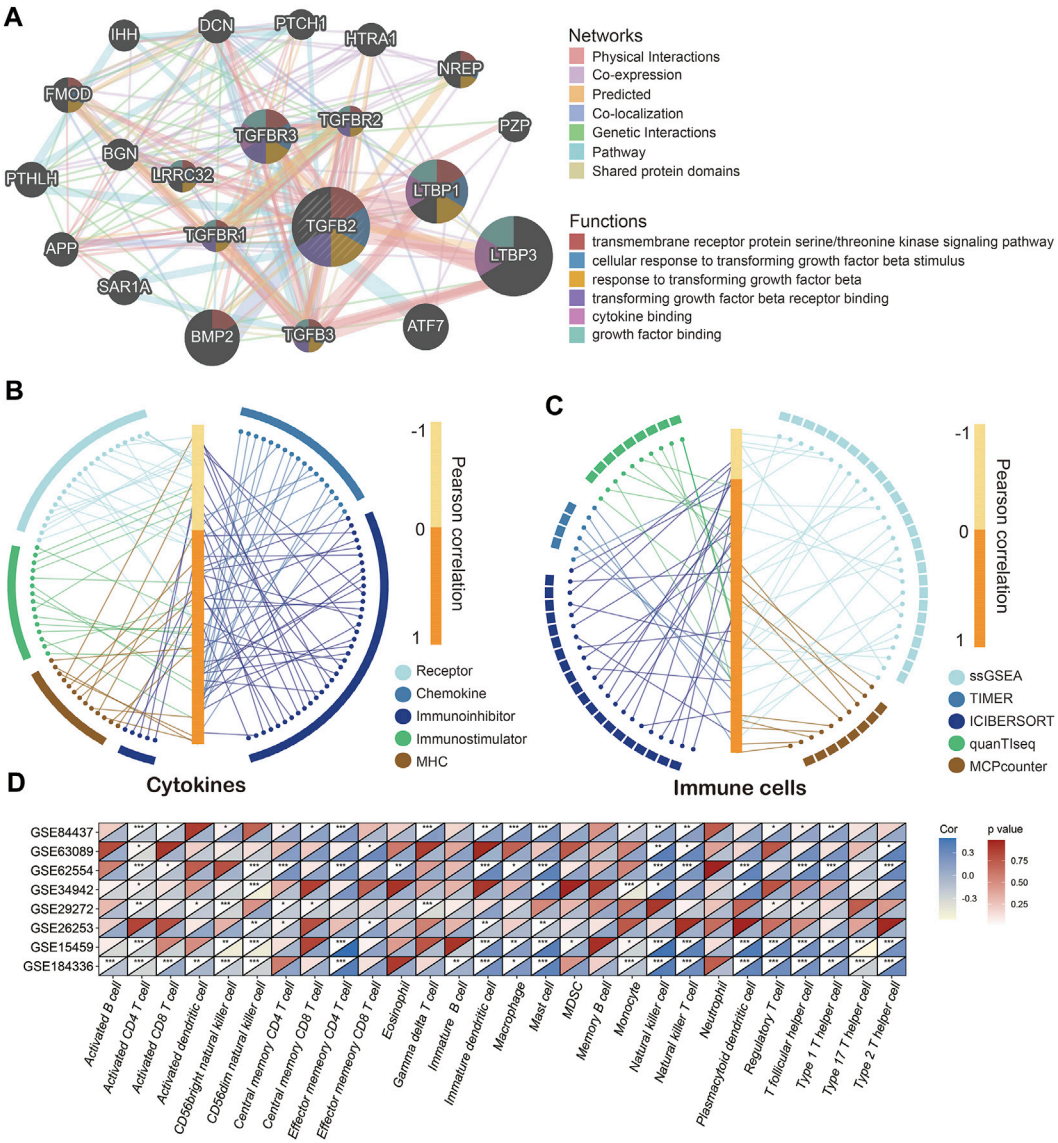


FIGURE 6 | TGFβ2 and immune cell infiltration and cytokine levels. **(A)** TGFβ2-related gene network. Correlation of TGFβ2 with immune cells **(B)** and cytokines **(C)**. Correlation of TGFβ2 with immune cells in multiple GEO datasets **(D)**. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

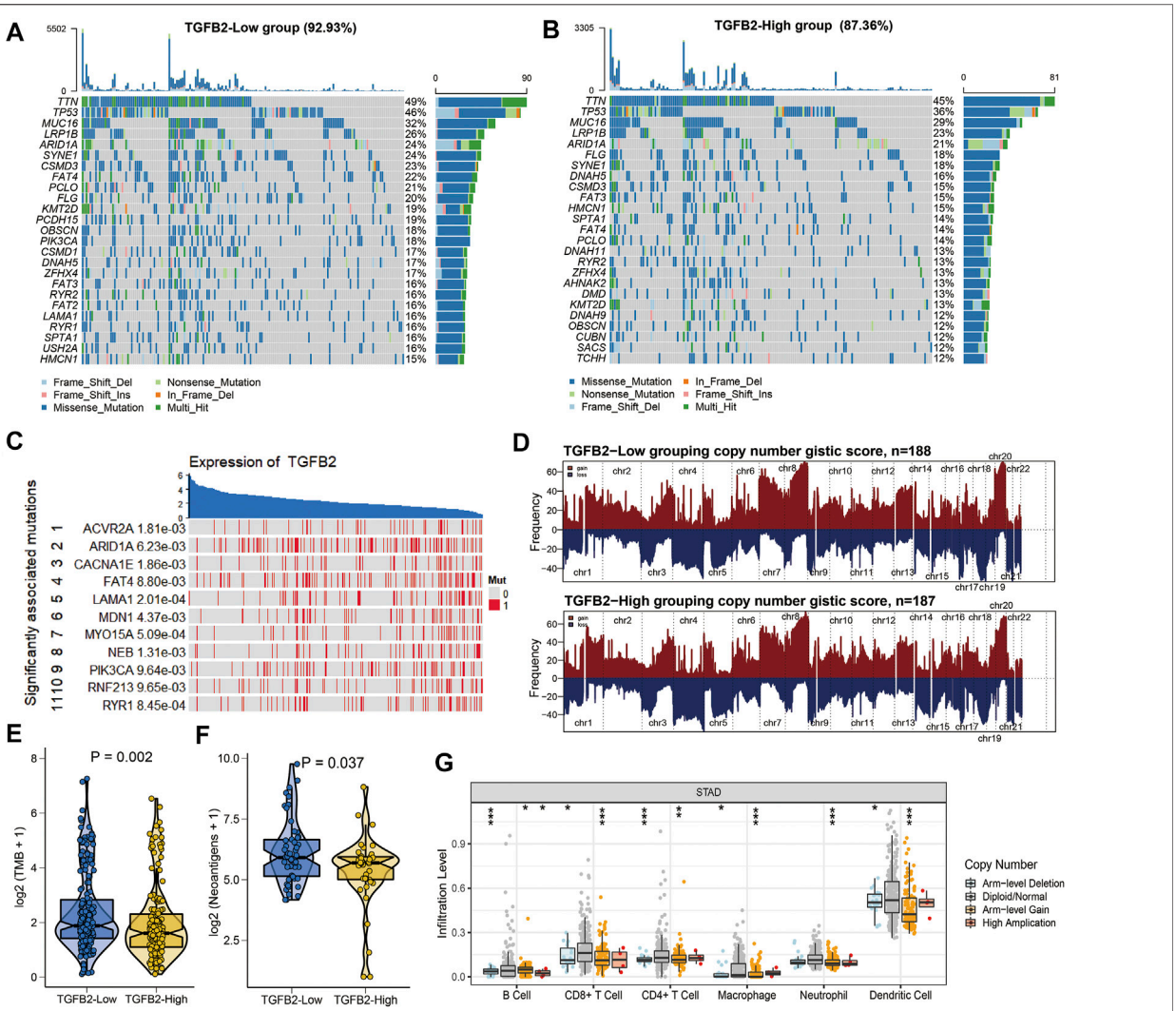


FIGURE 7 | TGFβ2 and mutations. **(A, B)** Waterfall plot of the 25 genes with the highest mutation frequencies in different TGFβ2 groups. **(C)** Mutations in genes affecting TGFβ2 expression. **(D)** Comparison of the frequency of copy number changes in different TGFβ2 groupings. Chromosomal locations of peaks of significantly recurring focal amplification (red) and deletions (blue) were presented. Comparison of TMB **(E)** and neoantigens **(F)** between different TGFβ2 groups. **(G)** Effect of the Genetic Alterations of TGFβ2 on the Immune Cell Infifiltration. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

between different pan-cancers. TGFβ2 was significantly negatively correlated with mRNAsi and significantly positively correlated with cancer stem cell markers DCLK1, Lgr5, CD133 and CD44. Also TGFβ2 was significantly positively correlated with EMT score and mesenchymal markers (CDH2, VIM and ZEB1), and in STAD, there was no significance in the analysis between TGFβ2 and CDH1. In addition, TGFβ2 was positively correlated with stromalscore, immunescore and ESTIMATEScore and negatively correlated with tumor purity in a variety of tumors including STAD (**Figure 5A**).

We also performed validation in the GEO gastric cancer dataset (GSE184336 (HMUCH), GSE84437, GSE63089, GSE62254, GSE34942, GSE29272, GSE26253 and GSE15459). The results showed that TGFβ2 was positively correlated with stromalscore and CAFs and negatively correlated with tumor purity in several datasets, which was consistent with the results of the analysis in STAD. In the correlation analysis with EMT, TGFβ2 was positively correlated with EMT and mesenchymal markers and negatively correlated with CDH1 in multiple datasets (**Figure 5B**). We also analyzed the relationship between TGFβ2 and M1 and M2 marker genes, and the results showed that TGFβ2 was negatively associated with most M1-related marker genes and positively associated with M2-related genes (**Figure 5C**). According to the differential analysis of different TGFβ2 groupings, GSEA analysis showed that the high TGFβ2 expression group promoted epithelial-mesenchymal transition, inflammatory response and angiogenesis, and inhibited interferon alpha response, oxidative phosphorylation and DNA repair (**Figures 5D,E**).

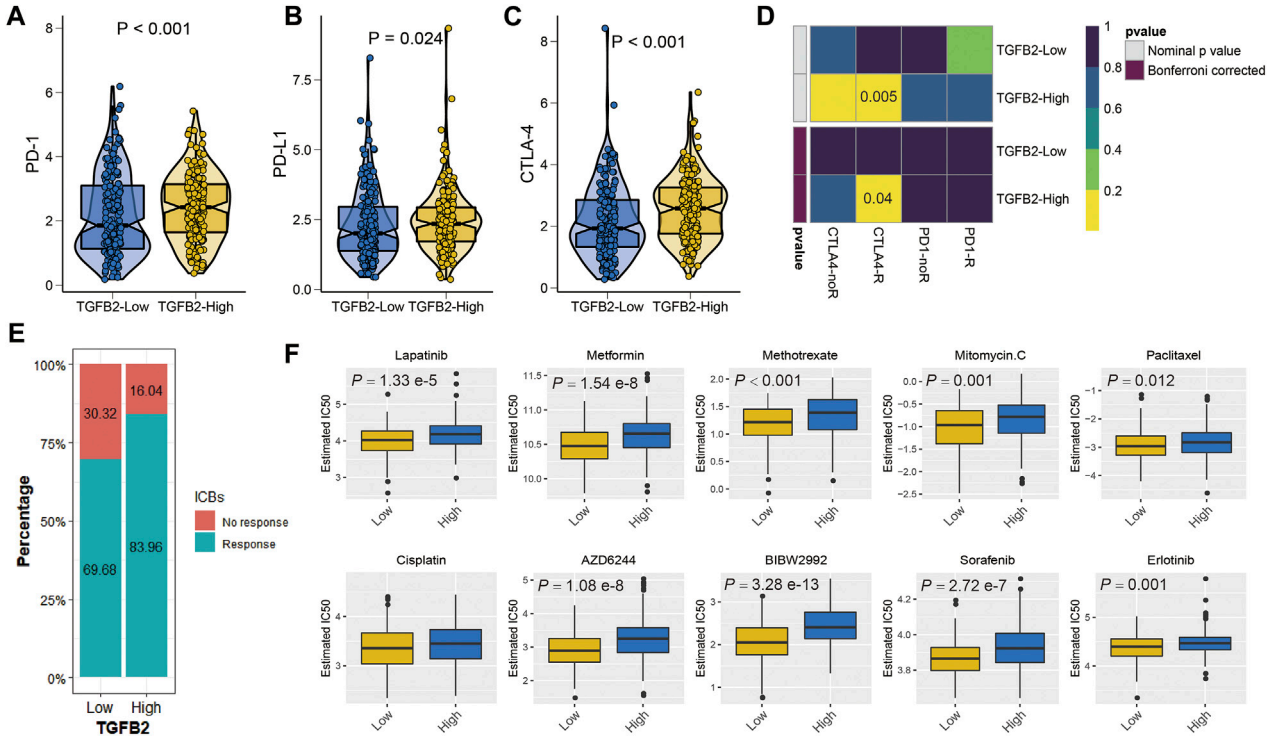


FIGURE 8 | TGFβ2 and immunotherapy and chemotherapy. Differences in immune checkpoints between high and low TGFβ2 groups, PD-1 **(A)**, PD-L1 **(B)** and CTLA-4 **(C)**. **(D)** Heatmap visualized the response to anti-CTLA4 and anti-PD1 therapies between the two groups. **(E)** The histogram showed the responsiveness of immunotherapy between high and low TGFβ2 groups, the height of each bar represents the frequency of change. **(F)** Boxplots depicted the differences in the estimated IC50 levels of Lapatinib, Metformin, Methotrexate, Mitomycin. C, Paclitaxel, Cisplatin, AZD6244, BIBW2992, Sorafenib and Erlotinib between the high and low TGFβ2 groups.

TGFβ2 and Immune Cell Infiltration and Cytokine Levels

We used the GeneMANIA cloud database to analyze other genes related to TGFβ2, and the gene network showed that TGFβ2 was more closely related to LTBP3, LTBP1 and BMP2 (**Figure 6A**) (Warde-Farley et al., 2010). Given the important role of TGFβ2 in the tumor microenvironment, we further analyzed the relationship between TGFβ2 and immune cell infiltration and cytokines (receptor, chemokine, immunoinhibitor, immunostimulator and MHC). The results showed that TGFβ2 was significantly and positively correlated with increased levels of multiple immune cell infiltration and cytokines (**Figures 6B,C**). Meanwhile, TGF-β2 had strong correlation with various immune cell ($CD8^+$ T cell, Monocyte, TAM, M1 Macrophage, M2 Macrophage and Treg) marker genes (**Supplementary Table S3**).

Similarly in the high TGFβ2 subgroup most immune cells infiltrated and cytokine expression levels were higher (**Supplementary Figures S4A–E** and **Supplementary Figure S5**). In contrast, most immune cells in the high TGFβ2 methylation score subgroup had low levels of infiltration (**Supplementary Figure S4F**). We also analyzed the correlation between TGFβ2 and immune cell infiltration in the GEO dataset, and the results showed that TGFβ2 positively correlated with activated CD4 T cell, CD56dim natural killer cell, immature dendritic cell, mast cell, natural killer cell and regulatory T cell in multiple GEO datasets (**Figure 6D**).

TGFβ2 and Mutations

We showed the 25 genes with the highest mutation frequency in different TGFβ2 subgroups and the difference of copy number variation between the two groups (**Figures 7A,B,D**). TGFβ2 expression was associated with mutations in genes (ACVR2A, ARID1A and CACNA1E) (**Figure 7C**). Tumor mutation burden and neoantigens were both higher in the TGFβ2 low expression group (**Figures 7E,F**). We also analyzed the effect of somatic copy number alterations (CNAs) of TGFβ2 on immune cell infiltration to elucidate the potential mechanism of TGFβ2 associated with infiltration of different immune cells, and the results showed that arm-level deletion and arm-level gain significantly affected the infiltration levels of B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages and dendritic cells in STAD (**Figure 7G**).

TGFβ2 and Immunotherapy and Chemotherapy

We explored the relationship between TGFβ2 and immunotherapy, and PD-1, PD-L1 and CTLA-4 expression levels were higher in the TGFβ2 high expression group (p <

TABLE 1 | The relationship between TGFβ2 expression and clinicopathological factors in HMUCH (GSE184336) and TCGA-STAD database.

Clinical Features	Total	TGFβ2 Expression (HMUCH)		*p*-Value	Total	TGFβ2 Expression (STAD)		*p*-Value
		Low (%)	High (%)			Low (%)	High (%)	
Age				0.948				0.864
<60	114	57 (49.1)	57 (49.6)		105	56 (29.8)	56 (30.6)	
≥60	117	59 (50.9)	58 (50.4)		236	132 (70.2)	127 (69.4)	
Gender				0.462				0.859
Female	83	39 (33.6%)	44 (38.3%)		122	68 (36.2)	66 (35.3)	
Male	148	77 (66.4%)	71 (61.7%)		222	120 (63.8)	121 (64.7)	
TNM stage				0.708				0.453
I	36	20 (17.2)	16 (13.9)		53	32 (17.5)	21 (11.9)	
II	49	27 (23.3)	22 (19.1)		112	58 (31.7)	54 (30.7)	
III	129	61 (52.6)	68 (59.1)		154	74 (40.4)	80 (45.5)	
IV	17	8 (6.9)	9 (7.8)		40	19 (10.4)	21 (11.9)	
T stage				0.217				0.176
T1	21	11 (9.5)	10 (8.7)		19	14 (7.4)	5 (2.8)	
T2	25	17 (14.7)	8 (7)		80	41 (21.8)	39 (21.8)	
T3	142	70 (60.3)	72 (62.6)		168	87 (46.3)	81 (45.2)	
T4	43	18 (15.5)	25 (21.7)		100	46 (24.5)	54 (30.2)	
N stage				0.241				0.494
N0	65	33 (28.4)	32 (27.8)		111	63 (34.6)	48 (27.3)	
N1	29	19 (16.4)	10 (8.7)		97	46 (25.3)	51 (29.0)	
N2	44	23 (19.8)	21 (18.3)		76	36 (19.8)	40 (22.7)	
N3	93	41 (35.3)	52 (45.2)		74	37 (20.3)	37 (21)	
Histologic Grade				0.606				**0.003**
G1	4	3 (2.6)	1 (0.9)		10	7 (3.8)	3 (1.7)	
G2	92	46 (39.6)	46 (40)		138	84 (45.2)	54 (29.8)	
G3	135	67 (57.8)	68 (59.1)		219	95 (51.0)	124 (68.5)	

Bold values indicate p-*value < 0.05.*

0.05) (**Figures 8A–C**). We evaluated immunotherapy response using two methods, firstly the submap algorithm results showed that CTLA-4 immunosuppressive treatment was meaningful for patients in the high TGFβ2 group and secondly a higher proportion of patients in the high TGFβ2 group were also predicted to respond to immunotherapy according to the ImmuCellAI database (chi-squared test = 10.724, $p = 0.001$) (**Figures 8D,E**).

As an important method of adjuvant treatment for gastric cancer, chemotherapy plays an important role in clinical treatment. Differences in chemotherapy (IC50) between different TGFβ2-expressing groups were predicted according to the GDSC database, which showed a higher sensitivity to Lapatinib, Metformin, Methotrexate, Mitomycin. C, Paclitaxel, AZD6244, BIBW2992, Sorafenib and Erlotinib in the low TGFβ2 group (**Figure 8F**).

TGFβ2 and TGFβ2 Methylation Scores and Clinicopathological Factors

To further assess the clinical relevance of TGFβ2 and TGFβ2 methylation scores, we divided patients into two groups based on median TGFβ2 and TGFβ2 methylation scores (**Table 1** and **Supplementary Table S4)**. In the STAD dataset, patients in the TGFβ2 high expression group had worse tumor grade, and the TGFβ2 methylation score was associated with T stage (**Figures 9A,B**). The time-dependent area under the ROC curve (AUC) for TGFβ2 at 1, 3 and 5 years in the STAD, GSE62254 and GSE15459 datasets was around 0.6 (**Figures 9C–E**). In addition, The AUCs of the 1-, 3- and 5-years time-dependent ROC curves for TGFβ2 methylation scores in STAD were 0.387, 0.394 and 0.439, respectively (**Figure 9F**).

To test whether TGFβ2 could be an independent prognostic factor, we performed multivariate Cox regression analysis based on the clinical characteristics of the patients, including age, T stage and TNM stage. We found that TGFβ2 was a reliable and independent prognostic marker for assessing patient prognosis in STAD (HR = 1.25, 95% CI 1.058-1.47, $p = 0.009$; **Figure 9G**). TGFβ2 remained an independent prognostic marker in the GSE62254 dataset (HR = 3.866, 95% CI 2.2-6.79, $p < 0.001$; **Figure 9H**). These results suggested that TGFβ2 was a valid predictor of prognosis for patients with gastric cancer.

Nomogram and Calibration

In order to quantify the influence of clinicopathological factors including TGFβ2 on the prognosis, we used a nomogram to establish a predictive model. We drew a nomogram based on the multivariate analysis of STAD patients (age, TNM and TGFβ2; $p < 0.05$) (**Figure 10A**). The calibration chart for the 5-years survival rate of the three cohorts was well predicted (C-index: 0.661 for STAD cohort, 0.726 for validation cohort GSE62254, and 0.762 for validation cohort GSE15459) (**Figures 10B–D**).

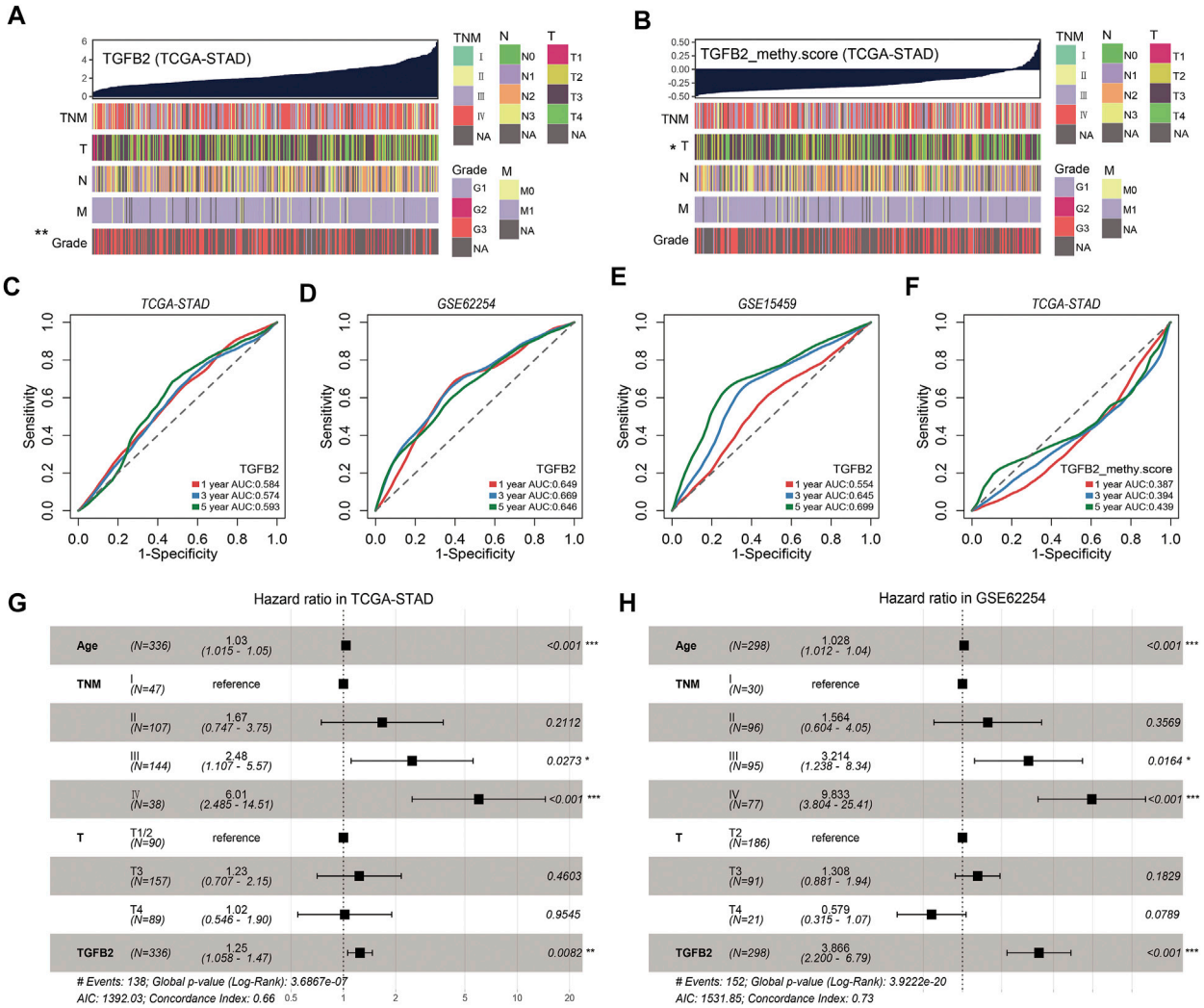


FIGURE 9 | TGFβ2 and TGFβ2 methylation scores and clinicopathological factors. **(A)** Heatmap of TGFβ2 and clinicopathological features in STAD. **(B)** Heatmap of TGFβ2 methylation score and clinicopathological features in STAD. The predictive value of TGFβ2 in patients in cohorts STAD **(C)**, GSE62254 **(D)** and GSE15459 **(E)**. The predictive value of TGFβ2 methylation scores in patients in cohorts STAD **(F)**. Multivariate Cox regression model analysis, which included the factors of age, TNM stage, T stage and TGFβ2 in the STAD **(G)** and GSE62254 **(H)** cohorts. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

DISCUSSION

Cancer cells are controlled by multiple regulatory signals during development, among which the TGFβ family plays an important role. Some studies have reported that the TGFβ2 signaling pathway plays a pro-cancer role in a variety of tumors, such as highly aggressive gliomas, breast cancers, and squamous cell carcinomas (Busch et al., 2015; Zhang et al., 2017; Abraham et al., 2018). However, there are fewer studies on TGFβ2 and its methylation in gastric cancer. The results of this study showed that the TGFβ2 expression level was significantly and negatively correlated with the TGFβ2 methylation promoter region and TGFβ2 methylation score. The expression levels of TGFβ2 in gastric cancer tissues were significantly higher in the datasets GSE184336, GSE29272 and STAD than in normal tissues adjacent to the cancer. Survival analysis showed that patients with high levels of TGFβ2 gastric cancer had shorter survival times and those with high TGFβ2 methylation scores had longer survival times. Multiple methylation sites acted as conservation roles in the univariate analysis of TGFβ2 methylation sites. TGFβ2 was an independent prognostic factor for patients with gastric cancer in both the STAD and GSE62254 datasets in a multifactorial survival analysis. In view of the relationship between TGFβ2 expression and TGFβ2 methylation, the results all emphasized that high TGFβ2 expression is a poor prognostic factor for gastric cancer, which is consistent with the results of previous research reports (Yang et al., 2020).

Methylation functional enrichment analysis revealed that methylation of different TGFβ2 methylation scores is mainly involved in participating in cell cycle regulation, extracellular matrix formation and energy metabolism regulation. The proportion of energy metabolism-related regulation was high

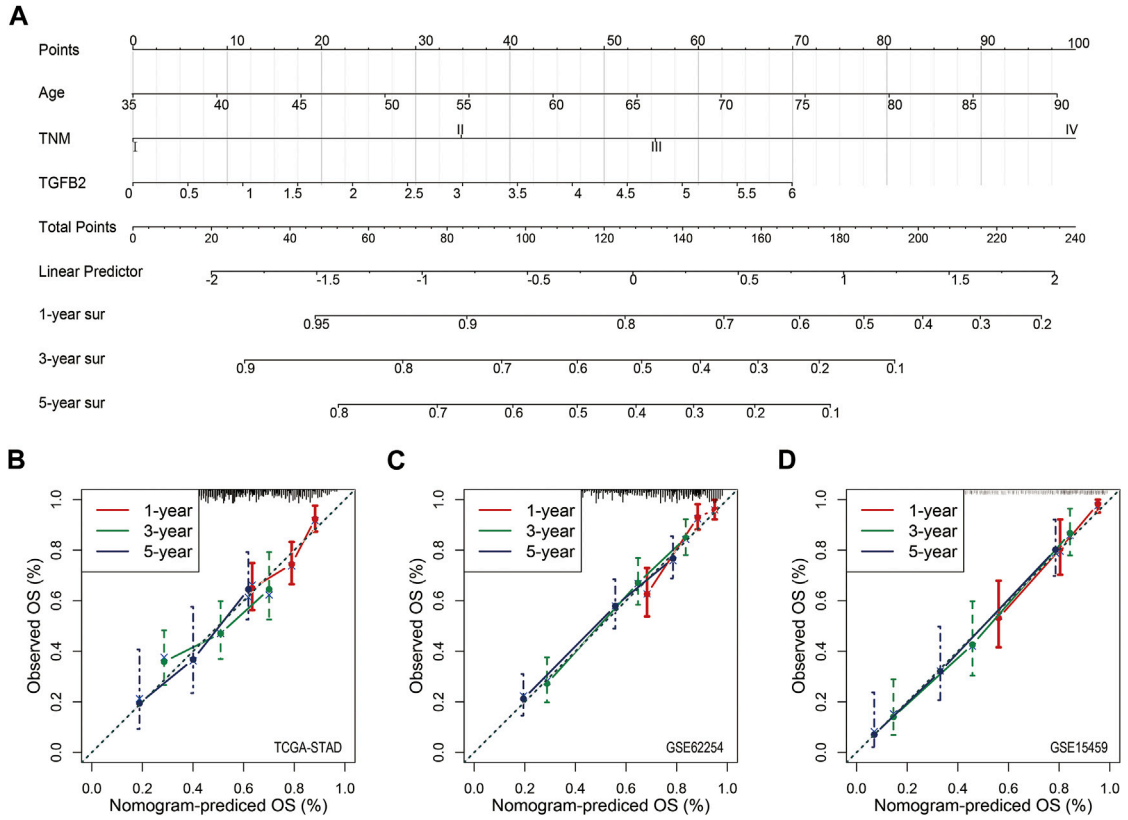


FIGURE 10 | Nomogram and Calibration. **(A)** Nomogram for predicting the probability of 1-, 3- and 5-years OS for STAD patients. Calibration curve for the nomogram for predicting the 1-, 3-, and 5-years OS probabilities for STAD **(B)**, GSE62254 **(C)** and GSE15459 **(D)**.

in the enrichment analysis (GO) results, and considering the importance of oxygen content in energy supply, we further analyzed the marker genes that correlate TGFβ2 with the oxygen content of the tissue microenvironment. The results showed high expression of HIF1A family and pro-angiogenesis related genes in the microenvironment of patients in the high TGFβ2 group, which can be inferred to be a hypoxic environment within the high TGFβ2 expressing tumor tissues. Studies have reported that in the myocardial ischemia experiment of mice, with the increase of hypobaric hypoxia time, the mRNA level of TGFβ increases (Xiao et al., 2016). Hypoxia, as one of the important features of tumors, plays an important role in cancer progression; therefore, high expression of TGFβ2 may also be indirectly involved in tumor progression by regulating hypoxia-related genes. Tumor cells induce hypoxia through various mechanisms, such as high metabolic rate and high oxygen consumption, creating a chronic hypoxic environment, activating hypoxia-inducible factors (HIF) signaling pathways, accelerating tumor growth, increasing tumor aggressiveness, and contributing to tumor metastasis (Andrysik et al., 2021).

In the gastric cancer microenvironment TGFβ2 is closely related to tumor stemness, EMT and stroma. Cancer stem cells (CSCs) a subpopulation of tumor cells with the ability to self-renew and differentiate, play an important role in cancer progression (Yang et al., 2015). mRNAsi is an indicator describing the degree of similarity of tumor cells to stem cells and can be considered as a quantification of CSCs. TGFβ2 was negatively correlated with mRNAsi and significantly positively correlated with stemness markers (DCLK1 and CD44) in this study, so TGFβ2 may be important factor in maintaining tumor stemness and promoting tumor differentiation in gastric cancer (Kalantari et al., 2017). EMT is the process by which polar epithelial cells convert to migratory mesenchymal cells and acquire the ability to invade and migrate, and it is present in several physiological and pathological processes in the human body. There are many regulatory factors of EMT, such as TGFβ, Wnt signaling pathway, microRNA and transcription factors (Vergara et al., 2019). Our results showed that TGFβ2 was significantly and positively correlated with EMT, CDH2, VIM and ZEB1. As cancer cells weaken their epithelial features during EMT, they may express fewer tumor-specific neoantigens to avoid recognition by immune cells, all of which contribute to cancer progression (Batlle and Massagué, 2019). Also high expression of TGFβ2 in the gastric cancer microenvironment was associated with higher levels of stromalscore and CAFs infiltration, and similar results were seen in the pan-cancer and multiple gastric cancer datasets. In functional enrichment analysis, the high TGFβ2 expression group promoted epithelial-mesenchymal transition and angiogenesis and inhibited oxidative phosphorylation and DNA repair.

Our findings suggested that TGFβ2 was strongly correlated with a variety of immune cell infiltrates, including $CD8^+$ T cells, monocytes, TAM, M1 macrophages, M2 macrophages and Treg. Treg is a subgroup of cells with significant immunosuppressive effects (Flemming, 2016). As an important secretion factor, TGFβ2 has a strong correlation with Treg, so Treg may inhibit protective immune cells through TGFβ2. TGFβ also affects the types of Myeloid-derived suppressor cells (MDSCs) in TME, including macrophages and neutrophils, prompting them to gradually transform into a tumor-promoting phenotype during cancer progression (Fridlender et al., 2009; Laoui et al., 2011). The results of this study showed that TGFβ2 was negatively correlated with M1 macrophage markers and positively correlated with M2 macrophage markers, showing an overall M2 macrophage phenotype that promotes immune escape of tumor cells. The relationship between TGFβ2 and immune cell infiltration was also demonstrated by genetic mutations, and the results showed that the somatic copy number alteration (Arm-level Gain) of the TGFβ2 gene was closely related to the level of STAD immune cell infiltration.

Given the strong correlation between TGFβ2 and immune cell infiltration and its important role in the regulation of immune cell function in gastric cancer, TGFβ2 is an important factor that cannot be ignored in gastric cancer immunotherapy. Stromal fibroblasts and other cells in tumor tissue shape the immunosuppressive environment of the tumor through TGFβ signaling, inhibiting the anti-tumor activity of immune cells and preventing or diminishing the effects of anti-cancer immunotherapy (Chakravarthy et al., 2018). Therefore, inhibition of TGFβ signaling is considered to be a prerequisite and an important way to improve the effectiveness of immunotherapy. Considering TGFβ, CTLA4 and PD-L1/PD-1 as parallel immunosuppressive pathways, combining TGFβ inhibitors with other immune checkpoint inhibitors may improve treatment outcomes. Combination therapy has been pre-evaluated in mouse cancer models where, depending on the model and experimental design, anti-PD-1 or anti-PD-L1 antibody treatment enhanced the antitumor effects of TGFβ inhibition and inhibited tumor metastasis (Mariathasan et al., 2018). The results of this study showed higher TMB and neoantigens in the low TGFβ2 expression group. At the same time, we also predicted the response of TGFβ2 expression level to ICB treatment, and patients with high TGFβ2 expression had a higher response rate from ICB therapy. However, this study has some limitations and still needs further validation by biological experiments and clinical data.

In conclusion, high expression of TGFβ2 and hypomethylation of TGFβ2 are factors of poor prognosis in gastric cancer. The high expression of TGFβ2 in gastric cancer tissue affects the tumor microenvironment and the level of immune cell infiltration by regulating DNA damage, angiogenesis, inflammation and EMT. The high responsiveness of ICB when TGFβ2 is highly expressed suggests that the detection of TGFβ2 expression can predict the response of gastric cancer patients to immune checkpoint inhibitors and may serve as a candidate target for gastric cancer immunotherapy.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by This study complied with the requirements of the Research Ethics Committee of the Harbin Medical University Cancer Hospital (2019-164-R). 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

BH and TF conceived the project and wrote the manuscript. BH and YW participated in data analysis. BH and YZ participated in discussion and language editing. YX reviewed the manuscript.

FUNDING

This study was supported by the Harbin Science and Technology Bureau Research and Development Project of Applied Technology (No. 2017RAXXJ054) and Nn 10 Program of Harbin Medical University Cancer Hospital (No. Nn 10 PY 2017-03).

ACKNOWLEDGMENTS

We are grateful for the data provided by the TCGA and GEO databases.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.808041/full#supplementary-material

REFERENCES

Abraham, C. G., Ludwig, M. P., Andrysik, Z., Pandey, A., Joshi, M., Galbraith, M. D., et al. (2018). ΔNp63α Suppresses TGFB2 Expression and RHOA Activity to Drive Cell Proliferation in Squamous Cell Carcinomas. *Cell Rep.* 24, 3224–3236. doi:10.1016/j.celrep.2018.08.058

Andrysik, Z., Bender, H., Galbraith, M. D., and Espinosa, J. M. (2021). Multi-omics Analysis Reveals Contextual Tumor Suppressive and Oncogenic Gene Modules within the Acute Hypoxic Response. *Nat. Commun.* 12, 1375. doi:10.1038/s41467-021-21687-2

Batlle, E., and Massagué, J. (2019). Transforming Growth Factor-β Signaling in Immunity and Cancer. *Immunity* 50, 924–940. doi:10.1016/j.immuni.2019.03.024

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

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, 394–424. doi:10.3322/caac.21492

Busch, S., Acar, A., Magnusson, Y., Gregersson, P., Rydén, L., and Landberg, G. (2015). TGF-beta Receptor Type-2 Expression in Cancer-Associated Fibroblasts Regulates Breast Cancer Cell Growth and Survival and Is a Prognostic Marker in Pre-menopausal Breast Cancer. *Oncogene* 34, 27–38. doi:10.1038/onc.2013.527

Chakravarthy, A., Khan, L., Bensler, N. P., Bose, P., and De Carvalho, D. D. (2018). TGF-β-associated Extracellular Matrix Genes Link Cancer-Associated Fibroblasts to Immune Evasion and Immunotherapy Failure. *Nat. Commun.* 9, 4692. doi:10.1038/s41467-018-06654-8

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, 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

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, 707–719. doi:10.1136/gutjnl-2013-306596

Cho, W., Ziogas, D. E., Katsios, C., and Roukos, D. H. (2012). Emerging Personalized Oncology: Sequencing and Systems Strategies. *Future Oncol.* 8 (6), 637–641. doi:10.2217/fon.12.44

Cho, W. C. S., Tse, K.-P., Ngan, R. K. C., Cheuk, W., Ma, V. W. S., Yang, Y.-T., et al. (2020). Genomic Characterization Reveals Potential Biomarkers in Nasopharyngeal Carcinoma Patients with Relapse. *Expert Rev. Mol. Diagnostics* 20, 1149–1159. doi:10.1080/14737159.2020.1835473

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, 449–456. doi:10.1038/nm.3850

David, C. J., and Massagué, J. (2018). Contextual Determinants of TGFβ Action in Development, Immunity and Cancer. *Nat. Rev. Mol. Cell Biol.* 19, 419–435. doi:10.1038/s41580-018-0007-0

Ding, W., Chen, J., Feng, G., Chen, G., Wu, J., Guo, Y., et al. (2020). DNMIVD: DNA Methylation Interactive Visualization Database. *Nucleic Acids Res.* 48, D856–D862. doi:10.1093/nar/gkz830

Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., et al. (2019). Molecular and Pharmacological Modulators of the Tumor Immune Contexture Revealed by Deconvolution of RNA-Seq Data. *Genome Med.* 11, 34. doi:10.1186/s13073-019-0638-6

Flemming, A. (2016). Cancer: Tumour-Specific Ablation of Treg Cells Induces Anticancer Response. *Nat. Rev. Drug Discov.* 15, 676–677. doi:10.1038/nrd.2016.198

Fridlender, Z. G., Sun, J., Kim, S., Kapoor, V., Cheng, G., Ling, L., et al. (2009). Polarization of Tumor-Associated Neutrophil Phenotype by TGF-β: "N1" versus "N2" TAN. *Cancer Cell* 16, 183–194. doi:10.1016/j.ccr.2009.06.017

Garon, E. B. (2017). Cancer Immunotherapy Trials Not Immune from Imprecise Selection of Patients. *N. Engl. J. Med.* 376, 2483–2485. doi:10.1056/NEJMe1705692

Geeleher, P., Hartnett, L., Egan, L. J., Golden, A., Raja Ali, R. A., and Seoighe, C. (2013). Gene-set Analysis Is Severely Biased when Applied to Genome-wide Methylation Data. *Bioinformatics* 29, 1851–1857. doi:10.1093/bioinformatics/btt311

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

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

Holmgaard, R. B., Schaer, D. A., Li, Y., Castaneda, S. P., Murphy, M. Y., Xu, X., et al. (2018). Targeting the TGFβ Pathway with Galunisertib, a TGFβRI Small Molecule Inhibitor, Promotes Anti-tumor Immunity Leading to Durable, Complete Responses, as Monotherapy and in Combination with Checkpoint Blockade. *J. Immunother. cancer* 6, 47. doi:10.1186/s40425-018-0356-4

Hoshida, Y., Brunet, J.-P., Tamayo, P., Golub, T. R., and Mesirov, J. P. (2007). Subclass Mapping: Identifying Common Subtypes in Independent Disease Data Sets. *PLoS One* 2, e1195. doi:10.1371/journal.pone.0001195

Kalantari, E., Asadi Lari, M. H., Roudi, R., Korourian, A., and Madjd, Z. (2017). Lgr5High/DCLK1High Phenotype Is More Common in Early Stage and Intestinal Subtypes of Gastric Carcinomas. *Cbm* 20 (4), 563–573. doi:10.3233/CBM-170383

Kalantari, E., Ghods, R., Saeednejad Zanjani, L., Rahimi, M., Eini, L., Razmi, M., et al. (2022). Cytoplasmic Expression of DCLK1-S, a Novel DCLK1 Isoform, Is Associated with Tumor Aggressiveness and Worse Disease-specific Survival in Colorectal Cancer. *Cbm* 33 (3), 277–289. doi:10.3233/CBM-210330

Laoui, D., Movahedi, K., Van Overmeire, E., Van den Bossche, J., Schouppe, E., Mommer, C., et al. (2011). Tumor-associated Macrophages in Breast Cancer: Distinct Subsets, Distinct Functions. *Int. J. Dev. Biol.* 55, 861–867. doi:10.1387/ijdb.113371dl

Li, W.-Q., Hu, N., Burton, V. H., Yang, H. H., Su, H., Conway, C. M., et al. (2014). PLCE1 mRNA and Protein Expression and Survival of Patients with Esophageal Squamous Cell Carcinoma and Gastric Adenocarcinoma. *Cancer Epidemiol. Biomarkers Prev.* 23, 1579–1588. doi:10.1158/1055-9965.EPI-13-1329

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, e108–e110. doi:10.1158/0008-5472.CAN-17-0307

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, 544–548. doi:10.1038/nature25501

Massagué, J. (1998). TGF-β Signal Transduction. *Annu. Rev. Biochem.* 67, 753–791. doi:10.1146/annurev.biochem.67.1.753

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

Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhim, R., and Getz, G. (2011). GISTIC2.0 Facilitates Sensitive and Confident Localization of the Targets of Focal Somatic Copy-Number Alteration in Human Cancers. *Genome Biol.* 12, R41. doi:10.1186/gb-2011-12-4-r41

Miao, Y. R., Zhang, Q., Lei, Q., Luo, M., Xie, G. Y., Wang, H., et al. (2020). ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy. *Adv. Sci.* 7, 1902880. doi:10.1002/advs.201902880

Oh, S. C., Sohn, B. H., Cheong, J.-H., Kim, S.-B., Lee, J. E., Park, K. C., et al. (2018). Clinical and Genomic Landscape of Gastric Cancer with a Mesenchymal Phenotype. *Nat. Commun.* 9, 1777. doi:10.1038/s41467-018-04179-8

Roh, W., Chen, P.-L., Reuben, A., Spencer, C. N., Prieto, P. A., Miller, J. P., et al. (2017). Integrated Molecular Analysis of Tumor Biopsies on Sequential CTLA-4 and PD-1 Blockade Reveals Markers of Response and Resistance. *Sci. Transl. Med.* 9, eaah3560. doi:10.1126/scitranslmed.aah3560

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

Schlingensiepen, K., Schlingensiepen, R., Steinbrecher, A., Hau, P., Bogdahn, U., Fischerblass, B., et al. (2006). Targeted Tumor Therapy with the TGF-β2 Antisense Compound AP 12009. *Cytokine Growth Factor Rev.* 17, 129–139. doi:10.1016/j.cytogfr.2005.09.002

Shariat, S. F., Karakiewicz, P. I., Suardi, N., and Kattan, M. W. (2008). Comparison of Nomograms with Other Methods for Predicting Outcomes in Prostate Cancer: a Critical Analysis of the Literature. *Clin. Cancer Res.* 14, 4400–4407. doi:10.1158/1078-0432.CCR-07-4713

Subhash, V. V., Yeo, M. S., Wang, L., Tan, S. H., Wong, F. Y., Thuya, W. L., et al. (2018). Anti-tumor Efficacy of Selinexor (KPT-330) in Gastric Cancer Is

Dependent on Nuclear Accumulation of P53 Tumor Suppressor. *Sci. Rep.* 8, 12248. doi:10.1038/s41598-018-30686-1

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, 15545–15550. doi:10.1073/pnas.0506580102

Tauriello, D. V. F., Palomo-Ponce, S., Stork, D., Berenguer-Llergo, A., Badia-Ramentol, J., Iglesias, M., et al. (2018). TGFβ Drives Immune Evasion in Genetically Reconstituted Colon Cancer Metastasis. *Nature* 554, 538–543. doi:10.1038/nature25492

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.e14. doi:10.1016/j.immuni.2018.03.023

Vagenas, K., Spyropoulos, C., Gavala, V., and Tsamandas, A. C. (2007). TGFβ1, TGFβ2, and TGFβ3 Protein Expression in Gastric Carcinomas: Correlation with Prognostics Factors and Patient Survival. *J. Surg. Res.* 139, 182–188. doi:10.1016/j.jss.2006.10.005

Vergara, D., Simeone, P., Damato, M., Maffia, M., Lanuti, P., and Trerotola, M. (2019). The Cancer Microbiota: EMT and Inflammation as Shared Molecular Mechanisms Associated with Plasticity and Progression. *J. Oncol.* 2019, 1–16. doi:10.1155/2019/1253727

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

Xiao, L., Jian, H., Guan, L., Yu, C., Li, X., Mi, B., et al. (2016). Roles of TGF-β and CTGF in Hypobaric Hypoxia-Induced Right Cardiac Ventricular Fibrosis in Rats. *J. Third Mil. Med. Univ.* 38, 927–931. doi:10.16016/j.1000-5404.201510109

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, C., Jin, K., Tong, Y., and Cho, W. C. (2015). Therapeutic Potential of Cancer Stem Cells. *Med. Oncol.* 32 (6), 619. doi:10.1007/s12032-015-0619-6

Yang, B., Bai, J., Shi, R., Shao, X., Yang, Y., Jin, Y., et al. (2020). TGFB2 Serves as a Link between Epithelial-Mesenchymal Transition and Tumor Mutation Burden in Gastric Cancer. *Int. Immunopharmacol.* 84, 106532. doi:10.1016/j.intimp.2020.106532

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, 314. doi:10.1186/s12885-020-06814-4

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

Zhang, X., Ni, Z., Duan, Z., Xin, Z., Wang, H., Tan, J., et al. (2015). Overexpression of E2F mRNAs Associated with Gastric Cancer Progression Identified by the Transcription Factor and miRNA Co-regulatory Network Analysis. *PLoS One* 10, e0116979. doi:10.1371/journal.pone.0116979

Zhang, C., Zhang, X., Xu, R., Huang, B., Chen, A.-J., Li, C., et al. (2017). RETRACTED ARTICLE: TGF-β2 Initiates Autophagy via Smad and Non-smad Pathway to Promote Glioma Cells' Invasion. *J. Exp. Clin. Cancer Res.* 36, 162. doi:10.1186/s13046-017-0628-8

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 Han, Fang, Wang, Zhang and Xue. 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: 13 May 2022
doi: 10.3389/fgene.2022.880387

Amino Acid Metabolism-Related lncRNA Signature Predicts the Prognosis of Breast Cancer

*Yin-wei Dai[1†], Zhi-kai Wen[2†], Zhi-xuan Wu[1], Hao-dong Wu[1], Lin-xi Lv[3], Cong-zhi Yan[1], Cong-hui Liu[1], Zi-qiong Wang[1] and Chen Zheng[1]**

[1]Department of Breast Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [2]Department of Hepatopancreatobiliary Surgery, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [3]Wenzhou Medical University, Wenzhou, China

Background and Purpose: Breast cancer (BRCA) is the most frequent female malignancy and is potentially life threatening. The amino acid metabolism (AAM) has been shown to be strongly associated with the development and progression of human malignancies. In turn, long noncoding RNAs (lncRNAs) exert an important influence on the regulation of metabolism. Therefore, we attempted to build an AAM-related lncRNA prognostic model for BRCA and illustrate its immune characteristics and molecular mechanism.

Experimental Design: The RNA-seq data for BRCA from the TCGA-BRCA datasets were stochastically split into training and validation cohorts at a 3:1 ratio, to construct and validate the model, respectively. The amino acid metabolism-related genes were obtained from the Molecular Signature Database. A univariate Cox analysis, least absolute shrinkage and selection operator (LASSO) regression, and a multivariate Cox analysis were applied to create a predictive risk signature. Subsequently, the immune and molecular characteristics and the benefits of chemotherapeutic drugs in the high-risk and low-risk subgroups were examined.

Results: The prognostic model was developed based on the lncRNA group including LIPE-AS1, AC124067.4, LINC01655, AP005131.3, AC015802.3, USP30-AS1, SNHG26, and AL589765.4. Low-risk patients had a more favorable overall survival than did high-risk patients, in accordance with the results obtained for the validation cohort and the complete TCGA cohort. The elaborate results illustrated that a low-risk index was correlated with DNA-repair–associated pathways; a low *TP53* and *PIK3CA* mutation rate; high infiltration of $CD4^+$ T cells, $CD8^+$ T cells, and M1 macrophages; active immunity; and less-aggressive phenotypes. In contrast, a high-risk index was correlated with cancer and metastasis-related pathways; a high *PIK3CA* and *TP53* mutation rate; high infiltration of M0 macrophages, fibroblasts, and M2 macrophages; inhibition of the immune response; and more invasive phenotypes.

Abbreviations: AAM, Amino acid metabolism; BRCA, breast cancer; C-index, Harrell's concordance index; DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; OS, overall survival; ROC, received operating characteristic; ssGSEA, single-sample gene set enrichment analysis.

OPEN ACCESS

Edited by:
Lan Zhao,
Stanford University, United States

Reviewed by:
Iman Mamdouh Talaat,
University of Sharjah, United Arab Emirates
Mahmood Yaseen Hachim,
Mohammed Bin Rashid University of Medicine and Health Sciences, United Arab Emirates

***Correspondence:**
Chen Zheng
haopeng2810@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: 22 February 2022
Accepted: 25 April 2022
Published: 13 May 2022

Citation:
Dai Y-w, Wen Z-k, Wu Z-x, Wu H-d, Lv L-x, Yan C-z, Liu C-h, Wang Z-q and Zheng C (2022) Amino Acid Metabolism-Related lncRNA Signature Predicts the Prognosis of Breast Cancer. Front. Genet. 13:880387. doi: 10.3389/fgene.2022.880387

Conclusion: In conclusion, we attempted to shed light on the importance of AAM-associated lncRNAs in BRCA. The prognostic model built here might be acknowledged as an indispensable reference for predicting the outcome of patients with BRCA and help identify immune and molecular characteristics.

Keywords: amino acid metabolism, breast cancer, long non-coding RNA, prognostic signature, prognostic model, immunity

INTRODUCTION

Breast cancer (BRCA) is the most frequent female malignancy and is potentially life threatening. Moreover, BRCA has one of the highest lethality rates among the female malignant tumors (Siegel et al., 2019; Siegel et al., 2020; Siegel et al., 2021). The BRCA incidence rates continue to increase by about 0.5% per year (Siegel et al., 2020; Siegel et al., 2021). Currently, dozens of treatments, including surgery, hormonal therapy, radiation therapy, and chemotherapy, are used to manage female BRCA. Nevertheless, most patients with BRCA are still at risk of having adverse outcomes, even for those who receive therapy in the early stage of the disease (Ciriello et al., 2015). In recent years, researchers have dedicated increased efforts to proving that the AAM is dramatically associated with BRCA development (Cha et al., 2018; Deyu Zhang et al., 2021; Morotti et al., 2021) and have demonstrated that BRCA is infiltrated by many types of immune cells and own a high level of immunogenicity, illustrating the hypothesis that immune cell infiltration plays an indispensable role in the clinical prognosis of BRCA (Lehmann et al., 2011; Loi et al., 2014; Denkert et al., 2018; Pruneri et al., 2018). Moreover, the tumor immune microenvironment participates significantly in the development of BRCA. Tumor-infiltrating lymphocytes have been shown to be correlated with the outcome of this disease (Pruneri et al., 2018).

Tumor-infiltrating lymphocytes have been shown to be correlated with the outcome of this disease (Adams et al., 2014; Ali et al., 2016; Stanton and Disis, 2016; Lee et al., 2018).

Altered metabolism is a hallmark of cancer, and the reprogramming of the energy metabolism has historically been acknowledged as a general phenomenon underlying tumors (Pei-Hsuan Chen et al., 2019; Faubert et al., 2020; Wang et al., 2018). One of the best-known alternative theories on cancer development is the "Warburg effect," which consists of the continued activation of aerobic glycolysis in cancer cells (Hanahan and Weinberg, 2011). Furthermore, the AAM has been shown to be strongly correlated with the evolution and progression of human malignancies. Glutamine, serine, and glycine, for instance, are vital nutrients for tumor growth and maintenance. Moreover, myc overexpression influences the cellular glutamine levels by activating the transcription of GLS1 and the glutamine transporter SLC1A5 (Gao et al., 2009). In turn, phosphoglycerate dehydrogenase (PHGDH), which is a crucial enzyme in the serine synthesis pathway, is greatly upregulated in breast cancer cells (Dias et al., 2019). Similarly, immune cells require amino acids as a vital source of nutrition. The AAM can regulate immune effector protein activity (Kelly and Pearce, 2020). For example, activated $CD8^+$ T cells exhibit higher levels of Slc7a5 and Slc1a5 on the cell surface compared with naïve $CD8^+$ T cells. Activated T cells need a considerable amount of amino acids to maintain growth by improving transporter expression (Kelly and Pearce, 2020). Recent studies have demonstrated that the combination of a glutamine antagonist with anti-PD-1 therapy had a more obvious anti-tumor effect than did the anti-PD-1 therapy alone, and did not cause immune cell failure (Leone et al., 2019). In addition, the combination of the ladiratuzumab vedotin LIV-1 directed which is a transmembrane protein with zinc transporter and metalloproteinase activity (Nagayama et al., 2020) and anti-PD-1 therapy have been highlighted in breast cancer, particularly Triple negative breast cancer (TNBC) (Rizzo et al., 2022). There are many TNBC patients who benefit from immune checkpoint blockade therapy (Rizzo et al., 2021; Rizzo and Ricci, 2021).

LncRNAs, which are RNA molecules with a length of about 200 nt, can adjust the expression of genes (Tao Zhang et al., 2021). With the exception of gene regulation, lncRNAs participate in numerous biological regulatory processes, including those involved in the appearance, development, and metastasis of tumors (Gupta et al., 2010). Moreover, lncRNAs have a great impact on the regulation of metabolism (Guo et al., 2021; Tan et al., 2021). lncRNAs can directly modulate the posttranslational modification of key metabolic enzymes, lncRNAs can also indirectly regulate metabolic pathways through posttranslational modifications (Tan et al., 2021). It has been proved by experiments that the lncRNA XLOC_006390 stablized c-Myc by preventing its ubiquitination, increasing the expression of glutamate dehydrogenase 1 (GDH1), and subsequently stimulating the production of alpha-ketoglutarate (α-KG). Excess α-KG supplied the tricarboxylic acid (TCA) cycle and facilitated glutamate metabolism, promoting pancreatic cancer growth (He et al., 2020). Meanwhile, LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2 (Liao et al., 2019). However, to date, AAM-related lncRNAs have not been used to predict overall survival (OS) in patients with BRCA. Whether amino acid metabolization-related lncRNAs participate in the immune regulation of BRCA remains ambiguous.

In this study, we used the TCGA BRCA database, from which genomic and transcriptome data (RNA-seq) are accessible. In *cross*-validation analyses, the dataset was stochastically divided into training and *validation* groups. Finally, nine lncRNAs that were correlated with BRCA outcomes were identified. According to the expression levels of these nine lncRNAs, a risk score prognostic model for BRCA was created according to the patients in the training cohort, and the vital prognostic values of this model were further acknowledged in the patients in the validation cohort and the patients in the whole cohort. The relationships between the risk score subtypes and immune checkpoints, the proportions of 22 immune cells [B cells naive,

B cells memory, plasma cells, T cells CD8, T cells CD4 naive, T cells CD4 memory resting, T cells CD4 memory activated, T cells follicular helper, T cells gamma delta, T cells regulatory (Tregs), NK cells resting, NK cells activated, monocytes, macrophages M0, macrophages M1, macrophages M2, dendritic cells resting, dendritic cells activated, mast cells resting, mast cells activated, eosinophils, and neutrophils], and intrinsic molecular subtypes were also illustrated (Newman et al., 2015; Becht et al., 2016). Lastly, the difference between the two sets regarding sensitivity to chemotherapy was predicted.

METHODS

Data Source and Preprocessing

To collect the messenger RNA (mRNA) expression profiles and clinical information of patients with BRCA, comprehensive computerized searches of TCGA datasets (https://portal.gdc.cancer.gov/repository) were conducted. Samples with a follow-up time of less than 1 month or male BRCA samples were excluded. A total of 1208 TCGA female patients with BRCA for whom lncRNA expression profiles were available were analyzed in the present study. The AAM-related gene sets (REACTOME_METABOLISM_OF_AMINO_ACIDS_AND_DERIVATIVES) were derived from the Molecular Signatures Database v5.1 (MSigDB) (http://www.broad.mit.edu/gsea/msigdb/), which incorporated 374 genes in total. After assessing the overlap with genes in the TCGA RNA-seq datasets, 374 genes associated with AAM remained in study. The AAM-related lncRNAs were selected based on the criteria of $p < 0.001$ and Pearson's correlation coefficient $|> 0.4$, as assessed using the limma R package (Schober et al., 2018). Moreover, the "limma" R package was used to identify the differentially expressed genes (DEGs), including lncRNAs, protein-coding genes, miRNAs, etc., between nontumor and tumor tissues, with a false discovery rate (FDR) < 0.05 and $|\log2FC| \geq 1$.

Here, we first explored the function of both downregulated and upregulated AAM-related DEGs. Subsequently, we applied a gene ontology (GO) analysis to assess the biological pathways related to the DEGs. A further functional analysis of biological processes (BPs), molecular functions (MFs) and cellular components (CCs) adjusted to the individually expressed AAM-related DEGs was performed based on Kyoto Encyclopedia of Genes and Genomes (KEGG) data using the R software, ggplot2 package.

Development and Validation of the Prognosis Model

The 1,005 patients who had a survival time ≥30 days and complete follow-up information were stochastically split into the training and validation cohorts at a 3:1 ratio, for the building and validation of the accuracy of the prognostic model. A univariate Cox regression analysis was run to screen prognostic lncRNAs that were correlated ($p < 0.05$) with the overall survival (OS) of the patients in the training cohort. All 17 of these lncRNAs were further enrolled into a LASSO analysis for dimension reduction in the "glmnet" R package. Subsequently, a multivariate Cox analysis further selected nine lncRNAs according to the lowest Akaike information criterion value obtained for the 17 AAM-related lncRNAs with prognostic significance described above. The risk score of each patient derived from this prognostic signature was calculated based on the normalized expression level of AAM-related lncRNAs and corresponding regression coefficients. The computational formula used here was as follows:

$$\text{Risk Score} = e^{\text{sum}\left(\text{each lncRNA's expression} \times \text{corresponding regression coefficient}\right)}$$

The patients in the training cohort were split into high-risk and low-risk sets based on the comparison of the median value of the risk score and the OS between the different groups using a Kaplan–Meier (K–M) analysis with a log-rank test. Subsequently, a time-dependent receiver operating curve (ROC) curve analysis was performed using the "survivalROC" R package, to assess the predictive veracity of the prognosis model. Furthermore, we applied univariate and multivariate Cox regression analyses to verify if the nine AAM-related lncRNAs were independent prognostic factors for BRCA. For the purpose of validating this model, the same regression coefficients, formula, and genes were applied in the analysis of the validation cohort and the complete cohort, to calculate the risk score.

Comparison of the AAM-Related lncRNA Signature With Other BRCA Prognostic Models

To examine whether the model constructed here is more referential than are other BRCA prognostic models, we utilized a ROC for comparison with an 11-lncRNA signature (Shen et al., 2020), another 11-lncRNA [11 (2)-lncRNA] signature (Xiaoying Li et al., 2021), and an eight-lncRNA signature (Zhu et al., 2021). We obtained the correlative lncRNAs in these models from the literature, and the 1-, 3-, and 5-year OS ROC curves for the complete TCGA cohort were created. Finally, these lncRNA-based prognostic models were compared to illustrate the merits and shortcomings of each of them.

Building of an lncRNA–mRNA Co-expression Network

For the purpose of demonstrating the interaction between the nine-lncRNA signature and their interrelated mRNAs, we applied the Cytoscape software (version 3.8.2,http://www.cytoscape.org/) to design and visualize an mRNA–lncRNA co-expression network.

Gene Set Enrichment Analysis

We ran a gene set enrichment analysis (GSEA) 4.2.1 (https://www.gsea-msigdb.org/gsea/msigdb) (Subramanian et al., 2005) to distinguish various functional phenotypes between the high-risk and low-risk sets. The mRNA expression profiles of BRCA samples in the TCGA datasets, which were separated into two

groups based on risk score, were applied to KEGG gene sets. The study included enriched gene sets with $p < 0.05$, and 1,000 random sample permutations were acknowledged as being statistically significant. Their default values were used for the other parameters.

Establishment of Predictive Nomograms

A nomogram was created on the ground that the results of the multivariate analysis in the R software package using the nomogram function from the "rms" library, to predict the 1-, 3-, and 5-year survival of patients with BRCA. Harrell's concordance index (C-index) and calibration curve illustrate the predicting value of the nomograms and their discrimination performance.

Immune-Related Features

The MCPcounter (Wang et al., 2019), ESTIMATE (Yoshihara et al., 2013), CIBERSORT (Newman et al., 2015; Charoentong et al., 2017), and single-sample gene set enrichment analysis (ssGSEA) (Yi et al., 2020) algorithms were compared to estimate the differences in cell immune responses or cellular components between the low-risk and high-risk groups. We used a heatmap and boxplots to illustrate differences in the immune response using the various algorithms. Furthermore, for the purpose of quantifying the tumor-infiltrating immune cell subgroups in the BRCA tumor microenvironment (TME) among the two groups, as well as for estimating their immune function, we used ssGSEA. We retrieved dozens of potential immune checkpoints from the literature.

Drug Susceptibility and Mutation Analysis

For the purpose of determining the somatic mutations of patients with BRCA in the high- and low-risk sets, the mutation annotation format from the TCGA database was created utilizing the "maftools" R package. The tumor mutation burden (TMB), which is described as mutations/megabase (mutations/Mb), is an effective biomarker for predicting the efficacy of immunotherapy. The TMB score for each sample with BRCA in the two groups was calculated. For the purpose of exploring the differences in the responses to chemotherapeutic drugs between the two groups, we analyzed the semi-inhibitory concentration (IC_{50}) values of the chemotherapeutic drugs that are usually employed to treat BRCA using the "pRRophetic" package.

Statistical Analysis

We applied R version 4.1.0. to run all statistical analyses. Significance was set at $p < 0.05$.

RESULTS

Data Processing and Clinicopathological Features

A flow chart of the data analysis and process used in this study was drawn (**Figure 1**). After data partitioning and preprocessing, 765 patients with BRCA were distributed into the training set and 255 patients with BRCA were distributed into the validation set. Their clinicopathological features were outlined (**Table 1**).

Enrichment Analysis of AAM-Related Genes

The differential gene expression analysis (1,096 tumors vs. 112 normal samples) identified 55 DEGs, 18 of which were downregulated and 37 of which were upregulated in the tumor samples compared with the normal samples (**Supplementary Table S1**). The DEGs in the BP category were involved in the production of alpha-amino acid and cellular amino acid metabolic processes, among others; the DEGs in the MF category mainly participated in the regulation of the production of dioxygenase and the binding of vitamins; the DEGs in the CC category were primarily upregulated in the mitochondrial matrix and mitochondrial inner membrane pathways. The KEGG-based analysis revealed that the DEGs primarily participated in arginine and tryptophan metabolism, the biosynthesis of amino acids, proline metabolism, cysteine and methionine metabolism, tyrosine metabolism, etc. (**Figures 2A,B**).

AAM-Related lncRNA-Based Prognostic Signature

We identified 316 AAM-based lncRNAs (**Supplementary Table S2**). The univariate Cox analysis selected 17 AAM-associated lncRNAs (**Figure 2A**) that were included in the multivariate Cox analysis. Ultimately, nine lncRNAs (LIPE-AS1, AC124067.4, LINC01655, AP005131.3, AC015802.3, USP30-AS1, SNHG26, and AL589765.4) were selected as independent prognosis predictors of BRCA. Thus, we constructed a prognostic index for training cohort cancer samples using the following formula: risk score = (−0.606072 × expression of LIPE-AS1) + (−0.28451 × expression of AC124067.4) + (0.6666797 × expression of LINC01655) + (−0.988819 × expression of AP005131.3) + (−0.140664 × expression of AC008115.3) + (−0.767441 × expression of AC015802.3) + (−0.277495 × expression of USP30-AS1) + (−0.765508 × expression of SNHG26) + (0.138989 × expression of AL589765.4) (**Table 2**).

Survival Results and Multivariate Analysis

The K-M survival analysis also showed that patients with BRCA in the high-risk group exhibited a lower OS ($p < 0.001$, **Figure 3B**). Concomitantly, The AUC of the signature lncRNAs was 0.881, illustrating a better predictive effect than that of the traditional clinicopathological features (**Figure 3C**). Interestingly, our heatmap showed that most of the novel lncRNAs exhibited a negative association with our risk model; additional experiments are necessary to address this issue (**Figure 3D**). The predictive value of the AUC of the novel lncRNA signature regarding the 1-, 3-, and 5-year survival rate was 0.881, 0.766, and 0.713, respectively (**Figure 3E**).

The hazard ratio and 95% CI of the risk score were 1.260 and 1.191–1.333 in the univariate Cox regression analysis ($p < 0.001$) and 1.242 and 1.174–1.315 in the multivariate Cox regression analysis ($p < 0.001$), respectively. These findings indicated that the nine-lncRNA signature was an independent prognosis factor of OS in patients with BRCA (**Figures 3F,G**).

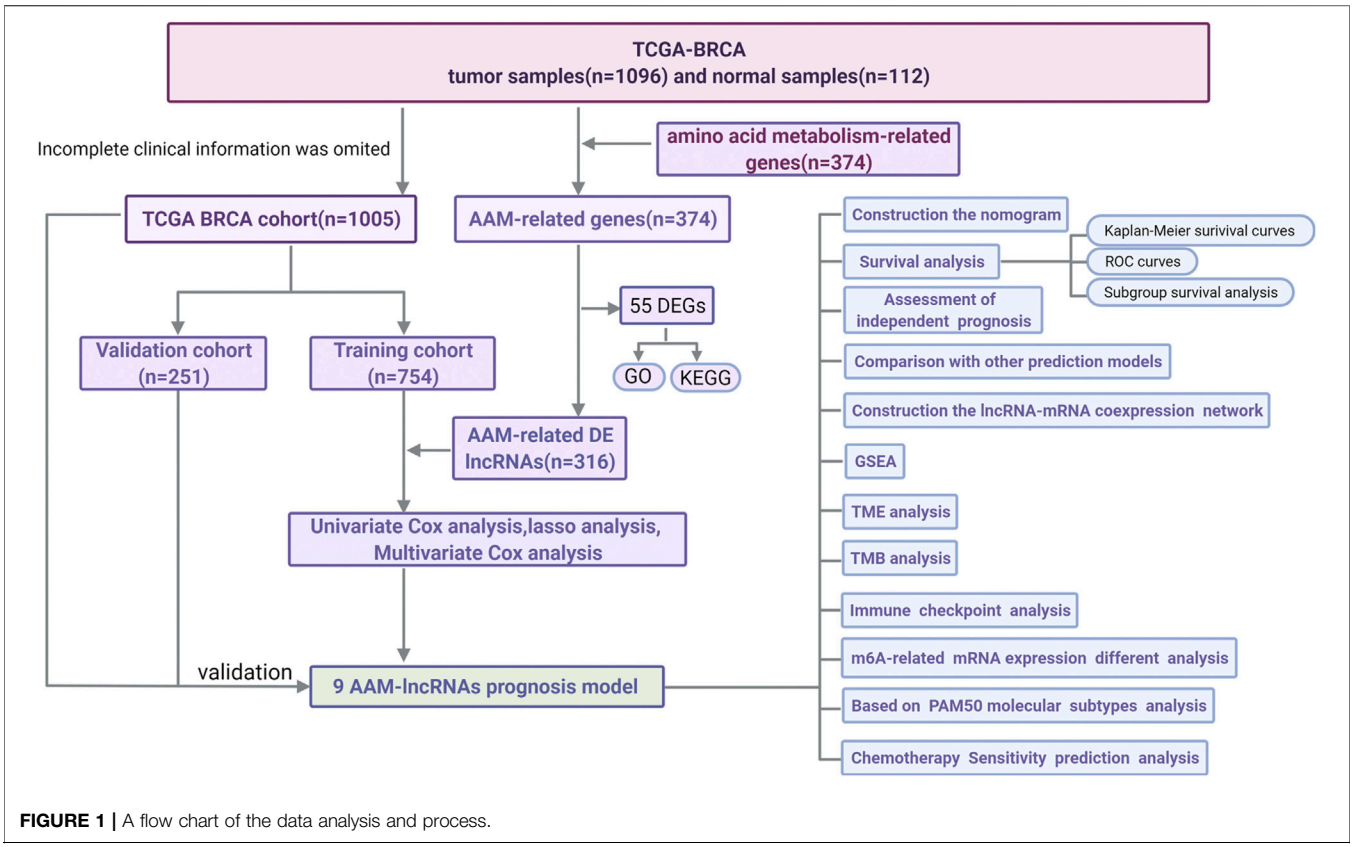


FIGURE 1 | A flow chart of the data analysis and process.

TABLE 1 | Clinicopathological features of the training set and the validation set.

Characteristic	TCGA set	Train set	Validation set	*p* value
Total cases	1,005	754	251	—
Age (years)	—	—	—	0.63
>65	278 (27.6%)	210 (27.9%)	68 (27%)	—
≤65	727 (72.3%)	544 (72.1%)	183 (72.9%)	—
Stage	—	—	—	0.21
I	186 (18.5%)	136 (18.0%)	50 (19.9%)	—
II	582 (57.9%)	433 (57.4%)	149 (59.3%)	—
III	222 (22.1%)	172 (22.8%)	50 (19.9%)	—
IV	15 (1.5%)	13 (1.7%)	2 (0.8%)	—
Fustat	—	—	—	0.058
Alive	892 (88.6%)	661 (87.7%)	231 (92.0%)	—
Dead	112 (11.4%)	93 (12.3%)	20 (8.0%)	—
Futime	—	—	—	0.3
>3 years	405 (40.3%)	315 (41.8%)	90 (35.9%)	—
≤3 years	600 (59.7%)	439 (58.2%)	161 (64.1%)	—

Figure 4A depicts the correlation between mRNAs and lncRNAs. The heatmap of the relationship between the clinicopathological manifestations and the AAM-related lncRNA prognostic index is also presented (**Figure 4B**). The hybrid nomogram (c-index = 0.754) encompassing the novel AAM-related lncRNA prognostic index and the clinicopathological characteristics is shown in **Figure 5A**. The calibration curve analysis revealed that the practical and the predicted 1-, 3-, and 5-year survival rates agreed with the reference curve (**Figure 5B**). These results suggested that the nomogram was accurate and stable; thus, it is suitable for implementation in the clinical management of patients with BRCA.

We ran a similar survival analysis by regulating the risk model using different physiological and clinical factors (e.g., age and tumor TNM stage). The K–M curves illustrated that the low-risk set had a better OS than the high-risk group in all subsets (**Figures 5C–L**). The K–M survival curves of the M1-stage subgroup were not statistically significant (p = 0.244). We considered that the fact that the number of patients was significantly low (only 20 samples) contributed to this observation; however, in general, the high-risk set had a worse OS than the low-risk set.

To confirm the prognostic accuracy of the risk score, the risk score each patient in the validation set and in the entire TCGA set was calculated and then split into two sets according to the median value. A survival analysis revealed a more favorable outcome in the low-risk set compared with the high-risk set (log-rank test; p < 0.001; **Supplementary Figures S1A,B**). An analysis of the 1-, 3-, and 5-year prognostic prediction classification efficiencies suggested that the risk score still had comparably high AUC values (**Supplementary Figures S1C,D**), suggesting that the risk model had an outstanding ability to predict the outcome of BRCA.

Comparison of the AAM-Related lncRNA Signature With Other BRCA Prognostic Models

To determine if our nine-lncRNA signature is more accurate than other BRCA prognostic models, we compared it with an 11-

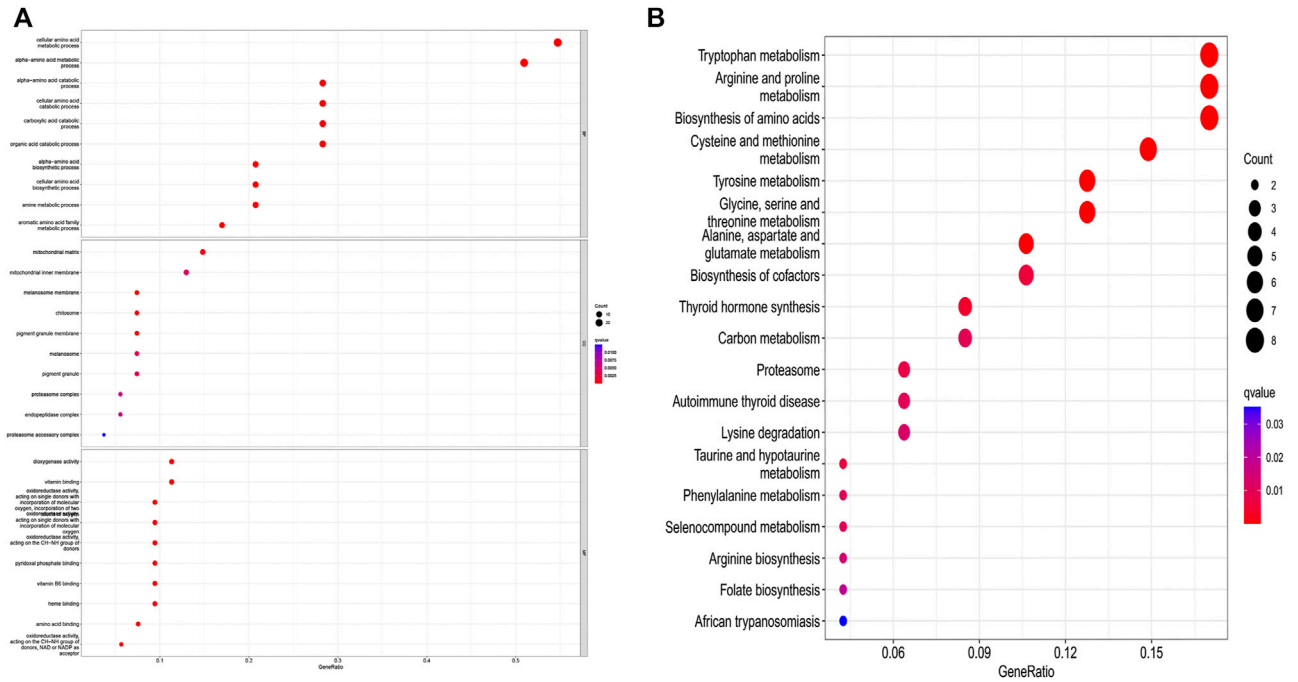


FIGURE 2 | GO and KEGG analysis of amino acid metabolism-associated DEGs. **(A)** GO and **(B)** KEGG.

TABLE 2 | Amino acid metabolism-related lncRNA-based prognostic signature.

Id	Coef	HR	HR.95L	HR.95H	*p* value
LIPE-AS1	−0.606072	0.5454891	0.2695349	1.10397	0.091994
AC124067.4	−0.28451	0.7523826	0.5595592	1.0116527	0.0596629
LINC01655	0.6666797	1.9477593	0.9994916	3.7956963	0.0501749
AP005131.3*	−0.988819	0.3720158	0.1479188	0.9356202	0.0356073
AC008115.3	−0.140664	0.8687813	0.7484818	1.0084158	0.0643486
AC015802.3	−0.767441	0.4641993	0.1671622	1.2890539	0.1408279
USP30-AS1*	−0.277495	0.7576792	0.5893708	0.9740519	0.0303807
SNHG26*	−0.765508	0.4650974	0.2367205	0.9138019	0.026313
AL589765.4***	0.138989	1.1491115	1.0659176	1.2387986	0.0002892

*p < *0.05*,**p < *0.01*,***p < *0.001*.

lncRNA signature (Shen et al., 2020), another 11-lncRNA (11 (2)-lncRNA) signature (Xiaoying Li et al., 2021), and an eight-lncRNA signature (Zhu et al., 2021) for the entire TCGA cohort. However, the predictive accuracy of the AAM-related nine-lncRNA signature was greater than that of the remaining three prognostic models (**Supplementary Figures S2A–H**).

Molecular Characteristics of the Different Risk Subgroups

GSEA was used to perform functional annotations in the two groups. The gene sets of the low-risk samples were enriched in immune-related pathways (**Figure 6B**), whereas the gene sets of the high-risk samples were enriched in lipid and glucose metabolism pathways (**Figure 6A**; $p < 0.05$).

Next, we found that the risk score was slightly correlated with the TMB ($r = 0.14$, $p < 0.001$), as shown in **Figures 6C,D**.

Moreover, to shed light on the immunologic nature of the risk subgroups, gene mutations were explored in the different risk subgroups was explored. We noticed an apparently more frequent mutation in the high-risk subgroup vs. the low-risk subgroup ($p = 0.0049$, t-test) (**Figure 6E**). Moreover, the most frequent mutation type was missense mutation, followed by frameshift deletion and nonsense mutation. The top 20 genes with the highest mutation rates in the subgroups are illustrated in **Figure 6E**. The mutation rates of the TP53, PIK3CA, TTN, and CDH1 genes were higher than 10% in both groups. Mutation of the MAPK3K1 gene was more prominent in the high-risk subgroup, whereas mutation of the MUC4 gene was more prominent in the low-risk subgroup.

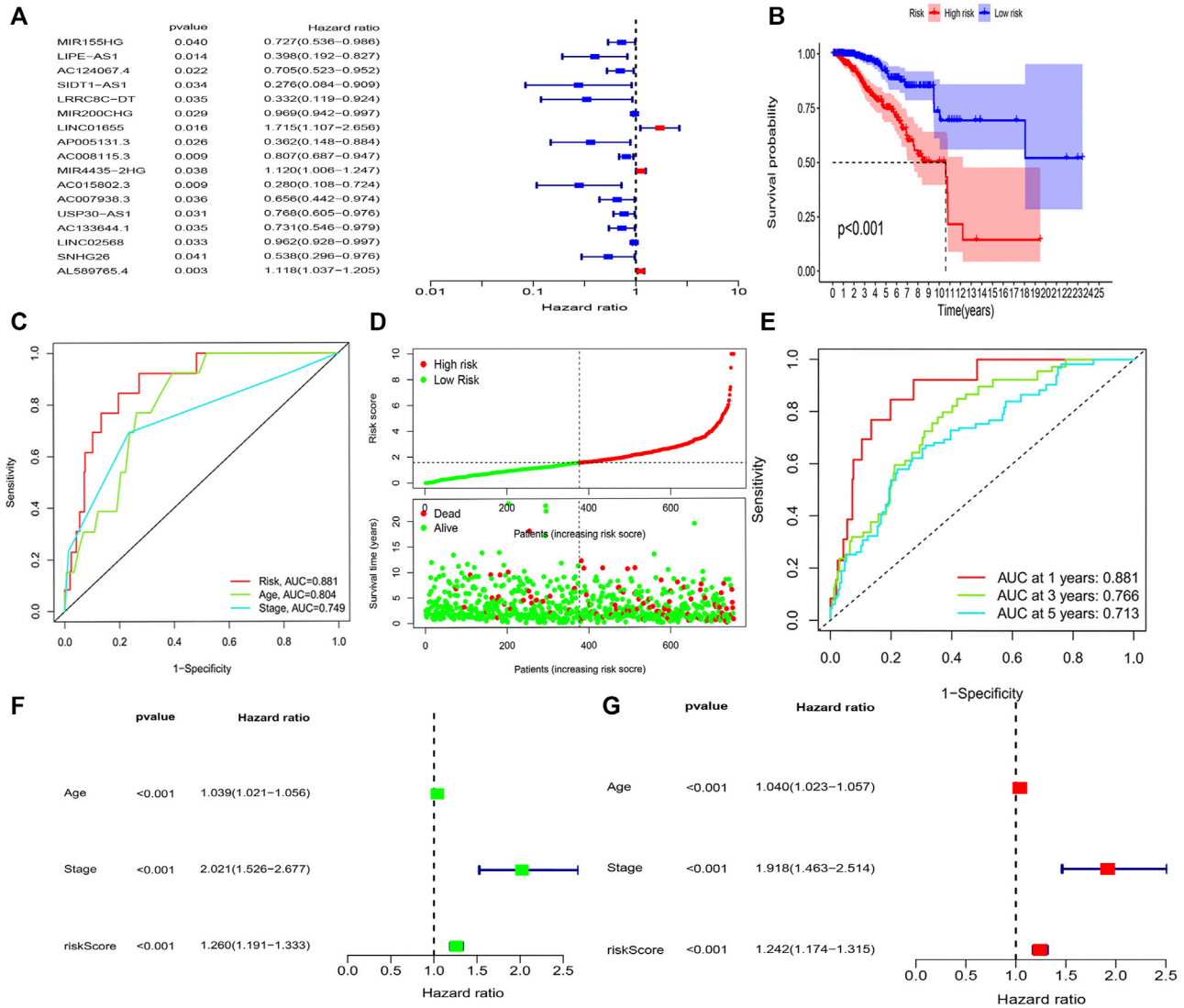


FIGURE 3 | Amino acid metabolism-associated lncRNA signature based on training sets. **(A)** Univariate cox analysis **(B)** Kaplan–Meier curves, **(C)** multi-index ROC analysis, **(D)** risk score, and **(E)** time-dependent ROC analysis. Univariate and multivariate Cox analyses of the expression of AAM-related lncRNAs. **(F)** Univariate and **(G)** multivariate analyses.

Immune Characteristics of the Different Risk Subgroups

To examine the composition of immune cells in the different risk subgroups, we used the Wilcoxon test to compare the distribution of immune cells among the different risk subgroups. We found that $CD8^+$ T cells, plasma cells, naïve $CD8^+$ T cells, M1 macrophages, memory B cells, and endothelial cells were more abundant in the low-risk subgroup, whereas M2 macrophages, M0 macrophages, and fibroblasts were more abundant in the high-risk subgroup (**Figure 7B**, **Supplementary Figure S3**). A low-Risk score was also deeply correlated with a high immune score (**Figure 7A**). We also investigated whether the prognostic value of risk scores stemmed from better immune control or from less-aggressive cancer growth. As shown in **Figure 8A**, we found that patients with a higher score on HLA, checkpoint, inflammation-promoting, parainflammation, T-cell co-inhibition, T-cell co-stimulation, type II IFN response, cytolytic activity, MHC-class-I, and type I IFN response had a better outcome. Therefore, we suggest that the prognostic value of the risk scores might result from both better immune control and less-aggressive cancer growth. The difference in the expression of immune checkpoints between the two subsets was further explored. The results of this analysis suggest a more abundant expression of PDCD-1 (PD-1), CD274 (PD-L1), BTLA, TIGIT, CTLA4, PDCD1LG2, and LAG3, among others, in the low-risk subsets compared with the high-risk subsets (**Figure 8C**). **Figure 8B** shows that, compared with the expression of m6A-related mRNAs between the low- and high-risk groups, the expression of RBM15, YTHDC2, WTAP, METTL14, YTHDC1, and METTL3 was differentiated.

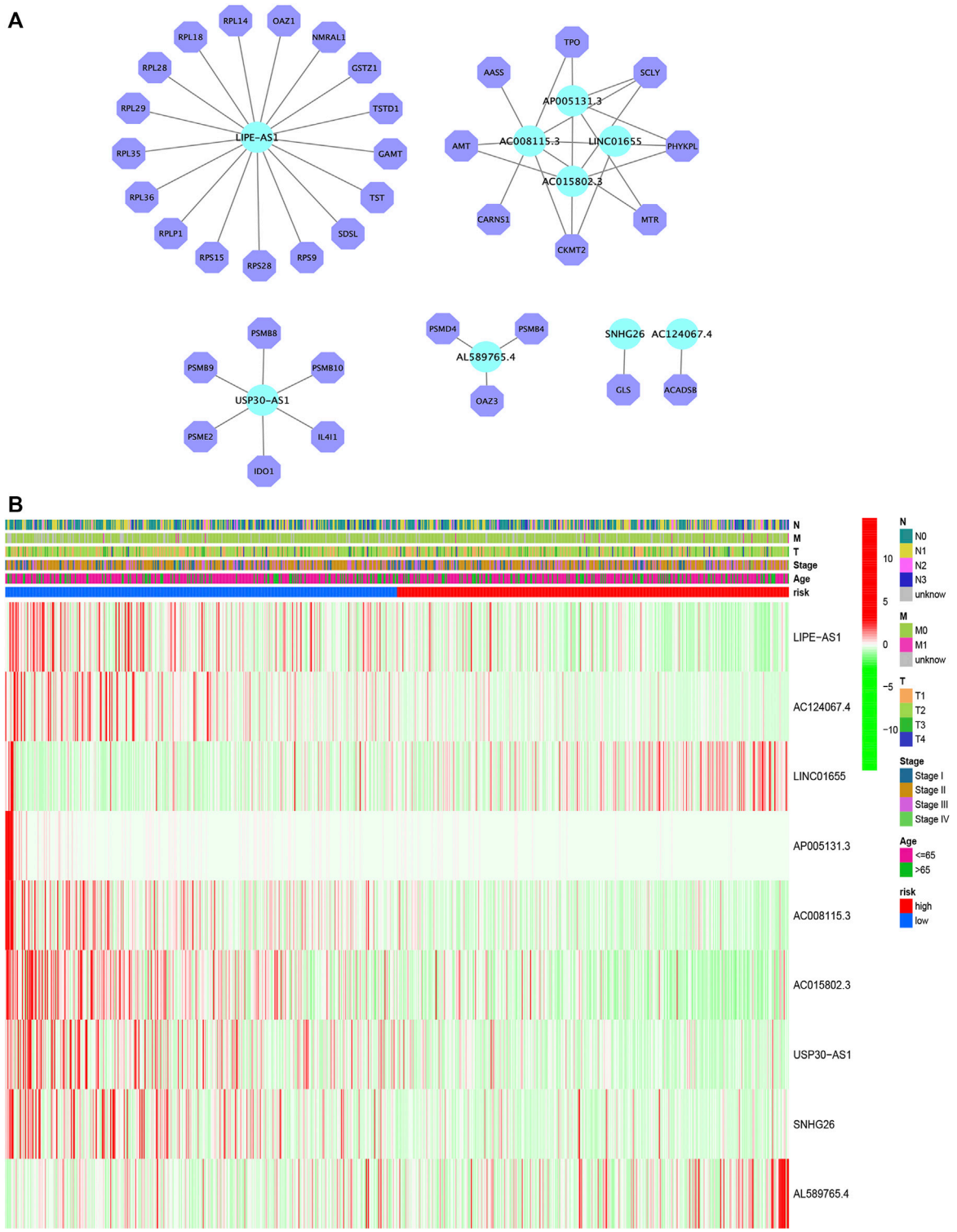


FIGURE 4 | Construction of the mRNA–lncRNA regulatory network **(A)**. Heatmap of the clinicopathological manifestations and AAM-related lncRNA prognostic signature **(B)**.

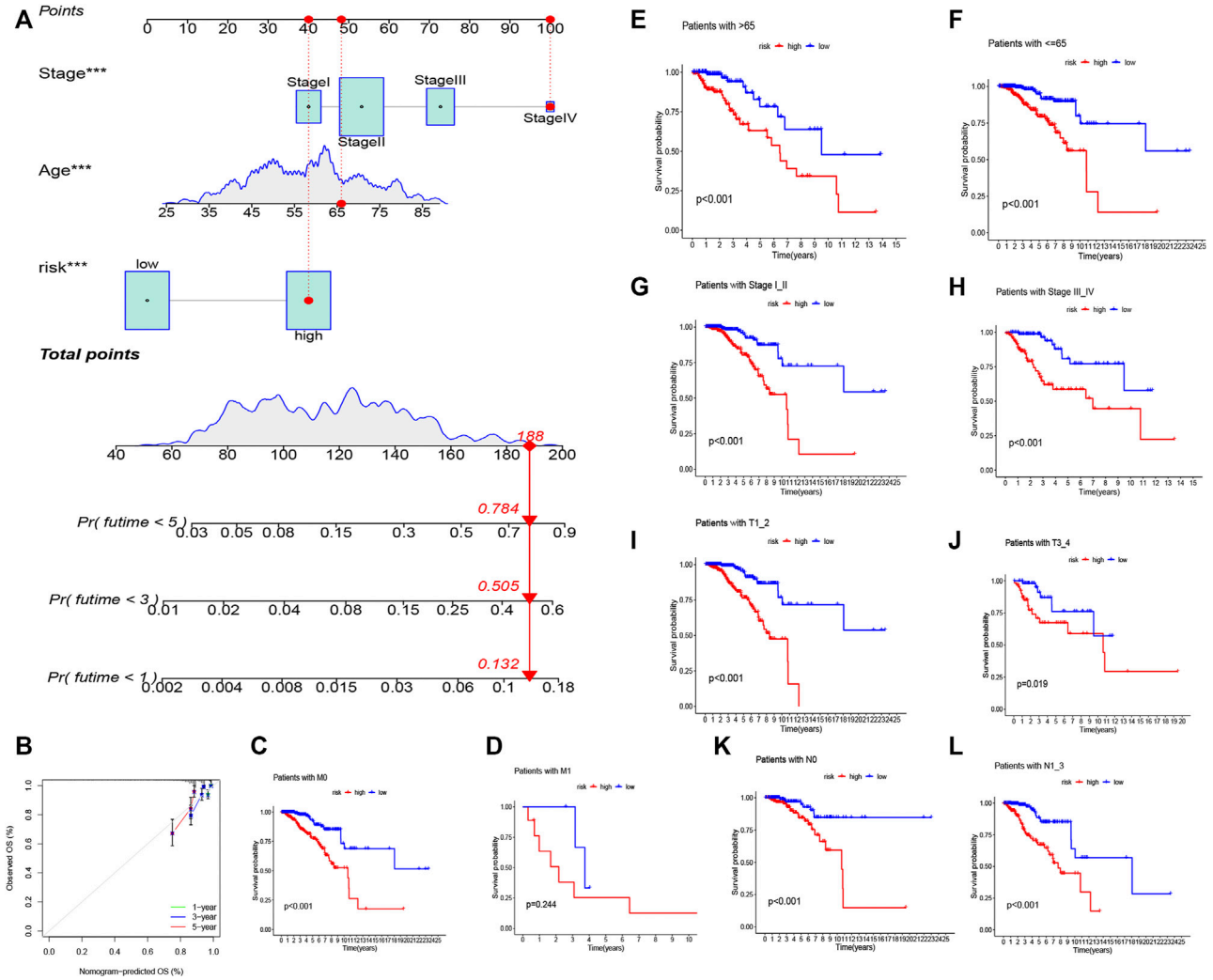


FIGURE 5 | Nomogram of both prognostic AAM-associated lncRNAs and clinical–pathological factors **(A)**. Calibration plot for the nomogram **(B)**. Stratification analysis of the risk score in BRCA. **(C,D)** Age (age >65 and age ≤65 years). **(E,F)** Tumor stage (I–II or III–IV). **(G,H)** Tumor T stage (T1–2 or T3–4). **(I,J)** Tumor M stage (M0 or M1). **(K,L)** Tumor N stage (N0 or N1–3).

Correlation Between Risk Grouping and PAM50 Molecular Subtypes

The high-risk subsets and low-risk subsets were split into PAM50 molecular subtypes (Parker et al., 2009), respectively, including basal, Her2+, luminal A, luminal B, and triple-negative breast cancer (TNBC), as shown in **Figure 9A**. In summary, the proportion of TNBC samples was almost equally distributed between the two groups, whereas there were more HER2+ samples and more luminal A samples in the high-risk subgroup ($p < 0.001$, $\chi 2$ test).

Chemotherapy Sensitivity Related to the Risk Score

The correlation between the sensitivity to chemotherapeutic drugs and this prognostic model was explored. The IC50 values of usual chemotherapeutic drugs were predicted and compared between the low- and high-risk groups. Patients in the low-risk group were more responsive to gefitinib, epothilone B, and doxorubicin, whereas patients in the high-risk set were more responsive to docetaxel and Lapatinib. However, no statistical significance was observed regarding the differences in the response to paclitaxel (**Figures 9B–G**).

DISCUSSION

In recent years, researchers have dedicated increased efforts toward proving that the amino acid metabolism (AAM) is dramatically associated with BRCA development (Cha et al., 2018; Deyu Zhang et al., 2021; Morotti et al., 2021). Glutamine, serine, glycine, etc. are critical nutrients for tumor growth and maintenance. Similarly, amino acids are vital nutrients for

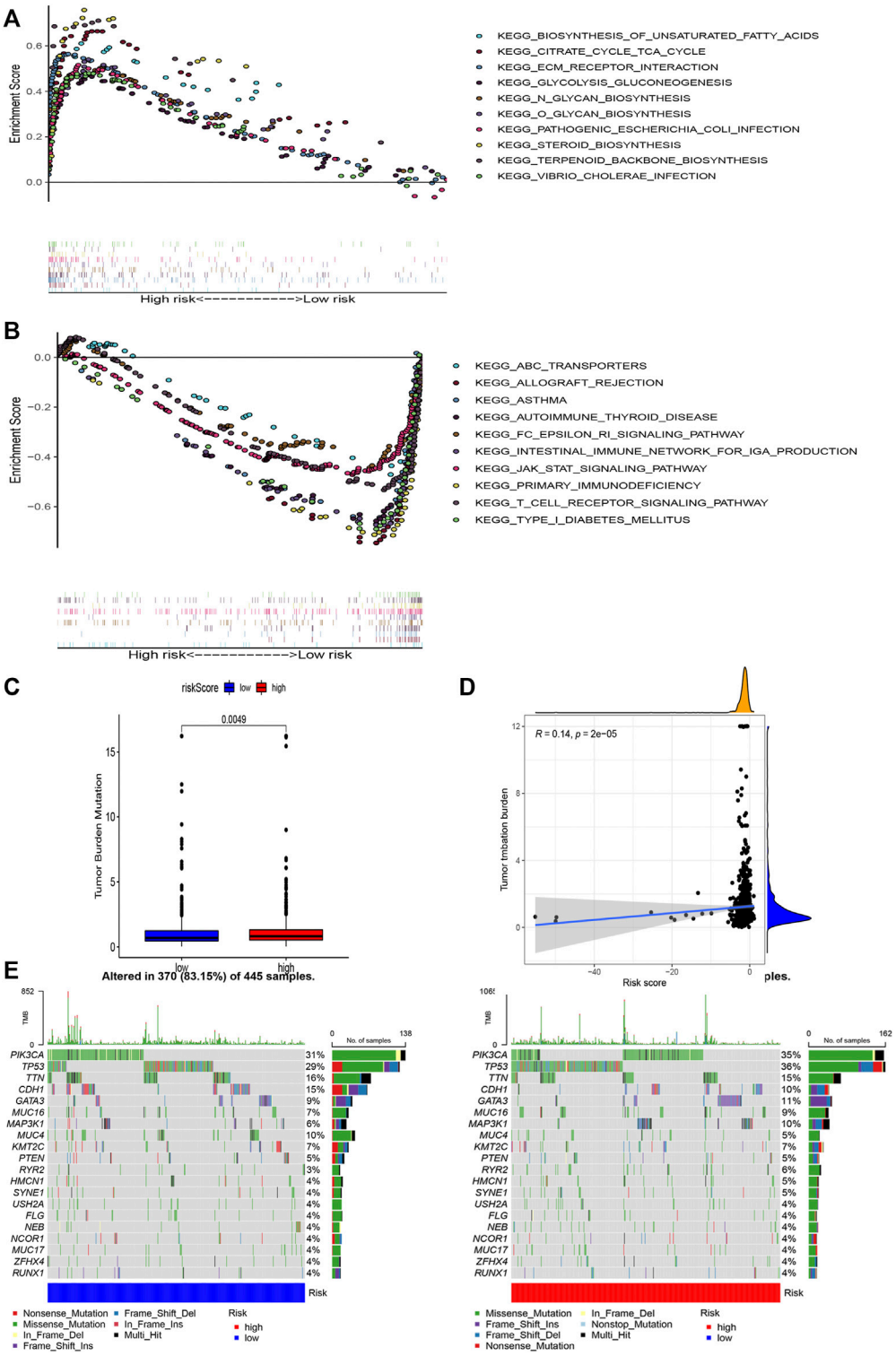


FIGURE 6 | Gene enrichment analysis for AAM-related lncRNAs based on TCGA in the high **(A)** and low **(B)** BRCA risk groups. Correlation between the TMB and the two risk subsets **(C)**. Association between the TMB and risk score **(D)**. Prominently mutated genes in the patients with BRCA in the different risk subgroups. The mutated genes (rows, top 20) are ranked according to mutation rate; samples (columns) are arranged to emphasize the mutual exclusivity among mutations. The right panels depicts the mutation percentage, and the top panel indicates the overall number of mutations. The color coding indicates the mutation type **(E)**.

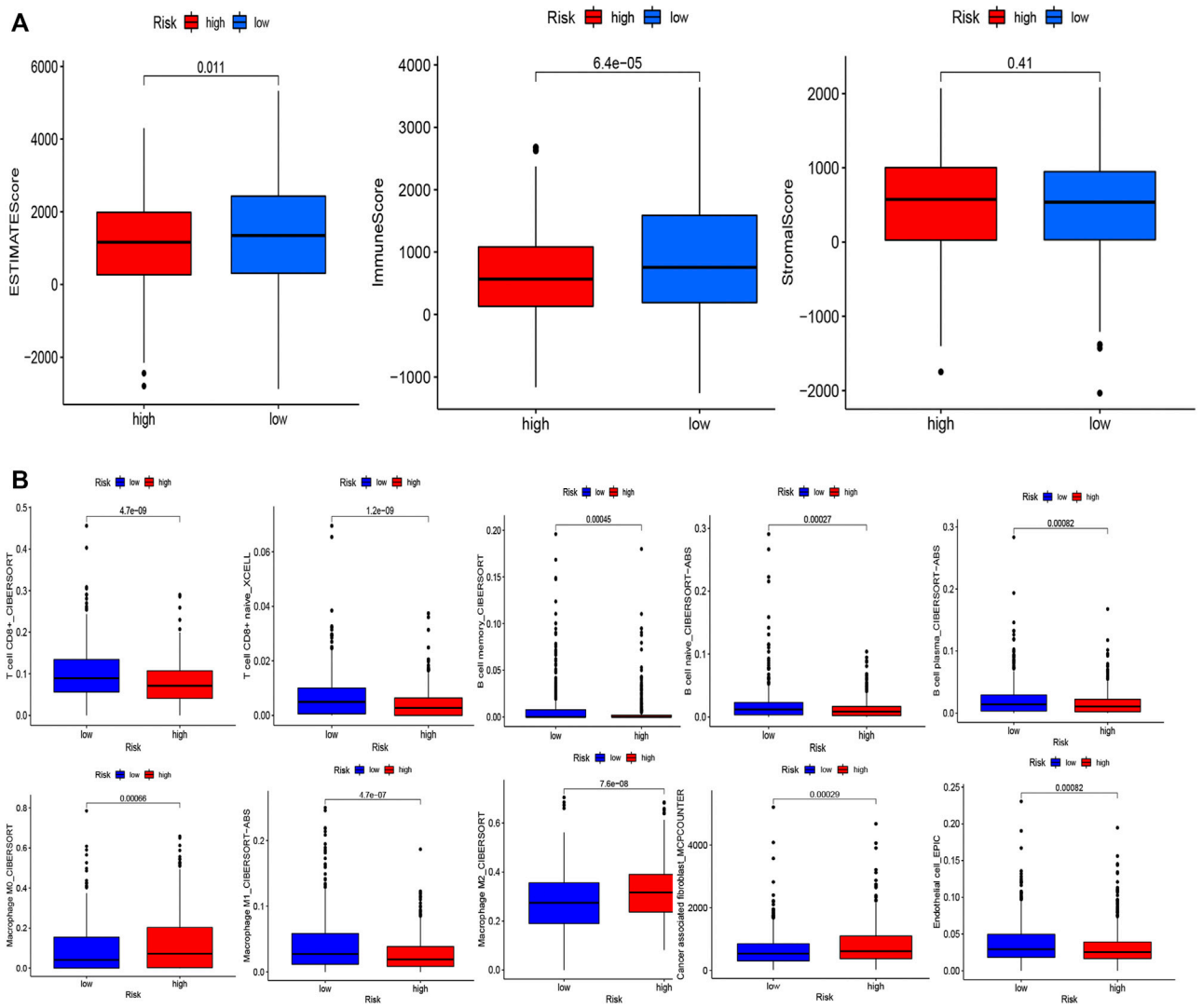


FIGURE 7 | Evaluation of the TME and levels of lymphocyte infiltration in the two groups. **(A)** Associations between the risk score and the immune and stromal scores. **(B)** Associations between the risk score and immune cell types.

immune cells. For example, Glutaminolysis is a major energy-producing process for proliferating cells, including activated T cells (Newsholme et al., 1999), by supplying a-ketoglutarate (aKG) to the TCA cycle, *via* glutamate. Glutamine is used to promote LPS induction of IL-1 production by macrophages (Wallace and Keast, 1992). Activated NK cells use some glutamine to replenish TCA cycle intermediates and increase oxidative phosphorylation (Lam et al., 2016). One function of the reduced form of glutathione (GSH) in Treg cells is to restrict serine metabolism in order to maintain their suppressive function. Macrophages also utilize serine to generate glycine for GSH, needed for LPS-induced IL-1b mRNA expression (Rodriguez et al., 2019). Consequently, increasing attention has been paid to AAM in this context (Jones et al., 2018; Butler et al., 2021). As expected, recent studies have reported that the combination of a glutamine antagonist and anti-PD-1 therapy had a more obvious anti-tumor effect than did the anti-PD-1 therapy alone, and did not cause immune cell failure (Leone et al., 2019). Moreover, previous studies have produced prognostic models of genes related with amino acid metabolism in glioma and hepatoma *via* bioinformatics analysis (Liu et al., 2019; Zhao et al., 2021). In recent years, lncRNAs have been proved that they have a great impact on the regulation of metabolism (Guo et al., 2021; Tan et al., 2021). To date, however, AAM-related lncRNAs have not been used to predict OS in patients with BRCA. Whether amino acid metabolization-related lncRNAs participate in the immune regulation of BRCA remains ambiguous.

In this study, we explored for the first time the characteristics of AAM-related lncRNAs in BRCA and established a risk signature associated with OS. First, we screened AAM-

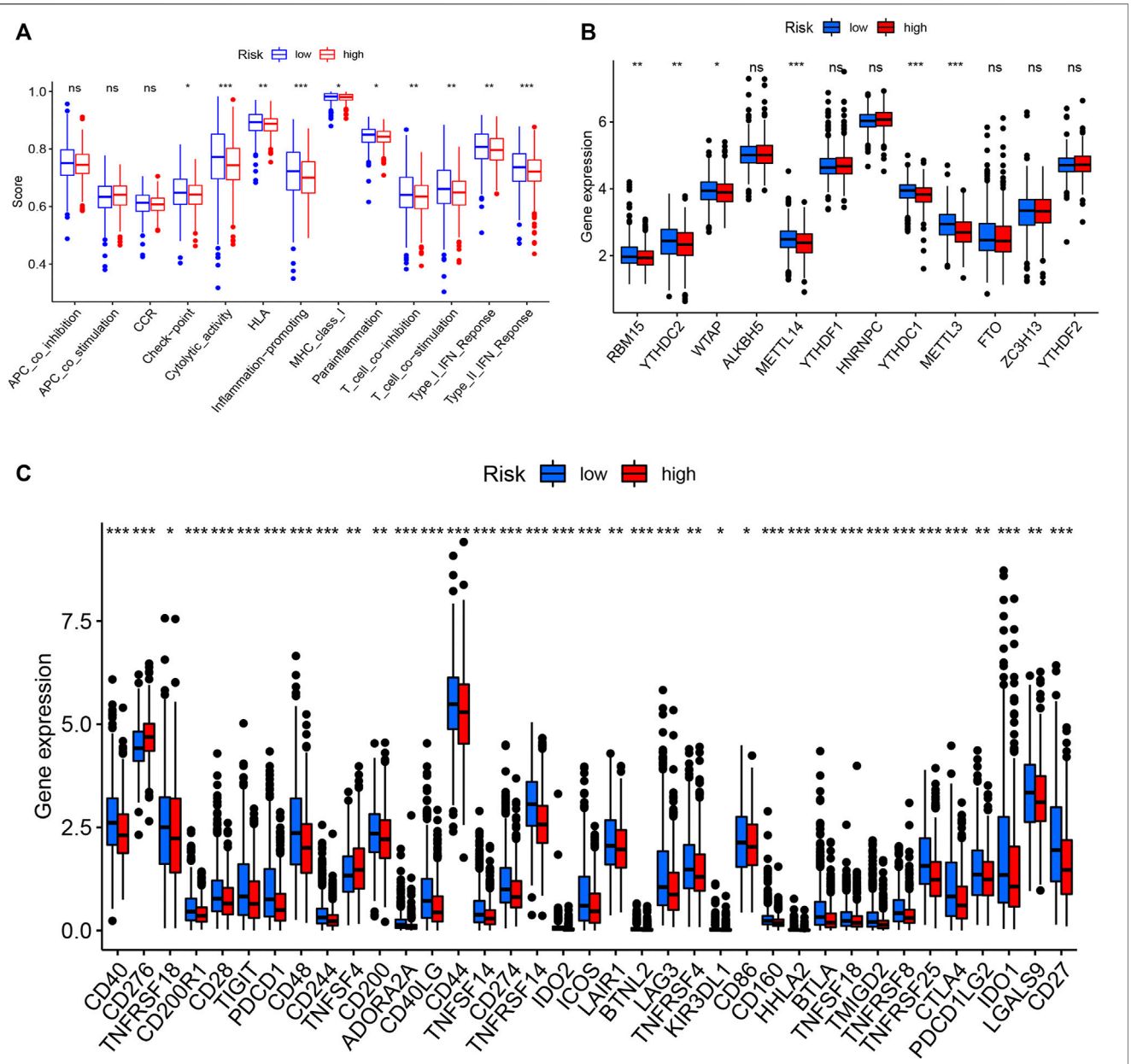


FIGURE 8 | Immune cell infiltration levels and corresponding function determined by ssGSEA **(A)**. Expression of m6A-related genes in both groups **(B)**. Expression of immune checkpoint-related genes in both groups **(C)**.

associated DEGs between BRCA and normal breast tissues based on RNA-seq data. Furthermore, we established a robust and effective prognostic signature using univariate Cox, LASSO regression, and multivariate Cox analyses. Nine lncRNAs were included in this signature (LIPE-AS1, AC124067.4, LINC01655, AP005131.3, AC015802.3, USP30-AS1, SNHG26, AL589765.4). Moreover, 1,005 samples in total were stochastically split into training and validation cohorts at a 3:1 ratio, for building and validating the AAM-related lncRNA signature. The training set of 754 samples, the validation set of 251 samples, and the whole TCGA data set of 1,005 samples all showed the feasibility of this model.

Several recent studies have demonstrated that LIPE-AS1 is prominently expressed in BRCA and cervical squamous cell carcinoma and is associated with a higher survival rate (Xu et al., 2021; Wang et al., 2022). In turn, AC124067.4-hsa-miR-92b-3p (hsa-miR-589-5p)-PHYHIPL both decrease the MSI and TMB in COAD, thus reducing the risk of genome instability and alterations (Ren et al., 2021), whereas USP30-AS1 promotes mitochondrial quality control in glioblastoma cells (Wang et al., 2021), the USP30-AS1/miR-299-3p/PTP4A1 pathway aggravates the malignant progression of cervical cancer (Wang et al., 2021), and SNHG26 promotes the metastasis, growth, and cisplatin resistance of tongue squamous cell carcinoma through the PGK1/Akt/mTOR signaling

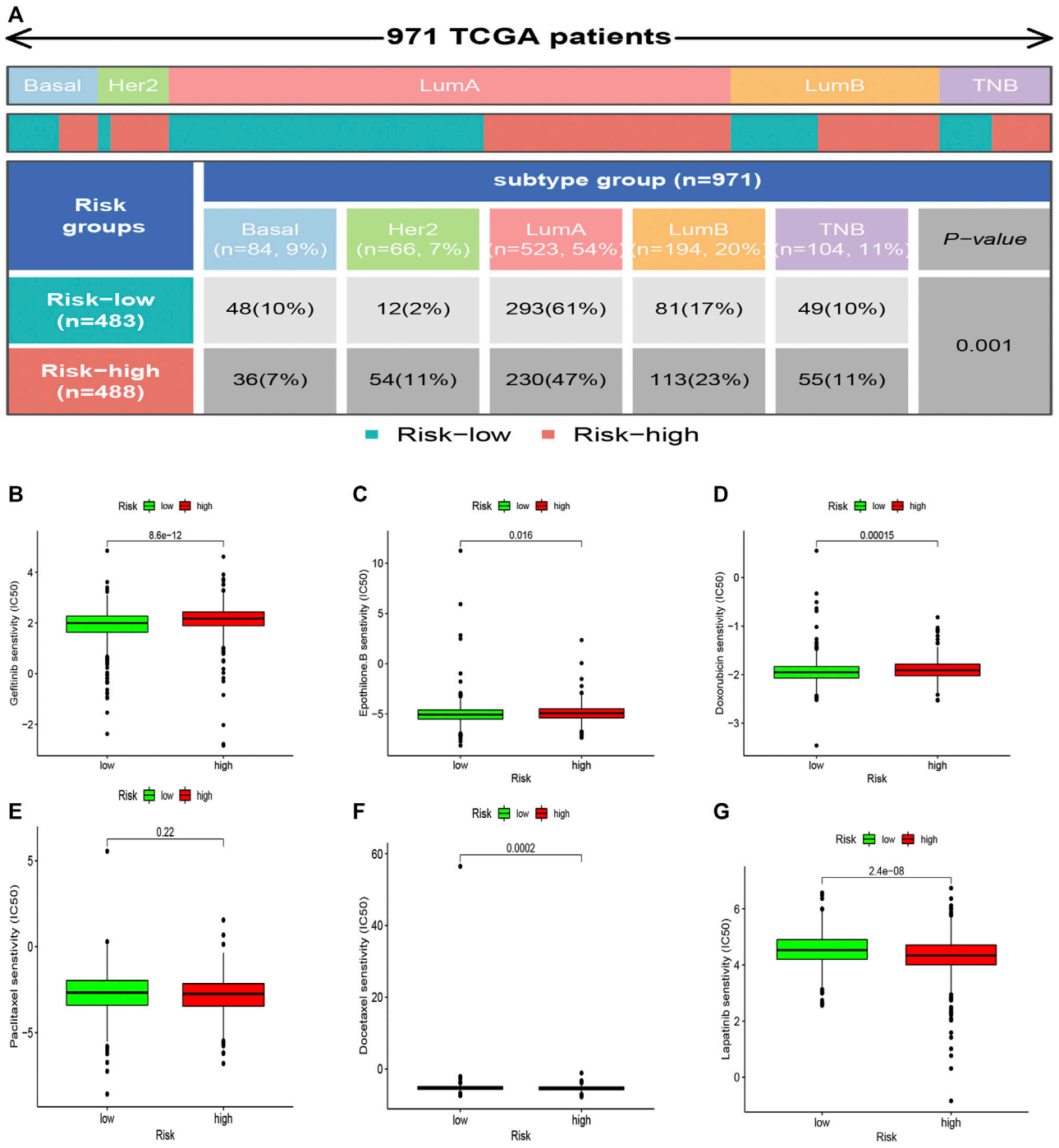


FIGURE 9 | Heatmap and table showing the distribution of the BRCA PAM50 molecular subtypes (basal, luminal A, luminal B, HER2$^+$, and TNBC) in the risk subgroups **(A)**. Relationships between the risk score and chemotherapeutic sensitivity **(B–G)**.

pathway (Jiang et al., 2022). Nevertheless, no studies have investigated the prognostic value of LINC01655, AP005131.3, AC015802.3, and AL589765.4 in patients with BRCA or other malignancies. The current findings show for the first time that these four lncRNAs were associated with the prognosis of BRCA. The potential role of the four lncRNAs needs to be further explored.

To shed light on the immunologic nature of the risk subgroups, gene mutations were explored in the different risk subgroups. We found that the most frequent mutation type was missense mutation, followed by frameshift deletion and nonsense mutation, as reported previously (2012). The most significant difference in mutations between the groups was observed for

TP53 mutations, which were more frequent in high-risk samples than in low-risk samples (77 vs. 64%). TP53 mutation is not only the single most significant genetic event in cancer, but also linked with poorer patient outcomes and more aggressive disease in many malignant tumors (Vousden and Prives, 2005), particularly BRCA (Olivier et al., 2006; Shahbandi et al., 2020). TP53 can affect the cancer cell cycle through the p53/TGF β signaling pathway. In a study of stage III breast cancer, patients with TP53 mutations were shown to have worse disease free survival (DFS) following treatment with paclitaxel ($p = 0.007$) (Chrisanthar et al., 2011). Olivier et al. noted in a study of 1794 breast cancer patients that those with tumors harboring TP53 mutations in exons 5–8 of the gene had a worse risk of dying of breast cancer within 10 years following surgery ($p < 0.0001$) (Olivier et al., 2006).

In addition, there was a higher rate of PIK3CA mutation in the high-risk subgroup compared with the low-risk subgroup, which could mean that tumor growth in high-risk BRCA cases is promoted through the PI3K–AKT signaling pathway (Xing et al., 2019). Hence, high-risk patients with high TP53 and PIK3CA mutations have a poorer prognosis than do low-risk patients with low TP53 and PIK3CA mutations, in accordance with our survival results. Next, we speculated that low-risk patients benefit more from the immune checkpoint inhibitor therapy, based on the results presented in **Figure 8C**, especially among patients with TNBC cancer in both groups (Lyons, 2019; Kwapisz, 2021).

Here, we concluded that the risk score had a slight positive relationship with the TMB (**Figure 6D**), which implies that the TMB can help explain why the risk score affects prognosis to a certain degree; however, other possible mechanisms may be involved in this relationship. Recent studies reported that TMB-high tumors not only did not exhibit a higher susceptibility to immune checkpoint blockade (ICB) vs. TMB-low tumors, but also exhibited a significantly lower susceptibility to ICB in BRCA, prostate cancer, etc. (McGrail et al., 2021). Interestingly, high expression levels of CTLA-4 and TIGIT were correlated with favorable prognosis in breast cancer (Fang et al., 2020). These are consistent with our results. Furthermore, we shed light on the tumor microenvironment and composition of immune cell infiltrates. The results of these analyses (**Figures 7A,B**) indicated that $CD8^+$ T cells, plasma cells, naïve $CD8^+$ T cells, M1 macrophages, memory B cells, and endothelial cells were more abundant in the low-risk subgroup, whereas M2 macrophages, M0 macrophages, and fibroblasts were more common in the high-risk subgroup. A substantial body of research has revealed that a high level of infiltration of T cells, especially cytotoxic $CD8^+$ T cells, predicts a beneficial outcome (Bindea et al., 2013; Gentles et al., 2015; Fridman et al., 2017). Considering the fibroblasts in the high risk group were significantly higher than those in the low risk group ($p < 0.001$), we speculated that the high risk group were more abundant in cancer-associated fibroblast (CAF) compared to the low risk group. CAF is the main cell component of tumor microenvironment (TMB). Studies have shown that CAFs can assist the immune escape of breast cancer cells, promote proliferation, invasion and metastasis of breast cancer cells, inhibit immune response (Shimura et al., 2018; Gok Yavuz et al., 2019;Ivy X Chen et al., 2019) and inhibit T cell infiltration (Lei Li et al., 2021). This result is in accordance with our study.

In most malignancies, M2 macrophages, which are a significant subtype of macrophages, have been shown to correlate with increased tumor cell proliferation, development of an invasive phenotype, and chronic inflammation, and these cells have been correlated with a poor outcome in breast, gastric, ovarian, bladder, and prostate cancers (Josephs et al., 2015; Ruffell and Coussens, 2015; Fridman et al., 2017). Conversely, a high density of M1 macrophages seems to be correlated with acute inflammation and imply a favorable prognosis among patients with HCC, NSCLC, gastric, or ovarian cancers (Josephs et al., 2015; Ruffell and Coussens, 2015; Fridman et al., 2017). The results of our study support these conclusion.

Moreover, we found that the low-risk samples had a more robust ability for damage repair, whereas the high-risk samples had more immunosuppressive cells and signals and tumor and metastasis-related signals, which implies that the high-risk subgroup exhibited characteristics of immunosuppression and active tumor progression.

Furthermore, there were different proportions of the PAM50 subtypes in the two subsets (**Figure 9A**), as we found that the high-risk subgroup possessed more HER2+ samples, which are more invasive. In contrast, the low-risk subgroup possessed more luminal A samples, which are less invasive. In summary, we concluded that the low-risk subgroup was characterized by lower tumor aggressiveness and an active immune response, whereas the high-risk subgroup was characterized by higher tumor aggressiveness and an immune-suppressive response.

Finally, we learned that patients in the low-risk group were more responsive to treatment with gefitinib, epothilone B, and doxorubicin. In contrast, patients in the high-risk group were more responsive to treatment with docetaxel (**Figures 9B–G**). This phenomenon might provide valuable clinical treatment recommendations for high- and low-risk groups.

There are still some limitations to this study. First, the data was obtained only from a single TCGA dataset. The analysis of multiple datasets would have been more convincing. Second, the associations were analyzed solely by statistical analysis and were not validated experimentally. Lastly, when exploring the immune microenvironment, we did not illustrate the signalling pathways of the target genes at a deeper level. We should investigate the specific mechanisms of the AAM-related prognostic lncRNAs and immune cells in the future. There is still a long way to go to considerably optimise personalised immunotherapy management.

CONCLUSION

In conclusion, we attempted to shed light on the importance of AAM-associated lncRNAs in BRCA. The prognostic model built here might be acknowledged as an indispensable reference for predicting the outcome of patients with BRCA and help identify their immune and molecular characteristics. However, further studies are needed to illustrate this point.

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

Y-wD and Z-kW designed the studies and finished most of the work. Z-xW, H-dW, and C-zY drafted the article. L-xL, Z-qW, and C-hL contributed to data collection and analyses. CZ revised and reviewed the articles. All authors read and approved the final article.

ACKNOWLEDGMENTS

Written informed consent was obtained from each participant. We are greatly thankful to CZ and xuanxuan Dai at the First Affiliated Hospital of Wenzhou Medical University for friendly guidance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.880387/full#supplementary-material

Supplementary Figure S1 | Validation of the risk score in the validation set and the complete TCGA data set. **(A)** K–M analysis in the validation set. **(B)** K–M analysis in the complete TCGA data set. **(C)** Time-dependent ROC analysis in the validation set. **(D)** Time-dependent ROC analysis in the complete TCGA data set.

Supplementary Figure S2 | AAM-related lncRNA signature, 11-lncRNA signature (Shen et al., 2020), 11 (2)-lncRNA signature(Li et al., 2021b) , and eight-lncRNA signature(Zhu et al., 2021), respectively, based on TCGA. **(A–D)** Respective Kaplan–Meier curve results, **(E–H)** respective AUC for the prediction of the 1-, 3-, and 5-year survival rate in BRCA.

Supplementary Figure S3 | Heatmap for immune responses based on CIBERSORT, ESTIMATE, MCPcounter, ssGSEA, and TIMER algorithms among high and low risk group.

REFERENCES

Adams, S., Gray, R. J., Demaria, S., Goldstein, L., Perez, E. A., Shulman, L. N., et al. (2014). Prognostic Value of Tumor-Infiltrating Lymphocytes in Triple-Negative Breast Cancers from Two Phase III Randomized Adjuvant Breast Cancer Trials: ECOG 2197 and ECOG 1199. *Jco* 32, 2959–2966. doi:10.1200/jco.2013.55.0491

Ali, H. R., Chlon, L., Pharoah, P. D. P., Markowetz, F., and Caldas, C. (2016). Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *PLoS Med.* 13, e1002194. doi:10.1371/journal.pmed.1002194

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

Butler, M., van der Meer, L. T., and van Leeuwen, F. N. (2021). Amino Acid Depletion Therapies: Starving Cancer Cells to Death. *Trends Endocrinol. Metab.* 32, 367–381. doi:10.1016/j.tem.2021.03.003

Cancer Genome Atlas Network (2012). Comprehensive Molecular Portraits of Human Breast Tumours. *Nature* 490, 61–70. doi:10.1038/nature11412

Cha, Y. J., Kim, E. S., and Koo, J. S. (2018). Amino Acid Transporters and Glutamine Metabolism in Breast Cancer. *Int. J. Mol. Sci.* 19, 907. doi:10.3390/ijms19030907

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, I. X., Chauhan, V. P., Posada, J., Ng, M. R., Wu, M. W., Adstamongkonkul, P., et al. (2019). Blocking CXCR4 Alleviates Desmoplasia, Increases T-Lymphocyte Infiltration, and Improves Immunotherapy in Metastatic Breast Cancer. *Proc. Natl. Acad. Sci. U.S.A.* 116, 4558–4566. doi:10.1073/pnas.1815515116

Chen, P.-H., Cai, L., Huffman, K., Yang, C., Kim, J., Faubert, B., et al. (2019). Metabolic Diversity in Human Non-Small Cell Lung Cancer Cells. *Mol. Cell* 76, 838–851.e5. doi:10.1016/j.molcel.2019.08.028

Chrisanthar, R., Knappskog, S., Løkkevik, E., Anker, G., Østenstad, B., Lundgren, S., et al. (2011). Predictive and Prognostic Impact of TP53 Mutations and MDM2 Promoter Genotype in Primary Breast Cancer Patients Treated with Epirubicin or Paclitaxel. *PloS one* 6, e19249. doi:10.1371/journal.pone.0019249

Ciriello, G., Gatza, M. L., Beck, A. H., Wilkerson, M. D., Rhie, S. K., Pastore, A., et al. (2015). Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* 163, 506–519. doi:10.1016/j.cell.2015.09.033

Denkert, C., von Minckwitz, G., Darb-Esfahani, S., Lederer, B., Heppner, B. I., Weber, K. E., et al. (2018). Tumour-infiltrating Lymphocytes and Prognosis in Different Subtypes of Breast Cancer: a Pooled Analysis of 3771 Patients Treated with Neoadjuvant Therapy. *Lancet. Oncol.* 19, 40–50. doi:10.1016/s1470-2045(17)30904-x

Dias, A. S., Almeida, C. R., Helguero, L. A., and Duarte, I. F. (2019). Metabolic Crosstalk in the Breast Cancer Microenvironment. *Eur. J. Cancer* 121, 154–171. doi:10.1016/j.ejca.2019.09.002

Fang, J., Chen, F., Liu, D., Gu, F., Chen, Z., and Wang, Y. (2020). Prognostic Value of Immune Checkpoint Molecules in Breast Cancer. *Biosci. Rep.* 40, BSR20201054. doi:10.1042/BSR20201054

Faubert, B., Solmonson, A., and DeBerardinis, R. J. (2020). Metabolic Reprogramming and Cancer Progression. *Science* 368, eaaw5473. doi:10.1126/science.aaw5473

Fridman, W. H., Zitvogel, L., Sautès–Fridman, C., and Kroemer, G. (2017). The Immune Contexture in Cancer Prognosis and Treatment. *Nat. Rev. Clin. Oncol.* 14, 717–734. doi:10.1038/nrclinonc.2017.101

Gao, P., Tchernyshyov, I., Chang, T.-C., Lee, Y.-S., Kita, K., Ochi, T., et al. (2009). c-Myc Suppression of miR-23a/b Enhances Mitochondrial Glutaminase Expression and Glutamine Metabolism. *Nature* 458, 762–765. doi:10.1038/nature07823

Gentles, A. J., Newman, A. M., Liu, C. L., Bratman, S. V., Feng, W., Kim, D., et al. (2015). The Prognostic Landscape of Genes and Infiltrating Immune Cells across Human Cancers. *Nat. Med.* 21, 938–945. doi:10.1038/nm.3909

Gok Yavuz, B., Gunaydin, G., Gedik, M. E., Kosemehmetoglu, K., Karakoc, D., Ozgur, F., et al. (2019). Cancer Associated Fibroblasts Sculpt Tumour Microenvironment by Recruiting Monocytes and Inducing Immunosuppressive PD-1(+) TAMs. *Sci. Rep.* 9, 3172. doi:10.1038/s41598-019-39553-z

Guo, Y., Lv, B., Liu, R., Dai, Z., Zhang, F., Liang, Y., et al. (2021). Role of LncRNAs in Regulating Cancer Amino Acid Metabolism. *Cancer Cell Int.* 21, 209. doi:10.1186/s12935-021-01926-8

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., et al. (2010). Long Non-coding RNA HOTAIR Reprograms Chromatin State to Promote Cancer Metastasis. *Nature* 464, 1071–1076. doi:10.1038/nature08975

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, J., Li, F., Zhou, Y., Hou, X., Liu, S., Li, X., et al. (2020). LncRNA XLOC_006390 Promotes Pancreatic Carcinogenesis and Glutamate Metabolism by Stabilizing C-Myc. *Cancer Lett.* 469, 419–428. doi:10.1016/j.canlet.2019.11.021

Jiang, Q., Wang, Z., Qi, Q., Li, J., Xin, Y., and Qiu, J. (2022). lncRNA SNHG26 Promoted the Growth, Metastasis, and Cisplatin Resistance of Tongue Squamous Cell Carcinoma through PGK1/Akt/mTOR Signal Pathway. *Mol. Ther. - Oncol.* 24, 355–370. doi:10.1016/j.omto.2021.12.021

Jones, C. L., Stevens, B. M., D'Alessandro, A., Reisz, J. A., Culp-Hill, R., Nemkov, T., et al. (2018). Inhibition of Amino Acid Metabolism Selectively Targets Human Leukemia Stem Cells. *Cancer Cell* 34, 724–740. e724. doi:10.1016/j.ccell.2018.10.005

Josephs, D. H., Bax, H. J., and Karagiannis, S. N. (2015). Tumour-associated Macrophage Polarisation and Re-education with Immunotherapy. *Front. Biosci. (Elite Ed.)* 7, 293–308. doi:10.2741/e735

Kelly, B., and Pearce, E. L. (2020). Amino Assets: How Amino Acids Support Immunity. *Cell metab.* 32, 154–175. doi:10.1016/j.cmet.2020.06.010

Kwapisz, D. (2021). Pembrolizumab and Atezolizumab in Triple-Negative Breast Cancer. *Cancer Immunol. Immunother.* 70, 607–617. doi:10.1007/s00262-020-02736-z

Lam, W. Y., Becker, A. M., Kennerly, K. M., Wong, R., Curtis, J. D., Llufrio, E. M., et al. (2016). Mitochondrial Pyruvate Import Promotes Long-Term Survival of Antibody-Secreting Plasma Cells. *Immunity* 45, 60–73. doi:10.1016/j.immuni.2016.06.011

Lee, K. H., Kim, E. Y., Yun, J. S., Park, Y. L., Do, S.-I., Chae, S. W., et al. (2018). The Prognostic and Predictive Value of Tumor-Infiltrating Lymphocytes and Hematologic Parameters in Patients with Breast Cancer. *BMC cancer* 18, 938. doi:10.1186/s12885-018-4832-5

Lehmann, B. D., Bauer, J. A., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., et al. (2011). Identification of Human Triple-Negative Breast Cancer Subtypes and Preclinical Models for Selection of Targeted Therapies. *J. Clin. Invest.* 121, 2750–2767. doi:10.1172/jci45014

Li, L., Wei, J. R., Dong, J., Lin, Q. G., Tang, H., Jia, Y. X., et al. (2021). Laminin γ2-mediating T Cell Exclusion Attenuates Response to Anti-PD-1 Therapy. *Sci. Adv.* 7, eabc8346. doi:10.1126/sciadv.abc8346

Leone, R. D., Zhao, L., Englert, J. M., Sun, I.-M., Oh, M.-H., Sun, I.-H., et al. (2019). Glutamine Blockade Induces Divergent Metabolic Programs to Overcome Tumor Immune Evasion. *Science* 366, 1013–1021. doi:10.1126/science.aav2588

Liao, M., Liao, W., Xu, N., Li, B., Liu, F., Zhang, S., et al. (2019). LncRNA EPB41L4A-AS1 Regulates Glycolysis and Glutaminolysis by Mediating Nucleolar Translocation of HDAC2. *EBioMedicine* 41, 200–213. doi:10.1016/j.ebiom.2019.01.035

Liu, Y.-Q., Chai, R.-C., Wang, Y.-Z., Wang, Z., Liu, X., Wu, F., et al. (2019). Amino Acid Metabolism-Related Gene Expression-Based Risk Signature Can Better Predict Overall Survival for Glioma. *Cancer Sci.* 110, 321–333. doi:10.1111/cas.13878

Li, X., Jin, F., and Li, Y. (2021). A Novel Autophagy-related lncRNA Prognostic Risk Model for Breast Cancer. *J. Cell. Mol. Med.* 25, 4–14. doi:10.1111/jcmm.15980

Loi, S., Michiels, S., Salgado, R., Sirtaine, N., Jose, V., Fumagalli, D., et al. (2014). Tumor Infiltrating Lymphocytes Are Prognostic in Triple Negative Breast Cancer and Predictive for Trastuzumab Benefit in Early Breast Cancer: Results from the FinHER Trial. *Ann. Oncol.* 25, 1544–1550. doi:10.1093/annonc/mdu112

Lyons, T. G. (2019). Targeted Therapies for Triple-Negative Breast Cancer. *Curr. Treat. Options Oncol.* 20, 82. doi:10.1007/s11864-019-0682-x

McGrail, D. J., Pilié, P. G., Rashid, N. U., Voorwerk, L., Slagter, M., Kok, M., et al. (2021). High Tumor Mutation Burden Fails to Predict Immune Checkpoint Blockade Response across All Cancer Types. *Ann. Oncol.* 32, 661–672. doi:10.1016/j.annonc.2021.02.006

Morotti, M., Zois, C. E., El-Ansari, R., Craze, M. L., Rakha, E. A., Fan, S.-J., et al. (2021). Increased Expression of Glutamine Transporter SNAT2/SLC38A2 Promotes Glutamine Dependence and Oxidative Stress Resistance, and Is Associated with Worse Prognosis in Triple-Negative Breast Cancer. *Br. J. Cancer* 124, 494–505. doi:10.1038/s41416-020-01113-y

Nagayama, A., Vidula, N., Ellisen, L., and Bardia, A. (2020). Novel Antibody-Drug Conjugates for Triple Negative Breast Cancer. *Ther. Adv. Med. Oncol.* 12, 1758835920915980. doi:10.1177/1758835920915980

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

Newsholme, P., Curi, R., Pithon Curi, T. C., Murphy, C. J., Garcia, C., and Pires de Melo, M. (1999). Glutamine Metabolism by Lymphocytes, Macrophages, and Neutrophils: its Importance in Health and Disease. *J. Nutr. Biochem.* 10, 316–324. doi:10.1016/s0955-2863(99)00022-4

Olivier, M., Langerød, A., Carrieri, P., Bergh, J., Klaar, S., Eyfjord, J., et al. (2006). The Clinical Value of Somatic TP53 Gene Mutations in 1,794 Patients with Breast Cancer. *Clin. Cancer Res.* 12, 1157–1167. doi:10.1158/1078-0432.ccr-05-1029

Parker, J. S., Mullins, M., Cheang, M. C. U., Leung, S., Voduc, D., Vickery, T., et al. (2009). Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *Jco* 27, 1160–1167. doi:10.1200/jco.2008.18.1370

Pruneri, G., Vingiani, A., and Denkert, C. (2018). Tumor Infiltrating Lymphocytes in Early Breast Cancer. *Breast* 37, 207–214. doi:10.1016/j.breast.2017.03.010

Ren, Z., Wang, Z., Gu, D., Ma, H., Zhu, Y., Cai, M., et al. (2021). Genome Instability and Long Noncoding RNA Reveal Biomarkers for Immunotherapy and Prognosis and Novel Competing Endogenous RNA Mechanism in Colon Adenocarcinoma. *Front. Cell Dev. Biol.* 9, 740455. doi:10.3389/fcell.2021.740455

Rizzo, A., and Ricci, A. D. (2021). Biomarkers for Breast Cancer Immunotherapy: PD-L1, TILs, and beyond. *Expert Opin. Investig. Drugs*, 1–7. doi:10.1080/13543784.2022.2008354

Rizzo, A., Ricci, A. D., Lanotte, L., Lombardi, L., Di Federico, A., Brandi, G., et al. (2021). Immune-based Combinations for Metastatic Triple Negative Breast Cancer in Clinical Trials: Current Knowledge and Therapeutic Prospects. *Expert Opin. Investig. Drugs*, 1–9. doi:10.1080/13543784.2022.2009456

Rizzo, A., Cusmai, A., Acquafredda, S., Rinaldi, L., and Palmiotti, G. (2022). Ladiratuzumab Vedotin for Metastatic Triple Negative Cancer: Preliminary Results, Key Challenges, and Clinical Potential. *Expert Opin. Investig. Drugs*, 1–4. doi:10.1080/13543784.2022.2042252

Rodriguez, A. E., Ducker, G. S., Billingham, L. K., Martinez, C. A., Mainolfi, N., Suri, V., et al. (2019). Serine Metabolism Supports Macrophage IL-1β Production. *Cell Metab.* 29, 1003–1011.e4. doi:10.1016/j.cmet.2019.01.014

Ruffell, B., and Coussens, L. M. (2015). Macrophages and Therapeutic Resistance in Cancer. *Cancer Cell* 27, 462–472. doi:10.1016/j.ccell.2015.02.015

Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *Anesth. Anal.* 126, 1763–1768. doi:10.1213/ane.0000000000002864

Shahbandi, A., Nguyen, H. D., and Jackson, J. G. (2020). TP53 Mutations and Outcomes in Breast Cancer: Reading beyond the Headlines. *Trends cancer* 6, 98–110. doi:10.1016/j.trecan.2020.01.007

Shen, Y., Peng, X., and Shen, C. (2020). Identification and Validation of Immune-Related lncRNA Prognostic Signature for Breast Cancer. *Genomics* 112, 2640–2646. doi:10.1016/j.ygeno.2020.02.015

Shimura, T., Sasatani, M., Kawai, H., Kamiya, K., Kobayashi, J., Komatsu, K., et al. (2018). Radiation-Induced Myofibroblasts Promote Tumor Growth via Mitochondrial ROS-Activated TGFβ Signaling. *Mol. Cancer Res.* 16, 1676–1686. doi:10.1158/1541-7786.mcr-18-0321

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer Statistics, 2019. *CA A Cancer J. Clin.* 69, 7–34. doi:10.3322/caac.21551

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer Statistics, 2020. *CA A Cancer J. Clin.* 70, 7–30. doi:10.3322/caac.21590

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

Stanton, S. E., and Disis, M. L. (2016). Clinical Significance of Tumor-Infiltrating Lymphocytes in Breast Cancer. *J. Immunother. cancer* 4, 59. doi:10.1186/s40425-016-0165-6

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, 15545–15550. doi:10.1073/pnas.0506580102

Tan, Y. T., Lin, J. F., Li, T., Li, J. J., Xu, R. H., and Ju, H. Q. (2021). LncRNA-mediated Posttranslational Modifications and Reprogramming of Energy Metabolism in Cancer. *Cancer Commun.* 41, 109–120. doi:10.1002/cac2.12108

Vousden, K., and Prives, C. (2005). P53 and Prognosis: New Insights and Further Complexity. *Cell* 120, 7–10. doi:10.1016/s0092-8674(04)01252-8

Wallace, C., and Keast, D. (1992). Glutamine and Macrophage Function. *Metabolism* 41, 1016–1020. doi:10.1016/0026-0495(92)90130-3

Wang, Y., Xia, Y., and Lu, Z. (2018). Metabolic Features of Cancer Cells. *Cancer Commun.* 38, 65. doi:10.1186/s40880-018-0335-7

Wang, Z., Wang, Y., Peng, M., and Yi, L. (2019). UBASH3B Is a Novel Prognostic Biomarker and Correlated with Immune Infiltrates in Prostate Cancer. *Front. Oncol.* 9, 1517. doi:10.3389/fonc.2019.01517

Wang, N., Li, J., Xin, Q., and Xu, N. (2021). USP30-AS1 Contributes to Mitochondrial Quality Control in Glioblastoma Cells. *Biochem. biophysical Res. Commun.* 581, 31–37. doi:10.1016/j.bbrc.2021.10.006

Wang, L., Kong, L., Xu, S., Wang, X., Huang, K., Wang, S., et al. (2022). Isoliquiritigenin-mediated miR-23a-3p Inhibition Activates PGC-1α to Alleviate Alcoholic Liver Injury. *Phytomedicine* 96, 153845. doi:10.1016/j.phymed.2021.153845

Xing, Y., Lin, N. U., Maurer, M. A., Chen, H., Mahvash, A., Sahin, A., et al. (2019). Phase II Trial of AKT Inhibitor MK-2206 in Patients with Advanced Breast Cancer Who Have Tumors with PIK3CA or AKT Mutations, And/or PTEN Loss/PTEN Mutation. *Breast Cancer Res.* 21, 78. doi:10.1186/s13058-019-1154-8

Xu, Z., Jiang, S., Ma, J., Tang, D., Yan, C., and Fang, K. (2021). Comprehensive Analysis of Ferroptosis-Related LncRNAs in Breast Cancer Patients Reveals Prognostic Value and Relationship with Tumor Immune Microenvironment. *Front. Surg.* 8, 742360. doi:10.3389/fsurg.2021.742360

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, 10258. doi:10.1038/s41598-020-66986-8

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

Zhang, D., Xu, X., and Ye, Q. (2021). Metabolism and Immunity in Breast Cancer. *Front. Med.* 15, 178–207. doi:10.1007/s11684-020-0793-6

Zhang, T., Li, K., Zhang, Z. L., Gao, K., and Lv, C. L. (2021). LncRNA Airsci Increases the Inflammatory Response after Spinal Cord Injury in Rats through the Nuclear Factor Kappa B Signaling Pathway. *Neural Regen. Res.* 16, 772–777. doi:10.4103/1673-5374.295335

Zhao, Y., Zhang, J., Wang, S., Jiang, Q., and Xu, K. (2021). Identification and Validation of a Nine-Gene Amino Acid Metabolism-Related Risk Signature in HCC. *Front. Cell Dev. Biol.* 9, 731790. doi:10.3389/fcell.2021.731790

Zhu, L., Cui, K., Weng, L., Yu, P., Du, Y., Zhang, T., et al. (2021). A Panel of 8-lncRNA Predicts Prognosis of Breast Cancer Patients and Migration of Breast Cancer Cells. *PloS one* 16, e0249174. doi:10.1371/journal.pone.0249174

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 Dai, Wen, Wu, Wu, Lv, Yan, Liu, Wang and Zheng. 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: 18 May 2022
doi: 10.3389/fgene.2022.864473

Comprehensive Analysis of the Immune Cell Infiltration Landscape and Immune-Related Methylation in Retinoblastoma

Peiyao Mao[1], *Yinchen Shen*[1], *Xun Xu*[1] *and Jiawei Zhong*[2]*

[1]*Shanghai Key Laboratory of Ocular Fundus Diseases, Department of Ophthalmology, Shanghai General Hospital, School of Medicine, Shanghai Jiao Tong University, National Clinical Research Center for Eye Diseases, Shanghai Engineering Center for Visual Science and Photomedicine, Shanghai Engineering Center for Precise Diagnosis and Treatment of Eye Diseases, Shanghai, China,* [2]*State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-Sen University, Guangzhou, China*

Retinoblastoma is a common pediatric intraocular cancer, originating from cone precursors. The development of immunotherapies can help eradicate the tumor without vision loss, which would largely improve the quality of life of patients with retinoblastoma. Investigation of the tumor immune microenvironment provides knowledge for developing novel immunotherapies in cancer. However, the immune cell infiltrative landscape of retinoblastoma is unknown. Here, we compared the relative expression of immune gene signatures among 59 patients with retinoblastoma. The patients were divided into two subgroups according to the 28 types of immune cell infiltration (ICI) scores. We found that a subgroup with high ICI scores had increased expression levels of late cone markers, while the other subgroup exhibited larger tumor size and metastasis propensity. Furthermore, hypermethylated genes in the high-ICI subgroup were associated with immune regulation in the tumor microenvironment, suggesting that DNA methylation may play a vital regulatory role in retinoblastoma immunity. Our study provides a comprehensive framework for the systemic analysis of the influences of epigenetic events on the tumor immune microenvironment. We anticipate that our assay can not only provide insights into tumor immune regulation but also open up the perspectives for the identification of novel immunotherapy targets for retinoblastoma.

Keywords: retinoblastoma, immune cell infiltration, DNA methylation, bioinformatics analysis, immunotherapy

OPEN ACCESS

Edited by:
William C. Cho,
QEH, Hong Kong SAR, China

Reviewed by:
Parvin Mehdipour,
Tehran University of Medical Sciences, Iran
Jian Wang,
Children's Hospital of Soochow University, China

****Correspondence:***
Jiawei Zhong
zhongjiawei1@hotmail.com

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

Received: *28 January 2022*
Accepted: *21 March 2022*
Published: *18 May 2022*

Citation:
Mao P, Shen Y, Xu X and Zhong J (2022) Comprehensive Analysis of the Immune Cell Infiltration Landscape and Immune-Related Methylation in Retinoblastoma. Front. Genet. 13:864473. doi: 10.3389/fgene.2022.864473

INTRODUCTION

Retinoblastoma is one of the most common primary ocular malignancies in children with an incidence of 1:16000–1:18000 (Dimaras et al., 2015). It is usually initiated by biallelic retinoblastoma gene (*RB1*) mutation. Despite the significant improvement in treatments including cryotherapy, radiotherapy, ophthalmic artery chemosurgery, and intravitreous chemotherapy, some retinoblastoma patients eventually develop metastases due to invasion of the central nervous system through the optic nerve and dissemination through the sclera to the orbit (Gündüz et al., 2006; Abramson et al., 2015; Lu et al., 2019). Over the past several decades, cancer immunotherapy, including immune checkpoint blockade, vaccination, and adoptive T-cell therapy, has brought significant improvement for patients in terms of survival and quality of life

(Esfahani et al., 2020). However, compared with other cancer types, few immunotherapies have been applied to patients with retinoblastoma (Schefler and Kim, 2021). Therefore, a systematic investigation of the tumor immune microenvironment is essential for the development of effective anti-tumor immunotherapies.

With the help of novel technologies such as single-cell RNA sequencing (scRNA-seq), the critical role of the tumor microenvironment (TME) in tumor genesis, invasion, metastasis, and relapse has been revealed (Schelker et al., 2017; Steele et al., 2020; Lim et al., 2021; Mani et al., 2022). The TME in retinoblastoma contains numerous immune cells, including dendritic cells, monocytes, macrophages, and T-lymphocyte cells (Sarver et al., 2021). For example, a previous study has shown that reduced retinoblastoma cell proliferation was associated with increased immune cell infiltration (Sarver et al., 2021). Moreover, bioinformatics algorithms are developed for an immune infiltration estimation of a series of cancer tissues based on their transcriptional data (Hänzelmann et al., 2013; Yoshihara et al., 2013; Newman et al., 2015). These methods have significantly promoted our understanding of the TME and have been applied to hepatocellular carcinoma (Liu S. et al., 2021), clear cell renal cell carcinoma (Zhang et al., 2021), pancreatic adenocarcinoma (Liu et al., 2020), and so on. However, the application of retinoblastoma has not been fully elucidated.

DNA methylation has proved its role as the crucial epigenetic regulator in cancer progression by regulating genome sequence stability and gene expression (Feinberg et al., 2016). It is commonly known that the inactivation of certain tumor-suppressor genes occurs as a consequence of hypermethylation within the promoter regions (Kulis and Esteller, 2010). Moreover, abnormal methylation events were observed in retinoblastoma (Stirzaker et al., 1997; Berdasco et al., 2017). However, the relationship between DNA methylation and the immune microenvironment of retinoblastoma has not been broadly interrogated.

In this study, based on the immune profile of 28 types of immune cells, we identified two immunological subgroups of retinoblastoma. These two subgroups of retinoblastoma patients have distinct clinical characteristics and gene expression profiles. Next, we systematically examined the distinct DNA methylation patterns between these two subgroups. Moreover, we screened 6 differentially methylated and expressed genes as hub genes, which may provide new insights into the molecular pathogenesis and the clinical immunotherapy of retinoblastoma.

MATERIALS AND METHODS

Data Collection From GEO Databases

Gene expression arrays from 59 samples diagnosed with retinoblastoma were obtained from GEO databases with accession code GSE58780. The DNA methylation array from retinoblastoma patients was obtained from GEO databases with accession code GSE58783. Clinical data of all samples were downloaded from https://static-content.springer.com/esm/art%3A10.1038%2Fs41467-021-25792-0/MediaObjects/41467_2021_25792_MOESM4_ESM.xlsx. scRNA-seq data from retinoblastoma patients were obtained from GEO databases with accession code GSE174200.

Immune Cell Infiltration Analysis

We performed a single-sample Gene Set Enrichment Analysis (ssGSEA) by using the *GSVA* (version 1.34.0) (Hänzelmann et al., 2013) R package based on the default parameters to calculate the immune infiltration level of 28 immune cell types (Charoentong et al., 2017). Among these immune cells, the activated $CD4^+$ T cell, activated $CD8^+$ T cell, central memory $CD4^+$ T cell, central memory $CD8^+$ T cell, effector memory $CD4^+$ T cell, effector memory $CD8^+$ T cell, type 1 T helper cell, type 17 T helper cell, activated dendritic cell, $CD56^{bright}$ natural killer cell, natural killer cell, and natural killer T cell are considered to have anti-tumor capacities. Regulatory T cell, type 2 T helper cell, $CD56^{dim}$ natural killer cell, immature dendritic cell, macrophage, MDSC, neutrophil, and plasmacytoid dendritic cell are considered to have pro-tumor capacities. We also used the Estimation of STromal and Immune cells in MAlignant Tumors using Expression data (ESTIMATE) algorithm of the *estimate* (version 1.0.13) (Yoshihara et al., 2013) R package to calculate the stromal and immune scores and tumor purity of each sample.

Clustering Analysis Based on Immune Cell Infiltration Analysis

By using the *stats* (version 3.6.0) R package, we performed an unsupervised hierarchical clustering (based on Euclidean distance and Ward's linkage) to cluster retinoblastoma samples based on a sample-signal matrix including 28 types of immune cells of 59 retinoblastoma samples. Fifty-nine samples were divided into high and low infiltration subgroups. The visualization of K-means clustering result was performed by the *pheatmap* (version 1.0.12) R package, and comparison between the two subgroups in terms of their signal enrichment score of 28 immune cell types was computed using a two-sided t-test and visualized by the *ggpubr* (version 0.3.0) R package.

Principal Component Analysis

We performed a principal component analysis (PCA) on the sample-signal matrix using *FactoMineR* (version2.4) R package with default parameters. The result from the PCA was visualized by the *factoextra* (version 1.0.7) R package.

Differential Expression Analysis

The sample-gene gene expression matrix was input into the *limma* (version 3.42.2) (Ritchie et al., 2015) R package for the identification of differentially expressed genes between the high-ICI subgroup and low-ICI subgroup. We determined differentially expressed genes (DEGs) with the criteria of absolute fold change >1.5 and false discovery rate (FDR) adjusted $p < 0.05$. Clusters of DEGs were identified by an unsupervised hierarchical cluster analysis (based on Euclidean distance and Ward's linkage).

Enrichment Analysis

An enrichment analysis was performed using the *clusterProfiler* (version 3.14.3) (Yu et al., 2012) R package with the "*enricher*" and "*GSEA*" function, and the FDR adjusted to $p < 0.05$ was considered as statistically significant. All gene sets were obtained from the Molecular Signatures Database (MSigDB) using the *msigdbr* (version 7.2.1) R package.

DNA Methylation Array Processing and Differential Methylation Analysis

After obtaining the microarray data from GSE58783, we used the "*champ.filter*" function of the *ChAMP* (version 2.16.2) (Tian et al., 2017) R package to remove probes which are located in sex chromosomes and near SNP to eliminate the influence of sex and SNP, respectively. We used the "*champ.DMP*" function of the *ChAMP* R package with the criteria of absolute $\Delta\beta > 0.2$ and FDR adjusted $p < 0.05$ for the identification of differentially methylated probes (DMPs) between the high-ICI subgroup and low-ICI subgroup. We next excluded the genes which had both hypermethylated and hypomethylated probes. Genes with either a hypermethylated probe or hypomethylated probe were considered as hypermethylated genes or hypomethylated genes, respectively.

Protein–Protein Interaction Network

We used *STRING* (version 11.5) (Szklarczyk et al., 2019) with default parameters to construct the protein–protein interaction (PPI) network. The generated PPI networks were visualized by *Cytoscape* software (version 3.9.0) (Shannon et al., 2003). In *Cytoscape*, we used *cytoHubba* (Chin et al., 2014) to screen hub genes by the Maximal Clique Centrality (MCC) method.

Statistical Analysis

All analyses were performed by R software (version 3.6.0). An unpaired two-tailed *t*-test was used to compare two subgroups of continuously distributed variables. The correlations of the retinoblastoma subgroups and clinical characteristics were analyzed using the chi-square test. $p \geq 0.05$ (n.s.), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), and $p < 0.0001$ (****).

Code Availability

All custom computer codes used in this study are freely available at https://github.com/jiawei-zhong/Mao_et_al_RB/

RESULTS

Identification of Immune-Related Gene Subtypes in Retinoblastoma Based on Immune Cell Infiltration

To study immune cell infiltration (ICI) of retinoblastoma, we performed the single-sample Gene Set Enrichment Analysis (ssGSEA) of 59 retinoblastoma patients from GSE58780 (Hänzelmann et al., 2013; Liu J. et al., 2021). Using gene sets which are related to 28 types of immune cells (Charoentong et al., 2017), the immune infiltration levels of all immune cell types were calculated (**Figure 1A**). Among these 28 types of immune cells, 12 types of immune cells, such as activated CD4 T cell and activated CD8 T cell, are considered to execute anti-tumor immunity; while 8 types of immune cells, such as regulatory T cell and type 2 T helper cell are considered having immune-suppressive functions (Jia et al., 2018). We used an unsupervised hierarchical clustering algorithm to assign the retinoblastoma samples into two clusters (high-ICI subgroup and low-ICI subgroup) based on immune infiltration levels (**Figure 1A**). The principal component analysis (PCA) of the retinoblastoma samples by immune infiltration levels confirmed the rationality of the result of hierarchical clustering (**Figure 1B**). The normalized enrichment score (NES) of each immune cell was then compared between the two subgroups, and the NES of 24 immune cells was significantly higher in the high-ICI subgroup (**Supplementary Figure S1A**). Subsequently, leveraging the ESTIMATE algorithm, we found that the stromal, immune, and ESTIMATE scores were relatively higher in the high-ICI subgroup (**Figures 1C–E**), whereas the tumor purity in the high-ICI subgroup was lower than that in the low-ICI subgroup (**Figure 1F**) (Yoshihara et al., 2013).

To prove that a higher ICI score represents a higher ICI level, we applied the ssGSEA to retinoblastoma scRNA-seq data from GSE174200 (Norrie et al., 2021), and the retinoblastoma samples were divided into two subgroups (**Supplementary Figure S2A**). Consistently, the high-ICI subgroup exhibited a higher immune cell percentage, indicating the robustness of the classification (**Supplementary Figure S2B**).

We next explored the differences in the clinical and pathological features between the two subgroups. The tumor diameter of the high-ICI subgroup was significantly smaller (mean diameter = 14.56 versus 15.8 mm, **Supplementary Figure S1B**). Patients with a high ICI score were significantly more likely to be hereditary forms (*RB1* mutation), whereas most of the patients with a low ICI score were non-hereditary (**Supplementary Figure S1C**). In the high-ICI subgroup, 60% of the patients were bilateral, higher than that in the low-ICI subgroup (**Supplementary Figure S1D**). Other characteristics (e.g., growth pattern, necrosis, optic nerve invasion, and choroid and sclera invasion) exhibited no statistical significance between the two subtypes (**Supplementary Figures S1E–H**).

Altogether, by estimating the immune infiltration levels in patients with retinoblastoma, we identified two different retinoblastoma subtypes with distinct immune features. We also found that the two subtypes exhibited significant differences in tumor size, *RB1* mutation, and laterality.

The Two Subgroups Displayed Differences in the Expression of Photoreceptor Markers and Proliferation Genes

To investigate the genes associated with immune cell infiltration, we performed a differential expression analysis to detect differentially expressed genes (DEGs) between the two subtypes by using the *limma* R package (Ritchie et al., 2015). Five-hundred fifty five and 320 genes were upregulated in the high-ICI subgroup and low-ICI subgroup, respectively

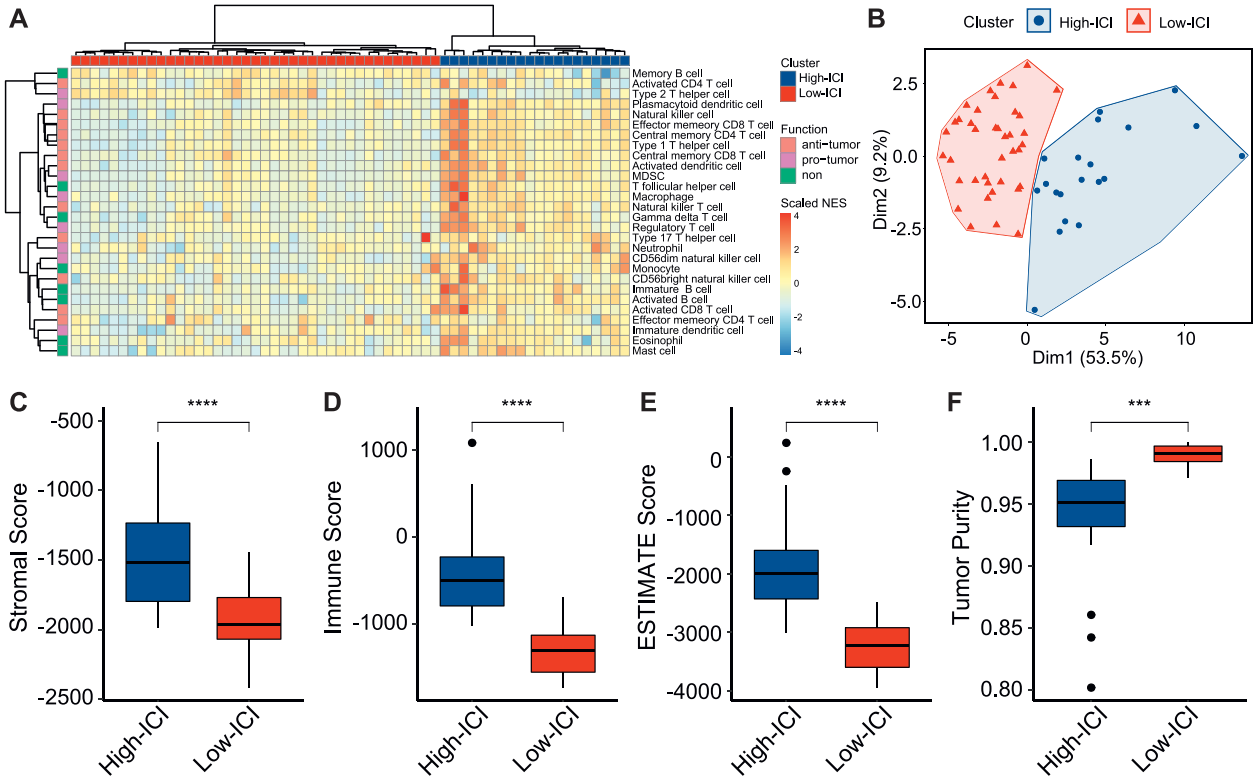


FIGURE 1 | Landscape of immune cell infiltration in retinoblastoma. **(A)** Heatmap showing the normalized enrichment scores of each retinoblastoma sample on 28 immune cell types. **(B)** Scatter plot showing the distinct subgroups of retinoblastoma divided by PCA. **(C–F)** Boxplot showing the difference of stromal score **(C)**, immune score **(D)**, ESTIMATE score **(E)**, and tumor purity **(F)** between two retinoblastoma subtypes.

(**Figure 2A**, **Supplementary Table S1**). Photoreceptor-specific genes, such as *OPN1SW* and *PDC*, were specifically expressed in the high-ICI subgroup, whereas proliferation markers (e.g., *MKI67*, *TOP2A*, *CENPE*, *CENPF*, and *TTL*) were highly expressed in the low-ICI subgroup (**Figure 2A**). By performing an unsupervised hierarchical clustering, these DEGs were classified into three main gene clusters, including two upregulated in the high-ICI subgroup and one upregulated in the low-ICI subgroup (**Figure 2B**). We next performed a Gene Ontology (GO) enrichment analysis of each gene cluster. The genes in cluster 1 were enriched for "visual perception," "sensory perception of light stimulus," and "phototransduction" (**Figure 2C**). The genes of cluster 2 were associated with immune/inflammation signature (e.g., "neutrophil activation," "response to interferon-gamma," and "MHC protein complex") (**Figure 2D**). The genes of cluster 3 were related to "chromosome segregation," "mitotic nuclear division," and "tubulin binding" (**Figure 2E**). The enrichment analysis of these three gene clusters using other databases (KEGG, HALLMARK, PID, and REACTOME) exhibited enrichments in similar pathways (**Supplementary Figures S3A–C**). Altogether, our data showed that the low-ICI subgroup presented a high proliferation potential.

Next, we used a gene set enrichment analysis (GSEA) to obtain deeper insights into the function of immune infiltration in retinoblastoma. For example, the high-ICI subgroup was enriched for "senescence and autophagy in cancer" and "retinal cone cell differentiation" (**Figures 2F,G**). The low-ICI subgroup was enriched for "retinoblastoma gene in cancer" (**Figure 2H**). Taken together, our results suggest that retinoblastoma in the high-ICI subgroup maintains a cone-differentiation state, and the overexpression of proliferation markers which we found in the low-ICI subgroup may result in a higher propensity for metastasis.

DNA Methylation Analysis Based on Different Immune Subtypes of Retinoblastoma

To interrogate the differences in the epigenome between the two immune subtypes of retinoblastoma, we analyzed DNA methylation with DNA methylation arrays from GSE58783. There were 16 and 37 cases corresponding to the high-ICI subgroup and low-ICI subgroup respectively. Leveraging the *ChAMP* R package, a total of 3,940 significantly differentially methylated probes (DMPs) were detected, including 3,217 and 723 hypermethylated probes in the high-ICI subgroup and low-ICI subgroup, respectively (**Figure 3A**, **Supplementary Table S2**) (Tian et al., 2017). Next, we explored the distributions of these DMPs. By considering the CpG island and the adjacent context,

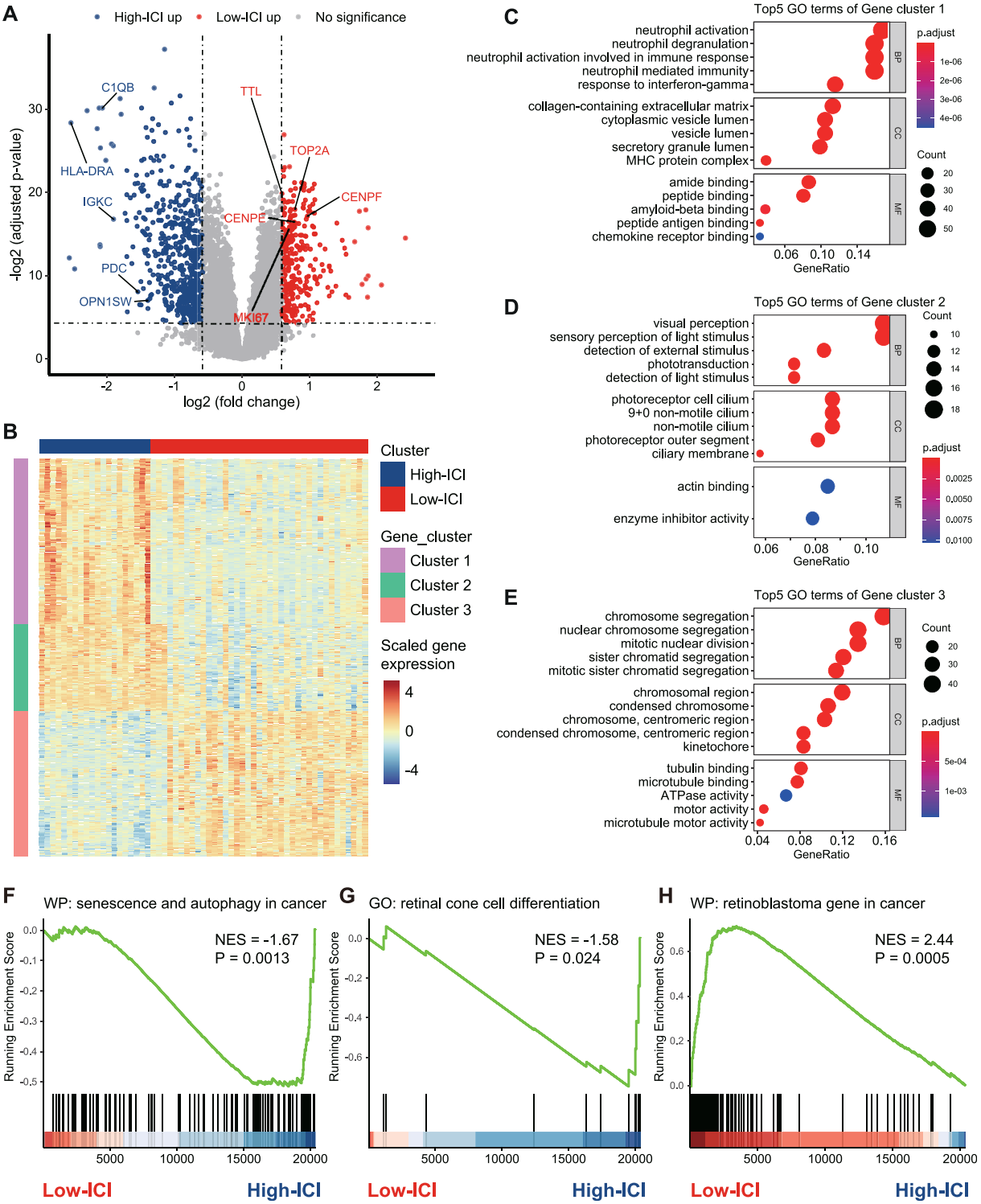


FIGURE 2 | Identification and functional annotation of DEGs between two subtypes. **(A)** Volcano plot depicting the DEGs in two subtypes. Specific genes were indicated. **(B)** Heatmap of the gene expression value of DEGs on each sample. **(C–E)** Dot plots of the top-ranked GO terms of genes in gene cluster 1 **(C)**, gene cluster 2 **(D)**, and gene cluster 3 **(E)**. **(F–H)** Enrichment plot showing upregulation of the "senescence and autophagy in cancer" **(F)**, "retinal cone cell differentiation" **(G)**, and downregulation of the "retinoblastoma gene in cancer" (**H**) in the high-ICI subgroup.

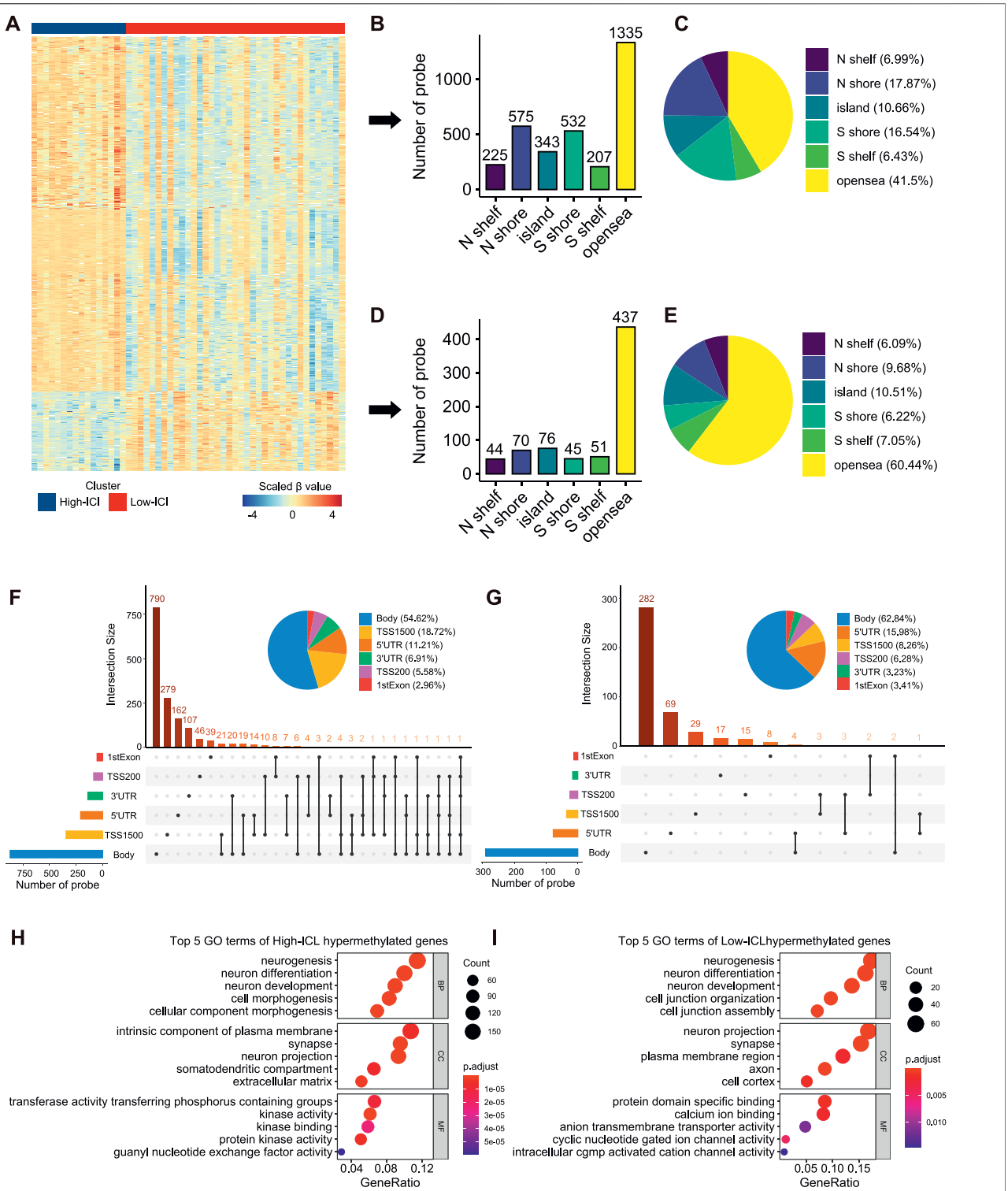


FIGURE 3 | Differential methylation analysis of DNA methylation array data based on two subtypes. **(A)** Heatmap of the β value of DMPs on each sample. **(B)** Bar plot showing the number of high-ICI hypermethylated probes distributed to different regions. **(C)** Pie plot representing the composition of high-ICI hypermethylated probes. **(D)** Bar plot showing numbers of low-ICI hypermethylated probes distributed to different regions. **(E)** Pie plot representing the composition of low-ICI hypermethylated probe. **(F–G)** Upset graph showing the distribution of high-ICI **(F)** and low-ICI **(G)** hypermethylated probes on different genomic regions. Dot plots of the top-ranked GO terms of high-ICI hypermethylated genes **(H)** and low-ICI hypermethylated genes **(I)** are shown.

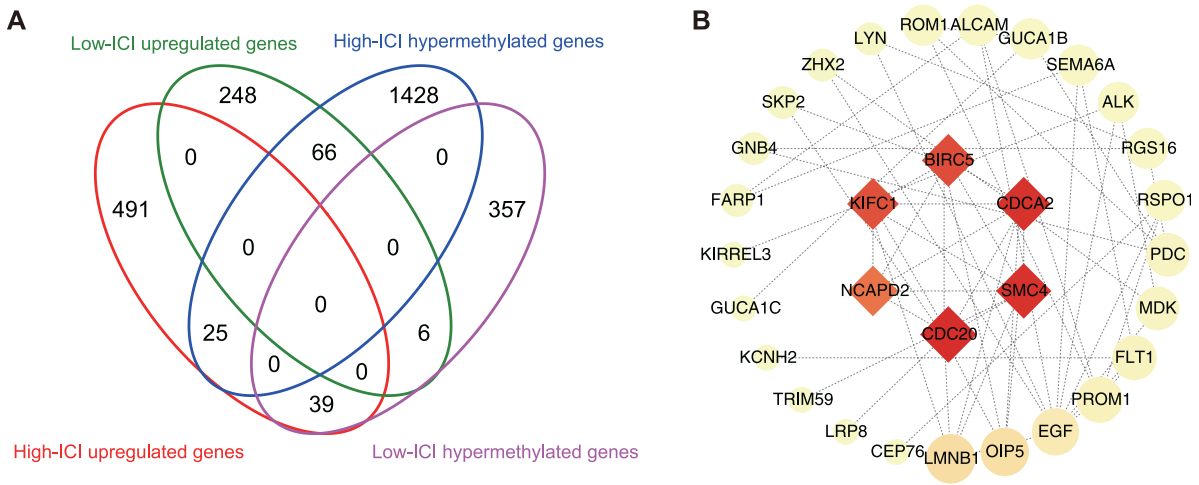


FIGURE 4 | Protein–protein interaction network on differentially methylated and expressed genes. **(A)** Flower plot showing the overlap of DEGs and DMGs. **(B)** Protein–protein interaction (PPI) network of differentially methylated DEGs. Hub genes were presented at the middle of network. The color and area of nodes reflect the score which is calculated by *cytoHubba*.

hypermethylated probes in the high-ICI subgroup were mostly located in the open sea (41.50%), followed by N-Shore (17.87%) and S-Shore (16.54%) (**Figures 3B,C**), while hypermethylated probes in the low-ICI subgroup were mostly located in the open sea (60.44%), followed by the CpG island (10.51%) and N-Shore (9.68%) (**Figures 3D,E**). After removing the probes in the intergenic region, the distributions of hypermethylated probes in the high-ICI subgroup in various functional genomic regions are shown in **Figure 3F**. The majority of the probes were discovered to be located in the gene body (54.62%), TSS1500 (200–1,500 bp upstream of the TSS, 18.72%) and 5′UTR (11.21%). The distributions of hypermethylated probes in the low-ICI subgroup were similar to those in the high-ICI subgroup, but some changes were observed. For example, the ratio of the probes across the gene body was higher (62.84%), but that across the 5′UTR area was lower (15.98%) (**Figure 3G**).

Next, we inquired about the potential biological processes related to DMPs. We mapped them to their corresponding genes and identified 1,519 hypermethylated genes in the high-ICI subgroup and 402 hypermethylated genes in the low-ICI subgroup (**Supplementary Figure S4A**). Surprisingly, by using an enrichment analysis, we observed that the hypermethylated genes in both subgroups were related to "neurogenesis," "neuron differentiation," and "neuron development" (**Figures 3H,J**). Moreover, some immune-related genes were included in the hypermethylated genes of the high-ICI subgroup, such as *CD83*, *HLA-DOA*, *IRF4*, *DOK3*, and *CXCR1*. Similarly, we also found that the hypermethylated genes of the high-ICI subgroup were significantly enriched for some biological processes related to immune regulation (e.g., T cell activation and B cell receptor signaling pathway), but no immune-related pathway was observed to be enriched by hypermethylated genes in the low-ICI subgroup (**Supplementary Figure S4B**). These results suggest that DNA hypermethylation is associated with the tumor immune microenvironment of retinoblastoma.

Identification of Differentially Methylated and Expressed Genes and PPI Networks

Given that DNA methylation is a major epigenetic factor influencing gene expression, we further investigated which DEGs underwent methylation or demethylation (Moore et al., 2013). Subsequently, the overlapping of DEGs and differentially methylated genes (DMGs) was performed, then 66 and 39 hypermethylated downregulated genes were found in the high-ICI subgroup and low-ICI subgroup, respectively (**Figure 4A**). To further understand the underlying functions of differentially methylated and expressed genes, these genes were inputted into the STRING to build protein–protein interaction (PPI) networks (Szklarczyk et al., 2019). The PPI network consisted of 48 nodes and 66 edges. Using *cytoHubba*, we removed the nodes with a low connectivity score (less than 2), and the biggest module was retained (**Figure 4B**) (Chin et al., 2014). In this module, six nodes (*BIRC5*, *CDCA2*, *SMC4*, *CDC20*, *NCAPD2*, and *KIFC1*) had more than 20 connectivity scores, which were further screened and identified as hub genes.

DISCUSSION

Over the past several years, immune cell infiltration in tumors has been found to be of crucial importance in precision medicine, which can be attributed to the deep understanding of the tumor microenvironment (Yang, 2015; Tamborero et al., 2018). Immunotherapy has been successfully applied to various types of cancers (Havel et al., 2019). A recent study has proved the efficacy of immunotherapy in eliminating retinoblastoma cells whilst preserving the mouse vision (Wang et al., 2020). However, the treatment of advanced retinoblastoma remains challenging, thus, further therapeutic development is needed (Abramson et al., 2015). Therefore, elucidation and understanding of the immune landscape of tumors may not only provide insights into tumor

immune dysregulation, but also lay the foundation for identifying novel immunotherapy targets.

In the present study, we characterized immune cell infiltration patterns in retinoblastoma using GEO databases. We used the ssGSEA method which is suitable for cross-platform evaluations of the landscape of 28 types of immune cells in retinoblastoma, and found two distinct subgroups. Patients with retinoblastoma in the high-ICI subgroup highly express late-stage cone markers (e.g. *GUCA1A* and *OPN1SW*) (Welby et al., 2017), whereas patients in the low-ICI subgroup highly express proliferation genes (e.g., *MKI67* and *TOP2A*) and retinoblastoma-related genes [e.g., *TTK* (Zeng et al., 2020) and *CDC25A* (Singh et al., 2015)], suggesting that immune cell infiltration is associated with retinoblastoma migration and metastatic progression. Consistent with the high expression of proliferation genes in patients with a low-ICI score, patients in the low-ICI subgroup exhibited a larger tumor diameter. Our study also showed that there was no significant difference in other clinical features (e.g. growth pattern, necrosis, optic nerve invasion, and choroid and sclera invasion) between the two subgroups. Further experiments are needed to confirm this result.

Many bioinformatics tools are being developed for the immune infiltration estimation of various cancer tissues based on their transcriptional data (Hänzelmann et al., 2013; Yoshihara et al., 2013; Newman et al., 2015). However, for validation, the application of a protein analysis on the single-cell level, including flow cytometry and immunostaining, on analyzed samples is required, which is essential to increase the reliability.

DNA methylation has proved its role as a significant epigenetic driving factor in cancer progression, development, and metastasis (Fleischer et al., 2014; Feinberg et al., 2016; Fleischer et al., 2017; Sina et al., 2019). Although evidence of DNA methylation regulating the immune microenvironment in breast cancer (Fleischer et al., 2017), glioma (Briand et al., 2019), and gingivo-buccal oral cancer (Das et al., 2019) has been presented, the role of DNA methylation in the retinoblastoma immune environment has not been completely explored. We systematically analyzed DNA methylation based on the two subgroups. Further analysis of DNA methylation differences between the two subgroups of retinoblastoma patients showed that a global DNA hypermethylation pattern was presented in the high-ICI subgroup (3,217 hypermethylated probes versus 723 hypomethylated probes). The enrichment analysis of these differentially methylated genes showed that these genes were remarkably related to the immune regulation in the tumor microenvironment including T cell activation, B cell receptor signaling pathway, and cytokine signaling in the immune system. These results implied that alterations in DNA methylation may play a crucial role in retinoblastoma immune cell infiltration.

The PPI network of differentially expressed and methylated genes provided a comprehensive observation of their functional connections, and screened hub genes. We identified a total of six hub genes: *BIRC5*, *CDCA2*, *SMC4*, *CDC20*, *NCAPD2,* and *KIFC1*. Among the hub genes, *BIRC5* (also named survivin) is a well-known cancer therapeutic target (Li et al., 2019). *BIRC5* immunotherapy-related clinical trials have been applied in patients with colorectal cancer (Tsuruma et al., 2004), malignant glioma (Fenstermaker et al., 2016), and melanoma (Becker et al., 2012). The role of *BIRC5* in retinoblastoma has also been investigated before. Exposure to carboplatin, topotecan, or radiation resulted in the elevated expression of *BIRC5* in the retinoblastoma cell line (Ferrario et al., 2016). A previous study showed that Dnmt1, Dnmt3a, and Dnmt3b can regulate the methylation status of *BIRC5* in glioblastoma multiforme (Hervouet et al., 2010). Our result extended and enriched the knowledge about the relationship between *BIRC5* and DNA methylation in tumors.

This study is not devoid of limitations. First, all our results were theoretical and validation based on patients or animal samples is lacking. Second, we had not found another dataset that includes DNA methylation data, so a validation cohort to confirm our conclusions is needed. Third, gene expression is a complex process involving numerous steps and many other regulatory elements, such as DNA methylation, nucleosome positioning and composition, 3D structural interactions, and histone modification can alter gene expression (Carter and Zhao, 2021). However, because of the lack of other multi-omics data, here we only elaborated on the relationship between gene expression and DNA methylation. Fourth, we only focused on hypermethylated downregulated genes and hypomethylated upregulated genes, which is a generally accepted regulative paradigm between DNA methylation and gene expression (Esteller, 2002; Ehrlich, 2009; Witte et al., 2014). However, recent efforts identified subtle changes in the relationship between DNA methylation and gene expression, beyond the classical dogma (Wan et al., 2015; Liu A. et al., 2021). Therefore, further analysis is required to evaluate contra-regulated genes. Fifth, we did not apply single-cell-level protein profiling, which is able to provide solid evidence to measure the infiltration level, on retinoblastoma to investigate the infiltration pattern.

In summary, by comprehensively assessing the immune cell infiltration in retinoblastoma, we highlight the differences between the two subgroups in gene expression and DNA methylation levels. The retinoblastoma immune landscape analysis may help clinicians develop novel immunotherapeutic targets.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found at: GSE58780 GSE58783 GSE174200.

AUTHOR CONTRIBUTIONS

PM and JZ conceived and designed the study, analyzed the data, and performed statistical analyses. All authors interpreted the data and drafted the manuscript.

FUNDING

This work was supported by the National Key R&D Program of China (No. 2016YFC0904800 and 2019YFC0840607), National Natural Science Foundation of China (No. 81800831),

Bethune•Lumitin Young and Middle-aged Ophthalmic Research Fund (No. BJ-LM2021010J), and Science and Technology Research Project of Songjiang District (No. 2020SJ307).

ACKNOWLEDGMENTS

The authors wish to thank Jing Liu and his/her colleagues for sharing the data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.864473/full#supplementary-material

Supplementary Figure S1 | Immune scores and clinical characteristics between two retinoblastoma subtypes. **(A)** Box plot showing the comparison of two subtypes in terms of their NES of different immune cell types. **(B-H)** Bar plot representing composition of **(B)** tumor diameter, **(C)** RB1 germ line, **(D)** laterality, **(E)** growth pattern, **(F)** necrosis, **(G)** optic nerve invasion, and **(H)** choroid and sclera invasion in the high-ICI and the low-ICI subgroup.

Supplementary Figure S2 | Immune cell infiltration landscape of scRNA-seq data from retinoblastoma patients. **(A)** Heatmap showing the normalized enrichment scores of scRNA-seq data from retinoblastoma patients on 28 immune cell types. **(B)** Box plot showing the difference in the immune cell percentage between two retinoblastoma subtypes.

Supplementary Figure S3 | Enrichment pathways of DEGs. **(A–C)** Dot plot showing the enriched pathways of gene clusters, **(A)** gene cluster 1, **(B)** gene cluster 2, and **(C)** gene cluster 3.

Supplementary Figure S4 | Functional annotation of DMGs. **(A)** Volcano plot depicting the DMGs in two subtypes. Specific genes were indicated. **(B)** Dot plot showing the enriched pathways of high-ICI hypermethylated genes, and immune-related pathways were selected.

REFERENCES

Abramson, D. H., Shields, C. L., Munier, F. L., and Chantada, G. L. (2015). Treatment of Retinoblastoma in 2015. *JAMA Ophthalmol.* 133, 1341–1347. doi:10.1001/jamaophthalmol.2015.3108

Becker, J. C., Andersen, M. H., Hofmeister-Müller, V., Wobser, M., Frey, L., Sandig, C., et al. (2012). Survivin-specific T-Cell Reactivity Correlates with Tumor Response and Patient Survival: a Phase-II Peptide Vaccination Trial in Metastatic Melanoma. *Cancer Immunol. Immunother.*, 61, 2091–2103. doi:10.1007/s00262-012-1266-9

Berdasco, M., Gómez, A., Rubio, M. J., Català-Mora, J., Zanón-Moreno, V., Lopez, M., et al. (2017). DNA Methylomes Reveal Biological Networks Involved in Human Eye Development, Functions and Associated Disorders. *Sci. Rep.* 7, 11762. doi:10.1038/s41598-017-12084-1

Briand, J., Nadaradjane, A., Bougras-Cartron, G., Olivier, C., Vallette, F. M., and Cartron, P.-F. (2019). Diuron Exposure and Akt Overexpression Promote Glioma Formation through DNA Hypomethylation. *Clin. Epigenet* 11, 159. doi:10.1186/s13148-019-0759-1

Carter, B., and Zhao, K. (2021). The Epigenetic Basis of Cellular Heterogeneity. *Nat. Rev. Genet.* 22, 235–250. doi:10.1038/s41576-020-00300-0

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. *Cel Rep.* 18, 248–262. doi:10.1016/j.celrep.2016.12.019

Chin, C. H., Chen, S. H., Wu, H. H., Ho, C. W., Ko, M. T., and Lin, C. Y. (2014). cytoHubba: Identifying Hub Objects and Sub-networks from Complex Interactome. *BMC Syst. Biol.* 8 Suppl 4 (Suppl. 4), S11. doi:10.1186/1752-0509-8-S4-S11

Das, D., Ghosh, S., Maitra, A., Biswas, N. K., Panda, C. K., Roy, B., et al. (2019). Epigenomic Dysregulation-Mediated Alterations of Key Biological Pathways and Tumor Immune Evasion Are Hallmarks of Gingivo-Buccal Oral Cancer. *Clin. Epigenet* 11, 178. doi:10.1186/s13148-019-0782-2

Dimaras, H., Corson, T. W., Cobrinik, D., White, A., Zhao, J., Munier, F. L., et al. (2015). Retinoblastoma. *Nat. Rev. Dis. Primers* 1, 15021. doi:10.1038/nrdp.2015.21

Ehrlich, M. (2009). DNA Hypomethylation in Cancer Cells. *Epigenomics* 1, 239–259. doi:10.2217/epi.09.33

Esfahani, K., Roudaia, L., Buhlaiga, N., Del Rincon, S. V., Papneja, N., and Miller, W. H., Jr. (2020). A Review of Cancer Immunotherapy: from the Past, to the Present, to the Future. *Curr. Oncol.* 27, S87–s97. doi:10.3747/co.27.5223

Esteller, M. (2002). CpG Island Hypermethylation and Tumor Suppressor Genes: a Booming Present, a Brighter Future. *Oncogene* 21, 5427–5440. doi:10.1038/sj.onc.1205600

Feinberg, A. P., Koldobskiy, M. A., and Göndör, A. (2016). Epigenetic Modulators, Modifiers and Mediators in Cancer Aetiology and Progression. *Nat. Rev. Genet.* 17, 284–299. doi:10.1038/nrg.2016.13

Fenstermaker, R. A., Ciesielski, M. J., Qiu, J., Yang, N., Frank, C. L., Lee, K. P., et al. (2016). Clinical Study of a Survivin Long Peptide Vaccine (SurVaxM) in Patients with Recurrent Malignant Glioma. *Cancer Immunol. Immunother.* 65, 1339–1352. doi:10.1007/s00262-016-1890-x

Ferrario, A., Luna, M., Rucker, N., Wong, S., Lederman, A., Kim, J., et al. (2016). Targeting Survivin Enhances Chemosensitivity in Retinoblastoma Cells and Orthotopic Tumors. *PloS one* 11, e0153011. doi:10.1371/journal.pone.0153011

Fleischer, T., Frigessi, A., Johnson, K. C., Edvardsen, H., Touleimat, N., Klajic, J., et al. (2014). Genome-wide DNA Methylation Profiles in Progression to in Situand Invasive Carcinoma of the Breast with Impact on Gene Transcription and Prognosis. *Genome Biol.* 15, 435. doi:10.1186/s13059-014-0435-x

Fleischer, T., Tekpli, X., Tekpli, X., Mathelier, A., Wang, S., Nebdal, D., et al. (2017). DNA Methylation at Enhancers Identifies Distinct Breast Cancer Lineages. *Nat. Commun.* 8, 1379. doi:10.1038/s41467-017-00510-x

Gündüz, K., Müftüoglu, O., Günalp, I., Unal, E., and Taçyildiz, N. (2006). Metastatic RetinoblastomaClinical Features, Treatment, and Prognosis. *Ophthalmology* 113, 1558–1566. doi:10.1016/j.ophtha.2006.03.039

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics* 14, 7. doi:10.1186/1471-2105-14-7

Havel, J. J., Chowell, D., and Chan, T. A. (2019). The Evolving Landscape of Biomarkers for Checkpoint Inhibitor Immunotherapy. *Nat. Rev. Cancer* 19, 133–150. doi:10.1038/s41568-019-0116-x

Hervouet, E., Vallette, F. M., and Cartron, P.-F. (2010). Impact of the DNA Methyltransferases Expression on the Methylation Status of Apoptosis-Associated Genes in Glioblastoma Multiforme. *Cell Death Dis* 1, e8. doi:10.1038/cddis.2009.7

Jia, Q., Wu, W., Wang, Y., Alexander, P. B., Sun, C., Gong, Z., et al. (2018). Local Mutational Diversity Drives Intratumoral Immune Heterogeneity in Non-small Cell Lung Cancer. *Nat. Commun.* 9, 5361. doi:10.1038/s41467-018-07767-w

Kulis, M., and Esteller, M. (2010). DNA Methylation and Cancer. *Adv. Genet.* 70, 27–56. doi:10.1016/b978-0-12-380866-0.60002-2

Li, F., Aljahdali, I., and Ling, X. (2019). Cancer Therapeutics Using Survivin BIRC5 as a Target: what Can We Do after over Two Decades of Study? *J. Exp. Clin. Cancer Res.* 38, 368. doi:10.1186/s13046-019-1362-1

Lim, Y. W., Coles, G. L., Sandhu, S. K., Johnson, D. S., Adler, A. S., and Stone, E. L. (2021). Single-cell Transcriptomics Reveals the Effect of PD-L1/tgf-β Blockade on the Tumor Microenvironment. *BMC Biol.* 19, 107. doi:10.1186/s12915-021-01034-z

Liu, A., Dai, Y., Mendez, E. F., Hu, R., Fries, G. R., Najera, K. E., et al. (2021a). Genome-Wide Correlation of DNA Methylation and Gene Expression in Postmortem Brain Tissues of Opioid Use Disorder Patients. *Int. J. Neuropsychopharmacol.* 24, 879–891. doi:10.1093/ijnp/pyab043

Liu, J., Ottaviani, D., Sefta, M., Desbrousses, C., Chapeaublanc, E., Aschero, R., et al. (2021b). A High-Risk Retinoblastoma Subtype with Stemness Features, Dedifferentiated Cone States and Neuronal/ganglion Cell Gene Expression. *Nat. Commun.* 12, 5578. doi:10.1038/s41467-021-25792-0

Liu, S., Yu, G., Liu, L., Zou, X., Zhou, L., Hu, E., et al. (2021c). Identification of Prognostic Stromal-Immune Score-Based Genes in Hepatocellular Carcinoma Microenvironment. *Front. Genet.* 12, 625236. doi:10.3389/fgene.2021.625236

Liu, X. S., Gao, Y., Liu, C., Chen, X. Q., Zhou, L. M., Yang, J. W., et al. (2020). Comprehensive Analysis of Prognostic and Immune Infiltrates for E2F Transcription Factors in Human Pancreatic Adenocarcinoma. *Front. Oncol.* 10, 606735. doi:10.3389/fonc.2020.606735

Lu, J. E., Francis, J. H., Dunkel, I. J., Shields, C. L., Yu, M. D., Berry, J. L., et al. (2019). Metastases and Death Rates after Primary Enucleation of Unilateral Retinoblastoma in the USA 2007-2017. *Br. J. Ophthalmol.* 103, 1272–1277. doi:10.1136/bjophthalmol-2018-312915

Mani, D. R., Krug, K., Zhang, B., Satpathy, S., Clauser, K. R., Ding, L., et al. (2022). Cancer Proteogenomics: Current Impact and Future Prospects. *Nat. Rev. Cancer*. doi:10.1038/s41568-022-00446-5

Moore, L. D., Le, T., and Fan, G. (2013). DNA Methylation and its Basic Function. *Neuropsychopharmacol*, 38, 23–38. doi:10.1038/npp.2012.112

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

Norrie, J. L., Nityanandam, A., Lai, K., Chen, X., Wilson, M., Stewart, E., et al. (2021). Retinoblastoma from Human Stem Cell-Derived Retinal Organoids. *Nat. Commun.* 12, 4535. doi:10.1038/s41467-021-24781-7

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

Sarver, A. L., Xie, C., Riddle, M. J., Forster, C. L., Wang, X., Lu, H., et al. (2021). Retinoblastoma Tumor Cell Proliferation Is Negatively Associated with an Immune Gene Expression Signature and Increased Immune Cells. *Lab. Invest.* 101, 701–718. doi:10.1038/s41374-021-00573-x

Schefler, A. C., and Kim, R. S. (2021). Recent Advancements in the Management of Retinoblastoma and Uveal Melanoma. *Fac. Rev.* 10, 51. doi:10.12703/r/10-51

Schelker, M., Feau, S., Du, J., Ranu, N., Klipp, E., MacBeath, G., et al. (2017). Estimation of Immune Cell Content in Tumour Tissue Using Single-Cell RNA-Seq Data. *Nat. Commun.* 8, 2032. doi:10.1038/s41467-017-02289-3

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303

Sina, A. A. I., Carrascosa, L. G., and Trau, M. (2019). DNA Methylation-Based Point-of-Care Cancer Detection: Challenges and Possibilities. *Trends Molecular Medicine* 25, 955–966. doi:10.1016/j.molmed.2019.05.014

Singh, L., Pushker, N., Sen, S., Singh, M. K., Bakhshi, S., Chawla, B., et al. (2015). Expression of CDC25A and CDC25B Phosphatase Proteins in Human Retinoblastoma and its Correlation with Clinicopathological Parameters. *Br. J. Ophthalmol.* 99, 457–463. doi:10.1136/bjophthalmol-2014-305830

Steele, N. G., Carpenter, E. S., Kemp, S. B., Sirihorachai, V. R., The, S., Delrosario, L., et al. (2020). Multimodal Mapping of the Tumor and Peripheral Blood Immune Landscape in Human Pancreatic Cancer. *Nat. Cancer* 1, 1097–1112. doi:10.1038/s43018-020-00121-4

Stirzaker, C., Millar, D. S., Paul, C. L., Warnecke, P. M., Harrison, J., Vincent, P. C., et al. (1997). Extensive DNA Methylation Spanning the Rb Promoter in Retinoblastoma Tumors. *Cancer Res.* 57, 2229

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

Tamborero, D., Rubio-Perez, C., Muiños, F., Sabarinathan, R., Piulats, J. M., Muntasell, A., et al. (2018). A Pan-Cancer Landscape of Interactions between Solid Tumors and Infiltrating Immune Cell Populations. *Clin. Cancer Res.* 24, 3717–3728. doi:10.1158/1078-0432.ccr-17-3509

Tian, Y., Morris, T. J., Webster, A. P., Yang, Z., Beck, S., Feber, A., et al. (2017). ChAMP: Updated Methylation Analysis Pipeline for Illumina BeadChips. *Bioinformatics (Oxford, England)* 33, 3982–3984. doi:10.1093/bioinformatics/btx513

Tsuruma, T., Hata, F., Torigoe, T., Furuhata, T., Idenoue, S., Kurotaki, T., et al. (2004). Phase I Clinical Study of Anti-apoptosis Protein, Survivin-Derived Peptide Vaccine Therapy for Patients with Advanced or Recurrent Colorectal Cancer. *J. Transl Med.* 2, 19. doi:10.1186/1479-5876-2-19

Wan, J., Oliver, V. F., Wang, G., Zhu, H., Zack, D. J., Merbs, S. L., et al. (2015). Characterization of Tissue-specific Differential DNA Methylation Suggests Distinct Modes of Positive and Negative Gene Expression Regulation. *BMC genomics* 16, 49. doi:10.1186/s12864-015-1271-4

Wang, K., Chen, Y., Ahn, S., Zheng, M., Landoni, E., Dotti, G., et al. (2020). GD2-specific CAR T Cells Encapsulated in an Injectable Hydrogel Control Retinoblastoma and Preserve Vision. *Nat. Cancer* 1, 990–997. doi:10.1038/s43018-020-00119-y

Welby, E., Lakowski, J., Di Foggia, V., Budinger, D., Gonzalez-Cordero, A., Lun, A. T. L., et al. (2017). Isolation and Comparative Transcriptome Analysis of Human Fetal and iPSC-Derived Cone Photoreceptor Cells. *Stem Cel. Rep.* 9, 1898–1915. doi:10.1016/j.stemcr.2017.10.018

Witte, T., Plass, C., and Gerhauser, C. (2014). Pan-cancer Patterns of DNA Methylation. *Genome Med.* 6, 66. doi:10.1186/s13073-014-0066-6

Yang, Y. (2015). Cancer Immunotherapy: Harnessing the Immune System to Battle Cancer. *J. Clin. Invest.* 125, 3335–3337. doi:10.1172/jci83871

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

Zeng, Y., He, T., Liu, J., Li, Z., Xie, F., Chen, C., et al. (2020). Bioinformatics Analysis of Multi-Omics Data Identifying Molecular Biomarker Candidates and Epigenetically Regulatory Targets Associated with Retinoblastoma. *Medicine* 99, e23314. doi:10.1097/md.0000000000023314

Zhang, S., Zheng, W., Jiang, D., Xiong, H., Liao, G., Yang, X., et al. (2021). Systematic Chromatin Accessibility Analysis Based on Different Immunological Subtypes of Clear Cell Renal Cell Carcinoma. *Front. Oncol.* 11, 575425. doi:10.3389/fonc.2021.575425

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 Mao, Shen, Xu and Zhong. 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: 08 June 2022
doi: 10.3389/fgene.2022.911223

A Systematic Analysis Reveals the Prognostic and Immunological Role of Reptin/RUVBL2 in Human Tumors

*Xiaoru Su, Gaoming Zheng, Zhifang Gui, Xiao Yang, Lahong Zhang and Feng Pan**

Department of Clinical Laboratory, Affiliated Hospital of Hangzhou Normal University, Hangzhou, China

Reptin/RUVBL2 is involved in the remodeling of chromatin, DNA damage repair, and regulation of the cell cycle, all of which help to play essential roles in cancer. However, relevant pan-cancer analysis of Reptin is lacking. This study first investigated the potential oncogenic roles of Reptin and revealed a relationship between Reptin with clinicopathological characteristics and immune infiltration based on big data. Here, we showed that Reptin is overexpressed in many cancers. A significant association exists between the expression of Reptin and the prognosis of cancer cases. Reptin had a meaningful interaction with the immune infiltration of $CD4^+$ Th1 cells and immune modulator genes in multiple cancer types. And negative correlation exists between Reptin and cancer-associated fibroblasts in BRCA, PRAD, TGCT, and THYM. A significant negative association exists between Reptin and regulatory T cells in TGCT and THCA. Moreover, Reptin is significantly associated with genomic heterogeneity, DNA mismatch repair genes, methyltransferase, and RNA modification genes in specific cancer types. Spliceosome, Hippo signaling pathway, DNA replication pathway, and acetyltransferase activity-associated functions were observed in the effect of Reptin on the tumor. This systematic analysis highlights Reptin as a vital cancer regulator among numerous genes and proved its potential prognosticator value and therapeutic target role for specific tumor types.

Keywords: reptin, cancer, prognosis, genomic heterogeneity, immune infiltration

OPEN ACCESS

Edited by:
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Lei Gu,
Max Planck Institute for Heart and Lung Research, Germany
Jean Rosenbaum,
Institut National de la Santé et de la Recherche Médicale (INSERM), France

****Correspondence:***
Feng Pan
panfeng1240@126.com

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

Received: *02 April 2022*
Accepted: *16 May 2022*
Published: *08 June 2022*

Citation:
Su X, Zheng G, Gui Z, Yang X, Zhang L and Pan F (2022) A Systematic Analysis Reveals the Prognostic and Immunological Role of Reptin/RUVBL2 in Human Tumors. Front. Genet. 13:911223. doi: 10.3389/fgene.2022.911223

INTRODUCTION

Cancer poses the highest social and economic burden of all human diseases, is the second worldwide cause of death, and is a significant barrier to increasing the average lifespan (Mattiuzzi and Lippi, 2019). Despite current therapeutics, including surgery, radiotherapy, chemotherapy, etc., showing certain clinical benefits, the survival rate and prognosis are still deficient. Searching for new tumor molecular markers for cancer diagnosis and therapy is necessary. It is crucial to conduct a comprehensive pan-cancer analysis of a specific gene of interest. The Cancer Genome Atlas (TCGA) database allows us to study the correlation between clinical prognosis and potential molecular mechanisms in pan-cancer.

AAA+ ATPases are a protein superfamily that power various cellular activities. Reptin, also known as RUVBL2, belongs to the AAA+ superfamily and was first reported in the late 1990s in various species, and they share homology with the bacterial RuvB helicase (Mao and Houry, 2017). Research has shown that Reptin is associated with DNA double-strand break repair (Raymond et al., 2015) and participates in the chromatin remodeling complexes INO80 within the nucleus, essential

for homologous recombination (Alatwi and Downs, 2015). The ATPase activity of Reptin is required for chromatin remodeling, and the chromatin state further manipulates cell fate (Tarangelo et al., 2015; Blanco et al., 2021). Depletion of Reptin impairs telomerase RNP accumulation (Venteicher et al., 2008). Reptin is also observed to play specific roles in mitosis (Gentili et al., 2015). Reptin is one of the components of the HSP90-interacting chaperone-like complex R2TP (Li et al., 2017b), which can regulate the assembly of small nuclear ribonucleoprotein particles (snRNPs) specific proteins (Malinova et al., 2017). The snRNPs can comprise the spliceosome and mediate the splicing of pre-mRNAs, and RNA undergoing different modifications and splicing could have different biological functions (Matera and Wang, 2014; Gu et al., 2020). As part of the histone acetyltransferase TIP60 complexes, Reptin can also promote E2F1 transcription and regulate the cell cycle (Tarangelo et al., 2015).

In addition, Reptin interacted with bona fide proto-oncogenes, such as c-MYC and β-catenin (Mikesch et al., 2018a; Armenteros-Monterroso et al., 2019). Reptin promotes tumor cell proliferation and progression (Javary et al., 2021; Nakamura et al., 2021). Moreover, *in vivo* silencing of Reptin led to the arrest of tumor progression in human hepatocellular carcinoma xenografts in mice (Menard et al., 2010).

Reptin is a critical and valuable gene. However, studies on the role of Reptin have been confined to specific types of cancer so far. No pan-cancer analysis was found here on the connection between Reptin and multiple cancer types, exploring the common ground based on big data.

The present study comprehensively described the prognostic value and immunological role of Reptin based on the Cancer Genome Atlas (TCGA), Genotype-Tissue Expression (GTEx), cBio Cancer Genomics Portal (cBioPortal), and Tumor Immune Estimation Resource (TIMER). Moreover, the relationship between Reptin and genomic heterogeneity, including Neoantigen (NEO), tumor mutational burden (TMB), microsatellite instability (MSI), DNA mismatch repair (MMR) genes, methyltransferase, RNA modification genes, and related signaling pathways were also evaluated to search for the possible biological mechanism of Reptin in the prognosis or immune within various cancer types.

MATERIALS AND METHODS

Reptin Differential Expression Level Analysis

Reptin expression information from 31 distinct tissues was obtained from the GTEx data (https://commonfund.nih.gov/GTEx). The Reptin level between different cancers and normal tissues was observed by integrating the dataset of TCGA and GTEx in UCSC XENA (https://xenabrowser.net/datapages/), in which clinical follow-up data for cancer patients and RNA sequencing data in TPM format were obtained. All data were processed uniformly by the Toil process (Vivian et al., 2017) and were analyzed after log2 conversion. R software (Version 3.6.3) was utilized to conduct the analysis. The R package "ggplot2 (Version 3.3.3) was used to draw box plots. The Clinical proteomic tumor analysis consortium (CPTAC) dataset from the UALCAN website (http://ualcan.path.uab.edu/) (Chandrashekar et al., 2017) was utilized to conduct Reptin protein-level analysis, in which ten types of tumors: breast cancer (BRCA), ovarian cancer (OV), colon cancer (COAD), clear cell renal cell carcinoma (RCC), uterine corpus endometrial carcinoma (UCEC), lung adenocarcinoma (LUAD), Head and neck squamous carcinoma (HNSC), Pancreatic adenocarcinoma (PAAD), Glioblastoma multiforme (GBM), and Hepatocellular carcinoma (LIHC) were available to conduct the compare between tumor and normal tissue. In addition, Reptin expression levels in pathological stages I-IV of tumors were visualized with violin plots *via* GEPIA.

Survival Prognostic Analysis

The pan-cancer dataset was downloaded from UCSC website, and the high-quality TCGA prognosis dataset was obtained (Liu et al., 2018). To ensure the reliability of the results, the cancer types for which sample number less than ten were excluded. The links between Reptin and prognosis of cancer cases, including overall survival (OS) and disease-specific survival (DSS) in all TCGA tumors, were investigated by forest plots and Kaplan-Meier survival curves. Univariate survival analysis was conducted to determine the hazard ratios and 95% confidence intervals. The R package maxstat (version: 0.7–25) was used to calculate the best cut-off value for Reptin. The minimum number of grouped samples was greater than 25%, the maximum number was less than 75%, and the patients were divided into high-expression and low-expression groups. All values were transformed with a log2 transformation. The R package (survival, version 3.2–7) was utilized, and a Log-rank test was used for statistical calculation.

Genetic Variant Analysis

A detailed analysis of genetic alteration frequency, copy number alterations (CNAs), and variations in the types of Reptin genes was investigated on the cBioPortal website (https://www.cbioportal.org/) (Cerami et al., 2012; Gao et al., 2013) across all TCGA tumor types. The mutated site of Reptin is shown in the protein structure sketch map. Additionally, the association between genetic variation of Reptin and the OS of cancer patients was calculated.

Correlation of Reptin Expression Levels With Neoantigen, Tumor Mutational Burden, Microsatellite Instability, Mismatch Repair, Methyltransferase, and RNA Modification Genes

The standardized pan-cancer data was obtained from the UCSC website. The level 4 simple nucleotide variation dataset of all TCGA tumors was downloaded from the GDC website (https://portal.gdc.cancer.gov/), which was pretreated by MuTect2 software (Beroukhim et al., 2010). R software package maftools (version 2.8.05) was utilized to calculate the TMB of each tumor. MSI scores were obtained from a previous study (Bonneville et al., 2017). All expression values have been

transformed by log2. The expression data were applied to study the association between Reptin and mutations in five types of MMRs (mutL homolog 1 (MLH1), mutS homolog 2 (MSH2), mutS homolog 6 (MSH6), PMS1 homolog 2, mismatch repair system component (PMS2), epithelial cell adhesion molecule (EPCAM)). The association between Reptin and DNA methyltransferase 1 (DNMT1), DNA methyltransferase-2 (DNMT2), DNA methyltransferase 3 alpha (DNMT3A), and DNA methyltransferase 3 beta (DNMT3B) were also evaluated. The data of Reptin and three types of RNA modification gene markers, including N1-methyladenosine (m1A), 5-methylcytosine (m5C), N6-methyladenosine (m6A), were extracted from each sample. All normalized values were log2 transformed. Correlation analysis was done by Spearman's method. Sangerbox tools (http://www.sangerbox.com/tool) were used.

Immune Infiltration and Immune Checkpoint Gene Analysis

Reptin expression data and the level of immune cell infiltrates were observed in different tumor types *via* TIMER (https://cistrome.shinyapps.io/timer/) (Li et al., 2016; Li et al., 2017a). Two types of immune checkpoint genes [Inhibitory (24 types), Stimulatory (36 types)] were derived from a previous report (Thorsson et al., 2018). All values were log2 transformed. Then Spearman correlation analysis proceeded, and a relevant heat map was created.

Gene Enrichment Analysis

Based on the STRING database (https://string-db.org/), we downloaded the experimentally determined Reptin-binding proteins. The parameter settings were the following: minimum required interaction score ["High confidence (0.700)"], max number of interactors ["custom value"], and max interactors ["100"] in the first shell. Cytoscape (version 3.7.2) MCODE plug-in was used to visualize the clustered sub-networks, and Cyto-Hubba was applied to determine the top 10 hub genes ranked by Maximal Clique Centrality. Subsequently, GEPIA displayed the top 100 Reptin-interacting genes and performed a Pearson correlation analysis between Reptin and the selected top 10 genes. We converted the expression values to log2 (TPM) for all dot plots. TIMER was utilized to observe the related heat map in the pan-cancer. The overlap between Reptin-binding genes and Reptin-interacting genes is displayed in the Venn diagram. Then the two groups of data, Reptin-binding and Reptin-interacting genes, were utilized to conduct the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and Gene Ontology (GO) analysis. Which GO analysis included three parts: Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). These results are shown in the bubble diagram. The "ClusterProfiler" (3.14.3) R (Yu et al., 2012) package was employed for KEGG pathway analysis, and GO annotations were obtained from the "org.Hs.eg.db" (3.10.0) Bioconductor packages for humans. The results were displayed using the R packages "ggplot2 (3.3.3)".

Statistical Analysis

The Kruskal-Wallis test was applied to calculate the differential expression of Reptin in normal human tissues. The expression level of Reptin in TCGA tumors and normal human tissues were calculated using Wilcoxon rank-sum test, and the HR and *p*-value of survival analysis were calculated with Univariate Cox regression methods. According to the Reptin expression level, Kaplan Meier analysis was utilized to explore the survival time of tumor patients. Correlation analysis was performed by Spearman's method or Pearson's method, and $p < 0.05$ was considered significant.

RESULTS

Reptin Differential Expression Analysis

The mRNA expression, protein level, and different pathological stage expressions of Reptin were analyzed. The results from GTEx showed that differences exist in Reptin expression in various human tissues. Reptin was expressed highest in the Testis tissues and lowest in the Blood (**Figure 1A**). Subsequently, TCGA and GTEx data were combined to study the expression of Reptin across all TCGA tumor tissues and normal human tissues. Reptin mRNA expression, illustrated in **Figure 1B**, was significantly higher in tumor tissues compared to normal tissues in the majority of cancers, including adrenocortical carcinoma (ACC), bladder urothelial carcinoma (BLCA), BRCA, cervical squamous cell carcinoma (CESC), cholangiocarcinoma (CHOL), COAD, lymphoid neoplasm diffuse large B cell lymphoma (DLBC), esophageal carcinoma (ESCA), GBM, HNSC, kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), brain lower grade glioma (LGG), LIHC, LUAD, lung squamous cell carcinoma (LUSC), OV, PAAD, prostate adenocarcinoma (PRAD), rectum adenocarcinoma (READ), skin cutaneous melanoma (SKCM), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), thymoma (THYM), UCEC and uterine carcinosarcoma (UCS) datasets. However, the mRNA expression of Reptin was significantly lower in acute myeloid leukemia (LAML) and testicular germ cell tumor (TGCT) datasets. At the same time, there was no difference in kidney chromophobe (KICH), and pheochromocytoma and paraganglioma (PCPG).

The comparison of Reptin total protein levels according to the CPTAC database indicated higher Reptin levels in the tumor tissues of OV, COAD, RCC, UCEC, LUAD, HNSC, PAAD, GBM, and LIHC than those in normal human tissues; however, there was no statistical difference in the BRCA (**Figure 1C**). GEPIA was applied to study the association between Reptin and different pathological stages of TCGA tumors. The results showed that Reptin expression in stages I–IV of cancer cases, including ACC, HNSC, KICH, KIRC, KIRP, and LIHC, all had *p* values of less than 0.05 (**Figure 1D**).

A Prognosis Analysis of Reptin in Pan-Cancer

We used the survival metrics of OS and DSS to estimate the link between Reptin and the prognosis of multiple cancer cases. Cox

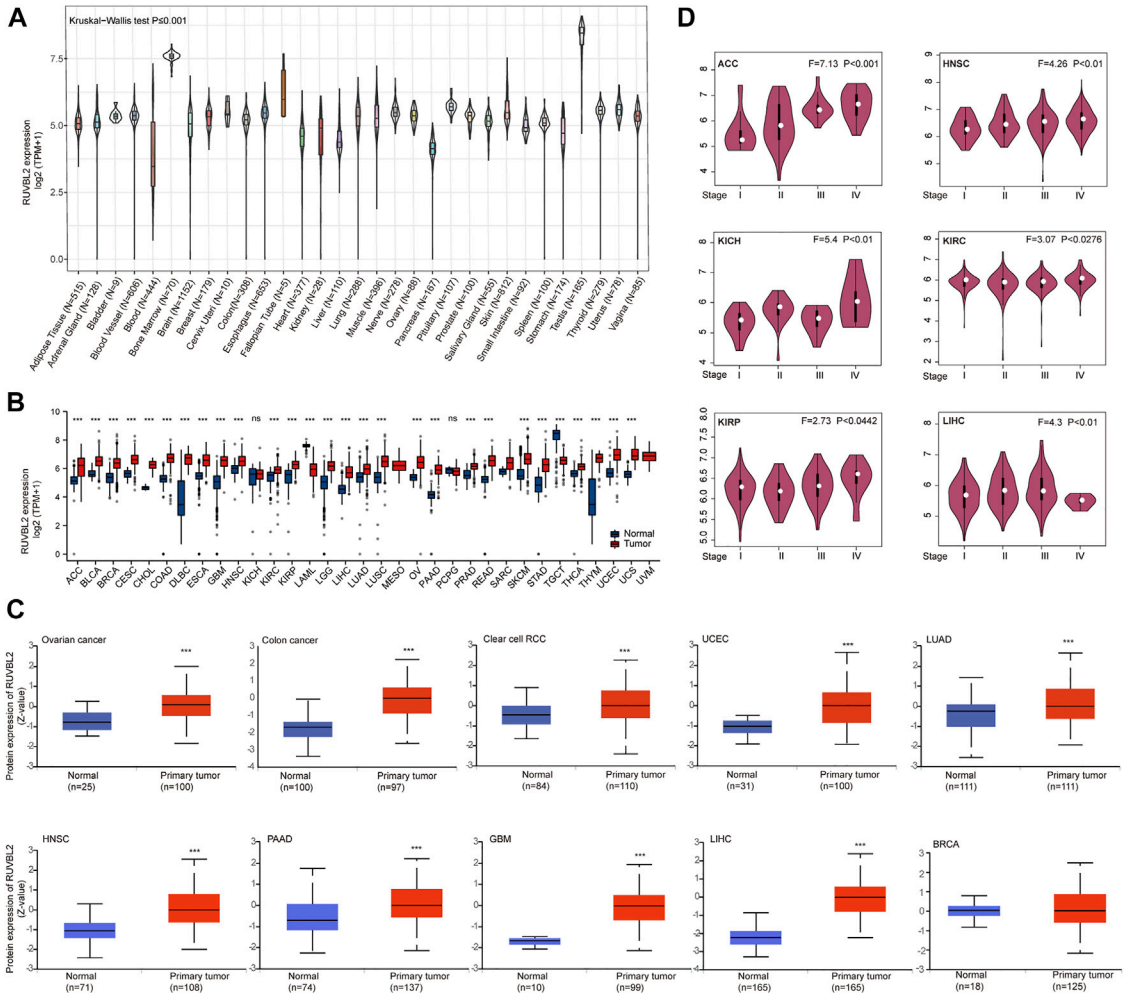


FIGURE 1 | Differential expression of Reptin. **(A)** The expression of Reptin in normal human tissues. **(B)** The expression of Reptin in 33 types of TCGA cancer. **(C)** The protein level of Reptin in BRCA, OV, COAD, RCC, UCEC, LUAD, HNSC, PAAD, GBM, and LIHC. **(D)** The expression of Reptin in pathological stages I-IV of ACC, HNSC, KICH, KIRC, KIRP, and LIHC. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

regression analysis suggested that Reptin is significantly related to OS in Glioma (GBMLGG), sarcoma (SARC), mesothelioma (MESO), uveal melanoma (UVM), LGG, LUAD, HNSC, LIHC, metastatic SKCM (SKCM-M), SKCM, ACC, KICH, THYM, and DLBC (**Figure 2A**). Besides, Kaplan-Meier survival analyses were conducted based on the expression of Reptin. The results indicated that the upregulation of Reptin was strongly related to worse OS in GBMLGG, LGG, LUAD, SARC, HNSC, LIHC, MESO, SKCM-M, SKCM, ACC, BLCA, and KIRP. In contrast, low expression of Reptin was remarkably related to worse OS in CESC, DLBC, READ, and THYM (**Figure 2B**). In addition, we explored the link between Reptin and DSS in these types of tumors. Cox regression analysis suggested that Reptin is significantly related to DSS in 16 types of cancer, including GBMLGG, LGG, LUAD, SARC, KIRP, KIPAN, PRAD, HNSC, LIHC, THCA, MESO, SKCM-M, SKCM, UVM, ACC, and KICH (**Figure 3A**). The Kaplan–Meier survival analyses demonstrated that the upregulation of Reptin was very strongly related to worse DSS in 20 types of cancer, including GBMLGG, LGG, LUAD, SARC, KIRP, KIPAN, PRAD, HNSC, LIHC, MESO, SKCM-M, SKCM, UVM, ACC, BLCA, BRCA, CHOL, KIRC, LUSC, and Stomach and Esophageal carcinoma (STES) (**Figure 3B**).

Reptin Genetic Variant Pattern Analysis

The cBioPortal database was applied to observe the genetic variant status of Reptin in TCGA tumors. The results are displayed in **Figure 4A**. As shown, UCEC patients had the highest frequency of Reptin alteration (~3.8%), in which mutation was the primary type. The second group was SKCM patients with an alteration frequency of Reptin of about 3.6%, in which every observed case had a mutation. It is worth emphasizing that all UCS cases with genetic variants had amplification of copy number alteration (CNA) (~3.5% frequency). The copy number deletion of Reptin was the primary type in LGG patients (~2% frequency). Thymoma cases showed an alteration frequency of less than 1%.

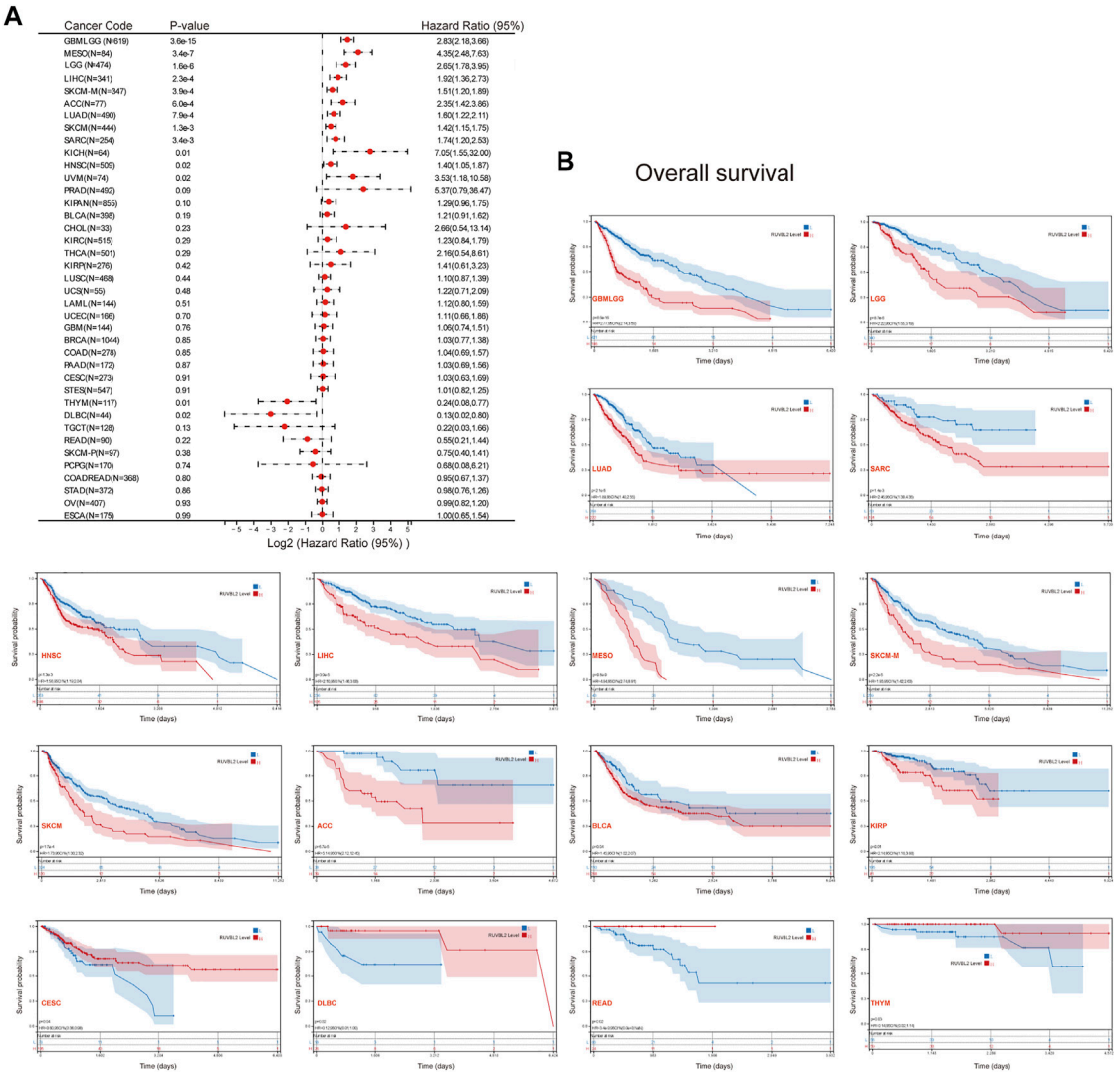


FIGURE 2 | Prognosis analysis of Reptin in pan-cancer. **(A)** Forest plot of the associations between Reptin and OS in 39 types of tumors. **(B)** Kaplan-Meier survival analysis of the relationship between Reptin and OS. Only positive results were presented.

Figure 4B displays the case number, types, and sites of the Reptin genetic variants. In short, missense mutation of Reptin was the primary type. Additionally, we explored the possible association between alterations of Reptin and the prognosis of various cancer patients. The results demonstrated that Reptin had no difference in OS between the patients with genetic alterations and patients without alterations (**Figure 4C**).

Correlation of Reptin Expression Levels With Neoantigen, Tumor Mutational Burden, Microsatellite Instability, Mismatch Repair, Methyltransferase, and RNA Modification Genes

Sangerbox tools were applied to calculate the number of new antigens and determine the association between Reptin and neoantigens. As shown in **Figure 5A**, the results demonstrated that Reptin is positively related to NEO in LUAD (R = 0.17, p = 0.0004), SARC (R = 0.24, p = 0.008), LUSC (R = 0.14, p = 0.003), and negatively related to NEO in READ (R = −0.24, p = 0.04) and PAAD (R = −0.26, p = 0.02). A highly significant positive association exists between Reptin and TMB in 15 types of tumors, including GBM (R = 0.19, p = 0.02), GBMLGG (R = 0.35, $p = 8.4e^{-20}$), LGG (R = 0.21, $p \leq 0.001$), LUAD (R = 0.23, $p = 1.1e^{-7}$), COAD (R = 0.12, p = 0.04), COADREAD (R = 0.12, p = 0.02), STES (R = 0.17, $p \leq 0.001$), SARC (R = 0.31, $p = 9.07e^{-7}$), KIPAN (R = 0.24, $p = 4.12e^{-10}$), STAD (R = 0.26, $p = 6.73e^{-8}$), LUSC (R = 0.12, p = 0.005), LIHC (R = 0.13, p = 0.01), PAAD (R = 0.36, $p \leq 0.001$), BLCA (R = 0.097, p = 0.049), and ACC (R = 0.40, $p \leq 0.001$), however, significant negative correlations exist between Reptin and TMB in THYM (R = −0.46, p = 1.15e-7) and THCA (R = −0.09, p = 0.045). The relationship between

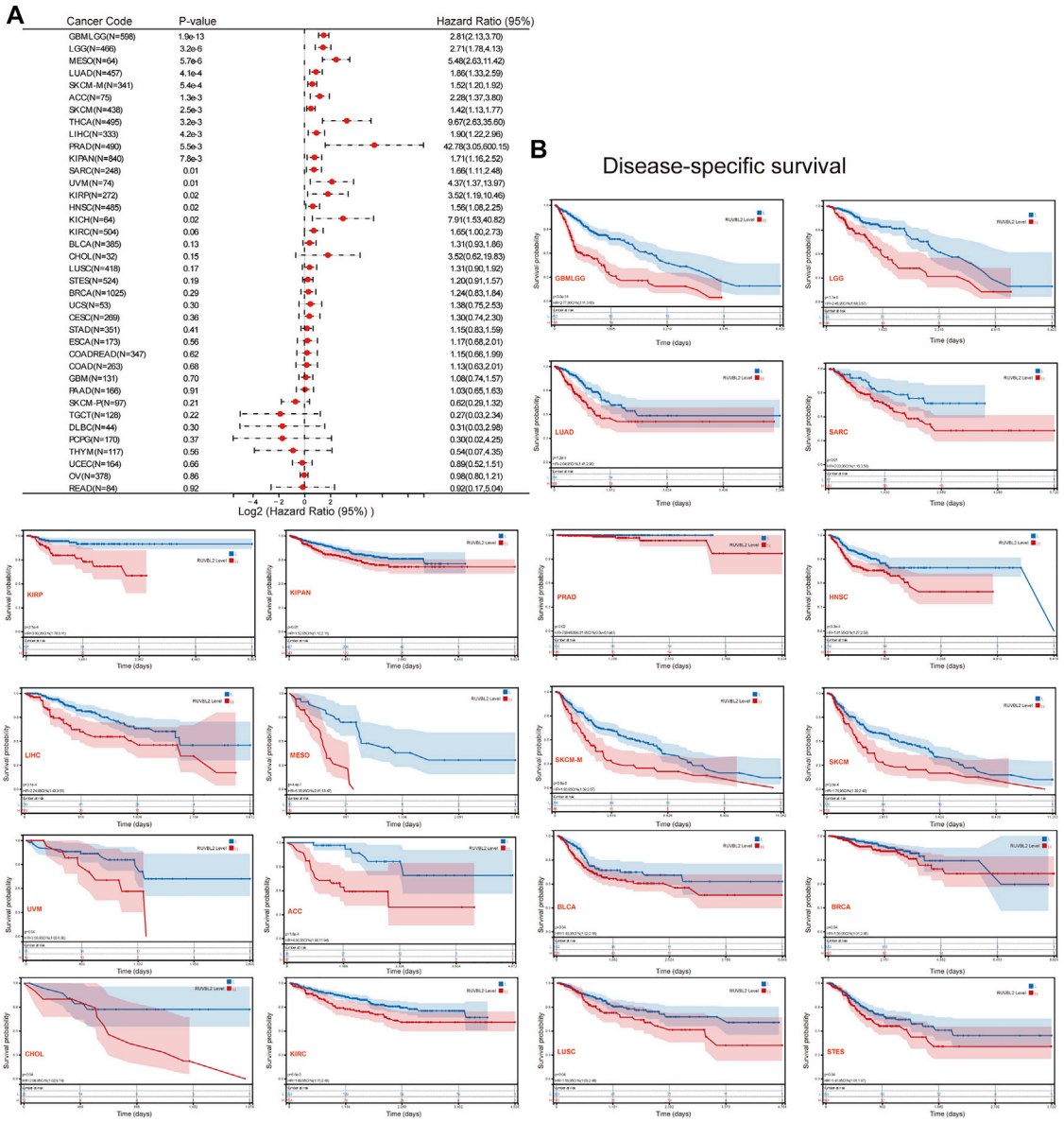


FIGURE 3 | Prognosis analysis of Reptin in pan-cancer. **(A)** Forest plot of the associations between Reptin and DSS in 38 types of tumors. **(B)** Kaplan-Meier survival analysis of the relationship between Reptin and DSS. The positive results are shown.

Reptin and MSI was investigated in TCGA cancer. The results indicated that there were significant positive correlations between Reptin and MSI in 11 types of cancers, including STES (R = 0.21, p = 2.06e-7), SARC (R = 0.31, $p = 4.53e^{-7}$), KIPAN (R = 0.098, p = 0.009), STAD (R = 0.24, $p \leq 0.001$), HNSC (R = 0.10, p = 0.02), KIRC (R = 0.21, $p \leq 0.001$), LIHC (R = 0.27, $p = 1.19e^{-7}$), UVM (R = 0.32, p = 0.003), BLCA (R = 0.10, p = 0.03), KICH (R = 0.26, p = 0.03) and DLBC (R = 0.31, p = 0.03), while there was negative correlation between Reptin and MSI in GBMLGG (R = −0.19, $p = 4.07e^{-7}$).

Moreover, the relationship between Reptin and five types of MMR mutation (MLH1, MSH2, MSH6, PMS2, and EPCAM) was evaluated. The results are shown in **Figure 5B**. Reptin is significantly correlated with different MMR mutations in 31 types of cancer, except in CHOL and DLBC. Moreover, the association between Reptin and methyltransferases (DNMT1, DNMT2, DNMT3A, and DNMT3B) was visualized in **Figure 5B**. We observed that Reptin is strongly related to methyltransferases in 32 types of TCGA cancer, except in UCS.

It is well known that RNA modification is critical for post-transcriptional gene regulation that can be classified as "writers" (installation), "readers" (recognition), and "erasers" (removal). Thus, the Spearman correlations

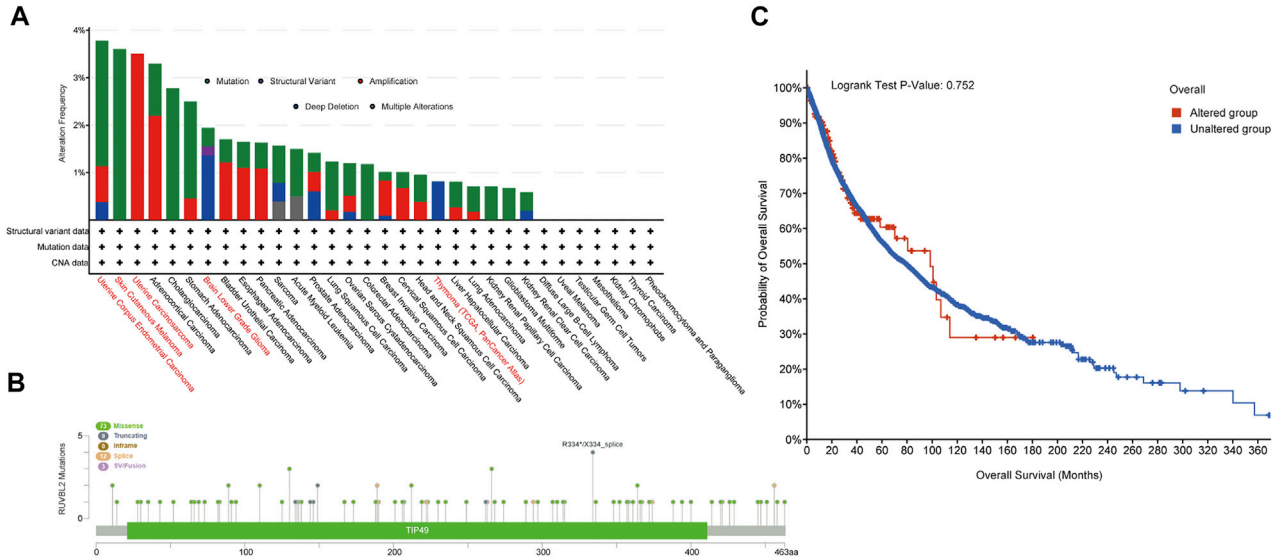


FIGURE 4 | Mutation characteristics of Reptin. **(A)** The frequencies of different genetic variant types in all TCGA tumor samples. **(B)** The distribution of Reptin mutation sites. **(C)** The comparison of clinical survival prognosis analysis between the Reptin altered group and the unaltered group.

between Reptin and these regulatory genes were calculated. As shown in **Figure 5C**, we found numerous meaningful associations. Significant positive correlations exist between Reptin and m1A, m5C, and m6A in multiple cancer types. The molecules that interacted the most with Reptin included TRMT61A, ALYREF, and ELAVL1. TRMT61A is positively associated with Reptin as an m1A writer in 35 types of cancer, and ALYREF, as the m5C reader, was positively associated with Reptin in 35 cancer types. ELAVL1 as an m6A reader positively correlates with Reptin in 32 types of cancer. Notably, in CHOL, Reptin was only significantly negatively correlated with m1A writer TRMT10C and m5C writer NSUN3.

Immune Infiltration and Immune Modulator Gene Analysis

We observed a statistically significant negative association between Reptin and the infiltration of cancer-associated fibroblasts in BRCA, PRAD, TGCT, and THYM based on all algorithms (**Figure 6A** and **Figure 6B**). Moreover, significant positive correlations between Reptin and infiltration of $CD4^+$ Th1 cells were found in most types of tumors, except in CHOL and UCEC, based on the XCELL algorithm (**Figure 6C**). Besides, there was a statistically negative association between Reptin and regulatory T cells (Tregs) in TGCT and THCA based on all algorithms (**Figure 6D**).

Furthermore, we extracted sixty immune checkpoint genes and evaluated their correlations with Reptin levels. The results indicated that Reptin levels were significantly negatively related to many immune regulators in most tumors, such as THYM, THCA, PAAD, DLBC, UCS, KIPAN, BRCA PRAD, and GBM. Notably, Reptin levels did not correlate with the sixty immune checkpoint genes in CHOL (**Figure 6E**).

Functional Enrichment Analysis of Reptin-Related Partners

We screened out the Reptin-binding proteins and Reptin-interacting genes to conduct the functional enrichment analysis. Based on the STRING database, 100 experimental supported Reptin-binding proteins were obtained, and the interaction network was shown in **Figure 7A**. Highly interconnected five sub-clusters from Cytoscape plug-in MCODE and the top 10 hub genes ranked by Maximal Clique Centrality from Cyto-Hubba were also shown in **Figure 7A**. Besides, GEPIA was utilized to download the top 100 genes that highly interacted with Reptin. **Figure 7B** showed that Reptin was significantly positively correlated with HSPBP1 (HSPA (Hsp70) binding protein 1) (R = 0.73), PRMT1 (protein arginine methyltransferase 1) (R = 0.72), PRPF31 (pre-mRNA processing factor 31) (R = 0.67), BCL2L12 (BCL2 like 12) (R = 0.62), GRWD1 (glutamate-rich WD repeat-containing 1) (R = 0.57), NOSIP (nitric oxide synthase interacting protein) (R = 0.65), TOMM40 (translocase of outer mitochondrial membrane 40) (R = 0.71), UBE2S (ubiquitin-conjugating enzyme E2 S) (R = 0.64), PIH1D1 (PIH1 domain containing 1) (R = 0.57), and EXOSC5 (exosome component 5) (R = 0.63), all p value less than 0.001. The heat map depicted significant positive correlations between Reptin and the ten genes mentioned above in most tumors (**Figure 7C**).

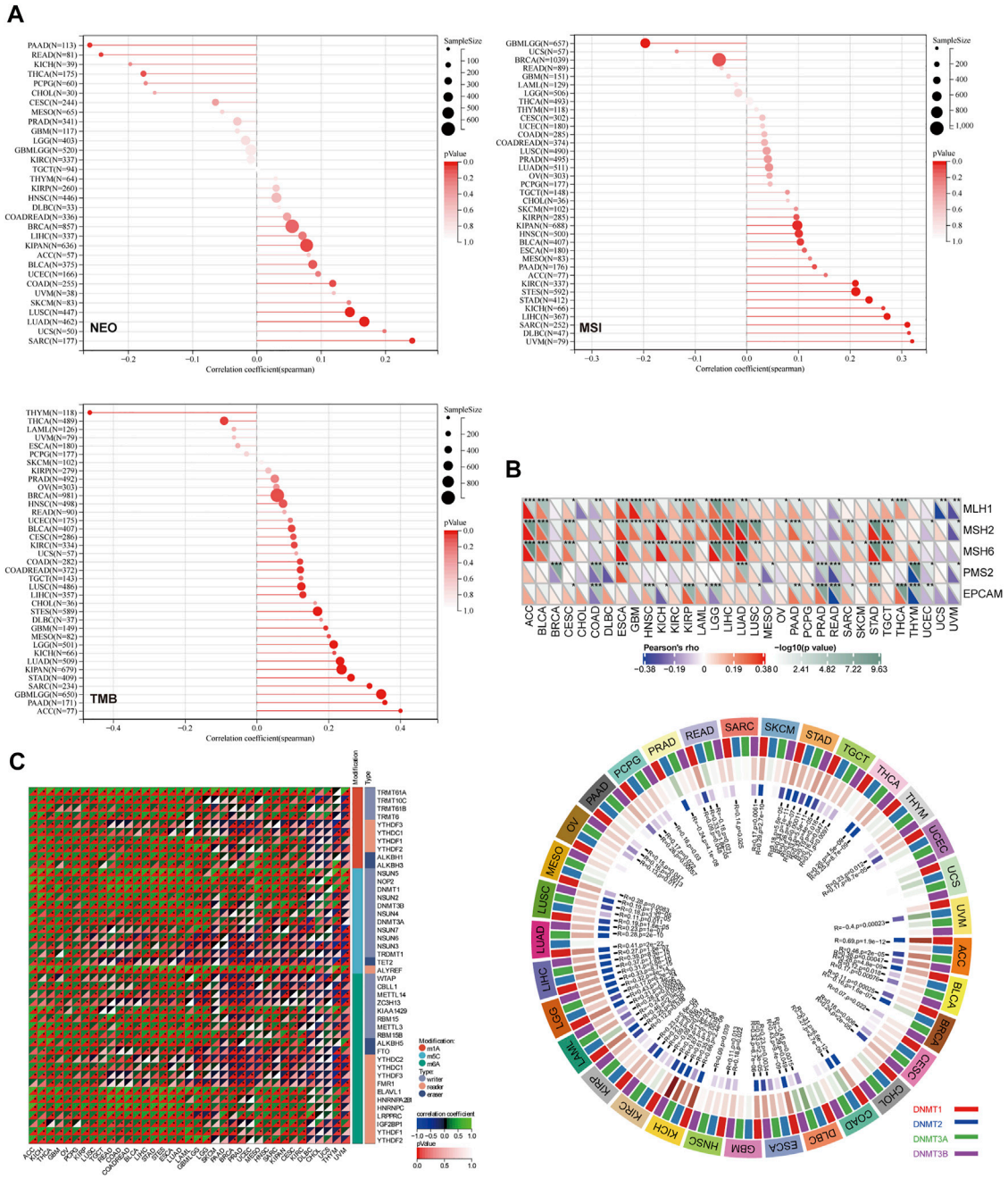


FIGURE 5 | The correlation analysis between Reptin and NEO, TMB, MSI, MMRs, methyltransferase, and RNA modification genes. **(A)** The correlation between Reptin and NEO, TMB, MSI in pan-cancer. **(B)** The correlation between Reptin and MMRs, methyltransferase. **(C)** The correlation between Reptin and RNA modification genes. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

The overlap analysis exhibited two common members, RUVBL1 and PIH1D1 (**Figure 7D**), indicating that RUVBL1 and PIH1D1 are genes that both bind and interact with Reptin. KEGG analysis showed that the spliceosome, Hippo signaling, and DNA replication pathways were involved in the effect of Reptin on the tumor, and the GO study suggested that the majority of these genes were related to acetyltransferase activity (**Figure 7E**).

DISCUSSION

The ATPase activity of Reptin is essential for most of its roles in cancer. The role of Reptin in cancer cell migration and invasion has been investigated (Breig et al., 2017). Reptin interacts with p53 and suppresses its anti-tumor activity (Maslon et al., 2010). Targeting Pontin/Reptin can potentially treat cancer (Menard et al., 2010; Mikesch et al., 2018b; Assimon et al., 2019). However,

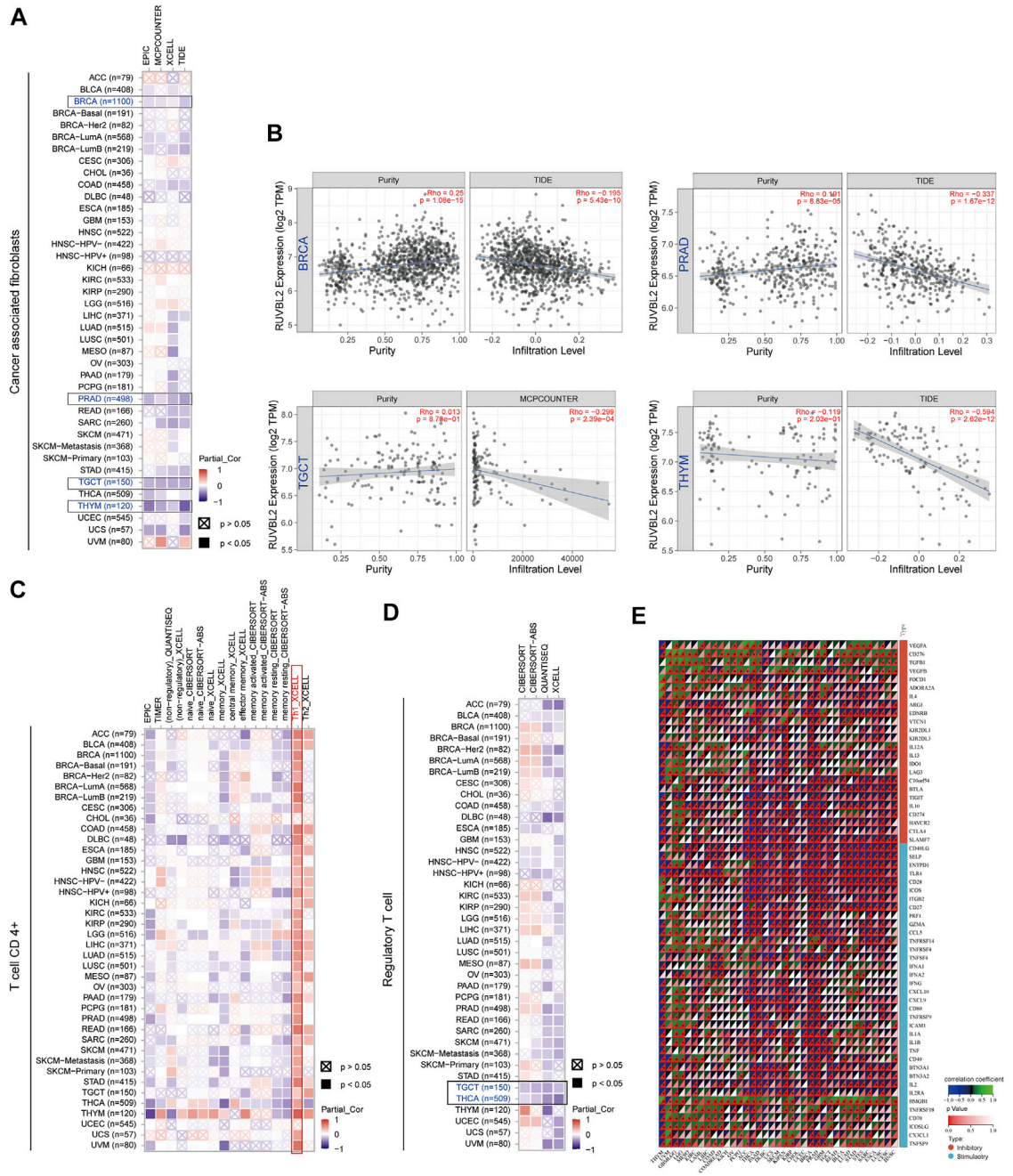


FIGURE 6 | Association between Reptin and immune cell infiltration across all TCGA cancers. **(A)** The association between Reptin and the infiltration of cancer-associated fibroblasts. **(B)** The relationship between Reptin and the infiltration level of cancer-associated fibroblasts in BRCA, PRAD, TGCT, and THYM. **(C)** The correlation between Reptin and the infiltration of $CD4^+$ T cell. **(D)** The correlation between Reptin and the infiltration of regulatory T cell. **(E)** The correlation between Reptin and immune modulator genes in pan-cancer. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

whether Reptin can play a role in the prognosis or immune of various cancers *via* the same mechanisms remains clarified.

To clarify the expression landscape of Reptin, we integrated multiple data across all TCGA tumor types. The study found that Reptin was overexpressed in numerous cancers but not every cancer. It appears downregulated in AML and TGCT. Upregulation of Reptin was significantly related to worse clinical prognosis in most tumors. However, overexpressed Reptin suggested a good prognosis in CESC, DLBC, READ, and THYM.

In vitro experiments have already demonstrated the pro-proliferative role of Pontin in oral squamous cell carcinoma (OSCC) (Nakamura et al., 2021), and Pontin overexpression has been described as a poor prognostic indicator for HNSC (Lin et al., 2020). However, relevant analysis about Reptin in

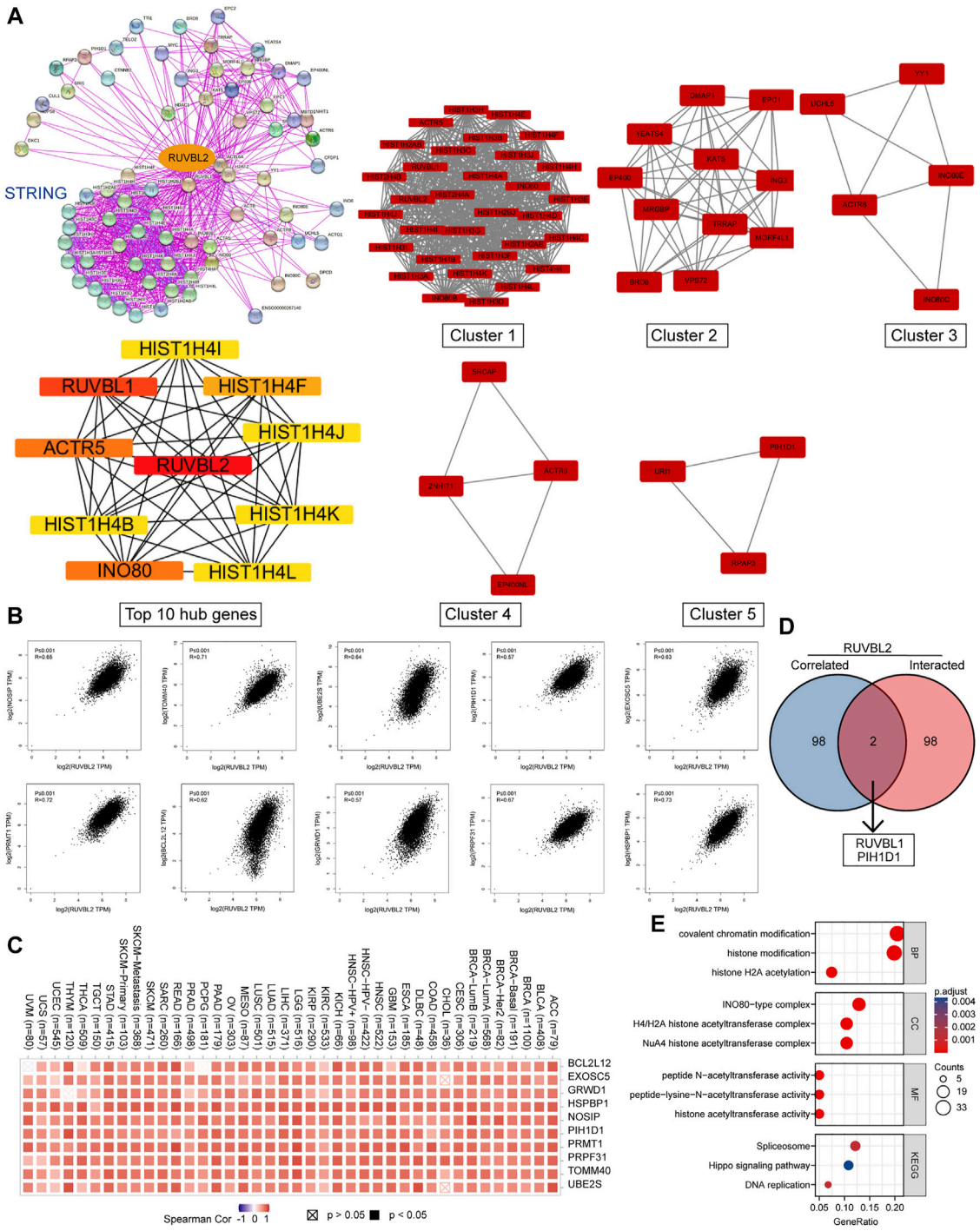


FIGURE 7 | The results of gene enrichment analysis. **(A)** 100 Reptin-binding proteins were shown *via* the STRING tool. Five sub-network modules (Clusters 1–5) and the top 10 hub genes were presented using the Cytoscape plug-in, MCODE, and cytoHubba. **(B)** The association between Reptin and the selected top 10 Reptin-interacting genes. **(C)** The relevant heatmap data in cancer. **(D)** The overlap analysis of Reptin-binding and interacting genes. **(E)** The results of KEGG and GO analysis were visualized.

HNSC is lacking. Based on the big data, we found that overexpression of Reptin was related to worse prognostic in HNSC. For lung cancer, studies have reported that Excavatolide B may exhibit cytotoxic effects against lung cancer A549 cells by affecting Reptin expression (Velmurugan et al., 2017). Reptin is highly expressed in NSCLC patient tumors, related to poor survival (Mikesch et al., 2018b). Reptin was overexpressed in gastric tissue (Li et al., 2010), and its high

expression was related to advanced nodal disease in colon cancer patients (Flavin et al., 2011). For hepatocellular liver carcinoma, overexpressed Reptin is necessary for hepatocyte proliferation (Javary et al., 2021), and it could be used to predict poor prognosis (Yan et al., 2019). Within the excretory system, a previous study showed that overexpressed Reptin is related to the poor differentiation of RCC cells and portends poor outcomes for patients (Ren et al., 2013). These findings outlined above were consistent with our results based on big data. In addition, no reports were found of Reptin on the prognosis of CESC, DLBC, READ, and THYM. Our big data-based study complements this aspect.

Here, we describe a decrease in Reptin expression in AML. Previous results found that Reptin is required for the oncogenic potential of leukemia (Osaki et al., 2013), and silencing Reptin resulted in AML cell apoptosis in engrafted mice by increased transcriptional activation c-MYB (Armenteros-Monterroso et al., 2019). Admittedly, Reptin plays an important role in AML. The level of Reptin in leukemia cells needs to be determined. Although the study found that Reptin is readily detected in MLL-AF9 mouse myeloid cells, they did not compare Reptin levels in leukemia cells and normal blood cells (Osaki et al., 2013).

It should be noted that AML is a "liquid" cancer; the choice of the control, normal tissue, to be compared with the leukemic cells is critical. Our study derived leukemia cells compared with normal blood cells from the GTEx database. Further experiment verification of this point may be required in the future.

In addition, Reptin expression correlated positively with the pathological stage in ACC, HNSC, KICH, KIRC, KIRP, and LIHC. Reptin levels at all four stages seem identical in KIRC and KIRP, with Reptin being most expressed in stage IV cases of KICH. However, there is even a decreased expression of Reptin in stage IV in LIHC, which may be worthy of further study.

We also analyzed the genetic variant pattern of Reptin. The missense mutation of Reptin was the primary type. UCEC patients had the highest frequency of Reptin alteration (~3.8%). All UCS cases with genetic variants had amplification of copy number, and the copy number deletion of Reptin was the primary type in LGG. Genetic alterations in Reptin did not impact the survival of tumor patients.

Moreover, we provide evidence for the potential functional link between Reptin and NEO, TMB, MSI, MMR, DNA methyltransferase, and RNA modification genes across all TCGA tumors.

Tumor neoantigen, encoded by the tumor cell mutant gene, is highly immunogenic. We can design a new vaccine to achieve the therapeutic effect based on the immune activity of the new antigen (Xu et al., 2020). Our results suggested significant positive correlations between Reptin and tumor neoantigen in LUAD, SARC, and LUSC, and significant negative correlations were found in READ and PAAD. TMB and MSI are related to the immunotherapy response. TMB can be employed in tumor cells to show the number of mutations, and it has been identified as a potential biomarker in advancing targeted therapy (Choucair et al., 2020). Moreover, it can predict the response to immune checkpoint protein PD- L1 treatment (Allgauer et al., 2018). At the same time, the expression of PD-L1 was reported to promote tumor progression and formation (Burks et al., 2019) and enhance the immune evasion potential of cancer (Chu et al., 2022). Studies have shown that high MSI can lead to many molecular changes, including increased expression of neoantigens, high TMB, and abundant tumor-infiltrating lymphocytes (Chang et al., 2018). In the present study, there were negative correlations between Reptin and TMB in THYM and THCA. Positive correlations were observed between Reptin and TMB in GBM, GBMLGG, LGG, LUAD, COAD, COADREAD, STES, SARC, KIPAN, STAD, LUSC, LIHC, PAAD, BLCA, and ACC. Besides, the high expressed Reptin was negatively associated with MSI in GBMLGG and positively related to MSI in 11 types of tumors. These findings indicate that Reptin expression can affect the NEO, TMB, and MSI in specific tumors, thereby changing the patient's response to immunotherapy.

It is well known that cancer is driven by various genetic and epigenetic alterations, and DNA methylation as a key regulator has a crucial role in epigenetic gene regulation (Gu et al., 2015). DNA methylation can control gene expression without changing the DNA sequence and is involved in phenotypic development and altered susceptibility to disease (Moore et al., 2013; Bauer et al., 2016). We found that Reptin expression was significantly related to methyltransferases expression in 32 tumor types. Reptin was positively correlated with m1A, m5C, and m6A in most tumors. TRMT61A, ALYREF, and ELAVL1 are the most interacting molecules with Reptin. Moreover, loss of the critical gene of the mismatch repair mechanism will cause higher somatic mutations and accelerate tumorigenesis (Maki-Nevala et al., 2019). We found that Reptin expression was significantly related to MMR mutation in 31 tumor types. Notably, there was no correlation between Reptin and MMR mutation in CHOL and DLBC, and there was no correlation between Reptin and DNMT1, DNMT2, DNMT3A, and DNMT3B methyltransferases in UCS. In CHOL, Reptin was only significantly negatively correlated with m1A writer TRMT10C and m5C writer NSUN3.

In summary, although the role of Reptin was dependent on the cancer type, it is closely related to epigenetic regulation and mismatch repair mechanism in most tumors.

In addition, we explored the association between Reptin and immune cell infiltrates. The immune system can recognize and clear tumor cells under ordinary circumstances, but tumor cells can manipulate immune cells to evade the immune system's surveillance (Muenst et al., 2016). In the tumor microenvironment, cancer-associated fibroblasts can modulate the function of tumor-infiltrating immune cells and affect cancer development (Chen and Song, 2019). Tumor-infiltrating lymphocytes have been identified as an independent predictor of survival and cancer sentinel node status (Santos et al., 2019). We found a negative correlation between Reptin and the infiltration of cancer-associated fibroblasts in BRCA, PRAD, TGCT, and THYM. There were significant negative correlations between Reptin and regulatory T cells in TGCT and THCA. We also found a positive association between Reptin and the infiltration of $CD4^+$ Th1 cells in most kinds of tumors. These findings were confirmed by previous research, which reported that Reptin is required for T-cell development (Arnold et al., 2012), and Reptin, associated with GATA-binding protein 3, inhibits the expression of the cyclin-

dependent kinase inhibitor 2c to help the cell proliferation of Th2 (Hosokawa et al., 2013). We also collected sixty immune checkpoint-associated genes in these immune cells. The results indicated significant negative correlations between Reptin and numerous immune regulators in most tumors. However, Reptin levels did not correlate with the sixty immune checkpoint genes in CHOL. Based on these findings, we infer that BRCA, PRAD, TGCT, THYM, and THCA may be suitable for anti-Reptin immunotherapy, while CHOL was not.

We also integrated the data on Reptin binding partners and Reptin interacting genes for enrichment analyses. We found that two critical genes, RUVBL1, and PIH1D1, both bind and interact with Reptin, indicating their uniqueness and importance for Reptin among numerous genes in tumors. This result was consistent with previous reports (Mao and Houry, 2017). We identified the spliceosome, Hippo signaling pathway, DNA replication pathway, and acetyltransferase activity-associated functions involved in the impact of Reptin on the tumor.

In short, we confirmed that the upregulation of Reptin is related to a worse prognosis in most tumors. It is associated with genome heterogeneity, MMRs, methyltransferase, RNA modification genes, immune cell infiltration, and immune checkpoint genes in certain tumors. Although these findings highlight Reptin as a vital cancer regulator among numerous genes and prove its potential prognosticator value and therapeutic target role for specific tumor types, further experimental validation, and mechanistic studies are needed.

DATA AVAILABILITY STATEMENT

The dataset supporting the conclusions of this article is available from the public databases. These data can be found in The Cancer Genome Atlas (TCGA, https://tcga-data.nci.nih.gov/tcga/) and The Genotype-Tissue Expression (GTEx, https://commonfund.nih.gov/GTEx).

AUTHOR CONTRIBUTIONS

XS and FP contributed to the study's conception and design. All authors were responsible for data acquisition, analysis, or interpretation. GZ, ZG, and XY performed the statistical analysis. The first draft of the manuscript was written by XS. LZ, and FP reviewed the manuscript and worked on the language. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Zhejiang Medical and Health Science and Technology Program Fund (Nos. 2019KY501, 2019KY131, and 2021KY888), Hangzhou Medical and Health Science and Technology Program Fund (No. A20210008), and Hangzhou Science and Technology Plan Development Project Fund (Nos. 20191203B106 and 20201231Y026).

REFERENCES

Alatwi, H. E., and Downs, J. A. (2015). Removal of H2A.Z by INO 80 Promotes Homologous Recombination. *EMBO Rep.* 16 (8), 986–994. doi:10.15252/embr.201540330

Allgäuer, M., Budczies, J., Christopoulos, P., Endris, V., Lier, A., Rempel, E., et al. (2018). Implementing Tumor Mutational Burden (TMB) Analysis in Routine Diagnostics-A Primer for Molecular Pathologists and Clinicians. *Transl. Lung Cancer Res.* 7 (6), 703–715. doi:10.21037/tlcr.2018.08.14

Armenteros-Monterroso, E., Zhao, L., Gasparoli, L., Brooks, T., Pearce, K., Mansour, M. R., et al. (2019). The AAA+ATPase RUVBL2 Is Essential for the Oncogenic Function of C-MYB in Acute Myeloid Leukemia. *Leukemia* 33 (12), 2817–2829. doi:10.1038/s41375-019-0495-8

Arnold, C. N., Pirie, E., Dosenovic, P., McInerney, G. M., Xia, Y., Wang, N., et al. (2012). A Forward Genetic Screen Reveals Roles for Nfkbid , Zeb1 , and Ruvbl2 in Humoral Immunity. *Proc. Natl. Acad. Sci. U.S.A.* 109 (31), 12286–12293. doi:10.1073/pnas.1209134109

Assimon, V. A., Tang, Y., Vargas, J. D., Lee, G. J., Wu, Z. Y., Lou, K., et al. (2019). CB-6644 Is a Selective Inhibitor of the RUVBL1/2 Complex with Anticancer Activity. *ACS Chem. Biol.* 14 (2), 236–244. doi:10.1021/acschembio.8b00904

Bauer, T., Trump, S., Ishaque, N., Thürmann, L., Gu, L., Bauer, M., et al. (2016). Environment-induced Epigenetic Reprogramming in Genomic Regulatory Elements in Smoking Mothers and Their Children. *Mol. Syst. Biol.* 12 (3), 861. doi:10.15252/msb.20156520

Beroukhim, R., Mermel, C. H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., et al. (2010). The Landscape of Somatic Copy-Number Alteration across Human Cancers. *Nature* 463 (7283), 899–905. doi:10.1038/nature08822

Blanco, M. A., Sykes, D. B., Gu, L., Wu, M., Petroni, R., Karnik, R., et al. (2021). Chromatin-state Barriers Enforce an Irreversible Mammalian Cell Fate Decision. *Cell. Rep.* 37 (6), 109967. doi:10.1016/j.celrep.2021.109967

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.*, 1–15. doi:10.1200/PO.17.00073

Breig, O., Yates, M., Neaud, V., Couchy, G., Grigoletto, A., Lucchesi, C., et al. (2017). Metalloproteinase Meprin α Regulates Migration and Invasion of Human Hepatocarcinoma Cells and Is a Mediator of the Oncoprotein Reptin. *Oncotarget* 8 (5), 7839–7851. doi:10.18632/oncotarget.13975

Burks, J., Fleury, A., Livingston, S., and Smith, J. P. (2019). ISG15 Pathway Knockdown Reverses Pancreatic Cancer Cell Transformation and Decreases Murine Pancreatic Tumor Growth via Downregulation of PDL-1 Expression. *Cancer Immunol. Immunother.* 68 (12), 2029–2039. doi:10.1007/s00262-019-02422-9

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 (5), 401–404. doi:10.1158/2159-8290.CD-12-0095

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* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Chang, L., Chang, M., Chang, H. M., and Chang, F. (2018). Microsatellite Instability: A Predictive Biomarker for Cancer Immunotherapy. *Appl. Immunohistochem. Mol. Morphol.* 26 (2), e15–e21. doi:10.1097/PAI.0000000000000575

Chen, X., and Song, E. (2019). Turning Foes to Friends: Targeting Cancer-Associated Fibroblasts. *Nat. Rev. Drug Discov.* 18 (2), 99–115. doi:10.1038/s41573-018-0004-1

Choucair, K., Morand, S., Stanbery, L., Edelman, G., Dworkin, L., and Nemunaitis, J. (2020). TMB: a Promising Immune-Response Biomarker, and Potential Spearhead in Advancing Targeted Therapy Trials. *Cancer Gene Ther.* 27 (12), 841–853. doi:10.1038/s41417-020-0174-y

Chu, Z., Gu, L., Hu, Y., Zhang, X., Li, M., Chen, J., et al. (2022). STAG2 Regulates Interferon Signaling in Melanoma via Enhancer Loop Reprogramming. *Nat. Commun.* 13 (1), 1859. doi:10.1038/s41467-022-29541-9

Flavin, P., Redmond, A., McBryan, J., Cocchiglia, S., Tibbitts, P., Fahy-Browne, P., et al. (2011). RuvBl2 Cooperates with Ets2 to Transcriptionally Regulate hTERT in Colon Cancer. *FEBS Lett.* 585 (15), 2537–2544. doi:10.1016/j.febslet.2011.07.005

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

Gentili, C., Castor, D., Kaden, S., Lauterbach, D., Gysi, M., Steigemann, P., et al. (2015). Chromosome Missegregation Associated with RUVBL1 Deficiency. *PLoS One* 10 (7), e0133576. doi:10.1371/journal.pone.0133576

Gu, L., Frommel, S. C., Frommel, S. C., Oakes, C. C., Simon, R., Grupp, K., et al. (2015). BAZ2A (TIP5) Is Involved in Epigenetic Alterations in Prostate Cancer and its Overexpression Predicts Disease Recurrence. *Nat. Genet.* 47 (1), 22–30. doi:10.1038/ng.3165

Gu, L., Wang, L., Chen, H., Hong, J., Shen, Z., Dhall, A., et al. (2020). CG14906 (Mettl4) Mediates m6A Methylation of U2 snRNA in Drosophila. *Cell. Discov.* 6, 44. doi:10.1038/s41421-020-0178-7

Hosokawa, H., Tanaka, T., Kato, M., Shinoda, K., Tohyama, H., Hanazawa, A., et al. (2013). Gata3/Ruvbl2 Complex Regulates T Helper 2 Cell Proliferation via Repression of Cdkn2c Expression. *Proc. Natl. Acad. Sci. U.S.A.* 110 (46), 18626–18631. doi:10.1073/pnas.1311100110

Javary, J., Allain, N., Ezzoukhry, Z., Di Tommaso, S., Dupuy, J. W., Costet, P., et al. (2021). Reptin/RUVBL2 Is Required for Hepatocyte Proliferation *In Vivo*, Liver Regeneration and Homeostasis. *Liver Int.* 41 (6), 1423–1429. doi:10.1111/liv.14886

Li, B., Severson, E., Pignon, J.-C., Zhao, H., Li, T., Novak, J., et al. (2016). Comprehensive Analyses of Tumor Immunity: Implications for Cancer Immunotherapy. *Genome Biol.* 17 (1), 174. doi:10.1186/s13059-016-1028-7

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017a). 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

Li, W., Zeng, J., Li, Q., Zhao, L., Liu, T., Björkholm, M., et al. (2010). Reptin Is Required for the Transcription of Telomerase Reverse Transcriptase and Over-expressed in Gastric Cancer. *Mol. Cancer* 9, 132. doi:10.1186/1476-4598-9-132

Li, Y., Zhao, L., Yuan, S., Zhang, J., and Sun, Z. (2017b). Axonemal Dynein Assembly Requires the R2TP Complex Component Pontin. *Development* 144 (24), 4684–4693. doi:10.1242/dev.152314

Lin, D., Lin, B., Bhanot, H., Riou, R., Abt, N. B., Rajagopal, J., et al. (2020). RUVBL1 Is an Amplified Epigenetic Factor Promoting Proliferation and Inhibiting Differentiation Program in Head and Neck Squamous Cancers. *Oral Oncol.* 111, 104930. doi:10.1016/j.oraloncology.2020.104930

Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell.* 173 (2), 400–e11. doi:10.1016/j.cell.2018.02.052

Mäki-Nevala, S., Valo, S., Ristimäki, A., Sarhadi, V., Knuutila, S., Nyström, M., et al. (2019). DNA Methylation Changes and Somatic Mutations as Tumorigenic Events in Lynch Syndrome-Associated Adenomas Retaining Mismatch Repair Protein Expression. *EBioMedicine* 39, 280–291. doi:10.1016/j.ebiom.2018.12.018

Malinová, A., Cvačková, Z., Matějů, D., Hořejší, Z., Abéza, C., Vandermoere, F., et al. (2017). Assembly of the U5 snRNP Component PRPF8 Is Controlled by the HSP90/R2TP Chaperones. *J. Cell. Biol.* 216 (6), 1579–1596. doi:10.1083/jcb.201701165

Mao, Y.-Q., and Houry, W. A. (2017). The Role of Pontin and Reptin in Cellular Physiology and Cancer Etiology. *Front. Mol. Biosci.* 4, 58. doi:10.3389/fmolb.2017.00058

Maslon, M. M., Hrstka, R., Vojtesek, B., and Hupp, T. R. (2010). A Divergent Substrate-Binding Loop within the Pro-oncogenic Protein Anterior Gradient-2 Forms a Docking Site for Reptin. *J. Mol. Biol.* 404 (3), 418–438. doi:10.1016/j.jmb.2010.09.035

Matera, A. G., and Wang, Z. (2014). A Day in the Life of the Spliceosome. *Nat. Rev. Mol. Cell. Biol.* 15 (2), 108–121. doi:10.1038/nrm3742

Mattiuzzi, C., and Lippi, G. (2019). Current Cancer Epidemiology. *J. Epidemiol. Glob. Health.* 9 (4), 217–222. doi:10.2991/jegh.k.191008.001

Ménard, L., Taras, D., Grigoletto, A., Haurie, V., Nicou, A., Dugot-Senant, N., et al. (2010). *In Vivo* silencing of Reptin Blocks the Progression of Human Hepatocellular Carcinoma in Xenografts and Is Associated with Replicative Senescence. *J. Hepatology* 52 (5), 681–689. doi:10.1016/j.jhep.2009.12.029

Mikesch, J.-H., Hartmann, W., Angenendt, L., Huber, O., Schliemann, C., Arteaga, M. F., et al. (2018a). AAA+ ATPases Reptin and Pontin as Potential Diagnostic and Prognostic Biomarkers in Salivary Gland Cancer - a Short Report. *Cell. Oncol.* 41 (4), 455–462. doi:10.1007/s13402-018-0382-8

Mikesch, J.-H., Schwammbach, D., Hartmann, W., Schmidt, L. H., Schliemann, C., Angenendt, L., et al. (2018b). Reptin Drives Tumour Progression and Resistance to Chemotherapy in Nonsmall Cell Lung Cancer. *Eur. Respir. J.* 52 (1), 1701637. doi:10.1183/13993003.01637-2017

Moore, L. D., Le, T., and Fan, G. (2013). DNA Methylation and its Basic Function. *Neuropsychopharmacol* 38 (1), 23–38. doi:10.1038/npp.2012.112

Muenst, S., Läubli, H., Soysal, S. D., Zippelius, A., Tzankov, A., and Hoeller, S. (2016). The Immune System and Cancer Evasion Strategies: Therapeutic Concepts. *J. Intern Med.* 279 (6), 541–562. doi:10.1111/joim.12470

Nakamura, A., Kakihara, Y., Funayama, A., Haga, K., Mikami, T., Kobayashi, D., et al. (2021). HEATR1, a Novel Interactor of Pontin/Reptin, Stabilizes Pontin/Reptin and Promotes Cell Proliferation of Oral Squamous Cell Carcinoma. *Biochem. Biophysical Res. Commun.* 557, 294–301. doi:10.1016/j.bbrc.2021.04.021

Osaki, H., Walf-Vorderwülbecke, V., Mangolini, M., Zhao, L., Horton, S. J., Morrone, G., et al. (2013). The AAA+ ATPase RUVBL2 Is a Critical Mediator of MLL-AF9 Oncogenesis. *Leukemia* 27 (7), 1461–1468. doi:10.1038/leu.2013.42

Raymond, A.-A., Benhamouche, S., Neaud, V., Di Martino, J., Javary, J., and Rosenbaum, J. (2015). Reptin Regulates DNA Double Strand Breaks Repair in Human Hepatocellular Carcinoma. *PLoS One* 10 (4), e0123333. doi:10.1371/journal.pone.0123333

Ren, J., Li, W., Liu, H., Yan, L., Jiao, W., Li, D., et al. (2013). Overexpression of Reptin in Renal Cell Carcinoma Contributes to Tumor Malignancies and its Inhibition Triggers Senescence of Cancer Cells. *Urologic Oncol. Seminars Orig. Investigations* 31 (7), 1358–1366. doi:10.1016/j.urolonc.2012.01.004

Santos, F. D. M. d., Silva, F. C. d., Pedron, J., Furian, R. D., Fortes, C., and Bonamigo, R. R. (2019). Association between Tumor-Infiltrating Lymphocytes and Sentinel Lymph Node Positivity in Thin Melanoma. *An. Bras. Dermatol.* 94 (1), 47–51. doi:10.1590/abd1806-4841.20197414

Tarangelo, A., Lo, N., Teng, R., Kim, E., Le, L., Watson, D., et al. (2015). Recruitment of Pontin/Reptin by E2f1 Amplifies E2f Transcriptional Response during Cancer Progression. *Nat. Commun.* 6, 10028. doi:10.1038/ncomms10028

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–e14. doi:10.1016/j.immuni.2018.03.023

Velmurugan, B. K., Yang, H.-H., Sung, P.-J., and Weng, C.-F. (2017). Excavatolide B Inhibits Nonsmall Cell Lung Cancer Proliferation by Altering Peroxisome Proliferator Activated Receptor Gamma Expression and PTEN/AKT/NF-Kβ Expression. *Environ. Toxicol.* 32 (1), 290–301. doi:10.1002/tox.22235

Venteicher, A. S., Meng, Z., Mason, P. J., Veenstra, T. D., and Artandi, S. E. (2008). Identification of ATPases Pontin and Reptin as Telomerase Components Essential for Holoenzyme Assembly. *Cell.* 132 (6), 945–957. doi:10.1016/j.cell.2008.01.019

Vivian, J., Rao, A. A., Nothaft, F. A., Ketchum, C., Armstrong, J., Novak, A., et al. (2017). Toil Enables Reproducible, Open Source, Big Biomedical Data Analyses. *Nat. Biotechnol.* 35 (4), 314–316. doi:10.1038/nbt.3772

Xu, P., Luo, H., Kong, Y., Lai, W.-F., Cui, L., and Zhu, X. (2020). Cancer Neoantigen: Boosting Immunotherapy. *Biomed. Pharmacother.* 131, 110640. doi:10.1016/j.biopha.2020.110640

Yan, T., Liu, F., Gao, J., Lu, H., Cai, J., Zhao, X., et al. (2019). Multilevel Regulation of RUVBL2 Expression Predicts Poor Prognosis in Hepatocellular Carcinoma. *Cancer Cell. Int.* 19, 249. doi:10.1186/s12935-019-0974-z

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 (5), 284–287. doi:10.1089/omi.2011.0118

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 Su, Zheng, Gui, Yang, Zhang and Pan. 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: 08 June 2022
doi: 10.3389/fgene.2022.911346

Unfolded Protein Response–Related Signature Associates With the Immune Microenvironment and Prognostic Prediction in Osteosarcoma

*Zhao Zhang†, Xincheng Liu†, Debin Cheng†, Jingyi Dang, Zhenzhou Mi, Yubo Shi, Lei Wang and Hongbin Fan**

Division of Musculoskeletal Cancer Service, Department of Orthopaedic Surgery, Xi-jing Hospital, The Fourth Military Medical University, Xi'an, China

Background: Osteosarcoma is a highly malignant bone tumor commonly occurring in adolescents with a poor 5-year survival rate. The unfolded protein response (UPR) can alleviate the accumulation of misfolded proteins to maintain homeostasis under endoplasmic reticulum stress. The UPR is linked to the occurrence, progression, and drug resistance of tumors. However, the function of UPR-related genes (UPRRGs) in disease progression and prognosis of osteosarcoma remains unclear.

Methods: The mRNA expression profiling and corresponding clinical features of osteosarcoma were acquired from TARGET and GEO databases. Consensus clustering was conducted to confirm different UPRRG subtypes. Subsequently, we evaluated the prognosis and immune status of the different subtypes. Functional analysis of GO, GSEA, and GSVA was used to reveal the molecular mechanism between the subtypes. Finally, four genes (*STC2*, *PREB*, *TSPYL2*, and *ATP6V0D1*) were screened to construct and validate a risk signature to predict the prognosis of patients with osteosarcoma.

Result: We identified two subtypes according to the UPRRG expression patterns. The subgroup with higher immune scores, lower tumor purity, and active immune status was linked to a better prognosis. Meanwhile, functional enrichment revealed that immune-related signaling pathways varied markedly in the two subtypes, suggesting that the UPR might influence the prognosis of osteosarcoma *via* influencing the immune microenvironment. Moreover, prognostic signature and nomogram models were developed based on UPRRGs, and the results showed that our model has an excellent performance in predicting the prognosis of osteosarcoma. qPCR analysis was also conducted to verify the expression levels of the four genes.

Conclusion: We revealed the crucial contribution of UPRRGs in the immune microenvironment and prognostic prediction of osteosarcoma patients and provided new insights for targeted therapy and prognostic assessment of the disease.

Keywords: unfolded protein response, prognostic prediction, immune infiltration, nomogram, osteosarcoma

OPEN ACCESS

Edited by:
William C. Cho,
QEH, Hong Kong SAR, China

Reviewed by:
Wweibin Zhang,
Ruijin Hospital, China
Weichun Guo,
Renmin Hospital of Wuhan University, China

****Correspondence:***
Hongbin Fan
fanhb@fmmu.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: *02 April 2022*
Accepted: *27 April 2022*
Published: *08 June 2022*

Citation:
Zhang Z, Liu X, Cheng D, Dang J, Mi Z, Shi Y, Wang L and Fan H (2022) Unfolded Protein Response–Related Signature Associates With the Immune Microenvironment and Prognostic Prediction in Osteosarcoma. Front. Genet. 13:911346. doi: 10.3389/fgene.2022.911346

INTRODUCTION

Osteosarcoma is one of the most frequent primary malignant bone tumors in adolescents, often occurring in fast-growing long bones (Tang et al., 2008; Arndt et al., 2012). With significant advances in surgical options, neoadjuvant chemotherapy, and diagnostic imaging, the overall survival of patients with osteosarcoma has been significantly prolonged (Li et al., 2015; Wakamatsu et al., 2019; Hao et al., 2021). Nevertheless, the outcome of metastatic and recurrent patients remains unsatisfactory due to its intensely aggressive feature, with a 5-year mortality ratio of below 20% (Ferrari et al., 2005). In addition, high tumor heterogeneity, leading to chemoresistance in some patients, remains challenging clinically (Xiao et al., 2015). Therefore, there is a need for the discovery of novel targets for targeted treatment and better prognostic markers for osteosarcoma.

The unfolded protein response (UPR) is a highly conserved homeostatic response in all eukaryotic cells and can help cells mitigate the accumulation of misfolded proteins in the endoplasmic reticulum (ER) (Hetz et al., 2020). Under ER stress conditions, such as hypoxia, nutritional deprivation, acidosis, and inflammatory stimuli, the UPR can activate three sensor proteins (IRE1α, PERK, and ATF6) to initiate specific signal transduction cascades which regulate the rate of protein production for maintaining ER homeostasis (Vanacker et al., 2017). Lately, extensive studies have indicated that the UPR acts a dual function in tumor occurrence and development. In earlier stages of tumorigenesis, the UPR is capable of exerting antitumor effects to hinder tumor transformation, while in the established tumors, tumor cells can induce chronic UPR to relieve ER stress-induced apoptosis and develop drug resistance to maintain tumor survival (Ma and Hendershot, 2004; Yang et al., 2019). Aberrant activation of the UPR was found in a wide range of tumors, including bladder cancer, cutaneous melanoma, and liver cancer (Houessinon et al., 2016; Wan et al., 2020; Zhu et al., 2021). Apart from a direct impact on tumor biology, the UPR also has the ability to remodel the tumor immune microenvironment (TIME) to regulate the crosstalk between immune cells, which serves an essential function in immune surveillance and immune escape (Zanetti et al., 2022). Mahadevan et al. demonstrated that the UPR can modulate the phenotype of dendritic cells and $CD8^+$ T cells to facilitate tumor growth. Notably, previous studies have confirmed the UPR to be relevant to the growth, prognosis, and drug resistance of osteosarcoma (Mahadevan et al., 2012; Yan et al., 2015; Shi et al., 2022). However, the role of the UPR-related gene (UPRRG) sets on prognostic prediction and immune infiltration in patients with osteosarcoma remains unclear.

Recently, following the advancement of the high-throughput sequencing technology for tumor genomics, probing new molecular patterns through bioinformatics approaches offers new insights for tumor treatment and prognosis evaluation (Qian et al., 2021). In the present study, we performed a comprehensive analysis of the prognosis and immune microenvironment of different molecular subtypes based on the expression of UPRRGs and revealed potential functional and signaling pathways. Moreover, we developed and validated a novel prognostic signature for predicting the prognosis of osteosarcoma patients *via* integrating risk scores and clinical features.

METHODS AND MATERIALS

Data Collection

The mRNA expression profiles and corresponding clinical features of 85 osteosarcoma patients were acquired from the TARGET database as a training cohort (https://ocg.cancer.gov/programs/target). UPRRGs were extracted from hallmark gene sets from the Molecular Signature Database (Zhang et al., 2021). In addition, mRNA expression profiles and relevant clinical information of 53 osteosarcoma patients from GSE21257 were obtained from the GEO database to serve as an external validation cohort (https://www.ncbi.nlm.nih.gov/geo/). The clinical information for osteosarcoma patients is presented in **Supplementary Table S1**.

Classification of Molecular Subtypes

To assess the biological functions of UPRRGs in OS, we first identified 15 prognosis-related UPRRGs for osteosarcoma based on a univariate Cox regression analysis. Subsequently, consensus clustering analysis was conducted based on the expression matrix of these 15 genes with the R package "ConsensusClusterPlus". For the analysis, 80% of the total samples of the target dataset are included in each iteration and reiterated 1,000 times to ensure cluster stability. The average consistency value and clinical significance within the clustering group were used together with the optimal number of clusters. Principal component analysis (PCA) was performed to confirm the gene expression patterns in the different clusters. KM curve and log-rank tests were employed to evaluate the prognosis of different clusters. The heat map was used to show the gene expression and clinical characteristics of the different clusters.

TIME Evaluation

The tumor microenvironment scores for individual samples in the different subtypes were assessed by the ESTIMATE algorithm (Yoshihara et al., 2013). The TIMER algorithm was conducted to comprehensively evaluate the abundance of immune infiltrating cells in each sample (Li et al., 2020). Moreover, the infiltration abundance of 28 immune cell types for an individual sample was tested by the single sample gene set enrichment analysis (ssGSEA) algorithm. The expression of immune checkpoint (ICP) genes in different subtypes was also evaluated.

Identification of DEG and Enrichment Analysis

Differentially expressed genes (DEGs) in different subtypes were analyzed *via* the R package "limma", and log2 (Foldchange) > 1.5 and FDR<0.05 were considered as the threshold. Gene Ontology (GO) terms of DEGs were enriched by using the R package "clusterProfiler" and visualized *via* Metascape (https://metascape.org) (Zhou et al., 2019). Gene set variation analysis (GSVA) was

conducted to verify functional pathway variations between the different subtypes according to GO: the biological process (BP). Also, Gene Set Enrichment Analysis (GSEA) was employed to analyze hallmark gene sets for different subtypes (Subramanian et al., 2005). |NES| > 1, NOM p-val < 0.05, and FDR < 0.25 were taken as the threshold.

Establishment and Validation of the Prognostic Signature

The previously obtained 15 prognosis-related UPRRGs were screened *via* least absolute shrinkage and selection operator (LASSO) Cox regression based on the R package "glmnet", and the minimum lamba is taken as the optimal value. Then, a multivariate Cox analysis was conducted to further optimize and establish the prognostic signature. The risk score in the training and validation cohorts was calculated with the following formula: Risk score = $\sum_{n}^{i}$ ($Coef_i$ * X_i). In the training cohort, all patients were classified into low- and high-risk groups by the median risk score, and the overall survival (OS) between the two groups was investigated using the KM curve and log-rank tests. In addition, receiver operating characteristic (ROC) curves were applied to measure the effectiveness of prognostic models. Furthermore, the abovementioned formula was also used in patients from GSE23257 to generate a risk model to validate the prognostic benefit. Finally, we integrated different clinical characteristics including gender, age, metastatic status, and disease site to assess if the risk score was an independent prognostic element in osteosarcoma patients based on univariate and multivariate Cox regression.

Construction and Calibration of the Nomogram

A nomogram model was established to forecast the 3-year and 5-year survival probability of osteosarcoma patients *via* integrating risk scores and clinical profiles. The C-index, ROC curves, and calibration plots were employed to survey the predictive performance of the constructed nomogram in both cohorts.

Patient Sample Collection and Cell Culture

A total of six osteosarcoma patient tissues and corresponding adjacent normal tissues were obtained from the operation and stored in liquid nitrogen banks. The study was approved by the institutional review board of Xijing Hospital, the Fourth Military Medical University, and all patients provided signed informed consent. Moreover, human osteosarcoma cell lines HOS and MG-63 were procured from Procell Life Science & Technology Co. Ltd and cultured in DMEM medium (Gibco, Shanghai, China) containing 10% fetal bovine serum (FBS; Gibco, Shanghai, China) and 1% (v/v) penicillin/streptomycin (Gibco, Shanghai, China). The osteoblast cell line hFOB 1.19 was provided by Dr. Jianping Bai of Xijing Hospital and cultured in DMEM/F12 medium with 10% FBS, 0.3 mg/ml G418 (Procell, Wuhan, China), and 1% (v/v) penicillin/streptomycin. At 37 C, 5% CO_2 environment, all cell lines were incubated.

TABLE 1 | Primer sequences of the candidate genes.

Gene	Sequence (5' -> 3')
STC2	Forward:GGGTGTGGCGTGTTTGAATG
	Reverse:TTTCCAGCGTTGTGCAGAAAA
TSPYL2	Forward:ACAGGTGCTGGCCGATATG
	Reverse:CCGACTCGATGGTAGAATCCC
PREB	Forward:ACGGGCCACCATGAACTTG
	Reverse:GGGTTTCCGCTCCACATTTCT
ATP6V0D1	Forward:TTCCCGGAGCTTTACTTTAACG
	Reverse:CAAGTCCTCTAGCGTCTCGC
GAPDH	Forward:GGAGCGAGATCCCTCCAAAAT
	Reverse:GGCTGTTGTCATACTTCTCATGG

Quantitative Real-Time PCR (qRT-PCR)

The TRIzol method was utilized to extract and purify RNA from tissues and cells. Then, a cDNA synthesis kit (Takara, China) was applied to reverse transcribe the RNA. The TB Green Premux Ex TaqII (Tli RNaseH Plus) was used for qRT-PCR with the Bio-Rad CFX96 Real-Time PCR system (Bio-Rad, USA). The internal control was GADPH. The primer sequences of the candidate genes are shown in **Table 1**, and the analysis was conducted three times for all genes.

Statistics

R software (version 4.0.5), SPSS 21.0 software, and GraphPad Prism 8 were carried out for all statistical analyses. The *t*-test was applied for two groups. One-way ANOVA was applied to three groups. $p < 0.05$ was taken as statistically significant. $*p < 0.05$; $**p < 0.01$; and $***p < 0.001$.

RESULTS

Identification of UPRRG Molecular Subtypes via Consensus Analysis

A total of 113 UPRRGs were retrieved from hallmark gene sets, out of which we identified 15 prognosis-related UPRRGs for osteosarcoma, according to univariate COX analysis (**Supplementary Table S2, 3**). According to the expression profile of these genes, consensus clustering analysis was used to ascertain subgroups of osteosarcoma patients in the training database. k = 2 is considered the best category number of clusters, depending on the average consistency value and clinical significance within the clustering group (**Figure 1A** and **Supplementary Figure S1**). PCA analysis revealed a relatively apparent distinction between the two subtypes (**Figure 1B**). The heat map illustrated the gene expression profile and clinical characteristics of the two subtypes (**Figure 1C**). Moreover, we noticed that patients in cluster 2 experienced a dismal prognosis to that in cluster 1 ($p < 0.001$, **Figure 1D**). Previous studies have proven that the UPR can coordinate the crosstalk between immune cells and tumor cells in the TIME to exert immunosurveillance and immunosuppressive functions to influence tumor prognosis (Vanacker et al., 2017). Thus, we then evaluated the differences in the TIME across different subtypes. The ESTIMATE algorithm indicated that cluster 1 had higher immune score (p = 8.4e-5), stromal score (p = 8.1e-7), ESTIMATE score (p = 9.6e-7), and lower

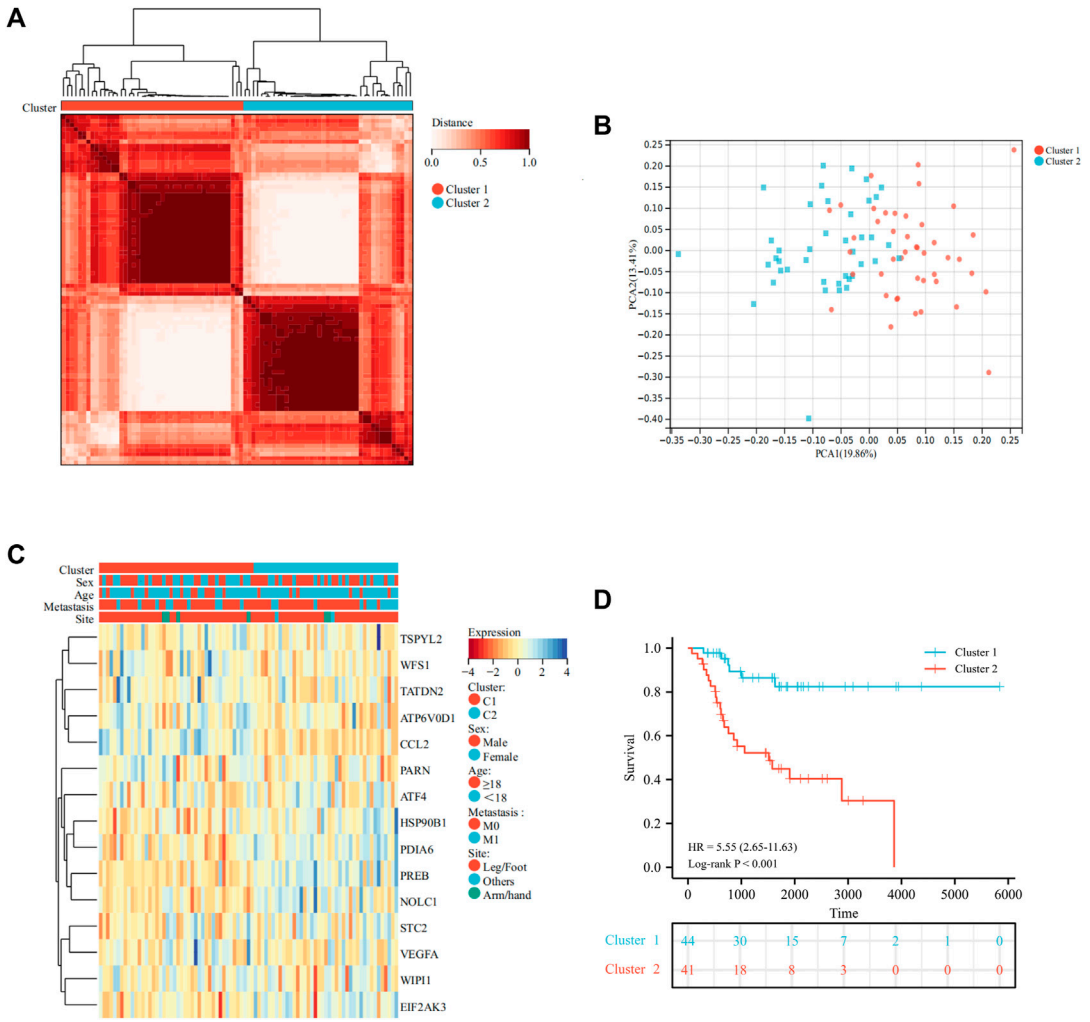


FIGURE 1 | Identification of molecular subtypes of UPRRGs by consensus clustering. **(A).** Clustering heat map at k = 2. **(B)** PCA plot between the two subtypes. **(C)** Heat map of the UPR-related gene expression and clinical features in the two subtypes. **(D)** Survival curves for the two subgroups.

tumor purity (p = 1.1e-6), as compared to cluster 2 (**Figure 2A**). In addition, the TIMER algorithm discovered that the abundance of most immune infiltrating cells in cluster1 was significantly increased than in cluster 2, including dendritic cells, neutrophil cells, $CD4^+$ T cells, neutrophil cells, and $CD8^+$ T cells, while B cells showed the opposite result (**Figure 2B**). As shown in **Figure 2C**, ssGSEA analysis found that the abundance of 20 immune cell types was significantly increased in cluster 1 compared to cluster 2. In addition, we also observed that *CD274*, *LAG3*, *HAVCR2*, and *PDCD1* were expressed at an elevated level in cluster 1 than in cluster 2 (**Figure 2D**). Our findings suggested that the prognosis of different subtypes may be affected by the TIME.

Functions and Pathway Annotations of DEGs for UPRRG Subtypes

To reveal the potential mechanisms regulating the TIME between different subtypes, we performed DEGs analysis on the two clusters. The results indicated that 121 genes were upregulated and 157 genes were downregulated in cluster 2 to cluster 1 (**Figure 3A**). GO analysis implied that these DEGs were primarily associated with inflammatory response, leukocyte activation, and extracellular matrix (**Figure 3B**). These findings implied that UPR could influence the TIME and prognosis of osteosarcoma *via* modulation of immune-associated pathways. We then used GSVA and GSEA analysis to explore the functional differences in the two clusters. GSEA analysis demonstrated that coagulation, inflammatory response, and IL6/JAK/STAT3 signaling were markedly upregulated in cluster 1 (**Figure 3C**). GSVA analysis showed that positive regulation of calcium ion import, regulation of the apoptotic process involved in the development, and some immune-related pathways were significantly upregulated in cluster 1 to cluster 2 (**Figure 3D**). Therefore, we speculated that the UPR plays an essential role in regulating the immune function, thus contributing to the prognosis of osteosarcoma.

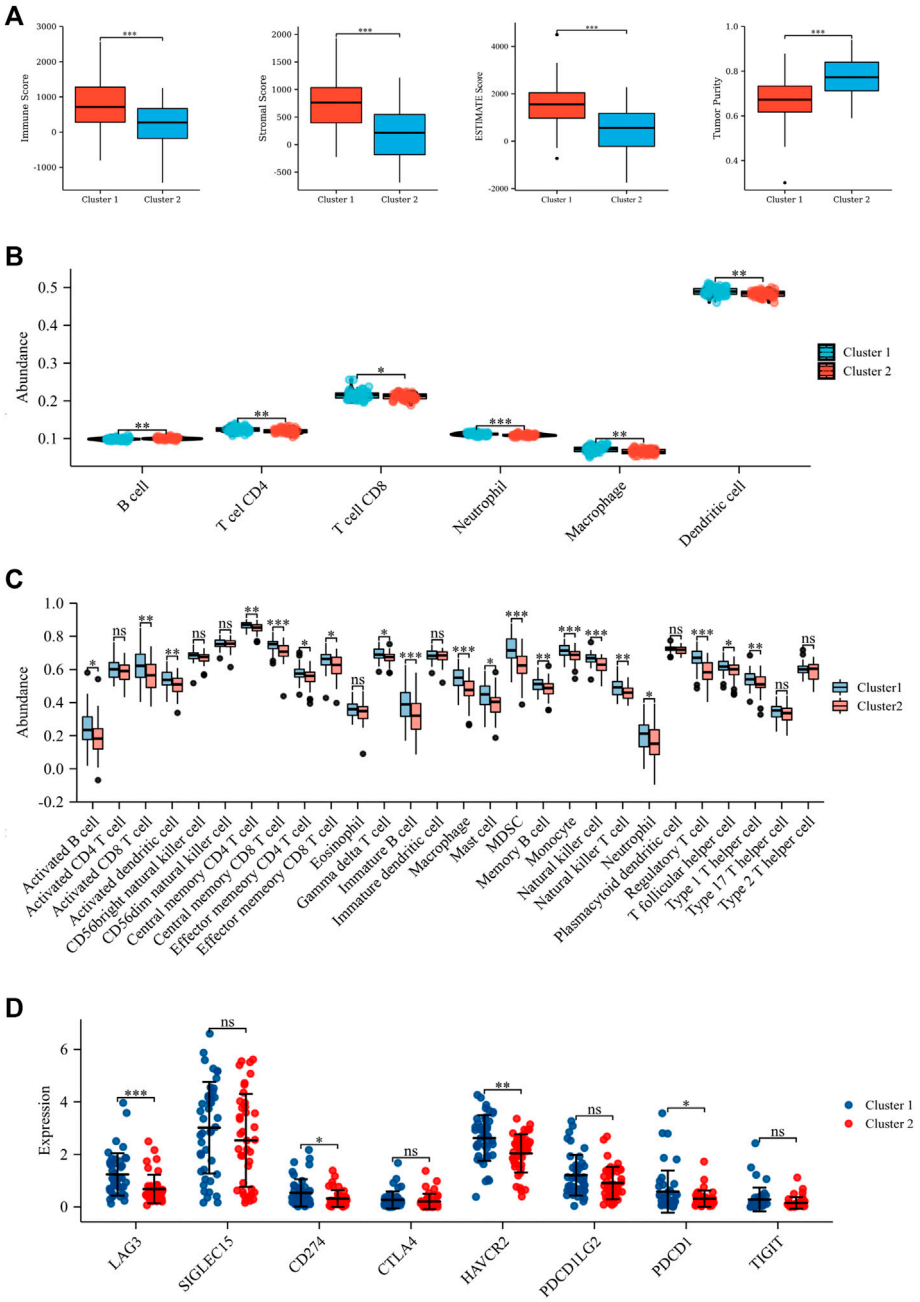


FIGURE 2 | Tumor microenvironment in the two subtypes. **(A)** Stromal score, immune score, ESTIMATE score, and tumor purity based on ESTIMATE algorithm. **(B)** Six immune cell abundance assessments by the TIMER algorithm. **(C)** Twenty-nine immune cell abundance assessments by the ssGSEA algorithm. **(D)** Differences in immune checkpoints between the two subtypes. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Construction and Validation of the UPRRG Risk Signature

To further elucidate the prognostic predictive role of UPRRGs in osteosarcoma patients, LASSO regression analysis was applied to screen for potential genes, and eight genes were identified by the minimal lambo value (**Figures 4A,B**). Finally, based on the results of previous screening, four UPRRGs (*STC2*, *PREB*, *TSPYL2*, and *ATP6V0D1*) were identified to establish a risk model for osteosarcoma *via* the multivariate Cox regression analysis (**Figure 4C** and **Supplementary Table S4**). The following formula was applied to generate the risk score for each sample: Risk score = -1.523 × ATP6V0D1 + 0.903× PREB +0.586 × STC2 -0.760 × TSPYL2. All patients in the training cohort were classified into high- and low-risk groups by the median risk scores, and

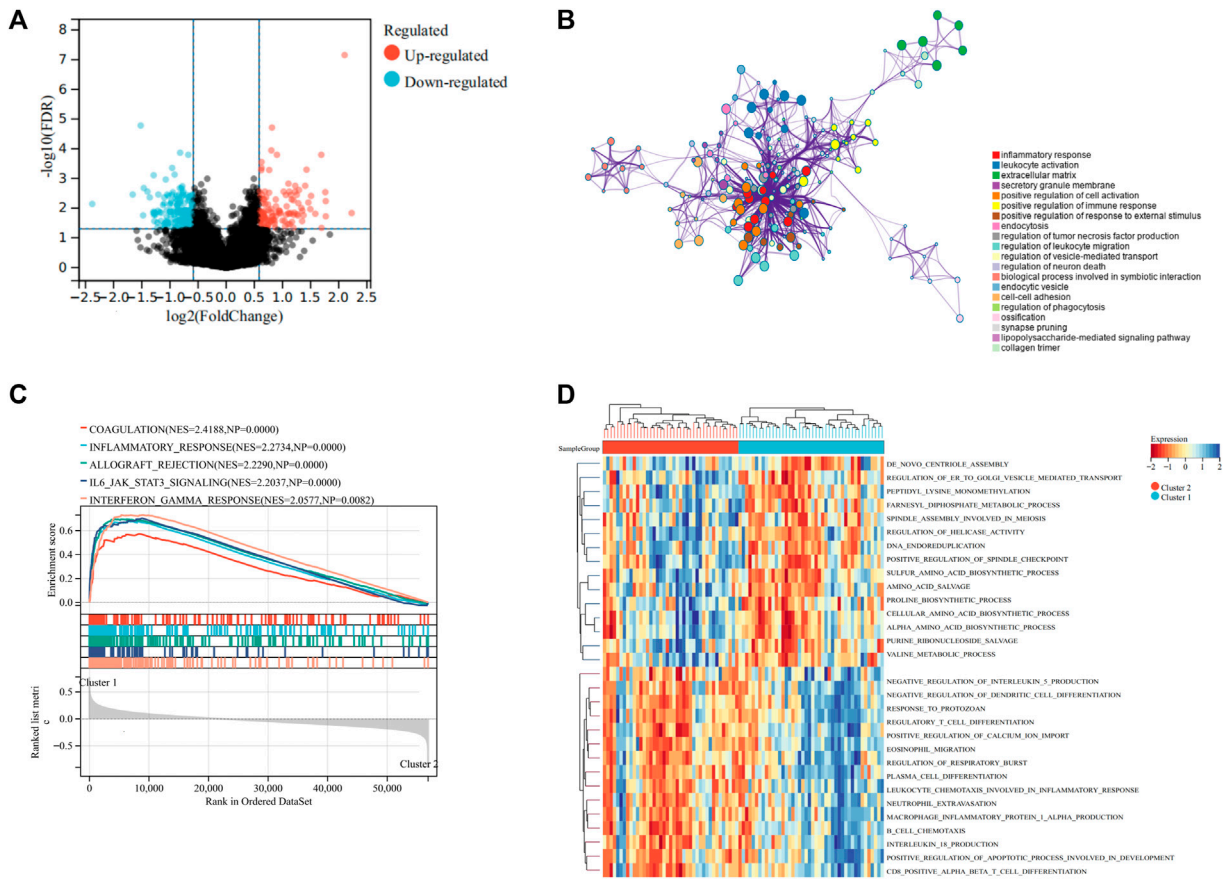


FIGURE 3 | Differentially expressed genes and functional enrichment analyses. **(A)** Volcano plot showing the DEGs between the two subgroups. **(B)** Bubble plot exhibited the functional enrichment of DEGs through GO analysis. **(C)** GSEA shows the hallmark gene sets in the two subgroups. **(D)** Heat map depicted the results of the GSVA analysis.

patients in the high-risk group showed shorter survival intervals than those in the low-risk group (**Figure 4D**). KM curves demonstrated a poor prognosis in the high-risk group to the low-risk group (**Figure 4E**). Moreover, time-dependent ROC analysis found the AUC values of 1, 3, and 5 years were 0.84, 0.87, and 0.83, respectively, which suggested our risk signature showed excellent predictive performance (**Figure 4F**). Notably, we also observed that the low-risk group saw higher immune scores (p = 3.5e-4), stromal scores (p = 1.1e-4), ESTIMATE scores (p = 3.3e-5), and lower tumor purity (p = 2.9e-5) relative to the high-risk group, which suggested that the TIME might be strongly linked to prognosis in different risk groups (**Figure 4G**). Then, the relevance between the risk signature and clinical features was also evaluated, and the results revealed metastatic patients had significantly higher risk scores than non-metastatic patients (p = 0.03), while no differences were found in any other clinical characteristics (**Figure 5A**). When patients were reclassified for metastatic status, there was a significantly improved prognosis for patients in the low-risk group over those in the high-risk group (**Figures 5B,C**). These findings supported that the UPRRG risk signature might be strongly correlated with the metastasis status in patients with osteosarcoma. Furthermore, multivariate Cox analysis demonstrated that the risk score and metastatic status were independent prognostic factors for osteosarcoma patients, which meant that the UPRRGs risk model was applicable to predict survival in osteosarcoma (**Figures 5D,E**).

Moreover, we further tested the prognostic performance of the UPRRG risk signature in a validation cohort. As shown in **Figure 6A**, the patients in the validation cohort were clearly separated into different risk groups *via* the abovementioned formula. The heat map demonstrated these four genes' expression profiles in subgroups (**Figure 6B**). The KM curve likewise showed that the low-risk group had a better prognosis (p = 0.04) (**Figure 6C**). ROC curves suggested the risk signature had better prediction accuracy at 1 and 3 years (**Figure 6D**). Similarly, the ESTIMATE algorithm obtained results consistent with the training cohort, which further confirmed the role of the TIME in the UPRRG risk signature (**Figure 6E**).

Construction and Validation of the Nomogram Prediction Model

A nomogram model was constructed to improve the accuracy of predicting the prognosis of osteosarcoma patients at 3 and 5 years by integrating risk scores and clinical characteristics (**Figure 7A**). Then,

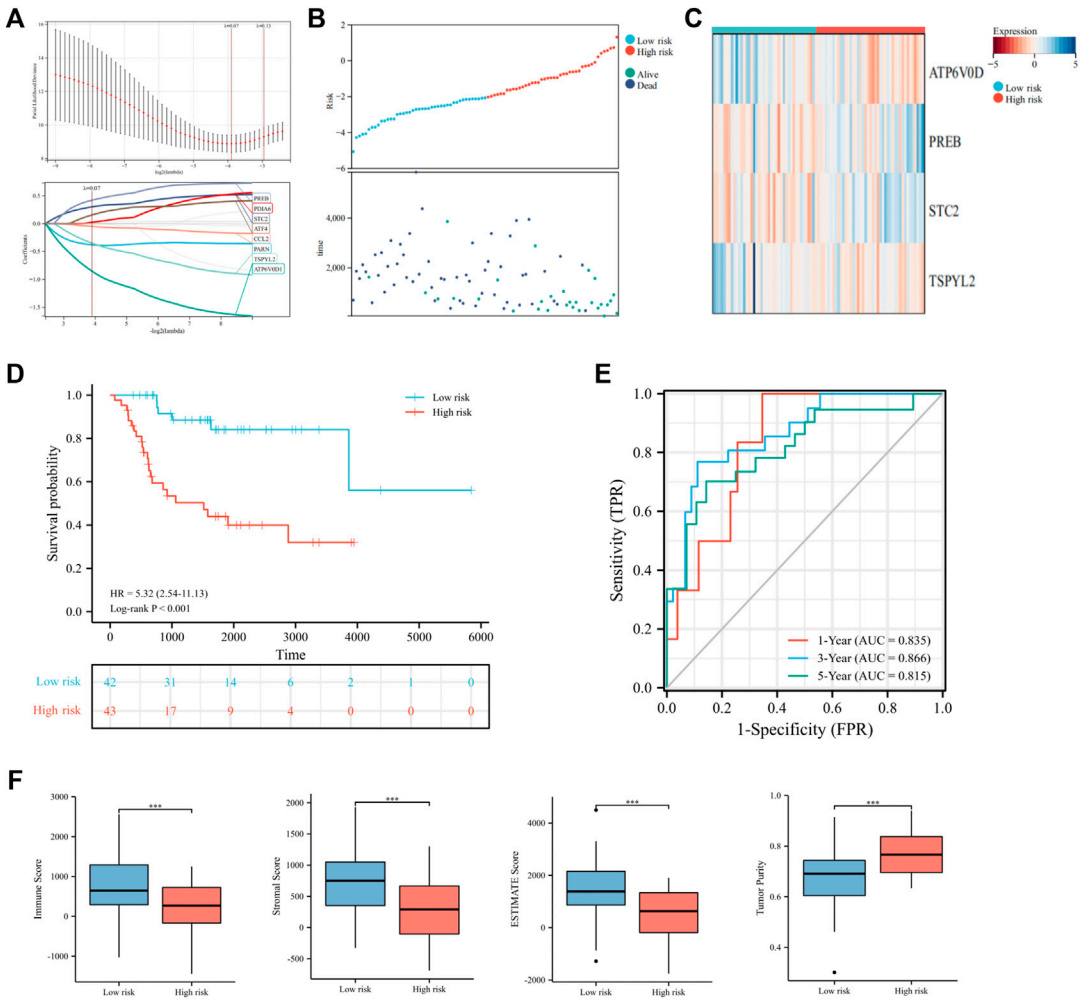


FIGURE 4 | Construction of the risk signature based on UPRRGs in the training cohort. **(A)** Eight optimal UPRRGs filtered by LASSO analysis. **(B)** Distribution of risk scores and patient status in the two risk groups; **(C).** Heat map showing the expressions of four candidate genes. **(D)** Survival curves for the two risk groups. **(E)** Time-dependent ROC curve of the risk model. **(F)** Tumor microenvironment analysis in the two risk groups through the ESTIMATE algorithm. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

we validated the predictive efficacy of the nomogram model in the two cohorts. The C-index values for the training and validation cohorts were 0.88 and 0.87, respectively. The ROC curves revealed AUC values of 0.93 and 0.90 for the training cohort at 3 and 5 years, respectively, and for the validation cohort, ROC curves also exhibited excellent prediction accuracy (**Figures 7B,C**). Moreover, the calibration curves for the training and validation cohort showed that the nomogram model has a strong predictive capacity for the prognosis of osteosarcoma patients at 3 and 5 years (**Figures 7D,E**). Collectively, the aforementioned results pointed out that the UPRRG nomogram model has high predictive accuracy and can be applied to predict the prognosis of osteosarcoma patients.

Verification of Candidate Genes by qRT-PCR Analysis

To certify the expression levels of these four candidate genes, we performed a qPCR analysis in patients' tissue and cell lines. The results revealed that the expression levels STC2 and PREB were elevated clearly in osteosarcoma tissues than in normal tissues, whereas the ATP6V0D1 expression level was significantly downregulated in tumor tissues compared with normal tissues (**Figure 8A**). The results of the cell further identified that the expression levels of PREB and STC2 were higher in osteosarcoma cells than in normal osteoblasts, while TSPYL2 and ATP6V0D1 exhibited the opposite results (**Figure 8B**).

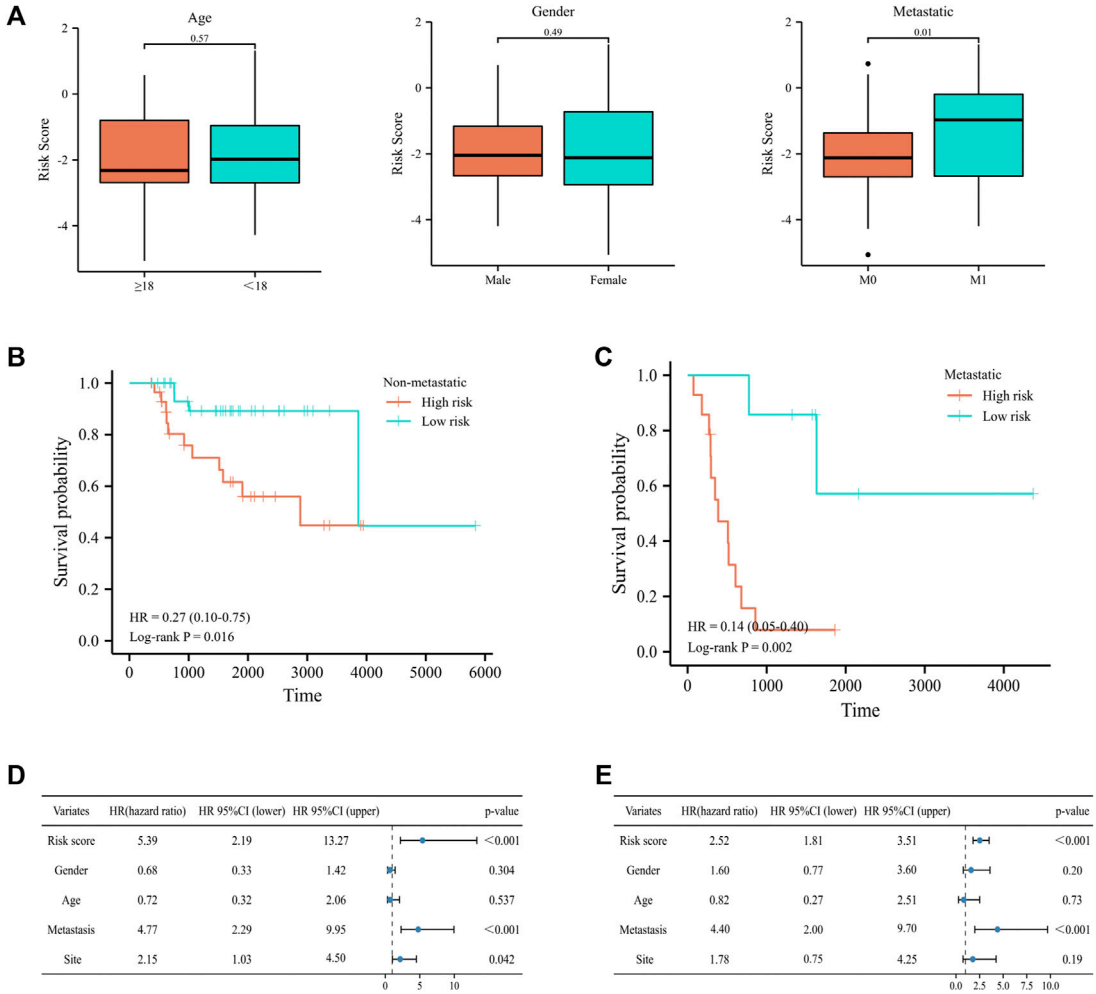


FIGURE 5 | Correlation of the risk signature with clinical features in the training cohort. **(A)** Differences in risk scores among osteosarcoma patients by age, gender, and metastatic status. **(B–C)** Survival curves for patients with osteosarcoma regrouped by metastatic status. **(D–E)** Univariate and multivariate cox regression analyses for integrating risk characteristics and clinical features.

DISCUSSION

The unfolded protein response is an adaptive signaling pathway that regulates the endoplasmic reticulum homeostasis (Hetz et al., 2020). The UPR has been extensively studied in tumorigenesis, and its abnormal activation is involved in various stages of tumorigenesis and progression (Ma and Hendershot, 2004). Chaiyawat et al. revealed that plenty of UPR-related proteins were highly expressed in osteosarcoma patients and closely associated with chemoresistance, suggesting that targeting UPR pathways might be promising for the treatment of osteosarcoma (Chaiyawat et al., 2019). Ji et al. found that PERK could induce autophagy in osteosarcoma by inhibiting the mTORC1 pathway to counteract ER stress-induced cell apoptosis (Ji et al., 2015). Yan et al. demonstrated that the UPR could inhibit cisplatin-mediated apoptosis by triggering the NF-κB pathway, contributing to drug resistance in osteosarcoma (Yan et al., 2015). Shi et al. identified aberrant activation of UPR-related pathways in osteosarcoma and built a prognostic model for differentially expressed genes (Shi et al., 2022). Although previous studies have shown the UPR was implicated in the development, progression, and treatment of osteosarcoma, the function of UPRRGs on immune infiltration and prognostic prediction in patients with osteosarcoma is not clear.

In this research, we determined two molecule subtypes of osteosarcoma based on prognosis-related UPRRGs and presented the relevance of different subtypes to clinical profiles. Our study found that different subtypes had different survival intervals and TIME. The TIME consists of various components, including immune and stromal cells, which together contribute to tumorigenesis, progression, and prognosis (Zhao et al., 2018). The UPR is a critical factor in regulating the balance of immune dynamics in the tumor microenvironment which can directly affect innate and adaptive immune responses to participate in tumor progression (Martins et al., 2016). Our results indicated that cluster 1 has a better prognosis with higher immune and

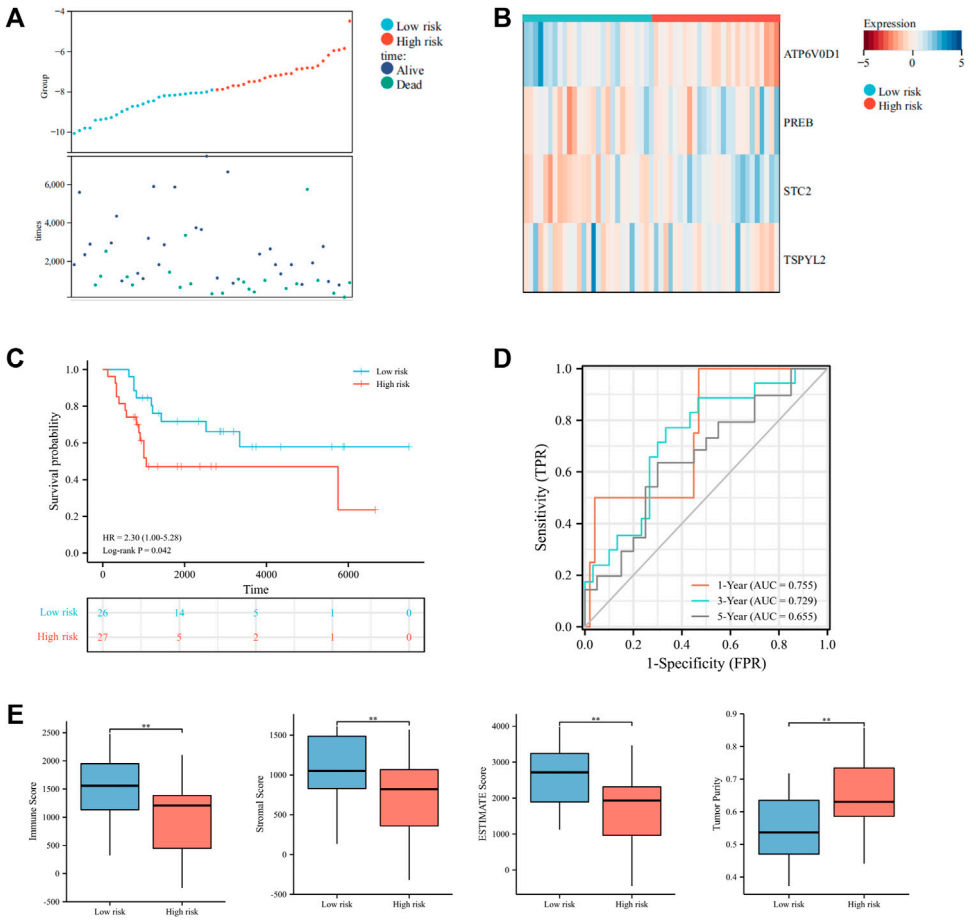


FIGURE 6 | Validation of the constructed risk signature in the verification cohort. **(A)** Distribution of risk scores and patient status in different risk groups. **(B)** Heat map displayed the expressions of four candidate genes in the verification cohort. **(C)** Survival curves of the two risk groups. **(D)** Time-dependent ROC curve in the verification cohort. **(E)** Tumor microenvironment analysis by the ESTIMATE algorithm. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

stromal scores compared to cluster 2. Previous studies have demonstrated that a higher immune score and stromal score were connected to a better prognosis in osteosarcoma, which is in agreement with our findings (Qian et al., 2021). Further analysis indicated that the abundance of multiple immune infiltrating cells was distinctly greater in cluster 1 than in cluster 2, according to the TIMER, indicating that the active immune status of the UPRRG subtypes might be closely better associated with osteosarcoma prognosis. ssGSEA analysis found that the abundance of a variety of tumor-infiltrating lymphocytes was expressed at obviously higher levels in cluster 1 than in cluster 2. Previous studies have demonstrated that the UPR acts as a critical mediator of tumor immunity as an appropriate UPR can induce immune cells to eliminate tumor cells, whereas the sustained UPR could induce immune cell apoptosis to enhance tumor proliferation and invasion (Vanacker et al., 2017). Hence, we believed that the elevated abundance of multiple TILs in cluster 1 might be due to appropriate UPR, which induced activation of TILs to deliver anti-tumor immunity. Nevertheless, the sustained UPR led to massive TIL exhaustion, causing a poor prognosis in patients with osteosarcoma in Cluster 2. Moreover, we also noted that as a consequence of a large number of immune infiltration cells in cluster 1, the expressions of several ICPs were correspondingly elevated in cluster 1 compared to Cluster 2, implying that patients in cluster 1 might be more sensitive to ICP inhibitors. These findings suggested that the identification of UPRRG subtypes may provide a new clinical strategy for prognostic evaluation and targeted therapy of osteosarcoma.

To reveal the molecular mechanisms underlying regulating TIME between different UPRRG subtypes, we carried functional enrichment analysis of the DEGs between the two subtypes. GO analysis suggested that the DEGs were mainly participating in immune-related pathways including inflammatory response and leukocyte activation. In addition, GSEA analysis also confirmed significant enrichment of some immune-related pathways in cluster 1, including coagulation, inflammatory response, and IL6/JAK/STAT3 signaling. The UPR is proven to not only affect the growth and survival of tumor cells but also play an essential role in remodeling the TIME (Zanetti et al., 2022). Batista et al. identified that the UPR participates in the macrophage polarization in the TIME by activating the IRE1α/XBP1 axis, resulting in upregulation of IL-6, IL-23, arginase 1, CD86, and

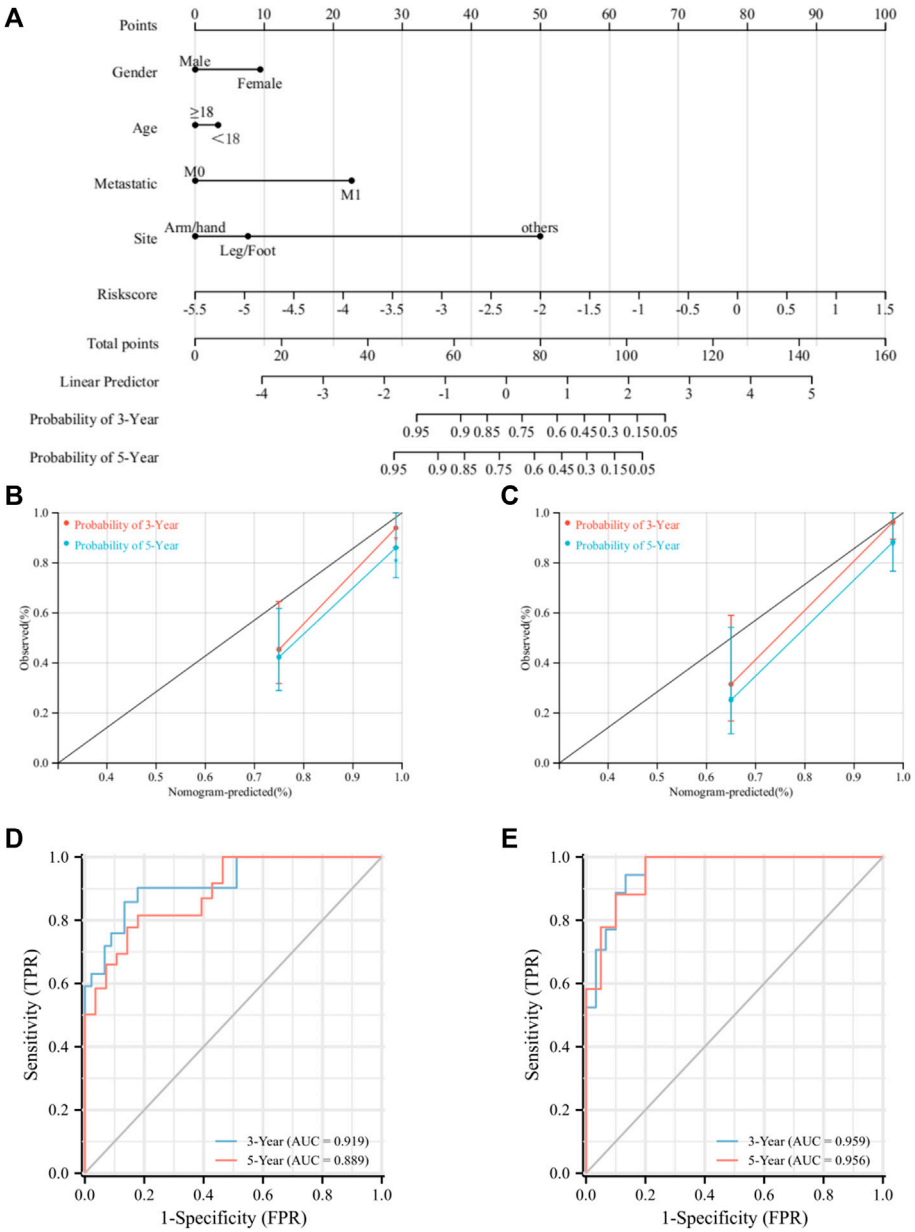


FIGURE 7 | Construction and evaluation of the nomogram. **(A)**. Nomogram for predicting the prognosis of patients with osteosarcoma. **(B)** Calibration for 3-and 5-year OS in the training cohort. **(C)** Calibration for 3-and 5-year OS in the verification cohort. **(D)** ROC analysis for 3-and 5-year OS in the training cohort. **(E)** ROC analysis for 3-and 5-year OS in the verification cohort.

PD-L1 that lead to local immune dysregulation (Batista et al., 2020). Zanetti et al. showed that tumor cells regulate immune cell phenotypes through UPR activation of dendritic cells and T cells to promote tumor proliferation (Mahadevan et al., 2012). In addition, Medel et al. also demonstrated that dendritic cells could enhance $CD8^+$ T-cell–specific responses through activation of the IRE1α/XBP1 axis to exert an anti-tumor effect (Medel et al., 2018). Taken together, these findings revealed that the UPR can modulate immune cell infiltration through various immune signaling pathways to improve the prognosis of patients with osteosarcoma. Moreover, GSVA analysis revealed that aside from immune-related pathways, apoptosis and calcium homeostasis-related pathways were also upregulated in cluster 1. These results suggested that the UPR may exert antitumor effects through multiple pathways between UPRRG subtypes. Overall, our data showed that patients with osteosarcoma can activate the UPR in the presence of ER stress to modulate tumorigenesis and progression through multiple pathways and that targeting UPR might be a promising treatment strategy for osteosarcoma.

Next, to assess the role of UPRRGs in predicting the prognosis of osteosarcoma, we constructed a prognostic signature to predict the survival of osteosarcoma patients *via* four genes (*STC2*, *PREB*,

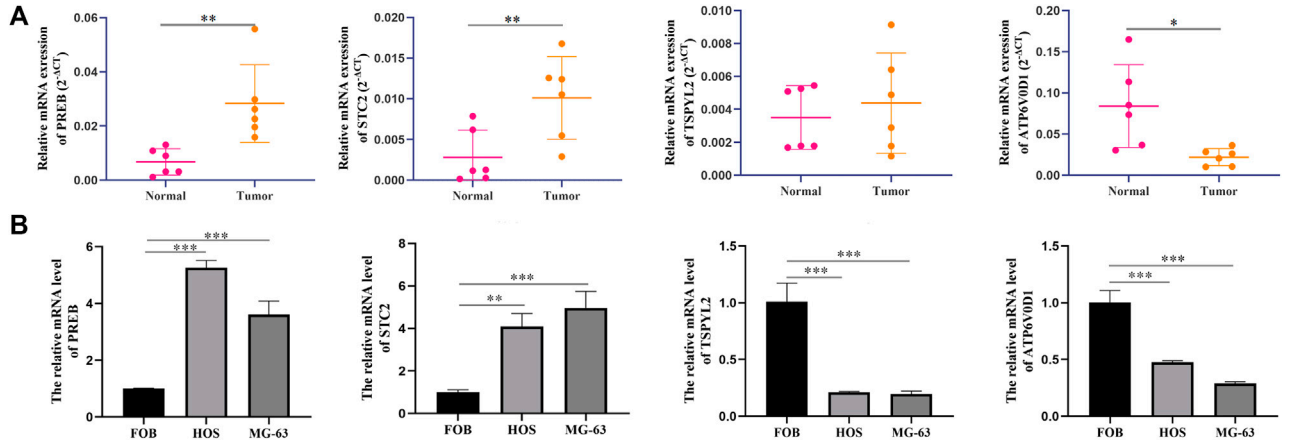


FIGURE 8 | Candidate gene validation. **(A)** Expression levels of candidate genes in tumor and normal tissues from osteosarcoma patients. **(B)** Expression levels of candidate genes in different cell lines. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TSPYL2, and *ATP6V0D1*). Our results found that high expressions of *STC2* and *PREB* were linked to high risk (risk factors), whereas high expressions of *TSPYL2* and *ATP6V0D1* were linked to low risk (protective factors). *STC2* encodes a glycoprotein that performs an essential function in the development and invasion of multiple tumors (Li S. et al., 2021). Previous studies have demonstrated endoplasmic reticulum stress could activate PERK-ATF4 to induce the upregulated expression of *STC2* to inhibit cell apoptosis (Ito et al., 2004). Chen et al. found that *STC2* could promote tumor proliferation by activating the AKT-ERK pathway, and increased STC2 was strongly correlated with poor prognosis in colorectal tumors (Chen et al., 2016). *PREB* can encode transcription factors that bind and activate the basal prolactin promoter activity. Murao et al. showed that *PREB* could act as a transcription factor for TNF-αand IL-1β to regulate the expression of monocyte chemoattractant protein-1, suggesting that *PREB* plays an active role in immune responses (Murao et al., 2009). In addition, *PREB* is a member of a eukaryotic family of WD-repeat proteins involved in many biological activities, including vesicle trafficking, RNA processing, and signal transduction (Taylor Clelland et al., 2000). *TSPYL2*, a member of the TSPY-L nucleosome assembly protein-1 superfamily, can exert anti-tumor effects by inhibiting the cell cycle and regulating DNA damage. Previous studies have found that *TSPYL2* maintains the G1 checkpoint function by inducing p21 transcription to modulate DNA damage and inhibit cellular growth (Tao et al., 2011). Liu et al. indicated that *TSPYL2* could inhibit SIRT1-mediated FOXO3 deacetylation to reduce gefitinib resistance and inhibit DNA damage, implying that *TSPYL2* is a promising therapeutic target (Liu et al., 2022). ATP6V0D1 is an encoded protein involved in vacuolar ATPase formation, which has a crucial role in the modulation of the acidic microenvironment (Lu et al., 2021). Numerous research studies have found that dysregulation of V-ATPase is related to tumor growth and invasion (Whitton et al., 2018). Targeting V-ATPase could upregulate ER stress-related markers and inhibit mTOR signaling to exert anticancer effects (Kitazawa et al., 2017).

Survival analysis and ROC curve presented that the UPRRG risk signature demonstrated satisfactory predictive accuracy in both cohorts and that the low immune status in the high-risk group was strongly connected with poor prognosis. Moreover, our results identified that risk score was an independent prognostic factor for osteosarcoma patients. Subsequently, we performed a nomogram model to better predict the prognosis of osteosarcoma patients *via* risk score and clinical characteristics. Previous studies have reported several nomogram models to predict the prognosis of osteosarcoma (Qian et al., 2021). Li et al. built a nomogram model by autophagy-related genes with AUC values of 0.735 and 0.726 at 3 and 5 years, respectively (Li J. et al., 2021). Wen et al. developed a 3-gene nomogram model with 3-year and 5-year AUC values of 0.853 and 0.818, respectively (Wen et al., 2020). Wu et al. established a hypoxic nomogram model with an AUC value of only 0.73 (Wu et al., 2021). In our study, the constructed nomogram model had 3-year and 5-year AUC values of 0.93 and 0.90, respectively, which were consistently superior to other models. We presented that the UPRRG nomogram model could better predict the prognosis of patients with osteosarcoma than existing models. Collectively, these findings offered new options for personalized treatment and prognostic prediction of osteosarcoma.

Despite the many strengths of the current study, there are notable limitations. First, the UPRRG risk signature was constructed based on the TARGET and GEO databases, which may be biased due to the limited number of patients. Second, this study lacked some clinical information relevant to the prognosis of osteosarcoma, such as tumor pathological grade, which constrained clinical variables that can be incorporated into the nomogram model. A larger, multicenter, prospective clinical cohort will be needed to further evaluate the clinical merit of our findings in the future.

CONCLUSION

Our study identified two molecular subtypes through consensus clustering based on prognosis-related UPRRGs, and the two subtypes exhibited different survival times and immune statuses. Functional analysis revealed that the UPRRG subtypes might influence the progression and prognosis of

osteosarcoma patients through immune-related pathways. Moreover, a novel prognostic model based on UPRRGs was constructed and validated to better predict the prognosis of patients with osteosarcoma. We elucidated the important function of UPRRGs in the development and prognosis of osteosarcoma, shedding light on new insights for targeted therapy and clinical decision-making in patients with osteosarcoma.

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-Jing Hospital, The Fourth Military Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HF and ZZ have designed the study. ZZ and XL have written the manuscript, and DC, JD, ZM, YS, and LW have analyzed the data. All authors have read and approved it for publication.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 31971272).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.911346/full#supplementary-material

REFERENCES

Arndt, C. A. S., Rose, P. S., Folpe, A. L., and Laack, N. N. (2012). Common Musculoskeletal Tumors of Childhood and Adolescence. *Mayo Clin. Proc.* 87 (5), 475–487. doi:10.1016/j.mayocp.2012.01.015

Batista, A., Rodvold, J. J., Xian, S., Searles, S. C., Lew, A., Iwawaki, T., et al. (2020). IRE1α Regulates Macrophage Polarization, PD-L1 Expression, and Tumor Survival. *PLoS Biol.* 18 (6), e3000687. doi:10.1371/journal.pbio.3000687

Chaiyawat, P., Sungngam, P., Teeyakasem, P., Sirikaew, N., Klangjorhor, J., Settakorn, J., et al. (2019). Protein Profiling of Osteosarcoma Tissue and Soft Callus Unveils Activation of the Unfolded Protein Response Pathway. *Int. J. Oncol.* 54 (5), 1704–1718. doi:10.3892/ijo.2019.4737

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

Ferrari, S., Smeland, S., Mercuri, M., Bertoni, F., Longhi, A., Ruggieri, P., et al. (2005). Neoadjuvant Chemotherapy with High-Dose Ifosfamide, High-Dose Methotrexate, Cisplatin, and Doxorubicin for Patients with Localized Osteosarcoma of the Extremity: a Joint Study by the Italian and Scandinavian Sarcoma Groups. *Jco* 23 (34), 8845–8852. doi:10.1200/jco.2004.00.5785

Hao, Y., An, R., Xue, Y., Li, F., Wang, H., Zheng, J., et al. (2021). Prognostic Value of Tumoral and Peritumoral Magnetic Resonance Parameters in Osteosarcoma Patients for Monitoring Chemotherapy Response. *Eur. Radiol.* 31 (5), 3518–3529. doi:10.1007/s00330-020-07338-y

Hetz, C., Zhang, K., and Kaufman, R. J. (2020). Mechanisms, Regulation and Functions of the Unfolded Protein Response. *Nat. Rev. Mol. Cell. Biol.* 21 (8), 421–438. doi:10.1038/s41580-020-0250-z

Houessinon, A., Gicquel, A., Bochereau, F., Louandre, C., Nyga, R., Godin, C., et al. (2016). Alpha-fetoprotein Is a Biomarker of Unfolded Protein Response and Altered Proteostasis in Hepatocellular Carcinoma Cells Exposed to Sorafenib. *Cancer Lett.* 370 (2), 242–249. doi:10.1016/j.canlet.2015.10.032

Ito, D., Walker, J. R., Thompson, C. S., Moroz, I., Lin, W., Veselits, M. L., et al. (2004). Characterization of Stanniocalcin 2, a Novel Target of the Mammalian Unfolded Protein Response with Cytoprotective Properties. *Mol. Cell. Biol.* 24 (21), 9456–9469. doi:10.1128/mcb.24.21.9456-9469.2004

Ji, G.-r., Yu, N.-c., Xue, X., and Li, Z.-g. (2015). PERK-mediated Autophagy in Osteosarcoma Cells Resists ER Stress-Induced Cell Apoptosis. *Int. J. Biol. Sci.* 11 (7), 803–812. doi:10.7150/ijbs.11100

Kitazawa, S., Nishizawa, S., Nakagawa, H., Funata, M., Nishimura, K., Soga, T., et al. (2017). Cancer with Low Cathepsin D Levels Is Susceptible to Vacuolar (H+)-ATPase Inhibition. *Cancer Sci.* 108 (6), 1185–1193. doi:10.1111/cas.13240

Li, J., Guo, Z., Wang, Z., Fan, H., and Fu, J. (2015). Does Microwave Ablation of the Tumor Edge Allow for Joint-Sparing Surgery in Patients with Osteosarcoma of the Proximal Tibia? *Clin. Orthop. Relat. Res.* 473 (10), 3204–3211. doi:10.1007/s11999-015-4447-y

Li, J., Tang, X., Du, Y., Dong, J., Zhao, Z., Hu, H., et al. (2021a). Establishment of an Autophagy-Related Clinical Prognosis Model for Predicting the Overall Survival of Osteosarcoma. *BioMed Res. Int.* 2021, 1–17. doi:10.1155/2021/5428425

Li, S., Huang, Q., Li, D., Lv, L., Li, Y., and Wu, Z. (2021b). The Significance of Stanniocalcin 2 in Malignancies and Mechanisms. *Bioengineered* 12 (1), 7276–7285. doi:10.1080/21655979.2021.1977551

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

Liu, Z., Li, C., Yu, C., Chen, Z., Zhao, C., and Ye, L. (2022). TSPYL2 Reduced Gefitinib Resistance and DNA Damage Repair via Suppressing SIRT1-Mediated FOXO3 Deacetylation. *Future Med. Chem.* 14 (6), 407–419. doi:10.4155/fmc-2021-0136

Lu, J., Ma, J., Hao, Z., and Li, W. (2021). HPS6 Regulates the Biogenesis of Weibel-Palade Body in Endothelial Cells through Trafficking V-ATPase to its Limiting Membrane. *Front. Cell. Dev. Biol.* 9, 743124. doi:10.3389/fcell.2021.743124

Ma, Y., and Hendershot, L. M. (2004). The Role of the Unfolded Protein Response in Tumour Development: Friend or Foe? *Nat. Rev. Cancer* 4 (12), 966–977. doi:10.1038/nrc1505

Mahadevan, N. R., Anufreichik, V., Rodvold, J. J., Chiu, K. T., Sepulveda, H., and Zanetti, M. (2012). Cell-Extrinsic Effects of Tumor ER Stress Imprint Myeloid Dendritic Cells and Impair CD8+ T Cell Priming. *PLoS One* 7 (12), e51845. doi:10.1371/journal.pone.0051845

Martins, A. S., Alves, I., Helguero, L., Domingues, M. R., and Neves, B. M. (2016). The Unfolded Protein Response in Homeostasis and Modulation of Mammalian Immune Cells. *Int. Rev. Immunol.* 35 (6), 457–476. doi:10.3109/08830185.2015.1110151

Medel, B., Costoya, C., Fernandez, D., Pereda, C., Lladser, A., Sauma, D., et al. (2018). IRE1α Activation in Bone Marrow-Derived Dendritic Cells Modulates Innate Recognition of Melanoma Cells and Favors CD8+ T Cell Priming. *Front. Immunol.* 9, 3050. doi:10.3389/fimmu.2018.03050

Murao, K., Imachi, H., Yu, X., Muraoka, T., Hosami, N., Dobashi, H., et al. (2009). The Transcriptional Factor PREB Mediates MCP-1 Transcription Induced by Cytokines in Human Vascular Endothelial Cells. *Atherosclerosis* 207 (1), 45–50. doi:10.1016/j.atherosclerosis.2009.03.051

Qian, H., Lei, T., Hu, Y., and Lei, P. (2021). Expression of Lipid-Metabolism Genes Is Correlated with Immune Microenvironment and Predicts Prognosis in Osteosarcoma. *Front. Cell. Dev. Biol.* 9, 673827. doi:10.3389/fcell.2021.673827

Shi, C., Zhao, F., Zhang, T., Xu, D., Hao, Z., Cui, F., et al. (2022). A Novel Prognostic Signature in Osteosarcoma Characterised from the Perspective of Unfolded Protein Response. *Clin. Transl. Med* 12 (3), e750. doi:10.1002/ctm2.750

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

Tang, N., Song, W.-X., Luo, J., Haydon, R. C., and He, T.-C. (2008). Osteosarcoma Development and Stem Cell Differentiation. *Clin. Orthop. Relat. Res.* 466 (9), 2114–2130. doi:10.1007/s11999-008-0335-z

Tao, K. P., Fong, S. W., Lu, Z., Ching, Y. P., Chan, K. W., and Chan, S. Y. (2011). TSPYL2 Is Important for G1 Checkpoint Maintenance upon DNA Damage. *PLoS One* 6 (6), e21602. doi:10.1371/journal.pone.0021602

Taylor Clelland, C. L., Craciun, L., Bancroft, C., and Lufkin, T. (2000). Mapping and Developmental Expression Analysis of the WD-Repeat Gene Preb. *Genomics* 63 (3), 391–399. doi:10.1006/geno.1999.6089

Vanacker, H., Vetters, J., Moudombi, L., Caux, C., Janssens, S., and Michallet, M.-C. (2017). Emerging Role of the Unfolded Protein Response in Tumor Immunosurveillance. *Trends Cancer* 3 (7), 491–505. doi:10.1016/j.trecan.2017.05.005

Wakamatsu, T., Kakunaga, S., Takenaka, S., Outani, H., Hamada, K., Imura, Y., et al. (2019). Prognostic Implication of Adjuvant/neoadjuvant Chemotherapy Consisting of Doxorubicin and Ifosfamide in Patients with Extraskeletal Osteosarcoma. *Int. J. Clin. Oncol.* 24 (10), 1311–1319. doi:10.1007/s10147-019-01475-1

Wan, Q., Jin, L., and Wang, Z. (2020). Comprehensive Analysis of Cancer Hallmarks in Cutaneous Melanoma and Identification of a Novel Unfolded Protein Response as a Prognostic Signature. *aging* 12 (20), 20684–20701. doi:10.18632/aging.103974

Wen, C., Wang, H., Wang, H., Mo, H., Zhong, W., Tang, J., et al. (2020). A Three-Gene Signature Based on Tumour Microenvironment Predicts Overall Survival of Osteosarcoma in Adolescents and Young Adults. *Aging* 13 (1), 619–645. doi:10.18632/aging.202170

Whitton, B., Okamoto, H., Packham, G., and Crabb, S. J. (2018). Vacuolar ATPase as a Potential Therapeutic Target and Mediator of Treatment Resistance in Cancer. *Cancer Med.* 7 (8), 3800–3811. doi:10.1002/cam4.1594

Wu, F., Xu, J., Jin, M., Jiang, X., Li, J., Li, X., et al. (2021). Development and Verification of a Hypoxic Gene Signature for Predicting Prognosis, Immune Microenvironment, and Chemosensitivity for Osteosarcoma. *Front. Mol. Biosci.* 8, 705148. doi:10.3389/fmolb.2021.705148

Xiao, X., Wang, W., Zhang, H., Gao, P., Fan, B., Huang, C., et al. (2015). Individualized Chemotherapy for Osteosarcoma and Identification of Gene Mutations in Osteosarcoma. *Tumor Biol.* 36 (4), 2427–2435. doi:10.1007/s13277-014-2853-5

Yan, M., Ni, J., Song, D., Ding, M., and Huang, J. (2015). Activation of Unfolded Protein Response Protects Osteosarcoma Cells from Cisplatin-Induced Apoptosis through NF-Kb Pathway. *Int. J. Clin. Exp. Pathol.* 8 (9), 10204–10215.

Yang, H., Liang, S.-Q., Xu, D., Yang, Z., Marti, T. M., Gao, Y., et al. (2019). HSP90/AXL/eIF4E-regulated Unfolded Protein Response as an Acquired Vulnerability in Drug-Resistant KRAS-Mutant Lung Cancer. *Oncogenesis* 8 (9), 45. doi:10.1038/s41389-019-0158-7

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

Zanetti, M., Xian, S., Dosset, M., and Carter, H. (2022). The Unfolded Protein Response at the Tumor-Immune Interface. *Front. Immunol.* 13, 823157. doi:10.3389/fimmu.2022.823157

Zhang, F., Feng, D., Wang, X., Gu, Y., Shen, Z., Yang, Y., et al. (2021). An Unfolded Protein Response Related Signature Could Robustly Predict Survival Outcomes and Closely Correlate with Response to Immunotherapy and Chemotherapy in Bladder Cancer. *Front. Mol. Biosci.* 8, 780329. doi:10.3389/fmolb.2021.780329

Zhao, Y., Shao, Q., Zhu, H., Xu, H., Long, W., Yu, B., et al. (2018). Resveratrol Ameliorates Lewis Lung Carcinoma-bearing Mice Development, Decreases Granulocytic Myeloid-derived Suppressor Cell Accumulation and Impairs its Suppressive Ability. *Cancer Sci.* 109 (9), 2677–2686. doi:10.1111/cas.13720

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 (1), 1523. doi:10.1038/s41467-019-09234-6

Zhu, K., Xiaoqiang, L., Deng, W., Wang, G., and Fu, B. (2021). Identification of a Novel Signature Based on Unfolded Protein Response-Related Gene for Predicting Prognosis in Bladder Cancer. *Hum. Genomics* 15 (1), 73. doi:10.1186/s40246-021-00372-x

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 Zhang, Liu, Cheng, Dang, Mi, Shi, Wang and Fan. 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: 08 June 2022
doi: 10.3389/fgene.2022.892177

The Crosstalk Between Immune Infiltration, Circulating Tumor Cells, and Metastasis in Pancreatic Cancer: Identification of HMGB3 From a Multiple Omics Analysis

*Hao-dong Tang[1], Yang Wang[2], Peng Xie[1], Si-yuan Tan[1], Hai-feng Li[1], Hao Shen[1], Zheng Zhang[1], Zheng-qing Lei[2] and Jia-hua Zhou[1,2]**

[1]Department of Surgery, School of Medicine, Southeast University, Nanjing, China, [2]Department of Hepato-Pancreatico-Biliary Surgery, Zhongda Hospital Southeast University, Nanjing, China

OPEN ACCESS

Edited by:
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Changzheng Du,
Southern University of Science and Technology, China
Fuhai Li,
Washington University in St. Louis, United States

***Correspondence:**
Jia-hua Zhou
zhoujh@seu.edu.cn

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

Received: *08 March 2022*
Accepted: *19 April 2022*
Published: *08 June 2022*

Citation:
Tang H-d, Wang Y, Xie P, Tan S-y, Li H-f, Shen H, Zhang Z, Lei Z-q and Zhou J-h (2022) The Crosstalk Between Immune Infiltration, Circulating Tumor Cells, and Metastasis in Pancreatic Cancer: Identification of HMGB3 From a Multiple Omics Analysis. Front. Genet. 13:892177. doi: 10.3389/fgene.2022.892177

Metastasis is the major cause of death in patients with pancreatic ductal adenocarcinoma (PDAC), and circulating tumor cells (CTCs) play an important role in the development of metastasis. However, few studies have uncovered the metastasis mechanism of PDAC based on CTCs. In this study, the existing bulk RNA-sequencing (bulk RNA-seq) and single-cell sequencing (scRNA-seq) data for CTCs in pancreatic cancer were obtained from the Gene Expression Omnibus (GEO) database. Analysis of tumor-infiltrating immune cells (TIICs) by CIBERSORT showed that the CTCs enriched from the peripheral blood of metastatic PDAC were found to contain a high proportion of T cell regulators (Tregs) and macrophages, while the proportion of dendritic cells (DCs) was lower than that enriched from localized PDAC. Through weighted gene co-expression network analysis (WGCNA) and the result of scRNA-seq, we identified the hub module (265 genes) and 87 marker genes, respectively, which were highly associated with metastasis. The results of functional enrichment analysis indicated that the two gene sets mentioned above are mainly involved in cell adhesion and cytoskeleton and epithelial–mesenchymal transition (EMT). Finally, we found that HMGB3 was the hub gene according to the Venn diagram. The expression of HMGB3 in PDAC was significantly higher than that in normal tissues (protein and mRNA levels). HMGB3 expression was significantly positively correlated with both EMT-related molecules and CTC cluster–related markers. Furthermore, it was also found that HMGB3 mutations were favorably related to tumor-associated immune cells through the TIMER2.0 online tool. We further demonstrated that PDAC patients with higher HMGB3 expression had significantly worse overall survival (OS) in multiple datasets. In summary, our study suggests that HMGB3 is a hub gene associated with EMT in CTCs, the formation of CTC clusters, and infiltration patterns of immune cells favorable for tumor progression and metastasis to distant organs.

Keywords: circulating tumor cells (CTCs), pancreatic ductal adenocarcinoma (PDAC), CIBERSORT, weighted gene co-expression network analysis (WGCNA), single-cell RNA-sequencing (scRNA-seq)

INTRODUCTION

Pancreatic cancer is an aggressive and fatal disease. According to the American Cancer Society (ACS), there are approximately 60,430 new cases of pancreatic cancer and approximately 48,220 deaths in the United States in 2021 (Siegel et al., 2021). Pancreatic cancer could become the second leading cause of cancer death by 2030 (Rahib et al., 2014). Pancreatic ductal adenocarcinoma (PDAC) is the most common pathological type, accounting for more than 90% of pancreatic cancer cases (Hidalgo and Von Hoff, 2012). At present, surgical resection is still the only method to cure PDAC. However, due to its insidious onset and lack of biological markers for early diagnosis, most patients already acquire distant metastasis by the time the diagnosis is confirmed, resulting in less than 9% of the 5-year survival rate (Chen et al., 2016; Siegel et al., 2020). Although localized PDAC patients undergo radical resection, many patients still suffer from early recurrence or even metastasis, contributing to more than 90% of cancer deaths worldwide. Metastasis of tumors has been widely concerning. In 1869, John Ashworth discovered that there were tumor cells in the blood of patients with metastatic cancer histologically similar to the primary tumor and proposed the concept of circulating tumor cell (CTC) (Ashworth, 1869). CTC is a complete tumor cell, which carries multiple omics information (genome, transcriptome, proteome, metabolome, etc.) and is considered a precursor cell of metastasis. It is detached from the primary tumor and enters peripheral blood circulation to form distant lesions (Keller and Pantel, 2019). A plethora of studies have shown that increased CTC number indicates malignant clinicopathological features of the tumor (Liao et al., 2014; Wang et al., 2017; Miyamoto et al., 2018). The detection rate of CTC was higher in patients with metastatic tumor than in those with primary tumor (Goodman et al., 2018). The number of CTCs could be used to evaluate the therapeutic effect and prognosis for various kinds of cancer (Goldkorn et al., 2014; Huang et al., 2015; Cui et al., 2020). A meta-analysis conducted by our research center also found the value of CTC in predicting the prognosis of patients with pancreatic cancer (Wang et al., 2020). However, with the development of RNA-sequencing technology, it is gradually found that it is insufficient to only focus on the number of CTCs. An increasing number of studies have found that CTC has various specific phenotypes (circulating cancer stem cells, CTC with epithelial–mesenchymal transition and CTC clusters, etc.,), which might be the key factor affecting tumor metastasis and prognosis (Barriere et al., 2014; Fabisiewicz and Grzybowska, 2017; Agnoletto et al., 2019).

In addition, the tumor microenvironment has been increasingly significant in immunotherapy for cancer in recent years, and multiple studies have found that the proportion and functional changes of tumor-infiltrating immune cells (TIICs) are related to the initiation, progression, and prognosis of PDAC (Chen et al., 2020; Chen et al., 2021). CTCs are released from the primary site and are more vulnerable to attack by immune cells as they circulate in the peripheral blood. However, previous research has revealed that CTCs induce a local microenvironment, which is conducive to CTCs evading immune surveillance to survive (Fabisiewicz and Grzybowska, 2017). Therefore, CTCs that could eventually survive and lead to metastasis might have a special microenvironment of immune infiltration and gene expression profile.

Unfortunately, few studies have been conducted to study immune cell infiltration, tumor metastasis, and proliferation based on CTCs. The aim of the present study is to investigate patterns of infiltrating immune cells and key gene and pathways associated with metastasis in pancreatic CTCs using CIBERSORT analysis, weighted gene co-expression network analysis (WGCNA), principal component analysis (PCA), and T-stochastic neighbor embedding (TSNE) analysis. Taking it as a new target to interfere with the metastasis of cancer could provide a new solution for cancer treatment.

MATERIALS AND METHODS

Data Sources

The gene sequencing data related to PDAC metastasis were retrieved from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Finally, GSE144561 (Franses et al., 2020) and GSE114704 (Dimitrov-Markov et al., 2020) were selected for subsequent analysis. GSE144561 included RNA-sequencing datasets of healthy donors (21 normal samples) and PDAC patients (42 localized PDAC samples and 18 metastatic PDAC samples) of CTC samples processed by using microfluidic CTC-iChip. All samples were generated on the Illumina NextSeq 500 (Homo sapiens) on the GPL18573 platform.

Another scRNA-seq dataset GSE114704 included 10 CTCs, 23 liver metastasis (LM) cells, and 37 primary tumor (PT) cells derived from metastatic PDAC patient–derived xenograft (PDX) mouse models. The enrichment and separation of CTCs were also conducted by using microfluidic CTC-iChip. All samples were based on the Illumina HiSeq 2500 (Homo sapiens) on the GPL16791 platform. Two sets of gene sequencing data can be accessed for free, and detailed information on both datasets is available in the Supplementary File.

Analysis of Infiltrating Immune Cell Patterns

CIBERSORT is an analytical tool developed by Newman et al. to quantify infiltrating immune cell portions based on normalized RNA gene expression profiles (Newman et al., 2015). We used "CIBERSORT" (R package) to analyze the 22 immune cell compositions of PDAC patients in the dataset GSE14456 (42 local PDAC and 18 metastatic PDAC). To improve the accuracy of the algorithm, we set 1000 aligned default signature matrices. The relative proportions of 22 infiltrating immune cells and CIBERSORT p-value were evaluated for each sample. Furthermore, samples with a CIBERSORT $p < 0.05$ were selected for subsequent analysis.

Acquisition of Differentially Expressed Genes

The genes with mean read counts over 1 in all samples were included in the following analysis. Then, DEGs were screened from read counts of 21,253 genes between healthy donors and

local PDAC patients (GSE14456) using the "DESeq2" R package (Love et al., 2014), with adjusted $p < 0.05$ & |log2FC|>2. The read counts of 23,049 genes between healthy donors and metastatic PDAC patients were also processed according to the same criteria. Finally, two DEG sets were merged to represent the DEGs between healthy donors and PDAC patients.

Weighted Gene Co-Expression Network Analysis and Identification of Modules Associated With Metastasis

We used the R package "WGCNA" to build the coexpression network of genes in localized PDAC and metastatic PDAC patients (Langfelder and Horvath, 2008). First, a scale-free network was constructed. Then, the function pick-soft threshold was used to select the appropriate soft threshold β. Subsequently, the soft threshold β was applied for the transformation of Pearson's correlation matrices to a weighted adjacency matrix. The adjacency matrix was transformed into a topological overlap matrix (TOM) to estimate network connectivity. In the next part, we used the dynamic tree-cut algorithm to cluster dendrogram branches into several modules and assigned them colors. The minimum module size was set at 20, with the module detection sensitivity deepSplit 2, and the modules with larger than 0.85 pairwise correlation were merged. Finally, Pearson's correlations between module eigengenes (MEs) and clinical trait were calculated to identify the module that was relevant to metastasis. Only the module with the highest significant positivity of correlation with metastasis was included in the subsequent study.

Principal Component Analysis and T-Stochastic Neighbor Embedding Analysis

Before analysis, "Seurat", a specific R toolkit, was used for quality control (QC) of single-cell transcriptomic data (GSE114704) (Stuart et al., 2019). We used a similar pipeline for following data analysis according to the previous study (Wu et al., 2020). To remove low-quality cells, we adopted strict filtering standards for each individual (for example, min. cells = 3; min. features = 50; nFeature_RNA > 50; percent. mt < 5%). Subsequently, the remaining transcriptomic data of each individual were log-normalized with a multiplied scale factor of 10,000 to perform a global-scaling normalization, and the "vst" method was then used to select the top 5,000 highly variable genes based on mean variance for downstream PCA analysis. Twenty principal components (PCs) were obtained by PCA analysis. Then, cluster analysis was performed on the cells in these 20 PCs to determine which PCs have the most significant difference. Then, these PCs with significant difference were subjected to TSNE analysis to define the different clusters. The genes among distinct clusters were then differentially analyzed for identification of marker genes (adjusted $p < 0.05$ and |log2FC|>0.5).

Functional Enrichment Analysis

The Metascape tool (https://metascape.org) (Zhou et al., 2019) was used to perform Gene Ontology (GO) biological processes (BP), cell components (CC), molecular functions (MF), Hallmark Gene Sets, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of module genes obtained by WGCNA analysis and marker genes identified by distinct clusters. In screening for enrichment information, $p < 0.05$ was considered statistically significant.

Definition of the Hub Gene

The genes coexpressed in module genes and marker genes were filtrated as hub genes correlated with metastasis through the Wayne diagram tool (http://bioinformatics.psb.ugent.be/webtools/Venn/). To confirm the importance of hub genes, we evaluated their expression levels in pancreatic cancer and normal tissues by The Gene Expression Profiling Interactive Analysis 2 (GEPIA2) (http://gepia.cancer-pku.cn/index.html). The protein expression levels of the hub gene in the pancreatic tumor and nontumor tissues were determined using the Human Protein Atlas (HPA) (https://www.proteinatlas.org/) (Asplund et al., 2012). We investigated whether the expression levels of hub genes correlated with clinical prognosis (overall survival, OS) in the pancreatic adenocarcinoma (PAAD) cohort by long-term outcome and gene expression profiling database of pan-cancers (LOGpc) (http://bioinfo.henu.edu.cn/DatabaseList.jsp) (Zhang etal., 2020). Furthermore, we analyzed the correlation between the key gene and EMT-related molecules and CTC cluster–related markers in the GEO and GEPIA2 databases, respectively. The correlation analysis between the key gene and immune cells was carried out through the TIMER2.0 database, which is a comprehensive and freely accessible data source (Li et al., 2020).

Statistical Analysis

Correlations of gene expression were established using Spearman's correlation coefficient. Associations between categorical variables were tested using the Wilcoxon test. All statistical analyses were performed using R software (version 4.1.0). When $p < 0.05$, the difference was considered statistically significant.

RESULTS

Immune Microenvironment of CTCs in PDAC Patients

The workflow of all analyses involved in this study is depicted in **Figure 1**. First, we explored proportions of the 22 TIICs in CTCs from 42 localized PDAC patients and 18 metastatic PDAC patients (GSE14456) using the CIBERSORT algorithm. Under screening criteria of CIBERSORT $p < 0.05$, only 37 samples were qualified in this study, and the landscape of immune infiltrations in CTCs is summarized in **Figure 2A**. Subsequently, we investigated the difference of TIICs in CTCs with localized PDAC and metastatic PDAC. Apparently, metastatic PDAC had more Tregs and M0 macrophages than localized PDAC ($p = 0001$; $p = 0001$). On the contrary, DCs were found to be significantly lower ($p = 0.019$). In addition, compared with localized PDAC, the proportions of $CD8^+$ T cells and NK cells were higher in metastatic PDAC, whereas infiltrations of plasma

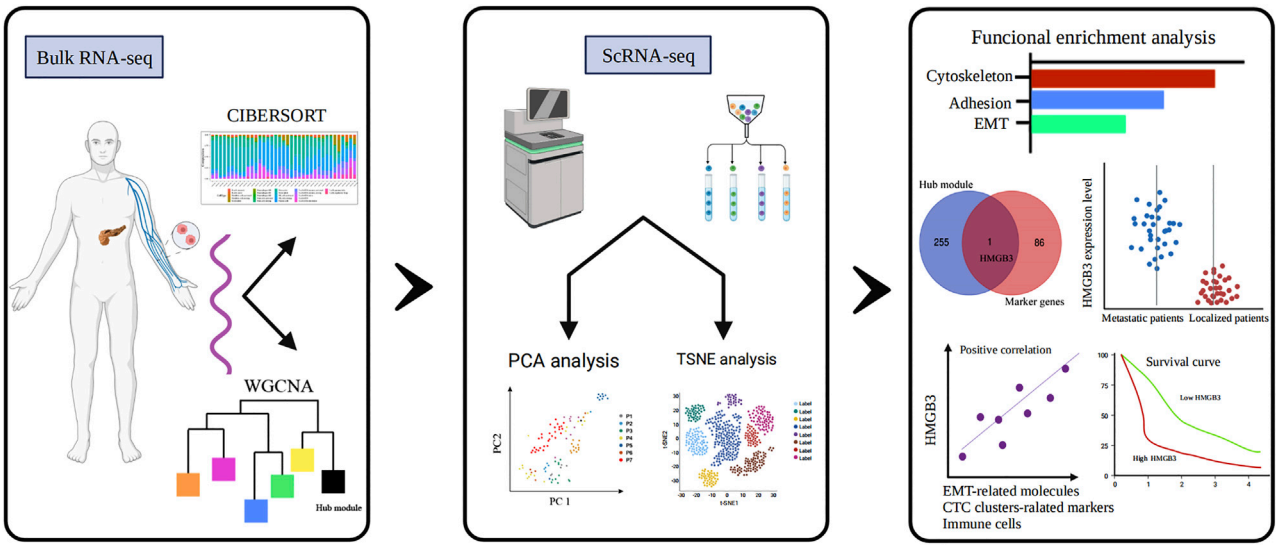


FIGURE 1 | Study flow diagram. WGCNA, weighted gene co-expression network analysis; PCA, principal component analysis; TSNE, T-stochastic neighbor embedding; HMGB3, high-mobility group box 3.

cells were lower. Unfortunately, there were no statistically significant differences (p = 0.062, p = 0.062, and p = 0.086) (**Figure 2B**). The high proportion of Tregs and macrophages infiltrating CTCs might be responsible for protecting CTCs from immune system attack in circulating blood and ultimately leading to metastasis.

Identification of DEGs Between Healthy Donors and PDAC Patients

The discrepancies in immune cell infiltration patterns of CTCs between localized PDAC and metastatic PDAC have been indicated, and then we further explored differences in RNA transcriptome expression between the two groups. First, with adjusted $p < 0.05$ and $|\log2FC| > 2$ as the screening conditions, 1699 DEGs were screened in localized PDAC patients (GSE14456, N = 42) compared with healthy donors (GSE14456, N = 21) (**Figure 3A**). Moreover, in metastatic PDAC patients (GSE14456, N = 18), 6,135 DEGs were identified (**Figure 3B**). Two DGE sets that were previously mentioned were merged, and then duplicates were removed to retain the unique genes (N = 7,452) for subsequent WGCNA analysis.

WGCNA and Identification of Modules Associated With Metastasis

First, the "Hclust" function was used for sample clustering analysis of PDAC (localized PDAC 42 and metastatic PDAC 18), and 13 outliers were removed according to the threshold value (cut height = 10,000) (**Supplementary Figure S1**). Then, the soft threshold β was determined by the "sft$powerEstimate" function and set to 26 (**Supplementary Figure S2**). After TOM network construction, the modules with larger than 0.85 pairwise correlation were merged (cut height as 0.15, **Supplementary Figure S3**) and 10 gene co-expression modules were detected and assigned colors with sizes between 43 and 3,745 genes (**Figure 4A**). The gray module is the set of genes that cannot be aggregated together. Then, Pearson's correlation between different modules and clinical trait (metastasis) was calculated, and the black module has the highest significant positive correlation with metastasis (Cor = 0.58, $p < 0.001$) (**Figure 4B**). There were 256 genes in the black module.

PCA and TSNE Analysis

Single-cell sequencing datasets of 10 CTCs, 23 LM cells, and 37 PT cells derived from metastatic PDAC patient–derived xenograft (PDX) mouse models were obtained from GSE114704. Quality control of read counts was implemented by "Seurat" (**Figures 5A–C**). As the expression of the mitochondrial gene is generally low, cells with nFeature > 50 were selected for subsequent analysis. Thereafter, a scatter diagram was plotted to find genes with larger expression alteration in different cells (**Figure 5D**), and the top 5,000 genes were selected for future analysis. Next, PCA was conducted and 20 principal components (PCs) were obtained. As revealed in the cell distribution in the top two PCs (PC1 and PC2) in **Figure 5E**, all cell populations could be arranged into two distinct clusters. One of the clusters largely comprised PT cells (in the blue circle), while the other mainly comprised CTCs and LM cells (in the red circle). In addition, cluster analysis was performed on the cells in these 20 PCs, and eventually the first three PCs with the most significant difference were identified (**Figure 5F**). These three PCs were adopted for TSNE analysis. Similarly, all 70 samples were grouped into two distinct clusters (clusters 0, 1) (**Figure 5G**). While cluster 0 mainly comprised PT cells (35/39), cluster 1 primarily contained CTCs and LM cells (28/30). The genetic

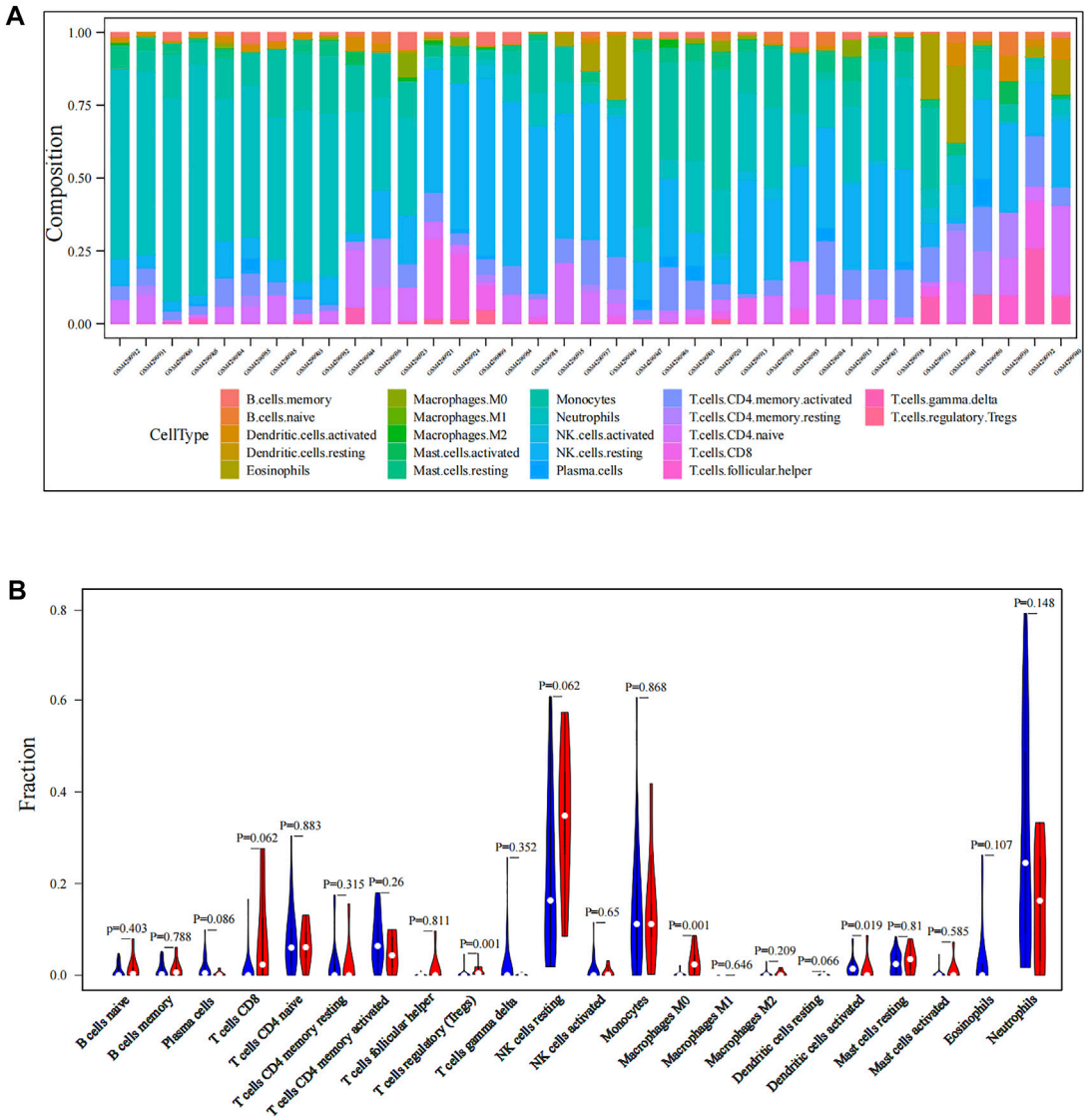


FIGURE 2 | Immune infiltration levels of CTCs in PDAC patients. **(A)** Landscape of immune infiltration in CTCs of PDAC patients. **(B)** Analysis of differential immune cells between localized PDAC and metastatic PDAC. CTCs, circulating tumor cells; PDAC, pancreatic ductal adenocarcinoma.

information carried by CTC from metastatic patients is highly similar to the information of metastatic tumor cells, which was consistent with the abovementioned dimensional reduction results (CTCs and metastatic tumor cells clustered together) (Chemi et al., 2019). Therefore, genes between cluster 0 and cluster 1 were differentially analyzed to find marker genes specifically expressed by CTC of metastatic PDAC (**Figure 6**). Finally, 87 marker genes were identified (|log2FC|>0.5, adjPval<0.05) (**Supplementary table S1**).

Functional Enrichment Analysis and Identification of the Hub Gene

The functional enrichment analysis of hub genes and marker genes was performed by Metascape. BP, CC, MF, Hallmark Gene Sets, and KEGG pathway enrichment analyses were performed separately. Black module genes were functionally enriched in epithelial–mesenchymal transition (Hallmark: M5930, p = 0.035), EGFR tyrosine kinase inhibitor resistance (KEGG: hsa01521, p = 0.013), and phospholipase D signaling pathway (KEGG: hsa04072, p = 0.013), which are crucial pathways for metastasis of PDAC. In addition, cell adhesion (GO:0098632; GO:0007156; GO:0005912; $p < 0.05$), protein kinase activity (GO:0004672, p = 0.024), and cytoskeleton (GO:0015629, p = 0.004; GO:0045104, p = 0.008; GO:0032970, p = 0.003) were all significantly enriched again (**Figure 7A**) While 87 marker genes were functionally enriched in epithelial–mesenchymal transition (Hallmark: M5930, p = 0.035), the P53 pathway (Hallmark: M5939, p = 0.020) and proteoglycans in cancer (KEGG:

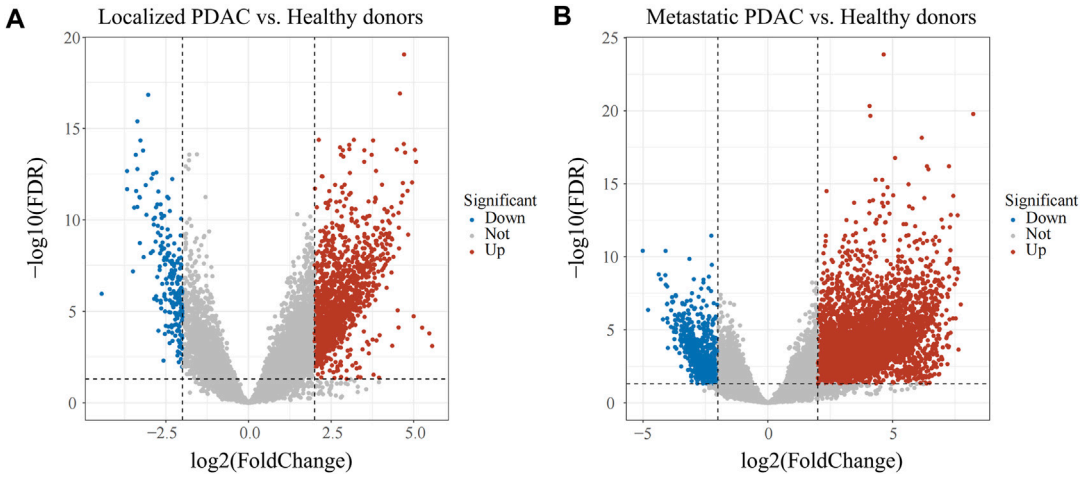


FIGURE 3 | Volcano plot of DEGs in GSE14456. **(A)** DEGs between localized PDAC patients and healthy donors. **(B)** DEGs between metastatic PDAC patients and healthy donors. PDAC, pancreatic ductal adenocarcinoma; DEGs, differentially expressed genes.

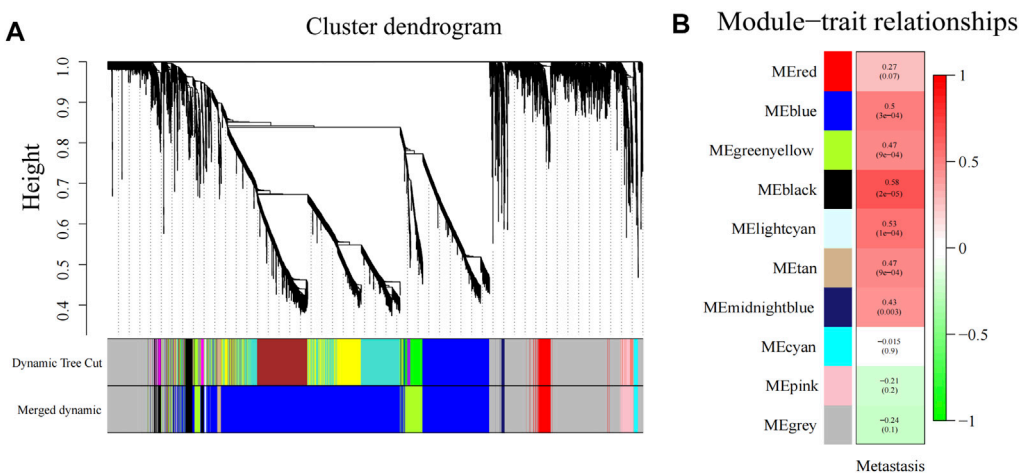


FIGURE 4 | Gene modules identified by WGCNA. **(A)** Gene dendrogram obtained by clustering the dissimilarity based on consensus topological overlap with the corresponding module colors indicated by the color row. **(B)** Heatmap of correlation between the module eigengenes and gene modules. WGCNA, weighted gene co-expression network analysis; ME, module eigengene.

hsa05205, $p < 0.001$) are indispensable signals for metastasis of PDAC. In addition, focal adhesion (GO: 0005925; $p < 0.001$), structural constituent of the cytoskeleton (GO:0005200, $p < 0.001$), blood vessel development (GO:0001568; $p < 0.001$; GO: 0002040; $p < 0.001$), regulation of the glucose metabolic process (GO:0010906, $p < 0.001$), and regulation of the transforming growth factor beta receptor signaling pathway (GO:0017015, $p < 0.001$) were all significantly enriched again (**Figure 7B**). Interestingly, the adhesion, cytoskeleton, and EMT pathway were enriched dramatically in both gene sets (GO:0098632; GO:0005925; GO:0015629; GO:0005200; Hallmark: M5930) (**Figure 7C**). This suggested that CTCs of metastatic PDAC patients might be more likely to undergo EMT or aggregate into CTC clusters. To identify hub genes correlated with metastasis, we screened the coexpressed genes of module genes and marker genes by the Wayne diagram tool, and HMGB3 was identified (**Figure 7D**).

Expression and Prognosis of HMGB3

HMGB3 from the black module has a significant positive correlation with metastasis ($p < 0.001$), and HMGB3 was significantly overexpressed in cluster 1 (CTCs and metastatic tumor cells) as compared to cluster 0 (PT cells) ($p < 0.001$) (**Figures 8A,B**). Furthermore, through the GEPIA 2 database, HMGB3 was also found to be highly expressed in pancreatic carcinoma compared with the matched normal tissue and the

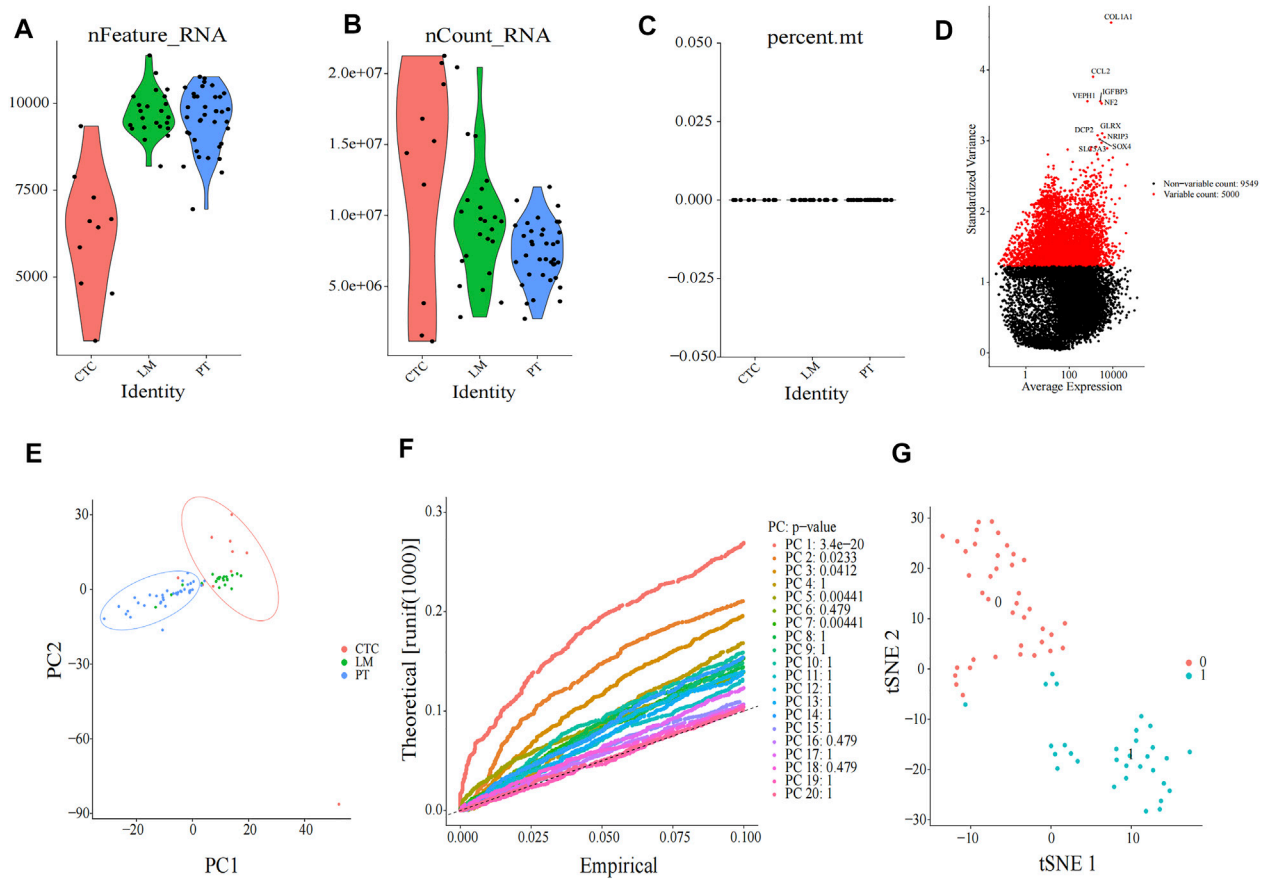


FIGURE 5 | Clustering of cells after single-cell sequencing. Quality check of 70 single-cell sequencing in **(A)** nFeature, **(B)** nCount, and **(C)** percent mt three aspects, as shown in violin plots. **(D)** Scatter diagram of larger expression alteration in different cells and the top 5,000 were marked red pots. **(E)** PCA of 70 cells. **(F)** JackStraw plot of the first 20 PCs. **(G)** t-SNE plot of 70 cells. All cells can be distributed into two clusters (cluster 0 and cluster 1). CTC, circulating tumor cell; LM, liver metastasis; PT, primary tumor; PCA, principal component analysis; PC, principal components; tSNE, t-Stochastic Neighbor Embedding.

normal pancreatic data from the Genotype-Tissue Expression (GTEx) portal (**Figure 8C**). Moreover, protein expression levels were similarly compared using the HPA database, and the low protein expression levels of HMGB3 were revealed in normal pancreatic tissues, while medium protein expression levels of HMGB3 were revealed in pancreatic cancer tissues (**Figure 8D**). We further investigated whether the expression level of HMGB3 correlated with OS in the PAAD by LOGpc. We found that patients with higher HMGB3 expression in GSE28735 and ICGC Array have significantly worse OS (HR = 2.83, p = 0.0427; HR = 1.95, p = 0.0056) (**Figures 8E,F**).

Correlation Analysis of HMGB3 With EMT, CTC Clusters, and Immune Cells

Ultimately, we evaluated the correlation between HMGB3 and EMT-related molecules (vimentin, FN1 and TWIST) (Wei et al., 2019; Lin et al., 2021), CTC cluster–related markers (JUP; TJP3) (Harris and Tepass, 2010; Aceto et al., 2014; Gkountela et al., 2019) and immune cells (T cell $CD8^+$ T cells, regulatory T cells, NK cell, and dendritic cell) to further confirm the key role of HMGB3 in CTC survival and metastasis. Surprisingly, the expression of HMGB3 was significantly positively correlated with the expression of FN1 (Cor = 0.23, p = 0.035), TWIST (Cor = 0.29, p = 0.0085), and JUP3 (Cor = 0.25, p = 0.026) (**Figures 9B,C,E**, respectively), but there was no statistical correlation among HMGB3, vimentin, and JUP in GSE14456 (All p value > 0.05) (**Figures 9A,D**, respectively). In addition, through the GEPIA 2 database, it was found that HMGB3 was also significantly positively correlated with vimentin (Cor = 0.55, p < 0.001), FN1 (Cor = 0.79, p < 0.001), TWIST (Cor = 0.45, p < 0.001), JUP (Cor = 0.48, p < 0.001), and TJP3 (Cor = 0.22, p < 0.001) (**Figures 9F–J**). Using TIMER2.0, we also attempted to investigate the relationship between HMGB3 mutation and immune cell infiltration. Notably, HMGB3 showed significant relationships with tumor-associated $CD8^+$ T cell, $CD4^+$ T cell, T cell regulatory, NK cell, B cell, and dendritic cell infiltrating levels, implying that HMGB3 was favorably related to tumor-associated immune cell infiltration in the CTC microenvironment (All p value < 0.05) (**Figures 9K–P**).

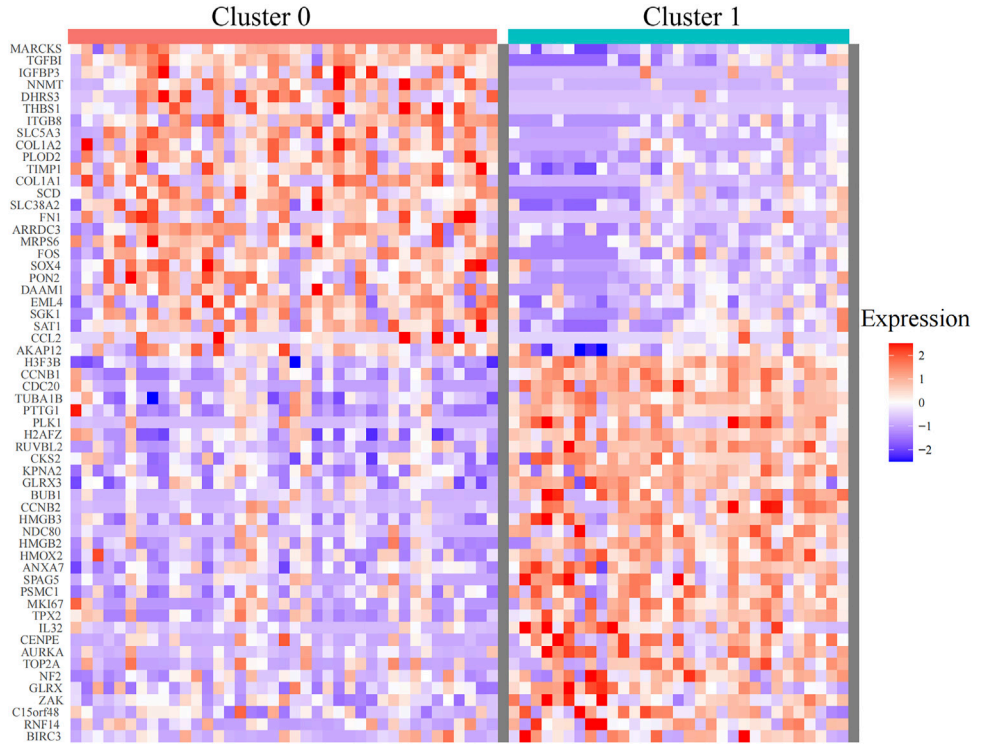


FIGURE 6 | Heatmaps of differentially expressed genes between cluster 0 and cluster 1 identified by TSNE analysis (with adjusted $p < 0.05$). TSNE, t-Stochastic Neighbor Embedding.

DISCUSSION

In PDAC, metastasis tends to occur early. Metastasis is the major cause of cancer-related death, and the mechanism of metastasis remains unclear. In recent years, CTCs detected and isolated from blood samples of cancer patients have attracted great attention. CTCs are considered to be responsible for tumor metastasis, released from the primary tumor and spread in the peripheral blood circulation between immune cells and red blood cells (Jeong et al., 2018). With the development of CTC detection and RNA-sequencing technology, it is gradually recognized that there is great heterogeneity among CTCs, and not all CTCs could survive in blood and consequently lead to distant metastasis. Previous studies demonstrated that specific phenotypes of CTCs, for example, CTCs with EMT and CTCs expressing tumor stemness genes and CTC clusters possessed stronger ability to survive in blood circulation (Barriere et al., 2014; Fabisiewicz and Grzybowska, 2017; Agnoletto et al., 2019). Those CTCs might be more relevant to metastasis and might have their own unique microenvironment (Fabisiewicz and Grzybowska, 2017). In addition, studies have reported that the genetic mutation information carried by CTCs from metastatic patients is 91% consistent with the information of metastatic tumor cells; hence, analysis based on CTCs may reveal the key targets that hinder tumor metastasis (Chemi et al., 2019). Furthermore, clinically, CTC is easy to be obtained and collected multiple times and could be used to evaluate the effect of chemotherapy for drug resistance screening in real time, which avoids the limitation of traditional pathological examination of solid tumor lesions. Therefore, it is critical for this study to explore the metastasis mechanism of PDAC based on CTCs.

For the comprehensiveness of our study, we selected GSE144561 and GSE114704, where CTC samples are all separated by negative selection (CTC-iChip) to preserve the diversity of CTC phenotypes. Although the Food and Drug Administration (FDA) approved the use of CellSearch (positive selection) in CTC enrichment, it could only enrich CTCs with epithelial cell adhesion molecules (EpCAM) (Bankó et al., 2019). Since CTCs exist in peripheral blood and are inactivated by immune cell attack, metastatic PDAC patients might have distinctive microenvironments that support CTC survival and eventually induce metastasis. To investigate the immune landscape of CTCs, we first analyzed the bulk CTC RNA-sequencing dataset GSE144561. The results intimated that the proportion of Tregs and macrophages in metastatic PDAC was significantly higher than that in localized PDAC, while the proportion of DCs was significantly lower (All $p < 0.05$). Tregs are a subgroup of T cells. As important immunosuppressive cells, Tregs participate in a variety of immune regulatory mechanisms to inhibit immune surveillance, promoting the occurrence, development, and metastasis of pancreatic cancer (Liu et al., 2016). Macrophages could also suppress tumor-associated immunogenicity, promote immune tolerance, and lead to poor prognosis (Zhong et al., 2018). Dendritic cells (DCs), as the

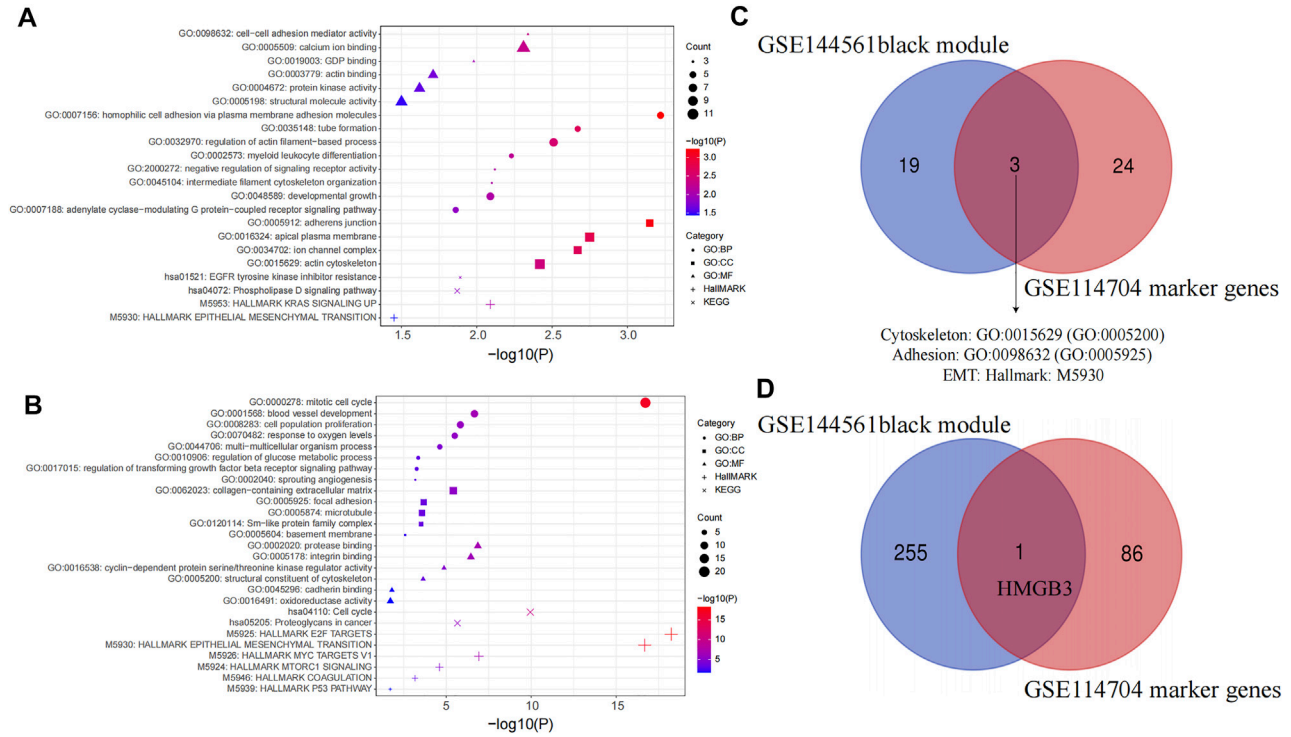


FIGURE 7 | Functional enrichment analysis. **(A)** Functional enrichment analysis of the black module in WGCNA, and **(B)** marker gene between cluster 0 and cluster 1. GO, Gene Ontology; BP, biological processes; CC, cell components; MF, molecular functions; KEGG, Kyoto Encyclopedia of Genes and Genomes. **(C)** Venn plot of enriched pathways in GSE144561 black module and GSE114704 marker genes. **(D)** Venn plot of GSE144561 black module genes and GSE114704 marker genes.

antigen-presenting cells (APCs), could recognize, process, and present tumor antigens to T cells, promoting immune antitumor response (Yamamoto et al., 2012). Hence, the higher proportion of Treg and macrophage infiltration and lower proportion of DC infiltration in CTCs might be responsible for protecting CTCs from immune system attack in circulating blood and ultimately leading to metastasis.

Next, we further explored differences in RNA transcriptome expression between localized PDAC and metastatic PDAC. First, we identified DEGs between PDAC patients and healthy donors. This could lower the potential risk that the enrichment of CTCs might be mixed up with nontumor cells (normal cells in the blood), which affects the accuracy of downstream study. Subsequently, DEGs were analyzed in WGCNA analysis based on localized PDAC and metastatic PDAC. We found that the black modules including 256 genes were significantly positively correlated with metastasis and the correlation coefficient was the highest (Cor = 0.58, $p < 0.001$). In addition, PCA and TSNE analysis were performed on scRNA-seq datasets GSE114704, and it was found that CTCs and metastasis cells could be clustered into cluster 1, which was consistent with the statement that CTCs and metastatic cells were homologous clones (Chemi et al., 2019). On the contrary, cluster 0 mainly comprised primary tumor cells. In order to explore genes correlated with metastasis, which are expressed by CTCs, the genes between cluster 0 and cluster 1 were then differentially analyzed and 87 marker genes were obtained.

Functional enrichment analyses were performed on black module genes and marker genes. Interestingly, the adhesion, cytoskeleton, and EMT pathways were enriched dramatically in both gene sets (GO:0098632; GO:0005925; GO:0015629; GO:0005200; Hallmark: M5930). This suggested that CTCs with metastatic characteristics might be more likely to be CTCs with EMT and CTC clusters. Formation of CTC clusters could protect tumor cells from fluid shear stress, anoikis (apoptotic cell death), and immune surveillance; hence, it had 23–50 times increased metastatic ability compared to that of single CTC (Aceto et al., 2014). The previous immune cell infiltration analysis also revealed that the CTC of patients with metastasis had a higher proportion of macrophages, which also suggested that metastatic patients formed more CTC clusters comprising CTC and macrophages. In terms of EMT, it has been considered an important factor leading to metastasis, promoting migration, invasive potential, and resistance to anoikis (Fabisiewicz and Grzybowska, 2017). Then, we intend to identify the key genes for CTC survival and ultimately metastasis. We found that HMGB3 was the only coexpressed gene of hub module genes and marker genes. Interestingly, we found that HMGB3 was, indeed, associated with immune cell infiltration ($CD8^+$ T cells, regulatory T cells, NK cell, and dendritic cell), EMT-related molecules (vimentin, FN1, and TWIST), and CTC cluster–related markers (JUP and TJP3) (All $p < 0.05$). In addition, the RNA and protein expression of HMGB3 in PDAC were significantly higher than those in normal

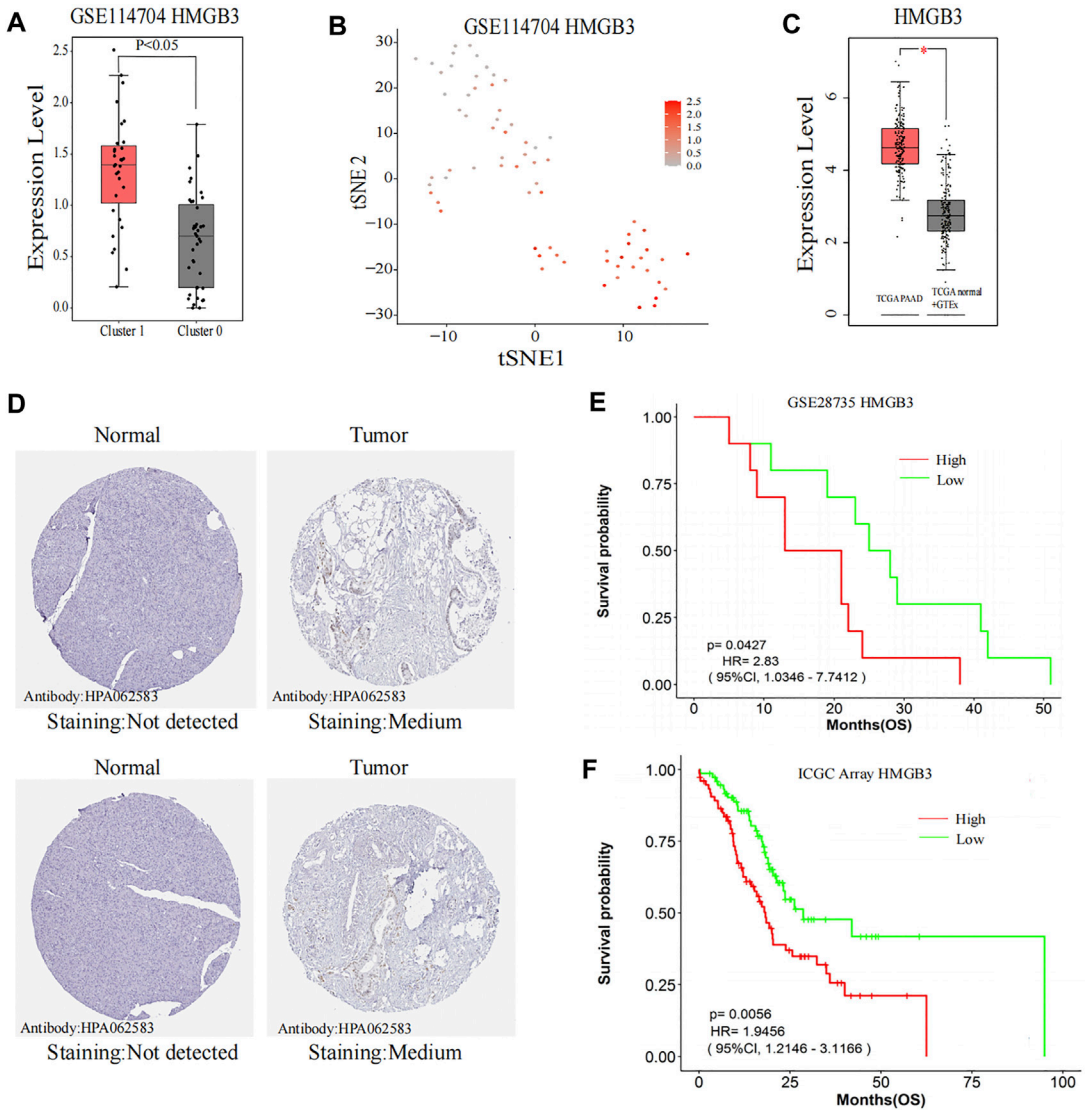


FIGURE 8 | Further research on the hub gene. **(A,B)** Comparison of HMGB3 expression between Cluster 0 and Cluster 1. **(C)** HMGB3 expression in the TCGA pancreatic tumors with corresponding matched normal tissue and GTEx pancreatic tissue. **(D)** Representative immunohistochemistry images of HMGB3 in PAAD and noncancerous pancreatic tissues derived from the HPA database. **(E)** Comparisons of OS curves between high and low expression of HMGB3 in GSE28735. **(F)** Comparisons of OS curves between high and low expression of HMGB3 in ICGC Array. TCGA, The Cancer Genome Atlas; ICGC, International Cancer Genome Consortium; GTEx, The Genotype-Tissue Expression; HPA, the Human Protein Atlas; HMGB3, high-mobility group box 3; PAAD, pancreatic adenocarcinoma; OS, overall survival; HR, hazard ratio; CI, confidence interval.

tissue ($p < 0.05$), and the OS of PDAC patients with higher HMGB3 expression was significantly worse in multiple datasets (GSE28735: HR = 2.83, $p = 0.0427$; ICGC Array: HR = 1.95, $p = 0.0056$).

High-mobility group box 3 (HMGB3) belongs to the high mobility group protein B (HMGB) family and can regulate the mechanisms of DNA replication, transcription, recombination, and repair. It also acts as cytokines to mediate responses to infection, injury, and inflammation (Yanai et al., 2011; Lv et al., 2019). HMGB3 has low or no expression in normal adult tissues but high expression in various kinds of tumor tissues, including esophageal cancer, lung cancer, breast cancer, gastric adenocarcinoma, bladder cancer, prostate cancer, and glioma (Gong et al., 2013; Li et al., 2015; Yamada et al., 2018; Gu et al., 2019; Lv et al., 2019; Xie et al., 2020; Yin and Liu, 2020). HMGB3 is considered to be an oncogene and a hub gene in tumor growth, participating in the growth, migration, invasion, immune escape, and even EMT of tumor cells (Fang et al., 2020; Chen et al., 2021). EMT was also found to be advantageous to CTC survival and metastasis in this study. HMGB3 also promotes self-renewal and colony formation of cancer stem cells and cancer cells (Song et al., 2019). In addition, it was shown that HMGB3 protein could bind to chemotherapeutic drug–induced damage DNA, leading to the activation of ATM-and Rad3-related protein, thus promoting the

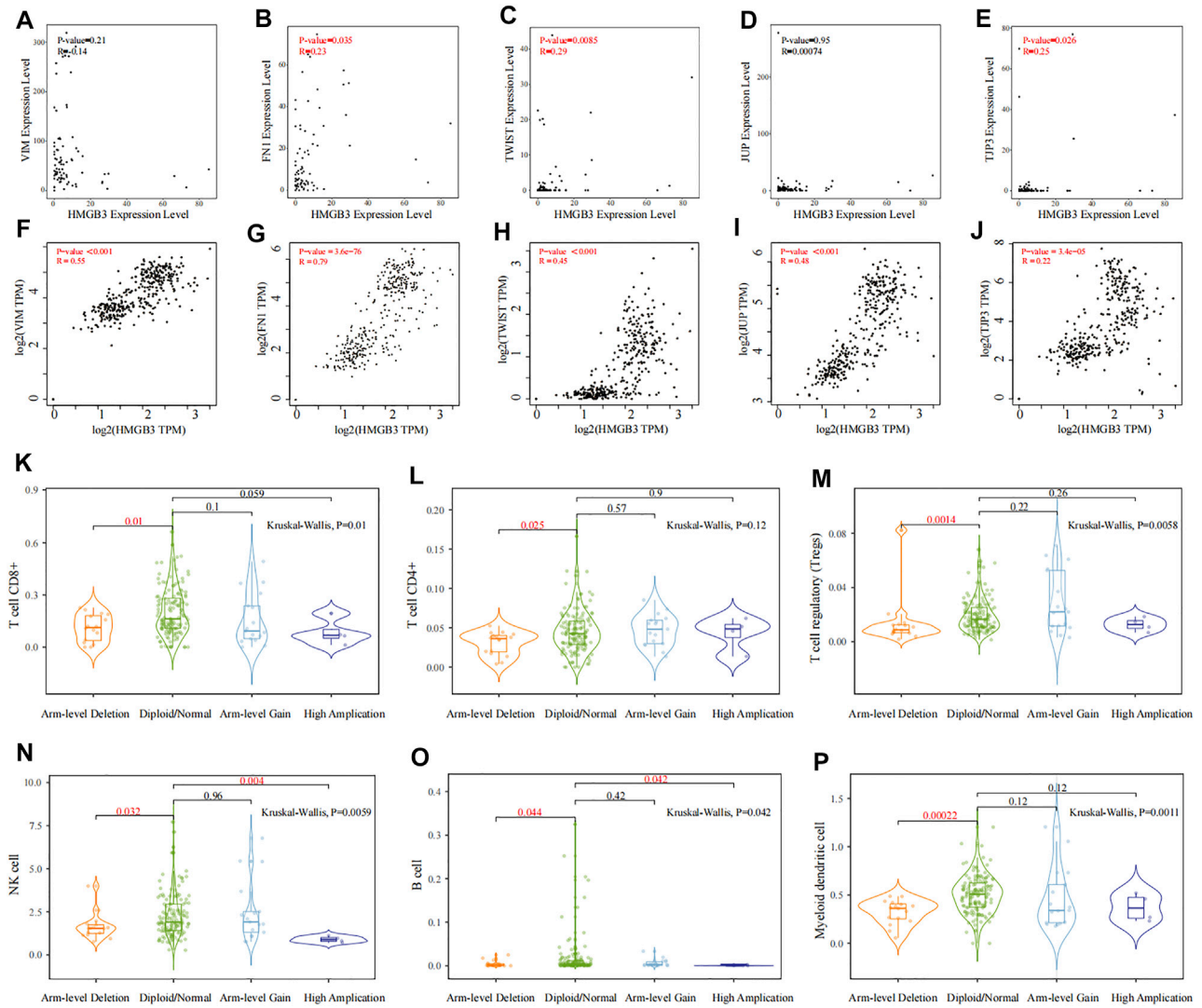


FIGURE 9 | Correlation analysis of HMGB3 with EMT, CTC clusters, and immune cells. **(A–E)** Correlation analysis between HMGB3 expression and VIM, FN1, TWIST, JUP, and TJP3 in GSE14456, respectively. **(F–J)** Relationship between HMGB3 expression and VIM, TWIST, JUP, and TJP3 through the GEPIA 2 database, respectively. **(K–P)** Relationship of HMGB3 mutation with tumor-associated $CD8^+$ T cell, $CD4^+$ T cell, T cell regulatory, NK cell, B cell, and myeloid dendritic cell infiltrating levels by TIMER2.0, respectively.

drug resistance of tumor cells (Mukherjee et al., 2019; Jiang et al., 2020). Surprisingly, it was consistent that the EGFR tyrosine kinase inhibitor resistance pathway (KEGG: hsa01521, $p = 0.013$) was significantly enriched in the present study. Moreover, the molecular function of HMGB3 is consistent with the conditions (resistance to fluid shear stress, anoikis, and immune surveillance) required for the survival and metastasis of CTC in blood circulation, suggesting that HMGB3 might be a key gene of CTCs in pancreatic cancer aggressiveness.

But, there are few studies on HMGB in pancreatic cancer, and our study is the first to investigate the metastasis mechanism of PDAC based on CTCs. It seems to be promising that the treatment with HMGB3 as an anticancer target could inhibit the survival and metastasis of CTCs and the proliferation of primary tumor cells, which puts forward a new direction for the treatment of PDAC. Our research center has enriched and isolated CTCs by ScreenCell® as previously described (a filtration method that allows the isolation of various CTCs by size, ensuring the diversity of CTCs) (Francescangeli et al., 2021) and is currently conducting research on primary CTC culture. Once we manage to cultivate primary pancreatic CTCs, subsequent cytological experiments would be carried out to further verify the role of HMGB3 in CTCs.

CONCLUSION

In conclusion, the results of our study suggest that CTCs of metastatic PDAC have specific immune landscape, RNA expression profile, and functional pathways. HMGB3 is a hub

gene associated with the survival and metastasis of CTCs. This study suggests a potentially worthwhile metastasis mechanism of PDAC based on CTCs, further deepening the understanding of PDAC development. Therapeutic agents targeted at HMGB3 in CTCs might provide a novel strategy for the treatment of PDAC.

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

Conception and design of the study: H-DT, YW, and J-HZ. H-DT, S-YT, H-FL, HS, and ZZ contributed to collection of data, data analysis, and interpretation. H-DT and YW drafted the manuscript. J-HZ, Z-QL, and PX critically revised the manuscript for important intellectual content. All authors read, revised, and approved the final draft.

FUNDING

This project was supported by the National Natural Science Foundation of China (81572408), the Natural Science Foundation of Jiangsu Province (BE2015712), and the Program of Medical Innovation Team and Leading Medical Talents in Jiangsu Province (2017ZXKJQW09).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.892177/full#supplementary-material

REFERENCES

Aceto, N., Bardia, A., and Miyamoto, D. T. (2014). Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell.* 158 (5), 1110–1122. doi:10.1016/j.cell.2014.07.013

Agnoletto, C., Corrà, F., and Minotti, L. (2019). Heterogeneity in Circulating Tumor Cells: The Relevance of the Stem-Cell Subset. *Cancers (Basel)* 11 (4), 483. doi:10.3390/cancers11040483

Ashworth, T. R. (1869). A Case of Cancer in Which Cells Similar to Those in the Tumours Were Seen in the Blood after Death. *Aust. Med. J.* 14, 146

Asplund, A., Edqvist, P. H., Schwenk, J. M., and Ponten, F. (2012). Antibodies for Profiling the Human Proteome-The Human Protein Atlas as a Resource for Cancer Research. *Proteomics* 12 (13), 2067–2077. doi:10.1002/pmic.201100504

Bankó, P., Lee, S. Y., and Nagygyörgy, V. (2019). Technologies for Circulating Tumor Cell Separation from Whole Blood. *J. Hematol. Oncol.* 12 (1), 48. doi:10.1186/s13045-019-0735-4

Barriere, G., Fici, P., Gallerani, G., Fabbri, F., Zoli, W., and Rigaud, M. (2014). Circulating Tumor Cells and Epithelial, Mesenchymal and Stemness Markers: Characterization of Cell Subpopulations. *Ann. Transl. Med.* 2 (11), 109. doi:10.3978/j.issn.2305-5839.2014.10.04

Chemi, F., Rothwell, D. G., and McGranahan, N. (2019). Pulmonary Venous Circulating Tumor Cell Dissemination before Tumor Resection and Disease Relapse. *Nat. Med.* 25 (10), 1534–1539. doi:10.1038/s41591-019-0593-1

Chen, Q., Pu, N., and Yin, H. (2020). CD73 Acts as a Prognostic Biomarker and Promotes Progression and Immune Escape in Pancreatic Cancer. *J. Cell. Mol. Med.* 24 (15), 8674–8686. doi:10.1111/jcmm.15500

Chen, W., Zheng, R., and Baade, P. D. (2016). Cancer Statistics in China, 2015. *CA Cancer J. Clin.* 66 (2), 115–132. doi:10.3322/caac.21338

Chen, Z., Pei, L., Zhang, D., Xu, F., Zhou, E., and Chen, X. (2021). HDAC3 Increases HMGB3 Expression to Facilitate the Immune Escape of Breast Cancer Cells via Down-Regulating microRNA-130a-3p. *Int. J. Biochem. Cell. Biol.* 135, 105967. doi:10.1016/j.biocel.2021.105967

Chen, K., Wang, Q., Li, M., et al. (2021). Single-Cell RNA-Seq Reveals Dynamic Change in Tumor Microenvironment During Pancreatic Ductal Adenocarcinoma Malignant Progression. *EBioMedicine* 66, 103315. doi:10.1016/j.ebiom.2021.103315

Cui, K., Ou, Y., Shen, Y., Li, S., and Sun, Z. (2020). Clinical Value of Circulating Tumor Cells for the Diagnosis and Prognosis of Hepatocellular Carcinoma (HCC): A Systematic Review and Meta-Analysis. *Med. Baltim.* 99 (40), e22242. doi:10.1097/md.0000000000022242

Dimitrov-Markov, S., Perales-Paton, J., and Bockorny, B. (2020). Discovery of New Targets to Control Metastasis in Pancreatic Cancer by Single-Cell Transcriptomics Analysis of Circulating Tumor Cells. *Mol. Cancer Ther.* 19 (8), 1751–1760. doi:10.1158/1535-7163.mct-19-1166

Fabisiewicz, A., and Grzybowska, E. (2017). CTC Clusters in Cancer Progression and Metastasis. *Med. Oncol.* 34 (1), 12. doi:10.1007/s12032-016-0875-0

Fang, J., Ge, X., and Xu, W. (2020). Bioinformatics Analysis of the Prognosis and Biological Significance of HMGB1, HMGB2, and HMGB3 in Gastric Cancer. *J. Cell. Physiol.* 235 (4), 3438–3446. doi:10.1002/jcp.29233

Francescangeli, F., Magri, V., and De Angelis, M. L. (2021). Sequential Isolation and Characterization of Single CTCs and Large CTC Clusters in Metastatic Colorectal Cancer Patients. *Cancers (Basel)* 13 (24). doi:10.3390/cancers13246362

Franses, J. W., Philipp, J., and Missios, P. (2020). Pancreatic Circulating Tumor Cell Profiling Identifies LIN28B as a Metastasis Driver and Drug Target. *Nat. Commun.* 11 (1), 3303. doi:10.1038/s41467-020-17150-3

Gkountela, S., Castro-Giner, F., and Szczerba, B. M. (2019). Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. *Cell.* 176 (1-2), 98–112 e114. doi:10.1016/j.cell.2018.11.046

Goldkorn, A., Ely, B., and Quinn, D. I. (2014). Circulating Tumor Cell Counts Are Prognostic of Overall Survival in SWOG S0421: a Phase III Trial of Docetaxel with or without Atrasentan for Metastatic Castration-Resistant Prostate Cancer. *J. Clin. Oncol.* 32 (11), 1136–1142. doi:10.1200/jco.2013.51.7417

Gong, Y., Cao, Y., Song, L., Zhou, J., Wang, C., and Wu, B. (2013). HMGB3 Characterization in Gastric Cancer. *Genet. Mol. Res.* 12 (4), 6032–6039. doi:10.4238/2013.december.2.1

Goodman, C. R., Seagle, B. L., and Friedl, T. W. P. (2018). Association of Circulating Tumor Cell Status with Benefit of Radiotherapy and Survival in Early-Stage Breast Cancer. *JAMA Oncol.* 4 (8), e180163. doi:10.1001/jamaoncol.2018.0163

Gu, J., Xu, T., Huang, Q. H., Zhang, C. M., and Chen, H. Y. (2019). HMGB3 Silence Inhibits Breast Cancer Cell Proliferation and Tumor Growth by Interacting with Hypoxia-Inducible Factor 1α. *Cancer Manag. Res.* 11, 5075–5089. doi:10.2147/cmar.s204357

Harris, T. J., and Tepass, U. (2010). Adherens Junctions: from Molecules to Morphogenesis. *Nat. Rev. Mol. Cell. Biol.* 11 (7), 502–514. doi:10.1038/nrm2927

Hidalgo, M., and Von Hoff, D. D. (2012). Translational Therapeutic Opportunities in Ductal Adenocarcinoma of the Pancreas. *Clin. Cancer Res.* 18 (16), 4249–4256. doi:10.1158/1078-0432.ccr-12-1327

Huang, X., Gao, P., and Song, Y. (2015). Meta-analysis of the Prognostic Value of Circulating Tumor Cells Detected with the CellSearch System in Colorectal Cancer. *BMC Cancer* 15, 202. doi:10.1186/s12885-015-1218-9

Jeong, K. Y., Kim, E. K., Park, M. H., and Kim, H. M. (2018). Perspective on Cancer Therapeutics Utilizing Analysis of Circulating Tumor Cells. *Diagn. (Basel)* 823 (2), 23. doi:10.3390/diagnostics8020023

Jiang, D., Wang, Y., and Liu, M. (2020). A Panel of Autoantibodies against Tumor-Associated Antigens in the Early Immunodiagnosis of Lung Cancer. *Immunobiology* 225 (1), 151848. doi:10.1016/j.imbio.2019.09.007

Keller, L., and Pantel, K. (2019). Unravelling Tumour Heterogeneity by Single-Cell Profiling of Circulating Tumour Cells. *Nat. Rev. Cancer* 19 (10), 553–567. doi:10.1038/s41568-019-0180-2

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

Li, M., Cai, Y., and Zhao, H. (2015). Overexpression of HMGB3 Protein Promotes Cell Proliferation, Migration and Is Associated with Poor Prognosis in Urinary Bladder Cancer Patients. *Tumour Biol.* 36 (6), 4785–4792. doi:10.1007/s13277-015-3130-y

Li, T., Fu, J., and Zeng, Z. (2020). TIMER2.0 for Analysis of Tumor-Infiltrating Immune Cells. *Nucleic Acids Res.* 48 (W1), W509–W514. doi:10.1093/nar/gkaa407

Liao, Y., Wang, S. Y., and Meng, X. Y. (2014). Circulating Tumor Cells in Breast Cancer and its Association with Tumor Clinicopathological Characteristics: a Meta-Analysis. *Med. Oncol.* 31 (12), 343. doi:10.1007/s12032-014-0343-7

Lin, D., Shen, L., and Luo, M. (2021). Circulating Tumor Cells: Biology and Clinical Significance. *Signal Transduct. Target Ther.* 6 (1), 404. doi:10.1038/s41392-021-00817-8

Liu, L., Zhao, G., and Wu, W. (2016). Low Intratumoral Regulatory T Cells and High Peritumoral CD8(+) T Cells Relate to Long-Term Survival in Patients with Pancreatic Ductal Adenocarcinoma after Pancreatectomy. *Cancer Immunol. Immunother.* 65 (1), 73–82. doi:10.1007/s00262-015-1775-4

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 (12), 550. doi:10.1186/s13059-014-0550-8

Lv, Y., Lv, M., and Ji, X. (2019). Down-regulated Expressed Protein HMGB3 Inhibits Proliferation and Migration, Promotes Apoptosis in the Placentas of Fetal Growth Restriction. *Int. J. Biochem. Cell. Biol.* 107, 69–76. doi:10.1016/j.biocel.2018.11.007

Miyamoto, D. T., Lee, R. J., and Kalinich, M. (2018). An RNA-Based Digital Circulating Tumor Cell Signature Is Predictive of Drug Response and Early Dissemination in Prostate Cancer. *Cancer Discov.* 8 (3), 288–303. doi:10.1158/2159-8290.cd-16-1406

Mukherjee, A., Huynh, V., Gaines, K., Reh, W. A., and Vasquez, K. M. (2019). Targeting the High-Mobility Group Box 3 Protein Sensitizes Chemoresistant Ovarian Cancer Cells to Cisplatin. *Cancer Res.* 79 (13), 3185–3191. doi:10.1158/0008-5472.can-19-0542

Newman, A. M., Liu, C. L., and Green, M. R. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., and Matrisian, L. M. (2014). Projecting Cancer Incidence and Deaths to 2030: the Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. *Cancer Res.* 74 (11), 2913–2921. doi:10.1158/0008-5472.can-14-0155

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

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer Statistics, 2020. *CA A Cancer J. Clin.* 70 (1), 7–30. doi:10.3322/caac.21590

Song, N., Wang, B., and Feng, G. (2019). Knockdown of High Mobility Group Box 3 Impairs Cell Viability and Colony Formation but Increases Apoptosis in A549 Human Non-small Cell Lung Cancer Cells. *Oncol. Lett.* 17 (3), 2937–2945. doi:10.3892/ol.2019.9927

Stuart, T., Butler, A., and Hoffman, P. (2019). Comprehensive Integration of Single-Cell Data. *Cell.* 177 (7), 1888, doi:10.1016/j.cell.2019.05.031

Wang, L., Wu, C., and Qiao, L. (2017). Clinical Significance of Folate Receptor-Positive Circulating Tumor Cells Detected by Ligand-Targeted Polymerase Chain Reaction in Lung Cancer. *J. Cancer* 8 (1), 104–110. doi:10.7150/jca.16856

Wang, Y., Yu, X., Hartmann, D., and Zhou, J. (2020). Circulating Tumor Cells in Peripheral Blood of P.ncreatic Cancer Patients and Their Prognostic Role: a Systematic Review and Meta-Analysis. *HPB Oxf.* 22 (5), 660–669. doi:10.1016/j.hpb.2019.11.003

Wei, T., Zhang, X., and Zhang, Q. (2019). Vimentin-positive Circulating Tumor Cells as a Biomarker for Diagnosis and Treatment Monitoring in Patients with Pancreatic Cancer. *Cancer Lett.* 452, 237–243. doi:10.1016/j.canlet.2019.03.009

Wu, Z., Shou, L., Wang, J., and Xu, X. (2020). Identification of the Key Gene and Pathways Associated with Osteoarthritis via Single-Cell RNA Sequencing on Synovial Fibroblasts. *Med. Baltim.* 99 (33), e21707. doi:10.1097/md.0000000000021707

Xie, Y., Wang, L., and Yang, D. (2020). CircEPSTI1 Promotes the Progression of Non-small Cell Lung Cancer through miR-145/HMGB3 Axis. *Cancer Manag. Res.* 12, 6827–6836. doi:10.2147/cmar.s252893

Yamada, Y., Nishikawa, R., and Kato, M. (2018). Regulation of HMGB3 by Antitumor miR-205-5p Inhibits Cancer Cell Aggressiveness and Is Involved in Prostate Cancer Pathogenesis. *J. Hum. Genet.* 63 (2), 195–205. doi:10.1038/s10038-017-0371-1

Yamamoto, T., Sohei Yanagimoto, H., and Fau - Toyokawa, S. (2012). Circulating Myeloid Dendritic Cells as Prognostic Factors in Patients with Pancreatic Cancer Who Have Undergone Surgical Resectionno. *J. Surg. Res.* 173, 299. doi:10.1016/j.jss.2010.09.027

Yanai, H., Chiba, S., and Ban, T. (2011). Suppression of Immune Responses by Nonimmunogenic Oligodeoxynucleotides with High Affinity for High-Mobility Group Box Proteins (HMGBs). *Proc. Natl. Acad. Sci. U. S. A.* 108 (28), 11542–11547. doi:10.1073/pnas.1108535108

Yin, K., and Liu, X. (2020). CircMMP1 Promotes the Progression of Glioma through miR-433/HMGB3 axis *In Vitro* and *In Vivo*. *IUBMB Life* 72 (11), 2508–2524. doi:10.1002/iub.2383

Zhang, G., Wang, Q., and Yang, M. (2020). OSpaad: An Online Tool to Perform Survival Analysis by Integrating Gene Expression Profiling and Long-Term Follow-Up Data of 1319 Pancreatic Carcinoma Patients. *Mol. Carcinog.* 59 (3), 304–310. doi:10.1002/mc.23154

Zhong, X., Chen, B., and Yang, Z. (2018). The Role of Tumor-Associated Macrophages in Colorectal Carcinoma Progression, *Cell. Physiol. Biochem.* 45, 1421–9778. doi:10.1159/000486816

Zhou, Y., Zhou, B., and Pache, L. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6

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 Tang, Wang, Xie, Tan, Li, Shen, Zhang, Lei 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: 21 June 2022
doi: 10.3389/fgene.2022.852708

The Immune Cell Infiltration Patterns and Characterization Score in Bladder Cancer to Identify Prognosis

Yongsheng Zhang[1,2], *Yunlong Wang*[1,3]*, *Jichuang Wang*[3] and *Kaixiang Zhang*[4]

[1]The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [2]Academy of Medical Science, Zhengzhou University, Zhengzhou, China, [3]Henan Bioengineering Research Center, Zhengzhou, China, [4]School of Pharmaceutical Sciences, Zhengzhou University, Zhengzhou, China

Background: Bladder cancer (BLCA) is among the most frequent types of cancer. Patients with BLCA have a significant recurrence rate and a poor post-surgery survival rate. Recent research has found a link between tumor immune cell infiltration (ICI) and the prognosis of BLCA patients. However, the ICI's picture of BLCA remains unclear.

Methods: Common gene expression data were obtained by combining the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) expression databases. Two computational algorithms were proposed to unravel the ICI landscape of BLCA patients. The R package "limma" was applied to find differentially expressed genes (DEGs). ICI patterns were defined by the unsupervised clustering method. Principal-component analysis (PCA) was used to calculate the ICI score. In addition, the combined ICI score and tumor burden mutation (TMB) were utilized to assess BLCA patients' prognosis. The predictive value of ICI scores was verified by different clinical characteristics.

Results: A total of 569 common gene expression data were retrieved from TCGA and GEO cohorts. $CD8^+$ T cells were found to have a substantial positive connection with activated memory $CD4^+$ T cells and immune score. On the contrary, $CD8^+$ T cells were found to have a substantial negative connection with macrophages M0. Thirty-eight DEGs were selected. Two ICI patterns were defined by the unsupervised clustering method. Patients of BLCA were separated into two groups. The high ICI score group exhibited a better outcome than the low ICI score one ($p < 0.001$). Finally, the group with a high tumor mutation burden (TMB) as well as a high ICI score had the best outcome. ($p < 0.001$).

Conclusions: Combining TMB and ICI scores resulted in a more accurate survival prediction, suggesting that ICI scores could be used as a prognostic marker for BLCA patients.

Keywords: bladder cancer, common gene expression samples data, immune cell infiltration, prognosis, overall survival rate

OPEN ACCESS

Edited by:
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Xianghong Zhou,
Sichuan University, China
Michael Ntim,
Kwame Nkrumah University of Science and Technology, Ghana

***Correspondence:**
Yunlong Wang
biowyl@126.com

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

Received: *11 January 2022*
Accepted: *16 May 2022*
Published: *21 June 2022*

Citation:
Zhang Y, Wang Y, Wang J and Zhang K (2022) The Immune Cell Infiltration Patterns and Characterization Score in Bladder Cancer to Identify Prognosis. Front. Genet. 13:852708. doi: 10.3389/fgene.2022.852708

INTRODUCTION

Bladder cancer (BLCA) is the world's 10th most prevalent cancer, accounting for around 549,000 new cases and 200,000 deaths in 2018 (Bray et al., 2018). As a highly heterogeneous tumor (Knowles and Hurst, 2015; Babjuk et al., 2017; Li et al., 2020), BLCA has a high recurrence rate (around 50%) and the five-year survival rate was around 60% after trimodal therapy (Cambier et al., 2016; Kamat et al., 2016; Sanli et al., 2017). Despite the rapid development of clinical imaging after chemotherapy and surgery, the method for evaluating the therapeutic effect of BLCA is not satisfactory. As a result, developing new diagnostic, therapeutic, and prognostic biomarkers for BLCA is critical.

Immune checkpoint inhibitor (ICI_S), a type of immunotherapy, could kill tumor cells. However, it only works for a few patients with advanced cancer (Johnson et al., 2017; D'Aiello et al., 2021; Topalian et al., 2015; Tøndell et al., 2021; Zou et al., 2016; Dermani et al., 2019). TMB is a predictive biomarker of immunotherapy because it reflects the overall load of new antigens (Rizvi et al., 2015; Hellmann et al., 2019; Zhao et al., 2021). However, the breakpoint between TMB-high and TMB-low is difficult to define (Samstein et al., 2019). Therefore, it is critical to find novel biomarkers that could predict the response of the tumor to immunotherapy.

Extensive research has established the crucial involvement of immune cell infiltration (ICI) in cancer proliferation, recurrence, and metastasis (Jiang et al., 2018; Zeng et al., 2018). The higher the proportion of immune score in the tumor microenvironment, the better prognosis in most patients (Na and Choi, 2018).In addition, tumor-infiltrating lymphocytes (TLSs), including CD4 and CD8 T cells, have been linked to increased survival rates (Vassilakopoulou et al., 2016; Zeng et al., 2019). In contrast, TAM exerts its tumor-promoting effect mainly through the following three ways (Noy and Pollard, 2014; Chen et al., 2019; Liang et al., 2020): 1) TAM can promote tumor growth, invasion, metastasis, and angiogenesis by secreting a variety of cytokines; 2) TAM is also immunosuppressive, inhibiting adaptive immune response and promoting tumor immune escape; 3) TAM can also induce drug resistance of tumor cells by promoting abnormal angiogenesis, affecting the transport of drugs in the blood, and weakening the signal of tumor cell apoptosis. Nevertheless, recognizing TLS cells is insufficient to characterize the complicated tumor microenvironment. TLSs and TAMs interact with each other, indicating that the link between the two sets of TME cells is more important than any single component (Sanli et al., 2017).

In this study, two approaches "CIBERSORT" and "ESTIMATE" were employed to unveil the patient's ICI picture. In addition, based on the ICI and DEGs, BLCA patients were classified into two subgroups. The ICI score was acquired by principal-component analysis (PCA). Finally, the ICI score was developed to describe distinct immune cell landscapes, which could exactly predict patient outcomes. "As a result, we discovered that ICI score could serve as a prospective prognostic marker that is different from TMB."

METHODS

BLCA Data Collection

TCGA and GEO databases were used to gather transcriptome and clinical data. In general, we collected two groups of cohort samples of BLCA: GSE13507 and TCGA-BLCA. The exclusion criteria were as follows: 1) Not tumor tissue sample. 2) The transcriptome sequencing data or clinical information of the samples were incomplete. 3) Not common gene expression sample data. Finally, 569 samples were included. We converted the fragments per kilobase million (FPKM) values to the transcripts per million (TPM) values by using the "limma" R package for the TCGA-BLCA database. We combined TCGA and GEO expression data to get new common gene expression samples' data for later analysis.

The Proportion of ICI Was Used to Categorize BLCA Patients

The "CIBERSORT" R package, the LM22 signature, and 1,000 permutations were used to evaluate infiltration levels for various immune cells in BLCA. ESTIMATE calculated the immune score and stromal score in BLCA patients. In addition, we acquired the correlation between different immune cells by using the "corrplot" R package. The hierarchical agglomerative clustering of BLCA was implemented by different ICI patterns of each sample. The number of clusters was determined by the consensus clustering algorithm. We performed the "ConsensusClusterPlus" R package and repeated it 1,000 times to ensure the stability of classification.

Acquisition of Differentially Expressed Genes Related to ICI Phenotype

In order to find genes linked with ICI patterns, we classified patients into distinct groups based on ICI. DEGs among different groups were screened by means of the R package "limma." The significant criteria of $|\log FC| > 1$ and p (adjust) < 0.05 were used to determine DEGs.

Generation of ICI Score

In order to further analysis, an unsupervised clustering method for DEG analysis was applied to divide the patients into different groups. Positive and negative DEG correlations with cluster signatures were classified as ICI gene signatures A and B, and the "Boruta" algorithm was applied to reduce their dimensionality. Using the PCA, the gene signature score of patients was derived. Finally, we used a procedure analogous to the grading index gene expression to determine the ICI score. ICI score = $\sum$PC1A-$\sum$PC1B.

Collection of Somatic Structural Variation Data

The correlation mutation information of patients in the TCGA-BLCA cohort was obtained from the TCGA data portal (https://portal.gdc.cancer.gov/repository). In order to determine tumor

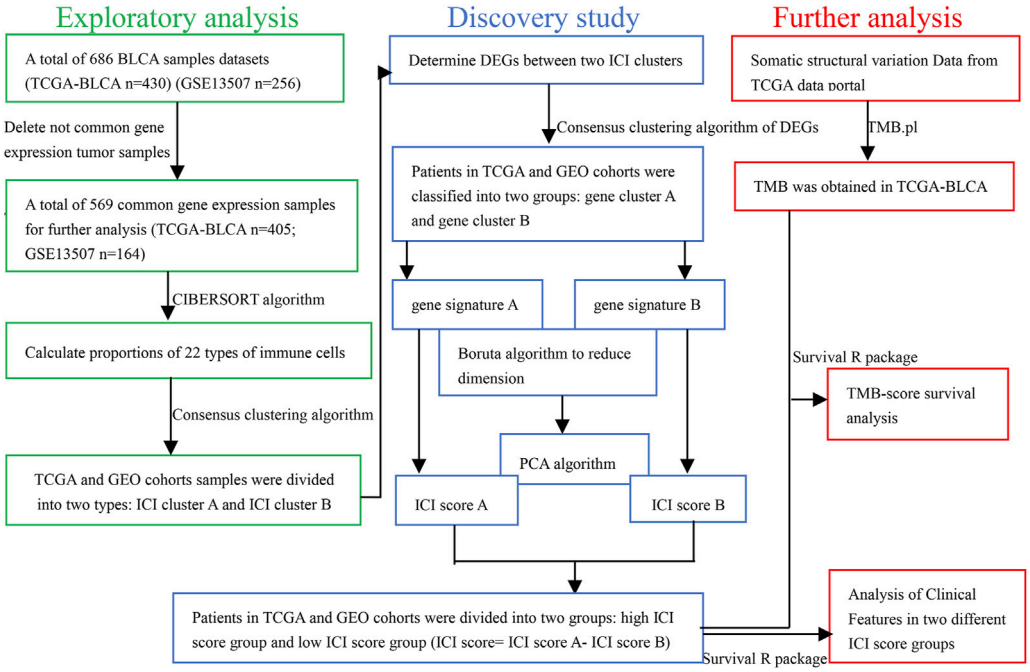


FIGURE 1 | Schematic diagram of research design.

mutation burden, we calculated the total number of non-synonymous mutations in BLCA. We got 20 driver genes through the R package "maftool," which had the highest mutation frequency in BLCA patients. Finally, we evaluated whether differences in the mutation frequency of genes between two ICI score groups.

Gene Ontology and Gene Set Enrichment Analysis

The "clusterprofiler" package was utilized for gene annotation and enrichment analysis of ICI distinctive genes. Gene ontology (GO) terms were screened by a stringent cut-off ($p < 0.05$). Furthermore, we used a gene set enrichment analysis (GSEA) to find pathways that were up- and down-regulated between two ICI score groups. The parameter settings were Gene sets database = "Kyoto Encyclopedia of Genes and Genomes (KEGG)," n Perm = 1,000, and $p < 0.05$.

Analysis of Clinical Features in Two ICI Score Groups

The corresponding clinical data from the TCGA and GEO databases were retrieved and manually organized. We verified the predictive value of the ICI score with distinct clinical features (such as age and gender) by the R package" survival."

Statistical Analysis

All data were analyzed by R software (version 4.0.4). The proportion of 22 types of immune cells in BLCA was calculated by the "CIBERSORT" algorithm. The "ConsensusClusterPlus" R package was used to divide BLCA patients into two types. DEGs between two ICI phenotypes were filtered through the R package "limma." The ICI score was calculated by the PCA algorithm. Further analysis, TMB was obtained in TCGA-BLCA by "TMB.pl." The prognosis of BLCA patients was evaluated by the R package "survival." The "clusterprofiler" package was used for gene annotation and enrichment analysis of ICI distinctive genes. The predictive value of ICI scores with different clinical characteristics (such as age and gender) was verified by the R package "survival." $p < 0.05$ was considered statistically significant.

RESULTS

The Pattern e of ICI in the TME of BLCA

The workflow is displayed in **Figure 1**. First, a total of 569 common gene expression data were extracted from TCGA and GEO cohorts. Then, the CIBERSORT and ESTIMATE algorithms were applied to evaluate the levels of immune cells (filter conditions: $p < 0.05$) in BLCA patients. (**Supplementary Tables 1, 2**). The correlation coefficient heatmap shows the significant positive correlation of CD8 T cells with activated memory CD4 T cells and immune score. On the contrary, there was a significant negative correlation between CD8 T cells with macrophages M0 (**Figure 3C**). Based on the results of immune cell infiltration, unsupervised clustering was performed by using the "ConsesusClusterPlus" package of R software to divide BLCA

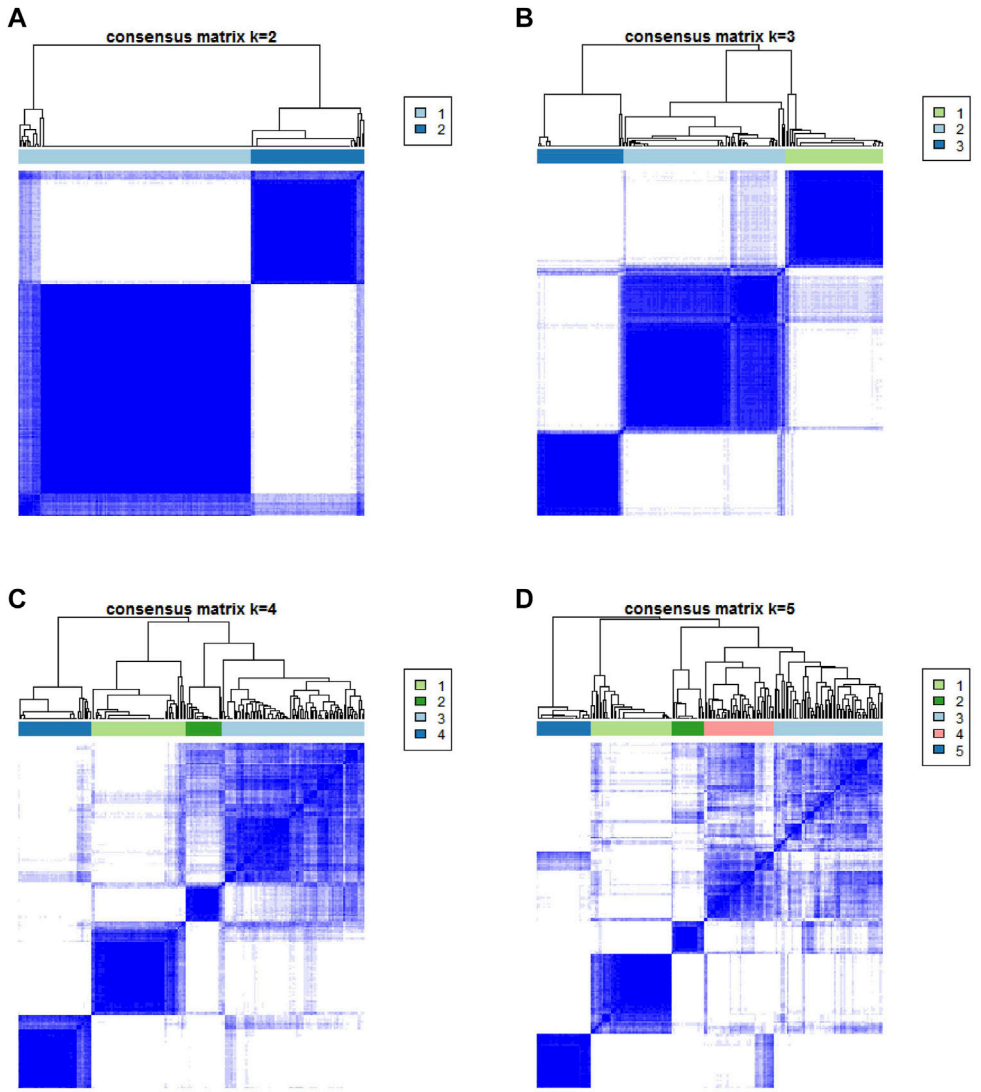


FIGURE 2 | Based on the proportion of infiltrating immune cells, consensus matrixes of all BLCA samples. **(A–D)** In consensus matrixes of all BLCA samples for each k (k = 2–5), the consensus matrix was the crispest when K = 2.

patients into two distinct ICI subtypes. The consensus matrix was the crispest when K = 2 (**Figures 2A–D**), namely, ICI clusters A and B (**Supplementary Table 3**). The heatmap enables visualization of the expression levels of immune cells of distinct ICI clusters (**Figure 3B**). Moreover, two independent ICI subtypes showed a significant difference in the overall survival rate (p = 0.002; **Figure 3A**).

To further describe and understand the biological and clinical distinctions among these inherent features, we analyzed the immune cell composition of two ICI subtypes. Between two ICI subtypes, ICI cluster A had a better outcome with a median duration of roughly five years. Meanwhile, it was marked by increased infiltration of CD8 T cells, activated memory CD4 T cells, resting mast cells, etc. In addition, the ICI cluster A has a higher immune score than the ICI cluster B. On the contrary, the ICI cluster B confirmed a poor prognosis (median survival duration of roughly three years) and showed a large rise in the number of macrophages M0 (**Figure 3D**).

Identified the Subtypes of Immune-Related Gene

In order to reveal the underlying biological properties of distinct immunophenotypes, we used the R package "limma" to carry out differential analysis to identify the transcriptome differences between two subtypes. Unsupervised clustering was implemented using the "limma" package of R software to obtain the differentially expressed genes (DEGs) (**Supplementary Table 4**). We classified BLCA patients into gene clusters A–B by DEGs (**Figures 4A–D**; **Supplementary Table 5**). The positive correlation of DEG values with the clusters signature was coded as ICI gene signature A, and the

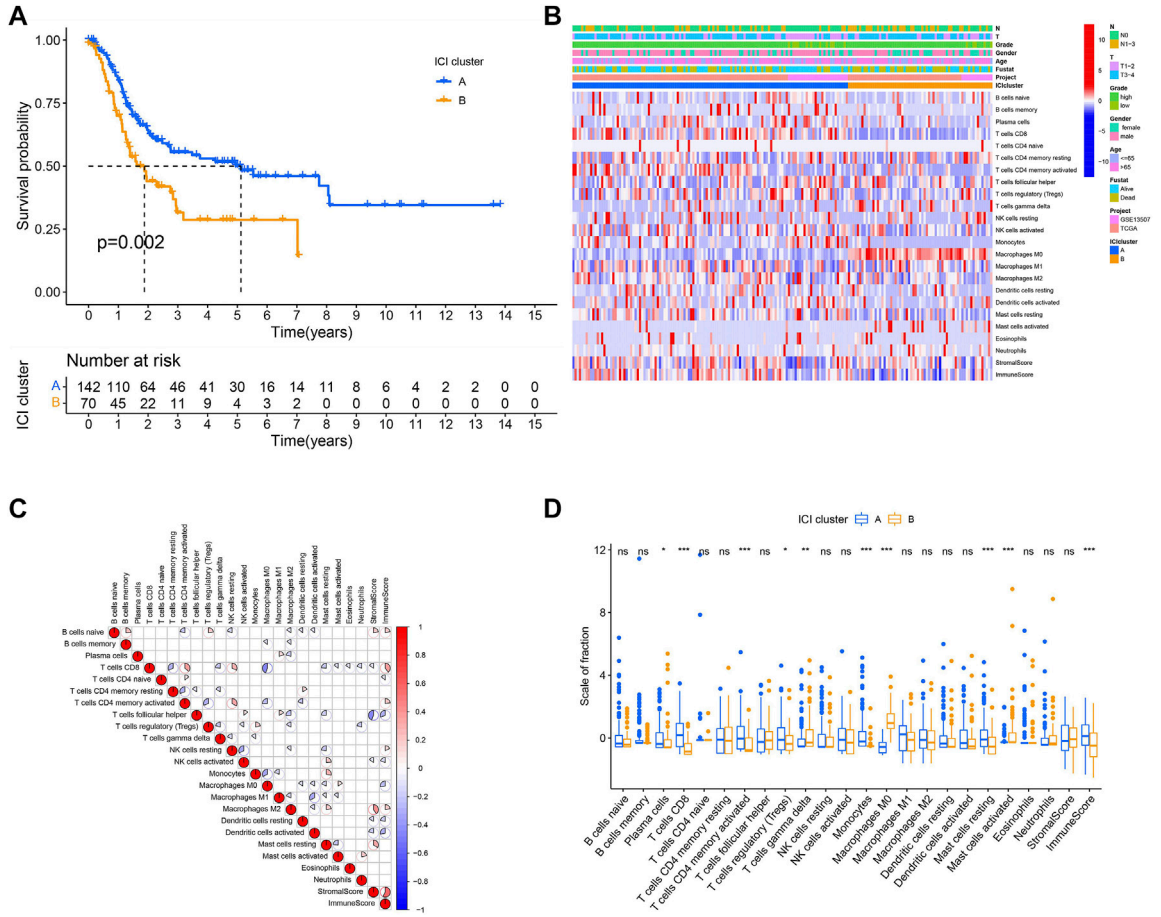


FIGURE 3 | The pattern of ICI in the TME of BLCA. **(A)** Survival time of patients in two independent ICI subtypes. **(B)** Visualized the expression levels of immune cells of distinct ICI clusters. **(C)** The correlation among different immune cells. **(D)** The proportion of tumor-infiltrating immune cells in two ICI clusters. We also depicted the immune score and stromal score of two ICI clusters. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

rest of the DEGs were coded as ICI gene signature B. At the same time, to reduce noise or redundant genes, we employed the "Boruta" method to reduce the dimension of gene signatures A and B. The transcriptome properties of DEGs are shown in a heatmap created with the R package "pheatmap (**Figure 5A**)." The R package "clusterProfiler" was applied to execute GO enrichment analysis on the signature genes. The significantly enriched biological processes are summarized in **Figures 5B–E**, and a detailed description is provided in **Supplementary Table 6**.

Next, we evaluated the prognosis of gene clusters A–B combined with survival information. It showed that two independent gene clusters had a significant difference in overall survival ($p < 0.001$; **Figure 5F**). The gene cluster A was characterized by a better outcome (median survival duration of roughly 5.5 years), whereas the gene cluster B had unfavorable outcomes (median survival duration of roughly two years). As displayed in **Figure 5G**, the gene cluster A showed an obvious increase in the infiltration of some immune cells, such as CD8 T cells and naive B cells. Meanwhile, the gene cluster B performed a higher macrophages M0 infiltration. Finally, some differentially expressed target genes were analyzed in two gene clusters by the "limma" package.

Generation of ICI Score

In order to obtain the quantitative index of the ICI landscape of BLCA patients, we computed two aggregate scores by the PCA algorithm: the ICI score A from ICI signature gene A and the ICI score B from ICI signature gene B. We obtained the sum of individual scores using ICI scores A and B of each sample in the study. Finally, we obtained the ICI score, which is a predictive signature score. The TCGA-BLCA and GSE13507 patients were separated into high and low ICI score groups using the "survival" package (**Supplementary Table 7**). The alluvial diagram described the correlation among the gene clusters, the ICI score, and survival outcomes (**Figure 6A**). CD274, CTLA4, HAVCR2, LAG3, and PDCD1 were chosen as immune-checkpoint–relevant signatures, and CD8A, CXCL10, CXCL9, GZMA, GZMB, IFNG, PRF1, TBX2, and TNF were selected as immune-activity–related signatures, to investigate the

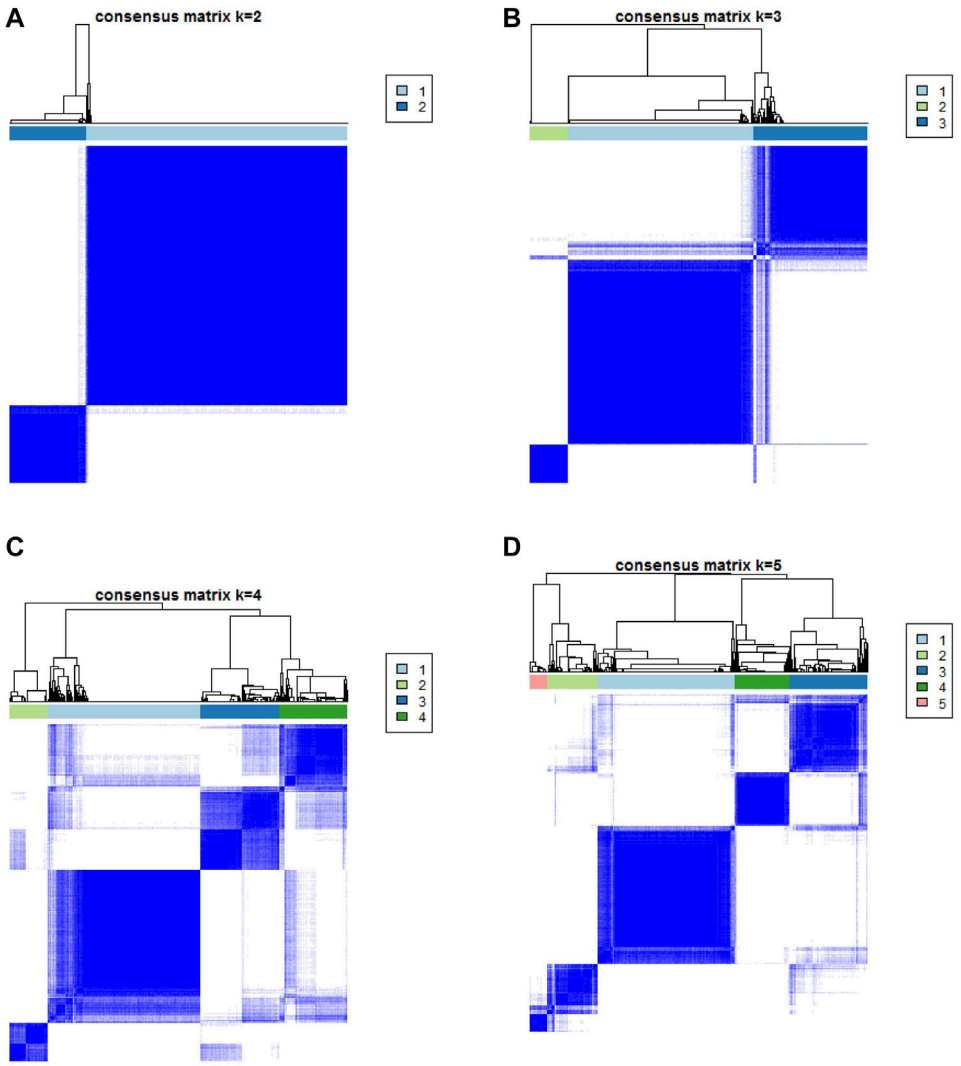


FIGURE 4 | Based on DEGs, consensus matrixes of all BLCA samples. **(A–D)** In consensus matrixes of all BLCA samples for each k (k = 2–5), the consensus matrix was the crispest when K = 2.

immunological activation and tolerant state of the TCGA-BLAC and GSE13507 cohorts. With the exception of TBX2, the ICI score was shown to have a substantial negative correlation with the expression quantity of immune-checkpoint–relevant and immune-activity–relevant genes (**Figure 6B**). Moreover, GSEA analysis results revealed that fatty acid metabolism and PPAR signaling pathways were considerably enriched in the high ICI score group, whereas proteasome and NOD-like receptor signaling pathways were substantially enriched in the low score one (**Figure 6C**). Detailed enriched information and description is provided in **Supplementary Table 8**. Next, we evaluated the impact of the ICI score on the prognosis of patients. It showed that two independent ICI score groups had remarkable difference in the overall survival rate ($p < 0.001$; **Figure 6D**). The high ICI score group showed a good prognosis (median survival duration of roughly 5.3 years), whereas the low one had an unfavorable outcome (median survival duration of roughly 1.2 years). Lastly, the "ggplot2" package was applied to evaluate the relation between ICI score and survival status. We found that two independent ICI score groups had a significant difference in survival status. The majority of BLCA in the high ICI score group were alive; on the contrary, the majority of BLCA in the low one was dead ($p = 0.0059$; **Figures 6E,F**).

TMB and ICI Score Were Applied to Evaluate the Prognosis of TCGA-BCLA Cohort Patients

Since BLCA was reported to have a high degree of somatic changes, subsequently, we determined the distribution of somatic mutations and combined it with the ICI score to evaluate the prognosis of patients. First, the total mutation

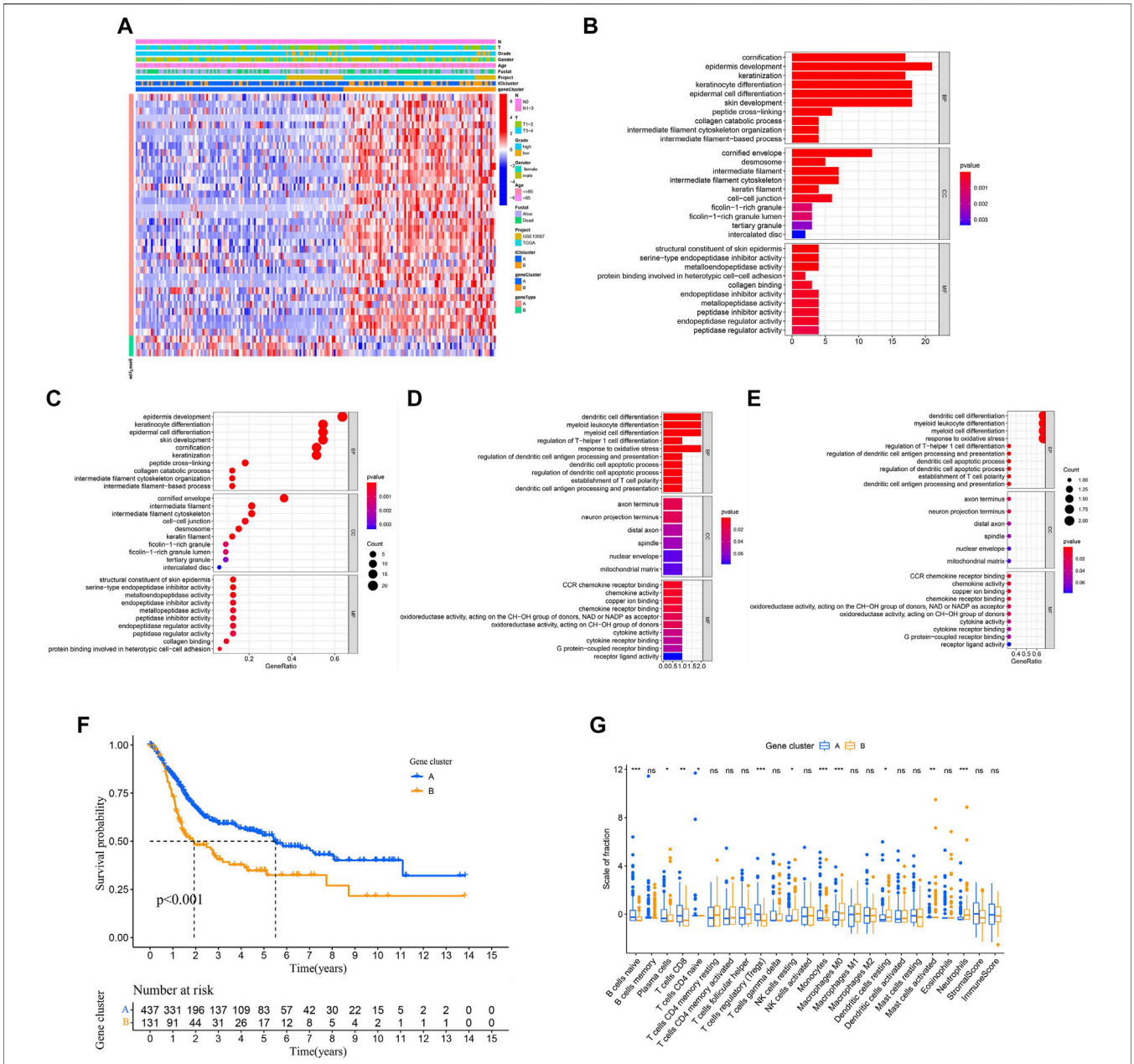


FIGURE 5 | Identification of the subtype of the immune gene. **(A)** Description of the transcriptomic profile of DEGs identified across the genomic clusters. **(B–E)** Functional and pathway enrichment analyses (Gene Ontology-Biological process) of ICI gene signatures A and B: ICI signature genes **A (B–C)** and **B (D–E)**. **(F)** Two independent gene clusters had a significant difference in the overall survival rate (log-rank test, $p < 0.001$). **(G)** The proportion of tumor-infiltrating immune cells in the two gene clusters. We also depicted the immune score and stromal score of two gene clusters. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

burden and mutation distribution of TCGA-BCLA were obtained by analyzing mutation annotation files. Meanwhile, we divided the patients into high- and low-TMB groups. As demonstrated in **Figure 7A**, we discovered that the high-TMB group was related to a better outcome than the low one ($p < 0.001$). Considering the contraindication value of TMB and ICI score for prognosis, we subsequently studied the synergistic effect of ICI score in the prognostic classification of BLCA. The results reveal that there was a substantial difference in survival between the high- and low-TMB groups depending on ICI score subtypes. Among them, the high TMB combined with a high ICI score had the best prognosis in BLCA ($p < 0.001$; **Figure 7B**). In conclusion, the ICI score could be utilized as a possible predictor different from TMB, which could effectively predict the response to immunotherapy. In addition, we screened out 20 driver genes with the highest mutation frequency for further

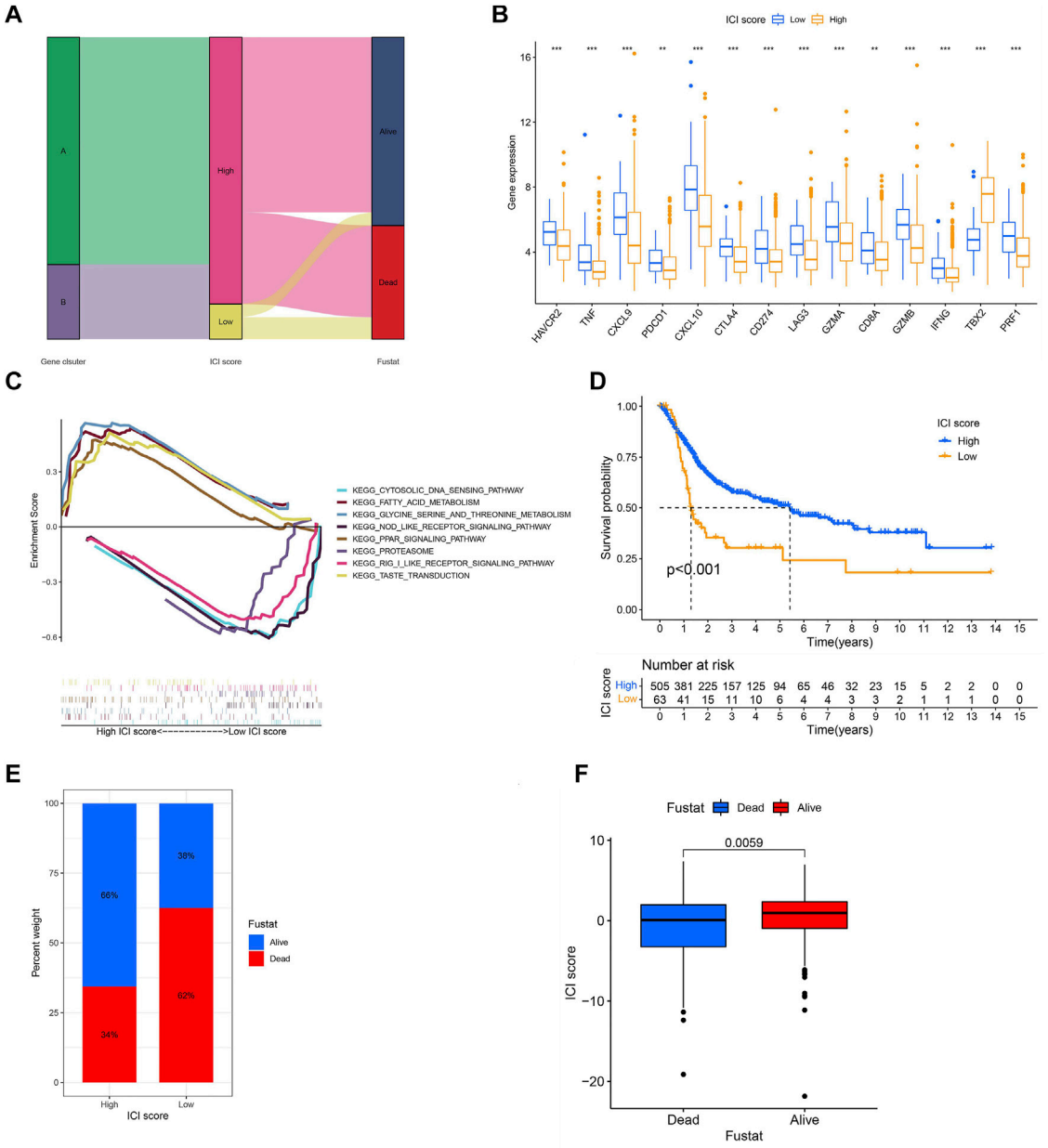


FIGURE 6 | Generation of ICI score. **(A)** The alluvial diagram described the correlation among the gene clusters, the ICI score, and survival outcomes. **(B)** The expression levels of immune-checkpoint–relevant signatures (CD274, CTLA4, HAVCR2, LAG3, and PDCD1) and immune-activity–related signatures (CD8A, CXCL10, CXCL9, GZMA, GZMB, IFNG, PRF1, TBX2, and TNF) in high and low ICI score groups. *$p < 0.05$; **$p < 0.01$, ***$p < 0.001$; **(C)** GSEA analysis results exhibited that some significantly enriched functions or pathways in high and low ICI score groups. **(D)** Survival time of patients in high and low ICI score groups (log-rank test, $p < 0.0010$. **(E,F)** Most of the patients in the high ICI score group were alive; on the contrary, most of the patients in the low ICI score group were dead (log-rank test, $p = 0.0059$).

analysis. We analyzed the distribution of driver genes in high and low ICI score groups. The result showed that the alteration frequency of TP53, KMT2D, PIK3CA, KMT2C, and FLG was considerably different between the two ICI score groups (**Figures 7C,D**). Moreover, we discovered that the TP53 mutation frequency was higher in the low ICI score group. The result proved once again that the group with the low ICI had a poor prognosis. These results may provide new ideas for the study of the mechanism of ICI in tumors (Şenbabaoğlu et al., 2016; Chen and Mellman, 2017).

Analysis of Clinical Features in Two ICI Score Groups

To be able to clarify the role of ICI score in BLCA, the relationship between ICI score and clinical characteristics was researched. The

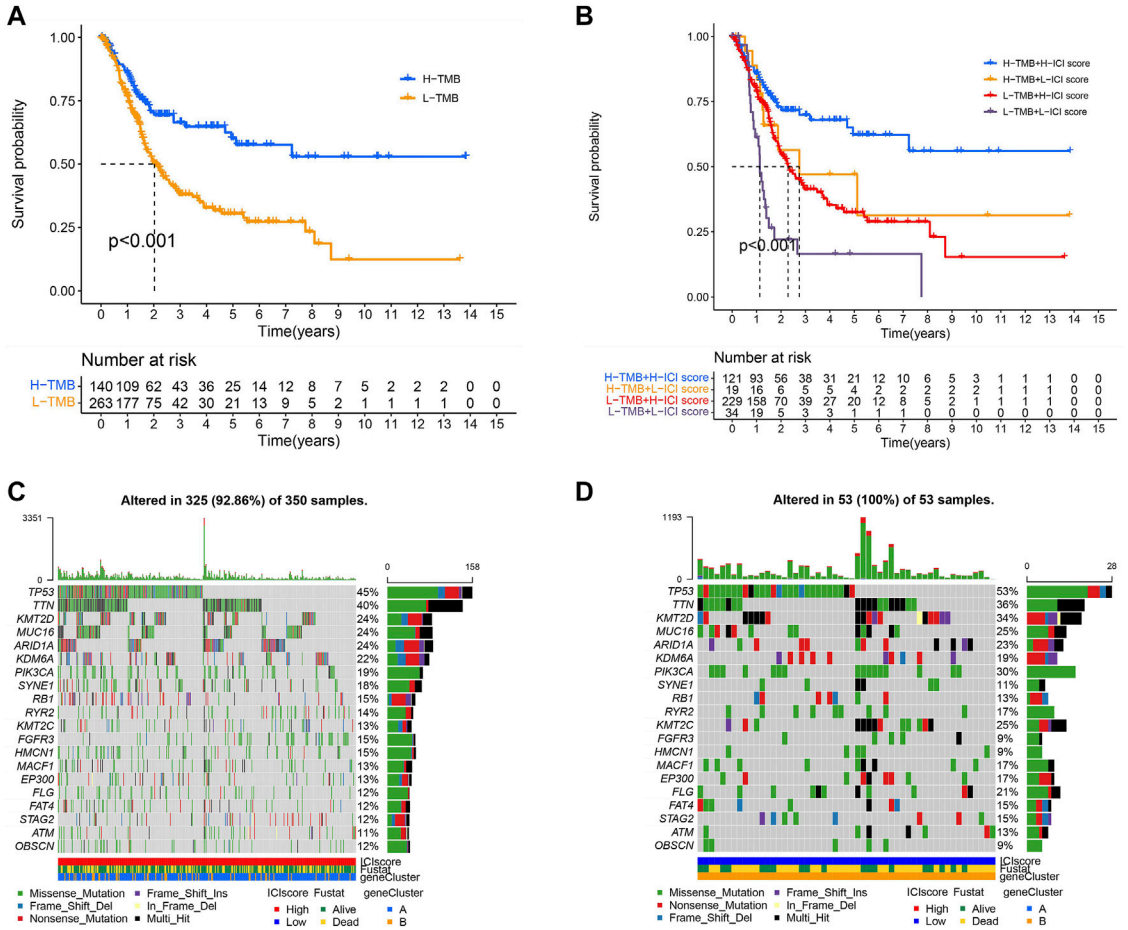


FIGURE 7 | TMB and ICI scores were used to evaluate the prognosis of TOGA-BOLA cohort patients. **(A)** Survival time of patients in high- and low-TMB groups (log-rank test, $p < 0.0010$. **(B)** The high-TMB group combined with the high ICI score group showed the best prognosis in BLCA patients (log-rank test, $p < 0.001$; **(C,D)** The distribution of driver genes in high and low ICI score groups: high ICI score group **(C)** and low ICI score group **(D)**.

stratified survival analysis was used to observe whether ICI scores could be applied to different clinicopathological features. Next, we analyzed patients' age and gender. Results showed that the ICI score could effectively forecast OS in all groups from the age and gender clinical features (**Figures 8A–D**).

DISCUSSION

At present, radical resection is still the main treatment method for localized BLCA, followed by intracavitary chemotherapy or immunotherapy (Cao et al., 2020). However, BLCA has the characteristic of a high recurrence rate and low survival rate (Cambier et al., 2016; Kamat et al., 2016; Sanli et al., 2017). Despite the fact that ICIs are effective against advanced urothelial malignancies, including BLCA, tumor reaction to ICIs is often poor and difficult to predict (Zou et al., 2016; Dermani et al., 2019). Besides, TMB is considered an important marker for predicting ICI response in a variety of tumor types. Nevertheless, the boundary between high and low TMB is yet to be properly defined (Samstein et al., 2019). As a result, finding a novel prognostic marker is critical. Instead of tumor cells, more and more attention has been paid to immune cell infiltration (ICI) recently. In this research, we combined TCGA-BLCA and GSE13507 to get common gene expression data, which contain 569 BLCA patient samples. Subsequently, based on the unsupervised clustering method, a total of 569 patient samples were divided into two different immune subtypes according to the proportion of ICI. Consensus clustering has been widely used in genome research (Şenbabaoğlu et al., 2014). Based on the DEGs between ICI cluster A and ICI cluster B, we classified the BLCA patients into two genomic clusters, namely, gene cluster A and gene cluster B, respectively. Anti-tumor cells and pro-tumor cells are two kinds of immune cells engaged in cancer local immune response (Yu and Ho, 2019; Riera -Domingo et al., 2020). Different immune cells may play different roles in different tumors (Foley et al., 2016). Our analysis results showed that the expression levels of CD8 T cells and naive B cells were up-regulated in gene cluster A, indicating a good outcome. Meanwhile, macrophage M0 was shown to be positively correlated with

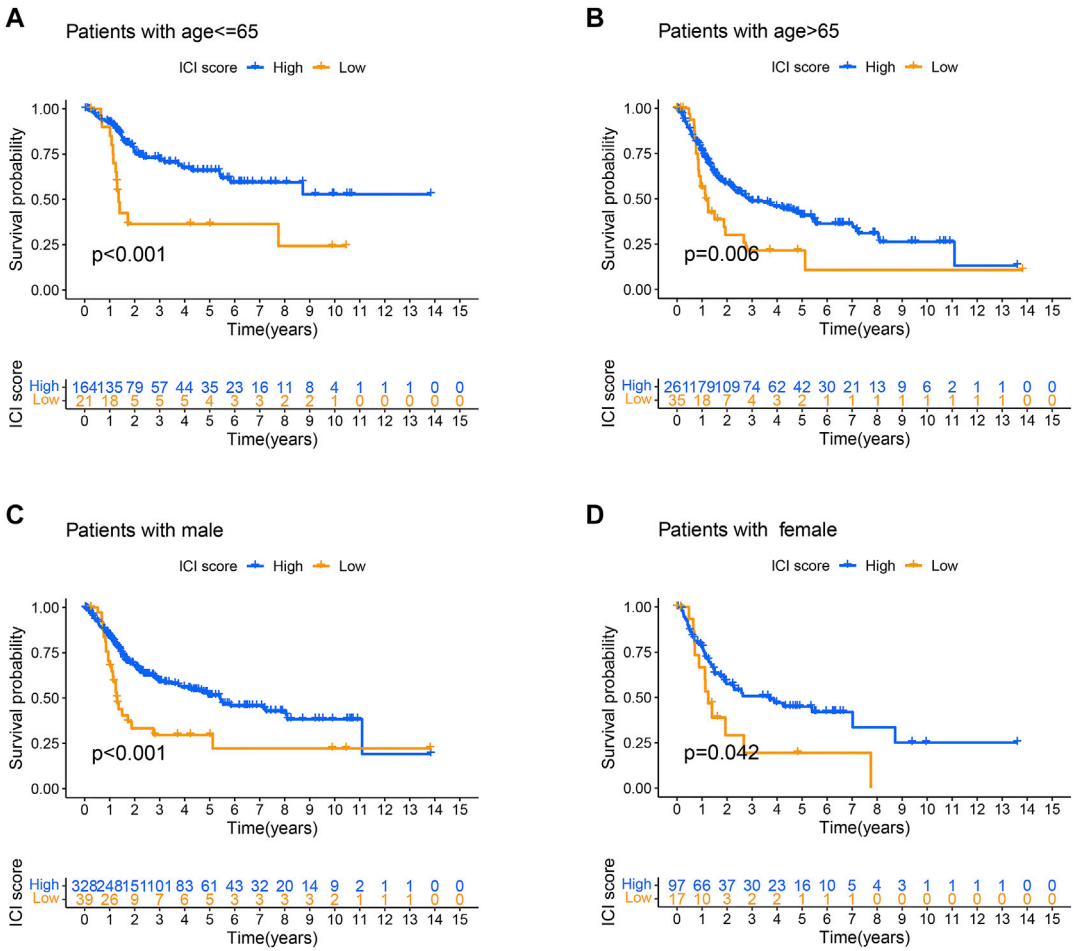


FIGURE 8 | Analysis of clinical features in high and low ICI score groups. **(A–D)** Survival time of patients with high and low ICI score groups in distinct clinical features: **(A)** age ≤ 65; **(B)** age > 65; **(C)** male patients; **(D)** female patients.

gene cluster B, which showed a poor prognosis. We obtained two gene signatures by different expression levels of DEGs in different gene clusters. Considering the individual heterogeneity of the immune environment, it is urgent to quantify the ICI model for individual tumors (Callari et al., 2016). In some cancers, individual-based models have been fully established to improve outcome forecasting (Bramsen et al., 2017; Komor et al., 2018). In this study, the PCA algorithm was used to separate the TCGA-BLCA and GSE13507 cohorts into two ICI score groups. The high score group has a better prognosis than the low one. Through GSEA, we found that the genes implicated in the immune activation pathway, such as fatty acid metabolism and PPAR signaling pathways, were significantly abundant in the high ICI score group. In various ICI score groups, we analyzed the levels of immune activation–related signal and immune-checkpoint–related signal. In the low ICI score group, the expression levels of CD274, CTLA4, HAVCR2, LAG3, PDCD1, CD8A, CXCL10, CXCL9, GZMA, GZMB, IFNG, PRF1, and TNF were up-regulated, except TBX2. Besides, we explored the mutation frequency of some driving genes in different ICI score groups. The results revealed that in the low ICI score group, the frequency of TP53 mutations was increased. All of these studies revealed that the ICI score was adversely linked with tumor malignancy from different perspectives. Since the neoantigen load could be easily detected and evaluated by TMB, it has been proved to be an indicator of clinical benefit and a prognostic factor for predicting ICI response. Our analysis shows that high TMB has better OS performance in BLCA. Finally, we found that the high TMB combined with the high ICI score has a higher survival rate than others. ICI score can effectively predict the OS of age and gender groups. However, all results of this study were obtained retrospectively based on public databases, which requires further prospective validation.

CONCLUSION

We comprehensively analyzed the BLCA ICI landscape, providing a clear picture of the anti-/pro-tumor immune response regulation in BLCA. The variation of ICI patterns is related to tumor heterogeneity. As a result, this discovery has significant clinical implications for the systematic evaluation of tumor ICI patterns. Our results revealed that the ICI score could be served as an efficient

predictive marker which is different from TMB. These findings may provide a new method to predict the prognosis of BLCA.

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

YZ conceived the article, conducted statistical analysis, and wrote the manuscript. YW conceived and reviewed the article. JW reviewed the article and data analysis. KZ collected samples and revised the manuscript.

FUNDING

This paper is supported by the National Natural Science Foundation of China. (Project Number: U19042).

ACKNOWLEDGMENTS

Thanks to the National Natural Science Foundation of China (U19042) for supporting this research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.852708/full#supplementary-material

REFERENCES

Babjuk, M., Böhle, A., Burger, M., Capoun, O., Cohen, D., Compérat, E. M., et al. (2017). EAU Guidelines on Non-muscle-invasive Urothelial Carcinoma of the Bladder: Update 2016. *Eur. Urol.* 71 (3), 447–461. doi:10.1016/j.eururo.2016.05.041

Bramsen, J. B., Rasmussen, M. H., Ongen, H., Mattesen, T. B., Ørntoft, M.-B. W., Árnadóttir, S. S., et al. (2017). Molecular-Subtype-Specific Biomarkers Improve Prediction of Prognosis in Colorectal Cancer. *Cell Rep.* 19 (6), 1268–1280. doi:10.1016/j.celrep.2017.04.045

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 Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492

Callari, M., Cappelletti, V., D'Aiuto, F., Musella, V., Lembo, A., Petel, F., et al. (2016). Subtype-Specific Metagene-Based Prediction of Outcome after Neoadjuvant and Adjuvant Treatment in Breast Cancer. *Clin. Cancer Res.* 22 (2), 337–345. doi:10.1158/1078-0432.ccr-15-0757

Cambier, S., Sylvester, R. J., Collette, L., Gontero, P., Brausi, M. A., van Andel, G., et al. (2016). EORTC Nomograms and Risk Groups for Predicting Recurrence, Progression, and Disease-specific and Overall Survival in Non-muscle-invasive Stage Ta-T1 Urothelial Bladder Cancer Patients Treated with 1-3 Years of Maintenance Bacillus Calmette-Guérin. *Eur. Urol.* 69 (6), e123–4. doi:10.1016/j.eururo.2015.06.045

Cao, R., Yuan, L., Ma, B., Wang, G., and Tian, Y. (2020). Immune-related Long Non-coding RNA Signature Identified Prognosis and Immunotherapeutic Efficiency in Bladder Cancer (BLCA). *Cancer Cell Int.* 20, 276. doi:10.1186/s12935-020-01362-0

Chen, D. S., and Mellman, I. (2017). Elements of Cancer Immunity and the Cancer-Immune Set Point. *Nature* 541 (7637), 321–330. doi:10.1038/nature21349

Chen, Y.-P., Wang, Y.-Q., Lv, J.-W., Li, Y.-Q., Chua, M. L. K., Le, Q.-T., et al. (2019). Identification and Validation of Novel Microenvironment-Based Immune Molecular Subgroups of Head and Neck Squamous Cell Carcinoma: Implications for Immunotherapy. *Ann. Oncol.* 30 (1), 68–75. doi:10.1093/annonc/mdy470

D'Aiello, A., Lin, J., Gucalp, R., Tabatabaie, V., Cheng, H., Bloomgarden, N. A., et al. (2021). Thyroid Dysfunction in Lung Cancer Patients Treated with Immune Checkpoint Inhibitors (ICIs): Outcomes in a Multiethnic Urban Cohort. *Cancers (Basel)* 13 (6). doi:10.3390/cancers13061464

Dermani, F. K., Samadi, P., Rahmani, G., Kohlan, A. K., and Najafi, R. (2019). PD-1/PD-L1 Immune Checkpoint: Potential Target for Cancer Therapy. *J. Cell Physiol.* 234 (2), 1313–1325. doi:10.1002/jcp.27172

Foley, K., Kim, V., Jaffee, E., and Zheng, L. (2016). Current Progress in Immunotherapy for Pancreatic Cancer. *Cancer Lett.* 381 (1), 244–251. doi:10.1016/j.canlet.2015.12.020

Hellmann, M. D., Callahan, M. K., Awad, M. M., Calvo, E., Ascierto, P. A., Atmaca, A., et al. (2019). Tumor Mutational Burden and Efficacy of Nivolumab Monotherapy and in Combination with Ipilimumab in Small-Cell Lung Cancer. *Cancer cell* 35 (2), 329. doi:10.1016/j.ccell.2019.01.011

Jiang, Y., Zhang, Q., Hu, Y., Li, T., Yu, J., Zhao, L., et al. (2018). ImmunoScore Signature. *Ann. Surg.* 267 (3), 504–513. doi:10.1097/sla.0000000000002116

Johnson, D. B., Sullivan, R. J., and Menzies, A. M. (2017). Immune Checkpoint Inhibitors in Challenging Populations. *Cancer* 123 (11), 1904–1911. doi:10.1002/cncr.30642

Kamat, A. M., Hahn, N. M., Efstathiou, J. A., Lerner, S. P., Malmström, P.-U., Choi, W., et al. (2016). Bladder Cancer. *Lancet* 388, 2796–2810. doi:10.1016/s0140-6736(16)30512-8

Knowles, M. A., and Hurst, C. D. (2015). Molecular Biology of Bladder Cancer: New Insights into Pathogenesis and Clinical Diversity. *Nat. Rev. Cancer* 15 (1), 25–41. doi:10.1038/nrc3817

Komor, M. A., Bosch, L. J., Bounova, G., Bolijn, A. S., Delis-van Diemen, P. M., Rausch, C., et al. (2018). Consensus Molecular Subtype Classification of Colorectal Adenomas. *J. Pathol.* 246 (3), 266–276. doi:10.1002/path.5129

Li, P., Cao, J., Li, J., Yao, Z., Han, D., Ying, L., et al. (2020). Identification of Prognostic Biomarkers Associated with Stromal Cell Infiltration in Muscle-invasive Bladder Cancer by Bioinformatics Analyses. *Cancer Med.* 9 (19), 7253–7267. doi:10.1002/cam4.3372

Liang, B., Tao, Y., and Wang, T. (2020). Profiles of Immune Cell Infiltration in Head and Neck Squamous Carcinoma. *Biosci. Rep.* 40 (2), BSR20192724. doi:10.1042/BSR20192724

Na, K. J., and Choi, H. (2018). Immune Landscape of Papillary Thyroid Cancer and Immunotherapeutic Implications. *Endocrine-related cancer* 25 (5), 523–531. doi:10.1530/erc-17-0532

Noy, R., and Pollard, J. W. (2014). Tumor-associated Macrophages: from Mechanisms to Therapy. *Immunity* 41 (1), 49–61. doi:10.1016/j.immuni.2014.06.010

Riera-Domingo, C., Audigé, A., Granja, S., Cheng, W.-C., Ho, P.-C., Baltazar, F., et al. (2020). Immunity, Hypoxia, and Metabolism-the Ménage à Trois of Cancer: Implications for Immunotherapy. *Physiol. Rev.* 100 (1), 1–102. doi:10.1152/physrev.00018.2019

Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., et al. (2015). Cancer Immunology. Mutational Landscape Determines Sensitivity to PD-1 Blockade in Non-small Cell Lung Cancer. *Science* 348 (6230), 124–128. doi:10.1126/science.aaa1348

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 (2), 202–206. doi:10.1038/s41588-018-0312-8

Sanli, O., Dobruch, J., Knowles, M. A., Burger, M., Alemozaffar, M., Nielsen, M. E., et al. (2017). Bladder Cancer. *Nat. Rev. Dis. Prim.* 3, 17022. doi:10.1038/nrdp.2017.22

Şenbabaoğlu, Y., Gejman, R. S., Winer, A. G., Liu, M., Van Allen, E. M., de Velasco, G., et al. (2016). Tumor Immune Microenvironment Characterization in Clear Cell Renal Cell Carcinoma Identifies Prognostic and Immunotherapeutically Relevant Messenger RNA Signatures. *Genome Biol.* 17 (1), 231. doi:10.1186/s13059-016-1092-z

Şenbabaoğlu, Y., Michailidis, G., and Li, J. Z. (2014). Critical Limitations of Consensus Clustering in Class Discovery. *Sci. Rep.* 4, 6207. doi:10.1038/srep06207

Tøndell, A., Subbannayya, Y., Wahl, S. G. F., Flatberg, A., Sørhaug, S., Børset, M., et al. (2021). Analysis of Intra-tumoral Macrophages and T Cells in Non-small Cell Lung Cancer (NSCLC) Indicates a Role for Immune Checkpoint and CD200-Cd200r Interactions. *Cancers* 13 (8), 1788. doi:10.3390/cancers13081788

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

Vassilakopoulou, M., Avgeris, M., Velcheti, V., Kotoula, V., Rampias, T., Chatzopoulos, K., et al. (2016). Evaluation of PD-L1 Expression and Associated Tumor-Infiltrating Lymphocytes in Laryngeal Squamous Cell Carcinoma. *Clin. Cancer Res.* 22 (3), 704–713. doi:10.1158/1078-0432.ccr-15-1543

Yu, Y.-R., and Ho, P.-C. (2019). Sculpting Tumor Microenvironment with Immune System: from Immunometabolism to Immunoediting. *Clin. Exp. Immunol.* 197 (2), 153–160. doi:10.1111/cei.13293

Zeng, D., Li, M., Zhou, R., Zhang, J., Sun, H., Shi, M., et al. (2019). Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. *Cancer Immunol. Res.* 7 (5), 737–750. doi:10.1158/2326-6066.cir-18-0436

Zeng, D., Zhou, R., Yu, Y., Luo, Y., Zhang, J., Sun, H., et al. (2018). Gene Expression Profiles for a Prognostic Immunoscore in Gastric Cancer. *Br. J. Surg.* 105 (10), 1338–1348. doi:10.1002/bjs.10871

Zhao, D.-Y., Sun, X.-Z., and Yao, S.-K. (2021). Mining the Cancer Genome Atlas Database for Tumor Mutation Burden and its Clinical Implications in Gastric Cancer. *Wjgo* 13 (1), 37–57. doi:10.4251/wjgo.v13.i1.37

Zou, W., Wolchok, J. D., and Chen, L. (2016). PD-L1 (B7-H1) and PD-1 Pathway Blockade for Cancer Therapy: Mechanisms, Response Biomarkers, and Combinations. *Sci. Transl. Med.* 8 (328), 328rv4. doi:10.1126/scitranslmed.aad7118

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 Zhang, Wang, 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

ORIGINAL RESEARCH
published: 29 June 2022
doi: 10.3389/fgene.2022.905051

Analysis of Pyroptosis-Related Immune Signatures and Identification of Pyroptosis-Related LncRNA Prognostic Signature in Clear Cell Renal Cell Carcinoma

Ming Zhong[1†], Xiaohua Wang[1†], Enyi Zhu[1†], Lian Gong[2], Lingyan Fei[1], Liang Zhao[3], Keping Wu[1], Chun Tang[1], Lizhen Zhang[4], Zhongli Wang[5] and Zhihua Zheng[1]**

[1]Department of Nephrology, Center of Kidney and Urology, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, China, [2]Department of Oncology, Third Xiangya Hospital, Central South University, Changsha, China, [3]National Clinical Research Center for Child Health, National Children's Regional Medical Center, The Children's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [4]Department of Urology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, [5]Department of Internal Medicine and Geriatrics, Zhongnan Hospital, Wuhan University School of Medicine, Wuhan, China

OPEN ACCESS

Edited by:
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Rajesh Kumar,
National Institutes of Health (NIH), United States
Wei Yang,
Cedars Sinai Medical Center, United States

***Correspondence:**
Zhongli Wang
1095854403@qq.com
Zhihua Zheng
zhzhihua@mail.sysu.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: 26 March 2022
Accepted: 16 May 2022
Published: 29 June 2022

Citation:
Zhong M, Wang X, Zhu E, Gong L, Fei L, Zhao L, Wu K, Tang C, Zhang L, Wang Z and Zheng Z (2022) Analysis of Pyroptosis-Related Immune Signatures and Identification of Pyroptosis-Related LncRNA Prognostic Signature in Clear Cell Renal Cell Carcinoma. Front. Genet. 13:905051. doi: 10.3389/fgene.2022.905051

Clear cell renal cell carcinoma (ccRCC) is a common urinary system malignant tumor with a high incidence and recurrence rate. Pyroptosis is a kind of programmed cell death caused by inflammasomes. More and more evidence had confirmed that pyroptosis plays a very significant part in cancer, and it is controversial whether pyroptosis promotes or inhibits tumors. Consistently, its potential role in ccRCC treatment efficacy and prognosis remains unclear. In this study, we systematically investigated the role of pyroptosis in the ccRCC samples from The Cancer Genome Atlas (TCGA) database. Based on the differentially expressed pyroptosis-related genes (DEPRGs), we identified three pyroptosis subtypes with different clinical outcomes, immune signatures, and responses to immunotherapy. Gene set variation analysis (GSVA), Gene Ontology (GO) analysis, and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis revealed that pyroptosis activation meant infiltration of more immune cells that is conducive to tumor progression. To further investigate the immunomodulatory effect of pyroptosis in ccRCC, we constructed a pyroptosis-score based on the common differential prognostic genes of the three pyroptosis subtypes. It was found that patients with high pyroptosis-score were in an unfavorable immune environment and the prognosis was worse. Gene set enrichment analysis suggested that immune-related biological processes were activated in the high pyroptosis-score group. Then, the least absolute shrinkage and selection operator (LASSO) Cox regression was implemented for constructing a prognostic model of eight pyroptosis-related long noncoding RNAs (PRlncRNAs) in the TCGA dataset, and the outcomes revealed that, compared with the low-risk group, the model-based high-risk group was intently associated with poor overall survival (OS). We further explored the relationship between high- and low-risk groups with tumor microenvironment (TME), immune infiltration, and drug therapy. Finally, we constructed and confirmed a robust and reliable PRlncRNA pairs prediction model of ccRCC, identified PRlncRNA, and verified it by experiments. Our findings suggested the potential role of pyroptosis in ccRCC,

offering new insights into the prognosis of ccRCC and guiding effectual targeted therapy and immunotherapy.

Keywords: clear cell renal cell carcinoma, pyroptosis, tumor microenvironment, long noncoding RNA, immune

INTRODUCTION

Renal cell carcinoma (RCC) is derived from renal tubular epithelium, accounting for 80%–90% of renal malignant tumors. The incidence of RCC ranks third in urinary system tumors (Sung et al., 2021). Clear cell RCC (ccRCC) is the most common histopathological type of RCC, accounting for about 60%–85% of RCC. Currently, the treatments of ccRCC are multidisciplinary comprehensive treatments including surgery, molecular targeted therapy, immunotherapy, chemotherapy, and radiotherapy. However, ccRCC patients often experience postoperative recurrence, insensitivity to medical treatment, or drug resistance after treatment, leading to poor prognosis, which is related to multiple factors, such as the excessive proliferation of malignant tumor cells and inhibition of cell death. Regulated cell death (RCD) is the defense mechanism against cancer, and it is also a way to drive tumorigenesis, including apoptosis, entosis, necroptosis, pyroptosis, and ferroptosis (Koren and Fuchs, 2021).

Pyroptosis represents an interesting modality of regulated necrosis and is a kind of programmed cell death caused by inflammasomes, which are manifested by the incessant cells swelling until the cells break, leading to the release of cellular contents and intense inflammation. The inflammasome is a key substance in the pyroptosis process. Under the stimulation of pathogens or lipopolysaccharides, it can promote the maturation of the precursor of IL-18 and IL-1β and trigger the pyroptosis process by activating Caspase-1 (Tang et al., 2020; Liu X. et al., 2021b). Current studies have proved that inflammasomes are present in a variety of tumor cells (Dunn et al., 2012). Inflammasome-associated proteins can promote or inhibit the growth of tumor cells in different tumor cells, showing the heterogeneousness of cancer and the complicacy of the immunity microenvironment. Pyroptosis can restrain the proliferation of tumor cells (Nagarajan et al., 2019) and also promote tumor growth by forming an environment suitable for tumor cell growth (Gao et al., 2018). Although there are studies (Jiang et al., 2021; Tang et al., 2021) exploring the potential role of pyroptosis in ccRCC, the data were not well optimized and therefore could not fully explain the effect of pyroptosis on ccRCC. Our study will optimize the data to investigate the influence of pyroptosis on the tumor microenvironment (TME) and drug sensitivity of ccRCC.

Long noncoding RNAs (lncRNAs) are RNA transcripts longer than 200 nucleotides, including natural antisense transcripts, overlapping transcripts, and intronic transcripts. Many studies have shown that lncRNAs have diverse phenotypes and mechanisms by regulating cell proliferation, replication, angiogenesis, cell death, and metastasis (Liu S. J. et al., 2021). LncRNAs have been proved to regulate the pyroptosis of tumor cells. For example, lncRNA-XIST can inhibit pyroptosis to promote non-small cell lung cancer (Liu et al., 2019) and lncRNA ADAMTS9-AS2 restrains gastric cancer and increases the drug sensitivity of cisplatin by promoting pyroptosis (Ren et al., 2020). Pyroptosis-related lncRNA (PRlncRNA) signature has been established as a model to predict the treatment effect and prognoses of various cancers (Hong et al., 2020; Bu et al., 2021; Lv et al., 2021; Ping et al., 2021; Song S. et al., 2021; Tang et al., 2021). However, as a result of the differences in data processing, it is impossible to directly compare the difference in absolute expression levels of lncRNAs among different data sets. Consequently, it is necessary to properly normalize and standardize the expression levels of lncRNAs. Fortunately, the researchers found a way to normalize based on the relative ranking of lncRNAs. For instance, an example of applying these methods was long noncoding RNA pairs (lncRNAPs), which have proved to be reliable (Hong et al., 2020; Song S. et al., 2021).

Given the above, we optimized the data to deeply survey the role of pyroptosis in the TME and targeted therapy effect of ccRCC and integrated a PRlncRNAPs prognostic model that eliminates differences in data processing to forecast the tumor immune infiltration, targeted therapy effect, and prognosis of ccRCC. Finally, we analyzed the lncRNA most related to pyroptosis and verified it by experiments, which provided a promising target for ccRCC.

MATERIALS AND METHODS

Data Sources and Processing

The RNA-seq data (counts value) and the corresponding clinical information of ccRCC from the KIRC project of TCGA_GDC (https://portal.gdc.cancer.gov/), including 539 ccRCCs and 72 normal samples, were downloaded. According to the sample quality annotations provided in The Cancer Genome Atlas (TCGA) database (https://gdc.cancer.gov/about-data/publications/pancanatlas), 100 samples were filtered to exclude patients whose pathological diagnosis was not consistent with ccRCC and who had more than one tumor and underwent radiotherapy and/or chemotherapy (**Supplementary Table S1**). Finally, we obtained a total of 439 patients for follow-up studies, containing 447 tumor samples and 64 normal samples (some patients had multiple samples; **Supplementary Table S2**). Then the counts values were converted into TPM values. The gencode.gene.info.v22 file was downloaded from the GDC reference file (https://gdc.cancer.gov/about-data/gdc-data-processing/gdc-reference-files) to do gene annotation and extract lncRNA. Using the "caret" R package, the TCGA cohort was randomly divided into training and validation sets in the ratio 7:3. GSE76207 downloaded from the GEO website (https://www.ncbi.nlm.nih.gov/geo/) was used as the external validation data. Subsequently, 52 pyroptosis-related genes (PRGs) were obtained from MSigDB

(REACTOME_PYROPTOSIS) (http://www.broad.mit.edu/gsea/msigdb/) and previously published articles (**Supplementary Table S3**).

Consensus Clustering Analysis of DEPRGs

Differential analysis was carried out by the "DESeq2" package (Love et al., 2014), and the counts value downloaded from TCGA was used as the input data (|log2FC| > 1, padj <0.05). The R package "ConsensusClusterPlus" was employed for consensus unsupervised clustering analysis (Wilkerson and Hayes, 2010), dividing patients into different molecular subtypes according to the expression levels of differentially expressed pyroptosis-related genes (DEPRGs). The clustering was based on the following criteria: First, the intragroup correlation was close, while the intergroup correlation was weak. Second, the area under the cumulative distribution function (CDF) curve did not increase significantly. Third, the number of samples in all groups should not be too small. Univariate Cox regression analysis was performed to screen the prognosis-related genes, and the Kaplan–Meier method was used to calculate the overall survival (OS) between different clusters.

The Differences in Tumor Microenvironment, Immune Infiltration, and Drug Therapy of Pyroptosis Subtypes

The "ESTIMATE" R package was used to evaluate the TME score, such as stromal score, immune score, estimate score, and tumor purity. We evaluated infiltrations of immune cells with "CIBERSORT (Newman et al., 2015)," "XCELL (Aran et al., 2017)," "GSVA (Hänzelmann et al., 2013)," "TIMER (Li et al., 2020)," "QUANTISEQ (Finotello et al., 2019)," "MCPCOUNTER (Becht et al., 2016)," "EPIC(Racle et al., 2017)," and "CIBERSORT-ABS" R packages. Four types of immunophenoscore (IPS), including CTLA4_negative + PD-1_negative, CTLA4_positive + PD-1_negative, CTLA4_negative + PD-1_positive, CTLA4_positive + PD-1_positive, were obtained from the TCIA Database (https://tcia.at/home). The high PD-1_positive IPS showed a well-predicted response to anti-PD-1 treatment. The R package "pRRophetic (Geeleher et al., 2014)" was applied to evaluate the drug sensitivity of first-line targeted therapy for ccRCC, including sunitinib, sorafenib, and axitinib (Numakura et al., 2021). Correlation analysis was performed using the SPEARMAN correlation test and the Wilcoxon signed-rank test.

Gene Set Enrichment Analysis

To determine whether there were differences in biological processes among different clusters, we downloaded "c2.cp.kegg.v7.4.symbols.gm" and "c5.go.v7.4.symbols.gmt" from MSigDB (http://www.gsea-msigdb.org/gsea/index.jsp). Then, we used "GSVA" package for analysis, and important pathways were shown in the form of heatmap. The "clusterProfiler" package was applied for Gene Ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analysis of differentially expressed genes (DEGs). An adjusted *p*-value (padj) < 0.05 was considered statistically significant.

Construction of Prognosis-Related Pyroptosis-Score

We obtained 403 common differential genes through pairwise difference analysis of the three pyroptosis subtypes by the "DESeq2″ R package (|log2FC| > 1, padj <0.01), and then univariate Cox analysis was performed to identify 183 prognosis-related genes. Subsequently, we used the principal component analysis to obtain PC1 and PC2 from feature genes, which were added as the pyroptosis-score of each patient. Then, we compared the survival and clinical characteristics of patients with different pyroptosis-score by Kaplan–Meier analysis. The "estimate," "CIBERSORT," and "pRRophetic" packages were utilized to evaluate the immune infiltration and immune microenvironment of patients with different pyroptosis-score, and to preliminarily explore therapeutic drugs. Finally, we used GSEA software for GSEA analysis to analyze the signal pathways that patients with different pyroptosis-score may participate in. The padj <0.05 was defined as statistically significant.

Construction of a Pyroptosis-Related Long Noncoding RNA Pairs-Based Prognostic Signature

First, we extracted the lncRNAs from the significantly different genes (|log2FC| > 1, padj <0.05) of patients with high or low pyroptosis-score and used co-expression analysis to obtain 76 lncRNAs co-expressed with 19 DEPRGs (|correlation coefficient| > 0.55, $p < 0.001$). These lncRNAs were paired to form PRlncRNAPs, and each PRlncRNAP was scored. Then, we compared the expression levels of these two lncRNAs. If the latter was lower than the former, the score was recorded as 1, otherwise the score was defined as 0. PRlncRNAPs with over 80% or under 20% of score 0 or 1 were excluded from further analysis (Xiong et al., 2020). Subsequently, the PRlncRNAPs were subjected to univariate Cox analysis, and PRlncRNAPs with $p < 0.001$ were used for further study. To address the multicollinearity effect between variables, we used the "glmnet" R package for least absolute shrinkage and selection operator (LASSO) Cox regression (Engebretsen and Bohlin, 2019). Then we constructed a prognostic model by multivariate Cox analysis. Finally, we worked out the riskScore of each sample based on the following formula:

$$riskScore = \sum_{i=1}^{n} Coef_i \times X_i$$

where *i* means the number of prognostic PRlncRNAPs, Coef is the regression coefficient, and X is the expression value of PRlncRNAPs, respectively. We evaluated the accuracy of riskScore in forecasting the prognoses of patients by the time-dependent receiver operating characteristic (ROC) curve (Heagerty et al., 2000). Then, the median value of riskScore was defined as the cutoff point to divide the patients into prone and low-risk groups in the training or validation group.

Clinical Value of riskScore

Kaplan–Meier survival analysis was performed to assess the differences in OS of patients with high or low risk. Univariate and multivariate Cox regression analyses were performed on riskScore and clinicopathological features to evaluate whether riskScore was an independent clinical prognostic factor. The hazard ratio (HR) was calculated by "survival" R package. Vesteinn Thomson's study showed that all TCGA tumors were divided into six immune subtypes, including wound healing, IFN-γ dominant, inflammatory, lymphocyte depleted, immunologically quiet, and TGF-β dominant (Thorsson et al., 2018). Chi-square test was performed to investigate the differences in immune subtypes between high- and low-risk groups.

Establishment and Validation of a Nomogram Scoring System

The "rms" package was employed for developing a predictive nomogram based on the results of the independent prognosis analysis. In the nomogram scoring system, each variable got a score and the scores of all variables were added to get the total score of each sample. The ability of the model to correctly classify the research event was evaluated by ROC curves and concordance index (C-index). The ROC curves and C-index) were also used to compare the models we built with others (He et al., 2021; Zhang F. et al., 2021; Chen et al., 2022; Sun et al., 2022; Zhang et al., 2022). The differences in the predicted survival events and the virtually observed outcomes were compared by calibration plots of the nomogram. Compared with the ROC curve, decision curve analysis (DCA) considers the clinical utility of a specific model, and we used it to depict the potential clinical effect of the prognostic model (Kerr et al., 2016).

Clinical Tissue Samples and Cells

The ccRCC tissues and matched adjacent normal tissues were obtained from 32 patients in The First Affiliated Hospital of Sun Yat-sen University (Guangzhou, China). All tissues were immediately frozen in liquid nitrogen and stored at −80°C until RNA was extracted. The samples used in this study were approved by the Ethics Committee of The First Affiliated Hospital of Sun Yat-sen University (Guangzhou, China). The above 32 patients signed the informed consents. HK2, 786-O, 769-P, CAKI-2, and OS-RC-2 in this study were purchased from the Chinese Academy of Science.

Quantitative Real-Time PCR (qRT-PCR)

Total cellular RNA was extracted using Trizol (Thermofisher Scientific, United States). The RNA was reverse transcribed into cDNA following the steps of the PrimeScrip Reverse Transcription Kit (Takara, Dalian, China). The PCR reaction system was configured and analyzed according to the instructions of SYBR Green Pro Taq HS premix (Accurate Biology, Changsha, China). The PCR primer sequences were: AC002331.1 forward: 5′-TGCTGCCAAAGTAGG AGGATTC-3′, reverse: 5′-GAAGGAAGTGCTCCACACAGTC-3'; GAPDH forward: 5′-GTCTCCTCTGACTTCAACAGCG-3′, reverse: 5′-ACCACCCTGTTGCTGTAGCCAA-3'.

TABLE 1 | Clinicopathological features of 439 ccRCC patients.

Type	Details	Total	Training	Validation	*p* value
Age	≤65	293 (66.74%)	201 (65.26%)	92 (70.23%)	0.3679
	>65	146 (33.26%)	107 (34.74%)	39 (29.77%)	—
Gender	Female	159 (36.22%)	111 (36.04%)	48 (36.64%)	0.9907
	Male	280 (63.78%)	197 (63.96%)	83 (63.36%)	—
Fustat	Alive	300 (68.34%)	211 (68.51%)	89 (67.94%)	0.9961
	Dead	139 (31.66%)	97 (31.49%)	42 (32.06%)	—
Grade	G1	13 (2.96%)	9 (2.92%)	4 (3.05%)	0.4889
	G2	180 (41%)	133 (43.18%)	47 (35.88%)	—
	G3	176 (40.09%)	117 (37.99%)	59 (45.04%)	—
	G4	63 (14.35%)	43 (13.96%)	20 (15.27%)	—
	Unknown	7 (1.59%)	6 (1.95%)	1 (0.76%)	—
Stage	Stage I	208 (47.38%)	145 (47.08%)	63 (48.09%)	0.4305
	Stage II	46 (10.48%)	37 (12.01%)	9 (6.87%)	—
	Stage III	109 (24.83%)	75 (24.35%)	34 (25.95%)	—
	Stage IV	73 (16.63%)	49 (15.91%)	24 (18.32%)	—
	Unknown	3 (0.68%)	2 (0.65%)	1 (0.76%)	—
T	T1	214 (48.75%)	150 (48.7%)	64 (48.85%)	0.1673
	T2	56 (12.76%)	46 (14.94%)	10 (7.63%)	—
	T3	160 (36.45%)	106 (34.42%)	54 (41.22%)	—
	T4	9 (2.05%)	6 (1.95%)	3 (2.29%)	—
M	M0	339 (77.22%)	242 (78.57%)	97 (74.05%)	0.6895
	M1	69 (15.72%)	47 (15.26%)	22 (16.79%)	—
	Unknown	31 (7.06%)	19 (6.17%)	12 (9.16%)	—
N	N0	196 (44.65%)	138 (44.81%)	58 (44.27%)	0.7167
	N1	13 (2.96%)	8 (2.6%)	5 (3.82%)	—
	Unknown	230 (52.39%)	162 (52.6%)	68 (51.91%)	—

Statistical Analyses

The unpaired *t*-test and Wilcoxon rank-sum test were used to evaluate the difference between normally distributed and non-normally distributed data, respectively. OS curves were obtained with Kaplan–Meier analysis and differences between groups were calculated with log-rank test. R software (version 4.0.3) and Adobe Illustrator (version 25.0) were employed for statistical analysis and drawing. *p*-value <0.05 indicates statistical significance.

RESULTS

Data Processing

This study was conducted according to the flow chart (**Supplementary Figure S1**). In order to make our research more trustworthy, TCGA data were randomly divided into a training set ($n = 308$) and a validation set ($n = 131$) in the ratio 7:3. There was no difference in various clinicopathological parameters between the two datasets, with p-value > 0.05 (**Table 1**).

Identification of Pyroptosis Subtypes in ccRCC

First, we obtained the DEGs by investigating the difference between the ccRCC group and the normal group using the "DESeq2" package. Then, 19 DEPRGs were acquired after the intersection with PRGs (**Figure 1A**), including 17 upregulated and two downregulated DEPRGs (**Figure 1B**). Univariate Cox regression was performed to reveal the prognosis of 19 DEPRGs in patients with ccRCC

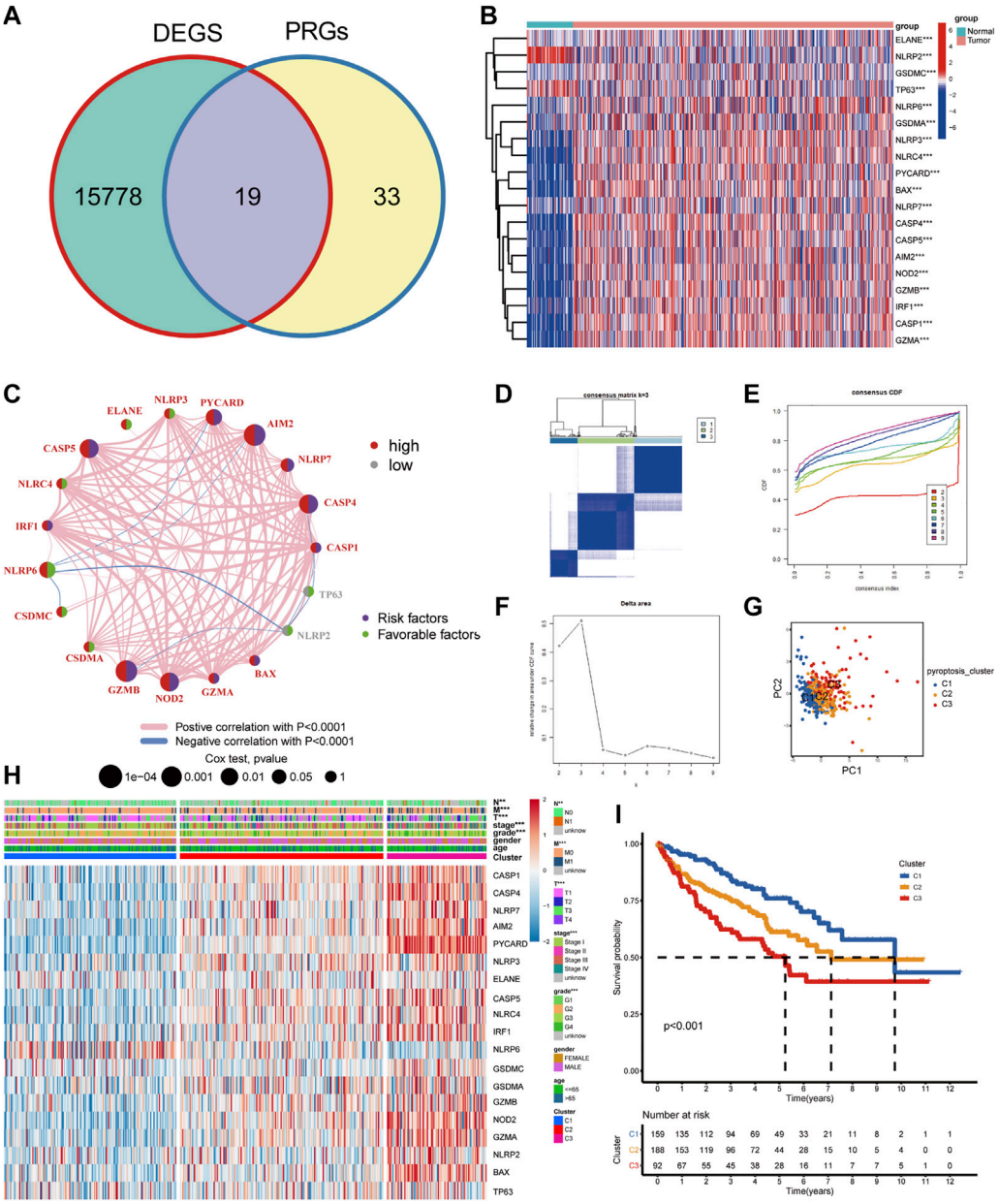


FIGURE 1 | The landscape of DEPRGs in ccRCC and identification of pyroptosis subtypes. **(A)** The Venn diagram showed overlapping genes. **(B)** A heatmap was used to show the differential expression of 19 DEPRGs expressions in ccRCC and noncancerous tissues. **(C)** Interaction between DEPRGs in ccRCC. The line connecting DEPRGs represented their interaction, and the strength of the correlation between DEPRGs was indicated by the thickness of the line. Pink and blue, respectively, represented positive and negative correlations. The color of the left half circle represented the gene expression in ccRCC, with high expression in red and low in gray. The color of the right half circle represented the influence of genes on prognosis. Purple was the risk factor of prognosis, green was the favorable factor of prognosis, and the size of the circle represented the *p*-value. **(D)** Consensus matrix when k was 3. **(E)** Consensus CDF when k was between 2 and 9. **(F)** Delta area showed the relative change of the area under the CDF curve comparing k and k−1 and it met our screening criteria when k = 3. **(G)** PCA showed marked differences in the transcriptome among the three pyroptosis subtypes. **(H)** The clinicopathological characteristics of the three subtypes and different expression levels of DEPRGs. **(I)** Survival curves of patients with three pyroptosis subtypes (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

(**Supplementary Table S4**). $p < 0.05$ was utilized as the screening threshold to screen 14 genes related to prognosis. Then, we divided patients into high and low groups based on the optimal cutoff value of the above 14 genes, and the survival analysis curves of 14 genes were obtained (**Supplementary Figure S2**). The pyroptosis network showed the interaction and prognostic value of the 19 DEPRGs (**Figure 1C**). Then, univariate Cox analysis was performed to identify seven prognosis-related DEPRGs. The "ConsensusClusterPlus" package was applied to cluster the TCGA-KIRC cohort into different groups through the consistent expression of the seven

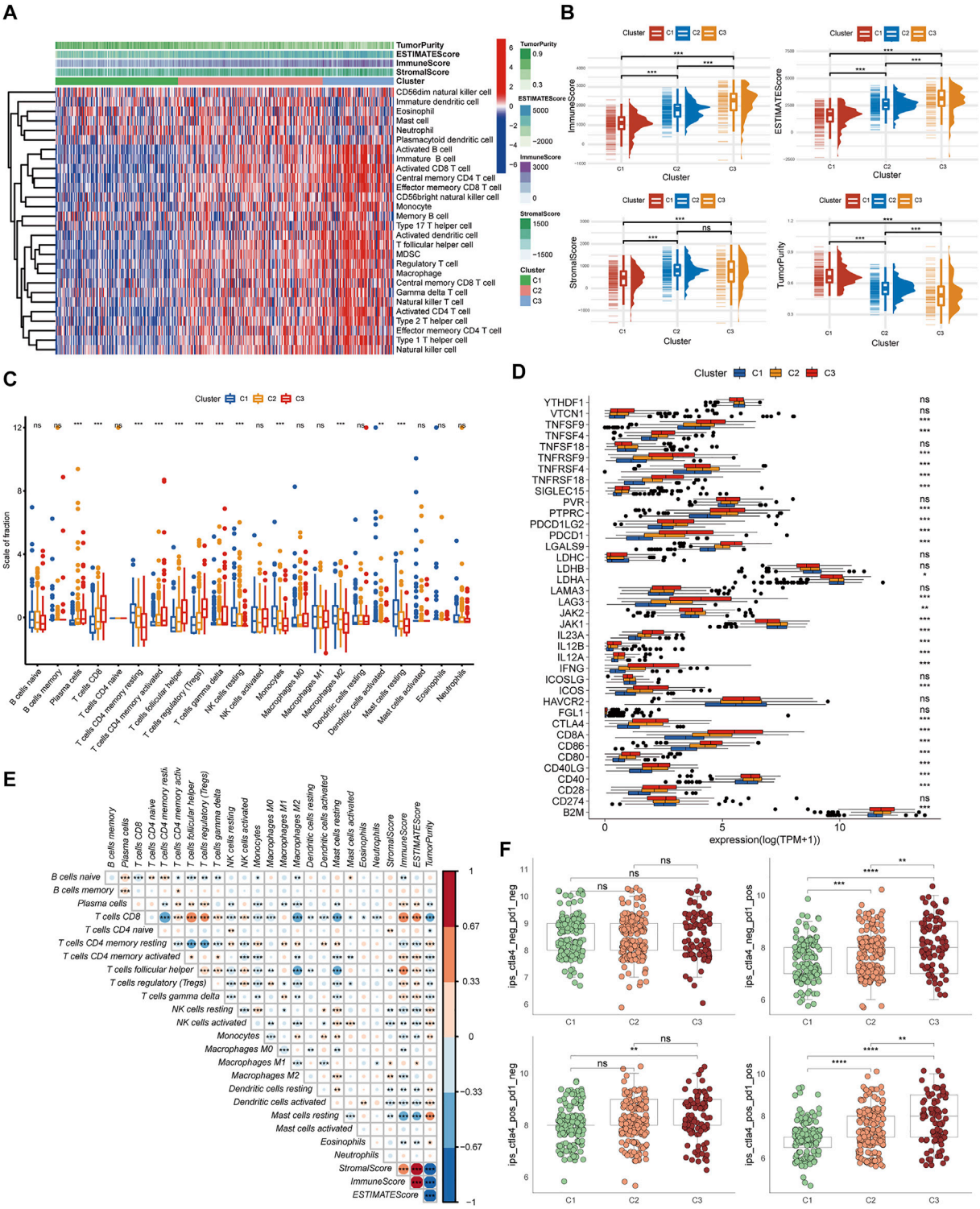


FIGURE 2 | Immune landscape and drug response of pyroptosis subtypes. **(A)** The heatmap of immune cell infiltration annotations and immune microenvironment scores in pyroptosis subtypes. **(B)** The levels of estimate score, immune score, stromal score, and tumor purity in pyroptosis subtypes. **(C)** The relative abundance of 22 immune cell types in pyroptosis subtypes. **(D)** The expression levels of common immune checkpoints among pyroptosis subtypes. **(E)** The heatmap of the interaction between immune cells. **(F)** IPS comparison among the three pyroptosis subtypes of the patients with ccRCC in the CTLA4 negative/positive or PD-1 negative/positive groups. CTLA4_positive or PD-1_positive, respectively, stood for anti-CTLA4 or anti-PD-1 therapy (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

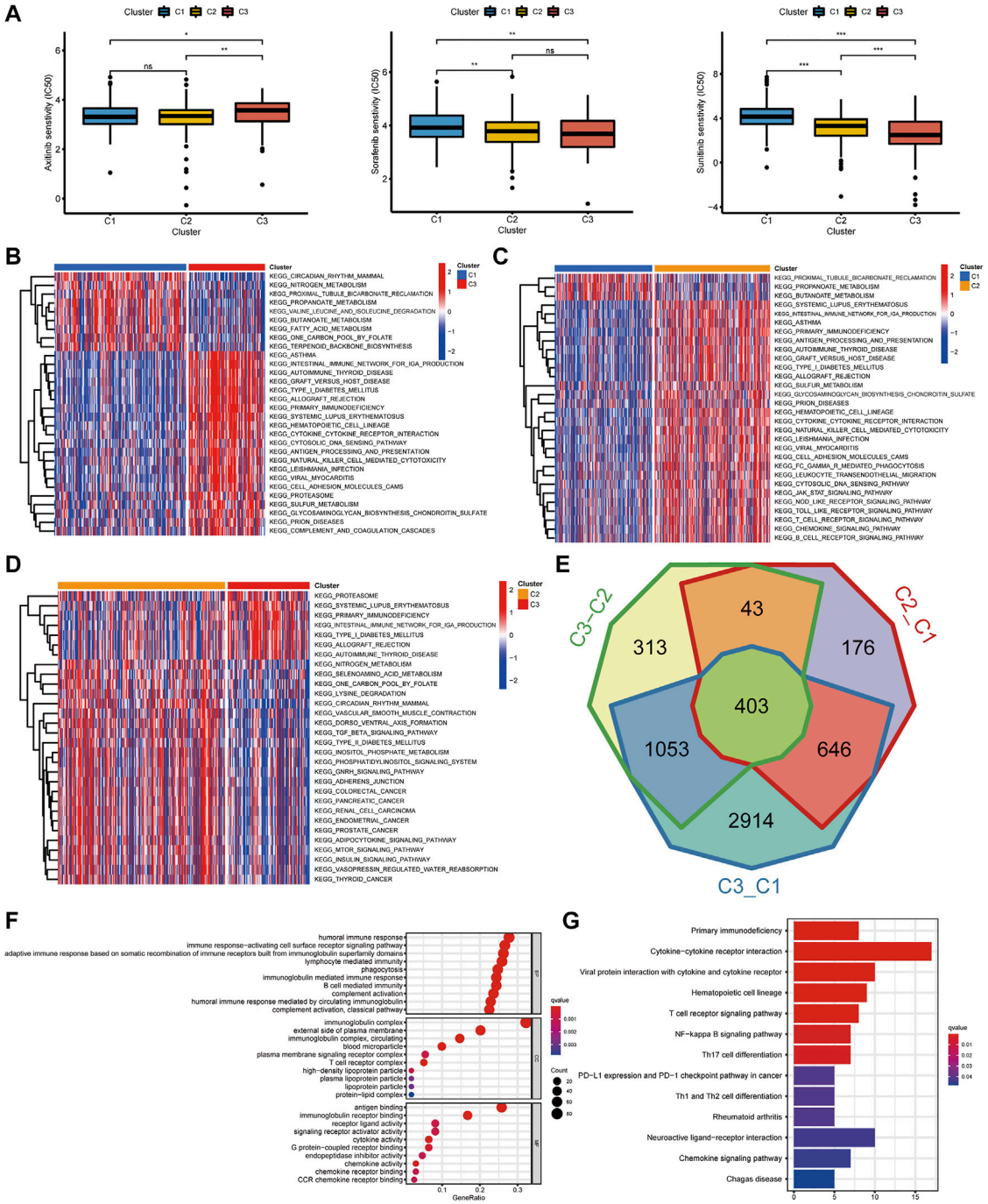


FIGURE 3 | The interaction and correlation among the pyroptosis subtypes. **(A)** The different IC50 of targeted drugs in the three pyroptosis subtypes. **(B–D)** GSVA enrichment analysis presents the activation states of biological pathways in the three pyroptosis subtypes. The heatmaps were utilized to visualize the biological processes. Red and blue, respectively, represented the activated and inhibited pathways. **(E)** A Venn diagram revealed the common differential genes of the three pyroptosis subtypes. **(F,G)** GO and KEGG analysis of DEGs in the three pyroptosis subtypes (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

DEPRGs. When the consensus matrix k value was 3, the crossover among ccRCC samples was the least, which met our screening criteria (**Figure 1D–F**). PCA demonstrated that the three subtypes were distributed in different clusters (**Figure 1G**). There were significant differences in clinicopathological parameters such as grade, stage, T, M, and N of the three pyroptosis subtypes of patients (**Figure 1H**). Kaplan–Meier survival analysis suggested that patients with different pyroptosis subtypes had significant differences in OS ($p < 0.001$), among which patients with C1 subtype had the best OS and C3 subtype was the worst (**Figure 1I**).

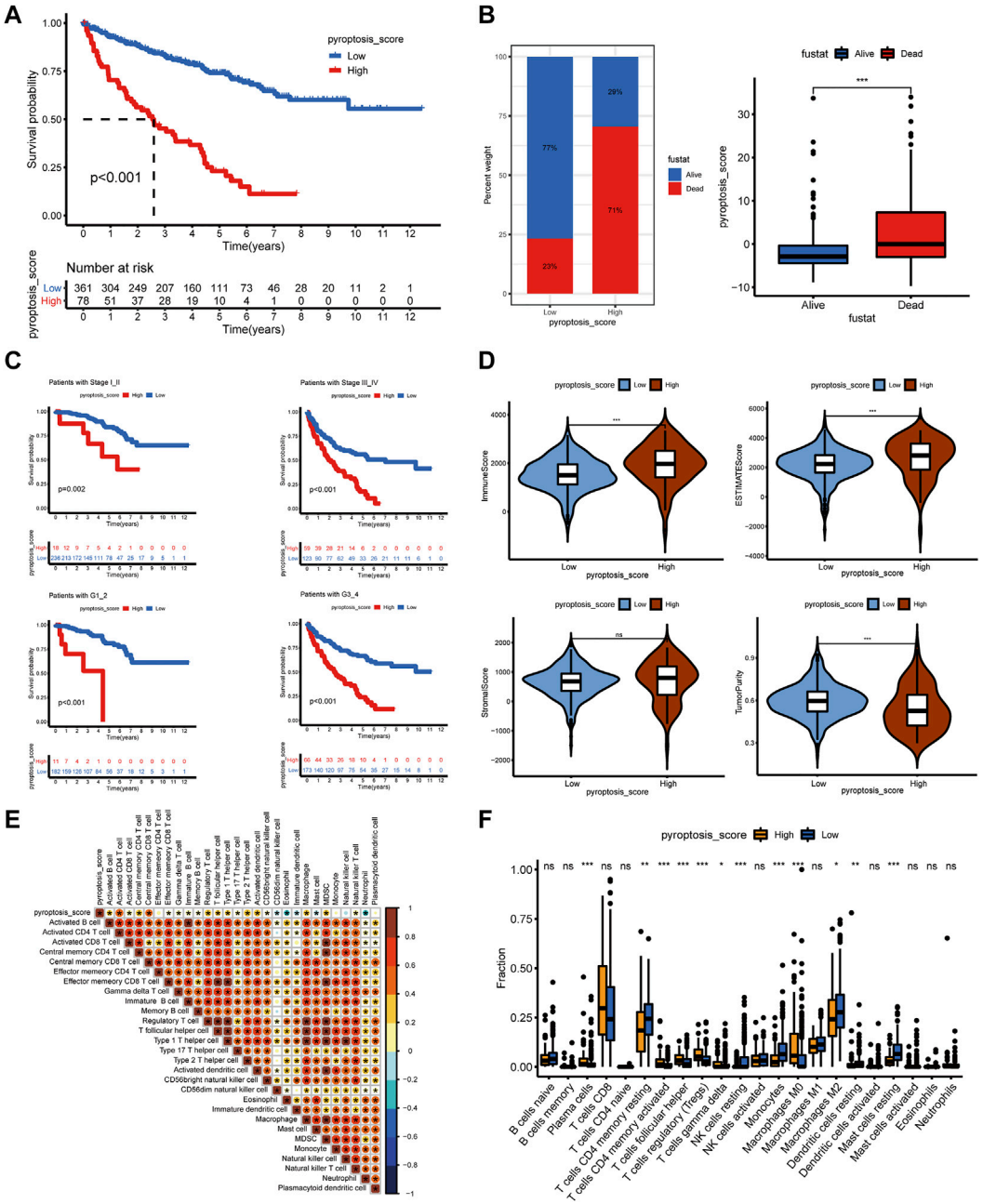


FIGURE 4 | The clinical characteristics and immune landscape of the pyroptosis-score groups. **(A)** Survival curves of patients with high and low pyroptosis-scores. **(B)** The ratio of survival status in patients with high and low pyroptosis-scores and the difference in pyroptosis-score level of different survival status. **(C)** The survival curves of patients with high and low pyroptosis-scores at different stages and grades. **(D)** The different levels of estimate score, immune score, stromal score, and tumor purity in high and low pyroptosis-score groups. **(E)** Correlation analysis between pyroptosis-score and immune cells. **(F)** Differences in relative abundance of 22 immune cell types in patients with high and low pyroptosis-scores (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

Immune Landscape and Drug Response of Pyroptosis Subtypes

To further investigate the relationship between pyroptosis subtypes and immune cells, we quantified the enrichment scores of single-sample gene set enrichment analysis (ssGSEA) for different immune cell subgroups. The results indicated that compared with subtype C1, activated dendritic cell, MDSC, macrophage, activated CD4 T cell, T follicular helper cell, activated CD8 T cell, and natural killer T cell infiltrated more in C2 and C3 (**Figure 2A**). Then "ESTIMATE" package was utilized to evaluate the TME scores of the three subtypes, containing stromal score, immune score, estimate score, and

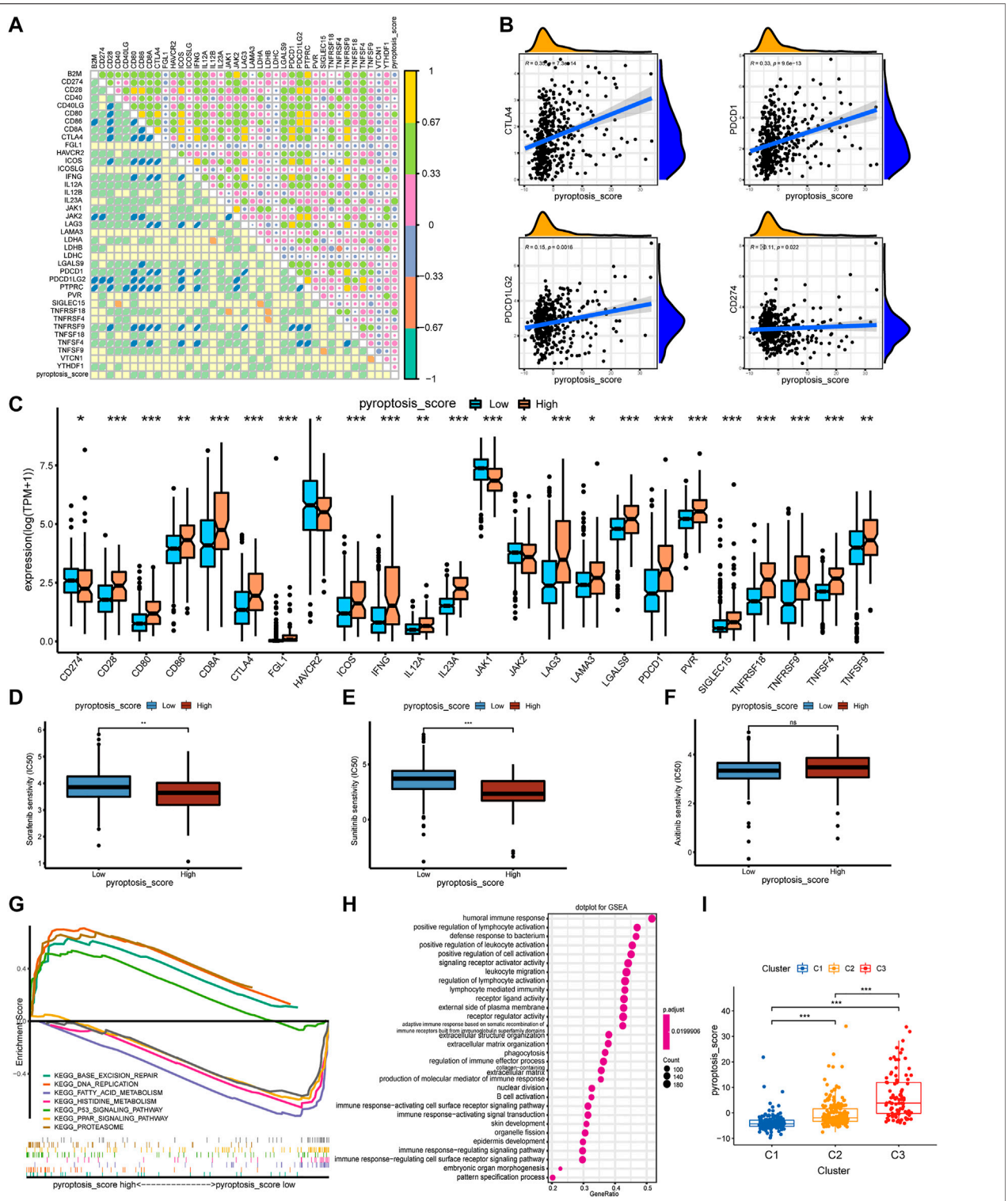


FIGURE 5 | The correlation of pyroptosis-score with immune checkpoints and drug response, and the biological process of grouping. **(A)** Spearman correlation coefficient diagram of pyroptosis-score and common immune checkpoints. **(B)** Spearman correlation scatter plot between pyroptosis-score and CTLA4, PDCD1, PDCD1LG2, and CD274. **(C)** The different expression levels of common immune checkpoints between the high and low pyroptosis-score groups. **(D–F)** The different IC50 of the targeted drugs among high and low pyroptosis-score. **(G,H)** GSEA enrichment analysis of high and low pyroptosis-score groups, including KEGG pathways **(G)**, and GO annotation **(H)** (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$). **(I)** The pyroptosis-score of the three pyroptosis subtypes.

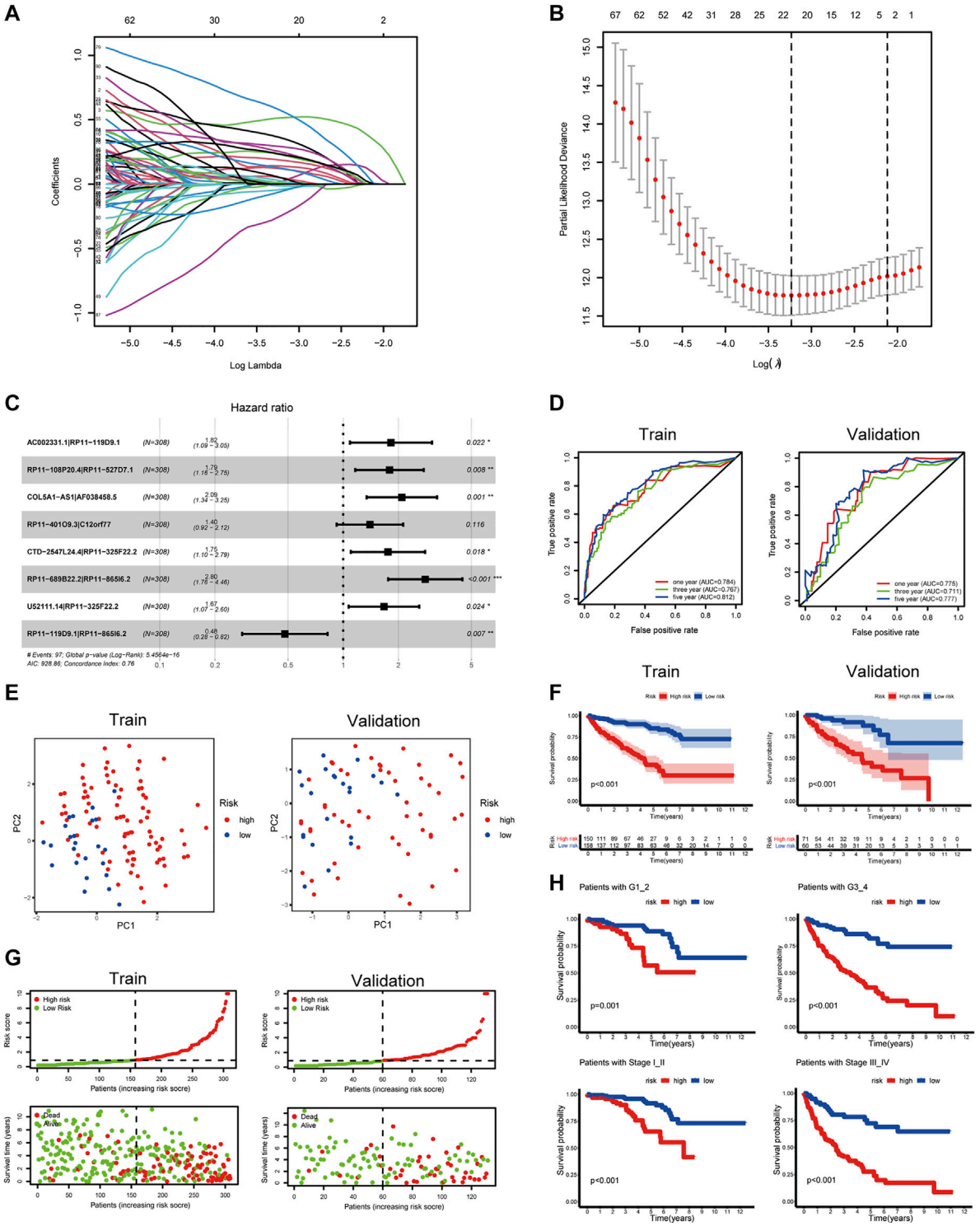


FIGURE 6 | Construction of a prognostic signature for ccRCC. **(A,B)** LASSO coefficient plot of PRlncRNAPs. **(C)** Forest plots of the eight OS-related PRlncRNAPs. **(D)** The ROC curve showed the diagnostic value of riskScore for the ccRCC prognosis. **(E)** PCA on the basis of prognostic characteristics. The red and blue dots, respectively, represented patients with high and low risk. **(F)** OS analysis of patients with high and low riskScore. **(G)** Distribution of riskScore and survival status of patients. **(H)** Survival curves of patients with the high- and low-risk groups in different grades and stages (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

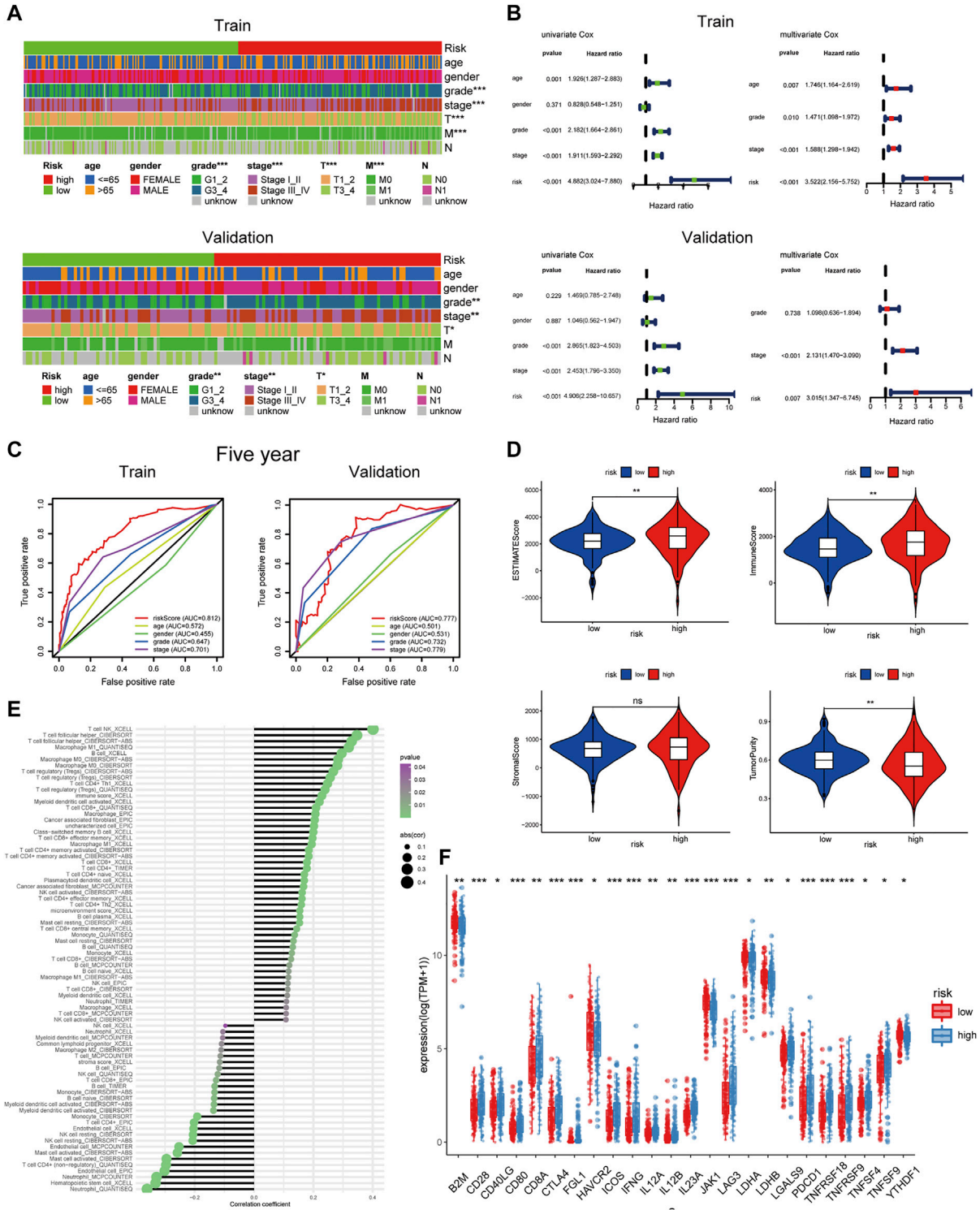


FIGURE 7 | Clinical characteristics and immune infiltration of risk groups based on riskScore. **(A)** The heatmap reveals the distribution of clinical characteristics in patients in the high- and low-risk groups. **(B)** Univariate and multivariate Cox regression analyses showed that risk was an independent prognostic factor. **(C)** The receiver operating characteristic curve of riskScore and clinical features. **(D)** Different levels of estimate score, immune score, stromal score, and tumor purity in different risk groups. **(E)** Lollipop chart of tumor-related infiltrating immune cells on the basis of TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, MCPCOUNTER, XCELL, and EPIC algorithms in different risk groups. **(F)** The expression levels of common immune checkpoints in different risk groups (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

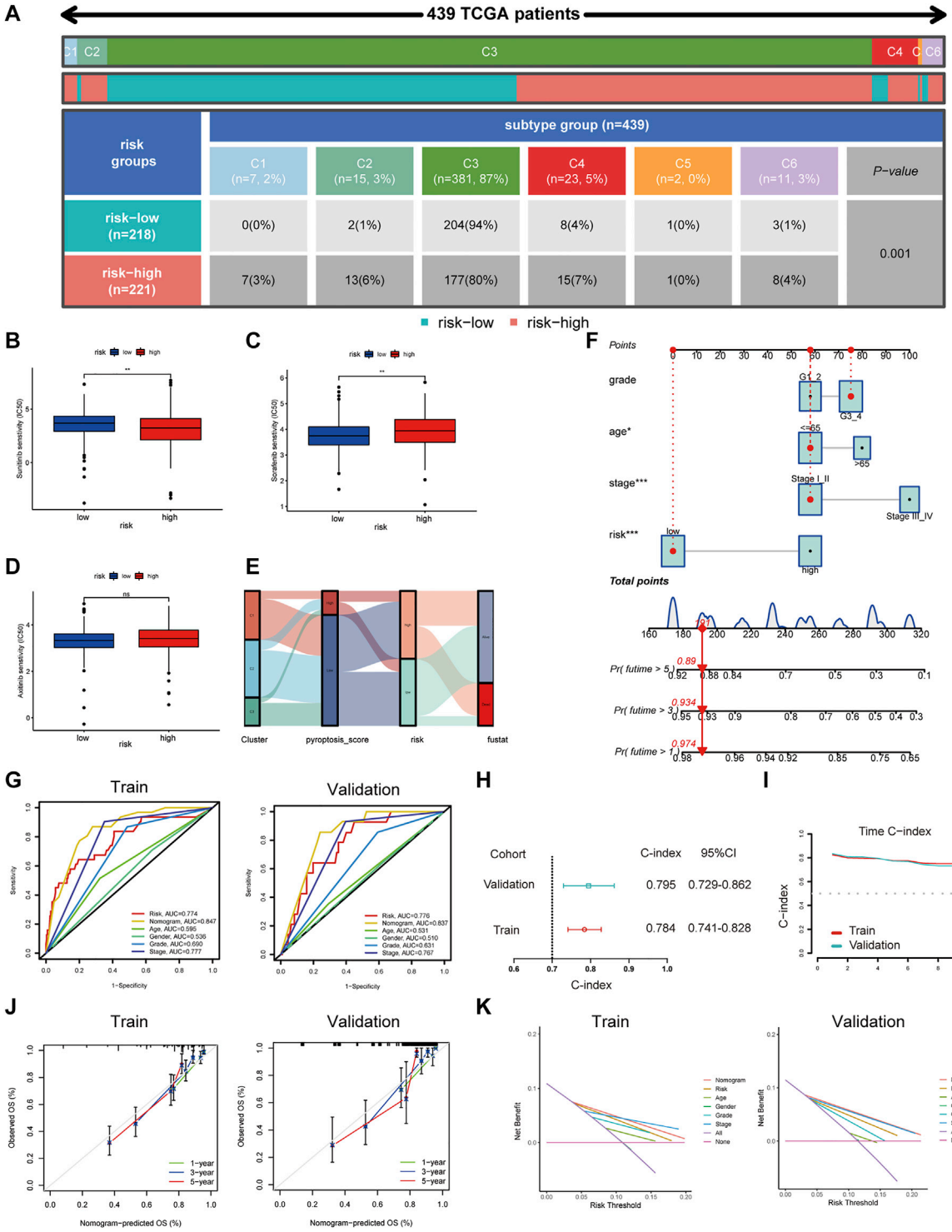


FIGURE 8 | Drug treatment response and nomogram construction and validation. **(A)** The distribution of six immune subtypes of patients in high- and low-risk groups. **(B–D)** Different IC50 of targeted drugs between the patients in high- and low-risk groups. **(E)** The alluvial diagram demonstrated the relationship between risk groups based on riskScore and molecular subtypes. **(F)** A nomogram containing clinical features and the risk groups based on riskScore to predict survival time. **(G)** The receiver operating characteristic curve of nomogram and clinical features. **(H)** The C-index for exhibiting the ratio of the predicted result to the actual result. **(I)** The C-index trend over time. **(J)** The calibration curve was used to predict the OS of ccRCC patients in the training and validation cohorts. **(K)** The NB assessing the outcome was illustrated by DCA.

tumor purity. Compared with C1, C2 and C3 subtypes showed that the tumor purity was lower and immune cells and stromal cells were higher. Estimate score suggested that the relative content of stromal cells and immune cells in C3 subtype was the highest (**Figure 2B**). CIBERSORT was employed for analyzing the relative abundance of 22 immune cells in each tumor sample, and the results suggested that there were more T cells regulatory (Tregs), T follicular helper cells, and CD8 T cells in C3 subtype (**Figure 2C**). Currently, immune checkpoint inhibitors are the first-line therapeutic drugs for advanced ccRCC. We had screened several targeted biomarkers that were essential for immunotherapy and further clarified whether the pyroptosis subtypes were related to them. We discovered that CTLA4, PDCD1, and PDCD1LG2 had the highest expression in the C3 subtype (**Figure 2D**). Survival analysis revealed that when FGL1, LAG3, TNFRSF18, and IL-23A were highly expressed, the prognosis of patients was worse, while JAK1, JAK2, and LDHA were on the contrary (**Supplementary Figure S3**), which may provide a new target for immunotherapy of ccRCC. The heatmap exhibiting the interaction of immune cells was plotted, which displayed that the immune score was highly negatively correlated with different types of macrophages but positively correlated with different types of T cells (**Figure 2E**). We performed immunophenogram analysis for analyzing the relationship between IPS and pyroptosis subtypes (**Figure 2F**). The outcomes revealed that in CTLA4_negative + PD-1_positive type and CTLA4_positive + PD-1_positive type, the IPS of C3 subtype was the highest. These results suggested that C3 subtype patients were more sensitive to anti-PD-1 therapy or a combination of anti-PD-1 and anti-CTLA4 therapies. The targeted therapies, including axitinib, sorafenib, and sunitinib, were the first-line therapeutic drugs for ccRCC. We explored the effect of pyroptosis subtypes on the sensitivity of these drugs in ccRCC (**Figure 3A**). Interestingly, we realized that the IC50 of axitinib was higher in C3 subtype, while those of sorafenib and sunitinib were lower. Based on the above analysis, we found that the state of pyroptosis may significantly inhibit or enhance the expression of specific immune cell types and then potentially affect the reaction to immunotherapy and targeted therapies.

Interaction and Correlation Between Pyroptosis Subtypes

The gene set variation analysis (GSVA) enrichment analysis was used for investigating the potential biological processes among the three pyroptosis subtypes (**Figure 3B–D**). It was found that metabolism-related pathways were significantly enriched in C1 subtypes, such as KEGG_FATTY_ACID_METABOLISM, while C2 and C3 subtypes presented enrichment pathways related to complete immune activation, including KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION, KEGG_PRIMARY_IMMUNODEFICIENCY, KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION, KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION, KEGG_T_CELL_RECEPTOR_SIGNALING_PATHWAY, and KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY. To further clarify the biological differences between the three pyroptosis subtypes, 403 common difference genes were obtained by pairwise difference analysis between pyroptosis subtypes (**Figure 3E**), and genes function enrichment analysis was carried out (**Figure 3F**). It was found that the common difference genes were significantly enriched in immune-related biological processes. KEGG enrichment pathway analysis further confirmed the activation of immune-related pathways, including cytokine–cytokine receptor interaction, T-cell receptor signaling pathway, and chemokine signaling pathway (**Figure 3G**). It suggested that pyroptosis played a vital role in the immune regulation of TME.

Construction of Pyroptosis-Score and the Relationship With TME and Drug Response

To further understand the immune regulation of pyroptosis on ccRCC, we made an effort to analyze the common differential genes of the three pyroptosis subtypes and established a pyroptosis-score based on the subtype-related DEGs for each patient. After univariate Cox regression analysis, 183 prognosis-related genes were obtained from 403 common DEGs of pyroptosis subtypes, and then PCA was performed. We add the scores of PC1 and PC2 to get the pyroptosis-score of each sample. After that, the optimal threshold was obtained by the R package "survminer" to divide the patients into high and low pyroptosis-score groups. Relative to the low pyroptosis-score patients, there were a shorter OS (**Figure 4A**) and more deaths (**Figure 4B**) in the high group. Kaplan–Meier analysis was applied to investigate the relationship between pyroptosis-score and the prognosis in different clinical groups. The results indicated that regardless of the patient's stage and grade, the prognosis of patients in the high pyroptosis-score group was worse (**Figure 4C**). Relative to low pyroptosis-score, patients in the high group had lower tumor purity and higher immune cells and stromal cells. The estimate score also revealed that the relative content of stromal cells and immune cells in the high pyroptosis-score group was higher (**Figure 4D**). Correlation analysis indicated that pyroptosis-score was positively correlated with activated CD4 T cell, activated CD8 T cell, T follicular helper cell, type 2 T-helper cell, activated dendritic cell, and MDSC (**Figure 4E**). Compared with the low pyroptosis-score group, the patients in the high group infiltrated more memory-activated CD4 T cell, T follicular helper cells, Tregs, and macrophage M0 cells (**Figure 4F**). Correlation analysis between scores and immune checkpoints presented that pyroptosis-score was positively correlated with CTLA4, PDCD1, and PDCD1LG2 (**Figure 5A,B**). The high pyroptosis-score group expressed higher CTLA4 and PDCD1 and lower CD274 (**Figure 5C**). Then, we performed immunophenogram analysis for analyzing the relationship between IPS and pyroptosis-score (**Supplementary Figure S4**). The outcomes revealed that in CTLA4_negative + PD-1_positive type, CTLA4_positive + PD-1_negative type, and CTLA4_positive + PD-1_positive type, the IPS of high pyroptosis-score group was higher. These results suggested that patients in high pyroptosis-score groups were more sensitive to immunity therapies. In terms of targeted drug treatment sensitivity, the high pyroptosis-score

group was more sensitive to sorafenib and sunitinib treatment (**Figure 5D,E**), but there was no significant difference in axitinib (**Figure 5F**). GSEA enrichment analysis indicated that the pathways of the high pyroptosis-score group were mainly related to DNA damage repairs, such as KEGG_BASE_EXCI SION_REPAIR, KEGG_DNA_REPLICATION, and KEGG_P53 _SIGNALING_PATHWAY. The pathways of the low pyroptosis-score group were enriched in the activation of metabolism pathways such as KEGG_FATTY_ACID_METABOLISM, KEGG_HISTIDINE_METABOLISM, and KEGG_TRYPTOPH AN_METABOLISM (**Figure 5G**). We also found that immune-related biological processes were activated in the high pyroptosis-score group (**Figure 5H**). There were significant differences in the pyroptosis-score among the pyroptosis subtypes. The outcomes revealed that the pyroptosis-score was the lowest in C1 and the highest in C3 (**Figure 5I**). The characteristics of patients with C3 subtype and high pyroptosis-score group were consistent. Both have immune activation characteristics, and the prognosis was worse than that in other groups.

Identification of OS-Related PRlncRNAPs and Establishment of the Prognostic Model

There were 76 lncRNAs co-expressed with 19 DEPRGs in the differential genes of the high and low pyroptosis-score groups. By recombining the 76 lncRNAs, we obtained 1,425 PRlncRNAPs and the corresponding relative expression levels; then 90 PRlncRNAPs related to prognosis were recognized through the univariate Cox regression analysis ($p < 0.001$). Subsequently, we extracted 22 OS-related PRlncRNAPs after 1,000 iterations by LASSO Cox regression analysis (**Figure 6A,B**). Ultimately, we obtained eight PRlncRNAPs to construct the risk model by multivariate regression analysis (**Figure 6C**), and the coefficient of PRlncRNAPs was employed for calculating the riskScore (**Supplementary Table S5**). The above method was carried out in the randomly selected training set, and the same coefficient was applied to the validation set. The ROC curves showed that the riskScore exhibited excellent prediction ability. In the training set, the 1-, 3-, and 5-year areas under the curve (AUC) were 0.784, 0.767, and 0.812, respectively. The AUCs in the validation set were, respectively, 0.775, 0.711, and 0.777 (**Figure 6D**). Based on the median value of riskScore in the training set, the patients with ccRCC were divided into high- and low-risk groups. PCA displayed distinguishable dimensions between the high- and low-risk groups in the training set while it did not seem to be so perfect in the validation set (**Figure 6E**).

Clinical Value of Risk Groups

To explore the clinical value of the risk groups, we plotted a survival curve to evaluate the survival of patients in the low- or high-risk groups. The outcomes showed that, relative to the patients in the low-risk groups, the prognosis of patients in the high-risk groups was worse in both the training and validation sets ($p < 0.001$, **Figure 6F**). According to the riskScore and survival status of patients, we noted an increase in the mortality rate of patients with high riskScore (**Figure 6G**). To explore whether the risk model was suitable for different clinical groups, we merged the riskScore and clinical information from the training and validation sets and sketched the survival curves of grades and stages at different stages. The outcomes showed that, compared to the patients in the low-risk groups, the prognosis of patients in the high-risk group was worse regardless of the grade and stage ($p < 0.001$, **Figure 6H**). Subsequently, we explored the relationship between riskScore and clinical characteristics. The heatmap displayed the distribution of age, gender, stage, and grade in the training and validation sets. It could be seen that riskScore was significantly correlated with the grade, stage, T, and M in the training set ($p < 0.05$) and grade, stage, and T in the validation set ($p < 0.05$, **Figure 7A**). The boxplots presented the relationship between the clinical characteristics and riskScore (**Supplementary Figure S5A,B**). External validation on the GEO dataset also confirmed that patients with stage III_IV had a higher riskScore (**Supplementary Figure S5C**). The univariate Cox regression analysis showed that risk groups based on riskScore were essential risk factors for ccRCC (HR > 1, $p < 0.001$). Furthermore, multivariate Cox regression analysis showed that it was also an independent prognostic factor of ccRCC (**Figure 7B**). The ROC curves indicated that, compared with other factors, riskScore exhibited better prediction ability in 5 years (**Figure 7C** and **Supplementary Figure S6A**). It meant that riskScore performed better in predicting long-term survival of patients compared with various clinical indicators. These outcomes indicated that the risk model could be applied as a vital indicator for evaluating the prognosis of ccRCC, which was confirmed in the validation set. Then we further investigated the correlation with the immune microenvironment of the signature which was based on PRlncRNAPs. Patients in the high-risk group had lower tumor purity and higher immune score than the low-risk group, but no difference in stromal score (**Figure 7D**). The lollipop chart presented that riskScore was positively correlated with Tregs, cancer-associated fibroblasts, T follicular helper cell, memory-activated $CD4^+$ T cell, and macrophage M0 and negatively correlated with NK cell resting, mast cell activated, etc. (**Figure 7E**). Differential analysis of immune checkpoints suggested that the expressions of CTLA4, PDCD1, LAG3, and TNFRSF4 were upregulated in high-risk patients, but JAK1 and HAVCR2 were downregulated (**Figure 7F**). Vesteinn Thomson's study (Thorsson et al., 2018) indicated that tumors were divided into six immune subtypes, including wound healing (C1), IFN-γ dominant (C2), inflammatory (C3), lymphocyte depleted (C4), immunologically quiet (C5), and TGF-β dominant (C6). Among the six immune subtypes, C3 has the best prognosis. Although there were many immune components, the prognosis of C1 and C2 was still poor. C4 and C6 encompassed mixed immune characteristics and underwent the worst prognosis. C3 accounted for the highest proportion of ccRCC patients, accounting for 87%. Relative to the patients in the low-risk group, the proportion of C3 subtypes with better prognosis was notably decreased in the high-risk group (high-risk group 80%, low-risk group 94%, $p < 0.05$), and C1, C2, C4, and C6 subtypes with poor prognosis were increased (**Figure 8A**).

Concerning drug treatment, the IC50 of sunitinib in the high-risk group was lower (**Figure 8B**), while the IC50 of sorafenib was higher (**Figure 8C**), but no difference in axitinib (**Figure 8D**). Finally, the alluvial diagram presented the distribution of pyroptosis subtypes, pyroptosis-score, and risk groups based on riskScore (**Figure 8E**).

Construction of Nomogram to Predict Survival

Considering the clinical practicability, we established a nomogram combining risk groups based on riskScore and clinicopathological parameters to predict 1-, 3-, and 5-year survival (**Figure 8F**). The predictive parameters were the results of our previous independent prognostic analysis, containing risk, stage, grade, and age. The AUCs of the nomogram were 0.847 and 0.837 (**Figure 8G**) and the C-indexes were 0.784 and 0.795, respectively, in the training and validation sets (**Figure 8H**). Time C-index showed that C-index changed over time and was maintained at about 0.8 (**Figure 8I**). Therefore, our nomogram was dependable in forecasting the prognosis of ccRCC (**Figure 8J**). Next, an ROC curve was plotted to observe the predicted value of the nomogram with or without the risk. The AUC displayed corresponding improvements in the training and validation sets after adding the risk to the predictive model (**Supplementary Figure S6B**). DCA results illustrated the net benefit (NB) assessing the ccRCC patients' outcomes by employing the risk groups based on riskScore, tumor stage, age, gender, grade, or a combination of some features (clinicopathological parameters in nomogram). The results showed that combining the risk groups based on riskScore with tumor stage, grade, and age significantly increased the NB (**Figure 8K**). Finally, based on the model built on the training set, we integrated the data of the TCGA training set and the validation set and compared it with the previous model. The results showed that our model had higher AUC value (**Supplementary Figure S7A**) and C-index (**Supplementary Figure S7B**).

Validation of the Pyroptosis-Related lncRNA

In our study, we established a prognostic model of eight PRlncRNAPs for predicting the prognosis of ccRCC. To further screen PRlncRNA targets, we intersected eight pairs of lncRNAs (13 lncRNAs) with differential genes in ccRCC and obtained five lncRNAs (**Supplementary Figure S8A**). The heatmap showed the expression of five lncRNAs in ccRCC and noncancerous tissues (**Supplementary Figure S8B**). Subsequently, we performed co-expression analysis of the above five lncRNAs with 19 DEPRGs, and got an lncRNA with the highest correlation, AC002331.1, also called LINC02195 (|correlation coefficient| > 0.6, $p < 0.001$, **Supplementary Table S6**). Then, we analyzed the expression of LINC02195 in different grades and stages of ccRCC by UALCAN database. The results suggested that compared to normal condition, all four stages and grades highly expressed LINC02195, showing a statistical significance except stage1 ($p < 0.05$, **Supplementary Figure S8C,D**). In the subtype analysis of ccRCC, we found that the expression level of LINC02195 was higher in ccB-type patients than ccA (**Supplementary Figure S8E**). The survival curves indicated that high expression of LINC02195 meant worse survival (**Supplementary Figure S8F**). Pan-cancer analysis displayed that LINC02195 was significantly upregulated in multiple tumors (**Supplementary Figure S9A**). Then, we detected the expression of LINC02195 in human renal cortical proximal tubule epithelial cell line HK2 and human renal clear cell carcinoma cell lines (786-O, 769-P, CAKI-2, OS-RC-2), and we found that LINC02195 in renal clear cell carcinoma cell line was significantly higher than that in HK2 (**Supplementary Figure S9B**). Finally, LINC02195 was highly expressed in 32 renal clear cell carcinoma tissues than matched adjacent tissues (**Supplementary Figure S9C**).

DISCUSSION

Pyroptosis is a kind of programmed cell death. Unlike apoptosis, pyroptosis is an inflammatory death of cells. When pyroptosis occurs, cells will release inflammatory mediators, which trigger the body's inflammatory response (Liu X. et al., 2021b). More and more studies have proved that pyroptosis may perform a dual function of stimulating or preventing cell growth or death in different tumor cells (Xia et al., 2019). Pyroptosis can promote tumor death to restrain the development of cancer. On the other hand, cells will activate numerous signaling pathways and release a large number of inflammatory mediators when pyroptosis occurs, which are associated with the occurrence and drug resistance of tumors (Zhou and Fang, 2019).

Previous studies have usually focused on a single pyroptosis molecule or a single type of tumor microenvironmental cell. Recently, attention has been paid to the effects of multiple genes on multiple tumor phenotypes. For example, Ye's study investigated the roles of numerous PRGs in colorectal cancer, and such analysis models allow to explore pyroptosis' function in cancer from a whole point of view (Song W. et al., 2021). There have also been studies discussing the role of pyroptosis in ccRCC, like Zhang's research (Zhang X. et al., 2021). They classified ccRCC into four types based on the expression of PRGs. Both subtypes B and C with high PRG expression were enriched in immune cells, but showed two different outcomes of good prognosis and poor prognosis. Zhang speculated that intricate cytokines secreted by cancer cells were the reason for this difference. Different from Zhang's study, we divided ccRCC into three subtypes based on the PRG expression after a comprehensive and systematic analysis. Among them, immune cells were significantly enriched and many immunosuppressive cells were activated in C2 and C3 with high PRG expression, both of which showed poor prognosis. The results of C1 with low PRG expression are opposite to C2 and C3. The results of our analysis were more uniform, which better explained the role of pyroptosis in mediating immune escape.

Among the three subtypes of pyroptosis we identified, it was found that activated dendritic cell, MDSC, macrophage, activated CD4 T cell, activated CD8 T cell, T follicular helper cell, and natural killer T cell were more infiltrated in C2 and C3 subtypes with poor prognosis, and there were more Tregs in C3 subtype. It has been reported that $CD8^+$ T cells infiltrated in RCC are in a state of disability and promote the formation of immune escape (Dai et al., 2021). Treg is a type of T cell with a significant immunosuppressive effect, which can inhibit the immune response of other cells (Hatzioannou et al., 2021). Therefore, it implied that pyroptosis may recruited immune cell infiltration and established an interference and inhibition state in the TME according to previous study and our results.

Immune checkpoints are a class of immunosuppressive molecules that express in immune cells and regulate the degree of immune activation, which prevents autoimmunity, but at the same time may also be the cause of tumor immune escape (Giraldo et al., 2015). Our results revealed that the expression of CTLA4, PDCD1, and PDCD1LG2 in patients with C2 and C3 subtypes was higher than that in C1. Cytotoxic T lymphocyte-associated antigen 4 (CTLA4 or CD152) is an immune checkpoint that negatively regulates T-cell-mediated immune responses by intrinsic and extrinsic mechanisms. CTLA4 delivers a negative signal to effector T cells directly and was mainly associated with functions of Tregs (Lisi et al., 2021). Both PDCD1 (also known as human PD-L1) and PDCD1LG2 (also known as human PD-L2) are ligands of PD-1 (also known as CD274), which inhibit the function of T lymphocytes by binding with PD-1, thereby inhibiting the autoimmune response (Nunes-Xavier et al., 2019). The results of the above analysis showed that the pyroptosis-activated state could regulate the expression of various immune checkpoints, inhibit the functions of various immune cells, and interfere with the clearance of tumors by immune cells, which further indicated that pyroptosis can mediate immune escape of tumors. In order to predict the effect of drug treatment in patients with different subtypes, we constructed the immunophenogram for forecasting anti- PD-1 therapy sensitivity of ccRCC, which presented that C3 patients were in an activated state of pyroptosis and were sensitive to anti-PD-1 therapy and a combination of anti-CTLA4 and anti-PD-1 therapy. However, the outcomes need further verification in the future. The drug treatment response showed that patients in the activated state of pyroptosis were more sensitive to sorafenib and sunitinib. GSVA enrichment analysis, GO, and KEGG analysis of common differential genes between pyroptosis subtypes illustrated that the enrichments of immune-related biological processes in the pyroptosis-activated state were more significant. In a word, targeted pyroptosis may reverse the immune interference and suppression state, thereby enhancing the effect of immunotherapy.

In order to further analyze the regulation of pyroptosis on immunity, we constructed a pyroptosis-score by performing PCA on common differential genes associated with prognosis in the three pyroptosis subtypes. We found that memory-activated CD4 T cell, T follicular helper cell, Tregs, and macrophage M0 cells were more infiltrated in patients with high pyroptosis-score, and pyroptosis-score was positively correlated with immune checkpoints CTLA4, PDCD1, and PDCD1LG2. The above results further illustrated that pyroptosis played an indispensable role in regulating immune activity and mediating tumor immune escape. Patients with high pyroptosis-score were more responsive to sorafenib and sunitinib, which indicated that patients with pyroptosis activation may benefit more from these two targeted therapy agents.

Nowadays, more and more functions of lncRNAs have been discovered (Statello et al., 2021), such as regulating gene expression, posttranscriptional modification, and splicing. It was reported that lncRNAs could regulate the pyroptosis in various tumors, and the researchers also tried to explore the regulatory role of lncRNAs on pyroptosis in ccRCC. For example, one study explored and established a prognosis model of PRlncRNAs (Tang et al., 2021). However, owing to the differences in data processing among different data sets, such a model cannot compare the absolute expression levels of lncRNA well. Therefore, we established a PRlncRNAPs prognostic model which can eliminate differences in data processing and explore its ability to forecast the survival of ccRCC patients. We developed a riskScore based on PRlncRNAPs and divided the patients into high- and low-riskScore groups. Survival analysis indicated that patients with higher riskScore have a worse prognosis. Clinical correlation analysis showed that riskScore was correlated with stage, grade, T, and M, which indicated that the riskScore has perfect prognostic value and clinical value. There were more immune cells infiltrated in patients with high riskScore, such as Tregs, cancer-associated fibroblasts (CAF), T follicular helper cell, memory-activated $CD4^+$ T cell, and macrophage M0 cells, which was consistent with pyroptosis subtypes and pyroptosis-score. Concerning the three targeted drug treatments, patients with high riskScore were more sensitive to sunitinib but less sensitive to sorafenib than patients with low riskScore. However, there was no difference with axitinib. Finally, we integrated various prognosis-related indicators to establish a nomogram. The ROC curve, calibration curve, DCA, and C-index were applied to confirm that the prognostic model can faultlessly forecast OS.

To further screen PRlncRNA targets, we found an lncRNA, LINC02195, which was intensely associated with 19 DEPRGs from eight pairs of lncRNAs related to prognosis by co-expression analysis. Pan-cancer analysis and survival analysis indicated that it was upregulated in various tumors and resulted in poor prognosis. qRT-PCR confirmed that it was highly expressed in ccRCC tissues and cells, providing a theoretical basis for mechanism exploration.

In conclusion, we performed comprehensive and systematic bioinformatics analysis to deeply analyze the function of pyroptosis in ccRCC, distinguished three pyroptosis subtypes, and constructed a PRlncRNAPs model. Pyroptosis activation meant worse prognoses and infiltration of more immunosuppressive cells that was conducive to tumor immune escape and tumor progression. We also predicted its response to current first-line treatment drugs, providing new ideas for clinically guiding ccRCC patients' personalized immune

and targeted therapy strategies. In addition, we developed a constructive and feasible prognostic model based on eight PRlncRNAPs, which performed well in forecasting the prognoses of ccRCC patients and assessing immune cell infiltration in ccRCC. However, several unidentified mechanisms remain to be explored for the interaction between pyroptosis and lncRNAs. Finally, we hope that the analysis of pyroptosis and lncRNAs in this study can provide new strategies for individualized therapy and immunotherapy of ccRCC.

DATA AVAILABILITY STATEMENT

The data sets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

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

MZ analyzed and interpreted the data. MZ, XW, EZ, and LG wrote the manuscript. LF, LZ, KW, and CT edited the manuscript. ZZ and ZW designed and edited the manuscript. All authors contributed to the article and approved the manuscript.

FUNDING

This work was supported by grants to ZZ from National Natural Science Foundation of China (82170690), and Shenzhen Science and Technology Innovation Committee Founding 2019 (33) and 2018 (327), and to CT by the Medical Scientific Research Foundation of Guangdong province (A2018071) and the Guangdong Basic and Applied Basic Research Foundation (2019A1515110488).

ACKNOWLEDGMENTS

Some materials in **Supplementary Figure S1** were modified from Servier Medical Art (http://smart.servier.com/), licensed under a Creative Common Attribution 3.0 Generic License (https://creativecommons.org/licenses/by/3.0/). We acknowledge the TCGA database and GEO database for providing their platforms and contributors for uploading their meaningful data sets.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.905051/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

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 (1), 218. doi:10.1186/s13059-016-1070-5

Bu, X., Liu, J., Ding, R., and Li, Z. (2021). Prognostic Value of a Pyroptosis-Related Long Noncoding RNA Signature Associated with Osteosarcoma Microenvironment. *J. Oncol.* 2021, 2182761. doi:10.1155/2021/2182761

Chen, S., Zhang, E., Guo, T., Shao, J., Wang, T., Zhang, N., et al. (2022). A Novel Ferroptosis-Related Gene Signature Associated with Cell Cycle for Prognosis Prediction in Patients with Clear Cell Renal Cell Carcinoma. *BMC Cancer* 22 (1), 1. doi:10.1186/s12885-021-09033-7

Dai, S., Zeng, H., Liu, Z., Jin, K., Jiang, W., Wang, Z., et al. (2021). Intratumoral CXCL13(+)CD8(+)T Cell Infiltration Determines Poor Clinical Outcomes and Immunoevasive Contexture in Patients with Clear Cell Renal Cell Carcinoma. *J. Immunother. Cancer* 9 (2), e001823. doi:10.1136/jitc-2020-001823

Dunn, J. H., Ellis, L. Z., and Fujita, M. (2012). Inflammasomes as Molecular Mediators of Inflammation and Cancer: Potential Role in Melanoma. *Cancer Lett.* 314 (1), 24–33. doi:10.1016/j.canlet.2011.10.001

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

Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., et al. (2019). Molecular and Pharmacological Modulators of the Tumor Immune Contexture Revealed by Deconvolution of RNA-Seq Data. *Genome Med.* 11 (1), 34. doi:10.1186/s13073-019-0638-6

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.* 40 (4), 1971–1984. doi:10.3892/or.2018.6634

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

Giraldo, N. A., Becht, E., Pagès, F., Skliris, G., Verkarre, V., Vano, Y., et al. (2015). Orchestration and Prognostic Significance of Immune Checkpoints in the Microenvironment of Primary and Metastatic Renal Cell Cancer. *Clin. Cancer Res.* 21 (13), 3031–3040. doi:10.1158/1078-0432.Ccr-14-2926

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

Hatzioannou, A., Boumpas, A., Papadopoulou, M., Papafragkos, I., Varveri, A., Alissafi, T., et al. (2021). Regulatory T Cells in Autoimmunity and Cancer: A Duplicitous Lifestyle. *Front. Immunol.* 12, 731947. doi:10.3389/fimmu.2021.731947

He, M., Li, M., Guan, Y., Wan, Z., Tian, J., Xu, F., et al. (2021). A New Prognostic Risk Score: Based on the Analysis of Autophagy-Related Genes and Renal Cell Carcinoma. *Front. Genet.* 12, 820154. doi:10.3389/fgene.2021.820154

Heagerty, P. J., Lumley, T., and Pepe, M. S. (2000). Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics* 56 (2), 337–344. doi:10.1111/j.0006-341x.2000.00337.x

Hong, W., Liang, L., Gu, Y., Qi, Z., Qiu, H., Yang, X., et al. (2020). Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of

Human Hepatocellular Carcinoma. *Mol. Ther. - Nucleic Acids* 22, 937–947. doi:10.1016/j.omtn.2020.10.002

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

Kerr, K. F., Brown, M. D., Zhu, K., and Janes, H. (2016). Assessing the Clinical Impact of Risk Prediction Models with Decision Curves: Guidance for Correct Interpretation and Appropriate Use. *J. Clin. Oncol.* 34 (21), 2534–2540. doi:10.1200/jco.2015.65.5654

Koren, E., and Fuchs, Y. (2021). Modes of Regulated Cell Death in Cancer. *Cancer Discov.* 11 (2), 245–265. doi:10.1158/2159-8290.Cd-20-0789

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

Lisi, L., Lacal, P. M., Martire, M., Navarra, P., and Graziani, G. (2022). Clinical Experience with CTLA-4 Blockade for Cancer Immunotherapy: From the Monospecific Monoclonal Antibody Ipilimumab to Probodies and Bispecific Molecules Targeting the Tumor Microenvironment. *Pharmacol. Res.* 175, 105997. doi:10.1016/j.phrs.2021.105997

Liu, J., Yao, L., Zhang, M., Jiang, J., Yang, M., and Wang, Y. (2019). Downregulation of LncRNA-XIST Inhibited Development of Non-small Cell Lung Cancer by Activating miR-335/SOD2/ROS Signal Pathway Mediated Pyroptotic Cell Death. *Aging* 11 (18), 7830–7846. doi:10.18632/aging.102291

Liu, S. J., Dang, H. X., Lim, D. A., Feng, F. Y., and Maher, C. A. (2021). Long Noncoding RNAs in Cancer Metastasis. *Nat. Rev. Cancer* 21 (7), 446–460. doi:10.1038/s41568-021-00353-1

Liu, X., Xia, S., Zhang, Z., Wu, H., and Lieberman, J. (2021b). Channelling Inflammation: Gasdermins in Physiology and Disease. *Nat. Rev. Drug Discov.* 20 (5), 384–405. doi:10.1038/s41573-021-00154-z

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 (12), 550. doi:10.1186/s13059-014-0550-8

Lv, W., Tan, Y., Zhao, C., Wang, Y., Wu, M., Wu, Y., et al. (2021). Identification of Pyroptosis-Related lncRNAs for Constructing a Prognostic Model and Their Correlation with Immune Infiltration in Breast Cancer. *J. Cell Mol. Med.* 25 (22), 10403–10417. doi:10.1111/jcmm.16969

Nagarajan, K., Soundarapandian, K., Thorne, R. F., Li, D., and Li, D. (2019). Activation of Pyroptotic Cell Death Pathways in Cancer: An Alternative Therapeutic Approach. *Transl. Oncol.* 12 (7), 925–931. doi:10.1016/j.tranon.2019.04.010

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

Numakura, K., Muto, Y., Naito, S., Hatakeyama, S., Kato, R., Koguchi, T., et al. (2021). Outcomes of Axitinib versus Sunitinib as First-Line Therapy to Patients with Metastatic Renal Cell Carcinoma in the Immune-Oncology Era. *Cancer Med.* 10 (17), 5839–5846. doi:10.1002/cam4.4130

Nunes-Xavier, C. E., Angulo, J. C., Pulido, R., and López, J. I. (2019). A Critical Insight into the Clinical Translation of PD-1/pd-L1 Blockade Therapy in Clear Cell Renal Cell Carcinoma. *Curr. Urol. Rep.* 20 (1), 1. doi:10.1007/s11934-019-0866-8

Ping, L., Zhang, K., Ou, X., Qiu, X., and Xiao, X. (2021). A Novel Pyroptosis-Associated Long Non-Coding RNA Signature Predicts Prognosis and Tumor Immune Microenvironment of Patients with Breast Cancer. *Front. Cell Dev. Biol.* 9, 727183. doi:10.3389/fcell.2021.727183

Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E., and Gfeller, D. (2017). Simultaneous Enumeration of Cancer and Immune Cell Types from Bulk Tumor Gene Expression Data. *Elife* 6, e26476. doi:10.7554/eLife.26476

Ren, N., Jiang, T., Wang, C., Xie, S., Xing, Y., Piao, D., et al. (2020). LncRNA ADAMTS9-AS2 Inhibits Gastric Cancer (GC) Development and Sensitizes Chemoresistant GC Cells to Cisplatin by Regulating miR-223-3p/NLRP3 Axis. *Aging* 12 (11), 11025–11041. doi:10.18632/aging.103314

Song, S., Liu, S., Wei, Z., Jin, X., Mao, D., He, Y., et al. (2021). Identification of an Immune-Related Long Noncoding RNA Pairs Model to Predict Survival and Immune Features in Gastric Cancer. *Front. Cell Dev. Biol.* 9, 726716. doi:10.3389/fcell.2021.726716

Song, W., Ren, J., Xiang, R., Kong, C., and Fu, T. (2021). Identification of Pyroptosis-Related Subtypes, the Development of a Prognosis Model, and Characterization of Tumor Microenvironment Infiltration in Colorectal Cancer. *Oncoimmunology* 10 (1), 1987636. doi:10.1080/2162402X.2021.1987636

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

Sun, Z., Li, T., Xiao, C., Zou, S., Zhang, M., Zhang, Q., et al. (2022). Prediction of Overall Survival Based upon a New Ferroptosis-Related Gene Signature in Patients with Clear Cell Renal Cell Carcinoma. *World J. Surg. Onc.* 20 (1), 120. doi:10.1186/s12957-022-02555-9

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 (3), 209–249. doi:10.3322/caac.21660

Tang, R., Xu, J., Zhang, B., Liu, J., Liang, C., Hua, J., et al. (2020). Ferroptosis, Necroptosis, and Pyroptosis in Anticancer Immunity. *J. Hematol. Oncol.* 13 (1), 110. doi:10.1186/s13045-020-00946-7

Tang, X., 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

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. e14. doi:10.1016/j.immuni.2018.03.023

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

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.* 10 (9), 650. doi:10.1038/s41419-019-1883-8

Xiong, Y., Liu, L., Bai, Q., Xia, Y., Qu, Y., Wang, J., et al. (2020). Individualized Immune-Related Gene Signature Predicts Immune Status and Oncologic Outcomes in Clear Cell Renal Cell Carcinoma Patients. *Urologic Oncol. Seminars Orig. Investigations* 38 (1), e1–e17. doi:10.1016/j.urolonc.2019.09.014

Zhang, F., Yu, S., Wu, P., Liu, L., Wei, D., and Li, S. (2021). Discovery and Construction of Prognostic Model for Clear Cell Renal Cell Carcinoma Based on Single-Cell and Bulk Transcriptome Analysis. *Transl. Androl. Urol.* 10 (9), 3540–3554. doi:10.21037/tau-21-581

Zhang, L., Su, Z., Hong, F., and Wang, L. (2022). Identification of a Methylation-Regulating Genes Prognostic Signature to Predict the Prognosis and Aid Immunotherapy of Clear Cell Renal Cell Carcinoma. *Front. Cell Dev. Biol.* 10, 832803. doi:10.3389/fcell.2022.832803

Zhang, X., Wei, X., Wang, Y., Wang, S., Ji, C., Yao, L., et al. (2021). Pyroptosis Regulators and Tumor Microenvironment Infiltration Characterization in Clear Cell Renal Cell Carcinoma. *Front. Oncol.* 11, 774279. doi:10.3389/fonc.2021.774279

Zhou, C.-B., and Fang, J.-Y. (2019). The Role of Pyroptosis in Gastrointestinal Cancer and Immune Responses to Intestinal Microbial Infection. *Biochimica Biophysica Acta (BBA) - Rev. Cancer* 1872 (1), 1–10. doi:10.1016/j.bbcan.2019.05.001

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 Zhong, Wang, Zhu, Gong, Fei, Zhao, Wu, Tang, Zhang, Wang and Zheng. 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: 30 June 2022
doi: 10.3389/fgene.2022.888764

Identification of a Prognostic Model Based on Fatty Acid Metabolism-Related Genes of Head and Neck Squamous Cell Carcinoma

Peiyu Du[1†], Yue Chai[2†], Shimin Zong[1], Jianxin Yue[1] and Hongjun Xiao[1]**

[1]Department of Otorhinolaryngology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [2]Department of Medical Oncology, National Cancer Cente, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

The fatty acid metabolism (FAM) is known to impact tumorigenesis, tumor progression and treatment resistance via enhancing lipid synthesis, storage and catabolism. However, the role of FAM in head and neck squamous cell carcinoma (HNSCC) has remained elusive. In the present study, we obtained a total of 69 differentially expressed FAM-related genes between 502 HNSCC samples and 44 normal samples from The Cancer Genome Atlas (TCGA) database. The HNSCC samples were divided into 2 clusters according to 69 differentially expressed genes (DEGs) via cluster analysis. Then DEGs in the two clusters were found, and 137 prognostic DEGs were identified by univariate analysis. Subsequently, combined with the clinical information of 546 HNSCC patients from TCGA database, a 12-gene prognostic risk model was established (FEPHX3, SPINK7, FCRLA, MASP1, ZNF541, CD5, BEST2 and ZAP70 were down-regulation, ADPRHL1, DYNC1I1, KCNG1 and LINC00460 were up-regulation) using multivariate Cox regression and LASSO regression analysis. The risk scores of 546 HNSCC samples were calculated. According to the median risk score, 546 HNSCC patients were divided into the high- and low-risk (high- and low score) groups. The Kaplan-Meier survival analysis showed that the survival time of HNSCC patients was significantly shorter in the high-risk group than that in the low-risk group ($p < 0.001$). The same conclusion was obtained in the Gene Expression Omnibus (GEO) dataset. After that, the multivariate Cox regression analysis indicated that the risk score was an independent factor for patients with HNSCC in the TCGA cohort. In addition, single-sample gene set enrichment analysis (ssGSEA) indicated that the level of infiltrating immune cells was relatively low in the high-risk group compared with the low-risk group. **In summary**, FAM-related gene expression-based risk signature could predict the prognosis of HNSCC independently.

Keywords: fatty acid metabolism, head and neck squamous cell carcinoma, prognosis, risk signature, lasso-cox regression

OPEN ACCESS

Edited by:
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

Reviewed by:
Hiroaki Inaba,
Okayama University, Japan
Francisco José Caramelo,
University of Coimbra, Portugal

****Correspondence:***
Jianxin Yue
yuejianxin@live.cn
Hongjun Xiao
xhjent_whxh@hust.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: *03 March 2022*
Accepted: *09 June 2022*
Published: *30 June 2022*

Citation:
Du P, Chai Y, Zong S, Yue J and Xiao H (2022) Identification of a Prognostic Model Based on Fatty Acid Metabolism-Related Genes of Head and Neck Squamous Cell Carcinoma. Front. Genet. 13:888764. doi: 10.3389/fgene.2022.888764

Abbreviations: FAM, fatty acid metabolism; HNSCC, head and neck squamous cell carcinoma; TCGA, The Cancer Genome Atlas database; GEO, Gene Expression Omnibus; DEGs, differentially expressed genes; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; AUC, areas under the curve; PCA, Principal Component Analysis; OS, overall survival; t-SNE, t-distributed stochastic neighbor embedding; ssGSEA, single-sample gene set enrichment analysis.

1 INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC), the most common pathological type of head and neck cancer, arises from the mucosal epithelium in the oral cavity, pharynx, and larynx (Johnson et al., 2020). HNSCC is the 6th most common cancer worldwide (Sung et al., 2021). The onset of HNSCC is increasing with approximately 600,000 new cases are diagnosed each year (Solomon et al., 2018). Oralcavity and larynx cancers are often connected with factors such as smoking and alcohol abuse, or both, whereas pharynx cancers are attributable to human papillomavirus (HPV) infection, especially HPV-16. Therefore, HNSCC is divided into 2 groups: HPV positive HNSCC with better prognosis and HPV negative with worse prognosis (Johnson et al., 2020). Early-stage HNSCC is typically asymptomatic and most HNSCC is already locally advanced or advanced when diagnosed. As a result of no early effective detection test or screening method for HNSCC being available currently, a careful physical examination has far remained the primary diagnostic tool in early detection of HNSCC. At present, treatment approaches for HNSCC are mainly comprehensive modalities with surgery, radiotherapy, chemotherapy, targeted therapy, immunotherapies, et al. However, the overall responses rates of patients with HNSCC to targeted therapeutic drugs, such as epidermal growth factor receptor (EGFR) inhibitors (cetuximab) and programmed cell death protein 1 inhibitors (pembrolizumab and nivolumab), were moderate (Zhang et al., 2021). Thus, finding a novel prognostic biomarker for the treatment of HNSCC is a major goal. The fatty acid metabolism (FAM) occupies a key role in the entire lipid metabolism, which is known to impact tumorigenesis, tumor progression and treatment resistance via enhancing lipid synthesis, storage and catabolism (Fernández et al., 2020; Hoy et al., 2021). The biosynthesis of fatty acids (FAs) is activated in cancer cells to fulfill lipid synthesis of membranes and signaling molecules, and energy storage (Fhu and Ali, 2020). Compared with glucose oxidation, fatty acid (FA) oxidation is more likely to be utilized by tumor cells due to higher energy production (Li and Zhang, 2016). Increased FA oxidation confers survival advantages for tumors not only to resist chemotherapeutic and radiation treatments but also to alleviate cellular stresses involved in the metastatic cascade (Corn et al., 2020). In addition, FAM could have an impact on cellular phenotype and function of tumor-infiltrating immune cells in tumor microenvironments, which may be associated with immunosuppression (Corn et al., 2020). Deregulating or blocking FAs levels through the FAM-related pathway in cancer might inhibit tumor cell growth and therefore that identifying the targets to regulate FAM is essential. Several studies have shown that FAM was closely related to the oncogenesis and progression of breast cancer, colorectal cancer and liver cancer (Wang Y.-n. et al., 2018; Madak-Erdogan et al., 2019; He et al., 2021; Hofmanová et al., 2021). A previous study has demonstrated that the FA synthase inhibitor orlistat could reduce the growth and metastasis of orthotopic tongue oral squamous cell carcinomas (Agostini et al., 2014). Su et al. found that the overexpression of long-chain acyl-CoA dehydrogenase, a type of FA oxidation-related enzyme, had a protective effect on OS in advanced HNSCC(Su et al., 2020). To date, the role of FAM in head and neck squamous cell carcinoma has remained elusive.

In the present research, bioinformatics analysis was performed to investigate FAM-related genes that may play an important role in HNSCC and constructed a prognostic model for patients with HNSCC. The association between the expression of FAM-related genes and the prognosis of HNSCC was further analyzed via this model to explore the potential prognostic value of FAM-related genes and provide additional evidence for potential therapeutic targets for HNSCC.

2 METHODS

2.1 Data Processing

Gene sets related to FAM were downloaded from the gene set enrichment analysis (GSEA, http://www.gsea-msigdb.org/gsea/msigdb/index.jsp) website. The RNA-sequencing (RNA-Seq) data of patients with HNSCC were searched in The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/repository) and Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo). Their clinicopathological, genetic, epigenetic, and survival data were downloaded for secondary analysis.

2.2 Identification of Differentially Expressed Fatty Acid Metabolism-Related Genes

The R language "limma"package was used to detect differentially expressed genes (DEGs) between tumor and normal tissues of patients with HNSCC. A p value <0.05 was used as the screening criterion. HNSCC patients were clustered into two subgroups based on prognosis-based FAM-related genes which were identified by univariate Cox regression analysis using the R "**limma**" "survival" "**ConsensusClusterPlus**" packages. After that, survival analyses were performed in the 2 subgroups using the **"survival**" "**survminer**" packages in R. The p values <0.001 were considered statistically significant. DEGs between the two subgroups were analyzed by R "**limma**" package, quantified as the logFC >1 by R "**sva**" package, and plotted by R "**ggplot**" package. Heatmaps of these DEGs between the two subgroups were drawn with R "**pheatmap**" package.

2.3 Development and Validation of the Fatty Acid Metabolism-Based Prognostic Model

The samples from the TCGA database was acted as the training set to construct a prognostic model while the samples from the GEO database were designated as the test set to verify the accuracy of the prognostic model. The key prognosis-related FAMs were selected by the LASSO Cox regression analysis and multivariate Cox regression analysis via the R "**glmnet**" package. The risk score formula for the prediction of prognosis of patients with HNSCC was as follows: The risk score calculating formula is:

$$\textit{Risk score} = \sum_{i=1}^{n} \textit{Coefi} * \textit{Expi},$$

where the n refers to the number of signature genes, Coefi refers to the coefficients, Expi refers to gene expression level. Then risk scores of HNSCC patients were obtained according to the risk model. Based on the median value of risk scores, HNSCC patients were divided into high- and low-risk subgroups. The R **"survival"** and **"survmine"** packages were used for survival analysis. Principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) were used for dimensionality reduction analysis via R "Rtsne" and "ggplot2" packages. The receiver operating characteristic (ROC) curve was drawn by R "timeROC" package to evaluate the accuracy of the risk model. The risk heat plot was plotted by the R "pheatmap" package. HNSCC cohort from the GEO database (GSE41613) was employed to validate the risk model. The differentially expressed FAM-related genes were obtained and the risk scores were calculated by the same methods used for the TCGA cohort. The patients with HNSCC in the cohort were also divided into high- or low-risk subgroups, and these subgroups were then compared to validate the risk model.

2.4 Human Protein Atlas

The protein expression levels of the hub genes were validated using the Human Protein Atlas (HPA) database (https://www. proteinatlas. org). HPA is a Swedish-based program initiated in 2003 for providing all human proteins in cells, tissues, and organs. Getting immunohistochemical data of patients with or without HNSCC from HPA helped us to further verify the protein expression levels of 12-genes that were identified in the risk model.

2.5 Independent Prognostic Analysis of the Risk Score

Univariate and multivariable Cox regression analyses including clinical characteristics and risk scores were performed to evaluate whether our ricsk score could serve as an independent prognostic factor for HNSCC and were visualized as forest plots and heatmaps using R "limma" and "pheatmap" packages. A p value <0.05 was considered significant.

2.6 Functional Enrichment Analysis

The DEGs of HNSCC patients in the high- and low-risk groups in the TCGA cohorts were calculated according to the risk model. Gene Ontology (GO) enrichment analysis of DEGs was performed to obtain the enrichment of DEGs regarding biological process, molecular function and cellular component, and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was carried out to obtain signal pathways of DEGs using the R "clusterProfiler" and "enrichplot" packages (p value > 0.05 and q value < 0.05) (Ashburner et al., 2000; Kanehisa et al., 2017). The functional enrichment results of GO and KEGG analyses were visualized as bubble charts and barplots.

2.7 Estimation of Immune Status

Single-sample GSEA (ssGSEA) enrichment scores were calculated by the "GSVA" package of R to quantified the immune infiltration levels of the DEGs of HNSCC patients in the high- and low-risk groups in the TCGA cohorts. Patients were scored according to the risk model that was established in our study. The median value of the risk score was as a cut-off point. The high-risk group consisted of patients with risk scores above the median value, while the low-risk group consisted of patients with risk scores below the median value.

2.8 Statistical Analysis

Categorical variables between normal and HNSCC tissues were compared by the Pearson chi-square test. Survival analyses were performed by Kaplan-Meier Log-rank analyses. The differences in immune cell infiltration and immune pathway activation between the two subgroups were ascertained by the Mann-Whitney test. Statistical analyses were performed using the R (version 4.1.1), R base packages and R Bioconductor packages.

3 RESULTS

3.1 Identification of Differentially Expressed Genes Between Normal and Tumor Tissues

The flow diagram was presented in **Figure 1**.103 FAM-related genes were obtained from GSEA website (gene set ID: GO_FATTY_ACID_CATABOLIC_PROCESS), and they were shown in **Supplementary Table S1**. In this study, the datasets of 502 tumor samples (including SCC of the oral cavity, larynx, oropharynx, hypopharynx, jaw, lip, and other sites of unclear origin) and 44 normal samples were from TCGA database. Based on our criteria of $p < 0.05$, a total of 69 DEGs were identified from 103 FAM-related genes (**Figure 2**).

3.2 Tumor Classification Based on the Fatty Acid Metabolism-Related Differentially Expressed Genes

Based on expression data of 69 FAM-related DEGs, 546 HNSCC patients in the TCGA cohort were robustly separated into 2 subgroups via a consensus cluster analysis with the clustering variable (k) of 2 (**Figure 3A**). The gene expression profile of 69 FAM-related genes and the clinical features including the TNM stage, T stage, N stage, M stage, grade (G1-G4), gender, and age (>65 or ≤65 years) were presented in a heatmap, and there were significant differences across 2 subgroups with respect to TNM stage, N stage and grade ($p < 0.05$)(**Figure 3B**). The overall survival (OS) was significantly different between the two clusters, the median survival values obtained by the Kaplan-Meier analysis is 2.57 years in cluster 1, while it is 7.04 years in cluster 2. There were differences in clinical phenotype, prognosis, and expression of FAM-related genes between Cluster 1 and Cluster 2. ($p < 0.001$, **Figure 3C**).

3.3 Construction of a Prognostic Model of Head and Neck Squamous Cell Carcinoma Based on Fatty Acid Metabolism-Related Genes

First, a total of 546 HNSCC samples' expression data from TCGA database and survival data were merged. Second, univariate Cox regression analysis showed that 137 FAM-related genes were related to the prognosis of HNSCC ($p < 0.05$). Among them, 30

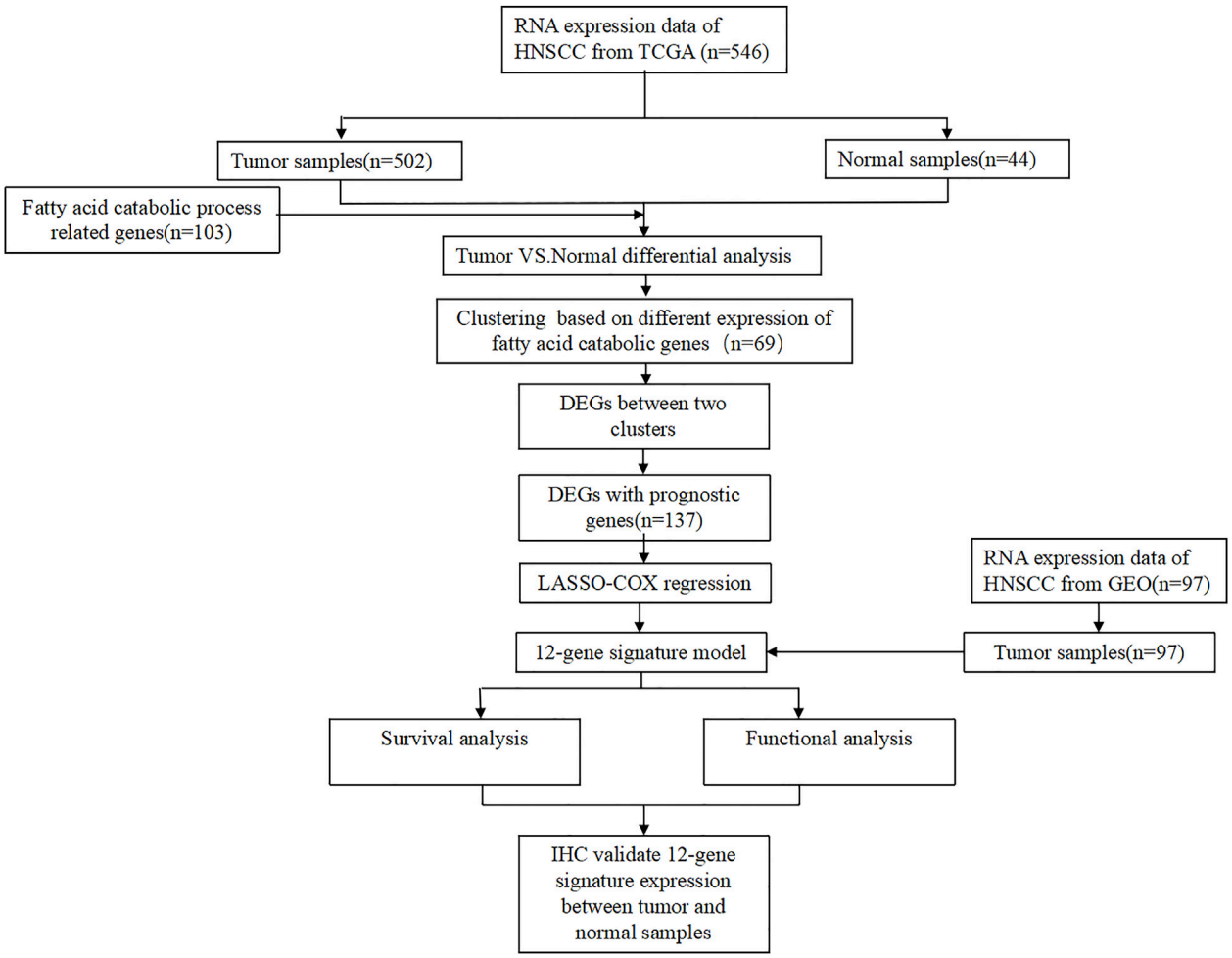


FIGURE 1 | 103 FAM-related genes were obtained from GSEA website.

genes were associated with increased risk with hazard ratios (HRs) > 1, while the other 131 genes were protective genes with HRs <1 (**Figure 4A**). Then, LASSO Cox regression analysis was conducted to identify the genes with the best prognostic value, and 12 optimal coefficients were obtained according to the optimum λ value at last (**Figures 4B,C**). The risk score was calculated based on the expression of the 12 genes and the risk score formula were as follows: $\boldsymbol{Risk\ score} = \sum_{i=1}^{n} \boldsymbol{Coefi} * \boldsymbol{Expi}$. 546 patients were divided into low- and high-risk subgroups by using the median score calculated by the risk score formula (**Figure 4D**). The PCA and t-SNE analyses further identified that HNSCC patients in different risk groups were well divided into two clusters (**Figure 4E**, **Figure 4F**). Patients in the high-risk group were more likely to have more deaths and a shorter survival time than those in the low-risk group (**Figure 4G**). The Kaplan-Meier survival analysis demonstrated notable survival discrepancies between the low- and high-risk subgroups ($p < 0.001$, **Figure 4H**). A ROC curve was applied to verify the sensitivity and specificity of the model. The areas under the curve (AUC) of the risk signature were 0.669 for 1-year, 0.718 for 2-years, and 0.662 for 5-years survival (**Figure 4I**).

3.4 Validation of the Risk Signatures

A total of 97 HNSCC patients from a GEO cohort were analyzed to verify the predictive performance of the risk model. Based on the median risk score in the training cohort, 97 patients in the GEO cohort were separated into the low-risk and high-risk subgroups (**Figure 5A**). The PCA and t-SNE analyses demonstrated significant separation between the two subgroups (**Figure 5B**, **Figure 5C**). Patients in the high-risk group were more likely to have more deaths and shorter survival time than those in the low-risk group. The median survival time was 2.02 years in the high-risk group and 5.04 years in the low-risk group (**Figure 5D**). Furthermore, the Kaplan-Meier analysis showed a significant difference in the OS between the low-and high-risk subgroups ($p = 0.002$, **Figure 5E**). In addition, ROC curve analysis of the GEO cohort revealed that the prognostic value of our model was high (AUC = 0.686 for 1-year, 0.673 for 2-years, and 0.688 for 5-years survival)

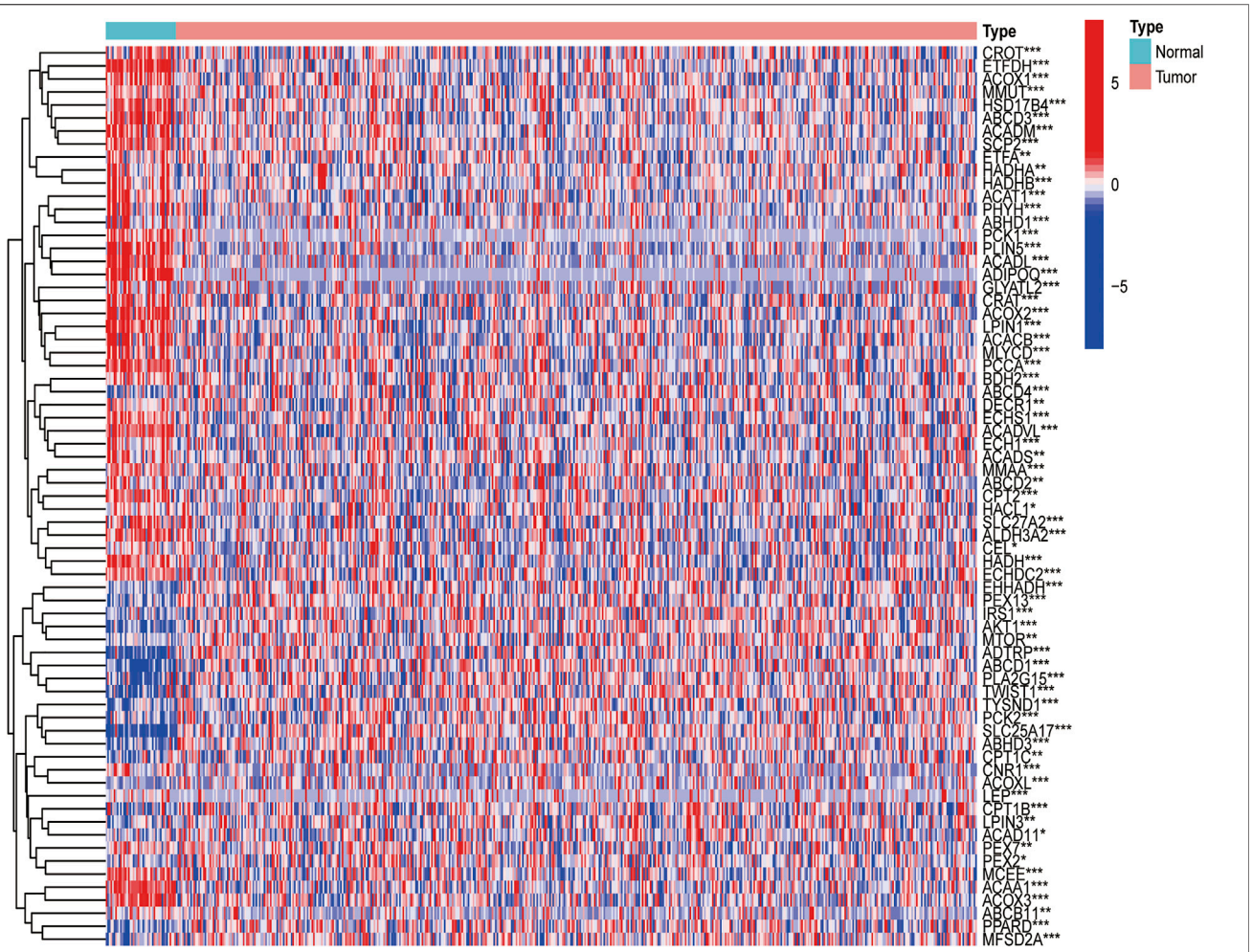


FIGURE 2 | Expressions of the 33 FAM-related genes and the interactions among them. Heatmap (green: low expression level; red: high expression level) of the FAM-related genes between the normal (N, brilliant blue) and the tumour tissues (T, red). *p* values were showed as: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

(**Figure 5F**). EPHX3, SPINK7, FCRLA, MASP1, and CD5 had lower expression levels in HNSCC tissues, DYNC1I1 and KCNG1 had the higher expression levels by immunohistochemistry (IHC) analysis. The expression of ZNF541, BEST2, ADPRHL1, and ZAP70 were negative in both HNSCC tissues and normal tissues. The expression of LINC00460 in HNSCC tissues and normal tissues was not available (**Figure 5G**).

3.5 Independent Prognostic Value of the Risk Signature

The univariate Cox regression analysis indicated that the risk score was a prognostic factor for patients with HNSCC in the TCGA cohort (HR = 7.181, 95% CI: 2.596–19.863, **Figure 6A**). The multivariate analysis showed that the N stage, M stage, and the risk score (HR = 4.650, 95% CI: 1.617–13.371) were independent prognostic factors for patients with HNSCC in the TCGA cohort ($p < 0.05$, **Figure 6B**). Additionally, the TNM stage, T stage and N stage were found to be significantly different between the low- and high-risk subgroups ($p < 0.05$) (**Figure 6C**).

3.6 Functional Analyses Based on the Risk Model

A total of 12 DEGs between the low- and high-risk subgroups in the TCGA cohort were found. Among them, 8 genes (FEPHX3, SPINK7, FCRLA, MASP1, ZNF541, CD5, BEST2 and ZAP70) were downregulated in the high-risk group, while the other 4 genes (ADPRHL1, DYNC1I1, KCNG1 and LINC00460)were upregulated. The results of GO functional enrichment analysis showed that these 12 DEGs were mainly enriched in the defense response to bacterium in the category of biological processes and the external side of plasma membrane in the category of cellular components. In terms of the category of molecular function, 12 DEGs were mainly enriched in the antigen binding, serine-type endopeptidase activity, serine-type peptidase activity, serine hydrolase activity, peptidase regulator activity and immunoglobulin receptor (**Figure 7A**, **Figure 7B**). KEGG enrichment analysis showed that hematopoietic cell lineage and cytokine-cytokine receptor interaction was the main pathway that DEGs involved, followed by Th1 and Th2 cell

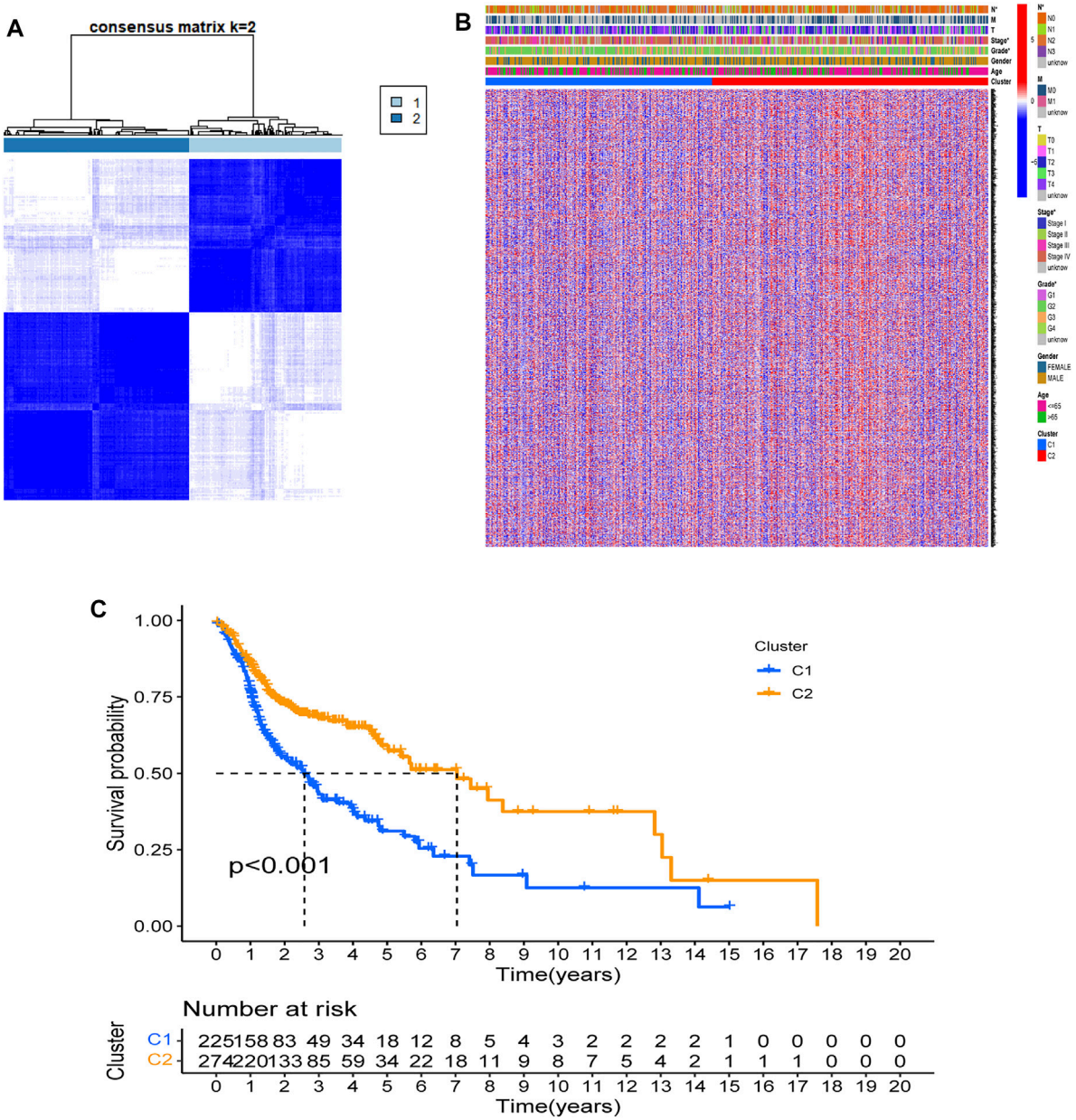


FIGURE 3 | Tumor classification based on the FAM-related DEGs. **(A)** 546 HNSCC patients were grouped into two clusters according to the consensus clustering matrix (k = 2). **(B)**. Heatmap and the clinicopathologic characters of the two clusters classified by these DEGs. Two samples were more likely to be grouped into the same cluster when there was a higher consensus score between them in different iterations. **(C)**. Kaplan-Meier OS curves for the two clusters.

differentiation, Th17 cell differentiation, cell adhesion molecules and human T-cell leukemia virus infection (**Figure 7C**, **Figure 7D**).

3.7 Comparison of the Immune Activity Between High- and Low-Risk Subgroups

The enrichment scores of 16 types of immune cells and the activity of 13 immune-related pathways between the low- and high-risk groups in both the TCGA and GEO cohorts were evaluated by the ssGSEA method. 16 immune cells showed significantly higher levels of infiltration in the low-risk subgroup than the high-risk subgroup in the TCGA cohort, while 13 immune cells, including activated dendritic cells (aDCs), B cells, $CD8^+$ T cells, dendritic cells (DCs), induced DCs (iDCs), mast cells, neutrophils, natural killer (NK) cells, plasmacytoid DCs (pDCs), T helper (Th) cells, tumor-infiltrating lymphocytes (TILs) and regulatory T (Treg) cells, were in the GEO cohorts ($p < 0.05$) (**Figure 8A**, **Figure 8C**). In the 13 immune pathways, except for the parainflammation and type-

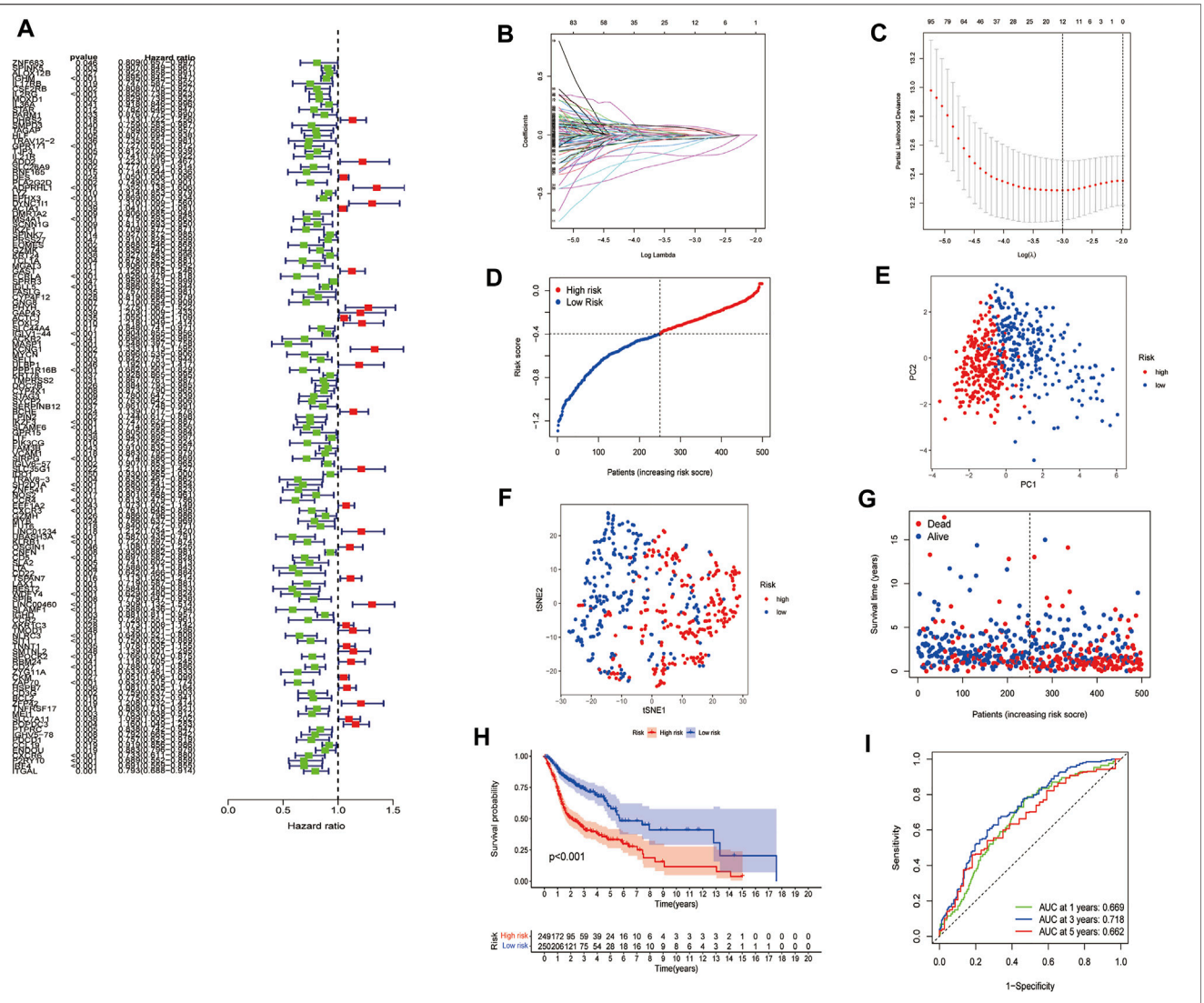


FIGURE 4 | Construction of risk signature in the TCGA cohort. **(A)**. Univariate cox regression analysis of OS for each FAM-related gene. **(B)**. LASSO regression of the 12 OS-related genes. **(C)**. Cross-validation for tuning the parameter selection in the LASSO regression. **(D)**. Distribution of patients based on the risk score. **(E)**. PCA plot for HNSCC based on the risk score. **(F)**. t-SNE plot for HNSCC based on the risk score. **(G)**. The survival status for each patient. **(H)**. Kaplan-Meier curves for the OS of patients in the high- and low-risk groups. **(I)**. ROC curves demonstrated the predictive efficiency of the risk score.

1 IFN response pathways, the remaining 11 pathways showed significantly higher activity in the low-risk group than in the high-risk group in the TCGA cohort ($p < 0.001$, **Figure 8B**). As for the GEO cohort, 9 out of 13 immune pathways illustrated higher activity in the low-risk group than in the high-risk group ($p < 0.05$). Among them, the cytolytic activity pathway, HLA pathway and T cell costimulation pathway showed significantly higher activity. ($p < 0.001$, **Figure 8D**).

DISCUSSION

In our research, we found 12 FAM-related genes related to the prognosis of HNSCC, and further constructed a prognostic risk model based on these genes to predict the prognosis of patients with HNSCC for the first time. Compared with single-gene biomarkers, the prognostic risk model composed of multiple genes was more accurate in predicting the prognosis of patients with HNSCC, thereby providing effective help for improving the therapeutic effect and prolonging survival of patients with HNSCC.

Studies have demonstrated that several FAM-related enzymes were increased in HNSCC compared with non-malignant tumors, but the relationship between FAM-related genes and the prognosis of HNSCC has remained unclear (Su et al., 2020). In the present study, using transcriptome data of 502 HNSCC samples and 44 normal samples released by TCGA database, we obtained a total of 69 DEGs FAM-related genes. The samples were divided into 2 clusters according to 69 DEGs via cluster analysis. Then DEGs in the two clusters were found, and 137

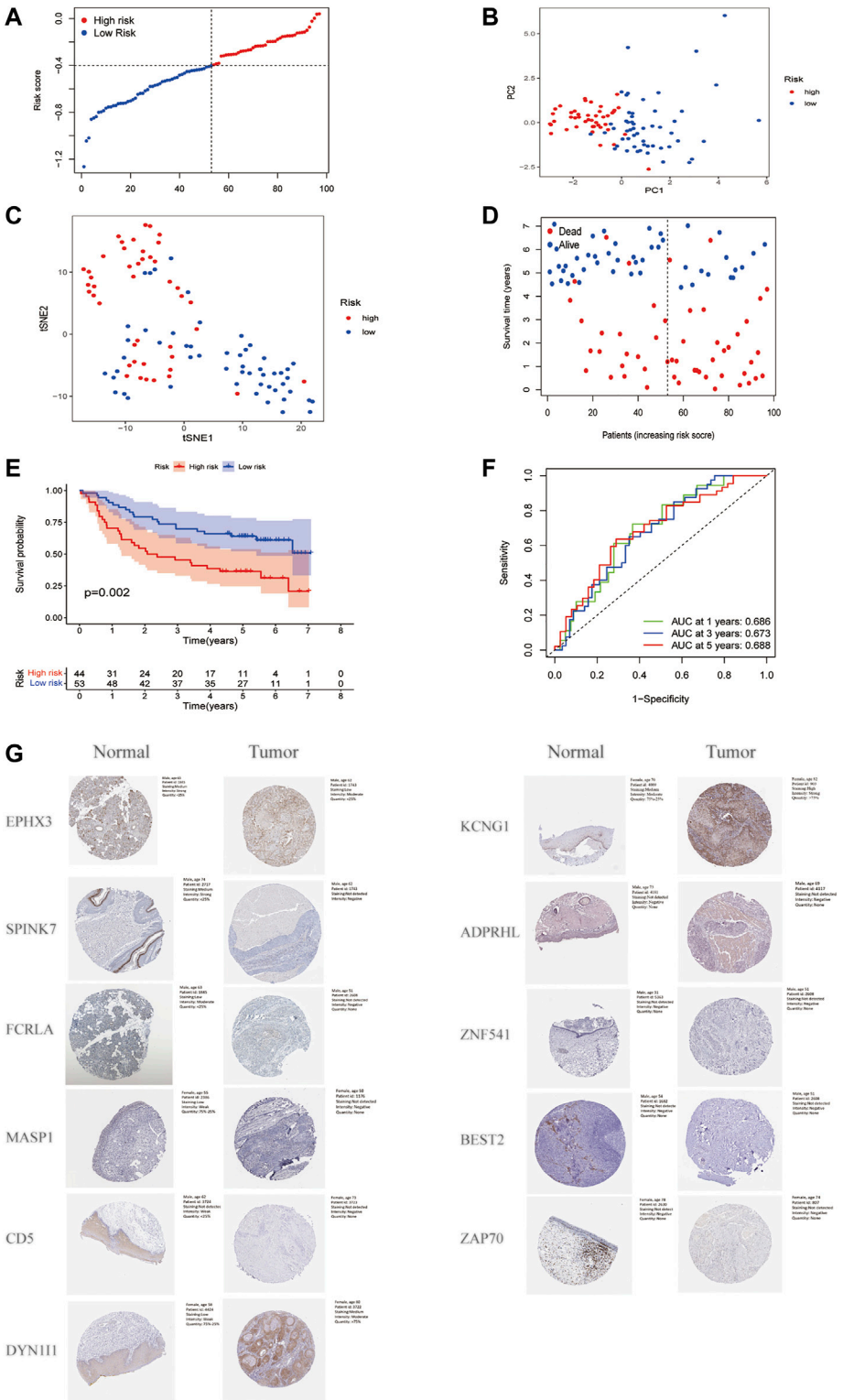


FIGURE 5 | Validation of the risk model in the GEO cohort. **(A)**. Distribution of patients in the GEO cohort based on the median risk score in the TCGA cohort. **(B)**. PCA plot for HNSCC. **(C)**. t-SNE plot for HNSCC. **(D)**. The survival status for each patient. **(E)**. Kaplan-Meier curves for comparison of the OS between low-and high-risk groups. **(F)**. Time-dependent ROC curves for HNSCC. **(G)**. Verification of FAM-related gene expressions in normal and tumor tissue utilizing the Human Protein Atlas (HPA).

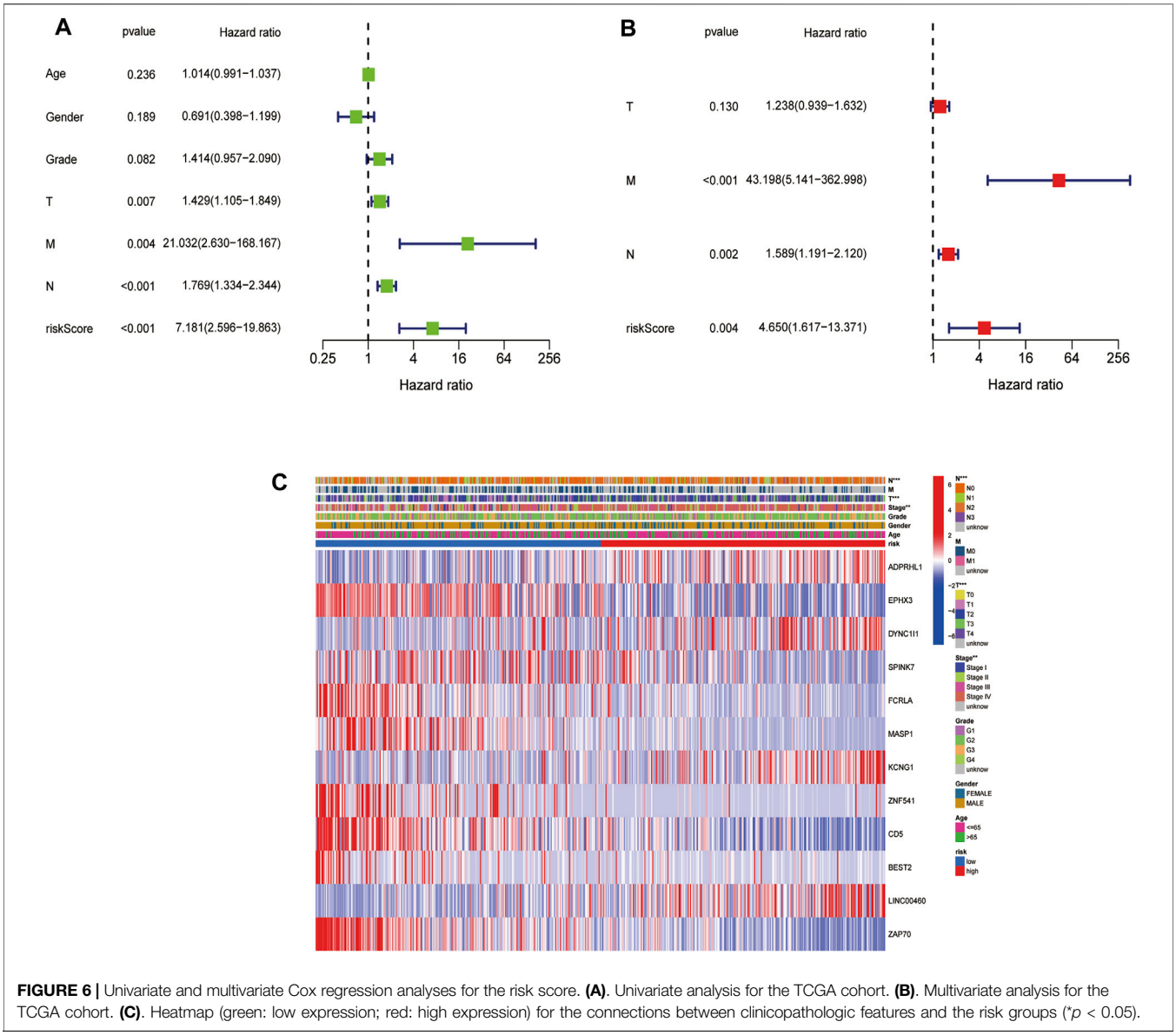


FIGURE 6 | Univariate and multivariate Cox regression analyses for the risk score. **(A)**. Univariate analysis for the TCGA cohort. **(B)**. Multivariate analysis for the TCGA cohort. **(C)**. Heatmap (green: low expression; red: high expression) for the connections between clinicopathologic features and the risk groups (*$p < 0.05$).

prognostic DEGs were identified by univariate analysis. Subsequently, combined with the clinical information of 546 HNSCC patients from TCGA database, a 12-gene prognostic risk model was established (FEPHX3, SPINK7, FCRLA, MASP1, ZNF541, CD5, BEST2 and ZAP70 were down-regulation, ADPRHL1, DYNC1I1, KCNG1 and LINC00460 were up-regulation) using multivariate Cox regression and LASSO regression analysis. The risk scores of 546 HNSCC samples were calculated. According to the median risk score, 546 HNSCC patients were divided into the high- and low-risk groups. The Kaplan-Meier survival analysis showed that the survival time of HNSCC patients was significantly shorter in the high-risk group than that in the low-risk group. The ROC curve (AUC value >0.60) verified the accuracy of the prognostic risk model. The same conclusion was obtained in the GEO dataset. After that, the multivariate Cox regression analysis indicated that the risk score was an independent factor for patients with HNSCC in the TCGA cohort. In addition, ssGSEA functional enrichment analysis was conducted to further investigate the relationship between the prognostic risk model and signaling pathways as well as immune status. The level of infiltrating immune cells was relatively low and the activity of immune-related pathways was decreased in the high-risk group compared with the low-risk group.

This study generated a signature featuring 12 FAM-related genes (ADPRHL1, DYNC1I1, KCNG1, LINC00460, EPHX3, SPINK7, FCRLA, MASP1, ZNF541, CD5, BEST2, ZAP70) and found that it could predict survival in patients with HNSCC. Among them, 8 of 12 hub genes (EPHX3, SPINK7, FCRLA, MASP1, ZNF541, CD5, BEST2, ZAP70) seemed to be cancer-promoting genes as they were downregulated in the high-risk group. Previous studies have suggested that the methylation of

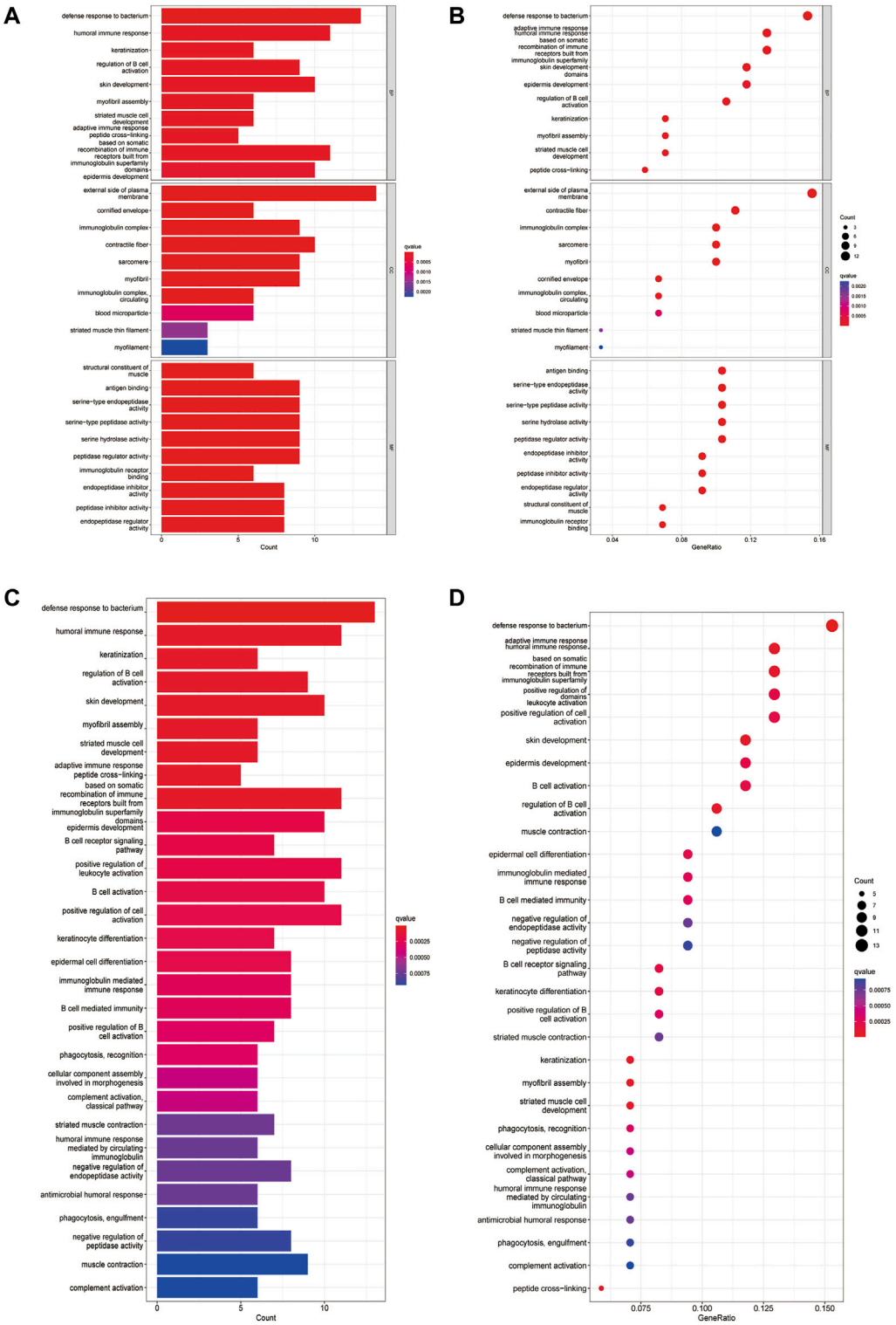


FIGURE 7 | Functional analysis based on the DEGs between the two-risk groups in the TCGA cohort. **(A)**. Barplot graph for GO enrichment (the longer bar means the more genes enriched, and the increasing depth of red means the differences were more obvious). **(B)**. Bubble graph for GO enrichment (the bigger bubble means the more genes enriched, and the increasing depth of red means the differences were more obvious; q-value: the adjusted *p*-value). **(C)**. Barplot graph for KEGG pathways (the longer bar means the more genes enriched, and the increasing depth of red means the differences were more obvious). **(D)**. Bubble graph for KEGG pathways (the bigger bubble means the more genes enriched, and the increasing depth of red means the differences were more obvious; q-value: the adjusted *p*-value).

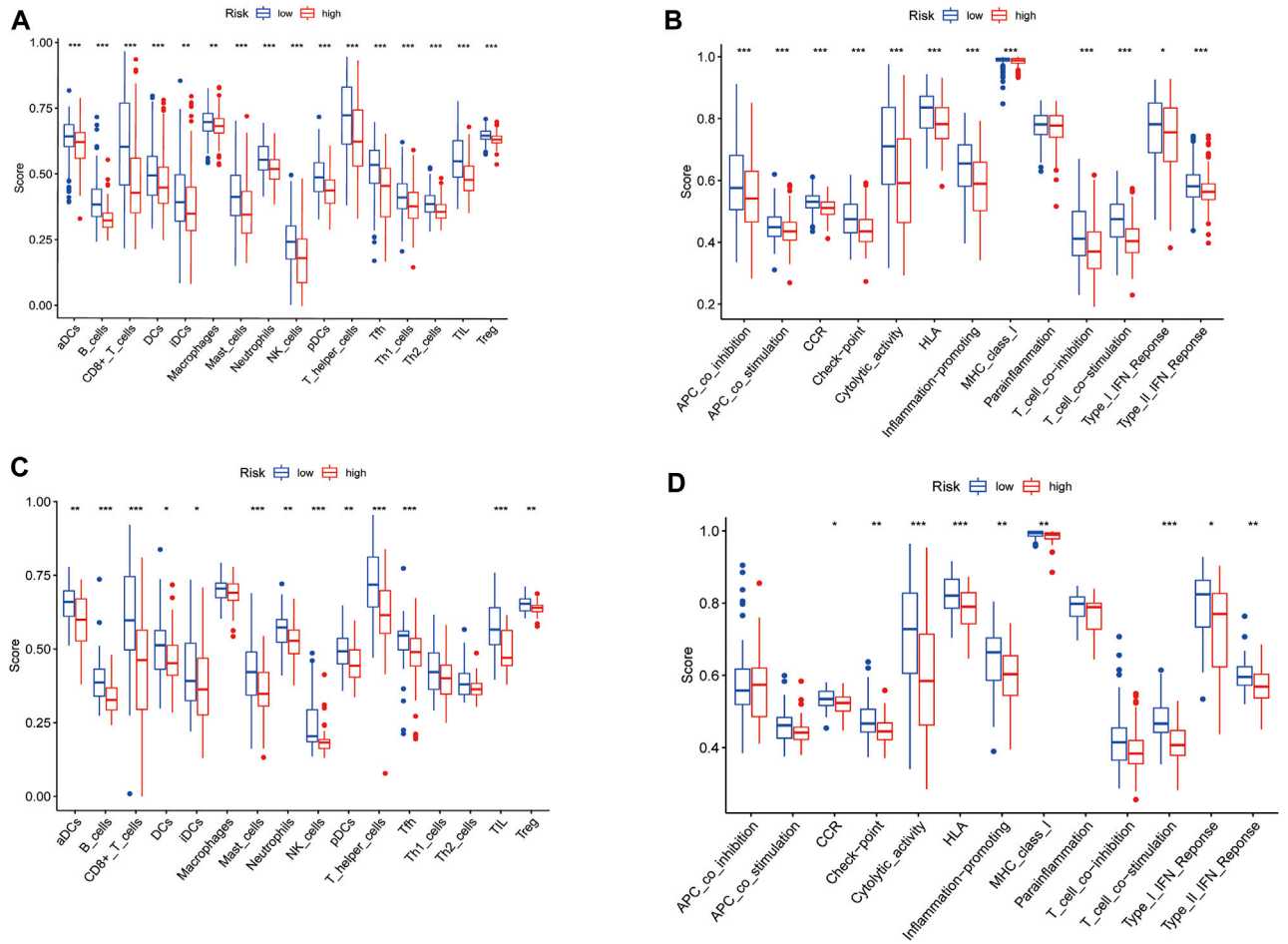


FIGURE 8 | Comparison of the ssGSEA scores for immune cells and immune pathways. **(A,C)** Comparison of the enrichment scores of 16 types of immune cells and 13 immune-related pathways between low- (green box) and high-risk (red box) group in the TCGA cohort. **(B,D)** Comparison of the tumour immunity between low- (blue box) and high-risk (red box) group in the GEO cohort. *p* values were showed as: ns not significant; **p* < 0.05; ***p* < 0.01; ****p* < 0.001.

EPHX3 (epoxide hydrolase 3), a metabolic enzyme that converted mutagenic epoxides into trans-dihydrodiols for detoxication, was considered as a predictive factor for the recurrence of prostate cancer, but as a favorable prognostic factor for HNSCC(Cottrell et al., 2007; Stott-Miller et al., 2014; Bai et al., 2019). The hypermethylation of EPHX3 was associated with salivary gland adenoid cystic carcinoma development and progression (Bell et al., 2011). High expression of EPHX3 increases ceramide linoleate epoxide hydrolysis and functions to control flux through the alternative and crucial route of metabolism by the dehydrogenation pathway of SDR9C7, which implicated a role for EPHX3 in tumor suppression (Edin et al., 2021). **SPINK7**, also named Kazal Type 7 or esophageal cancer-related gene 2 (ECRG2), was a type of protein from the family of serine peptidase inhibitor (Azouz et al., 2018). SPINK7 was identified as a tumor suppressor gene that could inhibit cell growth, suppress cell migration, invasion and metastasis, and promote cell apoptosis by inhibiting the binding of urokinase plasminogen-type activator (uPA) to uPA receptor (uPAR) and cleavage of uPAR(Huang et al., 2007; Cheng et al., 2010). A recent study has demonstrated that the expression of SPINK7 could be utilized to predict the molecular stage of the oral SCC lesions (Pennacchiotti et al., 2021). Another study showed that SPINK7 played an important role in skin homeostasis and inflammatory skin diseases (Chen H. et al., 2021). **FCRLA** (Fc receptor-like A) belongs to a family of Fc receptor like-molecules, and has been identified as a B cell-specific protein and may be involved in the development of lymphomas (Inozume et al., 2007; Fang et al., 2018). FCRLA has been confirmed to be related to the immune status and better prognosis of ovarian cancer (Fan et al., 2021). FCRLA was also found to be associated with the expression of CD19, CD20 and prognosis of laryngeal squamous cell carcinoma (Tagliabue et al., 2020). **MASP-1** (Mannose-binding lectin-associated serine protease 1) is a member of mannose-binding lectin-associated serine protease family and was first believed to up-regulate lectin pathway activation (Schwaner et al., 2017; Debreczeni et al., 2019). The role of MASP-1 in cervical cancer progression has been established (Maestri

et al., 2018). However, MASP-1 was found to be a protective factor for HNSCC in our study. Similarly, another study reported that MASP-1 was found positively correlated with a better prognosis of hepatocellular carcinoma (Xu et al., 2021). **ZNF541** (The zinc finger protein 541) is located on chromosome 19 which possesses the highest gene density of all human chromosomes. Gene mutations on chromosome 19 were usually found to be associated with the occurrence of malignant tumors (Rose, 2018). Low expression levels of ZNF541 were related to the radiosensitivity of breast cancer (Yan et al., 2021). However, ZNF541 expression level was independently associated with a better OS of HPV-positive oropharyngeal cancer, which is similar to our study (Camuzi et al., 2021). **CD5**, a T-Cell surface glycoprotein, was mainly expressed on T cells and a small subset of normal B cells and was considered as an immunoregulatory biomarker in resectable non-small cell lung cancer and other cancer types (Moreno-Manuel et al., 2020). Future studies could investigate the association between CD5 and prognosis in HNSCC. **BEST2** (Bestrophin 2), a part of the bestrophin gene family of anion channels, is expressed predominantly in the retinal pigment epithelium and colon (Yu et al., 2010). It has been reported that high expression of BEST2 gene in HNSCC may lead to a good prognosis (Qin et al., 2020). To the best of our knowledge, the association of BEST2 gene and the prognosis of HNSCC was identified for the first time in our study. **ZAP70** (Zeta chain of T cell receptor associated protein kinase 70) is expressed in a broad range of B cell malignancies (Sadras et al., 2021). Chen et al. demonstrated that ZAP70 could promote cell survival, microenvironment interactions, protein synthesis and further drive disease progression in chronic lymphocytic leukemia cells (Chen et al., 2021). ZAP-70 could also shape the immune microenvironment in B cell malignancies (Chen et al., 2020). Liu et al. proposed that ZAP70-deficiency may improve reverse cholesterol transport and decrease the inflammatory response of T cells (Liu et al., 2019). Gong et al. pointed that an altered methylation pattern of ZAP70 is associated with poor survival of HNSCC (Gong et al., 2020).

The remaining 4 FAM-related genes (LINC00460, ADPRHL1, DYNC1I1, KCNG1) in the prognostic risk model were up-regulated in HNSCC tissues, suggesting that they may play a role in tumorigenesis and development of HNSCC. **LINC00460** (long intergenic non-protein coding RNA 460) is located on chromosome 13q33.2 and transcribed as a 913-nt transcript. Some studies showed that LINC00460 could promote tumor growth and malignant progression and is correlated with survival in multiple tumor types, including nasopharyngeal carcinoma, papillary thyroid carcinoma, esophageal cancer, gastric cancer, lung cancer, breast cancer, colorectal cancer, renal cell carcinoma, pancreatic cancer, bladder and urothelial carcinoma, ovarian cancer, and meningioma (Kong et al., 2018; Liu W. et al., 2018; Liu X. et al., 2018; Wang F. et al., 2018; Xing et al., 2018; Wen et al., 2019; Ye et al., 2019; Cisneros-Villanueva et al., 2021; Hou et al., 2021; Qian et al., 2021; Wu et al., 2021; Yang et al., 2021). HNSCC is certainly no exception, with its identification as a prognostic lncRNA signature using orthogonal partial least squares discriminant analysis (OPLS-DA) which integrates RNA-Seq data from TCGA database and matching clinical information from a large cohort of patients with HNSCC(Cao et al., 2017). LINC00460 is a promising candidate potential target for cancer therapy for HNSCC(Jiang et al., 2019). **ADPRHL1** (ADP-ribosylhydrolase like 1) is a member of the ADP-ribosylhydrolase protein family and a reversible posttranslational modification which could regulate protein function (Wan et al., 2018). Pilot studies have noted **DYNC1I1** (dynein cytoplasmic 1 intermediate chain 1) was an adverse factor for the prognosis of patients with liver hepatocellular carcinoma, gastric cancer, colon cancer and glioblastoma (Gong L.-B. et al., 2019; Gong L. et al., 2019; Sakthikumar et al., 2020; Yin et al., 2020; Zhou J. et al., 2021). There were no published researches on the ADPRHL1, DYNC1I1, KCNG1 in HNSCC before to our knowledge.

A previous study found that high abundances of CD3 or CD8 TILs were independently associated with prolonged survival outcomes among HNSCC patients (Kim et al., 2016). Multiple pieces of evidence showed a general increase of both circulating and infiltrating Treg during HNSCC development. Furthermore, Treg levels increased accordingly with tumor staging and they were particularly elevated in patients with active disease (Maggioni et al., 2017). Another retrospective study showed that increased intratumoral $CD8^+$ T cell infiltration and an increased $CD8^+$ T cell/Treg ratio were linked to a better treatment response of recurrent/metastatic HNSCC patients who were treated with immunotherapy (Hanna et al., 2018). The high infiltration of pDCs in tumors was also verified to promote the progression of HNSCC (Zhou B. et al., 2021). In our study, although $CD8^+$ T cells and Treg cells were significantly higher in the low-risk group than those in the high-risk group, the number of Treg cells was lower than that in $CD8^+$ T cells in either training set or test set. Therefore, $CD8^+$ T cell/Treg ratio was better suited to predict the prognosis of patients with HNSCC. In addition, immune cell infiltration was significantly higher in the low-risk group than in the high-risk group in this study, which showed that the antigen presentation process, cellular immunity, and humoral immunity may be stronger in the low-risk group than in the high-risk group. It revealed that the risk prognostic model can partially reflect immune status in tumor microenvironment and has the potential to predict the efficacy of immunotherapy in patients with HNSCC.

There were still limitations in our study. Firstly, the prediction model was constructed based on the TCGA database and validated in the GEO database. But it has not been confirmed *in vivo* and *in vitro*. Secondly, only FAM-related genes were included in the prognostic model, thereby other important prognostic factors for HNSCC may have been excluded. This prognostic model deserves further in-depth studies.

CONCLUSION

In summary, our study constructed a prognostic model of HNSCC based on 12 FAM-related genes for the first time. The model was proved to be an independent prognostic factor for survival in patients with HNSCC in both training set and test set, providing a new potential biomarker for prognosis prediction for patients with HNSCC. In addition, this prognostic model may be useful for evaluating immune status in tumor microenvironment of HNSCC, further screening the dominant patients with HNSCC of 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.

AUTHOR CONTRIBUTIONS

PD and YC made contributions to the conception of this study. PD and YC analyzed the data and wrote the manuscript. HX, SZ, and JY helped with the analysis of the data and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant numbers 81771002 and 82071057).

ACKNOWLEDGMENTS

We would like to acknowledge the TCGA, HPA and the GEO (GSE41613) network for providing data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.888764/full#supplementary-material

REFERENCES

Agostini, M., Almeida, L. Y., Bastos, D. C., Ortega, R. M., Moreira, F. S., Seguin, F., et al. (2014). The Fatty Acid Synthase Inhibitor Orlistat Reduces the Growth and Metastasis of Orthotopic Tongue Oral Squamous Cell Carcinomas. *Mol. Cancer Ther.* 13, 585–595. doi:10.1158/1535-7163.mct-12-1136

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, 25–29. doi:10.1038/75556

Azouz, N. P., Ynga-Durand, M. A., Caldwell, J. M., Jain, A., Rochman, M., Fischesser, D. M., et al. (2018). The Antiprotease SPINK7 Serves as an Inhibitory Checkpoint for Esophageal Epithelial Inflammatory Responses. *Sci. Transl. Med.* 10, eaap9736. doi:10.1126/scitranslmed.aap9736

Bai, G., Song, J., Yuan, Y., Chen, Z., Tian, Y., Yin, X., et al. (2019). Systematic Analysis of Differentially Methylated Expressed Genes and Site-specific Methylation as Potential Prognostic Markers in Head and Neck Cancer. *J. Cell. Physiology* 234, 22687–22702. doi:10.1002/jcp.28835

Bell, A., Bell, D., Weber, R. S., and El-Naggar, A. K. (2011). CpG Island Methylation Profiling in Human Salivary Gland Adenoid Cystic Carcinoma. *Cancer* 117, 2898–2909. doi:10.1002/cncr.25818

Camuzi, D., Buexm, L. A., Lourenço, S., Esposti, D. D., Cuenin, C., Lopes, M., et al. (2021). HPV Infection Leaves a DNA Methylation Signature in Oropharyngeal Cancer Affecting Both Coding Genes and Transposable Elements. *Cancers (Basel)* 13, 3621. doi:10.3390/cancers13143621

Cao, W., Liu, J.-n., Liu, Z., Wang, X., Han, Z.-G., Ji, T., et al. (2017). A Three-lncRNA Signature Derived from the Atlas of ncRNA in Cancer (TANRIC) Database Predicts the Survival of Patients with Head and Neck Squamous Cell Carcinoma. *Oral Oncol.* 65, 94–101. doi:10.1016/j.oraloncology.2016.12.017

Chen, H., Zhou, Y., Han, X., and Shi, Y. (2021). The Changing Landscape of Anti-lymphoma Drug Clinical Trials in Mainland China in the Past 15 Years (2005-2020): A Systematic Review. *Lancet Regional Health - West. Pac.* 8, 100097. doi:10.1016/j.lanwpc.2021.100097

Chen, J., Moore, A., and Ringshausen, I. (2020). ZAP-70 Shapes the Immune Microenvironment in B Cell Malignancies. *Front. Oncol.* 10, 595832. doi:10.3389/fonc.2020.595832

Chen, J., Sathiaseelan, V., Moore, A., Tan, S., Chilamakuri, C. S. R., Roamio Franklin, V. N., et al. (2021). ZAP-70 Constitutively Regulates Gene Expression and Protein Synthesis in Chronic Lymphocytic Leukemia. *Blood* 137, 3629–3640. doi:10.1182/blood.2020009960

Cheng, X., Lu, S.-H., and Cui, Y. (2010). ECRG2 Regulates ECM Degradation and uPAR/FPRL1 Pathway Contributing Cell Invasion/migration. *Cancer Lett.* 290, 87–95. doi:10.1016/j.canlet.2009.09.001

Cisneros-Villanueva, M., Hidalgo-Pérez, L., Cedro-Tanda, A., Peña-Luna, M., Mancera-Rodríguez, M. A., Hurtado-Cordova, E., et al. (2021). LINC00460 Is a Dual Biomarker that Acts as a Predictor for Increased Prognosis in Basal-like Breast Cancer and Potentially Regulates Immunogenic and Differentiation-Related Genes. *Front. Oncol.* 11, 628027. doi:10.3389/fonc.2021.628027

Corn, K. C., Windham, M. A., and Rafat, M. (2020). Lipids in the Tumor Microenvironment: From Cancer Progression to Treatment. *Prog. Lipid Res.* 80, 101055. doi:10.1016/j.plipres.2020.101055

Cottrell, S., Jung, K., Kristiansen, G., Eltze, E., Semjonow, A., Ittmann, M., et al. (2007). Discovery and Validation of 3 Novel DNA Methylation Markers of Prostate Cancer Prognosis. *J. Urology* 177, 1753–1758. doi:10.1016/j.juro.2007.01.010

Debreczeni, M. L., Németh, Z., Kajdácsi, E., Schwaner, E., Makó, V., Masszi, A., et al. (2019). MASP-1 Increases Endothelial Permeability. *Front. Immunol.* 10, 991. doi:10.3389/fimmu.2019.00991

Edin, M. L., Yamanashi, H., Boeglin, W. E., Graves, J. P., DeGraff, L. M., Lih, F. B., et al. (2021). Epoxide Hydrolase 3 (Ephx3) Gene Disruption Reduces Ceramide Linoleate Epoxide Hydrolysis and Impairs Skin Barrier Function. *J. Biol. Chem.* 296, 100198. doi:10.1074/jbc.ra120.016570

Fan, L., Lei, H., Lin, Y., Zhou, Z., Shu, G., Yan, Z., et al. (2021). Identification of a Gene Set Correlated with Immune Status in Ovarian Cancer by Transcriptome-wide Data Mining. *Front. Mol. Biosci.* 8, 670666. doi:10.3389/fmolb.2021.670666

Fang, G., Zhang, Q. H., Tang, Q., Jiang, Z., Xing, S., Li, J., et al. (2018). Comprehensive Analysis of Gene Expression and DNA Methylation Datasets Identify Valuable Biomarkers for Rheumatoid Arthritis Progression. *Oncotarget* 9, 2977–2983. doi:10.18632/oncotarget.22918

Fernández, L. P., Gómez de Cedrón, M., and Ramírez de Molina, A. (2020). Alterations of Lipid Metabolism in Cancer: Implications in Prognosis and Treatment. *Front. Oncol.* 10, 577420. doi:10.3389/fonc.2020.577420

Fhu, C. W., and Ali, A. (2020). Fatty Acid Synthase: An Emerging Target in Cancer. *Molecules* 25, 3935. doi:10.3390/molecules25173935

Gong, F. C., Ji, R., Wang, Y. M., Yang, Z. T., Chen, Y., Mao, E. Q., et al. (2020). Identification of Potential Biomarkers and Immune Features of Sepsis Using Bioinformatics Analysis. *Mediat. Inflamm.* 2020, 3432587. doi:10.1155/2020/3432587

Gong, L.-B., Wen, T., Li, Z., Xin, X., Che, X.-F., Wang, J., et al. (2019). DYNC1I1 Promotes the Proliferation and Migration of Gastric Cancer by Up-Regulating IL-6 Expression. *Front. Oncol.* 9, 491. doi:10.3389/fonc.2019.00491

Gong, L., Wen, T., Li, Z., Wang, Y., Wang, J., Che, X., et al. (2019). TNPO2 Operates Downstream of DYNC1I1 and Promotes Gastric Cancer Cell Proliferation and Inhibits Apoptosis. *Cancer Med.* 8, 7299–7312. doi:10.1002/cam4.2582

Hanna, G. J., Lizotte, P., Cavanaugh, M., Kuo, F. C., Shivdasani, P., Frieden, A., et al. (2018). Frameshift Events Predict Anti-PD-1/l1 Response in Head and Neck Cancer. *JCI Insight* 3, e98811. doi:10.1172/jci.insight.98811

He, D., Cai, L., Huang, W., Weng, Q., Lin, X., You, M., et al. (2021). Prognostic Value of Fatty Acid Metabolism-Related Genes in Patients with Hepatocellular Carcinoma. *Aging* 13, 17847–17863. doi:10.18632/aging.203288

Hofmanová, J., Slavík, J., Ciganek, M., Ovesná, P., Tylichová, Z., Karasová, M., et al. (2021). Complex Alterations of Fatty Acid Metabolism and Phospholipidome Uncovered in Isolated Colon Cancer Epithelial Cells. *Int. J. Mol. Sci.* 22, 6650. doi:10.3390/ijms22136650

Hou, P., Meng, S., Li, M., Lin, T., Chu, S., Li, Z., et al. (2021). LINC00460/DHX9/IGF2BP2 Complex Promotes Colorectal Cancer Proliferation and Metastasis by Mediating HMGA1 mRNA Stability Depending on m6A Modification. *J. Exp. Clin. Cancer Res.* 40, 52. doi:10.1186/s13046-021-01857-2

Hoy, A. J., Nagarajan, S. R., and Butler, L. M. (2021). Tumour Fatty Acid Metabolism in the Context of Therapy Resistance and Obesity. *Nat. Rev. Cancer* 21, 753–766. doi:10.1038/s41568-021-00388-4

Huang, G., Hu, Z., Li, M., Cui, Y., Li, Y., Guo, L., et al. (2007). ECRG2 Inhibits Cancer Cell Migration, Invasion and Metastasis through the Down-Regulation of uPA/plasmin Activity. *Carcinogenesis* 28, 2274–2281. doi:10.1093/carcin/bgm140

Inozume, T., Mitsui, H., Okamoto, T., Matsuzaki, Y., Kawakami, Y., Shibagaki, N., et al. (2007). Dendritic Cells Transduced with Autoantigen FCRLA Induce Cytotoxic Lymphocytes and Vaccinate against Murine B-Cell Lymphoma. *J. Investigative Dermatology* 127, 2818–2822. doi:10.1038/sj.jid.5700909

Jiang, Y., Cao, W., Wu, K., Qin, X., Wang, X., Li, Y., et al. (2019). LncRNA LINC00460 Promotes EMT in Head and Neck Squamous Cell Carcinoma by Facilitating Peroxiredoxin-1 into the Nucleus. *J. Exp. Clin. Cancer Res.* 38, 365. doi:10.1186/s13046-019-1364-z

Johnson, D. E., Burtness, B., Leemans, C. R., Lui, V. W. Y., Bauman, J. E., and Grandis, J. R. (2020). Head and Neck Squamous Cell Carcinoma. *Nat. Rev. Dis. Prim.* 6, 92. doi:10.1038/s41572-020-00224-3

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* 45, D353–D361. doi:10.1093/nar/gkw1092

Kim, H. R., Ha, S.-J., Hong, M. H., Heo, S. J., Koh, Y. W., Choi, E. C., et al. (2016). PD-L1 Expression on Immune Cells, but Not on Tumor Cells, Is a Favorable Prognostic Factor for Head and Neck Cancer Patients. *Sci. Rep.* 6, 36956. doi:10.1038/srep36956

Kong, Y.-G., Cui, M., Chen, S.-M., Xu, Y., Xu, Y., and Tao, Z.-Z. (2018). LncRNA-LINC00460 Facilitates Nasopharyngeal Carcinoma Tumorigenesis through Sponging miR-149-5p to Up-Regulate IL6. *Gene* 639, 77–84. doi:10.1016/j.gene.2017.10.006

Li, Z., and Zhang, H. (2016). Reprogramming of Glucose, Fatty Acid and Amino Acid Metabolism for Cancer Progression. *Cell. Mol. Life Sci.* 73, 377–392. doi:10.1007/s00018-015-2070-4

Liu, J., Guo, K., Hu, L., Luo, T., Ma, Y., Zhang, Y., et al. (2019). ZAP70 Deficiency Promotes Reverse Cholesterol Transport through MAPK/ERK Pathway in Jurkat Cell. *Mol. Immunol.* 107, 21–28. doi:10.1016/j.molimm.2019.01.001

Liu, W., Zhang, Y., Chen, M., Shi, L., Xu, L., and Zou, X. (2018). A Genome-wide Analysis of Long Noncoding RNA Profile Identifies Differentially Expressed lncRNAs Associated with Esophageal Cancer. *Cancer Med.* 7, 4181–4189. doi:10.1002/cam4.1536

Liu, X., Wen, J., Wang, H., and Wang, Y. (2018). Long Non-coding RNA LINC00460 Promotes Epithelial Ovarian Cancer Progression by Regulating microRNA-338-3p. *Biomed. Pharmacother.* 108, 1022–1028. doi:10.1016/j.biopha.2018.09.103

Madak-Erdogan, Z., Band, S., Zhao, Y. C., Smith, B. P., Kulkoyluoglu-Cotul, E., Zuo, Q., et al. (2019). Free Fatty Acids Rewire Cancer Metabolism in Obesity-Associated Breast Cancer via Estrogen Receptor and mTOR Signaling. *Cancer Res.* 79, 2494–2510. doi:10.1158/0008-5472.CAN-18-2849

Maestri, C. A., Nisihara, R., Mendes, H. W., Jensenius, J., Thiel, S., Messias-Reason, I., et al. (2018). MASP-1 and MASP-2 Serum Levels Are Associated with Worse Prognostic in Cervical Cancer Progression. *Front. Immunol.* 9, 2742. doi:10.3389/fimmu.2018.02742

Maggioni, D., Pignataro, L., and Garavello, W. (2017). T-helper and T-Regulatory Cells Modulation in Head and Neck Squamous Cell Carcinoma. *Oncoimmunology* 6, e1325066. doi:10.1080/2162402x.2017.1325066

Moreno-Manuel, A., Jantus-Lewintre, E., Simões, I., Aranda, F., Calabuig-Fariñas, S., Carreras, E., et al. (2020). CD5 and CD6 as Immunoregulatory Biomarkers in Non-small Cell Lung Cancer. *Transl. Lung Cancer Res.* 9, 1074–1083. doi:10.21037/tlcr-19-445

Pennacchiotti, G., Valdés-Gutiérrez, F., González-Arriagada, W. A., Montes, H. F., Parra, J. M. R., Guida, V. A., et al. (2021). SPINK7 Expression Changes Accompanied by HER2, P53 and RB1 Can Be Relevant in Predicting Oral Squamous Cell Carcinoma at a Molecular Level. *Sci. Rep.* 11, 6939. doi:10.1038/s41598-021-86208-z

Qian, X., Tang, J., Li, L., Chen, Z., Chen, L., and Chu, Y. (2021). A New Ferroptosis-Related Gene Model for Prognostic Prediction of Papillary Thyroid Carcinoma. *Bioengineered* 12, 2341–2351. doi:10.1080/21655979.2021.1935400

Qin, L., Zeng, J., Shi, N., Chen, L., and Wang, L. (2020). Application of Weighted Gene Co-expression Network Analysis to Explore the Potential Diagnostic Biomarkers for Colorectal Cancer. *Mol. Med. Rep.* 21, 2533–2543. doi:10.3892/mmr.2020.11047

Rose, A. M. (2018). Cancer and the Junkyard Chromosome: How Repeat DNA Sequence on Chromosome 19 Influences Risk of Malignant Disease. *Oncotarget* 9, 31942–31944. doi:10.18632/oncotarget.25873

Sadras, T., Martin, M., Kume, K., Robinson, M. E., Saravanakumar, S., Lenz, G., et al. (2021). Developmental Partitioning of SYK and ZAP70 Prevents Autoimmunity and Cancer. *Mol. Cell* 81, 2094–2111. doi:10.1016/j.molcel.2021.03.043

Sakthikumar, S., Roy, A., Haseeb, L., Pettersson, M. E., Sundström, E., Marinescu, V. D., et al. (2020). Whole-genome Sequencing of Glioblastoma Reveals Enrichment of Non-coding Constraint Mutations in Known and Novel Genes. *Genome Biol.* 21, 127. doi:10.1186/s13059-020-02035-x

Schwaner, E., Németh, Z., Jani, P. K., Kajdácsi, E., Debreczeni, M. L., Doleschall, Z., et al. (2017). Transcriptome Analysis of Inflammation-Related Gene Expression in Endothelial Cells Activated by Complement MASP-1. *Sci. Rep.* 7, 10462. doi:10.1038/s41598-017-09058-8

Solomon, B., Young, R. J., and Rischin, D. (2018). Head and Neck Squamous Cell Carcinoma: Genomics and Emerging Biomarkers for Immunomodulatory Cancer Treatments. *Seminars Cancer Biol.* 52, 228–240. doi:10.1016/j.semcancer.2018.01.008

Stott-Miller, M., Zhao, S., Wright, J. L., Kolb, S., Bibikova, M., Klotzle, B., et al. (2014). Validation Study of Genes with Hypermethylated Promoter Regions Associated with Prostate Cancer Recurrence. *Cancer Epidemiol. Biomarkers Prev.* 23, 1331–1339. doi:10.1158/1055-9965.epi-13-1000

Su, Y. W., Wu, P. S., Lin, S. H., Huang, W. Y., Kuo, Y. S., and Lin, H. P. (2020). Prognostic Value of the Overexpression of Fatty Acid Metabolism-Related Enzymes in Squamous Cell Carcinoma of the Head and Neck. *Int. J. Mol. Sci.* 21, 6851. doi:10.3390/ijms21186851

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

Tagliabue, M., Maffini, F., Fumagalli, C., Gandini, S., Lepanto, D., Corso, F., et al. (2020). A Role for the Immune System in Advanced Laryngeal Cancer. *Sci. Rep.* 10, 18327. doi:10.1038/s41598-020-73747-0

Wan, Q., Tang, J., Han, Y., and Wang, D. (2018). Co-expression Modules Construction by WGCNA and Identify Potential Prognostic Markers of Uveal Melanoma. *Exp. Eye Res.* 166, 13–20. doi:10.1016/j.exer.2017.10.007

Wang, F., Liang, S., Liu, X., Han, L., Wang, J., and Du, Q. (2018). LINC00460 Modulates KDM2A to Promote Cell Proliferation and Migration by Targeting

miR-342-3p in Gastric Cancer. *Onco. Targets Ther.* 11, 6383–6394. doi:10.2147/ott.s169307

Wang, Y.-n., Zeng, Z.-l., Lu, J., Wang, Y., Liu, Z.-x., He, M.-m., et al. (2018). CPT1A-mediated Fatty Acid Oxidation Promotes Colorectal Cancer Cell Metastasis by Inhibiting Anoikis. *Oncogene* 37, 6025–6040. doi:10.1038/s41388-018-0384-z

Wen, L., Zhang, X., Bian, J., Han, L., Huang, H., He, M., et al. (2019). The Long Non-coding RNA LINC00460 Predicts the Prognosis and Promotes the Proliferation and Migration of Cells in Bladder Urothelial Carcinoma. *Oncol. Lett.* 17, 3874–3880. doi:10.3892/ol.2019.10023

Wu, J., Sun, S., Liao, W., Chen, E., Wang, X., Song, Y., et al. (2021). LINC00460 Promotes Pancreatic Cancer Progression by Sponging miR-491-5p. *J. Gene Med.* 23, e3333. doi:10.1002/jgm.3333

Xing, H., Wang, S., Li, Q., Ma, Y., and Sun, P. (2018). Long Noncoding RNA LINC00460 Targets miR-539/MMP-9 to Promote Meningioma Progression and Metastasis. *Biomed. Pharmacother.* 105, 677–682. doi:10.1016/j.biopha.2018.06.005

Xu, Y., Wang, Z., and Li, F. (2021). Survival Prediction and Response to Immune Checkpoint Inhibitors: A Prognostic Immune Signature for Hepatocellular Carcinoma. *Transl. Oncol.* 14, 100957. doi:10.1016/j.tranon.2020.100957

Yan, D., Shen, M., Du, Z., Cao, J., Tian, Y., Zeng, P., et al. (2021). Developing ZNF Gene Signatures Predicting Radiosensitivity of Patients with Breast Cancer. *J. Oncol.* 2021, 9255494. doi:10.1155/2021/9255494

Yang, H., Xiong, X., and Li, H. (2021). Development and Interpretation of a Genomic Instability Derived lncRNAs Based Risk Signature as a Predictor of Prognosis for Clear Cell Renal Cell Carcinoma Patients. *Front. Oncol.* 11, 678253. doi:10.3389/fonc.2021.678253

Ye, J.-J., Cheng, Y.-L., Deng, J.-J., Tao, W.-P., and Wu, L. (2019). LncRNA LINC00460 Promotes Tumor Growth of Human Lung Adenocarcinoma by Targeting miR-302c-5p/FOXA1 axis. *Gene* 685, 76–84. doi:10.1016/j.gene.2018.10.058

Yin, Z., Yan, X., Wang, Q., Deng, Z., Tang, K., Cao, Z., et al. (2020). Detecting Prognosis Risk Biomarkers for Colon Cancer through Multi-Omics-Based Prognostic Analysis and Target Regulation Simulation Modeling. *Front. Genet.* 11, 524. doi:10.3389/fgene.2020.00524

Yu, K., Lujan, R., Marmorstein, A., Gabriel, S., and Hartzell, H. C. (2010). Bestrophin-2 Mediates Bicarbonate Transport by Goblet Cells in Mouse Colon. *J. Clin. Invest.* 120, 1722–1735. doi:10.1172/jci41129

Zhang, Z., Wang, Y., Wang, Y., Wang, C., Shuai, Y., Luo, J., et al. (2021). BCAR3 Promotes Head and Neck Cancer Growth and Is Associated with Poor Prognosis. *Cell Death Discov.* 7, 316. doi:10.1038/s41420-021-00714-7

Zhou, B., Lawrence, T., and Liang, Y. (2021). The Role of Plasmacytoid Dendritic Cells in Cancers. *Front. Immunol.* 12, 749190. doi:10.3389/fimmu.2021.749190

Zhou, J., Zhu, Y., Ma, S., Li, Y., Liu, K., Xu, S., et al. (2021). Bioinformatics Analysis Identifies DYNC1I1 as Prognosis Marker in Male Patients with Liver Hepatocellular Carcinoma. *PLoS One* 16, e258797. doi:10.1371/journal.pone.0258797

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 Du, Chai, Zong, Yue and Xiao. 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.2022.874896

CCR5 as a Prognostic Factor in Lower-Grade Glioma is Involved in the Remodeling of the Tumor Microenvironment

Fang Wang[1,2†], Zhennan Tao[3†], Zhen Tian[4†], Jiaqi Jin[1,2], Jiawei Dong[1,2], Yuxiang Dai[3], Wanli Yu[5], Bin Tang[5*] and Shaoshan Hu[1,2*]*

[1]Department of Neurosurgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, China, [2]Department of Neurosurgery, Emergency Medicine Center, Zhejiang Provincial People's Hospital, Hangzhou Medical College, Hangzhou, China, [3]Department of Neurosurgery, the Affiliated Drum Tower Hospital, School of Medicine, Nanjing University, Nanjing, China, [4]Department of Minimally Invasive Interventional Oncology, Hubei Cancer Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [5]Department of Neurosurgery, The First Affiliated Hospital of Nanchang University, Nanchang, China

OPEN ACCESS

Edited by:
Lan Zhao,
Stanford University, United States

Reviewed by:
Ziheng Wang,
Affiliated Hospital of Nantong University, China
Dongdong Lin,
Zhejiang Hospital, China

***Correspondence:**
Wanli Yu
ywl_doctor@163.com
Bin Tang
tonytang19850815@sina.com
Shaoshan Hu
shaoshanhu421@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: 13 February 2022
Accepted: 25 May 2022
Published: 05 July 2022

Citation:
Wang F, Tao Z, Tian Z, Jin J, Dong J, Dai Y, Yu W, Tang B and Hu S (2022) CCR5 as a Prognostic Factor in Lower-Grade Glioma is Involved in the Remodeling of the Tumor Microenvironment. Front. Genet. 13:874896. doi: 10.3389/fgene.2022.874896

Background: Lower-grade gliomas (LGGs) carry a high risk of malignant transformation, leading to severe neurologic deterioration and ultimately, death. The tumor microenvironment (TME) plays an essential role in tumor maintenance, progression, and immunotherapy resistance. Therefore, the LGG TME deserves comprehensive exploration for a novel therapeutic target.

Methods: The ESTIMATE algorithm was used to estimate infiltrating stromal and immune cells of LGG patients obtained from the Cancer Genome Atlas (TCGA) database. Kaplan–Meier analysis was performed to classify survival differences. TME-related differentially expressed genes were identified between the low- and high-immune/stromal groups. Hub genes were screened by constructing protein–protein interaction networks and performing the Cox regression analysis. Differential analysis, survival analysis, gene set enrichment analysis, and clinical relevance analysis specific to hub genes were evaluated by using the TCGA and the Chinese Glioma Genome Atlas datasets, and the results were validated by qRT-PCR, Western blotting, and immunohistochemistry in tissues from LGG patients.

Results: The immune and stromal components in TME were negatively related to patient prognosis. Differentially expressed genes sharing immune score and stromal score were mainly involved in the immune response. C-C chemokine receptor type 5 (CCR5), as only a hub gene, was significantly higher in LGG patients than normal patients and negatively correlated with the prognosis of patients. High-expression CCR5 was positively related to

Abbreviations: LGG, lower-grade glioma; GBM, glioblastomas; TME, the tumor microenvironment; CCR5, C-C motif chemokine receptor type 5; TCGA, the Cancer Genome Atlas; CGGA, Chinese Glioma Genome Atlas; GTEx, the genotype-tissue expression; KM, Kaplan–Meier; DEGs, differentially expressed genes; FC, fold change; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein–protein interaction; OS, overall survival; GSEA, gene set enrichment analysis; TICs, tumor-infiltrating immune cells; PBT, peritumoral brain tissues; IHC, immunohistochemistry; IDH, isocitrate dehydrogenase; Tregs, T regulatory cells.

immune-related and tumor progression pathways. CCR5 protein expression was higher in LGG with isocitrate dehydrogenase wildtype. Validated results showed that CCR5 was upregulated in LGG tissues at mRNA and protein levels and could affect immune cell infiltration. These results suggested that CCR5 was a potential indicator for the status of TME.

Conclusion: Glioma cells remodel the immune microenvironment through the high expression of CCR5 and lead to a poor prognosis in patients with LGG. The inhibition of CCR5 may contribute to the efficacy of LGG immunotherapy.

Keywords: tumor microenvironment, LGG, glioma, CCR5, immunological biomarker

INTRODUCTION

Lower-grade glioma (LGG) is a subset of primary brain malignancy that causes severe neurological dysfunction and may progress toward glioblastoma (GBM) (Cuddapah et al., 2014; Lombardi et al., 2020). LGG has a better prognosis than GBM, with a median patient survival of 5.6–13.3 years (Louis et al., 2016). Modern comprehensive treatments are improving; however, LGG deterioration or recurrence still leads to an unsatisfactory outcome (Claus et al., 2015). Glioma recurrence relies on invasiveness and treatment resistance, which may partially ascribe to the peculiar tumor microenvironment (TME) (Hambardzumyan et al., 2016; Quail and Joyce 2017). Therefore, exploring the underlying mechanism of the interaction between glioma cells and the TME is helpful for developing effective therapeutic strategies.

The chemokine system has been widely demonstrated to be involved in the development of various tumors by promoting proliferation, metastasis, and angiogenesis, which can also modify the TME of tumor by recruiting leukocytes and enhancing angiogenesis (Balkwill 2012; Bachelerie et al., 2014; Korbecki et al., 2020). C-C motif chemokine receptor type 5 (CCR5), as a chemokine receptor, belongs to 7 transmembrane G-protein–coupled receptors and combines with different ligands to exert its effects. CCR5 is expressed on a broad array of tumors, and there is an increasing interest in developing CCR5 inhibitors as novel anticancer agents (Tan et al., 2009; Velasco-Velazquez et al., 2014; Jiao et al., 2019). A small number of studies have also reported the role of CCR5 in glioma. Some scholars found that CCR5 promoted invasion and metastasis of GBM through cell and animal experiments (Zhao et al., 2015; Novak et al., 2020). Another study revealed that pericytes augment glioblastoma cell resistance to temozolomide through CCL5–CCR5 paracrine signaling (Zhang et al., 2021). Interest is increasing on the role of CCR5 in glioma. However, since the current study focused solely on GBM, the roles of CCR5 in LGG still remain unclear. Not only that, there are no studies which used bioinformatics approaches to explore CCR5 in glioma.

In this study, with the bioinformatics and experimental analyses, we identified that CCR5 as a prognostic factor in lower-grade glioma is involved in the remodeling of the TME. The correlation between the expression of CCR5 with immune cell infiltration, immune-related signaling pathways, and biological processes was identified. This research provides new insights into the LGG-immune microenvironment, and our results offer further evidence for future clinical application of CCR5 inhibitors in glioma.

MATERIALS AND METHODS

Data Collection and Processing

mRNA-seq datasets and clinical information of LGG cases were obtained from the Cancer Genome Atlas (TCGA) database (https://cancergenome.nih.gov/) and the Chinese Glioma Genome Atlas (CGGA) database (https://www.cgga.org.cn). mRNA-seq data of healthy samples were downloaded from the Genotype-Tissue Expression (GTEx) database (https://commonfund.nih.gov/GTEx/). The illuminaHumanv4. db R package was used to convert the probe IDs to gene symbols.

Generation of Immune Score, Stromal Score and ESTIMATE Score

The R package "ESTIMATE" was used to assess the stromal score, immune score, and ESTIMATE score of tumor samples, which were correlated with the ratio of stromal and immune infiltration in the TME.

Analysis of Clinical Characteristics

The Kaplan–Meier (K-M) method was performed to plot the survival curve, the log-rank test was used to compare the subgroups, and the Kruskal–Wallis rank-sum test was applied to analyze the association between the scores and multiple subgroups of clinical variables.

Identification of Tumor Microenvironment–Related Differentially Expressed Genes and Enrichment Analysis

The R package "limma" was used to perform to screen Differentially Expressed Genes (DEGs) between high- and low-score groups regarding immune score and stromal score. The cutoff criteria were set as $|\log2 \text{ fold change (FC)}| > 1$ and $p < 0.05$. Then, Gene ontology (GO) functional and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway

enrichment analysis of DEGs were performed with the R packages "clusterProfiler."

Protein–Protein Interaction Network and Cox Regression Analysis

The STRING database (https://string-db.org) was used to construct the Protein-Protein Interaction (PPI) network encoded by DEGs, followed by reconstruction with Cytoscape software (version 3.8.2). Nodes with confidence score > 0.9 were used for building network. Univariate Cox regression was performed by using the R package "survival."

C-C Chemokine Receptor Type 5 Differential Expression and Survival Analyses

The Wilcoxon rank-sum test was applied to compare the difference in CCR5 expression between normal and tumor tissues. K-M analysis was performed to classify the association between the CCR5 gene and overall survival (OS).

Gene Set Enrichment Analysis and Analysis of Gene With Clinical Characteristics

To identify biological pathways between high and low CCR5 expression groups, Gene Set Enrichment Analysis (GSEA) was performed by using GSEA v4.1.0. Then, we applied "limma" package to assess the associations between CCR5 expression and clinical characteristics.

Analysis of Immune Cell Based on C-C Chemokine Receptor Type 5

The CIBERSORT algorithm was used to estimate the fraction of 22 tumor-infiltrating immune cells (TICs). Then, the correlation and differential analysis between CCR5 and immune cells were performed.

Human Tumor Samples

Eight tumor tissues and corresponding peritumoral brain tissues (PBT) from LGG patients were collected in the Second Affiliated Hospital of Harbin Medical University. This research was approved by all the patients and the Ethics Committee of the hospital. Clinical characteristics of patient cohort are displayed in **Supplementary Table S1**.

qRT-PCR

Total RNA was extracted from human tissues and cells by using TRIzol reagent (Invitrogen). According to manufacturer's instructions of the Nanodrop ND-2000 spectrophotometer (Thermo Scientific™, United States), 2 μg of the total RNA was transcribed into cDNA. The SYBR Green PCR kit (Takara, Japan) was used for qRT-PCR. The relative mRNA expression was standardized by GAPDH. The primers (Tsingke Biotechnology Co., Ltd, Beijing, China) used are displayed in **Supplementary Table S2**.

Western Blot

Tissues were lysed in the RIPA buffer, and protein concentration was detected with the BCA protein assay kit (Beyotime, China). The 20-μg proteins were fractionated by SDS-PAGE on 10% SDS-acrylamide gel. The membrane was blocked with 5% skim milk for 1 h. After culturing with primary (CCR5, Abcam and β-actin, Abcam) and secondary antibodies, the electro chemiluminescence kit (ECL; Solarbio, China) was used to detect the protein blots. The antibody information is listed in **Supplementary Table S3**.

Immunohistochemistry

Human tissues were processed into formalin-fixed and paraffin-embedded specimens. Immunohistochemistry (IHC) was performed by means of streptavidin coupled with peroxidase. The samples were incubated overnight with the primary antibodies (CCR5, Abcam; CD4, Abcam; and CD206, Proteintech), followed by incubation with the biotinylated secondary antibody. Finally, the signals were detected using an Olympus BX41 microscope. The antibody information is listed in **Supplementary Table S3**.

Statistical Analysis

All data were processed by GraphPad Prism Software (v 8.0.1, GraphPad Software, San Diego, CA, United States) and SPSS statistics software (v21.0, Chicago, United States). All measured data were expressed as mean ± SD. The comparison between the two groups used t-test. The statistical significance was established as $p < 0.05$.

RESULTS

Research Pipeline

The workflow and methodology of this study is shown schematically in **Figure 1**. The immune score, stromal score, and ESTIMATE score based on the transcriptomic data of 529 LGG cases were calculated. DEGs identified between the low and high immune score, and stromal score groups were used to construct the PPI network and flowed into the univariate Cox regression analysis. By intersecting the hub genes in the PPI network and the significant genes in univariate Cox regression analyses, CCR5 was identified. We next showed a series of analyses of CCR5, including differential analysis, survival analysis, GSEA, clinicopathological characteristics correlation analysis, correlation with TICs, and validation in patient tissues.

Immune Score and Stromal Score Were Correlated With Survival

The role of the TME in the progression and pathological features of several malignancies has been highlighted previously. To exhibit the significance of the TME in LGG, K-M analysis was performed to evaluate the correlation between the immune score–, stromal score–, and

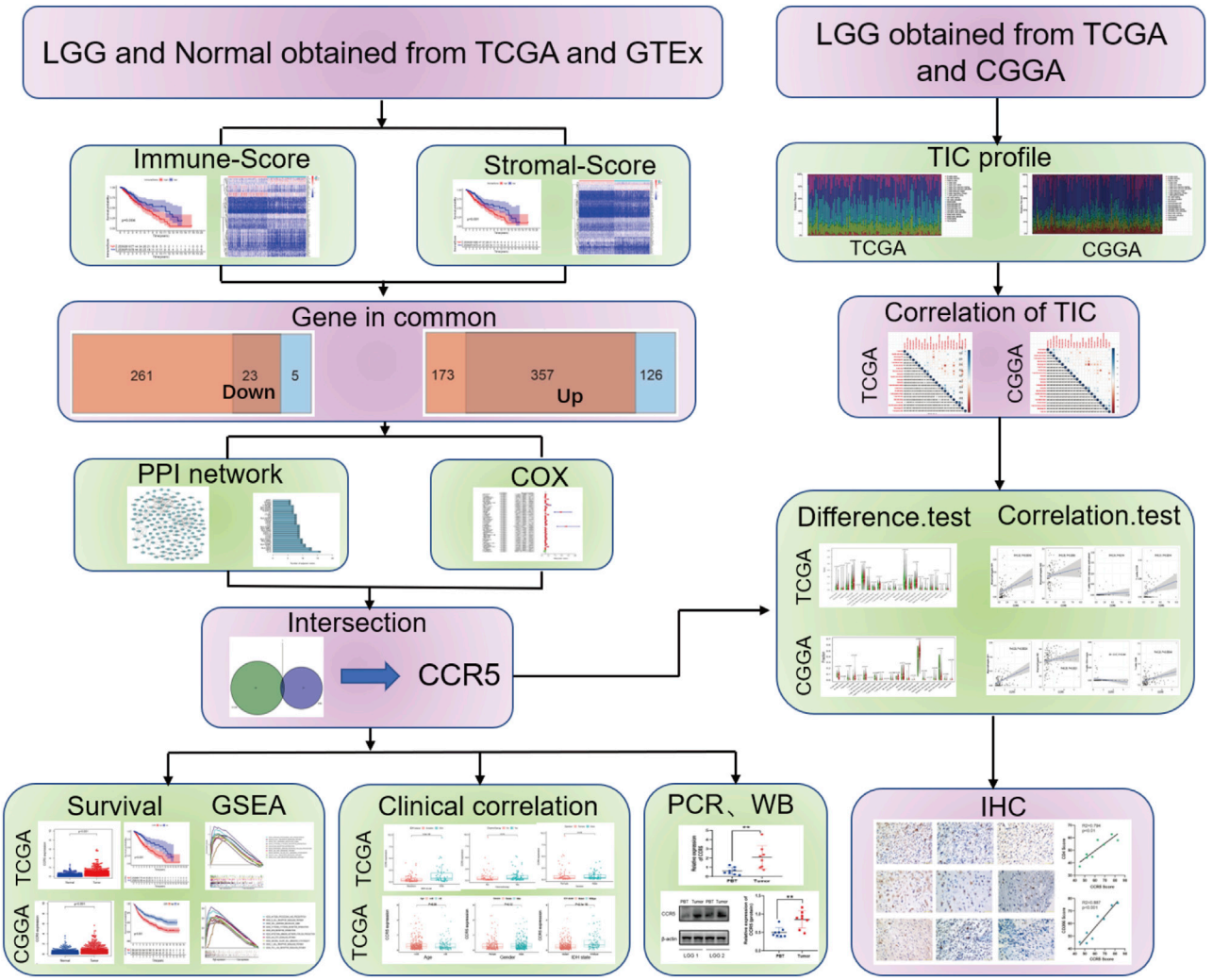


FIGURE 1 | Analysis workflow of this study.

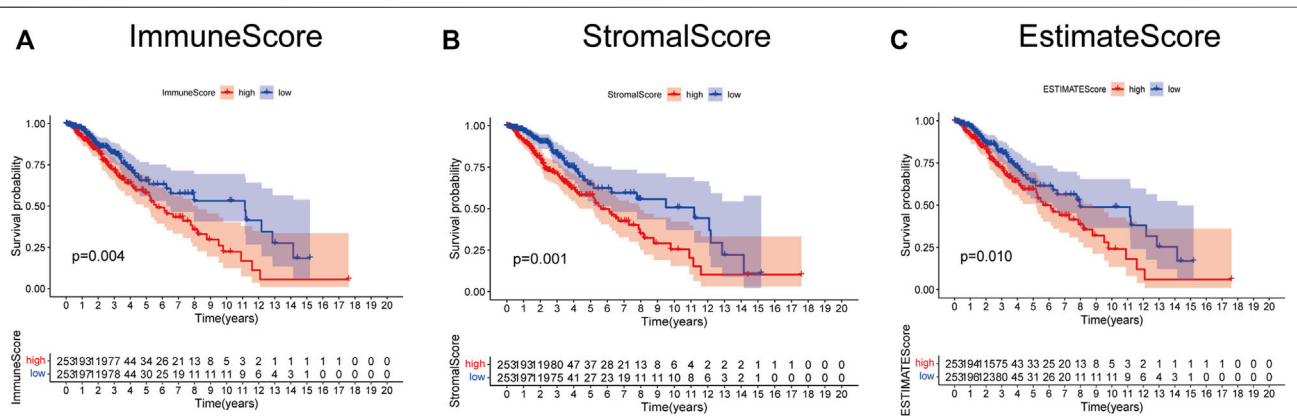


FIGURE 2 | Survival analysis based on scores. Patients in the low immune score **(A)**, stromal score **(B)**, and ESTIMATE score group **(C)** had better prognosis than those in the corresponding high-score group.

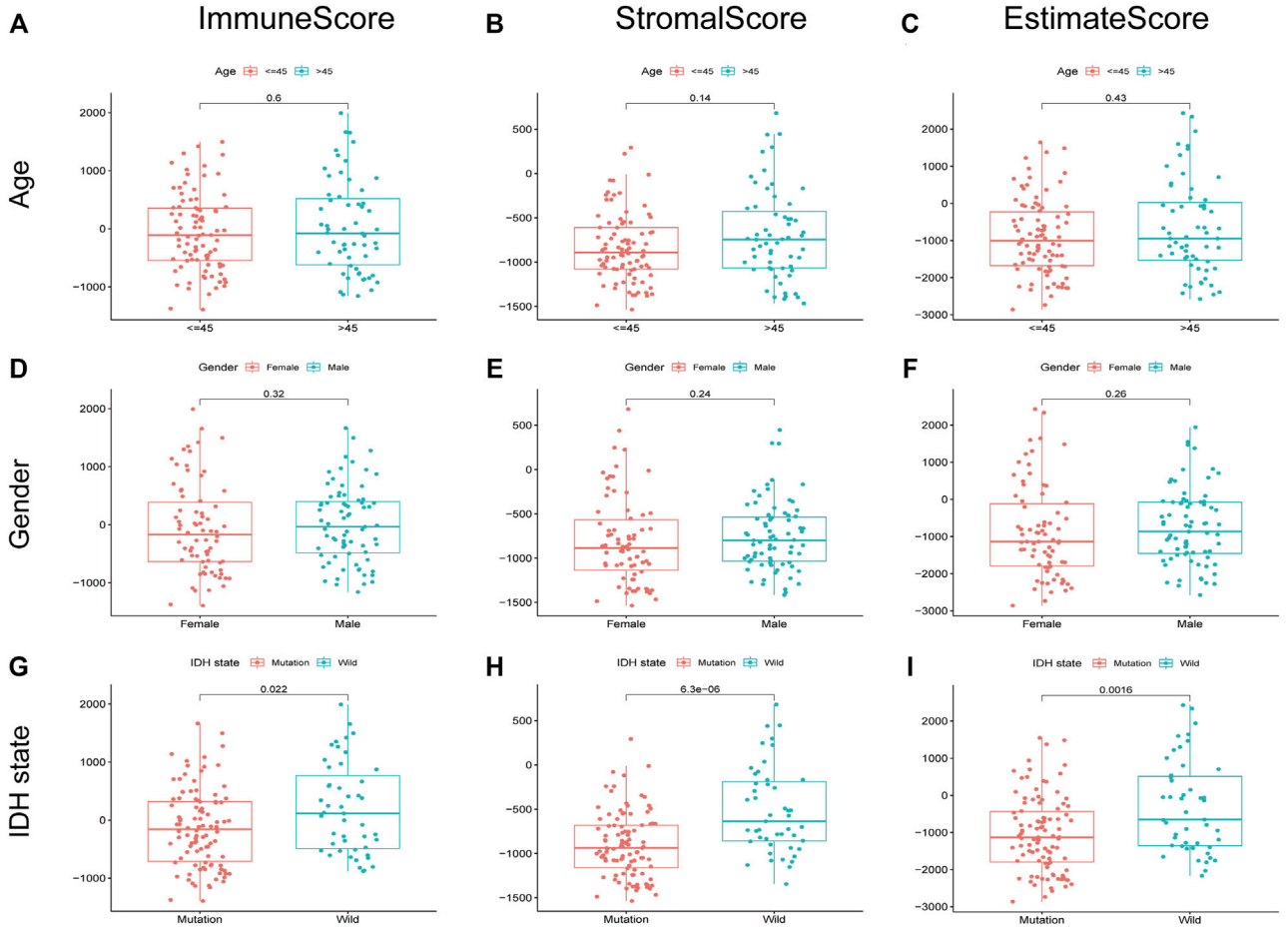


FIGURE 3 | Relationship between scores and clinical characteristics of LGG. immune score, stromal score, and ESTIMATE score were not significantly correlated with age **(A–C)** and gender **(D–F)**. The IDH state was significantly associated with immune score **(G)**, $p = 0.022$), stromal score **(H)**, $p = 6.3e\text{-}06$), and ESTIMATE score **(I)**, $p = 0.0016$).

ESTIMATE score–based groups, and sample survival. As a result, either the immune score–high, stromal score–high, and ESTIMATE score–high group had decreased OS (**Figures 2A–C**). These results indicated that the fraction of immune and stromal components in the TME was negatively correlated with the prognosis of LGG patients. In addition, the immune score, stromal score, and ESTIMATE score of each LGG sample were listed in **Supplementary Table S4**.

Association Between Immune Score, Stromal Score, and ESTIMATE Score With Clinicopathological Parameters of Lower-Grade Gliomas

We evaluated the relationship between the proportion of immune and stromal components with the clinicopathological parameters of LGG. As a result, the correlation between the scores with age and gender was insignificant (**Figure3A–F**). Nevertheless, isocitrate dehydrogenase (IDH) mutation robustly predicted decreased immune score, stromal score, and ESTIMATE score, indicating that LGG with IDH mutation may have less immune and stromal infiltrations (**Figures 3G–I**).

Differentially Expressed Genes Shared by Immune Score and Stromal Score Were Predominantly Involved in the Immune Response

TME-related DEGs were screened by comparing gene expression between high-score and low-score groups regarding immune score and stromal score. Consequently, we identified 814 immune score–related DEGs including 284 down-regulated genes and 530 upregulated genes (**Figure 4A**). Likewise, 511 DEGs were obtained from stromal score, consisting of 28 down-regulated genes and 483 upregulated genes (**Figure 4B**). We also observed that 380 shared DEGs (**Supplementary Table S5**), among which there were 357 upregulated DEGs in immune score and stromal score (**Figure 4D**) and 23 down-regulated DEGs in both the groups (**Figure 4C**). These genes may be functionally

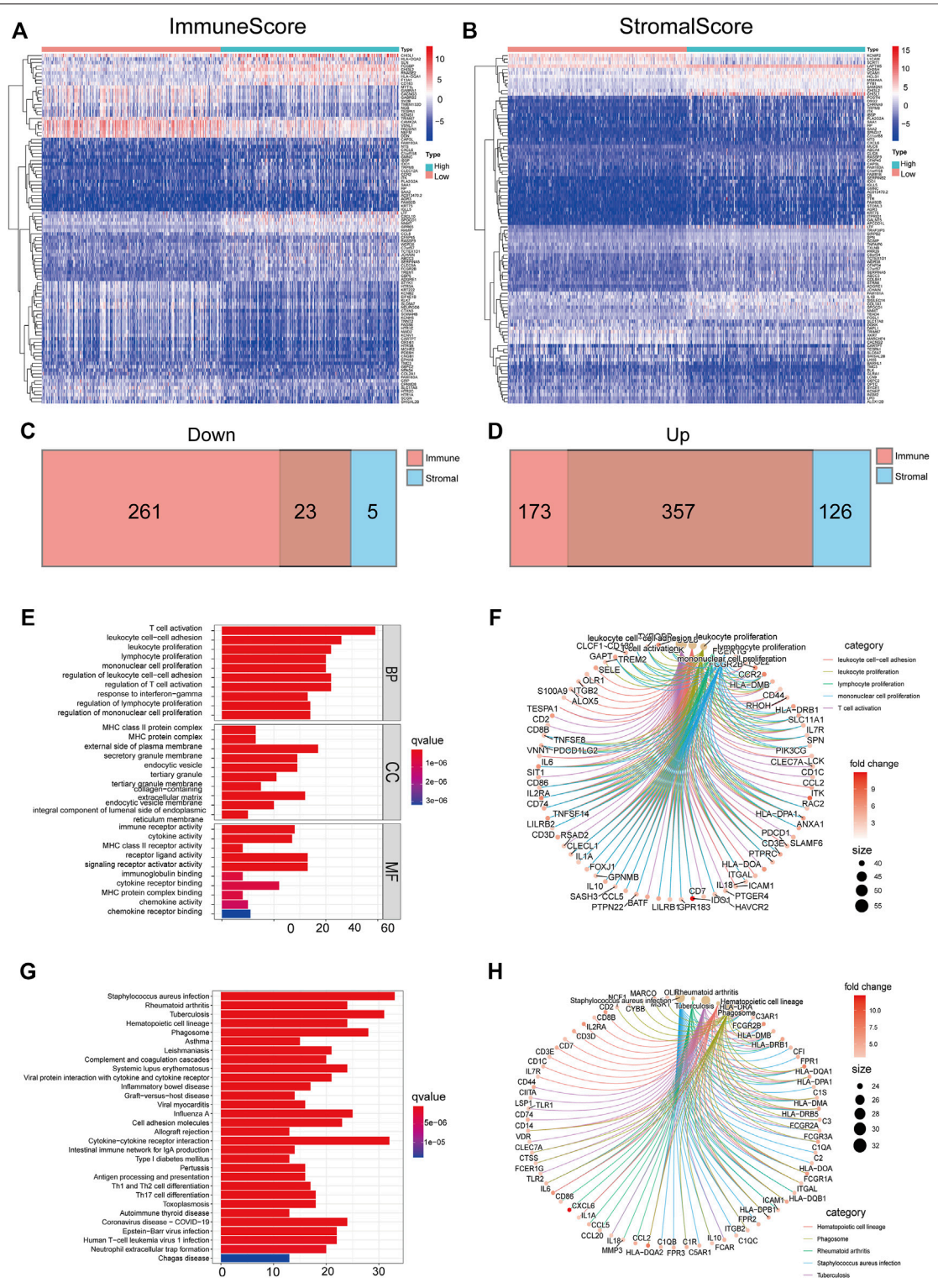


FIGURE 4 | DEGs shared by immune score and stromal score were mainly immune-related genes. The heatmaps of significant DEGs between high and low immune score **(A)** or stromal score **(B)** groups. Venn plots showing common down-regulated **(C)** and upregulated **(D)** DEGs sharing immune score and stromal score, red or blue, are immune or stromal DEGs, respectively. GO-enriched functions **(E,F)** and the KEGG pathway **(G,H)** analysis of DEGs.

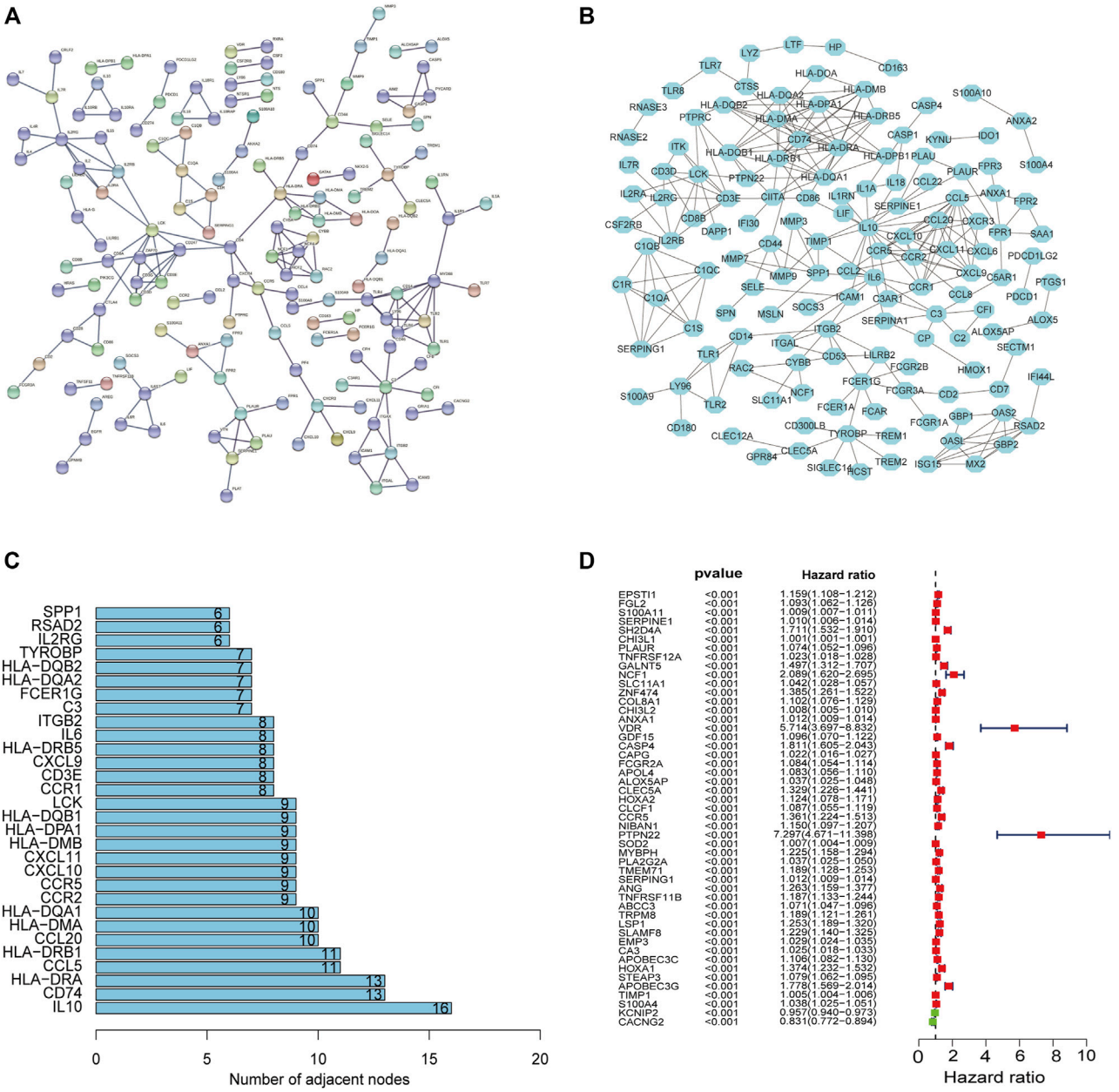


FIGURE 5 | PPI network and univariate Cox regression analyses. The PPI network based on the STRING confidence score > 0.9 **(A)**. The visualization of the PPI network **(B)**. The top 30 genes ordered by the number of nodes **(C)**. Univariate Cox analyses filtered 49 genes ($p < 0.001$) related to prognosis **(D)**.

important with regard to the status of the TME. In addition, functional enrichment analysis found that these genes were predominantly involved in the immune-related GO terms, including T-cell activation, leukocyte proliferation, and leukocyte cell–cell adhesion (**Figure 4E,F**), as well as KEGG pathways including Th1 and Th2 cell differentiation, chemokine signaling pathway, and IL-17 signaling pathway (**Figure 4G,H**). These results illustrated that shared DEGs are strongly related to immune activity, and immune factors play a crucial role in the TME of LGG patients.

Protein–Protein Interaction Network and Univariate Cox Regression Analyses

The PPI network was constructed to show the interactions among 380 DEGs based on the STRING database (confidence value > 0.9) (**Figure5A**). The results were visualized in

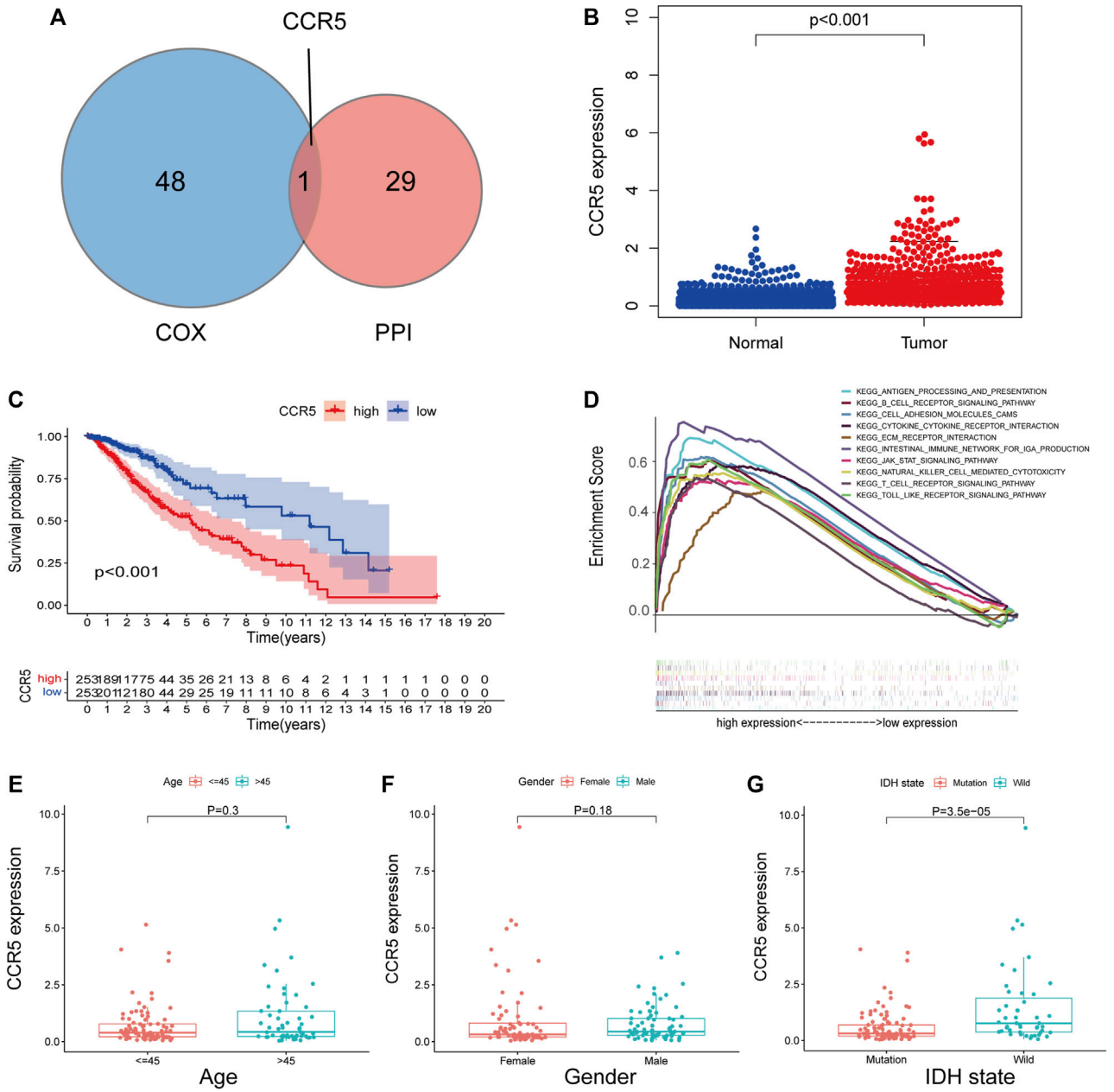


FIGURE 6 | Single-gene analysis of CCR5 in the training set based on the TCGA database. The Venn plot showed CCR5 was identified based on the intersection of the 49 significant genes of Cox regression analyses and the 30 hub genes in the PPI network **(A)**. The expression of CCR5 in normal and tumor tissues of LGG **(B)**. The survival analysis of LGG patients with low and high CCR5 expression **(C)**.The GSEA analysis of CCR5 showed the top ten signaling pathways associated with the high expression of CCR5 **(D)**. The CCR5 expression was not associated with age and gender of patients **(E,F)**. The CCR5 protein expression was higher in LGG with IDH wildtype compared to IDH mutation **(G)**.

Cytoscape v3.8.2 (**Figure 5B**), and then the bar plots were represented for the top 30 genes ranked by the number of nodes (**Figure 5C**), which may serve as hub nodes in the network. In addition, the 380 DEGs were subjected to univariate Cox regression analyses and 49 genes were of prognostic significance (**Supplementary Table S6**) (**Figure5D**).

C-C Chemokine Receptor Type 5 had Potential to be an Indicator of Tumor Microenvironment Modulation

Through the intersection of 49 genes of prognostic significance with the 30 hub genes in the PPI network, CCR5 was identified (**Figure 6A**). Subsequently, the expression and potential role of

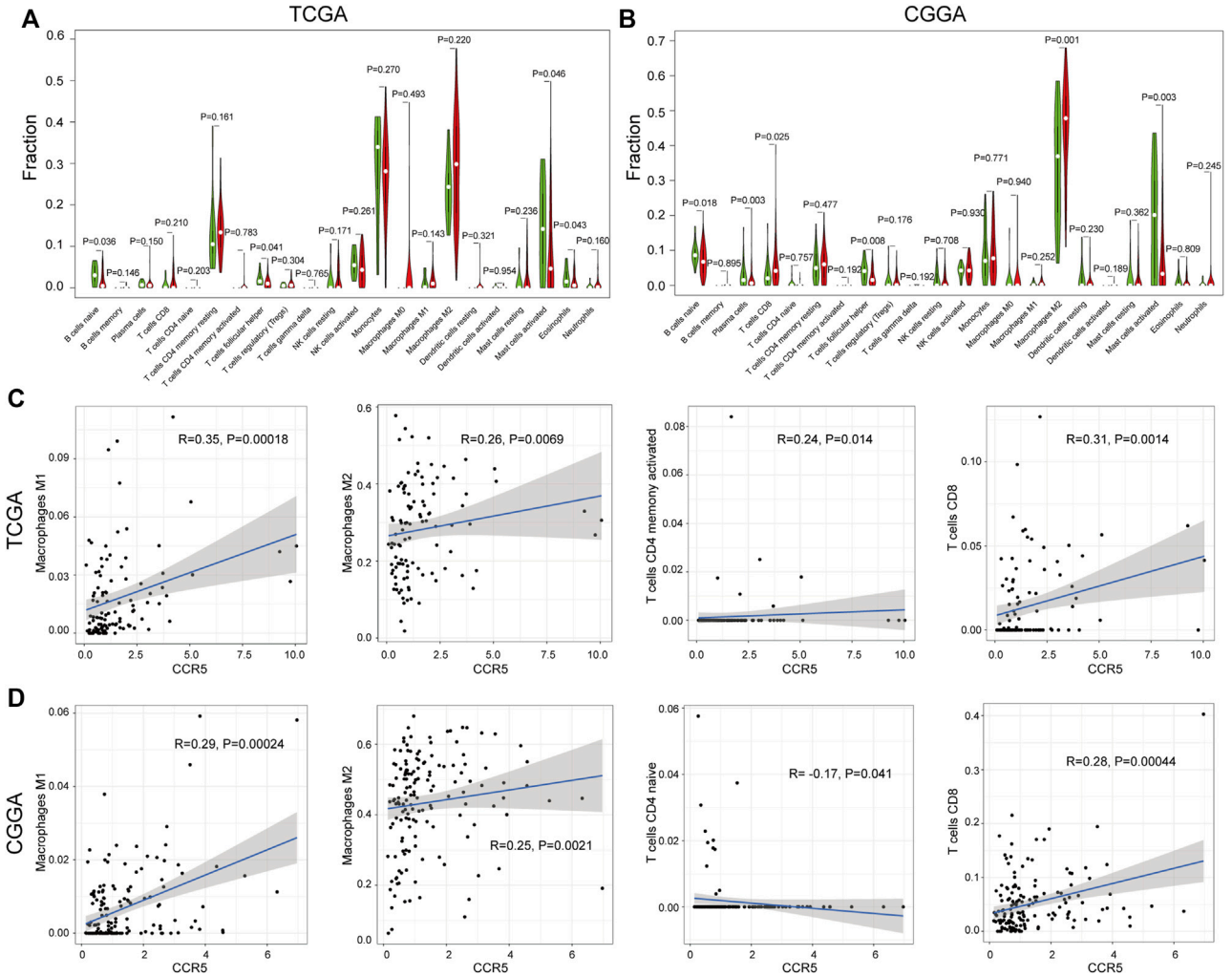


FIGURE 7 | Correlation analysis between CCR5 and TICs. The difference analysis of TICs based on the expression of CCR5 in the TCGA and CGGA database, respectively **(A,B)**. The correlation analysis of CCR5 with macrophages, $CD4^+$ memory-activated, and $CD8^+$ T cells based on the TCGA database **(C)**. The correlation analysis of CCR5 with macrophages, $CD4^+$ naive, and $CD8^+$ T cells based on the CGGA database **(D)**.

CCR5 were explored using the TCGA and CGGA datasets, respectively. The expression of CCR5 was significantly increased in LGG comparing normal tissue ($p < 0.001$, **Figure 6B**; **Supplementary Figure S1A**). K-M analysis showed that LGG patients with decreased expression of CCR5 had prolonged survival (**Figure 6C**; **Supplementary Figure S1B**). Additionally, the top ten signaling pathways enriched in the group with the increased expression of CCR5 were exhibited (**Figure 6D**; **Supplementary Figure S1C**). Notably, these pathways were mainly associated with immune function and tumor progression, such as cell adhesion molecules cams, cytokine–cytokine receptor interaction, T-cell receptor signaling pathway, and JAK/STAT signaling pathway. Similar to the immune score and stromal score, the expression of CCR5 was increased in LGG with wildtype of IDH (**Figure 6G**; **Supplementary Figure S1F**), while it was less associated with age and gender (**Figures 6E,F**; **Supplementary Figures S1D,E**). These results suggested that CCR5 may play a role in the status of TME.

Correlation of C-C Chemokine Receptor Type 5 With Tumor-Infiltrating Immune Cells

To explore the correlation of CCR5 with the TIC, the fraction of tumor-infiltrating immune cells was evaluated (**Supplementary Figures S2A,B**). The correlation analysis of the 22 TICs in LGG was shown in **Supplementary Figures S2C,D**. There were significant differences in naïve B cells, T-cell follicular helpers, and mast cells activated between the CCR5-high and -low groups both in the TCGA and CGGA datasets ($p < 0.05$, **Figures 7A,B**). Among them, macrophages M1, macrophages M2, CD4 memory-activated T cell, and CD8 T cell are positively correlated with the expression of CCR5 (**Figure7C,D**). These results further

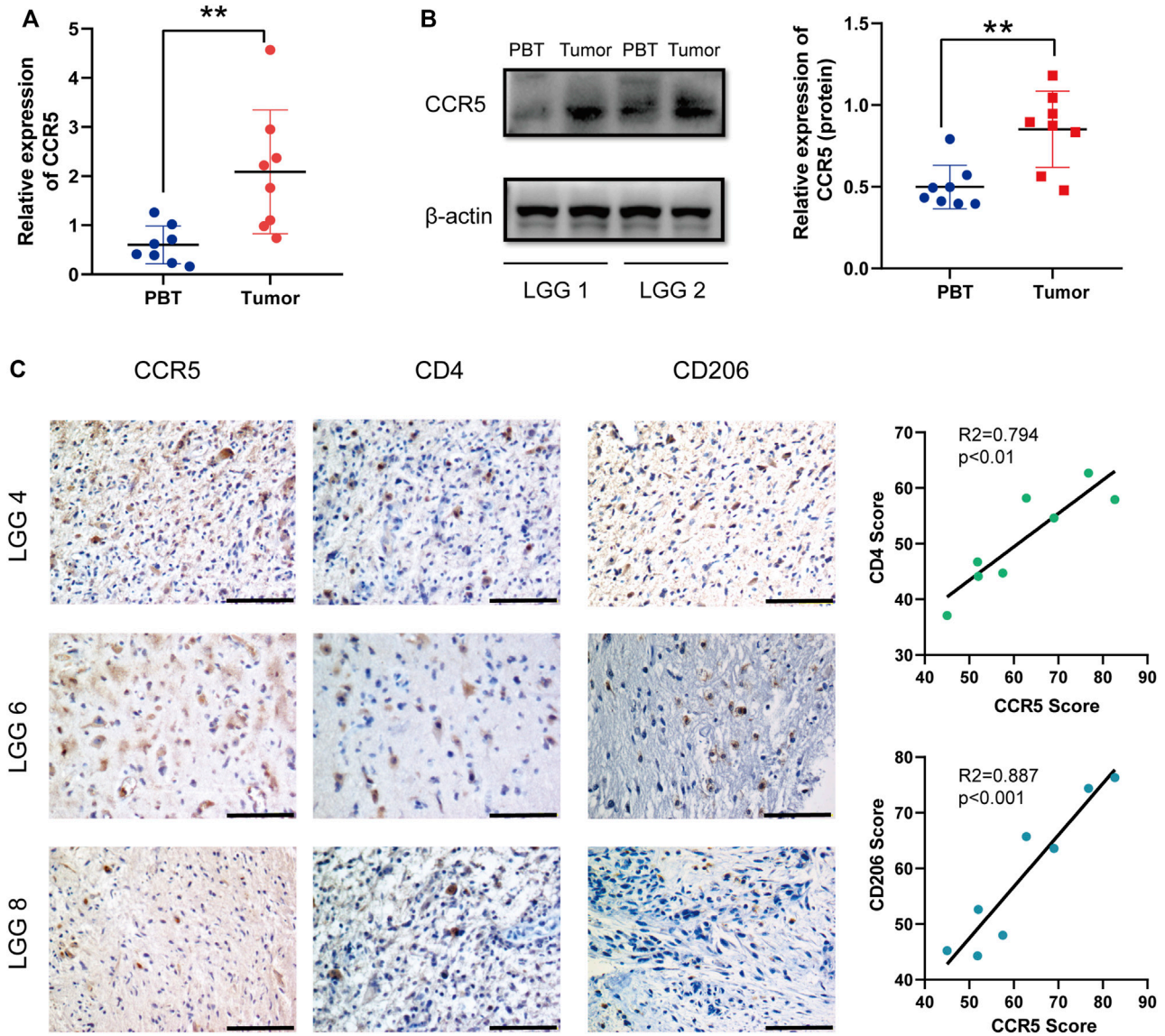


FIGURE 8 | CCR5 is upregulated in LGG tissues and may affect the immune cell infiltration. qRT-PCR **(A)** and Western blot **(B)** were used to detect the expression of CCR5 in PBT and tumor of LGG patients ($n = 8$), $^{**}p < 0.01$. Representative IHC staining images of the expression of CCR5 and the number of $CD4^+$ T cells and M2 macrophages. The CCR5 score was positively correlated with the CD4 score and CD206 score (scale bar, 100 μM) **(C)**.

corroborated our assumption that the expression of CCR5 is involved in the immune features of the TME.

C-C Chemokine Receptor Type 5 Was Upregulated in Lower-Grade Gliomas Tissues and May Affect Immune Cell Infiltration

The results of qRT-PCR of 8 LGG patients showed that the mRNA of CCR5 in tumors was significantly higher than that in PBT (**Figure8A**). To further verify the CCR5 expression from the protein level, Western blot was also conducted based on patient tissues and further verified the significant differences of CCR5 expression between tumor and PBT (**Figure 8B**). Then, we performed immunohistochemical experiments on the tissues of these 8 patients to detect the distribution and expression of CCR5, $CD4^+$ T cells, and macrophages. We found that the tumor of the same patient had higher CCR5 expression than PBT, and the infiltration of $CD4^+$ T cells and M2 macrophages were also higher (**Figure 8C**).

DISCUSSION

It is now broadly accepted that TME contributes to progression and aggravation of glioma (Quail and Joyce

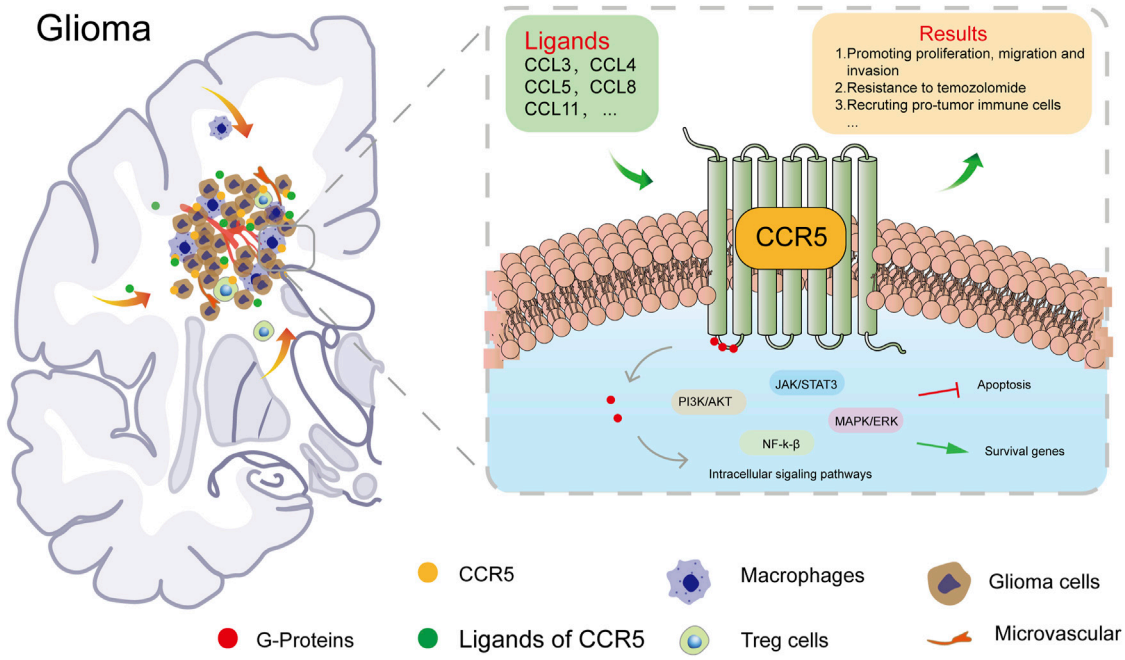


FIGURE 9 | Expression and role of CCR5 in glioma.

2017). Glioma obtains immunotherapeutic resistance through the reprogramming of TME (Bowman et al., 2016; Akhavan et al., 2019). In some cancers, through the binding of CCR5 to its cognate receptors, immune inhibitory signaling pathways were stimulated, resulting in an immune-suppressive state within the TME (Hemmatazad and Berger 2021). Clinical trials have recently started targeting CCR5 in the treatment of metastatic triple-negative breast cancer and colon cancer (Jiao et al., 2019). CCR5 has recently gained a lot of interest as a potential target in glioma. In the present study, we further explored the mechanism of action of CCR5 in glioma by using biological information and experimental methodology. Our findings suggest that CCR5 was associated with patient prognosis and identified to be involved in modulating the TME in LGGs.

The brain TME is composed of stromal components and various nonmalignant cells including immune cells (Oliveira et al., 2019). In recent years, TME has attracted increasing attention and plays a critical role in tumorigenesis (Jung et al., 2019; Xuan et al., 2021). TME is critical for the maintenance of cellular states in primary GBM (Pine et al., 2020). The ESTIMATE algorithm is a comprehensive algorithm for estimating tumor purity based on gene expression data (Chen et al., 2021). We calculated the immune, stromal, and ESTIMATE scores of each patient by using this algorithm. Results showed that these scores were negatively correlated with the OS rate and objectively reflected the effect of TME on the prognosis of LGG patients. Therefore, understanding the interaction between gliomas and noncancerous so-called stromal cells is crucial for the effectiveness of therapeutic strategies.

In this study, CCR5 is presented based on the intersection of the 49 significant genes of Cox regression analyses and the 30 hub genes in the PPI network. CCR5 has been confirmed to be overexpressed in a variety of tumors and affects patient prognosis (Hemmatazad and Berger 2021), such as esophageal, pancreatic, and neck cancers (Aldinucci and Colombatti 2014; Jiao et al., 2019). We confirmed experimentally and bioinformatically that the CCR5 expression was significantly higher in LGG tumor samples than in normal samples, and patients with high CCR5 expression had a worse prognosis. Furthermore, the median survival of LGG patients was correlated with IDH mutation status, and IDH wildtype LGGs have a more aggressive clinical behavior and worse outcome (Cancer Genome Atlas Research et al., 2015; Louis et al., 2016). Interestingly, our results indicated that the CCR5 expression was associated with the IDH mutation status.

The CCR5 in TME promotes tumor growth through different mechanisms (Hemmatazad and Berger 2021). It is a promiscuous receptor that binds with high affinity to CCL4, CCL5, CCL3, and CCL8; other ligands include CCL11 and CCL13 (Jiao et al., 2019). Through the binding of CCR5 to their cognate ligands, immune-activating and inhibitory signaling pathways are stimulated (Zlotnik and Yoshie 2012). The stimulation of recruitment of T regulatory cells (Tregs) and tumor-associated macrophages may promote an immunosuppressive state within the TME (Aldinucci and Colombatti 2014). Nie et al. have demonstrated that the CCR5 expression on macrophages was correlated with the M2 activation status in a mammary tumor mouse model (Nie et al., 2019). Another study found that CCR5 blockade exerts antitumor effects by promoting M1 polarization in an *in vitro* organotypic culture model (Halama et al., 2016).

Based on immunohistochemistry of human LGG samples, we found that M2 macrophages labeled by CD206 were positively correlated with the expression of CCR5. The number of CD206-labeled macrophages is positively correlated with the expression of CCR5, as was $CD4^+$ T cells (**Figure 8C**). Studies demonstrated that CCR5 enhances TGF-β–mediated killing of CD8 (+) T cells in colon cancer by recruiting Tregs (CD4 + FOXP3+) (Chang et al., 2012), and Blattner et al. reported that CCR5 induces the recruitment and homing of Tregs to the TME to stimulate immune evasion and tumor growth (Blattner et al., 2018). These observations suggest that CCR5 may affect tumor progression by mediating the infiltration of immune cells in LGGs.

In addition to influencing the immune infiltration of tumors, CCR5 can promote invasion and tumor cell migration by activating JAK/STAT3, MAPK/ERK, and PI3K/Akt signaling pathways (Huang et al., 2009). The GSEA analysis of CCR5 showed that cell-adhesion molecules, cytokine–cytokine receptor interaction, and JAK/STAT signaling pathway were positively associated with the CCR5 expression. Moreover, chemokines promote tumor angiogenesis by stimulating the secretion of vascular endothelial growth factor and mediating endothelial cell migration and proliferation, resulting in increased tumor cell invasion (Soria and Ben-Baruch 2008). A recent study in melanoma has convincingly demonstrated that the CCR5 expression was significantly elevated in murine melanoma cells and that CCR5 deficiency not only resulted in delayed tumor growth but also suppressed lung metastasis in mouse model (Liu et al., 2019) cancer. These findings imply that CCR5 in TME promotes tumor growth through different mechanisms. In this study, we confirmed that CCR5 could contribute to an immune-suppressive state within TME by influencing the recruitment of tumor-infiltrating immune cells such as tumor-associated macrophages in LGGs. We summarized the expression and role of CCR5 in glioma (**Figure 9**). Furthermore, more molecules and pathways working with CCR5 need to be deeply investigated by enrolling more experiments in our study in the future.

CONCLUSION

Overall, we used the TCGA and CGGA databases to identify the TME-related gene CCR5 as a potential prognostic and immunological biomarker of low-grade glioma. Our study first pointed to the immune-modulating ability of CCR5 and its potential to regulate malignant tumor behavior in LGG, which provides a potential target for LGG 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 authors.

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

All the authors contributed to the experimental design and the analysis of the data. Conception and design: SH, BT, WY, and YD; acquisition, analysis, and interpretation of data: FW, ZT, and ZT; writing, review, and/or revision of the manuscript: FW, ZT, WY, and ZT; technical and material support: JJ and JD; and study supervision: SH, BT, WY, and YD. All the authors have read and commented on the manuscript and approved the final version, experimental design, and the analysis of the data.

FUNDING

This work was supported by the Talent Introduction Project of Zhejiang Provincial People's Hospital (No. C-2021-QDJJ03-01) and the Project of Nanchang Science and Technology Support Plan of Jiangxi Province, China (HONGKE Zi [2021] 129).

ACKNOWLEDGMENTS

Thanks to the persons who facilitated our work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.874896/full#supplementary-material

REFERENCES

Akhavan, D., Alizadeh, D., Wang, D., Weist, M. R., Shepphird, J. K., and Brown, C. E. (2019). CAR T Cells for Brain Tumors: Lessons Learned and Road Ahead. *Immunol. Rev.* 290 (1), 60–84. doi:10.1111/imr.12773

Aldinucci, D., and Colombatti, A. (2014). The Inflammatory Chemokine CCL5 and Cancer Progression. *Mediat. Inflamm.* 2014, 1–12. doi:10.1155/2014/292376

Bachelerie, F., Graham, G. J., Locati, M., Mantovani, A., Murphy, P. M., Nibbs, R., et al. (2014). New Nomenclature for Atypical Chemokine Receptors. *Nat. Immunol.* 15 (3), 207–208. doi:10.1038/ni.2812

Balkwill, F. R. (2012). The Chemokine System and Cancer. *J. Pathol.* 226 (2), 148–157. doi:10.1002/path.3029

Blattner, C., Fleming, V., Weber, R., Himmelhan, B., Altevogt, P., Gebhardt, C., et al. (2018). CCR5(+) Myeloid-Derived Suppressor Cells Are Enriched and Activated in Melanoma Lesions. *Cancer Res.* 78 (1), 157–167. doi:10.1158/0008-5472.CAN-17-0348

Bowman, R. L., Klemm, F., Akkari, L., Pyonteck, S. M., Sevenich, L., Quail, D. F., et al. (2016). Macrophage Ontogeny Underlies Differences in Tumor-specific Education in Brain Malignancies. *Cell Rep.* 17 (9), 2445–2459. doi:10.1016/j.celrep.2016.10.052

Cancer Genome Atlas Research, NetworkBrat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., et al. (2015). Comprehensive, Integrative Genomic

Analysis of Diffuse Lower-Grade Gliomas. *N. Engl. J. Med.* 372 (26), 2481–2498. doi:10.1056/NEJMoa1402121

Chang, L.-Y., Lin, Y.-C., Mahalingam, J., Huang, C.-T., Chen, T.-W., Kang, C.-W., et al. (2012). Tumor-derived Chemokine CCL5 Enhances TGF-β-Mediated Killing of CD8(+) T Cells in Colon Cancer by T-Regulatory Cells. *Cancer Res.* 72 (5), 1092–1102. doi:10.1158/0008-5472.Can-11-2493

Chen, Y., MengZhang, Z. L., Zhang, L., and Liu, F. (2021). CD2 Is a Novel Immune-Related Prognostic Biomarker of Invasive Breast Carcinoma that Modulates the Tumor Microenvironment. *Front. Immunol.* 12, 664845. doi:10.3389/fimmu.2021.664845

Claus, E. B., Walsh, K. M., Wiencke, J. K., Molinaro, A. M., Wiemels, J. L., Schildkraut, J. M., et al. (2015). Survival and Low-Grade Glioma: the Emergence of Genetic Information. *Neurosurg. Focus* 38 (1), E6. doi:10.3171/2014.10.FOCUS12367

Cuddapah, V. A., Robel, S., Watkins, S., and Sontheimer, H. (2014). A Neurocentric Perspective on Glioma Invasion. *Nat. Rev. Neurosci.* 15 (7), 455–465. doi:10.1038/nrn3765

Halama, N., Zoernig, I., Berthel, A., Kahlert, C., Klupp, F., Suarez-Carmona, M., et al. (2016). Tumoral Immune Cell Exploitation in Colorectal Cancer Metastases Can Be Targeted Effectively by Anti-CCR5 Therapy in Cancer Patients. *Cancer Cell* 29 (4), 587–601. doi:10.1016/j.ccell.2016.03.005

Hambardzumyan, D., Gutmann, D. H., and Kettenmann, H. (2016). The Role of Microglia and Macrophages in Glioma Maintenance and Progression. *Nat. Neurosci.* 19 (1), 20–27. doi:10.1038/nn.4185

Hemmatazad, H., and Berger, M. D. (2021). CCR5 Is a Potential Therapeutic Target for Cancer. *Expert Opin. Ther. targets* 25 (4), 311–327. doi:10.1080/14728222.2021.1902505

Huang, C.-Y., Fong, Y.-C., Lee, C.-Y., Chen, M.-Y., Tsai, H.-C., Hsu, H.-C., et al. (2009). CCL5 Increases Lung Cancer Migration via PI3K, Akt and NF-kappaB Pathways. *Biochem. Pharmacol.* 77 (5), 794–803. doi:10.1016/j.bcp.2008.11.014

Jiao, X., Nawab, O., Patel, T., Kossenkov, A. V., Halama, N., Jaeger, D., et al. (2019). Recent Advances Targeting CCR5 for Cancer and its Role in Immuno-Oncology. *Cancer Res.* 79 (19), 4801–4807. doi:10.1158/0008-5472.CAN-19-1167

Jung, E., Alfonso, J., Osswald, M., Monyer, H., Wick, W., and Winkler, F. (2019). Emerging Intersections between Neuroscience and Glioma Biology. *Nat. Neurosci.* 22 (12), 1951–1960. doi:10.1038/s41593-019-0540-y

Korbecki, J., Grochans, S., Gutowska, I., Barczak, K., and Baranowska-Bosiacka, I. (2020). CC Chemokines in a Tumor: A Review of Pro-Cancer and Anti-Cancer Properties of Receptors CCR5, CCR6, CCR7, CCR8, CCR9, and CCR10 Ligands. *Int. J. Mol. Sci.* 21 (20), 7619. doi:10.3390/ijms21207619

Liu, J., Wang, C., Ma, X., Tian, Y., Wang, C., Fu, Y., et al. (2019). High Expression of CCR5 in Melanoma Enhances Epithelial-Mesenchymal Transition and Metastasis via TGFβ1. *J. Pathol.* 247 (4), 481–493. doi:10.1002/path.5207

Lombardi, G., Barresi, V., Castellano, A., Tabouret, E., Pasqualetti, F., Salvalaggio, A., et al. (2020). Clinical Management of Diffuse Low-Grade Gliomas. *Cancers* 12 (10), 3008. doi:10.3390/cancers12103008

Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., et al. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a Summary. *Acta Neuropathol.* 131 (6), 803–820. doi:10.1007/s00401-016-1545-1

Nie, Y., Huang, H., Guo, M., Chen, J., Wu, W., Li, W., et al. (2019). Breast Phyllodes Tumors Recruit and Repolarize Tumor-Associated Macrophages via Secreting CCL5 to Promote Malignant Progression, Which Can Be Inhibited by CCR5 Inhibition Therapy. *Clin. Cancer Res.* 25 (13), 3873–3886. doi:10.1158/1078-0432.Ccr-18-3421

Novak, M., Koprivnikar Krajnc, M., Hrastar, B., Breznik, B., Majc, B., Mlinar, M., et al. (2020). CCR5-Mediated Signaling Is Involved in Invasion of Glioblastoma Cells in its Microenvironment. *Int. J. Mol. Sci.* 21 (12), 4199. doi:10.3390/ijms21124199

Oliveira, M. N., Breznik, B., Pillat, M. M., Pereira, R. L., Ulrich, H., and Lah, T. T. (2019). Kinins in Glioblastoma Microenvironment. *Cancer Microenviron.* 12 (2-3), 77–94. doi:10.1007/s12307-019-00229-x

Pine, A. R., Cirigliano, S. M., Nicholson, J. G., Hu, Y., Linkous, A., Miyaguchi, K., et al. (2020). Tumor Microenvironment Is Critical for the Maintenance of Cellular States Found in Primary Glioblastomas. *Cancer Discov.* 10 (7), 964–979. doi:10.1158/2159-8290.Cd-20-0057

Quail, D. F., and Joyce, J. A. (2017). The Microenvironmental Landscape of Brain Tumors. *Cancer Cell* 31 (3), 326–341. doi:10.1016/j.ccell.2017.02.009

Soria, G., and Ben-Baruch, A. (2008). The Inflammatory Chemokines CCL2 and CCL5 in Breast Cancer. *Cancer Lett.* 267 (2), 271–285. doi:10.1016/j.canlet.2008.03.018

Tan, M. C. B., Goedegebuure, P. S., Belt, B. A., Flaherty, B., Sankpal, N., Gillanders, W. E., et al. (2009). Disruption of CCR5-dependent Homing of Regulatory T Cells Inhibits Tumor Growth in a Murine Model of Pancreatic Cancer. *J. Immunol.* 182 (3), 1746–1755. doi:10.4049/jimmunol.182.3.1746

Velasco-Velazquez, M., Xolalpa, W., and Pestell, R. G. (2014). The Potential to Target CCL5/CCR5 in Breast Cancer. *Expert Opin. Ther. Targets* 18 (11), 1265–1275. doi:10.1517/14728222.2014.949238

Xuan, W., Lesniak, M. S., James, C. D., Heimberger, A. B., and Chen, P. (2021). Context-Dependent Glioblastoma-Macrophage/Microglia Symbiosis and Associated Mechanisms. *Trends Immunol.* 42 (4), 280–292. doi:10.1016/j.it.2021.02.004

Zhang, X.-N., Yang, K.-D., Chen, C., He, Z.-C., Wang, Q.-H., Feng, H., et al. (2021). Pericytes Augment Glioblastoma Cell Resistance to Temozolomide through CCL5-CCR5 Paracrine Signaling. *Cell Res.* 31 (10), 1072–1087. doi:10.1038/s41422-021-00528-3

Zhao, L., Wang, Y., Xue, Y., Lv, W., Zhang, Y., and He, S. (2015). Critical Roles of Chemokine Receptor CCR5 in Regulating Glioblastoma Proliferation and Invasion. *Acta Biochim. Biophys. Sin.* 47 (11), 890–898. doi:10.1093/abbs/gmv095

Zlotnik, A., and Yoshie, O. (2012). The Chemokine Superfamily Revisited. *Immunity* 36 (5), 705–716. doi:10.1016/j.immuni.2012.05.008

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 Wang, Tao, Tian, Jin, Dong, Dai, Yu, Tang and Hu. 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 19 July 2022
DOI 10.3389/fgene.2022.900124

Check for updates

OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Xiaoying Xue,
Second Hospital of Hebei Medical University, China
Lina Qi,
Zhejiang University, China

*CORRESPONDENCE
Yingtai Chen,
yingtaichen@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 05 April 2022
ACCEPTED 29 June 2022
PUBLISHED 19 July 2022

CITATION
Zhao L, Wang W, Niu P, Luan X, Zhao D and Chen Y (2022), The molecular mechanisms of CTHRC1 in gastric cancer by integrating TCGA, GEO and GSA datasets.
Front. Genet. 13:900124.
doi: 10.3389/fgene.2022.900124

COPYRIGHT
© 2022 Zhao, Wang, Niu, Luan, Zhao 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.

The molecular mechanisms of CTHRC1 in gastric cancer by integrating TCGA, GEO and GSA datasets

Lulu Zhao†, Wanqing Wang†, Penghui Niu†, Xiaoyi Luan, Dongbing Zhao and Yingtai Chen*

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Collagen triple helix repeat containing-1 (CTHRC1), highly expressed in multiple human solid tumors, has been identified as a tumor associated protein. However, its specific role and mechanism with immune infiltrates in gastric cancer are still unclear. In this study, we systematically explored and validated the expression and prognostic value of CTHRC1 in gastric cancer by integrating the Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO) and Genome Sequence Archive (GSA) datasets. Compared to adjacent normal tissues, we observed that CTHRC1 was highly overexpressed in tumor sample of multiple cancers. It was revealed that CTHRC1 overexpression was positively correlated with the T stage in gastric cancer but not lymph nodes metastasis from TCGA dataset. In addition, CTHRC1 expression may induce tumor associated macrophage infiltration though GRN/TNFRSF1A and AnxA1/FPR1 pathways and also tumor angiogenesis in gastric cancer. In this context, our results indicate that CTHRC1 plays a pivotal role in regulating the angiogenesis and macrophage infiltration in tumor microenvironment, and also can predict poor prognosis in gastric cancer, suggesting that CTHRC1 might be a promising novel immunotherapy and angiogenesis target for gastric cancer.

KEYWORDS

CTHRC1, gastric cancer, tumor microenvironment, macrophage infiltration, angiogenesis

1 Introduction

As the fourth leading cause of cancer-related deaths and the fifth most common cancer globally, gastric cancer has poor long-term prognosis, despite of the continuous improvement of effective early screening tools and chemotherapy (Sung et al., 2021). Based on the better understanding of the cellular composition and their crosstalk in tumors in recent years, novel therapeutic targets like PD-L1 (programmed death one ligand-1), PD-1 (programmed cell death protein 1), FGFR2 (fibroblast growth factor receptor 2), and et al., are emerging to prolong the prognosis of gastric cancer patients (Vrána et al., 2018; Xie et al., 2019; Zeng et al., 2019). In this context, an increasing number

of specific genomic alterations can potentially serve as treatment targets, and thus assist in reducing the mortality of gastric cancer patients.

As an ECM (extracellular matrix) related protein, Collagen Triple Helix Repeat Containing-1 (CTHRC1) was initially identified in a screening study on differently expressed genes (DEGs) in balloon-injured versus normal rat arteries (Pyagay et al., 2005). Increasing studies have shown that CTHRC1 is aberrantly upregulated in multiple human solid tumors and functionally associated with tumorigenesis, proliferation, invasion, and metastasis (Mei et al., 2020). Specifically, it has been reported that CTHRC1 regulates tumor progression through the CTHRC1/Wnt/β-catenin pathways in breast cancer (Lai et al., 2017), non-small cell lung cancer (Ke et al., 2014), and oral cancer (Liu et al., 2013). Another newly published study showed that the CTHRC1 promotes liver metastasis by reshaping infiltrated macrophages through physical interactions with TGF-βreceptors in colorectal cancer (Zhang et al., 2021). Although CTHRC1 was a promising target for several cancer diseases, its specific role and mechanism with immune infiltrates in gastric cancer are still unclear.

Here, we systematically investigated the expression status and prognostic value of CTHRC1 in gastric cancer by integrating the Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO) and Genome Sequence Archive (GSA) datasets. In addition, we further explored the mechanisms of CTHRC1 in the progression of gastric cancer by focusing on the potential function of CTHRC1 in both macrophage infiltration and angiogenesis to identify a novel immunotherapy and angiogenesis target for gastric cancer treatment.

2 Methods

2.1 Data source and processing

2.1.1 TCGA dataset

A total of 375 gastric cancer cases with gene expression data (HTSeq-FPKM) were collected from TCGA (https://tcga-data.nci.nih.gov/tcga/). Next, Level 3 HTSeq-FPKM data of 375 gastric cancer patients were transformed into transcripts per million (TPM) reads for further analyses. In the TCGA dataset, unknown clinical features or unavailable data were regarded as missing values. These data were divided into two groups, the CTHRC1 high group and the CTHRC1 low group, which accords to the median level CTHRC1 expression.

2.2 GEO dataset

A high-throughput platform (GSE84437, Illumina HumanHT-12 V3.0 expression beadchip) was used for the external validation of the CTHRC1 gene in this study. The expression values at the probe level (probe ID) were converted to the corresponding gene symbol according to their annotation files without further standardization. If multiple probes matched with the same gene, the average value was calculated as the expression value of the gene. Gene expression data together with corresponding clinical information were downloaded from GEO dataset (https://www.ncbi.nlm.nih.gov/geo/). Finally, totally 433 patients with gastric cancer were enrolled for analysis.

2.3 Single-cell RNA sequencing (scRNA-seq) data from GSA

In our previous study, we performed the scRNA-seq on nine untreated non-metastatic gastric cancer patients (Li et al., 2022). A total of 47,304 cells with detectable expression of more than 200 genes were obtained after quality control. Here, in order to explore the CTHRC1 gene, we downloaded our previous scRNA-seq data in this study for more analysis (https://ngdc.cncb.ac.cn/gsa/, data number: CRA002586). Canonical correlation analysis (CCA) (Butler et al., 2018) in Seurat was first applied to the combined gene expression matrix of all samples to correct the batch effects among the experiments. Then, the CCA object was aligned with the *Align Subspace* function to return a new dimensional reduction that was used in cell clustering. Principle component analysis (PCA) was conducted to reduce the dimensionality of the expression matrix. T-SNE dimensionality reduction was performed on the first 30 principle components using the *FindCluster* function in the Seurat package with the default parameters. Canonical marker genes were used to identify the resulting cell clusters as known cell types.

2.4 Functional enrichment analysis

To elucidate the potential functions of CTHRC1, we analyzed the DEGs between the CTHRC1 high and CTHRC1 low groups with R package DESeq2 (Love et al., 2014), by setting a log-fold change larger than 1.5 and an adjusted p value less than 0.05. The identified DEGs were then processed for functional annotation on the Metascape database2 (http://metascape.org) (Zhou et al., 2019). Minimum counts larger than 3, enrichment factors larger than 1.5, and a p value less than 0.01 were set as analysis thresholds.

Gene set enrichment analysis (GSEA) (https://www.gsea-msigdb.org) as used to conduct a gene set enrichment analysis and determine whether a defined set of genes exhibits statistically significant concordant differences between two biological states. In this study, GSEA was performed to elucidate the significant function and pathway differences between the high and low CTHRC1 expression groups using R package clusterProfiler

(Yu et al., 2012). A pathway term with normalized enrichment score (NES) > 1, adjusted p value < 0.05, and false discovery rate (FDR) < 0.25 was considered to be significantly enriched.

2.5 TIMER database analysis

TIMER database (https://cistrome.shinyapps.io/timer/) is a comprehensive web platform, which contains 10,897 samples for the systematic analyses of immune infiltrates across many cancers from the TCGA dataset (Li et al., 2017). The gene module was used to visualize the correlation of CTHRC1 expression with immune infiltration level in gastric cancer.

2.6 Cell-cell interaction analysis

We used CellPhoneDB (v.1.1.0) to detect the pairwise interactions between cell clusters (Efremova et al., 2020). Only receptors and ligands whose expression was detected in more than 25% of cells were included in this analysis. The significance of a ligand and receptor pair in each cell-cell interaction was evaluated by 1,000 random permutations of the cell types. For each permutation, the total mean of the average receptor expression level and the average ligand expression level is calculated, and a null distribution is derived for each ligand-receptor pair.

2.7 Immunofluorescence staining

To confirm the association of CTHRC1 and blood vessel in gastric tumors, we performed immunofluorescence staining of CTHRC1 in the tumor samples. Serial sections (about 4 μm) were obtained from the formalin-fixed paraffin-embedded tumor tissues and stained by use of standard protocols. Anti-CD31 (mouse, 1:50, Proteintech, Ag1787, lot No. 66065-1-Ig) was used to stain endothelial cells, and the corresponding proteins were detected with anti-CTHRC1 (rabbit, 1:50, Proteintech, Ag9812, lot No. 16534-1-AP) and dilutions. DAPI was used to stain cell nuclei.

2.8 Statistical analyses

Spearman correlation analyses were used to assess the correlations between continuous variables. Differences in the variables between the CTHRC1 high and CTHRC1 low groups were evaluated with the Student t test, one-way ANOVA, or Pearson's chi-squared test. The survival data was obtained from TCGA and GSE84437 datasets, while Kaplan-Meier method and Cox proportional hazards regression analysis were used to evaluate prognostic factors. Results with a p value < 0.05 were considered significant. All statistical analyses were performed using R v3.6.1 (https://www.r-project.org/) and Prism 8.0.1 (GraphPad Software Inc.).

3 Results

3.1 The expression level of CTHRC1 in gastric cancer

The pan-cancer analyses were performed to compare the expression of CTHRC1 in the tumor samples of TCGA combined with normal samples of TCGA and the Genotype-Tissue Expression (GTEx) by Wilcoxon rank sum test (Supplementary Table S1). We found that CTHRC1 gene was highly expressed in 26 types of tumor (p < 0.05) (Figure 1A). Figure 1B showed the comparison of the CTHRC1 gene in 32 gastric cancer samples and 375 GC samples from TCGA dataset (p < 0.001). Furthermore, we compared the expression of CTHRC1 in 27 gastric cancer samples and matched normal samples (p < 0.01, Figure 1C).

Receiver operating characteristic curve (ROC) was used to analyze the distinguishing efficacy of CTHRC1 between tumor tissues and normal tissue (Figure 1D). The area under the curve (AUC) of CTHRC1 is 0.966, suggesting that CTHRC1 shows significant high sensitivity and specificity for the diagnosis of gastric cancer. To better understand the relevance and underlying mechanisms of CTHRC1 expression, we investigated the relationship between the CTHRC1 expression and clinical characteristics of gastric cancer samples. As shown in Table 1 and Figure 1E, increased CTHRC1 expression was enriched in T4 stage (p < 0.01).

3.2 Univariate and multivariate analysis of survival outcomes

According to the Kaplan-Meier survival analysis, the low CTHRC1 group (n = 187) showed a better prognosis than the high CTHRC1 group (n = 188) (Figure 2A, HR = 1.48, 95%CI: 1.06–2.05, p = 0.021). We further selected the top 20% patients (n = 75) with high or low CTHRC1 expression form TCGA dataset to analyses (Figure 2B), which showed similar OS results (HR = 2.49, 95CI:1.45–4.30 p = 0.021). To validate the association between CTHRC1 expression and survival outcome, we also used the GSE84437 dataset (n = 433) from the GEO database to verify it. The result demonstrated the same survival trend in both 50% cut-off value (216 vs. 217) and 20% cut-off value (43 vs. 44) groups (Figures 2C, D, p = 0.079 and p = 0.008, respectively).

Figure 2E showed the multivariate analysis after adjusting T stage, N stage, M stage, age and residual tumor, which produced an HR of 1.462 (95% CI: 1.011–2.115, p = 0.044), suggesting that

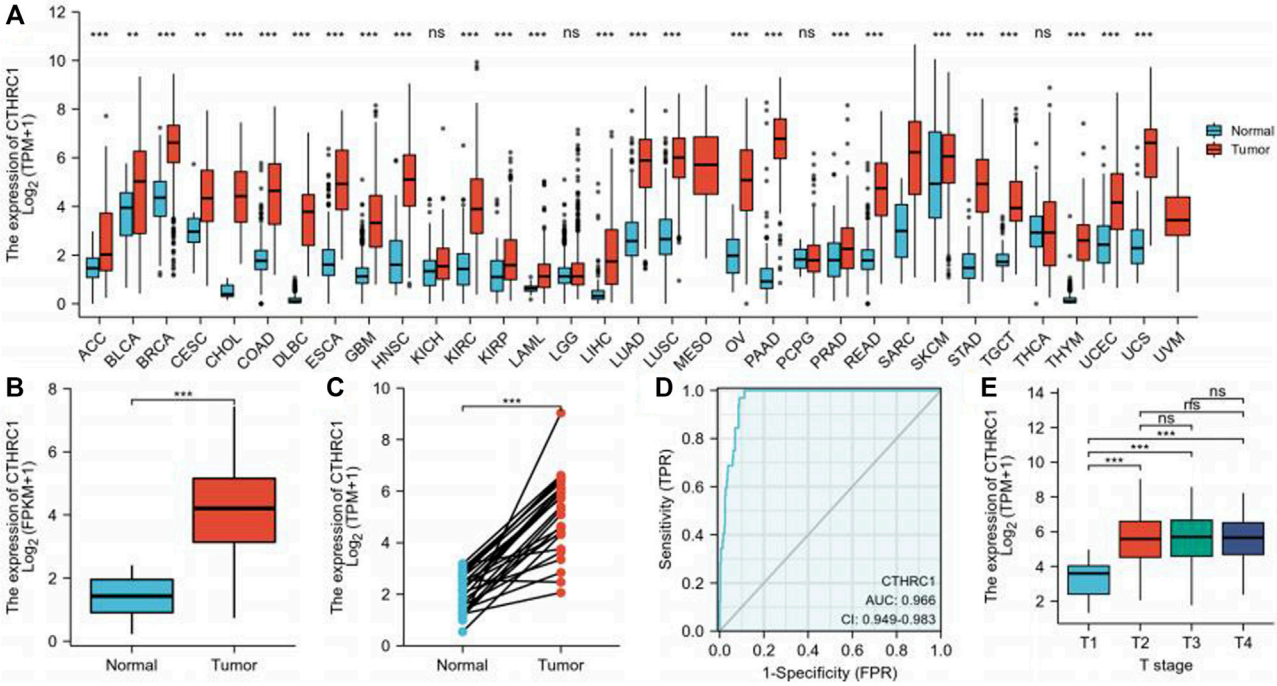


FIGURE 1
CTHRC1 Expression Levels in gastric cancer. **(A)** Higher or lower CTHRC1 expression in multiple tumors when compared with normal tissues across TCGA and the Genotype-Tissue Expression (GTEx). **(B)** CTHRC1 expression levels in gastric cancer and adjacent normal tissues across TCGA. **(C)** The expression of CTHRC1 in gastric cancer and its paired adjacent tissues. **(D)** Operating characteristic analysis (ROC) of CTHRC1 in gastric cancer. **(E)** CTHRC1 expression pattern by different T stage of gastric cancer. (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

CTHRC1 gene was independently associated with mortality of gastric cancer patients.

3.3 CTHRC1 related biological process

We further compared 188 gastric cancer CTHRC1 high samples with 187 CTHRC1 low controls in Table 1. A total of 297 DEGs, covering 145 upregulated genes and 152 downregulated genes, were identified to be statistically significant between the two cohorts (Fugure 3A; Supplementary Table S2). Relative expression values of the top 10 DEGs between the two cohorts were also showed in Figure 3A.

The software "Metascape" was adopted in order to evaluate the function of CTHRC1-associated DEGs in gastric cancer patients. We found that several pathways were enriched, including extracellular matrix organization and vascular development (Figure 3B). Further, GSEA was conducted between the high CTHRC1 and low CTHRC1 expression datasets to reveal significant differences in enrichment of MSigDB Collection (c2. cp.biocarta and hall. V6.1 symbols). The most significantly enriched signaling pathways based on their NES were selected. Moreover, the differentially enriched pathways in CTHRC1 high expression phenotype include pro-tumor and immune associated pathways (Figures 3C–J), such as hallmark epithelial mesenchymal transition (NES = 1.737, p = 0.011, FDR = 0.008), hallmark angiogenesis (NES = 1.598, p = 0.011, FDR = 0.008), go blood vessel morphogenesis (NES = 1.377, p = 0.018, FDR = 0.014), go positive of vasculature development (NES = 1.308, p = 0.018, FDR = 0.014), hallmark inflammatory response (NES = 1.253, p = 0.032, FDR = 0.024), go inflammatory response (NES = 1.290, p = 0.018, FDR = 0.014), go macrophage activation (NES = 1.362, p = 0.024, FDR = 0.018), and go positive regulation of macrophage activation (NES = 1.486, p = 0.049, FDR = 0.038).

3.4 CTHRC1 expression and macrophages infiltration

We next analyzed the correlation between the expression level of CTHRC1 and immune cell infiltration level quantified. As showed in Figure 4A, the expression of CTHRC1 was negatively correlated with the abundance of acquired immunocytes, such as helper T17 cell (Th17) and T helper cell, and meanwhile positively correlated with the abundance of innate immunocytes, such as macrophages, natural killer cell (NK), helper T1 cell (Th1), Dendritic Cells (DC), and DC

TABLE 1 CTHRC1 expression and clinical characteristics in gastric cancer.

Characteristic		CTHRC1 low (n = 187)	CTHRC1 high (n = 188)	p value
Age	≤65	83 (22.4%)	81 (21.8%)	0.737
	>65	100 (27%)	107 (28.8%)	
Age, meidan (IQR)	67 (58, 73)	68 (59, 74)	0.641	
Gender	Female	68 (18.1%)	66 (17.6%)	0.884
	Male	119 (31.7%)	122 (32.5%)	
H pylori infection	No	97 (59.5%)	48 (29.4%)	0.822
	Yes	11 (6.7%)	7 (4.3%)	
Histologic grade	G1	5 (1.4%)	5 (1.4%)	0.123
	G2	77 (21%)	60 (16.4%)	
	G3	99 (27%)	120 (32.8%)	
T stage	T1	19 (5.2%)	0 (0%)	<0.001
	T2	40 (10.9%)	40 (10.9%)	
	T3	80 (21.8%)	88 (24%)	
	T4	48 (13.1%)	52 (14.2%)	
N stage	N0	51 (14.3%)	60 (16.8%)	0.718
	N1	50 (14%)	47 (13.2%)	
	N2	40 (11.2%)	35 (9.8%)	
	N3	39 (10.9%)	35 (9.8%)	
M stage	M0	165 (46.5%)	165 (46.5%)	1
	M1	12 (3.4%)	13 (3.7%)	
Residual tumor	R0	154 (46.8%)	144 (43.8%)	0.933
	R1	8 (2.4%)	7 (2.1%)	
	R2	9 (2.7%)	7 (2.1%)	

interdigitating (iDC). Furthermore, the analyses showed that the CTHRC1 expression was positively correlated with the infiltration of macrophages, as well as M2 macrophages (R = 0.480, $p < 0.001$, Figures 4B, C).

Thus, we used TIMER software to explore the patients with different level of macrophages infiltration and CTHRC1 expression (Figure 4D). In the high CTHRC1 group, the patients with high macrophages infiltration had worse prognosis than those with low macrophages infiltration (p = 0.0144). This tendency was also showed in the low CTHRC1 group, but there was no statistical difference in this aspect between the two groups (p = 0.0937).

3.5 Validation of CTHRC1 expression in scRNA-seq data

To verify the correlation between CTHRC1 and macrophages, we downloaded our previous scRNA-seq data from GSA for analysis. After identifying major cell types, we found 11 fibroblast clusters (F01-F11) from gastric cancer patients (Li et al., 2022). In addition, F04 and F08 were significantly enriched in tumor samples, thus were denoted as three subsets of cancer associated fibroblasts (CAFs).

In this study, we found that CTHRC1 was mainly expressed in fibroblast cells, especially in F04 and F08 clusters (Figures 4E, F). This may indicate that CTHRC1 highly expressed in CAFs. We next performed predictive analyses of the cell-cell interactions based on the expression of ligand-receptor pairs using the statistical inference framework of CellPhoneDB (Efremova et al., 2020). Briefly, we observed more intensive cell-cell interactions in the macrophages than in the other immune cell types (Figure 4G), which was consistent with the results from TCGA dataset showed above. Furthermore, the interactions of CTHRC1 associated fibroblast cells (F04 and F08) and macrophages were mediated by a different set of ligand-receptor pairs (Figures 4H, I), including THY1_aXb2 complex, GRN_TNFRSF1A, CD99_PILRA, CD74_MIF, APP_CD74, ANXA1_FPR1, etc.

3.6 CTHRC1 expression and angiogenesis

To verify the correlation between CTHRC1 and angiogenesis, we constructed co-expression analysis between CTHRC1 and some angiogenesis associated markers in gastric cancer using the TCGA dataset. We found many angiogenesis associated genes showed consistent expression trend of CTHRC1 in gastric

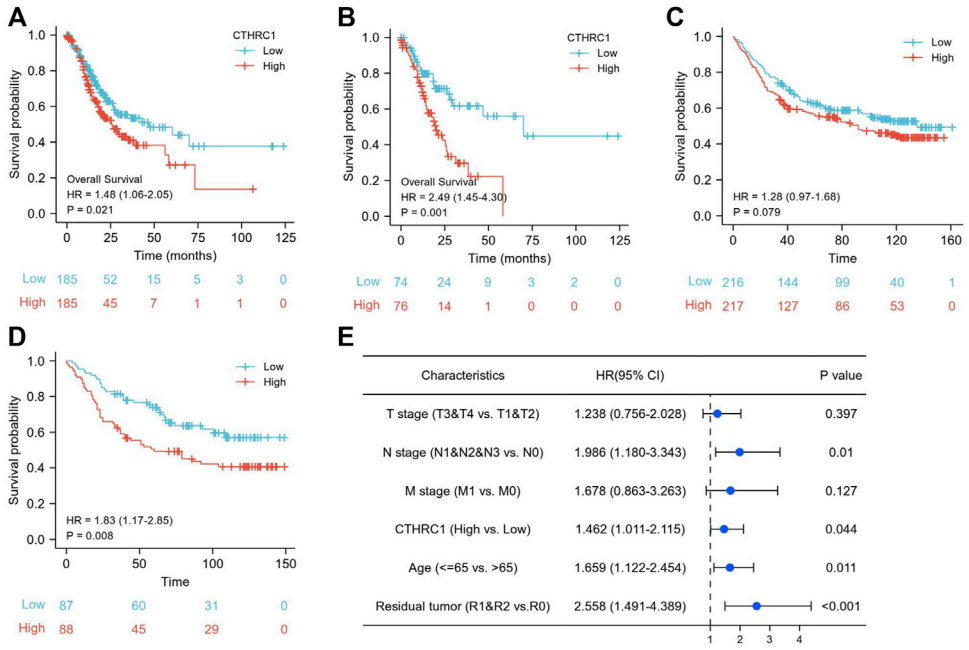


FIGURE 2
The prognostic value of CTHRC1 expression in gastric cancer. **(A)** Kaplan-Meier survival curves comparing the high and low expression of CTHRC1 in gastric cancer using TCGA dataset, and **(B)** also the outcomes with 20% cut off value. **(C)** Kaplan-Meier survival curves comparing the high and low expression of CTHRC1 in gastric cancer using GSE84437 dataset, and **(D)** also the outcomes with 20% cut off value. **(E)** The survival of multivariate analysis using TCGA dataset.

cancer, such as ACVRL1, ANGPTL3, ANGPTL4, NOTCH4, VEGFB, VEGFC etc. (Figure 5A) Immunofluorescence staining showed that CTHRC1 was present in the vascular tissue surrounding the gastric tumors (Figure 5B), suggesting that CTHRC1 may participate in extracellular matrix remodeling in gastric tumor angiogenesis.

4 Discussion

In this study, we presented a comprehensive catalog of CTHRC1 together with its prognostic value and related biological processes in gastric cancer to systematically investigate the expression pattern. We identified that CTHRC1 expression tends to be upregulated in gastric cancer at its later T stage, however, no correlation was found between CTHRC1 expression and lymph node metastasis. In addition, the CTHRC1 high group was associated with worse prognosis in gastric cancer patients in both univariate and multivariate analysis. We also found that CTHRC1 expression was significantly associated with immune response, as well as tumor progression associated pathways, like epithelial mesenchymal transformation and angiogenesis, indicating CTHRC1 gene multi-dimensionally participates the biological process of tumor microenvironment in gastric cancer.

The primary finding of our study was that higher CTHRC1 expression was significantly correlated with macrophages infiltration, although other innate immune cells, like NK cells, Th1 cells and DC cells showed the same trend. We further focus on the M2 like macrophages, and we found M2 macrophages infiltration were strongly associated with CTHRC1 high expression, which provides profound insight into mechanisms governing tumor macrophages infiltration and functional activation. As we all known, M2 macrophages can express a large number of scavenger receptors (Locati et al., 2020), which is related to the high-intensity expression of IL-10, IL-1β and matrix metalloproteins (MMPs) in the tumor microenvironment (Chen et al., 2017), so they have the function of promoting angiogenesis, tissue reconstruction, tumorigenesis and also tumor development.

In order to explore the mechanism of CTHRC1 in inducing macrophage infiltration, we used the scRNA-seq for cell-cell interaction analysis. We found CAFs which highly expressed CTHRC1 (from F04 and F08) and tumor associated macrophages were mediated by a different set of ligand-receptor pairs, including THY1/aXb2 complex, GRN/TNFRSF1A, CD99/PILRA, CD74/MIF, APP/CD74, ANXA1/FPR1, etc.

In our previous study (Li et al., 2022), F04 and F08 were significantly enriched in tumor samples, thus were denoted as

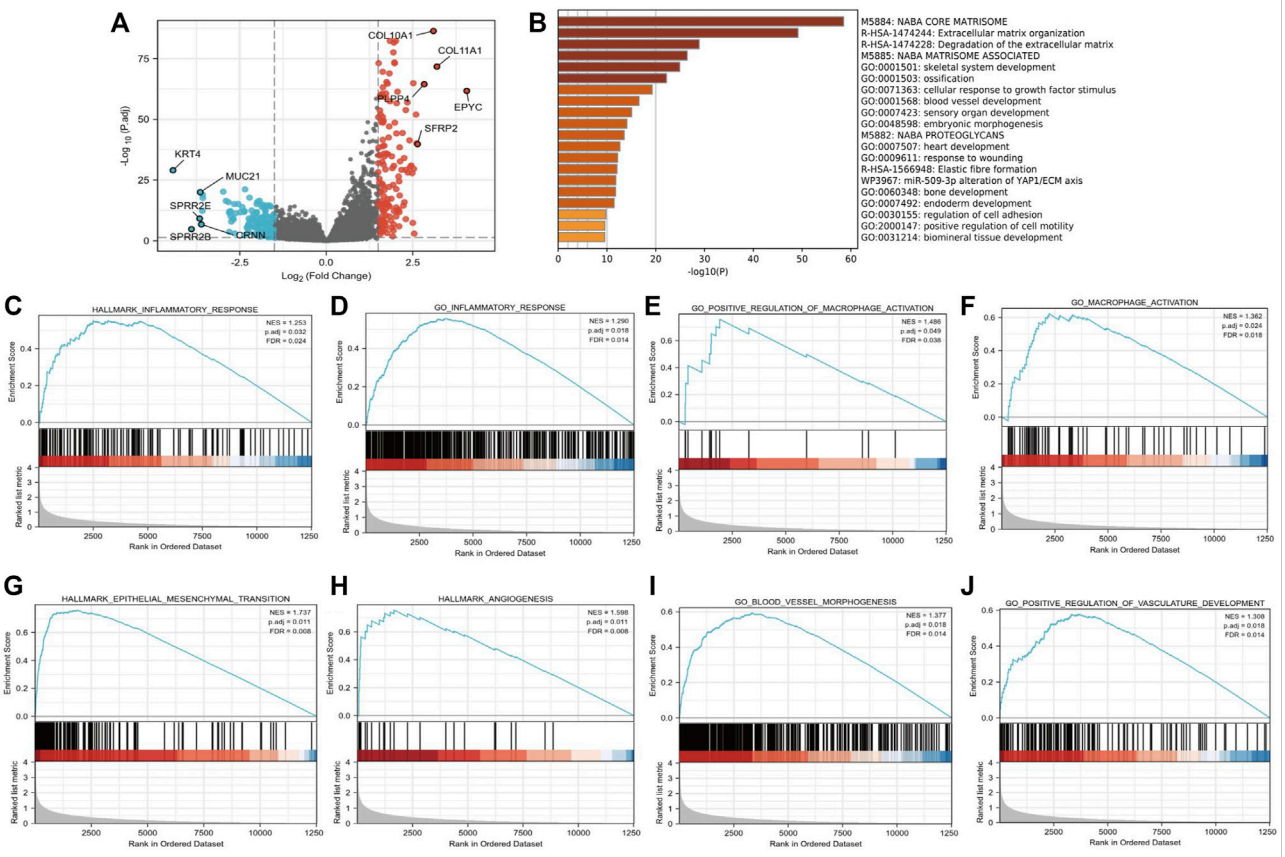


FIGURE 3
Functional annotation of differentially expressed genes (DEGs) in gastric cancer patients with distinct CTHRC1 levels. **(A)** Volcano plots of the DEGs with the top 10 DEGs. **(B)** Pathways enriched in CTHRC1 high groups using Metascape database. **(C–J)** Representative Gene Set Enrichment Analysis between CTHRC1 high and CTHRC1 low expression groups.

three subsets of CAFs. Tang et al. had indicated that GRN bound directly to tumor necrosis factor receptors (TNFRs) and disturbed the TNFα-TNFR interaction (Tang et al., 2011), while TNFα is responsible for a diverse range of signaling events within cells, leading to necrosis or apoptosis, and also important for resistance to cancers (So and Ishii, 2019). In addition, Vecchi has reported that inhibiting AnxA1/FPR1 autocrine axis can reduce breast cancer cell growth and aggressiveness both *in vitro* and *in vivo* (Vecchi et al., 2018). Thus, we may conclude that CTHRC1 may induce tumor associated macrophage infiltration though GRN/TNFRSF1A and AnxA1/FPR1 pathways.

Another concern is the mechanism of CTHRC1 in tumor angiogenesis. Previous studies have demonstrated that CTHRC1 activates HIF-α pathway and contributes to tumor angiogenesis in hepatocellular carcinoma and pancreatic cancer (Zhang et al., 2015; Lee et al., 2016). In gastric cancer, we found CTHRC1 was highly co-expressed with many factors, such as ACVRL1, ANGPTL3, ANGPTL4, NOTCH4, VEGFB, VEGFC etc., which have been proved to play an important role in angiogenesis in various cancer (Siamakpour-Reihani et al., 2015; Zheng et al., 2021). Immunofluorescence staining showed that CTHRC1 was present in the vascular tissue surrounding the gastric tumors, suggesting the high interaction for CTHRC1 and blood vascular development. Furthermore, stromal cells and cancer cells within the tumor microenvironment can be depicted as a highly complex signaling network to influence each other's function and impact on cancer progression and metastasis. Thus, we guess that tumor-associated macrophages induced by CTHRC1 also serve as angiogenesis promoting cells via the production of pro-angiogenic factors and MMPs and also vascular construction which guarantee the supply of oxygen and nutrients to solid tumor cells (Fu et al., 2020; Qian and Pollard, 2010). The angiogenesis progress is complex in gastric cancer, where inhibiting CTHRC1 may decrease tumor blood vascular development. Future studies are needed to focus on investigating the role of CTHRC1 in regulating the angiogenesis targeted animal models.

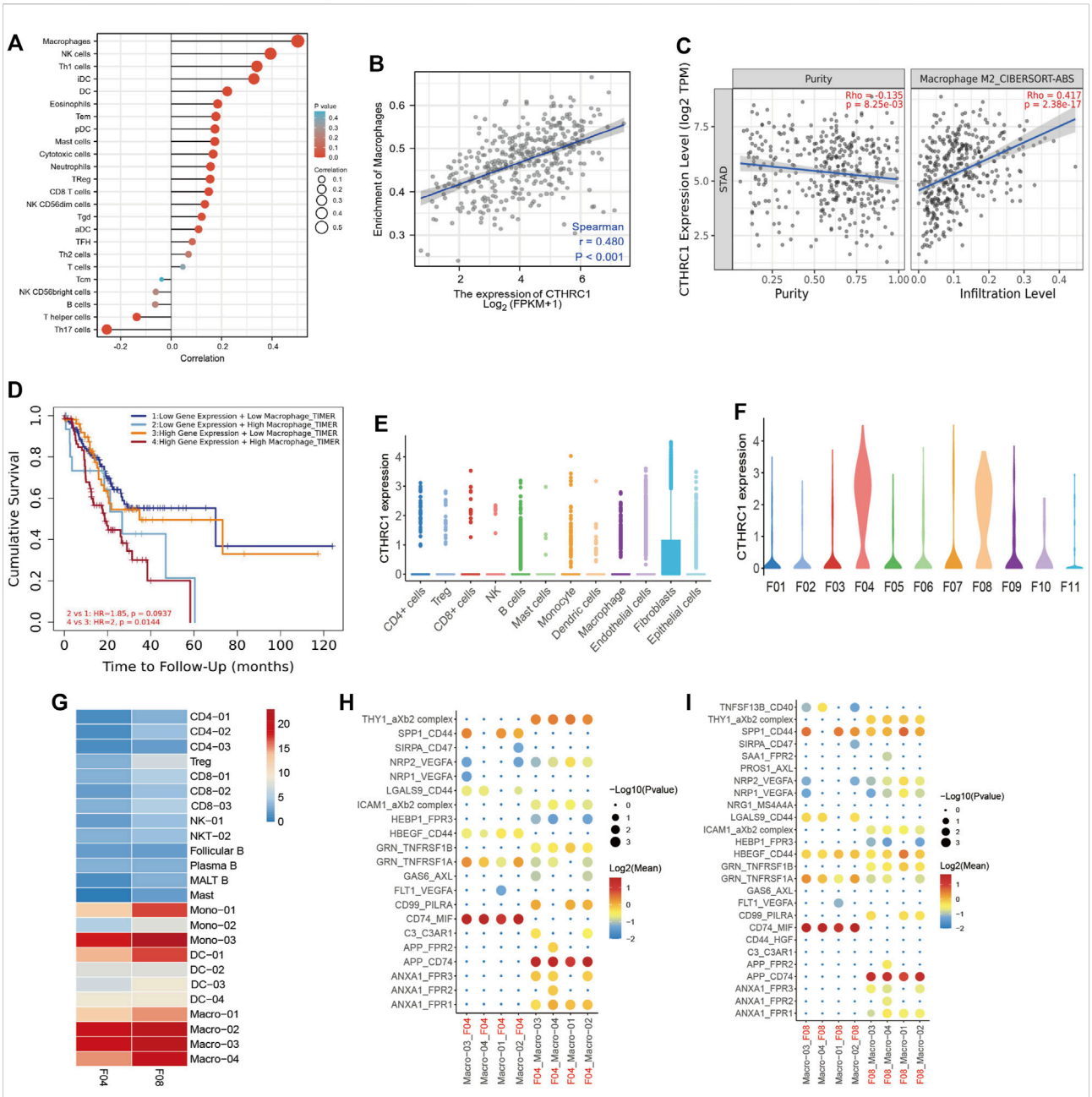


FIGURE 4
CTHRC1 expression and macrophages infiltration. **(A)** The expression level of CTHRC1 was associated with the immune infiltration in gastric cancer. **(B)** CTHRC1 expression significantly positively correlated with the macrophage infiltration, and **(C)** M2 macrophage infiltration. **(D)** Kaplan-Meier survival curves comparing the high and low expression of macrophages infiltration and CTHRC1 expression. **(E)** The CTHRC1 expression in different cell types in gastric cancer using single cell RNA sequencing. **(F)** The CTHRC1 expression in different fibroblast clusters in gastric cancer using single cell RNA sequencing. **(G)** The cell interaction between tumor cells from F04 and F08 and immune cells. **(H–I)** The ligand-receptor pairs between tumor cells from F04 and F08 and immune cells.

This study has several limitations. Firstly, our research does not include the verification of CTHRC1 gene *in vitro* and *in vivo* experiments, which would be our further work. Secondly, the sample size of these datasets, included TCGA, GEO and GSA datasets, is limited. Even though, the results of the present study reveal the molecular mechanisms of CTHRC1 in gastric cancer.

In conclusion, CTHRC1 expression may induce tumor associated macrophage infiltration though GRN/TNFRSF1A and AnxA1/

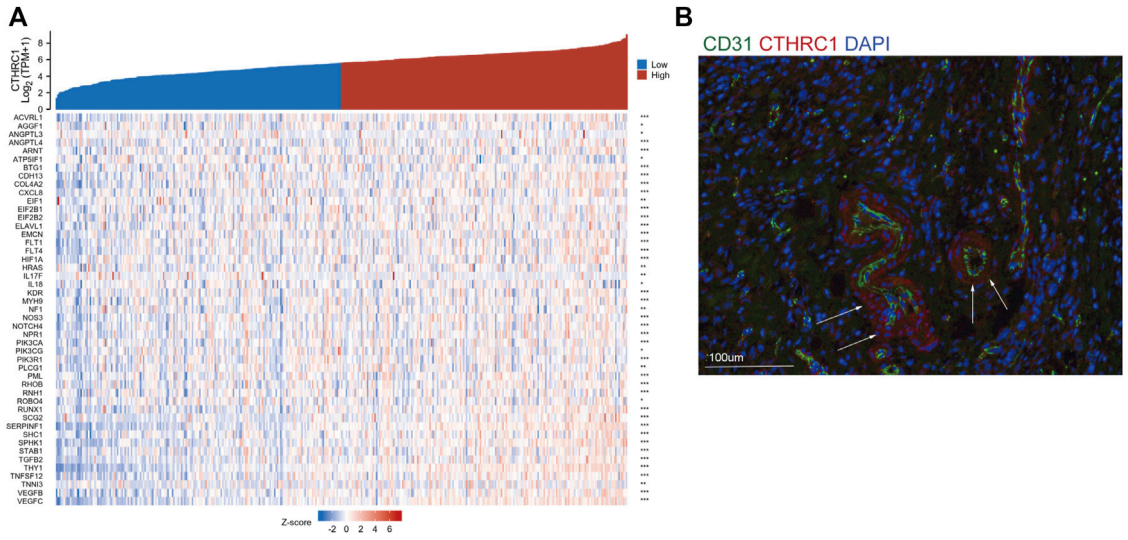


FIGURE 5
CTHRC1 expression and angiogenesis. **(A)** The co-expression between CTHRC1 and some angiogenesis associated genes. **(B)** Immunofluorescence staining of CTHRC1, CD31 and DAPI.

FPR1 pathways and also tumor angiogenesis in gastric cancer. Therefore, our results indicate that CTHRC1 plays a pivotal role in regulating the tumor microenvironment and predicts poor prognosis in gastric cancer, which suggests that CTHRC1 might be a promising novel immunotherapy and angiogenesis target for gastric cancer.

Data availability statement

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

Author Contributions

(I) Conception and design: YC and LZ; (II) Administrative support: DZ and YC (III) Provision of study materials or patients: PN and LZ; (IV) Collection and assembly of data: WW and XL; (V) Data analysis and interpretation: LZ, PN, and WW; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Funding

This work was supported by a grant from National Key R and D Program of China (No. 2017YFC0908300).

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.

Acknowledgments

All authors made substantial contributions to the intellectual content of this paper.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.900124/full#supplementary-material

References

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

Chen, Y., Zhang, S., Wang, Q., and Zhang, X. (2017). Tumor-recruited M2 macrophages promote gastric and breast cancer metastasis via M2 macrophage-secreted CHI3L1 protein. *J. Hematol. Oncol.* 10 (1), 36. doi:10.1186/s13045-017-0408-0

Efremova, M., Vento-Tormo, M., Teichmann, S. A., and Vento-Tormo, R. (2020). CellPhoneDB: Inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* 15 (4), 1484–1506. doi:10.1038/s41596-020-0292-x

Fu, L. Q., Du, W. L., Cai, M. H., Yao, J. Y., Zhao, Y. Y., Mou, X. Z., et al. (2020). The roles of tumor-associated macrophages in tumor angiogenesis and metastasis. *Cell. Immunol.*, 104119. doi:10.1016/j.cellimm.2020.104119

Ke, Z., He, W., Lai, Y., Guo, X., Chen, S., Li, S., et al. (2014). Overexpression of collagen triple helix repeat containing 1 (CTHRC1) is associated with tumour aggressiveness and poor prognosis in human non-small cell lung cancer. *Oncotarget* 5 (19), 9410–9424. doi:10.18632/oncotarget.2421

Lai, Y. H., Chen, J., Wang, X. P., Wu, Y. Q., Peng, H. T., Lin, X. H., et al. (2017). Collagen triple helix repeat containing-1 negatively regulated by microRNA-30c promotes cell proliferation and metastasis and indicates poor prognosis in breast cancer. *J. Exp. Clin. Cancer Res.* 36 (1), 92. doi:10.1186/s13046-017-0564-7

Lee, J., Song, J., Kwon, E. S., Jo, S., Kang, M. K., Kim, Y. J., et al. (2016). CTHRC1 promotes angiogenesis by recruiting Tie2-expressing monocytes to pancreatic tumors. *Exp. Mol. Med.* 48 (9), e261. doi:10.1038/emm.2016.87

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

Li, Y., Hu, X., Lin, R., Zhou, G., Zhao, L., Zhao, D., et al. (2022). Single-cell landscape reveals active cell subtypes and their interaction in the tumor microenvironment of gastric cancer. *Theranostics* 12 (8), 3818–3833. doi:10.7150/thno.71833

Liu, G., Sengupta, P. K., Jamal, B., Yang, H. Y., Bouchie, M. P., Lindner, V., et al. (2013). N-glycosylation induces the CTHRC1 protein and drives oral cancer cell migration. *J. Biol. Chem.* 288 (28), 20217–20227. doi:10.1074/jbc.M113.473785

Locati, M., Curtale, G., and Mantovani, A. (2020). Diversity, mechanisms, and significance of macrophage plasticity. *Annu. Rev. Pathol.* 15, 123–147. doi:10.1146/annurev-pathmechdis-012418-012718

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 (12), 550. doi:10.1186/s13059-014-0550-8

Mei, D., Zhu, Y., Zhang, L., and Wei, W. (2020). The role of CTHRC1 in regulation of multiple signaling and tumor progression and metastasis. *Mediat. Inflamm.* 2020, 9578701. doi:10.1155/2020/9578701

Pyagay, P., Heroult, M., Wang, Q., Lehnert, W., Belden, J., Liaw, L., et al. (2005). Collagen triple helix repeat containing 1, a novel secreted protein in injured and diseased arteries, inhibits collagen expression and promotes cell migration. *Circ. Res.* 96 (2), 261–268. doi:10.1161/01.RES.0000154262.07264.12

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

Siamakpour-Reihani, S., Owzar, K., Jiang, C., Turner, T., Deng, Y., Bean, S. M., et al. (2015). Prognostic significance of differential expression of angiogenic genes in women with high-grade serous ovarian carcinoma. *Gynecol. Oncol.* 139 (1), 23–29. doi:10.1016/j.ygyno.2015.08.001

So, T., and Ishii, N. (2019). The TNF-TNFR family of Co-signal molecules. *Adv. Exp. Med. Biol.* 1189, 53–84. doi:10.1007/978-981-32-9717-3_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. A Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Tang, W., Lu, Y., Tian, Q. Y., Zhang, Y., Guo, F. J., Liu, G. Y., et al. (2011). The growth factor progranulin binds to TNF receptors and is therapeutic against inflammatory arthritis in mice. *Sci. (New York, N.Y.)* 332 (6028), 478–484. doi:10.1126/science.1199214

Vecchi, L., Alves Pereira Zóia, M., Goss Santos, T., de Oliveira Beserra, A., Colaço Ramos, C. M., França Matias Colombo, B., et al. (2018). Inhibition of the AnxA1/FPR1 autocrine axis reduces MDA-MB-231 breast cancer cell growth and aggressiveness *in vitro* and *in vivo*. *Biochim. Biophys. Acta. Mol. Cell Res.* 1865 (9), 1368–1382. doi:10.1016/j.bbamcr.2018.06.010

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), E13. doi:10.3390/ijms20010013

Xie, F., Xu, M., Lu, J., Mao, L., and Wang, S. (2019). The role of exosomal PD-L1 in tumor progression and immunotherapy. *Mol. Cancer* 18 (1), 146. doi:10.1186/s12943-019-1074-3

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 (5), 284–287. doi:10.1089/omi.2011.0118

Zeng, D., Li, M., Zhou, R., Zhang, J., Sun, H., Shi, M., et al. (2019). Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol. Res.* 7 (5), 737–750. doi:10.1158/2326-6066.CIR-18-0436

Zhang, R., Cao, Y., Bai, L., Zhu, C., Li, R., He, H., et al. (2015). The collagen triple helix repeat containing 1 facilitates Hepatitis B virus-associated hepatocellular carcinoma progression by regulating multiple cellular factors and signal cascades. *Mol. Carcinog.* 54 (12), 1554–1566. doi:10.1002/mc.22229

Zhang, X. L., Hu, L. P., Yang, Q., Qin, W. T., Wang, X., Xu, C. J., et al. (2021). CTHRC1 promotes liver metastasis by reshaping infiltrated macrophages through physical interactions with TGF-β receptors in colorectal cancer. *Oncogene* 40 (23), 3959–3973. doi:10.1038/s41388-021-01827-0

Zheng, W., 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

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 (1), 1523. doi:10.1038/s41467-019-09234-6

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 10 August 2022
DOI 10.3389/fgene.2022.954840

Check for updates

OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Qi-Dong Xia,
Huazhong University of Science and Technology, China
Yicheng Liang,
National Cancer Center of China, China

*CORRESPONDENCE
Guanglan Xu,
xgl9592@126.com
Xiaohua Hong,
hongxiaohua613@163.com

†These authors share first authorship

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

RECEIVED 27 May 2022
ACCEPTED 27 June 2022
PUBLISHED 10 August 2022

CITATION
Wang G, Zhao M, Li J, Li G, Zheng F, Xu G and Hong X (2022), m7G-Associated subtypes, tumor microenvironment, and validation of prognostic signature in lung adenocarcinoma.
Front. Genet. 13:954840.
doi: 10.3389/fgene.2022.954840

COPYRIGHT
© 2022 Wang, Zhao, Li, Li, Zheng, Xu and Hong. 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.

m7G-Associated subtypes, tumor microenvironment, and validation of prognostic signature in lung adenocarcinoma

Guangyao Wang[1†], Mei Zhao[1†], Jiao Li[1], Guosheng Li[1], Fukui Zheng[2], Guanglan Xu[1*] and Xiaohua Hong[3*]

[1]Department of Respiratory and Critical Care Medicine, The First Affiliated Hospital of Guangxi University of Chinese Medicine, Nanning, China, [2]Department of Intensive Care Unit, The First Affiliated Hospital of Guangxi University of Chinese Medicine, Nanning, China, [3]Graduate School, Guangxi University of Chinese Medicine, Nanning, China

Background: 7-Methylguanosine (m7G) is an important posttranscriptional modification that regulates gene expression and is involved in tumorigenesis and development. Tumor microenvironment has been proven to be highly involved in tumor progression and prognosis. However, how m7G-associated genes affect the tumor microenvironment of patients with lung adenocarcinoma (LUAD) remains to be further clarified.

Methods: The genetic alterations of m7G-associated genes and their associations with the prognosis and tumor microenvironment in LUAD patients were systemically analyzed. An m7G-Riskscore was established and analyzed for its performance in disease prognosis and association with patient response to immunotherapy. Expression of the model genes at the protein level was investigated through *ex vivo* experiments. A nomogram was finally obtained based on the m7G-Riskscore and several significant clinical pathological features.

Results: m7G-Associated genes were obtained from five LUAD datasets from The Cancer Genome Atlas and Gene Expression Omnibus databases, and their expression pattern was determined. Based on the m7G-associated genes, three LUAD clusters were defined. The differentially expressed genes from the three clusters were screened and used to further divide the LUAD patients into two gene clusters. It was demonstrated that the alterations of m7G-associated genes were associated with the clinical pathological features, prognosis, and tumor immune infiltration in LUAD patients. An m7G-Riskscore including CAND1, RRM2, and SLC2A1 was obtained with robust and accurate prognostic performance. WB and cell immunofluorescence also showed significant dysregulation of CAND1, RRM2, and SLC2A1 in LUAD. In addition, a nomogram was established to improve the clinical feasibility of the m7G-Riskscore. Correlation analysis revealed that patients with a lower m7G-Riskscore had higher immune and stromal scores, responded well to chemotherapeutics and multiple targeted drugs, and survived longer. Patients with a higher m7G-Riskscore tended to suffer from a higher tumor

mutation burden. Furthermore, the m7G-Riskscore exhibited significant associations with immune cell infiltration and cancer stemness.

Conclusion: This study systemically analyzed m7G-associated genes and identified their potential role in tumor microenvironment and prognosis in patients with LUAD. The findings of the present study may help better understand LUAD from the m7G perspective and also provide a new thought toward the prognosis and treatment of LUAD.

KEYWORDS

m7G, lung adenocarcinoma (LUAD), tumor microenvironment, subtype, prognostic

Introduction

7-Methylguanosine (m7G) is an important posttranscriptional modification of messenger RNA with an addition of methyl in the guanine (G) site of the m7G motif under the action of methyltransferase. It plays a regulatory role in various functional processes that take place throughout the life cycle of mRNA (Tomikawa, 2018). Research found that m7G is important in the regulation of efficient gene expression and cell viability (Bi et al., 2020). Increasing evidence has suggested that abnormal RNA methylation contributes to the occurrence and progression of cancers in humans (Xie et al., 2020). m7G, being one of the modifications of RNA methylation, plays a vital role in lung cancer progression. Lung cancer ranks top in terms of mortality among the types of cancer globally (Sung et al., 2021) and is characterized by high aggressiveness, strong drug resistance, and active angiogenesis. Ma et al. (2021) reported that METTL1 and WDR4, which are components of m7G methyltransferase, exhibited an increase in expression in lung cancer tissues compared with normal tissues, which was associated with poor prognosis. However, knockout of METTL1 and WDR4 decreased the potential of cancer cells in proliferation, invasion, and oncogenesis. Another study found that METTL1 advanced the translation of VEGFA mRNA dependent on m7G methylation, resulting in increased angiogenesis (Zhao et al., 2021).

Tumor microenvironment (TME) provides a condition with diverse and complex compositions, including tumor and non-tumor cells (such as fibroblasts, endothelial cells, and immune cells). It plays a role in tumor angiogenesis under the coordination of the circulatory and lymphatic systems as well as related cells. In addition, TME is important in the induction of immune tolerance and is conducive to invasion advancement and tumor progression. Active angiogenesis is the distinct feature of lung cancer and is also an important process in TME. It was reported that m7G methylation can advance angiogenesis. However, the relationship between m7G methylation and TME has been less studied. A comprehensive analysis on such a relationship can help better understand m7G-mediated cell infiltration within the TME of different lung cancer subtypes and reveal the potential mechanisms to increase the prognosis of lung cancer and provide new insight into the treatment.

In the present study, patients with lung adenocarcinoma (LUAD), the most common histological subtype of lung cancer (Shiba-Ishii, 2021; Succony et al., 2021), were subclassified into three molecular subtypes based on the expression of m7G-associated genes. Differentially expressed genes (DEGs) among the three subtypes were screened and used to further stratify the patients into two gene-clusters. The molecular characteristics, TME, and prognosis of the m7G-associated subtypes were assessed. An m7G-Riskscore prognostic for the overall survival (OS) of LUAD patients was established, and a nomogram of clinical significance was correspondingly generated. Expression of the model genes at the protein level was examined through *ex vivo* experiments. Taken together, this study systemically analyzed the relationship between m7G-associated genes and the prognosis, TME, and drug sensitivity in LUAD patients, which may be beneficial for the accurate prognosis of the LUAD patients and provide potential therapeutic targets.

Materials and methods

Data source

TCGA-LUAD data in the FPKM format were downloaded on Nov. 23, 2021, from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer) and then converted into TPM. In the meantime, four LUAD datasets, GSE13213, GSE31210, GSE68465, and GSE30219, were obtained from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Data from the five datasets were combined and normalized using the R software to eliminate potential batch effects (Xia et al., 2022a). Corresponding clinical data were also extracted, and samples with missing data of OS or exhibiting OS as 0 were excluded. Ethical approval for the study was waived given that all data used in the study were from public databases. The workflow of the study is displayed in Figure 1(By Figdraw, ID:YPPUY42f39).

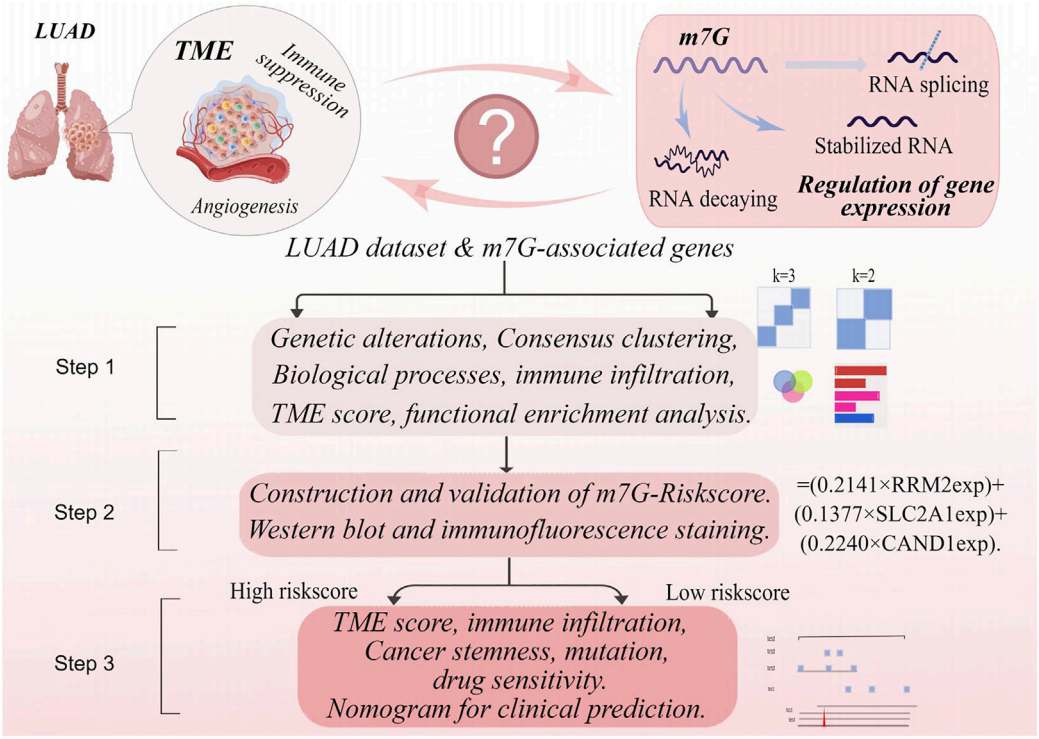


FIGURE 1
The workflow of the study.

Genetic alterations of m7G-associated genes in LUAD

m7G-Associated genes (n = 29) were obtained from the published review of Tomikawa (2018). Detailed information is shown in Supplementary Table S1. Somatic mutation occurs in case of DNA damage during cell activity or DNA replication error during cell division without correct repair. It is associated with aging and disease onset, whereas cancer is the most well-known (Martincorena and Campbell, 2015; Milholland et al., 2015; Li et al., 2021a). Here, R 4.1.1 was applied to visualize the genetic alterations of the 29 m7G-associated genes in LUAD.

Consensus clustering analysis and clinical prognosis

Consensus clustering analysis based on gene expression matrix is conducive to the diagnosis and treatment of cancer (Yu et al., 2015). Here, the R package "ConsensusClusterPlus" (Wilkerson and Hayes, 2010) was used to perform consensus clustering analysis based on the expression of m7G-associated genes to divide LUAD samples into different molecular subtypes. Reduced intergroup correlation while increased intragroup correlation is regarded as eligible. To study the clinical value of the m7G-based subtyping, clinical pathological features, including age, gender, and stage, were analyzed among the different subtypes. Moreover, the difference in OS among the different subtypes was explored using R packages "survival" and "survminer" (Qi-Dong et al., 2020; Li et al., 2021b).

Biological processes and immune infiltration in different subtypes

Gene set variation analysis (GSVA) (Hänzelmann et al., 2013) was performed using R packages "limma," "GSEABase," "GSVA," and "pheatmap" to study whether the biological processes among different LUAD subtypes are different. In addition, single-sample gene set enrichment analysis (ssGSEA) was conducted to estimate the immune infiltration using R packages "reshape2," "ggpubr," "limma," "GSEABase," and "GSVA" (Wang et al., 2021).

Identification of differentially expressed genes and functional enrichment analysis

DEGs among the three LUAD subtypes were screened using the R package "limma" following p.adj < 0.001 and then

combined and visualized on a Venn diagram using the package "VennDiagram." Detailed information of DEGs is presented in Supplementary Table S2. Packages "clusterProfiler," "org.Hs.eg.db," "enrichplot," and "ggplot2" were applied to explore the potential biological functions and signaling pathways of the DEGs (Xia et al., 2022b).

Construction and validation of m7G-Riskscore

DEGs were processed for univariate regression analysis using R packages "limma" and "survival" to screen genes of prognostic significance ($p < 0.05$). Then, the genes identified were used to divide LUAD samples into different gene-clusters using the package "ConsensusClusterPlus." All LUAD samples (n = 1,289) were equally grouped into training (n = 645) and test (n = 644) sets. The training set was applied to construct a prognostic model associated with m7G. During the modeling, the genes screened *via* univariate analysis were successively analyzed in LASSO regression model to eliminate overfitting and multivariate regression model to identify the most significant prognostic DEGs. The m7G-Riskscore was accordingly established and formulated as follows (Yu et al., 2015):

$$\text{m7G} - \text{Riskscore} = \sum_{i-1}^{n} (\text{coefi} \times \text{m7Gexpi})$$

where coef represents the multivariate regression coefficient and m7Gexp represents the expression of each gene involved in the model. On the basis of the m7G-Riskscore, patients in the training set (n = 645) were scored and grouped with the threshold set as the median value. Survival analysis and principal component analysis (PCA) were performed using R packages "survival" and "survminer," respectively. Distribution of the m7G-Riskscore and expression of the model genes in two groups were visualized using the R package "pheatmap." ROC curves were plotted using the package "timeROC" (Bai et al., 2021). In parallel, samples in the total (n = 1,289), test (n = 644), and GSE30219 (n = 85) sets were processed for the same analysis, and corresponding graphs were generated. Moreover, uni- and multivariate analyses were performed to study the independence of the m7G-Riskscore in the prognosis of LUAD using the package "survival." Subgroup analysis was also devised based on the clinical pathological features of LUAD in an attempt to explore whether the m7G-Riskscore remains to be powerful in prognosis in different subgroups.

Immune infiltration, cancer stemness, mutation, and drug sensitivity analysis

The CIBERSORT algorithm (Chen et al., 2018) was calculated to estimate the abundance of 23 immune infiltrates in high- and low-risk groups. The ESTIMATE algorithm was applied to obtain stromal, immune, and ESTIMATE scores (Wu et al., 2021). The association of immune infiltrates with the model genes was investigated. In addition, cancer stemness between the two groups was analyzed. Moreover, somatic mutation in the two groups was explored, and the results were visualized *via* a waterfall plot using the R package "maftools." At last, therapeutic response in two groups was investigated using the half-maximal inhibitory concentration (IC50) of multiple agents.

Establishment and validation of nomogram for clinical prediction

According to the Cox regression analysis, the variables with significant prognostic values were included and combined to establish a nomogram for clinical prediction. Each patient was conferred a score, which was predictive for the 1-, 3-, and 5-year OS. Furthermore, a calibration curve was generated to assess the consistency between the predicted result and the real clinical result. Packages "survival," "regplot," and "rms" were adopted.

Cell culture

Human normal lung cell line MRC-5 was cultured in MEM supplemented with 10% fetal bovine serum (FBS), whereas LUAD cell line A549 was maintained in F12K medium containing 10% FBS. The culture environment was an incubator with 5% CO_2 and a temperature of 37°C. After routine culture, the two cell lines were further cultured on a pretreated round cover slip until 60–70% confluency was reached. Afterward, the medium was absorbed. The cells were washed with PBS thrice, fixed with 4% paraformaldehyde, and finally washed again with PBS thrice.

Western blot and immunofluorescence staining

MRC-5 and A549 cells were collected using a cell scraper, followed by total protein extraction with an addition of cell lysis buffer (R0010, Solarbio). The protein concentration was determined using the BCA Protein Assay Kit (PC0020-50T, Solarbio). Western blot assay was applied to examine the protein levels of RRM2 (1:1000, A3424), SLC2A1 (1:1000, A11727), and CAND1 (1:1000, A14287). For immunofluorescence staining, decidual cells were obtained *via* 10-min culture with a permeabilization reagent. Dulbecco's phosphate-buffered saline (DPBS) washing was performed thrice. After 15 min of digestion, the cells were washed thrice with DPBS and then blocked for 30 min. Excessive solution was absorbed. Primary antibodies, including RRM2 (1:500, 11661-1-AP), SLC2A1 (1:200, PB9435), and CAND1 (1:500, 100869-T10),

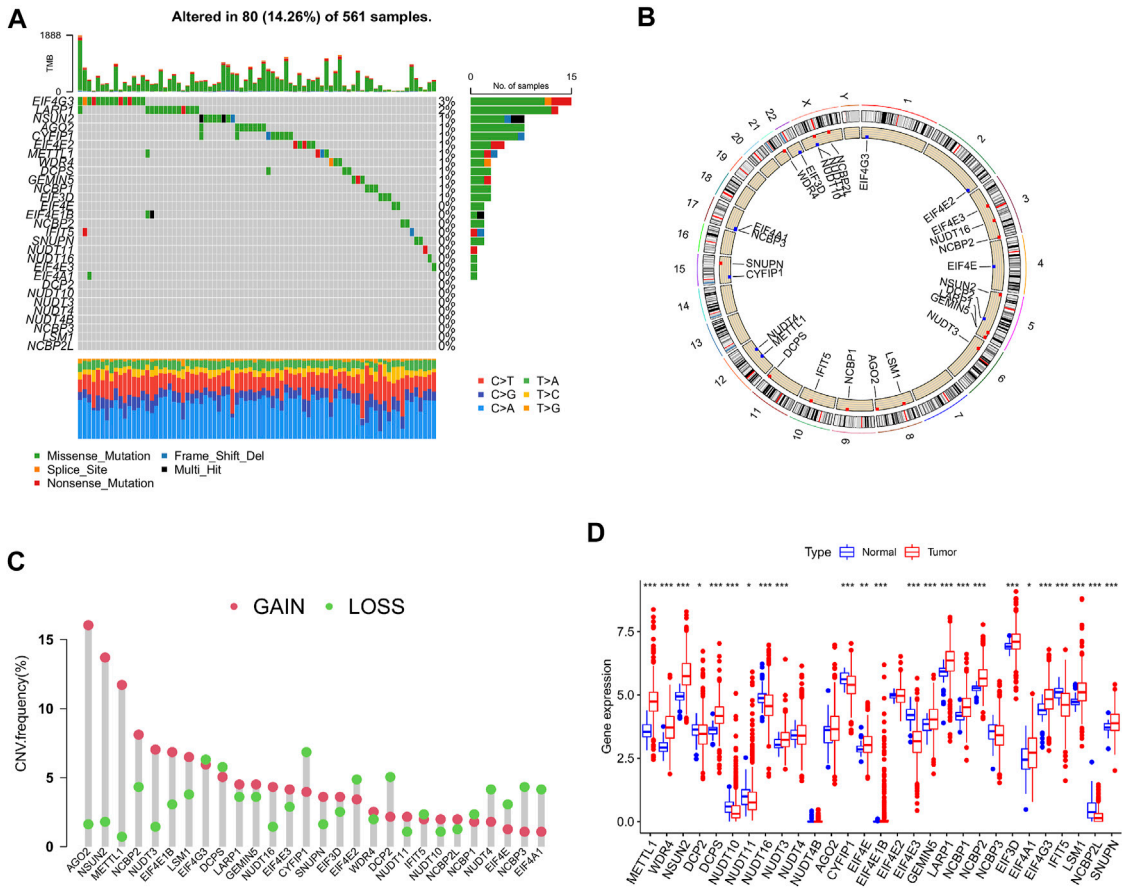


FIGURE 2
Genetic alterations of m7G-associated genes in lung adenocarcinoma (LUAD). **(A)** Mutation frequency of 29 m7G-associated genes in LUAD patients (*n* = 561). **(B)** Position of copy number variations (CNV) of the 29 m7G-associated genes on 23 chromosomes (red indicates gain of copy number; blue indicates loss of copy number). **(C)** CNV of the 29 m7G-associated genes (red indicates gain of copy number; green indicates loss of copy number). **(D)** Expression of the 29 m7G-associated genes in LUAD and normal tissue samples.

were added overnight, followed by DPBS washing thrice after rewarming. Secondary Antibody DyLight 488-labeled goat anti-rabbit (1:100, BA1127) was subsequently added for further incubation. After 45 min, the cells were stained with DAPI staining solution (AR1176) at room temperature for 3 min. DPBS washing was achieved four times. At last, an antifade mounting medium (AR1109) was used to seal the slides, and a fluorescence microscope was applied to observe the results and take photos.

Results

Genetic alterations of m7G-associated genes in lung adenocarcinoma

A total of 29 m7G-associated genes were included in this study and initially processed for somatic mutation analysis. The results revealed that 80 (14.26%) out of 561 LUAD patients experienced mutations, which were most frequent in the EIF4G3 gene, followed by LARP1, NSUN2, and AGO2. No mutations were found in the EIF4E and EIF4E1B genes. Missense mutation was the most common among all mutation types (Figure 2A). Copy number variations (CNV) were detected in almost all m7G-associated genes, except for EIF4E1B and NUDT4B. In the meantime, they were ubiquitous on the whole chromosomes apart from chromosomes 7, 13, 14, 18, 19, and 20. Increased copy number was demonstrated in 16 genes (such as AGO2, NSUN2, and METTL1), whereas decreased copy number was indicated in 11 genes (such as CYFIP1, EIF4G3, and DCPS) (Figures 2B, C). To clarify the association of somatic mutation with the expression of m7G-related genes, gene expression in LUAD and normal tissues was examined. It was found that partial genes exhibited a positive or negative association with CNV, whereas some genes showed no differential expression despite the increase in copy number (Figure 2D). In sum, the

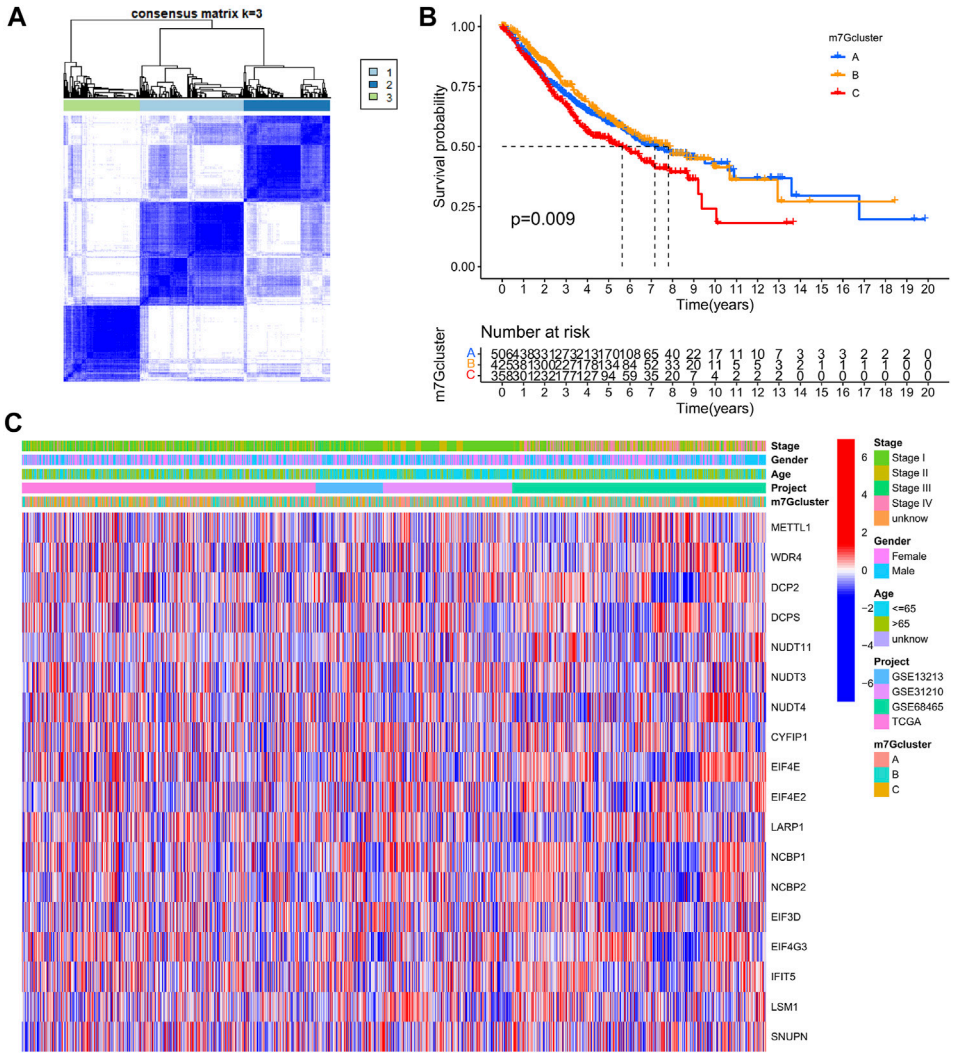


FIGURE 3
m7G-Associated LUAD subtypes and clinical prognosis. **(A)** Consensus clustering was used to divide the LUAD patients into 3 clusters (k = 3). **(B)** Kaplan–Meier OS curves of the 3 LUAD m7G-Clusters. **(C)** Clinical pathological features of the 3 LUAD m7G-Clusters.

results revealed that CNV might be one of the regulators of the expression of m7G-associated genes.

m7G-Associated lung adenocarcinoma subtypes and clinical prognosis

LUAD data derived from TCGA-LUAD, GSE13213, GSE31210, and GSE68465 were obtained and combined. Cox regression model and survival analysis were used to assess the prognostic value of the 29 m7G-associated genes for the OS of LUAD patients. The results demonstrated that 18 out of the 29 m7G-associated genes correlated to the OS of LUAD patients (Supplementary Table S3, Supplementary Figure S1). Based on the expression of the 18 genes, consensus clustering was performed to classify LUAD patients into 3 clusters following k = 3 (m7G-Clusters A, B, and C) (Figure 3A, Supplementary Figure S2). Survival analysis indicated that the OS of patients in m7G-Clusters A (p = 0.009) and B ($p < 0.05$) was better than that of patients in m7G-Cluster C (Figure 3B). The clinical pathological features of the three m7G-Clusters are shown in Figure 3C.

Biological and tumor microenvironment differences among the three lung adenocarcinoma subtypes

GSVA was performed to explore potential significant biological pathways in the three LUAD m7G-Clusters.

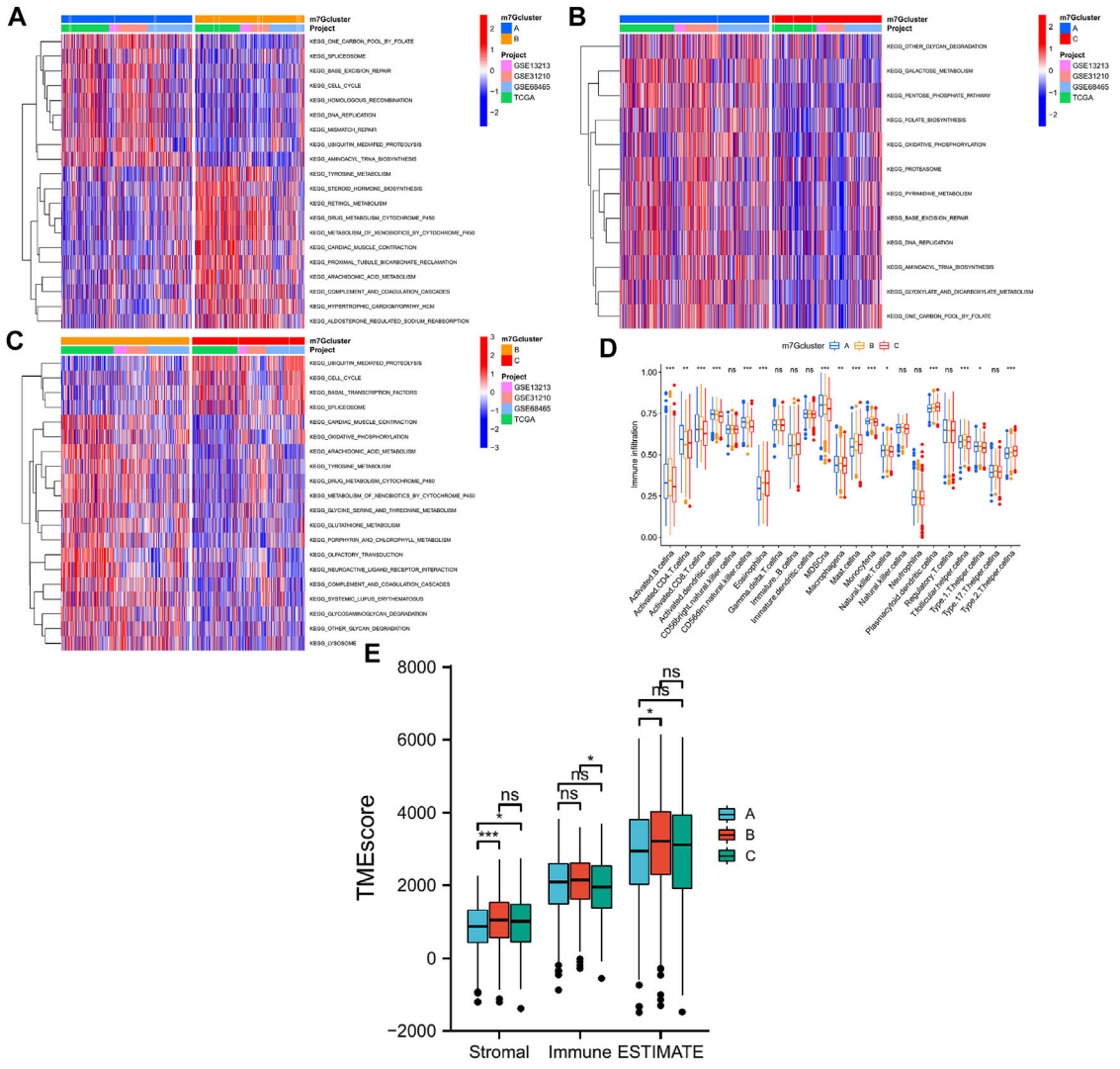


FIGURE 4
Biological processes and tumor immune microenvironment in 3 LUAD m7G-Clusters. **(A–C)** GSVA scores. **(D)** Infiltration of 23 immune cell subtypes. **(E)** Stromal, immune, and ESTIMATE scores.

Proliferation-related pathways exhibited significant enrichment scores in m7G-Clusters A and C, including pathways involved in cell cycle, DNA replication, ubiquitin-mediated proteolysis, and spliceosome. In parallel, the significantly enriched pathways in m7G-Cluster B were involved in metabolism (such as metabolism of xenobiotics by cytochrome P450, drug metabolism by cytochrome P450, arachidonic acid metabolism, oxidative phosphorylation, and sphingolipid metabolism), circulation (such as olfactory transduction and cardiac muscle contraction), and immunity (such as complement and coagulation cascades) (Figures 4A–C).

In addition, the TME difference in the three LUAD m7G-Clusters was analyzed by estimating the infiltration abundance of 23 immune cell subtypes using the CIBERSORT algorithm. The results demonstrated that there were 15 immune cell subtypes with a significant differential infiltration among the three m7G-Clusters (Figure 4D). In particular, Activated.B.cellna, Activated.CD8.T.cellna, Activated.dendritic.cellna, CD56dim.natural.killer.cellna, MDSCna, Macrophagena, Monocytena, Natural.killer.T.cellna, T.follicular.helper.cellna, and Type.1.T.helper.cellna were much more abundant in m7G-Clusters A and B, whereas Plasmacytoid.dendritic.cellna and Type.2.T.helper.cellna were highly abundant in m7G-Cluster C. However, m7G-Cluster B had a significantly lower infiltration level of Activated.CD4.T.cellna than m7G-Clusters A and C.

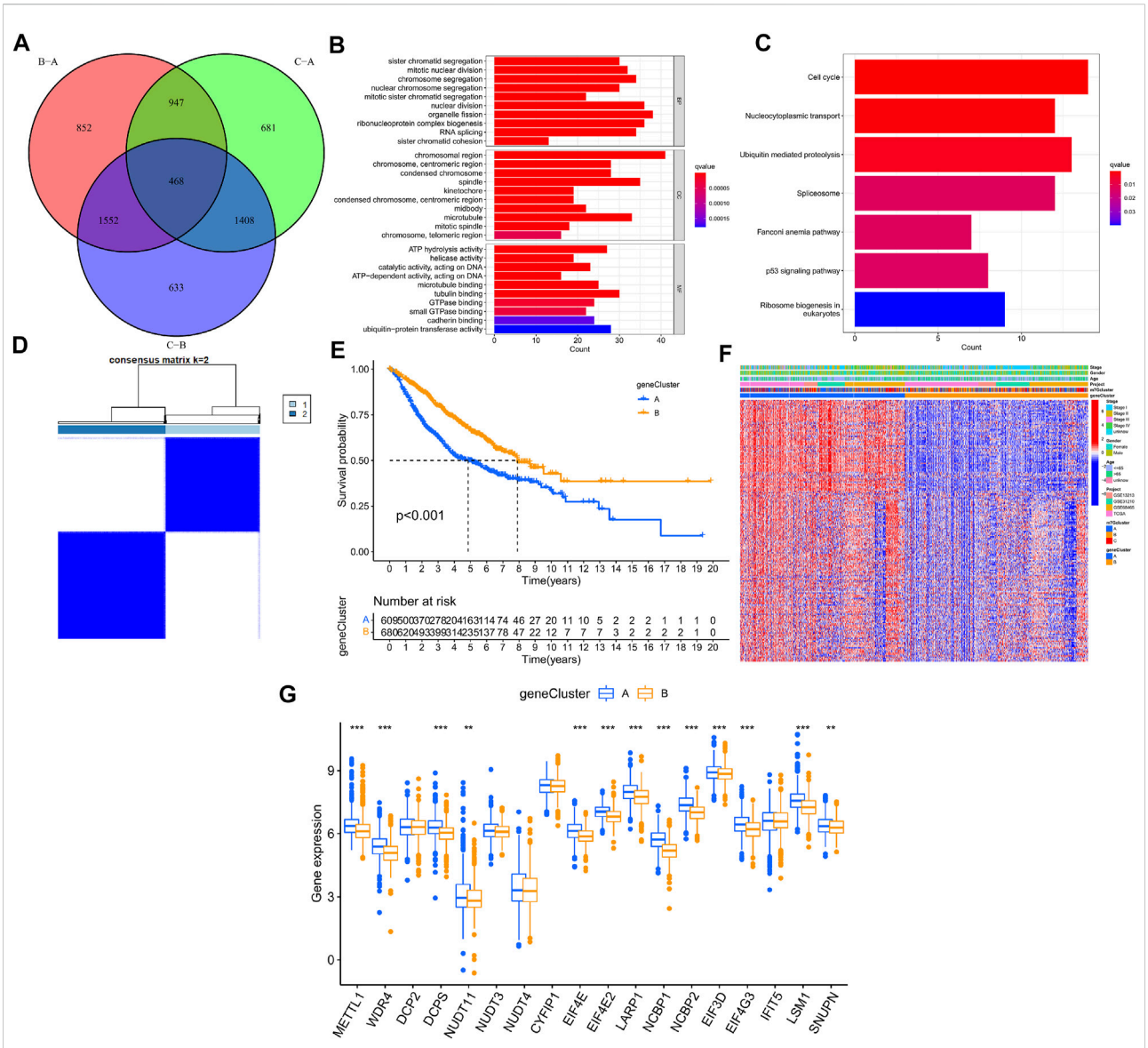


FIGURE 5
Genotyping based on m7G-associated DEGs in LUAD. **(A)** Venn diagram of DEGs from three LUAD m7G-Clusters. **(B–C)** GO and KEGG enrichment analysis for the 468 DEGs. **(D)** Consensus clustering was used to divide the LUAD patients into 2 gene clusters. **(E)** Kaplan–Meier OS curves of patients in the 2 gene clusters. **(F)** Clinical pathological features of the 2 gene clusters. **(G)** Expression of m7G-associated genes between the 2 gene clusters.

Moreover, the ESTIMATE algorithm was used to estimate the stromal, immune, and ESTIMATE scores of each cluster. The stromal score was significantly different in m7G-Cluster A vs. m7G-Cluster B and m7G-Cluster A vs. m7G-Cluster C and the immune score varied significantly between m7G-Cluster B and m7G-Cluster C, whereas the ESTIMATE score was distinct between m7G-Cluster A and m7G-Cluster B (all $p < 0.05$, Figure 4E). These results revealed that the TME of m7G-Cluster B had the highest score among the three clusters.

Identification of m7G-associated differentially expressed genes and genotyping

The R package "limma" was used to perform differential analysis, which resulted in the identification of 468 DEGs from the 3 LUAD m7G-Clusters (p.adj < 0.001) as shown in the Venn diagram (Figure 5A, Supplementary Table S2). The DEGs obtained were processed for functional enrichment analysis. The most enriched GO entries were biological processes

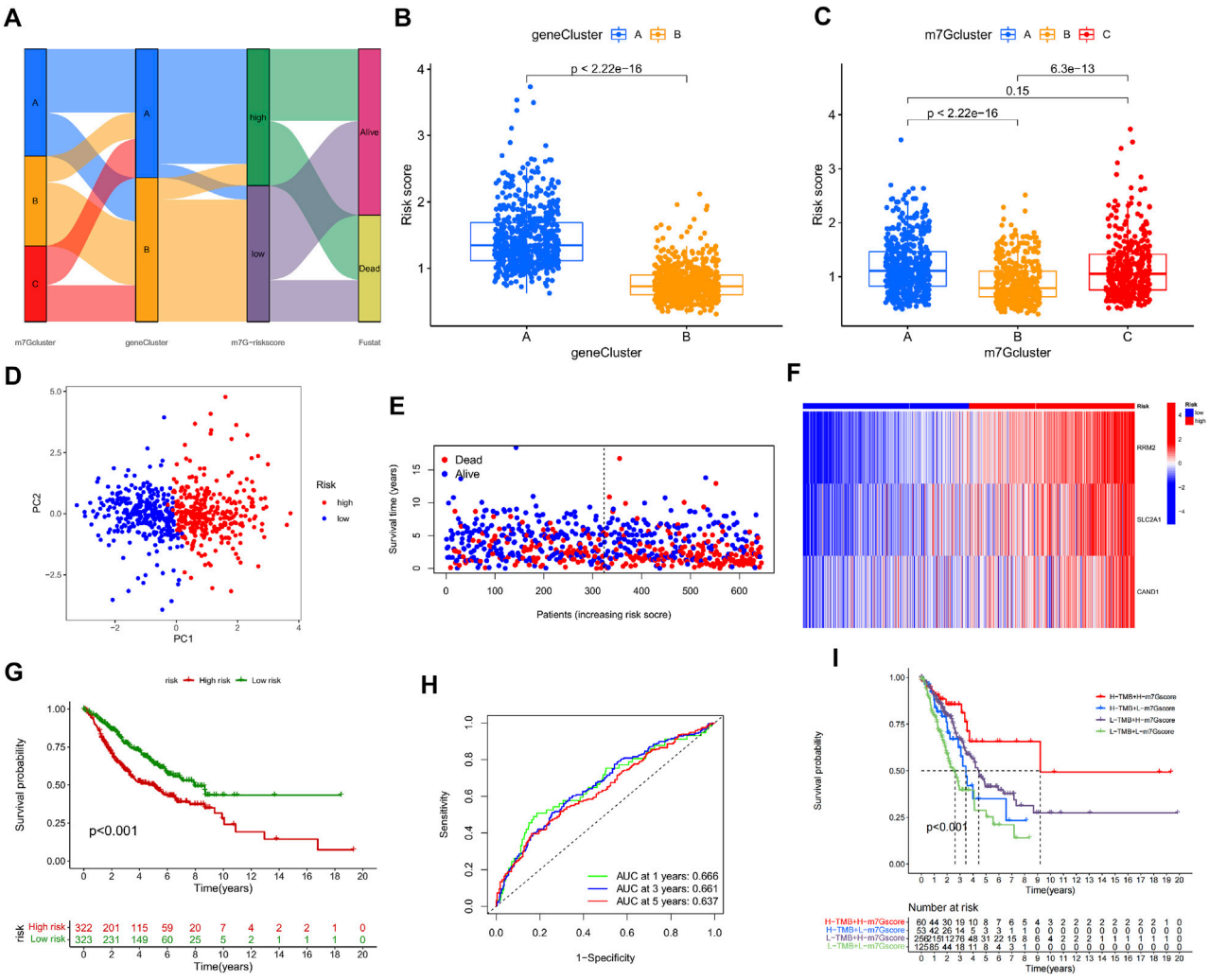


FIGURE 6
Construction and validation of m7G-Riskscore. **(A)** Sankey diagram of the clinical outcomes of patients in three m7G-Clusters, two gene clusters, and high- and low-risk groups. **(B)** m7G-Riskscore in two gene clusters. **(C)** m7G-Riskscore in 3 m7G-Clusters. **(D)** PCA. **(E)** Distribution of m7G-Riskscore. **(F)** Heatmap of the expression of the 3 signature genes (RRM2, SLC2A1, and CAND1). **(G)** Kaplan–Meier OS curves. **(H)** ROC curves for 1-, 3-, and 5-year OS. **(I)** Kaplan–Meier OS curves for LUAD patients stratified using m7G-Riskscore plus TMB.

involved in cell proliferation (Figure 5B), whereas the significantly enriched KEGG pathways were related to cancer cell proliferation, invasion, and death (Figure 5C). The results indicated that m7G played a vital part in tumor cell proliferation, invasion, and death. Thereafter, the 468 DEGs were subjected to univariate analysis, and 196 of them were found to be associated with the OS of LUAD ($p < 0.05$, Supplementary Table S4). Based on the 196 DEGs, consensus clustering was performed to divide the LUAD patients into two gene clusters (A, B) following $k = 2$ (Figure 5D). Further survival analysis demonstrated that patients in gene cluster B had better OS than patients in gene cluster A ($p < 0.001$, Figure 5E). In addition, gene cluster A was related to higher disease stages (III/IV) (Figure 5F). A significant difference in the expression of m7G-associated genes, such as METTL1, WDR4, and DCP2, was demonstrated between the two gene clusters (Figure 5G).

Construction and validation of m7G-Riskscore

On the basis of the three m7G-Clusters, an m7G-Riskscore was established. The information of patients stratified by m7G-associated genes, DEGs, and m7G-Riskscore is presented in Figure 6. LUAD patients were randomly categorized into training ($n = 645$) and test ($n = 644$) sets. LASSO and

multivariate Cox regression models were used to analyze the 196 DEGs identified in the univariate analysis (Supplementary Figure S3, Supplementary Table S4), and 3 of them, namely, RRM2, SLC2A1, and CAND1, were selected to construct an m7G-associated prognostic signature. The m7G-Riskscore was calculated as follows: m7G − Riskscore = (0.2141 × RRM2exp)+ (0.1377 × SLC2A1exp) + (0.2240 × CAND1exp).

Significantly higher scores were noted in gene cluster A than in gene cluster B. In addition, the scores in m7G-Clusters A and C were much higher than those in m7G-Cluster B. The findings collectively implied that the patients in m7G-Cluster B and gene cluster B had higher immune activities than those in the other groups (Figures 6B, C). Moreover, two groups of patients were defined based on the median m7G-Riskscore and validated to be well differentiated through PCA (Figure 6D). As shown in the distribution plot of the score in two groups, LUAD patients tended to survive shorter and die with an increase in the m7G-Riskscore (Figure 6E). The expression of the three signature genes in the two groups was visualized using a heatmap (Figure 6F). The patients in the high-risk group suffered from a lower survival rate than the patients in the low-risk group ($p < 0.001$, Figure 6G). The ROC curves of the m7G-Riskscore for 1-, 3-, and 5-year survival were generated, and the AUC scores were 0.666, 0.661, and 0.637, respectively, demonstrating that the m7G-Riskscore was capable of predicting the survival of LUAD patients (Figure 6H). Furthermore, the m7G-Riskscore and TMB were combined to study their effect on the survival of LUAD patients. The results showed that high TMB + high m7G-Riskscore was associated with the highest survival rate, suggesting that high TMB contributed to more immune antigens, making patients benefit more from immunotherapy (Figure 6I).

At last, the performance of the m7G-Riskscore was validated in the total, test, and GSE30219 sets. Using the same grouping strategy, LUAD patients in both sets were respectively divided into the high- and low-risk groups. The grouping quality was evaluated *via* PCA. The distributions of the m7G-Riskscore in both sets revealed that patients with a higher m7G-Riskscore had a shorter survival time. In addition, the survival rate of high-risk patients was much lower than that of low-risk patients in both the total, test, and GSE30219 sets (Supplementary Figures S4A–E, 5A–E, 6A–E). Furthermore, the ROC–AUC scores for 1-, 3-, and 5-year survival were 0.660, 0.660, and 0.629 in the total set; 0.651, 0.658, and 0.618 in the test set; and 0.974, 0.794, and 0.793 in the GSE30219 set, respectively (Supplementary Figures S4F, 5F, 6F). Taken together, the m7G-Riskscore is a robust tool with good prognostic performance for determining the OS of LUAD patients (Supplementary Figures S7, S8).

Protein expression of the 3 m7G-associated signature genes

Western blot was adopted to examine the protein levels of RRM2, SLC2A1, and CAND1. As shown in Figure 7, a distinctly increased expression was found in LUAD cell line A549 than in human normal lung cell line MRC-5. In addition, immunofluorescence staining was performed and revealed the same expression trend of the three genes (Figures 8A–C).

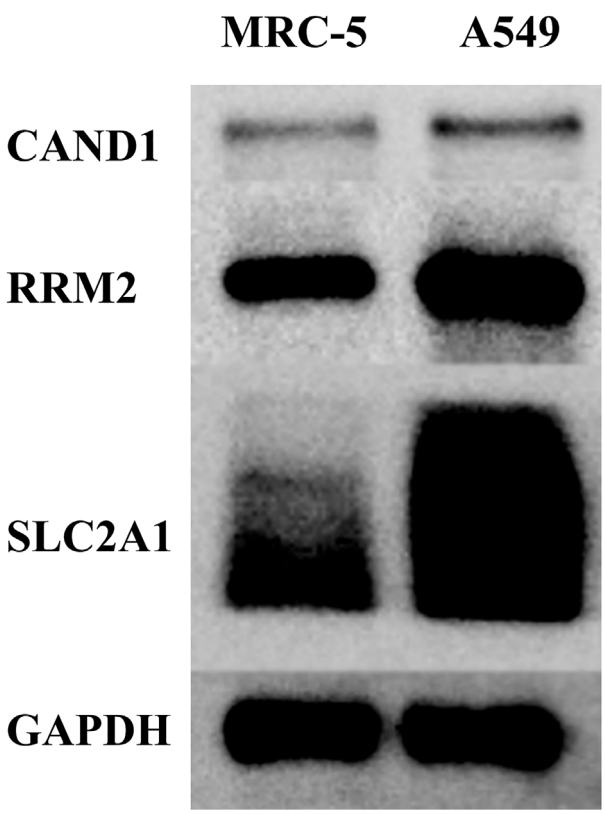


FIGURE 7
Western blot.

Tumor microenvironment of the high- and low-risk groups

The CIBERSORT algorithm was used to calculate the abundance of immune infiltration. It was found that the m7G-Riskscore was associated with the majority of immune infiltrates, including macrophages M0, macrophages M1, macrophages M2, mast cells activated, neutrophils, NK cells resting, CD4 memory T cells activated, and T cells CD8 (Figure 9A). In particular, negative associations were observed with the infiltration abundance of B cell memory, dendritic cells resting, mast cells resting, monocytes, NK cells activated, plasma cells, and CD4 memory T cell resting (Figure 9A). In addition, the m7G-Riskscore was noted to be negatively correlated with the immune and stromal scores (Figure 9B). Furthermore, correlation analysis revealed that expression of the three signature genes (CAND1, RRM2, and SLC2A1) was

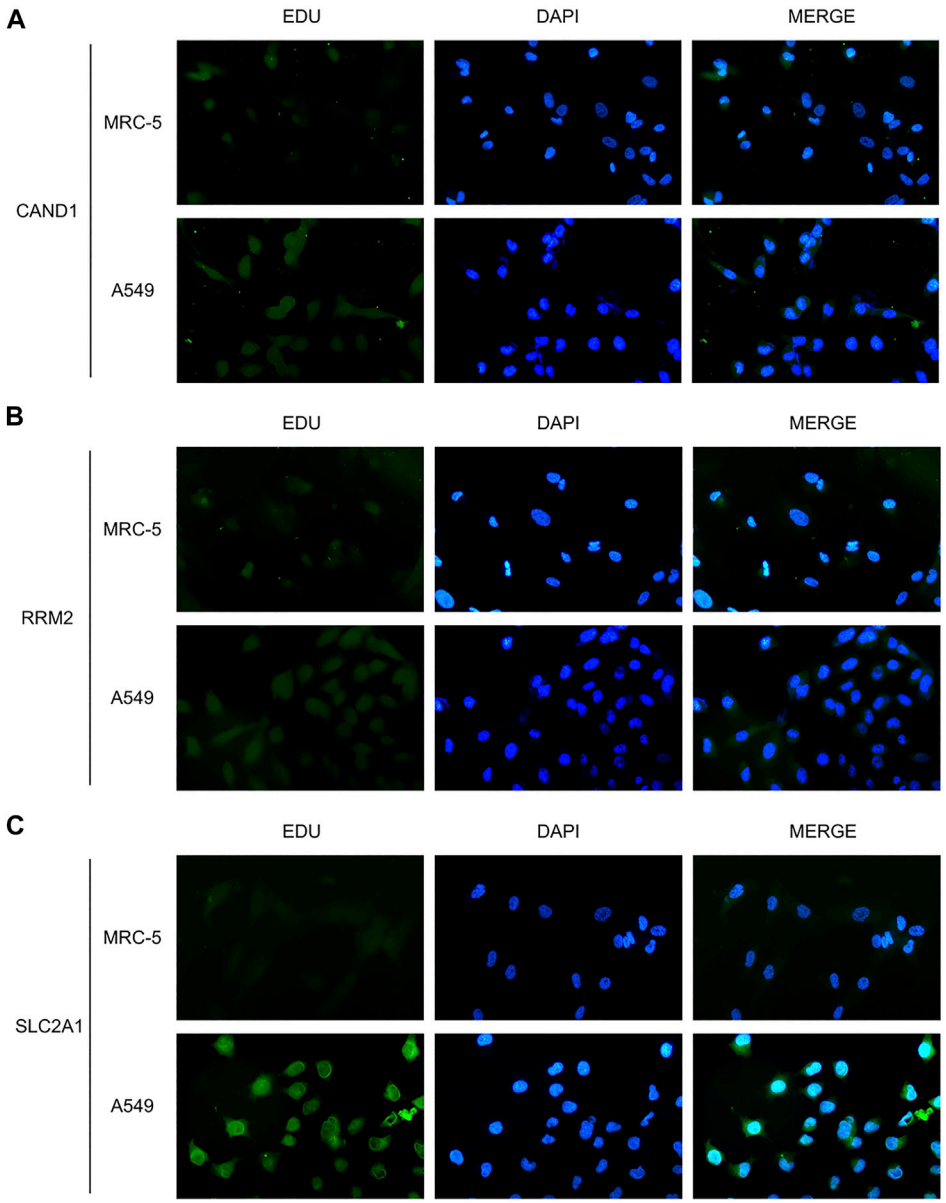


FIGURE 8
Immunofluorescence staining. **(A)** CAND1. **(B)** RRM2. **(C)** SLC2A1.

significantly associated with most of the immune infiltrates, especially CD4 memory T cells activated and neutrophils ($p < 0.001$, Figure 9C).

Cancer stemness, mutation, and drug sensitivity in high- and low-risk groups

A comprehensive analysis was devised to study the association of the m7G-Riskscore with cancer stemness, and a linear positive correlation was revealed ($R = 0.42$, $p < 2.2e\text{-}16$; Figure 10A). Research reported that TMB is conducive to the prediction of patient response to tumor immunotherapy. Here, we investigated the correlation between the m7G-Riskscore and TMB and found that the TMB increased with the increase in the m7G-Riskscore (Figure 10B). This result demonstrated that patients with a high m7G-Riskscore might benefit from immunotherapy. Furthermore, Spearman's correlation analysis revealed a positive association ($R = 0.29$, $p = 2.8e\text{-}11$; Figure 10C). Somatic mutation was then analyzed in the TCGA-LUAD dataset. As analyzed, TP53, TTN, MUC16, RYR2, and CSMD3 exhibited the highest frequency of mutation (>25%)

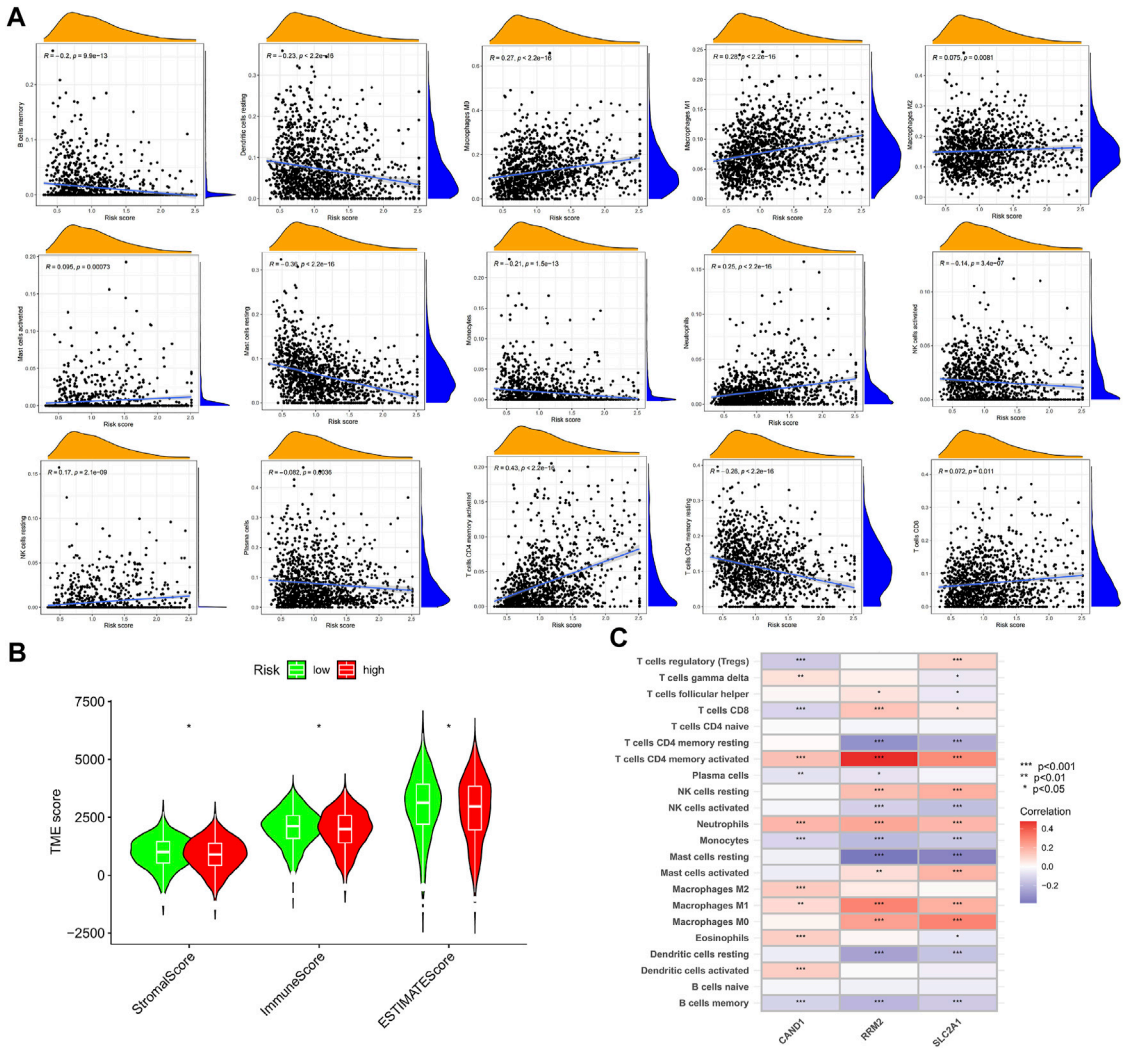


FIGURE 9
TME of the high- and low-risk groups. **(A)** Association of the m7G-Riskscore with immune infiltrates. **(B)** Association of the m7G-Riskscore with immune, stromal, and ESTIMATE scores. **(C)** Association of the three signature genes (CAND1, RRM2, and SLC2A1) with immune infiltrates.

(Figures 10D, E). When comparing the high- and low-risk groups, the rate of somatic mutation was much higher in the high-risk group.

To further explore the ability of the m7G-Riskscore to predict patient response to immunotherapy, the half-maximal inhibitory concentration (IC50) of multiple agents was calculated. Active responses to chemotherapeutics (such as cisplatin, docetaxel, etoposide, and gemcitabine) and multiple targeted drugs (such as afatinib, gefitinib, and erlotinib) were found in patients in the low-risk group, whereas active responses to lapatinib and metformin were observed in patients in the high-risk group (Figures 11A–O). Taken together, it was revealed that m7G-Riskscore is associated with drug sensitivity in LUAD patients.

Establishment and validation of nomogram for clinical prediction

A nomogram was established based on the m7G-Riskscore and clinical pathological features (including age, gender and stage) of LUAD to predict the 1-, 3- and 5-year OS of patients (Figure 12A). Calibration curve was made to show a high consistency between the predicted result and real result in the total, training, test and GSE30219 sets, demonstrating the favorable prognostic performance of the model (Figures 12B–E). In addition, ROC–AUC scores for 1-, 3- and 5-year OS were estimated and reached the optimal as 0.773, 0.772, 0.778, and 0.795, respectively (Figures 12F–I). The results showed that the nomogram was powerful in prediction of the survival outcome of LUAD patients.

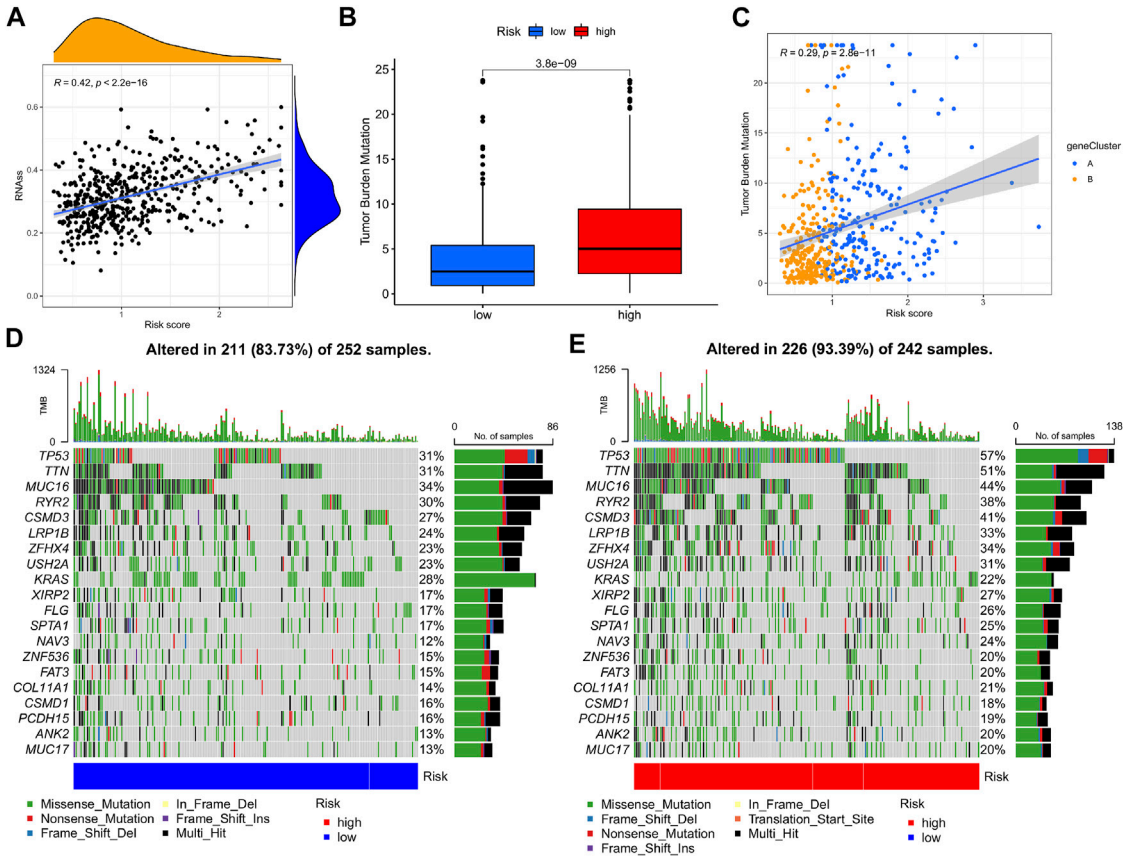


FIGURE 10
Cancer stemness and mutation analysis. **(A)** Association of the m7G-Riskscore with cancer stemness. **(B)** Association of the m7G-Riskscore with TMB. **(C)** Spearman's correlation between the m7G-Riskscore and TMB. **(D–E)** Waterfall plots of somatic mutation.

Discussion

Owing to the increasing studies on RNA methylation, m7G methylation has shown potential in the regulation of tumor. However, majority of the studies on m7G methylation in tumor regulation focus on single or two associated genes (Chen et al., 2021; Dai et al., 2021; Orellana et al., 2021). At present, growing evidence has proven that TME is important for tumor progression and that immune infiltrates are promising prognostic factors (Pitt et al., 2016). The multimediating effects of the combination of multiple m7G-associated genes and their role in TME need to be further studied. Research reported that m7G modification was oncogenic in lung cancer (Ma et al., 2021). In particular, it is necessary to clarify the specific effect of m7G modification on lung cancer and on tumor immune infiltration.

In the present study, we initially revealed the transcriptional expression pattern of m7G-associated genes in the TCGA-LUAD dataset. Then, prognostic m7G-associated genes were screened and used to divide LUAD patients into three m7G-Clusters *via* consensus clustering. It was found that patients in m7G-Cluster C had the worst OS than those in m7G-Clusters A and B. In addition, significant differences among the three m7G-Clusters were observed with respect to biological processes, including cell proliferation, metabolism, and immunity. The TME also significantly varied in the three m7G-Clusters. Afterward, DEGs from the three m7G-Clusters were further obtained. Functional enrichment analysis revealed that the DEGs were mainly involved in pathways associated with cancer cell proliferation, invasion, and death. On the basis of the DEGs, two gene clusters were defined, whereas gene cluster B was indicative of better OS than gene cluster A. In addition, the DEGs were processed for regression analysis to obtain prognostic genes. Three genes were identified, namely, RRM2, SLC2A1, and CAND1, and used to establish an m7G-Riskscore. Each patient was conferred an m7G-Riskscore and had poorer OS when

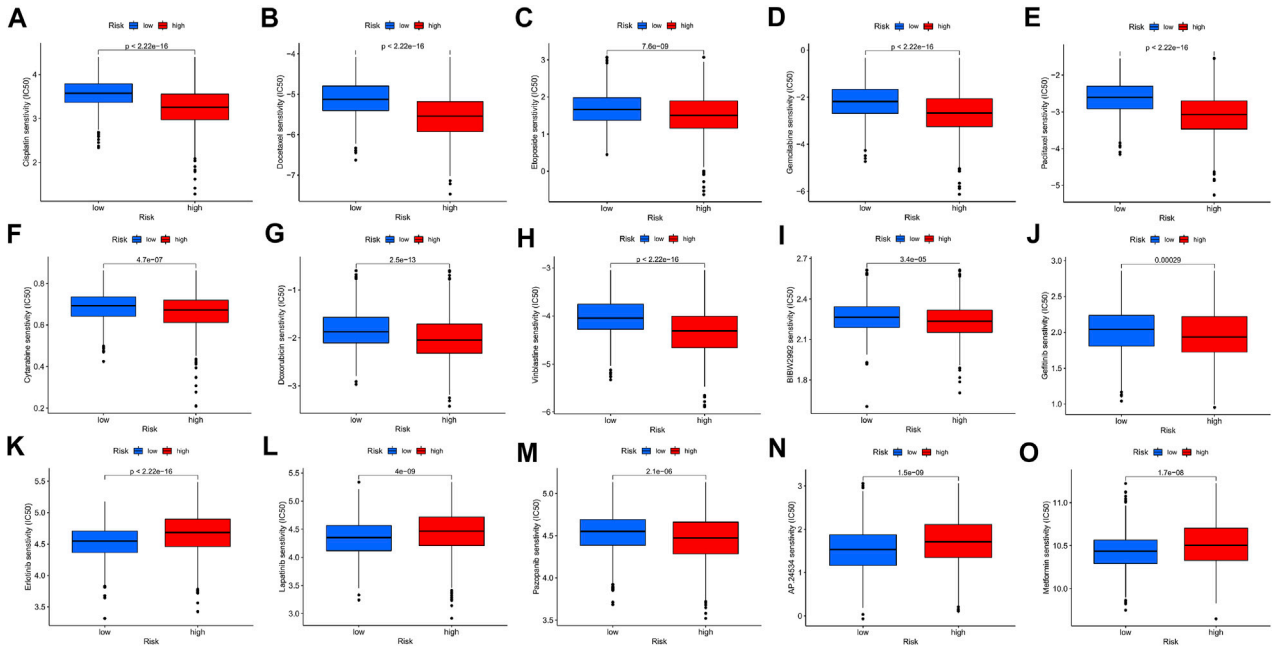


FIGURE 11
Association of the m7G-Riskscore with drug sensitivity.

classified as gene cluster A and scored higher. This finding demonstrated that a high m7G-Riskscore was prognostic for the poor OS of LUAD patients. It was reported that m7G-associated genes are involved in the proliferation, metastasis, and drug resistance of multiple tumors (Dai et al., 2021; Xia et al., 2021), including LUAD (Ma et al., 2021). The present study found that m7G-associated genes were significantly enriched in pathways involved in cancer cell proliferation and invasion. The protein expression of the three signature genes was examined *via ex vivo* cell experiments using western blot and immunofluorescence staining assays. Multivariate analysis revealed that the m7G-Riskscore was significantly associated with the OS of LUAD patients and could be used as an independent prognostic factor. Furthermore, the ROC–AUC scores demonstrated the potential of the m7G-Riskscore in the prognosis of LUAD.

This study also found that there were significant differences in the OS, CNV, TME, cancer stemness, TMB, and drug sensitivity among high- and low-risk patients stratified using the m7G-Riskscore. Previous study showed that a higher TMB was suggestive of a better prognosis (Samstein et al., 2019), which is consistent with the present study. TME is composed of immune and stromal cells, and their scores are highly associated with the clinical features and prognosis of tumor (Quail and Joyce, 2013; Belli et al., 2018). At present, LUAD patients still have an unsatisfied prognosis and tend to develop drug resistance after routine chemotherapy (Johnson and Patel, 2014). However, the prognosis of each patient is still unpredicted despite the great advancement of targeted and immune therapies, which might be due to the varying TMEs. In the present study, we found that the m7G-Riskscore was associated with the majority of immune infiltrates, and a lower score predicted higher immune and stromal scores. The results demonstrated that the TME might play an important role in the regulation of tumor cell proliferation, invasion, and progression by m7G. In addition, a higher m7G-Riskscore suggested a higher TMB, showing that patients who scored higher may benefit more from immunotherapy. Furthermore, drug sensitivity was assessed in the high- and low-risk groups. It was found that low-risk patients responded well to chemotherapeutics and multiple targeted drugs, whereas high-risk patients better responded to lapatinib and metformin. These findings may help formulate clinical strategy in drug use and thus reduce drug resistance. At last, we established a nomogram for clinical use by combining the m7G-Riskscore and several clinical pathological features. Taken together, the m7G-associated prognostic signature we constructed could be used to stratify the prognosis of LUAD patients, help understand the molecular mechanism underlying the

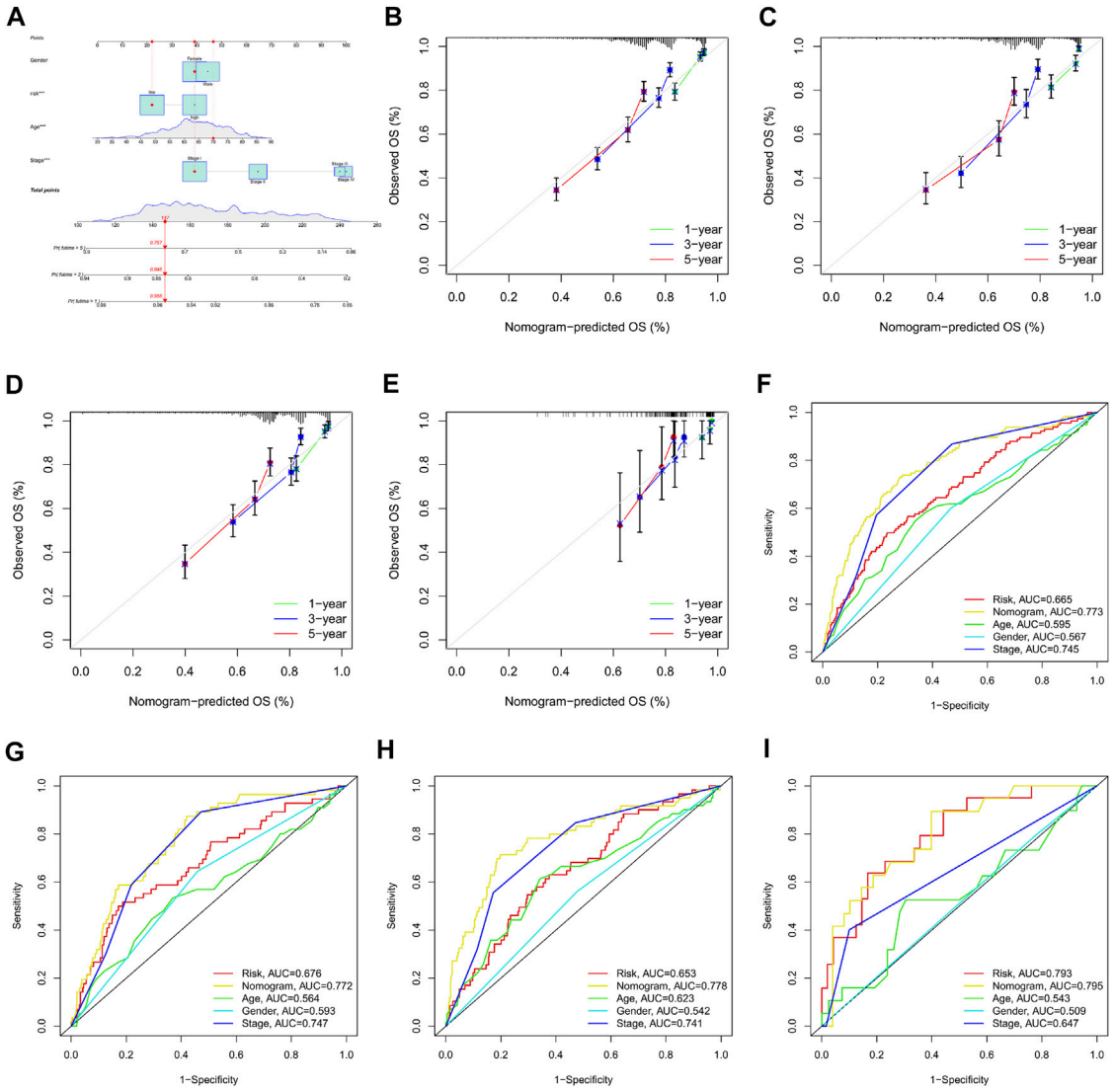


FIGURE 12
Establishment and validation of nomogram for clinical prediction. **(A)** Nomogram establishment. **(B–E)** Calibration curves for 1-, 3-, and 5-year OS in the total, training, test, and GSE30219 sets. **(F–I)** ROC curves for 1-, 3-, and 5-year OS in the total, training, test, and GSE30219 sets.

initiation and progression of LUAD, and provide new thoughts into cancer treatment.

The study still has some limitations. For example, the data included in this study were derived from public databases, resulting in certain selection bias and thus affecting the results. In addition, further *in vivo* experiments are required to validate the findings of the study, although we have preliminarily proven the good prognostic performance of the m7G-Riskscore and examined the *in vitro* expression of the three signature genes. Moreover, there are missing data for clinical variables, such as surgery and targeted therapy, which requires improvement and introduction of more related clinical variables to further validate the clinical value of m7G-Riskscore.

Conclusion

To sum up, the present study systemically analyzed m7G-associated genes and revealed the related regulatory mechanism involved in the TME, pathological features, and prognosis of LUAD patients. The findings of this study demonstrated the clinical value of m7G-associated genes

and provided a new thought for the clinical individualized treatment of LUAD.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://portal.gdc.cancer.gov/ and https://www.ncbi.nlm.nih.gov/geo/.

Author contributions

GW, MZ, and XH performed the study concept and design. XH and GX revised the manuscript and made the final approval of the version. GW, MZ, and JL analyzed the data and wrote the manuscript. GL and FZ helped write the manuscript. All authors contributed to the article and approved of the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (82160889), Guangxi Medical, and health key cultivation discipline construction project, The Basic Ability Enhancement Program for Young and Middle-Aged Teachers of Guangxi (2021KY0310) and Innovation Project of Guangxi Graduate Education (YCBZ2021077).

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.2022.954840/full#supplementary-material

References

Bai, H., Wang, Y., Liu, H., and Lu, J. (2021). Development of a four-mRNA expression-based prognostic signature for cutaneous melanoma. *Front. Genet.* 12, 680617. doi:10.3389/fgene.2021.680617

Belli, C., Trapani, D., Viale, G., D'Amico, P., Duso, B. A., Della Vigna, P., et al. (2018). Targeting the microenvironment in solid tumors. *Cancer Treat. Rev.* 65, 22–32. doi:10.1016/j.ctrv.2018.02.004

Bi, Y., Xiang, D., Ge, Z., Li, F., Jia, C., and Song, J. (2020). An interpretable prediction model for identifying N^7-methylguanosine sites based on XGBoost and SHAP. *Mol. Ther. Nucleic Acids* 22, 362–372. doi:10.1016/j.omtn.2020.08.022

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

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

Dai, Z., Liu, H., Liao, J., Huang, C., Ren, X., Zhu, W., et al. (2021). N^7-Methylguanosine tRNA modification enhances oncogenic mRNA translation and promotes intrahepatic cholangiocarcinoma progression. *Mol. Cell* 81 (16), 3339–3355. doi:10.1016/j.molcel.2021.07.003

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

Johnson, M. L., and Patel, J. D. (2014). Chemotherapy and targeted therapeutics as maintenance of response in advanced non-small cell lung cancer. *Semin. Oncol.* 41 (1), 93–100. doi:10.1053/j.seminoncol.2013.12.007

Li, R., Di, L., Li, J., Fan, W., Liu, Y., Guo, W., et al. (2021). A body map of somatic mutagenesis in morphologically normal human tissues. *Nature* 597 (7876), 398–403. doi:10.1038/s41586-021-03836-1

Li, R., Yin, Y. H., Ji, X. L., Liu, X., Li, J. P., and Qu, Y. Q. (2021). Pan-cancer prognostic, immunity, stemness, and anticancer drug sensitivity characterization of N6-methyladenosine RNA modification regulators in human cancers. *Front. Mol. Biosci.* 8, 644620. doi:10.3389/fmolb.2021.644620

Ma, J., Han, H., Huang, Y., Yang, C., Zheng, S., Cai, T., et al. (2021). METTL1/WDR4-mediated m^7G tRNA modifications and m^7G codon usage promote mRNA translation and lung cancer progression. *Mol. Ther.* 29 (12), 3422–3435. doi:10.1016/j.ymthe.2021.08.005

Martincorena, I., and Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science* 349 (6255), 1483–1489. doi:10.1126/science.aab4082

Milholland, B., Auton, A., Suh, Y., and Vijg, J. (2015). Age-related somatic mutations in the cancer genome. *Oncotarget* 6 (28), 24627–24635. doi:10.18632/oncotarget.5685

Orellana, E. A., Liu, Q., Yankova, E., Pirouz, M., De Braekeleer, E., Zhang, W., et al. (2021). METTL1-mediated m^7G modification of Arg-TCT tRNA drives oncogenic transformation. *Mol. Cell* 81 (16), 3323–3338. doi:10.1016/j.molcel.2021.06.031

Pitt, J. M., Marabelle, A., Eggermont, A., and Zitvogel, L. (2016). Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. *Ann. Oncol.* 27 (8), 1482–1492. doi:10.1093/annonc/mdw168

Qi-Dong, X., Yang, X., Lu, J. L., Liu, C. Q., Sun, J. X., Li, C., et al. (2020). Development and validation of a nine-redox-related long noncoding RNA signature in renal clear cell carcinoma. *Oxid. Med. Cell. Longev.* 2020, 6634247. doi:10.1155/2020/6634247

Quail, D. F., and Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* 19 (11), 1423–1437. doi:10.1038/nm.3394

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 (2), 202–206. doi:10.1038/s41588-018-0312-8

Shiba-Ishii, A. (2021). Significance of stratifin in early progression of lung adenocarcinoma and its potential therapeutic relevance. *Pathol. Int.* 71 (10), 655–665. doi:10.1111/pin.13147

Succony, L., Rassl, D. M., Barker, A. P., McCaughan, F. M., and Rintoul, R. C. (2021). Adenocarcinoma spectrum lesions of the lung: detection, pathology and treatment strategies. *Cancer Treat. Rev.* 99, 102237. doi:10.1016/j.ctrv.2021.102237

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

Tomikawa, C. (2018). 7-Methylguanosine modifications in transfer RNA (tRNA). *Int. J. Mol. Sci.* 19 (12), 4080. doi:10.3390/ijms19124080

Wang, X., Pan, L., Lu, Q., Huang, H., Feng, C., Tao, Y., et al. (2021). A combination of ssGSEA and mass cytometry identifies immune microenvironment in muscle-invasive bladder cancer. *J. Clin. Lab. Anal.* 35 (5), e23754. doi:10.1002/jcla.23754

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

Wu, J., Li, L., Zhang, H., Zhao, Y., Wu, S., Xu, B., et al. (2021). A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma. *Oncogene* 40 (26), 4413–4424. doi:10.1038/s41388-021-01853-y

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

Xia, Q. D., Sun, J. X., Xun, Y., Xiao, J., Liu, C. Q., Xu, J. Z., et al. (2022). SUMOylation pattern predicts prognosis and indicates tumor microenvironment infiltration characterization in bladder cancer. *Front. Immunol.* 13, 864156. doi:10.3389/fimmu.2022.864156

Xia, Q. D., Sun, J. X., Liu, C. Q., Xu, J. Z., An, Y., Xu, M. Y., et al. (2022). Ferroptosis patterns and tumor microenvironment infiltration characterization in bladder cancer. *Front. Cell Dev. Biol.* 10, 832892. doi:10.3389/fcell.2022.832892

Xie, S., Chen, W., Chen, K., Chang, Y., Yang, F., Lin, A., et al. (2020). Emerging roles of RNA methylation in gastrointestinal cancers. *Cancer Cell Int.* 20 (1), 585. doi:10.1186/s12935-020-01679-w

Yu, Z., Chen, H., You, J., Liu, J., Wong, H. S., Han, G., et al. (2015). Adaptive fuzzy consensus clustering framework for clustering analysis of cancer data. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 12 (4), 887–901. doi:10.1109/TCBB.2014.2359433

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

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 11 August 2022
DOI 10.3389/fgene.2022.911443

Check for updates

OPEN ACCESS

EDITED BY
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

REVIEWED BY
Joaquim Carreras,
Tokai University, Japan
Yang Cao,
Huazhong University of Science and Technology, China
Pradeep Reddy Cingaram,
University of Miami, United States

*CORRESPONDENCE
Wen-Da Luo,
luowd@enzemd.com
Min Zhu,
zhum@enzemed.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 02 April 2022
ACCEPTED 18 July 2022
PUBLISHED 11 August 2022

CITATION
Pan Y-B, Wang W, Cai H-K, Zhang J, Teng Y, Xue J, Zhu M and Luo W-D (2022), Integrative analysis of a necroptosis-related gene signature of clinical value and heterogeneity in diffuse large B cell lymphoma. *Front. Genet.* 13:911443. doi: 10.3389/fgene.2022.911443

COPYRIGHT
© 2022 Pan, Wang, Cai, Zhang, Teng, Xue, Zhu and Luo. 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 a necroptosis-related gene signature of clinical value and heterogeneity in diffuse large B cell lymphoma

Yu-Biao Pan[1,2†], Wei Wang[2,3†], Hong-Kai Cai[2,3], Jia Zhang[2,3], Ya Teng[3], Jiji Xue[3], Min Zhu[3*] and Wen-Da Luo[1,2,3*]

[1]Taizhou Hospital of Zhejiang Province, Zhejiang University, Hangzhoua, China, [2]Department of Hematology, Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, China, [3]Taizhou Hospital of Zhejiang Province Affiliated to Wenzhou Medical University, Linhai, China

Background: Diffuse large B-cell lymphoma (DLBCL), which is considered to be the most common subtype of lymphoma, is an aggressive tumor. Necroptosis, a novel type of programmed cell death, plays a bidirectional role in tumors and participates in the tumor microenvironment to influence tumor development. Targeting necroptosis is an intriguing direction, whereas its role in DLBCL needs to be further discussed.

Methods: We obtained 17 DLBCL-associated necroptosis-related genes by univariate cox regression screening. We clustered in GSE31312 depending on their expressions of these 17 genes and analyzed the differences in clinical characteristics between different clusters. To investigate the differences in prognosis across distinct clusters, the Kaplan-Meier method was utilized. The variations in the tumor immune microenvironment (TME) between distinct necroptosis-related clusters were investigated via "ESTIMATE", "Cibersort" and single-sample geneset enrichment analysis (ssGSEA). Finally, we constructed a 6-gene prognostic model by lasso-cox regression and subsequently integrated clinical features to construct a prognostic nomogram.

Results: Our analysis indicated stable but distinct mechanism of action of necroptosis in DLBCL. Based on necroptosis-related genes and cluster-associated genes, we identified three groups of patients with significant differences in prognosis, TME, and chemotherapy drug sensitivity. Analysis of immune infiltration in the TME showed that cluster 1, which displayed the best prognosis, was significantly infiltrated by natural killer T cells, dendritic cells, $CD8^+$ T cells, and M1 macrophages. Cluster 3 presented M2 macrophage infiltration and the worst prognosis. Importantly, the prognostic model successfully differentiated high-risk from low-risk patients, and could forecast the survival of DLBCL patients. And the constructed nomogram demonstrated a remarkable capacity to forecast the survival time of DLBCL patients after incorporating predictive clinical characteristics.

Conclusion: The different patterns of necroptosis explain its role in regulating the immune microenvironment of DLBCL and the response to R-CHOP treatment. Systematic assessment of necroptosis patterns in patients with DLBCL will help us understand the characteristics of tumor microenvironment cell infiltration and aid in the development of tailored therapy regimens.

KEYWORDS

diffuse large B-cell lymphoma, necroptosis, prognosis, TME, immunization

Introduction

DLBCL is an aggressive form of lymphoma that accounts for roughly 33 percent of all non-Hodgkin lymphomas (NHL) diagnosed each year (Liu and Barta, 2019; Siegel et al., 2019). Approximately two-thirds of patients with DLBCL are alleviated by the standard R-CHOP regimen that includes rituximab, cyclophosphamide, adriamycin, vincristine, and prednisolone. However, approximately one-third of patients present unsatisfactory results with this treatment; hence, long-term remission is achieved only in a small number of patients, implying that novel molecular targets for DLBCL therapy are urgently needed, implying that novel molecular indicators for the therapy of DLBCL patients are desperately needed (Coiffier et al., 2010; Sehn and Gascoyne, 2015).

Necroptosis is a new kind of programmed cell death that combines apoptosis with necrosis. It is primarily mediated by receptor-interacting protein kinase 1 (RIPK1), receptor-interacting protein kinase 3 (RIPK3), and mixed lineage kinase domain-like pseudokinase (MLKL) (Degterev et al., 2008; Pasparakis and Vandenabeele, 2015). Currently, necroptosis is found to be closely associated with tumors, and this relationship is two-sided. On the one hand, it can act as a programmed cell death to inhibit tumorigenesis and development; on the other hand, necroptosis can activate pro-inflammatory signals that strengthen cancer cells' proliferation and metastasis and enhance their invasiveness (McCormick et al., 2016; Strilic et al., 2016). One study found that in RIPK3 knockout mice, unregulated activation of some signaling pathways, such as the NF-kB and Wnt-β-catenin protein pathways, enhanced the capacity of intestinal epithelial cells (IEC) to multiply abnormally in a continuous inflammatory microenvironment, hence accelerating colorectal carcinogenesis (Bozec et al., 2016); Feng X et al. discovered that patients of colorectal cancer with low RIP3 expression had a lower overall survival (OS) and progression-free survival (PFS) than those that expressed high level of RIP3, insinuating that RIPK3 as a predictor of survival (Feng et al., 2015). Subsequent *in vitro* experiments demonstrated that overexpression of RIP3 significantly stunted the proliferation of cancer cells; similar to the above reports, Ertao et al. (2016), Sun et al. (2019) found that the down-regulation of MLKL was significantly associated with reduced OS in gastric cancer (GC) patients, implying that MLKL expression may be an independent predictive indicator for GC patients. Furthermore, Sun W et al. held that activated MLKL compromises the integrity of the cancer cell membrane, resulting in the discharge of intracellular pro-inflammatory molecules that could exert anti-tumor effects. Thus, by blocking MLKL-mediated necroptosis, gastric cancer cells might maintain tumor cell growth (Sun et al., 2019). In cancer cells, many critical molecules related with necroptosis are negatively regulated, raising the possibility that cancer cells may be able to resist necroptosis and hence survive. All malignancies, on the other hand, did not show downregulation of necroptosis-related molecules. In glioblastoma patients, increased RIP1 expression hindered p53 induction through activating the NF-κB pathway, and this upregulation was linked with a worse prognosis in this group of patients (Park et al., 2009). Necroptosis can also be involved in the tumor microenvironment to influence cancer progression. Seifert L et al. found in pancreatic ductal adenocarcinoma (PDA) that RIP3 downregulation *in vivo* did not promote tumor progression. Further studies found that RIP3 deletion induced an immunosuppressive tumor microenvironment with reduced infiltration of TAM and its M2-like Arg1+CD206+ subset, and meanwhile found that the lymphocyte infiltration in PDA increased. The microenvironment mediated by RIP3 deletion could suppress tumors (Seifert et al., 2016). Whereas alterations in the expression of these necroptotic molecules might cause changes in human immune surveillance against cancer. It was discovered that RIPK3 regulated NKT cell activity and promoted the generation of antitumor immune responses by these cells. The expression of RIPK1 and the activation of NF-κB were critical for the induction of $CD8^+$ T cell adaptive immunity (Newton et al., 2004; Yatim et al., 2015; Gong et al., 2019). This suggested that necroptosis might alter the expression of immune molecules in the microenvironment to affect cancer cell survival.

Changes in necroptosis-related molecules have been found not only in solid tumors but also in hematologic tumors. Höckendorf U et al. found in acute myeloid leukemia mice that cell death induced by RIPK3 and the release of interleukin-1b (IL-1b) which was mediated by inflammasome, limited myeloid leukemogenesis via eliminating transformed

cells and promoting differentiation of leukemia-initiating cells (Höckendorf et al., 2016); furthermore in the clinical 125 observed patients in chronic lymphocytic leukemia (CLL) found that the group of patients with low expression of the CYLD, which is a key mediator molecule in the necroptosis process, had a worse prognosis. This showed that CYLD might exert a critical role in the progression of CLL (Wu et al., 2014).

Nonetheless, in DLBCL, it remains to be further discussed whether and how necroptosis is involved in tumorigenesis development. Therefore, in the study, we analyzed DLBCL transcriptomic data to investigate the role of necroptosis in DLBCL. We looked into the mode of necroptosis in DLBCL and its relationship with the prognosis of DLBCL. And we specifically focused on its effect on the immune microenvironment of DLCBL. Our findings may add to our understanding of the role of necroptosis in malignancies and provide new information for the cure of DLBCL.

Materials and methods

Data download and processing

The datas in this study were from public GEO databases. We first searched by the search term "Diffuse large B cell lymphoma", and then by the restrictions 1) Organisms were restricted to "Homo sapiens"; 2) "Expression profiling by array" was the study type that we confined; 3) Sample counts >200; 4) Complete survival information; 5) Treatment with RCHOP regimen, and we finally included three datasets, GSE31312, GSE181063, and GSE10846. In addition, the dataset GSE31312 has the information about the treatment response assessment, where CR means complete remission, PR means partial remission, SD means stable disease, and PD stands for progressive disease.

We utilized the robust multiarray average (RMA) method to normalize the data after downloading the raw CEL files. And we used the "normalizebetweenarrays" function of the "limma" package to remove batch effect. After that, we cleaned the data even further, and the following criteria were employed in the data cleaning process for all three datasets: 1) data without complete survival time and survival status were excluded; 2) to ensure that death was due to tumor as much as possible, we excluded data with overall survival time <30 days; 3) to ensure comparability of patients, we excluded data treated with non-RCHOP regimens. Considering that GSE31312 covered the richest clinical information, we used it as the dataset for our main analysis, and to further perform data cleaning, we excluded sample with missing "IPI score", "GEP" information in GSE31312. Finally, GSE31312 included 421 patients with DLBCL, GSE181063 included 598 patients, and GSE10846 included 233 patients. Considering batch effects, combining datasets may cause unnecessary bias, so we kept the data independent.

Identification of DLBCL-associated necroptosis-related clusters

We collected 67 necroptosis-associated genes through "GOBP-NECROPTOTIC-SIGNALING -PATHWAY" and previous studies (Supplementary Material S1) (Zhao et al., 2021). The "GOBP-NECROPTOTIC-SIGNALING-PATHWAY" came from MSigDB (Molecular Signatures Database v7.4). First, we screened necroptosis-related genes with prognostic significance in the GSE31312 dataset by univariate cox regression, and genes with $p < 0.05$ were considered as DLBCL-associated necroptosis-related genes and included in the subsequent analysis. Subsequently, clustering analysis was performed in the GSE31312 dataset using the R package "ConsensusClusterPlus" (Wilkerson and Hayes, 2010) based on the expression of the above included genes. Resampling was performed 1000 times to ensure classification reliability. The clinical correlation study was carried out based on the clusters that had been identified. We used the survivfit function of the R package "survival" to analyze the prognostic differences among the three groups, and used the logrank test method to evaluate the significance of the prognostic differences between different groups of samples. And the proportion of cluster treatment response among different clusters was presented in the stacked histogram. The R package"ggalluvial" was used to visualize the relationship between the three clusters and IPI as well as GEP type, and we used Pearson chi square to test the above data.

Immune infiltration analysis

We first collected "NK cell mediated cytotoxicity" and "T cell receptor signaling pathway" from MSigDB (Molecular Signatures Database v7.4) and scored each patient in GSE31312 using ssGSEA. Subsequently, we analyzed the patients according to necroptosis-related clusters, and compared the differences of the above pathways between the groups, which were presented in the form of box plots.

After that, we tapped into the R package "ESTIMATE" (Yoshihara et al., 2013) to assess the "immune score", "stromal score" and "tumor purity" of the different clusters. We conducted Kruskal-Wallis test on the three groups and adopted Wilcox test between the two groups to evaluate the significance of the results. What's more, the differences between individual immune cells in different clusters were evaluated by both the "Cibersort" and "ssGSEA" methods. The B cells and associated cells were removed from the above two analyses in order to rule out relevant influence.

Identification of differential expression genes in necroptosis-related clusters

The R package "limma" was utilized to find differential genes among different clusters, and genes that met "adj.$p < 0.001$" as

well as "log FC = 0" were considered differentially significant for subsequent analysis. After completing the analysis of variance for all combinations, we took the intersection of the obtained results and showed them in the form of a Venn diagram. The similar method has been used to explore related genes in previously published article (Zhang et al., 2020).

We carried out a secondary clustering analysis on the GSE31312 dataset using the R package "ConsensusClusterPlus" based on the intersected differential genes, and resampled 1000 times to ensure classification reliability. Thereafter, we made on PCA analysis on the three clusters to visualize the differences in expression patterns among the three clusters. Similarly, we performed survival analysis on the three clusters to determine differences in patient prognosis between clusters to verify necroptotic functional cluster stability. We then integrated clinical characteristics including Gender, IPI, GEP, and treatment response to demonstrate the differences in gene expression and clinical characteristics among the three clusters in a heat map.

Construction and validation of a gene prognostic model and the evaluation of prognostic performance

We first applied the univariate Cox regression analysis to screen the cluster-related differential genes with prognostic significance in the GSE31312 dataset with $p < 0.05$ as the standard. Following that, we exploited Lasso-penalized Cox regression analysis to further screen for necroptosis-related genes with the greatest predictive performance and made use of these genes to build a risk score model. Multivariate Cox regression was adopted to further identify independent predictors and calculate regression coefficients. After collecting regression coefficients about every necroptosis-related gene that was significant to prognosis, using the following formula, we derived a risk score for each patient based on the expression of each gene (Carreras et al., 2020):

$$\begin{aligned}\text{Risk score} = &\left[\left(\text{Exp}_{\text{gene1}} \times \text{coefficient}_{\text{gene1}}\right)\right.\\ &+ \left(\text{Exp}_{\text{gene2}} \times \text{coefficient}_{\text{gene2}}\right)\\ &+ - + \left.\left(\text{Exp}_{\text{geneN}} \times \text{coefficient}_{\text{geneN}}\right)\right]\end{aligned}$$

The best cut-off value was calculated using the "survminer" package, and patients were separated into high and low groups based on this value. The Kaplan-Meier curves (Ranstam and Cook, 2017)were done to compare the OS of patients in both risk categories and the Time-ROC analysis (Kamarudin et al., 2017) was performed to determine the predictive potential of the model. The regression coefficients produced from GSE31312 were then applied to the test dataset, GSE181063 and GSE10846, which included entire clinical information, in order to level the risk scores of patients for external validation.

In view of the clinical characteristics, we integrated the genetic prognostic model grouping with clinical characteristics consisting of gender, IPI, GEP and treatment response, and performed lasso regression, and further selected the prognostic factors by stepwise regression based on the "lambda.min" value, and chose the final model based on the minimum AIC value. The final parameters obtained from the above analysis were used to construct prognostic line plots to forecast the OS of DLBCL patients at 1, 3, 5 and 7 years, and the stability of the model was appraised by time-ROC and calibration curve.

Results

Schematic Diagram of the Overall Flow of the study

Figure 1 is the workflow chart of this study, which basically describes the process of this study. First, we screened 17 genes through univariate cox expression analysis in the study. Based on these genes, cluster analysis was carried out to obtain cluster 1, 2 and 3. And then we carried out clinical survival analysis and clinical characteristics analysis on the three clusters. We found that the three groups had heterogeneity in clinical characteristics. So we then analyzed the TME of the three clusters (Figure 1A). In order to screen potentially necroptosis-related genes, we took the intersection of the obtained results through completing the analysis of variance for all combinations in the necroptosis-related clusters. And we performed secondary cluster analysis to obtain cluster A, B and C, and then performed survival analysis on the cluster A, B and C (Figure 1B). Through lasso penalized Cox analysis, 6 prognostic genes were included. Patients were divided into high-risk and low-risk groups according to these genes, and ROC time analysis was carried out. Finally, the clinical prognostic model was constructed, and the predictive nomogram was constructed and verified, and the calibration curve was analyzed (Figure 1C).

Necroptosis-related clusters identified in DLBCL

Firstly, 67 necroptotic genes were collected based on previous studies (Zhao et al., 2021), and necroptotic genes associated with DLBCL prognosis were screened using univariate cox regression analysis, whose resulte showed in the Supplementary Material S1, and genes with p-values less than 0.05 were deemed statistically significant. The final 17 DLBCL-related necroptotic genes were screened for the follow-up study. We regarded this group of genes as DLBCL-associated necroptosis-related genes, and the genes with risk ratio less than 1 were USP22, TNFRSF21, TNF, PANX1, MAP3K7, KLF9, IDH1, CYLD, BRAF, and ATRX. We believed that they were potential protective genes. TARDBP, SLC39A7, RNF31, MYC, EGFR, CASP8 and BCL2 were potential risk genes

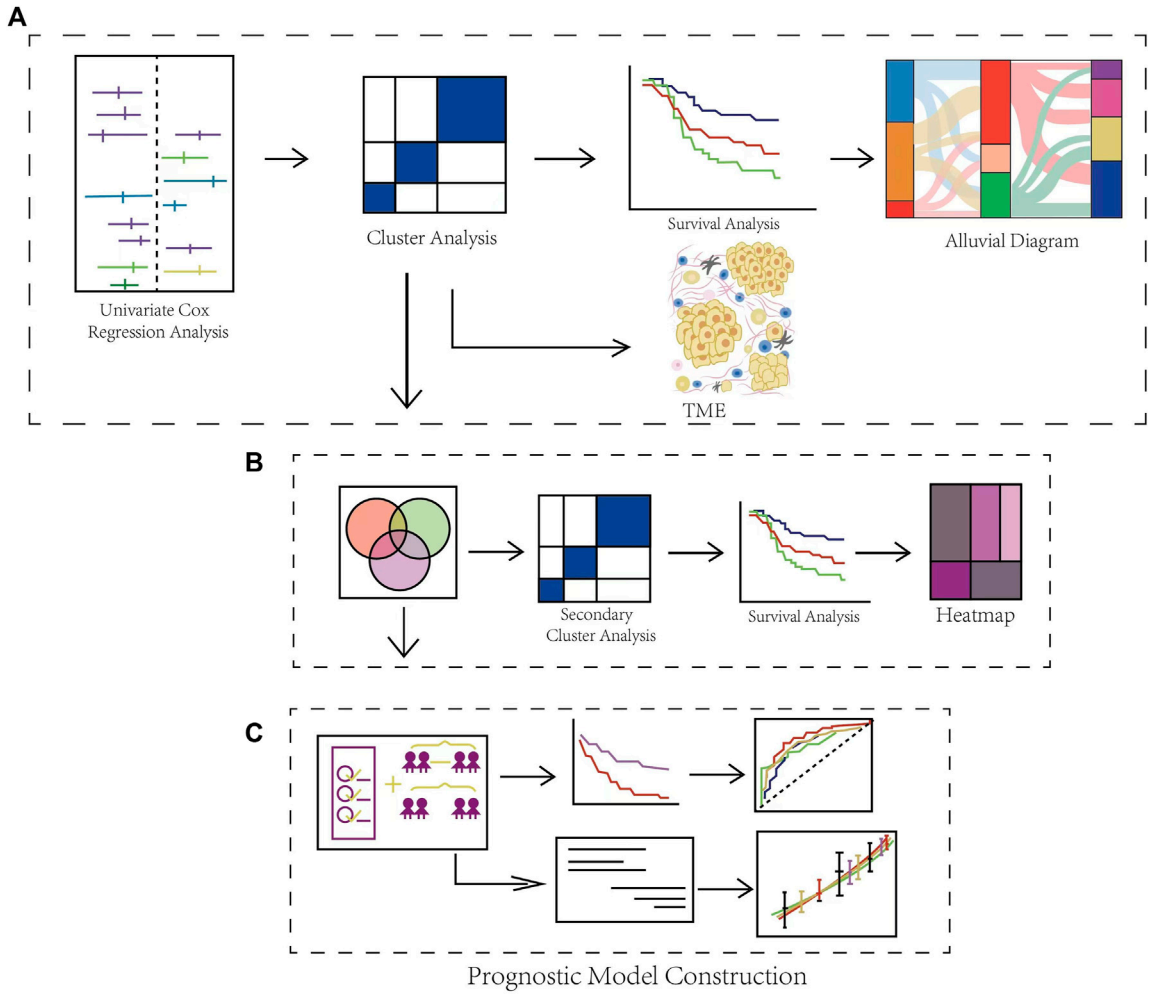


FIGURE 1
Schematic Diagram of the Overall Flow of the Study. **(A)** Identification of DLBCL-associated necroptosis-related clusters and immune infiltration analysis. **(B)** Identification of differential expression genes in necroptosis-related clusters. **(C)** Construction and validation of a gene prognostic model and the evaluation of prognostic performance.

(Figure 2A). According to previous studies, necroptosis had two sides in tumor (McCormick et al., 2016; Strilic et al., 2016), suggesting that necroptosis may also be heterogeneous in DLBCL. In the dataset GSE31312, we used the R package "ConsensusClusterPlus" to cluster 421 patients according to the expression of 17 DLBCL-associated necroptosis-related genes in DLBCL, and for the classification's reliability, 1000 resamplings were conducted. Finally, three distinctly different clusters, clusters 1, 2 and 3, were determined. Cluster 1 has 198 patients, cluster 2 has 94 patients, and cluster 3 has 129 patients (Supplementary Figure S1A). We further compared the expression of these 17 DLBCL-related necroptotic genes in three clusters. Most potential protective genes were highly expressed in cluster 1, while most potential risk genes were highly expressed in cluster 3 (Supplementary Figures S2D,E). This suggests that cluster 1 may be related to good clinical features and prognosis, while cluster 3 is the opposite. We then compared the prognosis of patients in these three clusters. Cluster 1 had the best prognosis, followed by cluster 2, and cluster 3 had the worst prognosis, indicating that necroptosis may have different modes of action in DLBCL (Figure 2B). Considering that the patients in GSE31312 were all treated with the same RCHOP regimen, the response to the regimen was directly related to the prognosis, so we analyzed the prognosis of the three clusters in which patients' response to RCHOP regimen treatment was found. The results showed that cluster 1 had the best response to R-CHOP treatment and had the highest CR rate, followed by cluster 2, while patients in cluster 3, who had the poorest prognosis, had the lowest CR rate. In contrast, the rate of patients with progressive disease (PD) after RCHOP treatment reflected an opposite trend, implying that necroptosis may be associated with RCHOP treatment sensitivity, and drug resistance (Figure 2C).

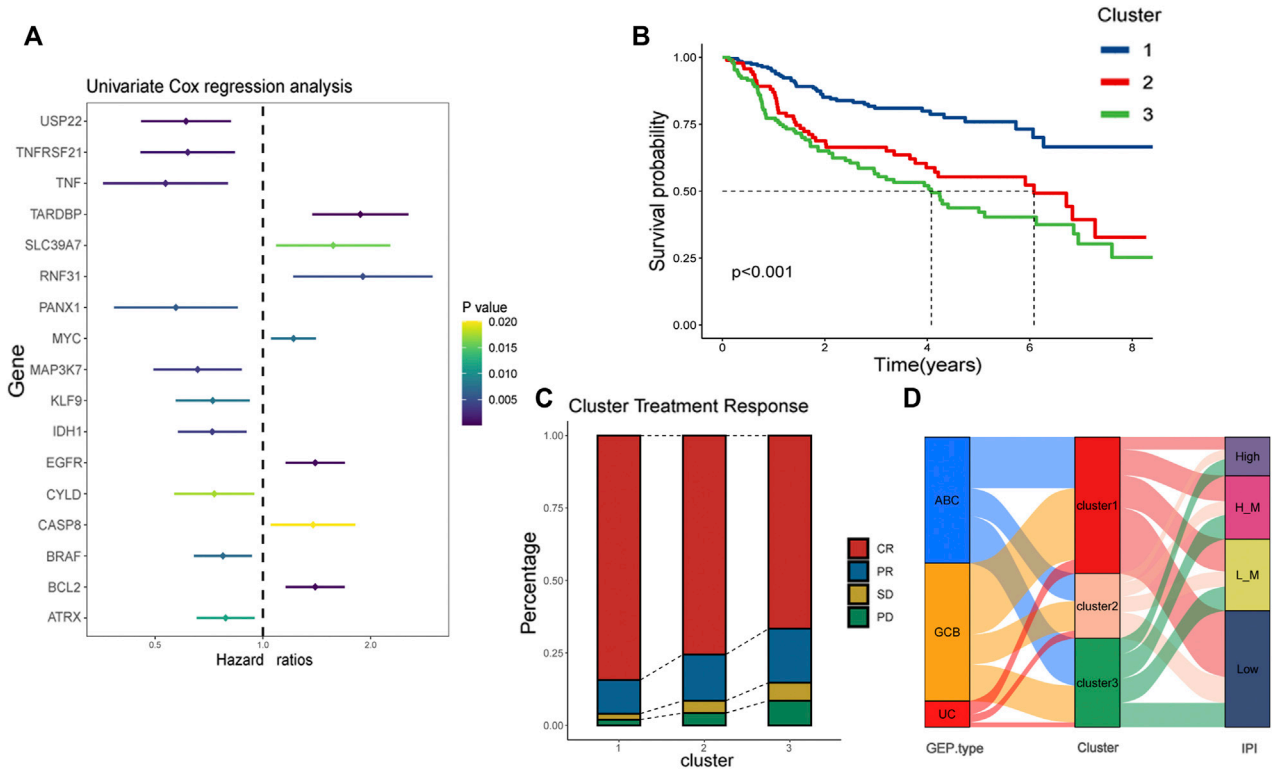


FIGURE 2
Identification of necroptosis-related clusters and clinical correlation analysis. **(A)** Risk ratios of 17 DLBCL prognosis-related necroptosis genes. Vertical coordinate is gene name, and horizontal coordinate represents risk ratio. Right side is *p*-value range symbolizing that the lighter the color, the larger the *p*-value. **(B)** Kaplan-Meier plots showing the prognosis of three necroptosis patterns in 421 patients from GSE31312. Blue line represents cluster 1, red cluster 2, and green cluster 3. Cluster 1 has the best prognosis. **(C)** Response of patients in the three clusters to RCHOP regimen treatment, with the vertical axis as a percentage. Red represents CR, yellow PR, green SD, and pink PD. **(D)** Alluvial is used to observe the relationship between cluster 1, cluster 2, and cluster 3 with IPI and GEP type. The red part of the middle bar represents cluster 1, pink cluster 2, and green cluster 3. "L_M"means low-intemediate, and "H_M" means high-intemediate.

To visualize the relationship between the three clusters and IPI as well as GEP type, we used the package "ggluvial" (Brunson, 2020) to visualize three. And we used Pearson chi square to test the above data. The results showed that the asymptotic significance was less than 0.05, indicating that cluster 1 had the highest proportion of patients with low IPI as well as GCB type among the three clusters (Supplementary Material S4). The results indicated that cluster 1's possess low IPI scores and GCB type; it has been reported that the prognosis of GCB type is superior to ABC type (Schmitz et al., 2018). This illustrates that cluster 1is associated with favorable survival characteristics (Figure 2D).

Evaluation of TME

The difference in prognosis suggests that there may be significant heterogeneity between our necroptosis-related clusters, so we first scored the necroptosis of the three clusters through "ssGSEA", which depicted that cluster 1 had the highest score, cluster 3 the second-highest, and cluster 2 the lowest (Figure 3A). This differs from the prognostic trend, with the best prognosis cluster 1 having the highest level of necroptosis and the worst prognosis cluster 3 scoring higher than cluster 2, explaining that necroptosis might exert a "two-sided" function in DLBCL. The tumor microenvironment (TME) is mainly composed of tumor cells, surrounding immune and inflammatory cells, tumor-related fibroblasts, and nearby interstitial tissues, microtubules, as well as various cell factors and chemokines. It is a complex integrated system, which can be divided into immune microenvironment dominated by immune cells and immune microenvironment dominated by fibroblasts (Fu et al., 2021). Earlier studies have shown that necroptosis is linked to TME, and we speculated that the heterogeneity of necroptosis in DLBCL was correlated with its involvement in influencing the tumor microenvironment. Subsequently, we analyzed the differential profile of several immune-related pathways among the three clusters. The results reflect, for example, that the NK cell mediated cytotoxicity, T cell receptor signaling pathway was significantly activated in

10.3389/fgene.2022.911443

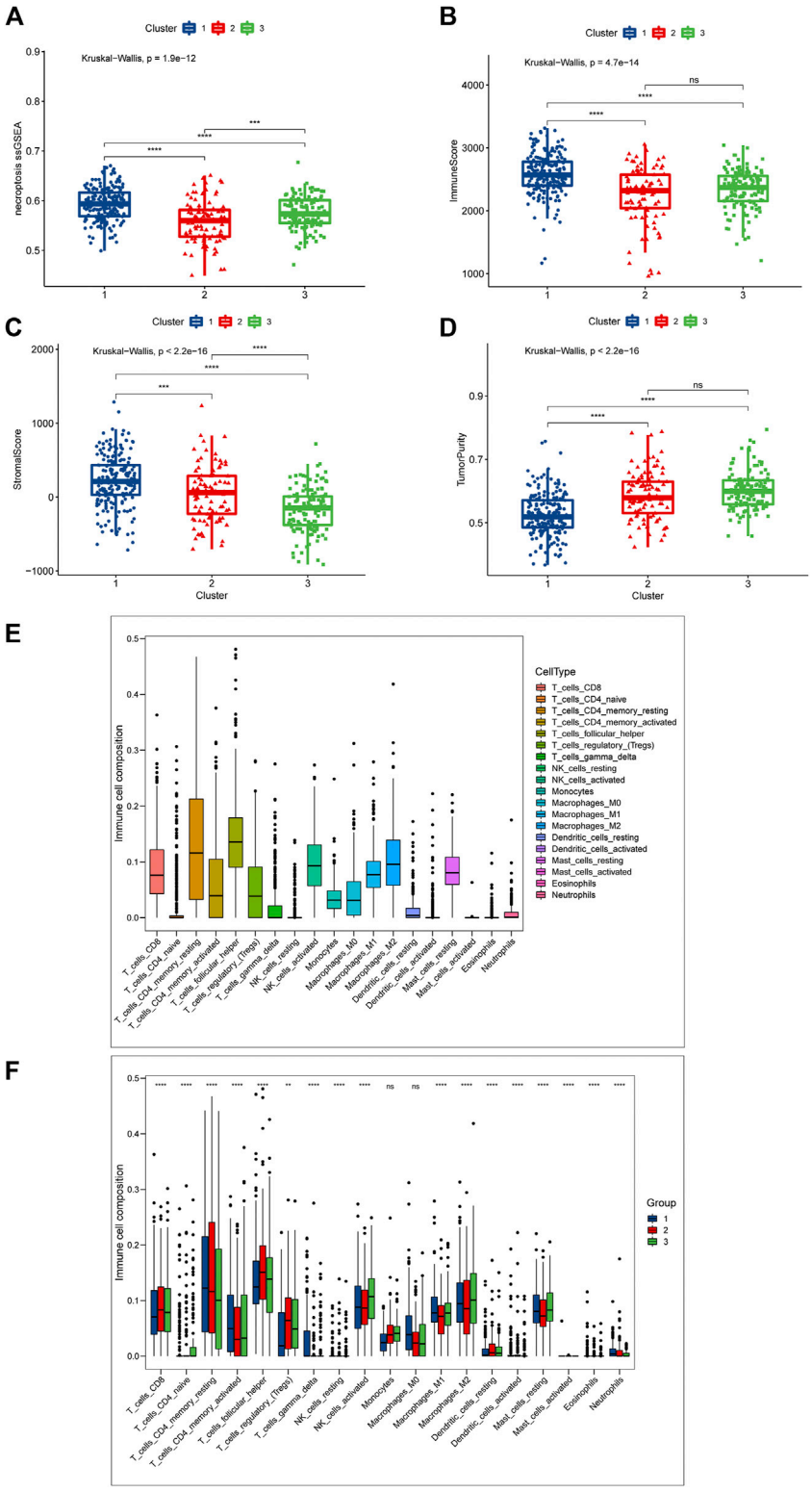


FIGURE 3
Differences in TME between the three necroptosis-related clusters (Cluster 1, Cluster 2 and Cluster 3). **(A)** Differences in necroptosis scores between the three clusters. **(B)** Differences in immune score. **(C)** Differences in stromal score between the three clusters; **(D)** Differences in tumor purity between the three clusters. **(E,F)** Cibersort was used to assess the infiltration of 19 immune cell types. **(E)** Overall infiltration of 19 immune cells. **(F)** Differences in 19 immune cells between the three clusters. Ns means "not statistically significant"; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$ (all significance designations that appear in this paper are minor criteria).

cluster 1 (Supplementary Figure S1B). Chances were that the immune microenvironment differed between different necroptosis-related clusters.

To give an insight into the immune microenvironment, we first valued the immune, stromal, and tumor purity between the three clusters by "ESTIMATE". As shown in Figure 2B, Cluster 1 had the highest immune score, and there was no statistically significant difference between clusters 2 and 3 (Figure 3B); as to the stromal score, cluster 1 had the highest stromal score, followed by cluster 2, and cluster 3 had the lowest stromal score (Figure 3C); cluster 1 had the lowest tumor purity, and there was no significant difference in tumor purity between cluster 2 and cluster 3 (Figure 3D). These data showed that the immunological microenvironment of the three clusters varied significantly, and we then employed the "Cibersort" package to further analyze the differences in immune cell composition among the three clusters. For the purpose of avoiding unneeded interference, we eliminated B cells and associated immune cells from the investigation. We first analyzed the proportion of 19 immune cells in the DLBCL patient population, and found that T cell follicular helper was the highest, followed by T cell CD4 memory resting, and Macrophages-M2 was the third (Figure 3E). Supplementary Figure S1C shows the different proportion of immune cells infiltrated in each patient. Next, we compared the differences in the immune components of the three clusters, and we focused on the aforementioned cells, and the results indicated that compared with other clusters, cluster 1 had the highest relative proportion in T cells CD4 memory resting, T cells CD4 memory activated, and Macrophages-M1, while cluster3 with the worst prognosis had the highest relative proportion in Macrophages-M2. For T cell follicular helper, it was different from the above cells and inconsistent with the trend of prognosis. Cluster 2 was the highest, cluster 3 was the second and cluster 1 was the lowest (Figure 3F).

In order to more rigorously evaluate the immune infiltration, we analyzed the abundance of different immune cells infiltration using "ssGSEA". What we discovered was that cluster1 displayed a high level of adaptive immune activation, such as $CD8^+$ T cells as well as $CD4^+$ T cells infiltrated abundantly; meanwhile, in the three clusters, the level of infiltration of many immune cells, including T cells CD4 memory activated, NK T cells, and NK cells, was consistent with the prognostic trend, manifesting the highest level of infiltration in cluster 1, followed by cluster 2, and the lowest in cluster 3 (Figure 4A); it was clear that we could notice differences in immune infiltration between the three clusters, and we also evaluated differences in antigen presentation mechanism (APM), which was shown to correlate with T cell infiltration scores (Şenbabaoğlu et al., 2016) and $CD8^+$ T cell effector. The results showed that cluster 1 attained a higher activation level than the other two clusters, which verified the high level of immune infiltration in cluster 1 (Figures 3B,C); in addition, we also noted that cluster 2 achieved the highest level of expression in mast cells, cluster1 got the lower level and cluster3 had the lowest. The above findings point towards the idea that necroptosis may have a heterogeneous effect on the immune microenvironment shaping of DLBCL among different functional clusters.

To further explore the correlation between immune cells and genes, we analyzed the correlation between M1, M2,CD8+T cells and 8 classical pathway molecules of necroptosis, as well as 6 modelled genes screened subsequently. The results showed that M1 cells were positively correlated with FAS, MLKL, RIPK1, RIPK3 and ACTB; M2 cells associated with poor prognosis were negatively correlated with SNRPD2,PAICS (Supplementary Figure S2B).

Identification of differentially expressed genes in necroptosis-related clusters

By taking the intersection of differentially expressed genes among the groups, we finally obtained 155 necroptosis-related cluster-associated genes (Figure 5A). We then carried out secondary clustering of GSE31312 based on these 155 genes and again obtained three clusters A, B, C, with 178 people in cluster A, 173 in cluster B, and 70 in cluster C (Supplementary Figure S1D). PCA analysis exhibited significant differences among the three clusters, and subsequent survival analysis observed similar survival differences to those of the necroptosis-related clusters (cluster 1, cluster 2, and cluster 3) — patients in cluster A had the best prognosis of the three, and cluster B had a better prognosis than cluster C (Figures 5B,C). A heat map of 155 genes suggested significantly different expression patterns among patients in the three clusters (Figure 5D). Among them, the expression patterns of Cluster A and Cluster C were almost opposite, and the genes with significantly elevated expression in Cluster A were mostly downregulated in Cluster C. We also analyzed the expression of 8 classical pathway genes of necroptosis among the three clusters, and we found that there was still significant inter-cluster heterogeneity. Most of the genes were expressed highest in cluster A and lowest in cluster C. It could be seen in Supplementary Figure S1E that TLR3, in contrast to the prognostic trend, experienced a gradual increase in expression level in the three clusters of A, B and C. Hence, we think that there may be stable differences in the mode of action of necroptosis in DLBCL.

Prognostic model construction of necroptosis-related genes

The study included 421 patients from the training dataset GSE31312.74 prognosis-related cluster differential genes were first screened by univariate cox regression and then further selected by Lasso-penalized Cox analysis. Eventually, 6 prognosis genes were

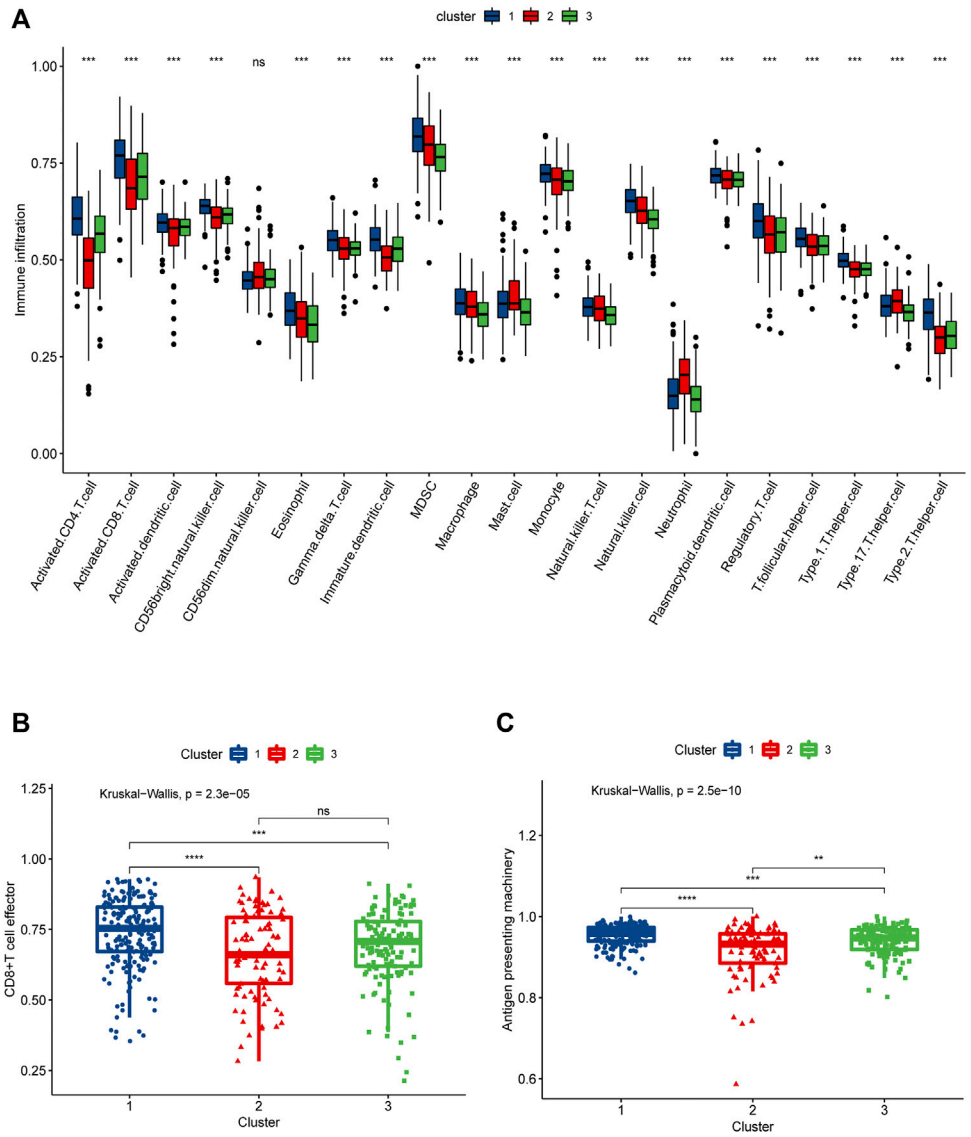


FIGURE 4
SsGSEA assessment of immune infiltration in three necroptosis-associated clusters (Cluster 1, Cluster 2 and Cluster 3). **(A)** Differences in the abundance of 21 infiltrating immune cells in the three necroptosis-related clusters, with Cluster 1 in blue, Cluster 2 in red and Cluster 3 in green. **(B)** Differences in $CD8^+$ T cell effector scores between the three necroptosis-related clusters. **(C)** Differences in antigen presenting machinery scores between three necroptosis-related clusters.

included and their regression coefficients—FSTL4, ACTB, SNRPD2, WHSC1L1, PAICS, and CLTC—were calculated. From the forest plot, FSTL4, SNRPD2 and PAICS could be potential oncogenes, while ACTB, WHSC1L1, and CLTC were thought as protective genes (Figure 6A). Next, each patient's risk score was computed using the following formula:

[(Exp FSTL4 × (0.807) + (Exp ACTB × (−0.474) + Exp SNRPD2 × (0.579) + Exp WHSC1L1 × (−0.26) + Exp PAICS × (0.207) + Exp CLTC × (−0.239)]. GSE31312 patients were classified into two groups according to their best cut-off values. The low-risk group had much superior OS than the high-risk group, as seen by Kaplan-Meier survival curves. (Figure 6B). The risk score was significantly connected with prognosis. As the score increased, the mortality rate of patients jumped (Figures 6C,D). We then evaluated the differential expression of 6 genes in the two risk groups and found that the expression of FSTL4, SNRPD2 and PAICS grew in the high-risk group, while the expression of ACTB, WHSC1L1 and CLTC went up in the low-risk group, which further confirmed the reliability of the selected prognostic genes (Figure 6E). In order to further verify the reliability of the genes we screened, we evaluated the differential expression of the 6 genes in clusters 1, 2 and 3. The results showed that SNRPD2 and PAICS,

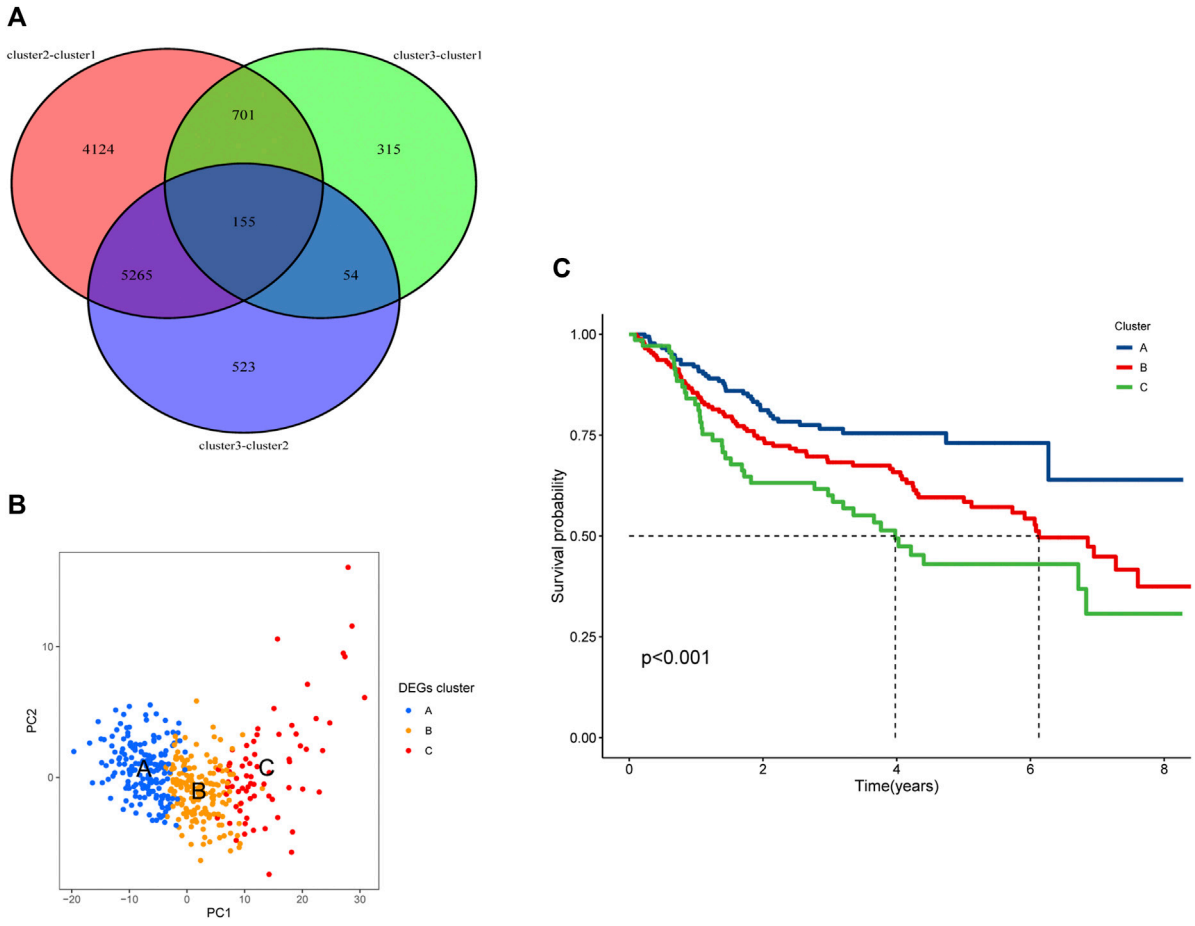


FIGURE 5
(Continued).

which represents poor prognosis in, creased in cluster3, while fstl4 had the lowest expression in cluster1 group with good prognosis. Besides, the expression of ACTB, WHSC1L1, and CLTC, which represent good prognosis, increased in cluster 1, further confirming the reliability of the screened prognostic genes (Supplementary Figure S2A). We then integrated the expression of necroptosis-related genes, the expression of 6 prognostic genes, and risk score data to construct a correlation matrix, as shown in Supplementary Figure S1F, 6 prognostic genes were firmly linked with the classical pathway molecules of necroptosis. Besides, ACTB, and FSTL4 showed a completely opposite relationship. FSTL4 was positively linked to FAD, FASLG, and TLR3, and negatively connected with FAS, MLKL, RIPK1, and RIPK3, while ACTB had a positive correlation with FAS, MLKL, RIPK1, and RIPK3, and inverse relation to TLR3 and FASLG. Consequently, we have reason to conjecture that ACTB and FSTL4 are mutually antagonistic necroptosis-related genes (Supplementary Figure S1F). The time-ROC curves depicted that the AUC of OS predicted by the genetic prognostic model was 0.74, 0.72, 0.71, and 0.74 at 1, 3, 5 and 7 years respectively, which were all greater than 0.7, demonstrating the good prognostic ability of the model (Figure 6F).As presented in the Kaplan Meier curves in the two external validation sets GSE10846 and GSE181063, compared to the low-risk group, the prognosis for the high-risk group was much poorer. (Figures 7A,B), suggesting that the model was stable.

Construction and validation of a predicted nomogram

Considering the excellent predictive prognostic ability of gene models, we further explored their role in clinical applications. We included several clinical features that were routinely considered prognostically significant in clinical practice and were available. Based on the results of lasso combined with stepwise analysis, we finally incorporated necroptosis-related genetic model risk groups, GEP type, RCHOP treatment response, and IPI scoring to construct a

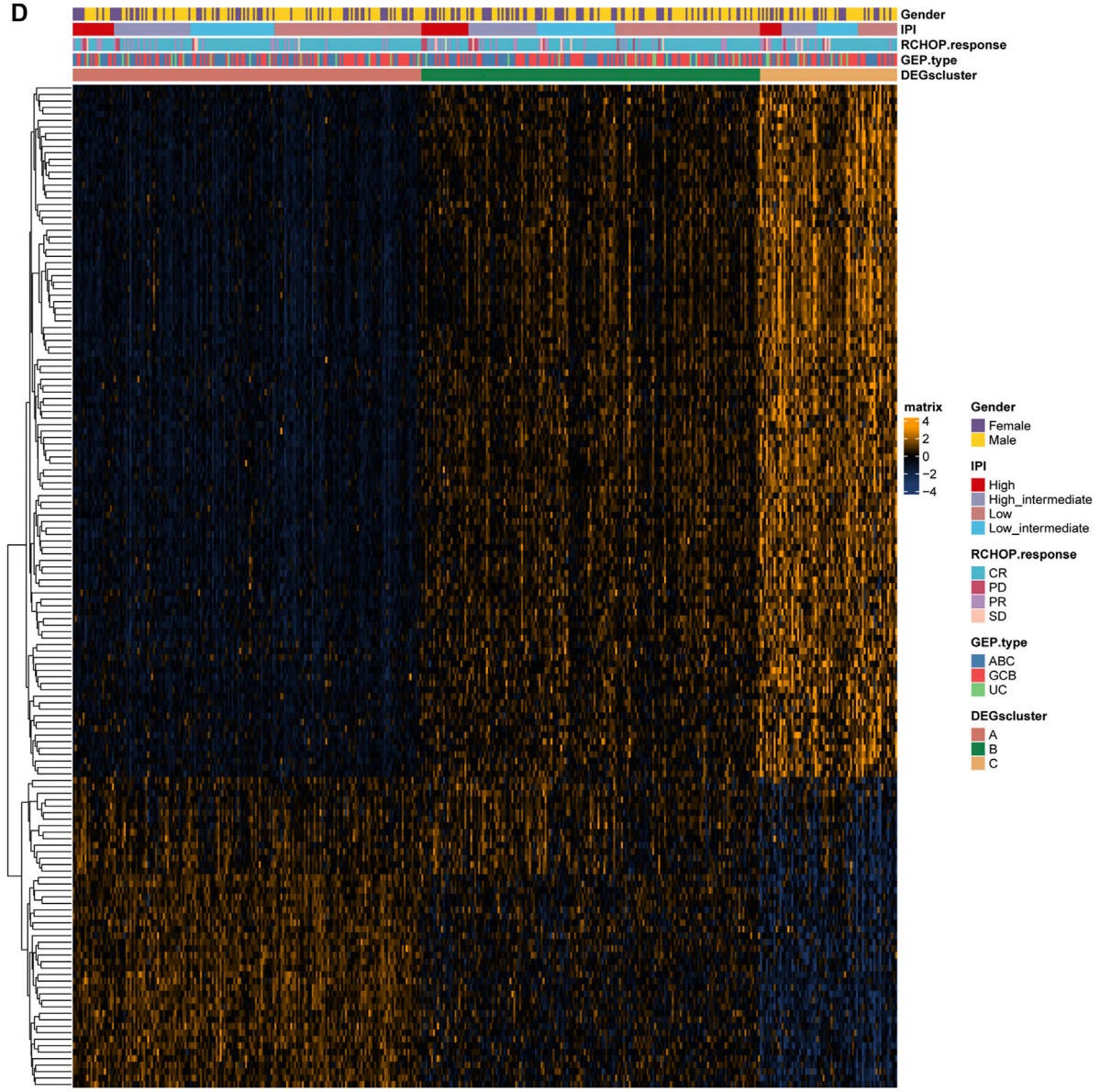


FIGURE 5
Cluster-related differentially expressed genes identification and secondary clustering. **(A)** Three circles represent the pairwise differentially expressed genes results between clusters, the middle dark blue represents common intersection, containing 155 genes in total. **(B)** PCA shows the differential situation of three differential gene related clusters. Blue dots represent cluster A, yellow dots cluster B, and red dots cluster C. **(C)** Survival analysis of clusters related to differentially expressed genes. Blue represents cluster A, red cluster B, and green cluster C. Cluster A has the best prognosis. **(D)** Heat map of 155 genes integrating clinical features, with yellow representing upregulated expression and blue indicating downregulation.

prognostic nomogram (Figure 7C). The C-index of the nomogram was 0.840. And the Time-ROC analysis showed that its AUCs for predicting OS at 1, 3, 5, and 7 years are 0.93, 0.87, 0.88, and 0.84, respectively (Figure 7D). According to a further inspection of the calibration curve, the predicted OS was generally consistent with the observed OS (Figure 7E). In conclusion, the above results indicated that the nomogram was successful in forecasting DLBCL patients' survival time.

Discussion

DLBCL, the most prevalent kind of NHL, is a heterogeneous group of diseases with varying biologic characteristics, clinical manifestations, and therapeutic responses. While 2/3 of DLBCL patients respond well to R-chop regimens, there is still a lack of effective treatment for patients with secondary or partial primary drug resistance. Therefore, it is essential to find new research

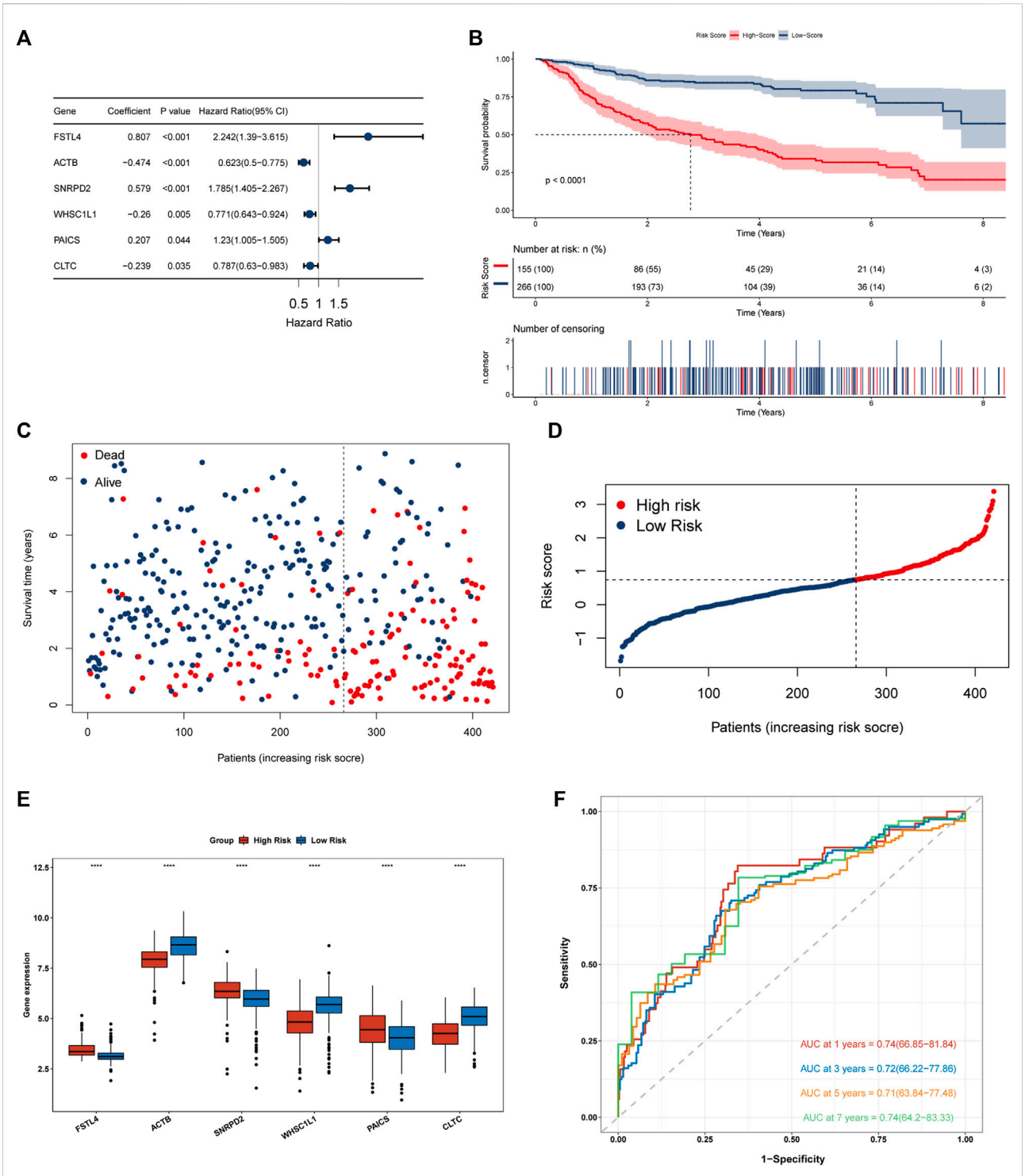


FIGURE 6
Construction and evaluation of a 6 necroptosis-related genes prognostic model. **(A)** Forest plot of 6 genes multivariate cox regression. **(B)** Kaplan-Meier survival analysis for the training set (GSE31312). **(C,D)** Heat map of risk scores and survival status of 421 DLBCL patients in the training dataset. **(E)** Differential expression of 6 modeled genes in the high and low risk groups in the training dataset, where red box represents high risk and blue represents low risk. **(F)** A Time-ROC curve analysis of the signature in training dataset.

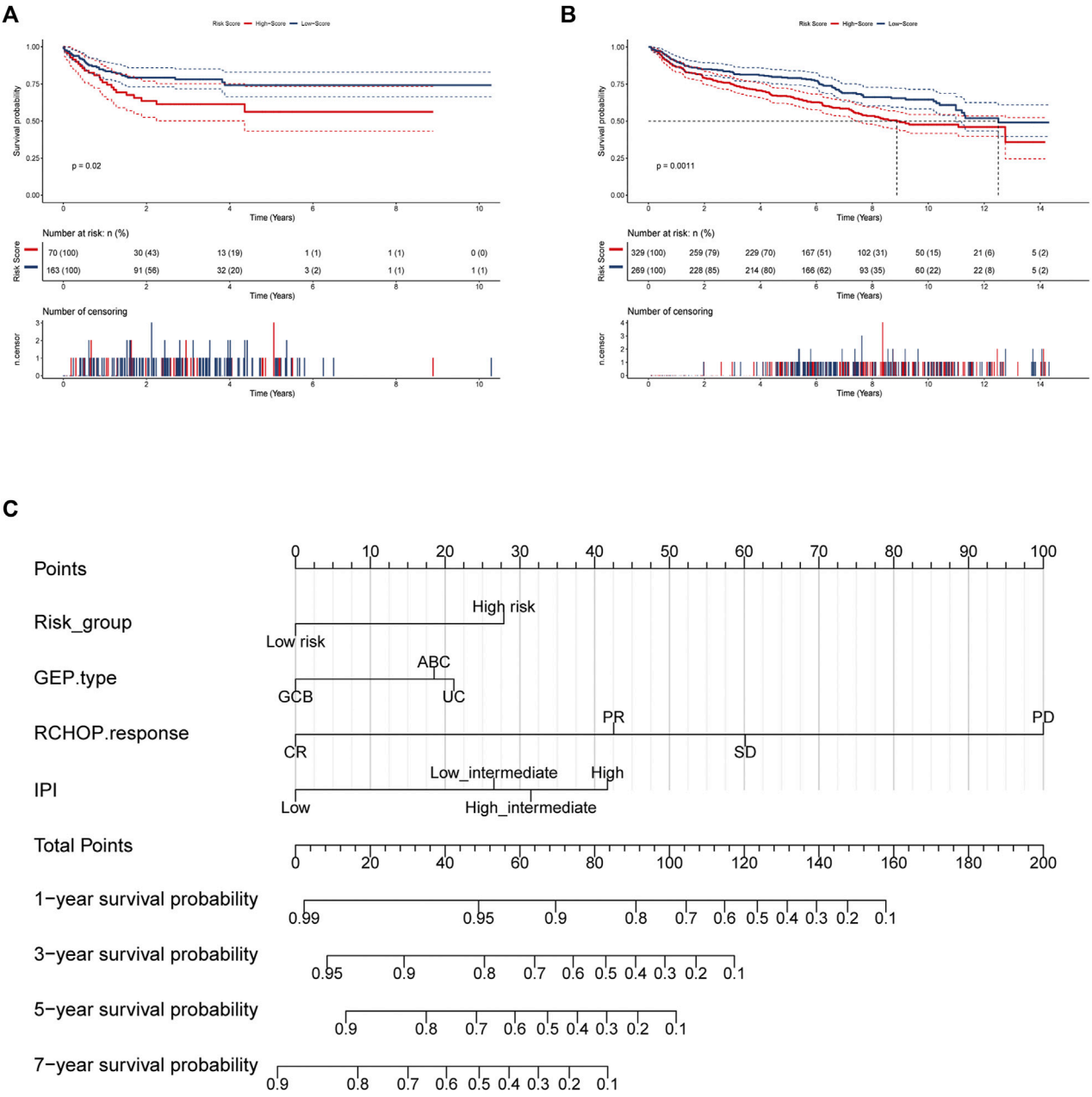


FIGURE 7
(Continued).

directions for DLBCL patients (Coiffier et al., 2010). Necroptosis, as a combination of apoptosis and necrosis (Tang et al., 2020), is a programmed cell death mediated mainly by RIPK1, RIPK3, and MLKL (Pasparakis and Vandenabeele, 2015; Gong et al., 2019). Currently, necroptosis has been found to be closely associated with tumors, such as colorectal cancer (Feng et al., 2015; Bozec et al., 2016), gastric cancer (Ertao et al., 2016; Sun et al., 2019), glioblastoma (Park et al., 2009), and pancreatic ductal adenocarcinoma (Seifert et al., 2016), whose involvement in the tumor microenvironment to influence the development of cancer has been demonstrated (Seifert et al., 2016). In addition, necroptosis, reportedly, played a role in the development of hematologic neoplasms. RIPK3 downregulation could accelerate leukemogenesis in acute myeloid leukemia mice, and afterward they found that RIPK3-mediated cell death could curtail the production of myeloid leukemia *via* eliminating transformed progenitor cells and promoting differentiation of leukemia-initiating cells (Höckendorf et al.,

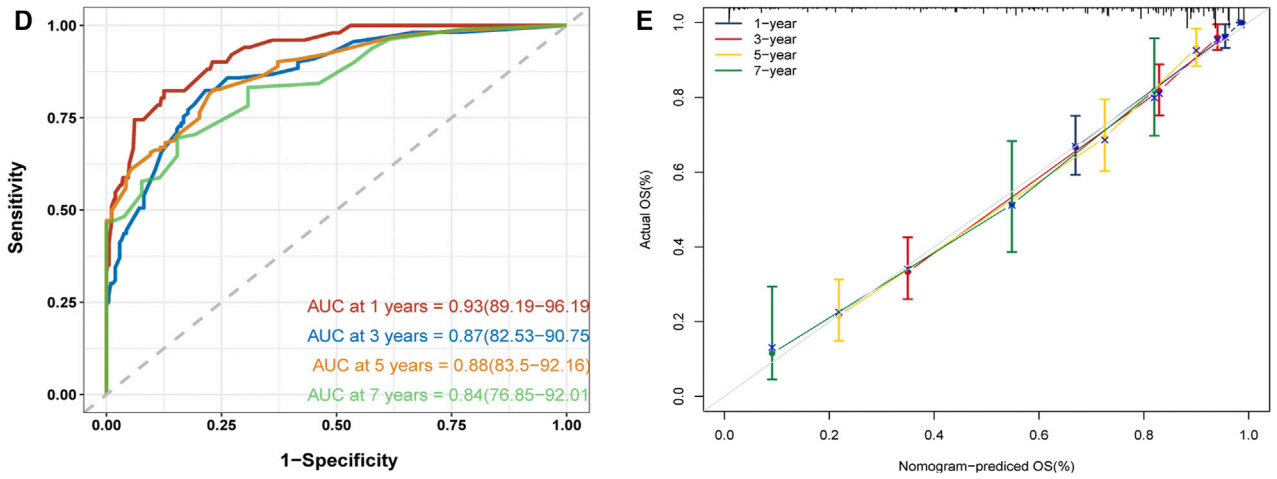


FIGURE 7
(Continued). External validation of signature and predicted prognostic nomogram construction. **(A,B)** Kaplan-Meier survival analysis of the external validation set, **(A)** GSE10846; **(B)** GSE181063. **(C)** Nomogram for predicting survival of DLBCL patients after 1, 3, 5 and 7 years **(D)** Time-dependent ROC assessment of the ability of nomogram to predict survival after 1,3,5,7 years **(E)** Calibration plots of predicted nomogram.

2016); in a clinical cohort study, chronic lymphocytic leukemia patients with low expression of the CYLD gene were observed to have a worse prognosis; before that, CYLD was noticed to interfere with several key tumor-associated signaling pathways, particularly the NF-κB pathway, by regulating the ubiquitination status of its components, and thus the low expression of CYLD impaired the ability to inhibit CLL tumor progression leading to a worse prognosis (Kovalenko et al., 2003; Wu et al., 2014). Yet, in DLBCL, whether and how necrotic apoptosis is involved in tumor development remains to be investigated, so our work was aimed to elucidate the function of necroptosis in DLBCL by analyzing DLBCL transcriptomic data.

In our research, three different necroptosis-related clusters were identified according to 17 necroptosis-related genes. We noted significant prognostic differences between the different clusters. The best prognosis was found in patients in cluster 1, while the poorest prognosis was seen in those in cluster 3. This reflected that heterogeneity may exist among the three clusters. Furthermore, cluster 1 was shown to have a GCB phenotype and a low IPI score compared to the other two clusters, which explained that cluster 1 may display favorable clinical characteristics. Following that, we re-obtained new three clusters, Cluster A, Cluster B, and Cluster C, based on the secondary clustering analysis of necroptosis-related differential genes in the three clusters mentioned above. The survival analysis results were similar to the first cluster analysis, and the expression of classical molecules of necroptosis in the three clusters was significantly heterogeneous, indicating that necroptosis may indeed manifest itself in distinct patterns in DLBCL.

RCHOP is the primary therapy for DLBCL, and a patient's reaction to medications is directly tied to their prognosis (Tilly et al., 2015). Our result of medication response is compatible with the distinct prognostic aspects of each cluster; whereas the rate of patients with PD showed an inverse tendency, suggesting that necroptosis may mediate RCHOP treatment sensitivity and resistance; in addition high necroptosis scores had lower mortality in cluster 1 than in cluster 2, yet cluster 3 had a higher mortality rate than cluster 2, which had the lowest necroptosis score. This demonstrated that necroptosis assumed a "double-edged sword" role in DLBCL.

The role of necroptosis in TME is increasingly recognized, as the immune microenvironment it created was linked to the development, metastasis, immunity and differentiation of many tumors (Tang et al., 2020). Therefore, by analyzing the immune infiltration of the three clusters, we noticed an abundant immune microenvironment component and a lower percentage of tumor cells in cluster 1while cluster 3 was opposite. Moreover, we observed that cluster 1 had the highest abundance of NK T cells, dendritic cells, and $CD8^+$ T cells among the three clusters by analyzing the immune components, and we noted that cluster 1 had a high relative proportion of Macrophages M1. Previously, necroptosis was found to promote the immunity of NKT cells by increasing RIP3 gene expression and activating PGAM5, which exerted a tumor suppressive function (Kang et al., 2015). After that, Paul S et al. also found that activation of NKT cells regulated the frequency of M1 macrophages and Th1 cells effector in secondary lymphoid tissues, further stunting tumor growth (Paul et al., 2019); reduced antigen presentation turned out a mechanism of tumor immune escape including inhibition of dendritic cell antigens, interference with antigen processing and presentation, but necroptosis, by the release of DAMPs, could activate dendritic cell releasing cytokines that activate adaptive immune to suppress tumors (Pasparakis and Vandenabeele, 2015; Jhunjhunwala et al., 2021). In addition,

Rosenberg et al. found through exploratory translational analysis that the gene expression of CD8 T cell effector was linked with PD-L1 immunohistochemical expression on tumor-infiltrating immune cells (Rosenberg et al., 2016). Thus we considered that necroptosis in cluster 1 acted as a tumor suppressor in DLBCL relying on inducing enhanced activation and infiltration of the immune component for improved prognosis, and analysis of APM and $CD8^+$ T cell effector indicated that cluster 1 may be responsive to anti-PD-1/PDL-1 therapy. We also noted in particular that in cluster 1 high levels of NK cell infiltration, and numerous experiments demonstrated that NK cells functioned effectively in fighting transformed and malignant cells (Hodgins et al., 2019), and in a clinical study by low NK cell counts could contributed to impaired R-CHOP response and increased risk of cancer recurrence Zare et al. (2020). In our analysis, patients in cluster 1 had the best therapeutic response to RCHOP, while patients in the other two clusters with less NK cell infiltration than cluster 1 had a worse therapeutic response. Therefore, it was reasonable to guess that necroptosis-induced high NK cell infiltration might potentially impact the responsiveness of R-chop treatment in individuals with DLBCL.

M2 macrophage was found to promote tumor cell survival, invasion, metastasis and angiogenesis, and their increased numbers gave rise to poor prognosis in patients (Pollard, 2004; Najafi et al., 2019). Some investigators found in patients with rectal cancer a potential link to increased tumor resistance to anticancer drugs by M2 (Lan et al., 2021). Such a prognostic relationship was seen in malignant lymphoma as well. Nam SJ, et al. found in a retrospective study of patients with follicular lymphoma that Tumor-associated macrophage (TAM) was observed at higher levels in the poor prognosis group than in the good prognosis group, suggesting that the downregulation of genes associated with macrophage activity in the mRNA transcriptome predicted a favorable outcome. What's more, such TAM had a phenotype and function that was similar to that of M2 macrophage (Nam et al., 2016). The team subsequently identified M2 in DLBCL as a possible significant predictor of poor patient prognosis (Nam et al., 2018). In our study, we found that the cluster 3 with the worst prognosis possessed the highest M2 infiltration, on the one hand further confirming that M2 took a negative role in the microenvironment of DLBCL, and on the other hand, notably, necroptosis in the cluster 3 characterized by relatively high levels of necroptosis seemed to play a role in contrast to cluster 1 which induced inhibition of the tumor microenvironment. Targeting TAM therapy including repolarization of TAM from M2 to M1 phenotype was gaining attention (Zheng et al., 2017) and in our study we found different levels of M1/M2 cell infiltration in different necroptotic functional clusters, and targeting necroptosis to induce M2 to M1 could be a promising therapeutic idea.

Meanwhile, we found the lowest abundance of mast cell infiltration in cluster 3 and the highest in cluster 2. Mast cells were reported to suppress immunity and promote tumor growth by releasing pro-angiogenic cytokines, interleukins and other cells in DLBCL patients, but previous studies showed that mast cell infiltration was a favorable prognostic factor in DLBCL (Hedström et al., 2007; Marinaccio et al., 2016). We believed that there was a certain association between mast cells and the reason why high levels of necroptosis in cluster 3 did not bring better survival than cluster 2. In other words, appropriate levels of mast cells could exert a positive effect in the TME microenvironment of DLBCL patients towards a good prognosis, while too high or too low levels of mast cells may play the opposite role. Notwithstanding this role in TME of DLBCL was not clear, it would a direction of our future research that deserved attention.

We also found that necroptosis-related genes were dependable predictors of prognosis. Our study identified six potential necroptosis genes of prognostic value, among which the ACTB-FSTL4 antagonistic relationship might be related to the mode of action of necroptosis in DLBCL, for patients with high expression of ACTB possessed a better prognosis, while upregulation of FSTL4 tended to indicate a poor prognosis. Previously, ACTB, in head and neck squamous carcinoma (HNSCC) and other cancers, was found to impact tumor metastasis as well as tumor invasion through NF-κB and Wnt-β-catenin protein pathways (Frontelo et al., 1998; Rubie et al., 2005; Popow et al., 2006); Wright A et al. reported that CYLD could limit the persistent activation of NF-κB signaling by deubiquitinating RIPK1, thus activating necroptosis-related pathways (Wright et al., 2007; Gong et al., 2019). Nonetheless, whether ACTB is engaged in the activation or inhibition of the necroptosis pathway in DLBCL by influencing RIPK1-mediated changes in the NF-κB pathway has not yet been investigated, and this is a direction worth exploring. Meanwhile, the 6 necroptosis-related gene prognostic model could successfully differentiate between high- and low-risk individuals of DLBCL. Such differences were validated in an external independent validation set, proving that the model was reliable. Further integration of clinical features could further appreciate the prognostic significance of the model, and the predictive prognostic nomogram constructed after combining with prognosis-related clinical features could effectively achieve individualized risk assessment.

However, our study has shortcomings. First, our study is based on a public database and lacks further validation of an independent prognostic cohort, and we included patients with baseline data prior to RCHOP treatment, and whether necroptosis is involved in tumor killing by the RCHOP regimen remains to be further investigated. Furthermore, the role of necroptosis in DLBCL still needs to be validated by further experiments.

Conclusion

In summary, our study found different modes of action of necroptosis in DLCBL, especially in the impact on TME. Clusters that induced abundant immune cell infiltration had a better prognosis, whereas clusters with a poorer immune

microenvironment component had a worse prognosis. Increased necroptosis-induced NK cell infiltration promised a better response to RCHOP treatment, while increased induction of M2 cell infiltration indicated a potential poor prognostic factor. The position of necroptosis in DLBCL could not be ignored, and a proper understanding of its role remained a worthwhile direction for our future studies. Systematic assessment of necroptosis patterns in DLBCL patients will facilitate our understanding of the cellular infiltration characteristics of the tumor microenvironment and the establishment of personalized therapy for DLBCL 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

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

Y-BP, WW, H-KC, JZ, YT, and JX participated in the design of the study and performed the statistical analysis. MZ and W-DL drafted the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported in part by Medical Science and Technology Project of Zhejiang Province. (2022KY1384).

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.2022.911443/full#supplementary-material

References

Bozec, D., Iuga, A. C., Roda, G., Dahan, S., and Yeretssian, G. (2016). Critical function of the necroptosis adaptor RIPK3 in protecting from intestinal tumorigenesis. *Oncotarget* 7 (29), 46384–46400. doi:10.18632/oncotarget.10135

Brunson, J. C. (2020). ggalluvial: Layered grammar for alluvial plots. *J. Open Source Softw.* 5 (49). doi:10.21105/joss.02017

Carreras, J., Hamoudi, R., and Nakamura, N. (2020). Artificial intelligence analysis of gene expression data predicted the prognosis of patients with diffuse large B-cell lymphoma. *Tokai J. Exp. Clin. Med.* 45 (1), 37–48.

Coiffier, B., Thieblemont, C., Van Den Neste, E., Lepeu, G., Plantier, I., Castaigne, S., et al. (2010). Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the groupe d'Etudes des lymphomes de l'Adulte. *Blood* 116 (12), 2040–2045. doi:10.1182/blood-2010-03-276246

Degterev, A., Hitomi, J., Germscheid, M., Ch'en, I. L., Korkina, O., Teng, X., et al. (2008). Identification of RIP1 kinase as a specific cellular target of necrostatins. *Nat. Chem. Biol.* 4 (5), 313–321. doi:10.1038/nchembio.83

Ertao, Z., Jianhui, C., Kang, W., Zhijun, Y., Hui, W., Chuangqi, C., et al. (2016). Prognostic value of mixed lineage kinase domain-like protein expression in the survival of patients with gastric caner. *Tumour Biol.* 37 (10), 13679–13685. doi:10.1007/s13277-016-5229-1

Feng, X., Song, Q., Yu, A., Tang, H., Peng, Z., and Wang, X. (2015). Receptor-interacting protein kinase 3 is a predictor of survival and plays a tumor suppressive role in colorectal cancer. *Neoplasma* 62 (4), 592–601. doi:10.4149/neo_2015_071

Frontelo, P., Gonzalez-GarriguesM.Vilaro, S., Gamallo, C., FAbrA, A., and QuintanillaM. (1998). Transforming growth factor beta 1 induces squamous carcinoma cell variants with increased metastatic abilities and a disorganized cytoskeleton. *Exp. Cell Res.* 244 (2), 420–432. doi:10.1006/excr.1998.4219

Fu, T., Dai, L. J., Wu, S. Y., Xiao, Y., Ma, D., Jiang, Y. Z., et al. (2021). Spatial architecture of the immune microenvironment orchestrates tumor immunity and therapeutic response. *J. Hematol. Oncol.* 14 (1), 98. doi:10.1186/s13045-021-01103-4

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 (1), 100. doi:10.1186/s12943-019-1029-8

Hedström, G., Berglund, M., Molin, D., Fischer, M., Nilsson, G., Thunberg, U., et al. (2007). Mast cell infiltration is a favourable prognostic factor in diffuse large B-cell lymphoma. *Br. J. Haematol.* 138 (1), 68–71. doi:10.1111/j.1365-2141.2007.06612.x

Höckendorf, U., Yabal, M., Herold, T., Munkhbaatar, E., Rott, S., Jilg, S., et al. (2016). RIPK3 restricts myeloid leukemogenesis by promoting cell death and differentiation of leukemia initiating cells. *Cancer Cell* 30 (1), 75–91. doi:10.1016/j.ccell.2016.06.002

Hodgins, J. J., Khan, S. T., Park, M. M., Auer, R. C., and Ardolino, M. (2019). Killers 2.0: NK cell therapies at the forefront of cancer control. *J. Clin. Invest.* 129 (9), 3499–3510. doi:10.1172/JCI129338

Jhunjhunwala, S., Hammer, C., and Delamarre, L. (2021). Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. *Nat. Rev. Cancer* 21 (5), 298–312. doi:10.1038/s41568-021-00339-z

Kamarudin, A. N., Cox, T., and Kolamunnage-Dona, R. (2017). Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med. Res. Methodol.* 17 (1), 53. doi:10.1186/s12874-017-0332-6

Kang, Y. J., Bang, B. R., Han, K. H., Hong, L., Shim, E. J., Ma, J., et al. (2015). Regulation of NKT cell-mediated immune responses to tumours and liver inflammation by mitochondrial PGAM5-Drp1 signalling. *Nat. Commun.* 6, 8371. doi:10.1038/ncomms9371

Kovalenko, A., Chable-Bessia, C., Cantarella, G., Israel, A., Wallach, D., and Courtois, G. (2003). The tumour suppressor CYLD negatively regulates NF-kappaB signalling by deubiquitination. *Nature* 424 (6950), 801–805. doi:10.1038/nature01802

Lan, H., Liu, Y., Liu, J., Wang, X., Guan, Z., Du, J., et al. (2021). Tumor-associated macrophages promote oxaliplatin resistance via METTL3-mediated m(6)A of TRAF5 and necroptosis in colorectal cancer. *Mol. Pharm.* 18 (3), 1026–1037. doi:10.1021/acs.molpharmaceut.0c00961

Liu, Y., and Barta, S. K. (2019). Diffuse large B-cell lymphoma: 2019 update on diagnosis, risk stratification, and treatment. *Am. J. Hematol.* 94 (5), 604–616. doi:10.1002/ajh.25460

Marinaccio, C., Ingravallo, G., Gaudio, F., Perrone, T., Ruggieri, S., Opinto, G., et al. (2016). T cells, mast cells and microvascular density in diffuse large B cell lymphoma. *Clin. Exp. Med.* 16 (3), 301–306. doi:10.1007/s10238-015-0354-5

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 (5), 522–529. doi:10.1093/carcin/bgw032

Najafi, M., Hashemi Goradel, N., Farhood, B., Salehi, E., Nashtaei, M. S., Khanlarkhani, N., et al. (2019). Macrophage polarity in cancer: a review. *J. Cell. Biochem.* 120 (3), 2756–2765. doi:10.1002/jcb.27646

Nam, S. J., Kim, S., Kwon, D., Kim, H., Kim, S., Lee, E., et al. (2018). Prognostic implications of tumor-infiltrating macrophages, M2 macrophages, regulatory T-cells, and indoleamine 2, 3-dioxygenase-positive cells in primary diffuse large B-cell lymphoma of the central nervous system. *Oncoimmunology* 7 (7), e1442164. doi:10.1080/2162402X.2018.1442164

Nam, S. J., Kim, S., Paik, J. H., Kim, T. M., Heo, D. S., Kim, C. W., et al. (2016). An increase in indoleamine 2, 3-dioxygenase-positive cells in the tumor microenvironment predicts favorable prognosis in patients with diffuse large B-cell lymphoma treated with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisolone. *Leuk. Lymphoma* 57 (8), 1956–1960. doi:10.3109/10428194.2015.1117610

Newton, K., Sun, X., and Dixit, V. M. (2004). Kinase RIP3 is dispensable for normal NF-kappa Bs, signaling by the B-cell and T-cell receptors, tumor necrosis factor receptor 1, and Toll-like receptors 2 and 4. *Mol. Cell. Biol.* 24 (4), 1464–1469. doi:10.1128/mcb.24.4.1464-1469.2004

Park, S., Hatanpaa, K. J., Xie, Y., Mickey, B. E., Madden, C. J., Raisanen, J. M., et al. (2009). The receptor interacting protein 1 inhibits p53 induction through NF-kappaB activation and confers a worse prognosis in glioblastoma. *Cancer Res.* 69 (7), 2809–2816. doi:10.1158/0008-5472.CAN-08-4079

Pasparakis, M., and Vandenabeele, P. (2015). Necroptosis and its role in inflammation. *Nature* 517 (7534), 311–320. doi:10.1038/nature14191

Paul, S., Chhatar, S., Mishra, A., and Lal, G. (2019). Natural killer T cell activation increases iNOS(+)CD206(-) M1 macrophage and controls the growth of solid tumor. *J. Immunother. Cancer* 7 (1), 208. doi:10.1186/s40425-019-0697-7

Pollard, J. W. (2004). Tumour-educated macrophages promote tumour progression and metastasis. *Nat. Rev. Cancer* 4 (1), 71–78. doi:10.1038/nrc1256

Popow, A., Nowak, D., and Malicka-Błaszkiewicz, M. (2006). Actin cytoskeleton and beta-actin expression in correlation with higher invasiveness of selected hepatoma Morris 5123 cells. *J. Physiol. Pharmacol.* 57, 111–123.

Ranstam, J., and Cook, J. A. (2017). Kaplan-Meier curve. *Br. J. Surg.* 104 (4), 442. doi:10.1002/bjs.10238

Rosenberg, J. E., Hoffman-Censits, J., Powles, T., van der Heijden, M. S., Balar, A. V., Necchi, A., et al. (2016). Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet* 387 (10031), 1909–1920. doi:10.1016/S0140-6736(16)00561-4

Rubie, C., Kempf, K., Hans, J., Su, T., Tilton, B., Georg, T., et al. (2005). Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues. *Mol. Cell. Probes* 19 (2), 101–109. doi:10.1016/j.mcp.2004.10.001

Schmitz, R., Wright, G. W., Huang, D. W., Johnson, C. A., Phelan, J. D., Wang, J. Q., et al. (2018). Genetics and pathogenesis of diffuse large B-cell lymphoma. *N. Engl. J. Med.* 378 (15), 1396–1407. doi:10.1056/NEJMoa1801445

Sehn, L. H., and Gascoyne, R. D. (2015). Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. *Blood* 125 (1), 22–32. doi:10.1182/blood-2014-05-577189

Seifert, L., Werba, G., Tiwari, S., Giao Ly, N. N., Alothman, S., Alqunaibit, D., et al. (2016). The necrosome promotes pancreatic oncogenesis via CXCL1 and Mincle-induced immune suppression. *Nature* 532 (7598), 245–249. doi:10.1038/nature17403

Şenbabaoğlu, Y., Gejman, R. S., Winer, A. G., Liu, M., Van Allen, E. M., de Velasco, G., et al. (2016). Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* 17 (1), 231. doi:10.1186/s13059-016-1092-z

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer statistics, 2019. *Ca. Cancer J. Clin.* 69 (1), 7–34. doi:10.3322/caac.21551

Strilic, B., Yang, L., Albarran-Juarez, J., Wachsmuth, L., Han, K., Muller, U. C., et al. (2016). Tumour-cell-induced endothelial cell necroptosis via death receptor 6 promotes metastasis. *Nature* 536 (7615), 215–218. doi:10.1038/nature19076

Sun, W., Yu, W., Shen, L., and Huang, T. (2019). MLKL is a potential prognostic marker in gastric cancer. *Oncol. Lett.* 18 (4), 3830–3836. doi:10.3892/ol.2019.10687

Tang, R., Xu, J., Zhang, B., Liu, J., Liang, C., Hua, J., et al. (2020). Ferroptosis, necroptosis, and pyroptosis in anticancer immunity. *J. Hematol. Oncol.* 13 (1), 110. doi:10.1186/s13045-020-00946-7

Tilly, H., GoMes da SilvaM.Vitolo, U., JAck, A., MeignanM.Lopez-Guillermo, A., et al. (2015). Diffuse large B-cell lymphoma (DLBCL): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 26, v116–25. doi:10.1093/annonc/mdv304

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

Wright, A., Reiley, W. W., Chang, M., Jin, W., Lee, A. J., Zhang, M., et al. (2007). Regulation of early wave of germ cell apoptosis and spermatogenesis by deubiquitinating enzyme CYLD. *Dev. Cell* 13 (5), 705–716. doi:10.1016/j.devcel.2007.09.007

Wu, W., Zhu, H., Fu, Y., Shen, W., Xu, J., Miao, K., et al. (2014). Clinical significance of down-regulated cylindromatosis gene in chronic lymphocytic leukemia. *Leuk. Lymphoma* 55 (3), 588–594. doi:10.3109/10428194.2013.809077

Yatim, N., Jusforgues-Saklani, H., Orozco, S., Schulz, O., Barreira da Silva, R., Reis e Sousa, C., et al. (2015). RIPK1 and NF-κB signaling in dying cells determines cross-priming of CD8$^+$ T cells. *Science* 350 (6258), 328–334. doi:10.1126/science.aad0395

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

Zare, N., Haghjooy Javanmard, S. H., Mehrzad, V., Eskandari, N., and Andalib, A. R. (2020). Effect of plasma-derived exosomes of refractory/relapsed or responsive patients with diffuse large B-cell lymphoma on natural killer cells functions. *Cell J.* 22 (1), 40–54. doi:10.22074/cellj.2020.6550

Zhang, B., Wu, Q., Li, B., Wang, D., Wang, L., and Zhou, Y. L. (2020). m(6 A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol. Cancer* 19 (1), 53. doi:10.1186/s12943-020-01170-0

Zhao, Z., Liu, H., Zhou, X., Fang, D., Ou, X., Ye, J., et al. (2021). Necroptosis-related lncRNAs: predicting prognosis and the distinction between the cold and hot tumors in gastric cancer. *J. Oncol.* 2021, 6718443. doi:10.1155/2021/6718443

Zheng, X., Turkowski, K., Mora, J., Brune, B., Seeger, W., Weigert, A., et al. (2017). Redirecting tumor-associated macrophages to become tumoricidal effectors as a novel strategy for cancer therapy. *Oncotarget* 8 (29), 48436–48452. doi:10.18632/oncotarget.17061

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 15 August 2022
DOI 10.3389/fgene.2022.896064

Check for updates

OPEN ACCESS

EDITED BY
Lan Zhao,
Stanford University, United States

REVIEWED BY
Shibiao Wan,
St. Jude Children's Research Hospital, United States
Austin Gillen,
University of Colorado Anschutz Medical Campus, United States

*CORRESPONDENCE
Jianchun Cai,
jianchunfh2@sina.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 14 March 2022
ACCEPTED 06 July 2022
PUBLISHED 15 August 2022

CITATION
Wei X, Liu J, Hong Z, Chen X, Wang K and Cai J (2022), Identification of novel tumor microenvironment-associated genes in gastric cancer based on single-cell RNA-sequencing datasets. *Front. Genet.* 13:896064. doi: 10.3389/fgene.2022.896064

COPYRIGHT
© 2022 Wei, Liu, Hong, Chen, Wang and Cai. 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.

Identification of novel tumor microenvironment-associated genes in gastric cancer based on single-cell RNA-sequencing datasets

Xujin Wei[1,2,3†], Jie Liu[1†], Zhijun Hong[2,3], Xin Chen[1,2,3], Kang Wang[2,3] and Jianchun Cai[1,2,3*]

[1]The Graduate School of Fujian Medical University, Fuzhou, China, [2]Department of Gastrointestinal Surgery, Zhongshan Hospital of Xiamen University, Institute of Gastrointestinal Oncology, School of Medicine, Xiamen University, Xiamen, China, [3]Xiamen Municipal Key Laboratory of Gastrointestinal Oncology, Xiamen, China

Tumor microenvironment and heterogeneity play vital roles in the development and progression of gastric cancer (GC). In the past decade, a considerable amount of single-cell RNA-sequencing (scRNA-seq) studies have been published in the fields of oncology and immunology, which improve our knowledge of the GC immune microenvironment. However, much uncertainty still exists about the relationship between the macroscopic and microscopic data in transcriptomics. In the current study, we made full use of scRNA-seq data from the Gene Expression Omnibus database (GSE134520) to identify 25 cell subsets, including 11 microenvironment-related cell types. The MIF signaling pathway network was obtained upon analysis of receptor–ligand pairs and cell–cell interactions. By comparing the gene expression in a wide variety of cells between intestinal metaplasia and early gastric cancer, we identified 64 differentially expressed genes annotated as immune response and cellular communication. Subsequently, we screened these genes for prognostic clinical value based on the patients' follow-up data from The Cancer Genome Atlas. *TMPRSS15, VIM, APOA1,* and *RNASE1* were then selected for the construction of LASSO risk scores, and a nomogram model incorporating another five clinical risk factors was successfully created. The effectiveness of least absolute shrinkage and selection operator risk scores was validated using gene set enrichment analysis and levels of immune cell infiltration. These findings will drive the development of prognostic evaluations affected by the immune tumor microenvironment in GC.

KEYWORDS

gastric cancer, single-cell RNA sequencing, tumor microenvironment, immune infiltration, prognostic biomarker

Introduction

Although the incidence and mortality rates have been declining worldwide, gastric cancer (GC) remains a common and lethal malignancy, especially in Asian countries (Smyth et al., 2020). In addition to traditional chemotherapy and surgery, adjuvant therapies, such as molecularly targeted therapy (Mundekkad and Cho, 2022) and immunotherapy, are gradually emerging as the staples of GC treatment. Uncovering the molecular mechanisms of the initiation and progression of GC is therefore critical for improving therapeutic efficacy. Molecular markers help deepen our understanding of GC subtypes (Bijlsma et al., 2017; Zhao et al., 2019), and with the technological advances in high-throughput sequencing, the focus on tumor heterogeneity is driving progress in precision medicine (Zeng and Jin, 2021) simultaneously. There is growing evidence that intra-tumoral heterogeneity includes not only genomic features but also the complex tumor microenvironment (TME). It is well recognized that TME comprises various stromal cells, abundant angiogenesis, and immune cell infiltration (Hanahan, 2022). Understanding the "soil" on which the "seed" grows into a tumor has essential implications for the diagnosis and treatment of GC.

Single-cell RNA sequencing (scRNA-seq) is a potent tool to obtain complete RNA transcripts at the level of single cells by RNA extraction, reverse transcription, amplification, and sequencing (Tang et al., 2009; Ramskold et al., 2012; Picelli et al., 2014). Compared with traditional sequencing methods of tumor tissue, scRNA-seq solves the problem that individual differences between cell types are ignored when the expression of all the genes in thousands of cells is averaged. It shows clear superiority in studying the diversity of tumor cell lineages and predicting interactions between cancer and the microenvironment (Muller and Diaz, 2017; Kumar et al., 2021). In terms of GC research, scRNA-seq shed light on the transcriptome network at different stages of the disease process, from atrophic gastritis, intestinal metaplasia (IM), and dysplasia, to early gastric cancer (EGC) (Zhang et al., 2019), as well as the spatial heterogeneity of microenvironment-related cells in diffuse-type GC (Jeong et al., 2021). Performing scRNA-seq of metastatic GC, the origins of transcriptomic heterogeneity in peritoneal carcinomatosis were analyzed (Wang R. et al., 2021), identifying *CLDN11* and *CDK12* as markers of lymph node metastasis (Wang B. et al., 2021). Nonetheless, combining large amounts of scRNA-seq data with multi-omics datasets and determining the clinical implications remains a challenge (Wang et al., 2022).

We data-mined existing single-cell transcriptome data from the Gene Expression Omnibus (GEO) to identify microenvironment-related cell types and draw signaling pathway networks according to the marker genes or intercellular communication-related genes. Enrichment analysis was performed to reveal the function of differentially expressed genes (DEGs) between IM and EGC. We leveraged clinical data from The Cancer Genome Atlas (TCGA) to establish a risk-scoring model by the LASSO-Cox regression algorithm. It contained the expression levels of *TMPRSS15, VIM, APOA1,* and *RNASE1,* which were highly correlated with clinical outcomes and immune cell infiltration. The results of our study may provide new options for prognostic biomarkers in GC.

Materials and methods

Data collection and pre-processing

ScRNA-seq data of normal cells and GC cells from the GSE134520 dataset were downloaded from the GEO database from the official website (https://www.ncbi.nlm.nih.gov/geo/) (Zhang et al., 2019), which had been pre-processed using CellRanger software. The dataset contains *Homo sapiens* samples sequenced by the HiSeq X Ten platform. A total of 13 biopsy specimens from the gastric antral mucosa of patients with non-atrophic gastritis, chronic atrophic gastritis, IM, and EGC were selected, and 56,440 cells were included in the dataset. Cells with an abnormal proportion of mitochondrial genes should be removed because it reflects the imbalance of cellular homeostasis and low cell quality. Given the potential existence of diploid cells, cells with genes < 200 or >5,800 were filtered out. Furthermore, the scDblFinder function in the Seurat R package was called to remove the double droplets. Finally, 48,566 cells were enrolled in our study.

The R package "TCGAbiolinks" (version 2.22.2) was used to obtain the gene expression in units of fragments per kilobase million (FPKM) for 407 patients with stomach adenocarcinoma (STAD) from TCGA. We also downloaded the clinical data for these patients, which include survival status (367 patients), stages, age, grades, and survival time. At the same time, the data of "Masked Somatic Mutation" and "Masked Copy Number Segment" called MuTect2 were downloaded.

ScRNA-seq analysis using Seurat

First, we installed the R packages in R (version 4.1) and Seurat (version 4.0.5) (Satija et al., 2015), and used the merge function to merge the created Seurat object. To reduce the influence of different sequencing depths in cells, the normalization for raw counts was performed using the NormalizeData function ("LogNormalize" method), the top 2,000 variable features were identified using the Find Variable Genes function ("vst" method), and the data were integrated and scaled using ScaleData function. Subsequently, we performed principal component analysis (PCA) with the variable genes as inputs and identified the significant principal components whose

TABLE 1 Cell types of 25 clusters.

Cell type	Marker gene	Number of cell (total = 48,566 (%))
Gastric epithelial cells	*MUC6*, *TFF2*, *MUC5AC*, and *TFF1*	26,870 (55.33%)
Metaplastic stem-like cell	*OLFM4*, *EPHB2*, and *SOX9*	7570 (15.59%)
Enterocytes	*FABP1* and *APOA1*	5340 (11.00%)
T cell	*CD2* and *CD3D*	2574 (5.30%)
B cell	*CD79A*	2307 (4.75%)
Fibroblast	*DCN* and *PDPN*	1431 (2.95%)
Endothelial cell	*VWF* and *ENG*	1002 (2.06%)
Macrophage	*CSF1R* and *CD68*	479 (0.99%)
Mast cell	*TPSAB1*	384 (0.79%)
Goblet cell	*MUC2* and *ITLN1*	357 (0.74%)
Smooth muscle cell	*ACTA2*	252 (0.52%)

p-value distributions were then visualized by the jackStraw function. In the FindNeighbors and FindClusters function with a resolution of 0.8, the Louvain algorithm was chosen for cell clustering. FindAllMarkers function with the Wilcoxon rank-sum test was carried out to identify specific marker genes, which compared expression values between cells in the cluster and all other cells. Finally, the results are represented with tSNE (t-distributed Stochastic Neighbor Embedding) dimension reduction by RunTSNE function.

Cell annotation and identification of differentially expressed genes

Cell types were identified based on marker gene sets listed in Table 1. Subsequently, we used the FindAllMarkers function in Seurat to detect DEGs among the different cell types, using a p-value < 0.05 and |log2FoldChange (FC)| > 1 as the thresholds. Additionally, we used the FindAllMarkers function to analyze the differences between the EGC and IM groups to extract DEGs using the same thresholds previously stated. The intersection of three methods ("wilcox," "t," and "roc" test) was taken and transformed into a list of the final DEGs. Finally, the expression patterns of these DEGs in the different cell types were shown using heat maps, and the corrplot R package (version 0.92) was then used to perform a correlation analysis between these DEGs.

Enrichment analysis of differentially expressed genes between early gastric cancer and IM groups

Gene Ontology (GO) enrichment analysis is commonly performed in large-scale functional enrichment analyses of genes on different dimensions and levels. It is generally performed on three levels, namely, biological process (BP), molecular function (MF), and cellular component (CC). Kyoto Encyclopedia of Genes and Genomes (KEGG) is a widely used database dealing with biological systems such as genomes, biological pathways, diseases, and drugs. The clusterProfiler (version 4.2.0) R package was used to perform GO functional annotation and KEGG pathway enrichment analysis for the DEGs between the EGC and IM groups to identify significantly enriched biological processes. The enrichment results were visualized using bar graphs and bubble graphs. An adjusted p-value of < 0.05 was defined as the significance threshold in the enrichment analysis.

Analysis of cell-cell communication

The CellChat (http://www.cellchat.org/) R package (Jin et al., 2021) was used to calculate the intensity of cell–cell interactions and communication based on single-cell gene expression profiles and known ligands, receptors, and their cofactors. Significant ligand-receptor pairs were further identified based on the probability of receptor–ligand interactions and the results of perturbation testing. We then built a cell–cell communication network by adding the number or the intensity of significantly interacting ligand–receptor pairs between cell types. In addition, we visualized multiple ligand–receptor pairs or intercellular communications mediated by signaling pathways using bubble graphs to investigate the intensity of ligand–receptor interactions between cell types or the characteristics of ligands and receptors in terms of gene expression levels, including their commonalities or differences. Finally, we performed a systemic analysis of the cell–cell communication network by identifying the pathways that contributed most to the incoming and outgoing signals for specific cell groups, performing a network centrality analysis of

the identified pathways, and then visualizing the network centrality scores.

Mutation analysis

With the help of GenePattern (https://cloud.genepattern.org), the data of somatic mutations and copy number variations (CNVs) downloaded from TCGA database were analyzed by GISTIC 2.0 to assess the CNV events at the chromosomal arm level and the minimum common region between samples. Then, we use the maftools R package (Mayakonda et al., 2018) to visualize the aforementioned analysis results in the mutation annotation format, and the plotmafSummary function was used to plot the summary file.

Construction of the prognostic model and clinical statistical analysis for model evaluation

To identify genes associated with the prognosis of gastric cancer, we selected the DEGs between the EGC and IM groups as candidate genes and performed univariate Cox regression analysis (p value < 0.01). Forest plots were created (R package: ggplot2) to display each variable's p-value, hazard ratio (HR), and 95% confidence interval. We then performed least absolute shrinkage and selection operator (LASSO) regression analysis based on these prognostic genes to construct a prognostic model (R package: Glmnet, survival). The risk score was calculated by the formula:

$$Riskscore = \sum_{k=1}^{n} coef\,(k) * x\,(k).$$

Here, the $coef\ (k)$ represents the LASSO-Cox regression coefficient, n represents the number of genes, and $x(k)$ represents the expression value of each gene. The TCGA GC cohort was divided into low- and high-risk groups based on the median risk score, and then Kaplan–Meier analysis (R package: survival and survminer) was performed to analyze and compare the overall survival (OS) between the two subgroups. The time-dependent receiver operating characteristic (ROC) curve was adopted (R package: timeROC) to analyze the predictive accuracy and risk scores. Considering the clinical features, we also created a prognostic nomogram (R package: rms and survival) to predict the 1-, 3-, and 5-year overall survival.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) is a computational method that determines whether a pre-defined set of genes shows statistically significant differences between two biological states and is usually performed to estimate the changes in pathways and biological process activities in gene expression dataset samples. To assess the differences in biological processes between the high- and low-risk groups based on the gene expression profiling datasets, we downloaded the reference gene sets (c2. cp.v7.2. symbols.gmt and c5. all.v7.2. symbols.gmt) from the MSigDB and used the GSEA method included in the clusterProfiler R package (Yu et al., 2012) to perform enrichment analysis of the datasets and visualize them. An adjusted p-value of <0.05 and a false discovery rate (FDR, q-value) < 0.25 were considered to suggest statistically significant differences.

Estimation of immune cell-type fractions

The CIBERSORT (Chen et al., 2018) computational method was adopted to quantitatively convert the transcriptomic data of tumor tissues into the absolute abundance of immune cells and stromal cells, to evaluate the changes in the proportion of 22 human immune cell subpopulations. For each tumor sample, the sum of all the estimated immune cell-type fractions equaled 1.

Results

Analysis of 25 cell clusters from biopsy specimens of the human gastric antral mucosa with scRNA-seq data revealed high levels of cellular heterogeneity

After filtration based on the quality control criteria and normalization of the scRNA-seq data, 48,566 cells were obtained (Supplementary Figure S1A). We selected 2,000 highly variable genes for downstream analysis and tagged the top ten (Figure 1B), such as *IGLL5*, *LIPF*, *TPSAB1*, and *APOA1*. PCA was performed to identify available dimensions and relevant genes, and 20 principal components (PC) were selected for subsequent analysis (Figure 1C). The tSNE algorithm was successfully applied to divide human cells of the gastric antral mucosa into 25 independent clusters (Figure 1D), which were identified by the marker genes for each cell type (Supplementary Figures S1B,C). Table 1 and Supplementary Figure S1D showed the numbers and percentages of each cell type identified with the 25 clusters. Specifically, gastric epithelial, metaplastic stem-like, and enterocytes accounted for 55.53%, 15.59%, and 11%, of the total cell count, respectively. We then identified DEGs between cell types (Supplementary Figure S1E) and used the top two genes with the most significant differential expressions to draw dot plots (Figure 1E). We also calculated the number of each cell type and the proportion of cells per cell type in the samples of the different disease groups (Figures 1F,G). The aforementioned

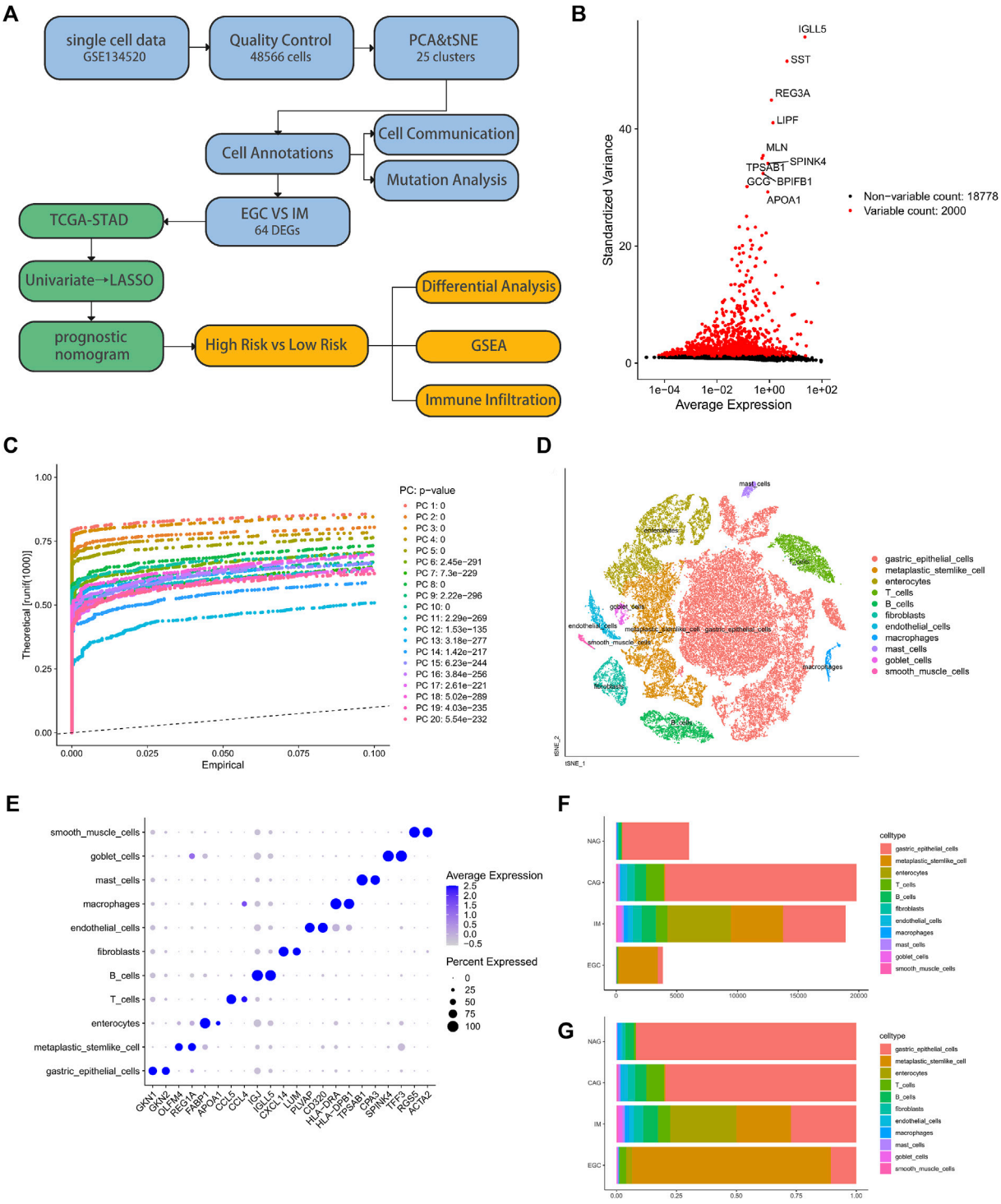


FIGURE 1
Analysis of the biopsy specimens from the gastric antral mucosa of patients with non-atrophic gastritis (NAG), chronic atrophic gastritis (CAG), intestinal metaplasia (IM), or early gastric cancer (EGC) based on single-cell RNA-seq data. **(A)** Flow chart of the study. **(B)** Scatter plot of standard deviation demonstrated the significantly differentially expressed genes between the cell types. **(C)** JackStrawPlot of 20 principal components (*p* value < 0.01) used to find clusters. **(D)** Cluster analysis based on the distribution of different cell types. **(E)** Prominent marker genes for each cell type. **(F)** and **(G)** Number of each cell type and the proportion of cells per cell type in the samples of the four different disease groups.

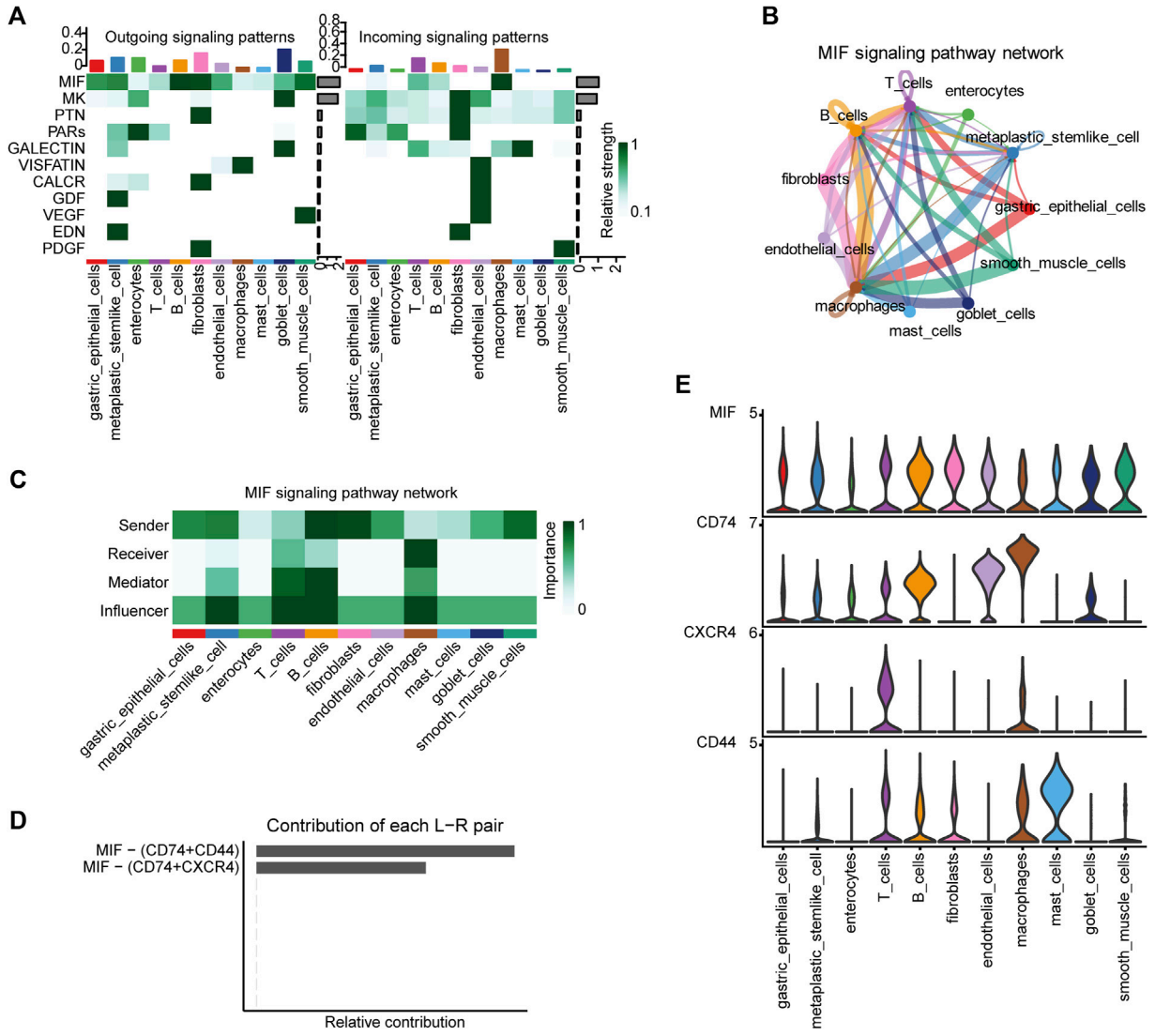


FIGURE 2
CellChat analyses of the intercellular communication network for the different cell types. **(A)** Contribution of the 11 pathways identified to the outgoing (left) and incoming (right) signals among the different cell types. **(B)** Intensity of cell–cell interactions in the MIF signaling pathway. **(C)** Network centrality scores of the MIF signaling pathway for each cell type. **(D)** Relative contribution of each ligand–receptor pair included in the MIF signaling pathway and **(E)** their expression patterns in the different cell types.

results portray the diverse landscape of the microenvironment between tumor and non-tumor samples.

Intercellular communication displayed locoregional immunomodulation in the carcinogenic process

We detected a total of 11 signaling pathways in the 11 cell types annotated in the single-cell data using CellChat, including MIF, MK, PTN, PARs, and GALECTIN. Heat maps were generated to illustrate the contribution of each pathway to the incoming or outgoing signals among the cell types (Figure 2A, Supplementary Figure S2). We selected the MIF signaling pathway as it contributed the most amongst the cell types, used a circle plot to illustrate the intensity of cell–cell interactions (Figure 2B) and performed a network centrality analysis (Figure 2C) for this pathway. Clearly, the macrophage is the most dominant recipient; B cell, T cell, and metaplastic stem-like cell perform key roles in the network. Moreover, we demonstrated the expression patterns of all ligand-receptor pairs included in the MIF signaling pathway (Figure 2D,E; Supplementary Figure S3) in different cell types: CD74 is mainly expressed in macrophages, endothelial cells, and

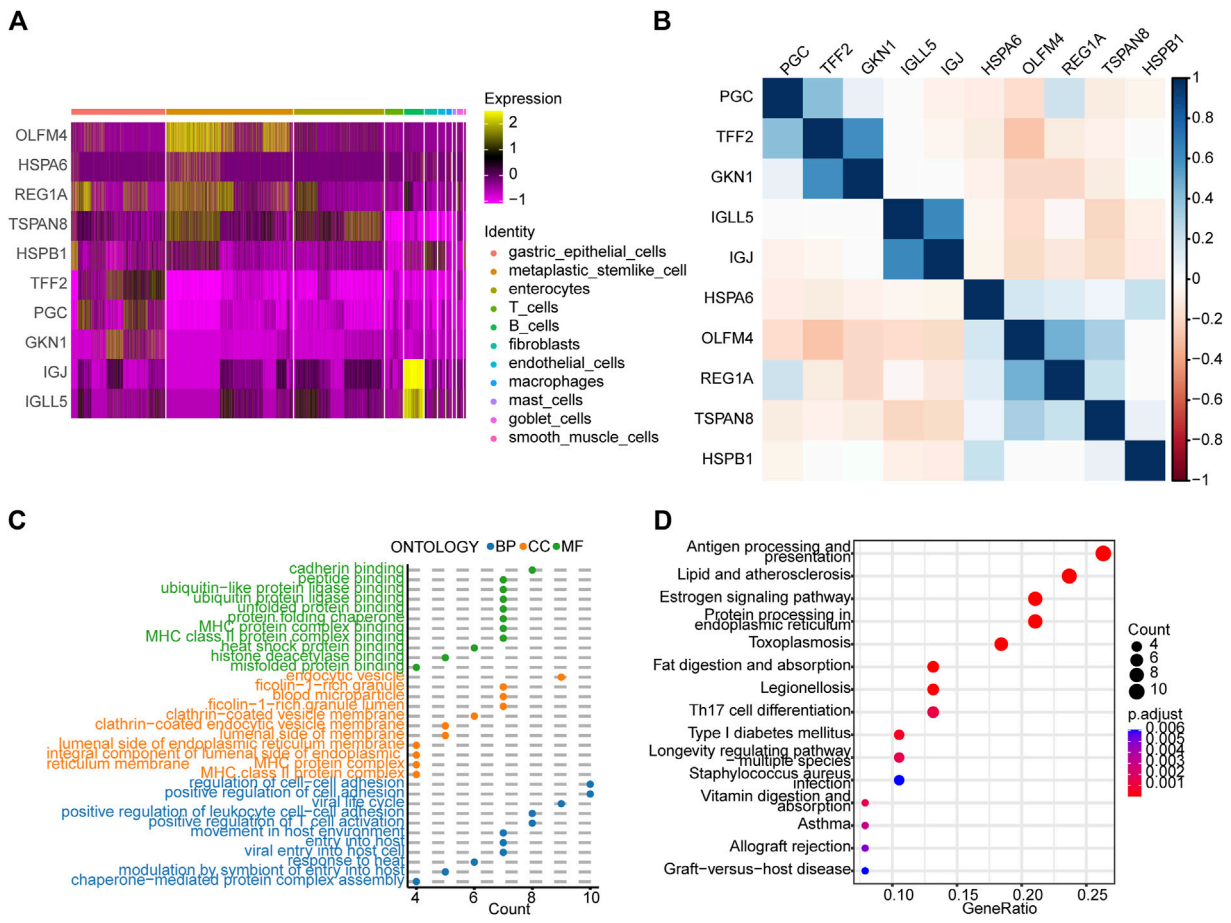


FIGURE 3
Analysis of differences between the early gastric cancer (EGC) and intestinal metaplasia (IM) groups and subsequent enrichment analyses. **(A)** Heat map showing the expression patterns of ten differentially expressed genes (DEGs) between the EGC and IM groups in the different cell types. **(B)** Heat map of the correlation between the DEGs. Dot plots showing the **(C)** GO and **(D)** KEGG pathway enrichment analyses of the DEGs (p-value < 0.01). BP, biological process; CC, cellular component; MF, molecular function.

B cells; CD44 is predominantly expressed in mast cells; and CXCR4 is primarily expressed in T cells.

Identification and enrichment analyses of differentially expressed genes between the early gastric cancer and IM groups

By analyzing the differences between the EGC and IM groups, we identified 64 DEGs in total (p-value < 0.05 and |log2FoldChange| > 1). Heat maps (Supplementary Figure S4) were drawn to visualize the DEGs, and a correlation analysis of these genes was performed (Supplementary Figure S5). As shown in Figures 3A,B, *GKN1* and *TFF2* are highly correlated in gastric epithelial cells; *IGJ* and *IGLL5* are highly correlated in B cells; and *OLFM4, REG1A*, and *TSPAN8* are highly correlated in gastric epithelial cells. GO and KEGG pathway enrichment analyses of the DEGs (adj. p value < 0.05) were visualized using dot plots (Figures 3C,D). These results prove that cell communication pathways (cadherin binding, adj. $p = 2.5 \times 10^{-4}$; ficolin−1−rich granule, adj. $p = 8.6 \times 10^{-6}$; regulation of cell−cell adhesion, adj. $p = 7.0 \times 10^{-5}$; and estrogen signaling pathways) and inflammatory response pathways (MHC protein complex, adj. $p = 1.6 \times 10^{-9}$; endocytic vesicle, adj. $p = 2.6 \times 10^{-5}$; antigen processing and presentation, and Th17 cell differentiation) were enriched.

Incidence of somatic mutations and copy number variations in 64 differentially expressed genes

We next examined the mutational landscape of 64 DEGs in TCGA patients with STAD. We found that the missense

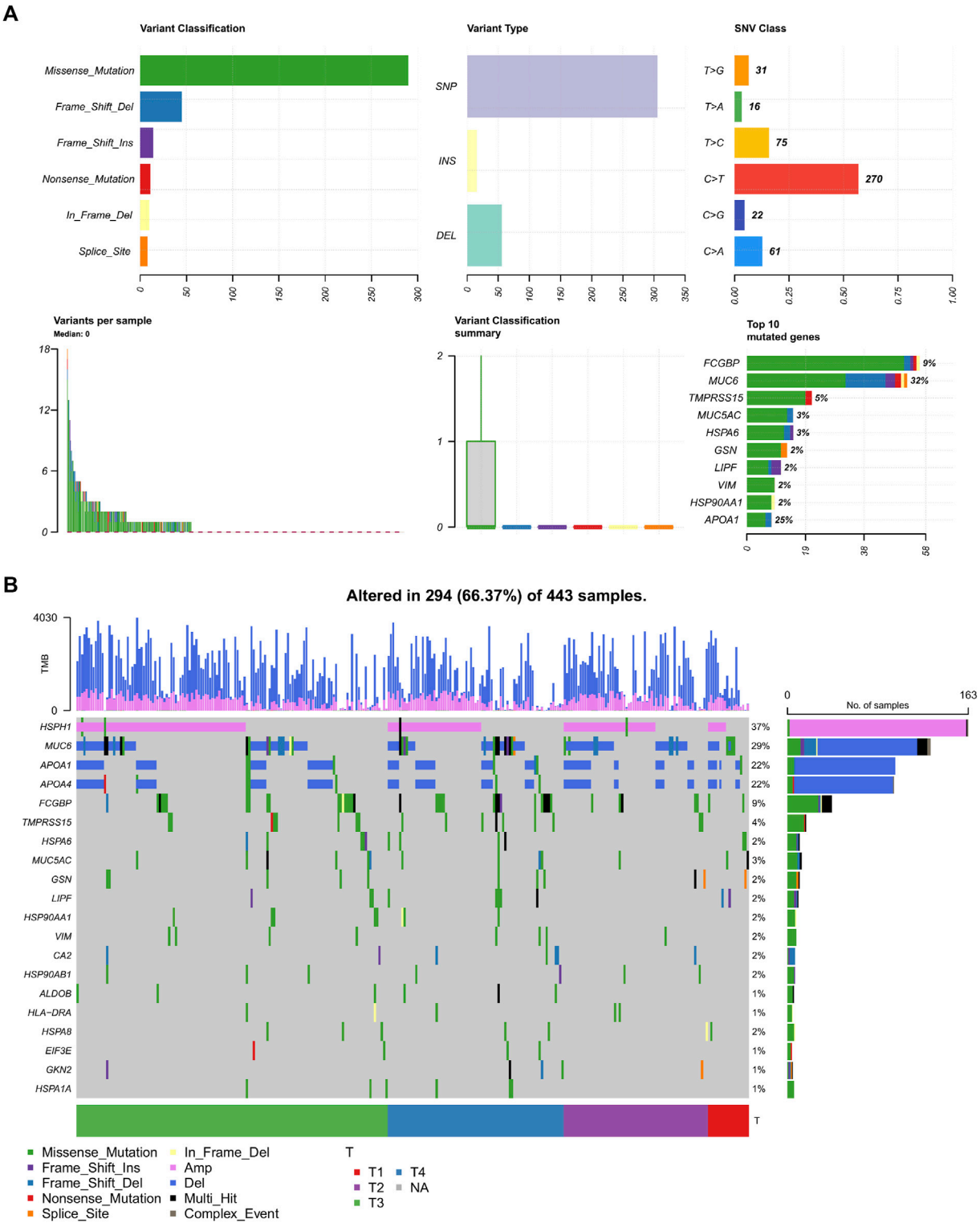


FIGURE 4
Mutational landscape in 64 DEGs. **(A)** Overview of the mutational frequency. **(B)** Mutation waterfall plot in the different T stages. SNP, single-nucleotide polymorphism; DEL, deletion; INS, insertion; SNV, single-nucleotide variant; Amp, amplification.

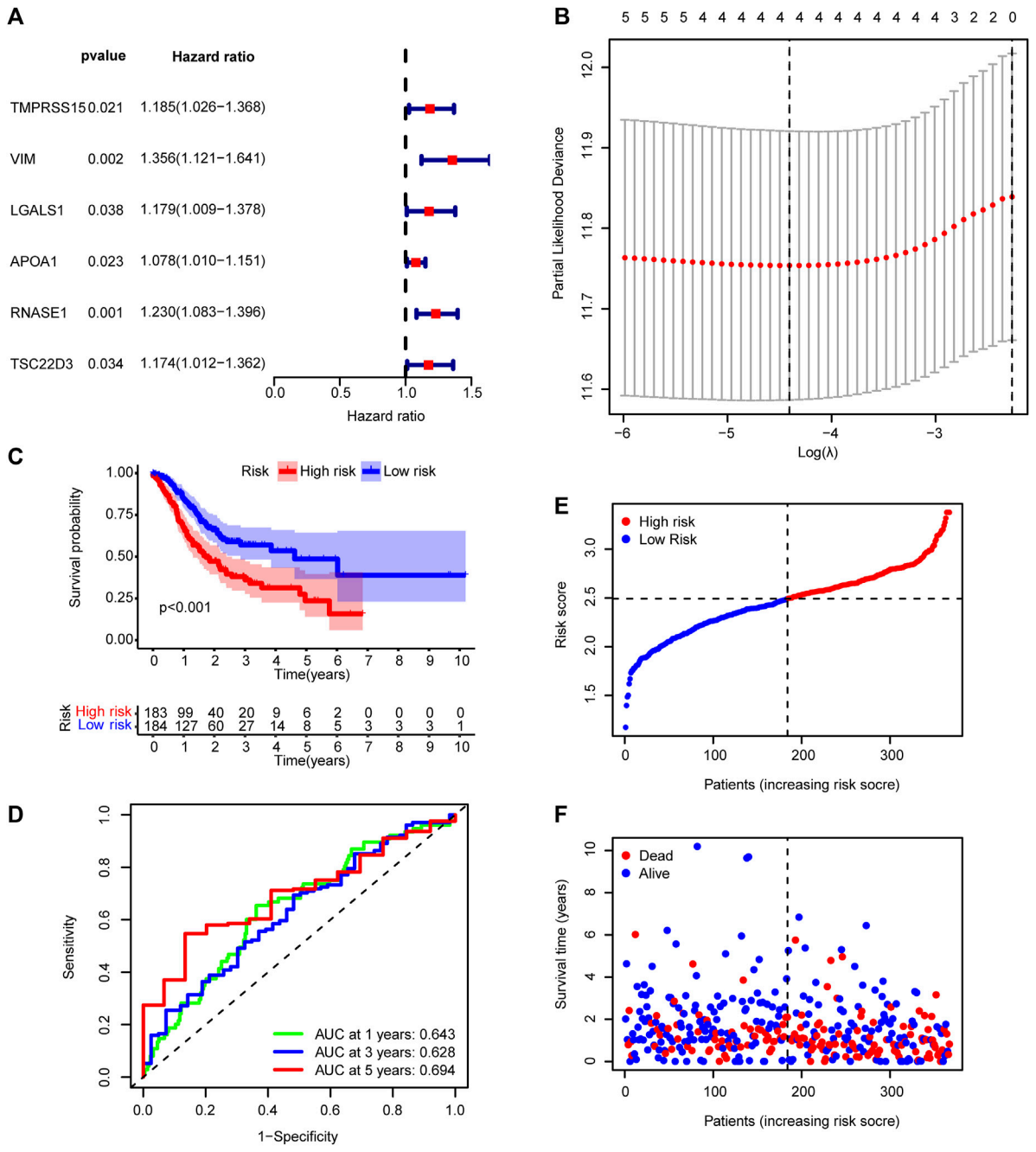


FIGURE 5
Prognostic model based on differentially expressed genes between the early gastric cancer and intestinal metaplasia groups. **(A)** Univariate Cox regression analysis showing the correlation between six genes and disease prognosis. **(B)** Diagram of error rates by 10-fold cross-validation. **(C)** Kaplan–Meier survival analysis for the high- and low-risk groups in TCGA cohort. **(D)** Time-dependent ROC curves showing the predictive accuracy of the prognostic model in TCGA cohort. **(E)** Risk score distribution and **(F)** survival status for patients in the TCGA dataset. AUC, area under the curve.

mutations have the highest mutation rate and the single-nucleotide polymorphism (SNP) is the most occurred variant type (Figure 4A). Among these SNPs, the highest proportion of mutations occurred in the C to T transition mutations. *FCGBP* and *MUC6* are dominant in the top 10 mutated genes, and *TMPRSS15*, *VIM*, and *APOA1* are also included. Figure 4B showed that the *HSPH1* amplification (37%) and *MUC6* (29%) and *APOA1* (22%) frameshift deletions are the most

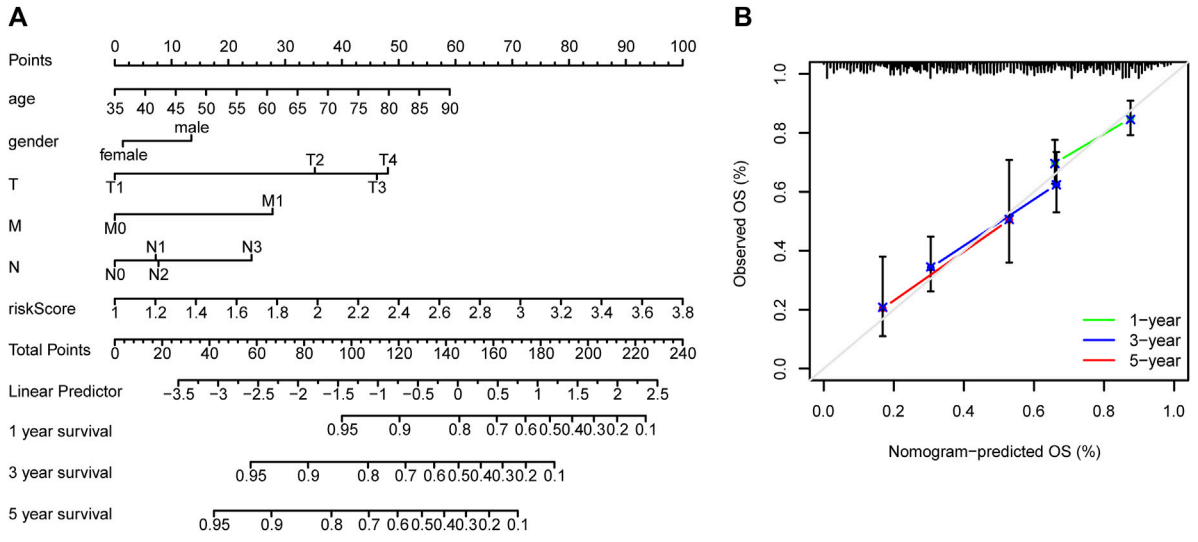


FIGURE 6
Prediction nomograms. **(A)** Nomogram model based on clinicopathological features and risk scores of patients with gastric cancer in TCGA dataset to predict their prognosis. **(B)** Calibration curve of the nomograms for predicting overall patient survival. The diagonal dotted line represents the ideal nomogram.

common alterations observed. The mutagenesis in these TME-related DEGs is illustrated as the pro-carcinogenic potential.

Construction of a TME-related prognostic model based on prognostic genes

To evaluate the correlation between the prognosis for patients with GC and the expression of the 64 DEGs, we used univariate Cox regression analysis. The results suggested that six of the genes (*TMPRSS15*, *VIM*, *LGALS1*, *APOA1*, *RNASE1*, and *TSC22D3*) are significantly correlated with the disease prognosis ($p < 0.05$) (Figure 5A). In addition, the LASSO-Cox regression algorithm was used to establish a prognostic model, and the results suggested that four genes (*TMPRSS15*, *VIM*, *APOA1*, and *RNASE1*) are highly correlated with disease prognosis (Figure 5B). Risk score = (*TMPRSS15* * 0.076) + (*VIM* * 0.225) + (*APOA1* * 0.066) + (*RNASE1* * 0.135). The Kaplan–Meier plotter is an online tool to find survival biomarkers in GC based on the meta-analysis of GEO, TCGA, and European Genome-phenome Archive databases (Szász et al., 2016). Using the Kaplan–Meier plotter that has 875 enrolled patients with GC (Supplementary Figure S6), we further validated that these four genes have diagnostic value in association with adverse outcomes. We also verified their protein expression in the Human Protein Atlas database (Supplementary Figure S7), which is corroborated by the transcript levels in Supplementary Figure S4.

TCGA cohort was then divided into low- and high-risk groups based on the median risk score. The Kaplan–Meier curve suggested a lower overall survival rate for patients with GC with high-risk scores than those with low-risk scores (Figure 5C). The area under the curve was more significant than 0.6 for the 1-, 3-, and 5-year survival ROC curves, which indicated that risk scores accurately predicted patient survival rates (Figure 5D). The dataset's risk score distribution and survival status are shown in Figures 4E,F, respectively. An increase in risk scores was accompanied by an increased risk of patient mortality and a shorter survival time (Figures 5E,F). Given the impact of clinicopathological features (including, age, gender, and TNM staging information) on prognosis, a prognostic nomogram was constructed to predict the survival of patients with GC in TCGA dataset (Figure 6A). The C-index of our nomogram model was 0.683, and the calibration plot (Figure 6B) showed a good fit with actual survival outcomes.

GSEA

GSEA was performed on all genes to analyze inter-group differences between the low- and high-risk groups. Figure 7 shows the nine most important functions or pathways based on normalized enrichment scores, such as GOBP: B cell receptor signaling pathway, GOCC: immunoglobulin complex circulating, and GOMF: antigen binding. As expected, increased risk scores are substantially associated with high-level immune response.

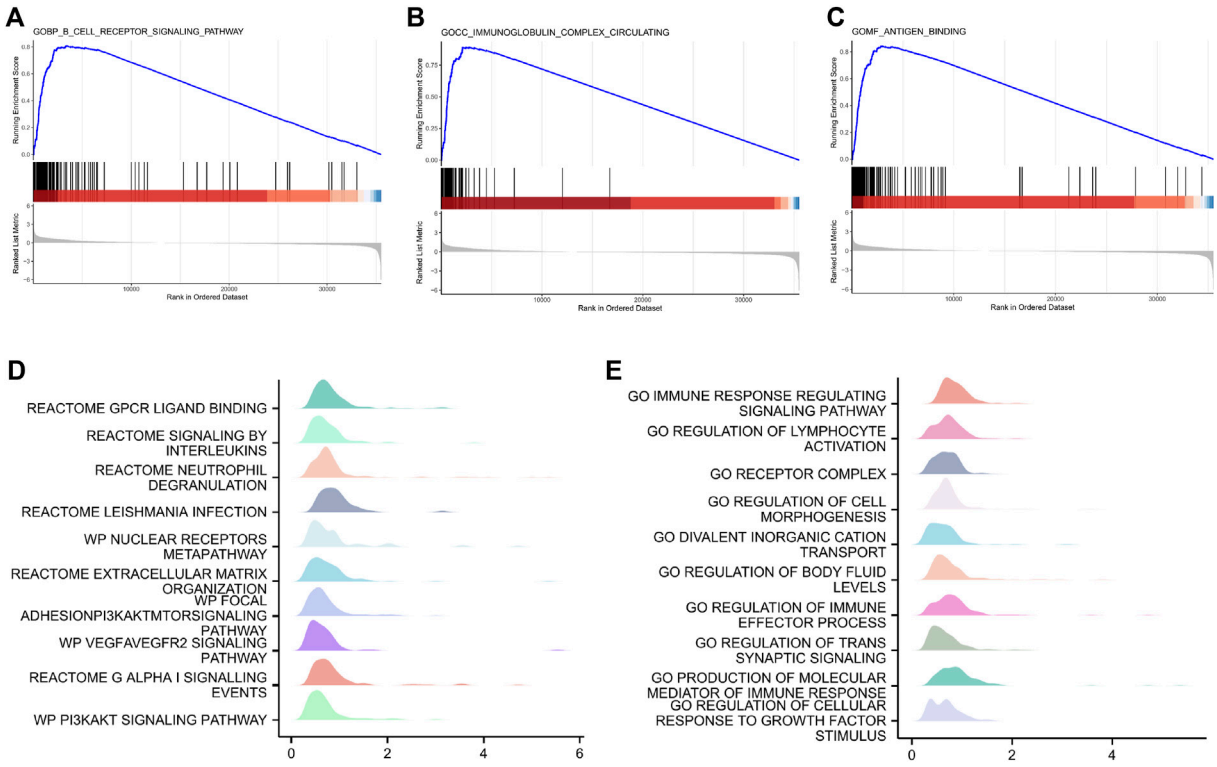


FIGURE 7
Gene set enrichment analysis showing the nine most important functions or pathways in low- and high-risk patients with gastric cancer in TCGA dataset. **(A)** GOBP: B-cell receptor signaling pathway. **(B)** GOCC: immunoglobulin complex. **(C)** GOMF: antigen binding. **(D,E)** GSEA mountain plot of representative enrichments in C2. all and C5. all MSigDB datasets. Significance was set at $p < 0.01$. GOBP, Gene Ontology biological process; GOCC, Gene Ontology cellular component; GOMF, Gene Ontology molecular function; NES, normalized enrichment scores.

Immune cell infiltration

Based on the previous results, we found that the survival status of the high-risk group was significantly worse than that of the low-risk group and hence speculated that there might exist differences in immune cell infiltration between the low- and the high-risk groups. The CIBERSORT algorithm was used to calculate immune cell infiltration in GC. The bar graph of immune cell infiltration and the boxplot comparing immune cells between the low- and high-risk groups are shown in Figures 8A,B, respectively. We found differences between the two groups in M0 and M2 macrophages, resting mast cells, monocytes, resting NK cells, and CD8 and follicular helper T cells. Interestingly, the expression levels of M0 macrophages, resting NK cells, and follicular helper T cells increased in the low-risk group, which expressed a relatively mild immune response. In contrast, the expression levels of tumor-associated immune cells increased in the high-risk group such as M2 macrophages, resting mast cells, monocytes, and CD8 T cells. Considering that the RNA-seq data of TCGA referred to the expression levels of tissue blocks instead of blood samples, these tumor-immune infiltrates often indicate higher malignancy and worse prognosis rather than protective immunity.

Figure 9 shows the correlations between the expression of prognostic genes (*TMPRSS15*, *VIM*, *APOA1*, and *RNASE1*) and immune cell infiltration. *APOA1* expression was positively correlated with the monocyte infiltration level; *RNASE1* expression was positively correlated with the infiltration levels of M2 macrophages and CD8 T cells, and negatively correlated with the infiltration levels of M0 macrophages, resting NK cells, and follicular helper T cells; *TMPRSS15* expression was negatively correlated with the M0 macrophages infiltration level; and *VIM* expression was positively correlated with the infiltration levels of M2 macrophages and resting mast cells, and negatively correlated with the follicular helper T cells infiltration level. This evidence suggested that the prognostic genes for GC are significantly associated with tumor immune infiltration.

Discussion

Among the numerous advances in the treatment of GC over the last decades, anti-PD-1/PD-L1 immunotherapy has raised

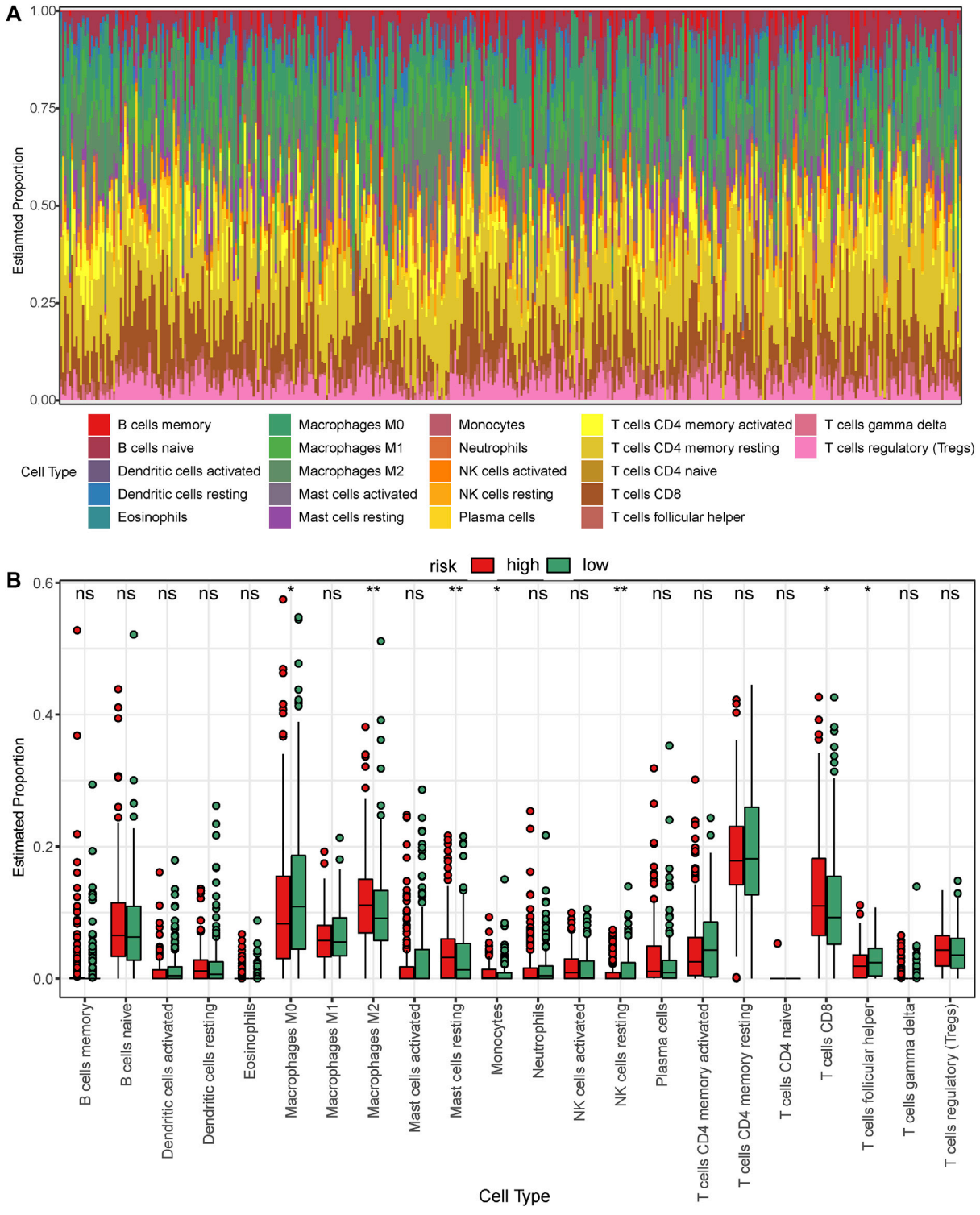


FIGURE 8
Analysis of immune cell infiltration in the low- and high-risk patients with gastric cancer in TCGA dataset. **(A)** Bar graph showing the proportion of each of the 22 immune cell types in the samples. **(B)** Boxplot comparing the proportion of each of the immune cell types between the low- and high-risk groups. The green and red blocks represent the high- and low-risk groups, respectively.

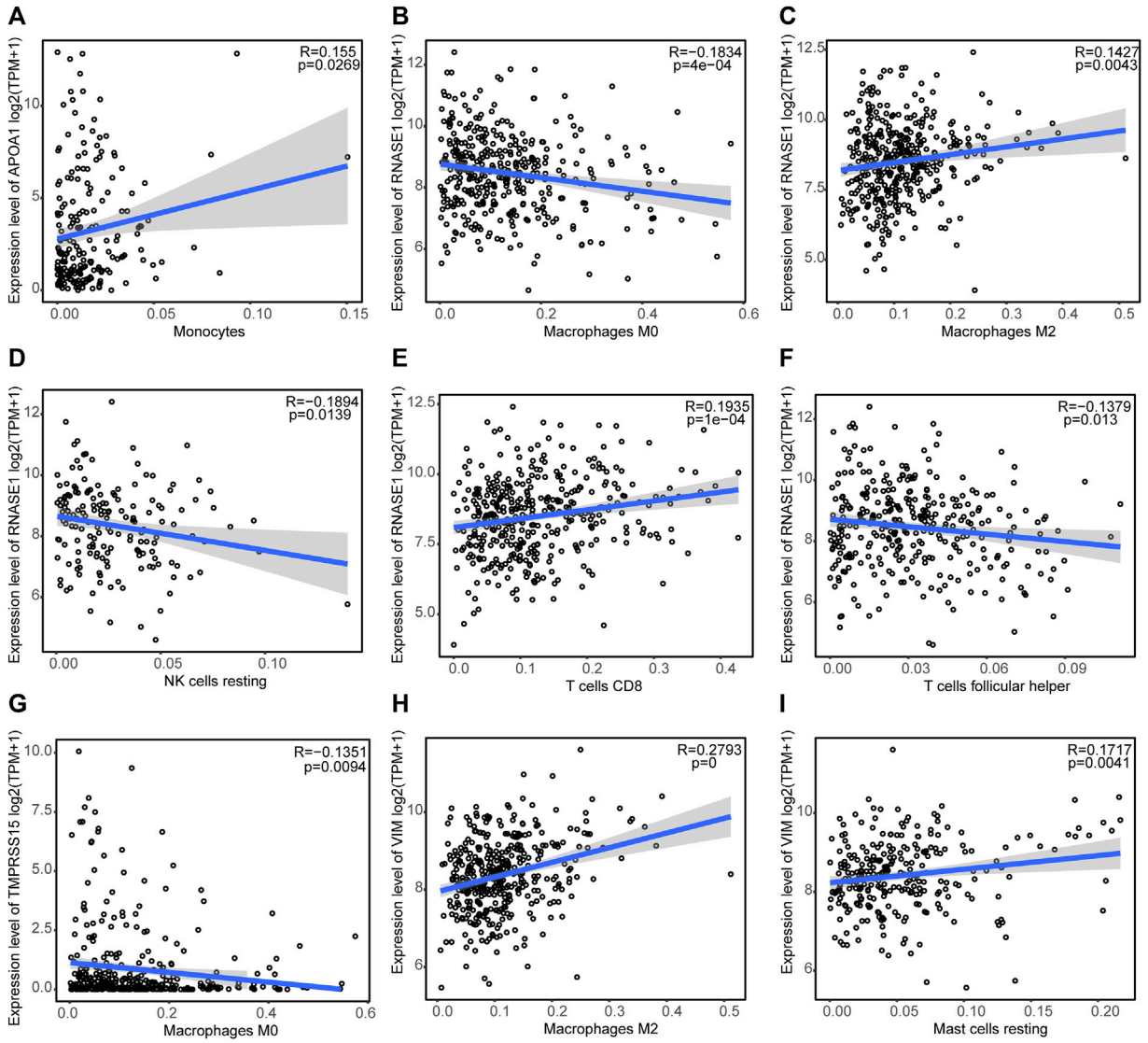


FIGURE 9
Correlations between the expression of prognostic genes and the levels of immune cell infiltration **(A)** Correlation between *APOA1* expression and the monocyte infiltration level. Correlations between *RNASE1* expression and the infiltration levels of **(B)** M0 macrophages, **(C)** M2 macrophages, **(D)** resting NK cells, **(E)** CD8 T cells, and **(F)** follicular helper T cells. **(G)** Correlation between *TMPRSS15* expression and the M0 macrophage infiltration level. Correlations between *VIM* expression and the infiltration levels of **(H)** M2 macrophages and **(I)** resting mast cells.

particular concerns (Joshi and Badgwell, 2021). TME in GC may influence the response of immunotherapy and affect prognosis (Moutafi and Rimm, 2021). As transcriptomics technologies continue to advance, scRNA-seq offers an excellent tool to research the immune microenvironment biomarkers with clinical translational potential.

The first step toward single-cell analysis is to cluster cell populations with characteristic genes, which is the basis for mapping cell interactions. In addition to the common methods used in this study (Andrews and Hemberg, 2018), a new ensemble random projection-based algorithm, SHARP, which can cluster 10 million cells in large-scale scRNA-seq data, has recently appeared (Wan et al., 2020). CNVs analysis is usually applied for the identification of malignant tumor cells (Patel et al., 2014), but it has unsatisfactory effectiveness in GC (Zhang M. et al., 2021). Our research avoided the controversial topic of separating benign and malignant epithelial cells and instead categorized the component of TME. We found significant differences in the percentage of infiltrating immune cells between early GC and non-cancerous samples and between the low- and high-risk groups in the TCGA. These findings are consistent with a previous comparison of samples before and after cisplatin

chemotherapy (Kim et al., 2021). To better clarify the humoral and cellular immune responses in GC, Sathe et al. (2020) concurrently sequenced matched peripheral blood mononuclear cells and revealed the immune remodeling of NK cells, dendritic cells, cytotoxic T cells, and plasma cells. Considering the loss of spatial dimension in cell isolation from tissues, the conjoint analysis of scRNA-seq and spatial transcriptomics in pancreatic cancer (Moncada et al., 2020) and squamous cell carcinoma (Ji et al., 2020) has been proposed, giving us a new perspective on tumor-infiltrating immune cells.

Some differentially expressed non-marker genes confirmed clustering results from another perspective. Overexpression of *GKN1* in epithelial cells (Yoon et al., 2013), *IGLL5* in B cells (Zheng et al., 2021), *TPSAB1* in mast cells (Konnikova et al., 2021), and *RGS5* in smooth muscle cells (Silini et al., 2012) are consistent with previous studies, while *CXCL14* in fibroblasts (Wu et al., 2020) suggests that the conclusion from bulk RNA-seq may not be precise or versatile enough. Together with marker genes, these genes constitute 64 DEGs between IM and EGC, which reflect transcriptional heterogeneity among different cell types. In addition, 9 of 64 DEGs are members of the heat-shock protein family involved in cellular stress responses and pro-tumor inflammation as molecular chaperones. For example, Hsp72 has been previously reported to promote the oxaliplatin resistance of GC cells by inhibiting SDF-2 degradation (Takahashi et al., 2016); Hsp90ab1 is known to facilitate the epithelial–mesenchymal transition in GC *by* preventing LRP5 ubiquitination (Wang et al., 2019). Both genes are highly expressed by enterocytes and can be perceived as indicators of carcinogenicity for extensive intestinal metaplasia.

Macrophage migration inhibitory factor (MIF) is a multipotent cytokine involved in both inflammatory processes and anti-tumor immune response, having the properties of an enzyme, chemokine, and hormone simultaneously (Sumaiya et al., 2021). As a proinflammatory mediator secreted by numerous immune cells, MIF promotes inflammation and autoimmune diseases mainly by binding with the receptor CD74 and co-receptor CD44, CXCR4, or CXCR2 (Liehn et al., 2013; Nakamura et al., 2021; Wallace et al., 2021). Previous studies have shown that the carcinogenic role MIF plays is related to the activation of p53, Ras/MAPK, and Akt pathways in the inflammation–cancer axis (Mittelbronn et al., 2011; Wang et al., 2017; Liu et al., 2018). This study observed the prominent MIF signaling pathway network in GC at the single-cell level. Comparisons and enrichment analysis for DEGs in IM and EGC supported the enriched pathway associated with cell communication and inflammatory response. Mechanistically, MIF-pathway overactivation is likely a manifestation of either the macrophage aggregation caused by *Helicobacter pylori* infection or accumulations of tumor-associated macrophages, which further predisposes epithelial cells to malignant transformation. Under various environmental stimuli, macrophages developing from differentiated monocytes are classified into classically activated (M1) or alternatively activated (M2) macrophages. Generally, IL-12 synthesized and secreted by M1 can induce the proliferation and differentiation of naïve T cells into Th1 cells and enhance the NK cell-mediated antitumor effect; IL-10 secreted by M2 has the opposite effect (Gambardella et al., 2020). Our prognostic model shows that the expression of pro-tumorigenic M2 in the high-risk group is increased, and prognostic genes are also positively correlated with M2 infiltration. A reasonable explanation is that MIF facilitates the M2 polarization of macrophages in GC (Huang et al., 2019; Chen et al., 2021). Recently scholars (Zhou et al., 2022) found that carfilzomib enables M2 to express M1 cytokines, showing great immunotherapeutic potential for solid tumors. For future applications, experimental validation and subgroup analysis of macrophages are required.

Physician-scientists have long grappled with the quantitative assessment of immune TME (Zhang Z. et al., 2021). In advanced GC, Zeng et al. (2021) built a TME scoring system for the effects of checkpoint immunotherapy. By implementing machine learning, Cai et al. (2020) established a prognostic classification model based on the expression of *VCAN*, *CLIP4*, and *MATN3* in a total of 1699 GC patients. In our study, four high-risk genes derived from scRNA-seq show the prognostication of poor outcomes. Vimentin (VIM), a type III intermediate filament protein, which characterizes the stromal component of TME in solid cancers, is often regarded as the epithelial–mesenchymal transition marker. It is closely related to GC cell invasion and metastasis, and is the most commonly used marker to detect the acquisition of these mesenchymal traits (Zeng et al., 2018; Zhu et al., 2019). Apolipoprotein A1 (APOA1), a component of high-density lipoprotein, exerts a favorable effect on the prevention of many cardiovascular diseases. However, its elevated level in the urine of bladder cancer patients (Peng et al., 2019) indicated shorter survival. For patients with colorectal cancer, increased APOA1 expression in the blood is concomitant with $CD3^+$ T cells aggregation in the core of the tumor as well as the invasive margin (Guo et al., 2019). *TMPRSS15,* known as an enzyme gene, is translated as a serine protease in enterocytes and goblet cells—enteropeptidase. The link between *TMPRSS15* and tumors has rarely been reported before, but other genes coding for the same family proteins were proven to be promising therapeutic targets. *TMPRSS*2 was found to favor the immune escape of COVID-19 (Shang et al., 2020), and *TMPRSS*4 was upregulated in malignancies of the stomach (Tazawa et al., 2022), liver (Dong et al., 2020), and prostate

(Lee et al., 2021). In the present study, we speculate that *TMPRSS15* and *APOA1* used for enterocyte identification might be a cellular trait of intestinal type-GC. Therefore, the high expression of these two genes heralded extensive metaplasia and worse outcome. Similarly, despite the little evidence for a direct relationship between RNase1 release in endothelial cells (Bedenbender and Schmeck, 2020) and cancer progression, *RNASE1* shows a good prediction performance based on raw data in TCGA (Li et al., 2020; Yang et al., 2022).

In conclusion, we analyzed the previously published data of scRNA-seq as well as TCGA, explored the inflammation–immunity–cancer axis, and developed a tumor microenvironment-associated risk score in GC. They well reflect the immune infiltration level and help construct a prognostic nomogram model that can assess overall long-term outcomes. Nevertheless, further in-depth studies are needed to confirm our results and broaden our view of TME in GC.

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

Obtained funding and provided study concept: JC; conceived and designed the study: XW and JC; analyzed the data and drafted the manuscript: XW and JL; validated the data and conclusion: ZH and XC; contributed to the literature search: KW. All authors revised the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (No. 81871979), the Natural Science Foundation of Fujian Province (No. 2021J02056), and the Medical and Health Sciences Foundation of Xiamen (Nos. 3502Z20199171 and 3502Z20204002).

Acknowledgments

The authors acknowledge TCGA database and contributors on the GEO database for providing 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.2022.896064/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Analysis of the biopsy specimens from the gastric antral mucosa of patients with non-atrophic gastritis (NAG), chronic atrophic gastritis (CAG), intestinal metaplasia (IM), or early gastric cancer (EGC) based on single-cell RNA-seq data. **(A)** Graphs showing the 48,566 cells included in this study after quality control. **(B,C)** Cluster analysis based on the disease group and cluster grouping. **(D)**Pie chart of cell types. **(E)** tSNE projections of DEGs in different cell types.

SUPPLEMENTARY FIGURE S2
CellChat analyses of the intercellular communication network for different cell types. **(A)** Chord diagram showing the input and output signal pathways in different cell types. **(B)** and **(C)** Number and intensity of cell–cell interaction signal pathways.

SUPPLEMENTARY FIGURE S3
Relative contribution of each ligand–receptor pair in different cell types.

SUPPLEMENTARY FIGURE S4
Complete heat map showing the expression patterns of ten differentially expressed genes (DEGs) between the EGC and IM groups in the different cell types.

SUPPLEMENTARY FIGURE S5
Complete heat map of the correlation between the DEGs.

SUPPLEMENTARY FIGURE S6
Kaplan–Meier (KM) survival curves of four genes in the Kaplan–Meier Plotter.

SUPPLEMENTARY FIGURE S7
Protein expression levels of four genes in HPA. **(A)** Vimentin is highly expressed in stromal cells of GC. **(B)** Enrichment of *APOA1* in tissue-specific cell types. **(C)** *TMPRSS15* is mainly expressed in proximal enterocytes. **(D)** Protein expression of *RNASE1* in the gastrointestinal tract.

References

Andrews, T. S., and Hemberg, M. (2018). Identifying cell populations with scRNASeq. *Mol. Asp. Med.* 59, 114–122. doi:10.1016/j.mam.2017.07.002

Bedenbender, K., and Schmeck, B. T. (2020). Endothelial ribonuclease 1 in cardiovascular and systemic inflammation. *Front. Cell Dev. Biol.* 8, 576491. doi:10.3389/fcell.2020.576491

Bijlsma, M., Sadanandam, A., Tan, P., and Vermeulen, L. (2017). Molecular subtypes in cancers of the gastrointestinal tract. *Nat. Rev. Gastroenterol. Hepatol.* 14 (6), 333–342. doi:10.1038/nrgastro.2017.33

Cai, W. Y., Dong, Z. N., Fu, X. T., Lin, L. Y., Wang, L., Ye, G. D., et al. (2020). Identification of a tumor microenvironment-relevant gene set-based prognostic signature and related therapy targets in gastric cancer. *Theranostics* 10 (19), 8633–8647. doi:10.7150/thno.47938

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

Chen, J., Wang, H., Jia, L., He, J., Li, Y., Liu, H., et al. (2021). Bufalin targets the SRC-3/MIF pathway in chemoresistant cells to regulate M2 macrophage polarization in colorectal cancer. *Cancer Lett.* 513, 63–74. doi:10.1016/j.canlet.2021.05.008

Dong, Z. R., Sun, D., Yang, Y. F., Zhou, W., Wu, R., Wang, X. W., et al. (2020). TMPRSS4 drives angiogenesis in hepatocellular carcinoma by promoting HB-egf expression and proteolytic cleavage. *Hepatology* 72 (3), 923–939. doi:10.1002/hep.31076

Gambardella, V., Castillo, J., Tarazona, N., Gimeno-Valiente, F., Martinez-Ciarpaglini, C., Cabeza-Segura, M., et al. (2020). The role of tumor-associated macrophages in gastric cancer development and their potential as a therapeutic target. *Cancer Treat. Rev.* 86, 102015. doi:10.1016/j.ctrv.2020.102015

Guo, G., Wang, Y., Zhou, Y., Quan, Q., Zhang, Y., Wang, H., et al. (2019). Immune cell concentrations among the primary tumor microenvironment in colorectal cancer patients predicted by clinicopathologic characteristics and blood indexes. *J. Immunother. Cancer* 7 (1), 179. doi:10.1186/s40425-019-0656-3

Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discov.* 12 (1), 31–46. doi:10.1158/2159-8290.CD-21-1059

Huang, W. C., Kuo, K. T., Wang, C. H., Yeh, C. T., and Wang, Y. (2019). Cisplatin resistant lung cancer cells promoted M2 polarization of tumor-associated macrophages via the Src/CD155/MIF functional pathway. *J. Exp. Clin. Cancer Res.* 38 (1), 180. doi:10.1186/s13046-019-1166-3

Jeong, H. Y., Ham, I. H., Lee, S. H., Ryu, D., Son, S. Y., Han, S. U., et al. (2021). Spatially distinct reprogramming of the tumor microenvironment based on tumor invasion in diffuse-type gastric cancers. *Clin. Cancer Res.* 27 (23), 6529–6542. doi:10.1158/1078-0432.CCR-21-0792

Ji, A. L., Rubin, A. J., Thrane, K., Jiang, S., Reynolds, D. L., Meyers, R. M., et al. (2020). Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* 182 (2), 1661–1662. e422. doi:10.1016/j.cell.2020.08.043

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

Joshi, S. S., and Badgwell, B. D. (2021). Current treatment and recent progress in gastric cancer. *Ca. Cancer J. Clin.* 71 (3), 264–279. doi:10.3322/caac.21657

Kim, R., An, M., Lee, H., Mehta, A., Heo, Y. J., Kim, K. M., et al. (2021). Early tumor–immune microenvironmental remodeling and response to first-line fluoropyrimidine and platinum chemotherapy in advanced gastric cancer. *Cancer Discov.* 12, 984–1001. doi:10.1158/2159-8290.CD-21-0888

Konnikova, L., Robinson, T. O., Owings, A. H., Shirley, J. F., Davis, E., Tang, Y., et al. (2021). Small intestinal immunopathology and GI-associated antibody formation in hereditary alpha-tryptasemia. *J. Allergy Clin. Immunol.* 148 (3), 813–821.e7. e817. doi:10.1016/j.jaci.2021.04.004

Kumar, V., Ramnarayanan, K., Sundar, R., Padmanabhan, N., Srivastava, S., Koiwa, M., et al. (2021). Single-cell Atlas of lineage states, tumor microenvironment, and subtype-specific expression programs in gastric cancer. *Cancer Discov.* 12, 670–691. doi:10.1158/2159-8290.CD-21-0683

Lee, Y., Yoon, J., Ko, D., Yu, M., Lee, S., Kim, S., et al. (2021). TMPRSS4 promotes cancer stem-like properties in prostate cancer cells through upregulation of SOX2 by SLUG and TWIST1. *J. Exp. Clin. Cancer Res.* 40 (1), 372. doi:10.1186/s13046-021-02147-7

Li, S., Jiang, L., Tang, J., Gao, N., and Guo, F. (2020). Kernel fusion method for detecting cancer subtypes via selecting relevant expression data. *Front. Genet.* 11, 979. doi:10.3389/fgene.2020.00979

Liehn, E., Kanzler, I., Konschalla, S., Kroh, A., Simsekyilmaz, S., Sönmez, T., et al. (2013). Compartmentalized protective and detrimental effects of endogenous macrophage migration-inhibitory factor mediated by CXCR2 in a mouse model of myocardial ischemia/reperfusion. *Arterioscler. Thromb. Vasc. Biol.* 33 (9), 2180–2186. doi:10.1161/atvbaha.113.301633

Liu, R., Sun, D., Jiao, Y., Wang, P., Zhang, J., Wang, M., et al. (2018). Macrophage migration inhibitory factor promotes tumor aggressiveness of esophageal squamous cell carcinoma via activation of Akt and inactivation of GSK3β. *Cancer Lett.* 412, 289–296. doi:10.1016/j.canlet.2017.10.018

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

Mittelbronn, M., Platten, M., Zeiner, P., Dombrowski, Y., Frank, B., Zachskorn, C., et al. (2011). Macrophage migration inhibitory factor (MIF) expression in human malignant gliomas contributes to immune escape and tumour progression. *Acta Neuropathol.* 122 (3), 353–365. doi:10.1007/s00401-011-0858-3

Moncada, R., Barkley, D., Wagner, F., Chiodin, M., Devlin, J. C., Baron, M., et al. (2020). Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat. Biotechnol.* 38 (3), 333–342. doi:10.1038/s41587-019-0392-8

Moutafi, M., and Rimm, D. (2021). Putting the microenvironment into the immunotherapy companion diagnostic. *Clin. Cancer Res.* 27 (14), 3812–3814. doi:10.1158/1078-0432.ccr-21-1238

Muller, S., and Diaz, A. (2017). Single-cell mRNA sequencing in cancer research: Integrating the genomic fingerprint. *Front. Genet.* 8, 73. doi:10.3389/fgene.2017.00073

Mundekkad, D., and Cho, W. C. (2022). Nanoparticles in clinical translation for cancer therapy. *Int. J. Mol. Sci.* 23 (3), 1685. doi:10.3390/ijms23031685

Nakamura, A., Zeng, F., Nakamura, S., Reid, K., Gracey, E., Lim, M., et al. (2021). Macrophage migration inhibitory factor drives pathology in a mouse model of spondyloarthritis and is associated with human disease. *Sci. Transl. Med.* 13 (616), eabg1210. doi:10.1126/scitranslmed.abg1210

Patel, A. P., Tirosh, I., Trombetta, J. J., Shalek, A. K., Gillespie, S. M., Wakimoto, H., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344 (6190), 1396–1401. doi:10.1126/science.1254257

Peng, C., Hua, M. Y., Li, N. S., Hsu, Y. P., Chen, Y. T., Chuang, C. K., et al. (2019). A colorimetric immunosensor based on self-linkable dual-nanozyme for ultrasensitive bladder cancer diagnosis and prognosis monitoring. *Biosens. Bioelectron.* 126, 581–589. doi:10.1016/j.bios.2018.11.022

Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S., Sandberg, R., et al. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* 9 (1), 171–181. doi:10.1038/nprot.2014.006

Ramskold, D., Luo, S., Wang, Y. C., Li, R., Deng, Q., Faridani, O. R., et al. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* 30 (8), 777–782. doi:10.1038/nbt.2282

Sathe, A., Grimes, S. M., Lau, B. T., Chen, J., Suarez, C., Huang, R. J., et al. (2020). Single-cell genomic characterization reveals the cellular reprogramming of the gastric tumor microenvironment. *Clin. Cancer Res.* 26 (11), 2640–2653. doi:10.1158/1078-0432.CCR-19-3231

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* 33 (5), 495–502. doi:10.1038/nbt.3192

Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., et al. (2020). Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. U. S. A.* 117 (21), 11727–11734. doi:10.1073/pnas.2003138117

Silini, A., Ghilardi, C., Figini, S., Sangalli, F., Fruscio, R., Dahse, R., et al. (2012). Regulator of G-protein signaling 5 (RGS5) protein: A novel marker of cancer vasculature elicited and sustained by the tumor's proangiogenic microenvironment. *Cell. Mol. Life Sci.* 69 (7), 1167–1178. doi:10.1007/s00018-011-0862-8

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

Sumaiya, K., Langford, D., Natarajaseenivasan, K., and Shanmughapriya, S. (2021). Macrophage migration inhibitory factor (MIF): A multifaceted cytokine regulated by genetic and physiological strategies. *Pharmacol. Ther.* 233, 108024. doi:10.1016/j.pharmthera.2021.108024

Szász, A., Lánczky, A., Nagy, Á., Förster, S., Hark, K., Green, J., et al. (2016). Cross-validation of survival associated biomarkers in gastric cancer using transcriptomic data of 1, 065 patients. *Oncotarget* 7 (31), 49322–49333. doi:10.18632/oncotarget.10337

Takahashi, K., Tanaka, M., Yashiro, M., Matsumoto, M., Ohtsuka, A., Nakayama, K. I., et al. (2016). Protection of stromal cell-derived factor 2 by heat shock protein 72 prevents oxaliplatin-induced cell death in oxaliplatin-resistant human gastric cancer cells. *Cancer Lett.* 378 (1), 8–15. doi:10.1016/j.canlet.2016.05.002

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* 6 (5), 377–382. doi:10.1038/nmeth.1315

Tazawa, H., Suzuki, T., Saito, A., Ishikawa, A., Komo, T., Sada, H., et al. (2022). Utility of TMPRSS4 as a prognostic biomarker and potential therapeutic target in patients with gastric cancer. *J. Gastrointest. Surg.* 26 (2), 305–313. doi:10.1007/s11605-021-05101-2

Wallace, D., Figueras, F., Wegener, W., and Goldenberg, D. (2021). Experience with milatuzumab, an anti-CD74 antibody against immunomodulatory macrophage migration inhibitory factor (MIF) receptor, for systemic lupus erythematosus (SLE). *Ann. Rheum. Dis.* 80 (7), 954–955. doi:10.1136/annrheumdis-2020-219803

Wan, S., Kim, J., and Won, K. J. (2020). SHARP: Hyperfast and accurate processing of single-cell RNA-seq data via ensemble random projection. *Genome Res.* 30 (2), 205–213. doi:10.1101/gr.254557.119

Wang, B., Zhang, Y., Qing, T., Xing, K., Li, J., Zhen, T., et al. (2021). Comprehensive analysis of metastatic gastric cancer tumour cells using single-cell RNA-seq. *Sci. Rep.* 11 (1), 1141. doi:10.1038/s41598-020-80881-2

Wang, C., Zhou, X., Li, W., Li, M., Tu, T., Ba, X., et al. (2017). Macrophage migration inhibitory factor promotes osteosarcoma growth and lung metastasis through activating the RAS/MAPK pathway. *Cancer Lett.* 403, 271–279. doi:10.1016/j.canlet.2017.06.011

Wang, H., Deng, G., Ai, M., Xu, Z., Mou, T., Yu, J., et al. (2019). Hsp90ab1 stabilizes LRP5 to promote epithelial-mesenchymal transition via activating of AKT and Wnt/β-catenin signaling pathways in gastric cancer progression. *Oncogene* 38 (9), 1489–1507. doi:10.1038/s41388-018-0532-5

Wang, R., Dang, M., Harada, K., Han, G., Wang, F., Pool Pizzi, M., et al. (2021). Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nat. Med.* 27 (1), 141–151. doi:10.1038/s41591-020-1125-8

Wang, R., Zheng, X., Wang, J., Wan, S., Song, F., Wong, M. H., et al. (2022). Improving bulk RNA-seq classification by transferring gene signature from single cells in acute myeloid leukemia. *Brief. Bioinform.* 23 (2), bbac002. doi:10.1093/bib/bbac002

Wu, D., Li, Z., Zhao, S., Yang, B., and Liu, Z. (2020). Downregulated microRNA-150 upregulates IRX1 to depress proliferation, migration, and invasion, but boost apoptosis of gastric cancer cells. *IUBMB Life* 72 (3), 476–491. doi:10.1002/iub.2214

Yang, H., Tian, W., and Zhou, B. (2022). Sarcopenia and a 5-mRNA risk module as a combined factor to predict prognosis for patients with stomach adenocarcinoma. *Genomics* 114 (1), 361–377. doi:10.1016/j.ygeno.2021.12.011

Yoon, J. H., Choi, Y. J., Choi, W. S., Ashktorab, H., Smoot, D. T., Nam, S. W., et al. (2013). GKN1-miR-185-DNMT1 axis suppresses gastric carcinogenesis through regulation of epigenetic alteration and cell cycle. *Clin. Cancer Res.* 19 (17), 4599–4610. doi:10.1158/1078-0432.CCR-12-3675

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

Zeng, D., Wu, J., Luo, H., Li, Y., Xiao, J., Peng, J., et al. (2021). Tumor microenvironment evaluation promotes precise checkpoint immunotherapy of advanced gastric cancer. *J. Immunother. Cancer* 9 (8), e002467. doi:10.1136/jitc-2021-002467

Zeng, S., Xie, X., Xiao, Y. F., Tang, B., Hu, C. J., Wang, S. M., et al. (2018). Long noncoding RNA LINC00675 enhances phosphorylation of vimentin on Ser83 to suppress gastric cancer progression. *Cancer Lett.* 412, 179–187. doi:10.1016/j.canlet.2017.10.026

Zeng, Y., and Jin, R. (2021). Molecular pathogenesis, targeted therapies, and future perspectives for gastric cancer. *Semin. Cancer Biol* 2021, 1044–579X. doi:10.1016/j.semcancer.2021.12.004

Zhang, M., Hu, S., Min, M., Ni, Y., Lu, Z., Sun, X., et al. (2021). Dissecting transcriptional heterogeneity in primary gastric adenocarcinoma by single cell RNA sequencing. *Gut* 70 (3), 464–475. doi:10.1136/gutjnl-2019-320368

Zhang, P., Yang, M., Zhang, Y., Xiao, S., Lai, X., Tan, A., et al. (2019). Dissecting the single-cell transcriptome network underlying gastric premalignant lesions and early gastric cancer. *Cell Rep.* 27 (6), 1934–1947. doi:10.1016/j.celrep.2019.04.052

Zhang, Z., Bao, S., Yan, C., Hou, P., Zhou, M., Sun, J., et al. (2021). Computational principles and practice for decoding immune contexture in the tumor microenvironment. *Brief. Bioinform.* 22 (3), bbaa075. doi:10.1093/bib/bbaa075

Zhao, L., Lee, V. H. F., Ng, M. K., Yan, H., and Bijlsma, M. F. (2019). Molecular subtyping of cancer: Current status and moving toward clinical applications. *Brief. Bioinform.* 20 (2), 572–584. doi:10.1093/bib/bby026

Zheng, B., Xie, F., Cheng, F., Wang, J., Yao, Z., He, W., et al. (2021). Integrative analysis of immune-related genes in the tumor microenvironment of renal clear cell carcinoma and renal papillary cell carcinoma. *Front. Mol. Biosci.* 8, 760031. doi:10.3389/fmolb.2021.760031

Zhou, Q., Liang, J., Yang, T., Liu, J., Li, B., Li, Y., et al. (2022). Carfilzomib modulates tumor microenvironment to potentiate immune checkpoint therapy for cancer. *EMBO Mol. Med.* 14 (1), e14502. doi:10.15252/emmm.202114502

Zhu, Z., Rong, Z., Luo, Z., Yu, Z., Zhang, J., Qiu, Z., et al. (2019). Circular RNA circNHSL1 promotes gastric cancer progression through the miR-1306-3p/SIX1/vimentin axis. *Mol. Cancer* 18 (1), 126. doi:10.1186/s12943-019-1054-7

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 17 August 2022
DOI 10.3389/fgene.2022.918486

Check for updates

OPEN ACCESS

EDITED BY
Lan Zhao,
Stanford University, United States

REVIEWED BY
Jiao Hu,
Central South University, China
Hualin Song,
Tianjin Medical University Cancer Institute and Hospital, China

*CORRESPONDENCE
Yuping Wang,
wangyuping@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 12 April 2022
ACCEPTED 06 July 2022
PUBLISHED 17 August 2022

CITATION
Zhang H, Huang T, Ren X, Fang X, Chen X, Wei H, Sun W and Wang Y (2022), Integrated pan-cancer analysis of CSMD2 as a potential prognostic, diagnostic, and immune biomarker. *Front. Genet.* 13:918486. doi: 10.3389/fgene.2022.918486

COPYRIGHT
© 2022 Zhang, Huang, Ren, Fang, Chen, Wei, Sun 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.

Integrated pan-cancer analysis of CSMD2 as a potential prognostic, diagnostic, and immune biomarker

Huiyun Zhang[1,2,3†], Taobi Huang[1,2,3†], Xiangqing Ren[1,2,3], Xidong Fang[1,2,3], Xia Chen[1,2,3], Hui Wei[1,2,3], Weiming Sun[4] and Yuping Wang[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, [4]Department of Endocrinology, The First Hospital of Lanzhou University, Lanzhou, China

The protein encoded by CUB and Sushi Multiple Domains 2 (CSMD2) is likely involved in regulating the complement cascade reaction of the immune system. However, current scientific evidence on the comprehensive roles of CSMD2 in pan-cancer is relatively scarce. Therefore, in this study, we explored the transcriptional level of CSMD2 in pan-caner using TCGA, GEO, and International Cancer Genome Consortium databases. Receiver operating characteristic curve analysis was used to investigate the diagnostic efficacy of CSMD2. The Kaplan-Meier Plotter and Oncolnc were used to investigate the correlation between CSMD2 expression and prognosis. Additionally, we analyzed the correlation between epigenetic methylation and CSMD2 expression in various cancers based on UALCAN, as well as, the correlation between CSMD2 and tumor mutational burden (TMB), microsatellite instability (MSI), and tumor neoantigen burden (TNB) in tumors. TIMER2.0 database was employed to investigate the correlation between CSMD2 and immune cells in the tumor microenvironment and immune checkpoints. Based on TISIDB, the correlation between CSMD2 and MHC molecules and immunostimulators was analyzed. Ultimately, we observed with a pan-cancer analysis that CSMD2 was upregulated in most tumors and had moderate to high diagnostic efficiency, and that high expression was closely associated with poor prognosis in patients with tumors. Moreover, hypermethylation of CSMD2 promoter and high levels of m6A methylation regulators were also observed in most cancers. CSMD2 expression was negatively correlated with TMB and MSI in stomach adenocarcinoma (STAD) and stomach and esophageal carcinoma (STES), as well as with tumor mutational burden, microsatellite instability, and TNB in head-neck squamous cell carcinoma (HNSC). In most cancers, CSMD2 might be associated with immune evasion or immunosuppression, as deficient anti-tumor immunity and upregulation of immune checkpoints were also observed in this study. In conclusion, CSMD2 could serve as a promising prognostic, diagnostic and immune biomarker in pan-cancer.

KEYWORDS

CSMD2, pan-cancer, prognosis, immune, microenvironment, methylation

1 Introduction

CUB and Sushi Multiple Domains 2 (CSMD2), located on the short arm of human chromosome 1 (1p35.1), is mainly expressed in the brain and gall bladder. Fourteen CUB domains at the N-terminal of CSMD2 are separated by a single complement control protein (CCP) domain, followed by 13 series of CCP domains. CCP is also called short consensus repeats (SCR) or Sushi domain, notably, multiple consecutive CCP domains are common characteristics of many complement inhibitors containing such domains (Rossi et al., 1998; Gialeli et al., 2018). CUB and sushi domains are considered as sites of protein-protein or protein-ligand interactions, indicating that CSMD proteins are either transmembrane receptors or adhesion proteins (Lau and Scholnick, 2003). CSMD2 has been associated with schizophrenia (Håvik et al., 2011). It is downregulated and is associated with the poor prognosis in colorectal cancer (Zhang and Song, 2014). However, current studies on the role of CSMD2 in tumorigenesis and the development of other tumors are lacking. In addition, genetic alterations in CSMD2 have been detected in primary lymphoma of the central nervous system and colorectal cancer and have been found to been associated with prognosis (Vater et al., 2015; Yang *p.*-S. et al., 2018). As a complement system regulator and receptor, its immune role in tumorigenesis and development remains unclear.

Cancer imposes a major burden on human society and was either the first or second leading cause of death before the age of 70 in 112 of 183 countries (Sung et al., 2021). In 2020, there were approximately 19.3 million new cases and 10 million cancer-related deaths worldwide (Sung et al., 2021). Therefore, early diagnosis and effective treatment are critical.

In this study, the expression of CSMD2 was examined and its diagnostic efficacy and prognostic value in pan-cancer were explored. In addition, this study investigated the association between CSMD2 expression and anti-tumor immunity and immune evasion in the tumor microenvironment, and the relationship with immune checkpoints, MHC molecules, and immunostimulators, which clarified the role of CSMD2 in suppressing anti-tumor immunity. Finally, epigenetic methylation analysis and functional enrichment analyses were performed, which provided ideas for further functional experiments.

2 Materials and methods

2.1 Differential CSMD2 expression analysis in cell lines, normal and tumor tissues

The RNAseq data of TCGA and GTEx was downloaded from UCSC XENA (https://xenabrowser.net/datapages/). The landscape of CSMD2 expression in 33 cancers and corresponding normal tissues were visualized using the ggplot2 package in the R version 4.0.3 program (The R Project for Statistical Computing). Wilcoxon rank sum test was used for analysis.

The Gene Expression Display Server (GEDS) (http://bioinfo.life.hust.edu.cn/web/GEDS/) was used to demonstrate the differential mRNA expression of CSMD2 in normal tissues. Data from Genotype-Tissue Expression (GTEx) was normalized by transcripts per million (TPM) (Xia et al., 2019). The data (log2 (TPM+1)) was from Cancer cell line encylopedia (CCLE) (https://depmap.org/portal/gene/CSMD2?tab=characterization) and visualized by R software for analyzing the expression of CSMD2 in cancer cell lines.

The transcriptome data used for subsequent analyses were exported and downloaded from the GEO database (http://www.ncbi.nih.gov/geo). The raw data were downloaded as MINiML files. Box plots are drawn by boxplot (Zhou et al., 2020). RNA sequencing expression (level 3) profiles and corresponding clinical information for live cancer (Japan) were obtained from the International Cancer Genome Consortium (ICGC) database (https://dcc.icgc.org/releases/current/Projects) (Zhang et al., 2019). Statistical analyses were performed using R software v4.0.3.

The human lung bronchial epithelial cell line BEAS-2B and human non-small cell lung cancer cell line A549 were purchased from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China) and Stem Cell Bank, Chinese Academy of Sciences, respectively. Total RNA from cells was extracted by using TRIzol (Thermo Fisher, Shanghai, China) RNA extraction protocol. Total RNA was reversely transcribed to cDNA using cDNA reverse transcription kits (TransGen Biotech, China). RT-qPCR was performed with TransStart® Top Green qPCR SuperMix (TransGen Biotech, China). GAPDH was used as the internal reference gene for normalization. The $2^{-\Delta\Delta Ct}$ method was used to analyze the qPCR results. The GraphPad Prism (version 8.0) was employed to visualize the relative gene expression levels in cell lines.

2.2 Exploring the diagnostic and prognostic potential of CSMD2

The RNAseq data and corresponding clinical data of 33 cancers were downloaded from the TCGA database (https://portal.gdc.cancer.gov/). Receiver operating characteristic (ROC) curves were used to evaluate the diagnostic efficacy of CSMD2. If there were no corresponding paracancerous data in the TCGA database, the tumor tissue data from TCGA and the corresponding normal tissue data from GTEx in UCSC XENA (https://xenabrowser.net/datapages/)

would be included. Statistical analysis and visualization were performed via the R software. The pROC package and ggplot2 packages were utilized to calculate the area under the curve (AUC) and visualize the ROC curve.

The closer the AUC is to 1, the better the diagnostic value. AUC between 0.5–0.7 indicates low accuracy, 0.7 to 0.9 indicates moderate accuracy, and greater than 0.9 indicates high accuracy. Univariate Cox regression analysis was performed using the survival package. The Kaplan-Meier Plotter (http://kmplot.com/analysis/) (Győrffy, 2021) was used to show statistically significant results. Oncolnc (http://www.oncolnc.org/) (Anaya, 2016) is a tool for exploring survival correlations, and the cut-off is 50%. The correlation between CSMD2 expression and the clinicopathological stage was visualized via the ggplot2 package.

2.3 cBioPortal

"TCGA Pan-Cancer Atlas Studies" in cBioportal (http://www.cbioportal.org) (Cerami et al., 2012; Gao et al., 2013) was employed to explore genetic alteration characteristics of CSMD2. The "Cancer Types Summary" module displayed CSMD2 alteration types and frequency in 32 cancer studies, and the "Mutations" module presented the mutation information of CSMD2. Kaplan-Meier plots with log-rank *p*-values were generated via the "Comparison" module, which can analyze the survival time of cancer patients with or without CSMD2 alterations.

2.4 SangerBox

Sangerbox 3.0 website (http://vip.sangerbox.com/home.html) is a visualization tool for bioinformatics analysis. The relationship between CSMD2 expression and tumor mutational burden (TMB), microsatellite instability (MSI), and tumor neoantigen burden (TNB) was analyzed. Their spearman's correlation was calculated in each type of tumor by the "single gene pan-cancer analysis" module in Sangerbox. CSMD2 expression data were obtained from the TCGA pan-cancer database (PANCAN, N= 10535). Simple Nucleotide variation data downloaded from GDC (https://portal.gdc.cancer.gov/) was used to calculate the TMB of samples via R-package "maftools" (version 2.8.05). MSI scores and TNB data for each tumor were obtained from previous studies (Bonneville et al., 2017; Thorsson et al., 2018). Samples with 0 expression levels were filtered, and each expression value was further transformed by log2 (x + 1). Finally, cancer types with fewer than three samples were also eliminated.

2.5 UALCAN

UALCAN (http://ualcan.path.uab.edu/analysis.html) (Chandrashekar et al., 2017) was used to analyze the promoter methylation level of CSMD2 between different cancers and the corresponding normal tissues of TCGA samples. The beta value indicated the level of DNA methylation, ranging from 0 (unmethylated) to 1 (fully methylated). Different beta value cut-off values are considered to indicate hypermethylation (beta value: 0.7–0.5) or hypomethylation (beta-value: 0.3–0.25). (Shinawi et al., 2013; Men et al., 2017).

2.6 TIMER2.0

TIMER2.0 (http://timer.comp-genomics.org/) (Li et al., 2020) is a comprehensive resource for investigating tumor immunological, clinical, and genomic features of tumors in TCGA. The "Gene_Corr" module was used to investigate the correlation between N6-methyladenosine (m6A) methylation regulators or CSMD2-related genes and CSMD2 expression level. The "Gene" module was used to analyze the relationship between CSMD2 expression and immune infiltration based on two kinds of immune deconvolutions, including CIBERSORT and XCELL. Since most immune cell types are negatively correlated with tumor purity, the "Purity Adjustment", which used the partial Spearman's correlation, was selected. The spearman's rho value indicates the degree of their correlation.

2.7 TISIDB

TISIDB (Ru et al., 2019) is an online tool for the network of tumor and immune system interactions. The data were obtained from the PubMed database, high-throughput screening data, exome and RNA sequencing dataset of patient cohorts with immunotherapy, and the TCGA database. The data of correlation between CSMD2 and immunostimulators, and MHC molecules were obtained from TISIDB, and visualized via R ("ggplot" package).

2.8 GEPIA2 and STRING

The "Similar Genes Detection" module of GEPIA2 (http://gepia2.cancer-pku.cn/#index) was used to obtain the top 100 CSMD2-related genes and the "Correlation Analysis" module to visualize the correlation between CSMD2 and the top 4 genes in 33 cancers. STRING (version 11.5) (https://www.string-db.org/) was used to predict proteins interacting proteins with CSMD2 and form protein-protein interaction (PPI) network. Functional enrichment analysis of CSMD2 was conducted using the clusterProfiler package, the org. Hs.eg.db package was used for gene ID conversion, and the ggplot2 package used for visualization.

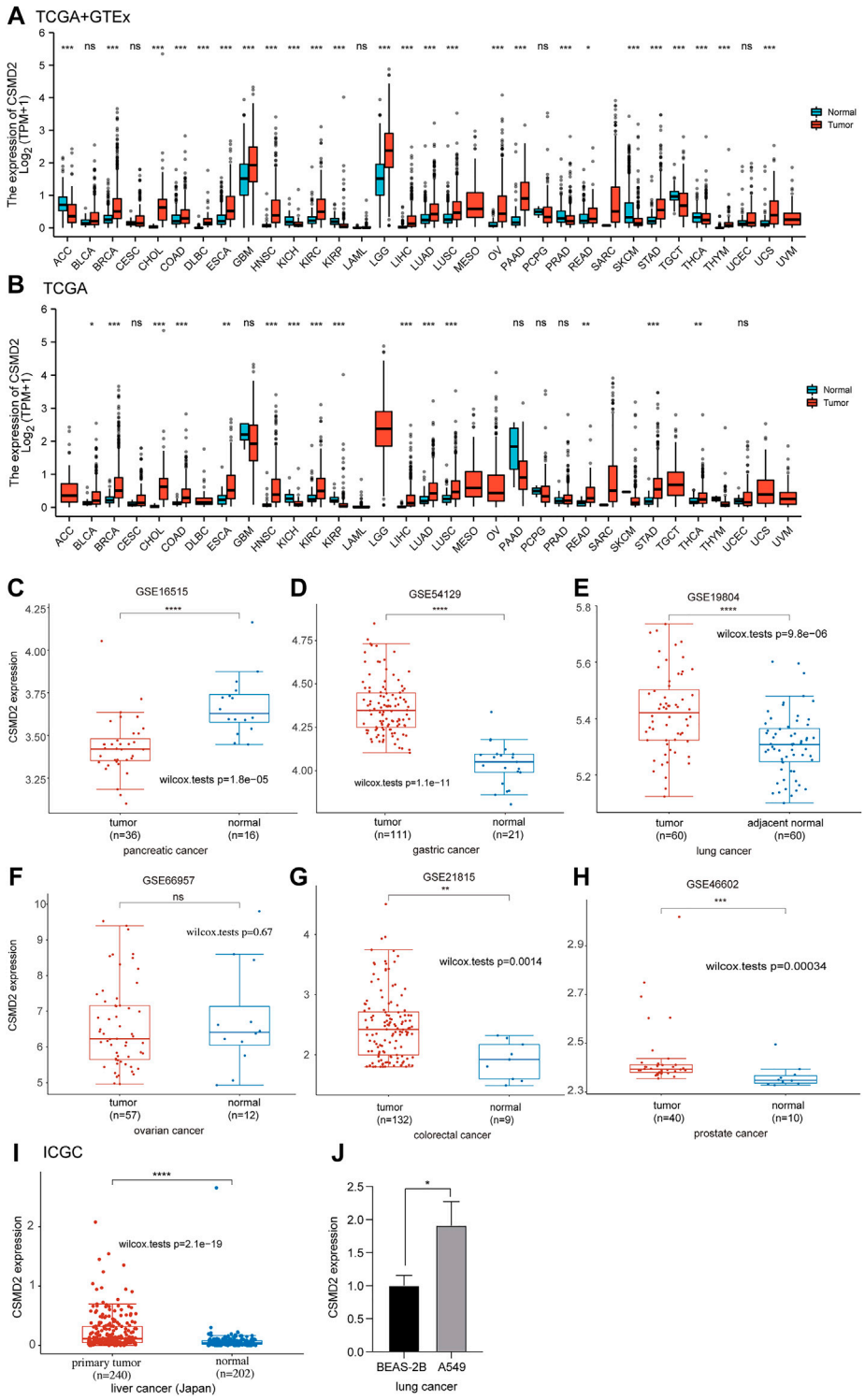


FIGURE 1
Differential mRNA expression of CSMD2. The boxplot shows CSMD2 expression in **(A,B)** 33 tumor types from TCGA and GTEx databases. The boxplot shows CSMD2 expression in pancreatic cancer **(C)**, gastric cancer **(D)**, lung cancer **(E)**, ovarian cancer **(F)**, colorectal cancer **(G)**, and prostate cancer **(H)** from GEO database. **(I)**The boxplot shows CSMD2 expression in live cancer (Japan) from ICGC databse. **(J)** The expression of CSMD2 in the human lung bronchinal epithelial cell line BEAS-2B and human non-small cell lung cancer cell line A549 (ns, $p \geq 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$).

3 Results

3.1 CSMD2 expression and clinical landscape in pan-cancer

As shown in Figure 1A, CSMD2 was differentially expressed in 25 of the 33 cancers (adrenocortical carcinoma [ACC], breast invasive carcinoma [BRCA], cholangiocarcinoma [CHOL], colon adenocarcinoma [COAD], lymphoid neoplasm diffuse large B-cell lymphoma [DLBC], esophageal carcinoma [ESCA], glioblastoma multiforme [GBM], head-neck squamous cell carcinoma [HNSC], kidney chromophobe [KICH], kidney renal 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], prostate adenocarcinoma [PRAD], rectum adenocarcinoma [READ], skin cutaneous melanoma [SKCM], stomach adenocarcinoma [STAD], testicular germ cell tumors [TGCT], thyroid carcinoma [THCA], thymoma [THYM], and uterine carcinosarcoma [UCS]). The expression levels of CSMD2 in the tumor tissues of BRCA, CHOL, COAD, DLBC, ESCA, GBM, HNSC, KIRC, LGG, LIHC, LUAD, LUSC, OV, PAAD, READ, STAD, THYM, and UCS were higher than those in normal tissues. Conversely, the expression levels of CSMD2 in the tumor tissues of ACC, KICH, KIRP, PRAD, SKCM, TGCT, and THCA were lower than normal tissues. Figure 1B illustrats the expression of *CSMD2* in 33 tumor types in TCGA database. By analyzing the GEO datasets, we found CSMD2 expressed highly in gastric cancer (p-value = 1.1e-11), lung cancer (p-value = 9.8e-06), colorectal cancer (p-value = 0.0014), and prostate cancer (p-value = 0.00034) (Figures 1C–H). As shown in Figure 1I, CSMD2 expressed highly in primary tumor in liver cancer. *In vitro* verification, we also found that CSMD2 expressed highly in human non-small cell lung cancer cell line A549 (Figure 1J). CSMD2 expression in the brain and central nervous system was highest in normal tissues and cancer cell lines (Supplementary Figures S4A,B).

Statistically significant differences in CSMD2 expression were observed based on different pathological stages in BLCA, ESCA, and KIRP, but not in other cancers (Supplementary Figure S1A). Furthermore, CSMD2 expression was significantly correlated with the T or N stage in ESCA, KIRP, and STAD (Supplementary Figure S1B).

We also explored the diagnostic value of CSMD2 as an independent biomarker for pan-cancers. ROC curves were used to evaluate the diagnostic sensitivity and specificity of CSMD2. ROC curves of CSMD2 expression in tumor and normal tissues showed that CSMD2 had high diagnostic efficiency in DLBC (AUC = 0.987, CI: 0.977–0.996), CHOL (AUC = 0.960, CI: 0.907–1.000), PAAD (AUC = 0.913, CI: 0.881–0.945), and THYM (AUC = 0.901, CI: 0.867–0.935), whereas, CSMD2 had low diagnostic efficiency in BLCA, CESC, GBM, LAML, PRAD, THCA, and UCEC, and had moderate diagnostic efficiency in other cancers (Figure 2).

Furthermore, as shown in the forest plots (Figure 3A), a negative association was observed between CSMD2 expression and overall survival (OS) in BLCA, KIRP, LIHC, STAD, and THYM. The results from the Kaplan–Meier plotter demonstrated that CSMD2 overexpression was significantly associated with poor prognosis in patients with BRCA, EAC, KIRP, LIHC, PAAD, SARC, STAD, and THYM patients (Figures 3B–I). Regarding CSMD2 and relapse-free survival (RFS), a significant negative association was found in patients with BRCA, ESCC, KIRP, PAAD, and STAD patients (Figures 3J–N). Patients with high expression of CSMD2 had poor survival in KIRP, LIHC, and STAD (Supplementary Figures S5A–C).

In summary, CSMD2 expression was upregulated in most tumors, with moderate to high diagnostic efficiency, and its high expression was associated with high stage and poor prognosis in tumor patients.

3.2 Genetic alteration characteristics of CSMD2

The frequency and types of genetic alterations of CSMD2 in 32 cancer studies were further investigated. Mutation was the most frequent alteration of CSMD2, whereas structural variation was less frequent (Figure 4A). We further analyzed the number, sites, types, and domains of the CSMD2 mutations. The percentage of samples with somatic mutations in CSMD2 was 6.7%. Missense mutations were the most frequent mutations in CSMD2. The site with the largest number of mutations was W1996*/R in the CUB domain, which was detected in six cases of SKCM and one case of LUSC (Figure 4B). Furthermore, we analyzed the potential relationship between genetic alterations in CSMD2 and the prognosis in patients with different cancer. The results showed that CSMD2 alteration was significantly associated with poor OS (p-value = 1.603e-3) and disease-specific survival (DSS) (p-value = 4.273e-3) in BRCA patients, favorable disease-free survival (DFS) (p-value = 0.0420) in OV patients, and favorable OS (p-value = 0.0101), progression-free survival (PFS) (p-value = 3.962e-3), and DSS (p-value = 0.0190) in UCEC patients (Figure 4C). In summary, the genetic alterations in CSMD2 are related to patient survival.

3.3 Correlation of CSMD2 expression with tumor mutational burden, microsatellite instability and tumor neoantigen burden

The correlations between CSMD2 expression and TMB, MSI, and TNB were explored and visualized using radar maps. Significant correlations between CSMD2 and TMB were

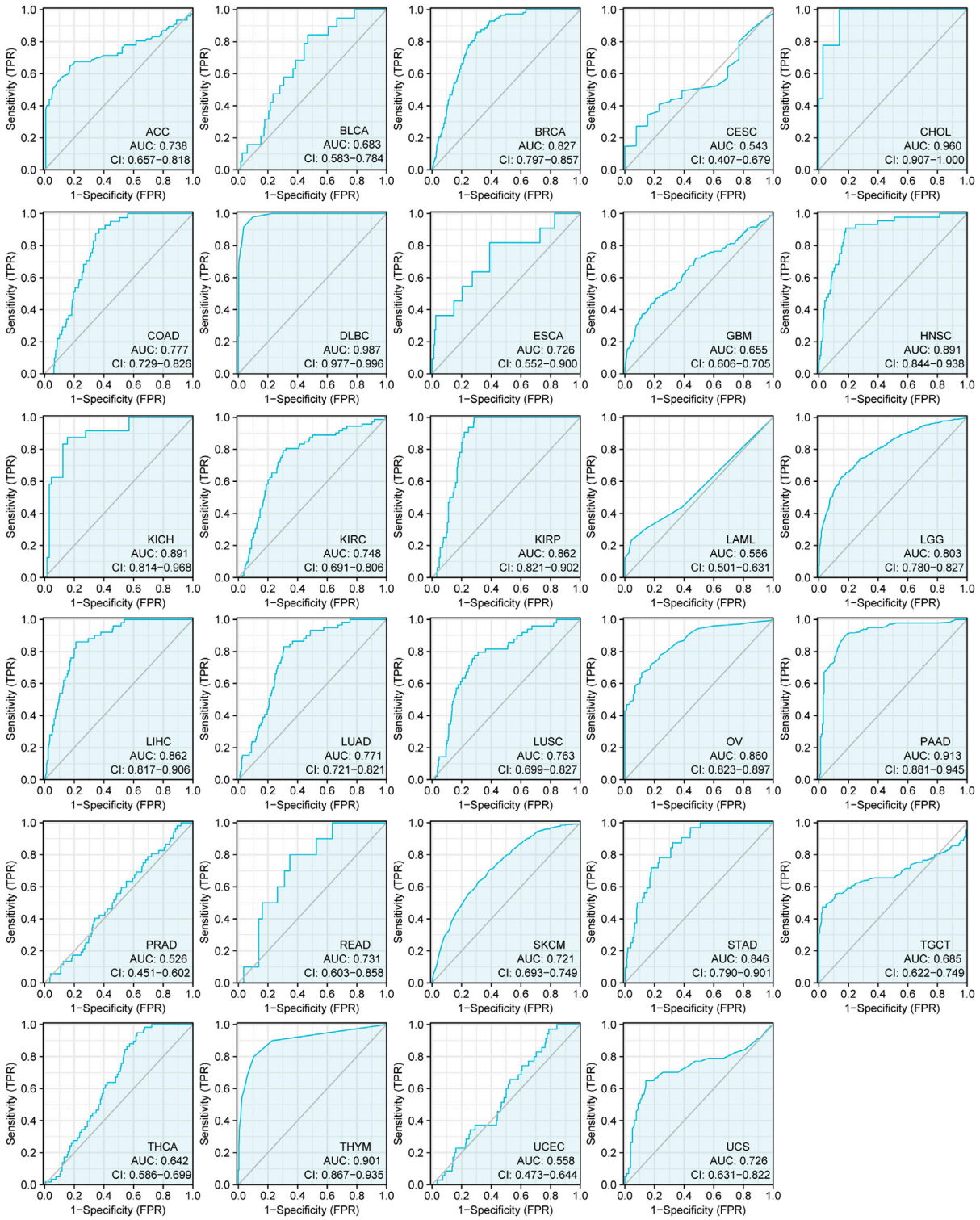


FIGURE 2
The ROC curves of CSMD2 expression as a diagnostic biomarker in tumor and normal tissues. AUC, area under curve.

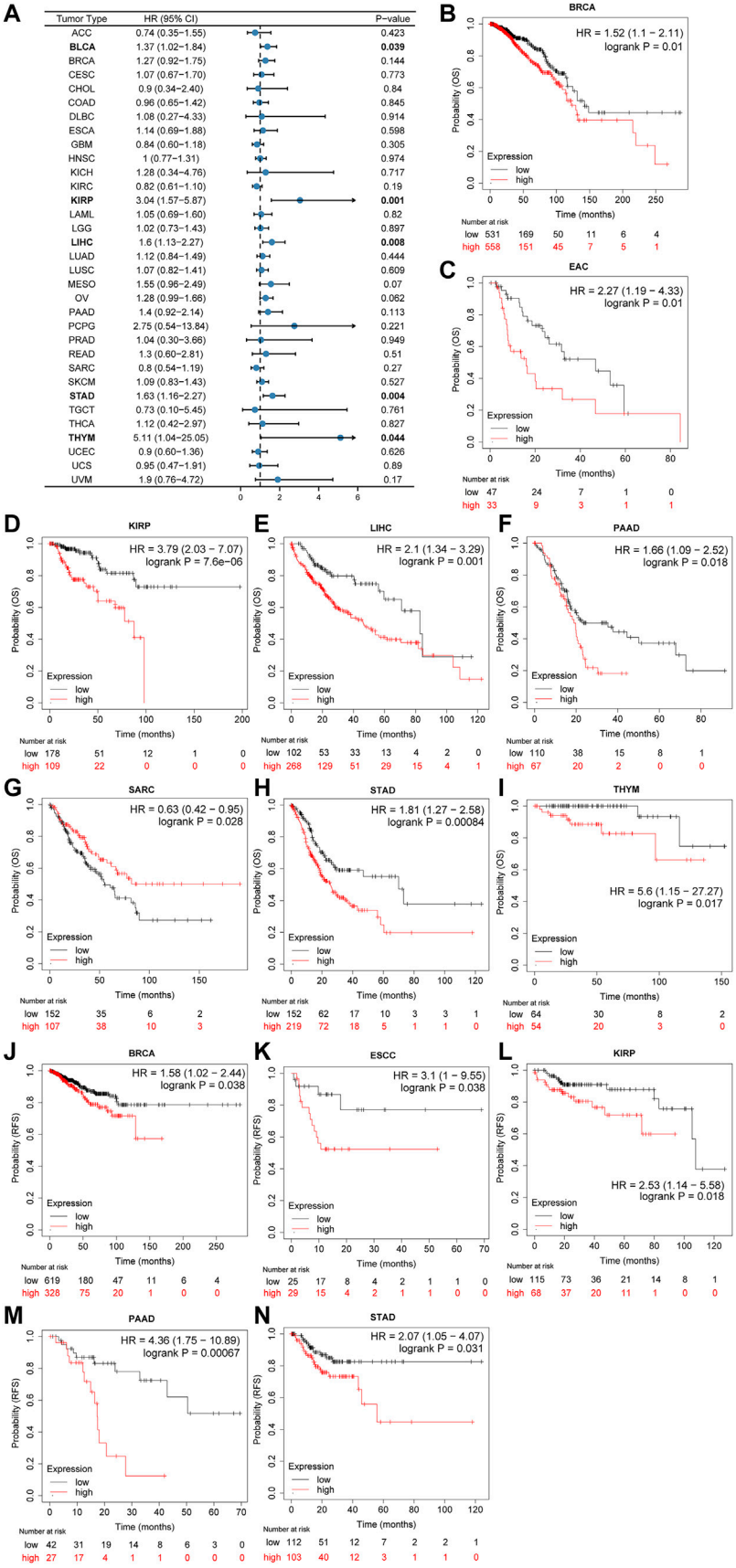


FIGURE 3
Association of prognosis with CSMD2 expression. **(A)** The forest plots of univariate cox regression analyses for OS. The bold items mean that CSMD2 expression was significantly correlated with prognosis in these types of cancers ($p < 0.05$). **(B–I)** Kaplan–Meier plots for overall survival. (*Continued*)

FIGURE 3 | CSMD2 overexpression was significantly associated with poor prognosis in BRCA, EAC, KIRP, LIHC, PAAD, SARC, STAD, and THYM patients. **(J–N)** Kaplan–Meier plots for RFS. CSMD2 overexpression was significantly associated with poor prognosis in BRCA, ESCC, KIRP, PAAD, and STAD patients. Items with a hazard ratio greater than 1 indicated that the CSMD2 expression is a promoting factor of death. **(B–N)** were from Kaplan–Meier plotter.

observed in eight tumors (Figure 5A, Supplementary Table S5), including a significant positive correlation in THYM ($p = 0.033$) and negative correlations in GBMLGG ($p = 0.0025$), CESC ($p = 0.041$), STES ($p = 0.0069$), STAD ($p = 0.006$), HNSC ($p = 0.0002$), MESO ($p = 0.038$), and UVM ($p = 0.0078$). Significant correlations between CSMD2 and MSI were observed in eight tumors (Figure 5B, Supplementary Table S4), including positive correlations in GBMLGG ($p = 0.00013$), COADREAD ($p = 0.023$), and ACC ($p = 0.0002$), and negative correlations in STES ($p = 0.018$), KIPAN ($p = 3.896e-15$), STAD ($p = 0.0188$), HNSC ($p = 0.0065$), and DLBC ($p = 0.001$). Significant correlations between CSMD2 and TNB were observed in four tumors (Figure 5C, Supplementary Table S3), including positive correlations in GBMLGG ($p = 0.0398$), LGG ($p = 0.023$), and READ ($p = 0.033$), and a negative correlation in HNSC ($p = 0.043$).

3.4 Epigenetic methylation analysis

Considering that m6A methylation plays an important role in tumorigenesis and development, the correlation between the expression of CSMD2 mRNA and m6A methylation regulatory factors was investigated for multiple cancers. A total of 21 key m6A methylation regulators, including seven writers (KIAA1429, METT10D, METTL14, METTL3, RBM15, WTAP, and ZC3H13), 11 readers (HNRNPA2B1, HNRNPC, IGF2BP1, IGF2BP2, IGF2BP3, RBMX, YTHDC1, YTHDC2, YTHDF1, YTHDF2, and YTHDF3) and three erasers (FTO, ALKBH3, ALKBH5) were selected. The heatmap indicated that CSMD2 mRNA was positively correlated with most m6A methylation regulatory factors in most cancers (Figure 6A). Additionally, promoter methylation levels of CSMD2 in normal tissues and tumors were compared. CSMD2 is hypermethylated in various cancers, including BLCA, BRCA, CESC, CHOL, COAD, HNSC, KIRP, LUAD, PAAD, PRAD, READ, and UCEC. In contrast, it is hypomethylated in LIHC and PCPG (Figure 6B).

In summary, hypermethylation of CSMD2 promoter and high levels of m6A methylation regulators have been observed in most cancers.

3.5 Immune infiltration, immune evasion, and immune checkpoints analysis in the tumor microenvironment

The correlation between CSMD2 expression and immune cells in the tumor microenvironment was investigated. The levels of cancer-associated fibroblasts (CAFs) and endothelial cells were positively correlated with CSMD2 expression in most cancers (Figure 7A). CSMD2 expression was positively correlated with the infiltration of anti-tumor immune cells ($CD8^+$ T cells, activated memory $CD4^+$ T cells, M1 macrophages, activated NK cells, follicular helper T cells, and gamma delta T cells) in CHOL, KIRP, THYM, and UVM, and negatively correlated with immune evasion associated cells (resting memory $CD4^+$ T cells, T cell regulatory (Tregs), M2 macrophages and resting NK cells). In contrast, it was negatively correlated with the infiltration of anti-tumor immune cells and positively correlated with immune evasion associated cells in most other cancers (Figure 7B). The top two strongest positive correlations with CSMD2 expression were the infiltration level of resting memory $CD4^+$ T cells in DLBC ($p = 1.05e-03$) and M2 macrophages in THYM ($p = 2.61e-07$). The top two strongest negative correlations with CSMD2 expression are the infiltration level of naive $CD4^+$ T cells in THYM ($p = 1.82e-08$) and neutrophils in DLBC ($p = 5.69e-03$) (Figure 7C).

Since tumor cells take advantage of immune checkpoints to evade immune responses, the relationships between the CSMD2 expression and the most common immune checkpoints, including TIGIT, CD274, PDCD1, LAG3, HAVCR2, CTLA4, IDO1, and PDCD1LG2 was also analyzed. The expression of CSMD2 was positively correlated with most immune checkpoints in ACC, BLCA, BRCA, COAD, ESCA, HNSC, KIRC, KIRP, LGG, LIHC, LUAD, LUSC, OV, PAAD, PRAD, READ, SKCM, STAD, THYM, and UVM. In contrast, the expression of CSMD2 negatively correlated with most immune checkpoints in GBM, PCPG, SARC, THCA, and UCEC. Notably, PDCD1LG2, HAVCR2, and CD274 showed the strongest positive correlations with CSMD2 expression in most tumors. (Figures 8A,B).

We observed that the expression of CSMD2 was negatively correlated with MHC-I molecules in CESC, GBM, LGG, PCPG, SARC, STAD, TGCT, and THCA but positively correlated with MHC-I molecule in BLCA, COAD, and LUSC (Supplementary Figure S2). The correlation between the expression of CSMD2 and MHC-II was positive in BLCA, BRCA, COAD, LUSC, OV, PAAD, READ, and SKCM but negative in GBM and THCA (Supplementary Figure S2).

Immunostimulators mainly work at various stages of lymphoid differentiation, development, and maturation, and

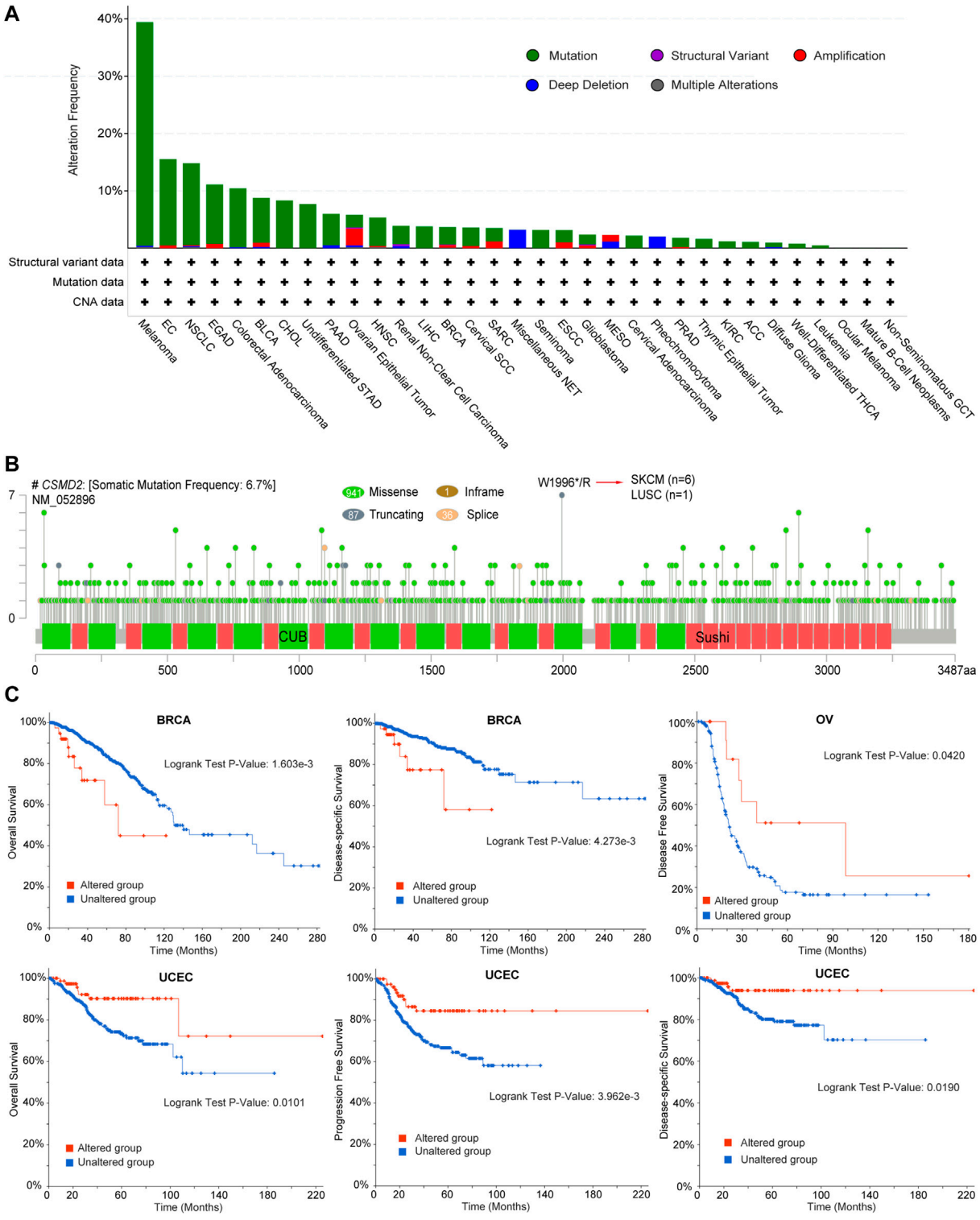


FIGURE 4
Genetic alteration characteristics of CSMD2. **(A)** Bar plot of CSMD2 alteration frequency and types across different cancer types **(B)** The landscape of CSMD2 mutation with the location, types, and number and their relationship with protein domains, and **(C)** Kaplan-Meier curves of differences in OS, DSS, DFS, and PFS between patients with cancer with and without CSMD2 alteration based on cBioportal. Sushi: Sushi repeat (SCR repeat), CUB: CUB domain.

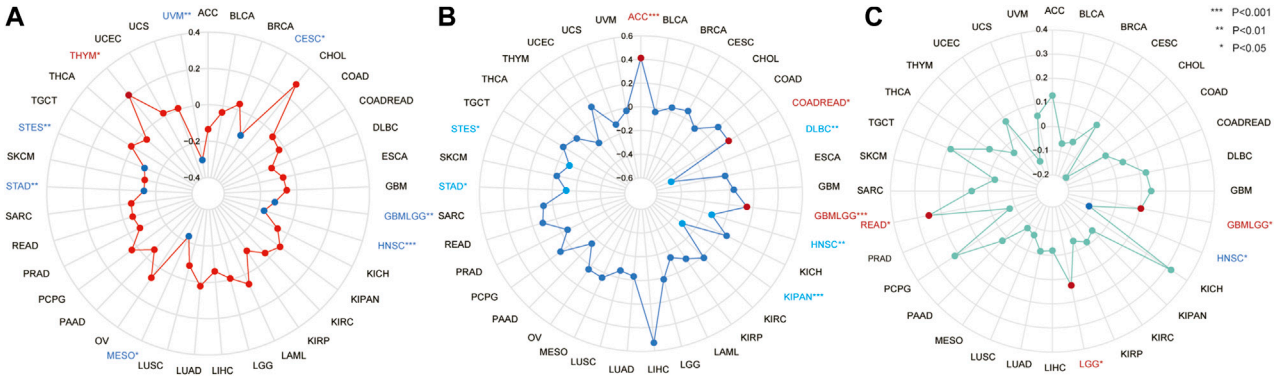


FIGURE 5
The correlation between the expression of CSMD2 and TMB, MSI, and TNB in tumors. **(A)** The correlation between CSMD2 and TMB is positive in THYM and negative in UVM, CESC, GBMLGG, HNSC, MESO, STAD, and STES **(B)** The correlation between CSMD2 and MSI is positive in ACC, COADREAD, and GBMLGG and negative in DLBC, HNSC, KIPAN, STAD, and STES. **(C)** The correlation between CSMD2 and TNB is negative in HNSC and positive in GBMLGG, LGG, and READ (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

regulate immune function, thereby enhancing the ability of the body to prevent and resist disease and exert an anti-tumor role. We observed that CSMD2 was positively correlated with immunostimulators in most tumors, whereas it was negatively correlated with immunostimulators in GBM and THCA (Supplementary Figure S3).

In most cancers, CSMD2 was associated with immune evasion or immunosuppression. Additionally, there was insufficient anti-tumor immunity and up-regulation of immune checkpoints.

3.6 Functional enrichment analysis

To further explore the biological function of CSMD2 in pan-cancer, a series of enrichment analyses were performed. The top 100 genes (Supplementary Table S1) associated with CSMD2 were obtained, and the top four genes including *KIF1B*, *NLGN2*, *QKI* and *CACNG7*, are displayed in Figures 9A,B. Additionally, 25 predicted proteins interacted directly with CSMD2 displayed in the PPI network, of which the interactions with CTSD, SCUBE3 and FOXP2 were experimentally determined (Figure 9C). Both the top 100 CSMD2-related genes and 25 CSMD2-interacting proteins were included in the functional enrichment analysis. We finally obtained 238 gene ontology (GO) terms including 133 biological processes (BP) terms, 75 cellular components (CC) terms, 34 molecular functions (MF) terms and 5 KEGG pathways (Supplementary Table S2). The KEGG analysis results confirmed the enrichment of Wnt signaling pathway and hepatocellular carcinoma pathway (Figure 9D). Furthermore, GO terms related to cancers were displayed. BP analysis showed that CSMD2 might be associated with the glutamate receptor signaling pathway, negative regulation of microtubule polymerization or depolymerization, cell junction organization, protein homooligomerization, and negative regulation of protein complex disassembly, regulation of extent of cell growth, and multicellular organismal signaling (Figure 9E). CC analysis revealed that CSMD2-related genes were enriched in the transmembrane transporter complex, transporter complex, cytoplasmic microtubule, extrinsic component of the plasma membrane, cell-cell adherens junction, and lysosomal lumen (Figure 9F). MF analysis showed that the related genes were associated with tubulin binding, ion channel activity, substrate-specific channel activity, HMG box domain binding, cell adhesion molecule binding, integrin binding, and tau-protein kinase activity (Figure 9G).

4 Discussion

Few studies have investigated the effects of CSMD2 on tumorigenesis and development and its molecular mechanisms. To the best of our knowledge, this is the first study to explore the role of CSMD2 in pan-cancers. The results showed that the expression of CSMD2 was inconsistent in 33 types of cancers, of which CSMD2 expression was upregulated in most tumors. Meanwhile, we analyzed CSMD2 expression in tumors from GEO and ICGC database. It had moderate or high diagnostic efficacy, and the high expression was related to a higher stage and poor prognosis, for example, the expression level of CSMD2 was high in gastric cancer, and patients with high CSMD2 expression had poor

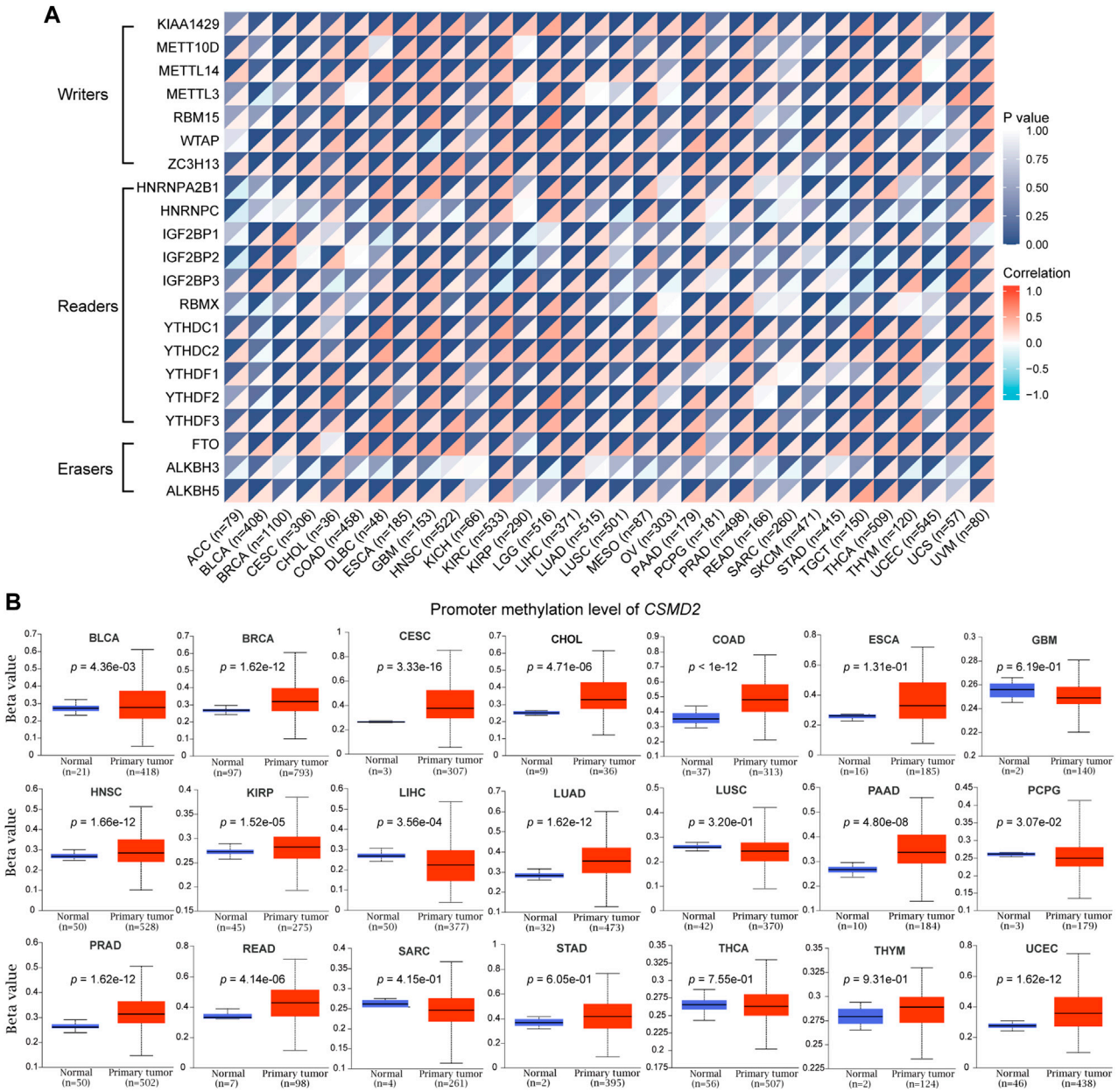


FIGURE 6
Epigenetic methylation analysis of CSMD2. **(A)** The correlation between the expression of CSMD2 mRNA and m6A methylation regulatory factors in multiple cancers. Correlations are depicted with Spearman's rho values and statistical significance. **(B)** Differential promoter methylation level (beta values) of CSMD2 in normal tissues and tumors based on UALCAN.

prognosis. These results suggested that CSMD2 may be an oncogenic molecule involved in tumorigenesis and development.

Missense mutations were the most common type of CSMD2 alterations. Seven mutations were detected in the W1996*/R site of the CUB domain, which were detected in six cases of SKCM and one case of LUSC. Whether this is a functional mutation site remains to be further verified. Further clinical correlation analysis showed that CSMD2 alterations were related to the survival of patients with cancer. CSMD2 alterations were associated with poor survival in BRCA and better survival in patients with OV and UCEC.

There is growing evidence that epigenetic modifications play a vital role in tumors through various mechanisms, in

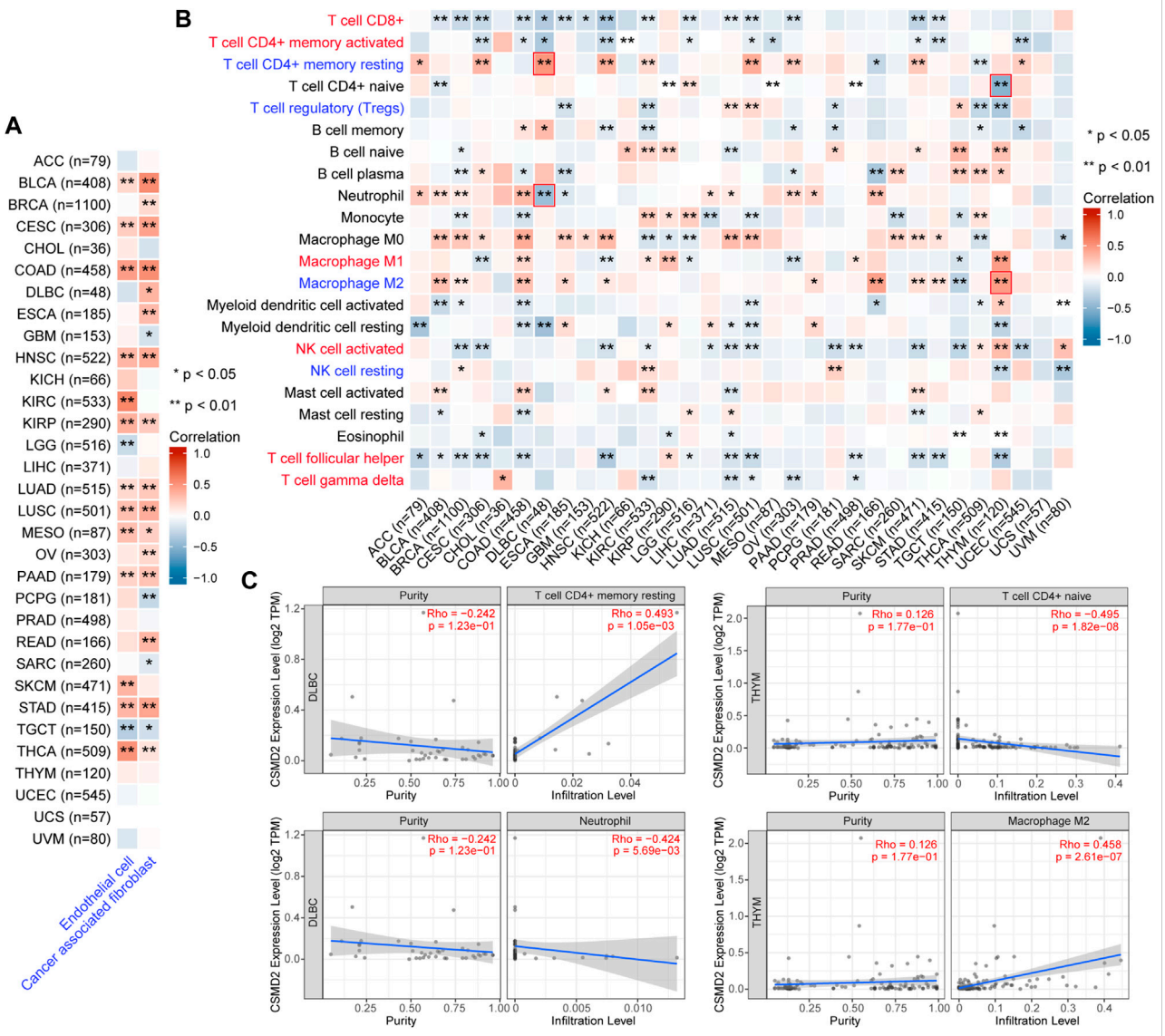


FIGURE 7
Correlation analysis between CSMD2 expression and immune infiltration in the tumor microenvironment. Heat maps display a correlation between CSMD2 expression level and infiltration level of **(A)** cancer-associated fibroblasts and endothelial cells based on XCELL, and **(B)** 22 immune cells based on CIBERSORT in pan-cancer. Anti-tumor immune cells are highlighted in red, and immune evasion-associated cells are highlighted in blue. **(C)** The scatter plots show the top two strongest positive and negative correlations marked in red boxes in **(B)**. Correlations are depicted with the partial Spearman's correlation and statistical significance based on TIMER2.0.

which m6A methylation is a common type of RNA modification. RNA methylation is regulated by different types of regulatory factors, including methyltransferases (writers), RNA-binding proteins (readers) and demethylases (erasers) (Yang Y. et al., 2018; Li et al., 2019). The level of m6A methylation was indirectly known by investigating the levels of these regulatory factors. Our results showed that CSMD2 expression was positively correlated with m6A methylation regulatory factors, and implied that m6A methylation levels might be positively correlated with CSMD2 expression in pan-cancer.

TMB and TNB are biomarkers for therapeutic benefits in many tumors (Hollern et al., 2019). The anti-tumor immune response is likely related to high TMB(Pakish et al., 2017). Usually, the higher the mutation burden, the higher the possibility of neoantigens, and hence, the higher the immunotherapy response rate (Turajlic et al., 2017). A defective DNA mismatch repair (MMR) system leads to

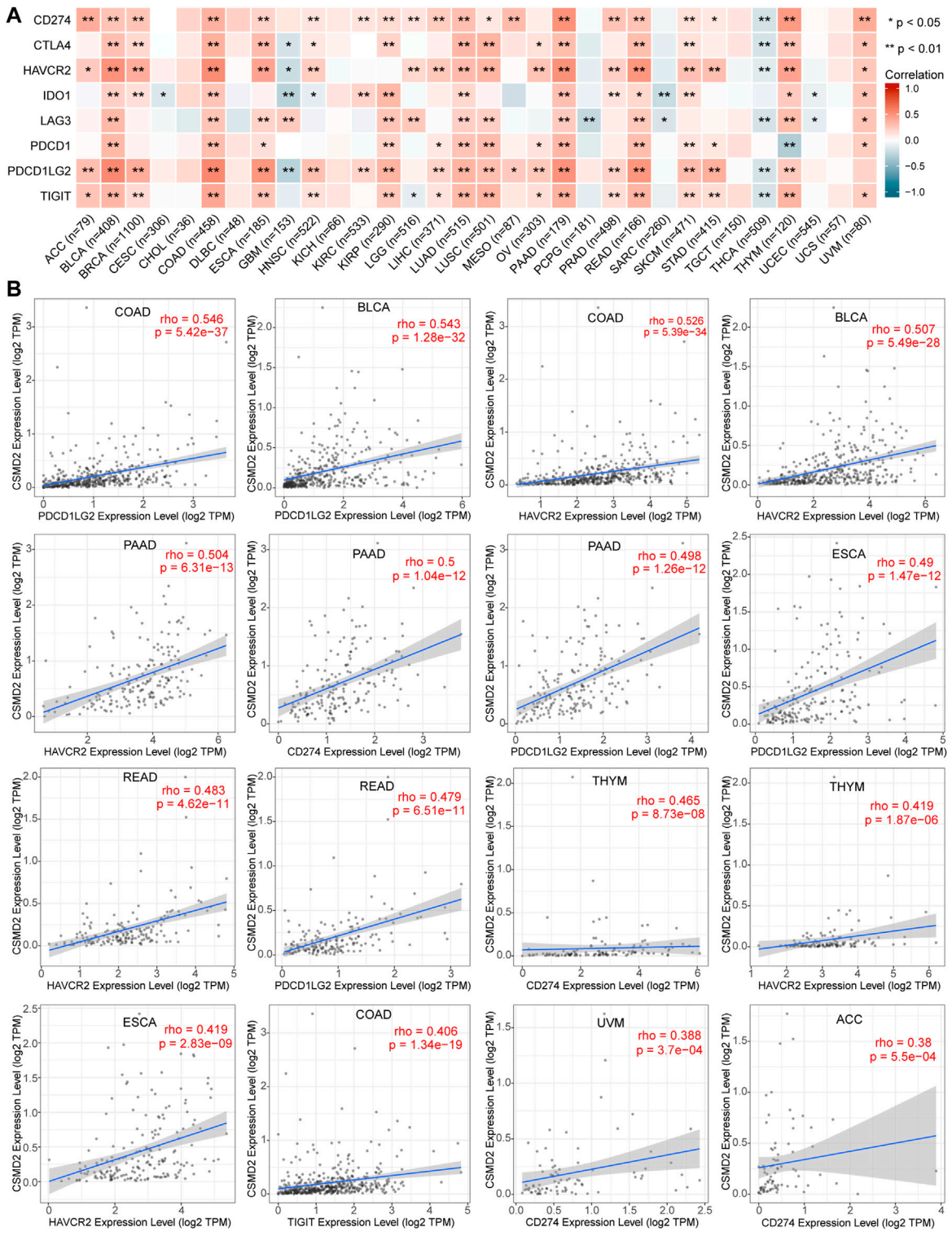


FIGURE 8
Correlation analyses of the CSMD2 expression with immune checkpoints. **(A)** Heat maps display a correlation between CSMD2 expression and immune checkpoints, including TIGIT, CD274, PDCD1, LAG3, HAVCR2, CTLA4, IDO1, and PDCD1LG2 in pan-cancer. **(B)** The top sixteen strongest correlations are displayed via scatter plots. Correlations are depicted with Spearman's correlation and statistical significance based on TIMER 2.0.

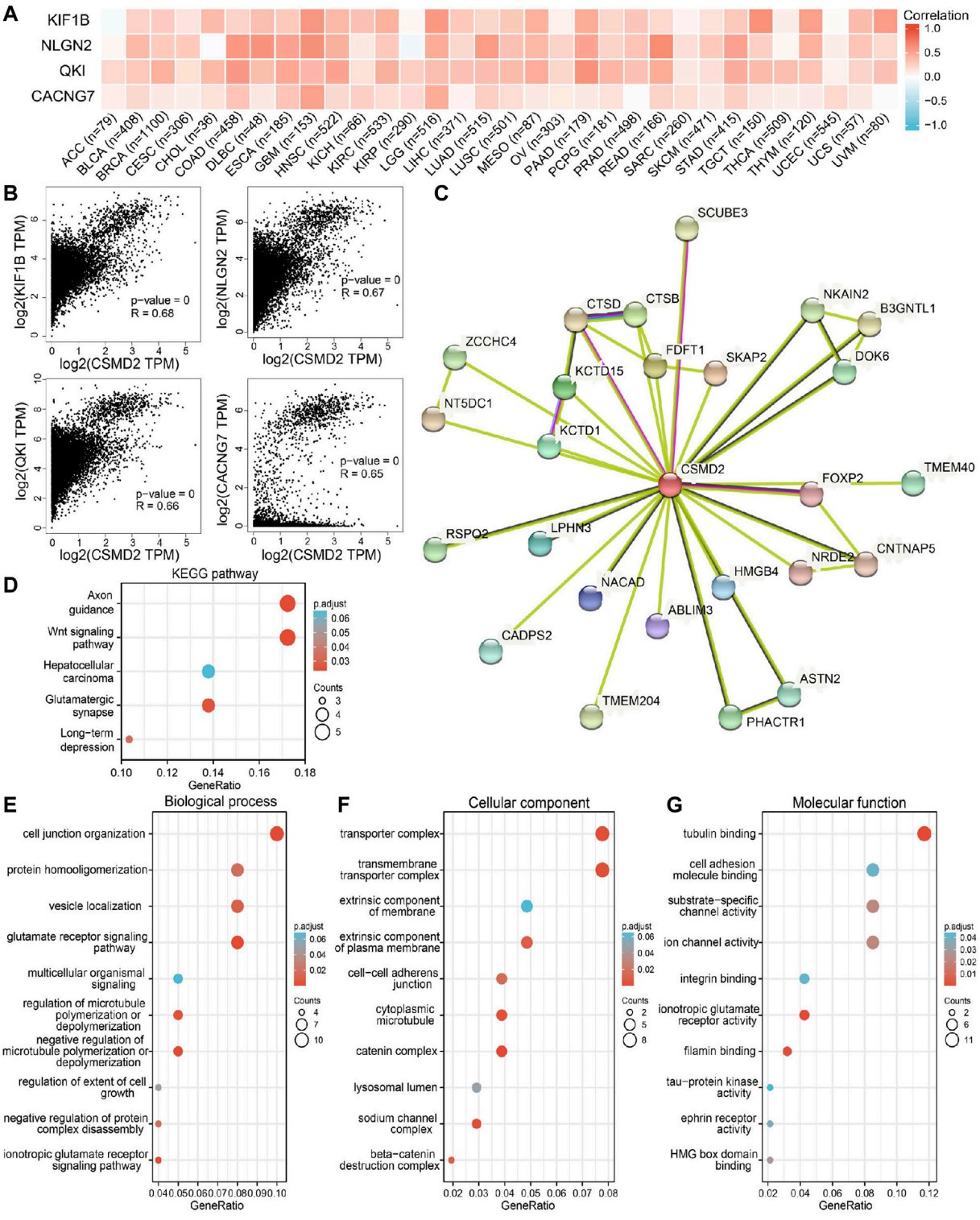


FIGURE 9
CSMD2-related genes, interacting proteins and functional enrichment analysis. **(A)** and **(B)** showed the correlation between CSMD2 and the top four genes related to it in 33 cancers. **(C)** The PPI network of CSMD2. GO analyses, including **(D)** KEGG pathway, **(E)** biological process, **(F)** cellular component, and **(G)** molecular function.

the accumulation of genetic errors while copying microsatellite loci, resulting in MSI(de Rosa et al., 2016). High microsatellite instability (MSI-H) had been used as a biomarker of the impaired function of the MMR system and is correlated with better efficacy of immunotherapy (Lee et al., 2002; Chen et al., 2019). Higher TMB is reportedly associated with better OS and better response to ICIs(Samstein et al., 2019). In this study, CSMD2 expression was found to be negatively correlated with TMB and MSI in both STAD and STES. CSMD2 expression negatively correlated with TMB, MSI, and TNB in HNSC. CSMD2 was highly expressed in the above three tumors, and high CSMD2 expression was correlated with poor prognosis in patients with STAD. Through the correlation analysis of CSMD2 with TMB, MSI, and TNB, we predicted that CSMD2 might play a role in immunotherapy.

MHC-I molecules present endogenous antigens and activate $CD8^+$ T-cells, which are then transformed into active cytotoxic T lymphocytes to kill target cells. MHC-II molecules are mainly involved in presenting exogenous antigenic peptides to $CD4^+$ T cells, which activate $CD4^+$ helper T cells, proliferates and express the corresponding lymphokines, and initiate humoral immune responses. In this study, CSMD2 was negatively correlated with MHC-I molecules, whereas, it was positively correlated with MHC-II molecules, and immunostimulators in most tumors. The mechanism of tumor immune regulation is highly complex, therefore, the relationship between CSMD2 and immunity requires further research.

Tumor microenvironment is the surrounding environment where tumor develops and survives. In addition to tumor cells, surrounding fibroblasts, immune and inflammatory cells, and microvessels are present in the TME (Hui and Chen, 2015). We observed that CSMD2 expression was negatively correlated with infiltration of anti-tumor immune cells, including $CD8^+$ T cells, activated memory $CD4^+$ T cells, M1 macrophages, activated NK cells, follicular helper T cells, and gamma delta T cells, as well as positively correlated with immune evasion- or suppression-associated cells, including CAFs, endothelial cells, Tregs, and M2 macrophages in most cancers. Therefore, CSMD2 is speculated to promote tumor cells proliferation, migration and invasion through immune escape or immunosuppression rather than anti-tumor immune infiltration. In addition, the upregulation of immune checkpoints helps explain this and may provide the possibility of promoting ICI effects in patients with cancer.

In this study, the correlation between CSMD2 expression and anti-tumor immune response and microenvironment was comprehensively analyzed. It was found that CSMD2 expression might be related to immune escape and promoting the occurrence and development of tumors. The major challenge of current cancer immunotherapy is specific tumor immune response. However, tumors of different types and sites vary in their response to immunotherapy, the mechanism is extremely complex, which needs to be studied in future.

Evaluating global methylation abnormalities by methylation load can predict the degree of tumor immunogenicity. The degree of abnormal methylation is negatively correlated with tumor immunogenicity (Park et al., 2021). High levels of promoter methylation of CSMD2 and m6A methylation were found in tumor tissues or high CSMD2 expression tissues, suggesting the importance of abnormal methylation in tumor evasion of immune surveillance.

Enrichment analyses showed that CSMD2 might be located on the cell membrane, constitute a component of channel proteins, and participate in signal transduction between tumor cells. Interestingly, pathway enrichment analyses revealed its relation to the Wnt signaling pathway and hepatocellular carcinoma pathway. Therefore, further experimental verification is required to confirm this finding. In addition, the predicted proteins interacting with CSMD2 need to be verified.

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

Conceiving and designing: HZ and TH. Analyses: HZ, TH, XR, XF, XC, HW, WS, and YW. Writing (original draft preparation): HZ and TH. Review and editing: HZ and YW. All authors have read and approved the final version of the article. HZ and TH have contributed equally to the study.

Funding

This work was supported by the National Key R&D Program of China (2016YFC1302201), the Natural Science Foundation of Gansu Province (21JR1RA116).

Acknowledgments

We acknowledge the data support from the TCGA database. We also thank the Xiantao (https://www.xiantao.love/), Assistant

for clinical bioinformatics (https://www.aclbi.com/static/index.html#/), GEO, ICGC, GTEx, GEDS, CCLE, cBioportal, Sangerbox, UALCAN, the Kaplan-Meier Plotter, Oncolnc, TIMER2.0, TISIDB, GEPIA2, and STRING database for their assistance. We would like to thank Editage (www.editage.cn) for English language editing. We also acknowledge Guozhi Wu 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.2022.918486/full#supplementary-material

References

Anaya, J. (2016). OncoLnc: Linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *PeerJ Comput. Sci.* 2, e67. doi:10.7717/peerj-cs.67

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

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. 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 (New York, N.Y.)* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Chen, M. H., Chang, S. C., Lin, P. C., Yang, S. H., Lin, C. C., Lan, Y. T., et al. (2019). Combined microsatellite instability and elevated microsatellite alterations at selected tetranucleotide repeats (EMAST) might Be a more promising immune biomarker in colorectal cancer. *Oncologist* 24 (12), 1534–1542. doi:10.1634/theoncologist.2019-0171

de Rosa, N., Rodriguez-Bigas, M. A., Chang, G. J., Veerapong, J., Borras, E., Krishnan, S., et al. (2016). DNA mismatch repair deficiency in rectal cancer: Benchmarking its impact on prognosis, neoadjuvant response prediction, and clinical cancer genetics. *J. Clin. Oncol.* 34 (25), 3039–3046. doi:10.1200/jco.2016.66.6826

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

Gialeli, C., Gungor, B., and Blom, A. M. (2018). Novel potential inhibitors of complement system and their roles in complement regulation and beyond. *Mol. Immunol.* 102, 73–83. doi:10.1016/j.molimm.2018.05.023

Győrffy, B. (2021). Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput. Struct. Biotechnol. J.* 19, 4101–4109. doi:10.1016/j.csbj.2021.07.014

Håvik, B., Le Hellard, S., Rietschel, M., Lybæk, H., Djurovic, S., Mattheisen, M., et al. (2011). The complement control-related genes CSMD1 and CSMD2 associate to schizophrenia. *Biol. Psychiatry* 70 (1), 35–42. doi:10.1016/j.biopsych.2011.01.030

Hollern, D. P., Xu, N., Thennavan, A., Glodowski, C., Garcia-Recio, S., Mott, K. R., et al. (2019). B cells and T follicular helper cells mediate response to checkpoint inhibitors in high mutation burden mouse models of breast cancer. *Cell* 179 (5), 1191–1206. e1121. doi:10.1016/j.cell.2019.10.028

Hui, L., and Chen, Y. (2015). Tumor microenvironment: Sanctuary of the devil. *Cancer Lett.* 368 (1), 7–13. doi:10.1016/j.canlet.2015.07.039

Lau, W. L., and Scholnick, S. B. (2003). Identification of two new members of the CSMD gene family. *Genomics* 82 (3), 412–415. doi:10.1016/s0888-7543(03)00149-6

Lee, H. S., Choi, S. I., Lee, H. K., Kim, H. S., Yang, H. K., Kang, G. H., et al. (2002). Distinct clinical features and outcomes of gastric cancers with microsatellite instability. *Mod. Pathol.* 15 (6), 632–640. doi:10.1038/modpathol.3880578

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), W509W514. doi:10.1093/nar/gkaa407

Li, Y., Xiao, J., Bai, J., Tian, Y., Qu, Y., Chen, X., et al. (2019). Molecular characterization and clinical relevance of m6A regulators across 33 cancer types. *Mol. Cancer* 18 (1), 137. doi:10.1186/s12943-019-1066-3

Men, C., Chai, H., Song, X., Li, Y., Du, H., Ren, Q., et al. (2017). Identification of DNA methylation associated gene signatures in endometrial cancer via integrated analysis of DNA methylation and gene expression systematically. *J. Gynecol. Oncol.* 28 (6), e83. doi:10.3802/jgo.2017.28.e83

Pakish, J. B., Zhang, Q., Chen, Z., Liang, H., Chisholm, G. B., Yuan, Y., et al. (2017). Immune microenvironment in microsatellite-instable endometrial cancers: Hereditary or sporadic origin matters. *Clin. Cancer Res.* 23 (15), 4473–4481. doi:10.1158/1078-0432.Ccr-16-2655

Park, C., Jeong, K., Park, J.-H., Jung, S., Bae, J. M., Kim, K., et al. (2021). Pan-cancer methylation analysis reveals an inverse correlation of tumor immunogenicity with methylation aberrancy. *Cancer Immunol. Immunother.* 70 (6), 1605–1617. doi:10.1007/s00262-020-02796-1

Rossi, V., Bally, I., Thielens, N. M., Esser, A. F., and Arlaud, G. J. (1998). Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J. Biol. Chem.* 273 (2), 1232–1239. doi:10.1074/jbc.273.2.1232

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 (20), 4200–4202. doi:10.1093/bioinformatics/btz210

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 (2), 202–206. doi:10.1038/s41588-018-0312-8

Shinawi, T., Hill, V. K., Krex, D., Schackert, G., Gentle, D., Morris, M. R., et al. (2013). DNA methylation profiles of long- and short-term glioblastoma survivors. *Epigenetics* 8 (2), 149–156. doi:10.4161/epi.23398

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

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

Turajlic, S., Litchfield, K., Xu, H., Rosenthal, R., McGranahan, N., Reading, J. L., et al. (2017). Insertion-and-deletion-derived tumour-specific neoantigens and the

immunogenic phenotype: A pan-cancer analysis. *Lancet. Oncol.* 18 (8), 1009–1021. doi:10.1016/s1470-2045(17)30516-8

Vater, I., Montesinos-Rongen, M., Schlesner, M., Haake, A., Purschke, F., Sprute, R., et al. (2015). The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing. *Leukemia* 29 (3), 677–685. doi:10.1038/leu.2014.264

Xia, M., Liu, C.-J., Zhang, Q., and Guo, A.-Y. (2019). Geds: A gene expression display server for mRNAs, miRNAs and proteins. *Cells* 8 (7), E675. doi:10.3390/cells8070675

Yang, P.-S., Hsu, H.-H., Hsu, T.-C., Chen, M.-J., Wang, C.-D., Yu, S.-L., et al. (2018). Genome-wide scan for copy number alteration association with relapse-free survival in colorectal cancer with liver metastasis patients. *J. Clin. Med.* 7 (11), E446. doi:10.3390/jcm7110446

Yang, Y., Hsu, P. J., Chen, Y.-S., and Yang, Y.-G. (2018). Dynamic transcriptomic m6A decoration: Writers, erasers, readers and functions in RNA metabolism. *Cell Res.* 28 (6), 616–624. doi:10.1038/s41422-018-0040-8

Zhang, J., Bajari, R., Andric, D., Gerthoffert, F., Lepsa, A., Nahal-Bose, H., et al. (2019). The international cancer Genome Consortium data portal. *Nat. Biotechnol.* 37 (4), 367–369. doi:10.1038/s41587-019-0055-9

Zhang, R., and Song, C. (2014). Loss of CSMD1 or 2 may contribute to the poor prognosis of colorectal cancer patients. *Tumour Biol.* 35 (5), 4419–4423. doi:10.1007/s13277-013-1581-6

Zhou, T., Cai, Z., Ma, N., Xie, W., Gao, C., Huang, M., et al. (2020). A novel ten-gene signature predicting prognosis in hepatocellular carcinoma. *Front. Cell Dev. Biol.* 8, 629. doi:10.3389/fcell.2020.00629

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 25 August 2022
DOI 10.3389/fgene.2022.953997

Check for updates

OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Siming Gong,
Leipzig University, Germany
Xin Wu,
Zhejiang Chinese Medical University, China

*CORRESPONDENCE
Feng Wang,
zzuwangfeng@zzu.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 07 June 2022
ACCEPTED 27 July 2022
PUBLISHED 25 August 2022

CITATION
Xin D, Man Y, Yang Y and Wang F (2022), A novel prognostic and therapeutic target biomarker based on necroptosis-related gene signature and immune microenvironment infiltration in gastric cancer.
Front. Genet. 13:953997.
doi: 10.3389/fgene.2022.953997

COPYRIGHT
© 2022 Xin, Man, Yang 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.

A novel prognostic and therapeutic target biomarker based on necroptosis-related gene signature and immune microenvironment infiltration in gastric cancer

Dao Xin[1†], Yuxin Man[2†], Yalan Yang[1†] and Feng Wang[1*]

[1]Department of Oncology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China, [2]Department of Medical Oncology, Sichuan Cancer Hospital, Medical School of University of Electronic Science and Technology of China, Chengdu, China

Background: Gastric cancer is a major global public health burden worldwide. Although treatment strategies are continuously improving, the overall prognosis remains poor. Necroptosis is a newly discovered form of cell death associated with anti-tumor immunity.

Methods: Gastric cancer (GC) data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) were downloaded. Bioinformatics analysis was performed to construct a necroptosis-related risk model and to establish cancer subtypes. Potential associations of the tumor immune microenvironment and immunotherapy response with necroptosis-related prognostic risk score (NRG risk score) were comprehensively explored. 16 GC and paired normal tissues were collected and RT-PCR was performed to examine expression of NRG related genes.

Results: GC samples were stratified into three subtypes according to prognostic necroptosis gene expression. A necroptosis risk model based on 12 genes (*NPC1L1, GAL, RNASE1, PCDH7, NOX4, GJA4, SLC39A4, BASP1, BLVRA, NCF1, PNOC*, and *CCR5*) was constructed and validated. The model was significantly associated with the OS and PFS of GC patients and the tumor immune microenvironment including immune cell infiltration, microsatellite instability (MSI) status, tumor mutational burden (TMB) score, immune checkpoint, and human leukocyte antigen (HLA) gene expression. A prognostic nomogram based on the NRG_score was additionally constructed. A low NRG risk score was correlated with high tumor immunogenicity and might benefit from immunotherapy.

Conclusion: We have identified a useful prognostic model based on necroptosis-related genes in GC and comprehensively the relationship between necroptosis and tumor immunity. Predicting value to immunotherapy response is promising, and further research to validate the model in clinical practice is needed.

KEYWORDS

necroptosis, prognostic model, cancer subtypes, gastric cancer, tumor immune microenvironment, immunotherapy

Introduction

Gastric cancer (GC), including gastroesophageal junction (GEJ) adenocarcinoma, is the fourth leading cause of cancer-related mortality worldwide (Sung et al., 2021). Adenocarcinoma of the stomach (STAD) is the most common pathological subtype, accounting for up to 95%, and other pathological types such as lymphoma and leiomyosarcoma are relatively rare. The majority of GC patients present with distant metastasis at first diagnosis. The treatment options for advanced GC have undergone significant evolution over recent years, developing from traditional chemotherapy to targeted therapy and immunotherapy, resulting in the progressive improvement of outcomes (Joshi and Badgwell, 2021). However, GC is a highly heterogeneous malignant tumor, and in clinical practice, a large number of patients do not benefit from targeted therapy or immunotherapy (Seeneevassen et al., 2021). Selection of the patient populations that potentially benefit from these treatments is therefore critical for the optimization of survival outcomes.

Necroptosis, formerly considered an unregulated accidental cell death process, is a caspase-independent form of cell death involving receptor-interacting protein kinase 1 (RIP1), RIP3, and mixed lineage kinase domain-like protein (MLKL) (Gong et al., 2019; Khoury et al., 2020). This newly discovered pathway regulates necrosis and is induced by death receptors, interferon, Toll-like receptors, intracellular RNA and DNA sensors, and other potential regulators (Pasparakis and Vandenabeele, 2015). Recent studies have provided exciting novel insights into the regulatory mechanisms of necrosis and their associations and suggest that necrosis is critical in the pathogenesis of multiple human diseases. Rather than the formation of apoptotic bodies, which occurs during cellular apoptosis, necroptosis is accompanied by rupture of the cell membrane and release of tumor neoantigens, which can trigger strong inflammatory and anti-tumor immune responses. This pathway is involved in the development and progression of various tumors and appears to serve as an effective biomarker of survival outcome or therapeutic effect (Gong et al., 2019; Sprooten et al., 2020; Tang et al., 2020). To date, limited studies have focused on the potential involvement and underlying molecular mechanisms and therapeutic response of the immunotherapy of necroptosis in gastric cancer, and few necroptosis-related cancer subtypes or prognostic models are currently available.

In this study, GC samples were collected from TCGA and GEO databases for comprehensive analysis of necroptosis-related gene (NRG) expression, mutation status, and copy number variations. Based on NRG expression patterns, all samples were classified into two necroptosis-associated subtypes. According to differentially expressed gene (DEG) patterns between the two groups, samples were classified into three different gene subtypes. We successfully constructed a predictive model and comprehensively evaluated the correlations between different risk layers and the tumor immune microenvironment (TIME). Our collective findings provide novel insights that could aid in the evolution of strategies for accurate classification and effective immunotherapy and innovative targeted therapy of gastric cancer.

Materials and methods

Data collection

The flow chart of the study is presented in Figure 1. RNA sequencing data (fragments per kilobase million; FPKM) and relevant clinical and follow-up information were obtained from TCGA and GEO datasets (GSE84433). GEO database are available in GPL6947 platform (Illumina HumanHT-12 V3. 0 expression beadchip). FPKM values of TCGA-stomach adenocarcinoma (STAD) data were converted and normalized to transcripts per kilobase million (TPM), which are more similar to those resulting from microarrays and more comparable between samples. Intersection genes of TCGA and GEO databases were selected out. The normal samples in TCGA database were removed and log2 was performed for GEO database. The "limma" R package was used for data normalization. The "ComBat" algorithm of "sva" R package was used to correct the batch effects due to the non-biological technical bias. Then, data from the two cohorts were combined, with the exclusion of cases with missing follow-up information or unknown survival status (Song et al., 2021; Yang et al., 2022). Clinical data, such as TNM stage, age, gender, follow-up time, and survival status, were collected. A list of 67 necroptosis-related genes (NRG) were selected for analysis (Supplementary Table S1) from the Gene Set Enrichment Analysis (GSEA) database and previous publications.

Somatic mutations and copy number alterations of NRGs

Somatic mutation status (workflow type: VarScan2 Variant Aggregation and Masking) and copy number variation (CNV) data were downloaded from the TCGA database. Summary analysis of somatic mutation frequency in the 12 necroptosis genes was performed to select genes with high mutational frequency. Additionally, somatic copy number alterations of necroptosis genes were analyzed, along with correlations of

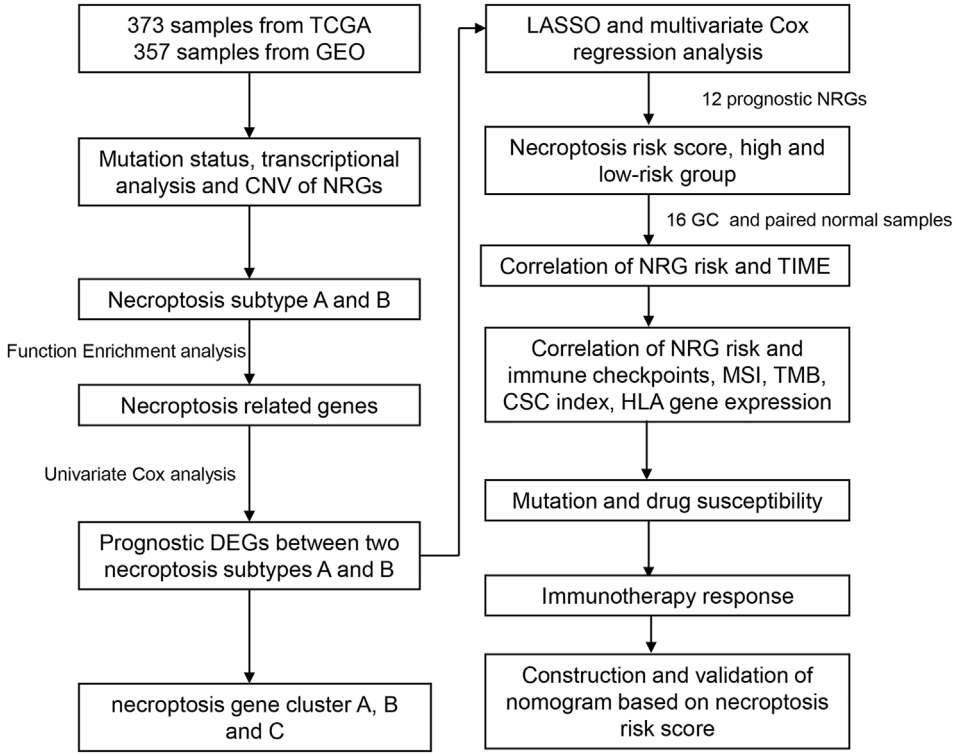


FIGURE 1
Flowchart of the study.

CNV and mRNA expression and chromosomal location of necroptosis genes.

Consensus clustering analysis of NRGs

The R package "ConsensusClusterPlus" was employed for subsequent consensus unsupervised sample clustering analysis. GC samples were classified into two different subtypes based on necroptosis genes expression. To achieve the optimal subtyping effect, cumulative distribution function (CDF) curves were increased gradually and smoothly, and sample sizes were relatively balanced among the different subgroups.

Correlation of necroptosis subtypes with clinicopathological characteristics and prognosis

We created a heatmap to evaluate the associations of the two necroptosis subtypes with prognosis and clinicopathological characteristics, including TNM stage, pathological grade, age, and sex. Kaplan–Meier analysis was performed to compare differential prognosis in overall survival (OS). Univariate Cox regression was conducted for preliminary analysis of the correlation between the expression of individual necroptosis genes and the prognosis of GC patients.

DEG identification and functional analysis

Differentially expressed genes (DEG) were selected between the different necroptosis subgroups using the R package "limma" at a threshold p-value<0.05 and a fold change of 1.5. Subsequent Gene set variation analysis (GSVA) was performed to identify the functions and biological processes of the different subgroups. GSVA was based on the hallmark gene set (c2.cp.kegg.v7.4.symbols.gmt) downloaded from the MSigDB database. Additionally, functional enrichment analysis was conducted based on DEGs using the R packages "clusterProfiler" and "enrichplot."

Necroptosis-related gene prognostic risk score

Necroptosis-related prognostic risk score (NRG_score) was created to quantify the necroptosis patterns of each GC sample. DEGs associated with OS were selected according to univariate

Cox regression. Next, all patients were divided into different necroptosis gene clusters (A, B, C) following the unsupervised clustering method according to prognostic DEGs. All GC ($n = 696$) cases from TCGA and GEO cohorts were combined and randomly divided into training ($n = 348$) and validation cohorts ($n = 348$) at a 1:1 ratio. The training cohort was used for subsequent analysis to construct the NRG_score. After univariate Cox regression analysis, Lasso regression was performed using the R package "glmnet" and multivariate Cox regression was eventually conducted to identify candidate prognostic necroptosis genes. The NRG_score was calculated using the formula: NRG_score = Σ (Exp * coefi). Patients were further divided into high-risk and low-risk groups based on the median NRG_score.

Correlation of prognostic NRG risk score with clinical factors

Patients in the validation and all other cohorts were also divided into high-risk and low-risk groups using the same criteria. The distribution plot revealed a correlation between overall survival status and NRG_score. Kaplan–Meier analysis was performed to compare the prognosis of the different risk groups in overall survival (OS) among the training, validation, and all patient cohort. A heatmap was plotted to compare the differential expression patterns of prognostic NRGs between the two groups.

Immune landscape, microsatellite instability status, cancer stem cell index, tumor mutation burden score, and HLA gene expression between high- and low-risk groups

The CIBERSORT algorithm was applied to explore the correlations of 22 infiltrating human immune cell types and NRG_scores. The ESTIMATE algorithm was performed to assess the immune and stromal scores between high-risk and low-risk groups. Correlations of expression of 12 prognostic NRGs and immune cells were additionally evaluated. We further examined the associations of the NRG_score and immune checkpoint gene expression, MSI status, CSC index, TMB score, and HLA gene expression. A mutation annotation format (MAF) was performed for comparison of GC patients in high-risk and low-risk groups using the maftools R package.

Drug susceptibility analysis and immunotherapy response prediction

To explore the application value of NRG_score in clinical drug selection, "pRRophetic" R package analysis was performed to assess drug susceptibility in the two risk groups and the half-maximal inhibitory concentration (IC50) values calculated for commonly used chemotherapy or targeted therapeutic drugs. The Cancer Immunome Atlas (https://tcia.at/) analyzed the immune landscapes and antigenomes of 20 solid tumors that were quantified by Immunophenoscore (IPS, a superior immune response molecular marker). The IPS value, which ranged from 0 to 10, was positively correlated to tumor immunogenicity and could predict the patients' response to immune checkpoint inhibitors therapy, including anti-PD-1/PD-L1 and anti-CTLA-4 immunotherapy. Imvigor210 was a transcriptome database including treatment response data of patients who received anti-PD-L1 immunotherapy. It was extracted to assess the predicting value of NRG risk score in immunotherapy response. Another public cohort GSE78220, was also used to assess the response and survival outcomes for patients receiving immunotherapy based on NRG risk score.

Construction and validation of an NRG-related nomogram

A prognostic nomogram was constructed by integrating the NRG risk level with common clinical variables. We additionally generated calibration curves for 1-, 3-, and 5-year OS to compare the model prediction values with actual outcomes. Decision curve analysis (DCA) analysis was conducted to estimate the predictive value of the nomogram in clinical decision-making practice. A receiver operating characteristic (ROC) curve was plotted to compare the prognostic power of the NRG_score risk group alone with the nomogram model.

Human tissues and quantitative real time-polymerase chain reaction

We obtained 16 cancer and their paired normal tissues from gastric cancer patients who underwent stomach surgery in the First Affiliated Hospital of Zhengzhou University. The study protocol was approved by the First Affiliated Hospital of Zhengzhou University Research Ethics Committee. Written informed consent was provided by each patient. Total RNA was extracted by Trizol reagent (Takara, Beijing, China) according to the manufacturer's protocol. The synthesis of cDNAs corresponding to the mRNAs of interest depended on PrimeScript RT reagent Kit with gDNA Eraser (Takara) and SYBR Green Premix (Cowin Biosciences, Jiangsu province, China) with specific PCR primers (Sangon Biotech Co., Ltd, Shanghai, China). The data were normalized with GAPDH. The primers used in PCR assays were listed in Supplementary Table S15.

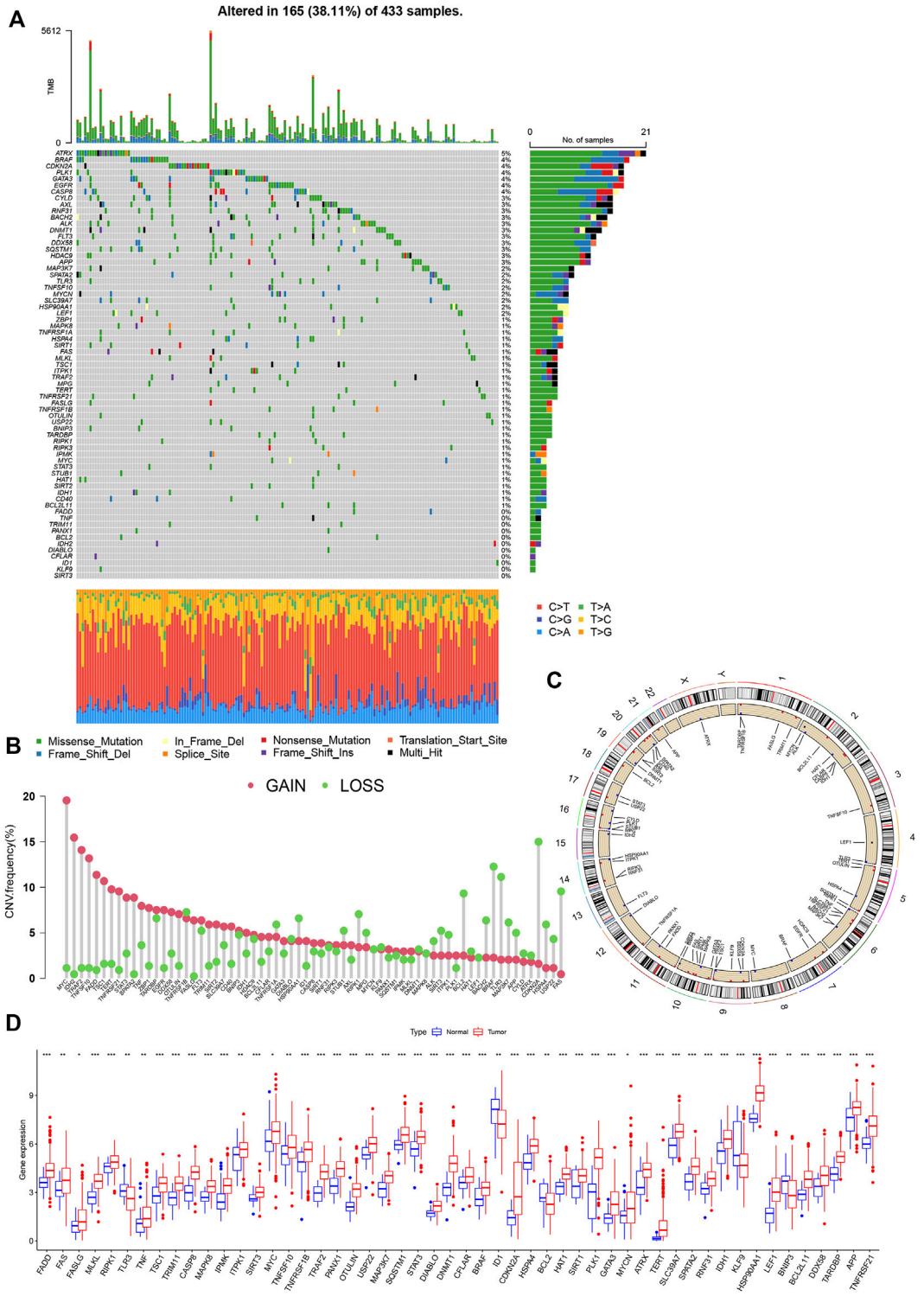


FIGURE 2
Genetic and transcriptional analysis of NRGs in GC. **(A)** Mutation frequencies of 67 NRGs in GC patients from TCGA cohort. **(B)** Copy number cariation of NRGs. **(C)** Locations of NRGs CNV on 23 chromosomes. **(D)** Expression differences of NRGs between normal and GC samples.

Statistical analysis

Statistical analyses were performed using R software version 4.1.1 (2021-08-10). Data were considered statistically significant at p-values<0.05. All R scripts have been added in the Supplementary Material.

Results

Mutation landscape and copy number alterations of NRGs in GC

A total of 67 necroptosis-related genes (NRG) were included for analysis. Analysis of mutations revealed relatively high mutational frequencies of NRGs in GC. Among the 465 GC samples, 165 (38.11%) displayed NRG mutations (Figure 2A). Overall, 56 (83.6%) genes showed different mutation frequencies and types. The gene with the highest mutational frequency was *ATRX* (5%), followed by *BRAF*, *CDKN2A*, *PLK1*, *GATA3*, *EGFR*, and *CASP8* (4%). No mutations were detected in 11 NRGs (*SIRT3*, *KLF9*, *ID1*, *CFLAR*, *DIABLO*, *IDH2*, *BCL2*, *PANX1*, *TRIM11*, *TNF*, and *FADD*). Assessment of copy number variations (CNV) disclosed different CNV frequencies of all NRGs. *MYC*, *IDH2*, *TRAF2*, *TNFSF10*, and *FADD* exhibited a significant increase in CNV while CNV was decreased in *CDKN2A*, *BRAF*, *TLR3*, *FAS*, and *RIPK1* (Figure 2B). The locations of 67 NRGs in the 23 chromosomes were further investigated. Chromosome 2 was the most common location, housing *MYCN*, *ALK*, *HAT1*, *CFLAR*, *CASP8*, and *IDH1* (Figure 2C).

The mRNA levels were different between GC and normal samples for the majority of NRGs (Figure 2D). Expression patterns of mRNA were consistent with CNV results. For instance, *MYC*, *FADD*, *TERT*, and *TNFSF10* showed high expression in GC samples with CNV gain status while *TLR3*, *KLF9*, and *BCL2* were expressed at low levels in GC samples with CNV loss status. Some NRGs showed upregulated mRNA expression with CNV loss, such as *CDKN2A* and *RIPK1*, while the *ID1* gene with CNV gain was downregulated. These findings suggest transcriptional regulation by other potential mechanisms in addition to CNV. The observed background of NRG gene expression and mutation in addition to CNV and chromosomal localization data in GC patients support an important role of NRGs in the oncogenesis of gastric cancer.

Necroptosis cluster identification in GC

Detailed information on GC patients from the TCGA and GSE84433 cohorts is presented in Supplementary Tables S2, S3. According to expression patterns of the 67 NRGs, GC patients were subdivided into different clusters using a consensus clustering algorithm (Supplementary Figure S1). The data indicate that k = 2 is an appropriate choice for the classification of patients into A (n = 340) and B (n = 360) subtypes (Figures 3A,B; Supplementary Table S4). At k = 2, the cumulative distribution function curve (CDF) increased gradually, and the sample size was relatively balanced between the two subgroups. In principal component analysis (PCA), subtypes A and B showed obvious differences in necroptosis transcription profiles (Figure 3C). Univariate Cox regression analysis was applied to determine the prognostic significance of the 67 NRGs in GC (Supplementary Table S5). Based on univariate Cox regression results, a necroptosis network was generated showing the interrelationship between each necroptosis gene and their predictive value in GC (Figure 3D). The Kaplan–Meier curve showed a relatively longer OS in subtype_A than subtype_B patients, but these differences were not significant (p = 0.604, Figure 3E). Most of the necroptosis genes were associated with the prognosis of gastric cancer patients. Kaplan–Meier survival analysis revealed an association of high expression of *PADD*, *FASLG*, *MLKL*, *RIPK3*, and *FAS* with better outcomes while low expression of *KLF9* was correlated with longer OS in GC (Figure 3F). Heatmap was performed to reveal relative differences in NRG expression and clinical features between the two subgroups (Figure 3G).

Identification and functional analysis of DEGs

To determine potential functional differences between the two necroptosis subgroups, we performed GSVA analysis, which showed significant enrichment of subtype_A in the steroid biosynthesis pathway (Figure 4A, Supplementary Table S6) and enrichment of subtype_B in essential immune-related pathways, including JAK-STAT signaling, B cell receptor signaling, T cell receptor signaling, natural killing cell-mediated cytotoxicity, and leukocyte transendothelial migration.

We identified 1239 necroptosis-related DEGs between subtypes A and B using the "limma" R package (Supplementary Table S7). Functional GO analysis revealed that these DEGs were mainly enriched in biological processes associated with tumor immunity, such as T cell activation, lymphocyte differentiation, regulation of T cell activation, and cell-cell adhesion, indicative of an important role of necroptosis in the tumor microenvironment (Figures 4B,C; Supplementary Table S8).

Necroptosis gene cluster based on prognostic DEGs

Univariate Cox regression analysis was performed based on 1239 DEGs and among which 606 genes related to OS for GC were selected (Supplementary Table S9). A new consensus

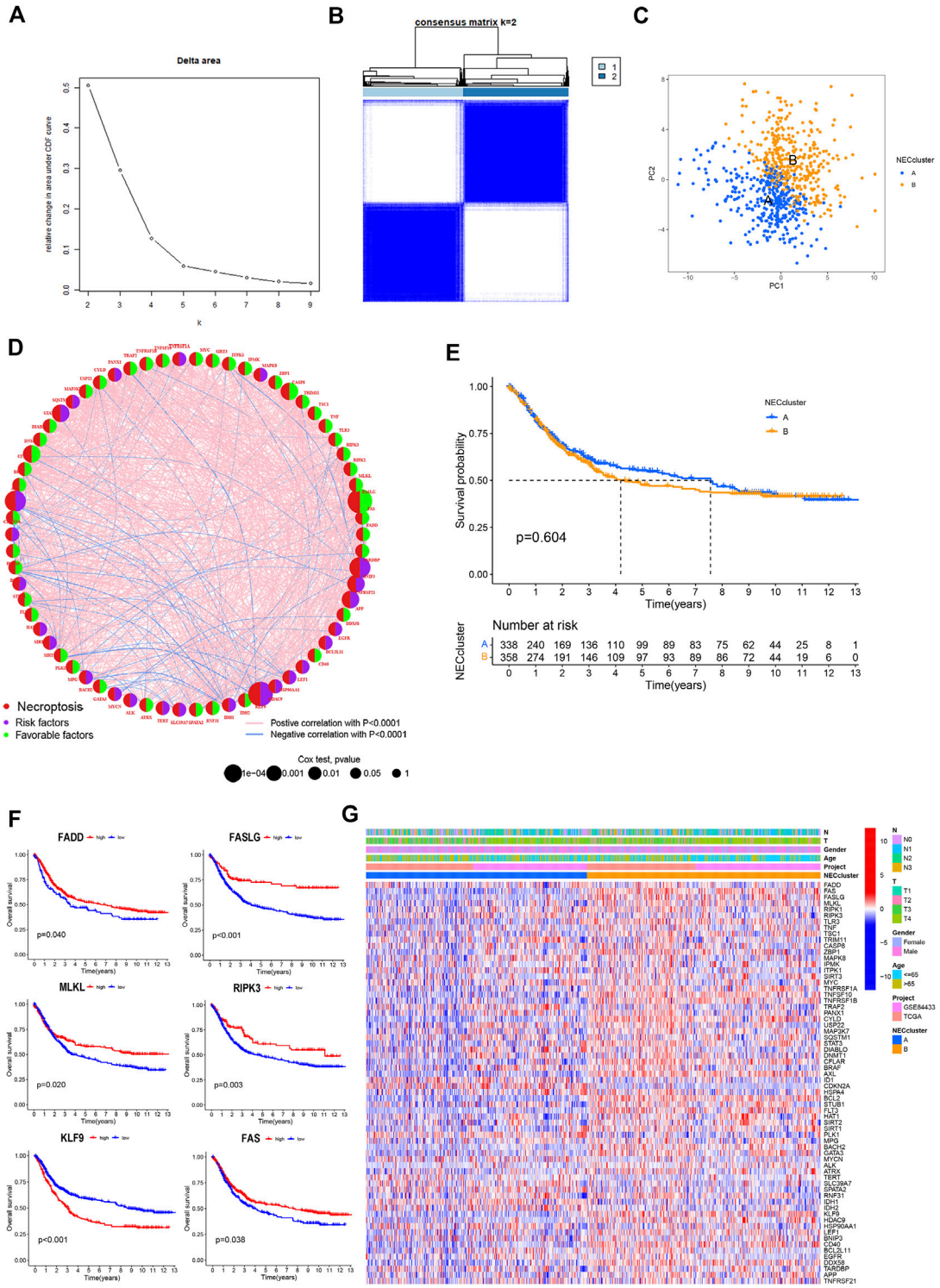


FIGURE 3
Necroptosis subtypes in GC. **(A,B)** Consensus matrix heatmap identifying 2 clusters (k = 2). **(C)** PCA analysis showing an obvious difference in necroptosis transcription profile. **(D)** Interactions among NRGs in GC. The line connecting the NRGs represents their interaction. Blue represents negative and pink represents positive correlations, respectively. **(E)** Survival analysis of NRG cluster according to OS in GC. **(F)** Kaplan-Meier curve showed that high expression of *PADD, FASLG, MLKL, RIPK3* and *FAS* were associated with better outcomes, while low expression of *KLF9* had a longer OS in GC. **(G)** Differences of the clinicopathological factors and NRGs expression between two necroptosis clusters.

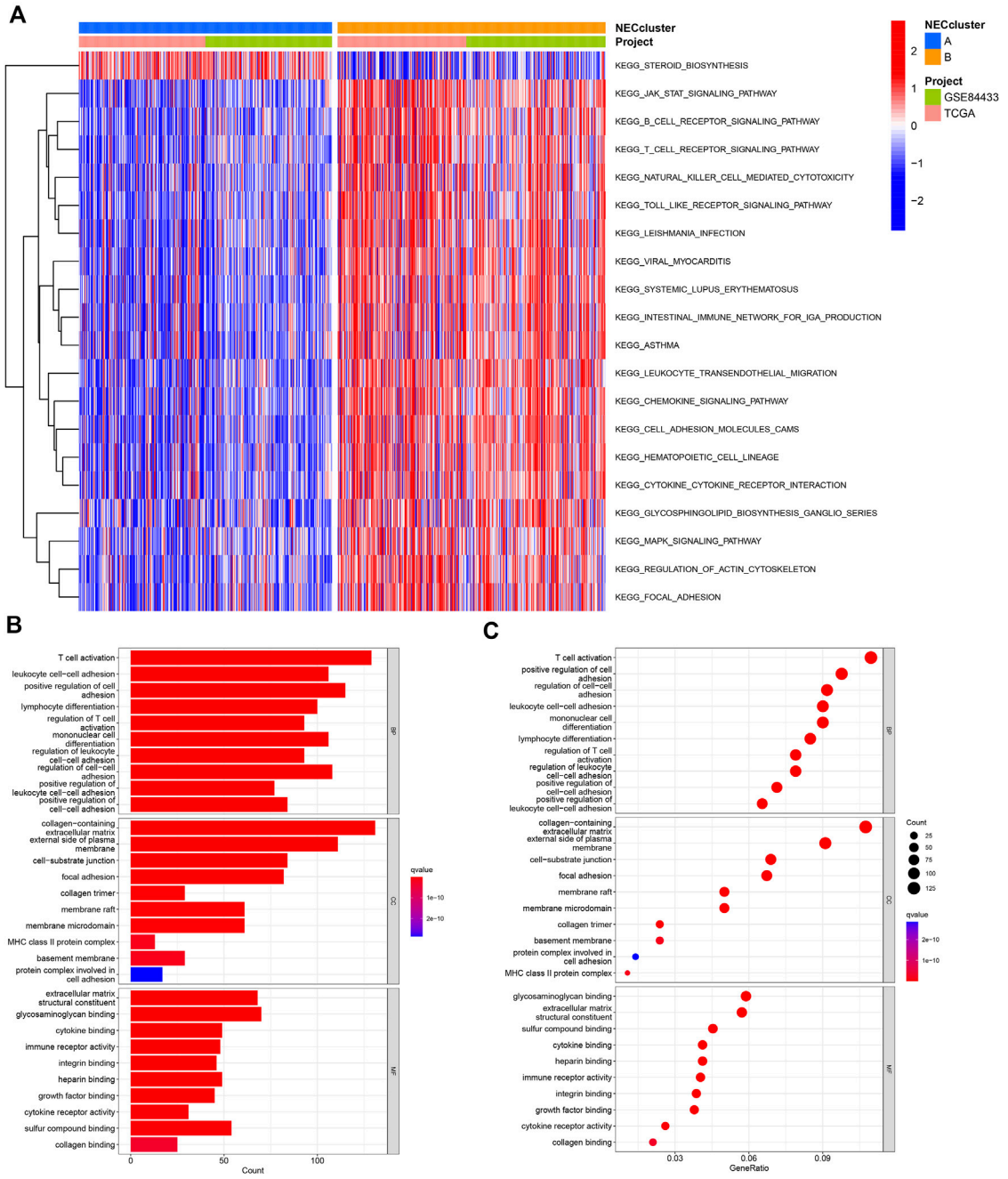


FIGURE 4
Functional analysis of necroptosis subtypes in GC. **(A)** CSVA analysis of the biological pathways between necroptosis subtypes. **(B,C)** GO and KEGG analysis of DEGs between two necroptosis subtypes.

clustering was conducted based on these prognostic NRGs and GC patients were divided into three necroptosis gene subtypes designated gene subtype_A, gene subtype_B, and gene subtype_C (k = 3, Figures 5A,B; Supplementary Figure S2; Supplemntary Table S10). PCA showed significant differences in the necroptosis transcription profiles of subtypes A, B, and C (Figure 5C). In Kaplan–Meier analysis, gene subtype_A had the highest OS while the shortest outcome was determined for subtype_C ($p < 0.001$, Figure 5D). The relative differences in T, N classification, gender, age, and necroptosis clustering between A, B, and C gene subtypes are presented in Figure 5E. In addition, significant differences in NRG expression among the three gene subtypes were observed (Figure 5F).

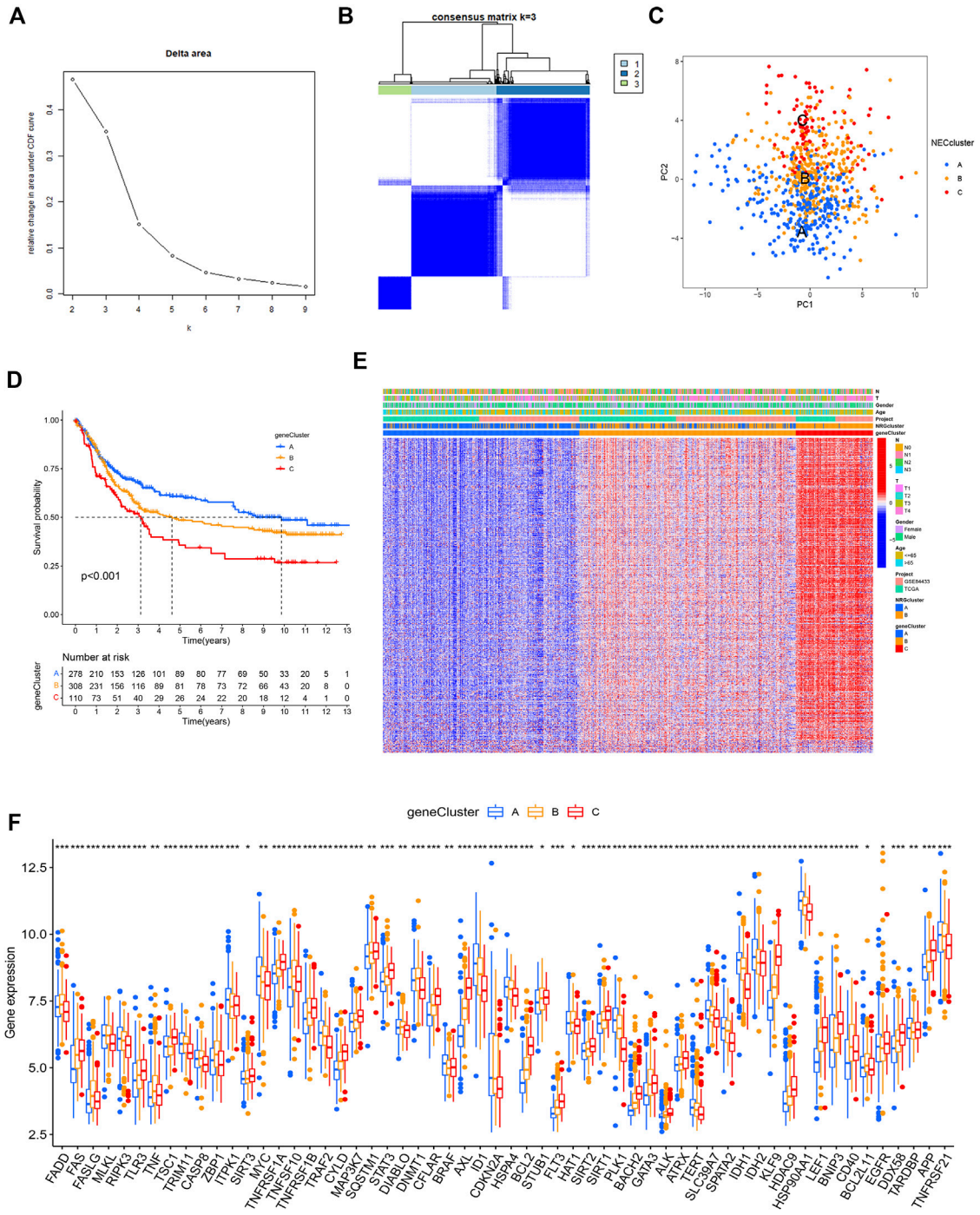


FIGURE 5
Gene clusters based on prognostic DEGs. **(A,B)** Consensus matrix heatmap identifying 3 clusters (k = 3). **(C)** PCA analysis showing an obvious difference in necroptosis transcription profile. **(D)** Survival analysis showing gene clusters were associated with OS in GC patients. **(E)** Differences of the clinicopathological factors and NRGs expression between three gene clusters. **(F)** Differences of NRGs expression between three gene clusters.

Prognostic NRG risk score

GC patients from TCGA and GEO cohorts were combined and randomly divided into training and validation cohorts at a 1:1 ratio. The training cohort was used for subsequent analyses. LASSO analysis was performed for 606 prognostic DEGs and 25 genes selected for subsequent multivariate Cox regression analysis (Figures 6A,B). Finally, 12 prognostic NRGs were

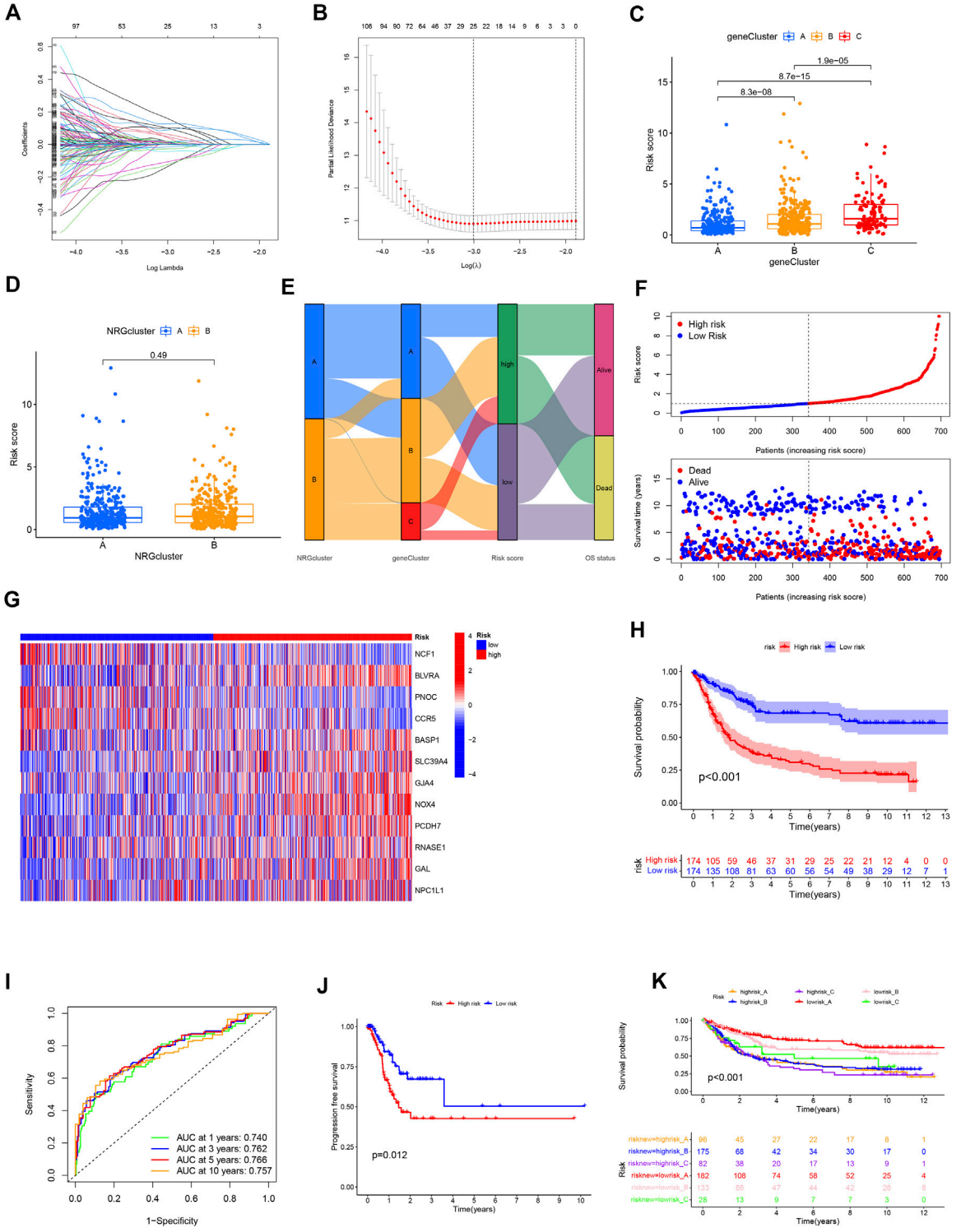


FIGURE 6
Construction of the NRG risk score in GC patients. **(A,B)** The LASSO regression analysis and partial likelihood deviance on the prognostic genes. **(C,D)** NRG risk score distribution in necroptosis and gene subtypes. **(E)** Correlation of necroptosis subtype, gene subtype, NRG risk score and survival outcomes. **(F)** Ranked dot and scatter plot showing the NRG risk score and survival status in all GC patient cohort. **(G)** The 12 prognostic NRG genes in the model between high and low-risk groups. **(H)** Survival analysis of the OS between the high and low-risk groups. **(I)** ROC analysis of the predictive ability of 1,3,5 and 10-year OS according to NRG risk score. **(J)** Survival analysis of the PFS between the high and low-risk groups. **(K)** Survival analysis according to the gene cluster and NRG risk level.

TABLE1 Patient Characteristics according to necroptosis risk score.

Characteristics	High risk N = 353	Low risk N = 342	*p*-value*
Age,years			0.413
Median (IQR)	64.0 (54, 70)	63.5 (56, 70)	
Range	30, 90	27, 90	
Gender			0.811
Female	124 (35%)	117 (34%)	
Male	229 (65%)	225 (66%)	
Grade			0.981
G1	4 (1.1%)	4 (1.2%)	
G2	61 (17%)	63 (18%)	
G3	102 (29%)	97 (28%)	
missing	186 (53%)	178 (52%)	
Stage			0.186
Stage I	18 (5.1%)	28 (8.2%)	
Stage II	54 (15%)	48 (14%)	
Stage III	63 (18%)	71 (21%)	
Stage IV	22 (6.2%)	12 (3.5%)	
missing	196 (56%)	183 (54%)	
T			0.006
T1	9 (2.5%)	19 (5.6%)	
T2	51 (14%)	56 (16%)	
T3	100 (28%)	123 (36%)	
T4	187 (53%)	142 (42%)	
missing	6 (1.7%)	2 (0.6%)	
N			0.002
N0	65 (18%)	104 (30%)	
N1	125 (35%)	119 (35%)	
N2	100 (28%)	69 (20%)	
N3	55 (16%)	42 (12%)	
missing	8 (2.3%)	8 (2.3%)	
M			0.159
M0	147 (42%)	154 (45%)	
M1	16 (4.5%)	7 (2.0%)	
missing	190 (54%)	181 (53%)	
NECcluster			0.081
A	160 (45%)	178 (52%)	
B	193 (55%)	164 (48%)	
geneCluster			<0.001
A	96 (27%)	181 (53%)	
B	175 (50%)	133 (39%)	
C	82 (23%)	28 (8.2%)	

*Welch Two Sample t-test; Fisher's exact test.

identified (Supplementary Table S11), including 9 high-risk genes (*NPC1L1*, *GAL*, *RNASE1*, *PCDH7*, *NOX4*, *GJA4*, *SLC39A4*, *BASP1*, and *BLVRA*) and 3 low-risk genes (*NCF1*, *PNOC*, and *CCR5*). Correlation of the 12 model genes expression with OS in the whole patients set was also presented in Supplementary Figure S3. Based on the results of multivariate Cox regression analysis, the NRG risk score was calculated as follows:

NRG risk score = (−0.43594* expression of *NFC1*) + (−0.38714* expression of *PNOC*) + (−0.27229* expression of *CCR5*) + (0.46847* expression of *BLVRA*) + (0.29272* expression of *BASP1*) + (0.28869* expression of *SLC39A4*) + (0.28408* expression of *GJA4*) + (0.24301* expression of *NOX4*) + (0.18931* expression of *PCDH7*) + (0.18843* expression of *RNASE1*) + (0.09485* expression of *GAL*) + (0.07200* expression of *NPC1L1*).

Based on the median NRG risk score, all GC patients were divided into high-risk and low-risk groups. The majority of necroptosis-related genes were differentially expressed between the two risk groups (Supplementary Figure S4). Correlations between risk level with clinicopathological factors were presented in Table1. It showed that high NRG risk score was positively related with advanced T stage, N stage and gene cluster C subtype. NRG risk score was significantly different among the three gene clusters. The NRG score was highest in gene cluster_C and lowest in gene cluster_A (Figure 6C) while no significant differences were observed between necroptosis clusters A and B (Figure 6D).

The Sankey plot allowed visualization of the interrelationships among two necroptosis clusters, three gene clusters, risk_score, and overall survival status (Figure 6E). The distribution plot showed that OS of GC decreased with increased NRG risk score (Figure 6F). For all patient cohorts, a heatmap was generated to establish the relationships of the 12 prognostic marker genes with NRG risk groups. *BLVRA*, *SLC39A4*, *GJA4*, *NOX4*, *PCDH7*, and *GAL* were highly expressed in the high-risk group while *NCF1*, *PNOC*, and *CCR5* were highly expressed in the low-risk group (Figure 6G). The correlation results in the training and validation sets are presented in Supplementary Figure S5. Kaplan–Meier curve analysis showed significantly longer OS in the low-risk relative to the high-risk group ($p < 0.001$, Figure 6H). ROC curves indicated that NRG risk score had promissing value in predicting OS of GC patients, with AUC at 1, 3, 5, and 10 years of 0.740, 0.762, 0.766, and 0.757, respectively (Figure 6I). Progression free survival (PFS) was also higher in low-risk than high-risk group GC patients (Figure 6J). We obtained similar results with the whole patient and validation sets (Supplementary Figures S6, S7). Risk group and gene cluster data were combined for the subclassification of GC patients into six different subgroups. The lowrisk_A group had better OS relative to the other subgroups. Low-risk groups could be further divided into three survival outcome groups while the high-risk group consistently showed the poorest survival rates within our patient population (Figure 6K).

Validation of the expression levels of the 12 NRGs in the prognostic model

The expression levels of 12 prognostic genes were measured in 16 GC tissues and 16 paired adjacent normal tissues by RT-qPCR. As shown in Supplementary Figure S8, all of the 12 NRGs were significantly differentially expressed between cancer and normal tissues. The expression of *PNOC* and *RNASE1* were downregulated and other genes (*BASP1*, *BLVRA*, *CCR5*, *GAL*, *GJA4*, *NCF1*, *NOX4*, *NPC1L1*, *PCDH7*, and *SLC39A4*) were upregulated in GC tissues than corresponding normal tissues. Results of RT-qPCR were presented in Supplementary Table S16.

Tumor microenvironment in the high- and low-risk groups

CIBERSORT algorithm results revealed positive associations of the prognostic NRG risk score with resting CD4 memory cells, resting mast cells, activated mast cells, and M2 macrophages, and negative correlations with $CD8^+$T cells, follicular helper T cells, activated memory $CD4^+$T cells, plasma cells, and memory B cells (Figure 7A, Supplementary Table S12). The ESTIMATE algorithm was additionally performed to simulate TIME. The results showed that high NRG risk score was associated with low immune score and high stromal score in GC samples (Figure 7B, Supplementary Table S13), suggesting that the high-risk group has a relatively good immune microenvironment. The relationship between 12 prognostic necroptosis genes and 22 human immune-related cells was further examined (Figure 7C). The majority of immune cells were significantly positively or negatively regulated with the 12 genes. The relationship between NRG prognostic risk_score and immune checkpoints was further assessed (Figure 8A). Analysis of 33 immune checkpoint genes led to the identification of 22 that were differentially expressed between the two risk groups.

Correlations of NRG risk score with MSI, CSC index, HLA gene expression, and TMB score

Experiments were performed to determine the relationship of NRG risk score with immunotherapeutic biomarkers, such as MSI, CSC index, HLA gene expression, and TMB score. Notably, a low NRG risk score was significantly correlated with MSI-H status. In the low-risk group, MSI-H status was 24%, which was higher than that in the high-risk group (Figure 8C), and the median risk_score of MSI-H was significantly lower than that of MSI-L/MSS groups (Figure 8D). The NRG risk score showed a negative linear correlation with the CSC index ($R = -0.32$, $p < 0.001$, Figure 8E), indicating that GC cells with high NRG risk score have less distinct stem cell properties and a higher level of cell differentiation. We additionally performed a correlation analysis between NRG risk score and HLA gene expression. Our results showed lower expression of *HLA-DRB5* in the high-risk relative to the low-risk group (Figure 8B, Supplementary Table S14). The TMB score was negatively associated with the NRG gene cluster in this study (Supplementary Figure S9) and significantly lower in the high-risk than the low-risk group (Figure 8F). Next, we analyzed the distribution of somatic mutations between the two necroptosis risk groups in the TCGA GC cohort. Low-risk samples showed relatively higher mutation frequency than the high-risk group (90.06% vs. 86.98%, Figures 7G,H). Multiple genes displayed mutations in gastric cancer cells but the types and frequencies were

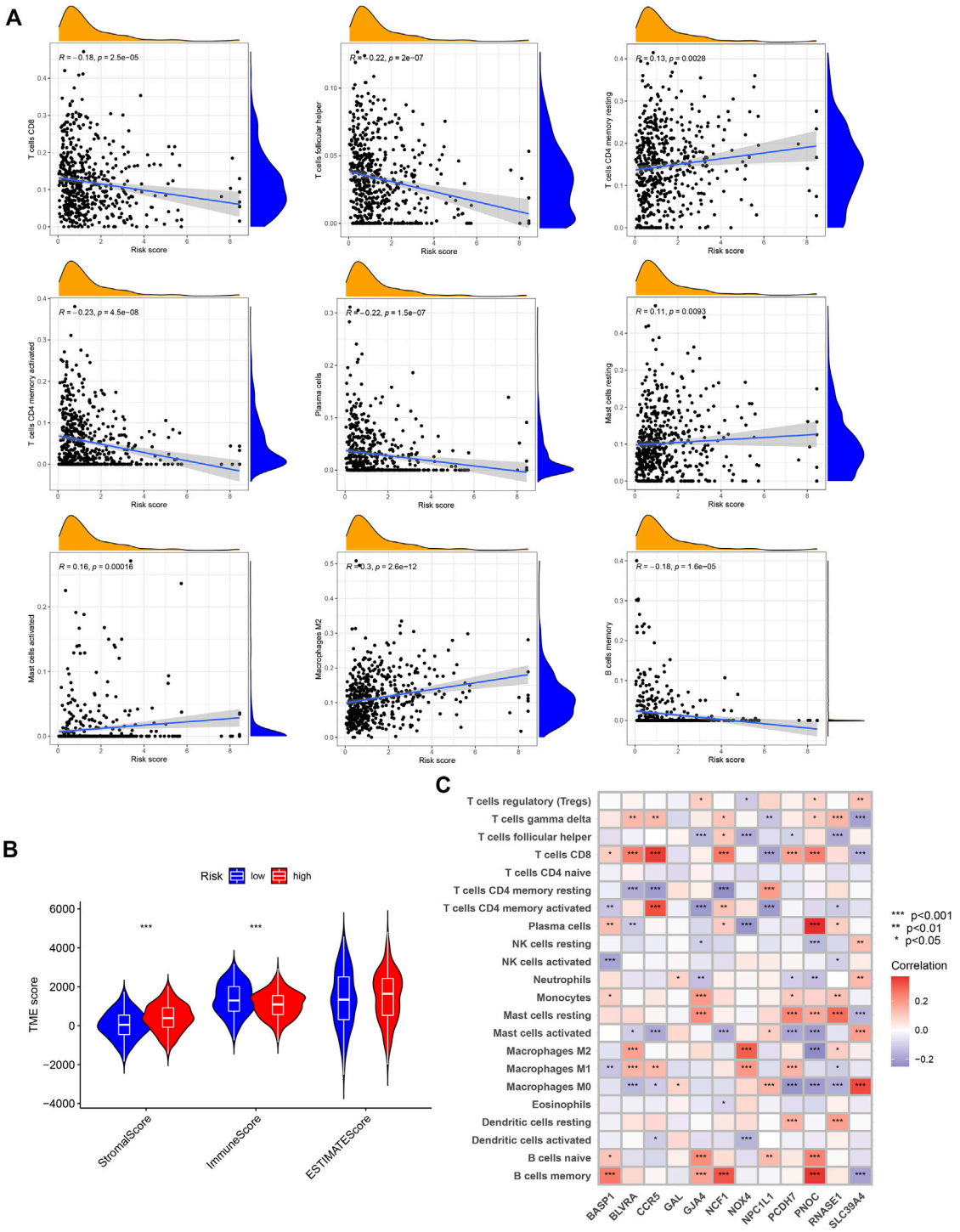


FIGURE 7
Association of TIME with risk_score. **(A)** Correlation of infiltrating immune cells with NRG risk score. **(B)** Correlation of NRG risk score with immune/stromal scores. **(C)** Correlation of immune cells with 12 NRGs.

distinct. For example, the frequency of *TTN* mutation was 47% in the high-risk group and 36% in the low-risk group. The top 5 mutated genes in the high-risk group were *TP53*, *TTN*, *MUC16*, *SYNE1*, and *LRP1B*, while *TTN*, *TP53*, *MUC16*, *ARID1A*, and *LRP1B* were the top 5 genes showing mutations in the low-risk group.

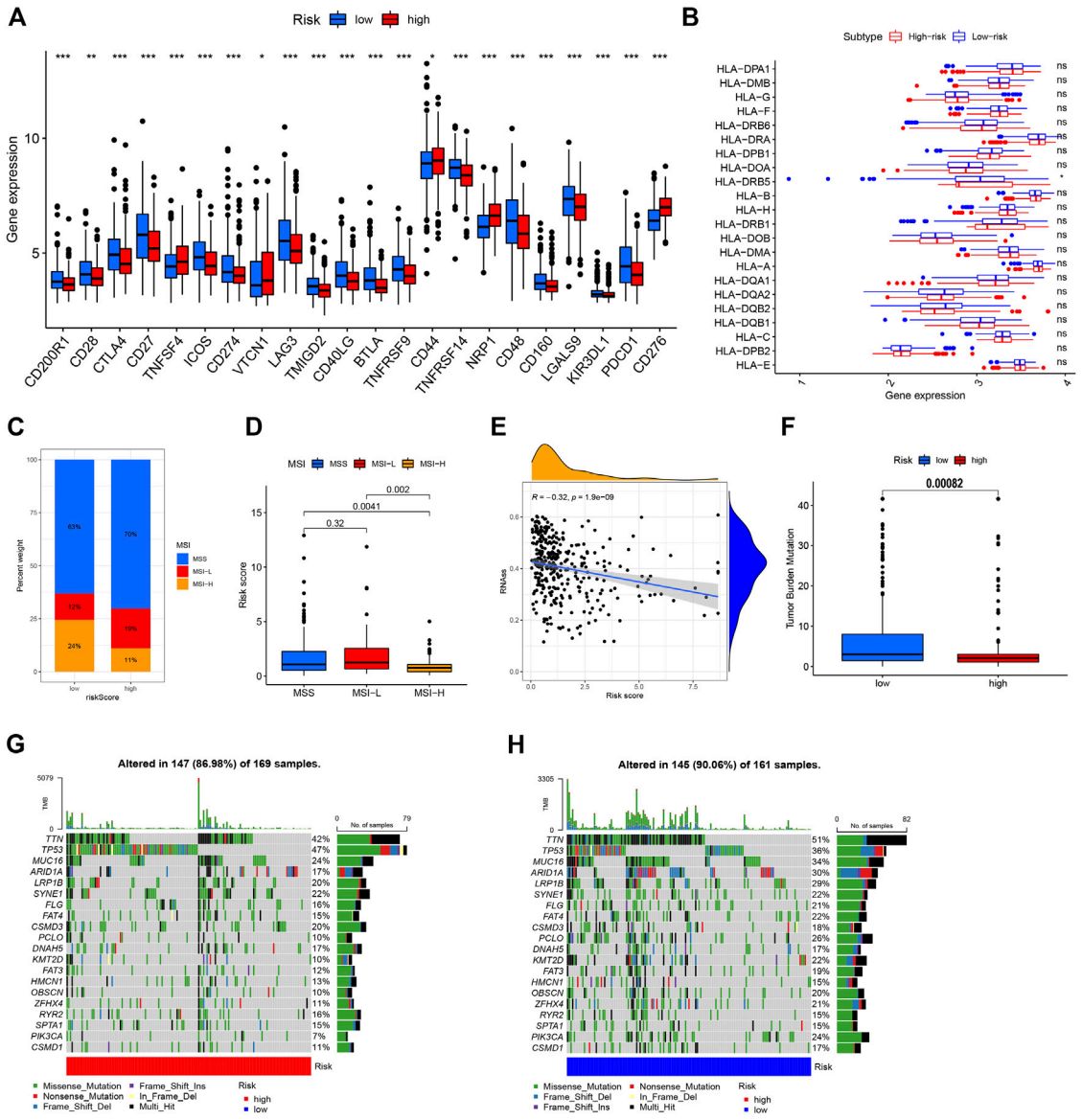


FIGURE 8
Comprehensive analysis of NRG risk score in GC. **(A)** Correlation of NRG risk score and immune checkpoint gene expression. **(B)** Correlation of NRG risk score and HLA gene expression. **(C,D)** Correlation of NRG risk score and MSI status. **(E)** Correlation of NRG risk score and CSC index. **(F)** Correlation of NRG risk score and TMB. **(G,H)** Somatic mutation features in the high- and low-risk groups.

Drug susceptibility analysis

We performed drug susceptibility analysis to select promising chemotherapy or targeted drugs for high- and low-risk groups of GC. Interestingly, patients in the high-risk group had lower IC_{50} values for docetaxel, lapatinib, pazopanib, dasatinib, and imatinib (Figures 9A–E) while those in the low-risk group had significantly lower IC_{50} values for gefitinib, metformin, bosutinib, lenalidomide, and salubrinal (Figures 9F–I). These results support the utility of NRG risk score in the prediction of drug sensitivity and selection of potential beneficiaries of specific treatment drugs among GC patients.

NRG risk score predicts response to immunotherapy

As the important biomarker for immunotherapy, we firstly performed K-M analysis based on TMB and MSI status in GC

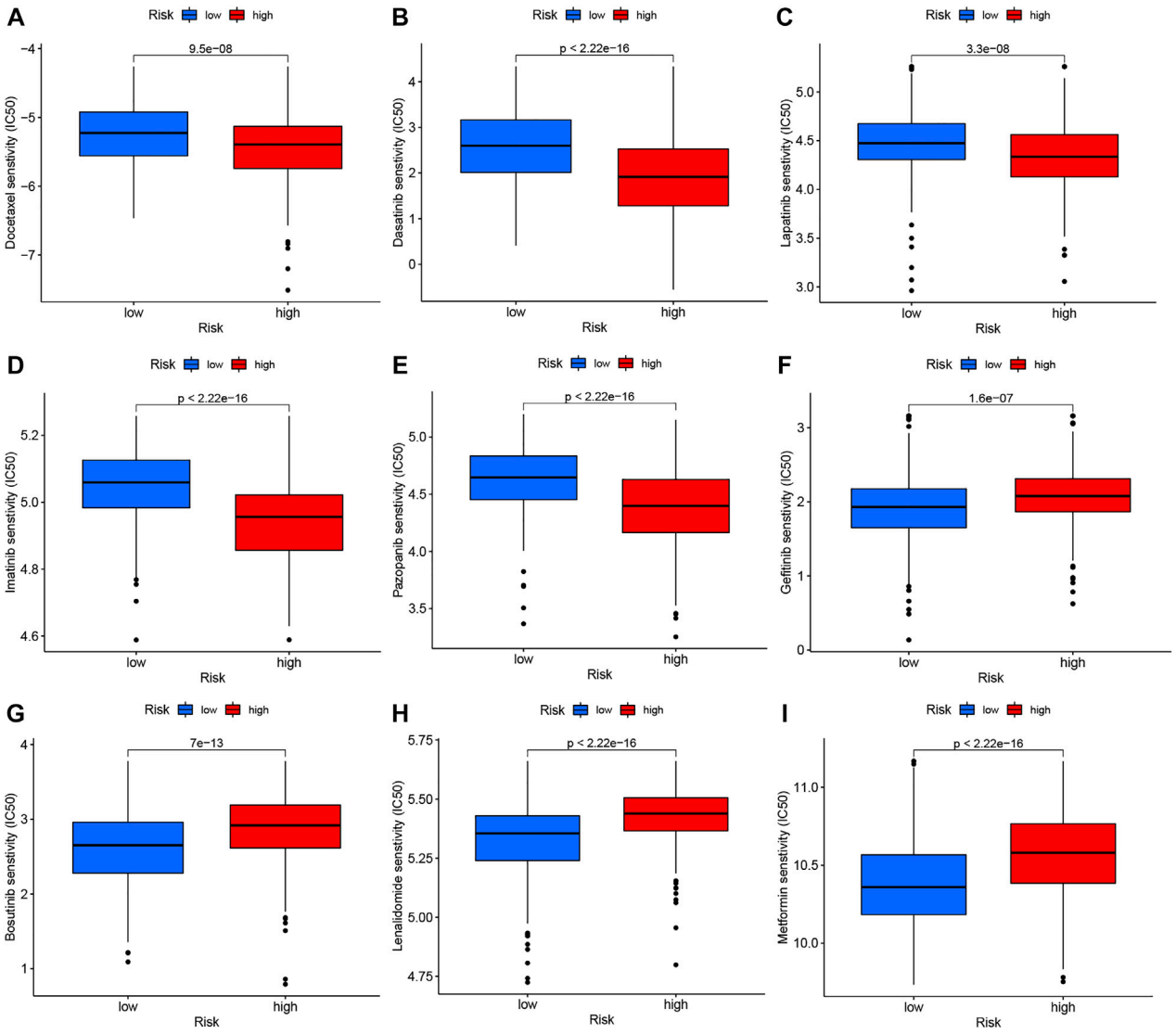


FIGURE 9
Drug susceptibility analysis of chemotherapy and targeted therapy in high- and low-risk groups. **(A)** Docetaxel, **(B)** Dasatinib, **(C)** Lapatinib, **(D)** Imatinib, **(E)** Pazopanib, **(F)** Gefitinib, **(G)** Bosutinib, **(H)** Lenalidomide, and **(I)** Metformin.

patients. It showed that High-TMB patients had a significantly better OS than low-TMB GC patients (Figure 10A, $p < 0.001$) and when combining with risk levels (Figure 10B). While, there was no difference between MSI-H and MSI-L/MSS patients (Figure 10C, $p = 0.144$), even combining with risk levels (Figure 10D, $p = 0.424$).

According to results of TCIA database analysis, IPS scores were download and correlation with NRG risk score was also evaluated. IPS-PD-1/PD-L1/PD-L2_pos and IPS-CTLA-4_pos was higher in low-risk group patients than in high-risk group patients (Figures 10E,F), indicating that low-NRG_risk gastric cancer patients might have higher tumor immunogenicity and benefit from anti- PD-1/PD-L1/PD-L2 and anti- CTLA-4 blocker immunotherapy. In Imvigor210 cohort, we also found that NRG risk score was significantly correlated with patients' objective response of immunotherapy. Risk score in the response group (CR/PR, complete response and partial response) group was obviously lower than none-response group (SD/PD, stable disease and progressive disease, $p = 0.046$, Figure 10G), expecially between CR and PD groups patients (Figure 10H, $p = 0.023$), indicating that NRG risk score could act as a potentially promising therapeutic biomarker of immunotherapy in gastric cancer.

Next, we evaluated the prognostic value of NRG risk score in a melanoma treated by anti-PD-1 cohort (GSE78220). Percentage of response to ICIs in the low-risk group was

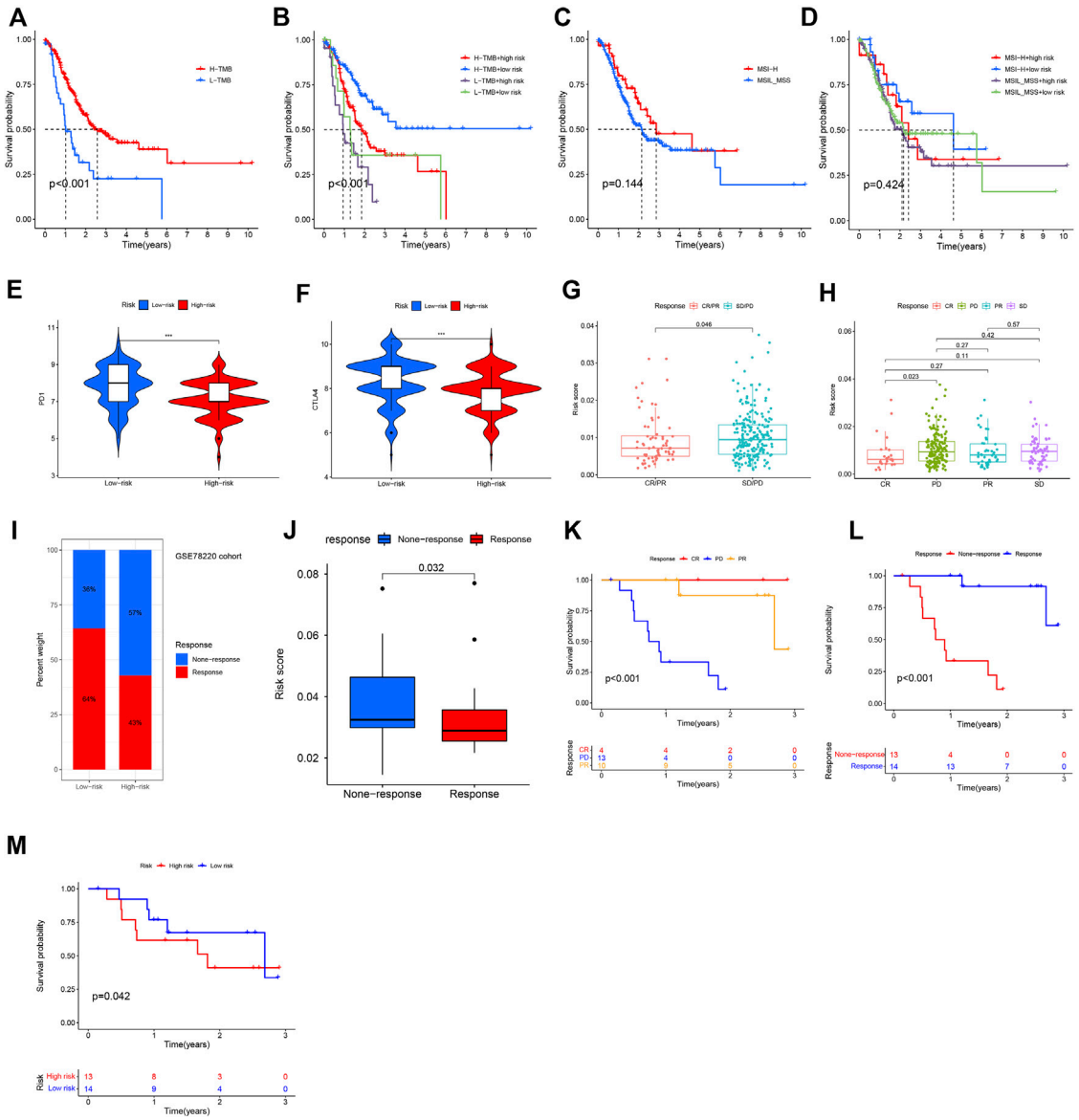


FIGURE 10
Immunotherapy response based on necroptosis risk score. **(A–D)** K-M analysis based on TMB, MSI status, or their combination with NRG risk score. **(E,F)** Predicting of response to anti-CTLA-4 and anti-PD-1/PD-L1 therapy (IPS score) in the low-risk score and high-risk score groups in TCGA-GC patients in this study. **(G,H)** Correlation of immunotherapy response with NRG risk score according to IMvigor210 database. **(I,J)** Correlation of NRG risk score with ICIs response. **(K,L)** Survival in CSE78220 cohort related to ICIs response. **(M)** Survival in CSE78220 cohort related to NRG risk score.

higher than high-risk group (Figure 10I), and the median NRG risk score in response group patients was significantly lower than none-response group patients (p = 0.038, Figure 10J). K-M analysis showed that patients with a response to ICIs (CR and PR) had a better OS than none-response (PD) group patients (Figures 10K,L, $p < 0.001$). At last, we assess the NRG risk score in survival of ICIs treated patients and the promising result showed that it was significantly correlated with OS in this study (p = 0.042, Figure 10M), indicating that NRG risk score might be a useful biomarker for immunotherapy.

Construction and validation of nomogram based on NRG risk score

We integrate NRG risk score and the common clinicalpathological factors, including age, gender, T stage, N

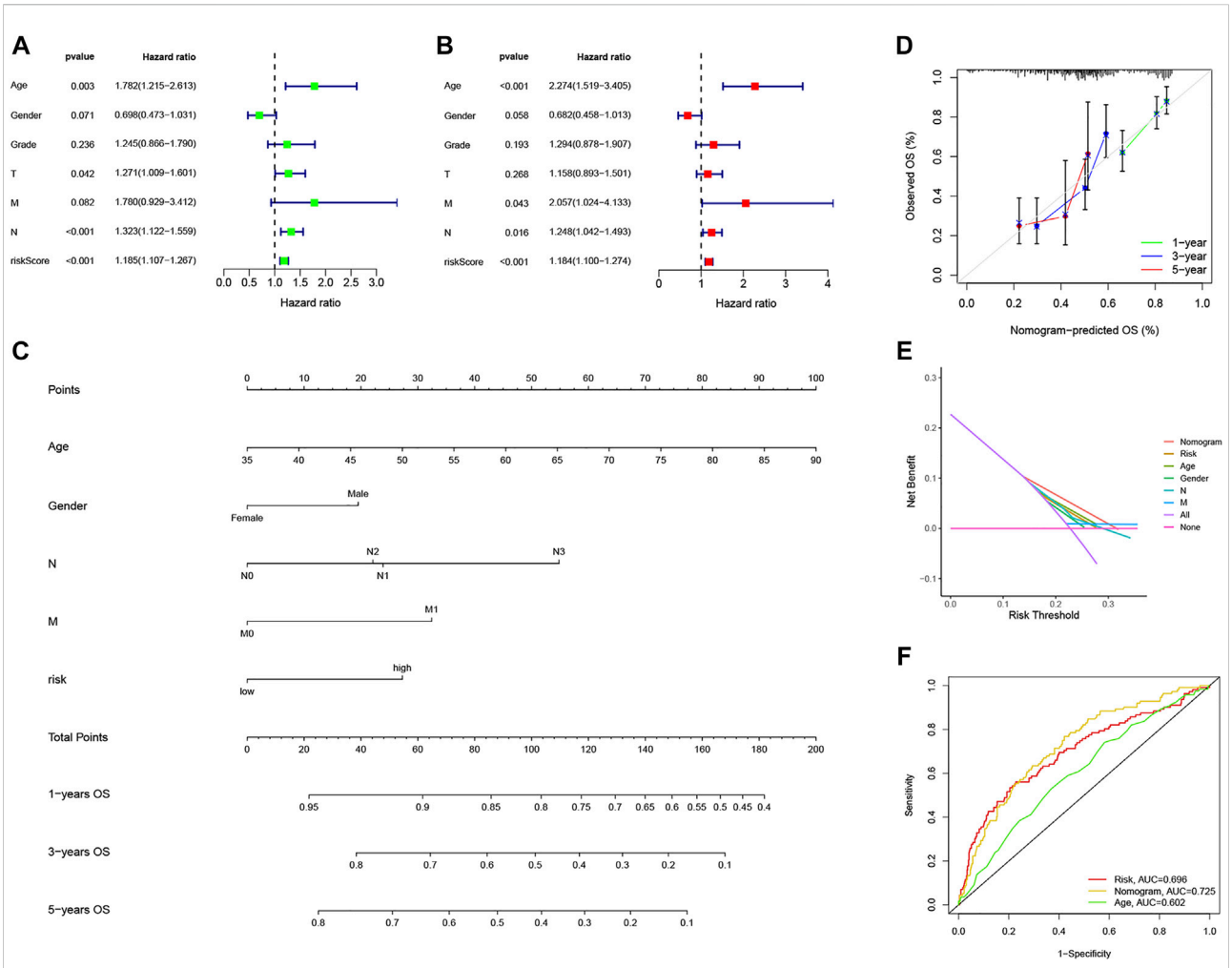


FIGURE 11
Construction and validation of the nomogram based on NRG risk score. **(A,B)** Forest plots of univariate and multivariate Cox Regression analysis. **(C)** Nomogram predicting 1-, 3-, and 5-year OS in GC patients. **(D)** Calibration curve for prediction of 1-,3- and 5-year OS. **(E)** Decision curve analyses of the nomogram. **(F)** ROC curves for predicting OS in GC patients.

stage, M stage and pathological grade, and performed univariate Cox regression analysis and subsequent multivariate Cox regression analysis. It revealed that NRG risk score was an independent prognostic factor for GC patients in this study ($p < 0.001$, Figures 11A,B). Factors showed significance in the multivariate analysis were incorporated into a nomogram. For gender acts as an easy and common factor in clinical practice, we also consider it in the model construction (Figure 11C). The calibration curves of the nomogram indicated excellent consistency with the standard curve between predicted and actual 1-, 3-, and 5-year OS rates in GC patients (Figure 11D). DCA was conducted to evaluate the predictive value of the nomogram in clinical decision-making (Figure 11E). Notably, the nomogram showed better reliability than the common clinicopathological factors or NRG risk score alone. The AUC value of the nomogram was 0.725, which was higher than the NRG risk score (0.696, Figure 11F), indicating a better predictive value than NRG risk score alone.

Discussion

Apoptosis is a common pathway of programmed cell death in the body. Traditional chemotherapeutic drugs mainly exert their effects by promoting apoptosis of tumor cells, with resistance to apoptosis identified as the predominant mechanism underlying tumor drug resistance (Mortezaee et al., 2019; Carneiro and El-Deiry, 2020). A comprehensive investigation of the pathways of tumor cell apoptosis has led to the gradual discovery of novel cells death modes, such as pyroptosis, ferroptosis, cuproptosis, and necroptosis. In contrast to apoptosis, necroptosis rarely presents with cytoplasmic shrinkage, chromatin condensation, nuclear

fragmentation, cell membrane blebbing, or shedding of apoptotic bodies (Pasparakis and Vandenabeele, 2015; Su et al., 2016). Necroptosis is reported to play potential roles in several diseases including inflammatory bowel disease, multiple dermatosis, acute kidney injury, inflammatory response syndrome, and atherosclerosis. Necroptosis is also implicated in cancer, with reports of its involvement in oncogenesis, adaptive immunity, cancer subtype, progression, and metastasis. Evidence from several studies supports the protective role of necroptosis against tumor development (Seifert et al., 2016; Strilic et al., 2016).

Gastric cancer is a fatal disease with poor overall survival statistics worldwide. Several factors (including low rates of early diagnosis, high intratumor heterogeneity, and drug resistance) contribute to the poor survival outcomes of GC patients (Johnston and Beckman, 2019; Thrift and El-Serag, 2020). The majority of patients present with distant metastasis at the time of diagnosis and therefore undergo chemotherapy. However, due to dose-limiting toxicity and limited responses, patients commonly experience disease progression following chemotherapy. The development of targeted therapy for advanced gastric cancer focuses on two aspects: *HER-2* and antiangiogenesis. Targeted therapy for *HER-2* as a first-line treatment is reported to improve survival in GC patients with positive *HER-2* expression (Sawaki et al., 2012). Ramucirumab, an antagonist of vascular endothelial growth factor receptor 2 (*VEGFR2*), has been shown to inhibit ligand-induced VEGFR2 activation, thereby suppressing ligand-induced proliferation and migration of human endothelial cells. Ramucirumab has been approved as a second-line treatment for gastric cancer and gastroesophageal junction carcinoma (Wilke et al., 2014). Although several new drugs have been identified, the availability of appropriate targeted therapy remains limited to a small number of GC patients. The advent of immunotherapy offers new options, and promising survival outcomes have been obtained for some patients receiving immunotherapy combined with chemotherapy according to PD-L1 expression or MSI status. In clinical practice, the identification of patient subgroups that could potentially benefit from immunotherapy is critical for improving outcomes. The field of gastric cancer therapy thus faces great opportunities and challenges (Zhao et al., 2019; Xie et al., 2021).

This study investigated the genetic features and transcriptional alterations in necroptosis-related genes in GC. Previous studies have indicated that cancer cells eradicate necroptosis by downregulating NRGs to evade necroptosis-induced cell death mainly involving RIP1, RIP3, and MLKL. In GC, the majority of NRGs were expressed at higher levels in tumor than normal samples. An earlier report demonstrated the downregulation of MLKL, a key regulator of the necroptosis axis, in PDAC, which was associated with lower OS (Colbert et al., 2013). Our experiments showed higher expression of MLKL in GC than in normal samples. Moreover, lower expression was associated with better survival outcomes, suggesting that MLKL has a promising prognostic ability. However, expression patterns of NRGs may differ among tumor types. Based on NRG expression, GC samples were divided into two cluster subtypes (A and B). While differences in OS were not significant between the two groups, we observed an extension of median survival time in group A. Function of enrichment analysis revealed enrichment of several immune activation-related signals in cluster B, including B cell receptor signaling, T cell receptor signaling, NK cell-mediated cytotoxicity, and JAK-STAT signaling, supporting potential correlation of necroptosis-related cancer subtypes with the tumor immune microenvironment. These findings provide novel insights and directions for future research. We further identified 1239 necroptosis-related DEGs between subtypes A and B, among which 606 were proven to be associated with OS in GC patients. According to the 606 prognostic DEGs, patients were further divided into three necroptosis gene clusters (gene subtype_A, gene subtype_B, and gene subtype_C). The gene cluster system was highly associated with OS of GC patients and most NRGs were differentially expressed among the three gene clusters. The collective results support the potential of NRGs as prognostic and immunotherapeutic biomarkers.

Next, we constructed and validated a useful prognostic model based on NRG risk score. NRG risk scores between the two cluster systems were additionally examined, revealing that survival was not significantly different between clusters A and B but differed among the three gene clusters. Based on NRG risk score, patients were eventually classified into high- and low-risk groups. Patients in the low-risk group had significantly better OS than those in the high-risk group. We subsequently explored the tumor immune microenvironment between the high- and low-risk groups. Several human immune cells were correlated with NRG risk score, including resting CD4 memory cells, resting mast cells, activated mast cells, M2 macrophages, $CD8^{+}$T cells, follicular helper T cells, and activated memory $CD4^{+}$T cells, plasma cells, and memory B cells. Marked differences in the immune score, stromal score, MSI status, TMB score, CSC index, HLA gene expression, and immune checkpoint expression levels were recorded between high- and low-risk groups. Drug susceptibility analysis may be used in clinical practice to select potentially effective drugs for GC based on necroptosis gene expression. Here, the first prognostic nomogram based on necroptosis in GC patients was constructed and its predictive ability assessed.

Current treatments for gastric cancer include resection of the tumor via gastroscopy or surgery, chemotherapy, targeted therapy, and radiotherapy. Despite considerable progress in the development of treatment strategies for gastric cancer, therapeutic outcomes remain unsatisfactory. Immunotherapy has been documented as an effective strategy for various tumor types, including gastric cancer, due to its precise effects on the tumor microenvironment and persistent response. While immunotherapy offers new hope, many patients do not benefit from this mode of treatment, mainly due to tumor heterogeneity and complex tumor immune microenvironments, which should be a focus of further research. The TIME is extremely complex and includes surrounding blood vessels, immune cells, fibroblasts,

bone marrow-derived inflammatory cells, various signaling molecules, and an extracellular matrix (Pitt et al., 2016; Cassim and Pouyssegur, 2019). The tumor microenvironment has been shown to play an important role in malignant progression, immune escape, and immunotherapy resistance by altering the ratio of immunosuppressive and cytotoxic responses in the vicinity of the tumor. Rather than forming apoptotic bodies during cell apoptosis, necroptosis is accompanied by rupture of the cell membrane and release of tumor neoantigens, which can trigger strong inflammatory and anti-tumor immune responses (Gong et al., 2019). Based on the necroptosis risk_score in this study, we showed that TIME and the majority of key immune cells were significantly different between high- and low-risk groups.

T cells, which are essential immune cells, play critical roles in antigen recognition, presentation, and tumor cell killing in gastric cancer (Thommen and Schumacher, 2018; Choi et al., 2020). In our experiments, $CD8^+$ and activated memory $CD4^+$ T cells were negatively correlated with NRG risk score, indicating a critical role of necroptosis in the anti-tumor immune response of gastric cancer. Limited research to date has focused on the anti-tumor role of B cells. Recent accumulating evidence suggests that B cells may serve as an important prognostic factor. B cells are active participants that fundamentally coordinate the immune response and resistance to tumors under certain conditions, mainly by generating tumor-specific antibodies. However, specific B cell subsets and antibody specificity can also inhibit anti-tumor immunity and promote tumor growth (Sharonov et al., 2020). Infiltrating B lymphocytes are important components of tertiary lymphoid structures (TLS) in tumor tissues. TLSs are ectopic lymphoid organs formed in non-lymphoid tissues during chronic inflammation and tumor growth that is composed of T, B, follicular dendritic, DC-$LAMP^+$ dendritic, and other cells. In multiple tumor types, B cell infiltration and TLS formation are positively correlated with patient response to immunotherapy, highlighting the critical role of B cells and TLSs in anti-tumor immunity and providing a basis for new theories and strategies of tumor immunotherapy (Sautès-Fridman et al., 2019; Helmink et al., 2020). Interestingly, in the current study, the level of infiltrating memory B cells was significantly negatively correlated with NRG risk score, indicating that necroptosis also participates in B cell anti-tumor processes. This result is consistent with our finding that necroptosis is correlated with B cell receptor signaling. To our knowledge, this is the first report to document the involvement of necroptosis in B cell anti-tumor processes.

Tumor-associated macrophages (TAMs), the major component of myeloid cells in tumors, constitute two major phenotypes: M1 (inhibiting cancer progression) and M2 (promoting cancer progression) (Pan et al., 2020). TAMs exert both pro-tumor and anti-tumor effects and may therefore serve as attractive potential targets for tumor therapy. Zhao and co-workers reported that TAMs isolated from gastric cancer tissues predominantly display an M2 phenotype (Li et al., 2019) and gastric cancer-derived mesenchymal stromal cells promote metastasis and epithelial-mesenchymal transition (EMT) by triggering M2 TAM polarization through the IL-6/IL-8-JAK2-STAT3 signaling pathway. Moreover, blockage of M2 TAMs could reactivate $CD8^+$ T cells against immunosuppressive tumors (Viitala et al., 2019) and infiltrating levels of M2 TAMs in gastric cancer were associated with the 5-year survival rate (Junttila et al., 2020). Here, we observed increasing infiltration of M2 TAMs in the high necroptosis risk_score group, suggesting that necroptosis may participate in the antitumor immune response via regulation of macrophages. However, further research is warranted to establish precise molecular mechanisms.

Immunotherapy has become an indispensable element of gastric cancer treatment. ICIs are an important aspect of immunotherapy, including anti-PD1 and anti-CTLA4 antibodies, which have continuously improved the survival of gastric cancer patients and progressed from back-line to front-line status in clinical practice. Several researchers have focused on the selection of effective immunotherapy biomarkers to date. MSI-H/dMMR is generally recognized as a good predictive biomarker in gastrointestinal tumors. In previous studies, MSI-H accounted for 19.09% of GC cases in a TCGA cohort and 5.75% cases in a Chinese cohort, which were higher relative to the proportion of MSI-H in other solid tumor types (Bonneville et al., 2017; Zang et al., 2019). Previous KEYNOTE-061 and KEYNOTE-062 clinical trials reported higher OS, progression-free survival (PFS), and objective response rate (ORR) with anti-PD1 therapy than chemotherapy for MSI-H gastric cancer patients (Shitara et al., 2018; Shitara et al., 2020). In the present investigation, MSI-H patients accounted for a higher percentage of the low-risk than high-risk group and the median risk_score of MSI-H was markedly lower than that of MSI-L/MSS groups. The majority of immune checkpoint genes were differentially expressed between the high- and low-risk necroptosis groups, including PD-1, PD-L1, CTLA4, and LAG3, confirming the value of exploring necroptosis in new immunotherapy approaches targeting other checkpoints. The tumor cell-killing function of immune cells is known to depend on efficient antigen presentation by human leukocyte antigen (HLA) molecules. Accumulating evidence suggests that HLA serves as a useful predictor of the efficacy of immunotherapy and HLA typing before treatment is an informative step for therapy (Chowell et al., 2018). In our experiments, HLA-DRB5 displayed distinct expression patterns among different necroptosis risk groups, supporting its potential utility in predicting response to ICB and designing neoantigen-based therapeutic vaccines in the future. We conclude that high necroptosis risk_score is correlated with low MSI-H percentage, low expression of immune checkpoints, low TMB score, and effector immune cells in GC. Precision therapy relies on accurate typing and a comprehensive understanding of the underlying mechanisms. Given the finding that necroptosis is a biological process with a relatively central role in the development of gastric cancer as well as regulation of TIME, exploration of the efficacy of targeted necroptosis therapy alone or

in combination with immunotherapy should be a focus of future studies.

Wang and Liu, (2021) have identified a prognostic signature based on necroptosis hub genes in GC and uncovered a lncRNA-miRNA regulatory axis related to necroptosis, but they did not assess the TIME and immunotherapy response in GC patients. In this study, we conducted a comprehensive investigation of the immune microenvironment in GC for the first time. However, our study had several limitations that should be taken into consideration. Data was collected from public databases and required verification in clinical samples. This was a retrospective study design, which may have led to selection bias in variables and samples, and patient sample volumes were limited. Prognostic necroptosis genes combined with clinical validation in the patients of GC prospective cohort is needed to prove its efficacy. Finally, immunotherapy response was predicted according to public database website and a urothelial carcinoma cohort (Imvigor210) or melanoma cohort (GSE78220), and GC cohorts treated by immunotherapy in future research is needed.

In conclusion, we have identified an efficacious prognostic model based on necroptosis-related genes in GC and comprehensively analyzed the relationship between necroptosis and tumor immunity-associated factors. Our results provide a promising way for exploring novel and innovative targeted and immunotherapy approaches for GC patients. In the future, we aim to examine the predictive value of necroptosis risk_score in immunotherapy for gastric cancer in clinical practice and further evaluate the molecular pathways by which necroptosis influences the immune microenvironment.

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. Additional data and the code used to analyze the data can be requested by contacting the corresponding author FW, zzuwangfeng@zzu.edu.cn upon reasonable request.

Ethics statement

The studies involving human participants were reviewed and approved by the First Affiliated Hospital of Zhengzhou University Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

DX, YM, and YY wrote the manuscript and performed bioinformatics analysis. DX contributed to the manuscript discussion. DX and FW designed the study, and YM researched the literature and contributed to figures and tables. FW supervised the study and contributed to data analysis. All authors read and approved the final manuscript.

Funding

This work was supported by Henan Provincial Science and Technology Research Project (No. SBGJ202002080), Leading Talent Training in Henan Province (No. YXKC2020017).

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.2022.953997/full#supplementary-material

References

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. doi:10.1200/po.17.00073

Carneiro, B. A., and El-Deiry, W. S. (2020). Targeting apoptosis in cancer therapy. *Nat. Rev. Clin. Oncol.* 17 (7), 395–417. doi:10.1038/s41571-020-0341-y

Cassim, S., and Pouyssegur, J. (2019). Tumor microenvironment: A metabolic player that shapes the immune response. *Int. J. Mol. Sci.* 21 (1), E157. doi:10.3390/ijms21010157

Choi, Y., Shi, Y., Haymaker, C. L., Naing, A., Ciliberto, G., and Hajjar, J. (2020). T-cell agonists in cancer immunotherapy. *J. Immunother. Cancer* 8 (2), e000966. doi:10.1136/jitc-2020-000966

Chowell, D., Morris, L. G. T., Grigg, C. M., Weber, J. K., Samstein, R. M., Makarov, V., et al. (2018). Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Sci. (New York, NY)* 359 (6375), 582–587. doi:10.1126/science.aao4572

Colbert, L. E., Fisher, S. B., Hardy, C. W., Hall, W. A., Saka, B., Shelton, J. W., et al. (2013). Pronecrotic mixed lineage kinase domain-like protein expression is a prognostic biomarker in patients with early-stage resected pancreatic adenocarcinoma. *Cancer* 119 (17), 3148–3155. doi:10.1002/cncr.28144

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 (1), 100. doi:10.1186/s12943-019-1029-8

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 (7791), 549–555. doi:10.1038/s41586-019-1922-8

Johnston, F. M., and Beckman, M. (2019). Updates on management of gastric cancer. *Curr. Oncol. Rep.* 21 (8), 67. doi:10.1007/s11912-019-0820-4

Joshi, S. S., and Badgwell, B. D. (2021). Current treatment and recent progress in gastric cancer. *Ca. Cancer J. Clin.* 71 (3), 264–279. doi:10.3322/caac.21657

Junttila, A., Helminen, O., Vayrynen, J. P., Ahtiainen, M., Kenessey, I., Jalkanen, S., et al. (2020). Immunophenotype based on inflammatory cells, PD-1/PD-L1 signalling pathway and M2 macrophages predicts survival in gastric cancer. *Br. J. Cancer* 123 (11), 1625–1632. doi:10.1038/s41416-020-01053-7

Khoury, M. K., Gupta, K., Franco, S. R., and Liu, B. (2020). Necroptosis in the pathophysiology of disease. *Am. J. Pathol.* 190 (2), 272–285. doi:10.1016/j.ajpath.2019.10.012

Li, W., Zhang, X., Wu, F., Zhou, Y., Bao, Z., Li, H., et al. (2019). Gastric cancer-derived mesenchymal stromal cells trigger M2 macrophage polarization that promotes metastasis and EMT in gastric cancer. *Cell Death Dis.* 10 (12), 918. doi:10.1038/s41419-019-2131-y

Mortezaee, K., Salehi, E., Mirtavoos-Mahyari, H., Motevaseli, E., Najafi, M., Farhood, B., et al. (2019). Mechanisms of apoptosis modulation by curcumin: Implications for cancer therapy. *J. Cell. Physiol.* 234 (8), 12537–12550. doi:10.1002/jcp.28122

Pan, Y., Yu, Y., Wang, X., and Zhang, T. (2020). Tumor-associated macrophages in tumor immunity. *Front. Immunol.* 11, 583084. doi:10.3389/fimmu.2020.583084

Pasparakis, M., and Vandenabeele, P. (2015). Necroptosis and its role in inflammation. *Nature* 517 (7534), 311–320. doi:10.1038/nature14191

Pitt, J. M., MArAbelle, A., Eggermont, A., Soria, J. C., Kroemer, G., and ZitvogeL, L. (2016). Targeting the tumor microenvironment: Removing obstruction to anticancer immune responses and immunotherapy. *Ann. Oncol.* 27 (8), 1482–1492. doi:10.1093/annonc/mdw168

Sautès-Fridman, C., Petitprez, F., Calderaro, J., and Fridman, W. H. (2019). Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* 19 (6), 307–325. doi:10.1038/s41568-019-0144-6

Sawaki, A., Ohashi, Y., Omuro, Y., Satoh, T., Hamamoto, Y., Boku, N., et al. (2012). Efficacy of trastuzumab in Japanese patients with HER2-positive advanced gastric or gastroesophageal junction cancer: A subgroup analysis of the trastuzumab for gastric cancer (ToGA) study. *Gastric Cancer* 15 (3), 313–322. doi:10.1007/s10120-011-0118-1

Seeneevassen, L., Bessede, E., Megraud, F., Lehours, P., Dubus, P., and Varon, C. (2021). Gastric cancer: Advances in carcinogenesis research and new therapeutic strategies. *Int. J. Mol. Sci.* 22 (7), 3418. doi:10.3390/ijms22073418

Seifert, L., Werba, G., Tiwari, S., Giao Ly, N. N., Alothman, S., Alqunaibit, D., et al. (2016). The necrosome promotes pancreatic oncogenesis via CXCL1 and Mincle-induced immune suppression. *Nature* 532 (7598), 245–249. doi:10.1038/nature17403

Sharonov, G. V., Serebrovskaya, E. O., Yuzhakova, D. V., Britanova, O. V., and Chudakov, D. M. (2020). B cells, plasma cells and antibody repertoires in the tumour microenvironment. *Nat. Rev. Immunol.* 20 (5), 294–307. doi:10.1038/s41577-019-0257-x

Shitara, K., Ozguroglu, M., Bang, Y. J., Di Bartolomeo, M., Mandala, M., Ryu, M. H., et al. (2018). Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): A randomised, open-label, controlled, phase 3 trial. *Lancet (London, Engl.* 392 (10142), 123–133. doi:10.1016/s0140-6736(18)31257-1

Shitara, K., Van Cutsem, E., Bang, Y. J., Fuchs, C., Wyrwicz, L., Lee, K. W., et al. (2020). Efficacy and safety of pembrolizumab or pembrolizumab plus chemotherapy vs chemotherapy alone for patients with first-line, advanced gastric cancer: The KEYNOTE-062 phase 3 randomized clinical trial. *JAMA Oncol.* 6 (10), 1571–1580. doi:10.1001/jamaoncol.2020.3370

Song, W., Ren, J., Xiang, R., Kong, C., and Fu, T. (2021). Identification of pyroptosis-related subtypes, the development of a prognosis model, and characterization of tumor microenvironment infiltration in colorectal cancer. *Oncoimmunology* 10 (1), 1987636. doi:10.1080/2162402x.2021.1987636

Sprooten, J., De Wijngaert, P., Vanmeerbeerk, I., Martin, S., Vangheluwe, P., Schlenner, S., et al. (2020). Necroptosis in immuno-oncology and cancer immunotherapy. *Cells* 9 (8), E1823. doi:10.3390/cells9081823

Strilic, B., Yang, L., Albarran-Juarez, J., Wachsmuth, L., Han, K., Muller, U. C., et al. (2016). Tumour-cell-induced endothelial cell necroptosis via death receptor 6 promotes metastasis. *Nature* 536 (7615), 215–218. doi:10.1038/nature19076

Su, Z., Yang, Z., Xie, L., DeWitt, J. P., and Chen, Y. (2016). Cancer therapy in the necroptosis era. *Cell Death Differ.* 23 (5), 748–756. doi:10.1038/cdd.2016.8

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

Tang, R., Xu, J., Zhang, B., Liu, J., Liang, C., Hua, J., et al. (2020). Ferroptosis, necroptosis, and pyroptosis in anticancer immunity. *J. Hematol. Oncol.* 13 (1), 110. doi:10.1186/s13045-020-00946-7

Thommen, D. S., and Schumacher, T. N. (2018). T cell dysfunction in cancer. *Cancer Cell* 33 (4), 547–562. doi:10.1016/j.ccell.2018.03.012

Thrift, A. P., and El-Serag, H. B. (2020). Burden of gastric cancer. *Clin. Gastroenterol. Hepatol.* 18 (3), 534–542. doi:10.1016/j.cgh.2019.07.045

Viitala, M., Virtakoivu, R., Tadayon, S., Rannikko, J., Jalkanen, S., and Hollmen, M. (2019). Immunotherapeutic blockade of macrophage clever-1 reactivates the CD8(+) T-cell response against immunosuppressive tumors. *Clin. Cancer Res.* 25 (11), 3289–3303. doi:10.1158/1078-0432.Ccr-18-3016

Wang, N., and Liu, D. (2021). Identification and validation a necroptosis-related prognostic signature and associated regulatory Axis in stomach adenocarcinoma. *Onco. Targets. Ther.* 14, 5373–5383. doi:10.2147/ott.S342613

Wilke, H., Muro, K., Van Cutsem, E., Oh, S. C., Bodoky, G., Shimada, Y., et al. (2014). Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): A double-blind, randomised phase 3 trial. *Lancet. Oncol.* 15 (11), 1224–1235. doi:10.1016/s1470-2045(14)70420-6

Xie, J., Fu, L., and Jin, L. (2021). Immunotherapy of gastric cancer: Past, future perspective and challenges. *Pathol. Res. Pract.* 218, 153322. doi:10.1016/j.prp.2020.153322

Yang, W., Niu, L., Zhao, X., Duan, L., Wang, X., Li, Y., et al. (2022). Pyroptosis impacts the prognosis and treatment response in gastric cancer via immune system modulation. *Am. J. Cancer Res.* 12 (4), 1511–1534.

Zang, Y. S., Dai, C., Xu, X., Cai, X., Wang, G., Wei, J., et al. (2019). Comprehensive analysis of potential immunotherapy genomic biomarkers in 1000 Chinese patients with cancer. *Cancer Med.* 8 (10), 4699–4708. doi:10.1002/cam4.2381

Zhao, Q., Cao, L., Guan, L., Bie, L., Wang, S., Xie, B., et al. (2019). Immunotherapy for gastric cancer: Dilemmas and prospect. *Brief. Funct. Genomics* 18 (2), 107–112. doi:10.1093/bfgp/ely019

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 30 August 2022
DOI 10.3389/fgene.2022.915372

Check for updates

OPEN ACCESS

EDITED BY
Lan Zhao,School of Medicine, Stanford University, United States

REVIEWED BY
Haitao Chen,
School of Public Health, Sun Yat-sen University, China
Anjali Lathwal,
Indraprastha Institute of Information Technology Delhi, India
Meng Zhou,
Wenzhou Medical University, China

*CORRESPONDENCE
Xiao-Liang Xing,
xiaoliangxinghnm@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 07 April 2022
ACCEPTED 06 July 2022
PUBLISHED 30 August 2022

CITATION
Han Z, Wang H, Liu Y and Xing X-L (2022), Establishment of a prognostic ferroptosis- and immune-related long noncoding RNAs profile in kidney renal clear cell carcinoma.
Front. Genet. 13:915372.
doi: 10.3389/fgene.2022.915372

COPYRIGHT
© 2022 Han, Wang, Liu and Xing. 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.

Establishment of a prognostic ferroptosis- and immune-related long noncoding RNAs profile in kidney renal clear cell carcinoma

Zhijun Han[1,2†], Hao Wang[1,3†], Yafei Liu[1] and Xiao-Liang Xing[1,2*]

[1]Department of Urology, Department of Ultrasonography, Zhuzhou Hospital Affiliated to Xiangya school of Medicine, Central South University, Zhuzhou, China, [2]School of Public Health and Laboratory Medicine, Hunan University of Medicine, Huaihua, China, [3]Department of Urology, The First Affiliated Hospital, Hengyang Medical School, University of South China, Huaihua, China

Background: Ferroptosis and immunity are novel treatments that target several cancers, including kidney renal clear cell carcinoma (KIRC). Long noncoding RNAs (lncRNAs) are an important class of gene expression regulators that play fundamental roles in the regulation of ferroptosis and immunity. We aimed to identify ferroptosis- and immune-related lncRNAs as biomarkers in patients with KIRC.

Methods: Corresponding data for each patient with KIRC were obtained from The Cancer Genome Atlas (TCGA) database. Univariate and multivariate Cox regression analyses were used to identify candidate biomarkers followed by least absolute shrinkage and selection operator (LASSO) regression analyses, weighted gene coexpression network analysis (WGCANA), and gene set enrichment analysis (GSEA).

Results: Three ferroptosis- and immune-related differentially expressed lncRNAs (FI-DELs) (AC124854.1, LINC02609, and ZNF503-AS2) were markedly and independently correlated with the overall survival (OS) of patients with KIRC. The area under the curve (AUC) value of the prognostic model in the entire group using the three FI-DELs was > 0.70. The sensitivity and specificity of the diagnostic model using the three FI-DELs were 0.8586 and 0.9583, respectively.

Conclusion: The present study found that AC124854.1, LINC02609, and ZNF503-AS2 were considerably and independently correlated with the OS of patients with KIRC, suggesting that the three FI-DELs could be used as prognostic and diagnostic biomarkers for patients with KIRC.

KEYWORDS

KIRC, ferroptosis, immune, lncRNAs, prognosis, diagnosis

1 Introduction

Kidney cancer is one of the most common malignant tumors originating in the kidneys. In 2020, approximately 180,000 deaths and 430,000 new cases of kidney cancer were reported globally (Sung et al., 2021). Almost 90% of kidney cancers are renal cell carcinoma (KIRC) (Singh, 2021). According to the morphology, renal cell carcinoma can be mainly divided into three subtypes: kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma, and kidney chromophobe (Owens, 2016). The most common subtype of kidney cancer is KIRC, and it accounts for approximately 80% of all primary malignant kidney tumors (Song et al., 2019; Zhou et al., 2020). Additionally, previous studies have demonstrated that 60% of patients with KIRC die within the first 2–3 years, and 30% of patients with KIRC are diagnosed with metastases (Mendoza-Alvarez et al., 2019). Moreover, numerous studies have demonstrated that the prognosis of patients with KIRC with metastases is poor (Chaffer and Weinberg, 2011; Forsea et al., 2012; Ljungberg et al., 2015). Currently, pathological tumors, nodes, and metastasis (TNM) are usually used as a biomarker to assess the risk of patients with KIRC (Wang et al., 2019a). Additionally, numerous studies have indicated that several molecular biomarkers, such as IL13RA2, SK1, and P4HB, can guide diagnosis, prognosis, and therapy (Salama et al., 2015; Shibasaki et al., 2015; Xie et al., 2020). However, the prognosis and diagnosis of KIRC remain poor. Therefore, it is necessary to obtain reliable biomarkers for the diagnosis and prognosis of patients with KIRC.

Ferroptosis is an iron-dependent form of cell death. Numerous studies have indicated that dysregulation of ferroptosis participates in the carcinogenesis and development of several cancers (Dixon et al., 2012; Nie et al., 2018). The regulation of ferroptosis contributes to anticancer therapy in various cancers, including drug-resistant cancers, such as non-small cell lung cancer and pancreatic cancer (Yamaguchi et al., 2013; Efferth, 2017). Immunotherapy targets specific cancer antigens and alerts the immune system to eradicate them through a concerted immune response (Johdi and Sukor, 2020). Immune cells and immune factors not only participate in antitumor immunity but also in the initiation and development of antitumor immunity (Berraondo et al., 2016; Chen et al., 2017). Cancer immunotherapy has been successfully used to treat many cancers (Im and Pavletic, 2017; Nixon et al., 2018). Interestingly, evidence indicates that there is a close link between ferroptosis and immunity. Ferroptosis cells can interact with immune cells such as NK cells and CD8+ T cells, among others, to release chemotaxis to regulate anticancer immunity (Wang et al., 2018a; Wang et al., 2019b; Stockwell and Jiang, 2019). Ferroptosis and immunity can regulate each other to participate in anticancer effects (Wang et al., 2019b; Stockwell and Jiang, 2019; Shi et al., 2021). The synergistic regulation of ferroptosis and immunity not only inhibits carcinogenesis but also stimulates immune responses (Li and Rong, 2020).

The lncRNAs are a group of transcriptional RNAs over 200 nucleotides in length that cannot encode proteins and are considered one of the most sensitive and specific cancer biomarkers (Fazal and Chang, 2016). Furthermore, they are involved in carcinogenesis and the development of various cancers (Wang et al., 2018b; Kim et al., 2018; Zhang et al., 2018; Slack and Chinnaiyan, 2019). Many lncRNAs can be used as prognostic biomarkers for several cancers, such as bladder cancer and lung cancer (Sun et al., 2020a; Zhang et al., 2021; Zhou et al., 2021). In this study, we evaluated FI-DELs as potential prognostic biomarkers using differential expression, Pearson correlation, and univariate/multivariate Cox regression analyses.

TABLE 1 Characteristics of KIRC patients.

Characteristic	Variable	Total (n = 530)	%
Age (years)	≤ 65	348	65.66
	> 65	182	34.34
Gender	Female	186	35.09
	Male	344	64.91
Stage	Stage I	265	50.00
	Stage II	57	10.75
	Stage III	123	23.21
	Stage IV	82	15.47
	Unknown	3	0.57
Tumor classification	T1	271	51.13
	T2	69	13.02
	T3	179	33.77
	T4	11	2.08
Lymph nodes	N0	239	45.09
	N1	16	3.02
	Unknown	275	51.89
Distant metastasis	M0	422	79.62
	M1	78	14.72
	Unknown	30	5.66
Survival status	Alive	357	67.36
	Death	173	32.64

2 Materials and methods

2.1 Acquisition of corresponding data

The read counts data of KIRC (72 normal individuals and 530 patients with KIRC) and their corresponding clinical information (Table 1) were downloaded from an open database The Cancer Genome Atlas (TCGA) which do not need the approval of the local ethics committee. The read counts data of KIRC (45 normal individuals and 91 patients with KIRC) were downloaded from another open database International Cancer Genome Consortium (ICGC). DEseq2 in R (3.6.1) was used to screen the differentially expressed genes (DEGs) as the following criteria: baseMean ≥ 100, |log2FoldChange| ≥ 1.00, and *p.adj* <

0.05. The recognized lncRNAs, ferroptosis-related genes, and recognized immune-related genes were downloaded from the GENCODE database (https://portal.gdc.cancer.gov/), FerrDb database (http://www.zhounan.org/ferrdb), and ImmPort database (http://www.immport.org).

Estimate in R (3.6.1) was used to evaluate the stromal, immune, and ESTIMATE scores and tumor purity. The evaluated infiltrating score of immune cells and immune factors was downloaded from Tumor IMmune Estimation Resource (TIMER) (https://cistrome.shinyapps.io/timer/).

2.2 Acquisition of ferroptosis- and immune-related signatures

After differentially expressed analyses, overlapping analyses were carried out for the differentially expressed genes (DEGs) with the recognized ferroptosis- and immune-related genes to screen the ferroptosis- and immune-related DEGs (FI-DEGs) and with the recognized lncRNAs to screen the DELs. Pearson correlation analyses were performed for the FI-DEGs and DELs to obtain the FI-DELs as the following criteria $r \geq 0.5$ and $p < 0.05$.

2.3 Acquisition of prognostic biomarkers

The median value of each gene expression was used to regroup the patients with KIRC into low- and high-expression groups. Univariate Cox regression analyses followed by least absolute shrinkage and selection operator (LASSO) regression analyses were used to investigate the relationship of the FI-DELs with their overall survival (OS). Then, multivariate Cox regression analyses were performed to screen the suitable FI-DELs as biomarkers.

2.4 WGCNA and gene set enrichment analysis

The TIMER algorithm was utilized to estimate the stromal score, immune score, tumor purity, and ESTIMATE score. We used WGCNA that can convert coexpression correlation into connection weights, to identify coexpressed genes in stromal and immune cells (Langfelder and Horvath, 2008). Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to assess the biological roles of the prognostic candidates by the "clusterProfiler" R package.

2.5 Construction of prognostic and diagnostic models

According to previous reports (Liu et al., 2020), the candidate biomarkers were used to construct a risk assessment model as follows: risk value = (0.831) × LINC02609 expression value + (−0.585) × ZNF503-AS2 expression value + (−0.530) × AC124854.1 expression value. A comprehensive index of ferroptosis and immune status (CIFI) was evaluated as follows: CIFI = (risk score—Min)/Max. The Youden index was used as the optimal cut-off value to regroup the patients with KIRC into low and high CIFI groups.

A diagnostic model was constructed as following after a stepwise logistic regression analyses: logit (P) = 0.747+ (0.212) × LINC02609 expression value + (−0.145) × ZNF503-AS2 expression value + (0.200) × AC124854.1 expression value (Liu et al., 2020). The Youden index was used as the optimal cut-off value to regroup the sample into a normal group and a KIRC group.

2.6 Statistical analyses

A repeated measure ANOVA followed by an unpaired two-tailed student's t-test was used as indicated. All results were expressed as the mean ± SEM. Principal component analyses were used to reduce the dimensions and visualize the distribution of KIRC patients with different CIFI scores.

3 Results

3.1 Identification of ferroptosis- and immune-related differentially expressed lncRNAs as candidate biomarkers

Through differential expression analyses, we obtained 3,978 DEGs, including 2,573 upregulated and 1,405 downregulated DEGs (Supplementary Figure S1A). Of these, there were 531 FI-DEGs (405 upregulated FI-DEGs and 126 downregulated FI-DEGs) (Supplementary Figure S1B) and 361 DELs (278 upregulated DELs and 83 downregulated DGLs) (Supplementary Figure S1C). Pearson correlation analyses for the 531 FI-DEGs and 361 DELs found that there were 3,483 FI-DEGs–DELs pairs involving 362 FI-DEGs and 261 DELs. These 261 DELs were renamed FI-DELs.

All patients with KIRC ($n = 530$) were randomly divided into training ($n = 265$) and validation ($n = 265$) groups to verify and obtain suitable biomarkers. Univariate Cox regression analyses followed by LASSO analyses for the 261 FI-DELs in the training group were performed, and results indicated that 17 FI-DELs were correlated with the OS of patients with KIRC (Supplementary Figure S1D–F). Through ESTIMATE analyses in R software (3.6.1), we found that the stromal, immune, and ESTIMATE scores were considerably increased, while tumor purity was markedly decreased in patients with KIRC (Supplementary Figure S2A–D). In the TCGA-KIRC cohort dataset, we obtained 38 modules using WGCNA (Figures

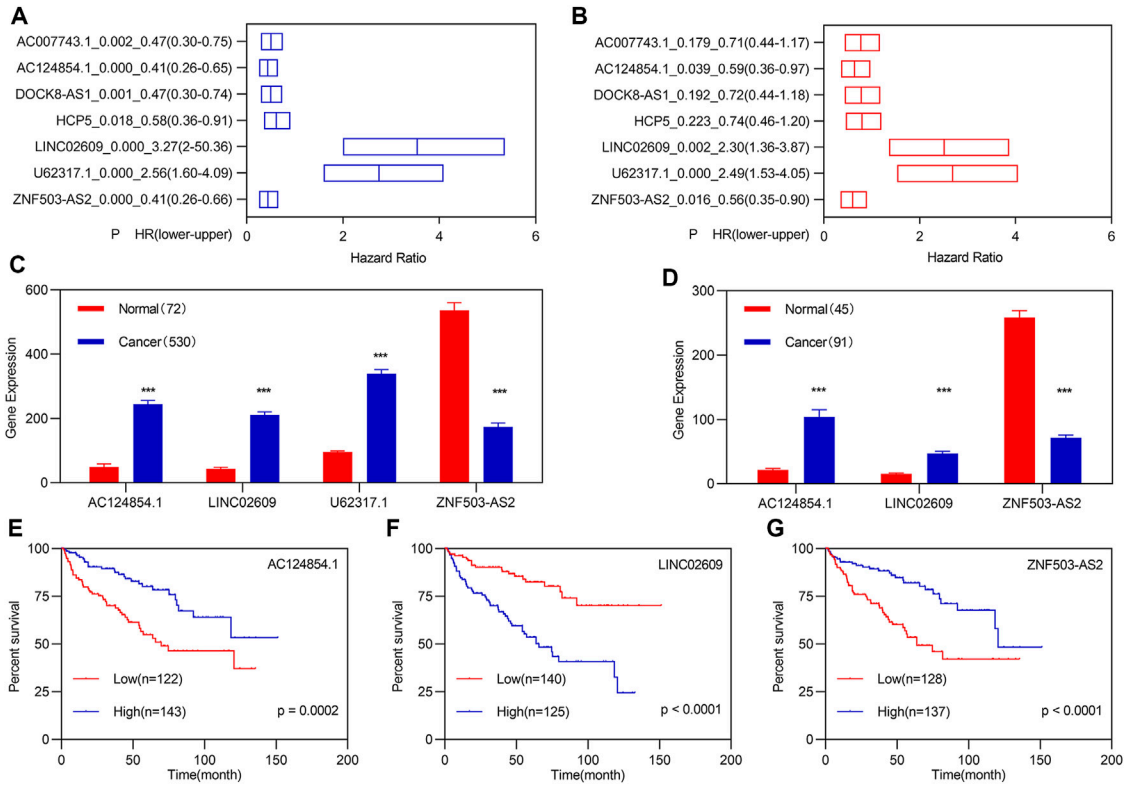


FIGURE 1
Identification of FI-DEGs as candidate biomarkers. **(A)** K-M analysis illustrated that seven FI-DELs were correlated with OS of patients with KIRC followed by LASSO analysis and WGCNA. **(B)** Multivariate Cox regression analysis illustrated that the four FI-DELs were correlated with the OS of patients with KIRC independently. **(C)** Expression of the four FI-DELs between normal and patients with KIRC in TCGA dataset. **(D)** Expression of the three FI-DELs between normal and patients with KIRC in the ICGC dataset. **(E–G)** K-M curve of the three FI-DELs in the training group. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

2E,F). Of which, six modules and seven models were highly correlated with stromal and immune scores, respectively, which enriched 2,858 lncRNAs. By overlapping analyses with the 17 FI-DELs, we obtained seven FI-DELs (Figure 1A). Next, multivariate Cox regression analyses for the seven FI-DELs were performed, and the results indicated that four FI-DELs (AC124854.1, LINC02609, U62317.1, and ZNF503-AS2) remained independently correlated with the OS of patients with KIRC (Figure 1B). The expression of AC124854.1, LINC02609, and U62317.1 was significantly increased, whereas the expression of ZNF503-AS2 was markedly decreased in patients with KIRC (Figure 1C). To clarify the expression of those four candidate prognostic biomarkers in KIRC, we selected other independent samples for verification. The expressions of AC124854.1, LINC02609, and ZNF503-AS2 in the ICGC dataset were consistent with that in the TCGA dataset (Figure 1D). There was no expression of U62317.1. Therefore, only the candidates AC124854.1, LINC02609, and ZNF503-AS2 were selected for the subsequent analyses. Patients with a high expression of LINC02609 displayed worse OS, while patients with high expression of AC124854.1 and ZNF503-AS2 displayed better OS (Figures 1E–G).

3.2 Construction and validation of CIFI as a prognostic model

Based on previous studies, we constructed a risk assessment model using the three FI-DELs. The Youden index was used as the optimal cut-off value to regroup patients with KIRC (Supplementary Figure S3). The CIFI value and survival status of each patient with KIRC are shown in Figure 2A. In patients with KIRC having high CIFI values, the expression of LINC02609 was markedly increased, the expression of AC124854.1 was considerably decreased, and there was no substantial difference for ZNF503-AS2 (Figure 2B). Patients with KIRC and high CIFI values displayed a decreased OS (Figure 2C). Receiver operating characteristic (ROC) curve analyses indicated that the AUC value of this prognostic model was 0.75 (Figure 2D). Time-dependent ROC analyses

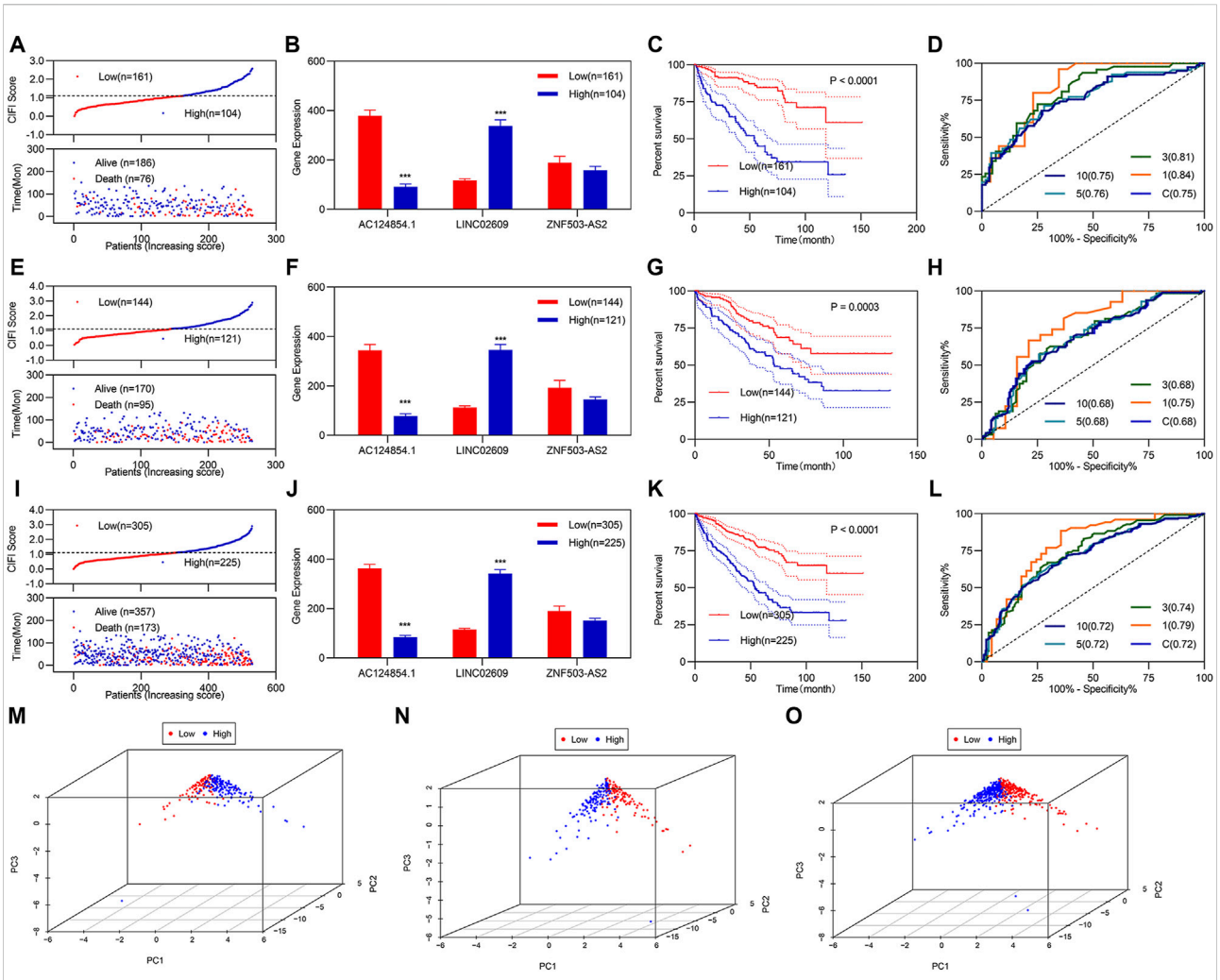


FIGURE 2
Development and validation of CIFI as a prognostic model. **(A–D)** and **(M)** Analysis of the prognostic model in the training group (n = 265). **(A)** For risk value (up) and survival status (down) for each patient. **(B)** Expression of the three candidate biomarkers. **(C)** K-M curve of KIRC patients with different CIFI values. **(D)** ROC curve of the risk assessment model. **(M)** Distribution of patients with different CIFI values. **(E–H)** and **(N)** Analysis of the prognostic model in the validation group (n = 265). **(E)** For risk value (up) and survival status (down) for each patient. **(F)** Expression of the three candidate biomarkers. **(G)** K–M curve of KIRC patients with different CIFI values. **(H)** ROC curve of the risk assessment model. **(N)** Distribution of patients with different CIFI values. i-l and o Analysis of the prognostic model in the entire group (n = 530). **(I)** For risk value (up) and survival status (down) for each patient. **(J)** Expression of the three candidate biomarkers. **(K)** **K–M** curve of KIRC patients with different CIFI values. **(L)** ROC curve of risk assessment model. **(O)** Distribution of patients with different CIFI values. *p < 0.05, **p < 0.01, ***p < 0.001.

indicated that the AUC values at 1, 3, 5, and 10 years were 0.84, 0.81, 0.76, and 0.75, respectively (Figure 2D). Principal component analyses (PCA) showed that patients with different CIFI values could be well distinguished using the three FI-DELs (Figure 2M).

Verification regarding the three FI-DELs as feasible biomarkers was performed by validation studies in the validation and entire groups; which indicated similar results for the validation and entire groups (Figures 2E–L). Particularly, the AUC values were 0.68 and 0.72 in both the validation group and the entire group respectively (Figures 2H,L). PCA showed that patients with different CIFI values could be well distinguished using the three FI-DELs in the validation and entire group (Figures 2N–O).

3.3 Independent prognostic model of the three ferroptosis- and immune-related differentially expressed lncRNAs

Previous studies have demonstrated that TNM classification and stage are usually used as predictors to assess the risk in patients with KIRC. To determine whether the CIFI score was an independent prognostic factor for patients with KIRC, we

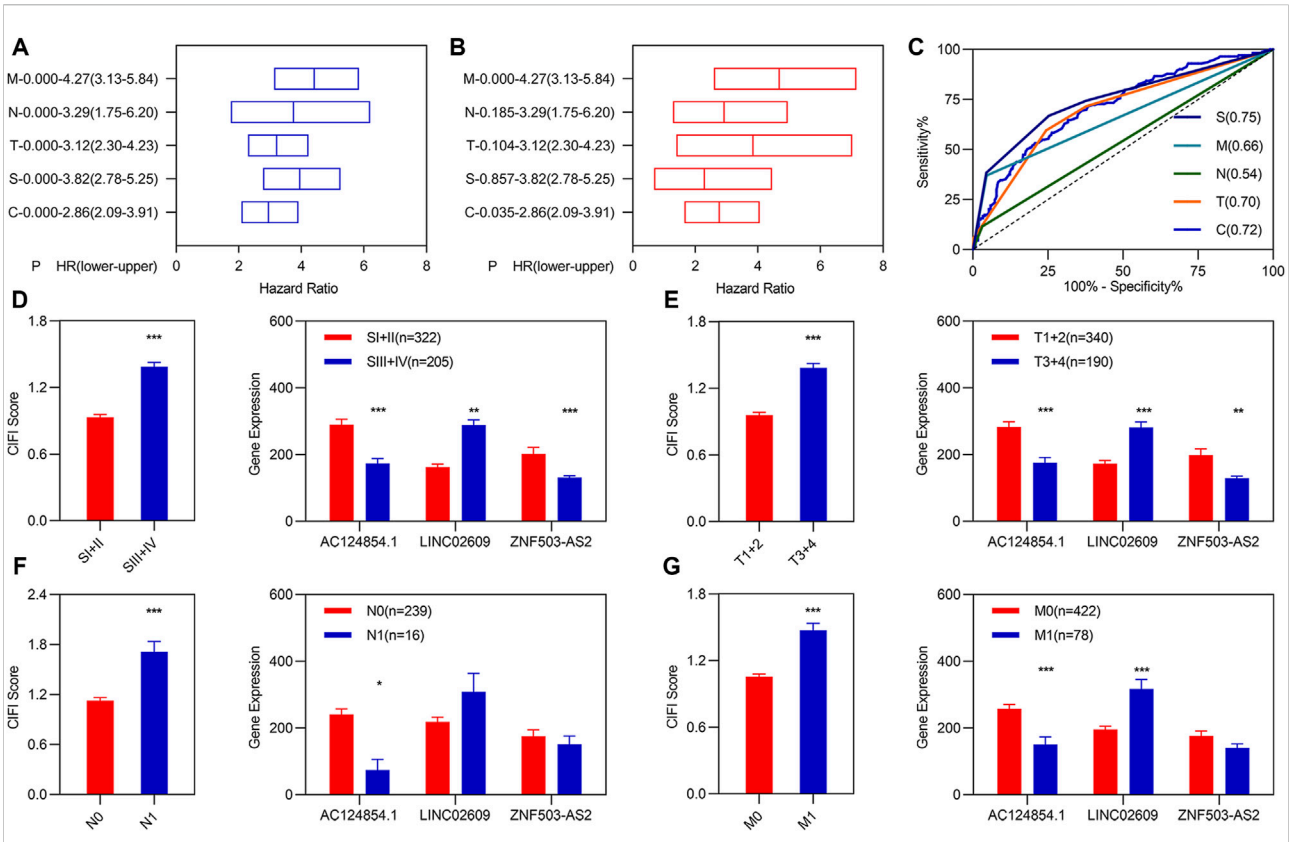


FIGURE 3
Independent prognostic values of the three FI-DELs. **(A)** K-M analysis illustrated different clinical characteristics, and CIFI values were correlated with OS of patients with KIRC. **(B)** Multivariate Cox regression analysis illustrated pathologic M, and CIFI values were correlated with the OS of patients with KIRC independently. **(C)** ROC curve of different clinical characteristics and CIFI value. **(D–G)** CIFI score and expression of the three FI-DELs within different clinical characteristics. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

performed univariate and multivariate Cox regression analyses among the clinical characteristics and CIFI scored in the entire group. Univariate Cox regression analyses showed that the CIFI score, pathologic TNM, and pathologic stage were significantly associated with the OS of KIRC (Figure 3A). Multivariate Cox regression analyses showed that the CIFI score and pathologic M were still significantly and independently associated with the OS of KIRC (Figure 3B). Subsequently, we plotted the ROC curve of the CIFI score and the different clinical characteristics and found that the AUC value of the CIFI score was higher than that of pathologic M (Figure 3C).

Correlation analyses showed that AC124854.1 and LINC02609 were significantly correlated with the CIFI values (Supplementary Figure S4). We carried out expression analyses to determine the relationship between the three FI-DEGs, CIFI scores, and clinical characteristics. The results are presented in Figures 3D–G. For example, the CIFI score was markedly associated with the pathologic TNM and stage (Figures 3D–G left). The expression of AC124854.1 was associated with the pathologic TNM, stage, and CIFI score (Figures 3D–G right).

3.4 Correlation analyses of CIFI with the immune cells and factors

After regrouping, we re-evaluated the stromal score, immune score, tumor purity, and ESTIMATE score between different clusters. There was no significant difference in stromal scores between patients with high and low CIFI values (Figure 4A). The immune and ESTIMATE scores were significantly increased, while tumor purity was significantly decreased in patients with KIRC with high CIFI values (Figures 4B–D). We performed correlation analyses for the estimated score with the three candidate biomarkers and CIFI values for patients with KIRC with high and low CIFI values (Figure 4E).

To determine which immune cells and immune factors were correlated with CIFI, differential expression analyses were first performed. There were 88 immune cells and immune factors that were markedly different between normal individuals and patients with KIRC (Supplementary Table S1). Of these, 56 immune cells and factors were considerably different between patients with KIRC with high and low CIFI values (Figures 4F–K). To further

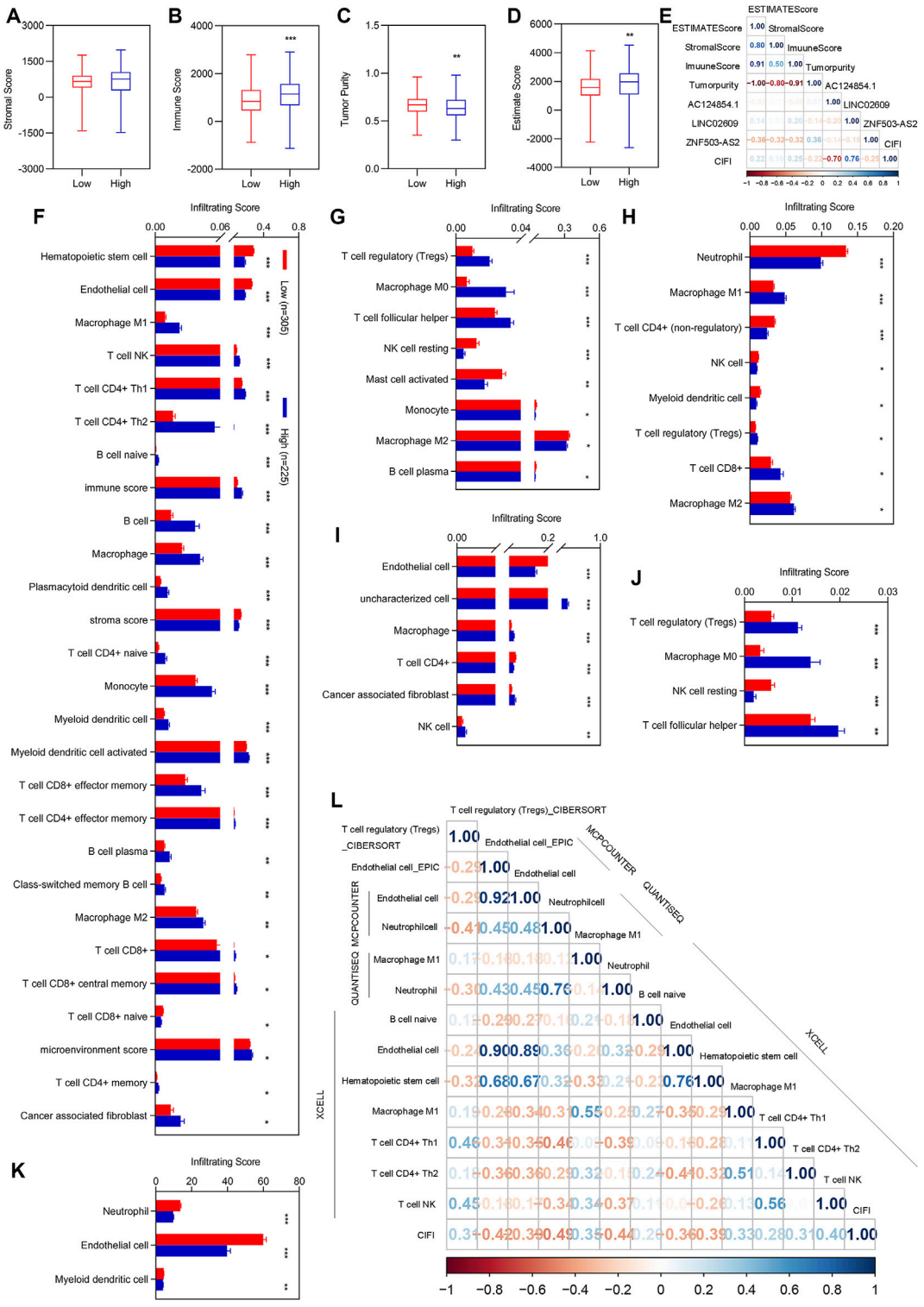


FIGURE 4
Correlation analysis of CIFI with the ESTIMATE score and immune infiltration. **(A–D)** Immune score between patients with KIRC with low CIFI value and patients with KIRC with high CIFI value. **(A)** Stromal score. **(B)** Immune score. **(C)** Tumor purity. **(D)** ESTIMATE score. **(E)** Correlation analysis of the three FI-DELs with the ESTIMATA Score. **(F–K)** Differentially expression analysis of the infiltrating score between patients with KIRC with low CIFI value and patients with KIRC with high CIFI value. **(F)** XCELL. **(G)** CIBERSORT. **(H)** quanTIseq. **(I)** EPIC. **(J)** CIBERSORT-ABS. **(K)** MCPCOUNTER. **(H)** Significantly associated immune cells and factors with CIFI. Correlation analysis of the prognostic model with the immune cells and factors. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

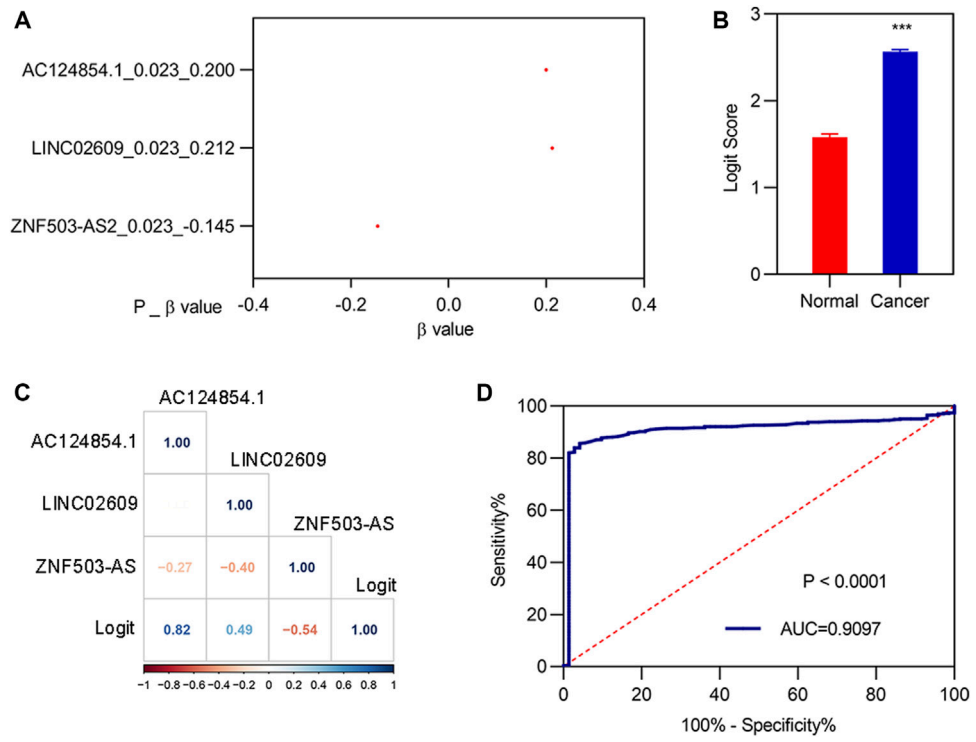


FIGURE 5
Construction of the diagnostic model. **(A)** β-value of the analysis of three FI-DELs by stepwise logistic regression. **(B)** Diagnosis values between normal and patients with KIRC. **(C)** Correlation analysis of the three FI-DELs with diagnosis values. **(D)** ROC curves of the diagnostic model. ***$p < 0.001$.

determine the relationship among these 56 immune cells and factors with CIFI values, Pearson correlation analyses were performed and 13 out of 56 immune cells and factors were correlated with the CIFI (Figure 4l).

3.5 Construction of logit (P) as a diagnostic model and enrichment analyses of Kyoto Encyclopedia of Genes and Genomes

To determine whether the three biomarkers could be used for the diagnosis of patients with KIRC, a diagnostic model was constructed and assessed. Stepwise logistic regression for the three FI-DELs (Figure 5A) was performed. The diagnosis score was markedly increased in patients with KIRC (Figure 5B) and was markedly correlated with AC124854.1, LINC02609, and ZNF503-AS2 (Figure 5C). The sensitivity and specificity of the diagnosis were 0.8566 and 0.9583, respectively (Table 2). The AUC value of this diagnostic model was 0.9097 (Figure 5D).

GSEA in R was used to compare the KEGG pathway between different clusters. We found that 60 signaling pathways were significantly enriched using the differentially expressed genes between normal and KIRC patients. The top 10 signaling pathways are shown in Figures 6A–J. We found that only five signaling pathways were significantly enriched using the differentially expressed genes between low CIFI and high CIFI patients (Figure 6K-6O). However, the signaling pathways they enriched were quite different.

TABLE 2 Sensitivity and specificity of the diagnostic model.

	Real KIRC	Real normal
Predicted KIRC	454	3
Predicted normal	76	69
Total	530	72
Correct	454	69
Sensitivity	0.8566	
Specificity		0.9583

4 Discussion

Kidney cancer is a heterogeneous disease. KIRC is one of the most common subtypes of kidney cancer and is characterized by

10.3389/fgene.2022.915372

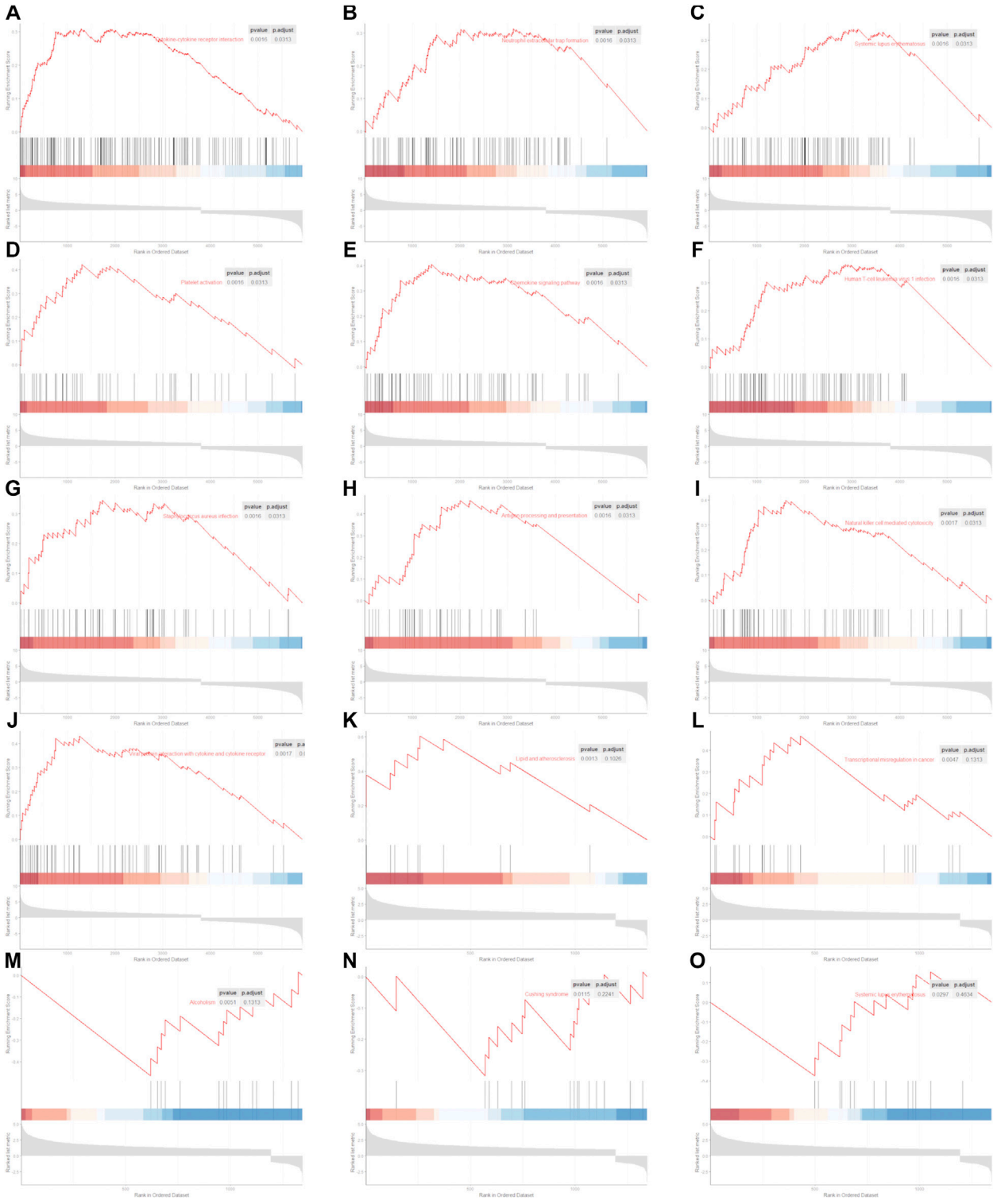


FIGURE 6
Significantly enriched KEGG pathway. **(A–J)** Top 10 significantly enriched signaling pathways between normal and KIRC. **(A)** Cytokine–cytokine receptor interaction. **(B)** Neutrophil extracellular trap formation. **(C)** Systemic lupus erythematosus. **(D)** Platelet activation. **(E)** Chemokine signaling pathway. **(F)** Human T-cell leukemia virus 1 infection. **(G)** *Staphylococcus aureus* infection. **(H)** Antigen processing and presentation. **(I)** Natural killer cell-mediated cytotoxicity. **(J)** Viral protein interaction with cytokine and cytokine receptor. **(K–O)** Five significantly enriched signaling pathways between patients with KIRC with low CIFI value and patients with KIRC with high CIFI value. **(K)** Lipid and atherosclerosis. **(L)** Transcriptional misregulation in cancer. **(M)** Alcoholism. **(N)** Cushing syndrome. **(O)** Systemic lupus erythematosus.

high morbidity, mortality, and poor prognosis worldwide (Zhao et al., 2018; Mendoza-Alvarez et al., 2019). Due to the limited biomarkers in prediction, some patients with KIRC may be diagnosed with inaccurate grades, which could influence their OS (Wang et al., 2019a). Therefore, identifying new suitable biomarkers with high sensitivity and specificity is crucial for the prognosis and diagnosis of KIRC.

In the current study, gene expression data and clinical information of KIRC were obtained from TCGA. We obtained 261 FI-DELs through differential expression and correlation analyses. For the 261 FI-DELs, we performed univariate Cox regression analyses, multivariate Cox regression analyses, and WGCNA and found that three FI-DELs could be used as biomarkers for patients with KIRC. We constructed prognostic and diagnostic models using the three FI-DELs. Univariate and multivariate Cox regression analyses indicated that the prognostic model using the three FI-DELs was an independent prognostic factor. In addition, we constructed a diagnostic model using the three FI-DELs. The AUC value of the diagnostic model was 0.9097, indicating that this diagnostic model may be feasible. Previous studies have indicated that the three biomarkers identified in the present study were related to OS in other cancers. For example, Wu et al. (2020) found that ZNF503-AS2 could be used as an independent prognostic biomarker for rhabdoid tumors of the kidney by univariate and multivariate Cox analyses. Xing et al. (2021) found that LINC02609 was associated with the OS of patients with KIRC. Cao et al. (2021) found that the glycolysis-related lncRNA AC124854.1 was markedly correlated with the OS of renal cancer by univariate and multivariate Cox regression analyses. Cheng et al. (2019) also found that AC124854.1 could be used as a prognostic and diagnostic biomarker for KIRC. Our results are consistent with those of previous studies that reinforced the feasibility of our results.

Previous studies have found that surgery is the primary treatment for patients with KIRC because most patients are resistant to radiation and chemotherapy (Zhao et al., 2018; Evelonn et al., 2019; Yin et al., 2019). Discovering recently, ferroptosis and immunity are new therapeutic targets for cancer. Additionally, some prognostic markers based on ferroptosis and immunity also exist. For example, the AUC value of the risk assessment model constructed by Sun et al. using immune-related lncRNA signatures was 0.71 [45], and Xing et al. found that the AUC value of the risk assessment model constructed using the ferroptosis-related lncRNA signatures was 0.72 [40]. We compared the present prognostic lncRNA features with published predictive models in patients with KIRC. Yin et al. (2018) found that nine lncRNAs (SLC25A5-AS1, COL18A1-AS1, WT1-AS, AC016773.1, LINC00460, LINC00313, HOTTIP, FGF14-AS1, and AC10502.1) could serve as independent biomarkers to predict survival in patients with KIRC. Sun et al. (2020b) found five immune-related lncRNA signatures (AC008105.3, LINC02084, AC243960.1, AC093278.2, and AC108449.2) with the ability to predict the prognosis of patients with KIRC. Liu et al. (2018) found that 19 lncRNAs (LOC606724, SCART1, SNORA8, LOC728024, HAVCR1P1, FCGR1CP, LINC00240, LINC00894, GK3P, SNHG3, KIAA0125, URB1-AS1, ZNF542P, TINCR, LINC00926, PDXDC2P, COL18A1-AS1, LINC00202-1, and LINC00937) that are potential biomarkers for the prognosis of KIRC. The AUC values of the three models were 0.684, 0.709, and 0.723, respectively, which were slightly lower than those of the present study (Liu et al., 2018; Yin et al., 2018; Sun et al., 2020b). Additionally, Zhang et al. (2019) found that 11 lncRNA signatures (AC245100.1, AP002761.1, LINC00488, AC017033.1, LINC-PINT, COL5A1-AS1, AC026471.4, AL009181.1, AL078590.3, LINC00524, and HOTTIP) could be potential biomarkers for KIRC. Yu et al. (2021) found that five prognostic-associated m6A-related lncRNAs (AC012170.2, AL157394.1, AP006621.2, AC025580.3, and AC124312.5) can be used as prognostic biomarkers for KIRC. The AUC values of these models were 0.781, and 0.809 respectively, which were slightly higher than those in the present study (Zhang et al., 2019; Yu et al., 2021). Comparatively, these two models used more prognostic biomarkers than those used in this study.

Although previous studies and our results indicate that the three FI-DELs may be used as prognostic and diagnostic biomarkers for patients with KIRC, the lack of cross-validation of other independent data and prospective clinical validation is an important shortcoming of our study. In addition, *in vitro* cell experiments and *in vivo* animal experiments are necessary to understand the functions of the three FI-DELs as biomarkers.

5 Conclusion

Through a series of bioinformatics analyses, we found that three FI-DELs (AC124854.1, LINC02609, and ZNF503-AS2) were independently significantly correlated with the OS of patients with KIRC. The prognostic and diagnostic models suggested that three FI-DELs could be used as prognostic and diagnostic biomarkers in patients with KIRC.

Data availability statement

Publicly available datasets were analyzed in this study. These data can be found at: https://portal.gdc.cancer.gov/.

Author contributions

ZH and HW conceived and designed the experiments; XX performed the analyses; YL helped in analyzing the data; and XX wrote 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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.915372/full#supplementary-material

References

Berraondo, P., Minute, L., Ajona, D., Corrales, L., Melero, I., and Pio, R. (2016). Innate immune mediators in cancer: Between defense and resistance. *Immunol. Rev.* 274, 290–306. doi:10.1111/imr.12464

Cao, H., 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

Chaffer, C. L., and Weinberg, R. A. (2011). A perspective on cancer cell metastasis. *Science* 331, 1559–1564. doi:10.1126/science.1203543

Chen, Y. P., Zhang, Y., Lv, J. W., Li, Y. Q., Wang, Y. Q., He, Q. M., et al. (2017). Genomic analysis of tumor microenvironment immune types across 14 solid cancer types: Immunotherapeutic implications. *Theranostics* 7, 3585–3594. doi:10.7150/thno.21471

Cheng, G., Liu, D., Liang, H., Yang, H., Chen, K., and Zhang, X. (2019). A cluster of long non-coding RNAs exhibit diagnostic and prognostic values in renal cell carcinoma. *Aging (Albany NY)* 11, 9597–9615. doi:10.18632/aging.102407

Dixon, S. J., Lemberg, K. M., Lamprecht, M. R., Skouta, R., Zaitsev, E. M., Gleason, C. E., et al. (2012). Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* 149, 1060–1072. doi:10.1016/j.cell.2012.03.042

Efferth, T. (2017). From ancient herb to modern drug: Artemisia annua and artemisinin for cancer therapy. *Semin. Cancer Biol.* 46, 65–83. doi:10.1016/j.semcancer.2017.02.009

Evelonn, E. A., Landfors, M., Haider, Z., Kohn, L., Ljungberg, B., Roos, G., et al. (2019). DNA methylation associates with survival in non-metastatic clear cell renal cell carcinoma. *BMC Cancer* 19, 65. doi:10.1186/s12885-019-5291-3

Fazal, F. M., and Chang, H. Y. (2016). lncRNA structure: Message to the heart. *Mol. Cell* 64, 1–2. doi:10.1016/j.molcel.2016.09.030

Forsea, A. M., Del Marmol, V., de Vries, E., Bailey, E. E., and Geller, A. C. (2012). Melanoma incidence and mortality in europe: New estimates, persistent disparities. *Br. J. Dermatol.* 167, 1124–1130. doi:10.1111/j.1365-2133.2012.11125.x

Im, A., and Pavletic, S. Z. (2017). Immunotherapy in hematologic malignancies: Past, present, and future. *J. Hematol. Oncol.* 10, 94. doi:10.1186/s13045-017-0453-8

Johdi, N. A., and Sukor, N. F. (2020). Colorectal cancer immunotherapy: Options and strategies. *Front. Immunol.* 11, 1624. doi:10.3389/fimmu.2020.01624

Kim, J., Piao, H. L., Kim, B. J., Yao, F., Han, Z., Wang, Y., et al. (2018). Long noncoding RNA MALAT1 suppresses breast cancer metastasis. *Nat. Genet.* 50, 1705–1715. doi:10.1038/s41588-018-0252-3

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

Li, Z., and Rong, L. (2020). Cascade reaction-mediated efficient ferroptosis synergizes with immunomodulation for high-performance cancer therapy. *Biomater. Sci.* 8, 6272–6285. doi:10.1039/d0bm01168a

Liu, T., Sui, J., Zhang, Y., Zhang, X. M., Wu, W. J., Yang, S., et al. (2018). Comprehensive analysis of a novel lncRNA profile reveals potential prognostic biomarkers in clear cell renal cell carcinoma. *Oncol. Rep.* 40, 1503–1514. doi:10.3892/or.2018.6540

Liu, Y., Zhang, X., Zhang, J., Tan, J., Li, J., and Song, Z. (2020). Development and validation of a combined ferroptosis and immune prognostic classifier for hepatocellular carcinoma. *Front. Cell Dev. Biol.* 8, 596679. doi:10.3389/fcell.2020.596679

Ljungberg, B., Bensalah, K., Canfield, S., Dabestani, S., Hofmann, F., Hora, M., et al. (2015). EAU guidelines on renal cell carcinoma: 2014 update. *Eur. Urol.* 67, 913–924. doi:10.1016/j.eururo.2015.01.005

Mendoza-Alvarez, A., Guillen-Guio, B., Baez-Ortega, A., Hernandez-Perez, C., Lakhwani-Lakhwani, S., Maeso, M. D. C., et al. (2019). Whole-exome sequencing identifies somatic mutations associated with mortality in metastatic clear cell kidney carcinoma. *Front. Genet.* 10, 439. doi:10.3389/fgene.2019.00439

Nie, J., Lin, B., Zhou, M., Wu, L., and Zheng, T. (2018). Role of ferroptosis in hepatocellular carcinoma. *J. Cancer Res. Clin. Oncol.* 144, 2329–2337. doi:10.1007/s00432-018-2740-3

Nixon, N. A., BlaisN.ErnSt, S., Kollmannsberger, C., Bebb, G., ButlerM., et al. (2018). Current landscape of immunotherapy in the treatment of solid tumours, with future opportunities and challenges. *Curr. Oncol.* 25, e373–e384. doi:10.3747/co.25.3840

Owens, B. (2016). Kidney cancer. *Nature* 537, S97. doi:10.1038/537S97a

Salama, M. F., Carroll, B., Adada, M., Pulkoski-Gross, M., Hannun, Y. A., and Obeid, L. M. (2015). A novel role of sphingosine kinase-1 in the invasion and angiogenesis of VHL mutant clear cell renal cell carcinoma. *FASEB J.* 29, 2803–2813. doi:10.1096/fj.15-270413

Shi, L., Liu, Y., Li, M., and Luo, Z. (2021). Emerging roles of ferroptosis in the tumor immune landscape: From danger signals to anti-tumor immunity. *FEBS J.* 289, 3655–3665. doi:10.1111/febs.16034

Shibasaki, N., Yamasaki, T., Kanno, T., Arakaki, R., Sakamoto, H., Utsunomiya, N., et al. (2015). Role of IL13RA2 in sunitinib resistance in clear cell renal cell carcinoma. *PLoS One* 10, e0130980. doi:10.1371/journal.pone.0130980

Singh, D. (2021). Current updates and future perspectives on the management of renal cell carcinoma. *Life Sci.* 264, 118632. doi:10.1016/j.lfs.2020.118632

Slack, F. J., and Chinnaiyan, A. M. (2019). The role of non-coding RNAs in oncology. *Cell* 179, 1033–1055. doi:10.1016/j.cell.2019.10.017

Song, J., Liu, Y. D., Su, J., Yuan, D., Sun, F., and Zhu, J. (2019). Systematic analysis of alternative splicing signature unveils prognostic predictor for kidney renal clear cell carcinoma. *J. Cell. Physiol.* 234, 22753–22764. doi:10.1002/jcp.28840

Stockwell, B. R., and Jiang, X. (2019). A physiological function for ferroptosis in tumor suppression by the immune system. *Cell Metab.* 30, 14–15. doi:10.1016/j.cmet.2019.06.012

Sun, J., Zhang, Z., Bao, S., Yan, C., Hou, P., Wu, N., et al. (2020). Identification of tumor immune infiltration-associated lncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. *J. Immunother. Cancer* 8, e000110. doi:10.1136/jitc-2019-000110

Sun, Z., 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

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

Wang, D., Xie, N., Gao, W., Kang, R., and Tang, D. (2018). The ferroptosis inducer erastin promotes proliferation and differentiation in human peripheral blood mononuclear cells. *Biochem. Biophys. Res. Commun.* 503, 1689–1695. doi:10.1016/j.bbrc.2018.07.100

Wang, Q., Zhang, H., Chen, Q., Wan, Z., Gao, X., and Qian, W. (2019). Identification of METTL14 in kidney renal clear cell carcinoma using bioinformatics analysis. *Dis. Markers* 2019, 5648783. doi:10.1155/2019/5648783

Wang, W., Green, M., Choi, J. E., Gijon, M., Kennedy, P. D., Johnson, J. K., et al. (2019). CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature* 569, 270–274. doi:10.1038/s41586-019-1170-y

Wang, Z., Yang, B., Zhang, M., Guo, W., Wu, Z., Wang, Y., et al. (2018). lncRNA epigenetic landscape analysis identifies EPIC1 as an oncogenic lncRNA that interacts with MYC and promotes cell-cycle progression in cancer. *Cancer Cell* 33, 706–720. doi:10.1016/j.ccell.2018.03.006

Wu, X., Liu, J., Zhu, C., Ma, M., Chen, X., Liu, Y., et al. (2020). Identification of potential biomarkers of prognosis-related long non-coding RNA (lncRNA) in pediatric rhabdoid tumor of the kidney based on ceRNA networks. *Med. Sci. Monit.* 26, e927725. doi:10.12659/MSM.927725

Xie, L., Li, H., Zhang, L., Ma, X., Dang, Y., Guo, J., et al. (2020). Autophagy-related gene P4HB: A novel diagnosis and prognosis marker for kidney renal clear cell carcinoma. *Aging (Albany NY)* 12, 1828–1842. doi:10.18632/aging.102715

Xing, X. L., 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, 591. doi:10.1186/s12935-021-02284-1

Yamaguchi, H., Hsu, J. L., Chen, C. T., Wang, Y. N., Hsu, M. C., Chang, S. S., et al. (2013). Caspase-independent cell death is involved in the negative effect of EGF receptor inhibitors on cisplatin in non-small cell lung cancer cells. *Clin. Cancer Res.* 19, 845–854. doi:10.1158/1078-0432.CCR-12-2621

Yin, H., Wang, X., Zhang, X., Wang, Y., Zeng, Y., Xiong, Y., et al. (2018). Integrated analysis of long noncoding RNA associated-competing endogenous RNA as prognostic biomarkers in clear cell renal carcinoma. *Cancer Sci.* 109, 3336–3349. doi:10.1111/cas.13778

Yin, L., Li, W., Wang, G., Shi, H., Wang, K., Yang, H., et al. (2019). NR1B2 suppress kidney renal clear cell carcinoma (KIRC) progression by regulation of LATS 1/2-YAP signaling. *J. Exp. Clin. Cancer Res.* 38, 343. doi:10.1186/s13046-019-1344-3

Yu, J., 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

Zhang, E., He, X., Zhang, C., Su, J., Lu, X., Si, X., et al. (2018). A novel long noncoding RNA HOXC-AS3 mediates tumorigenesis of gastric cancer by binding to YBX1. *Genome Biol.* 19, 154. doi:10.1186/s13059-018-1523-0

Zhang, J., Zhang, X., Piao, C., Bi, J., Zhang, Z., Li, Z., et al. (2019). A long non-coding RNA signature to improve prognostic prediction in clear cell renal cell carcinoma. *Biomed. Pharmacother.* 118, 109079. doi:10.1016/j.biopha.2019.109079

Zhang, Z., Yan, C., Li, K., Bao, S., Li, L., Chen, L., et al. (2021). Pan-cancer characterization of lncRNA modifiers of immune microenvironment reveals clinically distinct de novo tumor subtypes. *NPJ Genom. Med.* 6, 52. doi:10.1038/s41525-021-00215-7

Zhao, H., Cao, Y., Wang, Y., Zhang, L., Chen, C., Wang, Y., et al. (2018). Dynamic prognostic model for kidney renal clear cell carcinoma (KIRC) patients by combining clinical and genetic information. *Sci. Rep.* 8, 17613. doi:10.1038/s41598-018-35981-5

Zhou, J., Zhu, Y., Liu, Y., Niu, P., Chen, H., Deng, J., et al. (2020). High PRAS40 mRNA expression and its role in prognosis of clear cell renal cell carcinoma. *Transl. Androl. Urol.* 9, 1650–1660. doi:10.21037/tau-20-741

Zhou, M., Zhang, Z., Bao, S., Hou, P., Yan, C., Su, J., et al. (2021). Computational recognition of lncRNA signature of tumor-infiltrating B lymphocytes with potential implications in prognosis and immunotherapy of bladder cancer. *Brief. Bioinform.* 22, bbaa047. doi:10.1093/bib/bbaa047

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fgene.2022.933862

Check for updates

OPEN ACCESS

EDITED BY
Lan Zhao,
Stanford University, United States

REVIEWED BY
Yang Zhang,
Beijing Cancer Hospital, Peking University, China
Ming Sun,
The University of Chicago, United States

*CORRESPONDENCE
Zishao Zhong,
zhongzishao@qq.com
Suiping Huang,
gzdoctorhsp@126.com

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

RECEIVED 01 May 2022
ACCEPTED 12 August 2022
PUBLISHED 02 September 2022

CITATION
Yu W, He G, Zhang W, Ye Z, Zhong Z and Huang S (2022), INHBB is a novel prognostic biomarker and correlated with immune infiltrates in gastric cancer. *Front. Genet.* 13:933862. doi: 10.3389/fgene.2022.933862

COPYRIGHT
© 2022 Yu, He, Zhang, Ye, Zhong and Huang. 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.

INHBB is a novel prognostic biomarker and correlated with immune infiltrates in gastric cancer

Weifeng Yu[1,2], Guihua He[2], Wang Zhang[2], Zhenhao Ye[2], Zishao Zhong[2]* and Suiping Huang[1,2,3]*

[1]Gastroenterology Department, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China, [2]Gastroenterology Department, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou, China, [3]State Key Laboratory of Dampness Syndrome of Chinese Medicine, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China

Inhibin subunit beta B (INHBB) is a potential prognostic biomarker for a variety of cancers. However, its role in gastric cancer (GC) remains elusive. The differential expression data of *INHBB* in tumor and normal tissues were extracted from several databases and genetic alterations of *INHBB* were assessed by cBioPortal. Kaplan-Meier analysis was used to evaluate the survival rate of patients with GC with *INHBB* and association with clinical features in GC. Cox regression analysis was used to explore the prognostic value of clinical indicators and *INHBB* in GC, and a nomogram prognostic model was established. In addition, the predictive validity of the nomogram model was assessed by time-depended receiver operating characteristic (ROC) and calibration curves. Functional enrichment analyses were conducted to functionally annotate *INHBB*. Notably, we found that the quantitative assessment of immune cell subpopulation infiltration correlated with *INHBB* expression. *INHBB* expression is upregulated in GC and is correlated with several clinical features including prognostic indicators and a histological type. Genetic alterations were observed in *INHBB*, its DNA methylation level was negatively correlated with *INHBB* expression. High *INHBB* expression is associated with a poor prognosis and is an independent risk factor for prognosis in GC, along with age and residual tumor. The nomogram model showed a good prediction ability and was validated by time-depended ROC and calibration curves. Functional enrichment analysis indicated that *INHBB*-associated genes were enriched in tumor microenvironment Gene Ontology (GO) terms and were correlated with tumor-associated pathways. INHBB has a regulatory function in immune cell infiltration, especially macrophage infiltration in GC. Specifically, patients with GC with high *INHBB* expression and high macrophage infiltration have a worse prognosis. *INHBB* expression was negatively correlated with the expression of chemokines/chemokine receptors and plays a regulatory role in immunoinhibitor/immunostimulator-involved pathways. INHBB is a potential prognostic biomarker for GC and may drive the abnormal activity of critical cancer-associated pathways, potentially contributing to immune cell infiltration to promote GC development.

KEYWORDS

gastric cancer, prognostic marker, INHBB, immune cell infiltration, the cancer genome atlas, genotype-tissue expression

Introduction

Gastric cancer (GC) is a malignant tumor originating from the epithelium of the gastric mucosa. It ranks fifth in incidence and third in mortality among all tumors, with nearly 800,000 patients dying of GC worldwide in 2018 (Bray et al., 2018). The primary treatment in GC involves surgical resection of target lesions, followed by adjuvant chemotherapy (Sohn et al., 2017). Recently, advances in treatment such as inhibition of immune checkpoint or cancer stemness have changed the prognosis of patients with early GC. However, the overall survival rate is still unsatisfactory (Smyth et al., 2020), this may be attributed to occult development and non-specific clinical symptoms, and there is a lack of reliable specific markers. Traditional prognostic markers lack sensitivity because of which the patients with GC lose the early-stage treatment opportunities to treat malignant invasions, which may later develop into metastasis leading to a poor prognosis (Dicken et al., 2005). Therefore, it is pivotal to find reliable prognosis indicators to promote GC patient survival.

Activins, members of the transforming growth factor TGF-β superfamily of proteins, are synthesized as homo- or hetero-dimers of two highly related disulfide-linked inhibin beta subunits, inhibin subunit beta A (encoded by INHBA) and inhibin subunit beta B (encoded by INHBB). Thus, INHBA and INHBB can form three molecular species activin A, activin B, and activin AB (Chang et al., 2002; Rodgarkia-Dara et al., 2006). Activins are widely expressed and are functionally diverse; for example, they are involved in inflammation, cellular proliferation, and embryogenesis (Ethier and Findlay, 2001; Werner and Alzheimer, 2006; Xia and Schneyer, 2009). While the beta subunits have a similar sequence identity, the potency of each subunit in regulating biological processes shows a considerable variation (Muttukrishna et al., 1994). In the past, most research on INHBB focused on the reproductive system (Tong et al., 2003; Yao et al., 2006). Recently, INHBB was regarded as a valuable biomarker in various cancer types. Gutierrez et al. demonstrated that INHBB is regulated by methylation and closely associated with metastasis in colorectal cancer (Gutierrez et al., 2021). Xu et al. (2020) indicated that INHBB expression is upregulated in rectal cancer tissues and portends a poor prognosis. Moreover, Kita, Akihiro et al. suggested that high INHBB expression promotes cell migration and proliferation in oral squamous cell carcinoma and is associated closely with the tumor microenvironment (Kita et al., 2017). All these data suggested that INHBB may be a novel oncogene and associated with tumor progression, but the role of INHBB in GC remains unknown.

In this pilot study, we aimed to investigate the expression of INHBB in The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases and explore the prognostic value of INHBB in GC. Further, we verified the results using a GC clinical tissue microarray. The association between INHBB and immune infiltrates was also explored through bioinformatic analyses. In addition, correlation analyses between INHBB expression and immunomodulators as well as chemokines were performed to identify the potential immunoregulation of INHBB in GC.

Materials and methods

Data collection and analysis

INHBB expression data of the Genotype-Tissue Expression (GTEx) and TCGA were obtained from the UCSC Xena database (https://xenabrowser.net/datapages/) and have been uniformly processed using the Toil process (Vivian et al., 2017). Meanwhile, the GC data in TCGA and the corresponding healthy stomach tissue data in GTEx were extracted for further analyses. A scatter plot was used to show the difference in expression of INHBB in tumor and healthy tissues. The GSE26899 and GSE29272 datasets contained 108 and 268 tissue expression profiling samples, respectively, with a total of 146 normal samples and 230 gastric cancer samples. It was obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) to validate the expression of INHBB in GC.

Genetic alterations of inhibin subunit beta B

The cBioPortal for Cancer Genomics (http://cbioportal.org) (Gao et al., 2013) is a comprehensive genomics database which is widely used in the cancer field. Using cBioPortal, we visualized the genetic variation of INHBB in GC. Illumina HumanMethylation 450 BeadChip is a new platform with 450,000 DNA methylation detection sites; methylation profiles were downloaded from Illumina human methylation 450 in TCGA and used for correlation analysis with INHBB expression data.

Survival analysis

R package survival was used to analyze the overall survival of patients with high or low INHBB expression in GC using pan-

cancer level data from TCGA which was visualized using the survminer package (v.0.4.9). Kaplan-Meier plotter (www.kmplot.com) (Gyorffy, 2021) was used to assess the prognostic value of INHBB in GC data from the GEO database (GSE15459, GSE29272, GSE62254). We also explored the role of macrophage infiltration levels in the prognosis of patients with GC in the pan-cancer module.

Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed by R package org.Hs.eg.db (Bioconductor 3.2) and clusterProfiler (v.3.6.3) (Yu et al., 2012), and the result was visualized using ggplot2 (v.3.3.3). Gene set enrichment analysis (GSEA) was analyzed based on Hallmark gene set terms, the false discovery rate (FDR) q-value, normalized enrichment score (NES), and nominal *p*-value suggested the importance of the correlation between gene sets and pathways. Gene sets with p value <0.05 and FDR <0.25 were considered as significantly enriched.

Immune cell infiltration analysis

Gene set variation analysis (GSVA, v.1.34.0) was used to demonstrate the correlation of infiltration of 24 types of immune cells with INHBB expression in GC. Tumor Immune Estimation Resource (TIMER, https://cistrome.shinyapps.io/timer/), includes the data of 10,897 tumors from 32 cancer types and is a comprehensive analytical web tool to explore the molecular interactions of tumor immune infiltration and gene expression data (Li et al., 2017). Gene Expression Profiling Interactive Analysis (GEPIA, http://gepia.cancer-pku.cn/index.html) is an online platform application for gene expression analysis based on the data from the TCGA and the GTEx databases (Tang et al., 2017). We thus explored the association between INHBB and tumor-infiltrating immune cells. Meanwhile, we visualized six types of immune cell infiltration via TIMER and uncovered the prognostic impact of immune infiltrates in GC. Additionally, we visualized six types of immune cells involved in infiltration *via* TIMER and uncovered the prognostic impact of immune infiltrates in GC. We further explored the correlation between INHBB expression and immune cell markers using TIMER and GEPIA.

Immunomodulator analysis

TISIDB (http://cis.hku.hk/TISIDB/index.php) is a publicly open website for assessing tumor and immune system interaction (Ru et al., 2019). In the immunomodulator and chemokine modules, we explored the association between INHBB expression and immunoinhibitors/immunostimulators as well as chemokines/chemokine receptors at the GC and pan-cancer levels.

Clinical materials

GC tissue microarray samples from 97 patients with GC and paracarcinoma tissues from 83 patients were obtained from Shanghai Outdo Biotech (HStmA180Su08). According to the seventh edition of the UICC/AJCC TNM staging system, the patients were divided into different tumor clinical stages. All these patients were diagnosed with primary GC and they underwent surgery from December 2013 to September 2015, with a 5-year follow-up.

Immunohistochemistry

Microarray analysis of GC tissues and matched paracarcinoma tissues was performed. For Immunohistochemistry (IHC), the microarray samples were stained with an INHBB antibody (Invitrogen, PA5-119792). The steps are briefly described as follows. After paraffin-embedded sections of GC tissue specimens were dewaxed and hydrated with xylene as well as ethanol at different gradient concentrations, microwave antigen repair was performed, followed by incubation with hydrogen peroxide solution at room temperature for 10 min to block endogenous peroxidase activity. The primary antibody was added and incubated overnight at 4°C, washed three times with phosphate-buffered saline, and incubated with a secondary antibody at room temperature for 30 min. The sections were stained with 3,3′-diaminobenzidine (DAB) followed by hematoxylin re-staining, and finally, the slices were sealed with neutral gum.

The staining intensity and positive staining rate of cytoplasmic staining of INHBB were calculated independently for cancer and paracarcinoma tissues. The immunostaining index was based on the proportion of positively stained tumor cells and staining intensity. The proportion of positive staining rate was set from 0 to 100%, and the staining intensity score was divided into four categories: 0 (no immunostaining), 1 (weak), 2 (moderate), and 3 (strong). The staining intensity score multiplied by the staining positive score was finally calculated as the immunostaining index; tumors with indexes less than or equal to 140% were considered as immunostaining-low expression and those more than 140% were scored as immunostaining-high expression.

Statistical analysis

All data analyses were performed using R (v.3.6.3). A median threshold was employed to distinguish between high and low

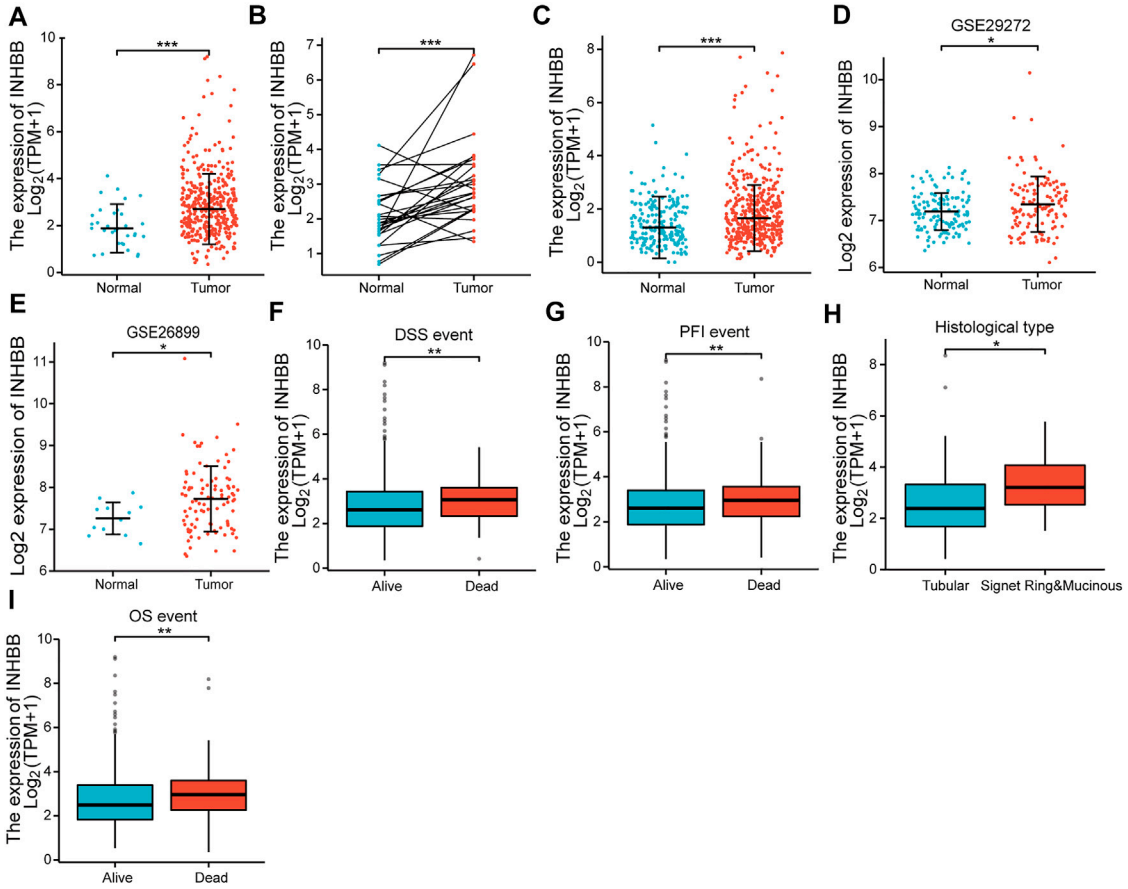


FIGURE 1
The expression of INHBB in **(A)** GC tissues and normal tissues data from TCGA, **(B)** GC tissues and paired paracarcinoma tissues data from TCGA, **(C)** GC tissues and normal tissues data from TCGA and GTEx, **(D,E)** GC tissues and normal tissues data from GSE29272 and GSE26899, and the association between INHBB expression and **(F)** DSS event, **(G)** PFI event, **(H)** histological type, **(I)** OS event.

expression of INHBB. The association between INHBB expression and clinical pathologic variables was analyzed using Wilcoxon rank-sum, Chi-square, and Fisher's exact tests. The Kaplan–Meier curve was used to estimate survival rates between the high and low INHBB-expressing groups. Univariate and multivariate Cox regression analyses were used to identify the importance of prognostic factors. *p*-values less than 0.05 were considered statistically significant.

Results

Inhibin subunit beta B is highly expressed in gastric cancer

First, we evaluated the mRNA expression levels of INHBB in patients with GC. The results suggested that INHBB expression in GC tissues was significantly higher than that in normal tissues ($p < 0.001$) (Figure 1A). The results were verified in GC and paired adjacent normal tissues ($p < 0.001$) (Figure 1B). The same results were obtained during the combined analysis of normal tissues in TCGA and GTEx databases (Figure 1C). In addition, we downland two GEO datasets (GSE26899, and GSE29272) to validate the transcription expression level of INHBB in cancer tissues and adjacent tissues. The results indicated that the expression of INHBB in GC lesions was significantly higher than that in adjacent noncancerous tissues from GEE26899 and GSE29272 datasets ($p < 0.05$) (Figures 1D, E).

Next, we analyzed the association between clinical characteristics and INHBB expression in patients with GC in the TCGA database. The clinical categories of patients with GC are summarized in Table 1, for example, it is clearly shown that the frequencies of T4 stage GC patients are 13.1% and 14.2% in low and high expression groups, respectively. The results indicated that high INHBB expression was significantly correlated with the histological type ($p < 0.05$), OS event ($p < 0.01$), DSS event ($p < 0.01$), and PFI event ($p < 0.01$) (Figures 1F–I). These results show that INHBB is significantly

10.3389/fgene.2022.933862

TABLE 1 Clinical characteristics of patients with GC based on TCGA.

Characteristic		Low expression of INHBB		High expression of INHBB		p
		N	%	N	%	
T stage	T1	15	4.1%	4	1.1%	0.060
	T2	41	11.2%	39	10.6%	
	T3	78	21.3%	90	24.5%	
	T4	48	13.1%	52	14.2%	
N stage	N0	56	15.7%	55	15.4%	0.841
	N1	48	13.4%	49	13.7%	
	N2	40	11.2%	35	9.8%	
	N3	34	9.5%	40	11.2%	
M stage	M0	165	46.5%	165	46.5%	1.000
	M1	13	3.7%	12	3.4%	
Pathologic stage	Stage I	32	9.1%	21	6%	0.228
	Stage II	48	13.6%	63	17.9%	
	Stage III	76	21.6%	74	21%	
	Stage IV	19	5.4%	19	5.4%	
Primary therapy outcome	PD	32	10.1%	33	10.4%	0.946
	SD	7	2.2%	10	3.2%	
	PR	2	0.6%	2	0.6%	
	CR	112	35.3%	119	37.5%	
Gender	Female	70	18.7%	64	17.1%	0.564
	Male	117	31.2%	124	33.1%	
Age	≤65	83	22.4%	81	21.8%	0.880
	>65	102	27.5%	105	28.3%	
Histological type	Diffuse type	32	8.6%	31	8.3%	**0.045**
	Mucinous type	6	1.6%	13	3.5%	
	Not otherwise specified	105	28.1%	102	27.3%	
	Papillary type	1	0.3%	4	1.1%	
	Signet ring type	2	0.5%	9	2.4%	
	Tubular type	41	11%	28	7.5%	
Residual tumor	R0	147	44.7%	151	45.9%	0.443
	R1	6	1.8%	9	2.7%	
	R2	10	3%	6	1.8%	
Histologic grade	G1	3	0.8%	7	1.9%	0.393
	G2	72	19.7%	65	17.8%	
	G3	110	30.1%	109	29.8%	
Anatomic neoplasm subdivision	Antrum/Distal	76	21.1%	62	17.2%	0.098
	Cardia/Proximal	21	5.8%	27	7.5%	
	Fundus/Body	59	16.3%	71	19.7%	
	Gastroesophageal junction	18	5%	23	6.4%	
	Other	4	1.1%	0	0%	
Reflux history	No	100	46.7%	75	35%	0.844
	Yes	21	9.8%	18	8.4%	
Antireflux treatment	No	81	45.3%	61	34.1%	0.888
	Yes	20	11.2%	17	9.5%	
Barretts esophagus	No	114	54.8%	79	38%	0.871
	Yes	8	3.8%	7	3.4%	

(Continued on following page)

TABLE 1 (*Continued*) Clinical characteristics of patients with GC based on TCGA.

Characteristic		Low expression of INHBB		High expression of INHBB		p
		N	%	N	%	
H pylori infection	No	82	50.3%	63	38.7%	0.782
	Yes	9	5.5%	9	5.5%	
OS event	Alive	130	34.7%	98	26.1%	**< 0.001**
	Dead	57	15.2%	90	24%	
DSS event	Alive	141	39.8%	122	34.5%	**0.004**
	Dead	32	9%	59	16.7%	
PFI event	Alive	138	36.8%	113	30.1%	**0.007**
	Dead	49	13.1%	75	20%	

Bold: N, number; %, percentage; p, *p*-value

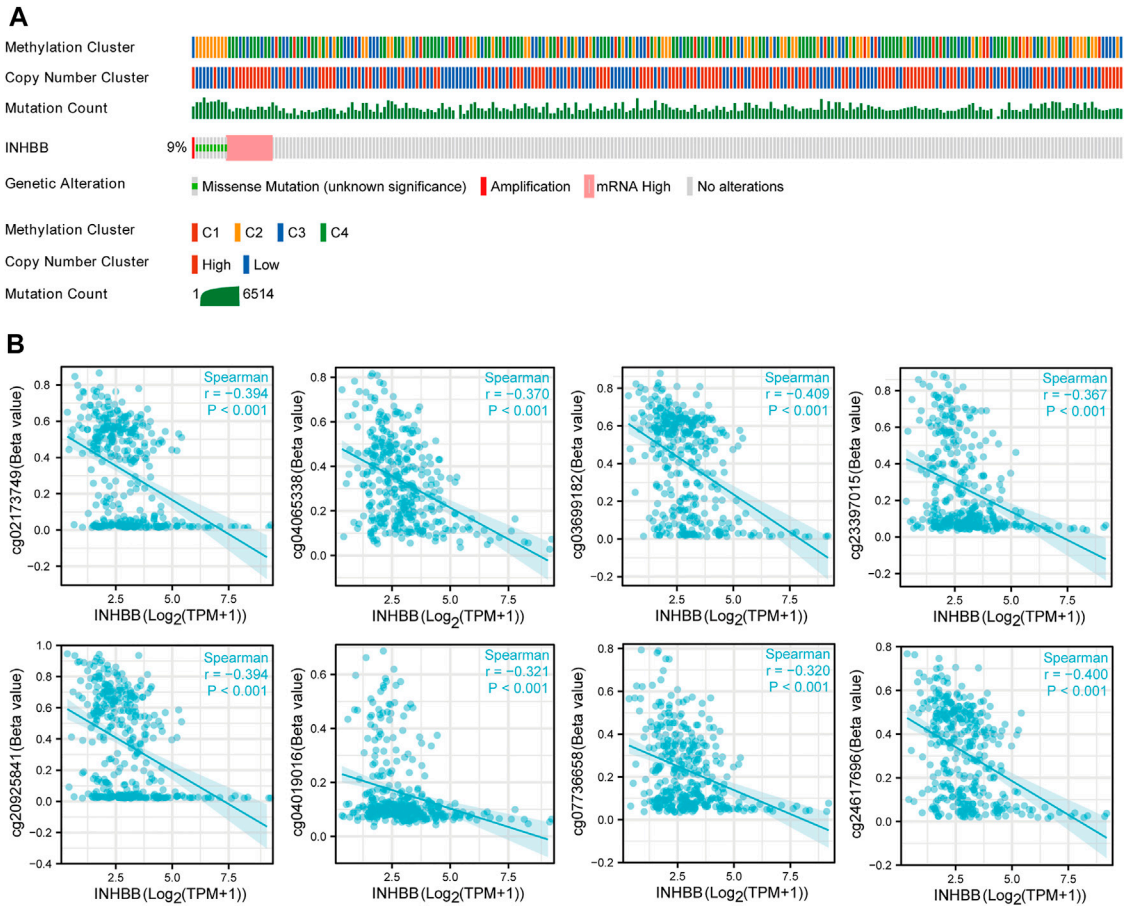


FIGURE 2
The epigenetic changes of INHBB in GC. **(A)** Genetic alteration of INHBB in GC, **(B)** Correlation of methylation and INHBB expression in GC.

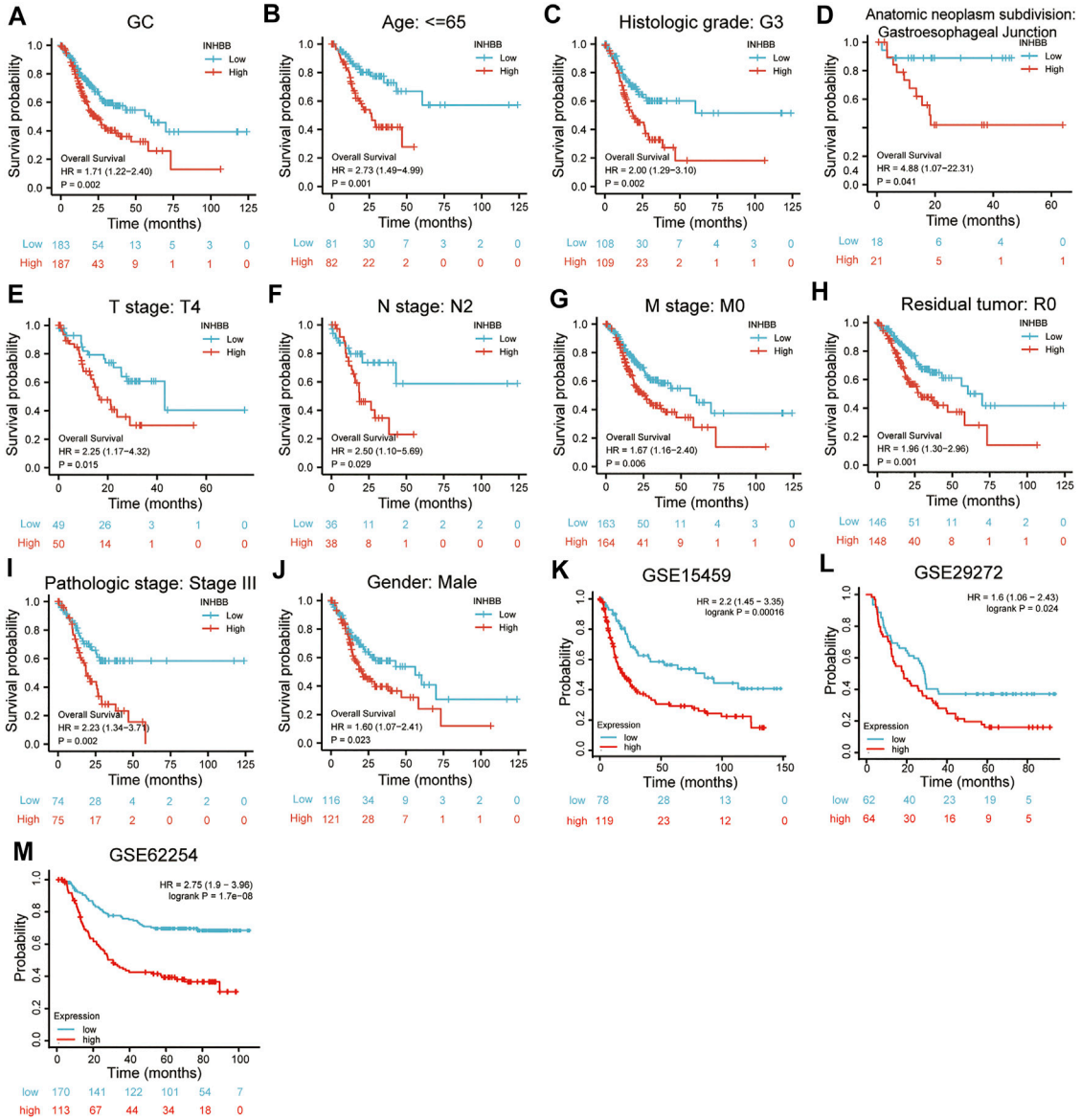


FIGURE 3
Kaplan-Meier survival analysis of INHBB expression and **(A)** prognosis in GC patient data from TCGA, **(B)** age, **(C)**histological grade, **(D)** anatomic neoplasm subdivision, **(E)** T stage, **(F)** N stage, **(G)** M stage, **(H)** residual tumor, **(I)** pathologic stage, **(J)** gender, prognosis in GC data from **(K)** GSE15459, **(L)** GSE29272, **(M)** GSE62254.

upregulated in GC and correlated with patient clinical characteristics.

To further clarify the potential mechanisms of abnormal INHBB expression, we explored the mutation alteration of INHBB in GC by using cBioPortal databases, as the result showed that INHBB had less than 9% missense mutations and gene amplifications in GC (Figure 2A). Furthermore, methylation data in TCGA databases were analyzed to evaluate the pre-transcriptional modification status of INHBB in GC. The results showed that the methylation levels were negatively correlated with INHBB expression in GC (Figure 2B).

High inhibin subunit beta B expression correlated to poor prognosis in gastric cancer

In order to identify the effect of INHBB expression on patient survival, we divided patients with GC into two groups of high and

low expression by the mean expression value of INHBB. The results of Kaplan-Meier analysis indicated that high INHBB expression was related to a poor prognosis of overall survival (OS) in seven cancer types, which is contrary to that in THCA, including BLCA, COAD, GBMLGG, HNSC, COADREAD MESO and UVM (Supplementary Figure S1). Next, we explored the relationship between high INHBB expression and disease-specific survival (DSS) as well as progression-free survival (PFI). The data showed that upregulated expression of INHBB has a worse DSS in BLCA, COAD, GBMLGG, COADREAD, KIRP, LIHC, UVM (Supplementary Figure S2) and a worse PFI in COAD, COADREAD, LUSC, UVM (Supplementary Figure S3).

In patients with GC, high INHBB expression correlated with poor OS (HR = 1.71, p = 0.002) (Figure 3A). Subgroup analysis indicated that high expression of INHBB was significantly correlated with poor OS in GC in the following categories: patients less than 65 years old (HR = 2.73, p = 0.001), male patients (HR = 1.60, p = 0.023), T4 (HR = 2.25, p = 0.015), N2 (HR = 2.5, p = 0.029), M0 (HR = 1.67, p = 0.006), G3 (HR = 2.00, p = 0.002), pathological stage III (HR = 2.23, p = 0.002), residual tumor R0 (HR = 1.96, p = 0.001), and those with gastroesophageal junction (HR = 4.88, p = 0.041). These data are shown in Figures 3B–J. Simultaneously, we conducted a Kaplan–Meier survival analysis of DSS and PFI to identify the prognostic role of INHBB in GC; the results showed that INHBB expression correlated with adverse prognosis of PFI (HR = 1.66, p = 0.006) and DSS (HR = 2.08, p = 0.001), subgroup analysis of DSS and PFI showed that T3/T4, N2, M0, R0, G3, age less than 65 years old, no barretts esophagus, no reflux history, and anatomic neoplasm subdivision cardia/proximal are associated poor DSS and PFI in GC (Supplementary Figures S4, S5).

Furthermore, we used three GEO datasets (GSE15459, GSE29272, and GSE62254) as the validation cohort to verify the transferability and reproducibility of the prognostic role of INHBB in GC. The results suggested that patients with increased INHBB expression in the three datasets had shorter OS (Figures 3K–M).

In addition, based on the patient clinical data from GC tissue microarray, Kaplan–Meier survival analysis and log-rank statistical test were used to validate the correlation between INHBB expression levels and OS in patients with GC. Surprisingly, the same result was obtained: high INHBB expression portends a poor OS in patients with GC (Figure 5E).

Higher inhibin subunit beta B expression is an independent prognostic factor in gastric cancer

To explore the association between INHBB expression and clinical features, we conducted Cox regression analyses. The univariate Cox analysis showed that high INHBB expression is significantly related to OS (p = 0.002) (Figure 4A). Then, we further analyzed the data using multivariate Cox analysis and found that INHBB expression had a significant relevance with OS (HR = 2.014, 95% CI = 1.365–2.974, p < 0.001), as well as age (HR = 1.519, 95% CI = 1.020–2.261, p = 0.039) and residual tumors (HR = 2.844, 95% CI = 1.637–4.940, p < 0.001) (Figure 4A); these results suggested that these three indicators may be independent risk factors in GC. Moreover, we constructed a nomogram prediction model based on the results of Cox analyses, indicators included in the model include residual tumor, INHBB, age, TNM stage, and pathologic stage. The C-index of the nomogram model was 0.663 indicating a good prediction ability, and the figure showed that the prognostic prediction value of INHBB was better than that of the age and TNM stage, both of which are classical traditional clinical indicators (Figure 4B). Subsequently, we evaluated the prediction ability and agreement of this prediction model using time-dependent receiver operating characteristic (ROC) and calibration curves. The 1-, 3-, and 5-year area under the curve (AUC) values of the nomogram model were 0.638, 0.75, and 0.786 (Figure 4C). Moreover, the calibration plots of the nomogram showed an excellent agreement in the 1- and 3-year OS rates compared with that of the ideal model (Figures 4D, E). These data indicated that INHBB plays a key role in the prognosis of GC.

Validation of inhibin subunit beta B expression and investigation of its prognostic role using gastric cancer tissue microarray

To further elucidate the role of INHBB in GC, we performed the expression and prognosis analysis of INHBB using a GC tissue microarray. First, we explored the protein expression level of INHBB in GC and normal tissues by IHC. The results indicated that the protein expression level of INHBB in GC tissues was higher than that in paracarcinoma tissues (p < 0.001) (Figures 5A–C). In addition, we performed univariate and multivariate Cox analyses using clinical microarray data and validated the results that high INHBB expression could be an independent risk factor in GC (Figure 5D). Kaplan–Meier survival analysis showed that patients with GC with high expression of INHBB have a worse prognosis, which is consistent with the results mentioned above (Figure 5E). These results showed that upregulated INHBB expression is associated with an adverse prognosis in GC.

Functional and pathway enrichment analysis

To elucidate the potential biological functions of INHBB in GC, GO and KEGG enrichment analysis was performed using

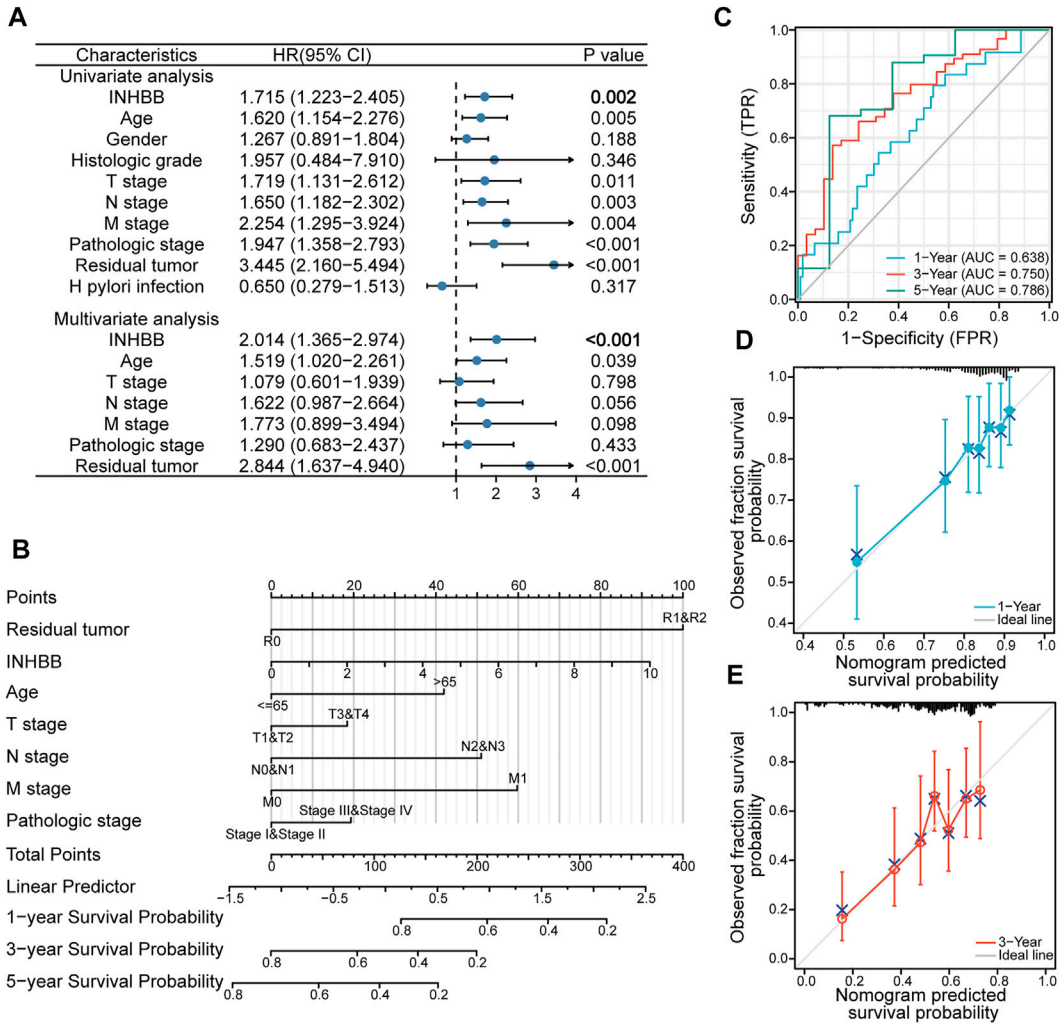


FIGURE 4
Clinical risk prediction models of INHBB in GC. **(A)** Cox regression analysis of INHBB in GC patient data from TCGA, **(B)** nomogram model based on Cox regression analysis, **(C)** Time-dependent analysis of ROC curve to evaluate nomogram performance, **(D,E)**1-year and 3-year nomogram calibration plot.

the R software. The results indicated that the most significant terms for GO enrichment were collagen−containing extracellular matrix, external encapsulating structure organization, extracellular matrix organization, synaptic membrane, and gated channel activity (Figure 6A). The results of KEGG enrichment analysis indicated that INHBB-associated genes were mostly enriched in the following pathways: calcium signaling pathway, PI3K-Akt signaling pathway, cAMP Signaling pathway, and Wnt signaling pathway (Figure 6B).

In addition, GSEA was performed on HALLMARK gene set terms and five pathways, including hallmark epithelial-mesenchymal transition, myogenesis, coagulation, KRAS signaling, and apical junction were identified as significantly enriched (Figure 6C). These results indicated that INHBB plays a key role in cancer promotion through multiple mechanisms.

Correlation of inhibin subunit beta B expression and immune infiltration

In the tumor microenvironment, different immune infiltration levels were significantly correlated with OS in patients with tumors. The findings mentioned above indicated that INHBB has a significant impact on the prognosis of patients with GC. Therefore, estimating the association between the expression of INHBB and immune infiltration level is reasonable. We first investigated the correlation between

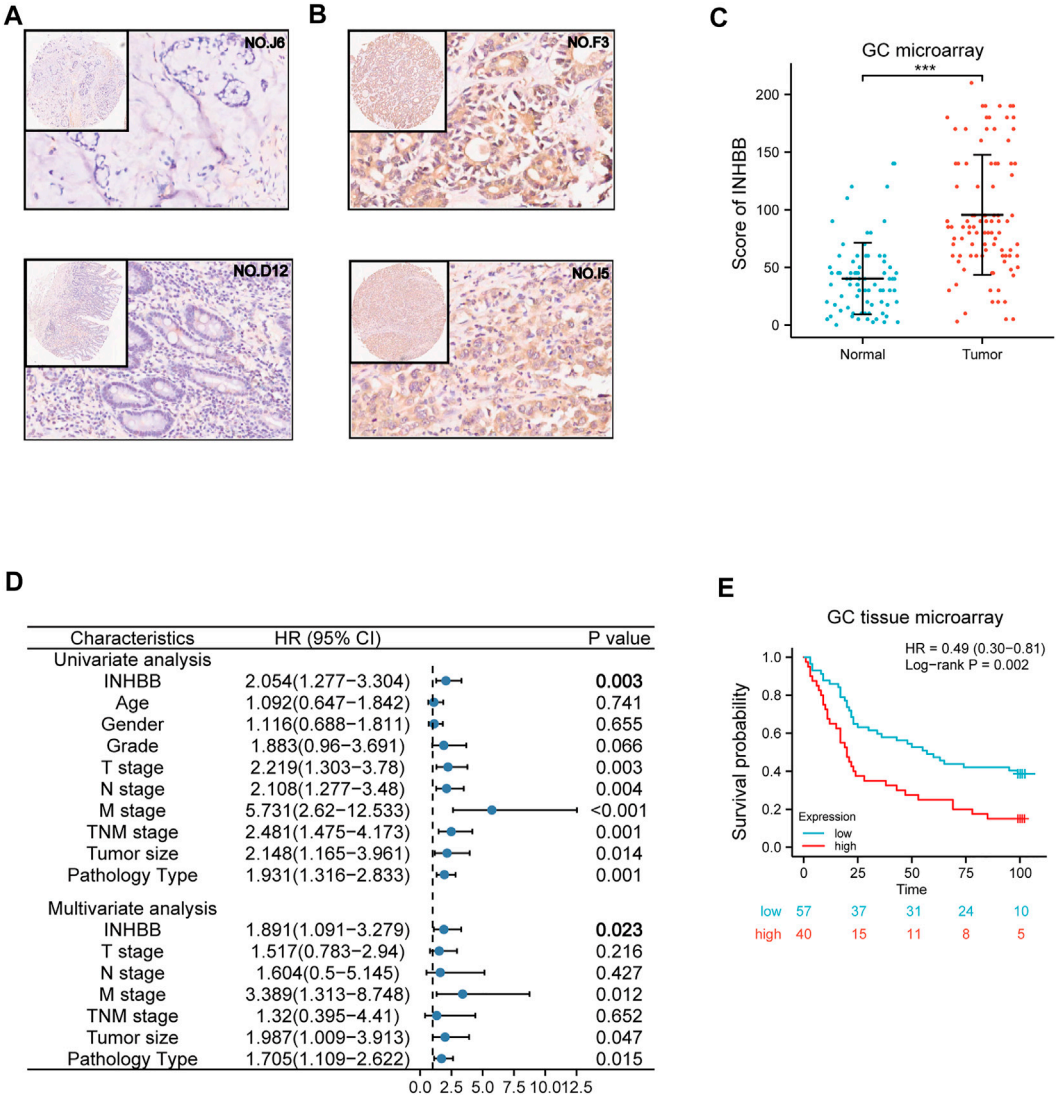


FIGURE 5
Validation of the role of INHBB in GC. **(A)** IHC of normal tissues (4X and 20X scopes), **(B)** IHC of GC tissues (4X and 20X scopes), **(C)** differential expression of INHBB in GC tissues and normal tissues, **(D)** Cox regression analysis of INHBB and clinical features in GC, **(E)** Kaplan-Meier survival analyses of INHBB in GC patient data form GC tissue microarray.

INHBB expression and infiltration level of 24 immune cell subtypes by single-sample GSEA (ssGSEA) (Figure 7A) and observed a significant correlation between INHBB expression and infiltration of immune cells such as Th2 cells ($p < 0.001$), NK cells ($p < 0.001$) and Tem ($p < 0.001$). In addition, we evaluated the association between the immune infiltration level and INHBB expression by various methods. The results suggested that INHBB expression was significantly correlated with macrophages (R = 0.2, p = 2e-05), endothelial cell (R = 0.38, $p < 2.2e-16$) and cancer-associated fibroblast (R = 0.25, $p < 9.3e-08$) in GC data from TCGA (Figure 7B, Supplementary Figure S6), these results were also validated using GSE15459 and GSE62254 datasets (Figure 7B). Furthermore, we further performed Kaplan-Meier analysis for immune infiltrates to visualize the survival difference in GC. As the figure shows, macrophage infiltration significantly correlated with GC prognosis ($p = 0.004$, Figure 7C). Thus, patients with GC have a worse prognosis with high INHBB expression and high macrophage infiltration compared with those with low macrophage infiltration (HR = 1.89, $p = 0.0144$, Figure 7D). Using the Kaplan–Meier plotter, we validated that high INHBB expression and macrophage-enriched infiltrates affect the prognosis of patients with GC (Figure 7E). Particularly, these results remind us whether the polarization of macrophages will

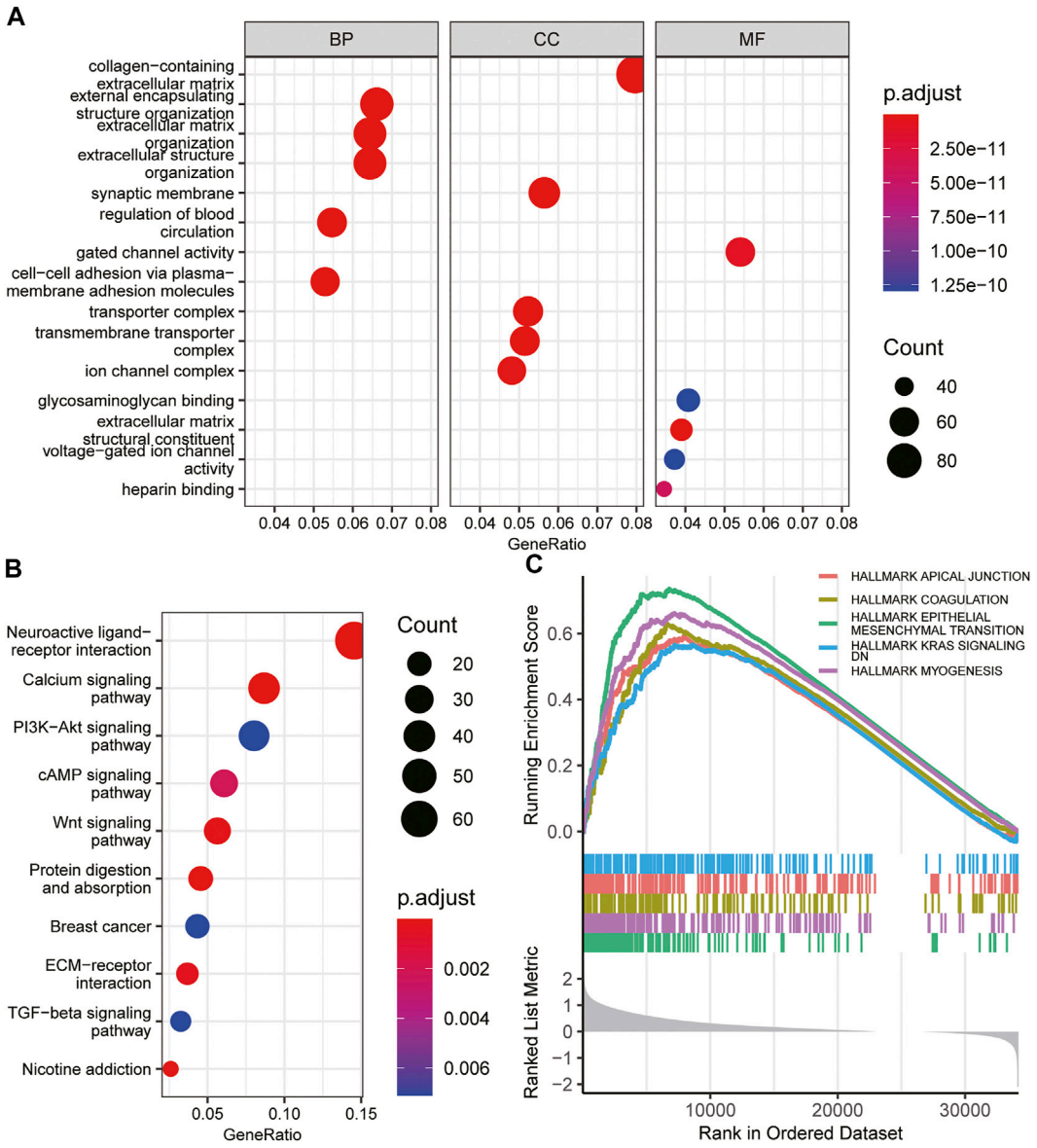


FIGURE 6
Functional enrichment of INHBB in GC. **(A)**GO enrichment analysis of INHBB in GC, **(B)** KEGG pathway analysis of INHBB in GC, and **(C)** GSEA enrichment analysis of INHBB in GC.

affect the prognosis of patients with GC, and further analysis indicated that high macrophage M2 infiltration in patients with GC with a high or low expression of INHBB may lead to a worse survival, but the results were statistically insignificant (Supplementary Figure S7). The analysis by TIMER (Table 2, Supplementary Table S1) and GEPIA (Table 3) database also showed a strong association between INHBB expression and several markers of immune cells, especially with macrophage infiltration. All these data suggested that INHBB plays a regulatory role in immune cell infiltration, especially in the macrophage infiltration of patients with GC.

Immune cell trafficking into the tumor microenvironment is mediated by chemokine/chemokine receptors (Nagarsheth et al., 2017). Therefore, we explored the association between INHBB expression and chemokines/chemokine receptors using the TISIDB database. The results of the heatmap demonstrated a significant correlation between several chemokines/chemokine receptors and INHBB expression in pan-cancers (Supplementary Figures S8A, B). In order to identify the association between INHBB expression and immune cell migration in GC, we further explored the association between INHBB expression and chemokines/chemokine receptors. The results suggested that

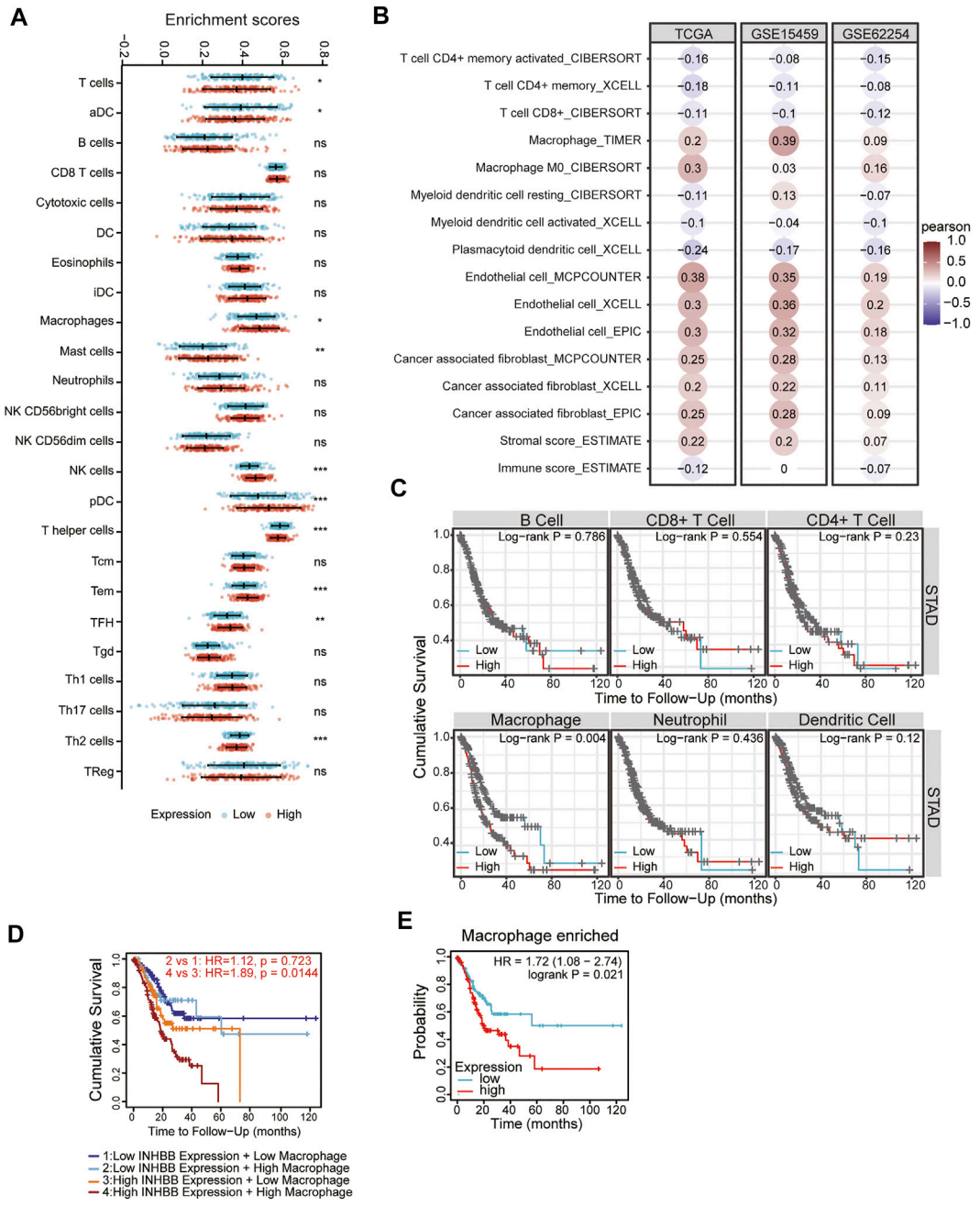


FIGURE 7
Correlation between immune infiltration and INHBB expression in GC. **(A)** INHBB expression and infiltration of 24 immune cell subtypes, **(B)** INHBB expression and immune infiltration data from TCGA and GEO, **(C)** Survival analysis for immune infiltrates and INHBB expression, **(D)** Survival analysis for differential expression of INHBB and macrophage infiltration, **(E)** Survival analysis for INHBB expression and macrophage enriched infiltration.

INHBB expression was negatively correlated with CXCL3 (r = −0.215, p = 1.02e-05), CXCL10 (r = −0.211, p = 1.51e-05), and CXCL11 (r = −0.228, p = 2.95e-06) (Figure 8A). These results suggested a negative correlation between the expression of INHBB and chemokines/chemokine receptors.

Immunotherapy for cancer is currently thriving and immune-checkpoint blockade is a new treatment option for cancers. Therefore, we explored the association between the expression of INHBB and immunoinhibitors/immunostimulators in several cancer types (Supplementary

TABLE 2 Correlation analysis between INHBB and macrophage associated markers in TIMER.

Cell type	Gene marker	None		Purity	
		Cor	p	Cor	p
Macrophage	CD68	0.013	0.79	0	0.992
	ITGAM	0.122	*	0.119	*
M1	NOS2	−0.117	*	−0.125	*
	ROS	0.138	**	0.149	**
	IRF5	0.169	***	0.168	**
	PTGS2	0.133	**	0.141	**
M2	ARG1	0.013	0.786	0.025	0.622
	MRC1	0.094	0.056	0.088	0.087
TAM	CCL2	0.234	***	0.234	***
	CCR5	0.002	0.973	−0.008	0.884
	CD80	−0.04	0.412	−0.038	0.465
	CD86	0.056	0.252	0.055	0.283
Monocyte	CD14	0.138	**	0.13	*
	CD16	0.048	0.327	0.05	0.331
	CD115	0.159	**	0.148	**

Figures S8C, D). The results indicated that INHBB expression were positively correlated with immunoinhibitors such as KDR (r = 0.251, p = 2.43e-07),TGFB1(r = 0.277, p = 1.17e-08), and TGFBR1(r = 0.218, p = 7.79e-06), whereas it was negatively correlated with LGALS9 (r = −0.254, p = 1.79e-07) (Figure 8B); as for immunostimulators, INHBB expression were positively associated with CD276 (r = 0.266, p = 4.36e-08), CXCL12 (r = 0.278, p = 9.86e-09),ENTPD1 (r = 0.222, p = 5.17e-06) and negatively correlated with that of ICOS(r = −0.205, p = 2.68e-05), KLRC1 (r = −0.216, p = 9.76e-06), and TNFSF13(r = −0.289, p = 2.22e-09) (Figure 8C). These results suggested that INHBB plays a regulatory role in tumor immunity.

Discussion

GC remains on the top list of cancers accounting for the highest number of deaths in the world (Bray et al., 2018). More than 30% of patient recurrence occurs within 5 years after treatment with chemotherapy, radiotherapy, and surgical resection, indicating that GC is a heterogeneous disease (Sohn et al., 2017). Protein glycosylation is a common modification occurring post-translationally in all animals with more than 90% of cell-surface proteins and lipid glycosylation helps to generate post-genomic diversity (Sweeney et al., 2018). Glycosylation alteration has been reported in several cancer types. As glycoproteins may be secreted or shed into the circulation, they can be regarded as potential biomarkers (Silsirivanit, 2019). These aberrantly expressed glycoproteins, including MUC1, MUC4, and MUC13 play important roles in tumor progression and treatment and are commonly referred to as tumor-related glycoproteins (Dhanisha et al., 2018). INHBB is a glycoprotein belonging to the TGF-β family. Recently, it has been reported that INHBB affects the development and prognosis in different tumors; however, there are few reports on the role of INHBB in GC. In the present study, we first explored the expression of INHBB using TCGA and GEO databases and discovered that INHBB was more highly expressed in GC tissues than in normal tissues, and the upregulated INHBB expression correlated with worse prognosis in patients with

TABLE 3 Correlation analysis between INHBB and macrophage associated markers in GEPIA.

Cell type	Gene marker	Tumor		Normal (TCGA)		Normal (TCGA + GTEx)	
		Cor	p	Cor	p	Cor	p
M1	NOS2	−0.12	*	0.36	*	0.28	***
	ROS	0.16	**	−0.003	0.99	−0.1	0.15
	IRF5	0.16	**	−0.21	0.23	0.13	0.056
	PTGS2	0.15	**	0.63	***	0.42	***
M2	ARG1	0.082	0.097	0.16	0.34	0.27	***
	MRC1	0.074	0.13	0.55	***	0.34	***
TAM	CCL2	0.22	***	0.38	*	0.32	***
	CCR5	0.029	0.56	−0.26	0.13	−0.14	*
	CD80	−0.051	0.31	−0.25	0.14	0.12	0.082
	CD86	0.032	0.51	−0.13	0.46	0.17	*
Monocyte	CD14	0.12	*	0.23	0.18	0.32	***
	CD16	0.033	0.51	0.091	0.6	0.38	***
	CD115	0.13	**	0.17	0.32	0.025	0.72

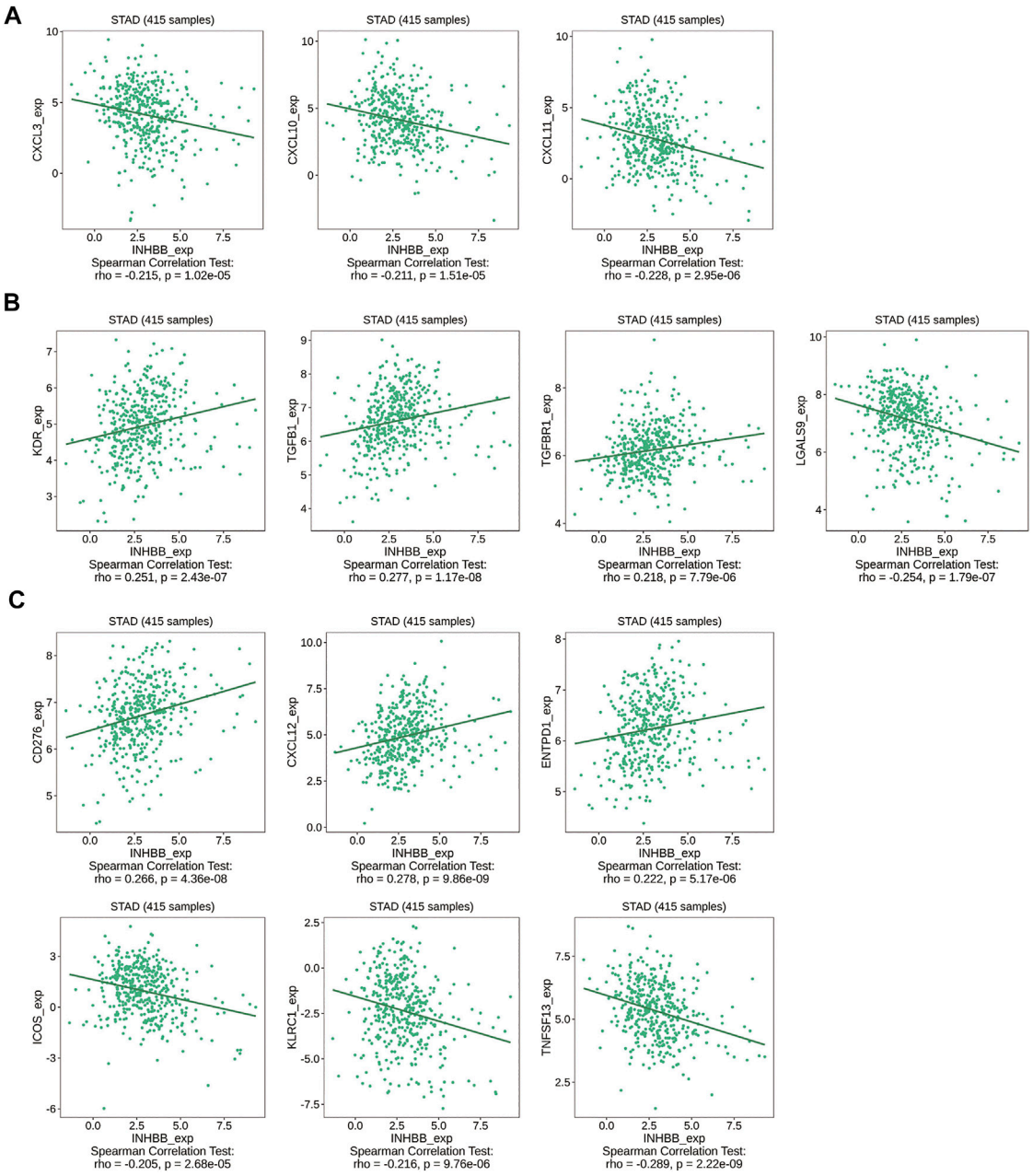


FIGURE 8
Correlation of INHBB expression and **(A)** chemokines/chemokine receptors, **(B)** immunoinhibitors, **(C)** immunostimulators.

GC. In addition, upregulated INHBB expression was found to be associated with poor clinicopathological features. Surprisingly, all these results were validated by our clinical patient data, suggesting that INHBB could be a novel prognostic biomarker for patients with GC.

Genetics and epigenetics play a crucial role in regulating cancer progression and tumor cell evasion (Chakravarthi et al., 2016). Historically, genetic alterations such as copy number alterations and somatic mutations have been used to assess tumor evolution; however, with the advancement of research in genetic alterations, several studies showed that cell lines with a high degree of genetic homogeneity are accompanied by an increased rate of cell-to-cell variability due to epigenetic alterations (Mazor et al., 2016). Analysis of genetics and epigenetics alterations is, therefore, helpful to understand the role of gene expression in cancer progression. Subsequently, we

conducted the analysis of INHBB expression by cBioPortal and found that INHBB somatic mutations and amplifications were found to be more frequent than deletions in GC. Methylation is the best studied epigenetic modification, and in general, methylation of CpG sites in the promoter is thought to silence the expression of the binding transcription factors. As expected, we found that the expression of INHBB was negatively associated with the methylation level in GC. However, it is worth noting that we only analyzed the correlation between INHBB expression and methylation sites, and whether the change in INHBB expression in GC is related to its own methylation modification still needs to be verified by subsequent experiments. Taken together, these results suggested that genetics or epigenetics alterations may affect the expression of INHBB and thus promote the progression of GC.

The tumor microenvironment is crucial for tumor progression; it contains extracellular matrix (ECM), fibroblasts, vasculature, and immune cells (Ahmed et al., 2016). ECM surrounds the tumor cells and supports their growth, survival, and eventually invasive capacity (Gordon-Alonso et al., 2017). The GO analysis showed that the structural component of ECM is the main biological function of INHBB-associated genes in GC. In addition, ECM-receptor interaction was the major enriched pathway according to the results of KEGG analysis, suggesting INHBB expression may have a specific correlation with the tumor microenvironment in GC. The other pathways observed in the KEGG analysis were involved in tumor development and progression. Ma et al. (2020) found that an imprinted gene regulates the function of cellular autophagy and epithelial–mesenchymal transition (EMT) via the PI3K/Akt pathway, thereby accelerating the deterioration progress of colorectal cancer. Calcium and cAMP are essential secondary messengers in signaling pathways in cells and both of them are central to tumorigenesis. EPAC1, one of the major downstream effectors of cAMP, regulates the metastatic properties such as proliferation and migration in triple-negative BRCA cells (Kumar et al., 2018). Similarly, the calcium signaling pathway has been found to mediate apoptosis and proliferation in non-small cell lung cancer cells (Wang et al., 2020). Moreover, pathway crosstalk of Wnt/calcium has been involved in the regulation of cell migration in cancer progression (Chuderland and Seger, 2008). Subsequently, the results of GSEA based on Hallmark gene set terms validated the oncogenetic property of INHBB in GC. There is a consensus on the key role of EMT and KRAS in cancers. Interestingly, the activation of coagulation correlated with tumor progression and invasion has been reported (Palumbo, 2008; Kvolik et al., 2010), and the potential mechanism may be related to the leaky tumor vasculature-enabling clotting factors in the blood entering the stroma (McEachron et al., 2010). Our data showed that INHBB is an oncogene and accelerates the development of GC in multiple ways.

As mentioned above, the tumor microenvironment is crucial for cancer progression and immune cells are key players in it. Previous studies have reported that immune cell infiltrates within a tumor are associated with disease prognosis and response to immunotherapy (Buisseret et al., 2018). Therefore, it is necessary to clarify the relationship between INHBB expression and immune infiltration in GC. In this pilot study, we identified that INHBB expression is positively correlated with the infiltration of macrophages, endothelial cells, and cancer-associated fibroblasts and is negatively associated with the infiltration of CD4+/CD8+ T cells and plasmacytoid dendritic cells. Further analyses showed that macrophage infiltration has an important impact on the prognosis of patients with GC. In tumors, macrophages are differentiated from monocytes and usually divided into two subtypes: M1, antitumor macrophages, and M2, tumor-promoting macrophages. Particularly, macrophage infiltration in tumors is heightened in the M2 subunit and not M1 (Mantovani et al., 2002). In turn, these macrophages foster a suitable microenvironment that supports tumor cell survival. An increased level of macrophage infiltration in tumors portends a worse prognosis (Qian and Pollard, 2010). Not surprisingly, our results showed that both the increased macrophage infiltration with upregulated INHBB expression correlated with poor prognosis in GC. Moreover, we found that high infiltration of M2 macrophages predicts worse survival, although the findings were not statistically significant. These findings suggested that INHBB may regulate the microenvironment of immune cell infiltration to promote GC progression.

Since chemokines secreted by tumor cells recruit immune cells to the tumor sites, we explored the association between INHBB expression and expression of chemokines/chemokine receptors in GC. The results suggested that the expression of CXCL3, CXCL10, and CXC11 were negatively correlated with INHBB expression in GC. CXCL3 is a known angiogenic chemokine and participates in the chemotaxis of neutrophils (Rainard et al., 2008), in addition, CXCL3 inhibits the growth of esophageal squamous cell carcinoma by attracting neutrophils (Chen et al., 2017). CXCL10 is induced by IFNγ and is responsible for the recruitment of T cells into tumors. Jiang, Zheng et al. found that the down-regulated expression of CXCL10 was a worse prognostic indicator in colorectal cancer (Jiang et al., 2010). Similar to CXCL10, CXC11 is also a downstream target of IFNγ. In T lymphocytes, it mediates antitumor immune responses (Hensbergen et al., 2005), which supports the result of this study that a low level of CXC11 is related to poor prognosis in GC. The immune system is complicated, and immunotherapy has changed the standard of care for multiple malignant tumors; among them, blocking inhibitory immune checkpoints has become an attractive antitumor strategy. Immunostimulators play an important role in immune activation, and the overexpression of immunoinhibitors is a clinical anti-checkpoint or combination treatment strategy. In this study, we explored the association between INHBB expression and expression of immunoinhibitors/immunostimulators and found several potential INHBB-associated targets of immunotherapy for GC.

These results suggest that INHBB plays an essential role in immune infiltration in GC.

This study also has some limitations. First, the data of expression and prognosis were obtained from multiple online databases, making it difficult to ensure authenticity and integrity; however, our microarray data included complete clinical information, and the expression of INHBB was validated by experiments. Second, we found that INHBB expression is associated with immune infiltration, although the data were mostly derived from the TCGA database. We used TCGA because it is one of the most commonly used oncology research-related portals with a complete set of data; in addition, we corroborated the results with those from the GEO database.

Conclusion

In summary, we found that INHBB was upregulated in GC tissues; high INHBB expression is an independent risk factor and is associated with poor prognosis in GC. Moreover, INHBB expression is significantly correlated with immune cell infiltration, especially macrophage infiltration. The alteration of INHBB-related immune cell infiltration may affect the prognostic outcome of GC patients, its clinical value deserves further exploration 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 authors.

Ethics statement

No human studies are presented in the manuscript.

Author contributions

SH, ZZ, and WY conceived and designed the research, WY contributed to the data collection, GH and ZY analyzed data and wrote the article, and WZ revised the article. All authors read and approved the final manuscript.

Funding

This work was supported by the Traditional Chinese Medicine Bureau of Guangdong Province (Grant No. 20201141) and The National Natural Science Foundation of China (82104604).

Conflict of interests

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.2022.933862/full#supplementary-material

References

Ahmed, M., Sottnik, J. L., Dancik, G. M., Sahu, D., Hansel, D. E., Theodorescu, D., et al. (2016). An osteopontin/CD44 Axis in RhoGDI2-mediated metastasis suppression. *Cancer Cell.* 30 (3), 432–443. doi:10.1016/j.ccell.2016.08.002

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. Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492

Buisseret, L., Pommey, S., Allard, B., Garaud, S., Bergeron, M., Cousineau, I., et al. (2018). Clinical significance of CD73 in triple-negative breast cancer: Multiplex analysis of a phase III clinical trial. *Ann. Oncol.* 29 (4), 1056–1062. doi:10.1093/annonc/mdx730

Chakravarthi, B. V., Nepal, S., and Varambally, S. (2016). Genomic and epigenomic alterations in cancer. *Am. J. Pathol.* 186 (7), 1724–1735. doi:10.1016/j.ajpath.2016.02.023

Chang, H., Brown, C. W., and Matzuk, M. M. (2002). Genetic analysis of the mammalian transforming growth factor-beta superfamily. *Endocr. Rev.* 23 (6), 787–823. doi:10.1210/er.2002-0003

Chen, C. L., Wang, Y., Huang, C. Y., Zhou, Z. Q., Zhao, J. J., Zhang, X. F., et al. (2017). IL-17 induces antitumor immunity by promoting beneficial neutrophil recruitment and activation in esophageal squamous cell carcinoma. *Oncoimmunology* 7 (1), e1373234. doi:10.1080/2162402x.2017.1373234

Chuderland, D., and Seger, R. (2008). Calcium regulates ERK signaling by modulating its protein-protein interactions. *Commun. Integr. Biol.* 1 (1), 4–5. doi:10.4161/cib.1.1.6107

Dhanisha, S. S., Guruvayoorappan, C., Drishya, S., and Abeesh, P. (2018). Mucins: Structural diversity, biosynthesis, its role in pathogenesis and as possible therapeutic targets. *Crit. Rev. Oncol. Hematol.* 122, 98–122. doi:10.1016/j.critrevonc.2017.12.006

Dicken, B. J., Bigam, D. L., Cass, C., Mackey, J. R., Joy, A. A., and Hamilton, S. M. (2005). Gastric adenocarcinoma: Review and considerations for future directions. *Ann. Surg.* 241 (1), 27–39. doi:10.1097/01.sla.0000149300.28588.23

Ethier, J. F., and Findlay, J. K. (2001). Roles of activin and its signal transduction mechanisms in reproductive tissues. *Reproduction* 121 (5), 667–675. doi:10.1530/rep.0.1210667

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

Gordon-Alonso, M., Hirsch, T., Wildmann, C., and van der Bruggen, P. (2017). Galectin-3 captures interferon-gamma in the tumor matrix reducing chemokine gradient production and T-cell tumor infiltration. *Nat. Commun.* 8 (1), 793. doi:10.1038/s41467-017-00925-6

Gutierrez, A., Demond, H., Brebi, P., and Ili, C. G. (2021). Novel methylation biomarkers for colorectal cancer prognosis. *Biomolecules* 11 (11), 1722. doi:10.3390/biom11111722

Gyorffy, B. (2021). Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. *Comput. Struct. Biotechnol. J.* 19, 4101–4109. doi:10.1016/j.csbj.2021.07.014

Hensbergen, P. J., Wijnands, P. G., Schreurs, M. W., Scheper, R. J., Willemze, R., and Tensen, C. P. (2005). The CXCR3 targeting chemokine CXCL11 has potent antitumor activity *in vivo* involving attraction of CD8+ T lymphocytes but not inhibition of angiogenesis. *J. Immunother.* 28 (4), 343–351. doi:10.1097/01.cji.0000165355.26795.27

Jiang, Z., Xu, Y., and Cai, S. (2010). CXCL10 expression and prognostic significance in stage II and III colorectal cancer. *Mol. Biol. Rep.* 37 (6), 3029–3036. doi:10.1007/s11033-009-9873-z

Kita, A., Kasamatsu, A., Nakashima, D., Endo-Sakamoto, Y., Ishida, S., Shimizu, T., et al. (2017). Activin B regulates adhesion, invasiveness, and migratory activities in oral cancer: A potential biomarker for metastasis. *J. Cancer* 8 (11), 2033–2041. doi:10.7150/jca.18714

Kumar, N., Prasad, P., Jash, E., Jayasundar, S., Singh, I., Alam, N., et al. (2018). cAMP regulated EPAC1 supports microvascular density, angiogenic and metastatic properties in a model of triple negative breast cancer. *Carcinogenesis* 39 (10), 1245–1253. doi:10.1093/carcin/bgy090

Kvolik, S., Jukic, M., Matijevic, M., Marjanovic, K., and Glavas-Obrovac, L. (2010). An overview of coagulation disorders in cancer patients. *Surg. Oncol.* 19 (1), e33–46. doi:10.1016/j.suronc.2009.03.008

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

Ma, Z., Lou, S., and Jiang, Z. (2020). PHLDA2 regulates EMT and autophagy in colorectal cancer via the PI3K/AKT signaling pathway. *Aging (Albany NY)* 12 (9), 7985–8000. doi:10.18632/aging.103117

Mantovani, A., Sozzani, S., Locati, M., Allavena, P., and Sica, A. (2002). Macrophage polarization: Tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends Immunol.* 23 (11), 549–555. doi:10.1016/s1471-4906(02)02302-5

Mazor, T., Pankov, A., Song, J. S., and Costello, J. F. (2016). Intratumoral heterogeneity of the epigenome. *Cancer Cell.* 29 (4), 440–451. doi:10.1016/j.ccell.2016.03.009

McEachron, T. A., Pawlinski, R., Richards, K. L., Church, F. C., and Mackman, N. (2010). Protease-activated receptors mediate crosstalk between coagulation and fibrinolysis. *Blood* 116 (23), 5037–5044. doi:10.1182/blood-2010-06-293126

Muttukrishna, S., Fowler, P. A., Groome, N. P., Mitchell, G. G., Robertson, W. R., and Knight, P. G. (1994). Serum concentrations of dimeric inhibin during the spontaneous human menstrual cycle and after treatment with exogenous gonadotrophin. *Hum. Reprod.* 9 (9), 1634–1642. doi:10.1093/oxfordjournals.humrep.a138765

Nagarsheth, N., Wicha, M. S., and Zou, W. (2017). Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. *Nat. Rev. Immunol.* 17 (9), 559–572. doi:10.1038/nri.2017.49

Palumbo, J. S. (2008). Mechanisms linking tumor cell-associated procoagulant function to tumor dissemination. *Semin. Thromb. Hemost.* 34 (2), 154–160. doi:10.1055/s-2008-1079255

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

Rainard, P., Riollet, C., Berthon, P., Cunha, P., Fromageau, A., Rossignol, C., et al. (2008). The chemokine CXCL3 is responsible for the constitutive chemotactic activity of bovine milk for neutrophils. *Mol. Immunol.* 45 (15), 4020–4027. doi:10.1016/j.molimm.2008.06.010

Rodgarkia-Dara, C., Vejda, S., Erlach, N., Losert, A., Bursch, W., Berger, W., et al. (2006). The activin axis in liver biology and disease. *Mutat. Res.* 613 (2-3), 123–137. doi:10.1016/j.mrrev.2006.07.002

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 (20), 4200–4202. doi:10.1093/bioinformatics/btz210

Silsirivanit, A. (2019). Glycosylation markers in cancer. *Adv. Clin. Chem.* 89, 189–213. doi:10.1016/bs.acc.2018.12.005

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

Sohn, B. H., Hwang, J. E., Jang, H. J., Lee, H. S., Oh, S. C., Shim, J. J., et al. (2017). Clinical significance of four molecular subtypes of gastric cancer identified by the cancer Genome Atlas project. *Clin. Cancer Res.* 23 (15), 4441–4449. doi:10.1158/1078-0432.ccr-16-2211

Sweeney, J. G., Liang, J., Antonopoulos, A., Giovannone, N., Kang, S., Mondala, T. S., et al. (2018). Loss of GCNT2/I-branched glycans enhances melanoma growth and survival. *Nat. Commun.* 9 (1), 3368. doi:10.1038/s41467-018-05795-0

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

Tong, S., Wallace, E. M., and Burger, H. G. (2003). Inhibins and activins: Clinical advances in reproductive medicine. *Clin. Endocrinol.* 58 (2), 115–127. doi:10.1046/j.1365-2265.2003.01686.x

Vivian, J., Rao, A. A., Nothaft, F. A., Ketchum, C., Armstrong, J., Novak, A., et al. (2017). Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* 35 (4), 314–316. doi:10.1038/nbt.3772

Wang, W. J., Mao, L. F., Lai, H. L., Wang, Y. W., Jiang, Z. B., Li, W., et al. (2020). Dolutegravir derivative inhibits proliferation and induces apoptosis of non-small cell lung cancer cells via calcium signaling pathway. *Pharmacol. Res.* 161, 105129. doi:10.1016/j.phrs.2020.105129

Werner, S., and Alzheimer, C. (2006). Roles of activin in tissue repair, fibrosis, and inflammatory disease. *Cytokine Growth Factor Rev.* 17 (3), 157–171. doi:10.1016/j.cytogfr.2006.01.001

Xia, Y., and Schneyer, A. L. (2009). The biology of activin: Recent advances in structure, regulation and function. *J. Endocrinol.* 202 (1), 1–12. doi:10.1677/joe-08-0549

Xu, Z., Li, Y., Cui, Y., and Guo, Y. (2020). Identifications of candidate genes significantly associated with rectal cancer by integrated bioinformatics analysis. *Technol. Cancer Res. Treat.* 19, 1533033820973270. doi:10.1177/1533033820973270

Yao, H. H., Aardema, J., and Holthusen, K. (2006). Sexually dimorphic regulation of inhibin beta B in establishing gonadal vasculature in mice. *Biol. Reprod.* 74 (5), 978–983. doi:10.1095/biolreprod.105.050286

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 (5), 284–287. doi:10.1089/omi.2011.0118

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 September 2022
DOI 10.3389/fgene.2022.938510

Check for updates

OPEN ACCESS

EDITED BY
Lan Zhao,
Stanford University, United States

REVIEWED BY
Michael Greenwood,
University of Bristol, United Kingdom
Weimin Lin,
Sichuan University, China

*CORRESPONDENCE
Tang Liu,
liutang1204@csu.edu.cn
Xianzhe Tang,
txztangxianzhe@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 07 May 2022
ACCEPTED 12 August 2022
PUBLISHED 09 September 2022

CITATION
Lin Z, Wu Y, Xiao X, Zhang X, Wan J, Zheng T, Chen H, Liu T and Tang X (2022), Pan-cancer analysis of CREB3L1 as biomarker in the prediction of prognosis and immunotherapeutic efficacy. *Front. Genet.* 13:938510. doi: 10.3389/fgene.2022.938510

COPYRIGHT
© 2022 Lin, Wu, Xiao, Zhang, Wan, Zheng, Chen, Liu 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.

Pan-cancer analysis of CREB3L1 as biomarker in the prediction of prognosis and immunotherapeutic efficacy

Zhengjun Lin[1†], Yanlin Wu[1†], XunGang Xiao[2], Xianghong Zhang[1], Jia Wan[1], Tao Zheng[1], Hongxuan Chen[1], Tang Liu[1*] and Xianzhe Tang[2*]

[1]Department of Orthopedics, The Second Xiangya Hospital, Central South University, Changsha, Hunan, China, [2]Department of Orthopedics, Chenzhou No. 1 People's Hospital, Chenzhou, Hunan, China

Background: CAMP response element binding protein 3-like 1 (CREB3L1) has been indicated as a critical biomarker and can modulate multifaced behaviors of tumor cells in diverse cancers. However, a systematic assessment of CREB3L1 in pan-cancer is of absence, and the predictive value of CREB3L1 in cancer prognosis, the tumor immune microenvironment and the efficacy of immunotherapy remains unexplored.

Methods: CREB3L1 expression in 33 different cancer types was investigated using RNAseq data from The Cancer Genome Atlas (TCGA) database. The characteristics of CREB3L1 alternations were illustrated in cBioPortal database. The prognostic and clinicopathological value of CREB3L1 was analyzed through clinical data downloaded from the TCGA database. The potential role of CREB3L1 in the tumor immune microenvironment was illustrated by utilizing CIBERSORT and ESTIMATE algorithms, and TISIDB online database. The associations between CREB3L1 expression and tumor mutation burden (TMB), and microsatellite instability (MSI) were assessed by spearman's rank correlation coefficient. Furthermore, Gene Set Enrichment Analysis (GSEA) was conducted to explore the potential biological functions and downstream pathways of CREB3L1 in different human cancers. The correlations of CREB3L1 expression with PD-1/PD-L1 inhibitors efficacy and drug sensitivity were also investigated.

Results: The expression of CREB3L1 was abnormally high or low in several different cancer types, and was also strictly associated with the prognosis of cancer patients. CREB3L1 expression levels have a strong relationship with infiltrating immune cells, including regulatory T cells, $CD8^+$ T cells, macrophages, B naïve cells, dendritic cells and mast cells. CREB3L1 expression was also correlated with the expression of multiple immune-related biomolecules, TMB, and MSI in several cancers. Moreover, CREB3L1 had promising applications in predicting the immunotherapeutic benefits and drug sensitivity in cancer management.

Conclusions: Our results highlight the value of CREB3L1 as a predictive biomarker for the prognosis and immunotherapy efficacy in multiple cancers, and CREB3L1 seems to play key roles in the tumor immune microenvironment, suggesting the role of CREB3L1 as a promising biomarker for predicting the prognosis and immune-related signatures in diverse cancers.

KEYWORDS

CREB3L1, pan-cancer, prognostic biomarker, tumor immune microenvironment, immunotherapy efficacy, drug sensitivity

Introduction

In the past decades, advances in cancer immunotherapy have evidently improved the therapeutic efficacy in the treatment of diverse cancers via the interactions between the human immune system and tumor (Yang, 2015; Kennedy and Salama, 2020). Despite immunotherapies have been successfully applied across a wide range of human malignancies, the majority of cancer patients possess limited or no responses to immunotherapies (Sharma et al., 2017; O'Donnell et al., 2019). Emerging evidence has confirmed that the tumor immune microenvironment, such as infiltrating immune cells, and the expression of immune checkpoints is critically modulated by multiple biomolecules, which may significantly influence the efficacy of immunotherapy (Lei et al., 2020; Kumagai et al., 2021). Hence, developing a promising biomarker to predict the tumor immune microenvironment-related signatures and immunotherapeutic responses of cancer patients at early stage is warranted.

The CREB3 family of transcription factors are endoplasmic reticulum (ER) localized proteins and belong to the large bZIP family (Sampieri et al., 2019). CREB3 family members play important roles in modulating tissue development, lipids metabolism, proteins secretion, cell differentiation, and tumorigenesis (Sampieri et al., 2019). Notably, recent studies have reported CREB3L1, a critical member of CREB3 family, participate in cancer initiation and progression, and can serve as a promising clinical biomarker for cancer patients (Vellanki et al., 2013; Sampieri et al., 2019; Morishita et al., 2021). For instance, it has been found that the combination of doxorubicin and oncolytic vaccinia virus can induce cell death through the activation of CREB3L1 in ovarian cancer (Mistarz et al., 2021). Several studies have indicated that CREB3L1 may function as the suppressor of cellular metastasis and proliferation through a similar mechanism to p53 (Denard et al., 2011). Mellor et al. have reported that CREB3L1 is a critical tumor suppressor that can inhibit metastasis, invasion, and angiogenesis *in vitro* (Mellor et al., 2013). On the contrary, several studies have reported opposite conclusions that CREB3L1 can contribute to cancer cells metastasis and invasion. CREB3L1 expression is significantly upregulated in distant metastasis of triple-negative breast cancers (TNBC) patients. Besides, cancer-specific PERK signaling induces cell invasion and metastasis through directly targeting CREB3L1 in breast cancer through inducing epithelial-mesenchymal transition (EMT) by an ATF4-Fra-1 interaction. Mechanistically, CREB3L1 can promote breast cancer cell invasion and metastasis through inducing ECM production by activating FAK, and a chemical inhibitor of proteases, AEBSF can inhibit breast cancer cell invasion by inhibiting CREB3L1 expression (Feng et al., 2017). Furthermore, it has been identified that CREB3L1 can function as a prognostic and diagnostic biomarker, and predict the chemotherapeutic responses in some cancer types (Denard et al., 2018; Liu et al., 2018; Morishita et al., 2021). These results indicate the critical roles of CREB3L1 in cancer development, however, different studies present controversial results which need to be clarified. Moreover, the role of CREB3L1 in modulating the tumor immune microenvironment is still largely unknown, and no systematic pan-cancer investigation has focused on CREB3L1.

In this pan-cancer analysis research, we analyzed the expression pattern of CREB3L1 and its relationship with the prognosis of cancer patients. Furthermore, we explored the relationship between CREB3L1 and the immune cells infiltration, the expression of immune-related biomarkers, and immunotherapeutic efficacy. Our study sheds highlight on the promising role of CREB3L1 as a prognostic biomarker, and an effective predictor for the immunotherapeutic efficacy and drug sensitivity in pan-cancer, as well as the potential mechanism by which CREB3L1 modulates the tumor immune microenvironment.

Materials and methods

Data acquisition

RNA sequencing (RNA-seq) data, somatic mutation, and related clinical data of 33 cancer types were downloaded from TCGA dataset, and RNA-seq data of 54 normal tissue samples in Genotype-Tissue Expression (GTEx) dataset were extracted from UCSC Xena (https://xena.ucsc.edu/) (Goldman et al., 2015). The Strawberry Perl software (Version 5.32.1.1, http://strawberryperl.com/) was used to extract the CREB3L1 expression data from these downloaded datasets, each expression value was

normalized by log2 transformation. Cellular CREB3L1 expression data in 24 different cancer cell lines were downloaded from the Broad Institute Cancer Cell Line Encyclopedia (CCLE) portal (https://portals.broadinstitute.org/ccle/about) (Barretina et al., 2012). The differential CREB3L1 expression between diverse cancer types and their corresponding normal samples were evaluated by student *t*-test and visualized via R package "ggplot2".

Genetic alternation analysis

Genetic alteration analysis of CREB3L1 was performed based on the "TCGA Pan-Cancer Atlas Studies" dataset in the cBioPortal (https://www.cbioportal.org), a dataset that includes diverse data types, such as DNA methylation data, transcription data, non−synonymous mutations, and DNA copy number data (Gao et al., 2013). We applied the data to investigate the mutation rate and characteristics of CREB3L1 alterations in pan-cancer.

Prognostic analysis

Survival and clinical phenotype data were acquired from the TCGA dataset. The Kaplan-Meier analysis and the univariate Cox regression analyses were performed to comprehensively investigate the relationship between the CREB3L1 expression levels and the prognosis of patients in pan-cancer, including overall survival (OS), progression-free survival (PFS), disease-free survival (DFS), and disease-specific survival (DSS) in pan-cancer by R-packages "survival", "survminer", "forestplot", "limma" and "ggpubr".

Relationship between CREB3L1 expression and immunity

Estimation of Stromal and Immune cells in Malignant Tumor tissues using Expression data (ESTIMATE), a bioinformatics algorithm that infers the degree of infiltration of immune cells and stromal cells in tumor samples was exploited to evaluate the stroma and immune scores of each sample using existing CREB3L1 expression data in pan-cancer via the R package "estimate" (Yoshihara et al., 2013). CIBERSORT, a bioinformatics algorithm that quantifies the cellular composition of tissue samples by their gene expression, was performed to calculate the correlation coefficient for diverse immune cells in 32 different cancer types (Chen et al., 2018). Associations between the immune cell infiltration degree and CREB3L1 expression in pan-cancer were investigated by R-packages "ggExtra", "ggplot2" and "ggpubr". The immune cells infiltration in pan-cancer was also evaluated utilizing MCPcounter (Dienstmann et al., 2019), XCELL (Aran, 2020), TIMER (Li et al., 2020), QUANTISEQ (Plattner et al., 2020), EPIC (Racle et al., 2017), and IPS (Charoentong et al., 2017) algorithms in Sangerbox (http://www.sangerbox.com/tool). Furthermore, we investigated the correlation of CREB3L1 expression with the expression of immune-related biomolecules, and immune molecular subtypes in the TISIDB database (http://cis.hku.hk/TISIDB/index.php). The TISIDB database is an online portal integrating abundant data types, such as high-throughput screening and RNA-seq data from the TCGA database for analyzing the tumor and immune system interaction (Ru et al., 2019).

Gene set enrichment analysis

To explore the underlying biological functions and signaling pathways of CREB3L1 in pan-cancer, we subsequently performed the GSEA. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene ontology (GO) gene database were extracted from the GSEA website (https://www.gsea-msigdb.org/gsea/downloads.jsp). GSEA was conducted utilizing the R packages "limma", "org.Hs.eg.db", "enrichplot" and "clusterProfiler" with the following parameters: nPerm = 100, and p-value-Cutoff = 1.

Correlation of CREB3L1 expression with TMB and MSI

The TMB and MSI scores were downloaded from TCGA pan-cancer mutation data (https://tcga.xenahubs.net). TMB, which represents the alteration frequency in a distinct cancer type, is a quantifiable biomarker assessing the immunotherapeutic response of PD-1 antibodies (Chan et al., 2019). MSI is a new microsatellite allele in a tumor due to a deficient mismatch repair system of DNA, and has been recognized as a critical tumor biomarker (Vilar and Gruber, 2010). A Perl script were performed to calculate TMB scores and revised by dividing the total length of exons. The correlation analysis between the CREB3L1 expression levels and TMB or MSI was elaborated using Spearman's rank correlation coefficient. R-package "fmsb" was used to draw Radar plots.

Relationship between CREB3L1 expression and the efficacy of immunotherapy

To evaluate the correlation between CREB3L1 expression and the efficacy of immunotherapy, we enrolled three databases including GSE78220 (melanoma) (Hugo et al., 2016), GSE67501 (renal cell carcinoma) (Ascierto et al., 2016), and IMvigor210 (metastatic urothelial cancer), which were downloaded from

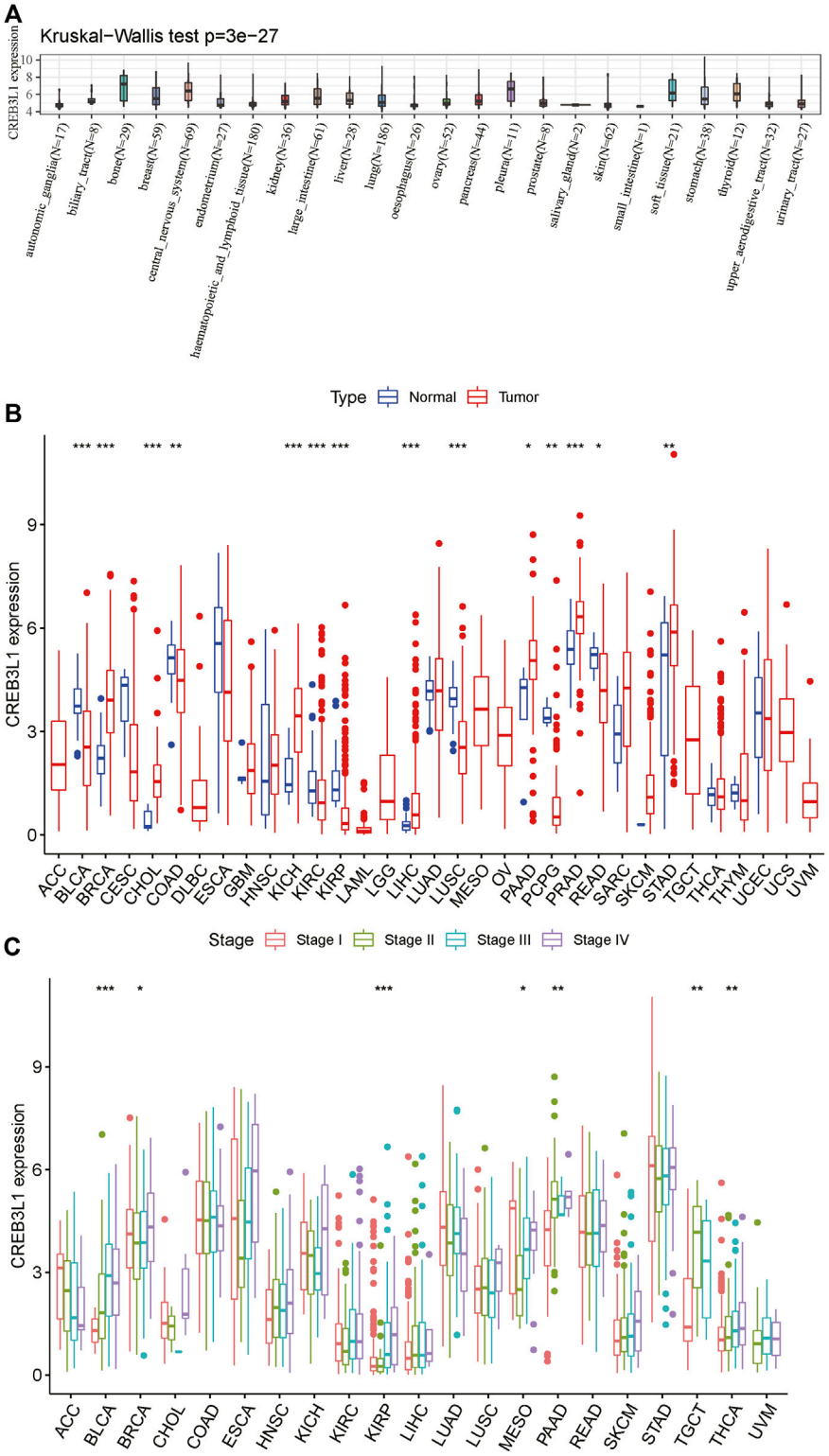


FIGURE 1
The expression pattern of CREB3L1. **(A)** CREB3L1 expression in 24 cancer cell lines from CCLE database. **(B)** Comparison of CREB3L1 expression between cancer and matched normal samples from TCGA database. **(C)** CREB3L1 expression in patients with different WHO stages in pan-cancer from TCGA database. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

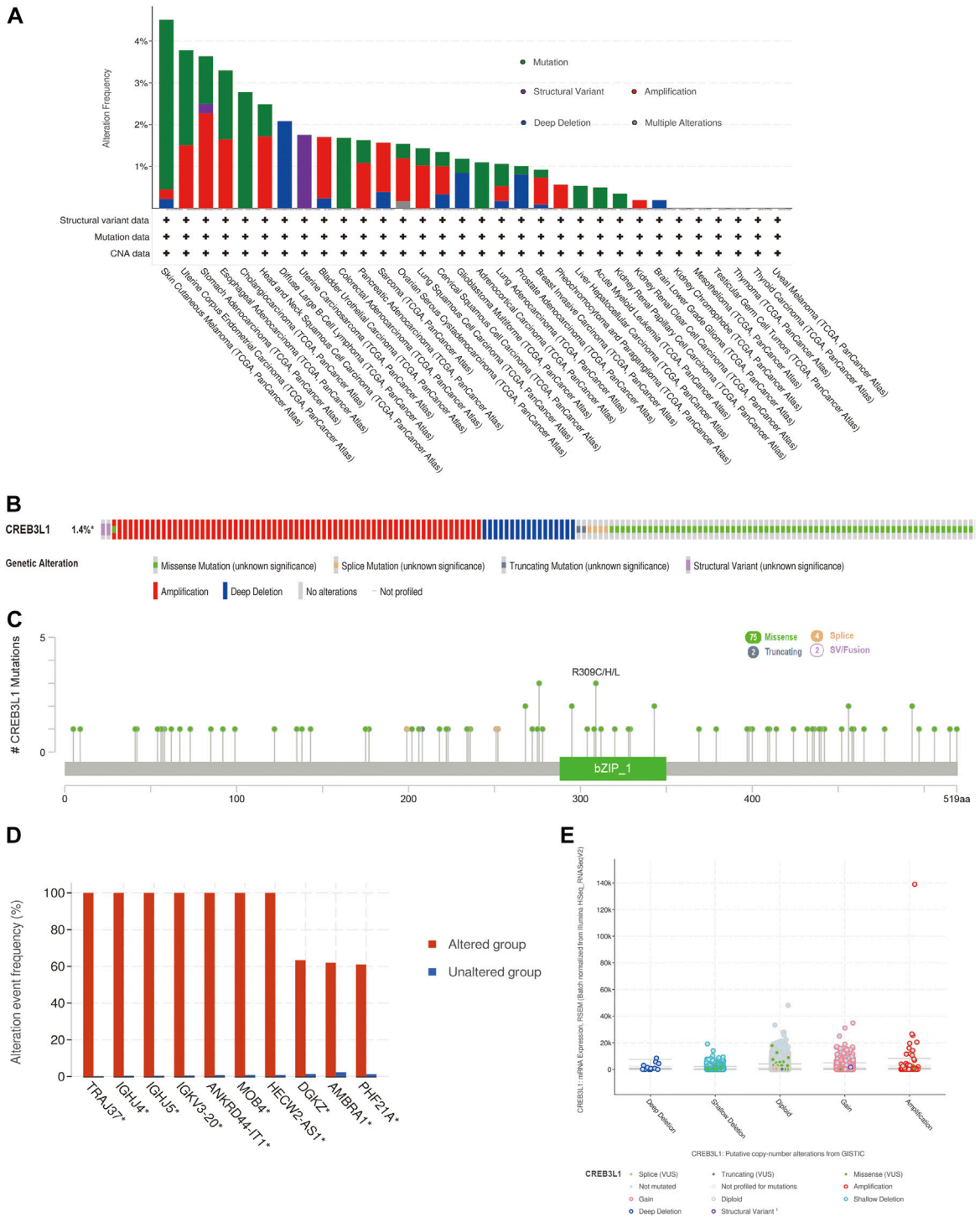


FIGURE 2
The genetic alteration characteristics of CREB3L1 in pan-cancer. **(A)** The alteration frequency of CREB3L1 with different types of mutations in various cancer types. **(B)** Summary of different genetic alteration types of CREB3L1 (Different colors refers to different types of CREB3L1 genetic alterations). **(C)** The mutation types, sites, and sample numbers of the CREB3L1 genetic alterations. **(D)** Co-occurrence of genetic mutations in tumors with CREB3L1 alterations. **(E)** The correlated alteration types and putative copy-number of CREB3L1 in pan-cancer. * Altered genes.

GEO (https://www.ncbi.nlm.nih.gov/geo/) (Charoentong et al., 2017). The procedure was conducted and the results were visualized by utilizing R-package "ggpubr" and "ggplot2".

Relationship between CREB3L1 expression and drug sensitivity

To elucidate the correlation between CREB3LI expression with drug sensitivity in pan-cancer, NCI-60 compound activity data with RNA-seq expression profiles were downloaded from the CallMiner™ online database (https://discover.nci.nih.gov/cellminer/home.do). The correlation between CREB3LI expression and the sensitivity of drugs approved by FDA was analyzed by utilizing R packages "impute", "limma", "ggplot2", and "ggpubr".

Statistical analysis

Student's *t*-test was used to analyze the differences on gene expression data between groups. For survival analyses, the Kaplan-Meier curve and Cox analysis were performed. Statistical analyses and visualization were performed on R software (Version 4.1.2, https://www.Rproject.org). $p < 0.05$ was indicated the statistical significance ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$).

Results

CREB3L1 expression analysis in pan-cancer

To start with, we explored CREB3L1 expression in diverse cancer cell lines in the CCLE database. The results showed that CREB3L1 expression varied from different cancer cell lines, and the expression levels of CREB3L1 were highest in bone, pleura and central nervous system cancer cell lines, while CREB3L1 was expressed lowly in small intestine, oesophagus and salivary gland cancer cell lines compared with other cancer cell lines (Figure 1A). Next, we compared CREB3L1 expression between normal and tumor samples in TCGA pan-cancer database, and we detected significant differences in 14 cancer types among 33 types of cancer excluding those without normal tissue data (Wilcoxon test $p < 0.05$) (Figure 1B). The results revealed that significantly higher CREB3L1 expression was observed in 7 cancer types than their corresponding adjacent non-cancerous tissues, including breast invasive carcinoma (BRCA), cholangiocarcinoma (CHOL), kidney chromophobe (KICH), liver hepatocellular carcinoma (LIHC), pancreatic adenocarcinoma (PAAD), prostate adenocarcinoma (PRAD) and stomach adenocarcinoma (STAD). In comparison, CREB3L1 expression levels were abnormally downregulated in tumor samples in 7 cancer types, including bladder urothelial carcinoma (BLCA), colon adenocarcinoma (COAD), kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), lung squamous cell carcinoma (LUSC), pheochromocytoma and paraganglioma (PCPG) and rectum adenocarcinoma (READ). To further illustrate the clinical significance of CREB3L1 in pan-cancer, we also assessed the CREB3L1 expression levels in different cancer clinical stages, and the results suggested that CREB3L1 expression was significantly related to the clinical stage in 7 cancer types, in which CREB3L1 expression levels were relatively higher in later clinical stages, including BLCA, BRCA, KIRP, mesothelioma (MESO), PAAD, testicular germ cell tumor (TGCT), thyroid carcinoma (THCA), while in other 14 caner types, CREB3L1 expression was not significantly modified during clinical stage progression (Figure 1C). Interestingly, CREB3L1 expression levels were higher in patients with stage I MESO and decreased in patients with stage II MESO. Then, the expression of CREB3L1 was gradually increased in stage III/IV MESO patients. Despite the increasing trend, CREB3L1 expression in stage IV MESO patients was still lower than those in stage I MESO patients. In addition, we also analyzed physiologic CREB3L1 gene expression levels in normal tissues based on the GTEx database. The results showed that CREB3L1 was highly expressed in the salivary gland, prostate and stomach tissues, while was lowly expressed in the blood, liver and muscle tissues in comparison with other normal tissues (Supplementary Figure S1A). We also evaluated CREB3L1 expression in pan-cancer, which was ranked from high to low (Supplementary Figure S1B). We found that CREB3L1 was expressed in 33 cancer types, in detail, CREB3L1 was expressed highest in PRAD while the lowest was in acute myeloid leukemia (LAML). Then, we also explored the differential CREB3L1 mRNA expression between cancer samples in TCGA and normal tissue samples in GTEx. A significant difference of CREB3L1 expression was detected in 25 cancer types (Supplementary Figure S1C). The results were basically consistent with the differential expression analysis in TCGA pan-cancer cohorts. In addition to cancer types mentioned above, the results showed that CREB3L1 was also abnormally upregulated in glioblastoma multiforme (GBM), lower grade glioma (LGG), lung adenocarcinoma (LUAD), ovarian serous cystadenocarcinoma (OV), TGCT, and uterine carcinosarcoma (UCS), whereas was abnormally downregulated in adrenocortical carcinoma (ACC), cervical squamous cell carcinoma (CESC), LAML, skin cutaneous melanoma (SKCM), and THCA. But in COAD, CREB3L1 expression was significantly upregulated in tumor samples compare with corresponding normal samples in GTEx, which was contrary to the result presented in Figure 1B.

Genetic alteration analysis of CREB3L1 in different cancers

We further explored the characteristics of CREB3L1 genetic alternations in different cancers by enrolling the TCGA Pan-Cancer Atlas studies in cBioPortal database. The results found

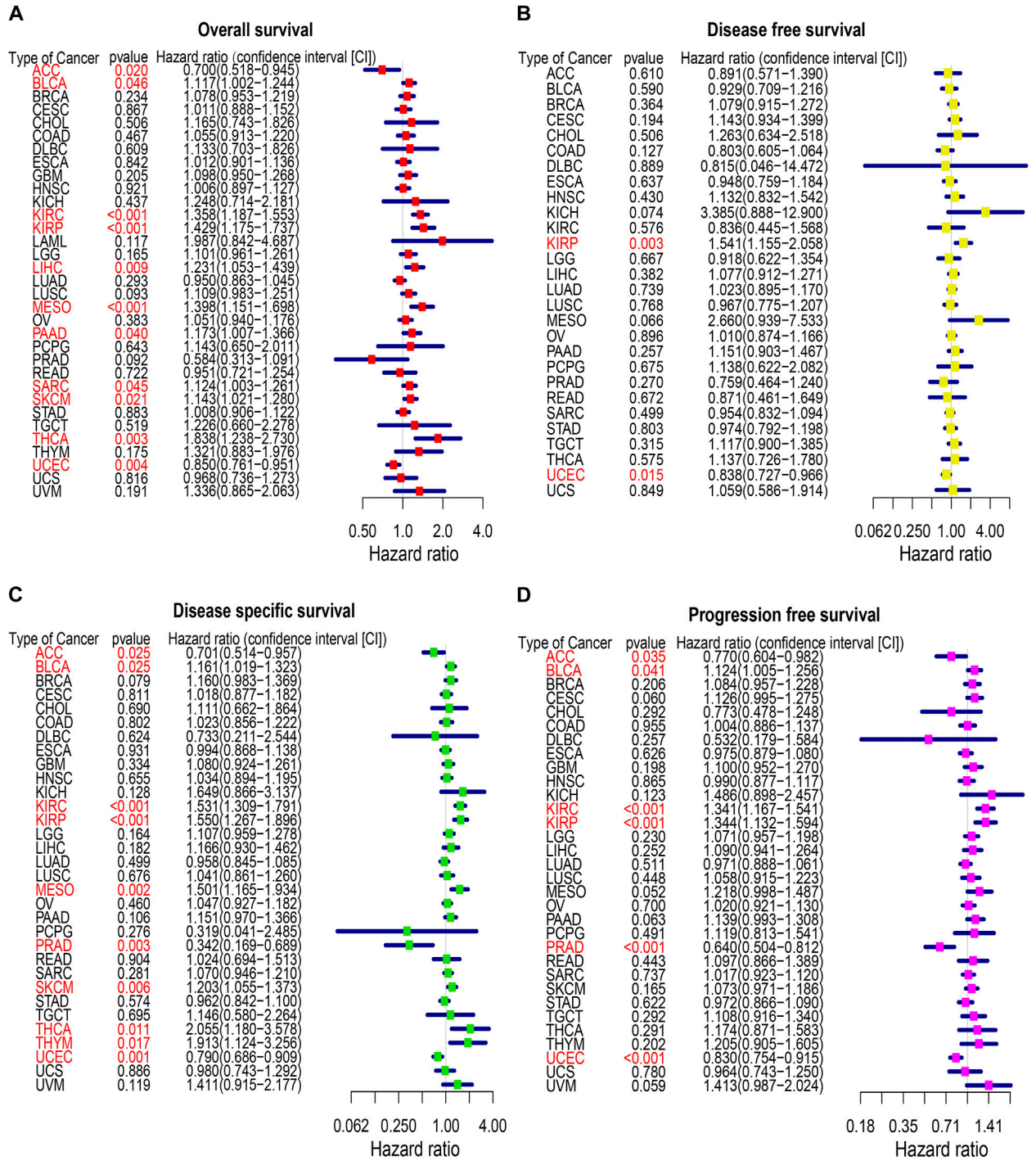


FIGURE 3
The forest map of univariate Cox regression analysis of CREB3L1. **(A)** Forest map shows the results of univariate cox regression analysis of CREB3L1 for OS in TCGA pan-cancer. **(B)** Forest map shows the results of univariate cox regression analysis of CREB3L1 for DFS in TCGA pan-cancer. **(C)** Forest map shows the results of univariate cox regression analysis of CREB3L1 for DSS in TCGA pan-cancer. **(D)** Forest map shows the results of univariate cox regression analysis of CREB3L1 for PFS in TCGA pan-cancer. Red items indicate the statistical significance.

that CREB3L1 was altered in 158 patients (1.4%) of 10953 patients. In patients with SKCM, uterine corpus endometrial carcinoma (UCEC), esophageal adenocarcinoma (ESCA), CHOL, and head and neck squamous cell carcinoma (HNSC), the high CREB3L1 gene alteration rate was occurred, in which the alteration frequency was higher than 2% (Figure 2A).

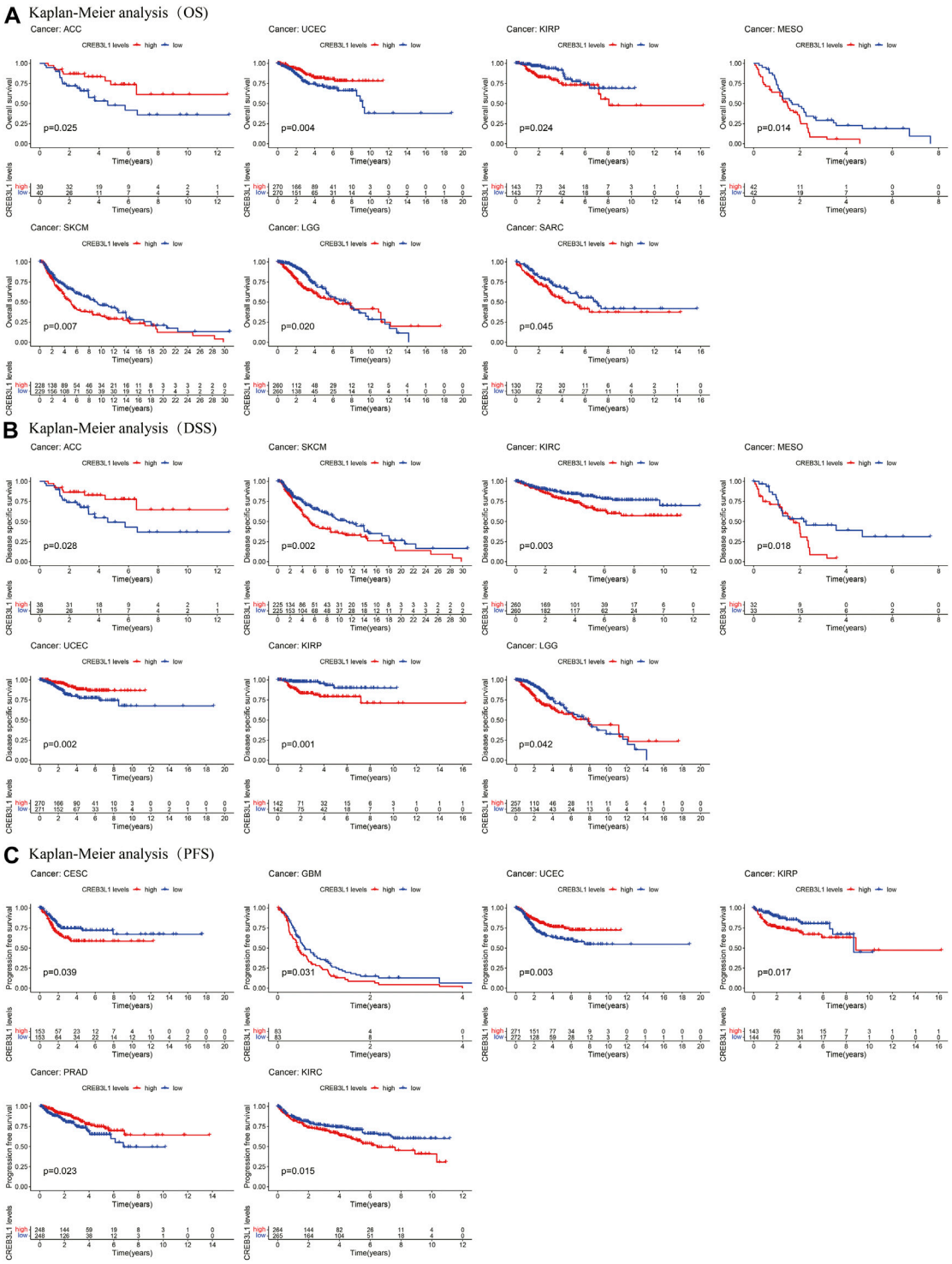


FIGURE 4
Kaplan-Meier survival curves of CREB3L1 in pan-cancer. **(A)** Kaplan–Meier analysis of the correlation between CREB3L1 expression and OS in 7 cancer types. **(B)** Kaplan–Meier analysis of the correlation between CREB3L1 expression and DSS in 7 cancer types. **(C)** Kaplan–Meier analysis of the correlation between CREB3L1 expression and PFS in six cancer types.

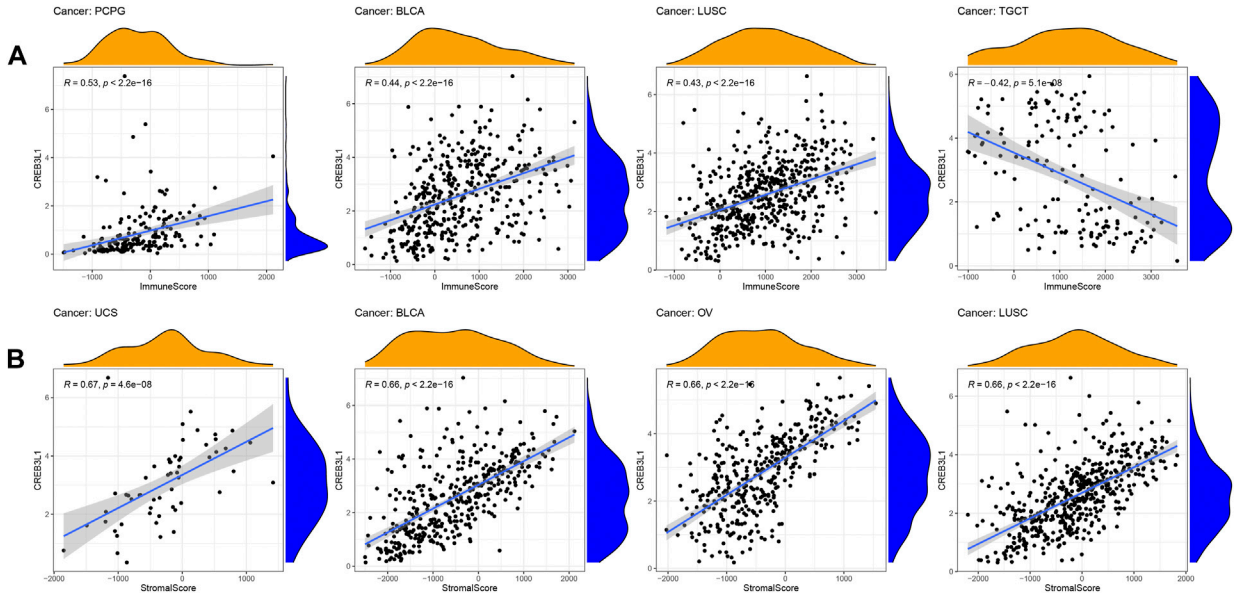


FIGURE 5
The correlation between CREB3L1 and immune and stromal scores in pan-cancer. **(A)** The association between CREB3L1 expression and immune score in PCPG, BLCA, LUSC, and TGCT. **(B)** The association between CREB3L1 expression and stromal score in top 4 cancer types, including UCS, BLCA, OV, and LUSC.

The primary genetic alteration types of CREB3L1 were amplification, miss mutation, and deep deletion (Figure 2B). As shown in Figure 2C, the mutation sites, types, and sample numbers of the CREB3L1 genetic alterations were presented. In addition, the main type alteration was CREB3L1 missense mutation, while R309C alteration was detected in LUSC, R309H alteration was detected in uterine papillary serous carcinoma and R309L alteration was detected in KIRP. Genomic alternations co-occurrence analysis indicated that alterations of several genes, including TRAJ37, IGHJ4, IGHJ5, IGKV3-20, ANKRD44-IT1, MOB4, HECW2-AS1, DGKZ, AMBRA1, and PHF21A were more commonly occurred in the CREB3L1-altered group (Figure 2D). Besides, the putative copy-number alterations of CREB3L1 from genomic identification of significant targets in cancer (GISTIC) included many types, such as amplification, deep deletion and gain function, resulting in the alternations of gene expression (Figure 2E).

Prognostic significance of CREB3L1

To identify the prognostic significance of CREB3L1 in cancer patients, we further performed a comprehensive prognosis analysis using the survival data of TCGA pan-cancer dataset, including OS, DFS, DSS and PFS. The univariate Cox regression analysis were performed and found CREB3L1 expression levels were correlated with OS in 11 cancers, DFS in two cancers, DSS in 10 cancers, and PFS in six cancers (Figures 3A–D). The results of univariate Cox regression analysis showed that the CREB3L1 expression levels were strongly related to OS in ACC (hazard ratio [HR], 0.700; 95% confidence interval [CI], 0.518-0.945; $p = 0.020$), BLCA (HR, 1.117; 95% CI, 1.002-1.244; $p = 0.046$), KIRC (HR, 1.358; 95% CI, 1.187-1.553; $p < 0.001$), KIRP (HR, 1.429; 95% CI, 1.175-1.737; $p < 0.001$), LIHC (HR, 1.231; 95% CI, 1.053-1.439; $p = 0.009$), MESO (HR, 1.398; 95% CI, 1.151-1.698; $p < 0.001$), PAAD (HR, 1.173; 95% CI, 1.007-1.366; $p = 0.040$), sarcoma (SARC) (HR, 1.124; 95% CI, 1.003-1.261; $p = 0.045$), SKCM (HR, 1.143; 95% CI, 1.021-1.280; $p = 0.021$), THCA (HR, 1.838; 95% CI, 1.238-2.730; $p = 0.003$) and UCEC (HR, 0.850; 95% CI, 0.761-0.951; $p = 0.004$) (Figure 3A). In the results, we found that CREB3L1 was a risk factor in BLCA, KIRC, KIRP, LIHC, SARC, SKCM and THCA, while it was a protective factor in ACC and UCEC. The DFS analysis revealed that CREB3L1 could act as a protective factor for patients with UCEC (HR, 0.838; 95% CI, 0.727-0.966; $p = 0.015$), while a risk factor for patients with KIRP (HR, 1.541; 95% CI, 1.155-2.058; $p = 0.003$) (Figure 3B). The DSS analysis indicated that CREB3L1 serves as a protective factor for patients with ACC (HR, 0.701; 95% CI, 0.514-0.957; $p = 0.025$), PRAD (HR, 0.342; 95% CI, 0.169-0.689; $p = 0.003$) and UCEC (HR, 0.790; 95% CI, 0.686-0.909; $p = 0.001$), and acts as a risk role in patients with BLCA (HR, 1.161; 95% CI, 1.019-1.323; $p = 0.025$), KIRC (HR, 1.531; 95% CI, 1.309-1.791; $p < 0.001$), KIRP (HR, 1.550; 95% CI, 1.267-1.896; $p < 0.001$), MESO (HR, 1.501; 95% CI, 1.165-1.934; $p = 0.002$), SKCM (HR, 1.203; 95% CI, 1.055-1.373; $p = 0.006$), THCA (HR,

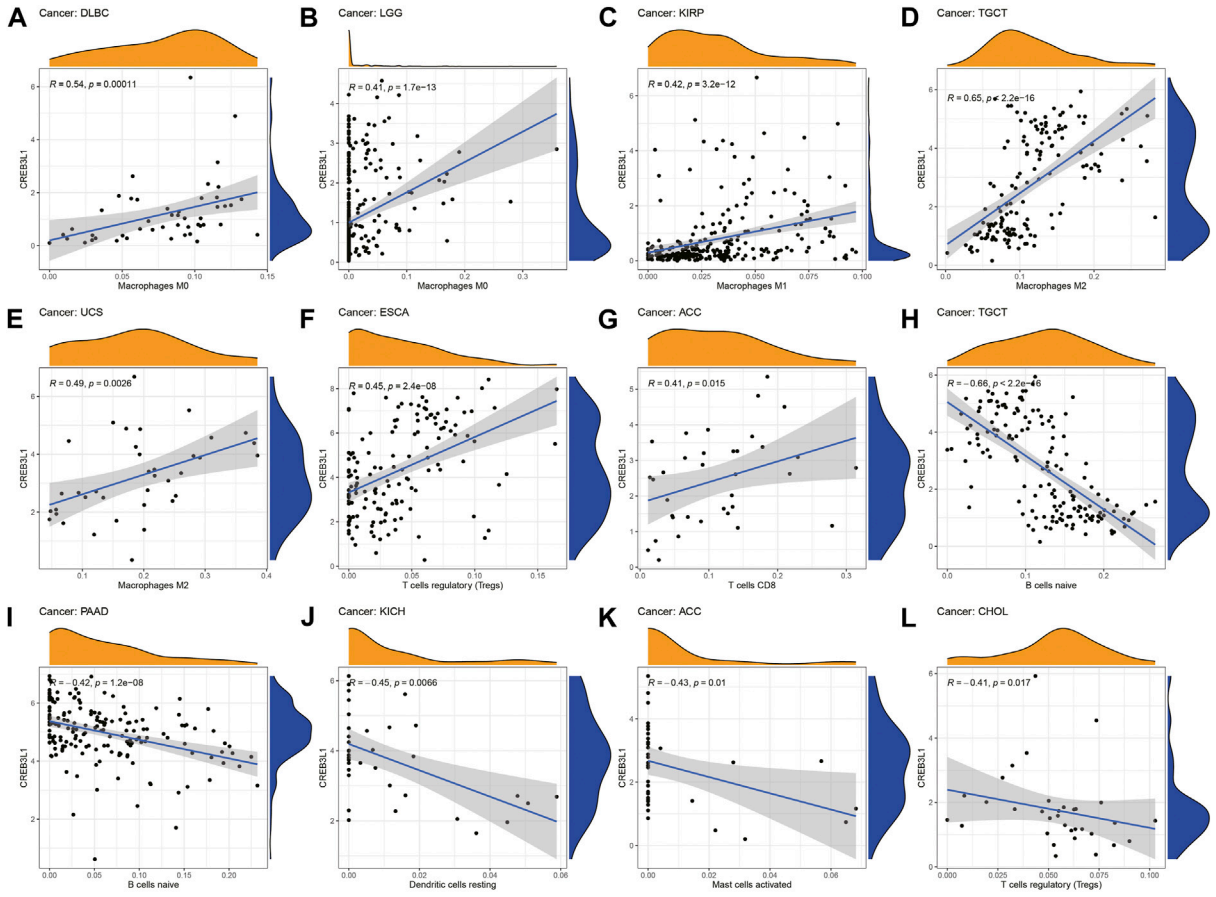


FIGURE 6
The correlation between CREB3L1 expression and the immune cells infiltration in pan-cancer. **(A–L)** CREB3L1 expression was associated with macrophages, regulatory T cells, CD8 + T cells, B cells naïve, dendritic cells and mast cells in different cancer types.

2.055; 95% CI, 1.180-3.578; $p = 0.011$) and thymoma (THYM) (HR, 1.913; 95% CI, 1.124-3.256; $p = 0.017$) (Figure 3C). In addition, the PFS analysis suggested that CREB3L1 plays a protective role in patients with ACC (HR, 0.770; 95% CI, 0.604-0.982; $p = 0.035$), PRAD (HR, 0.640; 95% CI, 0.504-0.812; $p < 0.001$) and UCEC (HR, 0.830; 95% CI, 0.754-0.915; $p < 0.001$) and a risk role in patients with BLCA (HR, 1.124; 95% CI, 1.005-1.256; $p = 0.041$), KIRC (HR, 1.341; 95% CI, 1.167-1.541; $p < 0.001$) and KIRP (HR, 1.344; 95% CI, 1.132-1.594; $p < 0.001$) (Figure 3D).

The Kaplan-Meier analysis indicated that the CREB3L1 expression levels were correlated with OS in 7 cancers, DSS in 7 cancers, and PFS in six cancers (Figures 4A–C). Our results of Kaplan-Meier OS analysis demonstrated that among patients with ACC ($p = 0.025$) and UCEC ($p = 0.004$), longer survival time was presented in those who with high levels of CREB3L1 expression, however, in patients with KIRP ($p = 0.024$), MESO ($p = 0.014$), SKCM ($p = 0.007$), LGG ($p = 0.020$) and SARC ($p = 0.045$), high CREB3L1 expression levels were associated with poor OS (Figure 4A). The analysis of DSS revealed correlations between high CREB3L1 expression levels and good prognosis in patients with ACC ($p = 0.028$), UCEC ($p = 0.002$) and LGG ($p = 0.042$), which was opposite in patients with SKCM ($p = 0.002$), KIRC ($p = 0.003$), MESO ($p = 0.018$) and KIRP ($p = 0.001$) (Figure 4B). The Kaplan-Meier curves of PFS indicated that in UCEC ($p = 0.003$) and PRAD ($p = 0.023$), patients with low CREB3L1 expression levels had a shorter survival time, while in CESC ($p = 0.039$), GBM ($p = 0.031$), KIRP ($p = 0.017$) and KIRC ($p = 0.015$), there were better prognoses in patients with low CREB3L1 expression levels (Figure 4C). These results indicated that CREB3L1 expression levels were significantly associated with cancer prognosis, and might exert different effects on prognoses in different cancer types.

Correlation of CREB3L1 expression with the tumor immune microenvironment

The tumor immune microenvironment plays important roles in tumor initiation and progression. Therefore, exploring the relationship between CREB3L1 expression and the tumor immune microenvironment in pan-cancer is vital. To calculate

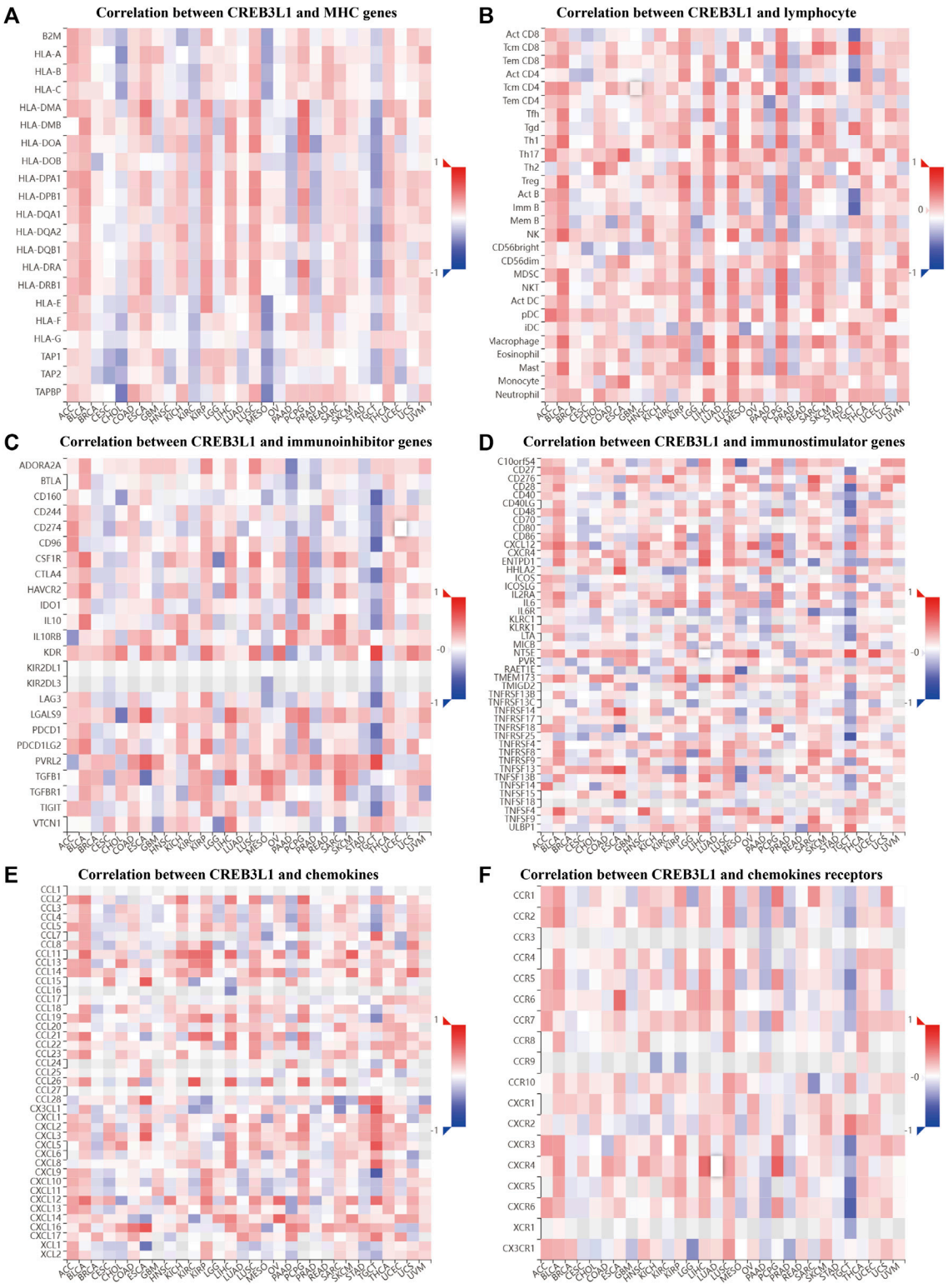


FIGURE 7
The correlation between CREB3L1 expression and immune-related biomarkers in TISIDB database. The co-expression heatmaps show the association between CREB3L1 expression and **(A)** MHC genes, **(B)** lymphocyte, **(C)** immunoinhibitor genes, **(D)** immunostimulator genes, **(E)** chemokines, and **(F)** chemokines receptors in different cancers.

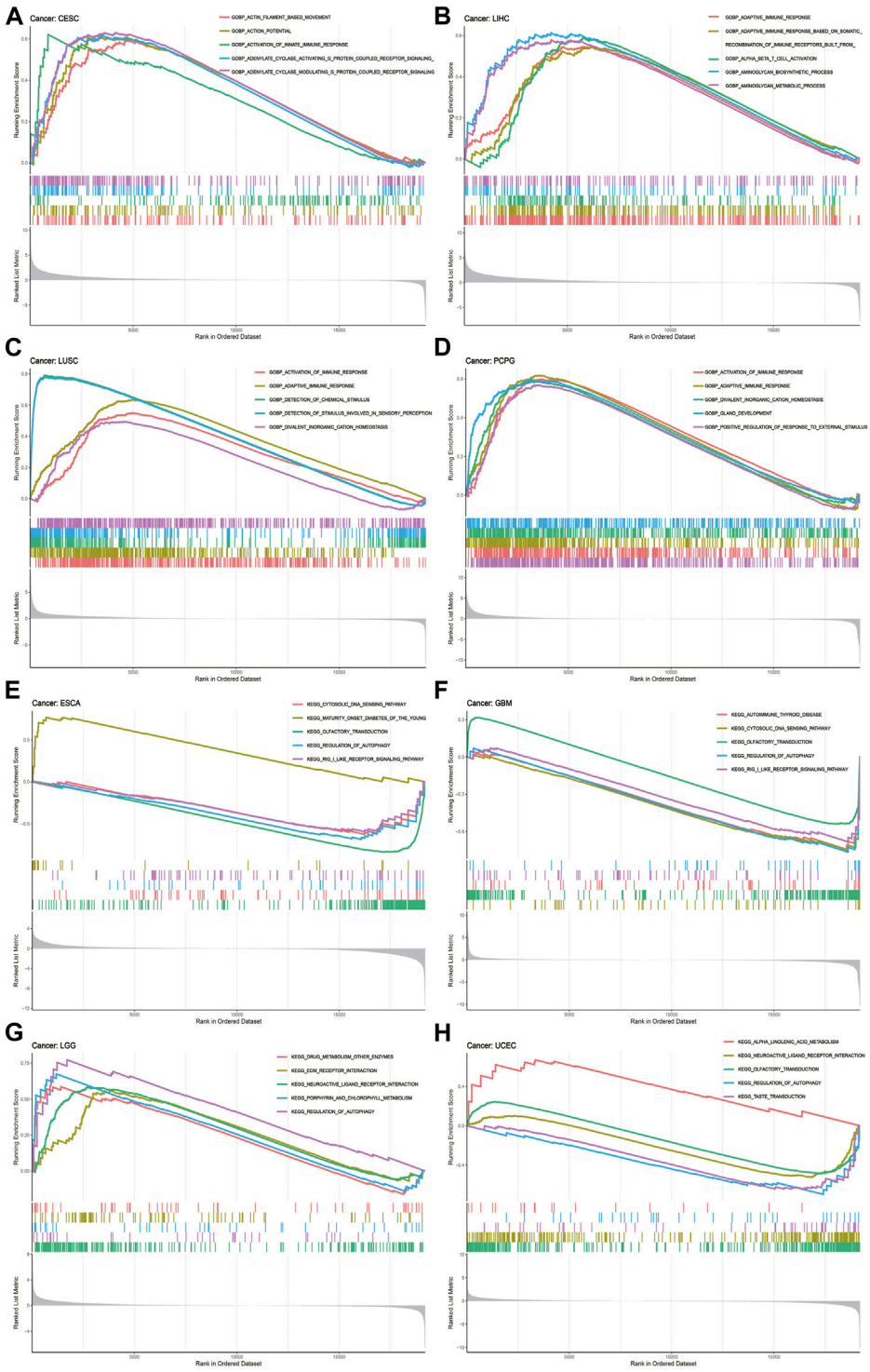


FIGURE 8
GSEA of CREB3L1. **(A–D)** GO functional annotation of CREB3L1 shows the activation of innate immune response in CESC, and adaptive immune response in LIHC, LUSC, and PCPG. **(E–H)** KEGG pathway analysis of CREB3L1 shows that regulation of autophagy was the most common signaling pathway of CREB3L1 participating in TCGA pan-cancer, including UCEC, LGG, GBM, and ESCA. Curves of different colors show different functions or pathways regulated in different cancers. Peaks on the upward curve indicate positive regulation and peaks on the downward curve indicate negative regulation.

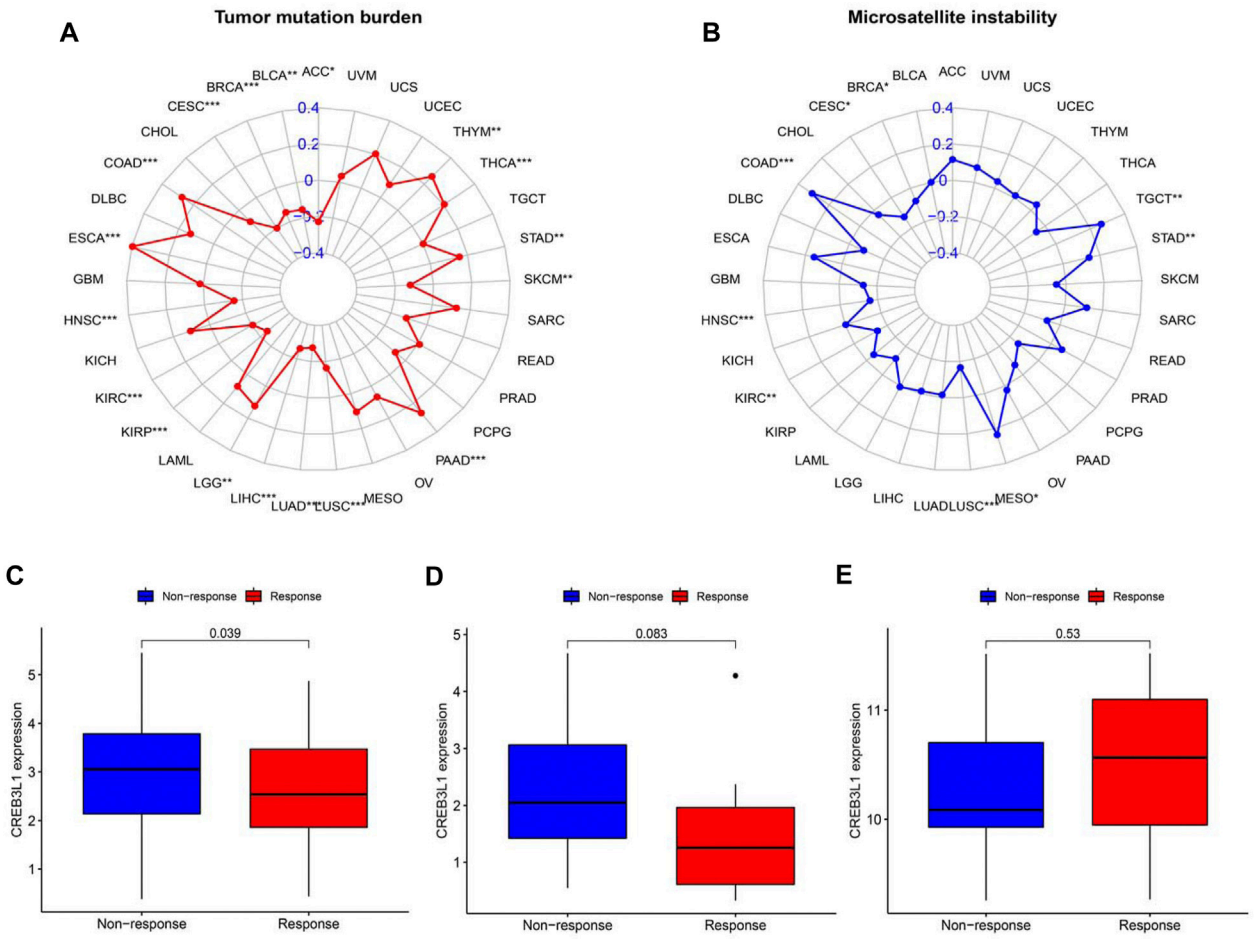


FIGURE 9
The correlation between CREB3L1 expression and TMB levels, MSI event and the immunotherapeutic efficacy. **(A)** Radar map of the relationship between CREB3L1 expression and TMB levels. **(B)** Radar map of the relationship between CREB3L1 expression and MSI event. **(C–E)** The relationship between CREB3L1 expression and the immunotherapeutic efficacy in IMvigor210 cohort **(C)**, GSE78220 **(D)**, and GSE67501 **(E)**. *$p < 0.05$, **$p < 0.01$; ***$p < 0.001$.

the stromal and immune scores in pan-cancer, and explore the association between these two scores and CREB3L1 expression levels, the ESTIMATE algorithm was employed for research. The results showed that CREB3L1 expression levels were positively correlated with immune scores in BLCA, LUSC, and PCPG, whereas were negatively related to immune scores in TGCT (Figure 5A). In addition, there was positive relationship between CREB3L1 expression and stromal scores in multiple cancers, particularly in UCS, BLCA, OV, and LUSC (Figure 5B). The relationship between CREB3L1 expression levels and stromal scores in other cancer types was shown in Supplementary Figure S2.

We further explored the relationship between CREB3L1 expression levels and the infiltration levels of 22 relevant immune cells. The results indicated that CREB3L1 expression levels were strongly related to tumor immune cell infiltration levels in several cancer types. We detected that CREB3L1 expression was positively associated with the infiltration levels of macrophages in several cancer types, including macrophage M0 in DLBC (Figure 6A) and LGG (Figure 6B), macrophage M1 in KIRP (Figure 6C), macrophage M2 in TGCT (Figure 6D) and UCS (Figure 6E), and regulatory T cells in ESCA (Figure 6F), $CD8^+$ T cells in ACC (Figure 6G), while was negatively associated with the abundance of B cells naïve in TGCT (Figure 6H) and PAAD (Figure 6I), dendritic cells resting in KICH (Figure 6J), mast cells activated in ACC (Figure 6K) and regulatory T cells in CHOL (Figure 6L). Furthermore, we also used the XCELL, TIMER, QUANTISEQ, EPIC, IPS and MCPcounter databases to analyze the association between CREB3L1 expression and the infiltrating levels of immune cells in pan-cancer (Supplementary Figure S3).

To further explore the potential function of CREB3L1 in regulating the tumor immune microenvironment in human cancers, we examined the associations between CREB3L1 expression and the

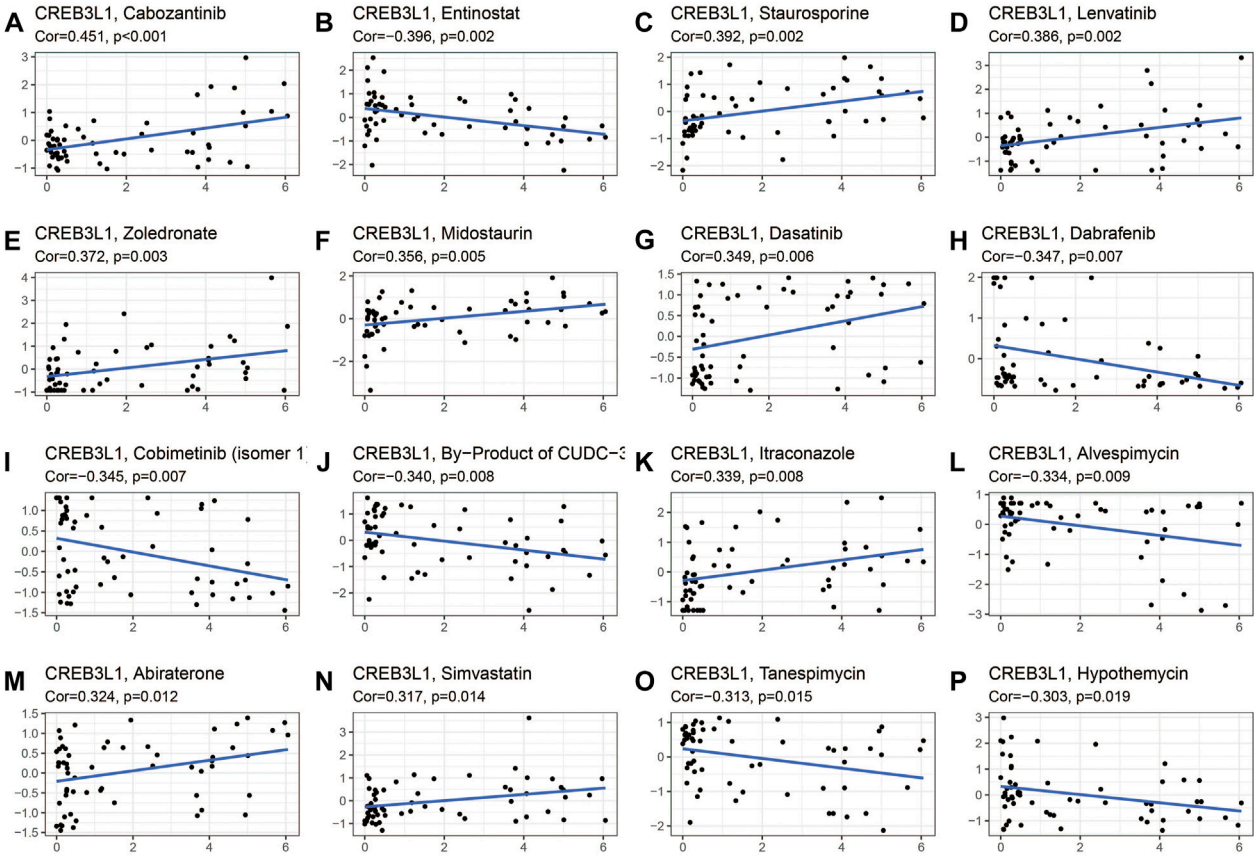


FIGURE 10
The correlation between CREB3L1 expression and drug sensitivity. The CREB3L1 was linked to the sensitivity of **(A)** cabozantinib, **(B)** entinostat, **(C)** staurosporine, **(D)** lenvatinib, **(E)** zoledronate, **(F)** midostaurin, **(G)** dasatinib, **(H)** dabrafenib, **(I)** cobimetinib (isomer1), **(J)** By-Product of CUDC-305, **(K)** itraconazole, **(L)** alvespimycin, **(M)** abiraterone, **(N)** simvastatin, **(O)** tanespimycin, and **(P)** hypothemycin.

expression of major histocompatibility complex (MHC) genes (Figure 7A), lymphocyte (Figure 7B), immunoinhibitor genes (Figure 7C), immunostimulator genes (Figure 7D), chemokine (Figure 7E) and chemokine receptors (Figure 7F) by gene co-expression analyses in TISIDB. Our study revealed that CREB3L1 expression was distinctly correlated with multiple immuneinhibitors genes, such as KDR, PVRL2, CD160, and CD96, and immunostimulators genes, such as NT5E, TNFRSF14, CD28, and CXCR4, and chemokines including CX3CL1, CXCL9, CCL15, and CXCL5 in multiple cancer types, suggesting the interactions between CREB3L1 and immune-related biomarkers expression in pan-cancer.

Gene set enrichment analysis

To deeply uncover the potential biological functions and signaling pathways of CREB3L1 in pan-cancer, GESA, including GO functional annotation and KEGG pathway analysis was conducted. The results suggested that CREB3L1 might exert various biological functions, especially immune modulatory functions in cancers, including the activation of innate immune response in CESC and adaptive immune response in LIHC, LUSC, PCPG (Figures 8A–D). Besides, CREB3L1 is also involved in immune response regulating signal pathway in READ, regulation of inflammatory response in UCS, interleukin 1 (IL-1) production in KICH, and DNA binding transcription activator activity in MESO (Supplementary Figure S4). In addition, KEGG analysis revealed that regulation of autophagy was the most common signaling pathway of CREB3L1 participating in multiple cancer types, including UCEC, LGG, GBM and ESCA (Figures 8E–H). Besides, chemokine signal pathway (in UCS, THCA, PCPG, LAML, and KIRP), pathway in cancer (in THYM, PCPG, and OV) and drug metabolism (in SKCM, LGG, and ACC) were also potential signaling pathways modulated by CREB3L1 in pan-cancer (Supplementary Figure S5). These results indicated that CREB3L1 might be critically involved in regulating the tumor immune microenvironment, and autophagy in diverse cancers.

Correlation of CREB3L1 expression with TMB and MSI

TMB and MSI have gradually emerged as biomarkers related to the immunotherapy response. Hence, we further investigated the relationship of CREB3L1 expression levels with TMB and MSI in pan-cancer, thereby speculating the potential role of CREB3L1 as a promising predictor for the immunotherapeutic efficacy in a specific cancer type. A remarkable correlation between CREB3L1 and TMB was identified in 18 cancers, a positive correlation was found in COAD, ESCA, LGG, PAAD, STAD, THCA, and THYM, while a negative correlation was suggested in ACC, BLCA, BRCA, CESC, HNSC, KIRC, KIRP, LIHC, LUAD, LUSC, and SKCM (Figure 9A). Similarly, we also found a significant association between CREB3L1 expression and MSI in nine cancers, which indicated that overexpression of CREB3L1 was positively linked to MSI in COAD, MESO, STAD, and TGCT, and was inversely correlated with MSI in BRCA, CESC, HNSC, KIRC, and LUSC (Figure 9B).

Correlation of CREB3L1 expression with immunotherapeutic efficacy

The potential role of CREB3L1 in predicting the immunotherapeutic efficacy in cancer patients was further investigated. We extracted three GEO datasets, including GSE78220, GSE67501, and IMvigor210, to compare the differences on CREB3L1 expression levels between immunotherapy-response and immunotherapy-nonresponse patients. The results showed that CREB3L1 expression was significantly higher in non-response group in IMvigor210 cohort (p = 0.039) (Figure 9C), and there was a statistical tendency towards high CREB3L1 expression in non-response group in GSE78220 cohort (p = 0.083) (Figure 9D), whereas no close relationship between CREB3L1 expression and the immunotherapeutic efficacy was detected in GSE67501 cohort (p = 0.53) (Figure 9E). The results suggested that CREB3L1 might effectively predict the efficacy of PD-1/PD-LI inhibitors treatment in some specific cancer types.

Correlation of CREB3L1 expression with drug sensitivity

The potential relationship between CREB3L1 expression and drug sensitivity in human cancers was also investigated (Figure 10). The results revealed that CREB3L1 was positively related to the sensitivity of several drugs, including cabozantinib (Figure 10A), staurosporine (Figure 10C), lenvatinib (Figure 10D), zoledronate (Figure 10E), midostaurin (Figure 10F), dasatinib (Figure 10G), itraconazole (Figure 10K), abiraterone (Figure 10M) and simvastatin (Figure 10N), while was negatively related to the sensitivity of entinostat (Figure 10B), dabrafenib (Figure 10H), cobimetinib (isomer1) (Figure 10I), By-Product of CUDC-305 (Figure 10J), alvespimycin (Figure 10L), tanespimycin (Figure 10O), and hypothemycin (Figure 10P). These results indicated that CREB3L1 could function as a promising predictor for the sensitivity of several anti-cancer agents, such as cabozantinib, lenvatinib, zoledronate, dasatinib, and dabrafenib, which have been commonly applied in clinical cancer management.

Discussion

The expression pattern of CREB3L1 has been evaluated in multiple human tissues, and the biological functions of CREB3L1, such as protein secretion, cell differentiation and tissue development have been illustrated in multiple studies (Keller et al., 2018; Sampieri et al., 2019; Kamikawa et al., 2021). Recently, several studies have illustrated CREB3L1 expression and functions in different cell lineages development through conducting single-cell RNA sequencing (scRNA-seq) analysis. It has been found that FB-1 cells, a major type of fibroblast in hypertrophic scar samples have higher transcriptional activities of CREB3L1 than those in normal samples, and hypertrophic scar samples have a higher proportion of the FB-1 cells expressing CREB3L1 through investigating online scRNA-seq data, suggesting that CREB3LI may be critically involved in hypertrophic scar formation (Zhang et al., 2021a). In addition, a recent study conducting scRNA-seq analysis has revealed the critical biomarkers and process of the differentiation of fibroblasts into myofibroblasts in systemic sclerosis, and CREB3L1 may be a critical upstream transcription factor in promoting systemic sclerosis myofibroblast differentiation (Tabib et al., 2021). Notably, it has been reported that CREB3L1 expression is frequently altered in many cancers, and complex functions and mechanisms of CREB3L1 in cancer development have been clarified in several studies (Feng et al., 2017; Liu et al., 2018; Bissonnette et al., 2021; Morishita et al., 2021). Xu et al. have recently explored the cellular composition of thyroid cancer and found that CREB3L1 was upregulated in anaplastic thyroid cancer (ATC)-derived thyroid cancer cells by scRNA-seq. Mechanistically, CREB3L1 overexpression can facilitate the dedifferentiation of papillary thyroid cancer cells into ATC cells, and promote thyroid cancer progression by regulating EMT process and the mTOR signaling pathway (Luo et al., 2021). Through scRNA-seq analysis of intrahepatic cholangiocarcinoma, S100P and SPP1 are regarded as two biomarkers for two intrahepatic cholangiocarcinoma molecular subtypes. CREB3L1, which is upregulated in S100P + SPP1− intrahepatic cholangiocarcinoma tumor cells, can facilitate the invasion of S100P + SPP1− perihilar large duct type tumor cells by directly targeting S100P (Song et al., 2022).

However, the investigation of CREB3L1 in human malignancies is still of absence to date, and the results of present studies are controversial. Besides, no pan-cancer analysis of CREB3L1 has been published to date. Thus, we conducted this first-time and comprehensive investigation to illustrate the roles of CREB3L1 across human cancers.

Firstly, we evaluated CREB3L1 expression pattern and its clinical value in different cancers, and the results suggested that compared with normal tissues, CREB3L1 expression levels were abnormally upregulated in several cancer types, such as BRCA, CHOL, and STAD, whereas was significantly decreased in some cancer types, such as BLCA, COAD, and LUSC. In addition, our results showed that high CREB3L1 expression was significantly associated with advanced clinical stage in several cancer types. Interestingly, our findings showed that CREB3L1 expression was abnormally decreased in BLCA, and CREB3L1 expression was increased during tumor stage progression in patients with BLCA, which seems to be contradictory. A previous study has reported that CREB3L1 was significantly downregulated in BLCA, which was consistent with our findings. CREB3L1 silenced by CREB3L1 promoter hypermethylation resulted in a more aggressive phenotype in BLCA, and no significant correlation between CREB3L1 and tumor stage was detected in this study (Rose et al., 2014). Future studies based on large-scale samples are required to illustrate the association between CREB3L1 expression and clinical stage progression in BLCA, and CREB3L1 alternations and epigenetic modifications during BLCA progression. Besides, the Kaplan–Meier survival showed that high CREB3L1 expression was significantly correlated with poor OS and DSS in KIRP, MESO, and SKCM, while was related to better OS and DSS in ACC, suggesting the prognostic significance of CREB3L1 in different human cancers. However, using OS as the endpoint of survival may decrease the utility of clinical trials, and non-cancer-induced death cannot effectively represent cancer biology. Additionally, longer follow-up time are required for OS or DSS analysis. Hence, in many clinical studies, using DFS or PFS can assess the influence of biomolecules on cancer patients more effectively. The Kaplan–Meier PFS curves showed that patients with high CREB3L1 expression had longer PFS time in ACC, PRAD and UCEC, while patients who expressed CREB3L1 highly had shorter PFS time in BLCA, KIRC and KIRP. Furthermore, the univariate Cox regression analysis results showed that CREB3L1 could serve as a protective factor in several cancer types, such as BLCA, KIRC and KIRP, while a risk factor in ACC and PRAD. These findings suggest that CREB3L1 may play different roles in cancer progression and clinical prognoses in distinct cancer types, which needs to be verified in future studies. Besides, we analyzed the characteristics of CREB3L1 mutations in pan-cancer. The results revealed that the total alteration rate of CREB3L1 was 1.4% in pan-cancer with various alternation types, such as amplification, miss mutation, and deep deletion, and patients with SKCM has the highest CREB3L1 mutation rate. To date, the biological functions of CREB3L1 mutations are mostly investigated in modulating osteogenesis process, and only few studies have reported that EWSR1-CREB3L1 fusion mutation in sarcomas (Keller et al., 2018; Bissonnette et al., 2021). The functions and mechanisms of CREB3L1 mutations are largely unknown in human neoplasms, and our investigation shed highlights on the potential roles of CREB3L1 mutations in cancer initiation and progression.

The tumor immune microenvironment fundamentally participates in cancer progression, and shows an important association with cancer patients' survival and the therapeutic efficacy (Pitt et al., 2016). Critical components and novel targets in the tumor immune microenvironment have been uncovered that can help improve the development and application of promising therapies in cancer management, notably immunotherapies (Pitt et al., 2016; Fu et al., 2021). To uncover the potential role of CREB3L1 in modulating the tumor immune microenvironment, we firstly confirmed that CREB3L1 expression levels were positively related to immune scores and stroma scores in multiple human cancers by utilizing ESTIMATE. The crosstalk between cancer cells and immune cells has been found to contribute to the microenvironment which facilitates tumor growth and metastasis, and immune cells presenting in the tumor microenvironment play a determinative role in cancer survival and development (Hinshaw and Shevde, 2019; Marzagalli et al., 2019; Dao et al., 2022). Hence, we further explored the correlation between CREB3L1 expression levels and the abundance of infiltrating immune cells in pan-cancer. A strongly positive correlation was found between CREB3L1 expression and T regulatory cells in ESCA, and macrophages, including M0 macrophages in DLBC and LGG, M1 macrophages in KIRP, and M2 macrophages in TGCT, while a negative relationship between CREB3L1 expression and B naïve cells was detected in PAAD and TGCT, indicating that group with high CREB3L1 expression had a higher infiltrating level of immune cells, which is in line with that high CREB3L1 expression levels being connected with higher immune scores in several cancer types. Tumor-infiltrating naïve B cells have been identified as a key component of adaptive immunity with various functions, and the abundance of naïve B cells is strongly correlated with cancer patients' prognosis (Chen et al., 2020; Downs-Canner et al., 2022). For instance, single-cell transcriptome analysis has found that the infiltrating levels of naïve-like B cells are lower in advanced NSCLC, which is closely related to worse prognosis of NSCLC patients (Chen et al., 2020). Previous studies have indicated that PAAD patients with decreasing infiltrating naïve B cell levels have worse prognose, which is in line with our results (Zhang et al., 2021b; Rajamanickam et al., 2021). These findings indicate that CREB3L1 may have regulatory effects on tumor-infiltrating naïve B cells and macrophage polarization in different cancer types, thereby influencing the cancer survival and progression. Moreover, the interaction between CREB3L1 and immune-related biomarkers was also investigated. Correlations were observed between CREB3L1 expression and several immune

biomarkers such as CD28, CXCR4 and KDR in several cancers. Previous studies have found that CD28 co-stimulation is fundamental for $CD8^+$ T cell rescue, and PD-1 can inhibit T cell function primarily by suppressing CD28. Moreover, CD28 co-stimulatory pathway plays key roles in modulating the tumor immune microenvironment and patients' responses to anti–PD-L1/PD-1 therapy (Hui et al., 2017; Kamphorst et al., 2017; Duraiswamy et al., 2021). CXCR4 has been found to exert activity in modulating immune cells infiltration, such as the inhibition of cytotoxic T lymphocytes, and increase of T regulatory cells, and blocking CXCR4 can effectively improve the immunotherapeutic efficacy (Chen et al., 2019; Daniel et al., 2020). These findings suggest that CREB3L1 may regulate these crucial immune-related biomarkers to remodel the tumor immune microenvironment and influence responses to immunotherapies of cancer patients. Overall, the biological functions of CREB3L1 in modulating the tumor immune microenvironment remains a research gap, which are worth exploring in future studies.

Furthermore, the results of GSEA analysis showed that CREB3L1 was obviously involved in modulating the immune response, immune relevant pathways and chemokine release in pan-cancer, further confirming the potential roles of CREB3L1 in the tumor immune microenvironment. Besides, regulation of autophagy was identified as the most common signaling pathway modulated by CREB3L1 in pan-cancer. Previous studies have identified that CREB3L1 is an ER localized protein, and can be transported from the ER to the Golgi, which plays critical roles in modulating ER and Golgi stress responses (Sampieri et al., 2019). Notably, ER stress and Golgi apparatus have been confirmed as important triggers of autophagy in eukaryotic cells, and autophagy under ER and Golgi stress have been investigated in multiple cancers (Lin et al., 2019; Deng et al., 2020). These results suggest that CREB3L1 may function as a critical modulator of autophagy and the tumor immune microenvironment in human cancers, thereby modulating the initiation and progression of human cancers.

Previous studies have demonstrated that CREB3L1 is significantly correlated with chemoresistance in some human malignancies (Denard et al., 2012; Denard et al., 2018). In our study, we confirmed that CREB3L1 is significantly correlated with the sensitivity of several anti-cancer drugs, especially targeted therapy agents, such as cabozantinib, lenvatinib, and dabrafenib. These findings indicate that CREB3L1 exhibits great potential as a predictive biomarker for anti-cancer drugs sensitivity and a novel therapeutic target in human cancers, and dynamic monitoring of CREB3L1 expression may be a valuable approach to effectively evaluate therapeutic responses of cancer patients, thus helping choose the most suitable treatment strategy for the individual patient. However, the mechanisms linked to the involvement of CREB3L1 in modulating drug sensitivity are still inconclusive, which may be a future direction for further research on CREB3L1. Furthermore, there is no related research on the correlation between CREB3L1 expression and the immunotherapy efficacy until now. Hence, we deduced that CREB3L1 may play critical roles in cancer immunotherapies, and influence the immunotherapeutic efficacy. TMB and MSI have been identified as critical regulators that affect the effectiveness of immune checkpoint inhibitors in the treatment of several tumors (Boland and Goel, 2010; Samstein et al., 2019). Higher TMB levels and MSI levels are generally acknowledged as important biomarkers related to the high therapeutic efficacy in multiple immune checkpoint inhibitors treated cancers (Boland and Goel, 2010; Samstein et al., 2019). Our study presented evidence for the possible relationship of CREB3L1 expression with MSI and TMB in different cancers. Our results showed that CREB3L1 expression was positively correlated with TMB and MSI in COAD and STAD, indicating that patients with COAD or STAD who express CREB3L1 highly may have better responses to immune checkpoint inhibitors, while a negative connection of CREB3L1 with TMB and MSI was detected in BRCA, CESC, HNSC, KIRC, and LUSC. These findings indicate that CREB3L1 may function as an effective predictor for the efficacy of immune checkpoint inhibitors treatment in several cancers, such as BRCA, COAD and STAD. Therapeutic antibodies that block the PD-1/PD-L1 pathway have shown efficacy and have been approved in multiple cancers (Yi et al., 2022). However, the responses of the majority of patients to PD-1/PD-L1 inhibitors are limited, and previous studies have suggested that PD-1/PD-L1 expression, MSI and TMB levels are not robust biomarkers for predicting the efficacy of immune checkpoint inhibitors. Developing novel predictive biomarkers for PD-1/PD-L1 inhibitors therefore plays a key role in maximizing the immunotherapeutic efficacy (Zeng et al., 2019; Yi et al., 2022). To further confirm the potential role of CREB3L1 as the predictor for the immunotherapeutic responses, immunotherapy-related cohorts were enrolled in our investigation. We found that patients with no-response to PD-1/PD-L1 inhibitors have a higher CREB3L1 expression in patients with urothelial carcinoma and renal cell carcinoma, indicating that CREB3L1 is a promising predictor of the anti-PD-1/PD-L1 treatment benefits, thereby providing a new biomarker for predicting the immunotherapeutic efficacy.

However, there are some limitations in this quality study. The data enrolled in our study is extracted from online databases instead of our own clinical samples. Thus, these findings in our present study should be further verified by our own patients. In addition, the roles of CREB3L1 in a specific cancer type remain unclear, and the mechanisms by which CREB3L1 modulates the tumor immune microenvironment are still unknown. Future studies are needed to verify these findings in experiments.

In conclusion, we performed a comprehensive pan-cancer analysis of CREB3L1 by assessing multifaced aspects of CREB3L1, including the expression pattern, prognostic significance, genetic mutations, and immune-related signature, revealing the promising role of CREB3L1 as a promising indicator for the prognosis and immunotherapeutic efficacy of cancer patients and its immunoregulatory effects in pan-cancer. Our research provides

highlights into the biological functions of CREB3L1 in pan-cancer, and the candidate effect of CREB3L1 in predicting the prognosis and immune-related phenotypes.

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

ZL and YW: writing–original draft, data curation, and formal analysis. XX and XZ: writing–original draft. XZ, and JW: formal analysis. HC and TZ: writing–review and editing, and visualization. TL, and XT: writing–review and editing, supervision, and funding acquisition. All authors contributed to the writing and revision of the manuscript, knew the content of it, and approved its submission.

Funding

This study was supported by the grants from National Natural Science Foundation of China (Grant No.81871783 and 82072441); Outstanding Youth Scientist Foundation of Hunan province (NO. 2022JJ10095); Chenzhou Municipal Science and Technology Bureau Science and Technology Development Program (Grant No.yfzx201910).

Acknowledgments

We acknowledge TCGA database for providing its platforms and contributors for uploading its meaningful datasets, and Sangerbox online dataset (http://sangerbox.com) for providing some 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.2022.938510/full#supplementary-material

References

Aran, D. (2020). Cell-type enrichment analysis of bulk transcriptomes using xCell. *Methods Mol. Biol.* 2120, 263–276. doi:10.1007/978-1-0716-0327-7_19

Ascierto, M. L., McMiller, T. L., Berger, A. E., Danilova, L., Anders, R. A., Netto, G. J., et al. (2016). The intratumoral balance between metabolic and immunologic gene expression is associated with anti-PD-1 response in patients with renal cell carcinoma. *Cancer Immunol. Res.* 4, 726–733. doi:10.1158/2326-6066.CIR-16-0072

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

Bissonnette, C., Shilo, K., Liebner, D., Rogers, A., Pollock, R. E., and Iwenofu, O. H. (2021). An EWSR1-creb3l1 fusion gene in extraskeletal undifferentiated round cell sarcoma expands the spectrum of genetic landscape in the "Ewing-Like" undifferentiated round cell sarcomas. *Int. J. Surg. Pathol.* 29, 109–116. doi:10.1177/1066896920929081

Boland, C. R., and Goel, A. (2010). Microsatellite instability in colorectal cancer. *Gastroenterology* 138, 2073–2087. e3. doi:10.1053/j.gastro.2009.12.064

Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., et al. (2019). Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. *Ann. Oncol.* 30, 44–56. doi:10.1093/annonc/mdy495

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, 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

Chen, I. X., Chauhan, V. P., Posada, J., Ng, M. R., Wu, M. W., Adstamongkonkul, P., et al. (2019). Blocking CXCR4 alleviates desmoplasia, increases T-lymphocyte infiltration, and improves immunotherapy in metastatic breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* 116, 4558–4566. doi:10.1073/pnas.1815515116

Chen, J., Tan, Y., Sun, F., Hou, L., Zhang, C., Ge, T., et al. (2020). Single-cell transcriptome and antigen-immunoglobin analysis reveals the diversity of B cells in non-small cell lung cancer. *Genome Biol.* 21, 152. doi:10.1186/s13059-020-02064-6

Daniel, S. K., Seo, Y. D., and Pillarisetty, V. G. (2020). The CXCL12-CXCR4/CXCR7 axis as a mechanism of immune resistance in gastrointestinal malignancies. *Semin. Cancer Biol.* 65, 176–188. doi:10.1016/j.semcancer.2019.12.007

Dao, V., Yuki, K., Lo, Y. H., Nakano, M., and Kuo, C. J. (2022). Immune organoids: From tumor modeling to precision oncology. *Trends Cancer* 1, 1. doi:10.1016/j.trecan.2022.06.001

Denard, B., Jiang, S., Peng, Y., and Ye, J. (2018). CREB3L1 as a potential biomarker predicting response of triple negative breast cancer to doxorubicin-based chemotherapy. *BMC Cancer* 18, 813. doi:10.1186/s12885-018-4724-8

Denard, B., Lee, C., and Ye, J. (2012). Doxorubicin blocks proliferation of cancer cells through proteolytic activation of CREB3L1. *Elife* 1, e00090. doi:10.7554/eLife.00090

Denard, B., Seemann, J., Chen, Q., Gay, A., Huang, H., Chen, Y., et al. (2011). The membrane-bound transcription factor CREB3L1 is activated in response to virus infection to inhibit proliferation of virus-infected cells. *Cell Host Microbe* 10, 65–74. doi:10.1016/j.chom.2011.06.006

Deng, S., Liu, J., Wu, X., and Lu, W. (2020). Golgi apparatus: A potential therapeutic target for autophagy-associated neurological diseases. *Front. Cell Dev. Biol.* 8, 564975. doi:10.3389/fcell.2020.564975

Dienstmann, R., Villacampa, G., Sveen, A., Mason, M. J., Niedzwiecki, D., Nesbakken, A., et al. (2019). Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. *Ann. Oncol.* 30, 1622–1629. doi:10.1093/annonc/mdz287

Downs-Canner, S. M., Meier, J., Vincent, B. G., and Serody, J. S. (2022). B cell function in the tumor microenvironment. *Annu. Rev. Immunol.* 40, 169–193. doi:10.1146/annurev-immunol-101220-015603

Duraiswamy, J., Turrini, R., Minasyan, A., Barras, D., Crespo, I., Grimm, A. J., et al. (2021). Myeloid antigen-presenting cell niches sustain antitumor T cells and license PD-1 blockade via CD28 costimulation. *Cancer Cell* 39, 1623–1642. e20. doi:10.1016/j.ccell.2021.10.008

Feng, Y. X., Jin, D. X., Sokol, E. S., Reinhardt, F., Miller, D. H., and Gupta, P. B. (2017). Cancer-specific PERK signaling drives invasion and metastasis through CREB3L1. *Nat. Commun.* 8, 1079. doi:10.1038/s41467-017-01052-y

Fu, T., Dai, L. J., Wu, S. Y., Xiao, Y., Ma, D., Jiang, Y. Z., et al. (2021). Spatial architecture of the immune microenvironment orchestrates tumor immunity and therapeutic response. *J. Hematol. Oncol.* 14, 98. doi:10.1186/s13045-021-01103-4

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., Craft, B., Swatloski, T., Cline, M., Morozova, O., Diekhans, M., et al. (2015). The UCSC cancer genomics browser: Update 2015. *Nucleic Acids Res.* 43, D812–D817. doi:10.1093/nar/gku1073

Hinshaw, D. C., and Shevde, L. A. (2019). The tumor microenvironment innately modulates cancer progression. *Cancer Res.* 79, 4557–4566. doi:10.1158/0008-5472.CAN-18-3962

Hugo, W., Zaretsky, J. M., Sun, L., Song, C., Moreno, B. H., Hu-Lieskovan, S., et al. (2016). Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* 165, 35–44. doi:10.1016/j.cell.2016.02.065

Hui, E., Cheung, J., Zhu, J., Su, X., Taylor, M. J., Wallweber, H. A., et al. (2017). T cell costimulatory receptor CD28 is a primary target for PD-1-mediated inhibition. *Science* 355, 1428–1433. doi:10.1126/science.aaf1292

Kamikawa, Y., Saito, A., Matsuhisa, K., Kaneko, M., Asada, R., Horikoshi, Y., et al. (2021). OASIS/CREB3L1 is a factor that responds to nuclear envelope stress. *Cell Death Discov.* 7, 152. doi:10.1038/s41420-021-00540-x

Kamphorst, A. O., Wieland, A., Nasti, T., Yang, S., Zhang, R., Barber, D. L., et al. (2017). Rescue of exhausted CD8 T cells by PD-1-targeted therapies is CD28-dependent. *Science* 355, 1423–1427. doi:10.1126/science.aaf0683

Keller, R. B., Tran, T. T., Pyott, S. M., Pepin, M. G., Savarirayan, R., McGillivray, G., et al. (2018). Monoallelic and biallelic CREB3L1 variant causes mild and severe osteogenesis imperfecta, respectively. *Genet. Med.* 20, 411–419. doi:10.1038/gim.2017.115

Kennedy, L. B., and Salama, A. K. S. (2020). A review of cancer immunotherapy toxicity. *Ca. Cancer J. Clin.* 70, 86–104. doi:10.3322/caac.21596

Kumagai, S., Koyama, S., and Nishikawa, H. (2021). Antitumour immunity regulated by aberrant ERBB family signalling. *Nat. Rev. Cancer* 21, 181–197. doi:10.1038/s41568-020-00322-0

Lei, X., Lei, Y., Li, J. K., Du, W. X., Li, R. G., Yang, J., et al. (2020). Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett.* 470, 126–133. doi:10.1016/j.canlet.2019.11.009

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

Lin, Y., Jiang, M., Chen, W., Zhao, T., and Wei, Y. (2019). Cancer and ER stress: Mutual crosstalk between autophagy, oxidative stress and inflammatory response. *Biomed. Pharmacother.* 118, 109249. doi:10.1016/j.biopha.2019.109249

Liu, L. Q., Feng, L. F., Nan, C. R., and Zhao, Z. M. (2018). CREB3L1 and PTN expressions correlate with prognosis of brain glioma patients. *Biosci. Rep.* 38, BSR20170100. doi:10.1042/BSR20170100

Luo, H., Xia, X., Kim, G. D., Liu, Y., Xue, Z., Zhang, L., et al. (2021). Characterizing dedifferentiation of thyroid cancer by integrated analysis. *Sci. Adv.* 7, eabf3657. doi:10.1126/sciadv.abf3657

Marzagalli, M., Ebelt, N. D., and Manuel, E. R. (2019). Unraveling the crosstalk between melanoma and immune cells in the tumor microenvironment. *Semin. Cancer Biol.* 59, 236–250. doi:10.1016/j.semcancer.2019.08.002

Mellor, P., Deibert, L., Calvert, B., Bonham, K., Carlsen, S. A., and Anderson, D. H. (2013). CREB3L1 is a metastasis suppressor that represses expression of genes regulating metastasis, invasion, and angiogenesis. *Mol. Cell. Biol.* 33, 4985–4995. doi:10.1128/MCB.00959-13

Mistarz, A., Graczyk, M., Winkler, M., Singh, P. K., Cortes, E., Miliotto, A., et al. (2021). Induction of cell death in ovarian cancer cells by doxorubicin and oncolytic vaccinia virus is associated with CREB3L1 activation. *Mol. Ther. Oncolytics* 23, 38–50. doi:10.1016/j.omto.2021.04.014

Morishita, S., Yasuda, H., Yamawaki, S., Kawaji, H., Itoh, M., Edahiro, Y., et al. (2021). CREB3L1 overexpression as a potential diagnostic marker of Philadelphia chromosome-negative myeloproliferative neoplasms. *Cancer Sci.* 112, 884–892. doi:10.1111/cas.14763

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

Pitt, J. M., Marabelle, A., Eggermont, A., Soria, J. C., Kroemer, G., and Zitvogel, L. (2016). Targeting the tumor microenvironment: Removing obstruction to anticancer immune responses and immunotherapy. *Ann. Oncol.* 27, 1482–1492. doi:10.1093/annonc/mdw168

Plattner, C., Finotello, F., and Rieder, D. (2020). Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. *Methods Enzymol.* 636, 261–285. doi:10.1016/bs.mie.2019.05.056

Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E., and Gfeller, D. (2017). Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife* 6, e26476. doi:10.7554/eLife.26476

Rajamanickam, V., Bui, C., Bernard, B., Pucilowska, J., Ballesteros-Merino, C., Schmidt, M., et al. (2021). Germinal center reactions in tertiary lymphoid structures associate with neoantigen burden, humoral immunity and long-term survivorship in pancreatic cancer. *Oncoimmunology* 10, 1900635. doi:10.1080/2162402X.2021.1900635

Rose, M., Schubert, C., Dierichs, L., Gaisa, N. T., Heer, M., Heidenreich, A., et al. (2014). OASIS/CREB3L1 is epigenetically silenced in human bladder cancer facilitating tumor cell spreading and migration *in vitro*. *Epigenetics* 9, 1626–1640. doi:10.4161/15592294.2014.988052

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

Sampieri, L., Di Giusto, P., and Alvarez, C. (2019). CREB3 transcription factors: ER-golgi stress transducers as hubs for cellular homeostasis. *Front. Cell Dev. Biol.* 7, 123. doi:10.3389/fcell.2019.00123

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

Sharma, P., Hu-Lieskovan, S., Wargo, J. A., and Ribas, A. (2017). Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* 168, 707–723. doi:10.1016/j.cell.2017.01.017

Song, G., Shi, Y., Meng, L., Ma, J., Huang, S., Zhang, J., et al. (2022). Single-cell transcriptomic analysis suggests two molecularly subtypes of intrahepatic cholangiocarcinoma. *Nat. Commun.* 13, 1642. doi:10.1038/s41467-022-29164-0

Tabib, T., Huang, M., Morse, N., Papazoglou, A., Behera, R., Jia, M., et al. (2021). Myofibroblast transcriptome indicates SFRP2(hi) fibroblast progenitors in systemic sclerosis skin. *Nat. Commun.* 12, 4384. doi:10.1038/s41467-021-24607-6

Vellanki, R. N., Zhang, L., and Volchuk, A. (2013). OASIS/CREB3L1 is induced by endoplasmic reticulum stress in human glioma cell lines and contributes to the unfolded protein response, extracellular matrix production and cell migration. *PLoS One* 8, e54060. doi:10.1371/journal.pone.0054060

Vilar, E., and Gruber, S. B. (2010). Microsatellite instability in colorectal cancer-the stable evidence. *Nat. Rev. Clin. Oncol.* 7, 153–162. doi:10.1038/nrclinonc.2009.237

Yang, Y. (2015). Cancer immunotherapy: Harnessing the immune system to battle cancer. *J. Clin. Invest.* 125, 3335–3337. doi:10.1172/JCI83871

Yi, M., Zheng, X., Niu, M., Zhu, S., Ge, H., and Wu, K. (2022). Combination strategies with PD-1/PD-L1 blockade: Current advances and future directions. *Mol. Cancer* 21, 28. doi:10.1186/s12943-021-01489-2

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

Zeng, D., Li, M., Zhou, R., Zhang, J., Sun, H., Shi, M., et al. (2019). Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol. Res.* 7, 737–750. doi:10.1158/2326-6066.CIR-18-0436

Zhang, Q., Lv, L., Ma, P., Zhang, Y., Deng, J., and Zhang, Y. (2021). Identification of an autophagy-related pair signature for predicting prognoses and immune activity in pancreatic adenocarcinoma. *Front. Immunol.* 12, 743938. doi:10.3389/fimmu.2021.743938

Zhang, S., Zhang, Y., and Min, P. (2021). Single-cell and bulk transcriptome data integration reveals dysfunctional cell types and aberrantly expressed genes in hypertrophic scar. *Front. Genet.* 12, 806740. doi:10.3389/fgene.2021.806740

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 14 September 2022
DOI 10.3389/fgene.2022.955355

Check for updates

OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong, SAR China

REVIEWED BY
Qingjia Chi,
Wuhan University of Technology, China
Yingying Cao,
Shanghai Jiao Tong University, China
Maryam Hassan Al-Zahrani,
King Abdulaziz University, Saudi Arabia

*CORRESPONDENCE
Huabing Zhang,
huabingzhang@ahmu.edu.cn
Wei Chen,
chenwei366@ahmu.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 28 May 2022
ACCEPTED 22 August 2022
PUBLISHED 14 September 2022

CITATION
Wang X, Sun R, Chan S, Meng L, Xu Y, Zuo X, Wang Z, Hu X, Han Q, Dai L, Bai T, Yu Z, Wang M, Yang W, Zhang H and Chen W (2022), PANoptosis-based molecular clustering and prognostic signature predicts patient survival and immune landscape in colon cancer. *Front. Genet.* 13:955355. doi: 10.3389/fgene.2022.955355

COPYRIGHT
© 2022 Wang, Sun, Chan, Meng, Xu, Zuo, Wang, Hu, Han, Dai, Bai, Yu, Wang, Yang, Zhang 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.

PANoptosis-based molecular clustering and prognostic signature predicts patient survival and immune landscape in colon cancer

Xu Wang[1†], Rui Sun[1†], Shixin Chan[1†], Lei Meng[1], Yuanmin Xu[1], Xiaomin Zuo[1], Zhenglin Wang[1], Xianyu Hu[1], Qijun Han[1], Longfei Dai[1], Tao Bai[1], Zhen Yu[1], Ming Wang[1], Wenqi Yang[1], Huabing Zhang[2,3]* and Wei Chen[1]*

[1]Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei, Anhui, China, [2]Affiliated Chuzhou Hospital of Anhui Medical University, First People's Hospital of Chuzhou, Chuzhou, Anhui, China, [3]Department of Biochemistry and Molecular Biology, Metabolic Disease Research Center, School of Basic Medicine, Anhui Medical University, Hefei, Anhui, China

PANoptosis is a newly-discovered cell death pathway that involves crosstalk and co-ordination between pyroptosis, apoptosis, and necroptosis processes. However, the roles of PANoptosis-related genes (PRGs) in prognosis and immune landscape of colon cancer remain widely unknown. Here, we performed a bioinformatics analysis of expression data of nineteen PRGs identified from previous studies and clinical data of colon cancer patients obtained from TCGA and GEO databases. Colon cancer cases were divided into two PRG clusters, and prognosis-related differentially expressed genes (PRDEGs) were identified. The patient data were then separated into two corresponding distinct gene clusters, and the relationship between the risk score, patient prognosis, and immune landscape was analyzed. The identified PRGs and gene clusters correlated with patient survival and immune system and cancer-related biological processes and pathways. A prognosis signature based on seven genes was identified, and patients were divided into high-risk and low-risk groups based on the calculated risk score. A nomogram model for prediction of patient survival was also developed based on the risk score and other clinical features. Accordingly, the high-risk group showed worse prognosis, and the risk score was related to immune cell abundance, cancer stem cell (CSC) index, checkpoint expression, and response to immunotherapy and chemotherapeutic drugs. Results of quantitative real-time polymerase chain reaction (qRT-PCR) showed that LGR5 and VSIG4 were differentially expressed between normal and colon cancer samples. In conclusion, we demonstrated the potential of PANoptosis-based molecular clustering and prognostic signatures for prediction of patient survival and tumor microenvironment (TME) in colon cancer. Our findings may improve our understanding of the role of PANoptosis in colon cancer, and enable the development of more effective treatment strategies.

KEYWORDS

PANoptosis, colon cancer, prognosis, tumor microenvironment, immunotherapy

Introduction

It is estimated that there are more than 1.9 million new cases of colorectal cancer (CRC). CRC also caused 935,000 deaths in 2020, accounting for approximately one-tenth of all cancer cases and deaths. Among all types of cancer, CRC ranks third in incidence rate but second in mortality (Sung et al., 2021). Patients with early colon cancer can be surgically treated. However, most patients with advanced colon cancer experience cancer recurrence and metastasis, and their 5-years survival rates are less than 10% (Bhandari et al., 2017; Doonan et al., 2017; Russo et al., 2019). With the development of chemotherapy and targeted drugs, the overall survival rate of patients with colon cancer has been significantly higher than that in the past. In recent years, progress in tumor immunotherapy and the application of immune checkpoint inhibitors have led to improvements in cancer treatment. Programmed cell death protein 1 (*PD-1*), first discovered in 1992, is a 288 amino acid protein expressed on the surface of T-cells and is involved in apoptosis (Kouo et al., 2015). In 2014, the FDA approved small cell blocking antibody (volumab) for the treatment of advanced lung cancer; volumab was further approved for the treatment of melanoma in 2015.

Cell death is a basic physiological process occurring in all organisms. Its role spans from embryonic development, organ maintenance, and aging, to coordinating immune responses and autoimmunity (Bertheloot et al., 2021). Among the proposed forms of programmed cell death (PCD), pyroptosis, apoptosis, and necroptosis are the most clearly defined. These processes involve complex molecular mechanisms responsible for the initiation, transduction, and execution of cell death (Galluzzi et al., 2018; Kesavardhana et al., 2020). Early studies on cell death have mainly focused on the unique procedures and biochemical functions under these separate mechanisms. However, recent studies have emphasized on the redundancies and crosstalk among them. PANoptosis is a newly discovered concept that highlights the crosstalk and coordination between pyroptosis, apoptosis, and necroptosis (Malireddi et al., 2019; Samir et al., 2020). Dysregulated cell death and inflammatory responses are related to tumorigenesis. Hanahan (2022) suggested that resistance to cell death is one of the basic hallmarks of cancer. Caspase-8 (*CASP8*) is a molecular switch that controls pyroptosis, apoptosis, and necroptosis (Fritsch et al., 2019). Jiang et al. reported that *CASP8* is a key protein in the crosstalk signaling pathway of PANoptosis in cancer (Jiang et al., 2021). A recent study (Karki et al., 2021a) showed that *ZBP1* activates PANoptosis through *RIPK3* signaling, *ADAR1* negatively regulates *ZBP1*-mediated apoptosis, and blocking ADAR1 activity contributes to apoptosis and inhibits tumorigenesis. Acquiring more knowledge about the effects of PANoptosis on cancer is vital for developing new strategies for cancer therapy.

Many studies have focused on constructing tumor classifications and prognostic signatures based on gene and non-coding RNA expression levels to predict the survival and immune landscape of patients with cancer. In recent years, genes and non-coding RNA related to various forms of cell death have been explored in many studies, including autophagy-, ferroptosis-, pyroptosis-, necroptosis-related genes and lncRNAs. Wang et al. (2021) identified six autophagy-related genes and developed a prognostic signature that can independently predict prognosis and reflect overall immune response intensity in the colon cancer microenvironment. Nie et al. (2021) constructed a novel ferroptosis-related gene signature to predict the prognosis and immune status of patients with colon cancer. Song et al. (2021) used pyroptosis-related genes to identify molecular subtypes and develop a prognostic model for characterizing tumor microenvironment infiltration in colorectal cancer. A recent study (Zhao et al., 2021) also used necroptosis-related lncRNAs to predict prognosis and identify molecular classifications to distinguish between cold and hot tumors in gastric cancer. However, genes related to the crosstalk and coordination between different types of cell death have not been well studied.

Our study demonstrated that PANoptosis-based molecular clustering and prognostic signatures could predict prognosis and the intratumoral immune landscape in patients with colon cancer. First, 951 colon cancer patients were divided into two discrete PRG clusters based on expression levels of PANoptosis-related genes (PRGs). Patients were then classified into two clusters based on differentially expressed genes (DEGs) identified from the two PRG clusters. A risk score was further calculated, and a prognostic signature was established to predict overall survival (OS) and response to immunotherapy in colon cancer patients.

Materials and methods

Data acquisition

Gene expression data (fragments per kilobase million, FPKM) and relevant clinical information of colon cancer patients were retrieved from the TCGA (https://portal.gdc.cancer.gov) and GEO (https://www.ncbi.nlm.nih.gov/geo/, ID: GSE39583) databases. Nineteen PANoptosis-related genes (PRGs) were identified from prior studies (Malireddi et al., 2019; Karki et al., 2020; Malireddi et al., 2020; Samir et al., 2020; Briard et al., 2021; Jiang et al., 2021; Lee et al., 2021; Place et al., 2021; Nguyen and Kanneganti, 2022), and the genes are listed in Supplementary Table S1. The FPKM values of TCGA colon adenocarcinoma (COAD) were transformed into transcripts per kilobase million (TPM) using the R software

(version 4.1.1). TCGA and GEO data were combined and batch effects were eliminated using the Combat algorithm of the *sva* R package. Patients with no follow-up data or incomplete clinical information were excluded from the study. Finally, 951 patients were included in the study. The clinical information of patients with colon cancer is presented in Supplementary Table S2.

Consensus clustering analysis of PANoptosis-related genes

A consensus clustering analysis based on PRGs expression was performed to investigate the connections between PRGs and colon cancer subtypes using the *ConsensusClusterPlus* R package. The classification with the highest intragroup correlations and the lowest intergroup correlations was identified. The prognosis of the two clusters was compared using the Kaplan-Meier (KM) method and log-rank test. Principal component analysis (PCA) was performed using the stats R package. Differences in clinical features between two clusters were analyzed using the Wilcoxon test, and DEGs between two clusters were identified with the criteria of |log fold change (FC)| >1 and p-value < 0.05, using *limma* package. To explore the differences in biological processes between the two PRG clusters, we performed gene set variation analysis (GSVA) using gsva R package. Single-sample gene set enrichment analysis (ssGSEA) was used to calculate the scores of infiltrating immune cells and evaluate the activity of immune-related functions.

Gene oncology and kyoto encyclopedia of genes and genomes analyses

To understand the biological functions and pathways related to the DEGs, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed using the *ggplot2*, *Bioconductor*, and *org. Hs.eg.db* R packages. p-values and q-values < 0.05 were considered statistically significant.

Construction of the PANoptosis-related prognostic signature

Prognosis-related DEGs (PRDEGs) were selected using a univariate Cox regression analysis. To identify additional PANoptosis-related genes for signature construction. We classified patients into two distinct clusters based on the expression of PRDEGs. Survival time, clinical features, and PRG expression were compared between the two gene clusters, and DEGs between gene clusters A and B were identified. Finally, seven genes were included to construct the prognostic signature after least absolute shrinkage and selection operator (LASSO) regression analysis and multivariate Cox regression analysis using the *survival*, *survminer*, and *glmnet* R packages. The risk score was calculated based on the expression levels of the seven genes, and the patients were divided into high- and low-risk groups using the median risk score. The KM analysis was used to evaluate survival differences between high-risk and low-risk groups, and the receiver operating characteristic (ROC) and area under the curve (AUC) were used to test the prediction efficiency of the risk score. A nomogram model was developed based on risk scores and clinical factors. Calibration graphs were constructed to show the differences between the nomogram-predicted and actual survival rates of colon cancer patients.

Evaluation of the tumor microenvironment between the high- and low-risk groups

To better understand the relationship between the risk score and tumor microenvironment (TME), CIBERSORT was used to quantify the abundance of infiltrating immune cells in heterogeneous samples from the high- and low-risk groups. The correlation between the abundance of immune cells and the seven genes was analyzed. TME scores, including stromal, immune, and ESTIMATE scores, of high- and low-risk groups were also compared using the Wilcoxon signed-rank test.

Analyses of mutations, microsatellite instability and cancer stem cell index between high- and low-risk groups

To explore the gene mutations in colon cancer patients in high- and low-risk groups, we generated the mutation annotation format (MAF) using the *maftools* R package. The tumor mutation burden (TMB) score of patients was calculated, and the correlation between the risk score and TMB was analyzed using the Spearman method. Furthermore, we analyzed the association between risk groups and MSI and CSC index.

Response to immunotherapy and chemotherapeutic drugs

The Cancer Immunome Atlas (TCIA, https://tcia.at/) is a dataset containing TCGA data of 20 solid cancers and more than 8,000 tumor samples. It can be used to calculate the immunophenotypic score (IPS) of tumor samples to predict the response to cytotoxic T lymphocyte antigen-4 (CTLA-4) and programmed cell death protein 1 (PD-1) blockers (Charoentong et al., 2017). IC50 is half of the maximum inhibitory concentration and represents the concentration of the drug required for 50% inhibition of cancer cells. To determine the relationship between the risk score and

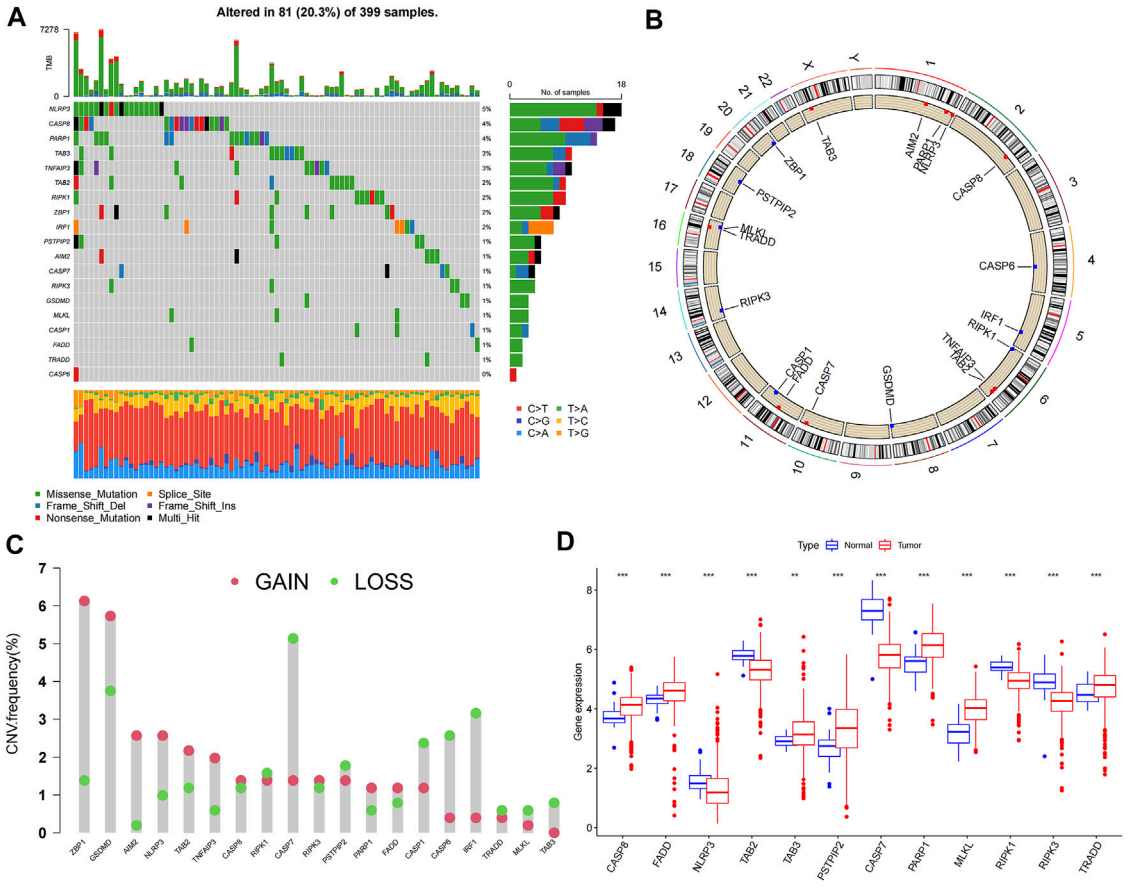


FIGURE 1
Genetic and transcriptional alterations of 19 PRGs in colon cancer. **(A)** Mutation frequencies of 19 PRGs in colon cancer patients from TCGA cohort; **(B)** Locations of CNV alterations in PRGs on 23 chromosomes; **(C)** Frequencies of CNV gain, loss, and non-CNV among PRGs; **(D)** Expression levels of PRGs between normal and tumor samples. $^{**}p < 0.01$; $^{***}p < 0.001$.

response to immunotherapy and chemotherapeutic drugs, checkpoint expression, immune subtypes, IPS of tumor samples, and the IC50 of drugs in the two risk groups were calculated and compared.

The verification of LGR5, VSIG4, GZMB, and ITLN1 by quantitative real-time polymerase chain reaction

Ten pairs of colon cancer and corresponding non-tumor tissues were collected from colon cancer patients in The First Affiliated Hospital of Anhui Medical University, the samples were preserved at −80°C. The study was approved by the Ethics Committee of The First Affiliated Hospital of Anhui Medical University. All participants signed an informed consent form. Total RNA was extracted using The HiPure Universal RNA Kit (Shanghai, Magen). Extracted RNA was reverse transcribed into cDNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, United States). The concentration of cDNA was measured using TB Green Premix Ex Taq II (GenStar, China) under the LightCycler480 System (Applied Biosystems, Waltham, MA, United States). The relative expression levels were computed using the $2^{-\Delta\Delta Ct}$ method, normalizing with *36B4*. The primer sequences for PCR amplification are shown in Supplementary Table S3. The differences of expression levels between colon cancer tissues and adjacent non-cancer tissues were compared using unpaired *t*-tests. The graphs were drawn using GraphPad Prism software (version 9.0.0).

Results

Landscape of genetic variation of PANoptosis-related genes in colon cancer

Expression data of 457 COAD patients were downloaded from the TCGA database, and expression levels of PRGs were

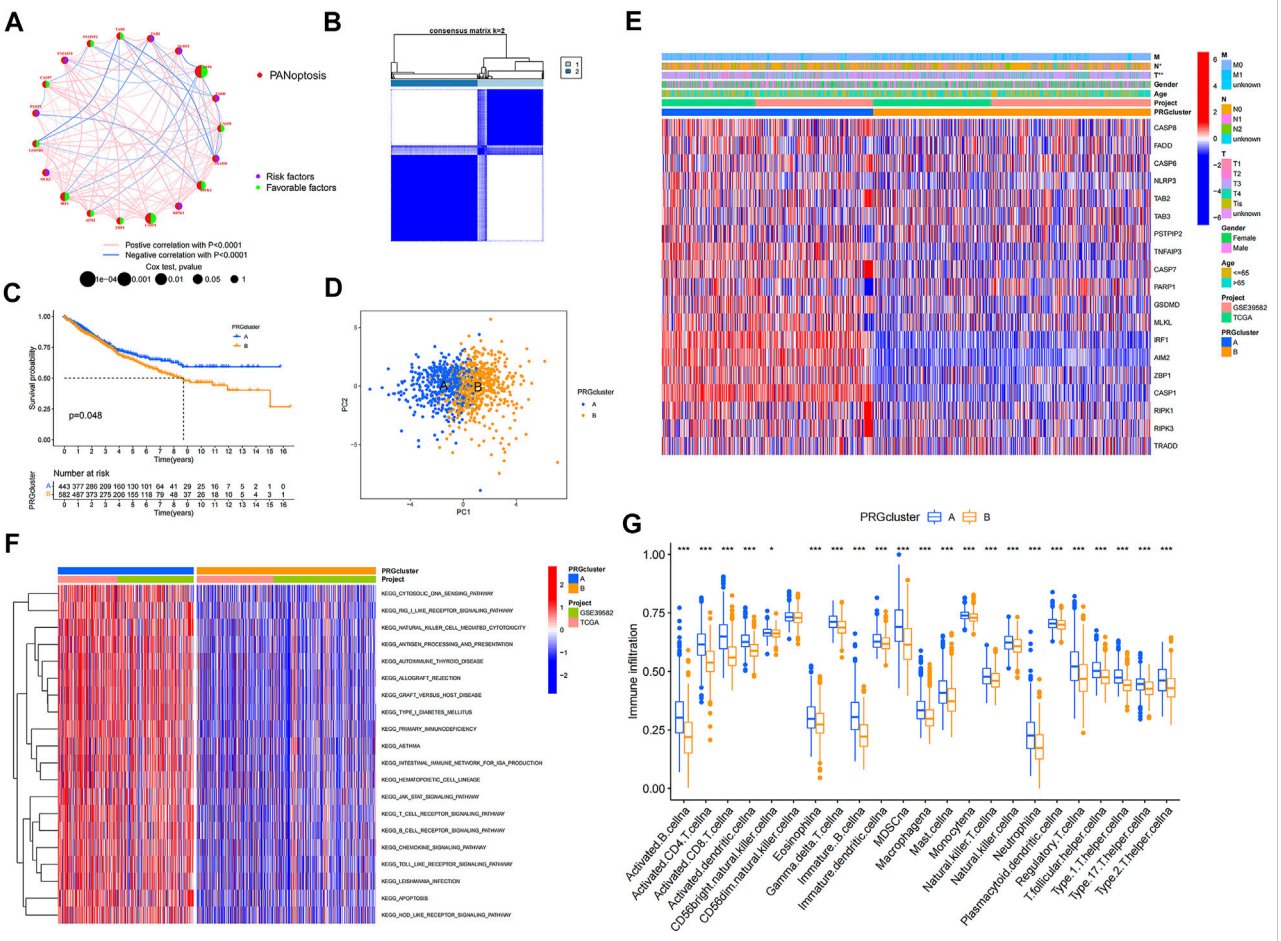


FIGURE 2
PRGclusters and clinical characteristics between colon cancer samples in two clusters. Relationship of tumor microenvironment in two PRGclusters. **(A)** Interactions among PRGs in colon cancer. The lines among the PRGs represents their interactions. Blue and red represent negative and positive correlations. **(B)** Two PRGclusters were defined using consensus clustering analyses. **(C)** KM curve indicated that PRGcluster a had longer survival time than PRGcluster B (p = 0.048). **(D)** PCA showed good distiction between two PRGclusters. **(E)** Heatmaps showed the relationship between PRGclusters and clinical features and PRGs expression in colon cancer patients. **(F)** GSVA showed the enriched pathways in PRGclusters. **(G)** ssGSEA investigated the differences of immune cell infiltration between two clusters. *p < 0.05; **p < 0.01.

compared between 41 normal and 473 tumor samples. Nineteen PRGs from previous studies were included in this study. The somatic mutation incidence in the 19 PRGs of colon cancer patients is shown in Figure 1A; 81 (20.3%) of the 399 samples had altered PRGs. Among the 19 PRGs, *NLRP3* showed the highest mutation frequency. Figure 1B shows the locations of the CNV alterations in PRGs on their chromosomes. Twelve PRGs were differentially expressed in colon cancer samples compared with their expression in normal samples. Somatic copy number alterations of the 19 PRGs were analyzed; *ZBP1, GSDMD, AIM2*, and *NLRP3* had the highest copy number variation (CNV), whereas *CASP7, CASP1, CASP6*, and *IRF3* showed significant CNV decreases (Figure 1C). Among these 12 PRGs, seven genes were upregulated in tumor samples, including *CASP8, FADD, TAB3, PSTPIP2, PARP1, MLKL*, and *TRADD*, whereas the other five genes, including *NLRP3, TAB2, CASP7, RIPK1*, and *RIPK3*, were downregulated in tumor samples (p < 0.05) (Figure 1D).

Identification of PANoptosis-related gene clusters in colon cancer

To explore the interactions between the 19 PRGs and their prognostic significance, a network was constructed, as shown in Figure 2A. Kaplan–Meier curves of the relationship between PRGs expression and the prognosis of colon cancer patients were shown in Supplementary Figure S1. Consensus clustering analysis was performed to explore the relationship between PRG expression and tumor classification (Supplementary

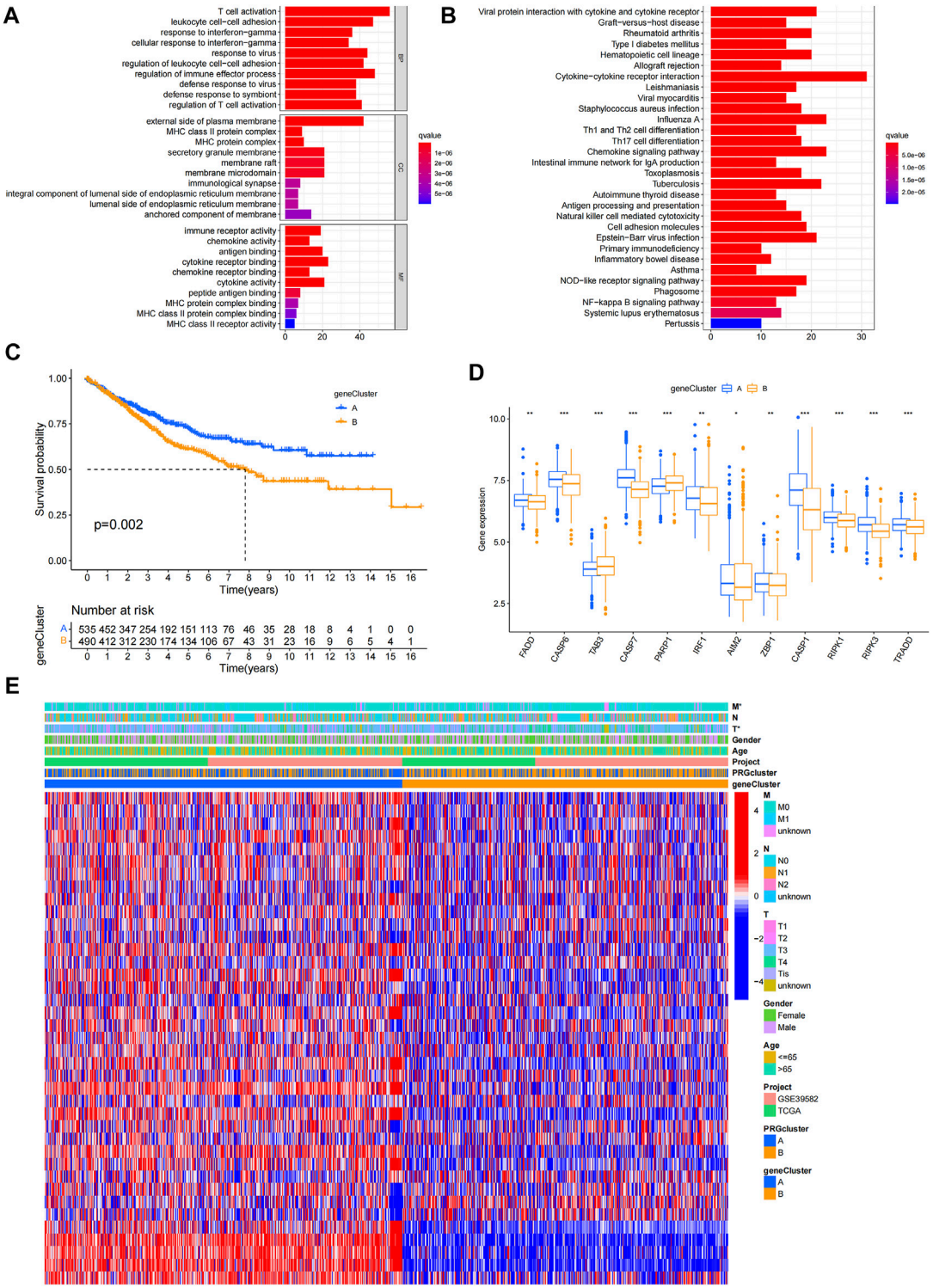


FIGURE 3
Identification of geneclusters based on DEGs. **(A–B)** GO and KEGG analyses showed the relevant biological processes (BP), cellular components (CC), molecular functions (MF) and pathways. **(C)** Heatmap showed the association between genecluster and clinical features. **(D)** KM curves showed genecluster A had a more favorable prognosis. **(E)** Expression levels of PRGs in two geneclusters. *$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$.

Figure S2). Clusters with the highest intragroup correlations and lowest intergroup correlations were identified. By increasing the clustering variable (k), we found that when $k = 2$, classification met the standard. Colon cancer patients were divided into two PRG clusters (A and B) based on PRG expression levels (Figure 2B). As shown in Figure 2C, patients in PRG cluster A had a significantly longer survival time than those in cluster B ($p = 0.048$). PCA showed a satisfactory separation between PRG cluster A and B (Figure 2D). Figure 2E shows the association between PRG clusters and clinical features and PRG expression in colon cancer patients. Tumor infiltration and lymph node metastasis correlated with PRG clusters ($p < 0.05$). GSVA showed that PRG cluster A was significantly enriched in immune-related pathways, including natural killer cell-mediated cytotoxicity, antigen processing and presentation, primary immunodeficiency, B cell, and T-cell receptor signaling pathways (Figure 2F). To evaluate the differences in immune cell infiltration between the two clusters, ssGSEA was performed, and the results showed that PRG cluster A had higher immune cell infiltration levels, including those of activated B cells, activated CD4 + T-cells, activated CD8 + T-cells, activated dendritic cells, macrophages, mast cells, and natural killer cells (Figure 2G).

Identification of gene clusters based on differentially expressed genes

DEGs were identified, and GO and KEGG analyses showed the relevant biological processes (BP), cellular components (CC), molecular functions (MF), and pathways (Figures 3A,B). These DEGs were mainly related to the BP of T cell activation, leukocyte cell-cell adhesion, and response to interferon-gamma, and were correlated with the CC such as the external side of the plasma membrane, major histocompatibility complex (MHC) class II protein complex, and MHC protein complex. Furthermore, they were involved in the MF of immune receptor activity, chemokine activity, and antigen binding. According to KEGG analysis, these DEGs participate in certain cancer-related pathways, including the chemokine signaling pathway, NOD-like receptor signaling pathway, and NF-κB signaling pathway. PRDEGs were identified using univariate Cox regression analysis. Patients were then divided into two clusters (gene cluster A and gene cluster B) based on PRDEG expression (Supplementary Figure S3; Figure 3C) shows that cluster A had higher survival rates than cluster B ($p = 0.002$). In Figure 3D, the boxplot shows that *FADD, CASP6, CASP7, IRF1, AIM2, ZBP1, CASP1, RIPK1, RIPK3,* and *TRADD* were upregulated in cluster A, whereas *TAB3* and *PARP1* were downregulated in cluster B ($p < 0.05$). The heatmaps show the association between gene clusters and clinical features and between PRDEGs expression and PRG clusters. Furthermore, gene clusters were significantly related to tumor infiltration and metastasis ($p < 0.05$) (Figure 3E).

Development abd validation of the PANoptosis-related prognostic signature

LASSO and Cox regression analyses were performed to screen for prognosis-related DEPRGs (Figures 4A,B). After selection, seven genes were included in the calculation of the risk score based on the following formula: Risk score = $\sum_{i=1}^{n} \beta i^{*} \lambda i$, where n represents the number of genes included to construct the signature, and βi and λi represent the regression coefficient and gene expression value, respectively. Boxplots showed that PRG cluster B and gene cluster B had higher risk scores than PRG cluster A and gene cluster A (Figures 4C,D). A Sankey diagram showed the associations among PRG cluster, gene cluster, risk groups, and survival status (Figure 4E). Fifteen of the nineteen PRGs were differentially expressed between the high- and low-risk groups (Figure 4F). Seven genes were used to construct the prognostic signature; Figure 5A shows the expression differences of these seven genes between the two risk groups. Based on the risk score, patients were divided into high- and low-risk groups, and patients with higher risk scores had a higher risk of mortality (Figure 5B). As shown in Figure 5C, the KM curve was plotted to show the survival differences between the two groups. Patients in the high-risk group had a significantly lower probability of survival than those in the low-risk group ($p < 0.001$). ROC curves were drawn to test the prediction efficiency of the risk score, and the AUCs for one-, three-, and 5-years survival were 0.612, 0.650, and 0.676, respectively (Figure 5D). The results of the risk score in TCGA (Supplementary Figure S4) and GSE39582 (Supplementary Figure S5) cohorts were also shown. The risk score and other clinical features were used to construct a nomogram model (Figure 5E). Calibration plots showing the differences between the nomogram-predicted and actual survival probabilities of colon cancer patients showed that the predicted survival probabilities were close to the actual survival probabilities (Figure 5F), indicating that this nomogram model accurately predicted the survival of colon cancer patients. The results for three independent validation cohorts, which were GSE17536 (Figure 5G, $p = 0.041$, 1-year AUC = 0.598, 3-years AUC = 0.624, 5-years AUC = 0.589), GSE17537 (Figure 5H, $p = 0.048$, 1-year AUC = 0.728, 3-years AUC = 0.624, 5-years AUC = 0.542), and GSE29621 (Figure 5I, $p = 0.011$, 1-year AUC = 0.763, 3-years AUC = 0.717, 5-years AUC = 0.702), revealed that the risk score could efficiently predict patient survival.

Comparative evaluation of the tumor microenvironments of high- and low-risk groups

Figure 6Ashows the correlation between the risk score and immune cell abundance: M0 macrophages, M2 macrophages, activated mast cells, and neutrophils were positively related to the

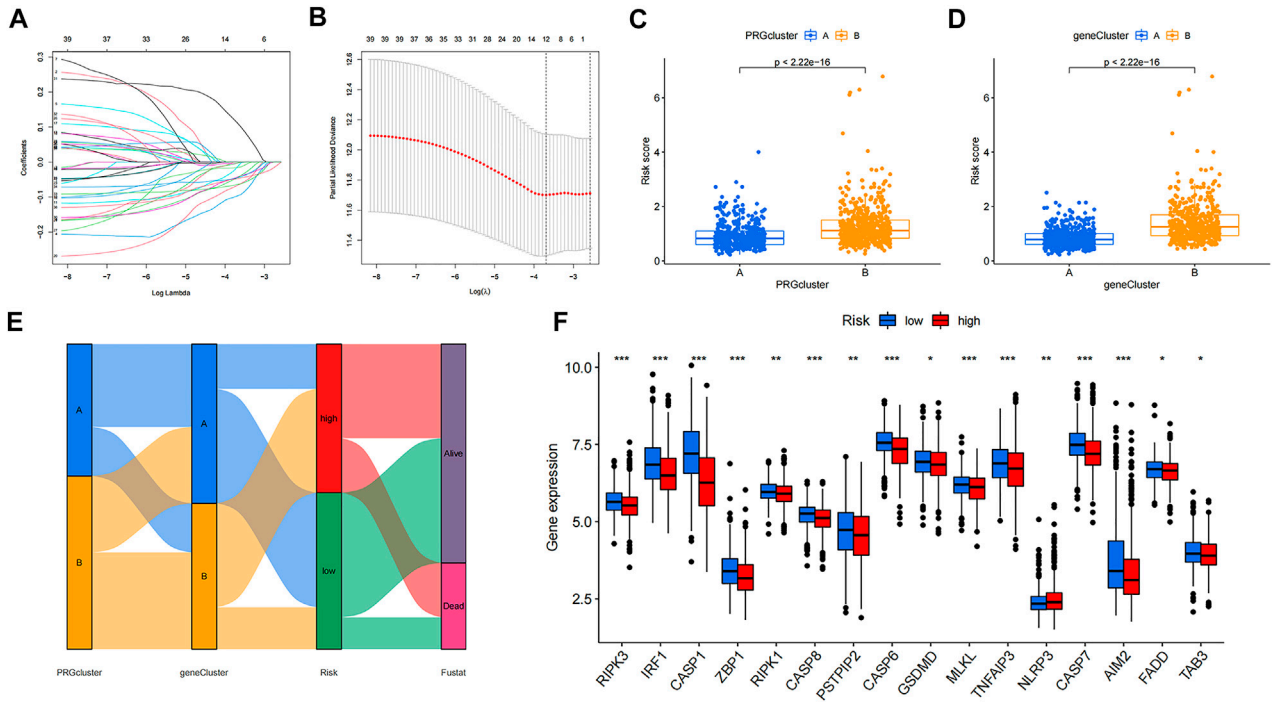


FIGURE 4
Identification of 7 genes for calculating the risk score and the relationship between molecular classifications, PRG expression levels and the risk score. **(A–B)** The LASSO regression analysis and partial likelihood deviance on the prognostic genes. **(C–D)** Association between risk score and molecular classifications. **(E)** Sankey plot showed the correlation between molecular classifications, risk groups and survival status in colon cancer patients. **(F)** Expression levels of PRGs in two risk groups. *$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$.

risk score, while naive B cells, activated dendritic cells, resting dendritic cells, M1 macrophages, resting mast cells, resting NK cells, plasma cells, activated memory CD4 + T-cells, resting memory CD4 + T-cells, CD8 + T-cells, and follicular helper T-cells were negatively related to the risk score. The relationship between the abundance of immune cells and seven genes in the prognostic signature was also evaluated (Figure 6B). The TME scores in the two risk groups were calculated, and the high-risk group was found to have a higher stromal and lower immune score (Figure 6C).

Comparative analysis of mutations, microsatellite instability and cancer stem cell index in high- and low-risk groups

Differences in somatic mutations between the two risk groups of colon cancer patients were analyzed; the five most mutated genes in the high- and low-risk groups were *APC, TP53, TTN, KRAS*, and *SYNE1* (Figures 7A,B). TMB (Figure 7C) and MSI (Figure 7D) did not show a significant relationship with the risk score, while CSC (Figure 7E) was negatively correlated with the risk score ($R = -0.15$, $p < 0.01$).

Response to immunotherapy and chemotherapeutic drugs

To analyze the ability of risk score to predict potential checkpoint blockade therapy, boxplots were drawn to show the differences in immune checkpoint gene expression between the high- and low-risk groups (Figure 8A). Checkpoint genes, including *CTLA4, LAG3, ID O 2, CD274*, and *PDCD1*, had higher expression levels in low-risk groups. In Figure 8B, cluster1 (C1), C2, C3, and C4 represent wound healing, IFN-gamma dominant, inflammatory, and lymphocyte depleted immune subgroups, respectively (Thorsson et al., 2019). The results showed that C3 samples were almost equally distributed between the two groups, but there were more C1 and C4 samples and fewer C2 samples in the high-risk subgroup than in the low-risk subgroup. Violin plots showed the relationship between IPSs and risk groups; a higher IPS represented a better response to *PD-1* and *CTLA-4* blockers (Figure 8C). We also found that eight drugs had lower IC50 values in the high-risk group, including bexarotene, bicalutamide, dasatinib, doxetacel, elesclomol, imatinib, midostaurin, and pazopanib (Figure 8D).

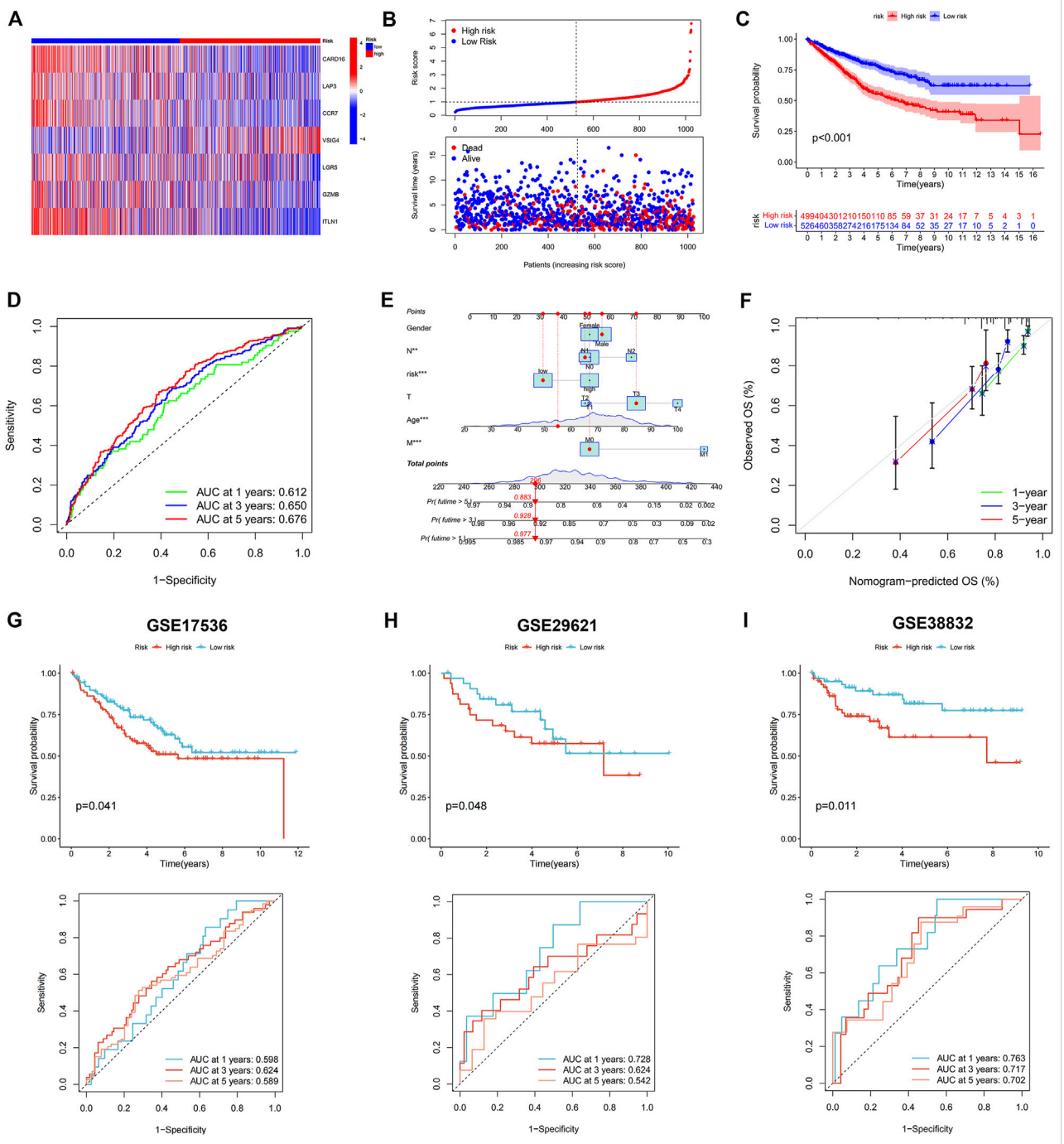


FIGURE 5
Construction and validation of the prognostic signature. **(A)** Heatmap showed the expression of 7 genes in two risk groups. **(B)** Risk score and survival outcome of each case. **(C)** KM curve showed that patients in high-risk group had a worse prognosis. **(D)** The AUC for 1-, 3- and 5-years survival were 0.612, 0.650, and 0.676, respectively. **(E)** Nomogram using risk score and other clinical features were constructed for predicting survival of colon cancer patients. **(F)** Calibration graphs investigated that the actual survival rates of colon cancer patients were close to the nomogram-predicted survival rates. The KM and ROC methods were used to evaluate the efficiency of the risk score at predicting patient survival in GSE17536 **(G)**, GSE29621 **(H)**, and GSE38832 **(I)** CRC datasets.

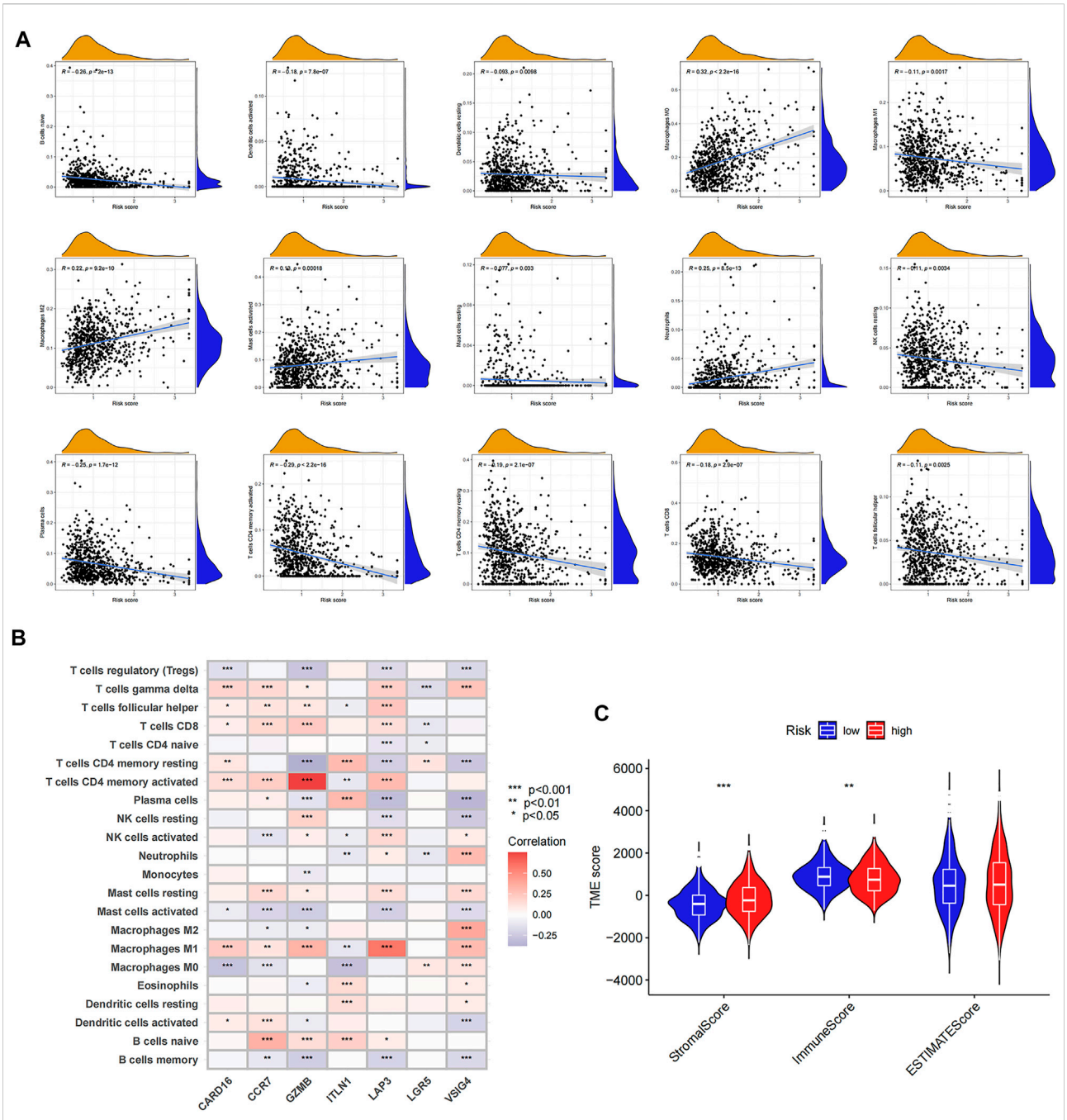


FIGURE 6
Evaluation of tumor microenvironment in high- and low-risk groups. **(A)** Relationship between risk score and different immune cell types. **(B)** Correlation between the abundance of immune cells and seven genes in the prognostic signature. **(C)** Correlation between risk score and immune-related scores. *$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$.

Validating expression levels of LGR5, VSIG4, GZMB, and ITLN1 *via* quantitative real-time polymerase chain reaction

Among the seven genes in the prognostic signature, *LGR5*, *VSIG4*, *GZMB*, and *ITLN1* were significantly differentially expressed in colon cancer samples from GEPIA database (Supplementary Figure S6). Expression levels of *LGR5*, *VSIG4*, *GZMB*, and *ITLN1* were tested in colon cancer and adjacent normal tissues *via* qRT-PCR method. Expression of *LGR5* was significantly higher in tumor tissues (Figure 9A) while *VSIG4* had higher expression levels in normal tissues ($p < 0.05$) (Figure 9B).

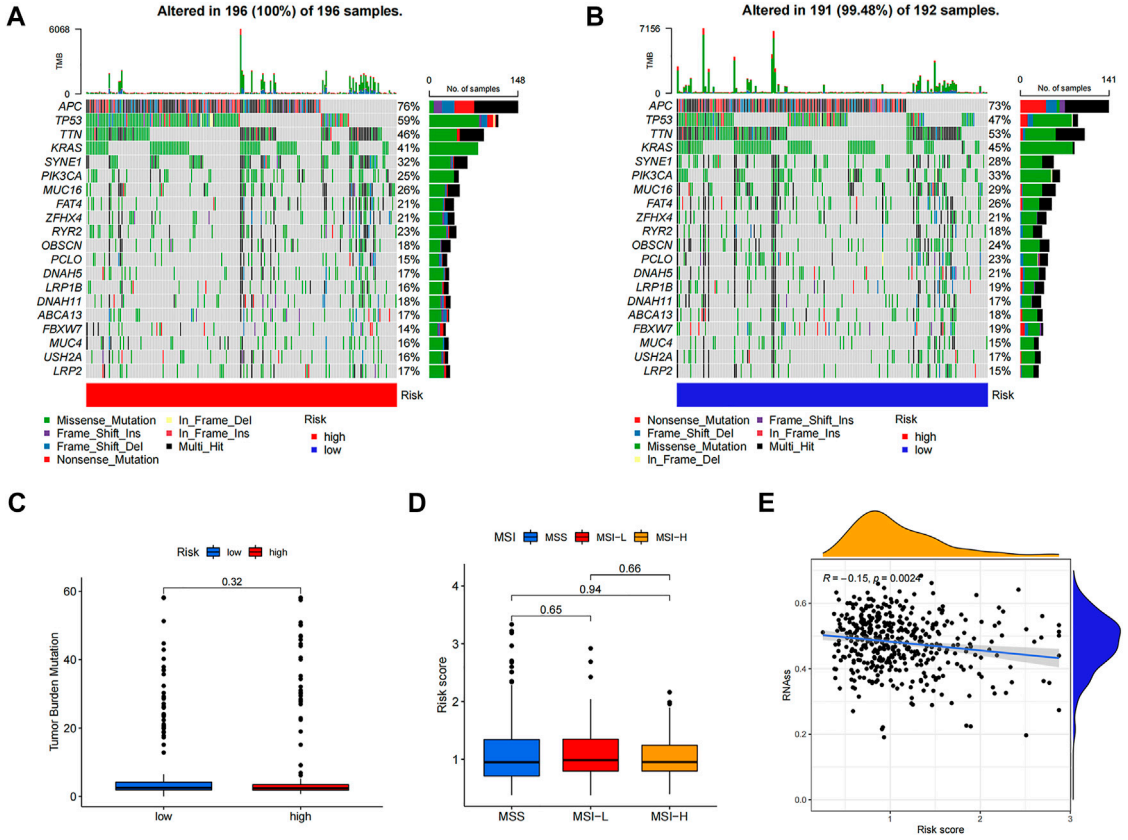


FIGURE 7
Comprehensive analyses of risk score in colon cancer. The somatic gene mutations in high-risk group **(A)** and low-risk group **(B)**. TMB **(C)** and MSI **(D)** did not show significant correlations with risk score while CSC **(E)** negatively correlated with risk score.

There was no significant differences in the expression levels of GZMB (Figure 9C) and ITLN1 (Figure 9D) between normal and tumor samples.

Discussion

Cell death usually does not occur independently but in a mixed form because cells can undergo extensive crosstalk under pathological conditions (Zheng and Kanneganti, 2020; Karki et al., 2021b). Previous research (Karki et al., 2021b) suggests that there exists a mixed form of cell death involving pyroptosis, apoptosis, and necroptosis, called PANoptosis. In recent years, many studies have investigated the effects of different forms of cell death on various human diseases, especially malignant tumors. Some studies have revealed molecular classifications of tumors and constructed prognostic models based on genes or non-coding RNAs relevant to different forms of cell death. However, the effects of PANoptosis in colon cancer have not been well studied.

In this study, 19 PRGs were identified. In previous studies, most of these 19 PRGs were found to be related to CRC. Yin et al. (2010) suggested that overexpression of the exogenous *FADD* gene can significantly improve the apoptosis-inducing effect of 5-fluorouracil on colorectal adenocarcinoma cells. Shi et al. (2021) demonstrated that low *NLPR3* expression is related to a better prognosis of CRC. Li et al. (2020) showed that *TAB3* was upregulated in CRC tissues and promoted CRC cell growth. It was also found that *AIM2* inhibits CRC cell proliferation and migration (Xu et al., 2020). Colon cancer cases from the TCGA and GEO databases were divided into two distinct PRG clusters. PRG cluster A had a better prognosis than PRG cluster B. Tumor infiltration and lymph node metastasis were correlated with the PRG clusters. Results of GSVA and ssGSEA showed that PRG cluster A was significantly enriched in immune-related pathways and had higher immune cell infiltration levels. Tumor-infiltrating immune cells can affect the response to anti-checkpoint blockade. Furthermore, tumor-infiltrating CD4 + T-cells can upregulate programmed cell death protein 1 (*PD-1*), T-cell immunoglobulin and mucin domain-3 (*TIM-3*), cytotoxic T lymphocyte-associated protein-4 (*CTLA-4*), and lymphocyte-activation-gene-3 (*LAG-3*) (Toor et al., 2019). PRDEGs between

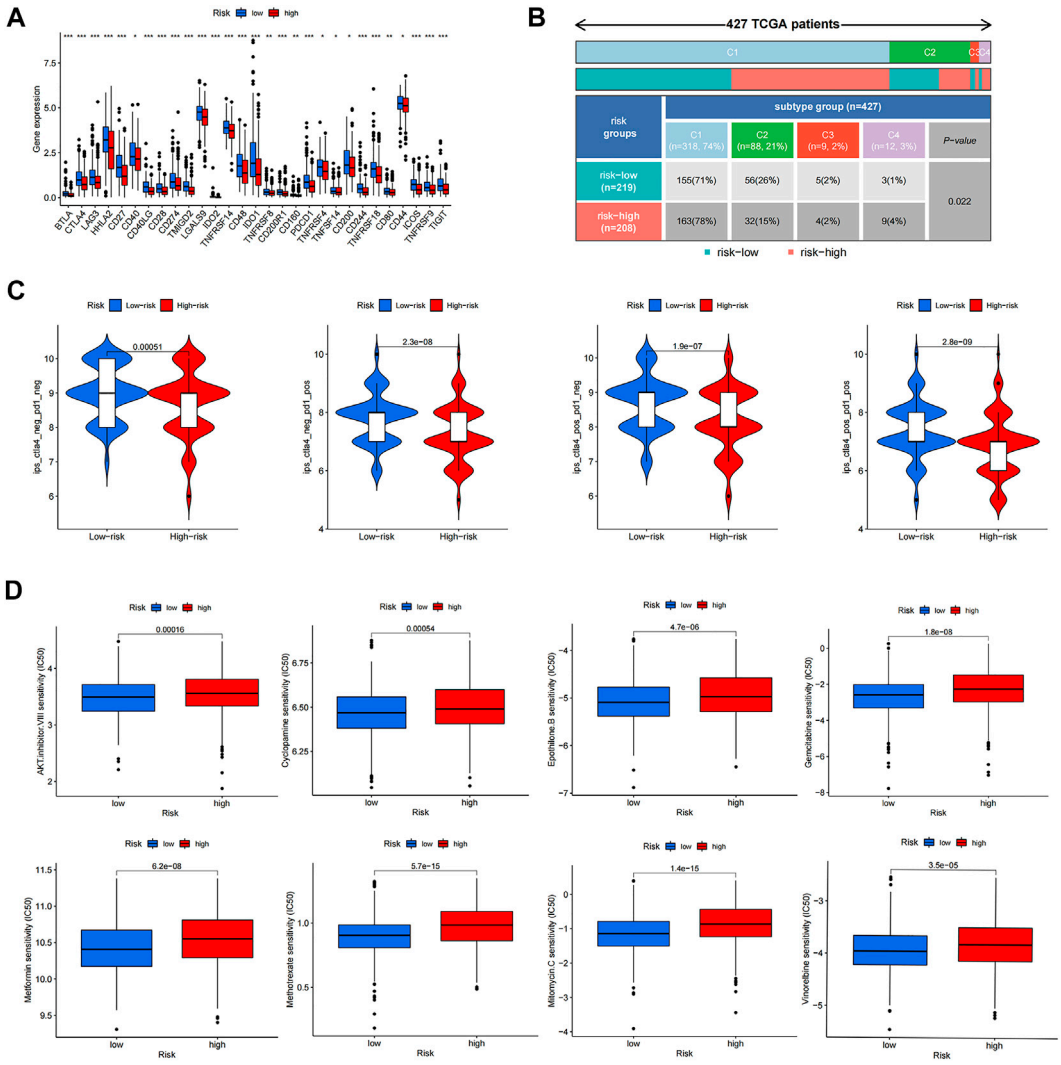


FIGURE 8
Response to anti-tumor tharapy of colon cancer patients in two risk groups. **(A)** The differences of immune checkpoint gene expression in high-risk and low-risk groups. **(B)** Heatmap and table showing the distribution of colon cancer immune subtypes between two risk groups. **(C)** Violin plots showed the relationship between IPSs and risk groups. **(D)** Eight therapeutic drugs showed significant IC50 differences. *$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$.

two PRG clusters were also identified, and patients were classified into two distinct clusters. GO and KEGG analyses revealed that these PRDEGs were associated with certain cancer-related biological functions and pathways, indicating that these PRDEGs were potentially associated with malignant tumors. Gene cluster A had a longer survival time than gene cluster B, and the two clusters were correlated with tumor infiltration and metastasis.

LASSO and multivariate Cox regression analyses were used to screen genes to construct a prognostic signature. Finally, the risk score was calculated based on the expression levels of *CARD16, LAP3, CCR7, VSIG4, LGR5, GZMB*, and *ITLN1*. Some of these seven genes have been found to be associated with various types of malignant tumors. *LAG3* can promote glioma progression by regulating the proliferation, migration, and invasion of glioma cells (He et al., 2015), and inhibition of *LAG3* suppresses the invasion of ovarian cancer (Wang et al., 2015). Bill et al. (2022) suggested that *CCR7* plays distinct roles in directing tumor cells to the lymph nodes, skin, and central nervous system. Zhu et al. (2018) demonstrated that downregulated VSIG4 expression was related to poor prognosis in hepatocellular carcinoma patients with hepatitis B infection. *LGR5* has been identified as a strong cancer stem cell biomarker in CRC (Kamakura et al., 2022). These results suggest that these seven genes could serve as potential biomarkers for cancer diagnosis and therapy. Patients were classified into high- and low-risk groups based on the risk score, and the KM curve showed that the prognosis of patients in the low-risk group was much better than

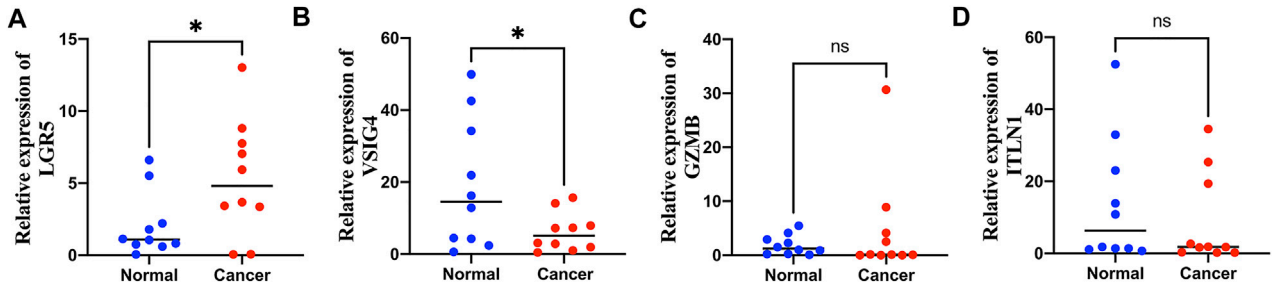


FIGURE 9
Quantitative real-time polymerase chain reaction (qRT-PCR) analyses of LGR5 **(A)**, VSIG4 **(B)**, GZMB **(C)** and ITLN1 **(D)** expression in 10 pairs of colon cancer tissues and adjacent non-cancer tissues. *$p < 0.05$; ns $p > 0.05$.

that of patients in the high-risk group. ROC analyses were performed to test the prediction efficiency of the risk score. Nomograms are widely used as prediction tools in oncology, particularly for survival prediction (Iasonos et al., 2008; Balachandran et al., 2015). A nomogram model was established according to the risk score and other clinical characteristics to accurately predict the survival time of patients, and the calibration plots showed that the actual survival rates were close to the nomogram-predicted survival rates. This indicated that the nomogram model had high accuracy in predicting patient survival.

The correlation between the risk score and immune cells was also analyzed; four types of immune cells were positively related to the risk score and the other 11 types of immune cells were negatively correlated with the risk score. The seven genes also showed significant associations with various types of immune cells. Dai et al. (2020) reported that an immune score based on immunogenomic analysis can indicate the efficacy of immunotherapy and chemotherapy. The high-risk group had higher stromal and lower immune scores, suggesting that the low-risk group might have a better response to antitumor therapy. CSCs are a subset of tumor cells associated with tumor metastasis, recurrence, and drug resistance. Similar to normal stem cells, CSCs exhibit self-renewal and differentiation abilities (Singh and Chellappan, 2014). The risk score was also related to the CSC index, indicating that the risk score may be related to colon cancer progression. The differences in immune checkpoint gene expression in the high-risk and low-risk groups were also analyzed, and the expression levels of checkpoints, including *CTLA4, LAG3, ID O 2, CD274*, and *PDCD1*, were found to be higher in the low-risk group. The correlation between risk groups and previously identified immune subtypes was analyzed; the results showed that the inflammatory samples were almost equally distributed between the two groups, but there was more wound healing and lymphocyte depletion and fewer IFN-gamma-dominant samples in the high-risk group than in the low-risk group. The IPSs of the two risk groups suggested that the low-risk group had a better response to *PD-1* and *CTLA-4* blockade therapy. IC50 values indicated that the low-risk group was more sensitive to immunotherapeutic and chemotherapeutic drugs, and the results confirmed our previous conclusion based on TME-related analyses. The findings of our study can be applied to guide clinical immunotherapy and chemotherapy in patients with colon cancer and help us to further understand the effects of PANoptosis on colon cancer. The expression levels of *LGR5*, *VSIG4*, *GZMB*, and *ITLN1* were further validated using qRT-PCR method, the results showed that *LGR5* was significantly upregulated in colon cancer while *VSIG4* was downregulated in colon cancer compared with normal tissues, indicating that *LGR5* and *VSIG4* may be potential biomarkers for diagnosis and therapy in colon cancer.

Nevertheless, our study has some limitations. Most analyses were based on data from public datasets, and all samples were obtained retrospectively, which may have caused an inherent case selection bias. In addition, limited molecular biology experiments were performed in the study, and further *in vitro* and *in vivo* experiments are needed to validate our findings. Finally, some valuable clinical features such as surgery, neoadjuvant chemotherapy, and tumor markers were not considered in our study. As such, clinical cases are needed to confirm our conclusions.

In summary, we constructed a PANoptosis-based molecular clustering and prognostic signature that plays a vital role in predicting survival, TMB, and guiding clinical therapy. The findings of this study may improve our understanding of PANoptosis in colon cancer and help develop more effective treatment strategies. However, this study has some limitations, and additional experiments and clinical cases are needed to validate our findings.

Data availability statement

The data presented in the study are deposited in the TCGA-COAD project and GEO database (accession number: GSE39582).

Author contributions

XW, RS, and SC are responsible for writing and submitting the manuscript. LM, YX, XZ, ZW, and XH are responsible for data collection and analysis. QH, LD, TB, ZY ,and MW are responsible for the production of pictures. WY, HZ, and WC are responsible for the ideas and guidance. All authors have read and approved the final manuscript.

Funding

The work was supported by the National Natural Science Foundation of China (No.81670517 and 81870402) and Research Fund of Anhui Institute of Translational Medicine (2021zhyx-C30).

Acknowledgments

We acknowledged TCGA and GEO database for providing their platform 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.2022.955355/full#supplementary-material

References

Balachandran, V. P., Gonen, M., Smith, J. J., and DeMatteo, R. P. (2015). Nomograms in oncology: More than meets the eye. *Lancet. Oncol.* 16 (4), e173–80. doi:10.1016/S1470-2045(14)71116-7

Bertheloot, D., Latz, E., and Franklin, B. S. (2021). Necroptosis, pyroptosis and apoptosis: An intricate game of cell death. *Cell. Mol. Immunol.* 18 (5), 1106–1121. doi:10.1038/s41423-020-00630-3

Bhandari, A., Woodhouse, M., and Gupta, S. (2017). Colorectal cancer is a leading cause of cancer incidence and mortality among adults younger than 50 years in the USA: A SEER-based analysis with comparison to other young-onset cancers. *J. Investig. Med.* 65 (2), 311–315. doi:10.1136/jim-2016-000229

Bill, C. A., Allen, C. M., and Vines, C. M. (2022). C-C chemokine receptor 7 in cancer. *Cells* 11 (4), 656. doi:10.3390/cells11040656

Briard, B., Malireddi, R. K. S., and Kanneganti, T. D. (2021). Role of inflammasomes/pyroptosis and PANoptosis during fungal infection. *PLoS Pathog.* 17 (3), e1009358. doi:10.1371/journal.ppat.1009358

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

Dai, G. P., Wang, L. P., Wen, Y. Q., Ren, X. Q., and Zuo, S. G. (2020). Identification of key genes for predicting colorectal cancer prognosis by integrated bioinformatics analysis. *Oncol. Lett.* 19 (1), 388–398. doi:10.3892/ol.2019.11068

Doonan, B. B., Schaafsma, E., Pinto, J. T., Wu, J. M., and Hsieh, T. C. (2017). Application of open-access databases to determine functional connectivity between resveratrol-binding protein QR2 and colorectal carcinoma. *Vitro Cell. Dev. Biol. Anim.* 53 (7), 575–578. doi:10.1007/s11626-017-0174-x

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 (7784), 683–687. doi:10.1038/s41586-019-1770-6

Galluzzi, L., Vitale, I., Aaronson, S. A., Abrams, J. M., Adam, D., Agostinis, P., et al. (2018). Molecular mechanisms of cell death: Recommendations of the nomenclature committee on cell death 2018. *Cell. Death Differ.* 25 (3), 486–541. doi:10.1038/s41418-017-0012-4

Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discov.* 12 (1), 31–46. doi:10.1158/2159-8290.CD-21-1059

He, X., Huang, Q., Qiu, X., Liu, X., Sun, G., Guo, J., et al. (2015). LAP3 promotes glioma progression by regulating proliferation, migration and invasion of glioma cells. *Int. J. Biol. Macromol.* 72, 1081–1089. doi:10.1016/j.ijbiomac.2014.10.021

Iasonos, A., Schrag, D., Raj, G. V., and Panageas, K. S. (2008). How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* 26 (8), 1364–1370. doi:10.1200/JCO.2007.12.9791

Jiang, M., Qi, L., Li, L., Wu, Y., Song, D., and Li, Y. (2021). Caspase-8: A key protein of cross-talk signal way in "PANoptosis" in cancer. *Int. J. Cancer* 149 (7), 1408–1420. doi:10.1002/ijc.33698

Kamakura, M., Uehara, T., Iwaya, M., Asaka, S., Kobayashi, S., Nakajima, T., et al. (2022). LGR5 expression and clinicopathological features of the invasive front in the fat infiltration area of pancreatic cancer. *Diagn. Pathol.* 17 (1), 21. doi:10.1186/s13000-022-01203-w

Karki, R., Sharma, B. R., Lee, E., Banoth, B., Malireddi, R. K. S., Samir, P., et al. (2020). Interferon regulatory factor 1 regulates PANoptosis to prevent colorectal cancer. *JCI Insight* 5 (12), e136720. doi:10.1172/jci.insight.136720

Karki, R., Sharma, B. R., Tuladhar, S., Williams, E. P., Zalduondo, L., Samir, P., et al. (2021). Synergism of TNF-α and IFN-γ triggers inflammatory cell death, tissue damage, and mortality in SARS-CoV-2 infection and cytokine shock syndromes. *Cell.* 184 (1), 149–168.e17. e17. doi:10.1016/j.cell.2020.11.025

Karki, R., Sundaram, B., Sharma, B. R., Lee, S., Malireddi, R. K. S., Nguyen, L. N., et al. (2021). ADAR1 restricts ZBP1-mediated immune response and PANoptosis to promote tumorigenesis. *Cell. Rep.* 37 (3), 109858. doi:10.1016/j.celrep.2021.109858

Kesavardhana, S., Malireddi, R. K. S., and Kanneganti, T. D. (2020). Caspases in cell death, inflammation, and pyroptosis. *Annu. Rev. Immunol.* 38, 567–595. doi:10.1146/annurev-immunol-073119-095439

Kouo, T., Huang, L., Pucsek, A. B., Cao, M., Solt, S., Armstrong, T., et al. (2015). Galectin-3 shapes antitumor immune responses by suppressing CD8+ T cells via LAG-3 and inhibiting expansion of plasmacytoid dendritic cells. *Cancer Immunol. Res.* 3 (4), 412–423. doi:10.1158/2326-6066.CIR-14-0150

Lee, S., Karki, R., Wang, Y., Nguyen, L. N., Kalathur, R. C., and Kanneganti, T. D. (2021). AIM2 forms a complex with pyrin and ZBP1 to drive PANoptosis and host defence. *Nature* 597 (7876), 415–419. doi:10.1038/s41586-021-03875-8

Li, Q., Chen, L., Luo, C., ChenYanGe, J., Zhu, Z., et al. (2020). TAB3 upregulates PIM1 expression by directly activating the TAK1-STAT3 complex to promote colorectal cancer growth. *Exp. Cell. Res.* 391 (1), 111975. doi:10.1016/j.yexcr.2020.111975

Malireddi, R. K. S., Kesavardhana, S., and Kanneganti, T. D. (2019). ZBP1 and TAK1: Master regulators of NLRP3 inflammasome/pyroptosis, apoptosis, and necroptosis (PAN-optosis). *Front. Cell. Infect. Microbiol.* 9, 406. doi:10.3389/fcimb.2019.00406

Malireddi, R. K. S., Tweedell, R. E., and Kanneganti, T. D. (2020). PANoptosis components, regulation, and implications. *Aging (Albany NY)* 12 (12), 11163–11164. doi:10.18632/aging.103528

Nguyen, L. N., and Kanneganti, T. D. (2022). PANoptosis in viral infection: The missing puzzle piece in the cell death field. *J. Mol. Biol.* 434 (4), 167249. doi:10.1016/j.jmb.2021.167249

Nie, J., Shan, D., Li, S., Zhang, S., Zi, X., Xing, F., et al. (2021). A novel ferroptosis related gene signature for prognosis prediction in patients with colon cancer. *Front. Oncol.* 11, 654076. doi:10.3389/fonc.2021.654076

Place, D. E., Lee, S., and Kanneganti, T. D. (2021). PANoptosis in microbial infection. *Curr. Opin. Microbiol.* 59, 42–49. doi:10.1016/j.mib.2020.07.012

Russo, A. G., Andreano, A., Sartore-Bianchi, A., Mauri, G., Decarli, A., and Siena, S. (2019). Increased incidence of colon cancer among individuals younger than 50 years: A 17 years analysis from the cancer registry of the municipality of milan, Italy. *Cancer Epidemiol.* 60, 134–140. doi:10.1016/j.canep.2019.03.015

Samir, P., Malireddi, R. K. S., and Kanneganti, T. D. (2020). The PANoptosome: A deadly protein complex driving pyroptosis, apoptosis, and necroptosis (PANoptosis). *Front. Cell. Infect. Microbiol.* 10, 238. doi:10.3389/fcimb.2020.00238

Shi, F., Wei, B., Lan, T., Xiao, Y., Quan, X., Chen, J., et al. (2021). Low NLRP3 expression predicts a better prognosis of colorectal cancer. *Biosci. Rep.* 41 (4), BSR20210280. doi:10.1042/BSR20210280

Singh, S., and Chellappan, S. (2014). Lung cancer stem cells: Molecular features and therapeutic targets. *Mol. Asp. Med.* 39, 50–60. doi:10.1016/j.mam.2013.08.003

Song, W., Ren, J., Xiang, R., Kong, C., and Fu, T. (2021). Identification of pyroptosis-related subtypes, the development of a prognosis model, and characterization of tumor microenvironment infiltration in colorectal cancer. *Oncoimmunology* 10 (1), 1987636. doi:10.1080/2162402X.2021.1987636

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

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2019). The immune landscape of cancer. *Immunity* 51 (2), 411–412. doi:10.1016/j.immuni.2019.08.004

Toor, S. M., Murshed, K., Al-Dhaheri, M., Khawar, M., Abu Nada, M., and Elkord, E. (2019). Immune checkpoints in circulating and tumor-infiltrating CD4+ T cell subsets in colorectal cancer patients. *Front. Immunol.* 10, 2936. doi:10.3389/fimmu.2019.02936

Wang, X., Shi, L., Deng, Y., Qu, M., Mao, S., Xu, L., et al. (2015). Inhibition of leucine aminopeptidase 3 suppresses invasion of ovarian cancer cells through down-regulation of fascin and MMP-2/9. *Eur. J. Pharmacol.* 768, 116–122. doi:10.1016/j.ejphar.2015.10.039

Wang, Y., Lin, K., Xu, T., Wang, L., Fu, L., Zhang, G., et al. (2021). Development and validation of prognostic model based on the analysis of autophagy-related genes in colon cancer. *Aging (Albany NY)* 13 (14), 19028–19047. doi:10.18632/aging.203352

Xu, M., Wang, J., Li, H., Zhang, Z., and Cheng, Z. (2020). AIM2 inhibits colorectal cancer cell proliferation and migration through suppression of Gli1. *Aging (Albany NY)* 13 (1), 1017–1031. doi:10.18632/aging.202226

Yin, A., Jiang, Y., Zhang, X., and Luo, H. (2010). Overexpression of FADD enhances 5-fluorouracil-induced apoptosis in colorectal adenocarcinoma cells. *Med. Oncol.* 27 (2), 397–405. doi:10.1007/s12032-009-9224-x

Zhao, Z., Liu, H., Zhou, X., Fang, D., Ou, X., Ye, J., et al. (2021). Necroptosis-related lncRNAs: Predicting prognosis and the distinction between the cold and hot tumors in gastric cancer. *J. Oncol.* 2021, 6718443. doi:10.1155/2021/6718443

Zheng, M., and Kanneganti, T. D. (2020). The regulation of the ZBP1-NLRP3 inflammasome and its implications in pyroptosis, apoptosis, and necroptosis (PANoptosis). *Immunol. Rev.* 297 (1), 26–38. doi:10.1111/imr.12909

Zhu, S., Tan, W., Li, W., Zhou, R., Wu, X., Chen, X., et al. (2018). Low expression of VSIG4 is associated with poor prognosis in hepatocellular carcinoma patients with Hepatitis B infection. *Cancer Manag. Res.* 10, 3697–3705. doi:10.2147/CMAR.S165822

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fgene.2022.927142

Check for updates

OPEN ACCESS

EDITED BY
Yunpeng Cai,
Shenzhen Institutes of Advanced Technology (CAS), China

REVIEWED BY
Siyi Wanggou,
Xiangya Hospital, Central South University, China
Quan Cheng,
Xiangya Hospital, Central South University, China

*CORRESPONDENCE
Huaan Cai,
601244097@qq.com
Fan Zhang,
zhangfan0913@csu.edu.cn

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

RECEIVED 23 April 2022
ACCEPTED 31 August 2022
PUBLISHED 20 September 2022

CITATION
Huang L, Zhang J, Gong F, Han Y, Huang X, Luo W, Cai H and Zhang F (2022), Identification and validation of ferroptosis-related lncRNA signatures as a novel prognostic model for glioma. *Front. Genet.* 13:927142. doi: 10.3389/fgene.2022.927142

COPYRIGHT
© 2022 Huang, Zhang, Gong, Han, Huang, Luo, Cai 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.

Identification and validation of ferroptosis-related lncRNA signatures as a novel prognostic model for glioma

Liang Huang[1], Juan Zhang[1], Fanghua Gong[2], Yuhua Han[3], Xing Huang[4], Wanxiang Luo[1], Huaan Cai[1]* and Fan Zhang[1]*

[1]Department of Rehabilitation Medicine, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, Hunan, China, [2]Department of Nursing, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, Hunan, China, [3]Department of Cadre Health Care, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, Hunan, China, [4]Department of General Surgery, Hunan Provincial People's Hospital, The First Affiliated Hospital of Hunan Normal University, Changsha, Hunan, China

Background: Ferroptosis is a newly discovered form of regulated cell death with distinct properties and recognizing functions involved in physical conditions or various diseases, including cancers. However, the relationship between gliomas and ferroptosis-related lncRNAs (FRLs) remains unclear.

Methods: We collected a total of 1850 samples from The Cancer Genome Atlas (TCGA) and Genotype Tissue Expression (GTEX) databases, including 698 tumor and 1,152 normal samples. A list of ferroptosis-related genes was downloaded from the Ferrdb website. Differentially expressed FRLs (DEFRLS) were analyzed using the "limma" package in R software. Subsequently, prognosis-related FRLs were obtained by univariate Cox analysis. Finally, a prognostic model based on the 3 FRLs was constructed using Cox regression analysis with the least absolute shrinkage and selection operator (LASSO) algorithm. The prognostic power of the model was assessed using receiver operating characteristic (ROC) curve analysis and Kaplan-Meier (K-M) survival curve analysis. In addition, we further explored the relationship of the immune landscape and somatic mutations to prognostic model characteristics. Finally, we validated the function of LINC01426 *in vitro*.

Results: We successfully constructed a 3-FRLs signature and classified glioma patients into high-risk and low-risk groups based on the risk score calculated from this signature. Compared with traditional clinicopathological features [age, sex, grade, isocitrate dehydrogenase (IDH) status], the prognostic accuracy of this model is more stable and stronger. Additionally, the model had stable predictive power for overall survival over a 5-year period. In addition, we found significant differences between the two groups in cellular immunity, the numbers of many immune cells, including NK cells, $CD4^+$, $CD8^+$ T-cells, and macrophages, and the expression of many immune-related genes. Finally, the two groups were also significantly different at the level of somatic mutations, especially in glioma prognosis-related genes such as IDH1 and ATRX, with lower mutation rates in the high-risk group leading to poorer prognosis. Finally, we

found that the ferroptosis process of glioma cells was inhibited after knocking down the expression of LINC01426.

Conclusion: The proposed 3-FRL signature is a promising biomarker for predicting prognostic features in glioma patients.

KEYWORDS

lncRNA, ferroptosis, glioma, prognostic signature, immune microenvironment

1 Introduction

Gliomas are the most common primary malignant brain tumors in adults, mainly in the brain and in glial tissue (Ostrom et al., 2013), accounting for 81% of malignant brain tumors. Although relatively rare relative to other cancers, they cause significant mortality and morbidity (Ostrom et al., 2014). Glioblastoma (GBM) is the most common and most clinically aggressive World Health Organization (WHO) grade IV glioma, with the highest degree of malignancy, the worst prognosis, and the lowest overall survival (OS) rate. The median OS of GBM is approximately 8 months, and the 5-year survival rate is 7.2% (Ostrom et al., 2020). Even with aggressive multimodal therapy, the median survival is only 12–15 months (Stupp et al., 2005). Glioma is also a highly heterogeneous tumor (Nicholson and Fine 2021; van den Bent et al., 2009), which makes it difficult to determine its prognostic effect and treatment response when treating glioma. Therefore, finding a biomarker and possible therapeutic target that can predict prognosis is crucial.

Ferroptosis is a novel cell death method first proposed in 2012 that is distinct from autophagy and apoptosis. Ferroptosis can be triggered by depleting the amino acid cysteine in the cell or by inhibiting the phospholipid hydroperoxidase glutathione peroxidase 4 (GPX4) (Dixon 2017). Ferroptosis is characterized by membrane lipid peroxidation in cells, which eventually leads to the loss of selective permeability of the plasma membrane and the occurrence of oxidative stress (Mou et al., 2019), resulting in rupture of the outer mitochondrial membrane, reduction or disappearance of the mitochondrial cristae, and condensation of the mitochondrial membrane, resulting in cell death (Xie et al., 2016). Recently, ferroptosis has also been proven to be involved in cancer immunotherapy. Due to its nonapoptotic nature, ferroptosis-based cancer therapy is expected to remedy the shortcomings of traditional therapies mediated by the apoptotic pathway (Liang et al., 2019). Therefore, screening ferroptosis-related genes (FRGs) based on clinical samples is beneficial for the diagnosis of glioma and provides possible therapeutic targets.

Long noncoding RNA (lncRNA) refers to a type of noncoding RNA more than 200 nucleotides in length. lncRNAs are involved in a wide range of cellular mechanisms, from almost all aspects of gene expression to protein translation and stability (Schmitz et al., 2016). Subsequent studies found that lncRNAs are dysregulated in tumors (Srikantan et al., 2000; Ji et al., 2003; Diederichs 2014). With the continuous in-depth understanding of lncRNAs, researchers have successively discovered the effect of lncRNAs on cancer, for example, lncRNAs can change epigenetics in glioma (Pop et al., 2018). Recent studies have found that lncRNAs can control the occurrence and development of tumors by affecting the process of ferroptosis. LINC00336 as a competing endogenous RNA inhibits ferroptosis in lung cancer (Wang et al., 2019; Mao et al., 2019), and the lncRNA GABPB1-AS1 regulates erastin-induced ferroptosis through GABPB1 in HepG2 hepatocellular carcinoma (Qi et al., 2019). Furthermore, there are 3 different ferroptosis-related lncRNAs(FRL) signatures were observed to be associated with glioma prognosis, containing 15 ferroptosis-related lncRNAs, 14 ferroptosis-related lncRNAs and 9 ferroptosis-related lncRNAs, respectively (He et al., 2021; Zheng et al., 2021; Shi et al., 2022). In machine learning models, a consistent cutoff value for different datasets enhances generalizability, and thereby increasing applicability in real world. Unfortunately, the previous studies didn't explore the best cutoff value for different datasets, which might lead to potential false positive results. Moreover, gene mutation has been proved to be an important factor affecting the survival and prognosis of glioma patients (Suzuki et al., 2015; Arita et al., 2020). For example, IDH mutation status has been shown to be closely related to the prognosis of glioma patients (Pirozzi and Yan 2021). The previous studies have not explored this aspect, too.

In this study, we obtained RNA-seq data from TCGA and GTEx databases and finally obtained three differentially expressed FRLs (DEFRLS) for constructing prognostic models. Then, the reliability of the model was verified by survival analysis, receiver operating characteristic (ROC) curve analysis and independent prognostic analysis. In addition, the mechanism of action of FRLs in glioma was further explored by gene set enrichment analysis (GSEA), mutated gene analysis, immune infiltration analysis and chemotherapeutic drug sensitivity analysis. Finally, our results provide a good predictive model and possible therapeutic targets for glioma patients.

2 Materials and methods

2.1 Data acquisition

A total of 1,850 samples from gliomas (GBM, LGG) in The Cancer Genome Atlas (TCGA) website (https://portal.gdc.

cancer.gov/) and Genotype-Tissue Expression (GTEx) website (https://www.gtexportal.org) were collected, including 698 tumors and 1,152 normal samples. Then, the data were $\log_2$-processed, and Ensembl IDs were converted to official gene symbols. lncRNAs and protein-coding genes were screened by the Genome Reference Consortium Human Build 38 (GRCh38).

2.2 Identification of ferroptosis-related lncRNAs

The ferroptosis-related dataset (FerrDb) was obtained from FerrDb (http://www.zhounan.org/ferrdb/index.html) website, resulting in a total of 176 validated human FRGs. Subsequently, Spearman correlation analysis ($|R^2| > 0.6$ and p value < 0.001) was performed according to the expression profiles of FRGs and lncRNAs, and 433 FRLs were obtained.

2.3 Differential expression analysis

The limma package (Ritchie et al., 2015) was used to perform differential analysis on the lncRNA expression matrix of LGG/GBM and normal samples, and a total of 2056 differentially expressed lncRNAs (DELs) were obtained. The criteria for DElncRNAs were $|\log_2$ (fold change) $|$ >1 and a false discovery rate (FDR) < 0.05 (Tu et al., 2020).

2.4 Construction of ferroptosis-related prognostic signature

A total of 433 FRLs intersected with 2056 DElncRNAs, and 52 lncRNAs were ultimately obtained. Then, univariate Cox analysis was performed based on the "survival" R package to define potential prognostic FRLs ($p < 0.001$), and a total of 35 prognosis-related lncRNAs were obtained. A total of 611 patients were randomly divided into training or validation groups in a 1:1 ratio. Subsequently, these prognostic candidates were included in least absolute shrinkage and selection operator (LASSO)-Cox regression analysis. Finally, by choosing the optimal penalty parameter λ associated with a minimum 10-fold cross-validation to construct the prognostic FRLS, we established a three-gene optimal prognostic model. The ferroptosis-related prognostic risk score for each patient was formulated as follows:

$$\text{Risk score} = \sum_1^n coef_i * x_i$$

where x_i and $coef_i$ represent the expression of each lncRNA and its corresponding coefficient, respectively. Based on the median risk score, we divided the training cohort patients into high-risk and low-risk groups. Kaplan-Meier curves were generated using the "survminer" R package with the log-rank test to compare OS between the high- and low-risk groups. ROC curve analysis was used to evaluate the prediction accuracy of FRLS by the R package "timeROC." To assess the model feasibility, all validations were performed simultaneously in the training and validation cohorts.

2.5 Functional enrichment analysis

Differentially expressed genes (DEGs) ($|\log_2$ (fold change)$| >$ 1 and FDR<0.05) between the high-risk and low-risk groups were identified using the "edgeR" (Robinson et al., 2010) R package and functionally annotated based on Gene Ontology (GO) and with the "clusterProfiler" (Wu et al., 2021).

R package of Kyoto Encyclopedia of Genes and Genomes (KEGG) (adjusted p value < 0.05).

2.6 Gene set enrichment analysis

To explore molecular and biological differences in high/low risk groups, the KEGG and HALLMARK gene sets in the Molecular Signature Database (https://www.gsea-msigdb.org/GSEA/Msigdb) were obtained. Gene set enrichment analysis (GSEA) between the two groups was performed by the "ClusterProfiler" R package ($p < 0.05$ and FDR < 0.25) (Subramanian et al., 2005). Subsequently, single-sample GSEA (ssGSEA) was performed on several representative genomes by the "GSVA" R package.

2.7 Assessment of immune cell infiltration and immune microenvironment

Immune infiltration in glioma patients was assessed using the ESTIMATE algorithm by the R package "estimate." The 22 immune cell subsets obtained from the CIBERSORT portal (http://CIBERSORT.stanford.edu/) were defined using CIBERSORT's LM22, and the differences in the infiltration of 22 immune cells were subsequently assessed using the CIBERSORT algorithm. Finally, Pearson correlation analysis was used to calculate the differences in the expression levels of immune cell markers between the two groups of patients (Tan et al., 2020).

2.7.1 Cell lines

Normal human astrocytes (HA 1800) and GBM cell lines U251, LN229, KNS-89, and T98G were purchased from .Cell lines were cultured in standard culture conditions (37°C, 95% humidity, 5% CO_2) in the culture in DMEM (Gibco BRL, United States) medium containing 10% fetal bovine serum (Gibco BRL, United States).

2.7.2 Construct stable cell lines

The shRNA targeting LINC01426 (sh#1, sh#2) from Cao et al. (2020).The sh#1, sh#2 and negative control (NC) viruses were obtained from Tsingke Biotech (Tsingke, China). Subsequently, U251 and KNS-89 were infected and selected after 48 h with 2 μg/ml puromycin (cat# A1113803, Thermo Fisher).

2.7.3 RNA extraction and quantitative real-time polymerase chain reaction

Total RNAs were extracted from cells by (cat# AG21024, Accurate Biology, China) following the manufacturer's instructions. 1,000 ng of total RNA was reverse transcribed with (cat# 11139ES10, Yeasen). Gene expression was quantified by Roche LightCycler 480 using SYBR Green Master Mix (cat#Q711-02,Vazyme). GAPDH was regarded as the reference gene. All primers are from Tsingke Biotech (Tsingke, China), and the primer sequences are shown in **List 1**.

2.7.4 Cell counting kit-8 assay

The cells were detached with 0.25% trypsin, centrifuged, and resuspended in complete culture medium at a density of 5×10^4 cells/ml. Each well of the 96-well plates was administered 100 μl cell suspension. Subsequently, at 24, 48, 72, and 96 h, the CCK8 kit was used for detection. After cultivating for 0, 24, 48, and 72 h, 10 μl CCK-8 solution (cat#A311-01, Vazyme) was added. Finally, the proliferation rate of the cells was detected by absorbance at 450 nm. All of the CCK-8 assays were repeated three times with the similar results and data represented with mean ± SD.

2.7.5 Reactive oxygen species detection

The levels of intracellular ROS were detected using a reactive oxygen species detection kit (cat# 50101ES01, Yeasen) following the manufacturer's instructions.Flow cytometer recording fluorescence intensity.

2.7.6 Determination of malondialdehyde and Fe^{2+} levels

MDA test kits (cell samples, E-BC-K028-M, Elabscience), are used to determine levels of MDA. Detection of Fe^{2+} levels by FerroOrange probe (F374, Dojindo).

2.8 Statistical analysis

R software (version 4.1.0) was used for all statistical analyses and graphical visualizations. Spearman correlation analysis was used to analyze the correlation between FRGs and FRLs. The proportion of tumor-infiltrating immune cells between the high- and low-risk groups was analyzed by the Wlicox test. The chi-square test was used to analyze the differences in clinical characteristics such as age and sex between the two groups. Cox univariate regression analysis and multivariate Cox regression analysis were used to define independent prognostic factors for OS in the two groups. Time-dependent ROC curve analysis was used to assess the predictive accuracy of the OS prognostic models. Two-tailed $p < 0.05$ was considered statistically significant.

3 Results

3.1 Identification of ferroptosis-related differentially expressed LncRNAs in glioma patients

The complete flow chart of the study is shown in Figure 1. We collected a total of 1,850 samples, of which 698 tumor samples (GBM, LGG) were from the TCGA database (https://portal.gdc.cancer.gov/repository), and 1,152 normal samples were from the TCGA and GTEx databases (https://www.gtexportal.org/home/datasets). A total of 13,230 lncRNAs were identified. Furthermore, based on the known ferroptosis-related dataset Ferrdb (http://www.zhounan.org/ferrdb/), we obtained 176 ferroptosis-related genes (FRGs). The specific details of these genes are recorded in Supplementary Table S1. To obtain ferroptosis-related lncRNAs (FRL), Spearman correlation analysis was conducted between lncRNAs in the TCGA database. An FRL was identified if it was significantly correlated with one or more FRGs ($|R^2| > 0.6$ and $p < 0.001$). In total, 433 FRLs were defined. We used a PCA map and bar plots to show the distribution of those samples, as shown in Supplementary Figures S1A,B. Then, we compared the expression of lncRNAs in tumor and normal tissues from the TCGA-GTEx database ($\log_2|FC| > 1$, FDR < 0.05) and identified 1,890 DELs, including 1,132 upregulated DELs and 758 downregulated DELs. We used volcano plots to show these data in Supplementary Figure S1C. Finally, we identified 52 ferroptosis-related DELs (FRDELs) between FRLs and DELs (Figure 2A).

3.2 Identification of prognostic ferroptosis-related differentially expressed lncRNAs

To further understand the prognostic potential of FRDELs, after obtaining OS data for GBM and LGG in TCGA, we predicted the prognostic potential of 52 FRDELs using univariate Cox regression analysis. Finally, 35 prognostic ferroptosis-related differentially expressed lncRNAs (PFRDELs) were obtained ($p < 0.001$) (Figure 2B; Supplementary Figure S1D). The coexpression relationship between the 35 PFRDELs and 176 FRGs is shown in Figure 2C. Thirteen PFRDELs were considered protective

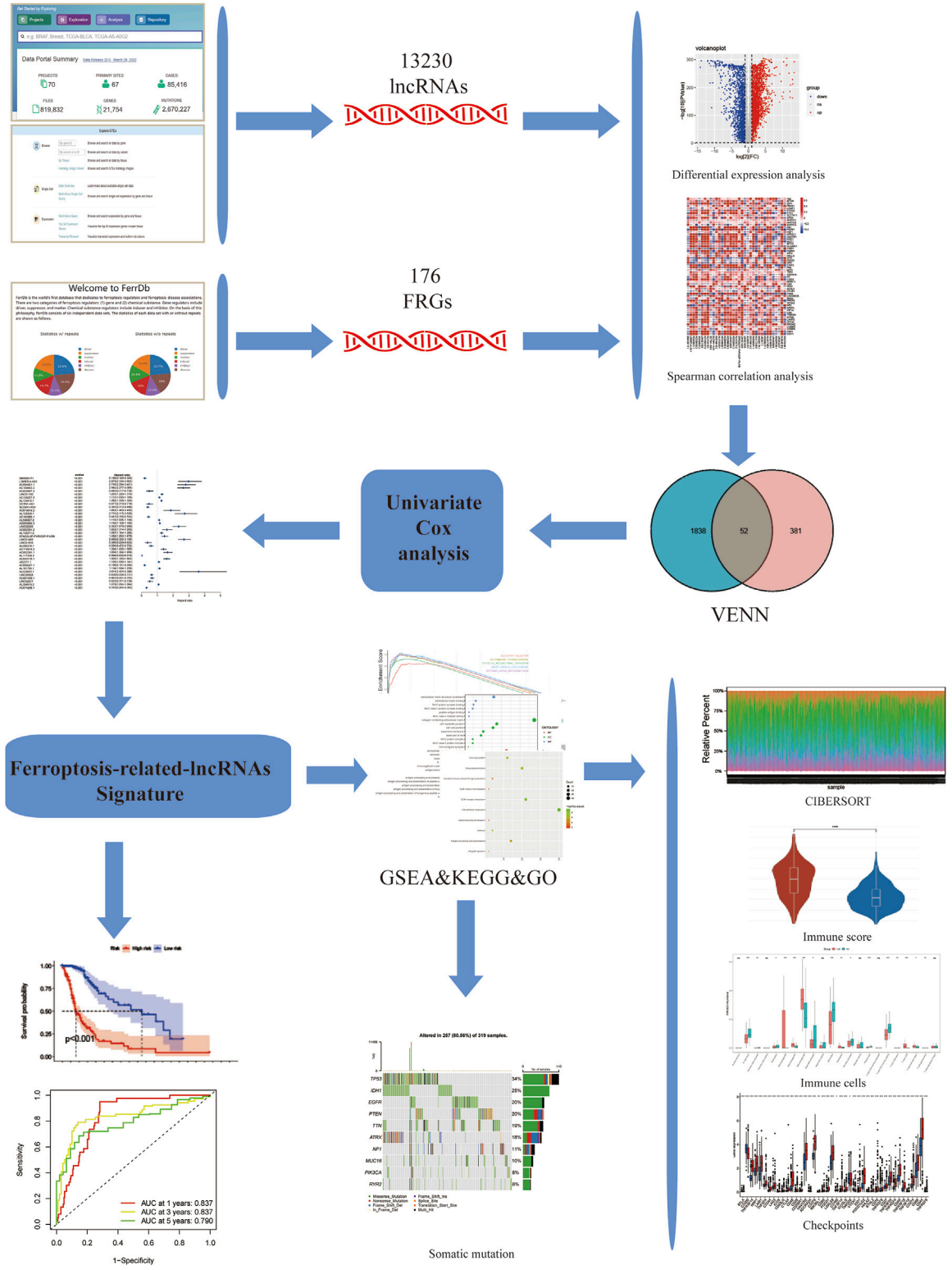


FIGURE 1
Study flowchart. 13,230 LncRNAs were obtained from TCGA and GTEx databases. 176 ferroptosis-related genes (FRGs) were obtained from the FerrDb database. Then, 433 ferroptosis-related lncRNAs (FRLs) were identified according to Spearman correlation analysis. Next, univariate COX analysis was applied to construct a 3-FRL signature. Finally, GSEA, KEGG, GO analysis, immune correlation analysis, somatic mutation analysis were applied to determine the potential function of this feature.

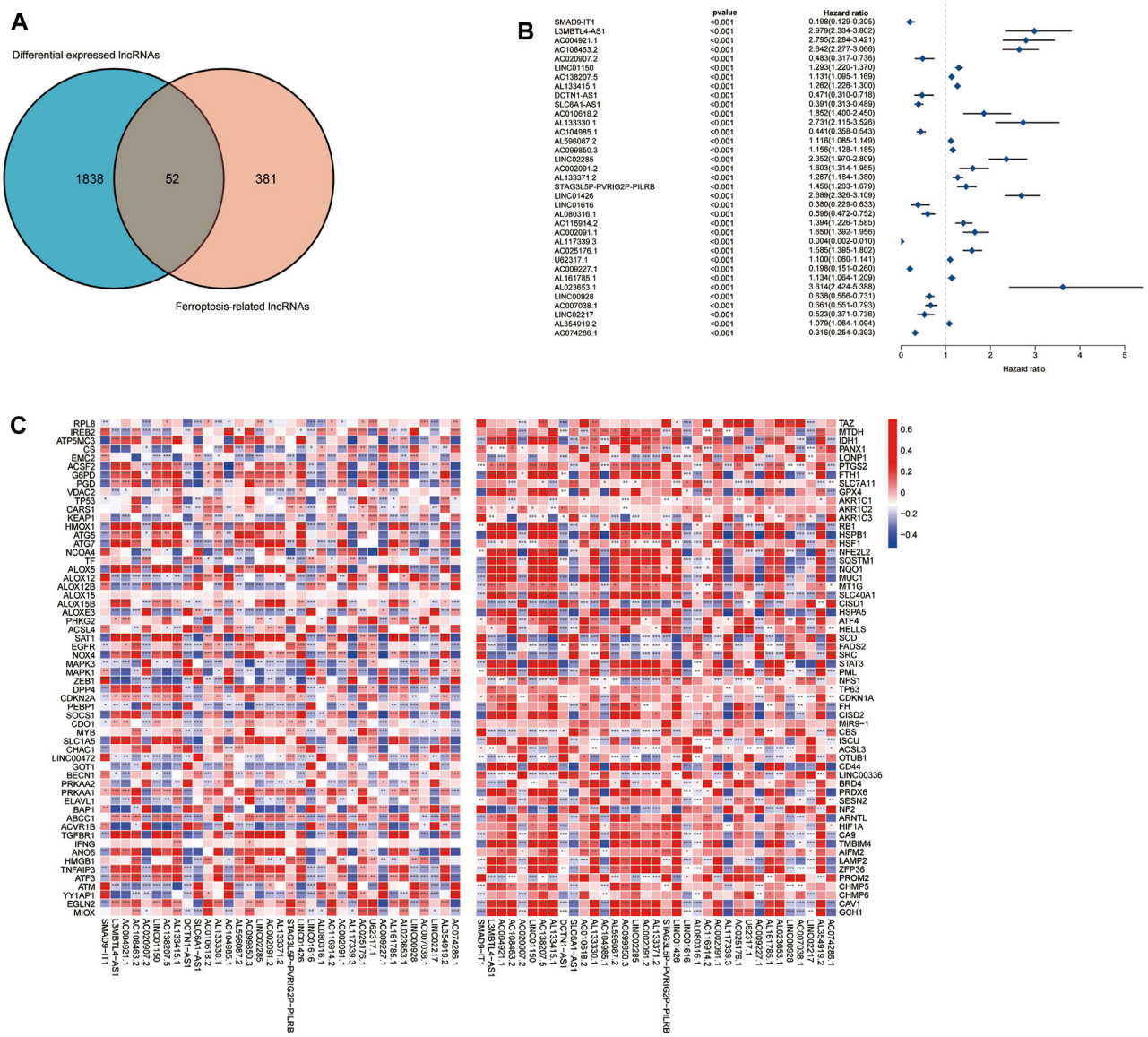


FIGURE 2
Prognostic analysis of differentially expressed ferroptosis-related lncRNAs. **(A)** Differentially expressed FRLs obtained from differentially expressed lncRNAs and FRLs by Venn diagram. **(B)** Forest plots showing the results of the Cox univariate regression analysis approximately 35 prognostic differentially expressed FRLs. **(C)** The correlation between 35 prognostic FRLs and 176 FRGs in the TCGA-LGG/GBM cohort. The colour of each unit shows the degree of corelation. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

factors, and 22 were considered risk factors (the list of these lncRNAs is shown in Supplementary Table S2).

3.3 Construction and validation of a ferroptosis-related lncRNAs prognostic model

To check the prognostic value of these DEFRLS. We collected clinical data from TCGA-GBM/LGG and randomly divided them into two groups: a training group and a validation group. The clinical characteristics of those samples in the above two groups are shown in Table 1.

These 35 PFRDELs in the training group were incorporated into the least absolute shrinkage and selection operator (LASSO) regression. As a result, 3 PFRDELs stood out for the construction of the prognostic FRLS, including AL133415.1, LINC01426 and AC009227.1. Then, based on the optimal penalty parameters (λ) of the LASSO model, a prognostic risk evaluation model for 3-FRLs was constructed. The cvfit and

TABLE 1 The clinical characteristics of glioma patients in the training and validation group.

Clinical parameters	Group		*p* Value
	Training	Validation	
Gender			
Female	265	260	>0.05
Male	182	188	
Age (years)			
≤65	362	364	>0.05
>65	85	84	
Grade			
Stage II	120	123	>0.05
Stage III	132	131	
Stage IV	195	194	
IDH Status			
Mutant	219	217	>0.05
Wild type	228	231	

lambda curves are shown in Figures 3A,B. In this model, each patient with GBM/LGG in the TCGA database was calculated with a risk score by summing the product of the expression level of each selected ferroptosis-related lncRNA and the corresponding coefficient. [Risk Score = AL133415.1*0.01732 + LINC01426*0.15269 + AC009227*(−0.10944)].

To evaluate the independent predictive potential of this signature, the OS-related factors were identified by univariate and multivariate Cox regression analyses. The results of both univariate and multivariate Cox regression analyses indicated that the PFRDLS-based risk score was always an independent prognostic factor for the OS rate of GBM/LGG patients (Figures 3C,D). Predictive nomograms were then constructed and the associated factors' scores on the scales were summed to calculate the likelihood of survival for these patients. The 1-, 3- and 5-year OS rates could be predicted accurately when compared with those of the ideal predictive model (Figures 3E,F). To evaluate the prognostic value of this 3-FRLs model. Then, the samples in the training group were stratified into the high-risk group and low-risk groups using the median risk score as the cutoff value. Subsequently, the risk score distribution and OS status distribution of the above two samples were determined, and the results showed that the distributions of the two samples were reasonable (Figure 4A). Kaplan-Meier analysis of the samples showed that the OS rate of GBM/LGG patients in the high-risk group was worse than that in the low-risk group (Figure 4D). Then, an ROC curve was performed in the training group and found that the prognostic accuracy of the 3-FRLs model was better than that of other clinicopathological characteristics. Since both disease state and factor values change over time, a time-dependent ROC curve was also performed in the training group. The AUCs for 1-, 3-, and 5-year OS in the training group were 0.837, 0.837, and 0.790, respectively (Figure 4G). We then constructed ROC curves for comparison with other clinicopathological features (Figure 4J). To determine whether the prognostic significance of FRLS persisted in other groups, the validation group and overall group were validated in heatmaps, distribution figures, Kaplan-Meier survival analysis and time-dependent ROC analysis. The distribution of the above two risk group samples in the validation group and the overall group is shown in Figures 4B–L. Since molecular subtype and IDH state contribute to the outcome and classification of glioma patients, it is hard to exclude the bias of these factors. Therefore, we performed the Kaplan-Meier analysis to verify whether the signature is functional in both LGG and GBM/IDH wt and IDH mut patients. The results show that the signature is functional in patients with LGG and IDHmut. However, GBM and IDH wt patients did not have statistical significance because of the large difference in sample size (Supplementary Figures S2A–D).All results agree that the mortality rate of the low-risk groups is lower than that of the high-risk groups, and the FRLS prognosis can accurately and stably predict the survival outcome of GBM/LGG patients.

3.4 Relationship between the 3-FRLs signature and the clinicopathological characteristics in glioma patients

In the TCGA-GBM/LGG cohort, two lncRNAs in our model were considered risk lncRNAs and upregulated in the high-risk group. Only AC009227.1 was considered a protective lncRNA that was upregulated in the low-risk group (Figure 5A). Next, we compared the differences in clinical characteristics between the two risk subgroups in terms of age, sex, glioma grade, and IDH status. We found that with the increase in glioma grade, the expression of risk lncRNAs was upregulated, and the expression of protective lncRNAs was downregulated, which ultimately led to the improvement of the risk score. Similar results were observed for age and IDH status, and the clinical features are also compared in Figures 5B–D. Studies have shown that the status of IDH has a significant relationship with the prognosis of glioma (Yan et al., 2009). Taken together, these results suggest that our 3-lncRNA signature has a significant potential to predict the prognosis of glioma patients by assessing risk scores through correlated gene expression levels.

3.5 Discovery of molecule function and pathways by GESA, gene ontology and kyoto encyclopedia of genes and genomes analysis

We further performed GSEA to explore potential differences in biological functions and signaling pathways between different risk groups classified by the 3-FRL signature. Many tumor metastasis pathways are enriched in high-risk populations, such as epithelial-mesenchymal

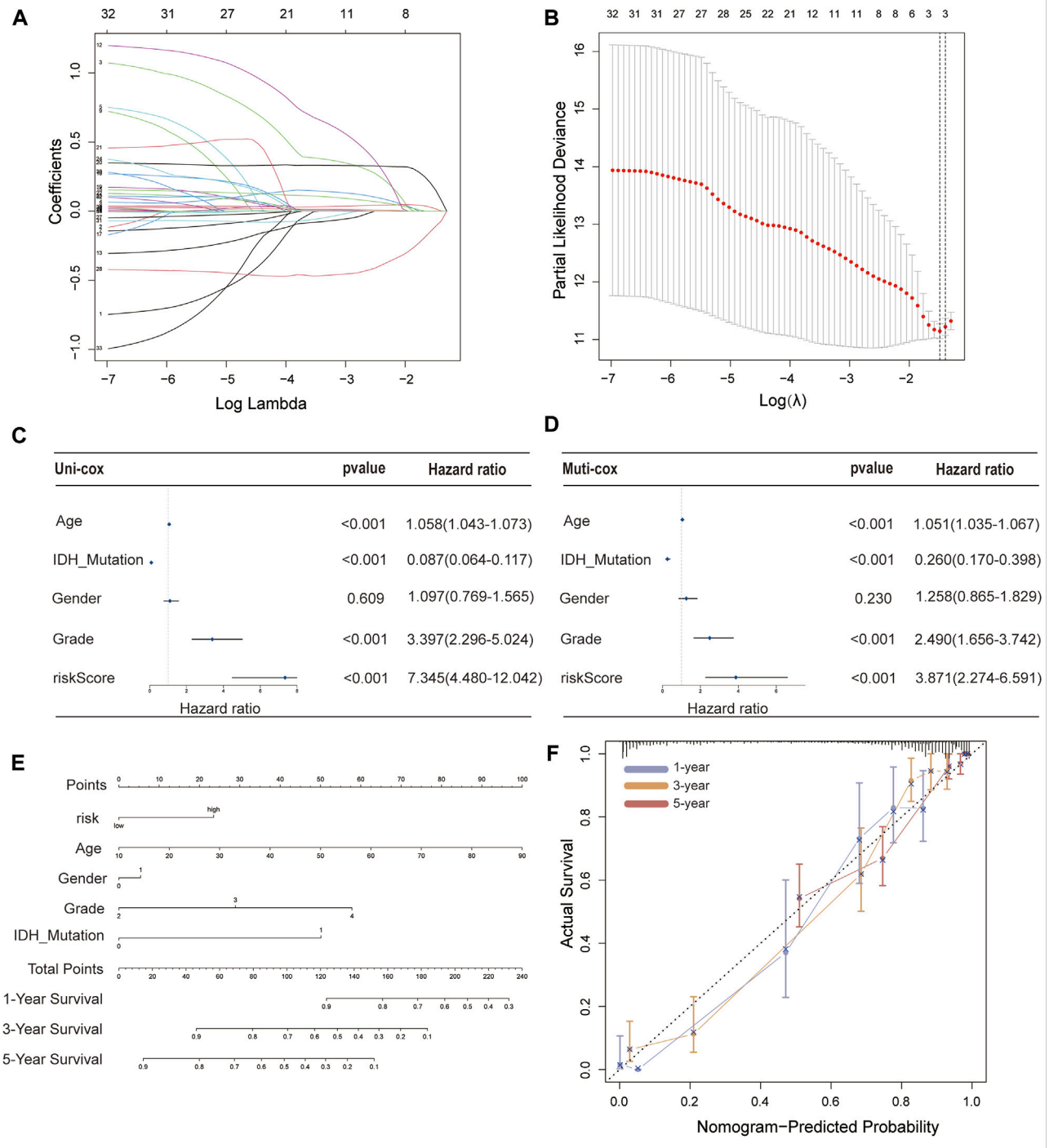


FIGURE 3
Construction of a 3-FRL signature and the analysis of independent prognostic potential. **(A,B)** The least absolute shrinkage and selection operator (LASSO) regression was performed with the minimum criteria. **(C)** Results of the univariate Cox regression analysis and multivariate Cox regression analysis regarding OS of the 3-FRLs signature. **(D)** Nomogram of OS over time for glioma patients. **(E)** This calibration curve is used to assess the accuracy of the nomogram model, and the dashed line represents the ideal nomogram. **(F)** The calibration curve for evaluating the accuracy of the nomogram model. The dashed diagonal line in grey colour represents the ideal nomogram.

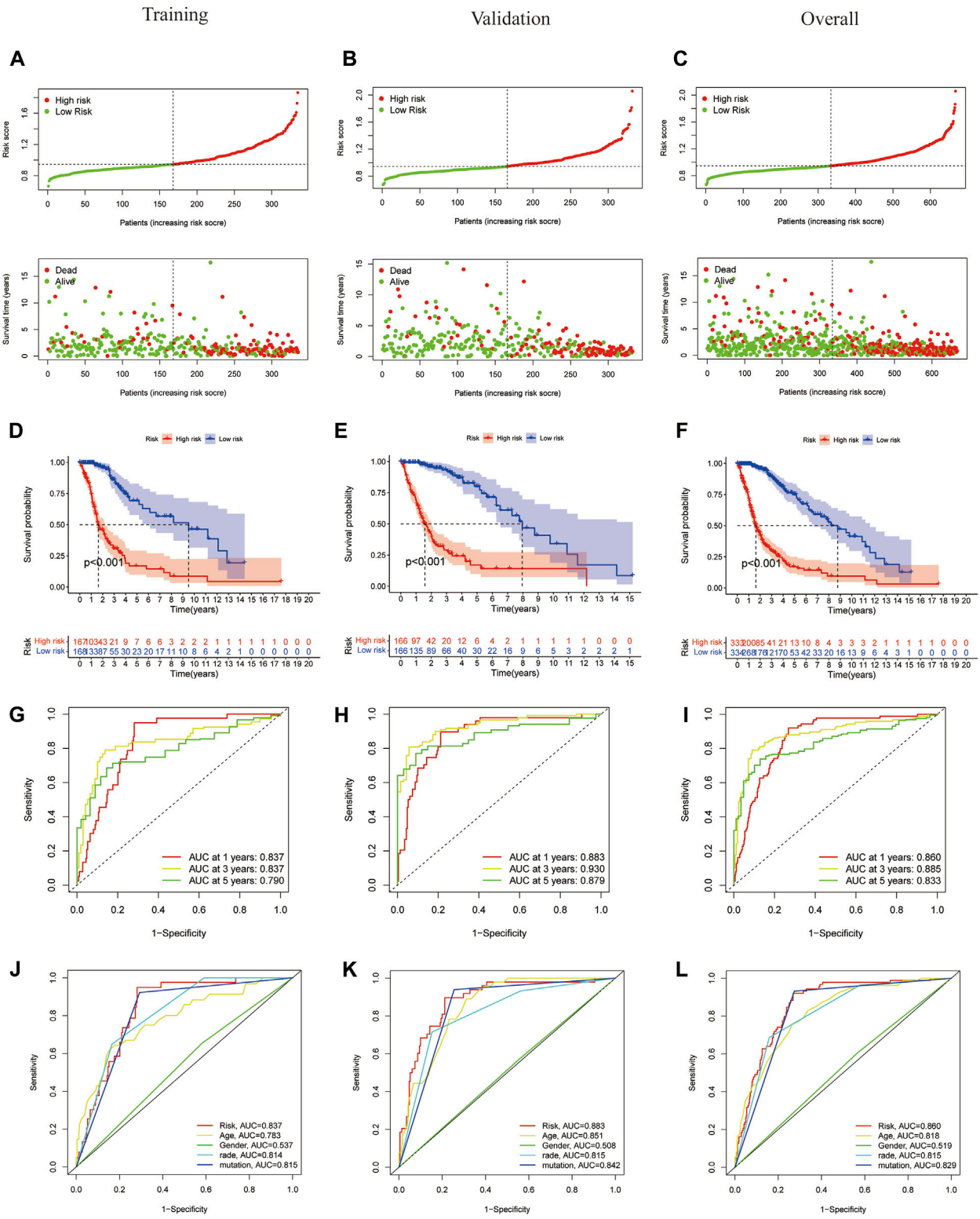


FIGURE 4
Construction and validation of 3-FRLs models in training cohorts, validation and overall groups. **(A–C)** The distribution plots of the risk score and survival status in training cohorts, validation and overall groups. **(D–F)** The Kaplan-Meier curves for survival status and survival time in the training, validation and overall groups. **(G–I)** The receiver operating characteristic (ROC) curve analyses of the prognostic FRLS in predicting 1-, 3-, and 5-year overall survival in training cohorts, validation and overall groups. **(J–L)** Risk scores and other prognostic features of the 3-FRLS model were compared using AUC of ROC curves in the training cohort, validation and overall groups.

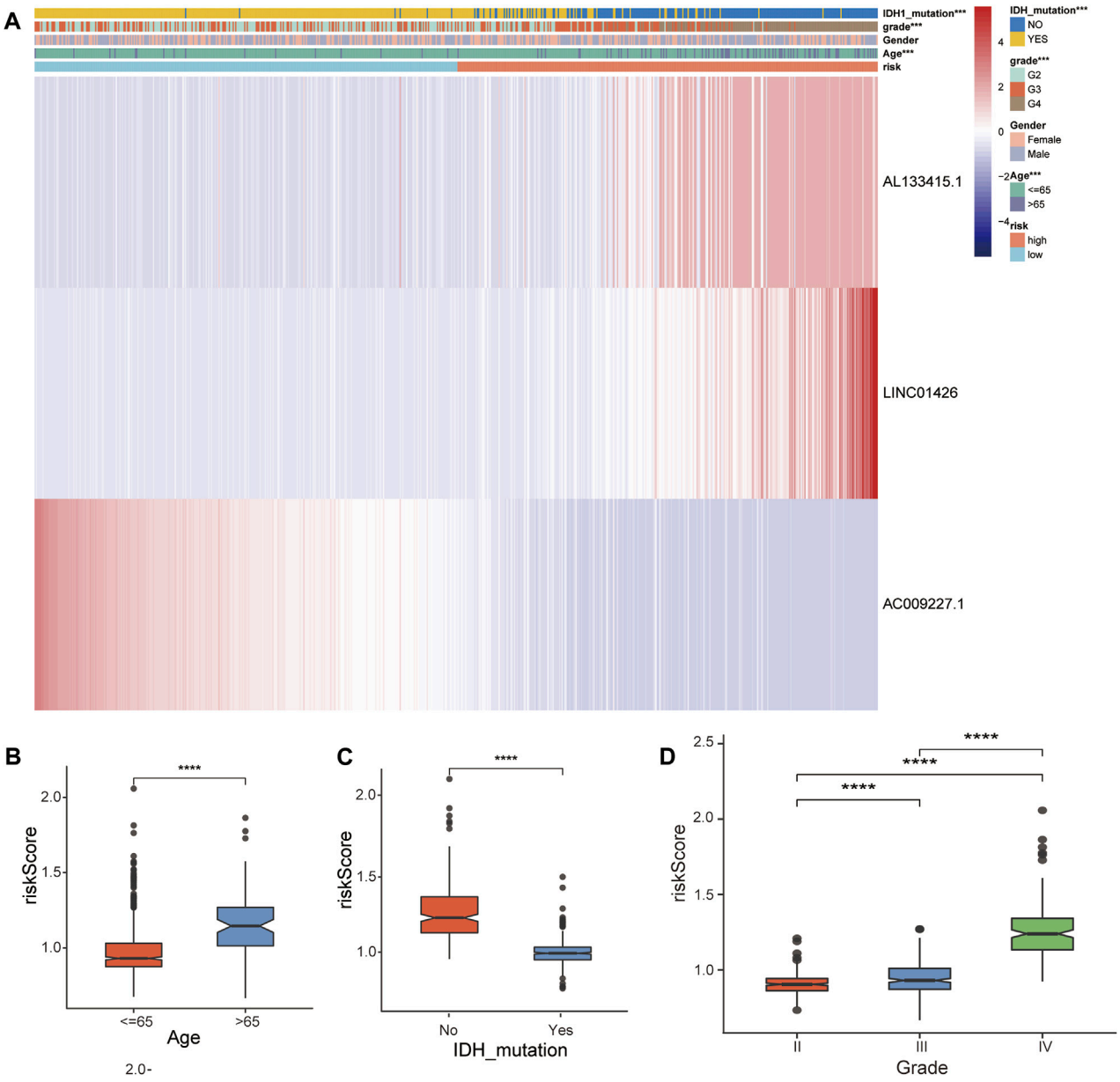


FIGURE 5
Correlation analysis between the prognostic FRLS and clinicopathological characteristics in the TCGA cohort. **(A)** Heatmaps depict the distribution of FRLS expression levels and clinicopathological features in high-risk and low-risk groups. **(B–D)** Different risk score levels in glioma patients stratified by age, sex, grade, IDH mutation status. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

transition (EMT). At the same time, many immune-related pathways were also enriched in high-risk groups, such as systemic lupus erythematosus, autoimmune thyroid disease, graft versus host disease and allograft rejection (Figure 6A). In addition, many signaling-related pathways were enriched in the low-risk group, such as the calcium signaling pathway, phosphatidylinositol signaling system, and hedgehog signaling (Figure 6B). Interestingly, some pathways related to metabolism and proliferation, such as angiogenesis-related pathways, glutathione metabolism and drug metabolism (amino sugar and nucleotide sugar metabolism), were also enriched. The details of the GSEA results are listed in Supplementary Table S3. To explore the biological functions characterizing DEGs between different risk groups. DEGs between the high-risk group and the low-risk group were determined by the cutoff of $\log_2|FC| > 1$ and

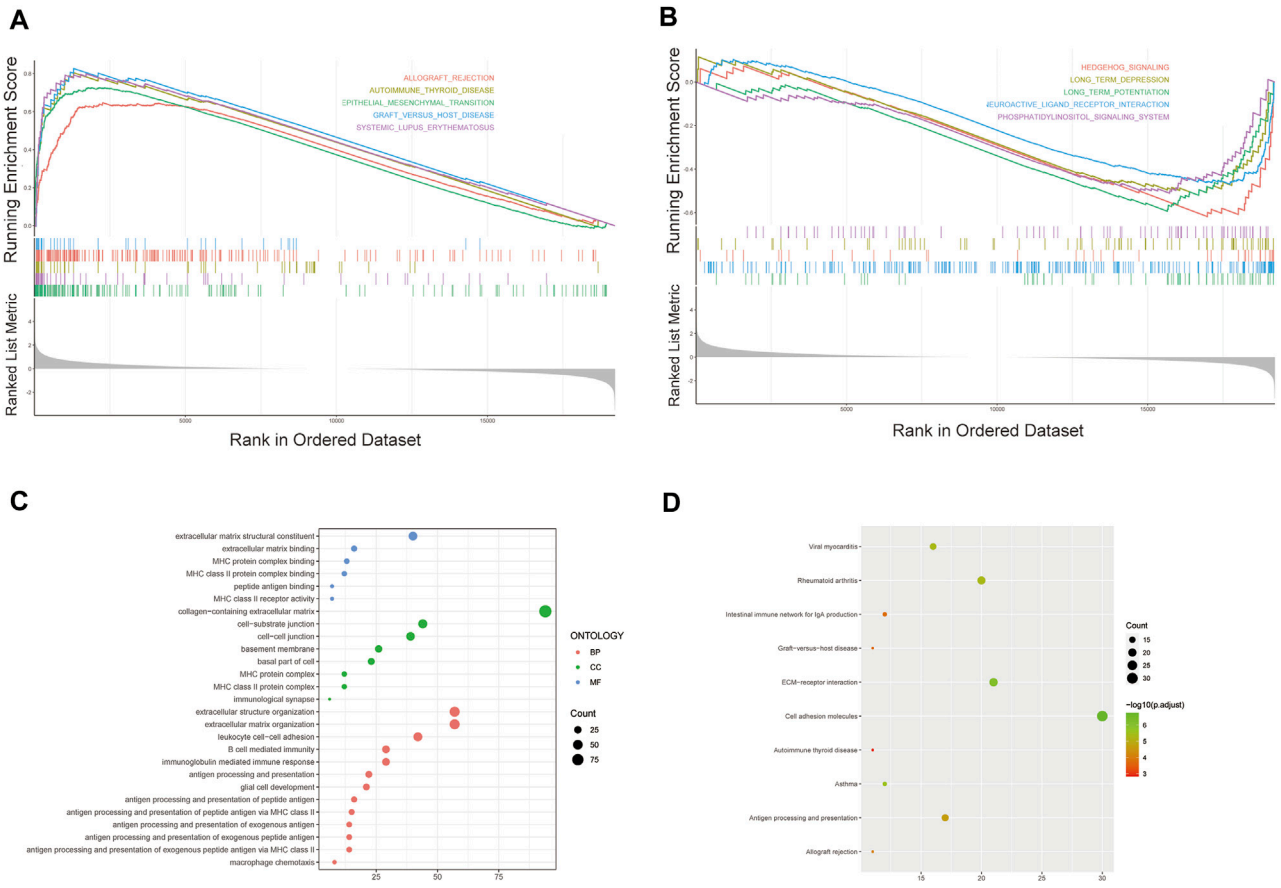


FIGURE 6
Gene biological function and pathway enrichment analysis of prognostic signatures of FRLs in high-risk group and low-risk group. **(A)** GSEA shows significant enrichment of immune-related and metastasis-related pathways in high-risk glioma patients. **(B)** GSEA shows significant enrichment of cancer-related signaling pathways in low-risk glioma patients. **(C)**GO analysis revealed enrichment of many immune-related processes and tumor metastasis-related processes. **(D)**KEGG analysis revealed that many immune-related processes and tumor metastasis-related processes were enriched.

FDR < 0.05, and annotation GO enrichment analysis and KEGG pathway analysis were then performed ($p < 0.05$). GO analysis shows enrichment of biological process (BP), molecular function (MF), and cell component (CC) in Figure 6C. Expectedly, the GO analysis revealed a significant enrichment of immune-related functions, especially in relation to MHC protein complex binding, immunological synapse and antigen processing and presentation of peptide antigen. Similarly, the KEGG analysis indicated the enrichment of metastasis-related pathways, including cell adhesion molecules and ECM–receptor interactions. In addition, many immune-related pathways were significantly enriched, including antigen processing and presentation, rheumatoid arthritis and asthma (Figure 6D). The above two bioinformatics analyses are similar to the GSEA results. In conclusion, these results suggested that the risk score of the 3-FRLs signature was associated mainly with tumor immunity, tumor metastasis and biological metabolism in glioma.

3.6 Immune-related analysis of glioma patients using the prognostic signature

To investigate the correlation of ferroptosis-related features and antitumor immunity in glioma patients. We used the CIBERSORT algorithm to identify the immune cell infiltration landscape of all patients with GBM/LGG from the TCGA database and calculated the proportion of each typical immune cell (Figure 7A). We compared the differences in immune cells in the low-risk group and high-risk group from the stromal score (substrate cells in the tumor tissue), immune score (immune cell infiltration in the tumor tissue) and estimate score (the summation of stromal and immune scores from

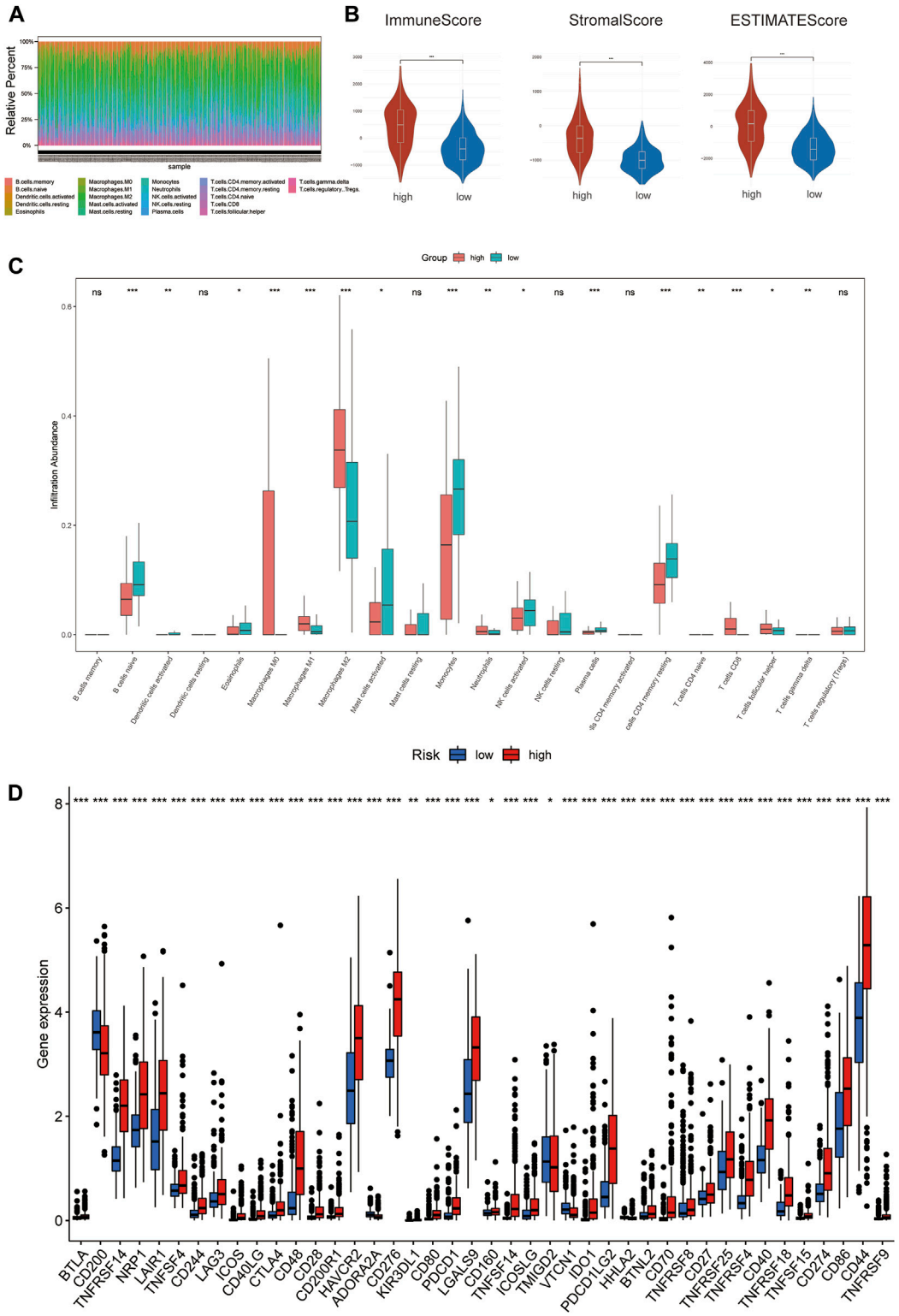


FIGURE 7
The degree of immune infiltration in glioma patients. **(A)** Immune cell distribution in high-risk and low-risk groups in the 3-FRLs model. **(B)** Stroma, immune, and ESTIMATE scores in the high-risk and lowrisk groups in glioma patients. **(C)** Boxplot of comparison of immune cells in high-risk and low-risk groups. **(D)** Boxplots comparing immune checkpoint genes in high- risk and low-risk groups.

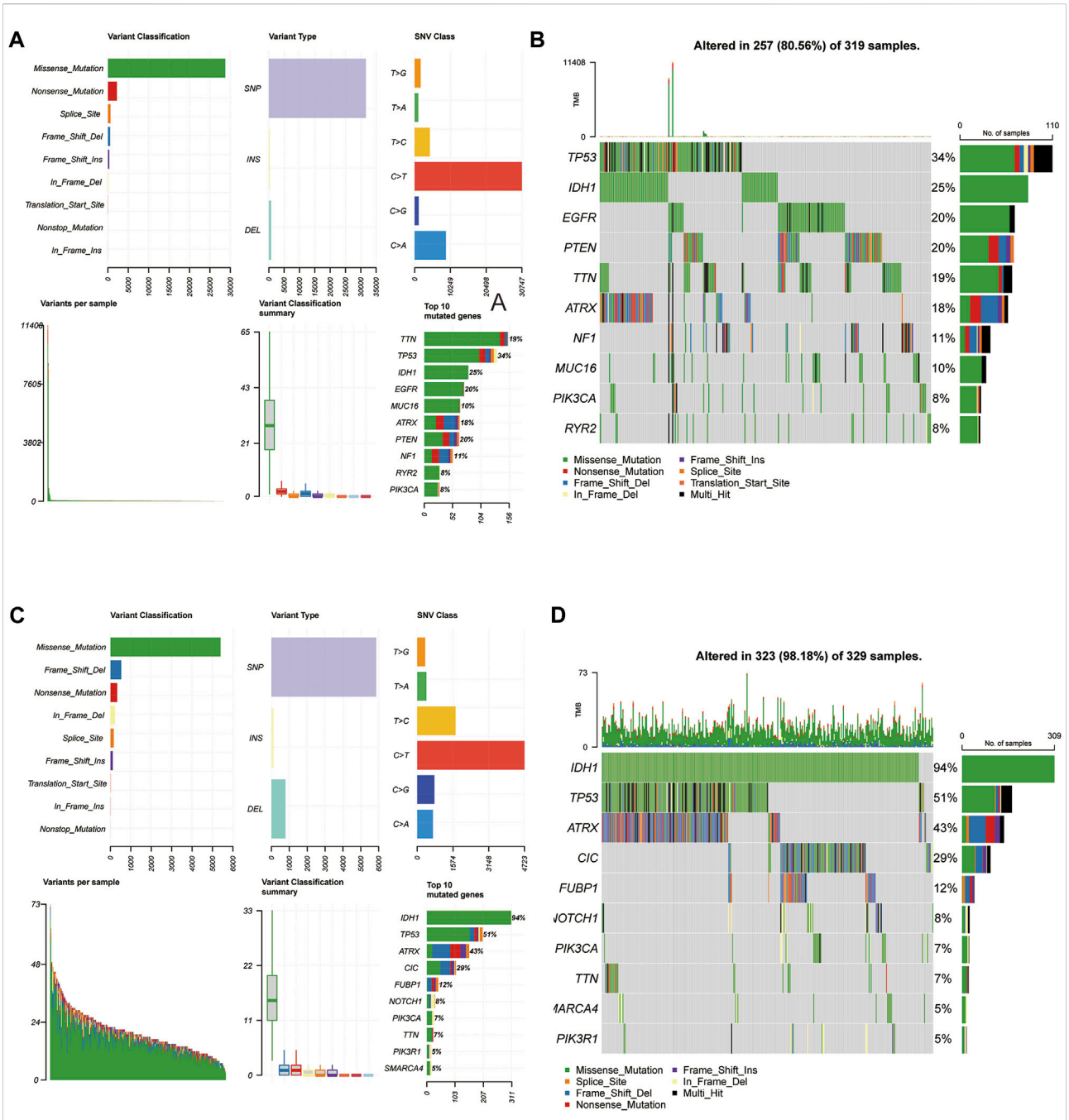


FIGURE 8
Somatic mutation analysis in high-risk and low-risk groups. **(A,B)** MAF-summary plots and oncoplots of somatic mutations in high-risk groups. **(C,D)** MAF-summary plots and oncoplots of somatic mutations in low-risk groups.

individual cases and defined as tumor purity). The results showed that the scores of the high-risk group were higher than those of the low-risk group ($p < 0.001$) (Figure 7B). Meanwhile, there were differences in the proportion of each immune cell between the high-risk group and the low-risk group, including memory B cells, naive B cells, resting dendritic cells, eosinophils, activated mast cells, monocytes, neutrophils, activated NK cells, resting NK cells, plasma cells, CD4 memory activated T-cells, naïve CD4 T-cells, CD8 T-cells, gamma delta T-cells, regulatory T-cells and M0, M1, and M2 macrophages (Figure 7C). Meanwhile, we found statistically significant differences in 42 checkpoint genes between the high- and low-

risk groups (Figure 7D). Among these genes, 39 genes, including PDCD1 (PD-1), CD274 (PD-L1), CTLA4 and LAG3, were highly expressed in the high-risk group, many of which are validated effective immunotherapy targets. In addition, only CD200 had lower expression than the low-risk group (Pardoll 2012; Anderson et al., 2016; Postow et al., 2018; Li et al., 2019). In conclusion, by comparing the relationship between risk scores calculated from 3-FRLs signatures and immune infiltrating cells, the results suggest that the risk level of glioma patients is related to immune cell infiltration.

3.7 Cancer-related gene mutation in the 3-FRLs signature

Mutations arise from replication errors or from DNA damage that is either repaired incorrectly or left unrepaired. The transition from normal cells to tumor cells is often accompanied by genetic mutations. The rates of different mutational processes vary among tumors and cancer types (Martincorena and Campbell 2015).Therefore, to further analyze whether the gene mutation levels of the 3-FRLs signature differed, we sorted out cancer-related gene mutations between the high-risk and low-risk groups separately (Figures 8A–D). Genes such as TP53 (34%), IDH-1 (25%), EGFR (20%), PTEN (20%), TTN (19%), and ATRX (18%) had the top six mutation frequencies in the high-risk group. IDH-1 (94%), TP53 (51%), ATRX (43%), CIC (29%), FUBP1 (12%), and NOTCH1 (8%) were the top six genes with the highest mutation frequencies in the low-risk group. In conclusion, IDH-1 (25% vs. 94%), TP53 (34% vs. 51%), and ATRX (18% vs. 43%) had relatively lower mutation rates in the high-risk group. However, mutations in these genes have been shown to be more frequently found in patients with low-grade gliomas, further demonstrating the predictive power of the 3-FRLs signature for glioma patients.

3.7.1 Validation of ferroptosis-related lncRNAs expression and LINC01426 regulated erastin-induced ferroptosis

PhyloCSF is a comparative genomics method to distinguish protein coding and non-coding regions (Lin et al., 2011). Therefore, we used PhyloCSF to determine whether these FRLs have protein-coding ability. As shown in Supplementary Figure S2E, LINC01426 with negative scores was retained as potential noncoding RNAs (Wang et al., 2020), AL133415.1 and AC009227.1 may have the potential to encode short peptides. Thus, these FRLs do not have the ability to encode complete proteins. We further observed the expression levels of these FRLs in cell lines, as shown in Figure 9A, compared with HA 1800, AL133415.1, and LINC01426 were expressed at relatively higher levels in glioma cell lines (including U251, LNS229, KNS- 89, and T98G), but AC009227.1 exhibited the opposite trend. These results further verified the correctness of the above bioinformatics research (Supplementary Figure S1D). Subsequently, we chose the LINC01426 with the highest scoring coefficient to further analyze. We use the short hairpin RNAs to achieve the stable knockdown of LINC01426 in U251 and KNS-89 . The qRT-PCR results of knockdown efficiency are shown in Figure 9B. As shown in Figure 9C, the Cell Counting Kit-8 (CCK8) assay indicates that the knockdown of LINC01426 significantly inhibited cell proliferation in U251 and KNS-89 cells.In order to further study the impact of LINC01426 on Ferroptosis, we use MDA and FerroOrange assay kits to detect malondialdehyde (MDA) and Fe^{2+} level. As shown in Figures 9D,E, after processing of 12 μM erastin (ferroptosis activator), compared with the control group, the knockdown of LINC01426 have a significant increase in the MDA and Fe^{2+} levels in U251 and KNS89 cells. The occurrence of ferroptosis has a close relationship with the accumulation of ROS (Li et al., 2020). ROS levels were clearly observed after U251 and KNS-89 cells were treated with 12 μM erastin. The erastin-induced ROS level has a significant increase after the knockdown of LINC01426 (Figure 9F). In conclusion, all results suggest that LINC01426 can inhibit the occurrence of ferroptosis in glioma.

4 Discussion

The reason for the high malignancy and drug resistance observed in glioma has been found to be that these tumors can effectively evade ferroptosis. Currently, many studies on glioma have focused on the relationship between lncRNAs and ferroptosis (Deng, et al., 2020; Huang et al., 2021a). The identification of FRLs is essential for finding potential therapeutic targets. However, the exploration of FRLs in gliomas is still limited. Therefore, it is important to construct a predictive model of ferroptosis-related lncRNAs. In this study, we analyzed glioma tumor samples and normal samples from TCGA and GTEx databases and obtained DELs. Then, the 176 ferroptosis-related genes obtained from the online FerrDb database were intersected, and the differentially expressed FRLs were finally screened. Subsequently, we obtained the clinical information and FRL expression profile of each patient from the TCGA database. The results identified 35 prognostic FRLs. Finally, we established a risk assessment model based on 3 FRLs. Compared to ROC curves of published literature (He et al.2021; Shi et al. 2022), our signature has more predictive ability of prognosis and contains fewer lncRNAs. The lncRNA prognosis assessment kits that have been commercialized at present are composed of only 3-

10.3389/fgene.2022.927142

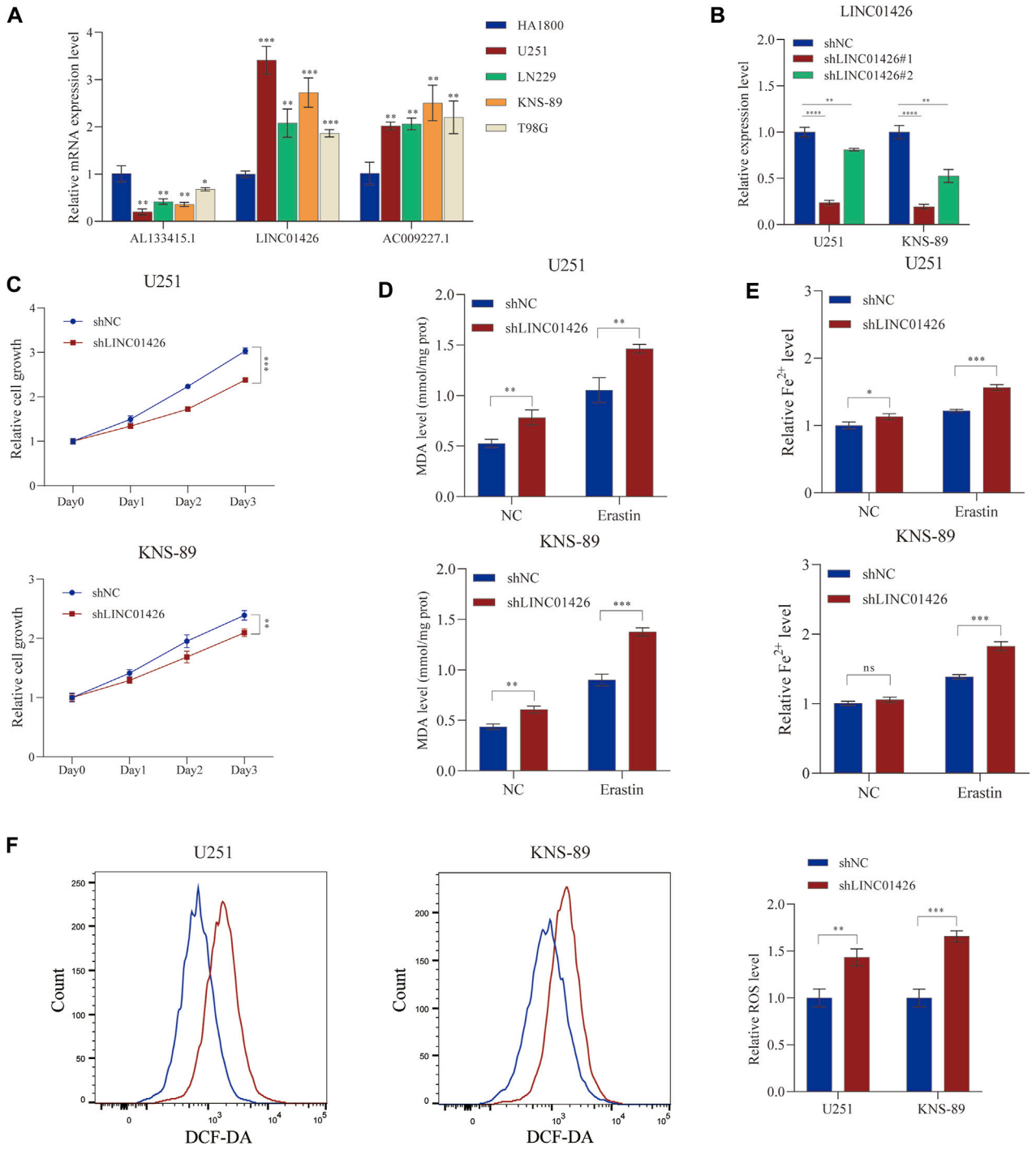


FIGURE 9
Validation of the expression level of the 3-FRLs in cell lines and ferroptosis regulation. **(A)** Expression analysis of 3-FRLs in four glioma cell lines (U251, LN229, KNS-89, T98G) with HA1800 lines (normal astrocytes). **(B)** Relative expression level of LINC01426 after transfection with the corresponding shRNA. **(C)** Cell proliferation levels of U251 and KNS-89 after knocking down LINC01426. **(D,E)** The ferroptosis process was evaluated by detecting MDA and Fe^{2+} levels in the non-erastin-induced and erastin-induced groups. **(F)** The comparison of erastin-induced ROS in the treatment and control groups.*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns, No significance.

5 lncRNAs (Patent No. CN201710998995.9; No. CN201810764922.8). Therefore, our signature is more clinically feasible and has potential for clinical translation.

Interestingly, in our constructed 3-ferroptosis-related lncRNA signature, LINC01426 is an oncogene that has been validated by many cancer researchers. LINC01426 promotes the development of lung cancer (Dai et al., 2020; Han et al., 2020; Liu et al., 2021; Zhu et al., 2022), clear cell renal cell carcinoma (Jiang et al., 2021), and osteosarcoma (Zhang et al., 2021). Furthermore, the role of LINC01426 in glioma has also received increasing attention, including sponging miR-345-3p and upregulating VAMP8 to promote glioblastoma (Cao et al. 2020). Mechanistic investigation showed that LINC01426 exhibited its tumor promoter role by modulating the PI3K/Akt signaling pathway (Wang et al., 2018) and PI3K has been shown to be targeted therapy for glioma (Cruceru et al., 2013). However, how LINC01426 is involved in regulating ferroptosis still needs further exploration. This study indicates that the ferroptosis process of glioma cells was inhibited after knocking down the expression of LINC01426, which fills the gap in this field.

In addition, a novel prognostic 3-lncRNA model was created. Compared with many other identified signatures, this model contains only 3 lncRNAs. Clinically, the model also has good predictive power for patient outcomes. We divided glioma patients into a high-risk group and a low-risk group based on their risk scores calculated by the formula of this prognostic model. To further explore the mechanism by which this signature regulates gliomas, we performed GSEA. The results showed that cancer metastasis pathways, such as epithelial-mesenchymal transition (EMT) and ECM-receptor interaction, were highly ranked in the high-risk group. Among these pathways, EMT can not only enhance tumor invasiveness but also be associated with enhanced stem cell properties and drug resistance (Aiello and Kang 2019). Angiogenesis-related pathways and glutathione metabolism are also enriched; angiogenesis (the formation of new blood vessels) has been shown to be an integral part of cancer development (Viallard and Larrivée 2017), and glutathione is an important component against reactive oxygen species (ROS), which are key substances in the ferroptosis process (Liu et al., 2022). Interestingly, many immune-related pathways were also enriched, including primary immunodeficiency, IL6-JAK-STAT3 signaling and the IL2-STAT5 pathway. Although the relationship between ferroptosis and the immune microenvironment remains controversial (Friedmann Angeli et al., 2019), it is reasonable to hypothesize that there is a link between ferroptosis and tumor immunity in glioma. Subsequently, KEGG enrichment analysis and GO enrichment analysis were also performed, including BP, MF and CC, and the enrichment pathway results were similar to the GSEA results. As we all know, the major barriers to effective treatment of GBM are their high proliferation, progressive spread, and invasiveness, but the underlying mechanisms for controlling gliomas are still far from understood (Groothuis 2000).Taken together, we can infer from the above results that the high-risk group suppressed the occurrence of ferroptosis through immune- and metabolic-related pathways.

Previous studies have shown that ferroptosis is closely related to tumor immunity (Xie et al. 2016; Tang et al., 2019), but direct evidence of the connection between ferroptosis and antitumor immunity was not available until Wang et al., 2019a reported that $CD8^+$ T cells induce ferroptosis in tumor cells *in vivo* (Wang et al., 2019b, Green et al., 2019). In addition, studies have shown that the increased intratumor production of prostaglandin E2 (PGE2) facilitates tumor evasion of immune surveillance (Kalinski 2012; Veglia et al., 2019). In terms of immunotherapy, studies have shown that $CD8^+$ T-cells are involved in radiotherapy-induced ferroptosis in human fibrosarcoma cells and melanoma cells (Lang et al., 2019). However, no study has reported a direct link between ferroptosis and immune cell infiltration in glioma. Since GSEA was enriched in many immune-related pathways, we further calculated the proportions of different types of tumor-infiltrating immune cells in gliomas. The high-risk group had higher immune, stromal, and estimated scores, as calculated by CIBERSORT from the TCGA database. Compared with the low-risk group, the high-risk group had higher expression levels of $CD8^+$ T-cells and macrophages and lower expression levels of monocytes or dendritic cells. High immune and stromal scores and high macrophage infiltration are associated with poor prognosis, which is consistent with our results (Deng et al., 2020). Subsequently, we found that among the immune checkpoints, forty genes, including PDCD1 (PD-1), CD274 (PD-L1), CTLA4, and LAG3, were highly expressed in the high-risk group. Therefore, these patients might benefit from many immune checkpoint blockades (Cristescu et al., 2018), which also provides a possible modality for ferroptosis immunotherapy in the future (Tang et al., 2020).

Next, we analyzed the cancer-related gene mutation status of the two risk subgroups to further explore the relationship between risk scores and cancer-related gene mutations. We found that mutations in IDH-1, TP53 and ATRX were significantly different between the two groups. Numerous studies have shown that IDH1 mutations lead to better overall survival in glioma patients and a better response to therapies (Yan et al. 2009; Franceschi et al., 2021). Recently, an ATRX-deficient genetically engineered glioma model demonstrated that loss of ATRX reduces median survival and increases genetic instability (Koschmann et al., 2016). Moreover, TP53 mutations are frequent in low-grade gliomas and secondary glioblastomas derived therefrom (Ohgaki and Kleihues 2005). These studies have shown that mutations in certain key genes in glioma have a greater impact on prognosis. In our 3-FRLs signature, the risk

scores are also strongly associated with the mutation status of these genes. Interestingly, lncRNA has recently been shown to be required for maintaining genomic stability (Lee et al., 2016; Munschauer et al., 2018).In addition, the lncRNA signatures of genome instability can Predict Survival in Patients (Huang,et al., 2021b; Xie, et al., 2016).The studies could be corroborated with our results. Not only that, our signature provides new genes for studying the relationship between lncRNAs and gene mutation. In the future, we can use these genes to explore the effect of the PUMILIO protein or the topoisomerase complex, as researchers did with NORAD.

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

LH constructed this study. LH, HC, FG, YH, XH, and WL performed the data analysis, figures plotted, and writing. JZ and FZ were responsible for the critical reading of the manuscript. HC and FZ provided guidance on the experimental section of the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

We all authors sincerely acknowledge the contributions from the TCGA and GTEx projects.

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.2022.927142/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Identififfication of differentially expressed ferroptosis related lncRNAs in GBM/LGG. **(A,B)** The PCA map and the bar plots showing the distribution of glioma samples. **(C)** The volcano map of lncRNAs differentially expressed in GBM/LGG. **(D)** The heatmap of these 35 lncRNAs showing the expression level of each lncRNA in each patient.

SUPPLEMENTARY TABLE S1
The details of all of the ferroptosis-related genes.

SUPPLEMENTARY TABLE S2
The table of differentially expressed lncRNAs and ferroptosis-related lncRNAs.

SUPPLEMENTARY TABLE S3
The list of pathways enriched in GSEA.

References

Aiello, N. M., and Kang, Y. (2019). Context-dependent EMT programs in cancer metastasis. *J. Exp. Med.* 6, 1016–1026. Epub 2019/04/13. doi:10.1084/jem.20181827

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, 989–1004. Epub 2016/05/19. doi:10.1016/j.immuni.2016.05.001

Arita, H., Matsushita, Y., Machida, R., Yamasaki, K., Hata, N., Ohno, M., et al. (2020). TERT promoter mutation confers favorable prognosis regardless of 1p/19q status in adult diffuse gliomas with IDH1/2 mutations. *Acta Neuropathol. Commun.* 23 (8), 201. Epub 2020/11/25. doi:10.1186/s40478-020-01078-2

Cao, J., Tang, Z., and Su, Z. (2020). Long non-coding RNA LINC01426 facilitates glioblastoma progression via sponging miR-345-3p and upregulation of VAMP8. *Cancer Cell Int.* 20, 327. Epub 2020/07/24. doi:10.1186/s12935-020-01416-3

Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., et al. (2018). Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 362, eaar3593. Epub 2018/10/13. doi:10.1126/science.aar3593

Cruceru, M. L., Enciu, A. M., Popa, A. C., Albulescu, R., Neagu, M., Tanase, C. P., et al. (2013). Signal transduction molecule patterns indicating potential glioblastoma therapy approaches. *Onco. Targets. Ther.* 6, 1737–1749. Epub 2013/12/19. doi:10.2147/OTT.S52365

Dai, J., Wang, B., Zhao, Y., Zuo, X., Cui, H., Chen, X., et al. (2020). Long noncoding RNA LINC01426 sequesters microRNA-519d-5p to promote non-small cell lung cancer progression by increasing ETS1 expression. *Cancer Manag. Res.* 12, 12697–12708. Epub 2020/12/19. doi:10.2147/CMAR.S277113

Deng, X., Lin, D., Zhang, X., Shen, X., Yang, Z., Yang, L., et al. (2020). Profiles of immune-related genes and immune cell infiltration in the tumor microenvironment of diffuse lower-grade gliomas. *J. Cell. Physiol.* 235, 7321–7331. Epub 2020/03/13. doi:10.1002/jcp.29633

Diederichs, S. (2014). The four dimensions of noncoding RNA conservation. *Trends Genet.* 30, 121–123. doi:10.1016/j.tig.2014.01.004

Dixon, S. (2017). Ferroptosis: Bug or feature? *Immunol. Rev.* 277, 150–157. doi:10.1111/imr.12533

Franceschi, E., De Biase, D., Di Nunno, V., Pession, A., Tosoni, A., Gatto, L., et al. (2021). IDH1 non-canonical mutations and survival in patients with glioma. *Diagn. (Basel)* 11, 342. Epub 2021/03/07. doi:10.3390/diagnostics11020342

Friedmann Angeli, J. P., Krysko, D. V., and Conrad, M. (2019). Ferroptosis at the crossroads of cancer-acquired drug resistance and immune evasion. *Nat. Rev. Cancer* 19, 405–414. Epub 2019/05/19. doi:10.1038/s41568-019-0149-1

Groothuis, D. R. (2000). The blood-brain and blood-tumor barriers: A review of strategies for increasing drug delivery. *Neuro. Oncol.* 2, 45–59. Epub 2001/04/17. doi:10.1093/neuonc/2.1.45

Han, X., Jiang, H., Qi, J., Li, J., Yang, J., Tian, Y., et al. (2020). Novel lncRNA UPLA1 mediates tumorigenesis and prognosis in lung adenocarcinoma. *Cell Death Dis.* 21 (11), 999. Epub 2020/11/23. doi:10.1038/s41419-020-03198-y

He, Y., Ye, Y., Tian, W., and Qiu, H. (2021). A novel lncRNA panel related to ferroptosis, tumor progression, and microenvironment is a robust prognostic indicator for glioma patients. *Front. Cell Dev. Biol.* 9, 788451. Epub 2021/12/25. doi:10.3389/fcell.2021.788451

Huang, L., Xie, Y., Jiang, S., Han, W., Zeng, F., and Li, D. (2021a). The lncRNA signatures of genome instability to predict survival in patients with renal cancer. *J. Healthc. Eng.* 2021, 1090698. Epub 2021/12/18. doi:10.1155/2021/1090698

Huang, R., Dong, R., Wang, N., He, Y., Zhu, P., Wang, C., et al. (2021b). Adaptive changes allow targeting of ferroptosis for glioma treatment. *Cell. Mol. Neurobiol.* 42, 2055–2074. Epub 2021/04/25. doi:10.1007/s10571-021-01092-5

Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P., et al. (2003). MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* 22, 8031–8041. doi:10.1038/sj.onc.1206928

Jiang, Y., Zhang, H., Li, W., Yan, Y., Yao, X., and Gu, W. (2021). LINC01426 contributes to clear cell renal cell carcinoma progression by modulating CTBP1/miR-423-5p/FOXM1 axis via interacting with IGF2BP1. *J. Cell. Physiol.* 236, 427–439. Epub 2020/06/26. doi:10.1002/jcp.29871

Kalinski, P. (2012). Regulation of immune responses by prostaglandin E2. *J. Immunol.* 188, 21–28. Epub 2011/12/22. doi:10.4049/jimmunol.1101029

Koschmann, C., Calinescu, A. A., Nunez, F. J., Mackay, A., Fazal-Salom, J., Thomas, D., et al. (2016). ATRX loss promotes tumor growth and impairs nonhomologous end joining DNA repair in glioma. *Sci. Transl. Med.* 8, 328ra28. 328ra328. Epub 2016/03/05. doi:10.1126/scitranslmed.aac8228

Lang, X., Green, M. D., Wang, W., Yu, J., Choi, J. E., Jiang, L., et al. (2019). Radiotherapy and immunotherapy promote tumoral lipid oxidation and ferroptosis via synergistic repression of SLC7A11. *Cancer Discov.* 9, 1673–1685. Epub 2019/09/27. doi:10.1158/2159-8290.CD-19-0338

Lee, S., Kopp, F., Chang, T. C., Sataluri, A., Chen, B., Sivakumar, S., et al. (2016). Noncoding RNA NORAD regulates genomic stability by sequestering PUMILIO proteins. *Cell* 14 (164), 69–80. Epub 2016/01/05. doi:10.1016/j.cell.2015.12.017

Li, B., Chan, H. L., and Chen, P. (2019). Immune checkpoint inhibitors: Basics and challenges. *Curr. Med. Chem.* 26, 3009–3025. Epub 2017/08/08. doi:10.2174/0929867324666170804143706

Li, J., Cao, F., Yin, H. L., Huang, Z. J., Lin, Z. T., Mao, N., et al. (2020). Ferroptosis: Past, present and future. *Cell Death Dis.* 11, 88. Epub 2020/02/06. doi:10.1038/s41419-020-2298-2

Liang, C., Zhang, X., Yang, M., and Dong, X. (2019). Recent progress in ferroptosis inducers for cancer therapy. *Adv. Mat.* 31, e1904197. Epub 2019/10/09. doi:10.1002/adma.201904197

Lin, M. F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: A comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* 27, i275–282. Epub 2011/06/21. doi:10.1093/bioinformatics/btr209

Liu, T., Sun, L., Zhang, Y., Wang, Y., and Zheng, J. (2022). Imbalanced GSH/ROS and sequential cell death. *J. Biochem. Mol. Toxicol.* 36, e22942. Epub 2021/11/03. doi:10.1002/jbt.22942

Liu, X., Yin, Z., Xu, L., Liu, H., Jiang, L., Liu, S., et al. (2021). Upregulation of LINC01426 promotes the progression and stemness in lung adenocarcinoma by enhancing the level of SHH protein to activate the hedgehog pathway. *Cell Death Dis.* 12, 173. Epub 2021/02/12. doi:10.1038/s41419-021-03435-y

Martincorena, I., and Campbell, P. (2015). Somatic mutation in cancer and normal cells. *Sci. (New York, NY)* 349, 1483–1489. doi:10.1126/science.aab4082

Mou, Y., Wang, J., Wu, J., He, D., Zhang, C., Duan, C., et al. (2019). Ferroptosis, a new form of cell death: Opportunities and challenges in cancer. *J. Hematol. Oncol.* 12, 34. Epub 2019/03/31. doi:10.1186/s13045-019-0720-y

Munschauer, M., Nguyen, C. T., Sirokman, K., Hartigan, C. R., Hogstrom, L., Engreitz, J. M., et al. (2018). The NORAD lncRNA assembles a topoisomerase complex critical for genome stability. *Nature* 561, 132–136. Epub 2018/08/29. doi:10.1038/s41586-018-0453-z

Nicholson, J., and Fine, H. (2021). Diffuse glioma heterogeneity and its therapeutic implications. *Cancer Discov.* 11, 575–590. doi:10.1158/2159-8290.CD-20-1474

Ohgaki, H., and Kleihues, P. (2005). Epidemiology and etiology of gliomas. *Acta Neuropathol.* 109, 93–108. doi:10.1007/s00401-005-0991-y

Ostrom, Q., Bauchet, L., Davis, F., Deltour, I., Fisher, J., Langer, C., et al. (2014). The epidemiology of glioma in adults: A "state of the science" review. *Neuro. Oncol.* 16, 896–913. doi:10.1093/neuonc/nou087

Ostrom, Q. T., Gittleman, H., Farah, P., Ondracek, A., Chen, Y., Wolinsky, Y., et al. (2013). CBTRUS statistical report: Primary brain and central nervous system tumors diagnosed in the United States in 2006-2010. *Neuro. Oncol.* 15, ii1–i56. Epub 2013/10/30. doi:10.1093/neuonc/not151

Ostrom, Q. T., Patil, N., Cioffi, G., Waite, K., Kruchko, C., and Barnholtz-Sloan, J. S. (2020). CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2013-2017. *Neuro. Oncol.* 22, iv1–iv96. Epub 2020/10/31. doi:10.1093/neuonc/noaa200

Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* 12, 252–264. Epub 2012/03/23. doi:10.1038/nrc3239

Pirozzi, C. J., and Yan, H. (2021). The implications of IDH mutations for cancer development and therapy. *Nat. Rev. Clin. Oncol.* 18, 645–661. Epub 2021/06/17. doi:10.1038/s41571-021-00521-0

Pop, S., Enciu, A. M., Necula, L. G., and Tanase, C. (2018). Long non-coding RNAs in brain tumours: Focus on recent epigenetic findings in glioma. *J. Cell. Mol. Med.* 22, 4597–4610. Epub 2018/08/18. doi:10.1111/jcmm.13781

Postow, M. A., Sidlow, R., and Hellmann, M. D. (2018). Immune-related adverse events associated with immune checkpoint blockade. *N. Engl. J. Med.* 378, 158–168. Epub 2018/01/11. doi:10.1056/NEJMra1703481

Qi, W., Li, Z., Xia, L., Dai, J., Zhang, Q., Wu, C., et al. (2019). LncRNA GABPB1-AS1 and GABPB1 regulate oxidative stress during erastin-induced ferroptosis in HepG2 hepatocellular carcinoma cells. *Sci. Rep.* 7 (9), 16185. Epub 2019/11/09. doi:10.1038/s41598-019-52837-8

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. Epub 2015/01/22. doi:10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. Epub 2009/11/17. doi:10.1093/bioinformatics/btp616

Schmitz, S., Grote, P., and Herrmann, B. (2016). Mechanisms of long noncoding RNA function in development and disease. *Cell. Mol. Life Sci.* 73, 2491–2509. doi:10.1007/s00018-016-2174-5

Shi, J., Lai, D., Zuo, X., Liu, D., Chen, B., Zheng, Y., et al. (2022). Identification of ferroptosis-related biomarkers for prognosis and immunotherapy in patients with glioma. *Front. Cell Dev. Biol.* 10, 817643. Epub 2022/02/18. doi:10.3389/fcell.2022.817643

Srikantan, V., Zou, Z., Petrovics, G., Xu, L., Augustus, M., Davis, L., et al. (2000). PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc. Natl. Acad. Sci. U. S. A.* 97, 12216–12221. doi:10.1073/pnas.97.22.12216

Stupp, R., Mason, W. P., van den Bent, M. J., Weller, M., Fisher, B., Taphoorn, M. J., et al. (2005). Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N. Engl. J. Med.* 352, 987–996. Epub 2005/03/11. doi:10.1056/NEJMoa043330

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, 15545–15550. Epub 2005/10/04. doi:10.1073/pnas.0506580102

Suzuki, H., Aoki, K., Chiba, K., Sato, Y., Shiozawa, Y., Shiraishi, Y., et al. (2015). Mutational landscape and clonal architecture in grade II and III gliomas. *Nat. Genet.* 47, 458–468. Epub 2015/04/08. doi:10.1038/ng.3273

Tan, Y., Zhang, S., Xiao, Q., Wang, J., Zhao, K., Liu, W., et al. (2020). Prognostic significance of ARL9 and its methylation in low-grade glioma. *Genomics* 112, 4808–4816. Epub 2020/09/04. doi:10.1016/j.ygeno.2020.08.035

Tang, D., Kang, R., Berghe, T. V., Vandenabeele, P., and Kroemer, G. (2019). The molecular machinery of regulated cell death. *Cell Res.* 29, 347–364. Epub 2019/04/06. doi:10.1038/s41422-019-0164-5

Tang, R., Xu, J., Zhang, B., Liu, J., Liang, C., Hua, J., et al. (2020). Ferroptosis, necroptosis, and pyroptosis in anticancer immunity. *J. Hematol. Oncol.* 13, 110. doi:10.1186/s13045-020-00946-7

Tu, Z., Wu, L., Wang, P., Hu, Q., Tao, C., Li, K., et al. (2020). N6-Methylandenosine-Related lncRNAs are potential biomarkers for predicting the overall survival of lower-grade glioma patients. *Front. Cell Dev. Biol.* 8, 642. Epub 2020/08/15. doi:10.3389/fcell.2020.00642

van den Bent, M. J., Dubbink, H. J., Sanson, M., van der Lee-Haarloo, C. R., Hegi, M., Jeuken, J. W., et al. (2009). MGMT promoter methylation is prognostic but not predictive for outcome to adjuvant PCV chemotherapy in anaplastic oligodendroglial tumors: A report from EORTC brain tumor group study 26951. *J. Clin. Oncol.* 27, 5881–5886. Epub 2009/11/11. doi:10.1200/JCO.2009.24.1034

Veglia, F., Tyurin, V. A., Blasi, M., De Leo, A., Kossenkov, A. V., Donthireddy, L., et al. (2019). Fatty acid transport protein 2 reprograms neutrophils in cancer. *Nature* 569, 73–78. Epub 2019/04/19. doi:10.1038/s41586-019-1118-2

Viallard, C., and Larrivée, B. (2017). Tumor angiogenesis and vascular normalization: Alternative therapeutic targets. *Angiogenesis* 20, 409–426. Epub 2017/07/01. doi:10.1007/s10456-017-9562-9

Wang, J., Koganti, P. P., and Yao, J. (2020). Systematic identification of long intergenic non-coding RNAs expressed in bovine oocytes. *Reprod. Biol. Endocrinol.* 18, 13. Epub 2020/02/23. doi:10.1186/s12958-020-00573-4

Wang, M., Mao, C., Ouyang, L., Liu, Y., Lai, W., Liu, N., et al. (2019a). Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. *Cell Death Differ.* 26, 2329–2343. Epub 2019/02/23. doi:10.1038/s41418-019-0304-y

Wang, S. J., Wang, H., Zhao, C. D., and Li, R. (2018). Long noncoding RNA LINC01426 promotes glioma progression through PI3K/AKT signaling pathway and serves as a prognostic biomarker. *Eur. Rev. Med. Pharmacol. Sci.* 22, 6358–6368. Epub 2018/10/20. doi:10.26355/eurrev_201810_16047

Wang, W., Green, M., Choi, J. E., Gijón, M., Kennedy, P. D., Johnson, J. K., et al. (2019b). CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature* 569, 270–274. Epub 2019/05/03. doi:10.1038/s41586-019-1170-y

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. (N Y).* 2, 100141. Epub 2021/09/25. doi:10.1016/j.xinn.2021.100141

Xie, Y., Hou, W., Song, X., Yu, Y., Huang, J., Sun, X., et al. (2016). Ferroptosis: Process and function. *Cell Death Differ.* 23, 369–379. Epub 2016/01/23. doi:10.1038/cdd.2015.158

Yan, H., Parsons, D. W., Jin, G., McLendon, R., Rasheed, B. A., Yuan, W., et al. (2009). IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.* 360, 765–773. Epub 2009/02/21. doi:10.1056/NEJMoa0808710

Zhang, P., Xu, K., Wang, J., Zhang, J., and Quan, H. (2021). Identification of N6-methylandenosine related LncRNAs biomarkers associated with the overall survival of osteosarcoma. *BMC Cancer* 21, 1285. Epub 2021/12/03. doi:10.1186/s12885-021-09011-z

Zheng, J., Zhou, Z., Qiu, Y., Wang, M., Yu, H., Wu, Z., et al. (2021). A prognostic ferroptosis-related lncRNAs signature associated with immune landscape and radiotherapy response in glioma. *Front. Cell Dev. Biol.* 9, 675555. Epub 2021/06/08. doi:10.3389/fcell.2021.675555

Zhu, X., Zhao, J., and Xu, J. (2022). Long noncoding RNA LINC01426 promotes the progression of lung adenocarcinoma via regulating miRNA-125a-5p/casein kinase 2 alpha 1 axis. *Bioengineered* 13, 7020–7033. Epub 2022/03/11. doi:10.1080/21655979.2022.2044251

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 07 November 2022
DOI 10.3389/fgene.2022.905988

Check for updates

OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Jiali Du,
Fudan University, China
Shudi Li,
University of Texas Health Science Center at Houston, United States

*CORRESPONDENCE
Zengnan Mo,
mozengnan@gxmu.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 28 March 2022
ACCEPTED 26 October 2022
PUBLISHED 07 November 2022

CITATION
Zou X, Guo Y and Mo Z (2022), TLR3 serves as a novel diagnostic and prognostic biomarker and is closely correlated with immune microenvironment in three types of cancer.
Front. Genet. 13:905988.
doi: 10.3389/fgene.2022.905988

COPYRIGHT
© 2022 Zou, Guo and Mo. 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.

TLR3 serves as a novel diagnostic and prognostic biomarker and is closely correlated with immune microenvironment in three types of cancer

Xiong Zou[1,2,3,4†], Yi Guo[2,3,4,5†] and Zengnan Mo[1,2,3,4,5*]

[1]Department of Urology, The First Affiliated Hospital of Guangxi Medical University, Nanning, China, [2]Center for Genomic and Personalized Medicine, Guangxi Medical University, Nanning, China, [3]Guangxi Collaborative Innovation Center for Genomic and Personalized Medicine, Nanning, China, [4]Guangxi Key Laboratory for Genomic and Personalized Medicine, Guangxi Key Laboratory of Colleges and Universities, Nanning, China, [5]Department of Gynecology and Obstetrics, The First Affiliated Hospital of Guangxi Medical University, Nanning, China

Background: Toll-like receptor 3 (TLR3) plays an important role in both innate and adaptive immunity, but the prognostic value of TLR3 in heterogeneous tumors and the correlations between TLR3 expression and immune infiltration of heterogeneous tumors remain unclear.

Methods: We investigated the expression of TLR3 in a variety of tumors and focused on the diagnostic and prognostic values of TLR3 in kidney renal clear cell carcinoma (KIRC), pancreatic adenocarcinoma (PAAD) and brain lower grade glioma (LGG) by GEPIA, DriverDBv3, UALCAN, TIMER, LinkedOmics, STRING, GeneMANIA and FunRich, as well as the possible mechanisms of TLR3 affecting tumor prognosis were discussed. Additionally, real-time fluorescence quantitative polymerase chain reaction (RT-qPCR) was used to validate TLR3 expression in early KIRC. We also compared the expression of TLR3 in the plasma of early KIRC patients and normal controls by enzyme linked immunosorbent assay (ELISA).

Results: TLR3 expression was significantly different in multiple tumors compared with paracancerous nontumor tissues. Elevated expression of TLR3 contributed to the prolonged survival outcome in KIRC patients. Suppressed expression of TLR3 contributed to the prolonged survival outcome in LGG and PAAD patients. Moreover, TLR3 was significantly elevated in stage1, grade1 and N0 of KIRC. The expression and function of TLR3 in KIRC, LGG and PAAD were closely related to tumor immune microenvironment. TRAF6 was a key gene in the interactions between TLR3 and its interacting genes. Finally, the results of RT-qPCR and ELISA indicated that TLR3 expression levels were significantly raised in renal tissue and plasma of early KIRC patients.

Conclusion: TLR3 has the potential to be a diagnostic biomarker of KIRC, LGG and PAAD as well as a biomarker for evaluating the prognosis of KIRC, LGG and PAAD, particularly for the early diagnosis of KIRC. TLR3 affects tumors mainly by

acting on the immune microenvironment of KIRC, LGG and PAAD. These findings could lead to new insights into the immunotherapeutic targets for KIRC, LGG, and PAAD.

KEYWORDS

TLR3, KIRC, LGG, PAAD, biomarker

1 Introduction

Toll-like receptors (TLRs) are the key to identifying pathogenic microorganisms and controlling immune response (Marin-Acevedo et al., 2018). Some studies have shown that the release of damage associated molecular patterns (DAMPs) by dead cancer cells can exert an effective and persistent anti-cancer immune response by activating TLRs on host immune cells (Galluzzi et al., 2017; Garg and Agostinis, 2017; Galluzzi et al., 2020). Toll-like receptor 3 (TLR3), a crucial member of TLRs (Bianchi et al., 2017), whose absence can lead to autoimmune diseases, septicemia, chronic inflammation and cancer, among other pathological conditions (Le Naour et al., 2020). TLR3 is expressed on the endosome membrane and is a key molecule for the recognition of viral double-stranded RNA (dsRNA) (Wang et al., 2020). TLR3 is not only a target for antiviral therapy, but also a potential target for antitumor therapy, because TLR3 agonists can help initiate adaptive immunity (Cheng and Xu, 2010). Previous studies on TLR3 were mainly focused on inflammation and infection, but there were few studies on the expression of TLR3 in tumors, the effect of TLR3 on tumor prognosis and the related biological functions of TLR3.

In this study, GEPIA, DriverDBv3, UALCAN and TIMER databases were used to analyze the relationships between the expression of TLR3 in different tumor types and tumor prognosis. Moreover, we investigated the associations of TLR3 expression and somatic copy number alterations (SCNA) with the immune cells infiltration levels in kidney renal clear cell carcinoma (KIRC), brain lower grade glioma (LGG) and pancreatic adenocarcinoma (PAAD). Gene set enrichment analysis (GSEA) was performed on TLR3 of KIRC, LGG and PAAD using the LinkedOmics database. What's more, we constructed the protein-protein interaction (PPI) network and analyzed the functions of TLR3 and its interacting genes. At the same time, we also obtained the key gene connecting TLR3 and its interacting genes by FunRich. Finally, we employed real-time fluorescence quantitative polymerase chain reaction (RT-qPCR) and enzyme linked immunosorbent assay (ELISA) to verify TLR3 expression in early KIRC patients.

2 Materials and methods

2.1 GEPIA

The functions of the GEPIA website are varied, such as analyzing RNA expression in tumor and paracancer tissues and assessing cancer prognosis by gene expression (Tang et al., 2017). Patient data in GEPIA are obtained from TCGA and GTEx. We studied the expression levels of TLR3 in 33 cancers relative to normal tissues by GEPIA, and plotted the survival curve of tumors with differential expression of TLR3 (Log-rank test).

2.2 DriverDBv3

DriverDBv3, a cancer multi-omics database, contains clinical and gene-level data such as survival curve, RNA expression, miRNA expression level, somatic mutation and methylation (Liu et al., 2020). Using DriverDBv3 database, we studied the effect of TLR3 expression on the prognosis of three types of cancers (Log-rank test).

2.3 UALCAN

UALCAN is a website for analyzing the TCGA database. Users can verify the levels of gene expression in different cancer types, chart patient survival and gene expression information, and evaluate the expression of specific genes in different pathological states of cancer (Chandrashekar et al., 2017). In this study, TLR3 expression in different grades, stages and lymph node metastases of KIRC, LGG and PAAD was determined by UALCAN.

2.4 TIMER

TIMER is an immune-related cancer web server that allows users to evaluate tumor characteristics based on specific functional parameters (Li et al., 2017). By means of TIMER, we investigated the expression of TLR3 in different tumors and the correlations of TLR3 expression with immune cells infiltration levels of LGG, KIRC and PAAD (Spearman's correlation). In addition, we also studied the relationship between the somatic copy number alterations of TLR3 and the levels of tumor immune infiltration through the "SCNA" module of TIMER (Wilcoxon rank-sum test).

2.5 LinkedOmics

Linkedomics is a tool for exploring multi-omics data from multiple cancer types in the TCGA (Vasaikar et al., 2018). Using

the "Gene Set Enrichment Analysis" of LindkeOmics, "KEGG Pathway enrichment analysis" and "GO analysis" of TLR3 were performed. The "Rank Criteria", "Minimum Number of Genes (Size)" and "Simulations" were set as "meta *p*-value", "3" and "500", respectively.

2.6 STRING

STRING is an online software that can conduct direct or indirect comprehensive analysis of related genes of selected genes (Szklarczyk et al., 2019). The interacting genes of TLR3 were obtained by STRING.

2.7 GeneMANIA

GeneMANIA can be used to analyze co-expression, PPI and related functions between genes (Warde-Farley et al., 2010). In our study, the main functions of TLR3 and its interacting genes were understood through GeneMANIA.

2.8 FunRich

FunRich (3.1.3 exe), an independent software tool, is mainly used for functional enrichment and interaction network analysis of proteins and genes (Fonseka et al., 2021). We identified the most critical genes related to TLR3 and its interacting genes using FunRich. Then we made an in-depth analysis of the most critical gene in order to further understand TLR3 and its i interacting genes.

2.9 Real-time fluorescence quantitative polymerase chain reaction analysis

The First Affiliated Hospital of Guangxi Medical University provided KIRC and paracancerous tissues, which were preserved at -80°C. Total RNA Kit I (R6834, Omega) was used to extract Total RNA. RNA was reverse transcribed into cDNA in line with the instructions using the PrimeScript RT reagent kit (RR036a, Takara, Kyoto, Japan), and then we detected the cDNA using LightCycler® 96 Instrument (06924204001, Roche) and FastStart Essential DNA Green Master (Roche). Each sample was set up to repeat the determination three times. The internal reference was glyceraldehyde-3-phosphate dehydrogenase (GADPH), and the relative expression of TLR3 mRNA was determined using the $2^{-\Delta\Delta Ct}$ algorithm. (TLR3 forward primer: 5′-TTGCCTTGTATC TACTTTTGGGG-3'; TLR3 reverse primer: 5′-TCAACACTG TTATGTTTGTGGGT-3′).

2.10 Enzyme linked immunosorbent assay

Blood samples were collected from 7 patients with early KIRC before and 3 days after operation by EDTA anticoagulant tube in the First Affiliated Hospital of Guangxi Medical University. Blood samples from 7 healthy adults were collected from the physical examination department. All blood samples were centrifuged at 3,000 rpm for 20 min, and the upper plasma was collected and stored at -80°C. The expression of TLR3 protein in plasma was detected according to ELISA kit (mlbio, ml027584) instruction.

2.11 Statistical analysis

The expression of TLR3 mRNA and protein between the two groups was analyzed by *t*-test using GraphPad Prism 7. All statistical results in this study were considered to be statistically significant if $p < 0.05$.

3 Results

3.1 TLR3 expression levels in different tumor types

First of all, we investigated the expression of TLR3 in 33 kinds of tumors compared with normal tissues by TIMER. Elevated expression levels of TLR3 were observed in KIRC, while suppressed expression levels of TLR3 were observed in bladder urothelial carcinoma (BLCA), head and neck squamous cell carcinoma (HNSC), breast invasive carcinoma (BRCA), liver hepatocellular carcinoma (LIHC), colon adenocarcinoma (COAD), kidney chromophobe (KICH), lung squamous cell carcinoma (LUSC), kidney renal papillary cell carcinoma (KIRP), lung adenocarcinoma (LUAD), rectum adenocarcinoma (READ), prostate adenocarcinoma (PRAD), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA) and uterine corpus endometrial carcinoma (UCEC) (Figure 1A). Then, we used GEPIA to study the expression levels of TLR3 in 33 cancer types. The results suggested that the TLR3 expression levels in glioblastoma multiforme (GBM), LGG, KIRC, PAAD and STAD were significantly elevated compared with those in adjacent nontumor tissues, but the expression levels of TLR3 in testicular germ cell tumors (TGCT) were significantly decreased (Figure 1B). There were differences between the results generated by the TIMER and GEPIA. Because the TIMER analysis was only based on TCGA, while the GEPIA analysis was based on TCGA and GTEx data.

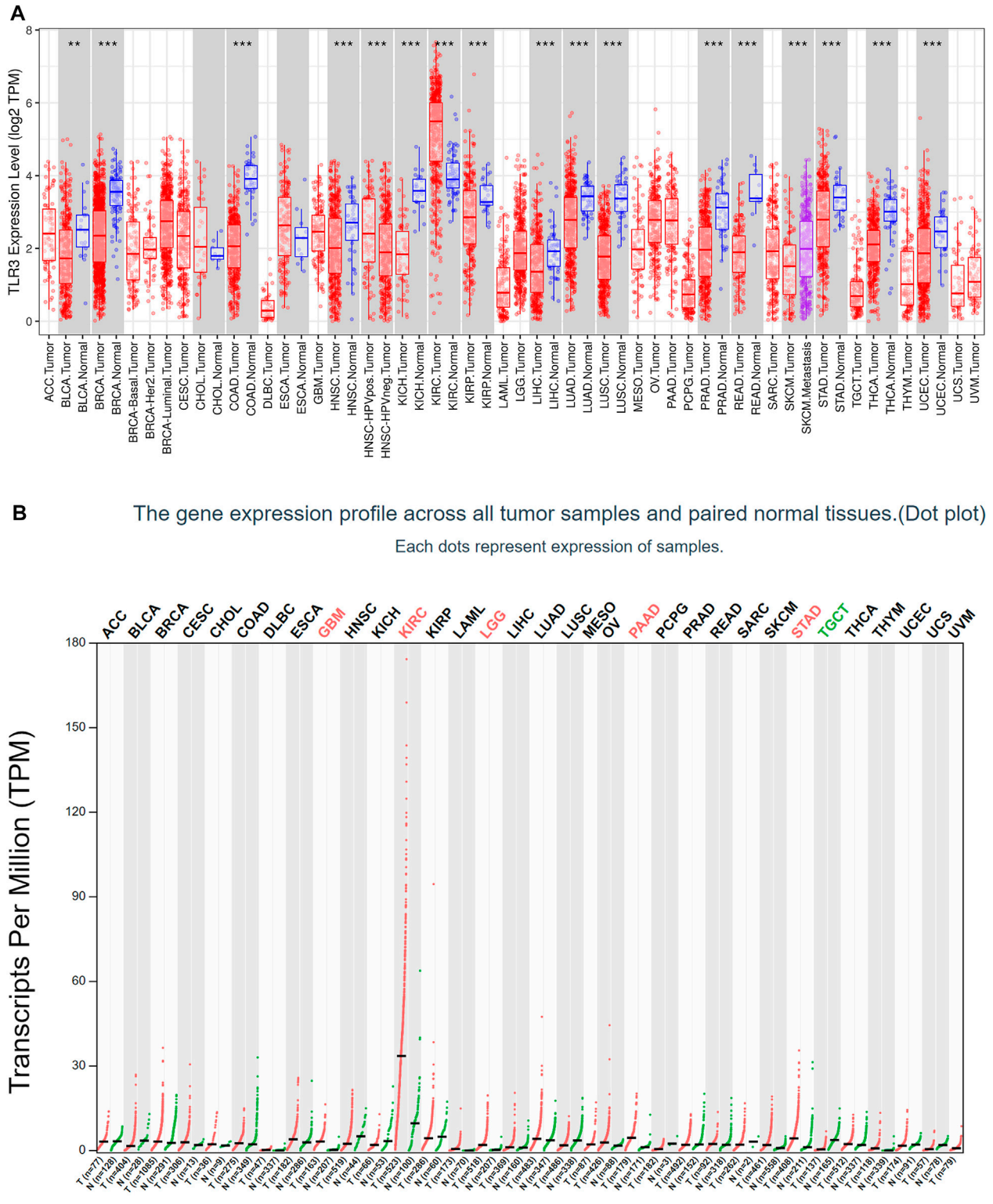


FIGURE 1
The expression of TLR3 in different cancers. **(A)** The expression levels of TLR3 gene in different cancer tissues and adjacent normal tissues were obtained by TIMER. **(B)** Expression levels of TLR3 transcript obtained from GEPIA in different cancer tissues and paired normal tissues. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

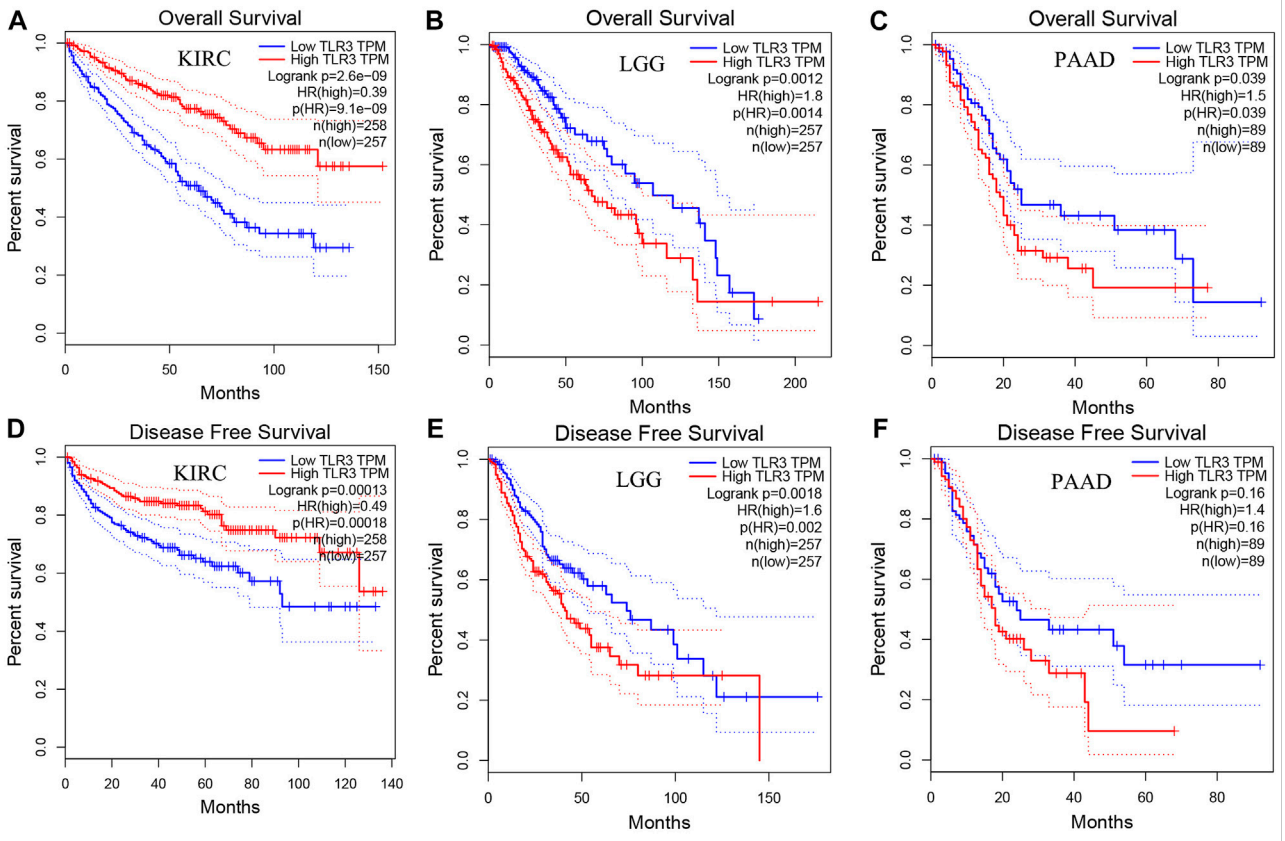


FIGURE 2
Comparing the effects of high and low expression of TLR3 on KIRC, LGG and PAAD survival outcomes by GEPIA. The effects of high and low expression of TLR3 on OS **(A–C)** and DFS **(D–F)** of KIRC, LGG and PAAD.

3.2 A biomarker of potential prognosis of KIRC, LGG and PAAD: TLR3

Next, we used GEPIA to conduct survival analysis on the above tumors with differential expression of TLR3 in order to understand the influence of TLR3 expression on their prognosis. To our surprise, although TLR3 was differentially expressed in multiple tumors compared with the corresponding paracancerous nontumor tissues, it only affected the overall survival (OS) of KIRC, LGG and PAAD. Elevated expression of TLR3 was beneficial to prolong the OS of KIRC (Figure 2A), while suppressed expression of TLR3 was beneficial to prolong the OS of LGG (Figure 2B) and PAAD (Figure 2C). The effect of TLR3 on the OS of other tumors with TLR3 differential expression was shown in Supplementary Figure S1. Next, we further investigated the associations between TLR3 and these three tumors. We found that elevated expression of TLR3 was beneficial for prolonging disease-free survival (DFS) in KIRC patients (Figure 2D), while suppressed expression of TLR3 was beneficial for prolonging DFS in LGG (Figure 2E), but TLR3 expression had no significant effect on DFS in PAAD (Figure 2F). In order to confirm the prognostic value of TLR3 expression in three kinds of tumors, we further investigated the associations of TLR3 expression with the prognosis of KIRC, LGG and PAAD by DriverDBv3. The results indicated that TLR3 overexpression was beneficial to prolong the OS (Figure 3A), platinum-free treatment interval (PFI) (Figure 3B) as well as disease specific survival (DSS) (Figure 3C) of KIRC, while suppressed expression of TLR3 was beneficial for prolonging OS (Figure 3D), PFI (Figure 3E) and DSS (Figure 3F) of LGG. Moreover, suppressed expression of TLR3 was beneficial for prolonging OS (Figure 3G) and disease-free interval (DFI) (Figure 3H) of PAAD. Therefore, all these results demonstrated that TLR3 expression was associated with the prognosis of KIRC, LGG, and PAAD. Additionally, the expression of TLR3 exhibited different effects on the survival outcome of KIRC, LGG and PAAD.

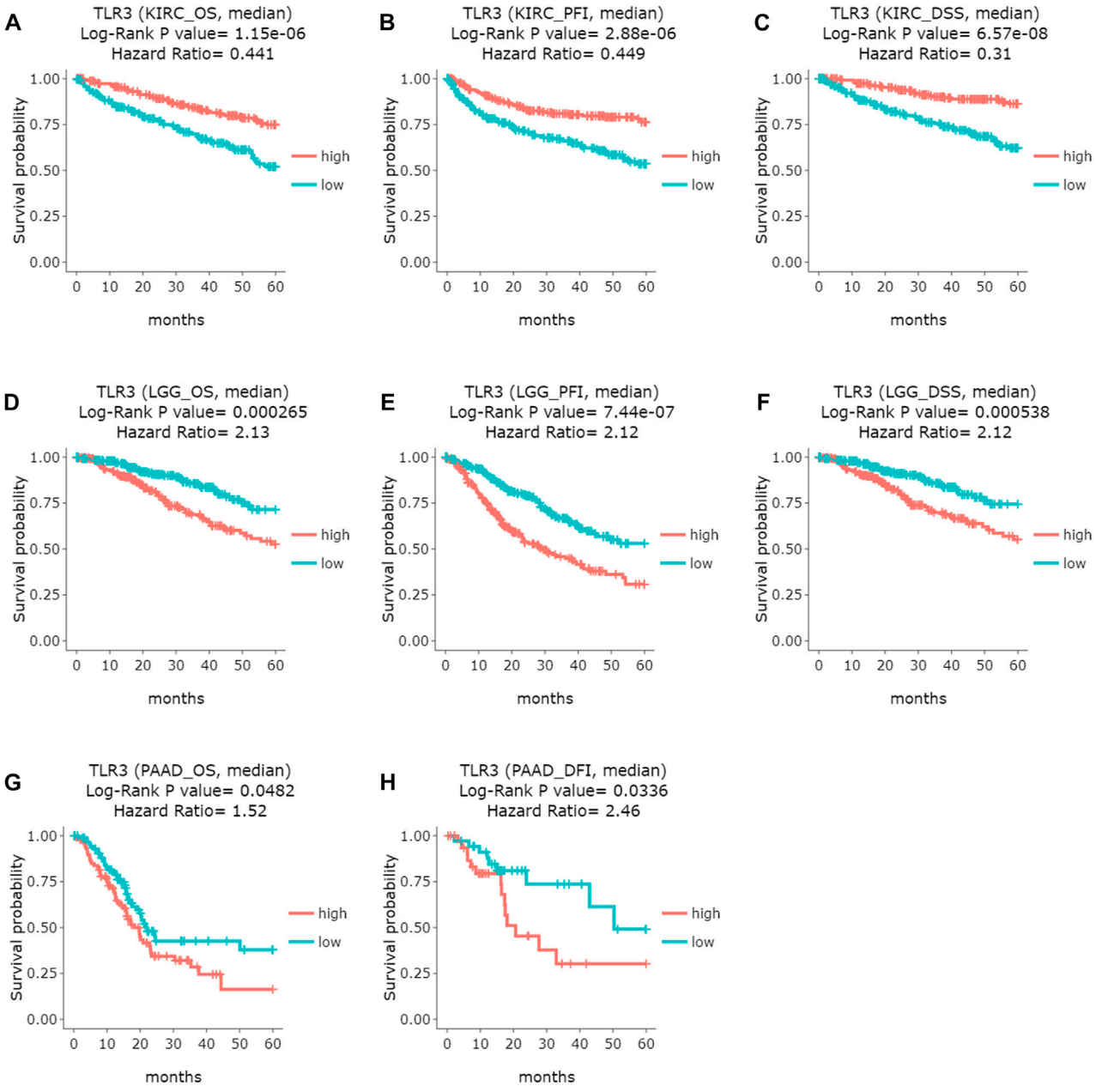


FIGURE 3
Comparing the effects of high and low expression of TLR3 on KIRC, LGG and PAAD survival outcomes by DriverDBv3. **(A–C)** The effects of high and low expression of TLR3 on OS, PFI and DFS of KIRC. **(D–F)** The effects of high and low expression of TLR3 on OS, PFI and DSS of LGG. **(G,H)** The effects of high and low expression of TLR3 on OS and DFI of PAAD.

3.3 The expression of TLR3 in different stages, grades, and lymph node metastases of KIRC, LGG, PAAD

Next, we continued to explore the expression of TLR3 in different stages, grades, and lymph node metastases of KIRC, LGG, and PAAD by UALCAN. To our surprise, the expression of TLR3 in different stages, grades and lymph node metastases of KIRC was significantly different from those in adjacent nontumor tissues. Moreover, the TLR3 expression in stage1 (Figure 4A), grade1 (Figure 4B) and N0 (Figure 4C) of KIRC was significantly increased compared with that in paracancerous nontumor tissues, which demonstrated that TLR3 was likely to be an early diagnostic biomarker for KIRC. In addition, the

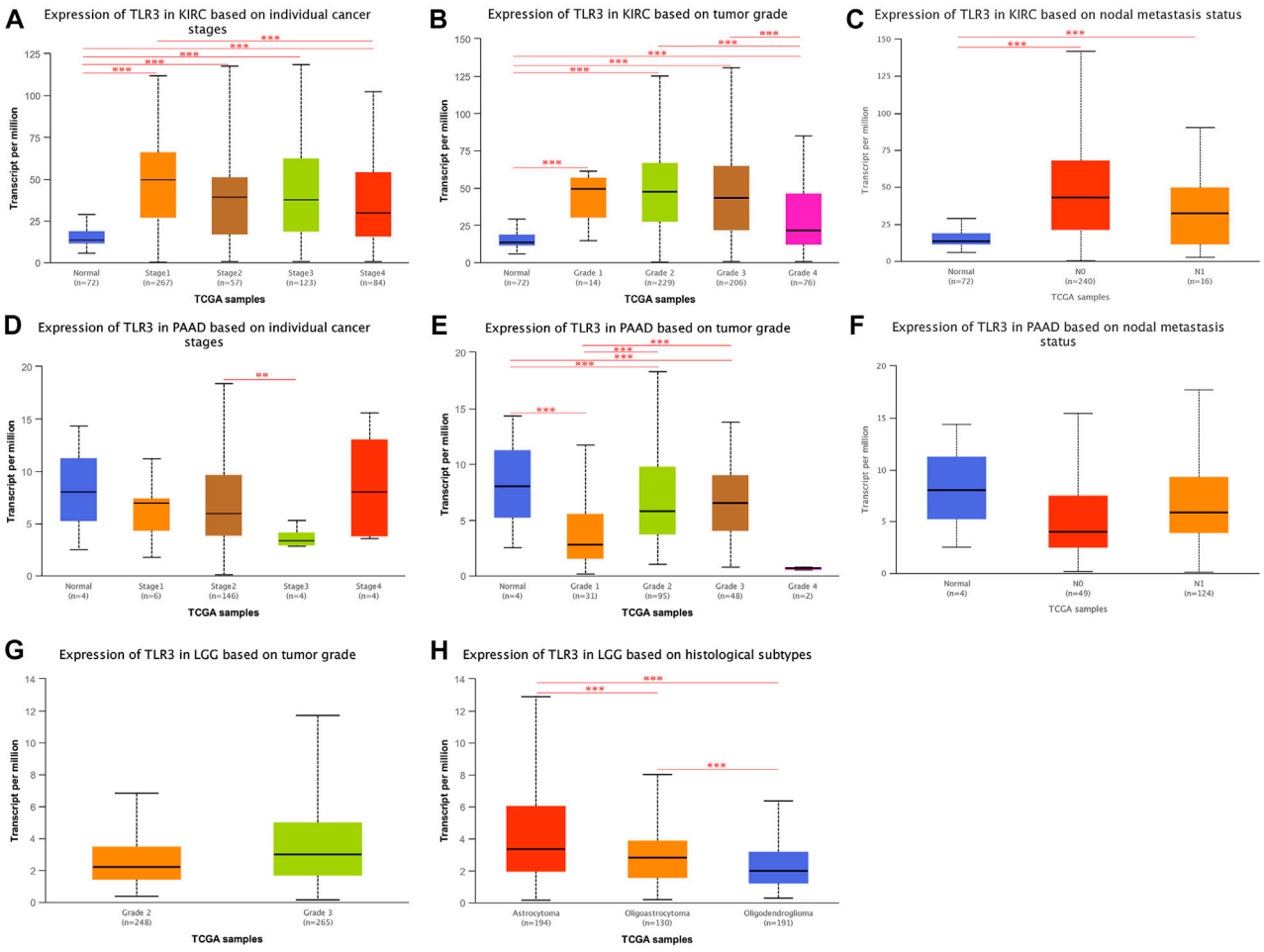


FIGURE 4
The expression of TLR3 transcript at different phases of KIRC, PAAD and LGG. **(A–C)** Box plots showing the expression of TLR3 transcript based on cancer stages, tumor grades and lymph nodal metastasis status of KIRC, respectively. **(D–F)** Box plots showing the expression of TLR3 transcript based on cancer stages, tumor grades and lymph nodal metastasis status of PAAD, respectively. **(G)** Box plot showing the expression of TLR3 transcript based on tumor grades of LGG. **(H)** Box plot showing the expression of TLR3 transcript based on histological subtypes of LGG. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TLR3 expression in stage1 of KIRC was significantly higher than that in stage4 of KIRC (Figure 4A), and the TLR3 expression in grade2 and grade3 of KIRC was significantly higher than that in grade4 of KIRC (Figure 4B). However, the expression of TLR3 in different stages and lymph node metastases of PAAD was no significant difference compared with normal tissues, but the expression of TLR3 in grade3 of PAAD was significantly higher than that in grade1 and grade2 of PAAD (Figures 4D–F). However, we considered that the number of adjacent normal samples of PAAD in UALCAN analysis was only 4, so the comparative expression of TLR3 in PAAD and adjacent normal tissues was lack of representativeness. Additionally, we found that there was no significant difference in the expression of TLR3 in different grades of LGG (Figure 4G), and there was a lack of relevant data on the expression of TLR3 in different stages and lymph node metastases of LGG. Nevertheless, we found that there were significant differences in the expression of TLR3 in different histological subtypes of LGG. These results suggested that TLR3 had the potential to be used as a diagnostic biomarker of KIRC, LGG and PAAD, especially as a biomarker for early diagnosis of KIRC and for pathological classification of LGG.

3.4 The associations of TLR3 expression and somatic copy number alterations with the immune cells infiltration levels in KIRC, LGG and PAAD

The correlations of TLR3 expression with immune cells infiltration levels in KIRC, LGG and PAAD were investigated by

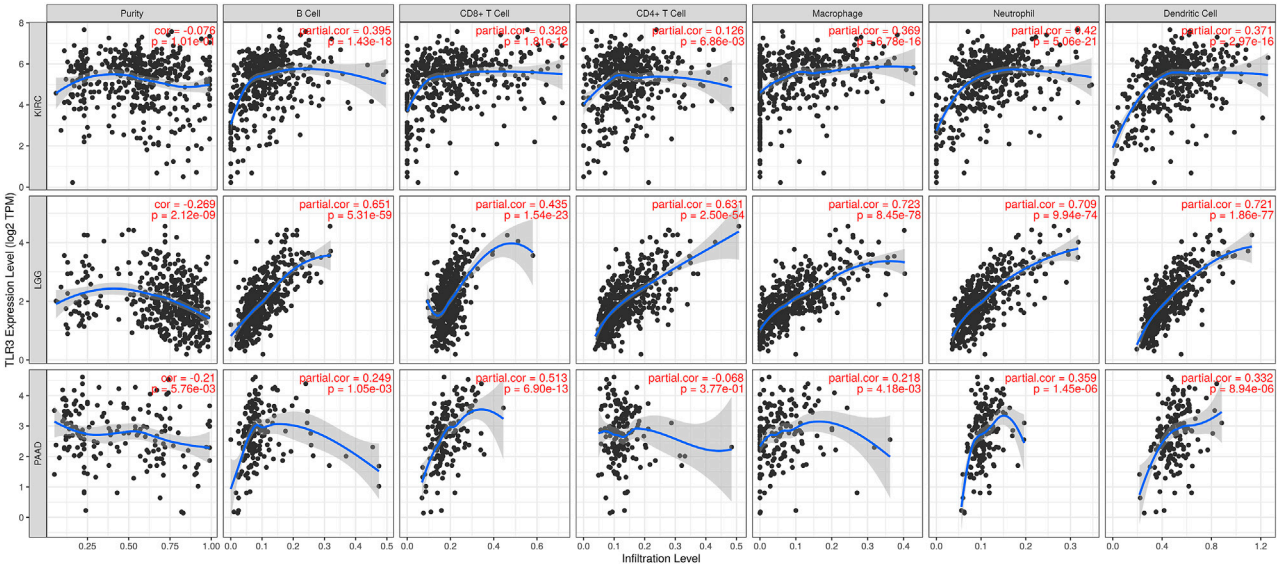


FIGURE 5
Associations of TLR3 expression with infiltration levels of immune cells in KIRC, LGG and PAAD.

TIMER. We discovered the TLR3 expression in KIRC and LGG was positively correlated with the infiltrations of CD8+T cells, B cells, neutrophils, CD4+T cells, macrophages and dendritic cells (DCs). However, TLR3 expression in PAAD was positively associated with the infiltrations of B cells, macrophages, CD8+T cells, neutrophils and DCs, but TLR3 expression was not detectable association with CD4+T cells (Figure 5). We also explored the relationships between somatic copy number alterations (SCNA) of TLR3 and the levels of immune cells infiltration in KIRC, LGG, and PAAD. Our study showed that arm-level gain of TLR3 in KIRC was significantly associated with the infiltrations of neutrophils, B cells and macrophages, while arm-level deletion of TLR3 in KIRC was significantly associated with the infiltrations of CD4+T cells and CD8+T cells. Moreover, arm-level deletion of TLR3 in LGG was apparent association with the infiltrations of DCs, macrophages, B cells, neutrophils, CD4+T cells and CD8+T cells, while deep deletion of TLR3 in LGG was obviously associated with the infiltrations of DCs and CD8+T cells. Additionally, arm-level deletion of TLR3 in PAAD was significantly associated with the infiltrations of B cells, while arm-level gain of TLR3 in PAAD was significantly associated with the infiltrations of CD4+T cells and B cells (Figure 6).

3.5 Gene set enrichment analysis of TLR3 in KIRC, LGG and PAAD

We used LinkedOmics to select data sets of KIRC, LGG and PAAD for enrichment analysis of TLR3. We explored GO and KEGG analyses of TLR3-related. Cytokine-cytokine receptor interaction, NF-kappa B signaling pathway, Th17 cell differentiation, toll-like receptor signaling pathway and natural killer cell mediated cytotoxicity were KEGG pathways mainly related to TLR3 in KIRC (Figure 7A). TLR3 in KIRC was involved in many biological processes (BP), mainly including leukocyte cell-cell adhesion, response to interferon-gamma, lymphocyte mediated immunity, cellular defense response and mast cell activation, (Figure 7B). In KIRC, the main cellular components (CC) involved in TLR3 were MHC protein complex, immunological synapse and receptor complex (Figure 7C). Antigen binding, cytokine receptor binding and pattern recognition receptor activity were the main molecular functions (MF) of TLR3 in KIRC (Figure 7D). In LGG and PAAD, the results of GO and KEGG analyses of TLR3 were shown in Figures 7E-L.

3.6 PPI network and functional analysis of TLR3 and its interacting genes

We constructed a PPI network of TLR3 through STRING. The PPI network showed a complex association of TLR3 with other genes (Figure 8A). Next, we performed functional analysis on TLR3 and its interacting genes through GeneMANIA. The results demonstrated that the functions of TLR3 and its interacting genes were mainly focused on pattern recognition receptor signaling pathway, I-kappaB kinase/NF-kappaB signaling, toll-like receptor signaling pathway, tumor necrosis factor superfamily cytokine production, positive regulation of

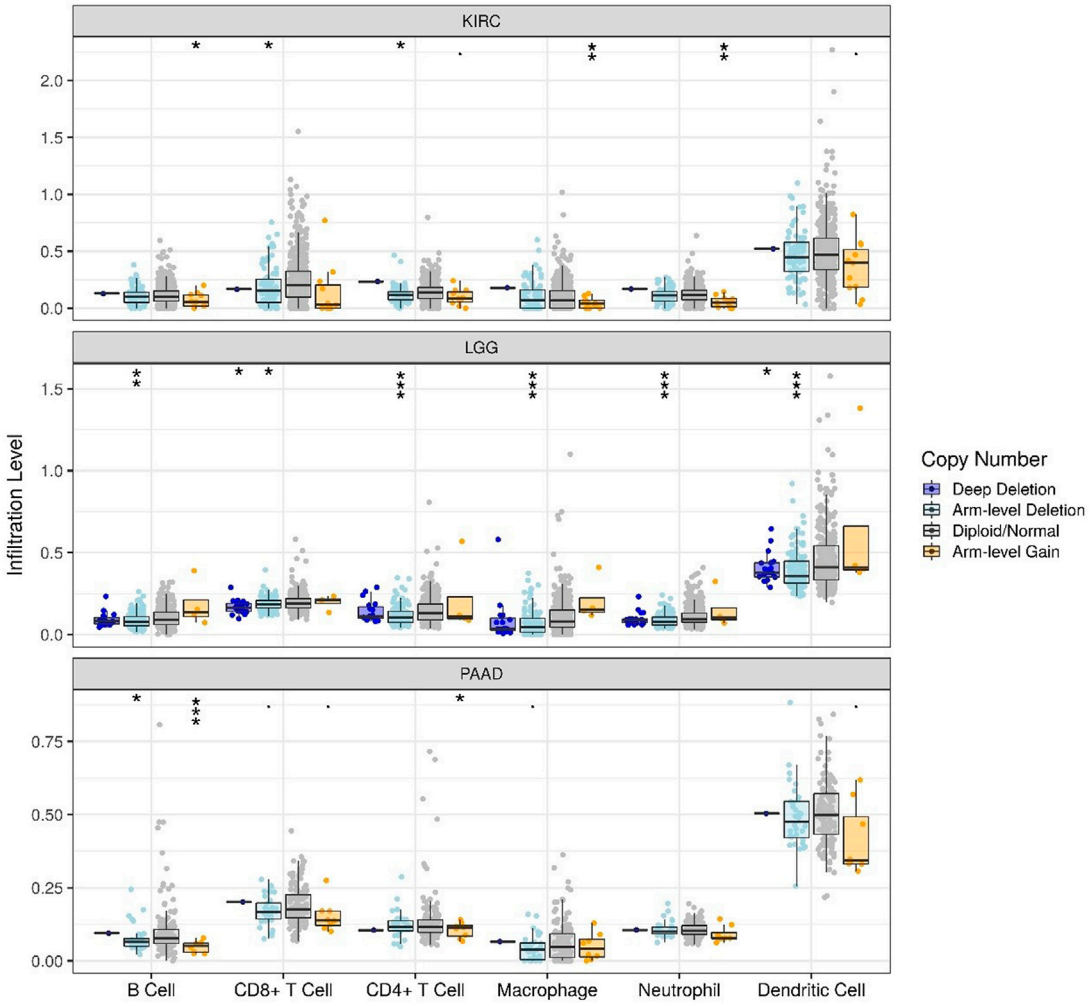


FIGURE 6
Associations of immune cells infiltration levels with somatic copy number alterations for TLR3 in KIRC, LGG and PAAD. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

defense response, programmed necrotic cell death and activation of protein kinase activity (Figure 8B). These results suggested that TLR3 and its interacting genes were closely related to immune and anti-tumor effects. Next, we used FunRich to obtain the key gene connecting TLR3 and its interacting genes. These results demonstrated that TRAF6 was a key gene in the interactions between TLR3 and its interacting genes (Figure 8C).

3.7 Verification the expression levels of TLR3 in early KIRC patients

Last but not least, we verified the TLR3 mRNA expression levels between tumor tissues and paracancerous nontumor tissues of early KIRC patients. The characteristics of patients used to study the expressions of TLR3 mRNA were shown in Table1. According to RT-qPCR data, the relative expression levels of TLR3 mRNA in early KIRC were considerably higher than that in paracancerous nontumor tissues (Figure 9A). The melting peaks of RT-qPCR were shown in Figure 9B, which demonstrated that the designed primers were specific. Moreover, we compared the expression of TLR3 protein in plasma of patients with early KIRC and healthy adults by ELISA. The basic information of KIRC patients and healthy adults used in the ELISA was shown in Supplementary Table S1. The results of ELISA showed that the expression of TLR3 protein in plasma of patients with KIRC before operation was significantly higher than that of healthy adults, and the expression of TLR3 protein in plasma of patients with KIRC 3 days after operation was significantly lower than that before operation, but still higher than that of healthy adults (Figure 9C).

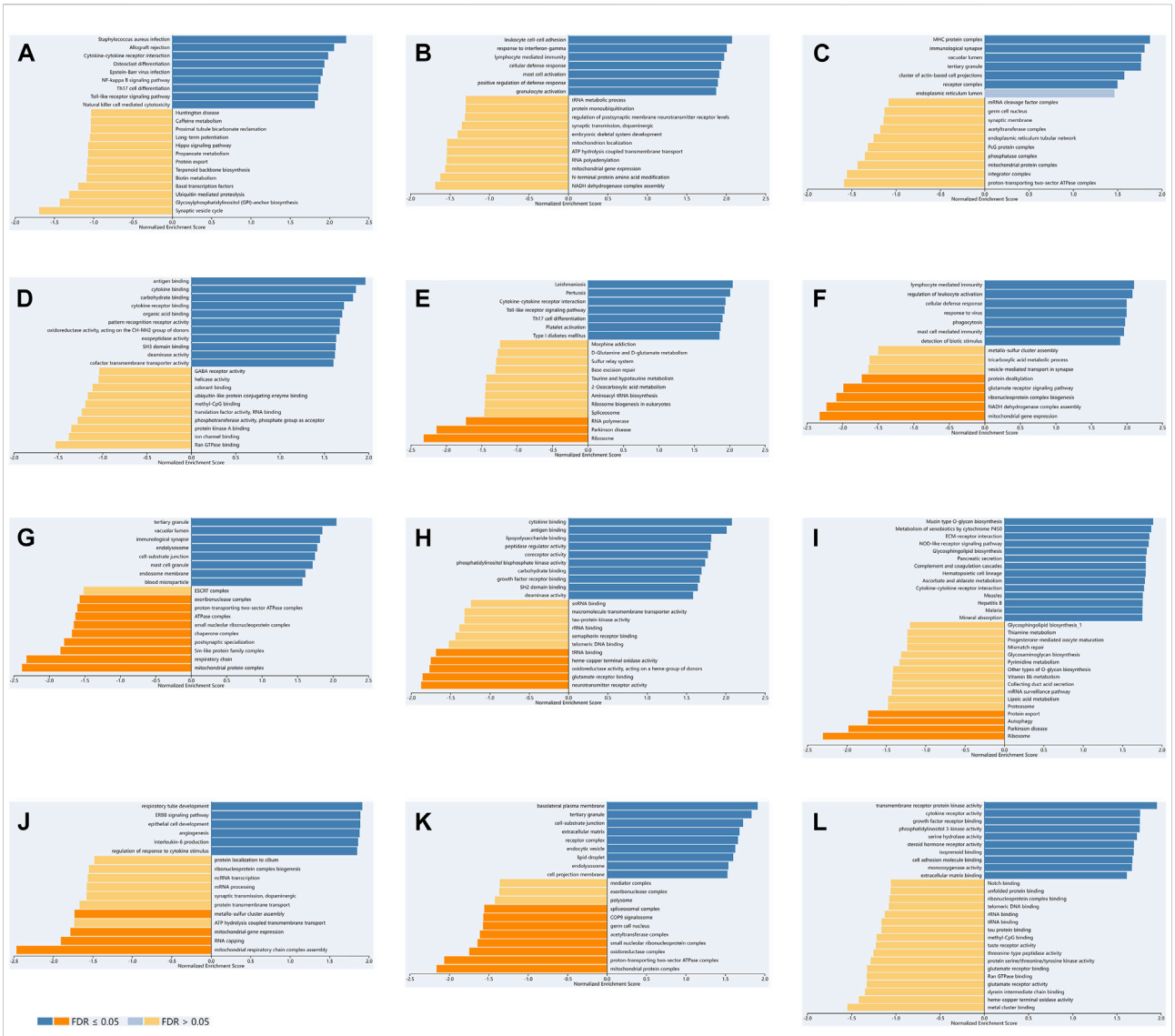
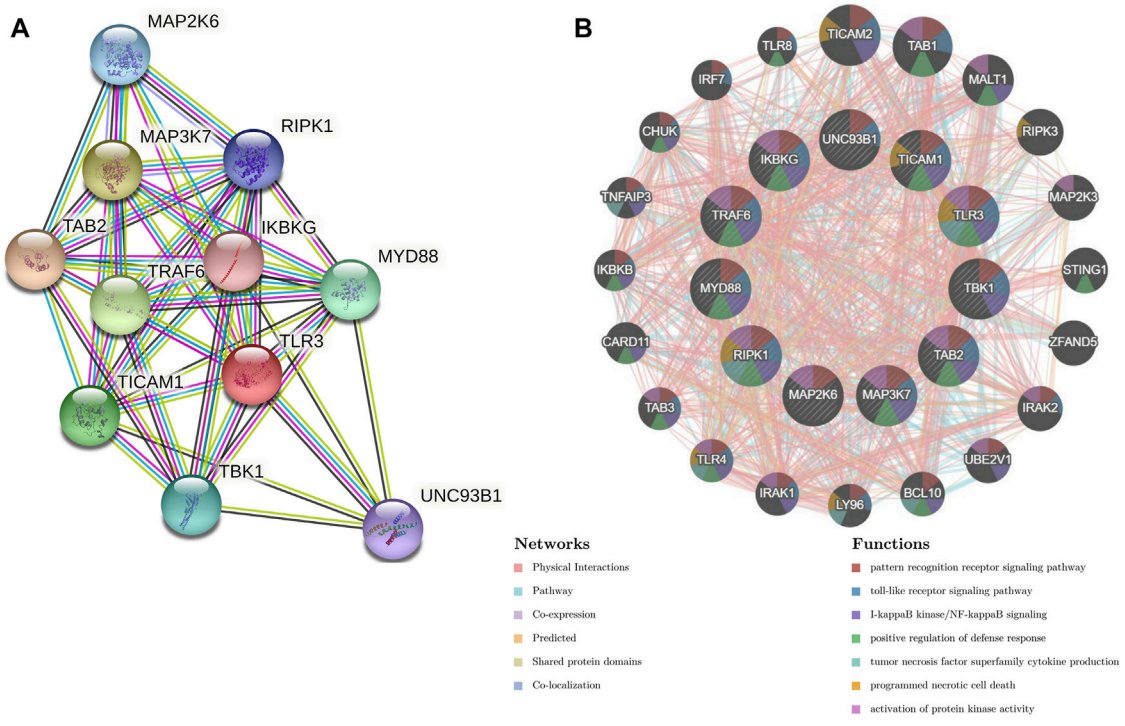


FIGURE 7
The GO and KEGG analyses of TLR3 in KIRC, LGG and PAAD based on LinkedOmics. KEGG **(A)**, BP **(B)**, CC **(C)** and MF **(D)** of TLR3 in KIRC. KEGG **(E)**, BP **(F)**, CC **(G)** and MF **(H)** of TLR3 in LGG. KEGG **(I)**, BP **(J)**, CC **(K)** and MF **(L)** of TLR3 in PAAD.

4 Discussion

The most common cancer in renal cell carcinoma (RCC) is KIRC, which accounts for around 75% of all RCC (Linehan and Ricketts, 2019). Early stage of KIRC can be surgically excised to achieve excellent results, however, treating advanced KIRC continues to be a serious challenge (Petitprez et al., 2021). In addition, clinicians currently rely on the stage of tumor lymph node metastasis (TNM) to estimate the prognosis of KIRC patients (Pichler et al., 2013; Qin et al., 2013; Zheng et al., 2017), but the prognosis of patients with the same TNM stage may vary greatly (Wei et al., 2019). However, the accuracy of prognosis prediction can be improved by stratifying the prognosis using biomarkers (such as the expressions of specific genes) [(Tamayo et al., 2011); (Zou and Mo, 2021)], therefore, it is necessary to search for specific biomarkers for prognosis and early diagnosis of KIRC, as well as for treatment of KIRC.

Glioma, a highly heterogeneous tumor (Poff et al., 2019), is one of the most prevalent primary brain tumors, accounting for about 70% of malignant brain tumors (Gusyatiner and Hegi, 2018). At present, more and more young people suffer LGG of WHO grade II and III (Gusyatiner and Hegi, 2018). Although the malignant features of LGG are less than those of high-grade gliomas (WHO grade IV), and the clinical prognosis of LGG is relatively better (Claus et al., 2015), there are no biomarkers that



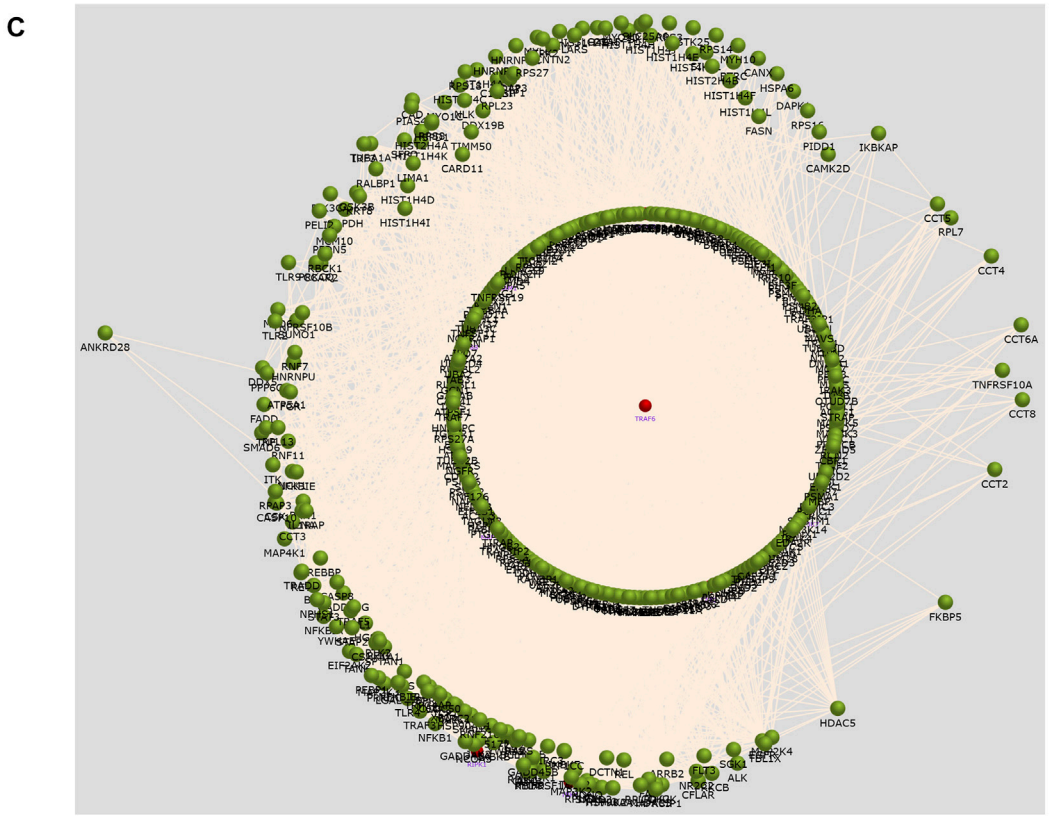


FIGURE 8
The PPI network and function of TLR3 and its interacting genes and the key gene that regulates their expression. **(A)** The PPI network of TLR3. **(B)** The PPI network and functions of TLR3 and its interacting genes. **(C)** Interaction plot showing that the key gene affecting TLR3 and its interacting genes is TRAF6.

TABLE 1 The characteristics of patients for RT-qPCR.

Patients	Sex	Years of age	Tumor location	Tumor size (cm)	TNM stage	Histological type
Patient1	Male	40	Right	4.8 × 4 × 3.5	$T_1N_0M_0$	KIRC
Patient2	Male	50	Right	2.5 × 2 × 2	$T_1N_0M_0$	KIRC
Patient3	Female	41	Left	4.5 × 4 × 4	$T_1N_0M_0$	KIRC

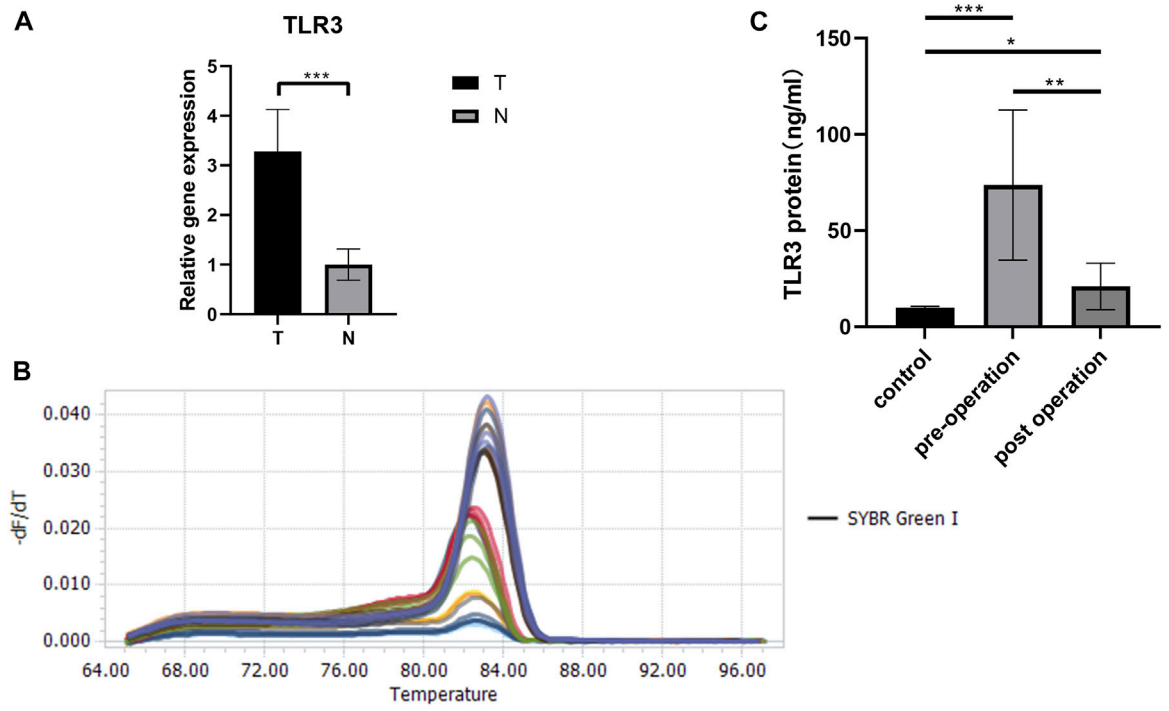


FIGURE 9
The results of RT-qPCR and ELISA. **(A)** The expression levels of TLR3 mRNA in early stage of KIRC tissues and adjacent normal tissues. **(B)** The melting peaks of RT-qPCR. **(C)** Comparison of TLR3 protein levels in plasma of patients with early KIRC before operation, 3 days after operation and healthy adults. T: KIRC tissues; N: adjacent normal tissues; control: healthy adults; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

can accurately determine the prognosis of LGG. If we can find a biomarker that can accurately judge the prognosis of LGG patients, it will be of great help to the management and treatment of LGG.

PAAD is a devastating tumor disease that is becoming more common and has the lowest 5-year survival rate of cancers (Siegel et al., 2017). There are similarities between chronic pancreatitis and PAAD in connective tissue hyperplasia and inflammatory infiltration (Zheng et al., 2013). At the same time, there is growing evidence that the evolution and maintenance of PAAD are influenced by complicated inflammatory and immunosuppressive environments (Wormann et al., 2014). Therefore, further studies on the association between PAAD and immunity can improve our understanding of the pathogenesis of PAAD and contribute to the development of immunotherapy for PAAD.

TLR3, mainly located in the endoplasmic reticulum and intracellular endocytic compartment, can recognize oligonucleotides from itself and microorganisms (Hussein et al., 2014). TLR3 is expressed in a variety of immune cells, such as dendritic cells (DCs), natural killer cells (NK) cells and macrophages, and TLR3 exerts a momentous effect in the activation of innate immunity (Iwasaki and Medzhitov, 2004). However, TLR3 can also be activated in non-immune cells, such as fibroblasts and vascular endothelial cells (Pirher et al., 2017). Interestingly, there is growing evidence that TLR3 is also generated in cancer cells, and the generation of TLR3 in cancer cells may exert an opposite effect in different cancer

progression (Zheng et al., 2021). Therefore, we investigated TLR3 expression in different cancers and its effect on cancer prognosis. At the same time, we studied the interacting genes and related functions of TLR3 in order to understand the possible mechanisms of TLR3 affecting cancers.

First of all, we explored TLR3 expression levels in multiple cancers using GEPIA and TIMER databases. Compared with paracancerous nontumor tissues, the expression levels of TLR3 were significantly elevated in GBM, KIRC, LGG, PAAD and STAD, while suppressed in TGCT, COAD, BLCA, KICH, BRCA, HNSC, LUAD, PRAD, READ, KIRP, LIHC, THCA, LUSC, STAD and UCEC. Secondly, we found that TLR3 was differentially expressed in varieties of tumors compared with paracancerous nontumor tissues, but it only affected the survival time of KIRC, LGG and PAAD. Overexpression of TLR3 was associated with prolonged PFI, DFS, OS and DSS in patients with KIRC, while suppressed expression of TLR3 was associated with prolonged PFI, DFS, OS and DSS in patients with LGG, and suppressed expression of TLR3 was associated with prolonged DFI and OS in patients with PAAD. These results suggested that TLR3 played different biological roles in different tumors. Similar conclusions could be drawn from other cancer studies. For example, high expression of TLR3 prolonged the survival time of patients with non-small-cell lung cancer by inducing apoptosis of cancer cells (Bianchi et al., 2020), but high expression of TLR3 could promote the migration of breast cancer cells by increasing the expression of E-cadherin (Bondhopadhyay et al., 2015). Next, we found that the TLR3 expression in stage1, grade1 and N0 of KIRC was significantly increased compared with that in paracancerous nontumor tissues, and the TLR3 expression in stage1 of KIRC was significantly higher than that in stage4 of KIRC, and the TLR3 expression in grade2 and grade3 of KIRC was significantly higher than that in grade4 of KIRC. These results indicated that the TLR3 expression in early and low grade KIRC was significantly higher than that in advanced and high grade KIRC. In addition, the expression of TLR3 in grade3 of PAAD was significantly higher than that in grade1 and grade2 of PAAD. And there were significant differences in the expression of TLR3 in different histological subtypes of LGG. We also verified that there was a significant elevation in the expression of TLR3 mRNA in early KIRC by RT-qPCR. To our excitement, the results of ELISA showed that the expression of TLR3 protein in plasma of patients with early KIRC before operation was significantly higher than that of healthy adults, and the expression of TLR3 protein in plasma of patients with early KIRC 3 days after operation was significantly lower than that before operation, but still higher than that of healthy adults, which indicated that the occurrence of KIRC could directly lead to the increase of TLR3 protein in plasma. Taking all of the results together, we concluded that TLR3 had the potential to be used as a prognostic biomarker of KIRC, LGG and PAAD, especially as a biomarker for early diagnosis of KIRC and for pathological classification of LGG. The prognostic value of TLR3 in KIRC was similar to the results of Liao et al. (Liao et al., 2021), but our results emphasized the changes of TLR3 in early KIRC and the potential role of TLR3 as an early diagnostic biomarker of KIRC.

For the sake of further exploring the possible mechanisms of TLR3 affecting the prognosis of cancers, we also explored the correlations of TLR3 expression with the immune cells infiltration levels in KIRC, LGG and PAAD. The expression of TLR3 in KIRC and LGG was positively correlated with the infiltration levels of CD4+T cells, CD8+T cells, B cells, neutrophils, macrophages and dendritic cells (DCs). On the other hand, TLR3 expression in PAAD was positively correlated with the infiltration levels of B cells, macrophages, CD8+T cells, neutrophils and DCs, but TLR3 expression was not significantly correlated with CD4+T cells. Therefore, the associations of TLR3 with levels of immune cells infiltration suggested that TLR3 played a vital role in regulating tumor immunology of KIRC, LGG and PAAD. Some studies had shown that tumor immune cells infiltration had different prognostic value in different cancers (Zhang et al., 2019; Zuo et al., 2020). In KIRC patients, B cells infiltration prolonged tumor-specific survival (Stenzel et al., 2020), CD8+T cells infiltration prolonged patients' OS outcome (Zhang et al., 2019), elevated neutrophils abundance was associated with favorable prognosis (Niu et al., 2021), what's more, infiltration of dendritic cells and T cells and elevated adaptive immune response could effectively inhibit tumor recurrence and metastasis [(Lim et al., 2007); (Ghatalia et al., 2019)]. Therefore, it is reasonable to speculate that the high expression of TLR3 in KIRC tissues may be an active immune response, which is consistent with the conclusion of a recent study on esophageal squamous cell carcinoma (Su et al., 2022). The study of esophageal squamous cell carcinoma demonstrated that a gene highly expressed in tumors and positively related to prognosis recruited a variety of anti-tumor immune cells for the tumor immune microenvironment. However, the specific reason of up-regulated TLR3 expression in KIRC is still unclear, which needs further study. A study on colon cancer may provide us with some inspiration (Cao et al., 2021). For example, further study on the interaction between TLR3 overexpression cells and other cells through single cell RNA sequencing data may help us to understand the specific mechanism of TLR3 up-regulation in KIRC and the relationship between the special expression pattern of TLR3 and the prognosis of KIRC. In LGG patients, elevated levels of immune cells infiltration were associated with poorer prognosis [(Feng et al., 2020); (Zhang et al., 2020)]. In PAAD patients, elevated levels of immune cells infiltration were associated with decreased OS (Ju et al., 2020), and patients at high risk had high levels of immune cells infiltration (Xu et al., 2021). Our results were similar to the results of these

predecessors, but further confirmed the link between three types of cancer and immunity. Moreover, we also explored the correlations between SCNA of TLR3 and the immune cells infiltration levels in KIRC, LGG, and PAAD. Our study showed that arm-level gain of TLR3 in KIRC was significantly associated with the infiltrations of neutrophils, B cells and macrophages, while arm-level deletion of TLR3 in KIRC was significantly associated with the infiltrations of CD4+T cells and CD8+T cells. Moreover, arm-level deletion of TLR3 in LGG was considerably associated with the infiltrations of DCs, macrophages, B cells, neutrophils, CD8+T cells and CD4+T cells, while deep deletion of TLR3 in LGG was significantly associated with the infiltration levels of DCs and CD8+T cells. Additionally, arm-level deletion of TLR3 in PAAD was considerably associated with the infiltrations of B cells, while arm-level gain of TLR3 in PAAD was significantly correlated with the infiltrations of CD4+T cells and B cells. These results further supported a strong association of TLR3 with the levels of immune cells infiltration in three types of cancer.

For a better understanding of TLR3 in KIRC, LGG, and PAAD, we explored GO and KEGG analyses of TLR3-related. Although the results of GO and KEGG analyses of TLR3 were different in KIRC, LGG and PAAD, one common feature of all results was that TLR3 was closely related to the immune process. TLR3 was involved in a variety of BP, CC and MF, mainly including response to interferon-gamma, mast cell activation, macrophage activation, pattern recognition receptor activity and cytokine binding. Interferon is a kind of molecule with multiple effects, including anti-tumor effect and cancer-promoting effect, which plays a vital role in tumor immune microenvironment [(Castro et al., 2018); (Snell et al., 2017)]. Activation of mast cells and macrophages can be used in cancer treatment (Eissmann et al., 2019). KEGG analysis identified the signaling pathways involved by TLR3, mainly including Th17 cell differentiation, toll-like receptor signaling pathway and NOD-like receptor signaling pathway. The imbalance of NOD-like receptor activation is involved in the pathogenesis of tumors, but its role in the development and progression of different cancers is quite different (Moossavi et al., 2018). The results of our study and previous researches demonstrated that TLR3 was closely related to the tumor immune microenvironment (Zou et al., 2022), and this study provided an explanation for the different prognostic outcome of TLR3 in different cancers.

Additionally, we investigated the TLR3 interacting genes and performed functional analysis on TLR3 and its interacting genes, the results demonstrated that the functions of TLR3 and its interacting genes were mainly focused on pattern recognition receptor signaling pathway, toll-like receptor signaling pathway, I-kappa B kinase/NF-kappa B signaling, positive regulation of defense response, tumor necrosis factor superfamily cytokine production, programmed necrotic cell death and activation of protein kinase activity. NF-kappa B is associated with the origin and progression of malignancies [(Jin et al., 2019); (Wang et al., 2019)]. Tumor necrosis factor has an effect on tumor-associated macrophages and cancer cells, thus affecting the prognosis of patients (Cassetta et al., 2019). The effect of programmed necrotic cell death can not only confer on tumor growth advantage, but also make tumor cells necrotic and vulnerable (Lin et al., 2020). Protein kinase activity exerts a critical role in antioxidant stress and cancer development [(Xu et al., 2019a); (Xu et al., 2020)]. These results once again provide a theoretical basis for TLR3 to play a different prognostic value in different cancers. Encouragingly, our study found that TRAF6 was a key gene in the interactions between TLR3 and its interacting genes. TRAF6, a key inflammatory mediator, mediates tumor growth and metastasis and effectively reduces the migration and activation of macrophages [(Park et al., 2020); (Seijkens et al., 2018)]. Additionally, TRAF6 plays an important role in the activation of TLR3 (Jiang et al., 2003). Studies have reported that immune checkpoint proteins affect tumor occurrence and development by regulating TRAF6, regulating toll-like receptor signal pathway and stimulating tumor immune microenvironment (Xu et al., 2019b). These results comprehensively demonstrated the important role of TLR3 in tumor immune microenvironment.

Previous studies on TLR3 were mainly focused on infectious diseases and inflammation-related diseases, but there were few studies on cancers. Our study enriched the associations between TLR3 and cancers and proved that TLR3 had the potential to be a diagnostic and prognostic biomarker for KIRC, LGG and PAAD, especially an early diagnosis biomarker of KIRC. However, the role and molecular mechanisms of TLR3 in cancer are quite complex, but there were significant correlations between the expression and SCNA of TLR3 and levels of immune cells infiltration of KIRC, LGG and PAAD. Moreover, GSEA of TLR3 in KIRC, LGG and PAAD showed that TLR3 was closely related to tumor immune microenvironment. The PPI network and functional analysis of TLR3 and its interacting genes also proved the close relationship between TLR3 and tumor immune microenvironment. We have every reason to speculate that TLR3 affects the prognosis of KIRC, LGG and PAAD by affecting tumor immune microenvironment, although TLR3 is also likely to influence the prognosis of patients through other factors. On the other hand, it is undeniable that our study still has some limitations. For example, we did not verify the association of PAAD and LGG with TLR3, and the sample size of comparing the expression levels of TLR3 protein in plasma of early KIRC patients and healthy adults needs to be further increased. The expression of TLR3 mRNA in early KIRC and adjacent normal tissues also needs to be further verified. More experimental and clinical researches are necessary to determine our findings about the effects of TLR3 on KIRC, LGG and PAAD. Next, we

will collect blood and tissue samples from KIRC, LGG, and PAAD patients. With the continuous collection of samples, we will increase the sample size to verify our results in future studies.

5 Conclusion

Compared with the corresponding normal tissues, TLR3 expression levels were significantly increased in KIRC, LGG and PAAD. Moreover, the expression of TLR3 was significantly higher in early KIRC compared with that in paracancerous nontumor tissues. The occurrence of KIRC could directly lead to the increase of TLR3 protein in plasma. Elevated expression of TLR3 was beneficial to prolong the OS, DFS, PFI and DSS of KIRC. Suppressed expression of TLR3 was beneficial to prolong the OS, DFS, PFI and DSS of LGG. Suppressed expression of TLR3 was beneficial for prolonging OS and DFI of PAAD. The expression and function of TLR3 in KIRC, LGG and PAAD were closely related to tumor immune microenvironment. TRAF6 was a key gene in the interactions between TLR3 and its interacting genes. Considering the changes of TLR3 expression levels in KIRC, LGG and PAAD, as well as the influence of TLR3 expression on the prognosis of these three cancers, and the close correlation between TLR3 and the tumor immune microenvironment, TLR3 may become a potential therapeutic target and an important molecular biomarker for judging the prognosis of KIRC, LGG and PAAD, as well as a diagnostic biomarker of KIRC, LGG and PAAD, especially an early diagnostic biomarker of KIRC.

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 The First Affiliated Hospital of Guangxi Medical University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

XZ, YG, and ZM contributed to conception and design of the study. XZ and YG performed RT-qPCR and ELISA. XZ, YG, and ZM performed the statistical analysis. XZ and YG collected the samples. XZ and YG wrote the manuscript. ZM revised the manuscript. All authors have read and approved the manuscript.

Funding

This work was supported by the grants from the National Natural Science Foundation of China (81770759), the National Key Research and Development Program of China (2017YFC0908000), Major Project of Guangxi Innovation Driven (AA18118016), Guangxi key Laboratory for Genomic and Personalized Medicine (grant number 16-380-54, 17-259-45, 19-050-22, 19-185-33, 20-065-33), Guangxi Science and Technology Base and Talent Project (2019AC17009) and Guangxi Clinical Research Center for Urology and Nephrology (2020AC03006).

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.2022.905988/full#supplementary-material

References

Bianchi, F., Alexiadis, S., Camisaschi, C., Truini, M., Centonze, G., Milione, M., et al. (2020). TLR3 expression induces apoptosis in human non-small-cell lung cancer. *Int. J. Mol. Sci.* 21 (4), E1440. doi:10.3390/ijms21041440

Bianchi, F., Pretto, S., Tagliabue, E., Balsari, A., and Sfondrini, L. (2017). Exploiting poly(I:C) to induce cancer cell apoptosis. *Cancer Biol. Ther.* 18 (10), 747–756. doi:10.1080/15384047.2017.1373220

Bondhopadhyay, B., Moirangthem, A., and Basu, A. (2015). Innate adjuvant receptor Toll-like receptor 3 can promote breast cancer through cell surface. *Tumour Biol.* 36 (2), 1261–1271. doi:10.1007/s13277-014-2737-8

Cao, Y., Jiao, N., Sun, T., Ma, Y., Zhang, X., Chen, H., et al. (2021). CXCL11 correlates with antitumor immunity and an improved prognosis in colon cancer. *Front. Cell Dev. Biol.* 9, 646252. doi:10.3389/fcell.2021.646252

Cassetta, L., Fragkogianni, S., Sims, A. H., Swierczak, A., Forrester, L. M., Zhang, H., et al. (2019). Human tumor-associated macrophage and monocyte transcriptional landscapes reveal cancer-specific reprogramming, biomarkers, and therapeutic targets. *Cancer Cell* 35 (4), 588–602. doi:10.1016/j.ccell.2019.02.009

Castro, F., Cardoso, A. P., Goncalves, R. M., Serre, K., and Oliveira, M. J. (2018). Interferon-gamma at the crossroads of tumor immune surveillance or evasion. *Front. Immunol.* 9, 847. doi:10.3389/fimmu.2018.00847

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., 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

Cheng, Y. S., and Xu, F. (2010). Anticancer function of polyinosinic-polycytidylic acid. *Cancer Biol. Ther.* 10 (12), 1219–1223. doi:10.4161/cbt.10.12.13450

Claus, E. B., Walsh, K. M., Wiencke, J. K., Molinaro, A. M., Wiemels, J. L., Schildkraut, J. M., et al. (2015). Survival and low-grade glioma: The emergence of genetic information. *Neurosurg. Focus* 38 (1), E6. doi:10.3171/2014.10.FOCUS12367

Eissmann, M. F., Dijkstra, C., Jarnicki, A., Phesse, T., Brunnberg, J., Poh, A. R., et al. (2019). IL-33-mediated mast cell activation promotes gastric cancer through macrophage mobilization. *Nat. Commun.* 10 (1), 2735. doi:10.1038/s41467-019-10676-1

Feng, P., Li, Z., Li, Y., and Zhang, Y. (2020). Molecular dynamics simulations of a chimeric androgen receptor protein (SPARKI) confirm the importance of the dimerization domain on DNA binding specificity. *Front. Mol. Biosci.* 7, 4. doi:10.3389/fmolb.2020.00004

Fonseka, P., Pathan, M., Chitti, S. V., Kang, T., and Mathivanan, S. (2021). FunRich enables enrichment analysis of OMICs datasets. *J. Mol. Biol.* 433 (11), 166747. doi:10.1016/j.jmb.2020.166747

Galluzzi, L., Buque, A., Kepp, O., Zitvogel, L., and Kroemer, G. (2017). Immunogenic cell death in cancer and infectious disease. *Nat. Rev. Immunol.* 17 (2), 97–111. doi:10.1038/nri.2016.107

Galluzzi, L., Vitale, I., Warren, S., Adjemian, S., Agostinis, P., Martinez, A. B., et al. (2020). Consensus guidelines for the definition, detection and interpretation of immunogenic cell death. *J. Immunother. Cancer* 8 (1), e000337. doi:10.1136/jitc-2019-000337

Garg, A. D., and Agostinis, P. (2017). Cell death and immunity in cancer: From danger signals to mimicry of pathogen defense responses. *Immunol. Rev.* 280 (1), 126–148. doi:10.1111/imr.12574

Ghatalia, P., Gordetsky, J., Kuo, F., Dulaimi, E., Cai, K. Q., Devarajan, K., et al. (2019). Prognostic impact of immune gene expression signature and tumor infiltrating immune cells in localized clear cell renal cell carcinoma. *J. Immunother. Cancer* 7 (1), 139. doi:10.1186/s40425-019-0621-1

Gusyatiner, O., and Hegi, M. E. (2018). Glioma epigenetics: From subclassification to novel treatment options. *Semin. Cancer Biol.* 51, 50–58. doi:10.1016/j.semcancer.2017.11.010

Hussein, W. M., Liu, T. Y., Skwarczynski, M., and Toth, I. (2014). Toll-like receptor agonists: A patent review (2011 - 2013). *Expert Opin. Ther. Pat.* 24 (4), 453–470. doi:10.1517/13543776.2014.880691

Iwasaki, A., and Medzhitov, R. (2004). Toll-like receptor control of the adaptive immune responses. *Nat. Immunol.* 5 (10), 987–995. doi:10.1038/ni1112

Jiang, Z., Zamanian-Daryoush, M., Nie, H., Silva, A. M., Williams, B. R., and Li, X. (2003). Poly(I-C)-induced Toll-like receptor 3 (TLR3)-mediated activation of NFkappa B and MAP kinase is through an interleukin-1 receptor-associated kinase (IRAK)-independent pathway employing the signaling components TLR3-TRAF6-TAK1-TAB2-PKR. *J. Biol. Chem.* 278 (19), 16713–16719. doi:10.1074/jbc.M300562200

Jin, X., Ding, D., Yan, Y., Li, H., Wang, B., Ma, L., et al. (2019). Phosphorylated RB promotes cancer immunity by inhibiting NF-κB activation and PD-L1 expression. *Mol. Cell* 73 (1), 22–35. doi:10.1016/j.molcel.2018.10.034

Ju, Q., Li, X., Zhang, H., Yan, S., Li, Y., and Zhao, Y. (2020). NFE2L2 is a potential prognostic biomarker and is correlated with immune infiltration in brain lower grade glioma: A pan-cancer analysis. *Oxid. Med. Cell. Longev.* 2020, 3580719. doi:10.1155/2020/3580719

Le Naour, J., Galluzzi, L., Zitvogel, L., Kroemer, G., and Vacchelli, E. (2020). Trial watch: TLR3 agonists in cancer therapy. *Oncoimmunology* 9 (1), 1771143. doi:10.1080/2162402X.2020.1771143

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

Liao, G., Lv, J., Ji, A., Meng, S., and Chen, C. (2021). TLR3 serves as a prognostic biomarker and associates with immune infiltration in the renal clear cell carcinoma microenvironment. *J. Oncol.* 2021, 3336770. doi:10.1155/2021/3336770

Lim, D. S., Kim, J. H., Lee, D. S., Yoon, C. H., and Bae, Y. S. (2007). DC immunotherapy is highly effective for the inhibition of tumor metastasis or recurrence, although it is not efficient for the eradication of established solid tumors. *Cancer Immunol. Immunother.* 56 (11), 1817–1829. doi:10.1007/s00262-007-0325-0

Lin, C. C., Mabe, N. W., Lin, Y. T., Yang, W. H., Tang, X., Hong, L., et al. (2020). RIPK3 upregulation confers robust proliferation and collateral cystine-dependence on breast cancer recurrence. *Cell Death Differ.* 27 (7), 2234–2247. doi:10.1038/s41418-020-0499-y

Linehan, W. M., and Ricketts, C. J. (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

Liu, S. H., Shen, P. C., Chen, C. Y., Hsu, A. N., Cho, Y. C., Lai, Y. L., et al. (2020). DriverDBv3: A multi-omics database for cancer driver gene research. *Nucleic Acids Res.* 48 (1), D863–D70. doi:10.1093/nar/gkz964

Marin-Acevedo, J. A., Dholaria, B., Soyano, A. E., Knutson, K. L., Chumsri, S., and Lou, Y. (2018). Next generation of immune checkpoint therapy in cancer: New developments and challenges. *J. Hematol. Oncol.* 11 (1), 39. doi:10.1186/s13045-018-0582-8

Moossavi, M., Parsamanesh, N., Bahrami, A., Atkin, S. L., and Sahebkar, A. (2018). Role of the NLRP3 inflammasome in cancer. *Mol. Cancer* 17 (1), 158. doi:10.1186/s12943-018-0900-3

Niu, X., Zhu, Z., Shao, E., and Bao, J. (2021). ACE2 is a prognostic biomarker and associated with immune infiltration in kidney renal clear cell carcinoma: Implication for COVID-19. *J. Oncol.* 2021, 8847307. doi:10.1155/2021/8847307

Park, Y., Pang, K., Park, J., Hong, E., Lee, J., Ooshima, A., et al. (2020). Destablilization of TRAF6 by DRAK1 suppresses tumor growth and metastasis in cervical cancer cells. *Cancer Res.* 80 (12), 2537–2549. doi:10.1158/0008-5472.CAN-19-3428

Petitprez, F., Ayadi, M., de Reyniès, A., Fridman, W. H., Sautès-Fridman, C., and Job, S. (2021). Review of prognostic expression markers for clear cell renal cell carcinoma. *Front. Oncol.* 11, 643065. doi:10.3389/fonc.2021.643065

Pichler, M., Hutterer, G. C., Chromecki, T. F., Jesche, J., Kampel-Kettner, K., Groselj-Strele, A., et al. (2013). Predictive ability of the 2002 and 2010 versions of the Tumour-Node-Metastasis classification system regarding metastasis-free, cancer-specific and overall survival in a European renal cell carcinoma single-centre series. *BJU Int.* 111 (4), E191–E195. doi:10.1111/j.1464-410X.2012.11584.x

Pirher, N., Pohar, J., Mancek-Keber, M., Bencina, M., and Jerala, R. (2017). Activation of cell membrane-localized Toll-like receptor 3 by siRNA. *Immunol. Lett.* 189, 55–63. doi:10.1016/j.imlet.2017.03.019

Poff, A., Koutnik, A. P., Egan, K. M., Sahebjam, S., D'Agostino, D., and Kumar, N. B. (2019). Targeting the Warburg effect for cancer treatment: Ketogenic diets for management of glioma. *Semin. Cancer Biol.* 56, 135–148. doi:10.1016/j.semcancer.2017.12.011

Qin, C., Sun, L. J., Cui, L., Cao, Q., Zhu, J., Li, P., et al. (2013). Application of the revised tumour node metastasis (TNM) staging system of clear cell renal cell carcinoma in eastern China: Advantages and limitations. *Asian J. Androl.* 15 (4), 550–557. doi:10.1038/aja.2012.161

Seijkens, T. T. P., van Tiel, C. M., Kusters, P. J. H., Atzler, D., Soehnlein, O., Zarzycka, B., et al. (2018). Targeting CD40-induced TRAF6 signaling in macrophages reduces atherosclerosis. *J. Am. Coll. Cardiol.* 71 (5), 527–542. doi:10.1016/j.jacc.2017.11.055

Siegel, R. L., Miller, K. D., and Jemal, A. (2017). Cancer statistics. *Ca. Cancer J. Clin.* 67 (1), 7–30. doi:10.3322/caac.21387

Snell, L. M., McGaha, T. L., and Brooks, D. G. (2017). Type I interferon in chronic virus infection and cancer. *Trends Immunol.* 38 (8), 542–557. doi:10.1016/j.it.2017.05.005

Stenzel, P. J., Schindeldecker, M., Tagscherer, K. E., Foersch, S., Herpel, E., Hohenfellner, M., et al. (2020). Prognostic and predictive value of tumor-infiltrating leukocytes and of immune checkpoint molecules PD1 and PDL1 in clear cell renal cell carcinoma. *Transl. Oncol.* 13 (2), 336–345. doi:10.1016/j.tranon.2019.11.002

Su, R., Cai, L., Xiong, P., Liu, Z., Chen, S., Liu, X., et al. (2022). TLR3 expression is a potential prognosis biomarker and shapes the immune-active tumor microenvironment in esophageal squamous cell carcinoma. *J. Inflamm. Res.* 15, 1437–1456. doi:10.2147/JIR.S348786

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 (1), D607–D13. doi:10.1093/nar/gky1131

Tamayo, P., Cho, Y. J., Tsherniak, A., Greulich, H., Ambrogio, L., Schouten-van Meeteren, N., et al. (2011). Predicting relapse in patients with medulloblastoma by integrating evidence from clinical and genomic features. *J. Clin. Oncol.* 29 (11), 1415–1423. doi:10.1200/JCO.2010.28.1675

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

Vasaikar, S. V., Straub, P., Wang, J., and Zhang, B. (2018). LinkedOmics: Analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res.* 46 (1), D956–D63. doi:10.1093/nar/gkx1090

Wang, X., Liu, R., Qu, X., Yu, H., Chu, H., Zhang, Y., et al. (2019). α-Ketoglutarate-Activated NF-κB signaling promotes compensatory glucose uptake and brain tumor development. *Mol. Cell* 76 (1), 148–162. doi:10.1016/j.molcel.2019.07.007

Wang, Y., Zhang, S., Li, H., Wang, H., Zhang, T., Hutchinson, M. R., et al. (2020). Small-molecule modulators of toll-like receptors. *Acc. Chem. Res.* 53 (5), 1046–1055. doi:10.1021/acs.accounts.9b00631

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, J. H., Feng, Z. H., Cao, Y., Zhao, H. W., Chen, Z. H., Liao, B., et al. (2019). Predictive value of single-nucleotide polymorphism signature for recurrence in localised renal cell carcinoma: A retrospective analysis and multicentre validation study. *Lancet. Oncol.* 20 (4), 591–600. doi:10.1016/S1470-2045(18)30932-X

Wormann, S. M., Diakopoulos, K. N., Lesina, M., and Algul, H. (2014). The immune network in pancreatic cancer development and progression. *Oncogene* 33 (23), 2956–2967. doi:10.1038/onc.2013.257

Xu, D., Li, X., Shao, F., Lv, G., Lv, H., Lee, J. H., et al. (2019). The protein kinase activity of fructokinase A specifies the antioxidant responses of tumor cells by phosphorylating p62. *Sci. Adv.* 5 (4), eaav4570. doi:10.1126/sciadv.aav4570

Xu, D., Wang, Z., Xia, Y., Shao, F., Xia, W., Wei, Y., et al. (2020). The gluconeogenic enzyme PCK1 phosphorylates INSIG1/2 for lipogenesis. *Nature* 580, 530–535. doi:10.1038/s41586-020-2183-2

Xu, F., Zhang, Z., Yuan, M., Zhao, Y., Zhou, Y., Pei, H., et al. (2021). M6A regulatory genes play an important role in the prognosis, progression and immune microenvironment of pancreatic adenocarcinoma. *Cancer Invest.* 39 (1), 39–54. doi:10.1080/07357907.2020.1834576

Xu, W., Dong, J., Zheng, Y., Zhou, J., Yuan, Y., Ta, H. M., et al. (2019). Immune-checkpoint protein VISTA regulates antitumor immunity by controlling myeloid cell-mediated inflammation and immunosuppression. *Cancer Immunol. Res.* 7 (9), 1497–1510. doi:10.1158/2326-6066.CIR-18-0489

Zhang, M., Wang, X., Chen, X., Zhang, Q., and Hong, J. (2020). Novel immune-related gene signature for risk stratification and prognosis of survival in lower-grade glioma. *Front. Genet.* 11, 363. doi:10.3389/fgene.2020.00363

Zhang, S., Zhang, E., Long, J., Hu, Z., Peng, J., Liu, L., et al. (2019). Immune infiltration in renal cell carcinoma. *Cancer Sci.* 110 (5), 1564–1572. doi:10.1111/cas.13996

Zheng, J., Wang, L., Peng, Z., Yang, Y., Feng, D., and He, J. (2017). Low level of PDZ domain containing 1 (PDZK1) predicts poor clinical outcome in patients with clear cell renal cell carcinoma. *EBioMedicine* 15, 62–72. doi:10.1016/j.ebiom.2016.12.003

Zheng, L., Xue, J., Jaffee, E. M., and Habtezion, A. (2013). Role of immune cells and immune-based therapies in pancreatitis and pancreatic ductal adenocarcinoma. *Gastroenterology* 144 (6), 1230–1240. doi:10.1053/j.gastro.2012.12.042

Zheng, X., Li, S., and Yang, H. (2021). Roles of toll-like receptor 3 in human tumors. *Front. Immunol.* 12, 667454. doi:10.3389/fimmu.2021.667454

Zou, X., Guo, B., Ling, Q., and Mo, Z. (2022). Toll-like receptors serve as biomarkers for early diagnosis and prognosis assessment of kidney renal clear cell carcinoma by influencing the immune microenvironment: Comprehensive bioinformatics analysis combined with experimental validation. *Front. Mol. Biosci.* 9, 832238. doi:10.3389/fmolb.2022.832238

Zou, X., and Mo, Z. (2021). CYP2J2 is a diagnostic and prognostic biomarker associated with immune infiltration in kidney renal clear cell carcinoma. *Biomed. Res. Int.* 2021, 3771866. doi:10.1155/2021/3771866

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 15 November 2022
DOI 10.3389/fgene.2022.949457

Check for updates

OPEN ACCESS

EDITED BY
William C Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Md Khadem Ali,
Stanford University, United States
Emil Bulatov,
Kazan Federal University, Russia

*CORRESPONDENCE
Cong Ye,
sphyc1989@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 21 May 2022
ACCEPTED 31 October 2022
PUBLISHED 15 November 2022

CITATION
Wang W, Chang G, Zhuo R and Ye C (2022), Identification of a ferroptosis-related long noncoding RNA signature with a prognostic value in adrenocortical carcinoma. *Front. Genet.* 13:949457. doi: 10.3389/fgene.2022.949457

COPYRIGHT
© 2022 Wang, Chang, Zhuo 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.

Identification of a ferroptosis-related long noncoding RNA signature with a prognostic value in adrenocortical carcinoma

Weixi Wang[1†], Guilin Chang[1†], Ran Zhuo[2] and Cong Ye[3*]

[1]Department of Geriatrics, Zhongshan Hospital, Fudan University, Shanghai, China, [2]Department of Urology, Zhongshan Hospital, Fudan University, Shanghai, China, [3]Department of Thoracic Surgery, Shanghai Pulmonary Hospital, Tongji University, Shanghai, China

Background: Adrenocortical carcinoma (ACC) is an uncommon endocrine malignancy associated with poor clinical outcome. As a novel form of cell death, ferroptosis is reliant on the accumulation of iron and reactive oxygen species and is involved in the pathogenesis of various tumors, including ACC. Our study aimed to identify and characterize the prognostic ferroptosis-related lncRNA signature (FerRLSig) in ACC.

Methods: A regulatory network of ferroptosis-related lncRNAs (FerRLs) and mRNAs was constructed based on The Cancer Genome Atlas (TCGA). Univariate and multivariate Cox regression assays were performed to construct the FerRLSig.

Results: Twenty-four FerRLs were identified in the prognostic model, and the high-risk FerRLSig was related to the worse overall survival (OS) in ACC [hazard ratio (HR): 1.936 (1.484–2.526), $p < 0.001$]. The area under the curve (AUC) value of the FerRLSig was 0.936 according to the receiver operating characteristic (ROC) analyses, superior to other traditional clinicopathological features, further supported the utility in prognosis prediction of ACC. We further established a prognostic nomogram combining clinical factors with the FerRLSig, which showed favorable efficacy for survival prediction. Next, gene set enrichment analysis (GSEA) revealed that gene sets were involved in many immune regulatory biological processes related to malignancies. T-cell function of type II INF response and the immune checkpoints, including CD40, CD276, IDO2, NRP1, and CD80, were expressed with a significant difference between the low- and high-risk groups.

Conclusion: This study offered new insights into the pathogenesis of ACC. The novel FerRLSig could be useful in predicting survival and may provide information of immunological research and treatment for ACC patients.

KEYWORDS

adrenocortical carcinoma (ACC), ferroptosis, lncRNA, immune infiltration, data mining

Introduction

Adrenocortical carcinoma (ACC) is an aggressive endocrine malignancy with a rare incidence but high mortality (Else et al., 2014; Zheng et al., 2016). Although patients could benefit from complete surgical resection or mitotane treatment, the prognosis of ACC is dismal with the 5-year survival of 20%–59% (Fassnacht et al., 2011; Zheng et al., 2016). In addition, the prognosis of ACC varies with the patient's age, surgical scope, mitotic intensity, and secretion of hormones. Due to the tumor heterogeneity of ACC, the current tumor, lymph node, and metastasis (TNM) staging system is not reliable in clinical outcome prediction. It is challenging to make a prediction of the survival in ACC patients owing to the low incidence rate, diverse pathogenic determinants, high heterogeneity, and dismal prognosis. Therefore, it is particularly critical to uncover new and reliable prognostic biological markers for patients with ACC.

Past decades have seen a growing number of research studies to investigate the function of ferroptosis in malignancies. As a novel form of cell death, ferroptosis is reliant on the accumulation of intracellular iron and reactive oxygen species, distinguished from other forms of cell death such as apoptosis and autophagy (Dixon et al., 2012; Huang et al., 2020). Imbalance of iron metabolism could facilitate tumor growth and also act as a risk factor for tumorigenesis. Compared with normal cells, malignant cells excessively depend on iron for cell proliferation, that is, iron addiction (Manz et al., 2016). One recent study uncovered that ACC is remarkably sensitive to ferroptosis in human, and compared to mitotane, the ferroptosis inducer may be more specific, more effective, and less toxic for the treatment of ACC patients (Belavgeni et al., 2019). Weigand et al. (2020) demonstrated that adrenocortical cells are extraordinary sensitive to ferroptosis owing to the active steroid synthesis pathway. In addition, Chen and colleagues identified a ferroptosis-related signature which could be useful in prognosis prediction and immunotherapy or targeted therapy screening (Chen et al., 2021). Therefore, it is of vital importance to further investigate the changing profiles of ferroptosis-related genes and the underlying mechanisms in ACC.

As a subset of RNA molecules, a long noncoding RNA (lncRNA) has the length of over 200 nt and is involved in gene regulation (Zhang et al., 2021a). In addition, lncRNA is involved in various biological process and plays an important role in the development and progression of malignancies (Gupta et al., 2010), including ACC (Buishand et al., 2020). At present, few studies have focused on the functions of ferroptosis-related lncRNAs (FerRLs) in tumors. Recently, Chao et al. reported that cytosolic lncRNA P53RRA suppressed tumor progression by promoting ferroptosis. The P53RRA–G3BP1 interaction in the cytoplasm leads to the accumulation of p53 in the nucleus, conferring to ferroptosis and cell-cycle arrest. In patients with breast cancers or lung cancers harboring the wild-type p53, the decreased expression of P53RRA was significantly related to a poor prognosis (Mao et al., 2018), and p53 is considered as one of the most promising molecular targets for cancer therapy (Chasov et al., 2020). In another related study, Zhang et al. (2021b) demonstrated that the chronic cadmium exposure could significantly increase the expression of lncRNA OIP5-AS1 in prostate cancer. OIP5-AS1 regulates the expression of SLC7A11 (an alternative marker of ferroptosis) by sponging miR-128-3p, leading to enhanced ferroptosis and decreased cell viability. Similarly, LINC00336 serves as a competing endogenous sponge of miR6852 to inhibit ferroptosis in lung cancer (Wang et al., 2019a). However, no research has been performed to identify the function of FerRLs in predicting the overall survival (OS) for ACC patients. Thus, we constructed a prognostic ferroptosis-related lncRNAs signature (FerRLSig) according to The Cancer Genome Atlas (TCGA) database. Additionally, functions and biological pathways of ferroptosis-related genes and the N6-methyladenosine (m6A) mRNA status and immune checkpoint in the survival prediction of ACC patients were further explored.

TABLE 1 Clinical characteristics of patients in TCGA dataset.

Variable	Number of samples
Gender	
Male/Female	32/60
Age at diagnosis	
≤65/>65	81/11
Stage	
I/II/III/IV/NA	9/44/19/18/2
T	
T1/T2/T3/T4/NA	9/49/11/21/2
M	
M0/M1/NA	69/21/2
N	
N0/N1/NA	80/10/2

Materials and methods

Data collection

On 5 February 2022, we extracted the RNA-sequence data of 92 patients from the TCGA-ACC data set (https://portal.gdc.cancer.gov/). Table 1 lists the clinical features, and the identified lncRNAs were annotated using the GTF annotation files of human lncRNAs retrieved from the GENCODE (https://www.gencodegenes.org/). We obtained 382 ferroptosis-related genes

(n = 150 for the drivers, n = 109 for the suppressors, and n = 123 for the markers) from the FerrDb database, which is a data network with comprehensive and timely updates, covering the latest progress of ferroptosis-related genes along with their regulatory molecules and related diseases (Zhou and FerrDb, 2020) (http://www.zhounan.org/ferrdb, Supplementary Table S1). Information of the tumor grade in ACC was missing; thus, we collected clinicopathological data including gender, age, TNM stage, and survival information. Since the study was based on a public database, ethical approval is not required.

Identification of a prognostic FerRLSig

The relationship between the FerRLs and ferroptosis-related genes was assessed by Pearson correlation. Statistical significance was achieved at a correlation coefficient $|R^2|>0.8$ and $p < 0.001$. Next, a FerRLSig was constructed based on the univariate and multivariate Cox regression analyses. Information of age, gender, stage, and risk score was included in the multivariate Cox regression analysis. The risk score was calculated as the sum of (coefficient of lncRNA1 × lncRNA1 expression) + (coefficient of lncRNA2 × lncRNA2 expression) + + (coefficient of lncRNAn × lncRNAn expression). Taking the median value as the cut-off point, the lncRNAs were divided into two groups: high-risk group and low-risk group. Differentially expressed genes (DEGs) between the low- and high-risk groups were identified and statistically significant was considered at a false discovery rate (FDR) < 0.05 and $|\log_2 FC|\geq 1$. A Kaplan–Meier survival analysis was then performed to evaluate the clinical correlation of risk stratification and patient prognosis. The FerRLSig was present as a risk plot, including the distribution of risk scores and survival status in ACC patients. We further calculated the area under the curve (AUC) value in the receiver operating characteristic (ROC) curve to evaluate the prediction accuracy of OS in ACC patients. Sensitivity and specificity of the FerRLSig were compared with other clinical and pathological factors *via* ROC curves and decision curve analysis (DCA) (Vickers and Elkin, 2006). Cytoscape 3.9.0 was used to visualize the co-expression network of ferroptosis-related genes, lncRNAs, and signaling pathways.

Functional analyses of DEGs

Functions of DEGs were compared between the low-risk group and the high-risk group. We evaluated the associated biological functions using gene ontology (GO), including biological processes (BPs), molecular functions (MFs), and cellular components (CCs). Based on the data in the Kyoto Encyclopedia of Genes and Genomes (KEGG), we further analyzed the biological pathways using the ggplot2 package in R 4.1.2.

Gene set enrichment analysis and the predictive nomogram

To explore potential molecular mechanisms of the FerRLSig, we performed a gene set enrichment analysis (GSEA) to define the enriched terms in the KEGG. Significant gene sets were achieved at p-value <0.05 (GSEA 4.2.2). Statistical significance was expressed by the normalized enrichment score (NES) and FDR (Subramanian et al., 2005). A nomogram was developed by integrating the FerRLSig and clinical characteristics to predict the 1-, 3-, and 5-year OS in ACC.

Immune infiltration and gene expression analyses

Potential correlation of the FerRLSig and immune cells infiltration was analyzed using the CIBERSORT algorithms (Newman et al., 2015; Charoentong et al., 2017) and TIMER website (http://timer.comp-genomics.org/) (Li et al., 2017). The differences of immune responses between the two groups were illustrated on a cluster heatmap. Next, a single-sample gene set enrichment analysis (ssGSEA) was performed to analyze the scores of tumor-infiltrating immune cells along with their functions in TCGA. Furthermore, we retrieved the previous literature to extract potential immune checkpoints (Charoentong et al., 2017).

Statistical analysis

R software (version 4.1.2), along with its appropriate packages, was used to perform the statistical analysis. The Wilcoxon test and unpaired Student's t-test were used to analyze the non-normally and normally distributed variables, respectively. Differentially expressed genes between the low- and high-risk groups were identified using the Benjamini–Hochberg method. The association between FerRLs and clinicopathological characteristics of ACC patients was assessed by logistic regression analysis, presented in a cluster heatmap. Statistical significance was achieved at $p < 0.05$.

Results

Identification of a prognostic FerRLSig in adrenocortical carcinoma

According to the Pearson correlation analysis, a total of 85 FerRLs were identified from the 382 ferroptosis-related genes in TCGA-ACC dataset (Supplementary Table S2). Subsequently, 32 significant FerRLs were screened by the univariate analysis and then included in the multivariate Cox analysis. Overall, 24 FerRLs (AC046143.1, AC099850.3,

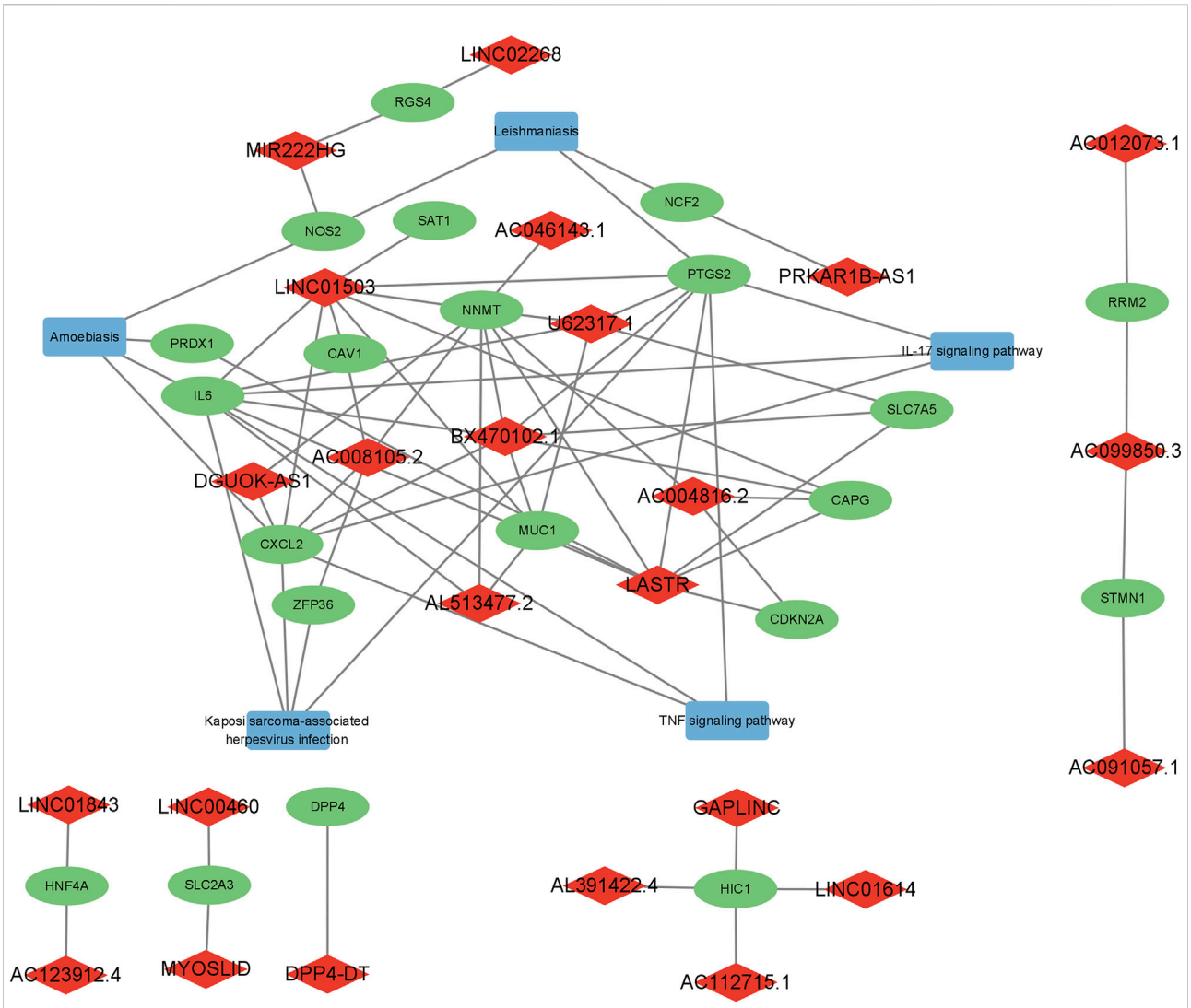


FIGURE 1
Regulatory network of the ferroptosis-related lncRNAs, mRNAs, and signaling pathways. Diamonds represent lncRNAs in red, rectangles indicate signaling pathways in blue, and ellipses represent mRNAs in green.

LINC01614, LASTR, DPP4-DT, AL391422.4, LINC01503, LINC01843, AC008105.2, AC004816.2, AC112715.1, AC012073.1, LINC00460, U62317.1, BX470102.1, LINC02268, GAPLINC, MYOSLID, AC123912.4, AC091057.1, DGUOK-AS1, PRKAR1B-AS1, MIR222HG, and AL513477.2) were identified to be independent prognostic factors for ACC (Supplementary Table S3). Annotations and potential mechanisms of these identified FerRLs are presented in Supplementary Table S4. Next, a prognostic FerRLSig was constructed based on the calculated risk scores.

The regulatory network of FerRLs, mRNAs, and pathways is shown in Figure 1, which indicated that these FerRLs regulate multiple ferroptosis-related pathways by targeting the mRNAs. For example, lncRNA BX470102.1 is involved in amoebiasis, Kaposi sarcoma-associated herpesvirus infection, TNF signaling pathway, and IL-17 signaling pathway by targeting multiple target genes such as IL-6, CXCL2, and PTGS2. The target genes of LINC01503, U62317.1, AC008105.2, LASTR, and AC004816.2 are also involved in the aforementioned pathways, which have been proved to be related to ferroptosis in multiple tumors (Liu et al., 2021).

Enrichment analysis of the identified ferroptosis-related genes

Overall, 540 genes were identified to be differentially expressed between the low- and high-risk groups (n = 103 for

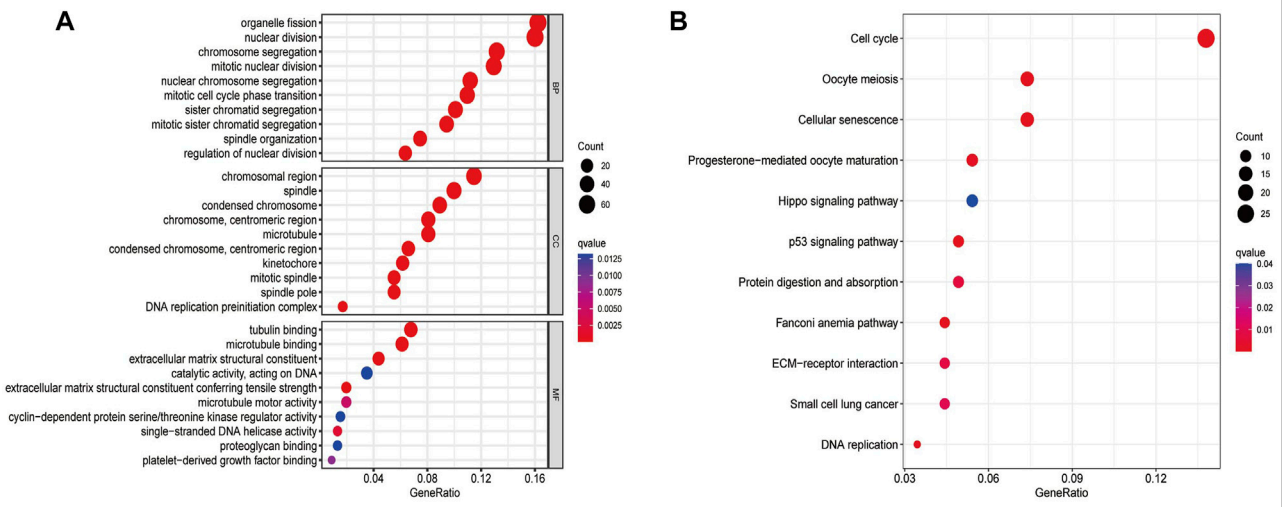


FIGURE 2
GO **(A)** and KEGG **(B)** analyses for ferroptosis-related genes differentially expressed between the low- and high-risk groups.

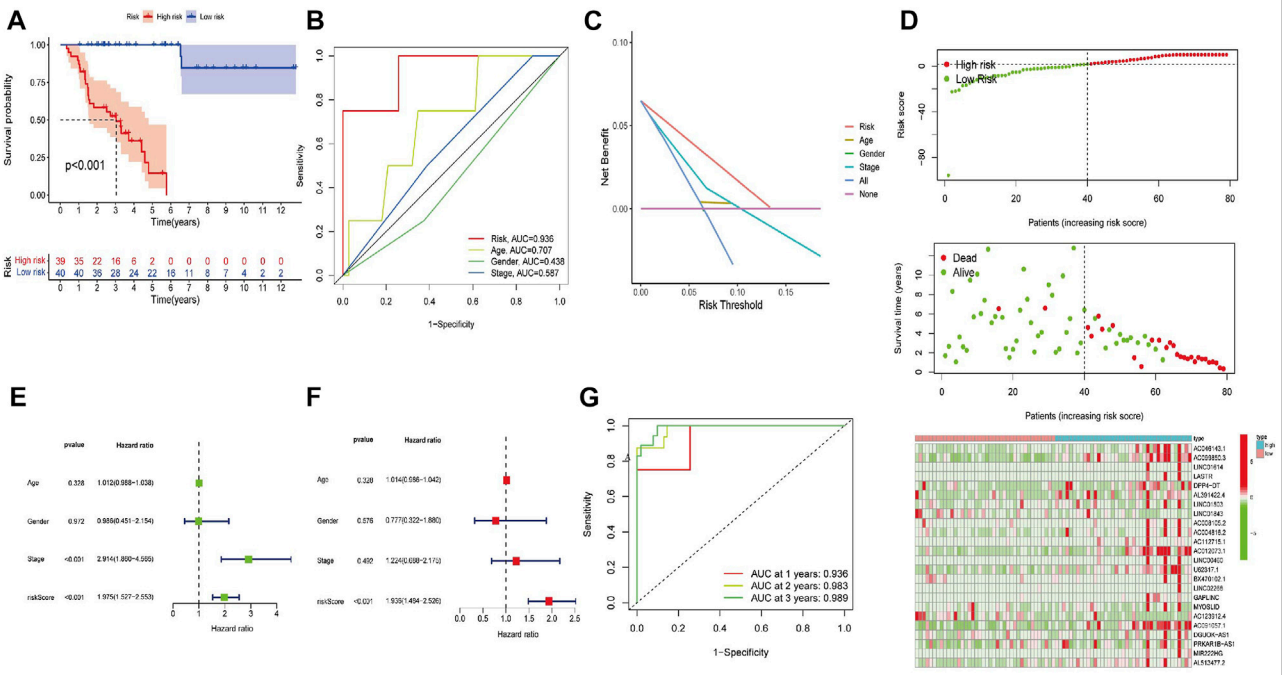


FIGURE 3
Construction of the ferroptosis-related lncRNA signature (FerRLSig). **(A)** Kaplan–Meier survival analysis stratified by the risk scores. **(B)** AUC values of the FerRLSig and other clinicopathological factors. **(C)** Decision curve analysis (DCA) of the risk factors. **(D)** Distribution of ACC patients based on the FerRLSig. Red and green dots represent the status of death and survival, respectively. The cluster heatmap represents the expressions of ferroptosis-related lncRNAs (FerRLs) stratified by the risk scores. **(E)** Univariate analysis of the risk score and other traditional clinicopathological factors. **(F)** Multivariate Cox analysis of the risk score and other traditional clinicopathological factors. **(G)** AUC value for the prediction of the 1-, 2-, and 3-year OS of ACC patients in the FerRLSig.

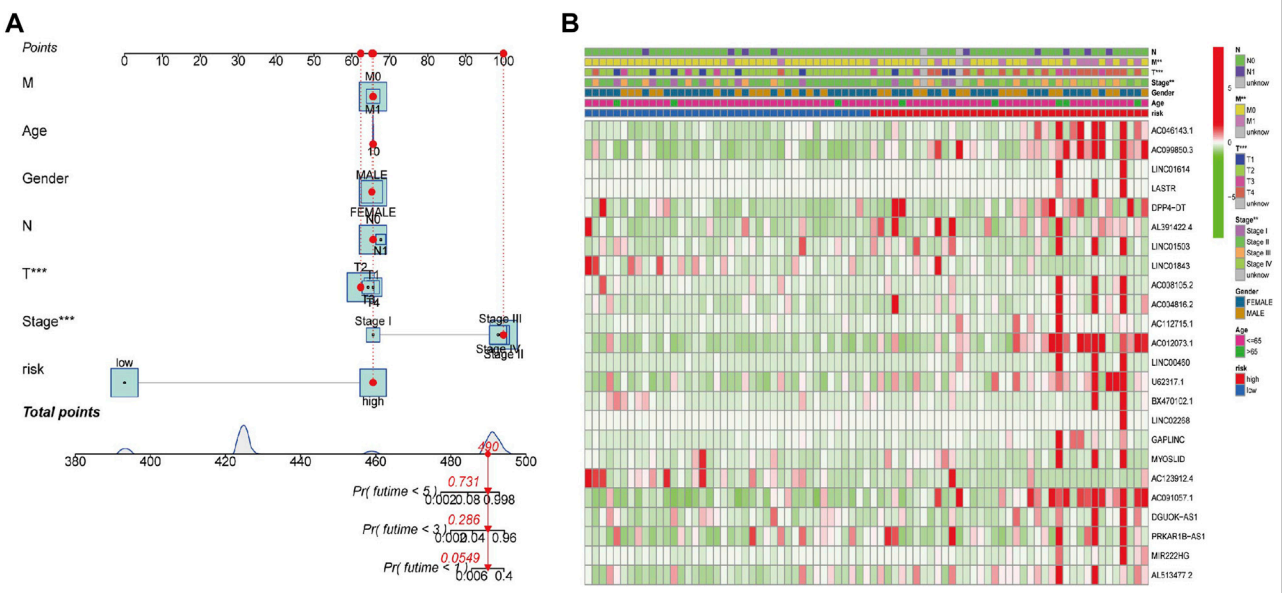


FIGURE 4
(A) A nomogram for the identified FerRLs and other clinicopathological factors. **(B)** Cluster heatmap for the association of clinicopathological manifestations and the prognostic FerRLSig.

the low-risk group and n = 437 for the high-risk group, Supplementary Table S5). BP participated in organelle fission, mitotic nuclear division, chromosome segregation, and mitotic cell cycle phase transition. CC were mainly involved in the chromosomal region, mitotic spindle, and DNA replication preinitiation complex pathways. MF mainly regulated the single-stranded DNA helicase activity, cyclin-dependent protein serine/threonine kinase regulator activity, tubulin binding, and extracellular matrix structural constituent (Figure 2A). The KEGG pathway analyses revealed that the genes in the high-risk group were particularly involved in oocyte meiosis, cell cycle, DNA replication, protein digestion and absorption, Hippo signaling pathway, small-cell lung cancer, cellular senescence, and p53 signaling pathway (Figure 2B; Supplementary Table S6).

Prognostic value of the FerRLSig

High-risk FerRLs were associated with a worse OS according to the Kaplan–Meier analyses (Figure 3A, $p < 0.001$). In addition, the AUC value for the FerRLSig was 0.936 according to the ROC curve, significantly higher than that for traditional clinicopathological factors such as gender, age, and tumor stage (Figures 3B,C), further supporting the potential superior performance of the FerRLSig in predicting the survival in ACC. As shown in the prognostic curve and scatter plot, the risk scores of ACC patients were negatively correlated with their OS. Meanwhile, the cluster heatmap revealed the elevated expressions of candidate FerRLs in the high-risk group (Figure 3D), which needs to be further explored.

To explore whether the generated FerRLSig could independently predict the OS for patients with ACC, we then performed univariate and multivariate Cox regression analyses. The hazard ratios (HRs) of the risk scores in the univariate analysis and multivariate Cox regression analysis were 1.975 (1.527–2.533) and 1.936 (1.484–2.526), respectively ($p < 0.001$, Figures 3E,F). Additionally, the predictive accuracy of this novel signature was evaluated by a ROC analysis, with the AUC values of 0.936, 0.983, and 0.989 for the 1-, 2-, and 3-year survival rates, respectively (Figure 2G). To make the model more applicable in clinical management, a nomogram incorporating the FerRLSig and clinicopathological features was constructed to predict the 1-,3-, and 5-year OS of ACC patients based on TCGA database (Figure 4A). The cluster heatmap for the association of clinicopathological manifestations and FerRLSig is illustrated in Figure 4B.

Identification of biological pathways in the prognostic FerRLSig

A total of 178 enriched KEGG pathways were identified in the GSEA (GSEA 4.2.2). Most of the identified FerRLs were involved in tumor-related and immunological pathways including cell cycle (NES = 2.07; $p < 0.001$), p53 signaling

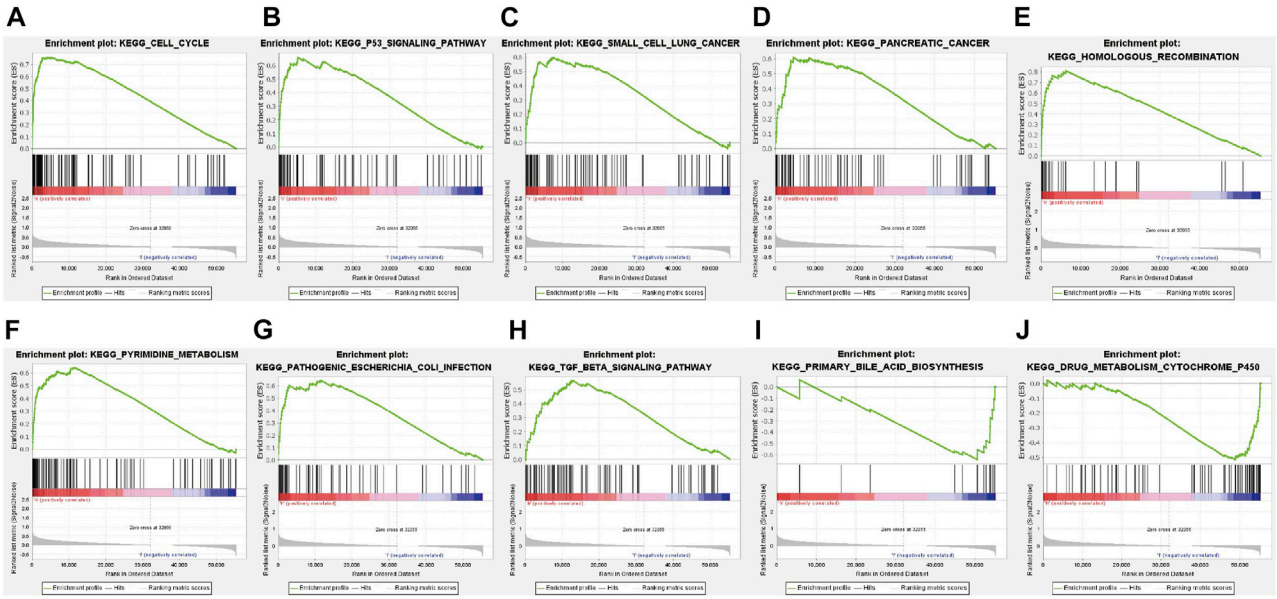


FIGURE 5
GSEA of the ferroptosis-related signaling pathways according to TCGA database. **(A)** Cell cycle, **(B)** p53 signaling pathway, **(C)** Small cell lung cancer, **(D)** Pancreatic cancer, **(E)** Homologous recombination, **(F)** Pyrimidine metabolism, **(G)** Pathogenic Escherichia coli infection, **(H)** TGF beta signaling pathway, **(I)** Primary bile acid biosynthesis, **(J)** Drug metabolism cytochrome P450 signaling pathway.

pathway (NES = 1.92; $p < 0.001$), small-cell lung cancer (NES = 1.75; $p = 0.002$), pancreatic cancer (NES = 1.74; $p < 0.001$), homologous recombination (NES = 1.91; $p < 0.001$), pyrimidine metabolism (NES = 1.89; $p < 0.001$), pathogenic *Escherichia coli* infection (NES = 1.78; $p = 0.002$), TGF-beta signaling pathway (NES = 1.67; $p = 0.002$), primary bile acid biosynthesis (NES = −1.66; $p = 0.013$), and drug metabolism cytochrome P450 (NES = −1.49; $p = 0.027$) signaling pathways (Figure 5 and Supplementary Table S7). Taken together, the novel prognostic signature may be involved in the microenvironment of tumor immune.

Ferroptosis-related genes and immunity

Figure 6 illustrates the immune responses clustering heatmap based on TIMER and CIBERSORT algorithms, which showed that immune infiltration was increased in cases with high-risk scores. Based on the ssGSEA of TCGA-ACC data, correlation analyses were performed to evaluate the association of immune cell subsets and corresponding functions, which revealed that T-cell function of type II INF response decreased significantly in the high-risk group (Figure 7A), suggesting that high-risk patients might be related to a suppressive immune microenvironment. Given the clinical importance of the checkpoint blocking therapy in tumors, we further analyzed the expression differences of immune checkpoints between the two groups, and the expressions of CD40, CD276, IDO2, NRP1, and CD80 were of significant differences (Figure 7B). The comparison of m6A-related mRNA expression between the two risk groups showed that RBM15, WTAP, YTHDF1, and HNRNPC were differentially expressed (Figure 7C). The aforementioned results indicate that the generated FerRLSig might be a candidate marker for the immune checkpoint inhibitor therapy in ACC patients.

Discussion

ACC is an uncommon endocrine malignancy associated with high mortality (Wang et al., 2019b). Early diagnosis and more accurate therapies are of great importance to improve the prognosis. Some nomograms and risk models have been constructed to predict the survival in ACC (Wang et al., 2021a; Chen et al., 2021); however, no prognostic FerRLSig has been identified. Ferroptosis is associated with the development and progression of malignancies and might be a potentially novel approach for tumor therapy; additionally, its pivotal role in ACC has recently been revealed (Belavgeni et al., 2019). In this study, a FerRLSig that can evaluate the prognosis of ACC was constructed based on TCGA-ACC dataset. We further explored the association of survival and immune infiltrating cells and immune checkpoint inhibitors, which could provide potential biomarkers and immune targets for ACC treatment.

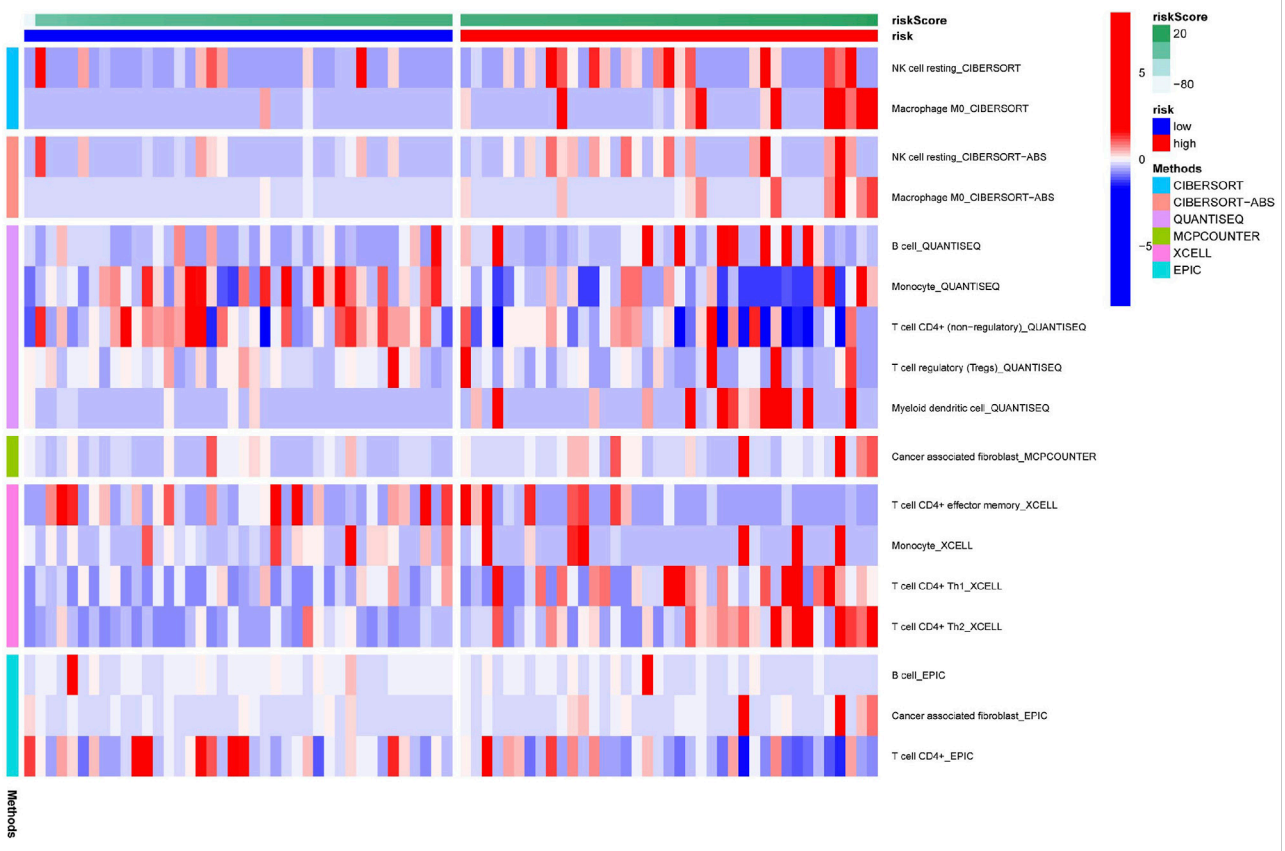


FIGURE 6
Immune response correlated to the ferroptosis-related lncRNA signature, illustrated in a cluster heatmap.

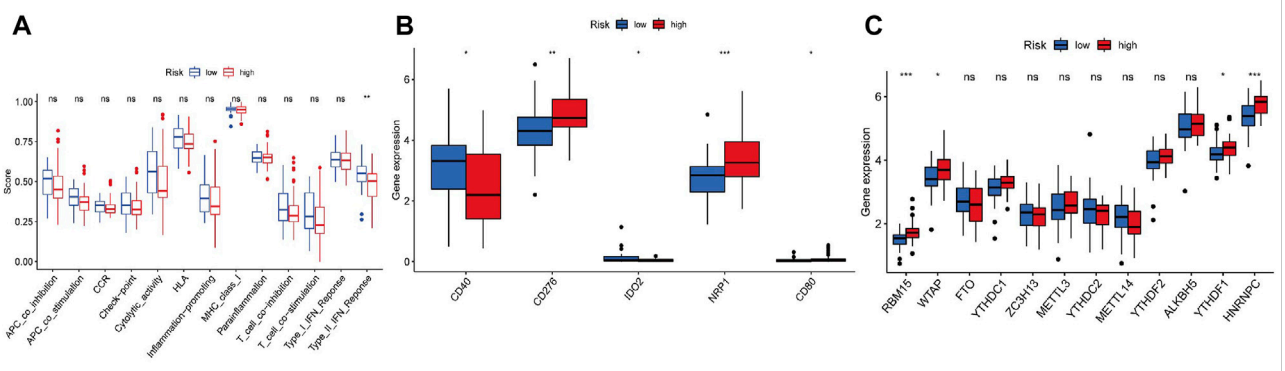


FIGURE 7
(A) Correlation analysis for immune cell subsets and the related functions by ssGSEA. **(B)** Comparisons of immune checkpoints expressions between the low- and high-risk groups. **(C)** Expressions of m6A-related genes based on the FerRLSig.

Overall, 24 FerRLs were identified in the prognostic FerRLSig. In review of the literature, cell types that predominantly express these FerRLs along with the potential mechanisms are presented in Supplementary Table S4. The AUC value of the ROC analysis indicated that the FerRLSig might have a superior predictive performance for OS in ACC. Additionally, a nomogram

incorporating the FerRLSig and clinicopathological features of ACC was constructed to facilitate the development of treatment strategies and might guide clinical decision-making. One recent study revealed that overexpression of lncRNA AC099850.3 could promote cell proliferation and invasion *via* the PRR11/PI3K/AKT pathway, predicting a poor prognosis in hepatocellular carcinoma (Zhong et al., 2022). In a related study, LINC01614 was significantly upregulated in osteosarcoma and could facilitate tumor progression through the miR-520a-3p/SNX3 axis (Cai et al., 2021). In addition, the overexpression of LINC01614 in bladder cancer could promote tumor proliferation, migration, and invasion through the miR-217/RUNX2 and Wnt/β-Catenin pathway (Wang et al., 2021b). Upregulated expression of lncRNA LASTR was found in several epithelial tumors and could facilitate cancer cell fitness in hypoxic breast cancer through SART3 (De Troyer et al., 2020). LINC01503 is involved in a variety of malignancies, including hepatocellular carcinoma, cholangio-carcinoma, non-small-cell lung cancer, cervical cancer, and gastric cancer (Qu et al., 2019; Shen et al., 2020; Wang et al., 2020; Zhang et al., 2020; Feng et al., 2021; Ma et al., 2021). Wu et al. (2021) reported that the overexpression of LINC00460 in pancreatic ductal adenocarcinoma could promote tumor aggressiveness by targeting miR-491-5p, conferring a poor clinical outcome. Similarly, LINC00460 in head and neck squamous cell carcinoma facilitates peroxiredoxin-1 entry into the nucleus, thereby promoting cell proliferation and metastasis (Jiang et al., 2019). Knockdown of lncRNA GAPLINC repressed the tumor growth in renal cell cancer, while overexpression of GAPLINC facilitated tumorigenesis *via* the miR-135b-5p/CSF1 axis, indicating poor survival (Wang et al., 2021c). On the other hand, Yang and colleagues revealed that lncRNA MYOSLID played a critical role in the progression of osteosarcoma through the miR-1286/RAB13 axis (Yang et al., 2019a). In cervical cancer, upregulated expression of LncRNA DGUOK-AS1 could sponge miR-653-5p to repress DNA repair and promote cell proliferation (Wu et al., 2020). However, to the best of our knowledge, there has been no research conducted to explore whether FerRLs participate in the pathogenesis and prognosis of ACC.

Next, the present study identified 540 ferroptosis-related DEGs and stratified them by the novel FerRLSig. KEGG-based analyses further revealed that the genes of the high-risk group mainly participated in the Hippo signaling pathway, p53 signaling pathway, cancer-related pathway, ECM–receptor interaction, cellular senescence, and cell cycle. Recently, Lin and colleagues revealed that recurrent breast tumors are highly sensitive to ferroptosis, and upregulated expression of DDR2 could promote ferroptosis through the Hippo signaling pathway (Lin et al., 2021). Wen-Hsuan et al. reported that TAZ (a Hippo pathway effector) mediates the cell density-regulated ferroptosis by regulating the EMP1-NOX4 axis in renal cell carcinoma (Yang et al., 2019b). Another related study identified an ACSL4-independent and ALOX12-mediated pathway of ferroptosis that is essential for p53-dependent tumor suppression (Chu et al., 2019). Taken together, the aforementioned signaling pathways are closely related to ferroptosis in tumors.

Another significant contribution of the current study is the uncovering of the association between the FerRLSig and tumor immune microenvironment. Increasing evidence suggests that T-cells could promote tumor ferroptosis to enhance antitumor activity, which might be a potential therapeutic target for cancers combining with an immune checkpoint blockade (Wang et al., 2019c). Functional enrichment analyses revealed that FerRLs mainly participated in tumor-related and immune pathways. Type II INF response in T-cell function was significantly decreased in the high-risk group, suggesting that FerRLs may be involved in the regulation of tumor immune infiltration.

As a novel form of cell death, ferroptosis provides a new way for tumor development and treatment (Hassannia et al., 2019; Liang et al., 2019). However, the exact mechanism of ferroptosis in ACC remains unclear. To date, we are the first to construct a FerRLsig for ACC, which could shed light on novel treatment modalities for this rare tumor. Nonetheless, several limitations exist in the present study. First and most importantly, our findings are based on a single database and have not been validated by a clinical cohort. We explored the Gene Expression Omnibus (GEO) and the International Cancer Genome Consortium (ICGC); however, ACC is extremely rare and only 26 ACCs were found in the GEO, and no case of ACC was found in the ICGC. Thus, we could not compare or validate our findings with other data sets. In addition, the retrospective nature of this study made it less applicable. In addition, potential predictors such as tumor markers and tumor grade were not included in the research due to the incomplete data. Furthermore, functional experimental validation of the molecular mechanisms of the generated FerRLSig is currently lacking. In general, the role of the prognostic signature established in this study needs further exploration, and international multicenter studies with a larger sample size are expected.

In conclusion, a FerRLSig was established to predict the survival of ACC. Additionally, this model might be related to immune infiltration, providing potential immune targets for the control of ACC.

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

WW and CY conceived and designed the study. WW and GC wrote and revised the manuscript. RZ collected and interpreted

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

Funding

The present study was supported by the National Natural Science Foundation of China (Grant Number: 82000084 and 82002828).

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.2022.949457/full#supplementary-material

References

Belavgeni, A., Bornstein, S. R., von Massenhausen, A., Tonnus, W., Stumpf, J., Meyer, C., et al. (2019). Exquisite sensitivity of adrenocortical carcinomas to induction of ferroptosis. *Proc. Natl. Acad. Sci. U. S. A.* 116 (44), 22269–22274. doi:10.1073/pnas.1912700116

Buishand, F. O., Liu-Chittenden, Y., Fan, Y., Tirosh, A., Gara, S. K., Patel, D., et al. (2020). Adrenocortical tumors have a distinct, long, non-coding RNA expression profile and LINC00271 is downregulated in malignancy. *Surgery* 167 (1), 224–232. doi:10.1016/j.surg.2019.04.067

Cai, Q., Zhao, X., Wang, Y., Li, S., Wang, J., Xin, Z., et al. (2021). LINC01614 promotes osteosarcoma progression via miR-520a-3p/SNX3 axis. *Cell. Signal.* 83, 109985. doi:10.1016/j.cellsig.2021.109985

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

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, X., Yan, L., Jiang, F., Lu, Y., Zeng, N., Yang, S., et al. (2021). Identification of a ferroptosis-related signature associated with prognosis and immune infiltration in adrenocortical carcinoma. *Int. J. Endocrinol.* 2021, 4654302. doi:10.1155/2021/4654302

Chu, B., Kon, N., Chen, D., Li, T., Liu, T., Jiang, L., et al. (2019). ALOX12 is required for p53-mediated tumour suppression through a distinct ferroptosis pathway. *Nat. Cell Biol.* 21 (5), 579–591. doi:10.1038/s41556-019-0305-6

De Troyer, L., Zhao, P., Pastor, T., Baietti, M. F., Barra, J., Vendramin, R., et al. (2020). Stress-induced lncRNA LASTR fosters cancer cell fitness by regulating the activity of the U4/U6 recycling factor SART3. *Nucleic Acids Res.* 48 (5), 5198–5199. doi:10.1093/nar/gkaa280

Dixon, S. J., Lemberg, K. M., Lamprecht, M. R., Skouta, R., Zaitsev, E. M., Gleason, C. E., et al. (2012). Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell* 149 (5), 1060–1072. doi:10.1016/j.cell.2012.03.042

Else, T., Kim, A. C., Sabolch, A., Raymond, V. M., Kandathil, A., Caoili, E. M., et al. (2014). Adrenocortical carcinoma. *Endocr. Rev.* 35 (2), 282–326. doi:10.1210/er.2013-1029

Fassnacht, M., Libe, R., Kroiss, M., and Allolio, B. (2011). Adrenocortical carcinoma: A clinician's update. *Nat. Rev. Endocrinol.* 7 (6), 323–335. doi:10.1038/nrendo.2010.235

Feng, J., Gao, F. Y., Li, Y. Y., Xiao, X. Y., and Xi, J. (2021). Upregulation of LINC01503 promotes cervical cancer progression by targeting the miR-615-3p/CCND1 axis. *J. Cancer* 12 (15), 4552–4560. doi:10.7150/jca.54148

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464 (7291), 1071–1076. doi:10.1038/nature08975

Hassannia, B., Vandenabeele, P., and Vanden Berghe, T. (2019). Targeting ferroptosis to iron out cancer. *Cancer Cell* 35 (6), 830–849. doi:10.1016/j.ccell.2019.04.002

Huang, L., McClatchy, D. B., Maher, P., Liang, Z., Diedrich, J. K., Soriano-Castell, D., et al. (2020). Intracellular amyloid toxicity induces oxytosis/ferroptosis regulated cell death. *Cell Death Dis.* 11 (10), 828. doi:10.1038/s41419-020-03020-9

Jiang, Y., Cao, W., Wu, K., Qin, X., Wang, X., Li, Y., et al. (2019). LncRNA LINC00460 promotes EMT in head and neck squamous cell carcinoma by facilitating peroxiredoxin-1 into the nucleus. *J. Exp. Clin. Cancer Res.* 38 (1), 365. doi:10.1186/s13046-019-1364-z

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

Liang, C., Zhang, X., Yang, M., and Dong, X. (2019). Recent progress in ferroptosis inducers for cancer therapy. *Adv. Mat.* 31 (51), 1904197. doi:10.1002/adma.201904197

Lin, C. C., Yang, W. H., Lin, Y. T., Tang, X., Chen, P. H., Ding, C. C., et al. (2021). DDR2 upregulation confers ferroptosis susceptibility of recurrent breast tumors through the Hippo pathway. *Oncogene* 40 (11), 2018–2034. doi:10.1038/s41388-021-01676-x

Liu, T., Li, X., Cui, Y., Meng, P., Zeng, G., Wang, Y., et al. (2021). Bioinformatics analysis identifies potential ferroptosis key genes in the pathogenesis of intracerebral hemorrhage. *Front. Neurosci.* 15, 661663. doi:10.3389/fnins.2021.661663

Ma, Z., Gao, X., Shuai, Y., Wu, X., Yan, Y., Xing, X., et al. (2021). EGR1-mediated linc01503 promotes cell cycle progression and tumorigenesis in gastric cancer. *Cell Prolif.* 54 (1), 12922. doi:10.1111/cpr.12922

Manz, D. H., Blanchette, N. L., Paul, B. T., Torti, F. M., and Torti, S. V. (2016). Iron and cancer: Recent insights. *Ann. N. Y. Acad. Sci.* 1368 (1), 149–161. doi:10.1111/nyas.13008

Mao, C., Wang, X., Liu, Y., Wang, M., Yan, B., Jiang, Y., et al. (2018). A G3BP1-interacting lncRNA promotes ferroptosis and apoptosis in cancer via nuclear sequestration of p53. *Cancer Res.* 78 (13), 3484–3496. doi:10.1158/0008-5472.CAN-17-3454

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

Qu, Y. K., Qu, X. S., Chen, G., Feng, Y., Teng, X. L., Liu, W. X., et al. (2019). LINC01503 promotes cell proliferation, invasion and EMT process in cholangiocarcinoma. *Eur. Rev. Med. Pharmacol. Sci.* 23 (15), 6445–6452. doi:10.26355/eurrev_201908_18526

Shen, Q., Sun, Y., and Xu, S. (2020). LINC01503/miR-342-3p facilittes malignancy in non-small-cell lung cancer cells via regulating LASP1. *Respir. Res.* 21 (1), 235. doi:10.1186/s12931-020-01464-3

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

Vickers, A. J., and Elkin, E. B. (2006). Decision curve analysis: A novel method for evaluating prediction models. *Med. Decis. Mak.* 26 (6), 565–574. doi:10.1177/0272989X06295361

Wang, M., Mao, C., Ouyang, L., Liu, Y., Lai, W., Liu, N., et al. (2019). Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. *Cell Death Differ.* 26 (11), 2329–2343. doi:10.1038/s41418-019-0304-y

Wang, M. R., Fang, D., Di, M. P., Guan, J. L., Wang, G., Liu, L., et al. (2020). Long non-coding RNA LINC01503 promotes the progression of hepatocellular carcinoma via activating MAPK/ERK pathway. *Int. J. Med. Sci.* 17 (9), 1224–1234. doi:10.7150/ijms.45256

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

Wang, W., Chang, G., Sun, Y., Zhuo, R., Li, H., Hu, Y., et al. (2021). Nomograms for individualized evaluation of prognosis in adrenocortical carcinomas for the elderly: A population-based analysis. *J. Invest. Surg.* 35, 1153–1160. doi:10.1080/08941939.2021.1968981

Wang, W., Green, M., Choi, J. E., Gijon, M., Kennedy, P. D., Johnson, J. K., et al. (2019). CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature* 569 (7755), 270–274. doi:10.1038/s41586-019-1170-y

Wang, W., Han, R., Ye, L., Xie, J., Tao, B., Sun, F., et al. (2019). Adrenocortical carcinoma in patients with MEN1: A kindred report and review of the literature. *Endocr. Connect.* 8 (3), 230–238. doi:10.1530/EC-18-0526

Wang, Z., Yan, H., Cheng, D., Xu, L., Shen, T., Chen, Y., et al. (2021). Novel lncRNA LINC01614 facilitates bladder cancer proliferation, migration and invasion through the miR-217/RUNX2/Wnt/β-Catenin Axis. *Cancer Manag. Res.* 13, 8387–8397. doi:10.2147/CMAR.S330019

Weigand, I., Schreiner, J., Rohrig, F., Sun, N., Landwehr, L. S., Urlaub, H., et al. (2020). Active steroid hormone synthesis renders adrenocortical cells highly susceptible to type II ferroptosis induction. *Cell Death Dis.* 11 (3), 192. doi:10.1038/s41419-020-2385-4

Wu, J., Sun, S., Liao, W., Chen, E., Wang, X., Song, Y., et al. (2021). LINC00460 promotes pancreatic cancer progression by sponging miR-491-5p. *J. Gene Med.* 23 (6), 3333. doi:10.1002/jgm.3333

Wu, N., Song, H., Ren, Y., Tao, S., and Li, S. (2020). DGUOK-AS1 promotes cell proliferation in cervical cancer via acting as a ceRNA of miR-653-5p. *Cell biochem. Funct.* 38 (7), 870–879. doi:10.1002/cbf.3506

Yang, S., Chen, M., and Lin, C. (2019). A novel lncRNA MYOSLID/miR-1286/RAB13 Axis plays a critical role in osteosarcoma progression. *Cancer Manag. Res.* 11, 10345–10351. doi:10.2147/CMAR.S231376

Yang, W. H., Ding, C. C., Sun, T., Rupprecht, G., Lin, C. C., Hsu, D., et al. (2019). The hippo pathway effector TAZ regulates ferroptosis in renal cell carcinoma. *Cell Rep.* 28 (10), 2501–2508. doi:10.1016/j.celrep.2019.07.107

Zhang, M. L., Zhao, T. T., Du, W. W., Yang, Z. F., Peng, W., and Cui, Z. J. (2020). C-MYC-induced upregulation of LINC01503 promotes progression of non-small cell lung cancer. *Eur. Rev. Med. Pharmacol. Sci.* 24 (21), 11120–11127. doi:10.26355/eurrev_202011_23599

Zhang, T., Li, K., Zhang, Z. L., Gao, K., and Lv, C. L. (2021). LncRNA Airsci increases the inflammatory response after spinal cord injury in rats through the nuclear factor kappa B signaling pathway. *Neural Regen. Res.* 16 (4), 772–777. doi:10.4103/1673-5374.295335

Zhang, Y., Guo, S., Wang, S., Li, X., Hou, D., Li, H., et al. (2021). LncRNA OIP5-AS1 inhibits ferroptosis in prostate cancer with long-term cadmium exposure through miR-128-3p/SLC7A11 signaling. *Ecotoxicol. Environ. Saf.* 220, 112376. doi:10.1016/j.ecoenv.2021.112376

Zheng, S., Cherniack, A. D., Dewal, N., Moffitt, R. A., Danilova, L., Murray, B. A., et al. (2016). Comprehensive pan-genomic characterization of adrenocortical carcinoma. *Cancer Cell* 29 (5), 723–736. doi:10.1016/j.ccell.2016.04.002

Zhong, F., Liu, S., Hu, D., and Chen, L. (2022). LncRNA AC099850.3 promotes hepatocellular carcinoma proliferation and invasion through PRR11/PI3K/AKT axis and is associated with patients prognosis. *J. Cancer* 13 (3), 1048–1060. doi:10.7150/jca.66092

Zhou, N., and FerrDb, Bao J. (2020). A manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database J. Biol. databases curation*, 2020. doi:10.1093/database/baaa021

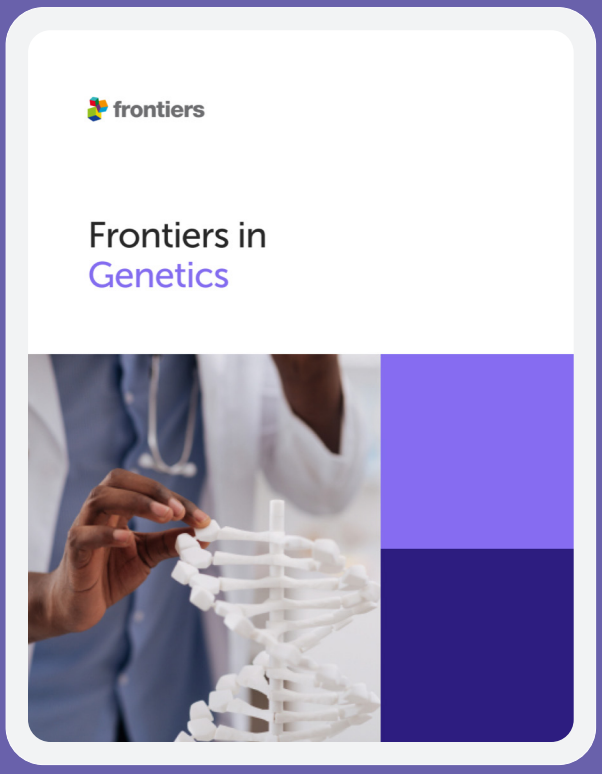
Frontiers in Genetics

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