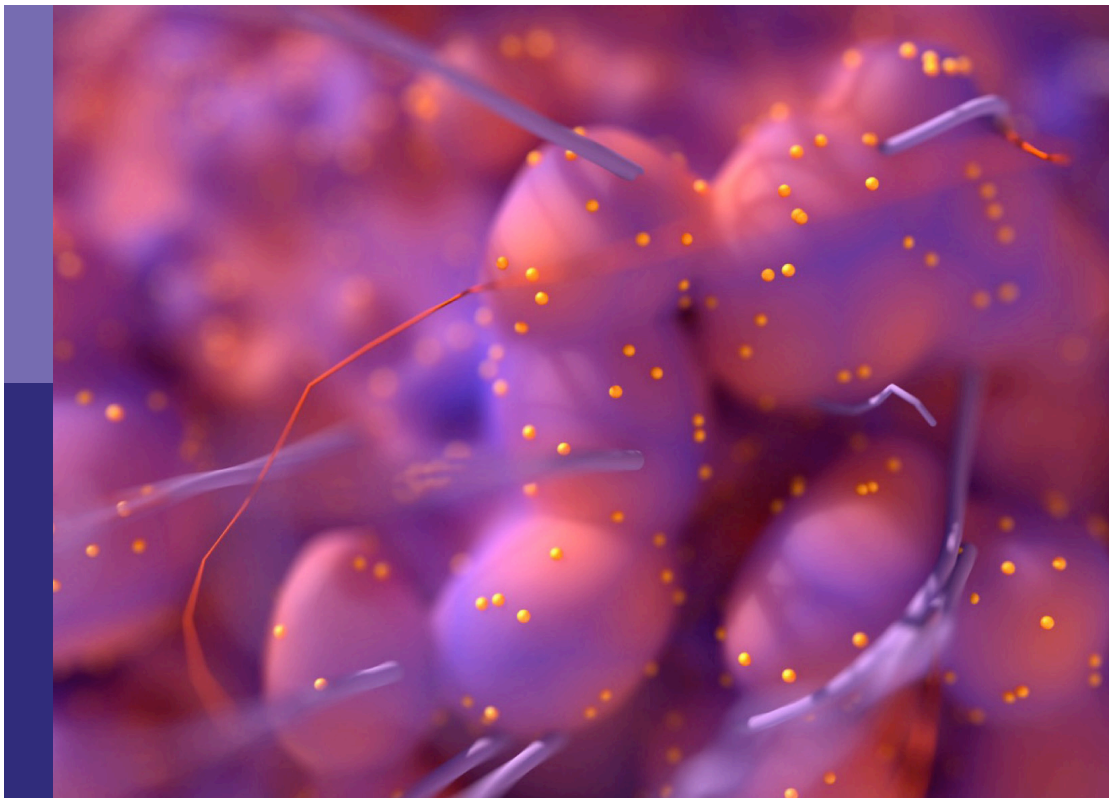
frontiers | Research Topics

The interconnection between epigenetic modifications and the tumor microenvironment

Edited by
Mingyue Tan, Fu Wang and Jian Song

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-83251-928-8
DOI 10.3389/978-2-83251-928-8

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

The interconnection between epigenetic modifications and the tumor microenvironment

Topic editors

Mingyue Tan — Shanghai University of Traditional Chinese Medicine, China

Fu Wang — Xi'an Jiaotong University, China

Jian Song — University Hospital Münster, Germany

Topic Coordinator

Feng Jiang — Fudan University, China

Citation

Tan, M., Wang, F., Song, J., eds. (2023). *The interconnection between epigenetic modifications and the tumor microenvironment*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-928-8

Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 05 April 2023
DOI 10.3389/fonc.2023.1166676

OPEN ACCESS

EDITED AND REVIEWED BY
Luisa Lanfrancone,
European Institute of Oncology (IEO), Italy

*CORRESPONDENCE
Feng Jiang
dxyjiang@163.com

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 15 February 2023
ACCEPTED 17 March 2023
PUBLISHED 05 April 2023

CITATION
He S and Jiang F (2023) Editorial: The interconnection between epigenetic modifications and the tumor microenvironment.
Front. Oncol. 13:1166676.
doi: 10.3389/fonc.2023.1166676

Editorial: The interconnection between epigenetic modifications and the tumor microenvironment

Siyuan He[1] and Feng Jiang[2]*

[1]School of Basic Medicine, Youjiang Medical University for Nationalities, Baise, China, [2]Department of Neonatology, Obstetrics and Gynecology Hospital of Fudan University, Shanghai, China

KEYWORDS

epigenetic modification, tumor microenvironment, immunotherapy, cancer biomarkers, oncology

Editorial on the Research Topic
The interconnection between epigenetic modifications and the tumor microenvironment

In the 1940s, Conrad Waddington coined the term 'epigenetics' to explain how the same combination of genes could produce different phenotypes in certain specific environments throughout animal development. The term epigenetics was later embraced by a wider range of disciplines and expanded to include the study of covalent and non-covalent changes in DNA and histones, as well as general alterations in chromatin structure in any biological or pathological process. For example, DNA methylation, post-translational changes in histones, chromatin remodeling, and the effects of non-coding RNAs on ribosome structure all fall within the field of epigenetic research. These epigenetic modifications translate environmental input signals into different gene combinations, allowing a limited number of transcription factors (TFs) to produce more diverse transcriptional patterns. The expression levels and biological activities of enzymes and regulators involved in epigenetic modifications may also be altered by environmental signals. Such heritable epigenetic changes with intertwined DNA/RNA/protein linkages provide a basis for studying environmental adaptations at the cellular level.

The tumor microenvironment (TME) is composed of cellular and non-cellular components, including stromal cells, immune cells, and chemokines (1). The biological importance of the TME as a response platform regulating various aspects of tumor initiation, development, metastatic progression, altered immune response, fulminant disease, and cancer recurrence is undeniable and constantly confirmed, as highlighted by numerous studies. In addition, as epigenetic alterations are associated with the control of the TME, DNA methylation may influence cancer growth by regulating immune infiltration and immune checkpoints of the TME (2, 3). In addition, histone acetylation may attenuate the immune destructive potential of the TME and promote tumor development. The main role of RNA modifications in tumor formation is to regulate angiogenesis, immune activity, and the infiltration of immune cells into the TME (4). ncRNAs released by certain cells in the TME are thought to influence the behavior of cancer cells, including invasion, metastasis, and treatment resistance (5). Furthermore, ncRNAs in tumor cells may be implicated in the

immune regulation of the TME and promote tumor formation. In conclusion, epigenetic alterations hold therapeutic promise in controlling elements of the tumor microenvironment and may be a target for cancer therapy.

The objective of the Research Topic titled "*The interconnection between epigenetic modifications and the tumor microenvironment*" is to discuss recent advances in the interaction between epigenetic modifications and the tumor microenvironment and to identify potential prognostic markers and specific components that may affect the efficacy of immunotherapy and other tumor treatments. Ultimately, a total of 11 papers, contributed by more than 60 authors as experts in the field, were accepted in 30 submissions, providing new comprehensive insights for future cancer therapies.

Han et al. identified prognostic genes significantly associated with metabolic changes in hepatocellular subpopulations at the single-cell level and examined the heterogeneity of the subpopulation and its interrelationship with other cells in the tumor microenvironment. A prognostic model for predicting overall survival (OS) in patients with hepatocellular carcinoma was established, validated, and found to show good predictive ability. In addition, differences in chemosensitivity between high- and low-risk groups were assessed and five drugs were focused on that could potentially reverse the risk score.

Kahlert et al.'s original research established COL10A1, a short-chain protein belonging to the collagen family, an important component of the stromal extracellular matrix, as a diagnostic marker to predict the development of colorectal cancer, expanding on previous studies on this protein. The authors found that the abundance of COL10A1 in CRC tissue predicts the metastatic and immunogenic potential of CRC and that COL10A1 transcription may mediate the interaction between tumor cells and the stromal microenvironment.

Cui et al. investigated the function of cuproptosis-related lncRNAs in colorectal cancer (COAD). They identified six cuproptosis-associated prognostic lncRNAs in COAD and constructed a prognostic model based on cuproptosis-associated lncRNAs, providing new insight into the risk classification and possible biomarkers for patients with colorectal cancer. Analysis of the immune microenvironment, mutations, and sensitivity to chemotherapy suggests that this signature may serve as a reference for immunotherapeutic and chemotherapeutic approaches.

Yu et al. constructed a pyroptosis-related lncRNA prognostic model for predicting prostate cancer using a machine-learning approach. The researchers explored the association between the prognostic model and patients' clinical characteristics, immune environment, immune checkpoints, gene mutations, and drug sensitivity, and constructed diagnostic and prognostic biomarkers for prostate cancer. *In vitro* experiments showed that silencing lncRNA AC005253.1 affected the expression of the AIM2 gene in prostate cancer and inhibited the proliferation, migration, and invasion of DU145 and PC-3 cells. In addition, silencing of AC005253.1 promoted the expression of pyroptosis inflammasome AIM2, and the pyroptosis-related gene AC005253.1 may be a valuable oncogene related to the prognosis of prostate cancer.

Xu et al. discussed the mechanism of MARCH1 in lung adenocarcinoma (LUAD). As a member of the E3 family, E3s were dysregulated in LUAD and were positively correlated with most immunological features, suggesting that MARCH1 may activate inflammatory TME in LUAD. Patients with LUAD with reduced MARCH1 expression had a poorer prognosis and were not sensitive to immune checkpoint inhibitors. In pan-cancer studies, MARCH1 was associated with most immunological features, suggesting that MARCH1 may be a new and promising biomarker as an indication of the immune status and effectiveness of immunotherapy in patients with LUAD.

Zeng et al. used a computational algorithm to screen out the fatty acid metabolism (FAM)-related genes associated with cervical cancer (CC) from the public databases. The FAM model (PLCB4, FBLN5, TSPAN8, CST6, and SERPINB7) risk score was an independent factor affecting the prognosis of patients with cervical cancer. This model had a high prognostic value, meaning that the FAM-related genes can be used as prognostic markers and potential immunotherapy targets for patients with CC.

The abovementioned Research Topic "*The interconnection between epigenetic modifications and the tumor microenvironment*" gathers studies focusing on the discovery of interactions between epigenetic modifications and the tumor microenvironment and the mechanisms of epigenetic modifications in immunotherapy against cancer. Several prognostic and predictive models have also been constructed that are useful for clinicians. It is hoped that this Research Topic will contribute to the understanding of the mechanisms of tumor development and provide new and broader insights into future cancer treatment.

Author contributions

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

Acknowledgments

We thank all the authors of the manuscripts included in this Research Topic for their contributions. We also thank all the staff of Frontiers in Oncology who manage and support this Research Topic.

Conflict of interest

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

Publisher's note

References

1. Kim J, Bae JS. Tumor-associated macrophages and neutrophils in tumor microenvironment. *Med Inflamm* (2016) 2016:6058147. doi: 10.1155/2016/6058147

2. Xiao Y, Yu D. Tumor microenvironment as a therapeutic target in cancer. *Pharmacol Ther* (2021) 221:107753. doi: 10.1016/j.pharmthera.2020.107753

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

4. Xu Y, Liao W, Luo Q, Yang D, Pan M. Histone acetylation regulator-mediated acetylation patterns define tumor malignant pathways and tumor microenvironment in hepatocellular carcinoma. *Front Immunol* (2022) 13:761046. doi: 10.3389/fimmu.2022.761046

5. Sang LJ, Ju HQ, Liu GP, Tian T, Ma GL, Lu YX, et al. LncRNA CamK-a regulates Ca(2+)-Signaling-Mediated tumor microenvironment remodeling. *Mol Cell* (2018) 72:71–83. doi: 10.1016/j.molcel.2018.08.014

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 23 September 2022
DOI 10.3389/fonc.2022.1007918

OPEN ACCESS

EDITED BY
Jian Song,
University Hospital Münster, Germany

REVIEWED BY
Shizhi Wang,
Southeast University, China
Liang Chen,
Fuyang Hospital Affiliated to Anhui Medical University, China

*CORRESPONDENCE
Fei Wei
weifei@njucm.edu.cn
Zhiguang Sun
zhiguangsun@njucm.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 31 July 2022
ACCEPTED 23 August 2022
PUBLISHED 23 September 2022

CITATION
Cui G, Liu J, Wang C, Gu R, Wang M, Sun Z and Wei F (2022) Comprehensive analysis of the prognostic signature and tumor microenvironment infiltration characteristics of cuproptosis-related lncRNAs for patients with colon adenocarcinoma.
Front. Oncol. 12:1007918.
doi: 10.3389/fonc.2022.1007918

Comprehensive analysis of the prognostic signature and tumor microenvironment infiltration characteristics of cuproptosis-related lncRNAs for patients with colon adenocarcinoma

Guoliang Cui[1†], Jinhui Liu[2†], Can Wang[3†], Renjun Gu[4†], Manli Wang[5], Zhiguang Sun[1*] and Fei Wei[6*]

[1]Department of Gastroenterology, The Second Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China, [2]Department of Gynecology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China, [3]Department of Colorectal Surgery, Jiangsu Province Hospital of Chinese Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China, [4]School of Traditional Chinese Medicine and School of Integrated Chinese and Western Medicine, Nanjing University of Chinese Medicine, Nanjing, China, [5]The First Clinical Medical College, Nanjing University of Chinese Medicine, Nanjing, China, [6]Department of Physiology, School of Medicine & Holistic Integrative Medicine, Nanjing University of Chinese Medicine, Nanjing, China

Background: Cuproptosis, a newly described method of regulatory cell death (RCD), may be a viable new therapy option for cancers. Long noncoding RNAs (lncRNAs) have been confirmed to be correlated with epigenetic controllers and regulate histone protein modification or DNA methylation during gene transcription. The roles of cuproptosis-related lncRNAs (CRLs) in Colon adenocarcinoma (COAD), however, remain unknown.

Methods: COAD transcriptome data was obtained from the TCGA database. Thirteen genes associated to cuproptosis were identified in published papers. Following that, correlation analysis was used to identify CRLs. The cuproptosis associated prognostic signature was built and evaluated using Lasso regression and COX regression analysis. A prognostic signature comprising six CRLs was established and the expression patterns of these CRLs were analyzed by qRT-PCR. To assess the clinical utility of prognostic signature, we performed tumor microenvironment (TME) analysis, mutation analysis, nomogram generation, and medication sensitivity analysis.

Results: We identified 49 prognosis-related CRLs in COAD and constructed a prognostic signature consisting of six CRLs. Each patient can be calculated for a risk score and the calculation formula is: Risk score =TNFRSF10A-AS1 * (-0.2449) + AC006449.3 * 1.407 + AC093382.1 *1.812 + AC099850.3 * (-0.0899) + ZEB1-AS1 * 0.4332 + NIFK-AS1 * 0.3956. Six CRLs expressions were investigated by qRT-PCR in three colorectal cancer cell lines. In three cohorts, COAD patients were identified with different risk groups, with the high-risk group having a worse prognosis than the low-risk group.

Furthermore, there were differences in immune cell infiltration and tumor mutation burden (TMB) between the two risk groups. We also identified certain drugs that were more sensitive to the high-risk group: Paclitaxel, Vinblastine, Sunitinib and Elescloml.

Conclusions: Our findings may be used to further investigate RCD, comprehension of the prognosis and tumor microenvironment infiltration characteristics in COAD.

KEYWORDS

colon adenocarcinoma, cuproptosis, prognosis, tumor immune microenvironment, bioinformatics

Introduction

Colorectal cancer (CRC) is the third most frequent cancer and the second major cause of cancer-related death worldwide (1, 2). In 2018, 1.8 million new instances of CRC were diagnosed, with over 800,000 deaths (3). The pathogenesis involves a chronic process, including precancerous lesions, activation of tumor stem cells, accumulation of genetic and epigenetic changes (4). CRC is a heterogeneous disease with widespread chromosomal instability and microsatellite instability (5). CRC morbidity and mortality are declining in most developed countries due to early screening and prevention of early risk factors (6). However, the situation of CRC in developing countries is still very serious (7). The pathogenesis of CRC involves a series of multi-step changes, including histological, morphological, and genetic changes (8, 9). Unhealthy diet, obesity, smoking and alcohol consumption are considered risk factors for CRC (10, 11). The 5-year overall survival (OS) rate of localized and regionalized CRC patients is impressively high, but decreases to 14% once metastasis occurs (12). In the past decade, immunotherapy has become a hot topic in refractory solid tumors due to its long-term response. Immunotherapy significantly inhibit the progression of advanced malignant tumors and prolong the survival of patients, which brings hope to CRC patients (13). Colon adenocarcinoma (COAD) is the ordinary histological subtype of CRC, therefore, it is of great value to explore a new prognosis assessment protocol and to establish a predictive signature for immunotherapy and immune microenvironment of COAD.

In recent years, regulatory cell death (RCD) plays an important role in maintaining normal homeostasis of body development and inhibiting rapid proliferation of tumor cells, which is considered as a new direction of tumor therapy (14, 15). In recent years, the most widely studied types of RCD are apoptosis, pyroptosis, necroptosis and ferroptosis (16, 17). Different from the known mechanism of cell death, Tsvetkov et al. found that cuproptosis is a new form of cell death, namely the existence of a copper-dependent, regulated cell death in human cells (18). Cuproptosis relies on the effect of copper ions on mitochondrial tricarboxylic acid metabolism, resulting in abnormal aggregation of lipoacylated proteins and loss of iron-sulfur (Fe-S) cluster proteins, which leads to the proteotoxic stress response of tumor cells and cell death (18, 19). However, the mechanism of RCD and its role in tumor microenvironment have not been thoroughly studied, which may play a double-edged sword role in tumors (17, 20). On the one hand, inducing tumor cell death can cure tumors; On the other hand, when the inflammatory response caused by cell death reaches a certain level, many signaling pathways can be activated, leading to tumor progression (21). Therefore, it is of great significance to explore the role of RCD in tumors, and many studies have established RCD-associated prognostic models to assess prognosis and immune microenvironment (22–24). However, the clinical significance of cuproptosis and cuproptosis-associated prognostic model have not been reported, especially in COAD.

Long noncoding RNAs (lncRNAs) can regulate gene expression through epigenetic regulation, transcriptional regulation and post transcriptional regulation, so as to participate in a variety of biological processes such as tumor cell proliferation, differentiation and apoptosis (25). Therefore, lncRNAs are considered as promising biomarkers and potential therapeutic targets for the diagnosis and treatment of various diseases, including COAD (26). More and more attention has been paid to the role and molecular mechanism of regulating RCD-related lncRNAs in tumor pathology (27). Therefore, as a new type of RCD, the identification of lncRNAs related to cuproptosis is of great significance for understanding the pathogenesis of tumor and providing new targets for prevention and treatment.

Through bioinformatics research, we investigated the importance of cuproptosis-related lncRNAs (CRLs) and a prognostic signature based on CRLs was constructed in COAD. A risk score for each COAD patient might be determined based on cuproptosis-associated model, which can be used for prognosis assessment, immunological prediction, and mutation analysis. Our findings may be useful in determining the prognosis and therapy of patients with COAD.

Methods and materials

Download of data

The TCGA database was used to obtain the data of COAD RNA sequencing and clinical information. We collected data from 473 tumor samples and 41 healthy controls. From the previously published publications (18, 28), thirteen cuproptosis-related genes were obtained, including FDX1, LIPT1, LIAS, DLD, DBT, GCSH, DLST, DLAT, PDHA1, PDHB, SLC31A1, ATP7A, and ATP7B.

Screening lncRNAs associated with cuproptosis

From the TCGA database, 1053 CRLs were identified using Pearson correlation analysis and a co-expression network was created based on the cutoff (Pearson R > 0.4 and $P < 0.001$) (29). Then, using univariate Cox regression analysis and forest maps, 49 CRLs with potential prognostic significance for COAD were identified. The "limma", "pheatmap", "reshape2", and "ggpubr" programs were used to create heat maps and boxplots to show the differential expression of CRLs in COAD and normal tissues, with following criteria: $|\log_2$ fold change (FC)| >1 and false discovery rate (FDR) < 0.05.

Consensus clustering analysis

To preliminarily understand the underlying the mechanism of the biological function of cuproptosis-related lncRNAs, The "ConsensusClusterPlus" package was used to construct a consensus cluster with 49 CRLs (K represents cluster count) (30). The cluster exhibited the best stability when K = 3 based on the similarity of expression levels of CRLs and the proportion of fuzzy clustering measures. As a result, 417 CRC patients were divided into three clusters: cluster 1 (n = 139), cluster 2 (n = 202), and cluster 3 (n = 76). The variations in survival, CRLs expression, and clinical characteristics were then compared among the three clusters. Immune checkpoint inhibitors co-expression (PD-L1, CTLA-4), immune cell content differences, and immunological score (including ESTMATE score, immune score and stromal score) were also investigated.

Construction and evaluation of the prognostic model

All COAD patients were randomly assigned to one of two groups: training cohort or validation cohort (testing cohort and entire cohort). No significant difference was observed in three cohort for the clinical-pathological factors (Supplementary Table 1). In the training cohort, LASSO and multivariate Cox regression analysis were used to identify prognostic model based on CRLs. The risk score was calculated using the following formula: coef (lncRNA1) × expr (lncRNA1) + coef (lncRNA2) × EXPR (lncRNA2) +… + coef (lncRNAn) × expr (lncRNAn), coef stands for coefficient, coef (lncRNAn) stands for coefficient of survival linked lncRNA, and expr (lncRNAn) stands for lncRNA expression. Patients in the training set were separated into two groups based on their median risk score: high-risk and low-risk. For survival analysis, the R packages "survival" and "survminer" were used by Kaplan-Meier curve, and a ROC curve was plotted (31). Finally, we run the above analyses in a validation cohort to verify the predictive power of the results.

Independent prognostic value assessment of the prognostic model

When paired with other clinical factors, univariate and multivariate Cox regression analysis were performed to determine whether risk score was an independent predictive factor in COAD patients (32).

Cell culture

Three human colorectal cancer cell lines (Caco-2, HT-29, HCT116) were all purchased from the China Center for Type Culture Collection (CCTCC, Wuhan, China). The normal colon epithelial cell line (FHC) was obtained from the Cell Bank of Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China). Caco-2, HT-29 cells, HCT116 and FHC were cultured in McCoy's 5A, RPMI-1640, high-glucose DMEM medium (Gibco, Shanghai, China) respectively, which were supplemented with 10% fetal bovine serum (FBS, Gibco, Shanghai, China) and 1% antibiotics. All cells were incubated at 37°C with 5% CO_2.

Quantitative RT-PCR

Total RNA from the cell lines was isolated with TRIZOL reagent (Thermo Fisher Scientific, USA). Complementary DNA (cDNA) was synthesized and quantitative RT-PCR was performed using SYBR qPCR Master Mix (Vazyme, China). The relative expression of the target gene was analyzed using the $2^{-\Delta\Delta CT}$ method and β-actin was chosen as the internal reference. The primer sequences are listed in Supplementary Table 2.

Construction and evaluation of the nomogram

Based on risk scores and patient clinical information, a nomogram was created to predict 1-year, 3-year, and 5-year OS (33). The Hosmer-Lemeshow test was used to construct modified curves to show the agreement between the actual and anticipated outcomes. The accuracy of the nomogram was assessed using ROC curves (34).

Gene set enrichment analysis

The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified using gene set enrichment analysis (GSEA) (35). The GSEA website (http://software.broadinstitute.org/gsea/index.jsp) was utilized to identify gene-level enrichment. Based on the risk score model, COAD samples from the entire set were separated into high-risk and low-risk groups. The underlying biological functions of the two groups were compared. The molecular signature database (MSigDB, http://software.broadinstitute.org/gsea/msigdb/index.jsp) collection of annotated gene sets was chosen as a reference gene set in the GSEA software. The cut-off criterion was set at a notional $P < 0.05$. As a reference document, we use "c2.cp.kegg.v7.4.symbols.gmt."

Evaluation of the immune microenvironment

For immune score, stromal score, estimated score, each sample was evaluated using a "estimated" R-package. The proportion of immune to stromal components in the tumor microenvironment is represented by these scores. Pearson correlation coefficient approach was used to assess the correlations between immune score, stromal score, estimated score, and risk score. Based on TCGA RNA sequencing data, the CIBERSORT tool was utilized to quantify 22 types of immune cell components (36). TIMER, CIBERSORT, Cibersort-ABS, QUANTISEQ, MCPCOUNTER, XCELL, and EPIC databases were all used to calculate immune cell infiltration. The Pearson correlation coefficient approach was used to assess the link between immune cell infiltration and CRLs expression level, risk score. One-class logistic regression (OCLR) machine-learning algorithm was used to quantify the stemness of tumor samples by calculating cancer stem cell indices (37).

Mutation analysis

TCGA provided mutation data (data category = copy number variation; "Maf" file). The top 20 mutant genes were visualized using waterfall diagrams created by the R software package "MAftools" (38). The tumor mutation burden (TMB), which is the number of somatic mutations per Megabyte genome sequence, can be used to identify patients who will respond better to immune checkpoint inhibitors (ICIs) (39). The differences of TMB between the two risk groups were investigated, as well as their correlation with risk score. The m^6A-related genes and human leukocyte antigen (HLA) genes were compared between the two risk groups using the "limma" package (40).

Drug sensitivity analysis

We compared the IC50 differences of the four chemotherapeutic drugs between the two risk groups using the R software package "PRROPHIC" (41). Using the R package "ggplot2," researchers discovered a link between six CRLs and chemotherapeutic sensitivity (42). The relationship between CRLs expression and drug susceptibility was investigated using Pearson correlation analysis.

Statistical analysis

The continuous variables in normal distribution are analyzed by Student's t-test, which is presented as mean ± standard deviation, and the continuous variables in abnormal distribution are presented as median (range). A *p*-value less than 0.05 was considered as statistical significance.

Results

Screening of cuproptosis-related lncRNAs with prognostic value

To identify lncRNAs associated with cuproptosis-related genes (CRGs), we performed co-expression analysis to reveal the correlation. Firstly, the co-expression network demonstrated the interaction between CRGs and CRLs (Figure 1A). Following that, using univariate COX regression analysis, 49 CRLs with prognostic value were identified (Figure 1B). The heat map and box plots indicated the expression difference of 49 CRLs between COAD and normal tissues (Figures 1C, D).

Consensus clustering analysis

The 49 CRLs were then put through a consensus clustering analysis to determine whether they might be used to stratify COAD patients. Based on the "ConsensusClusterPlus" program, a consensus cluster consisting of 49 CRLs was built (K represents cluster count, Supplementary Figure 1). The clustering exhibited

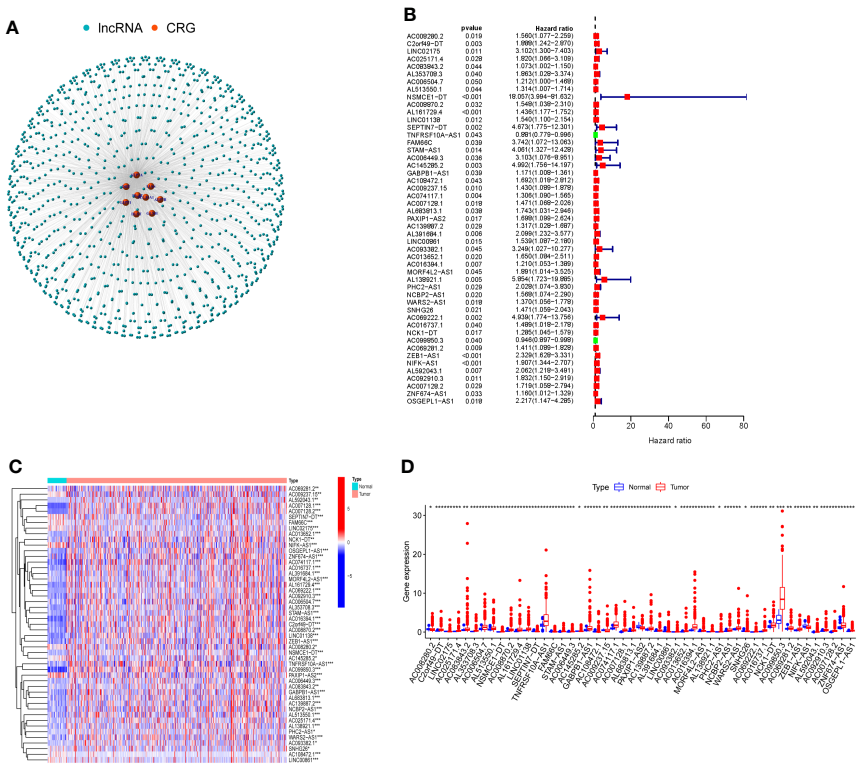


FIGURE 1
Cuproptosis-related lncRNAs (CRLs) with prognostic value were screened. **(A)** Co-expression network of cuproptosis-related genes and CRLs by Pearson correlation analysis. **(B)** There were 49 CRLs found to have prognostic value by COX regression analysis with one variable. **(C)** Heat map of prognosis-related CRLs expression in COAD and normal tissues. **(D)** The expression of prognosis-related CRLs in COAD and normal tissues was shown in box plots. (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).

the best stability when K = 3 based on the similarity of CRL expression levels and the proportion of fuzzy clustering measures (Figure 2A). As a result, 417 COAD patients were split into three clusters: cluster 1 (n = 139), cluster 2 (n = 202), and cluster 3 (n = 76). The prognosis of the three clusters was significantly different in survival analysis (P=0.020), with Cluster 1 having the worst prognosis (Figure 2B). In the form of a heatmap, Figure 2C depicted the differences in CRLs expression and clinical features between the three groups. The immune checkpoint genes PD-L1 and CTLA-4, as well as these lncRNAs, were found to have a co-expression relationship (Figures 2C, D). Following that, the analysis revealed the difference of Stromal score, Immunological score, ESTIMATE score, as well as the abundance of T cells CD4 memory activated, T cells regulatory, T cells gamma delta and NK cell resting in three clusters (Figures 2E–L).

Construction and evaluation of the prognostic model

The 49 CRLs described above were then further examined in order to reduce the scope and build a predictive model. COAD patients were separated into two cohort: training and validation (testing and entire). Univariate Cox regression analysis was performed on the training cohort, 12 CRLs (NSMCE1-DT, AL161729.4, LINC01138, SEPTIN7-DT, TNFRSF10A-AS1, AC006449.3, AC093382.1, PHC2-AS1, AC099850.3, AC069281.2, ZEB1-AS1, NIFK-AS1) with prognosis value were identified in COAD (Table 1). The LASSO Cox regression model was used to narrow the most robust lncRNAs for prognosis and build prognostic models in the training cohort. Ten-fold cross-validation was applied to overcome the over-fitting. To generate a prognostic CRLs signature model, multivariate Cox regression analysis was applied to evaluate the connection between CRLs and OS in the training set. The model is more stable when LAMDA = 6. (Supplementary Figure 2). TNFRSF10A-AS1, AC006449.3, AC093382.1, AC099850.3, ZEB1-AS1, and NIFK-AS1 were included in this model. The calculation formula is: Risk score =TNFRSF10A-AS1 * (-0.2449) + AC006449.3 * 1.407 + AC093382.1 *1.812 + AC099850.3 * (-0.0899) + ZEB1-AS1 * 0.4332 + NIFK-AS1 * 0.3956. The model was also tested in two validation cohort: testing, and the entire cohort. To begin, Figure 3A depicted the patient's risk score, survival status, and six CRLs expression level in the training cohort. Patients were split into high-risk and low-

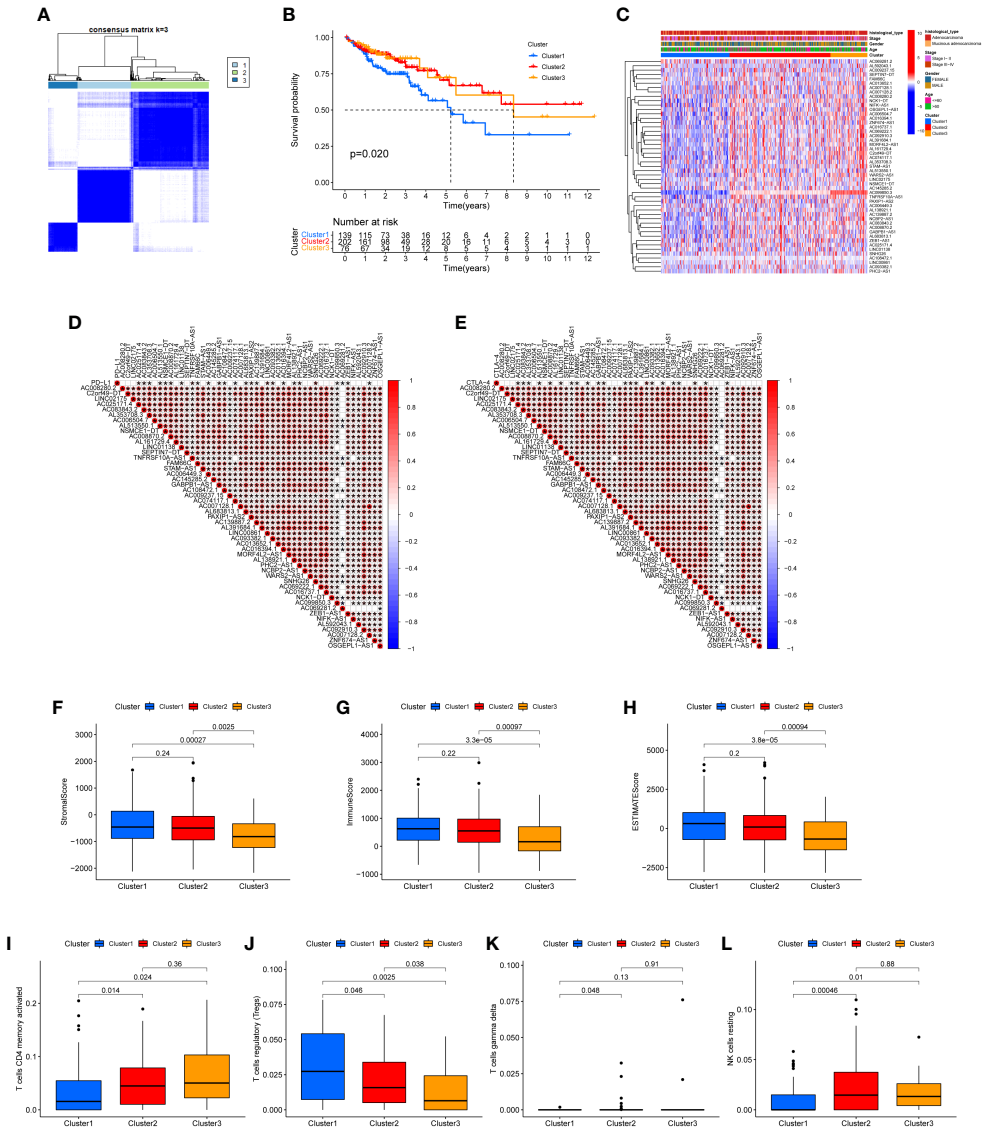


FIGURE 2
Clustering analysis *via* consensus. **(A)** When K = 3, the clustering was the most stable. **(B)** Survival analysis of the three clusters. The prognosis of Cluster 1 is the poorest. **(C)** A heat map depicting the differences in CRL expression and clinical features between the three clusters. **(D, E)** The immune checkpoints genes PD-L1 and CTLA-4, as well as prognosis-related CRLs, have a co-expression relationship. Stromal score **(F)**, Immune score **(G)**, ESTIMATE score **(H)**, the abundance of T cells CD4 memory activated **(I)**, T cells regulatory **(J)**, T cells gamma delta **(K)** and NK cell resting **(L)** in three clusters were shown.

risk groups based on median risk score, and survival analysis revealed that the high-risk group's prognosis was significantly worse (Figure 3B, $P<0.001$). We further validated the expression of six CRLs in colorectal cancer cell lines. As shown in Supplementary Figure 3, the expressions of TNFRSF10A-AS1, AC099850.3, ZEB1-AS1 and NIFK-AS1 were significantly higher in tumor cells compared to those in FHC cells. Meanwhile, AC006449.3 expression was upregulated in HT-29 cells, but downregulated in HCT-116 cells. Analogously, AC093382.1 expression was significantly higher in Caco-2 and HT-29 cells but lower in HCT-116 cells (Supplementary Figures 3A–F). The AUC values for 1, 3, and 5 years in the training cohort were 0.700, 0.691, and 0.807 respectively, according to the results of ROC curve (Figure 3C). According to the same risk score calculation formula, different risk scores and survival status of patients in the testing and entire cohort were identified, and the difference of six CRLs expression level between the two risk groups was also analyzed (testing cohort, Figure 3D; entire

TABLE 1 Univariate Cox analysis generated 12 CRLs that are significantly related to the overall survival (OS).

lncRNA	HR	Lower 95% CI	Higher 95% CI	*P* value
NSMCE1-DT	18.1956	1.6439	201.3981	0.0180
AL161729.4	1.4211	1.0226	1.9749	0.0363
LINC01138	1.7057	1.0521	2.7654	0.0303
SEPTIN7-DT	13.2654	1.4138	124.4671	0.0236
TNFRSF10A-AS1	0.8024	0.6682	0.9635	0.0184
AC006449.3	7.1223	1.6987	29.8617	0.0073
AC093382.1	9.8668	2.0216	48.1573	0.0047
PHC2-AS1	2.5447	1.0419	6.2149	0.0404
AC099850.3	0.9164	0.8494	0.9886	0.0241
AC069281.2	1.5789	1.0899	2.2873	0.01573
ZEB1-AS1	2.3421	1.4098	3.8907	0.0010
NIFK-AS1	1.7646	1.0950	2.8437	0.0197

cohort Figure 3G). In addition, the outcomes of patients with higher risk score in the validation cohort were significantly worse (Figure 3E, H). The 5-year AUC value of testing and entire cohort were 0.683 and 0.748 respectively, according to the ROC curve (Figures 3F, I). Moreover, risk score was found to be an independent prognostic factor for COAD patients in the above three cohorts using both univariate and multivariate COX regression (Supplementary Figures 4A–F). Survival analysis revealed the prognostic value of risk score in COAD patients with different ages, different stages, different genders and different histological types (Supplementary Figures 5A–H).

Construction and evaluation of the nomogram

Following that, nomogram was utilized to combine risk score and other clinical parameters to better evaluate the prognosis of COAD patients. We established a nomogram to assess COAD patients' 1-, 3-, and 5-year survival rates (Figure 4A). The calibration curves revealed that the nomogram was accurate in predicting 1-, 3-, and 5-years survival rates (Figure 4B). Compared with the AUC value of the clinical features, risk score could be used to predict the OS of COAD patients (Figure 4C).

Gene set enrichment analysis

GSEA was used to investigate the variations in pathway enrichment between high-risk and low-risk groups. Allograft rejection, asthma, autoimmune thyroid disease, cell adhesion molecules CAMs, and systemic lupus erythematosus were among the enriched KEGG pathways in the high-risk group (Figure 5A). However, aminoacyl tRNA biosynthesis, dna replication, nucleotide excision repair, O-glycan biosynthesis, and oocyte meiosis were among the enriched KEGG pathways in the low-risk group (Figure 5B). GSEA results revealed that patients with higher risk score were related to immune-related pathways while lower risk score patients were associated with tumor-related pathway, which maybe explained the survival different in two risk groups.

Analysis of immune microenvironment

The occurrence and development of tumor are affected by immune microenvironment and the study of immune microenvironment can provide reference for immunotherapy of tumor. Stromal, Immune and ESTIMATE scores were significantly different between high-risk and low-risk groups, and may be higher in high-risk groups (Figures 5C–E). Pearson correlation analysis revealed a strong positive correlation between Stromal score, ESTIMATE score, and Immune score and risk score (Figures 5F–H). Figure 6A depicted the immunological landscape of high-risk and low-risk groups as heatmap, using CIBERSORT, QUANTISEQ, MCPCOUNTER, XCELL, and EPIC algorithms. Between the two groups, different amounts of immune cell infiltration were detected and subsequent correlation analysis revealed the relation between six CRLs expression and immune cells. Except ZEB1-AS1 expression has no correlation with immune cells, the other five CRLs have different degrees of correlation with immune cells, among which AC099850.3 had the highest positive correlation with resting NK cells and highest negative connection between TNFRSF10A-AS1 and Treg cells and macrophage M0 (Figure 6B). In the shape of a box diagram, Figure 6C depicted the differences in immune cell infiltration levels and the levels of infiltration of T cells regulatory and dendritic cells were found to be substantially different between the high-risk and low-risk groups (Figure 6C). A Pearson correlation analysis revealed the correlation between different immune cells and risk scores (Figures 6D–K). Finally, DNA stem cell score (DNAss) was shown to be unrelated to risk score (Figure 6L), however, RNA stem cell score (RNAss) was found to be significantly inversely associated to risk score (R= -0.38, P = 2.5E-10, Figure 6M).

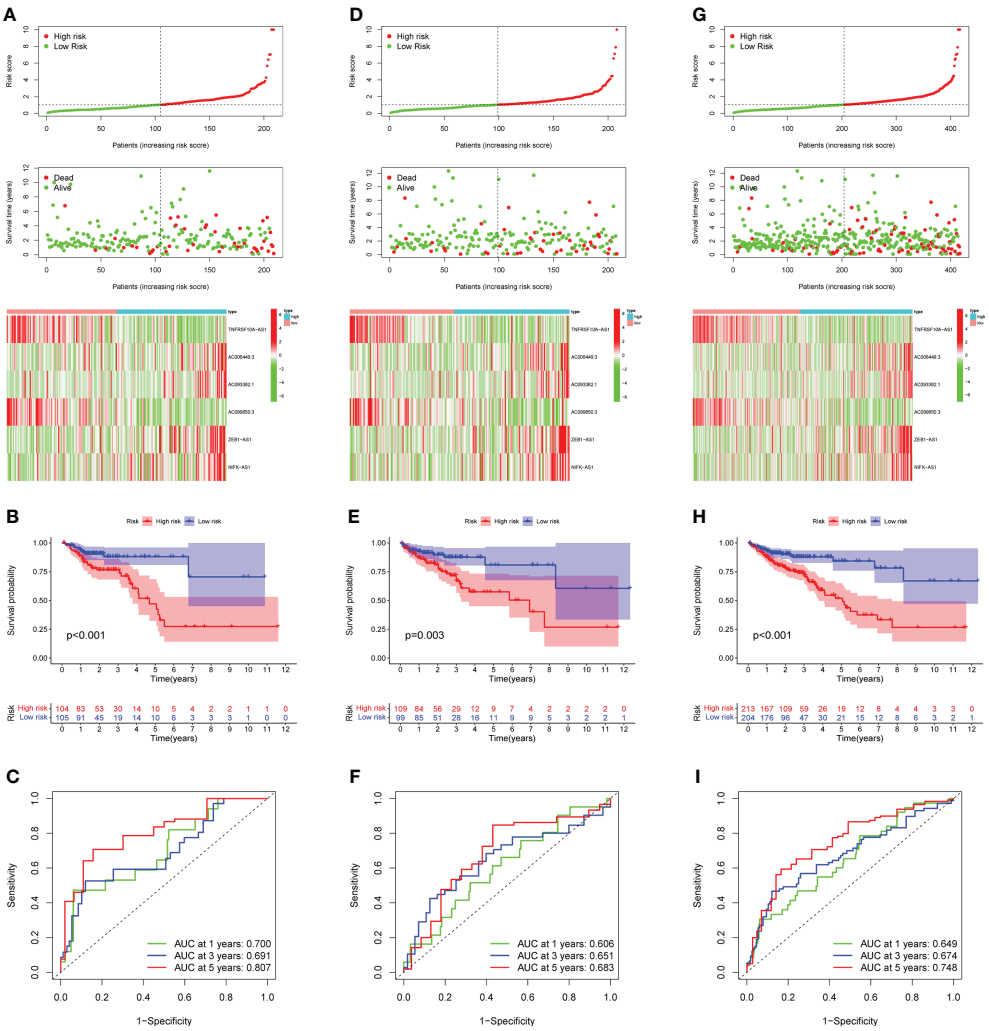


FIGURE 3
Prognostic model construction and evaluation. **(A)** In the training cohort, the patient with different risk score, survival status, and CRLs expression were shown. **(B)** Survival analysis of training cohort and the prognosis of high-risk group was significantly worse. **(C)** ROC curve revealed that in the training cohort, the AUC values for 1, 3, and 5-years OS were 0.700, 0.691, and 0.807, respectively. **(D, G)** In the testing and entire cohort, the patient's risk score, survival status, and CRLs expression were shown. **(E)**The high-risk group's prognosis in the testing cohort was also significantly worse. **(F)** In the testing cohort, AUC values of 1, 3, and 5-year OS were 0.606, 0.651, and 0.683, respectively, according to the ROC curve. **(H)** In the entire cohort, survival analysis revealed worse outcomes in the high-risk group. **(I)** In the entire cohort, the AUC values of the model in 1, 3, and 5-years OS were 0.649, 0.674 and 0.748, respectively.

Expression analysis of m^{6}A-related genes and human leukocyte antigen genes

N^6-methyladenosine (m^{6}A) is a type of tumor epigenetics that plays an important function in tumor progression and human leukocyte antigen (HLA) has been related to tumor immunotherapy (43). Therefore, it is necessary to investigate the differences in expression of m^{6}A-related genes and HLA-related genes between high-risk and low-risk groups. First, the expression of methylation-related genes HNRNPC, RBM15, YTHDC1, YTHDF3, YTHDF2, METTL14, WTAP, HNRNPA2B1, FMR1 was shown to differ between the high-risk and low-risk groups (Figure 7A). HLA-related genes including HLA-DQA1, HLA-DRB6, HLA-DQB1, HLA-DRB1, HLA-DPB1, HLA-L, HLA-DOA, HLA-DPA1, HLA-J, HLA-DQB2, HLA-DMA, HLA-E, HLA-DQA2, and HLA-G were shown to have varied levels of expression in high-risk and low-risk groups (Figure 7B).

Mutation analysis

Mutations in numerous genes are associated with tumor formation, and the tumor mutation burden (TMB) is thought to be a helpful signal for evaluating immune checkpoint-related

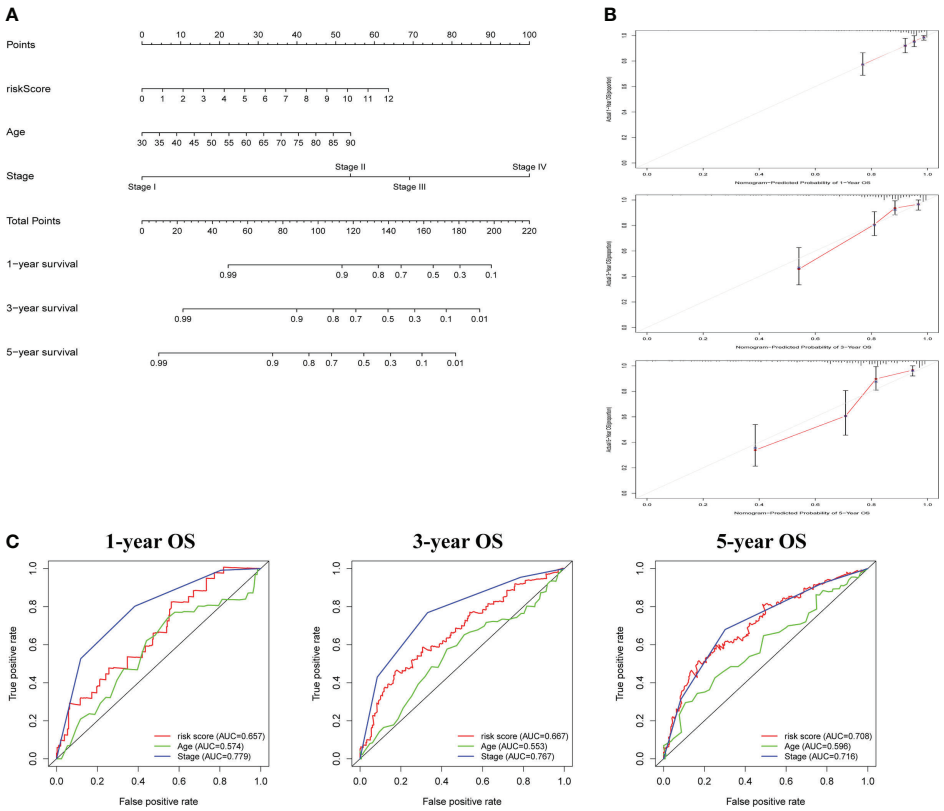


FIGURE 4
The construction and assessment of prognostic nomogram. **(A)** A nomogram for assessing 1-, 3-, and 5-year survival rates. **(B)**The calibration curves of the nomogram. **(C)**The ROC curve of risk score and clinical characteristics was investigated based on 1-, 3-, and 5-year OS.

therapy. The mutation landscape of the high-risk group was depicted in Figure 7C. APC has the highest mutation rate, as can be observed. Figure 7D depicted the mutation landscape of the low-risk group, with APC having the highest mutation rate. TMB was also different across the two groups, with TMB being higher in the low-risk group (Figure 7E). A Pearson correlation study revealed that TMB and risk score had a substantial negative correlation (Figure 7F). Next, we focus on the expression of mismatch repair (MMR) protein, as it plays a key role in the process of COAD and is a major cause of gene mutations and microsatellite instability (MSI) (44). The expressions of MMR-related proteins MLH1, MSH2, MSH6, and EPCAM were substantially up-regulated in the low-risk group ($P<0.05$), while PMS2 expression did not differ statistically between the two risk groups (Figures 7G–K).

Drug sensitivity analysis

Following the risk classification of COAD patients, medication sensitivity analysis can be used to identify effective treatments for different risk groups patients in order to individualized treatment. To begin with, paclitaxel, Vinblastine, Sunitinib, and Elesclomol had lower IC50 values in the high-risk group, indicating that the patients with higher risk score were more responsive to these medications (Figure 8A). Following that, correlation analysis was used to identify medicines that were significantly correlated with the expression of CRLs. For instance, the analysis demonstrated that up-regulated ZEB1-AS1 expression was associated with increased drug sensitivity of tumor cells to nelarabine, palbociclib, fluphenazine, asparaginase, LEE-011, ifosfamide, hydroxyurea and dexrazoxane, while increased ZEB1-AS1 expression was related to the increased resistance to vemurafenib in COAD patients (Figure 8B).

Discussion

COAD is still one of the most common cancer-related deaths in the world (45). Screening strategies, such as fecal occult blood test, screening colonoscopies and fecal immunochemical tests, can greatly reduce the incidence and mortality of COAD, but there are still many limitations of screening tests, and a large

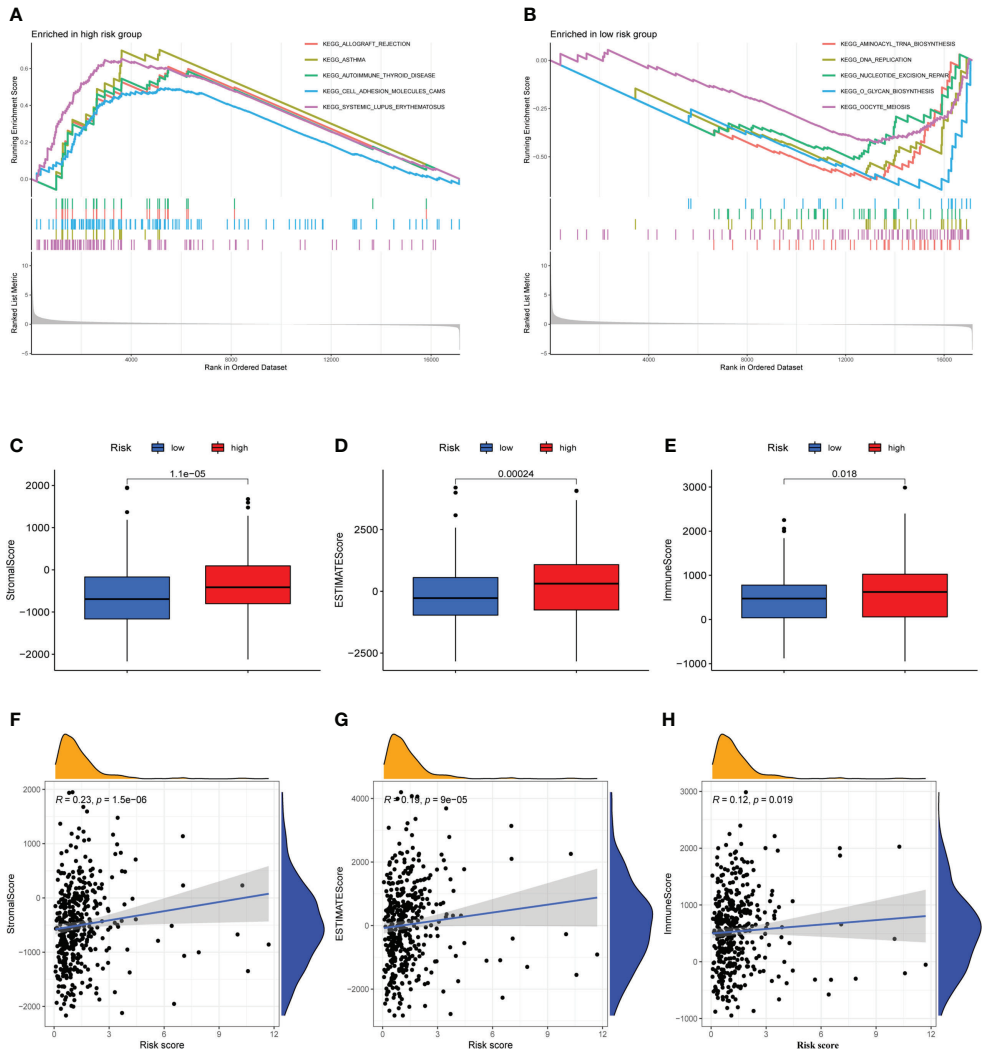


FIGURE 5
Immune microenvironment analysis and gene set enrichment analysis (GSEA). **(A)** Allograft rejection, asthma, autoimmune thyroid disease, cell adhesion molecules CAMs, and systemic lupus erythematosus are among the enriched KEGG pathways in the high-risk group. **(B)** Aminoacyl tRNA biosynthesis, dna replication, nucleotide excision repair, O-glycan biosynthesis, and oocyte meiosis are among the enriched KEGG pathways in the low-risk group. **(C–E)** The high-risk and low-risk groups had considerably different Stromal, ESTIMATE, and Immune scores, with the high-risk group having higher scores. **(F–H)** A Pearson correlation analysis revealed a strong positive relation between Stromal score, ESTIMATE score, Immune score and risk score.

number of people eligible for screening miss the opportunity of screening (46, 47). With the innovations of risk stratification and development of personalized screening, the burden of COAD might be further reduced. At present, the treatment of advanced COAD remains a challenge due to stubborn drug resistance, metastasis and recurrence (48). Therefore, there is an urgent need to explore novel signatures for patients with COAD to assess prognosis, identify high-risk populations and guide personalized treatment. In recent years, regulatory cell death (RCD) has also been recognized as a promising target for cancers (49). Among them, cuproptosis is a copper-dependent and regulated new cell death mode, which is different from other known cell death regulation mechanisms (18). Further research on copper dependent cell death can provide a basis for the intervention of copper metabolism dysfunction related diseases and the potential application of anti-tumor. Therefore, cuproptosis may have complex crosstalk with metabolic reprogramming in cancers. While a number of RCD-related prognostic models have been developed to assess prognosis and immune microenvironment, this study is mainly report prognostic signature associated with cuproposis-related lncRNA (CRLs), which helps us understand the roles of cuproposis and CRLs in COAD.

Here, we performed a comprehensive bioinformatics analysis to explore the significance of CRLs in COAD, co-

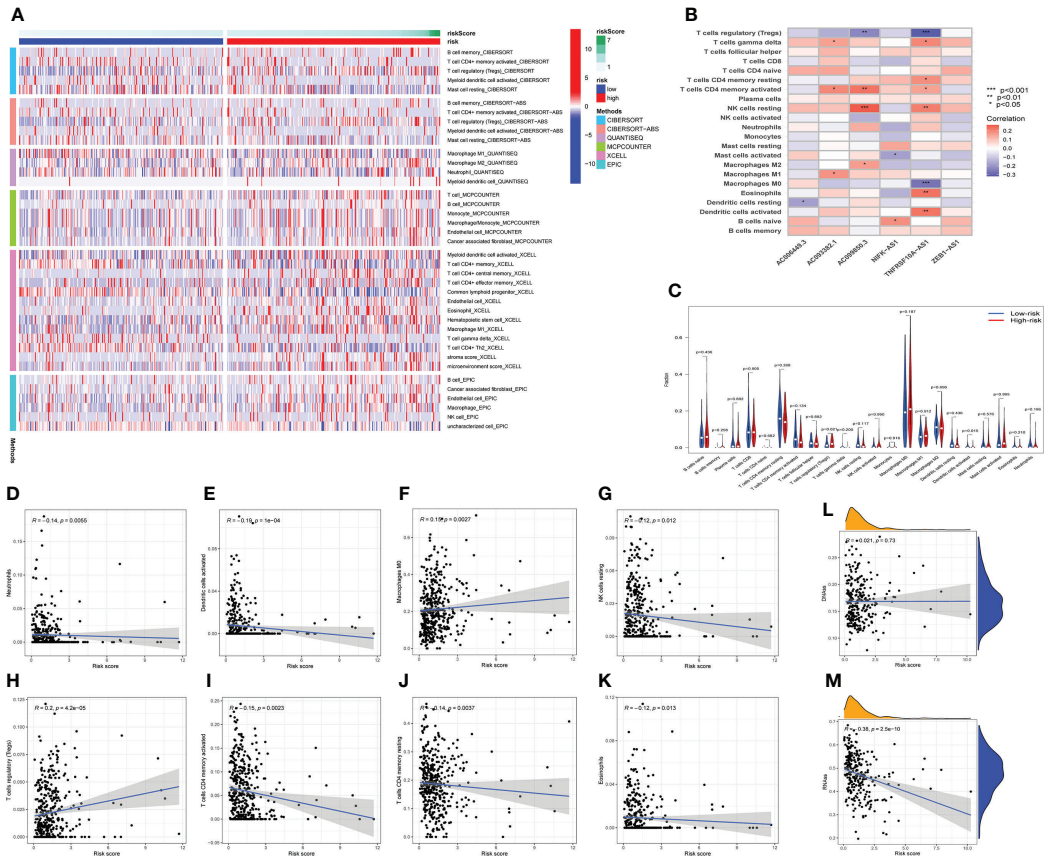


FIGURE 6
The immune landscape analysis. **(A)** The immune landscapes of high-risk and low-risk groups. **(B)** The correlation between CRLs and different immune cells, AC099850.3 had the highest positive correlation with resting NK cells, and TNFRSF10A-AS1 had the highest negative relation with T cells regulatory and macrophages M0. **(C)** The variations of immune cell infiltration levels between the high-risk and low-risk groups. **(D–K)** Analysis of the Pearson correlation between distinct immune cells and risk scores. **(L, M)** The correlation between DNA stem cell score (DNAss), RNA stem cell score (RNAss), and risk score. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

expression analysis and COX regression analysis were used to identify CRLs with prognostic significance. Subsequently, consensus cluster analysis showed that prognosis-related CRLs could divide patients with COAD into 3 clusters, which showed significant differences in prognosis and immune microenvironment. Following that, Lasso regression analysis was used to establish a prognostic signature with six CRLs. COAD patients could be separated into high-risk and low-risk groups according to median risk score, with the high-risk group having a much worse prognosis. This CRLs related signature gave a strategy of prognostic stratification for COAD patients. Finally, immune microenvironment, mutation and chemotherapeutic sensitivity analyses showed that this predictive signature could be used to provide evidence for immunotherapy and chemotherapy option.

Although screening and diagnosis of COAD have become more and more advanced, prognostic typing methods and sensitive genetic markers are still lacking (46, 50). Our study can provide CRLs-based prognostic model for patients with COAD, with an AUC value greater than 0.7 at 5 years OS, which can well classify patients into different prognostic groups, facilitate identification and early intervention of high-risk groups. There were also disparities in immune cell infiltration and TMB between high-risk and low-risk groups, providing some guidance for immunotherapy (51).

Our study identified six COAD prognostic markers correlated with cuproposis: TNFRSF10A-AS1, AC006449.3, AC093382.1, AC099850.3, ZEB1-AS1, and NIFK-AS1, which have been demonstrated to be associated with cancers in preliminary studies. First, Wei et al. discovered that TNFRSF10A-AS1 is a novel prognostic marker for colorectal cancer and may be related to autophagy (52). A regulatory network of lncRNA-miRNA-mRNA ceRNA was constructed for squamous cell carcinoma of tongue, and AC099850.3 was found to be strongly associated with the overall survival rate of patients (53). AC099850.3 has been confirmed to promote hepatocellular

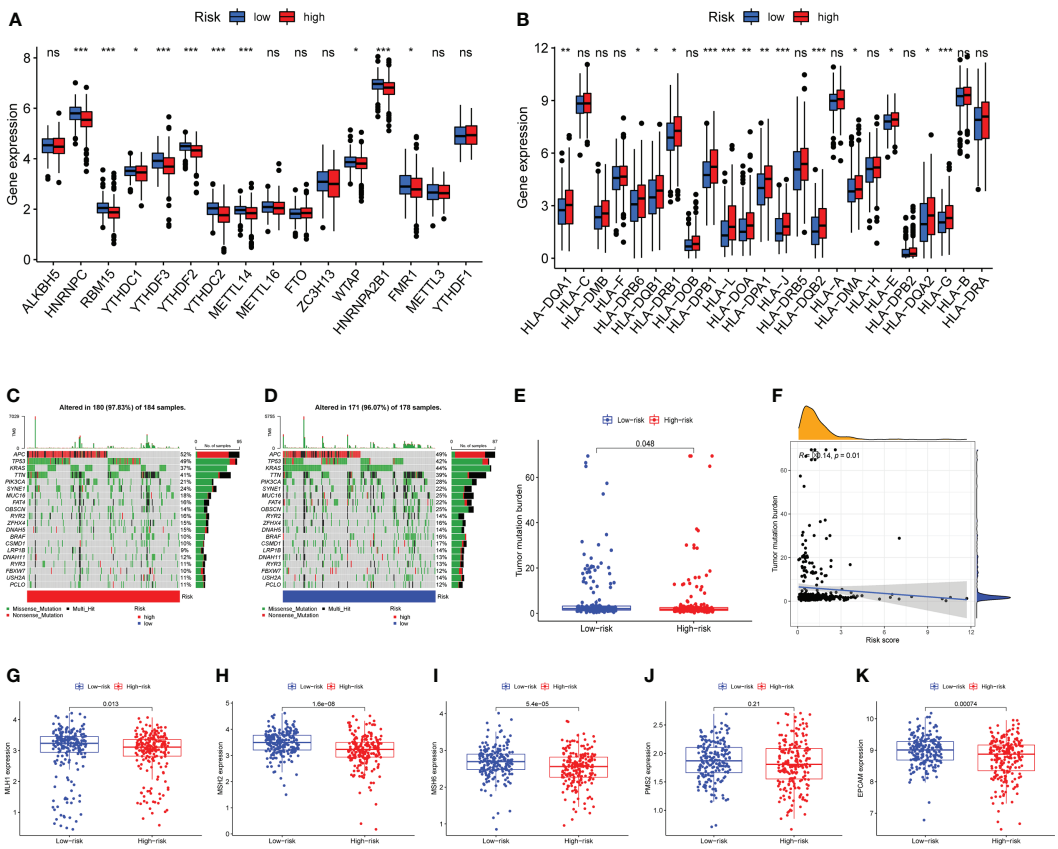


FIGURE 7
Expression analysis of m^{6}A-related genes and human leukocyte antigen (HLA) related genes, mutation analysis. **(A)** Differential expression of m^{6}A-related genes HNRNPC, RBM15, YTHDC1, YTHDF3, YTHDF2, YTHDC2, METTL14, WTAP, HNRNPA2B1, FMR1 were observed between the high-risk group and the low-risk groups. **(B)** The human leukocyte antigen gene analysis. HLA-DQA1, HLA-DRB6, HLA-DQB1, HLA-DRB1, HLA-DPB1, HLA-L, HLA-DOA, HLA-DPA1, HLA-J, HLA-DQB2, HLA-DMA, HLA-E, HLA-DQA2, and HLA-G were observed to be differentially expressed between high-risk and low-risk groups. **(C)** The mutation landscape of the high-risk group and draw the results into a waterfall diagram. **(D)**The waterfall diagram shows the mutation of patients in the low-risk group, and the mutation rate of APC was also highest. **(E)** Tumor mutation burden (TMB) analysis between the two risk groups. **(F)** Pearson-correlation analysis showed that there was a significant negative correlation between TMB and risk score. **(G–K)** Expression analysis of mismatch repair (MMR) protein. The expressions of MLH1, MSH2, MSH6 and EPCAM were significantly up-regulated in the low-risk group, while the expression of PMS2 showed no statistical difference between the two risk groups. ns, not significant,*P < 0.05, **P < 0.01, ***P < 0.001.

carcinoma (HCC) proliferation and invasion *via* the PRR11/PI3K/AKT axis and is a prognostic marker for HCC (54, 55). In addition, AC099850.3 was discovered as a predictive marker for non-small cell lung cancer (NSCLC) (56). Zinc finger E-box-binding homeobox 1 antisense 1 (ZEB1-AS1) facilitates the growth and metastasis of COAD cells, providing a new target for the diagnosis and treatment of COAD patients (57). Furthermore, ZEB1-AS1 can be used as one of the key lncRNAs in the construction of RCD-related prognostic signature (58). Consistent with previous studies, our study also included ZEB1-AS1 as the key lncRNA in RCD-related prognostic signature, which may reveal the important function of ZEB1-AS1 in RCD. Upregulation of NIFK-AS1 promote progression of HCC and Increased resistance to chemotherapy drugs through m6A methylation (59). Furthermore, NIFK-AS1 was discovered to suppress M2-like polarization of macrophages in endometrial cancer (60). However, there are currently few cancer research on AC006449.3 and AC093382.1, especially in COAD. In this study, we discovered a possible association between the six lncRNAs and cuproptosis, and offered evidences for their importance in the prognosis of COAD. Among six CRLs, TNFRSF10A-AS1 and AC099850.3 were protective factors while AC006449.3, AC093382.1, ZEB1-AS1 and NIFK-AS1 were adverse prognostic factors for COAD in this signature.

Immunotherapy, particularly immune checkpoint inhibitors, has been utilized to treat colorectal cancer in the past (61). However, “cold” tumors with low mutation rates and low microsatellite instability are not sensitive to immune checkpoint inhibitors (62, 63). As a result, it is critical to

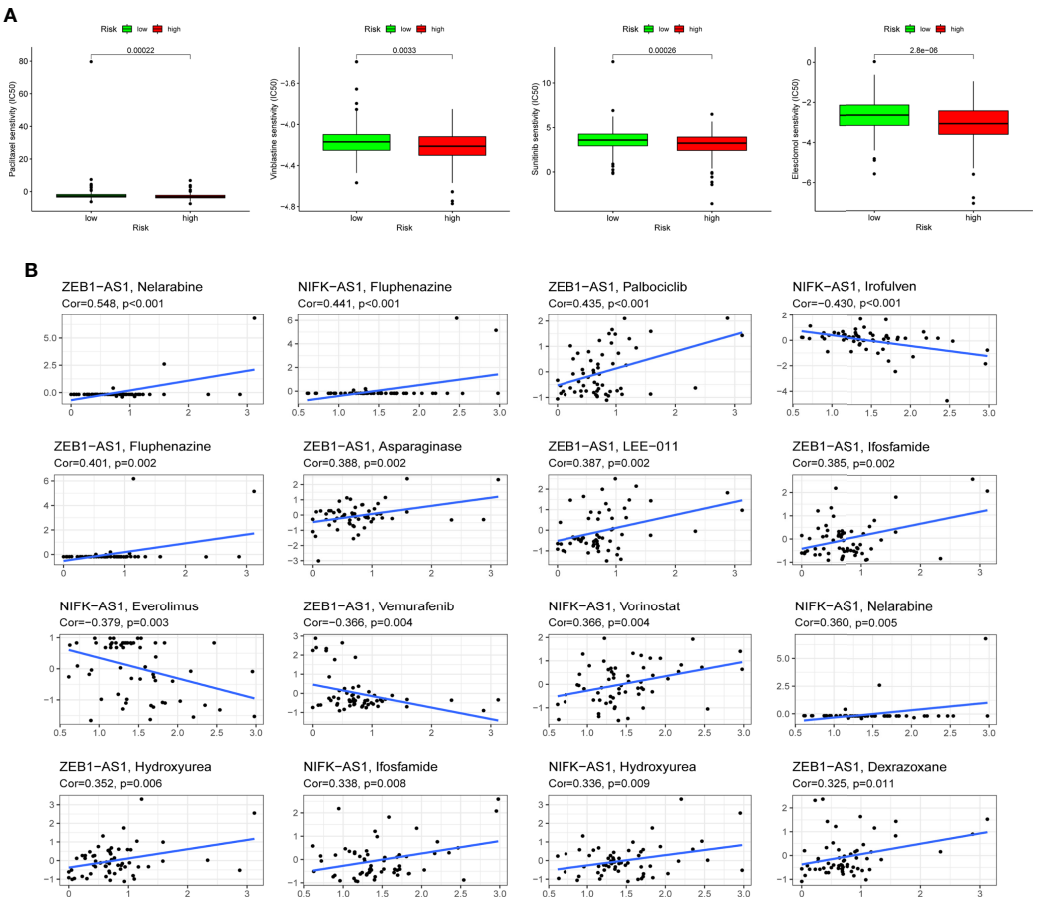


FIGURE 8
Drug sensitivity analysis. **(A)** Paclitaxel, Vinblastine, Sunitinib and Elesclomol were observed to have lower IC50 values in the high-risk group, that is, the high-risk group was more sensitive to these drugs. **(B)** Scatter plot of correlation between CRLs expression and drug sensitivity.

investigate the function of the predictive signature we developed in assessing mutation and expression of immune checkpoint-related genes in COAD. We discovered considerable disparities in tumor mutation burden (TMB), immune cell infiltration, HLA-related genes and mismatch repair proteins expression between the two risk groups based on signature constructed by cuproptosis-related lncRNAs, which might guide the immunotherapy for COAD patients. It also provided reference for understanding the potential association between tumor immunity and cuproptosis in colorectal cancer. Not only that, we identified more certain sensitive drugs for the COAD patients with higher risk score: Paclitaxel, Vinblastine, Sunitinib and Elesclomol, which was conducive to the early intervention and precision treatment for COAD.

Previous bioinformatics studies have revealed the role of other types of CRD in COAD (64, 65). Cuproptosis, novel types of cell death, has not been explored in COAD and our study is the first to highlight the function of cuproptosis-related lncRNAs. These findings help us understand the interaction of many regulatory cell death patterns, and provide a reference for precise treatment of COAD. However, there are some limitations in our study. Although the mechanism of copper inducing cell death has similar markers and characteristics of different forms of RCD, cuproptosis has not been confirmed in cell death nomenclature (66–68). The AUC value of our signature is not very high, which is not greater than 0.8, and it may be limited by the sample size. And we lack relevant functional experiments to verify the function of cuproptosis-related genes and CRLs in the model, which will be improved in the future.

Conclusions

Overall, our study is the first to develop a predictive signature based on the cuproptosis-associated lncRNA,

providing a novel approach to risk stratification and potential biomarkers for COAD patients. This signature is valuable for assessing prognosis, immune infiltration and chemotherapy sensitivity, which may help provide guidance for detections and treatments in patients with COAD.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repositories and accession numbers can be found in the article / Supplementary Material.

Author contributions

ZS and FW conceived the study and participated in the study design and performance. GC, and JL conducted the bioinformatics analysis and manuscript writing. CW, RG and MW revised the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

We would like to extend our gratitude to the researchers and study patients for their contributions.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1007918/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Consensus cluster analysis. **(A)** The cumulative distribution function (CDF) from *k* =2 to 9. **(B)** Relative change in area under CDF curve for *k* =2 to 9. **(C)** Tracking plot from *k* =2 to 9.

SUPPLEMENTARY FIGURE 2
The results of LASSO analysis. **(A)** The coefficient profile of 6 Cuproptosis-related lncRNAs. **(B)** 10-fold cross-validation of variable selection in LASSO models.

SUPPLEMENTARY FIGURE 3
The expression level of 6 Cuproptosis-related lncRNAs. **(A–F)** The expression level of TNFRSF10A-AS1, AC006449.3, AC093382.1, AC099850.3, ZEB1-AS1 and NIFK-AS1in colorectal cancer cell lines.

SUPPLEMENTARY FIGURE 4
Independent predictive factor analysis. Univariate and multivariate Cox regression analyses were used to evaluate whether risk score and clinical characteristics were independent predictors of COAD patients in three cohorts.

SUPPLEMENTARY FIGURE 5
Subgroup analysis of prognostic value of risk score. Survival analysis revealed the prognostic value of risk score in COAD patients with different ages **(A, B)**, different stages **(C, D)**, different genders **(E, F)** and different histological types **(G, H)**.

References

1. Thanikachalam K, Khan G. Colorectal cancer and nutrition. *Nutrients* (2019) 11(1):164. doi: 10.3390/nu11010164

2. Brody H. Colorectal cancer. *Nature* (2015) 521(7551):S1. doi: 10.1038/521S1a

3. Baidoun F, Elshiwy K, Elkeraie Y, Merjaneh Z, Khoudari G, Sarmini MT, et al. Colorectal cancer epidemiology: Recent trends and impact on outcomes. *Curr Drug Targets* (2021) 22(9):998–1009. doi: 10.2174/1389450121999201117115717

4. Du W, Frankel TL, Green M, Zou W. IFNγ signaling integrity in colorectal cancer immunity and immunotherapy. *Cell Mol Immunol* (2022) 19(1):23–32. doi: 10.1038/s41423-021-00735-3

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

6. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol* (2019) 16(12):713–32. doi: 10.1038/s41575-019-0189-8

7. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492

8. De Rosa M, Pace U, Rega D, Costabile V, Duraturo F, Izzo P, et al. Genetics, diagnosis and management of colorectal cancer (Review). *Oncol Rep* (2015) 34 (3):1087–96. doi: 10.3892/or.2015.4108

9. Huang Z, Yang M. Molecular network of colorectal cancer and current therapeutic options. *Front Oncol* (2022) 12:852927. doi: 10.3389/fonc.2022.852927

10. He X, Wu K, Ogino S, Giovannucci EL, Chan AT, Song M. Association between risk factors for colorectal cancer and risk of serrated polyps and conventional adenomas. *Gastroenterology* (2018) 155(2):355–373.e318. doi: 10.1053/j.gastro.2018.04.019

11. Sninsky JA, Shore BM, Lupu GV, Crockett SD. Risk factors for colorectal polyps and cancer. *Gastrointestinal endoscopy Clinics North America* (2022) 32 (2):195–213. doi: 10.1016/j.giec.2021.12.008

12. Lichtenstern CR, Ngu RK, Shalapour S, Karin M. Immunotherapy, inflammation and colorectal cancer. *Cells* (2020) 9(3):618. doi: 10.3390/cells9030618

13. Giannone G, Ghisoni E, Genta S, Scotto G, Tuninetti V, Turinetto M, et al. Immuno-metabolism and microenvironment in cancer: Key players for immunotherapy. *Int J Mol Sci* (2020) 21(12):4414. doi: 10.3390/ijms21124414

14. Fuchs Y, Steller H. Programmed cell death in animal development and disease. *Cell* (2011) 147(4):742–58. doi: 10.1016/j.cell.2011.10.033

15. Carneiro BA, El-Deiry WS. Targeting apoptosis in cancer therapy. *Nat Rev Clin Oncol* (2020) 17(7):395–417. doi: 10.1038/s41571-020-0341-y

16. Tang D, Kang R, Berghe TV, Vandenabeele P, Kroemer G. The molecular machinery of regulated cell death. *Cell Res* (2019) 29(5):347–64. doi: 10.1038/s41422-019-0164-5

17. Bedoui S, Herold MJ, Strasser A. Emerging connectivity of programmed cell death pathways and its physiological implications. *Nat Rev Mol Cell Biol* (2020) 21(11):678–95. doi: 10.1038/s41580-020-0270-8

18. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. *Sci (New York NY)* (2022) 375(6586):1254–61. doi: 10.1126/science.abf0529

19. Tang D, Chen X, Kroemer G. Cuproptosis: A copper-triggered modality of mitochondrial cell death. *Cell Res* (2022) 32(5):417–8. doi: 10.1038/s41422-022-00653-7

20. Mou Y, Wang J, Wu J, He D, Zhang C, Duan C, et al. Ferroptosis, a new form of cell death: Opportunities and challenges in cancer. *J Hematol Oncol* (2019) 12(1):34. doi: 10.1186/s13045-019-0720-y

21. Mishra AP, Salehi B, Sharifi-Rad M, Pezzani R, Kobarfard F, Sharifi-Rad J, et al. Programmed cell death, from a cancer perspective: An overview. *Mol diagnosis Ther* (2018) 22(3):281–95. doi: 10.1007/s40291-018-0329-9

22. Liu J, Geng R, Ni S, Cai L, Yang S, Shao F, et al. Pyroptosis-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with UCEC. *Mol Ther Nucleic Acids* (2022) 27:1036–55. doi: 10.1016/j.omtn.2022.01.018

23. Liu J, Wang Y, Meng H, Yin Y, Zhu H, Ni T. Identification of the prognostic signature associated with tumor immune microenvironment of uterine corpus endometrial carcinoma based on ferroptosis-related genes. *Front Cell Dev Biol* (2021) 9:735013. doi: 10.3389/fcell.2021.735013

24. Zhao Z, Liu H, Zhou X, Fang D, Ou X, Ye J, et al. Necroptosis-related lncRNAs: Predicting prognosis and the distinction between the cold and hot tumors in gastric cancer. *J Oncol* (2021) 2021:6718443. doi: 10.1155/2021/6718443

25. Moran VA, Perera RJ, Khalil AM. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. *Nucleic Acids Res* (2012) 40(14):6391–400. doi: 10.1093/nar/gks296

26. Qian Y, Shi L, Luo Z. Long non-coding RNAs in cancer: Implications for diagnosis, prognosis, and therapy. *Front Med* (2020) 7:612393. doi: 10.3389/fmed.2020.612393

27. Gao J, Chen X, Wei P, Wang Y, Li P, Shao K. Regulation of pyroptosis in cardiovascular pathologies: Role of noncoding RNAs. *Mol Ther Nucleic Acids* (2021) 25:220–36. doi: 10.1016/j.omtn.2021.05.016

28. Tsvetkov P, Detappe A, Cai K, Keys HR, Brune Z, Ying W, et al. *et al*: Mitochondrial metabolism promotes adaptation to proteotoxic stress. *Nat Chem Biol* (2019) 15(7):681–9. doi: 10.1038/s41589-019-0291-9

29. Liu J, Mei J, Wang Y, Chen X, Pan J, Tong L, et al. Development of a novel immune-related lncRNA signature as a prognostic classifier for endometrial carcinoma. *Int J Biol Sci* (2021) 17(2):448–59. doi: 10.7150/ijbs.51207

30. Tao C, Huang K, Shi J, Hu Q, Li K, Zhu X. Genomics and prognosis analysis of epithelial-mesenchymal transition in glioma. *Front Oncol* (2020) 10:183. doi: 10.3389/fonc.2020.00183

31. Ranstam J, Cook JA. Kaplan-Meier Curve. *Br J Surg* (2017) 104(4):442. doi: 10.1002/bjs.10238

32. Sun Z, Li T, Xiao C, Zou S, Zhang M, Zhang Q, et al. Prediction of overall survival based upon a new ferroptosis-related gene signature in patients with clear cell renal cell carcinoma. *World J Surg Oncol* (2022) 20(1):120. doi: 10.1186/s12957-022-02555-9

33. Hoshino N, Hida K, Sakai Y, Osada S, Idani H, Sato T, et al. Nomogram for predicting anastomotic leakage after low anterior resection for rectal cancer. *Int J colorectal Dis* (2018) 33(4):411–8. doi: 10.1007/s00384-018-2970-5

34. Hoo ZH, Candlish J, Teare D. What is an ROC curve? *Emergency Med J EMJ* (2017) 34(6):357–9. doi: 10.1136/emermed-2017-206735

35. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-p: a desktop application for gene set enrichment analysis. *Bioinf (Oxford England)* (2007) 23(23):3251–3. doi: 10.1093/bioinformatics/btm369

36. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337

37. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. *et al*: Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* (2018) 173(2):338–354.e315. doi: 10.1016/j.cell.2018.03.034

38. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118

39. Merino DM, McShane LM, Fabrizio D, Funari V, Chen SJ, White JR, et al. *et al*: Establishing guidelines to harmonize tumor mutational burden (TMB): in silico assessment of variation in TMB quantification across diagnostic platforms: phase I of the friends of cancer research TMB harmonization project. *J immunotherapy Cancer* (2020) 8(1):e000147. doi: 10.1136/jitc-2019-000147

40. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi: 10.1093/nar/gkv007

41. Sebaugh JL. Guidelines for accurate EC50/IC50 estimation. *Pharm Stat* (2011) 10(2):128–34. doi: 10.1002/pst.426

42. Hurwitz SN, Rider MA, Bundy JL, Liu X, Singh RK, Meckes DGJr. Proteomic profiling of NCI-60 extracellular vesicles uncovers common protein cargo and cancer type-specific biomarkers. *Oncotarget* (2016) 7(52):86999–7015. doi: 10.18632/oncotarget.13569

43. Hogg SJ, Beavis PA, Dawson MA, Johnstone RW. Targeting the epigenetic regulation of antitumour immunity. *Nat Rev Drug Discovery* (2020) 19(11):776–800. doi: 10.1038/s41573-020-0077-5

44. Kim J, Lee J, Oh JH, Sohn DK, Shin A, Kim J, et al. Dietary methyl donor nutrients, DNA mismatch repair polymorphisms, and risk of colorectal cancer based on microsatellite instability status. *Eur J Nutr* (2022) 61(6):3051–66. doi: 10.1007/s00394-022-02833-y

45. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet (London England)* (2019) 394(10207):1467–80. doi: 10.1016/S0140-6736(19)32319-0

46. Ladabaum U, Dominitz JA, Kahi C, Schoen RE. Strategies for colorectal cancer screening. *Gastroenterology* (2020) 158(2):418–32. doi: 10.1053/j.gastro.2019.06.043

47. Meester RG, Doubeni CA, Lansdorp-Vogelaar I, Goede SL, Levin TR, Quinn VP, et al. Colorectal cancer deaths attributable to nonuse of screening in the united states. *Ann Epidemiol* (2015) 25(3):208–213.e201. doi: 10.1016/j.annepidem.2014.11.011

48. Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: A review. *Jama* (2021) 325(7):669–85. doi: 10.1001/jama.2021.0106

49. Obeng E. Apoptosis (programmed cell death) and its signals - a review. *Braz J Biol* (2021) 81(4):1133–43. doi: 10.1590/1519-6984.228437

50. Issa IA, Noureddine M. Colorectal cancer screening: An updated review of the available options. *World J Gastroenterol* (2017) 23(28):5086–96. doi: 10.3748/wjg.v23.i28.5086

51. Fancello L, Gandini S, Pelicci PG, Mazzarella L. Tumor mutational burden quantification from targeted gene panels: Major advancements and challenges. *J immunotherapy Cancer* (2019) 7(1):183. doi: 10.1186/s40425-019-0647-4

52. Wei J, Ge X, Tang Y, Qian Y, Lu W, Jiang K, et al. *et al*: An autophagy-related long noncoding RNA signature contributes to poor prognosis in colorectal cancer. *J Oncol* (2020) 2020:4728947. doi: 10.1155/2020/4728947

53. Zhou RS, Zhang EX, Sun QF, Ye ZJ, Liu JW, Zhou DH, et al. Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of tongue. *BMC Cancer* (2019) 19(1):779. doi: 10.1186/s12885-019-5983-8

54. Zhong F, Liu S, Hu D, Chen L. LncRNA AC099850.3 promotes hepatocellular carcinoma proliferation and invasion through PRR11/PI3K/AKT axis and is associated with patients prognosis. *J Cancer* (2022) 13(3):1048–60. doi: 10.7150/jca.66092

55. Wang Y, Ge F, Sharma A, Rudan O, Setiawan MF, Gonzalez-Carmona MA, et al. Immunoautophagy-related long noncoding RNA (IAR-lncRNA) signature predicts survival in hepatocellular carcinoma. *Biology* (2021) 10(12):1301. doi: 10.3390/biology10121301

56. Zhou J, Zhang M, Dong H, Wang M, Cheng Y, Wang S, et al. Comprehensive analysis of acetylation-related lncRNAs and identified AC099850.3 as prognostic biomarker in non-small cell lung cancer. *J Oncol* (2021) 2021:4405697. doi: 10.1155/2021/4405697

57. Ni X, Ding Y, Yuan H, Shao J, Yan Y, Guo R, et al. Long non-coding RNA ZEB1-AS1 promotes colon adenocarcinoma malignant progression *via* miR-455-3p/PAK2 axis. *Cell proliferation* (2020) 53(1):e12723. doi: 10.1111/cpr.12723

58. Chen W, Chen Y, Liu L, Wu Y, Fu P, Cao Y, et al. Comprehensive analysis of immune infiltrates of ferroptosis-related long noncoding RNA and prediction of colon cancer patient prognoses. *J Immunol Res* (2022) 2022:9480628. doi: 10.1155/2022/9480628

59. Chen YT, Xiang D, Zhao XY, Chu XY. Upregulation of lncRNA NIFK-AS1 in hepatocellular carcinoma by m(6)A methylation promotes disease progression and sorafenib resistance. *Hum Cell* (2021) 34(6):1800–11. doi: 10.1007/s13577-021-00587-z

60. Zhou YX, Zhao W, Mao LW, Wang YL, Xia LQ, Cao M, et al. Long non-coding RNA NIFK-AS1 inhibits M2 polarization of macrophages in endometrial cancer through targeting miR-146a. *Int J Biochem Cell Biol* (2018) 104:25–33. doi: 10.1016/j.biocel.2018.08.017

61. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol* (2019) 16(6):361–75. doi: 10.1038/s41575-019-0126-x

62. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discovery* (2019) 18 (3):197–218. doi: 10.1038/s41573-018-0007-y

63. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. *et al*: Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Sci (New York NY)* (2017) 357(6349):409–13. doi: 10.1126/science.aan6733

64. Song W, Ren J, Xiang R, Kong C, Fu T. Identification of pyroptosis-related subtypes, the development of a prognosis model, and characterization of tumor microenvironment infiltration in colorectal cancer. *Oncoimmunology* (2021) 10 (1):1987636. doi: 10.1080/2162402X.2021.1987636

65. Shao Y, Jia H, Huang L, Li S, Wang C, Aikemu B, et al. *et al*: An original ferroptosis-related gene signature effectively predicts the prognosis and clinical status for colorectal cancer patients. *Front Oncol* (2021) 11:711776. doi: 10.3389/fonc.2021.711776

66. Li Y. Copper homeostasis: Emerging target for cancer treatment. *IUBMB Life* (2020) 72(9):1900–8. doi: 10.1002/iub.2341

67. Shanbhag VC, Gudekar N, Jasmer K, Papageorgiou C, Singh K, Petris MJ. Copper metabolism as a unique vulnerability in cancer. *Biochim Biophys Acta Mol Cell Res* (2021) 1868(2):118893. doi: 10.1016/j.bbamcr.2020.118893

68. Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, et al. Molecular mechanisms of cell death: Recommendations of the nomenclature committee on cell death 2018. *Cell Death differentiation* (2018) 25(3):486–541. doi: 10.1038/s41418-017-0012-4

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fonc.2022.1008770

OPEN ACCESS

EDITED BY
Fu Wang,
Xi'an Jiaotong University, China

REVIEWED BY
Xiulin Jiang,
Kunming Institute of Zoology (CAS), China
Yaoming Li,
Second Xiangya Hospital, Central South University, China

*CORRESPONDENCE
Fei Li
Taishanlongtanfei@163.com

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

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 01 August 2022
ACCEPTED 24 August 2022
PUBLISHED 26 September 2022

CITATION
Yin Y, Du W and Li F (2022) The construction of a hypoxia-based signature identified CA12 as a risk gene affecting uveal melanoma cell malignant phenotypes and immune checkpoint expression.
Front. Oncol. 12:1008770.
doi: 10.3389/fonc.2022.1008770

The construction of a hypoxia-based signature identified CA12 as a risk gene affecting uveal melanoma cell malignant phenotypes and immune checkpoint expression

Yan Yin[1†], Wei Du[2†] and Fei Li[3*]

[1]Department of Ophthalmology, The Second Affiliated Hospital of Shandong First Medical University, Taian, China, [2]Department of Ophthalmology, The Shandong Second Rehabilitation Hospital, Taian, China, [3]Department of Medicine, Shandong First Medical University, Taian, China

Uveal melanoma (UM) is a deadly intraocular neoplasm in the adult population and harbors limited therapeutic effects from the current treatment. Here, we aimed to investigate the role of hypoxia in UM progress. We adopted the Cancer Genome Atlas data set as a training cohort and Gene Expression Omnibus data sets as validating cohorts. We first used consensus clustering to identify hypoxia-related subtypes, and the C1 subtype predicted an unfavorable prognosis and exhibited high infiltration of immunocytes and globally elevated immune checkpoint expression. Besides this, the patients with the C1 subtype were predicted to respond to the PD-1 treatment. By the least absolute shrinkage and selection operator algorithm, we constructed a hypoxia risk score based on the hypoxia genes and identified 10 genes. The risk score predicted patient survival with high performance, and the high-risk group also harbored high immunocyte infiltration and immune checkpoint expression. Furthermore, we confirmed that the risk genes were upregulated under hypoxia, and knockdown of CA12 inhibited the epithelial–mesenchymal transition process, clone formation ability, and G1/S phase transformation of the UM cells. The CD276 was also downregulated when CA12 knockdown was performed. These results validate the prognostic role of the hypoxia signature in UM and demonstrate that CA12 is a critical factor for UM cell progression as well as a target to improve immunotherapeutic effects. We believe our study contributes to the understanding of hypoxia's roles in UM and provides a novel target that will benefit future therapeutic strategy development.

KEYWORDS

uveal melanoma, hypoxia, immune, carbonic anhydrase 12, risk score

Introduction

Uveal melanoma (UM) is a rare tumor type among the population, but it is the most common intraocular neoplasm in the adult population with high malignancy. UM harbors a morbidity of approximately five cases per million per year, and 90% of UM originates from the choroid (1). Nearly 50% of UM metastases to the liver during tumor development, and early stage intervention, such as chemoembolization and surgical excision, controls tumor progression, but mostly prognoses of UM returns unfavorable due to the limited therapeutic strategy effects. More efficient approaches for improving the therapeutic effects or prognostic management are urgently required.

Molecular pathogenesis and targeted therapy have been novel research topics and promising strategies to prevent UM processes or improve patient survival (2). Currently, many molecular features have been applied to indicate the patient prognostic diversity, such as that active mutation of the Gα11/Q pathway drives the tumorigenesis of UM and *BAP1, SF3B1*, and *EIF1AX* mutant precited metastatic progression (3). Moreover, various studies are still carried on to discover novel molecular targets for UM (4–6). Many targeted therapy–based clinical trials have been conducted, whereas no mature approach has been proven for application, indicating the urgent need for more effective strategies for UM treatment (7).

Hypoxia, characterized by insufficient tissue oxygenation, is a critical risk factor in cancer, for its connection with various hallmarks of cancers, including angiogenesis, metabolism programming (8), and immunosuppression (9), contributing to the progression of cancer and poor prognosis. To overcome the hypoxia-related signaling in cancers, many hypoxia-targeted therapies were developed (10). In UM, hypoxia has been suggested to correlate with angiogenesis, invasion, and autophagy, indicating its significant role in UM. Besides this, many drugs were discovered to gain therapeutic effects, mostly based on HIF and angiogenesis (11). Interestingly, a study has demonstrated the effects of hypoxia stress on monocyte migration and characteristics (12); this implies the association between hypoxia and immunity of UM, whereas their intercorrelation is far from understood.

Abbreviations: UM, Uveal melanoma; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; TCGA, The Cancer Genomes Atlas; GEO, Gene Expression Omnibus; TIDE, Tumor Immune Dysfunction and Exclusion; GSVA, Gene Set Variation Analysis; PCA, Principal component analysis; GSEA, Gene Set Enrichment Analysis; DEGs, Differentially expressed genes; BP, biological process;, CC, cellular component;, and MF, molecular function; LASSO, Least absolute shrinkage and selection operator; GDSC, Genomics of Drug Sensitivity in Cancer; PI, propidium iodide; ROC, Receiver operating characteristics curve; MDSC, Myeloid-derived suppressor cells; ECM-associated, Extracellular matrix; CA12, Carbonic Anhydrase 12.

In this study, we used a hypoxia-based signature to identify novel hypoxia-based subtypes and build a prediction model for patient prognosis, and the biological involvement of hypoxia in UM was investigated by functional enrichment, and its correlation with immunocyte, immune checkpoints were explored. Moreover, we conducted a series of experiments to validate hypoxia's effects on tumor cell phenotype and immune checkpoint expression. We hope this study will reveal a novel pathological mechanism of hypoxia in UM and provide alternative therapeutic targets for UM patient treatment.

Materials and methods

Data acquisition

The UM sample RNA expression and clinical information were obtained from the Cancer Genomes Atlas and were used as a training cohort, and UM samples from the Gene Expression Omnibus, GSE22138 and GSE84976, were downloaded as validating cohorts. The hypoxia, gene ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets were retrieved from the Gene Set Enrichment Analysis (GSEA) online database. The compounds used for potential drug identification were obtained from the GDSC database. The tumor immune dysfunction and exclusion (TIDE)-related calculation was performed on the TIDE online web tool.

Consensus clustering of UM samples by the hypoxia gene sets

We used the hypoxia gene sets to cluster the UM training cohort by the "ConsensusCluster" R package with the best *k* value and visualized the results by a principal component analysis (PCA) plot. For the clusters obtained, we used survival analysis to evaluate their prognostic difference. In a heat map, an expression of the hypoxia genes in all samples divided by the clusters were also presented. Subsequently, we investigate the diversity of cancer hallmarks between the clusters using gene set variation analysis (13) and "Hallmark" gene sets downloaded from GSEA.

Immune diversity between clusters

To investigate the immune diversity between the clusters of UM, we analyzed the 28 types of immunocyte infiltration of all samples using single sample GSEA (ssGSEA). We also compared the immune checkpoint expression differences between clusters. The results were presented in heat maps and box plots.

Drug IC50 calculation and immunotherapy analyses

To identify novel drugs for hypoxia cluster–based targets, we downloaded the compound information, cell line expression matrix, and cell line testing results to predict the drug IC50 for all samples by the R package "pRRophetic" (14), and the IC50 values were compared between clusters in box plots. Besides this, we analyzed the immunotherapeutic effects of the samples by TIDE analyses, including the dysfunction, exclusion, IFNG, and TIDE score calculation conducted on the TIDE web tool. We also analyzed the correlation between PD-1, CTLA4 response, and clusters to explore the immunotherapeutic potential of the clusters.

Biological diversity comparison between clusters

To compare the biological diversity between samples in different clusters, we first used the "limma" R package to filter the differentially expressed genes (DEGs) between clusters. Then, the DEGs were annotated according to the gene set annotation downloaded from GSEA, including KEGG and biological process (BP), cellular component (CC), and molecular function (MF) in GO.

Hypoxia Least absolute shrinkage and selection operator (LASSO) risk score construction

The hypoxia gene sets were applied to the LASSO regression (15) to reduce the number of parameters and construct a prognostic model. The risk score of each sample was calculated as the sum of the coefficient multiplied by the expression of each gene. We divided the samples into high- and low-risk groups according to the median risk score of the cohort. The survival time and risk score gene expression were presented in order of the risk scores. We checked the prognostic value of the risk score by survival analysis and receiver operating characteristics curve (ROC) and, finally, built a nomogram integrating risk score, age, gender, and stage, and its prognostic value was estimated by ROC and calibration curve.

The immune diversity between risk scores and potential drug identification

To discover the immune diversity between the risk groups, we used ssGSEA to analyze the immunocyte infiltration differences as well as the immune checkpoint expression variation (16), the results were presented by box plots. The correlation between CA12 and CD276 was quantified on GEPIA2.0, an online tool for cancer investigation. Besides this, we compared the Genomics of Drug Sensitivity in Cancer (GDSC) (17) compound IC50 values of the two groups, and the cMap was also utilized to filter potential drugs.

Cell culture and siRNA transfection

The highly aggressive MUM2B cells were cultivated in DMEM with 10% FBS, maintained under 5% CO_2 and 37°C and digested when they were 80% confluent. For the hypoxia culture, we cultivated the cells in the hypoxia incubator for 24 h. Then, the cells were planted into a 12-well plate and transfected with CA12 siRNAs and lip3000 transfection reagent. After 48 h transfection, the cells were harvested and counted for further experiments. The sequence of siRNA1 and siRNA2 was provided by Zhao et al. and Huang et al, respectively (18, 19).

RT-qPCR detection of the mRNA levels

The cells were washed with PBS and lyzed in trizol for 10 min. The RNA was collected and extracted using chloroform. After centrifugation and supernatant collection, the RNA was precipitated with isopropanol, followed by sequential washing with 80% ethanol and absolute ethanol. When the RNA was dried naturally, a quantification by a microplate reader was then conducted. Subsequently, the genomic DNA was removed, and RNA reverse transcription was performed. Finally, the 10 μl mixed system per well containing 1μl cDNA, 0.4μl primers, 5μl SYBP, and 3.2μl RNA-free water was prepared, and the RT-qPCR was performed.

Transwell migration assay

The harvested cells were resuspended with 200 μl FBS (1%) and planted into the upper chamber of the transwell with 1×10^4 cells per well (Corning Incorporated, Corning, NY, USA). The lower chamber was filled with 500 μl DMEM with FBS (20%). After cultivation for 24 h, the upper chamber was slightly washed with PBS three times and fixed with crystal violet. After 30 min, the unmigrated cells were wiped off using a cotton swab and left to dry for microscopy.

CCK8 assay

The harvested cells were resuspended using DMEM with FBS (10%) and seeded in a 96-well plate for the CCK8 assay. After cell adherence, the previous medium was replaced by DMEM with a 10% CCK8 reagent (GK3607-500T, GeneView,

DingGuo Company, Changsha, China) containing no FBS. After 2 h of cultivation, the OD value was detected using the microplate reader. The absorbance was measured at 450 nm.

Clone formation assay

As described, we used DMEM containing FBS (10%) to resuspend the cells and seeded them in the six-well plate, with 1000 cells per well. After 14 days of cultivation, the medium was removed and washed with PBS. Subsequently, the cells were fixed with crystal violet for 30 min.

Flow cytometric analysis

The digested cells were collected in EP tubes. We centrifuged the cells to remove the medium and washed the cells with precooled PBS. Then, the cells were fixed using 70% ethanol under 4°C overnight. Subsequently, the ethanol was removed and washed by precooled PBS again and stained with propidium iodide (PI, 20X), RNase (50X), and staining buffer (C1052, Beyotim Biotechnology Co. Ltd,Nantong, China). The mixed system was incubated from light under 37°C for 30 min and detected by the flow cytometer. The data were analyzed by FlowJ software.

Western blot detection

The cells were lysed using the RIPA lysis buffer and centrifuged to obtain the protein supernatants. The nuclear in the samples were then further lysed by ultrasound. After being mixed with loading buffer, the samples were loaded, and electrophoresis and membrane transformation were sequentially performed. The membrane was then blocked with skim milk powder (5%) and incubated under 4°C overnight with the primary antibodies (sources listed in Supplementary Table S1). The next day, the membrane was washed with TBST three times and incubated with the second antibodies for 1 h. Finally, the protein bands on the membrane were detected using the chemiluminescence detection system after TBST washing. The results were quantified by the ImageJ software.

Statistical analyses

The bioinformatic analyses were performed on the R software. The Kaplan–Meier curve and log-rank test were used for survival analyses. Cox regression calculated the HR of each factor. ROC and the calibration curve estimated the predictive discrimination and calibration, respectively. Student's *T*-test compared the expression differences between groups. Two-way ANOVA tested the CCK8 results. Spearman's correlation coefficient was used to quantify expression correlation between genes. $P < .05$ was defined as statistically significant.

Results

Unsupervised clustering identified a cluster with a worse prognosis and upregulated hallmark features

To identify the hypoxia-associated clusters in UM. We used "ConsensusCluster" to conduct the unsupervised clustering of the training cohort. We selected the *k* value as 2, the samples were perfectly divided into two clusters (Figure 1A), and the PCA plot exhibited that the two clusters were separated into two groups (Figure 1B).

We then performed survival analysis to test the prognostic differences between the two clusters. As a result, cluster C1 exhibited a worse survival rate (Figure 1C). Moreover, most of the hypoxia-related genes were highly expressed in cluster C1 (Figure 1D). The ssGSEA results of the "Hallmark" gene sets of all samples showed that most of the cancer hallmark pathways were upregulated in cluster C1. Notably, many immune-related pathways were enriched in cluster C1, including IL6-JAK-STAT3, IL2-STAT5, TGF-β, interferon-α/γ-response, and TNFA-related signaling pathways (Figure 1E). These results indicated that hypoxia played a critical role in UM development, and these were associated with cancer immunity.

Hypoxia-divided clusters presented a diverse immune status

To investigate the immunological diversity between the clusters, we ran ssGSEA to analyze the immunocyte infiltration levels of each sample, and most of the immunocytes were highly infiltrated in cluster C1 (Figure 2A), including several immunosuppressive cells, such as regulatory T cells and myeloid-derived suppressor cells (MDSC). Further, we compared the expression of the immune checkpoint between clusters C1 and C2. Surprisingly, most of the immune checkpoints were upregulated in cluster C1 (Figure 2B). These results were quantified in box plots (Figures 2C, D) and demonstrated the immunosuppressive environment in high-hypoxia UM samples. Hence, we then conducted TIDE analyses. The results presented that cluster C1 exhibited a lower TIDE score, indicating that it may respond to immunotherapy (Figure 3A). We further analyzed the expression similarity between the training cohort and the previous immunotherapy cohort, and we noticed that cluster C1 samples showed similar expression signatures with the PD1-therapy response cohort though the *p*-value increased to

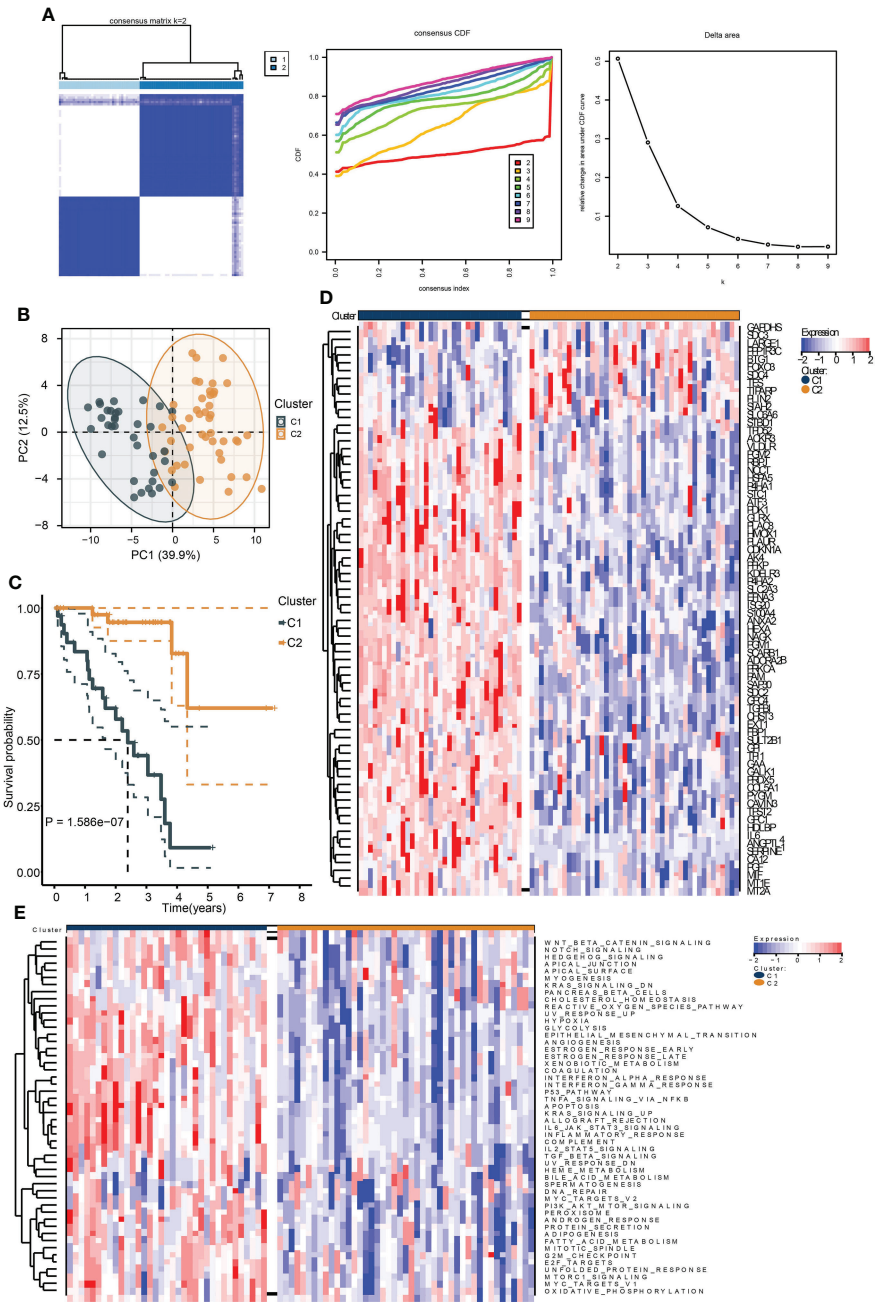


FIGURE 1
Identification of hypoxia-related subtypes of UM. **(A)** Consensus clustering heat map (left), CDF (middle), and the relative change in area under CDF curve (right) of the TCGA UM samples. **(B)** PCA plot showing the division results of consensus clustering. **(C)** Kaplan–Meier curve indicates the ability of subtypes to separate patient survival rate. **(D)** The expression heat map of the genes in hypoxia signature. **(E)** The enrichment heat map of the "hallmark" gene sets based on the DEGs between the subtypes.

0.055 after Bonferroni correction (Figure 3B). Additionally, we also sought possible chemotherapeutic drugs for UM patients. The GDSC drug IC50 was predicted for each sample, and we obtained two drugs, Methotrexate and Mitomycin C, with lower IC50 in cluster C1 (Figure 3C). The immunotherapeutic analyses indicated that UM cancer with high hypoxia harbored immunosuppression and may benefit from immunotherapy.

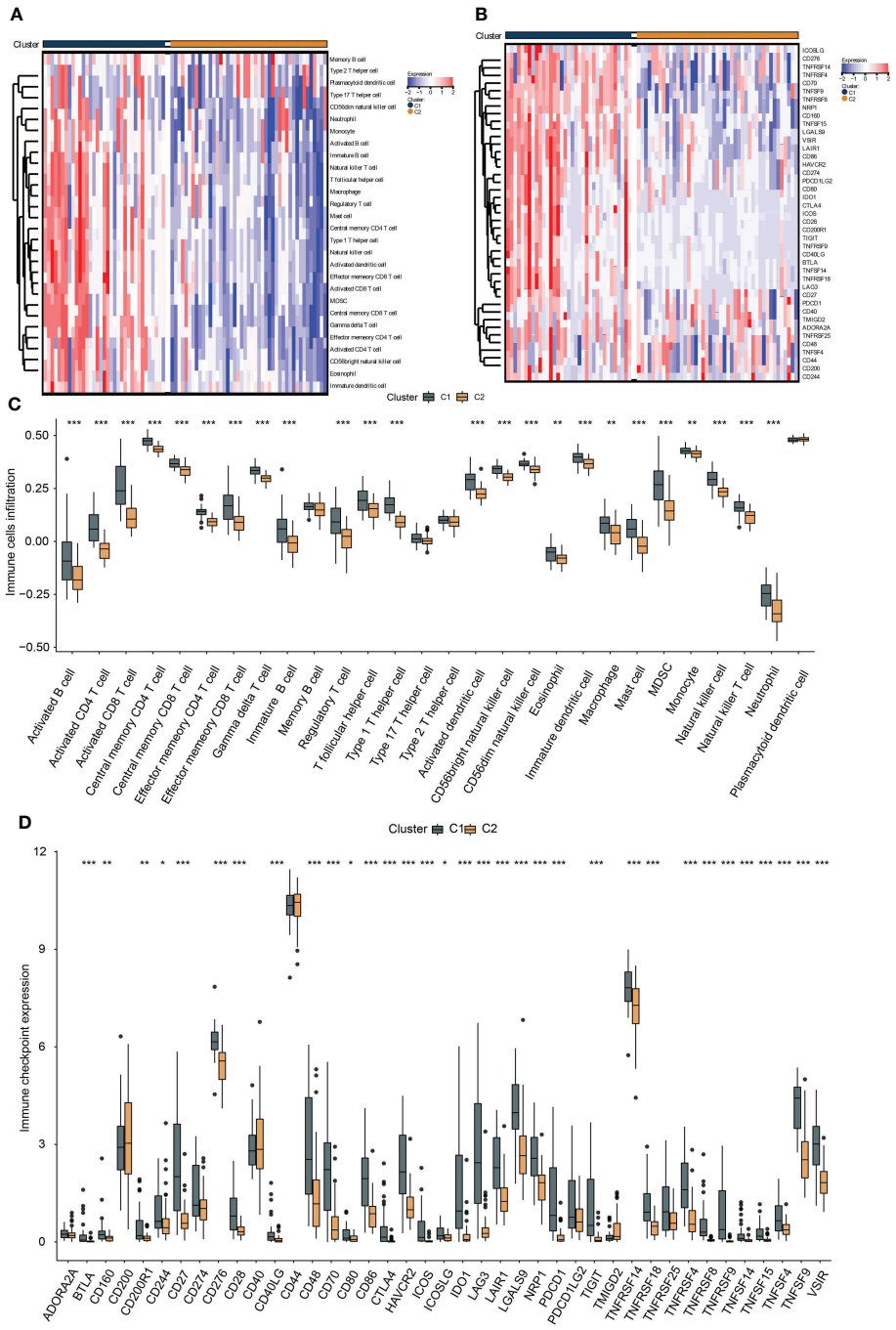


FIGURE 2
The association between hypoxia and immunocyte and immune checkpoints. The heat map presents the immunocyte infiltration difference **(A)** and the expression variation of the immune checkpoints **(B)** between subtypes, and their quantification results of immunocyte infiltration **(C)** and immune checkpoint expression **(D)** were exhibited in the box plots, *, **, *** represents p-value < 0.05, 0.01, and 0.001, respectively.

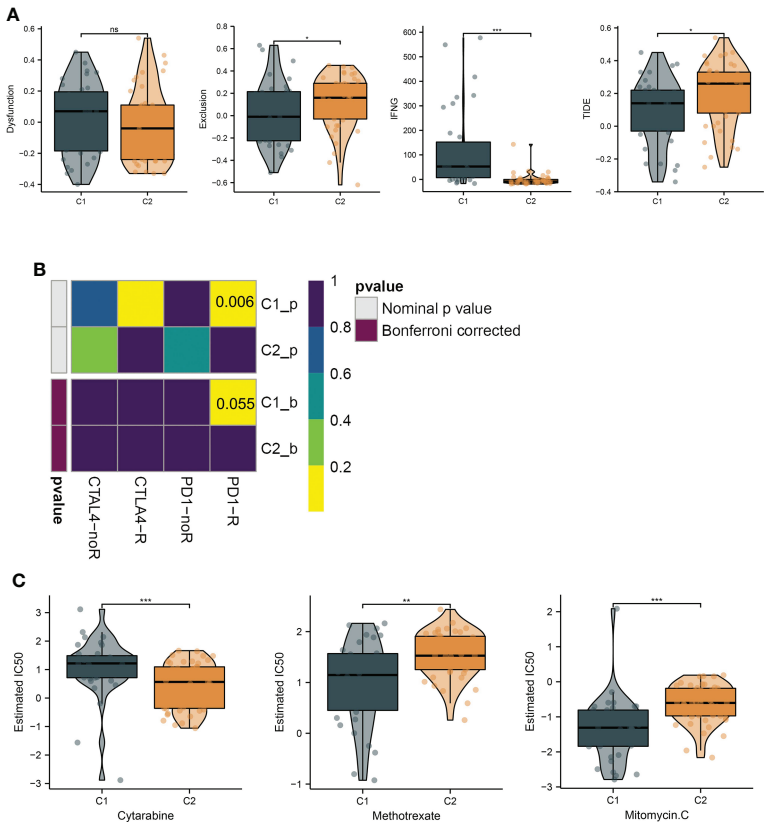


FIGURE 3
Hypoxia subtypes correlate with immunosuppression and their chemotherapeutic drugs development. **(A)** The TIDE estimation showing the dysfunction, exclusion, IFNG, and TIDE scores of the UM samples. **(B)** The submap of the expression similarity compared with immunotherapy response/nonresponse cohort. **(C)** Box plots present the IC50 differences of GDSC drugs between the subtypes. *, **, ***, and ns represents p-value<0.05, 0.01, 0.001, and not significant, respectively.

The clusters mainly differed in immune- and Extracellular matrix (ECM)-associated biological activities

To investigate the biological activity differences between clusters, we first filter the DEGs between them, and many upregulated genes were identified in cluster C1 (Figures 4A, B). To annotate their function, we performed functional enrichment analyses by the KEGG and GO gene sets. The results exhibited that their DEGs were mainly enriched in immune-related pathways. For instance, the cytokine–cytokine receptor interaction, chemokine signaling pathways in KEGG, cellular response to cytokine in GOBP, MHC protein complex in GOCC, and GO MHC class II in GOMF were identified, similar to the immune analyses results. Besides this, ECM-associated pathways, such as the ECM-receptor interaction in KEGG, extracellular region in GOCC, and extracellular matrix structural constituent in GOMF, were also enriched (Figures 4C–F), demonstrating the close correlation between hypoxia and the microenvironment.

Establishment of a hypoxia-based risk score

To identify the critical genes in the effects of hypoxia, we used the hypoxia gene set to establish a risk score. The LASSO algorithm reduced the gene number to 10, and five risky and five protective genes were finally obtained with their coefficients, respectively (Figures 5A, B). The patients were divided into the high- and low-risk groups according to the median risk score of the training cohort. The PCA plot showed that the risk groups separated obviously (Figure 5C). The risk-survival plot exhibited that the five risk genes' expression elevated as the risk score increased, and the opposite trend was observed for the five protective genes (Figure 5D). For the

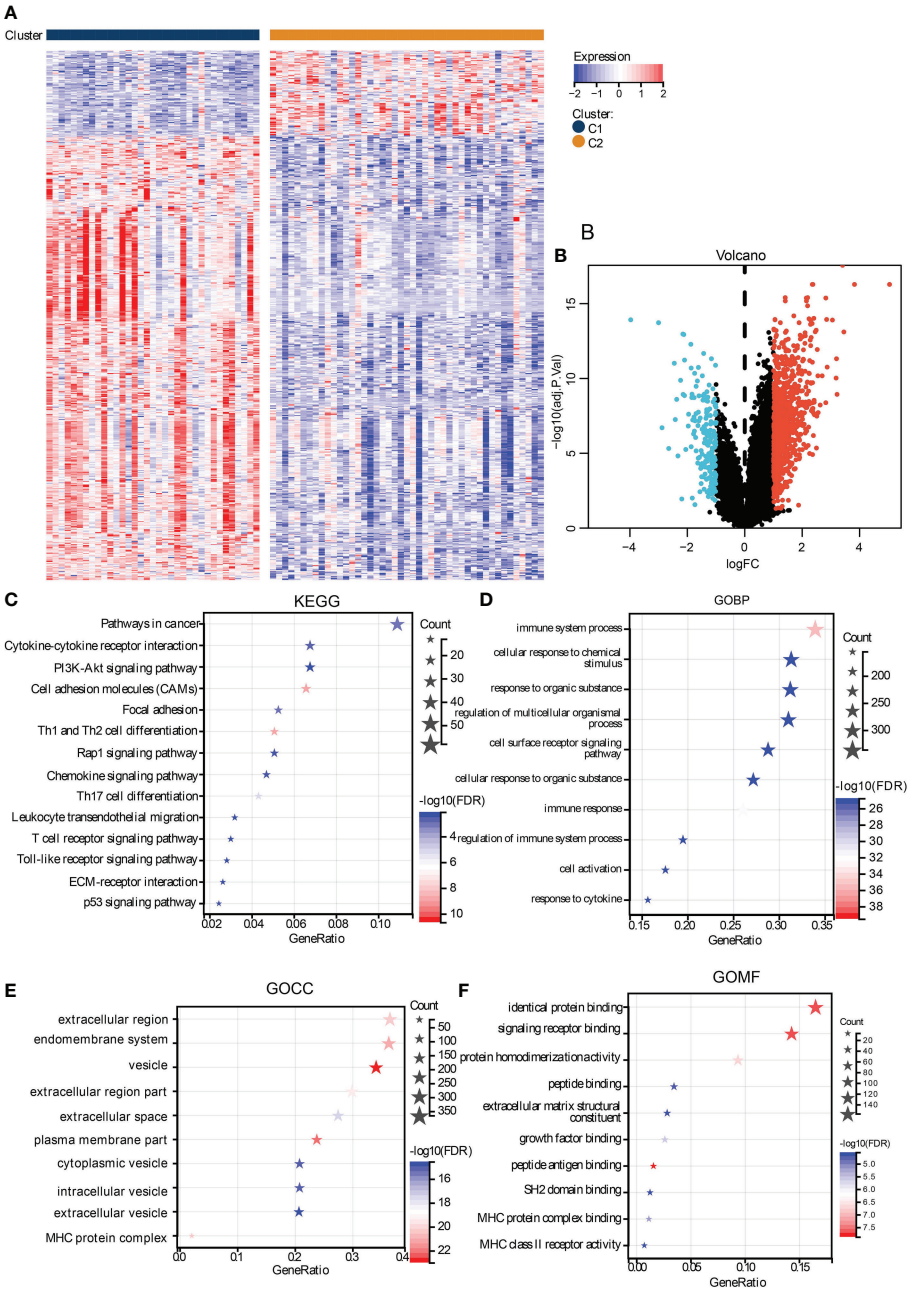


FIGURE 4
Functional analysis of DEGs between subtypes. The heat map **(A)** and volcano plot **(B)** shows the DEGs between subtypes. The functional analysis results were presented in bubble plots presenting the enriched KEGG **(C)**, GOBP **(D)**, GOCC **(E)**, and GOMF **(F)**, gene sets.

prognostic value, the high-risk group patients suffered significantly lower survival rates, and the ROC results demonstrated the risk score predicted patient overall survival with a high accuracy (Figures 5E, F). When validated in the validating cohorts, the same expression trend of the 10 genes was observed, and the risk score can predict the patient survival with high accuracy (Supplementary Figures S1A–S1F). We performed univariate and multivariate Cox analysis to select the prognostic clinical predictors apart from the risk score and age, and risk score passed the univariate test though it failed in the multivariate test (Figures 6A,B). When integrating the risk scores and age to form a nomogram, the nomogram also presented good performance in predicting patient survival (Figures 6C–E).

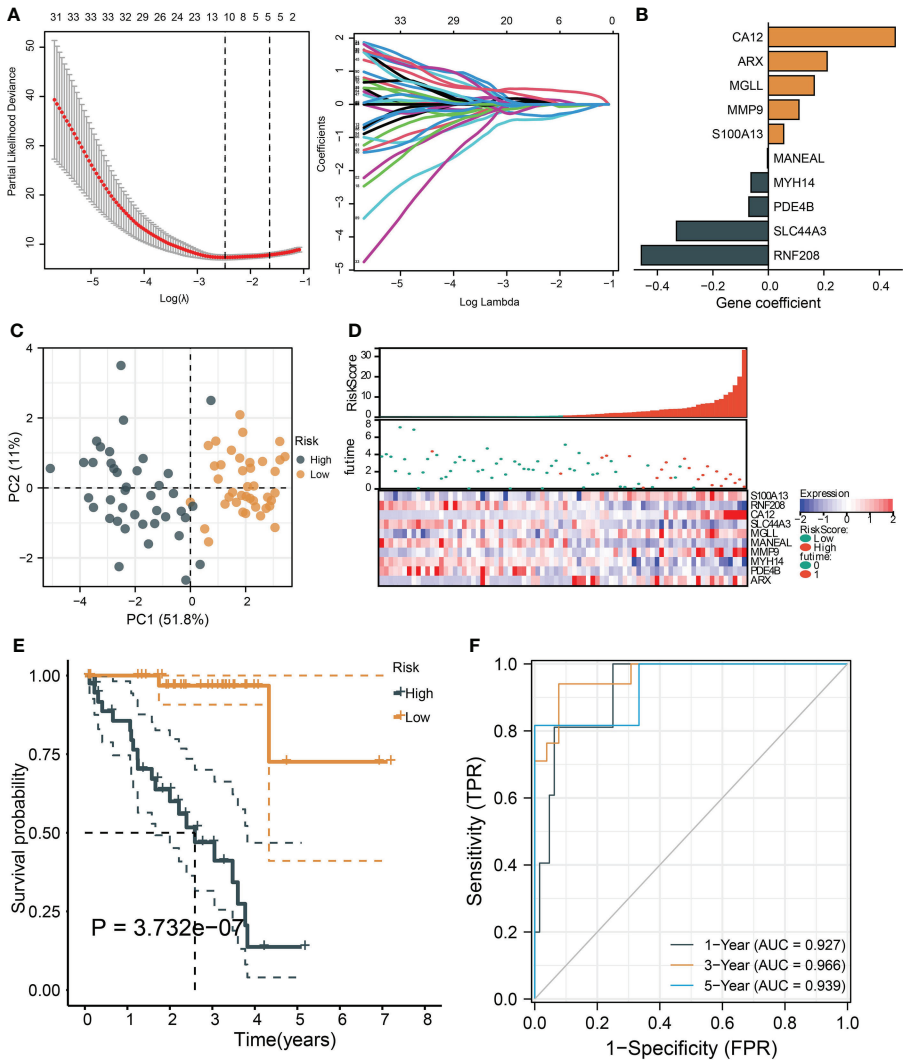


FIGURE 5
|Construction and validation of a hypoxia-based risk score. **(A)** The LASSO partial likelihood deviance plot and the coefficient profiles of all hypoxia gene. **(B)** The coefficients of the finally selected factors were presented in a bar chart. **(C)** The PCA plot shows the distance between the high- and low-risk group samples. **(D)** The risk survival table depicts the risk score gene expression pattern and patient survival status ranged by their risk score. **(E)** The Kaplan–Meier curve tests risk score's ability to predict patient survival. **(F)** The ROC for estimating the model prediction accuracy of 1-, 3-, and 5-year survival.

The risk groups presented diverse immunological characteristics and sensitivity to some newly identified drugs

We performed ssGSEA to compare the differences of immunocyte infiltration between the risk groups, and we noticed that many immunocytes were highly infiltrated in the high-risk group, including the immunosuppressive cells (regulatory T cell and MDSC) (Figure 7A). Also, the majority of the immune checkpoints were upregulated in the high-risk group, including CD276, CTLA4, and PDCD1, et al. (Figure 7B). We then searched GEPIA2.0 and discovered that CD276 and CA12 were significantly and positively correlated as their correlation coefficient reaches 0.55 (Figure 7C). Finally, we identified two drugs from GSDC and five drugs from cMap with therapeutic potential for high-risk patients (Figures 7D, E).

The risky genes were upregulated under hypoxia and CA12-knockdown affects EMT, cell cycle, and immune checkpoint expression

To experimentally validate whether the risky genes were correlated with hypoxia, we performed hypoxia cultivation, and

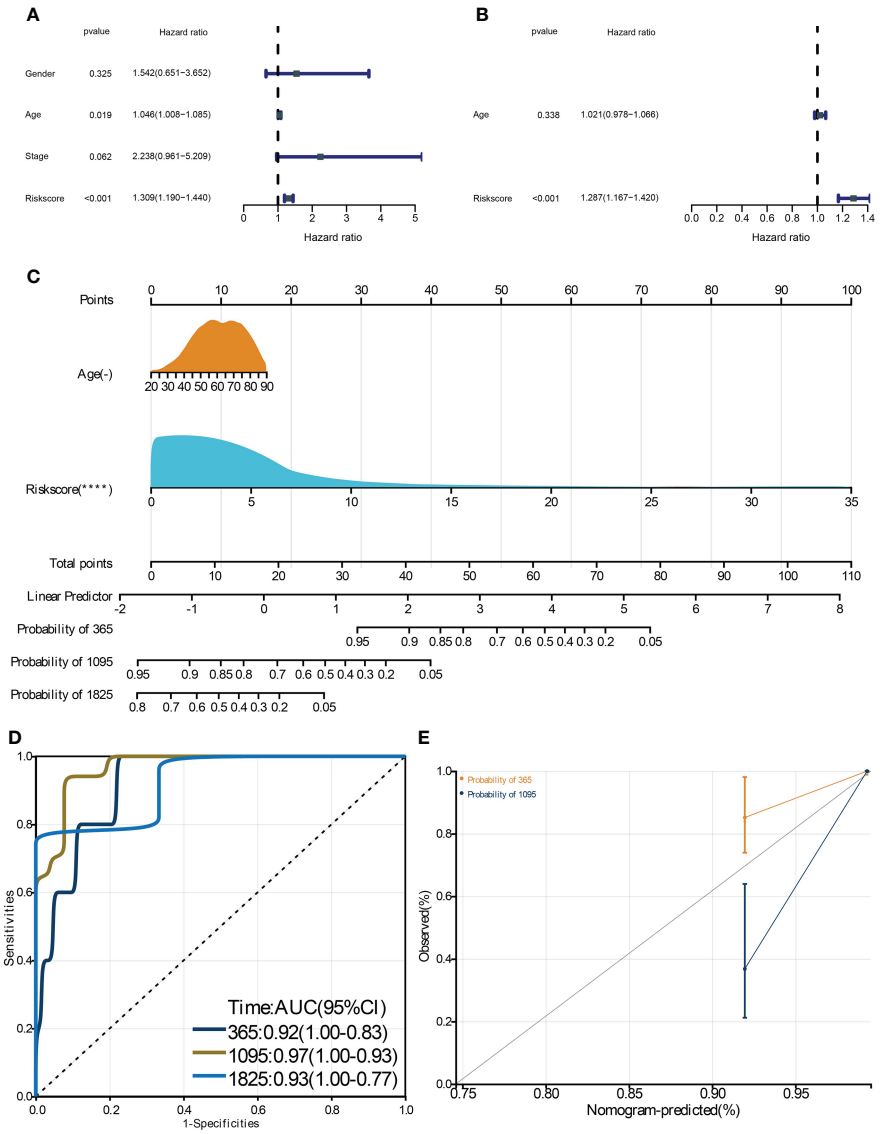


FIGURE 6
Construction and validation of a clinical nomogram. The univariate **(A)** and multivariate **(B)** Cox regression analysis to filter the prognostic predictors. **C:** The nomogram integrating the risk score and age to predict patient 1-, 3-, and 5-year survival. The ROC **(D)** and calibration curve **(E)** were used to evaluate the performances of the nomogram for predicting patient survival. **** means p-value < 0.0001.

we noticed that CA12, ARX, MGLL, and MMP9 were significantly upregulated under hypoxia; S100A13 was not significantly upregulated but also showed a similar trend (Figures 8A, B). Subsequently, we analyzed the effects of the top risky gene CA12 knockdown on cell phenotypes. The RT-qPCR results validated that the CA12 mRNA levels decreased significantly in both CA12-knockdown groups (Figure 8C). The Transwell results demonstrated that knockdown of CA12 significantly inhibited cell migration (Figure 8D), indicating a depressed EMT processes. Hence, we detected the E-cadherin, N-cadherin, and Vimentin protein expressions, representing the EMT process; the upregulated Vimentin, N-cadherin, and downregulated E-cadherin suggested an activated EMT process of MUM2B cells to enhance their migration though the N-cadherin upregulation of the siRNA-2 group was not statistically significant (Figures 8E, F).

Besides this, the clone formation and CCK8 assay results presented that CA12 knockdown also depressed cell viability and clone-formation ability (Figures 9A, B). Further, we performed flow cytometric analysis to investigate whether CA12 knockdown affected the cell cycle cells, and we found that CA12 knockdown increased the proportion of the G1 phase

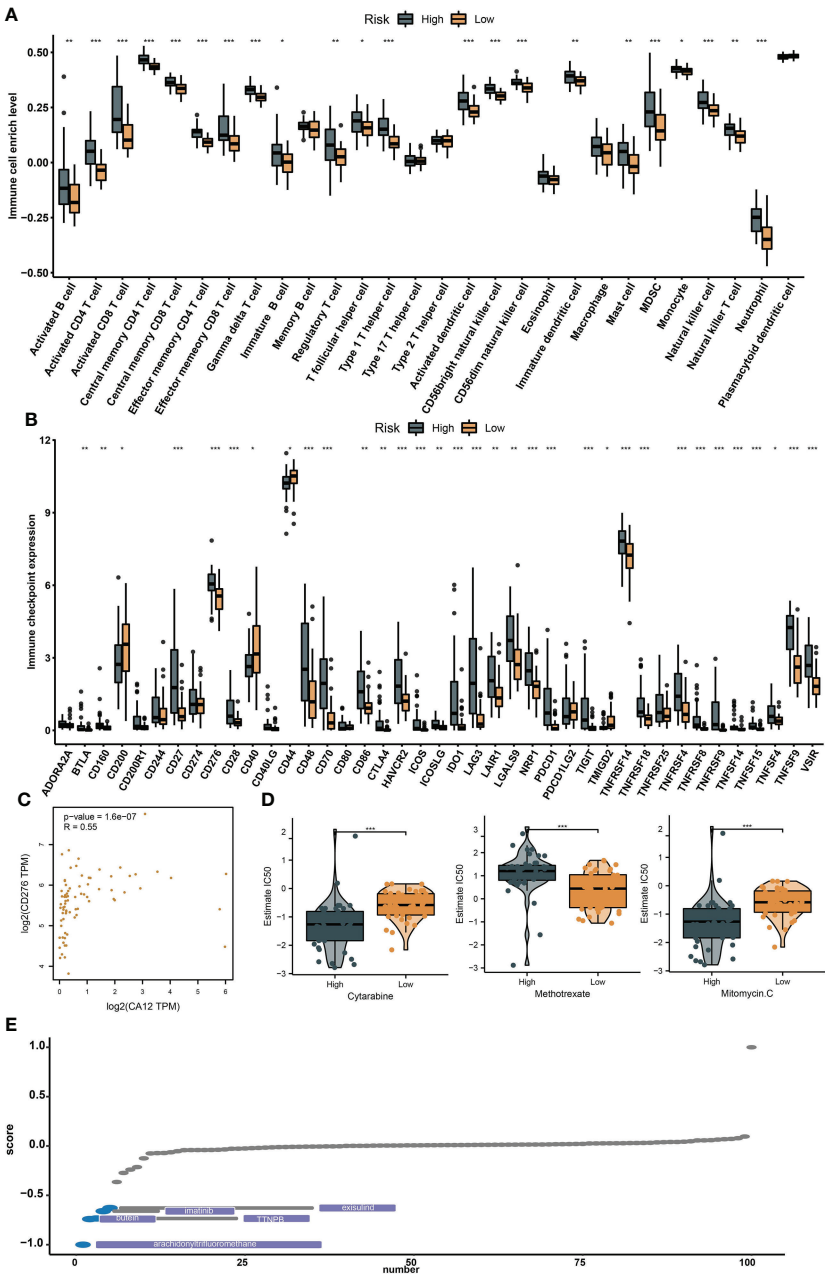


FIGURE 7
The immunological diversity between the risk groups and drug development. The box plots show the infiltration differences of immunocytes **(A)** and immune checkpoint differences **(B)** between the risk groups. **(C)** The correlation between CD12 and CD276 expression. **(D)** The IC50 level differences between the risk groups for the GDSC drugs. **(E)** Identification of potential effective drugs from the cMap database, *, **, *** represents p-value < 0.05, 0.01, and 0.001, respectively.

and decreased the S phase of cells (Figures 9C, D), implying the G1 phase arrest in MUM2B cells. The G1 phase–related proteins (cyclinD1, CDK4, and CDK6) were detected, and the decreased cyclinD1, CDK4, and CDK6 were observed in the CA12-knockdown group (Figures 9E, F), demonstrating that CA12 knockdown induced the G1 phase arrest.

Finally, we explored whether CA12 was associated with immune checkpoints. We detected the protein expression of CD276, and the results exhibited that knockdown of CA12 significantly decreased the expression of CD276 (Figures 9G, H). These results manifested that CA12 was a critical risky factor of UM for its association with EMT, cell cycle, and immune checkpoint CD276.

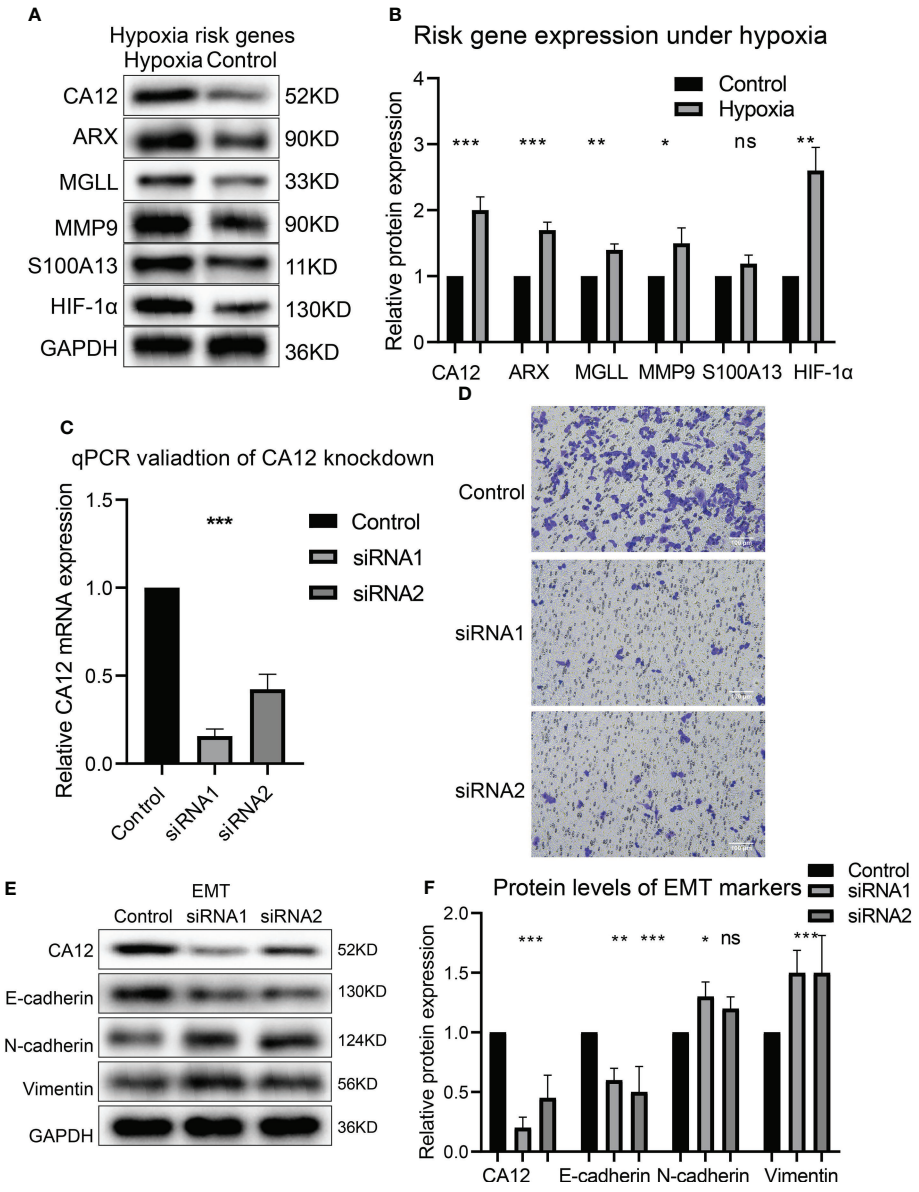


FIGURE 8
The risk gene expressions under hypoxia and the role of CA12 in UM cell EMT process. The protein bands of the risk gene expressions under normal and hypoxia conditions **(A)** and the quantification of the protein expression results in the box plot **(B)**. **(C)** The statistical comparison of the RT-qPCR results. **(D)** The image of transferred cells in control, CA12 knockdown-siRNA1, and siRNA2 groups. The bands of EMT pathway protein expressions **(E)** and their quantification in the box plots **(F)** between control, siRNA1, and siRNA2 groups, ns, *, **, *** represents not statistically significant, p-value < 0.05, 0.01, and 0.001, respectively.

Discussion

Hypoxia has various effects on cancer progression and correlates with multiple cancer hallmark features (8). In UM, hypoxia also affects cancer cell behavior like angiogenesis (11), but the extensive mechanism remained unknown, such as its interaction with cancer immunity. Also, hypoxia-based therapeutic research, except for HIF-targeted treatment, is still empty. Here, we identified hypoxia-related subtypes and constructed a hypoxia-related risk score for patient survival prediction. The subtypes and the risk score can significantly separate the patients' survival rates, and particularly, the risk score predicted patient overall survival with a high accuracy in both the training and validating cohorts according to a criteria for prediction models (20), indicating that our model was well-designed with critical hypoxia factors in cancer cell fate. To our knowledge, this is the first model focusing on the hypoxia signature's influence on UM patient prognosis.

10.3389/fonc.2022.1008770

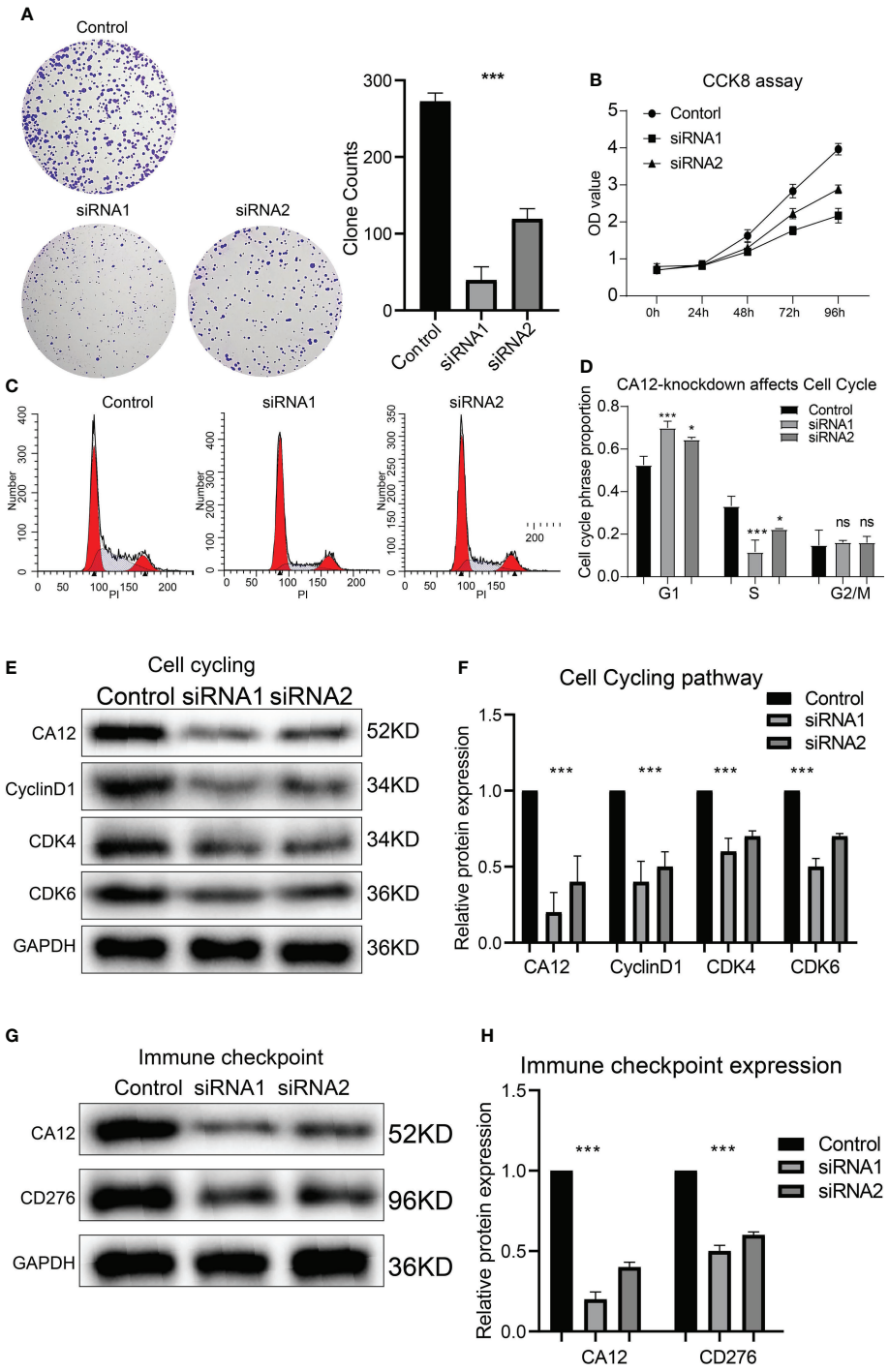


FIGURE 9
Roles of CA12 knockdown in UM cell growth and immune checkpoint expression. The images and the quantification results of clone formation assay **(A)** and the results of CCK8 assay **(B)** for the CA12-knockdown experiments. The cell cycle flow cytomatrix results **(C)** and the comparison of G1, S, and G2/M phases proportion **(D)** between control and siRNA-knockdown groups. The bands of G1 phase protein expressions **(E)** and their quantification comparison **(F)** between control and siRNA-knockdown groups. The bands of CD276 protein expression **(G)** and its quantification comparison **(H)** between control and siRNA-knockdown groups ns, *, *** means not statistically significant, p-value<0.05, and 0.001, respectively.

During the construction of the hypoxia model, we identified 10 prognostic genes, including five risk and five protective factors. Among the 10 key genes, CA12, MMP9, SLC44A3, and RNF208 have been reported to associate with UM patient survival. Carbonic anhydrase 12 (CA12) belongs to the zinc metalloenzymes that catalyzed the carbon dioxide reversible hydration. The role of CA12 in cancers remained paradoxical for it promoted pancreatic cancer apoptosis (21) while accelerating the EMT progression of glioma (22), suggesting its complex effects on cancers. In our study, we identified it as a risk gene and confirmed that CA12 knockdown arrests the cell cycle and inhibited the EMT transformation of UM cells; this is novel in UM study since no reports concerning the role of CA12 in UM have been reported. MMP9 is a matrix metalloproteinase, which degrades the extracellular matrix proteins. MMP9 has been discovered as a risk gene and predicts a worse prognosis for UM patients (23). SLC44A3 and RNF208 were found to be protective predictors (24); these discoveries are consistent with our findings in this study. Besides this, the other six genes (ARX, MGLL, S100A13, MANEALM, MYH14, and PDE4B) have not been presented in UM so far, and we first identified these genes as novel prognostic factors in UM tumors.

Hypoxia has been suggested to affect immunotherapy by tumor cell anaerobic glycolysis; the metabolite adenosine secreted to the ECM suppressed T cell activation, and thus, excused the tumor cells from immune attack (25), indicating the promising therapeutic strategy developed from cancers with hypoxia. Currently, hypoxia has presented association with immunotherapy response or immune checkpoint effects in many cancers (26–28), whereas only a clinical trial consisting of mixed melanoma patients suggested that the PD-1 blockage responders of immunotherapy has a reduced hypoxia transcriptomic change (29). In this study, we observed many upregulated immune checkpoints in the high-risk group and C1 subtype and transcription similarity with the samples that responded to the PD-1 therapy, demonstrating the potential of hypoxia as the target to improve immunotherapeutic effects. Most importantly, we validated that CD276 was downregulated when CA12 knockdown was conducted. CD276 was first derived from dendritic cells and the immune checkpoint in cancers; it impaired T cell–mediated anticancer immunity in ovarian cancer and destroyed the anti-PD-1 therapy in non-small cell lung cancer (30). Whereas in UM, no study has been reported. Hence, we first identified the relationship between CD276 and hypoxia and suggested that targeting CA12 may be a potential approach to restore CD276-mediated immunotherapeutic effects. Moreover, elevated regulatory T cell and MDSC were observed in the high-risk group and C1 subtype since CD276 was expressed on regulatory T cells and MDSC (31, 32); this indicated the involvement of MDSC or regulatory T cells in CD276-mediated immunosuppression in UM.

Comprehensively, our study identified hypoxia-related UM subtypes and risk groups, which accurately predicted the UM patient prognosis. The subtype or risk group with high hypoxia signature expression exhibited highly infiltrated immunocytes and immune checkpoints and presented transcriptional similarity with those responding to PD-1 therapy. Further, we confirmed the upregulation of the risk gene under hypoxia and validated that knockdown of CA12 inhibited UM cell EMT, clone formation, and G1/S phase transformation. Besides this, the CD276 expression decreased with CA12 knockdown. This study discovered the association between hypoxia and cancer immunity in UM and will shed light on novel therapeutic strategies development.

Data availability statement

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

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

FL contributes to the data collection, concept, and study design; YY and WD performed the data analysis and wrote the first manuscript draft. All authors approved the submission of this manuscript.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1008770/full#supplementary-material

References

1. Lamas NJ, Martel A, Nahon-Estève S, Goffinet S, Macocco A, Bertolotto C, et al. Prognostic biomarkers in uveal melanoma: The status quo, recent advances and future directions. *Cancers* (2021) 14(1):96. doi: 10.3390/cancers14010096

2. Xu TT, Moser JC, Dalvin LA. Uveal melanoma: laboratory advances and new frontiers in patient care. *Curr Opin ophthalmology.* (2021) 32(3):301–8. doi: 10.1097/ICU.0000000000000744

3. Smit KN, Jager MJ, de Klein A, Kiliç E. Uveal melanoma: Towards a molecular understanding. *Progress Retial Eye Res.* (2020) 75:100800. doi: 10.1016/j.preteyeres.2019.100800

4. Lv X, Ding M, Liu Y. Landscape of infiltrated immune cell characterization in uveal melanoma to improve immune checkpoint blockade therapy. *Front Immunol* (2022) 13:848455. doi: 10.3389/fimmu.2022.848455

5. Zhou W, Li J. Integrated analysis of genes associated with immune microenvironment and distant metastasis in uveal melanoma. *Front Cell Dev Biol* (2022) 10:874839. doi: 10.3389/fcell.2022.874839

6. Liu Y, Du H, Wan Q, He Y, Lu W, Wang W, et al. A novel four genes of prognostic signature for uveal melanoma. *J Oncol* (2022) 2022:8281067. doi: 10.1155/2022/8281067

7. Seedor RS, Orloff M, Sato T. Genetic landscape and emerging therapies in uveal melanoma. *Cancers* (2021) 13(21):5503. doi: 10.3390/cancers13215503

8. de Heer EC, Jalving M, Harris AL. HIFs, angiogenesis, and metabolism: elusive enemies in breast cancer. *J Clin Invest* (2020) 130(10):5074–87. doi: 10.1172/JCI137552

9. Jing X, Yang F, Shao C, Wei K, Xie M, Shen H, et al. Role of hypoxia in cancer therapy by regulating the tumor microenvironment. *Mol cancer.* (2019) 18(1):157. doi: 10.1186/s12943-019-1089-9

10. Masoud GN, Li W. HIF-1α pathway: Role, regulation and intervention for cancer therapy. *Acta Pharm Sin B* (2015) 5(5):378–89. doi: 10.1016/j.apsb.2015.05.007

11. D'Aguanno S, Mallone F, Marenco M, Del Bufalo D, Moramarco A. Hypoxia-dependent drivers of melanoma progression. *J Exp Clin Cancer Res CR.* (2021) 40(1):159. doi: 10.1186/s13046-021-01926-6

12. Bronkhorst IH, Jehs TM, Dijkgraaf EM, Luyten GP, van der Velden PA, van der Burg SH, et al. Effect of hypoxic stress on migration and characteristics of monocytes in uveal melanoma. *JAMA ophthalmology.* (2014) 132(5):614–21. doi: 10.1001/jamaophthalmol.2014.43

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

14. Geeleher P, Cox N, Huang RS. pRRophetic: An r package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS One* (2014) 9(9):e107468. doi: 10.1371/journal.pone.0107468

15. Tibshirani R. The lasso method for variable selection in the cox model. *Stat Med* (1997) 16(4):385–95. doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3

16. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. *Immunity* (2018) 48(4):812–30.e14. doi: 10.1016/j.immuni.2018.03.023

17. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* (2013) 41(Database issue): D955–61. doi: 10.1093/nar/gks1111

18. Huang T, Tang L, Wang H, Lin L, Fu J. Carbonic anhydrase 12 gene silencing reverses the sensitivity of paclitaxel in drug-resistant breast cancer cells. *Bioengineered* (2021) 12(2):9806–18. doi: 10.1080/21655979.2021.1995575

19. Zhao X, Shen P, Li H, Yang Y, Guo J, Chen S, et al. Carbonic anhydrase 12 protects endplate cartilage from degeneration regulated by IGF-1/PI3K/CREB signaling pathway. *Front Cell Dev Biol* (2020) 8:595969. doi: 10.3389/fcell.2020.595969

20. Liang X, Wang Z, Dai Z, Zhang H, Cheng Q, Liu Z. Promoting prognostic model application: A review based on gliomas. *J Oncol* (2021) 2021:7840007. doi: 10.1155/2021/7840007

21. Du Y, Xin Z, Liu T, Xu P, Mao F, Yao J. Overexpressed CA12 has prognostic value in pancreatic cancer and promotes tumor cell apoptosis *via* NF-κB signaling. *J Cancer Res Clin Oncol* (2021) 147(5):1557–64. doi: 10.1007/s00432-020-03447-9

22. Li G, Chen TW, Nickel AC, Muhammad S, Steiger HJ, Tzaridis T, et al. Carbonic anhydrase XII is a clinically significant, molecular tumor-subtype specific therapeutic target in glioma with the potential to combat invasion of brain tumor cells. *OncoTargets Ther* (2021) 14:1707–18. doi: 10.2147/OTT.S300623

23. Wang T, Zhang Y, Bai J, Xue Y, Peng Q. MMP1 and MMP9 are potential prognostic biomarkers and targets for uveal melanoma. *BMC cancer.* (2021) 21 (1):1068. doi: 10.1186/s12885-021-08788-3

24. Luo H, Ma C, Shao J, Cao J. Prognostic implications of novel ten-gene signature in uveal melanoma. *Front Oncol* (2020) 10:567512. doi: 10.3389/fonc.2020.567512

25. Graham K, Unger E. Overcoming tumor hypoxia as a barrier to radiotherapy, chemotherapy and immunotherapy in cancer treatment. *Int J nanomedicine.* (2018) 13:6049–58. doi: 10.2147/IJN.S140462

26. Jayaprakash P, Ai M, Liu A, Budhani P, Bartkowiak T, Sheng J, et al. Targeted hypoxia reduction restores T cell infiltration and sensitizes prostate cancer to immunotherapy. *J Clin Invest* (2018) 128(11):5137–49. doi: 10.1172/JCI96268

27. Cao R, Ma B, Wang G, Xiong Y, Tian Y, Yuan L. Characterization of hypoxia response patterns identified prognosis and immunotherapy response in bladder cancer. *Mol Ther oncolytics.* (2021) 22:277–93. doi: 10.1016/j.omto.2021.06.011

28. Wu Q, Zhou W, Yin S, Zhou Y, Chen T, Qian J, et al. Blocking triggering receptor expressed on myeloid cells-1-Positive tumor-associated macrophages induced by hypoxia reverses immunosuppression and anti-programmed cell death ligand 1 resistance in liver cancer. *Hepatol (Baltimore Md).* (2019) 70 (1):198–214. doi: 10.1002/hep.30593

29. Storkus WJ, Maurer D, Lin Y, Ding F, Bose A, Lowe D, et al. Dendritic cell vaccines targeting tumor blood vessel antigens in combination with dasatinib induce therapeutic immune responses in patients with checkpoint-refractory advanced melanoma. *J Immunotherapy Cancer* (2021) 9(11):e003675. doi: 10.1136/jitc-2021-003675

30. Zhou WT, Jin WL. B7-H3/CD276: An emerging cancer immunotherapy. *Front Immunol* (2021) 12:701006. doi: 10.3389/fimmu.2021.701006

31. Yim J, Koh J, Kim S, Song SG, Ahn HK, Kim YA, et al. Effects of B7-H3 expression on tumour-infiltrating immune cells and clinicopathological characteristics in non-small-cell lung cancer. *Eur J Cancer (Oxford Engl 1990).* (2020) 133:74–85. doi: 10.1016/j.ejca.2020.03.033

32. Wang L, Cao NN, Wang S, Man HW, Li PF, Shan BE. Roles of coinhibitory molecules B7-H3 and B7-H4 in esophageal squamous cell carcinoma. *Tumour Biol J Int Soc Oncodevelopmental Biol Med* (2016) 37(3):2961–71. doi: 10.1007/s13277-015-4132-5

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 29 September 2022
DOI 10.3389/fonc.2022.991165

OPEN ACCESS

EDITED BY
Mingyue Tan,
Shanghai University of Traditional Chinese Medicine, China

REVIEWED BY
Zewei Zhuo,
Guangdong Provincial People's Hospital, China
Bo Peng,
Second Affiliated Hospital of Harbin Medical University, China
Marwa Matboli Sayed,
Ain Shams University, Egypt

*CORRESPONDENCE
JinYu Li
haijiaoxiaoyu@csu.edu.cn

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 11 July 2022
ACCEPTED 31 August 2022
PUBLISHED 29 September 2022

CITATION
Yu J, Tang R and Li J (2022) Identification of pyroptosis-related lncRNA signature and AC005253.1 as a pyroptosis-related oncogene in prostate cancer.
Front. Oncol. 12:991165.
doi: 10.3389/fonc.2022.991165

Identification of pyroptosis-related lncRNA signature and AC005253.1 as a pyroptosis-related oncogene in prostate cancer

JiangFan Yu[1], Rui Tang[2] and JinYu Li[3]*

[1]Department of Dermatology, Second Xiangya Hospital, Central South University, Changsha, China, [2]Department of Rheumatology and Immunology, Second Xiangya Hospital, Central South University, Changsha, China, [3]Department of General Surgery, Second Xiangya Hospital, Central South University, Changsha, China

Background: Pyroptosis and prostate cancer (PCa) are closely related. The role of pyroptosis-related long non-coding RNAs (lncRNAs) (PRLs) in PCa remains elusive. This study aimed to explore the relationship between PRL and PCa prognosis.

Methods: Gene expression and clinical signatures were obtained from The Cancer Genome Atlas and Gene Expression Omnibus databases. A PRL risk prediction model was established by survival random forest analysis and least absolute shrinkage and selection operator regression. Functional enrichment, immune status, immune checkpoints, genetic mutations, and drug susceptibility analyses related to risk scores were performed by the single-sample gene set enrichment analysis, gene set variation analysis, and copy number variation analysis. PRL expression was verified in PCa cells. Cell Counting Kit-8, 5-ethynyl-2′-deoxyuridine, wound healing, transwell, and Western blotting assay were used to detect the proliferation, migration, invasion, and pyroptosis of PCa cells, respectively.

Results: Prognostic features based on six PRL (AC129507.1, AC005253.1, AC127502.2, AC068580.3, LIMD1-AS1, and LINC01852) were constructed, and patients in the high-score group had a worse prognosis than those in the low-score group. This feature was determined to be independent by Cox regression analysis, and the area under the curve of the 1-, 3-, and 5-year receiver operating characteristic curves in the testing cohort was 1, 0.93, and 0.92, respectively. Moreover, the external cohort validation confirmed the robustness of the PRL risk prediction model. There was a clear distinction between the immune status of the two groups. The expression of multiple immune checkpoints was also reduced in the high-score group. Gene mutation proportion in the high-score group increased, and the sensitivity to drugs increased significantly. Six PRLs were upregulated in PCa cells. Silencing of AC005253.1 inhibited cell proliferation, migration, and invasion in DU145 and

PC-3 cells. Moreover, silencing of AC005253.1 promoted pyroptosis and inflammasome AIM2 expression.

Conclusions: Overall, we constructed a prognostic model of PCa with six PRLs and identified their expression in PCa cells. The experimental verification showed that AC005253.1 could affect the proliferation, migration, and invasion abilities of PCa cells. Meanwhile, AC005253.1 may play an important role in PCa by affecting pyroptosis through the AIM2 inflammasome. This result requires further research for verification.

KEYWORDS

prostate cancer, lncRNA, pyroptosis, immunity, tumor biomarkers, machine learning

Introduction

Prostate cancer (PCa) is the second most common cancer in the male population worldwide and is one of the top five causes of cancer-related death (1). Epidemiological examinations and clinical studies have found that the incidence of PCa is still rising (2). The symptoms of patients with early-stage PCa are non-specific, so PCa is often found in the middle and late stages, and surgical treatment at this time has a poor prognosis and a low survival rate (3). With several important recent discoveries in immune mechanisms and advanced molecular diagnostic platforms, immunotherapy is emerging as a viable option for PCa, especially castration-resistant PCa, to stimulate antitumor immunity (4, 5). Different patient responses to the same immunotherapy have been observed in patients with different types and stages of cancer (6). Moreover, the patient response depends on multiple factors, including intratumoral heterogeneity and prior treatment history, suggesting that the need for individualized and combined therapy is an important direction for future successful immunotherapy (7, 8). Diagnostic, prognostic, and predictive biomarkers enable patient-specific management of PCa (9). Specific biomarkers to facilitate the clinical selection of immunotherapy patients include programmed death ligand 1 (PD-L1) and prostatic acid phosphatase (PAP), but these approaches are limited by tumor heterogeneity or small percentage populations (10, 11). Therefore, there is an urgent need to identify new and effective biomarkers to establish a prognostic model of PCa.

Pyroptosis, distinct from apoptosis, is a type of programmed cell death induced by the inflammasome and carried out by gasdermin proteins (12). It is characterized by cell rupture and many pro-inflammatory factors being discharged (13, 14). Pyroptosis affects tumor cell invasion, multiplication, and migration, affecting cancer prognosis (15). The association between pyroptosis and cancer is highly intricate because as a way of cell death, pyroptosis could inhibit cancer occurrence and development. Meanwhile, the release of inflammatory mediators and various signaling pathways in pyroptosis is associated with tumorigenesis and resistance to chemotherapy (16, 17). Due to the close association between pyroptosis and cancer progress and prognosis, various prognostic biomarker studies based on pyroptosis genes have been identified and used to construct gene signatures with predictive power. For example, risk signatures based on five pyroptosis-related genes (PRGs) were biomarkers to predict the immunological condition and the outcome of lung adenocarcinoma (18). Furthermore, the expression of four PRG features strongly predicted a breast cancer patient's prognosis (19). Nonetheless, the predictive merit of pyroptosis gene signatures in the prostate has not been completely clarified.

Long non-coding RNA (lncRNA) does not have the protein-coding capacity, and its length exceeds 200 nucleotides (20). LncRNA has been shown to play key functions in a variety of biological and disease processes, including cancer (21, 22). Growing evidence supports the involvement of lncRNA in PCa progression, including cell proliferation, apoptosis, metastasis, and invasion (23, 24). For example, overexpression of lncRNA PCAT14 inhibits the invasion of PCa cells and correlates with a good prognosis of PCa, which can be a diagnostic marker (25, 26). However, the role of pyroptosis-related lncRNAs (PRLs) in PCa still requires further exploration. Thus, exploring lncRNA biomarkers associated with pyroptosis in PCa is of clinical importance.

Machine learning is a branch of artificial intelligence that has been rapidly developed and applied in the field of medicine (27). Predictive models of diseases based on machine learning have been extensively mined (28). For example, Wu et al. used an ensemble of machine learning to develop a novel pyroptosis scoring system based on six lncRNAs to predict the prognosis of patients with low-grade glioma (29). In the present study, we

used a machine learning approach to construct a PRL risk signature for PCa prognosis. Then, we investigated the associations of risk score models and clinical features, immune microenvironment, immune checkpoints, genetic mutations, and drug sensitivity to provide potential diagnostic and prognostic biomarkers for PCa. This study may help to understand the effect of PRL on PCa prognosis.

Materials and methods

Data sets and pretreatments

The Cancer Genome Atlas (TCGA)-PRAD (Prostate Adenocarcinoma) dataset was downloaded from UCSC Xena (https://xenabrowser.net/). The data processing was performed to obtain FPKM data directly from TCGA and convert it into TPM value, normalized by $\log_2$ (TPM+1). The GSE116918 (GPL25318) dataset was from Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/gds/). For this validation cohort, 248 localized/locally advanced PCa patients commencing radical radiotherapy (with androgen deprivation therapy (ADT)) were included. The Affymetrix platform was utilized to generate raw data from the GSE116918 (GPL25318) dataset. The robust multi-chip averaging (RMA) algorithm was used to achieve background correction and normalization.

Construction of a machine learning prognostic model for pyroptosis-related lncRNA

Forty-four pyroptosis genes (AIM2, APIP, CASP1, CASP3, CASP4, CASP5, CASP6, CASP8, CASP9, DHX9, DDX58, ELANE, GSDMA, GSDMB, GSDMC, GSDMD, GSDME, IFI16, IL18, IL1B, MAPK8, MAPK9, NAIP, NFKB1, NFKB2, NLRC3, NLRC4, NLRP1, NLRP12, NLRP2, NLRP3, NLRP6, NLRP7, NLRP9, NOD1, NOD2, PJVK, PLCG1, PRKACA, SCAF11, TIRAP, TNF, GPX4, and IL6) were obtained from the literature (30–37). The names and abbreviations of the 44 pyroptosis-related genes are shown in Table 1. Gene set variation analysis (GSVA) was used to derive the pyroptosis score. LncRNA was then used to do correlation analysis with pyroptosis score and select the genes with $|\text{correlation coefficient}| > 0.3$ and $p < 0.05$. These genes were subjected to univariate analysis, $p < 0.05$, and the single-factor meaningful genes were selected. Then, survival random forest was used to perform dimensionality reduction analysis to screen important genes. The screening criteria were rel. importance > 0.2. Next, important gene variables were screened out, and these important gene variables were used for the least absolute shrinkage and selection operator (Lasso) regression to construct a risk score model. The risk score was the sum of gene expression values × coefficients. The flowchart of this study is presented in Figure S1.

TABLE 1 Pyroptosis gene members.

Genes	Full names
AIM2	Absent in melanoma 2
CASP1	Cysteine-aspartic acid protease-1
CASP3	Cysteine-aspartic acid protease-3
CASP4	Cysteine-aspartic acid protease-4
CASP5	Cysteine-aspartic acid protease-5
CASP6	Cysteine-aspartic acid protease-6
CASP8	Cysteine-aspartic acid protease-8
CASP9	Cysteine-aspartic acid protease-9
ELANE	Elastase, neutrophil expressed
GPX4	Glutathione peroxidase 4
GSDMA	Gasdermin A
GSDMB	Gasdermin B
GSDMC	Gasdermin C
GSDMD	Gasdermin D
GSDME	Gasdermin E
IL18	Interleukin 18
IL1B	Interleukin 1 beta
IL6	Interleukin 6
NLRC4	NLR family CARD domain containing 4
NLRP1	NLR family pyrin domain containing 1
NLRP2	NLR family pyrin domain containing 2
NLRP3	NLR family pyrin domain containing 3
NLRP6	NLR family pyrin domain containing 6
NLRP7	NLR family pyrin domain containing 7
NOD1	Nucleotide-binding oligomerization domain containing 1
NOD2	Nucleotide-binding oligomerization domain containing 2
PJVK	Pejvakin/deafness, autosomal recessive 59
PLCG1	Phospholipase C gamma 1
PRKACA	Protein kinase cAMP-activated catalytic subunit alpha
SCAF11	SR-related CTD-associated factor 11
TIRAP	TIR domain-containing adaptor protein
TNF	Tumor necrosis factor
APIP	Apoptotic protease activating factor 1-interacting protein
DHX9	DExH-box helicase 9
NLRP9	NLR family pyrin domain containing 9
NAIP	NLR family apoptosis inhibitory protein
IFI16	Interferon gamma inducible protein 16
NFKB1	Nuclear factor kappa B subunit 1
DDX58	Retinoic acid-inducible gene I
MAPK8	Mitogen-activated protein kinase 8
NLRC3	NLR family CARD domain containing 3
NLRP12	NLR family pyrin domain containing 12
MAPK9	Mitogen-activated protein kinase 9
NFKB2	Nuclear factor kappa B subunit 2

Pathway and immune-infiltration evaluation

The single-sample gene set enrichment analysis (ssGSEA) algorithm was used to quantify the abundance of 28 immune

cells (38) in PRAD and to compare immune infiltration with prognostic scores. Stromal score, Immune Score, ESTIMATE Score, and Tumor Purity were assessed with the ESTIMATE package, and the relationship between these scores and prognostic scores was compared. The GSVA package was used for GSVA analysis of Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). The correlation analysis was performed with prognostic and functional enrichment pathway scores. Correlation analysis was performed with prognostic scores and all genes, and then gene set enrichment analysis (GSEA) was performed by the clusterProfiler package.

Gene mutation and copy number variation analysis

The gene mutations of the two groups were compared, and the gene mutation patterns were checked. The somatic mutations and somatic copy number variation (CNV) profiles were gathered from the TCGA-PRAD datasets. The Genomic Identification of Significant Targets in Cancer (GISTIC) analysis was performed to evaluate the genomic features. The CNV landscape based on ITGA5 levels and the copy number gains or losses at the amplified or deleted peaks were assessed by GISTIC 2.0 analysis (https://gatk.broadinstitute.org).

Drug prediction

Information on the sensitivity of tumor cell lines to potential drugs was downloaded from Cancer Therapeutics Response Portal 2 (CTRP v2) and Profiling Relative Inhibition Simultaneously in Mixtures (PRISM). The lower the area under the curve (AUC) of the cell line, the higher the sensitivity to the potential drug. Expression of cancer cell lines was downloaded from Cancer Cell Line Encyclopedia (CCLE). Predictions were made using the R package of pRRophetic.

Cell culture and quantitative reverse transcription PCR

DU145 cells (BLUEFBIO, Shanghai, China) were grown in Dulbecco's modified Eagle's medium (DMEM) (complemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin (P/S)). PC-3 cells (Pricella, Wuhan, China) were grown in Ham's F-12K media (complemented with 10% FBS and 1% P/S). RWPE1 cells (Abiowell, Changsha, China) were cultured in keratinocyte serum-free medium (K-SFM) (complemented with 50 mg/ml of bovine pituitary extract, 5 ng/ml of epidermal growth factor (EGF), and 1% P/S).

Total RNA was obtained using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) from RWPE1, PC-3, and DU145 cells. RNA was reverse transcribed into cDNA using an mRNA reverse transcription kit (CW2569, CWBIO, Beijing, China). LncRNA expression was detected using the SYBR method (CW2601, CWBIO, China) and quantitative reverse transcription PCR (RT-qPCR) analysis with GAPDH as an internal reference. The primer sequences are shown in Table 2. The relative expression levels of genes were investigated by $2^{-\Delta\Delta Ct}$.

Cell transfection

The small interference RNA (siRNA) specifically targeting AC005253.1 (si-AC005253.1-1: 5′-CCGCAAGAAGAAGUGUGGUCATT-3′, 5′-UGACCACACUUCUUCUUGCGGTT-3′, si-AC005253.1-2: 5′-GCGUCCCAAGAAGAAGGUCAATT-3′, 5′-UUGACCUUCUUCUUGGGACGCTT-3′ and si-AC005253.1-3: 5′-GCGUCUGAUAUUUGCCGGCAATT-3′, 5′-UUGCCGGCAAAUAUCAGACGCTT-3′) and the corresponding negative controls (si-NC: 5′-UUCUCCGAACGUGUCACGUTT-3′, 5′-ACGUGACACGUUCGGAGAATT-3′) were obtained from Sangon Biotech (Shanghai, China). According to the manufacturer's protocol, cells were transfected with Lipofectamine 3000 reagent (Thermo Fisher, Waltham, MA, USA) (39).

Cell counting kit-8 assay

Cells were digested, counted, and seeded in a 96-well plate (5×10^3 cells/well, 100 µl). After adherent cell culture, 10 µl of Cell Counting Kit-8 (CCK-8) solution (NU679, Dojindo, Tokyo, Japan) was added. The cells were incubated at 37°C with 5% CO_2

TABLE 2 Primer sequences.

Genes	Sequences (5′–3′)
GAPDH	F: ACAGCCTCAAGATCATCAGC
	R: GGTCATGAGTCCTTCCACGAT
AC005253.1	F: AAGCCTTCCCTGATTACTGC
	R: CATGGTCAAACAGCCTACCTC
AC068580.3	F: CACAGCCAAAACCAAACTCCT
	R: TGGGTTGCCATTCACTGACT
AC127502.2	F: CTTCTGAATCTTTCCGGCGAAC
	R: GCGAACAACCTTCCTTGCAAA
AC129507.1	F: CTTCACTCGCACGGAGCAAC
	R: CCTCCTTGCTGCCGAGTCA
LIMD1-AS1	F: TTTGATGCCGCTTTGCTCAC
	R: TGCCACTTTTCCAGGTGTGT
LINC01852	F: GCCGGAGAACGAATGTGATG
	R: TCTTTTTGTTTACCGGAGTTCCA

for 4 h, and the optical density (OD) value at 450 nm was measured with a Bio-Tek microplate reader (MB-530, HEALES) (40).

5-Ethynyl-2′-deoxyuridine assay

The 5-ethynyl-2′-deoxyuridine (EDU) assay kit (Guangzhou RiboBio, Guangzhou, China) was used to monitor cell proliferation. The cells were inoculated into 96-well plates (1×10^4/well) until 80% confluence. Each well was added with 100 µl of EDU solution and incubated for 2 h. The cells were incubated with 4% paraformaldehyde at room temperature for 30 min. Then, the cells were treated with 100 µl of 1× Apollo® staining reaction solution for 30 min. Next, 100 µl of Hoechst 33342 reaction solution was added to each well and incubated for 30 min. A microscope (DSZ2000X, Beijing Cnmicro Instrument Co., Ltd., Beijing, China) was used to observe and take pictures.

Wound healing assay

Cells (1×10^5/well) were plated in 6-well plates until they achieved about 90% confluence. A 1-ml pipette tip was used to create scratch wounds, and photographs of the wounds (time 0 h) were immediately taken. Then, the cells were cultured in a serum-free medium. After incubation with 5% CO_2 at 37°C for 48 h, photographs were taken again.

Transwell assay

The invasion ability of cells was evaluated using the transwell assay. Transwell chambers (3428, Corning, New York, NY, USA) were pre-cooled overnight at 4°C one day in advance. Then, 100 µl of Matrigel dilute in serum-free medium was added to each well. The transfected cells were suspended in a serum-free medium and added to the transwell chamber; 500 µl of 10% fetal bovine serum complete medium was placed into the lower chamber. The cells were incubated at 37°C for 48 h. The upper chamber was removed and washed three times with phosphate-buffered saline (PBS), and the cells in the upper layer of the membrane were wiped off with a cotton ball. Cells were fixed with 4% paraformaldehyde for 20 min and stained with 0.1% crystal violet solution for 5 min. Cells were observed under an inverted microscope, and three fields of view were taken. The number of invasive cells was recorded.

Western blotting

Total protein was extracted from the cells by radioimmunoprecipitation assay (RIPA) lysate (AWB0136, Abiowell, China). Then, the protein was transferred to the polyvinylidene fluoride membrane after 10% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) treatment. The membrane was sealed with 5% skim milk (AWB0004, Abiowell) at room temperature for 2 h. AIM2 (1:1,500, 20590-1-AP, proteintech, Chicago, IL, USA), NLRC4 (1:1,000, ab201792, abcam, Cambridge, UK), NLRP3 (1:1,000, 19771-1-AP, proteintech), GSDMD-N (1:1,000, ab215203, abcam), ASC (1:2,000, 10500-1-AP, proteintech), caspase-1 (1:1,000, ab179515, abcam), IL-18 (1:8,000, 10663-1-AP, proteintech), IL-1β (1:1,000, 16806-1-AP, proteintech), and β-actin (1:5,000, 66009-1-Ig, proteintech) were incubated with the membrane at 4°C overnight. Then, the corresponding secondary antibodies were incubated with the membrane at room temperature for 2 h. The membrane was incubated with SuperECL Plus (AWB0005, abiowell), and then the protein bands were visualized by a chemiluminescence imaging system (ChemiScope 6100, Clinx, Shanghai, China).

Statistical analysis

The data were mainly visualized using the R package ggplot2. The Shapiro–Wilk normality test was used to evaluate for normality of variables. For normally distributed variables, significant quantitative differences were determined by two-tailed t-tests or one-way ANOVA. For non-normally distributed variables, significant quantitative differences were determined by the Wilcoxon test or the Kruskal–Wallis test. The Benjamini–Hochberg method was used, which converts p-values to false discovery rate (FDR) to identify significant genes. The log-rank test was used to determine the statistical differences in each dataset. The Survminer R package was used to generate survival curves. Receiver operating characteristic (ROC) curves were drawn using the pROC package. All heatmaps were generated based on pheatmap. All statistical analyses were performed in R (https://www.r-project.org/). $p < 0.05$ was considered statistically significant.

Results

Construction of pyroptosis-related lncRNA signatures

Forty-five pyroptosis genes were obtained from the literature, and the pyroptosis score was calculated by the GSVA method. The correlation between lncRNAs and pyroptosis scores was analyzed by Spearman's correlation test. The lncRNAs with |correlation coefficient| > 0.3 and $p < 0.05$ were selected, and 553 lncRNAs were obtained. These lncRNAs were subjected to univariate analysis, and 27 lncRNAs ($p < 0.05$) were screened (Figure 1A). The single-factor meaningful genes were selected for survival random forest analysis, and six important lncRNAs were

screened (Figure 1B, Figure S2A). These important gene variables were then used for Lasso analysis to build a risk scoring model (Figure 1C). The risk score was 0.4242 × AC005253.1 + 0.8249 × LIMD1-AS1 − 1.6456 × LINC01852 + 0.3408 × AC127502.2 + 1.5518 × AC068580.3 − 1.7451 × AC129507.1. With the median risk score as the cutoff value, PCa patients were categorized into high-score and low-score groups. PCa patients in the high-score group had poorer overall survival than those in the low-score group (Figure 1D). Furthermore, the AUC values of the 1-, 3-, and 5-year ROC curves were 1, 0.93, and 0.92, respectively (Figure 1E). The clinical characteristic score showed that the risk score of patients aged ≥45 was higher than that of patients aged <45, but there was no significant difference (p = 0.15). Patients with stage N1 had a higher risk score than patients with stage N0 (p = 0.00039). T4 stage patients had higher risk scores than T3 stage and T2 stage patients, and T3 stage patients had higher risk scores than T2 stage patients (p = 0.0014). Moreover, patients with the status alive had a lower risk score than patients with the status dead (p = 0.001) (Figure 1F). A publicly available dataset (GSE116918) was used to validate the reliability of the constructed risk scoring model. Consistent with the findings from the TCGA-PRAD cohort, survival analysis showed that patients in the high-score group had lower survival rates than those in the low-score group (Figure S2B). In addition, the AUC values of the 1-, 3-, and 5-year ROC curves were 0.7, 0.71, and 0.77, respectively (Figure S2C). To investigate whether the constructed risk scoring model was independent of clinicopathological parameters, univariate and multivariate Cox regression analyses were performed on age, T stage, N stage, and risk score. Risk score was the parameter independently predicting overall survival (Figures S2D, E). The predictive model could be considered an independent prognostic factor in PCa patients.

Correlation of risk scores with pyroptosis genes and immune infiltration

We surveyed the relationship between model genes and risk scores. Risk scores were favorably associated with LIMD1-AS1,

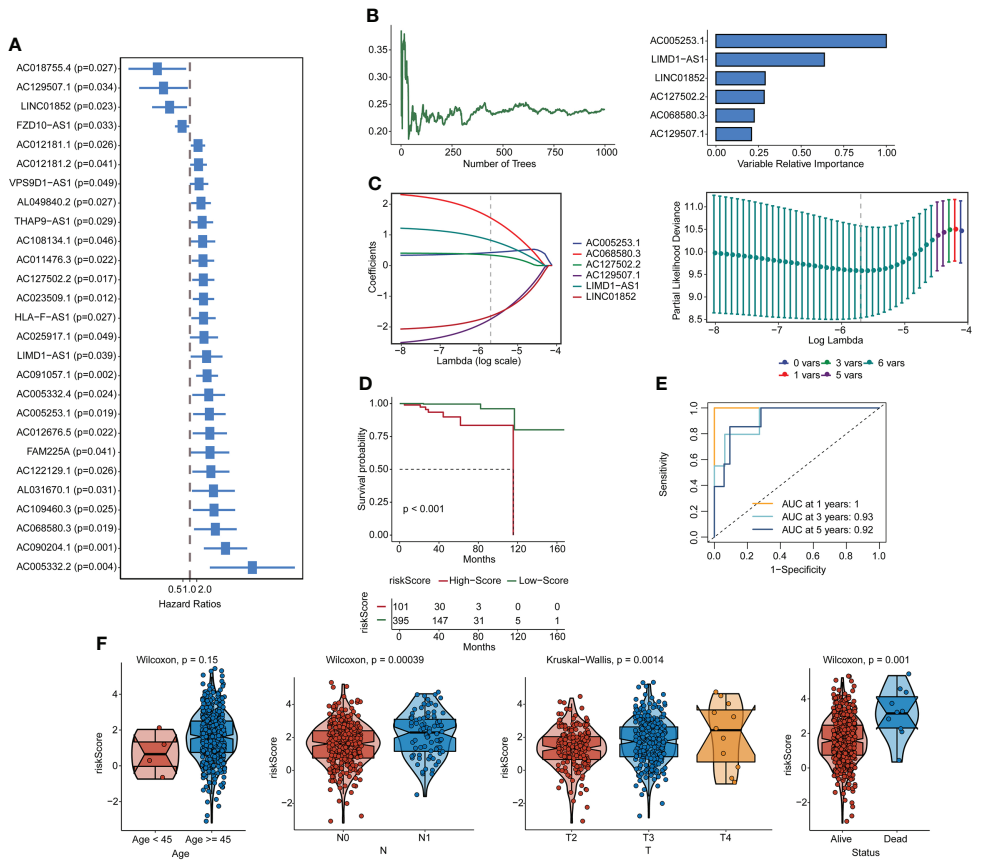


FIGURE 1
Construction of pyroptosis-related lncRNA signatures. **(A)** Correlation gene screens for genes of univariate significance. **(B, C)** Six PRL signatures were constructed through a random forest and Lasso analysis. **(D)** Survival curves. **(E)** ROC curves. **(F)** Clinical feature scores. lncRNA, long non-coding RNA; PRL, pyroptosis-related lncRNA; Lasso, least absolute shrinkage and selection operator; ROC, receiver operating characteristic.

AC127502.2, AC005253.1, and AC06850.3 and negatively associated with LINC01852 and AC129507.1 (Figures 2A, B). Next, we constructed a heatmap of risk scores and pyroptosis gene correlations. The results showed that ARRDC1-AS1, GPX4, GSDMD, GSDME, and NLRP3 were substantially associated with the risk score (Figure 2C). We used the ESTIMATE package and ssGSEA algorithm to evaluate immune infiltration. Among them, Stromal score, Immune Score, and ESTIMATE Score were negatively associated with risk score, while Tumor Purity was positively associated with risk score (Figure 2D). B cell, T cell, Macrophage, and Natural killer cell were significantly associated with risk score (Figure 2D). For example, Activated B cell, Activated CD8 T cell, and Activated dendritic cell were negatively associated with risk scores. Activated CD4 T cell was positively associated with risk score.

Immune checkpoint

We determined the expression of seven classes of immune checkpoint molecules in low- and high-risk-scoring populations. As shown in the Antigen present classification, HLA-A, HLA-B, HLA-DPA1, HLA-DPB1, HLA-DQB2, HLA-DRB1, and MICA were expressed at a low level in the high-score group. In Cell adhesion, SELP was expressed at a high level in the low-score group. In Ligand, CCL5, CX3CL1, and TGFB1 were expressed at a high level in the low-score group. In Receptor, CD27, CD40, EDNRB, and TLR4 were expressed at a high level in the low-score group. In Co-inhibitor, CD276 was expressed at a low level in the low-score group, and PDCD1LG2 and VTCN1 were expressed at a high level in the low-score group. In addition, HMGB1 was expressed at a low level in the low-score group, while ENTPD1 and PRF1 were expressed at a high level in the low-score group (Figure 3).

Functional analysis of risk score

We used the GSVA package for GO and KEGG enrichment analyses. Most samples were enriched for pathways closely related to tumorigenesis. Examples included DNA replication, cell cycle, and mTOR signaling pathway. These pathways were positively associated with risk scores (Figure 4A). Correlation analysis with risk score and functional enrichment pathway

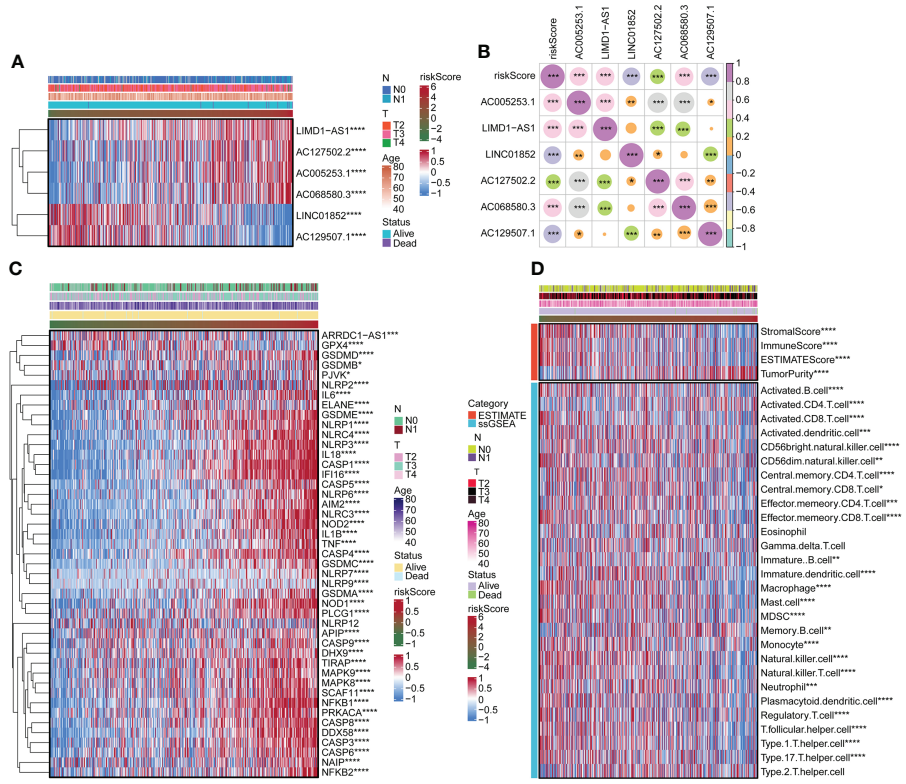


FIGURE 2
Correlation of risk scores with pyroptosis genes and immune infiltration. **(A)** The expression of LIMD1-AS1, AC127502.2, AC005253.1, AC068580.3, LINC01852, and AC129507.1. **(B)** Expression correlation plots of risk scores and model genes. **(C)** Heatmap of risk score associated with pyroptosis genes. **(D)** Heatmap of the relationship of the risk score to immune infiltration. *p < 0.05. **p < 0.01. ***p < 0.001 ****p < 0.0001.

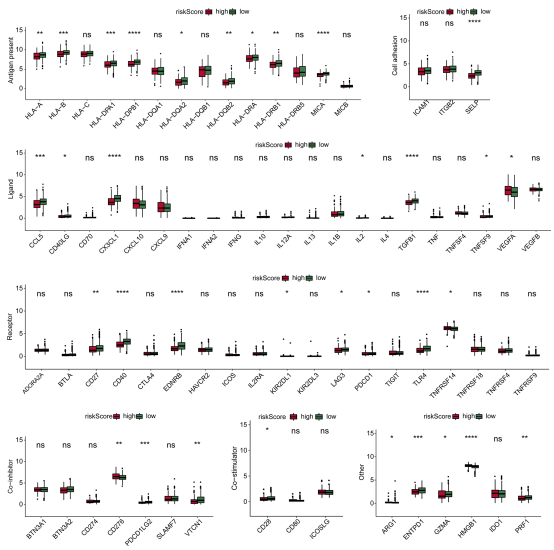


FIGURE 3
Immune checkpoint. Immune checkpoint molecule expression in low- and high-score groups. *p < 0.05. **p < 0.01. ***p < 0.001 ****p < 0.0001. ns, not significant.

score was performed. In addition to the immune checkpoint, risk scores were significantly associated with 16 other pathways (Figure 4B). Risk scores were positively correlated with cell cycle, DNA replication, DNA damage repair, and WNT target while negatively correlated with CD8 T effector and antigen processing machinery. We conducted a relevant analysis with the risk score and all genes and then performed a GSEA with the clusterProfiler package (Figure S3). The Hippo signaling pathway was downregulated.

Copy number variation and mutation analysis in high- and low-score groups

In this study, we compared the gene mutation status of two groups. Amplification frequency was mainly concentrated in 2p, 2q, 3p, 3q, 5p, 8p, 8q, 9p, 14p, 19p, and 20p, while deletion frequency was mainly concentrated in 4p, 4q, 5q, 8p, 8q, 10p, 10q, 12p, 15q, 17p, 17q, 18p, 18q, 21q, and 22q. In addition, there was a significant difference between the high-risk and low-score groups at the focal somatic copy number alterations (SCNA) level (Figure 5A). A waterfall plot was used to visualize the mutation frequency and type of the top 30 genes with the highest gene mutation frequency. The results showed that in the high-risk group, the top five genes with the highest mutation frequency were TP53 (17%), TTN (16%), FOXA1 (14%), SPOP (11%), and SPTA1 (10%), while in the low-risk group the top five genes were SPOP (11%), TP53 (10%), TTN (9%), MUC16 (5%), and KMT2D (5%) (Figure 5B).

Drug sensitivity analysis of risk scores in two groups

We downloaded information on the susceptibility of tumor cell lines to potential drugs from the CTRP v2 and PRISM. The lower the AUC of the cell line, the higher the sensitivity to the potential drug. These data revealed that among anticancer drugs, including ML258, 16-beta-bromoandrosterone, VU0155056, BRD-K02251932, BRD-K85133207, imiquimod, temoporfin, SGI-1027, and eptifibatide, the sensitivity of patients in the high-score group to the drugs was significantly increased than in the low-score group (Figure 6).

Risk score gene expression identification

Expression of AC129507.1, AC005253.1, AC127502.2, AC068580.3, LIMD1-AS1, and LINC01852 was verified by RT-qPCR in RWPE1 cells and PCa cell lines (PC-3 and DU145). The results showed that AC129507.1, AC005253.1, AC068580.3, and LIMD1-AS1 were upregulated in PC-3 and DU145 cells relative to RWPE1 cells. However, AC127502.2 and LINC01852 were only upregulated in DU145 cells (Figure 7).

Silencing of AC005253.1 affected prostate cancer cell proliferation, migration, and invasion

To explore the role of AC005253.1 in the development of PCa, we transfected si-AC005253.1-1, si-AC005253.1-2, si-AC005253.1-3, and si-NC in PC-3 and DU145 cells. RT-qPCR results showed that si-AC005253.1-1, si-AC005253.1-2, and si-AC005253.1-3 could reduce the expression of AC005253.1 in PC-3 and DU145 cells, among which si-AC005253.1-2 had a best silencing effect (Figure 8A). Therefore, si-AC005253.1-2 was used as a follow-up experiment. Silencing of AC005253.1 decreased the cell viability and proliferation of PC-3 and DU145 cells (Figures 8B, C). The wound healing assay and transwell assay results showed that after silencing AC005253.1, the migration and invasion abilities of PC-3 and DU145 cells were reduced (Figures 8D, E).

Silencing of AC005253.1 promoted pyroptosis of prostate cancer cells

We further tested the effect of AC005253.1 on the pyroptosis of PC-3 and DU145 cells by Western blotting experiment. The results showed that after silencing AC005253.1, the expression of the inflammasomes (AIM2, NLRC4, and NLRP3) was altered (Figure 9A). Compared with the si-NC group, the expression of

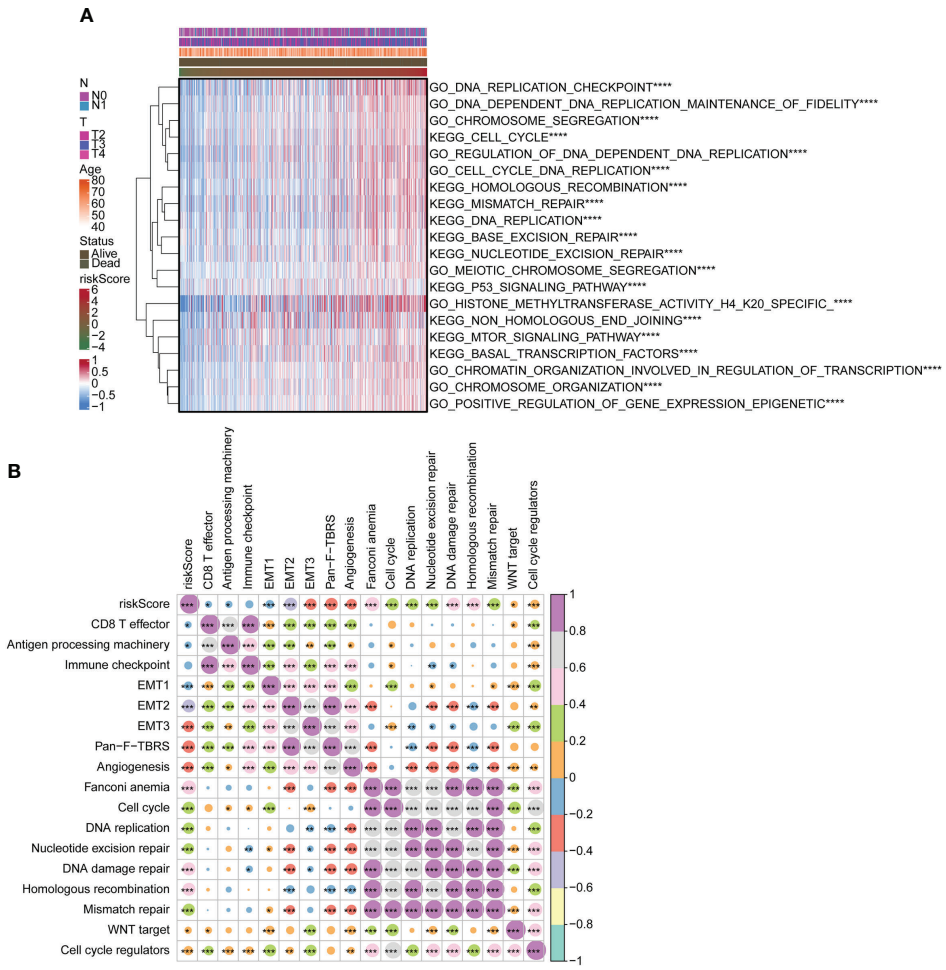


FIGURE 4
Functional analysis of risk score. **(A)** Heatmap for GO and KEGG analyses using GSVA package. **(B)** Correlation of risk score with functional enrichment pathways. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSVA, gene set variation analysis. *p < 0.05. **p < 0.01. ***p < 0.001 ****p < 0.0001.

AIM2 was increased in the si-AC005253.1 group, and the difference was most obvious (Figure 9A). Furthermore, after silencing AC005253.1, the expressions of GSDMD-N, ASC, cleaved caspase-1, IL-18, and IL-1β proteins were increased in PC-3 and DU145 cells (Figure 9B). These results suggested that silencing of AC005253.1 promoted pyroptosis in PCa cells.

Discussion

PCa is one of the most common tumors in men. Due to its heterogeneity and progressive nature, it remains incurable (41). Valid prognostic models based on specific biomarkers can accurately predict survival outcomes for the effective management of PCa patients (42). Pyroptosis-related lncRNA risk prediction models have been reported to be expected to assist in the treatment and management of various tumors (43, 44). Our study used a novel risk model of six PRLs (AC129507.1, AC005253.1, AC127502.2, AC068580.3, LIMD1-AS1, and LINC01852) developed using Lasso analysis. AC129507.1, AC005253.1, AC127502.2, AC068580.3, LIMD1-AS1, and LINC01852 were identified as PRLs for the first time. The results showed that the risk score model had the best ability to distinguish clinical characteristics between the high-risk and low-score groups significantly. In our proposed model, the AUC values of the ROC curves for 1-, 3-, and 5-year PCa were 1, 0.93, and 0.92, respectively. In addition, the 1-, 3-, and 5-year AUC values in the test set also had desirable results. Our risk model had excellent predictive power compared to other published pyroptosis-based prognostic models in PCa (45, 46). Different from the direct use of Lasso to build a prognostic model of eight pyroptosis-related genes in the study of

10.3389/fonc.2022.991165

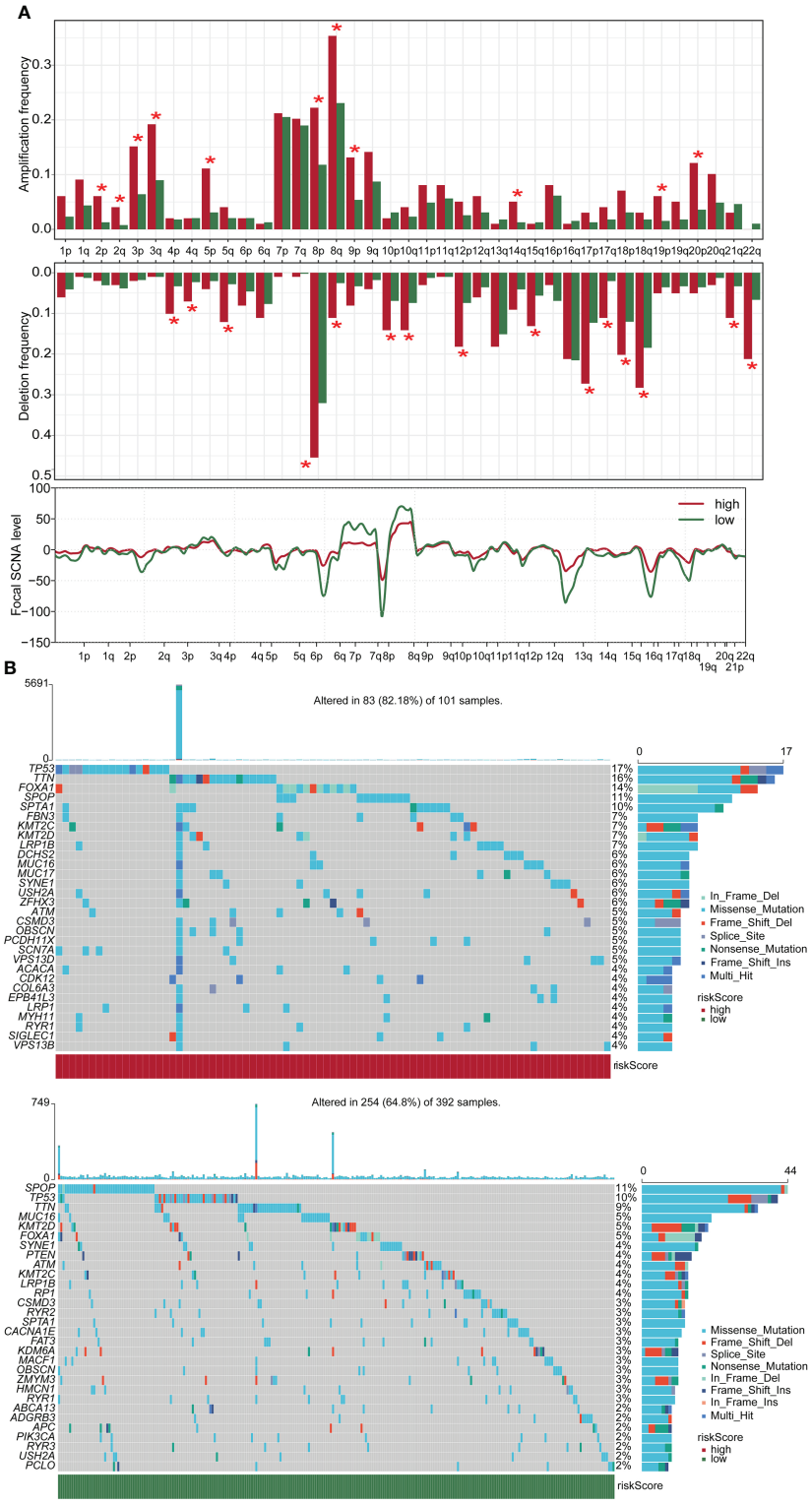


FIGURE 5
CNV and mutation analysis in high- and low-score groups. **(A)** CNV maps for groups with high and low score. **(B)** Top 30 gene mutation frequencies in two groups. CNV, copy number variation. *p < 0.05

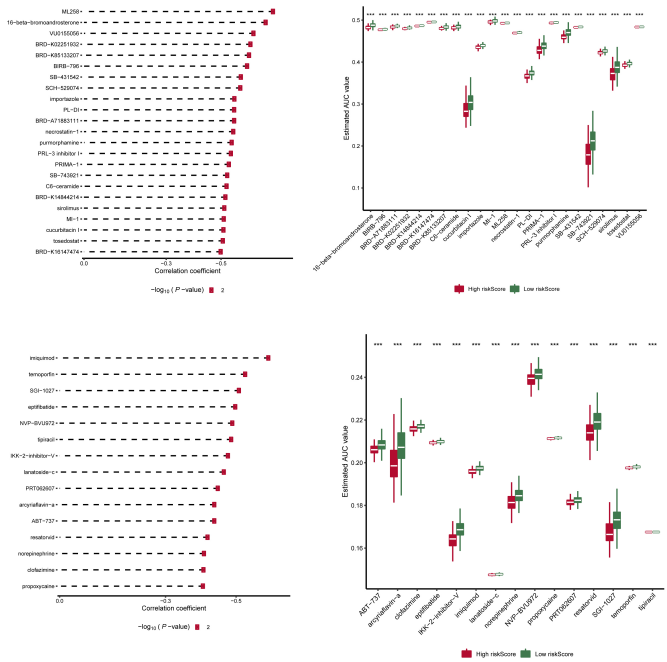


FIGURE 6
Drug sensitivity analysis of risk scores in two groups. Information on the sensitivity of tumor cell lines to potential drugs was downloaded from CTRP v2 and PRISM. Prediction of drug susceptibility in two groups using pRRophetic. The lower the AUC of the cell line, the higher the sensitivity to the potential drug. AUC, area under the curve. ***p < 0.001

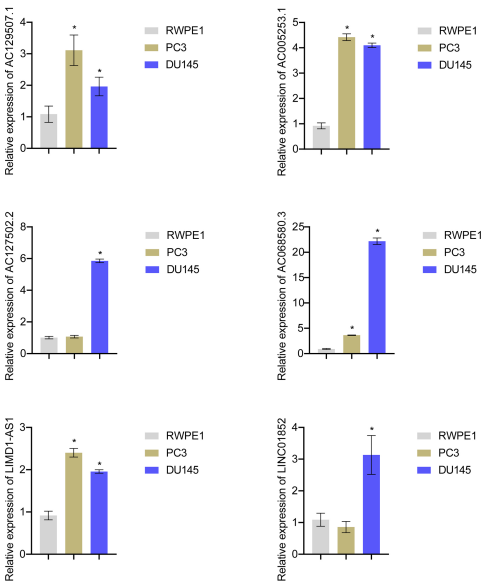


FIGURE 7
Risk score gene expression identification. RT-qPCR to detect the AC129507.1, AC005253.1, AC127502.2, AC068580.3, LIMD1-AS1, and LINC01852 expression in RWPE1, PC-3, and DU145 cell. *p < 0.05, *vs.* RWPE1.

Wang et al. (47), we used random forest dimensionality reduction and screening methods and further used Lasso analysis to build a predictive model. This analytical approach may help improve the predictive accuracy of the risk model signature.

It has been reported that pyroptosis is closely related to the tumor immune microenvironment. The release of inflammatory factors is caused by pyroptosis triggers powerful antitumor immunity (48). The infiltration of CD8+ T cells and natural killer cells in the pyroptosis-activated immune microenvironment can promote pyroptosis and form a positive feedback loop (49). Previous studies have shown that increased CD8+ T-cell infiltration is independently associated with improved survival after radical prostatectomy (50). Men with more CD4+ T cells in the prostate tumor environment have an increased risk of dying from PCa (51). B-cell activation is thought to be a driver of the PCa immune response and improves postoperative survival (52). Overall, the poor prognosis and outcome of PCa are closely related to pyroptosis-triggered immune cell infiltration, which is consistent with our results. Our results showed that the risk score was negatively correlated with activated B cells, CD8 T cells, and dendritic cells while positively correlated with activated CD4 T cells. Infiltration levels of B cells, CD8 T cells, and dendritic cells decreased with increasing risk scores, consistent with shorter survival times in patients with high scores.

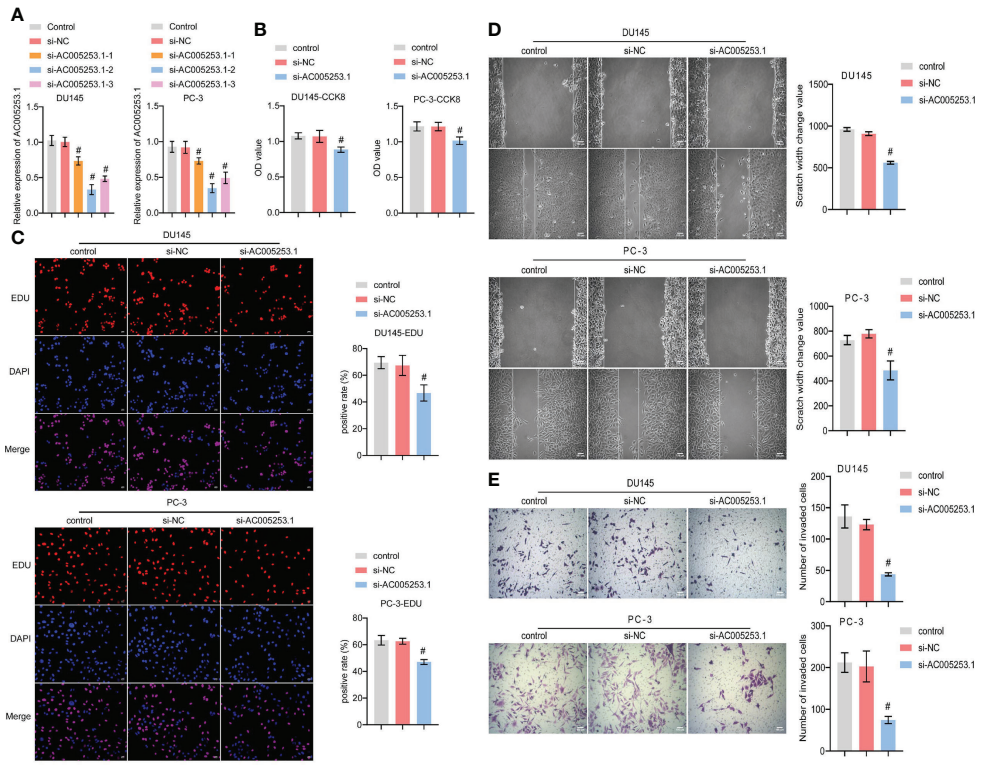


FIGURE 8
Silencing of AC005253.1 affected PCa cell proliferation, migration, and invasion. **(A)** RT-qPCR detection of AC005253.1 expression in PC-3 and DU145 cells. **(B)** CCK-8 assay was used to measure the cell viability in PC-3 and DU145 cells. **(C)** EDU assay results showed the effect of si-AC005253.1 on cell proliferation. **(D)** Wound healing assay was performed to detect the migration in PC-3 and DU145 cells. **(E)** Transwell assay was used to detect the invasion of PC-3 and DU145 cells. #p < 0.05, *vs.* si-NC group. PCa, prostate cancer; CCK-8, Cell Counting Kit-8.

We compared the expression of immune checkpoints in the high-risk and low-risk groups and found that some immune checkpoints such as HLA and MICA were expressed at a high level in the low-risk group. PCa downregulated the expression of the HLA-1 antigen processing machinery (APM) and had defects in the antigen presentation pathway (53). Low expression of MICA is associated with poorer overall survival in PCa and is associated with aggressiveness (54). This suggests that our signature could effectively identify the status of immune checkpoints in different PCa patients, providing new ideas for their treatment.

CNVs were regions of the genome with integer copy number changes, including amplifications and deletions of DNA sequences, that could drive cancer's rapid adaptive evolution and progression (55). The CNV results uncovered significant differences in mutation status between the high-score and low-score groups. In the high-score group, the gene with the highest mutation frequency was TP53 (17%), while in the low-risk group, it was SPOP (11%). TP53 mutation was the most common genetic alteration that played a major role in the pathogenesis of PCa (56, 57). SPOP mutations were associated with improved overall survival, whereas TP53 mutations were associated with poorer survival in secondary metastatic hormone-sensitive PCa (58). These data implicated that the high-score group might have more tumorigenic gene mutations.

To better assess the risk model's clinical feasibility, we analyzed information on the sensitivity of tumor cell lines to potential drugs. The results showed that the high-risk group cell lines were significantly less sensitive to drugs such as importazole and imiquimod. Importazole, a specific inhibitor that alters the interaction of KPNB1 with RanGTP, has a good inhibitory effect on PCa progression (59). Imiquimod (also known as a TLR7 agonist) inhibits the growth of mouse (TRAMP C2) and human PCa cells and can be used as an alternative therapy for locally generated PCa (60). Our findings suggest that the low-risk group is more likely to benefit from these drugs. Collectively, these findings may provide prospective treatment options for PCa patients.

Notably, RT-qPCR analysis confirmed high expression of AC129507.1, AC005253.1, AC127502.2, AC068580.3, LIMD1-AS1, and LINC01852 in PCa cell lines. AC129507.1,

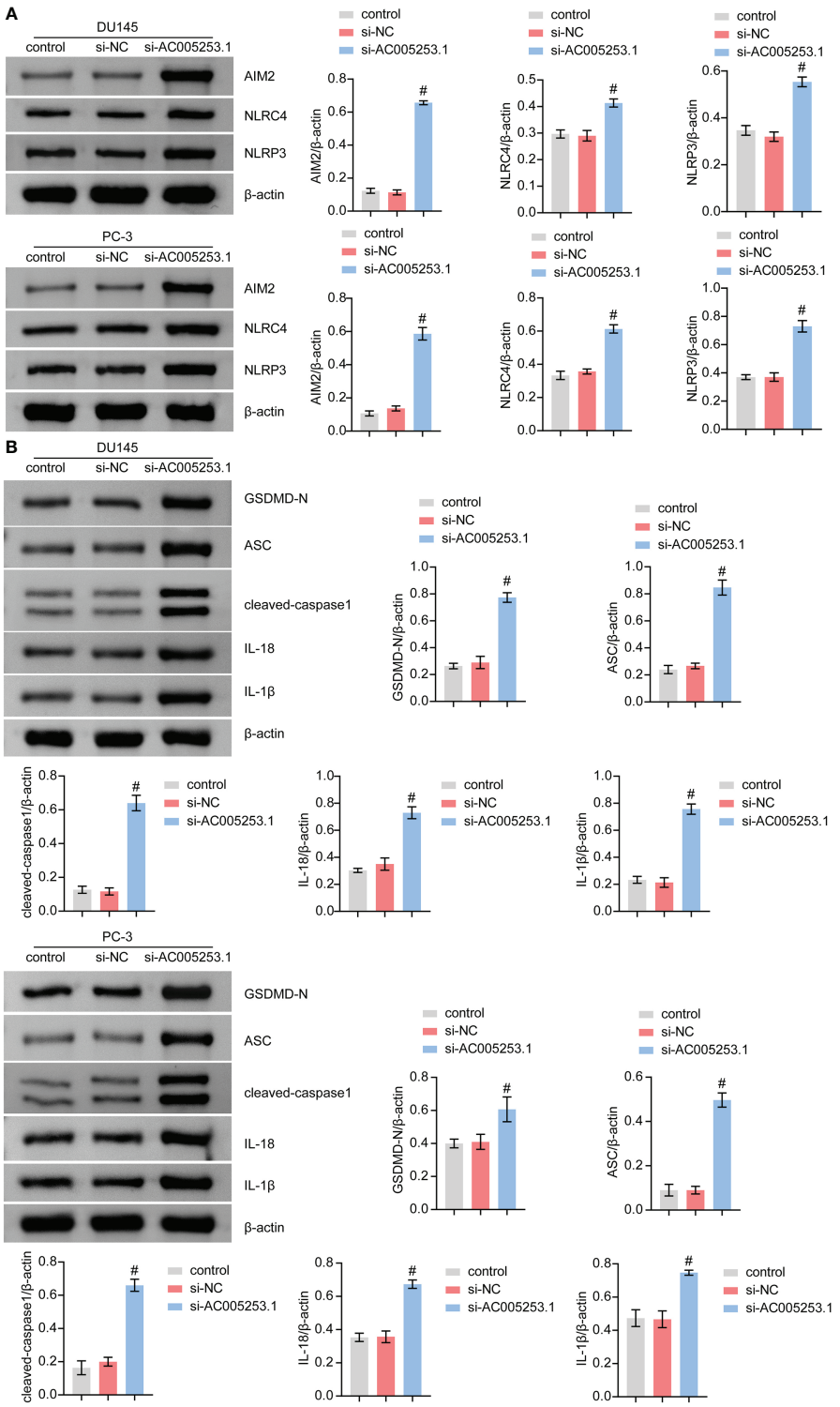


FIGURE 9
Silencing of AC005253.1 promoted pyroptosis of PCa cells. **(A)** AIM2, NLRC4, and NLRP3 levels were identified by Western blotting. **(B)** GSDMD-N, ASC, caspase-1, IL-18, and IL-1β proteins were identified by Western blotting in PC-3 and DU145 cells. #p < 0.05, *vs.* si-NC group. PCa, prostate cancer.

AC068580.3, and LIMD1-AS1 were thought to play important roles in different cancers, while AC005253.1, AC127502.2, and LINC01852 were identified for the first time. AC129507.1 was identified as an on-risk gene in risk models for prognosis patients with gastric adenocarcinoma (61). AC068580.3 was identified as an autophagy-related lncRNA as an indication of prognosis for colon adenocarcinoma (62). LIMD1-AS1 inhibited lung cancer progression by inhibiting cell multiplication and promoting apoptosis (63). Our study demonstrated that inhibiting the expression of AC005253.1 could inhibit cell viability, migration, and invasion. We further detected the expression of the inflammasome (AIM2, NLRC4, and NLRP3) and found that silencing of AC005253.1 could significantly increase the expression of the AIM2 inflammasome. Activation of the AIM2 inflammasome can promote pyroptosis (64). We also found that inhibition of AC005253.1 could promote pyroptosis in PCa cells. Therefore, we speculate that AC005253.1 may affect pyroptosis through the AIM2 inflammasome in PCa. In the present study, we report for the first time the relationship between AC005253.1 and pyroptosis in PCa.

In conclusion, we successfully established an efficient forecast PCa model based on six PRLs, including AC129507.1, AC005253.1, AC127502.2, AC068580.3, LIMD1-AS1, and LINC01852. This well-validated model built on these six PRLs will provide new insights into identifying PCa prognosis. Through *in vitro* experiments, we verified that silencing of AC005253.1 could inhibit the proliferation, migration, and invasion of PCa cells. In addition, silencing of AC005253.1 might promote pyroptosis by affecting the expression of AIM2 in PCa.

Data availability statement

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

Author contributions

JL conceived the project. RT performed the experiments. JY analyzed the data. JY and JL wrote the paper. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Hunan Provincial Natural Science Foundation of China (2021JJ40868) and the National Natural Science Foundation of China (no. 82001738).

Acknowledgments

We thank the authors who contributed to this study.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.991165/full#supplementary-material

SUPPLEMENTARY FIGURE S1
A flowchart of the study.

SUPPLEMENTARY FIGURE S2
Analysis of prognostic characteristics of risk score. (A) Univariate analysis of lncRNAs in TCGA. (B) Survival analysis of GSE116918. (C) ROC curves of GSE116918. (D, E) Univariate and multivariate Cox analysis.

SUPPLEMENTARY FIGURE S3
Enrichment analysis of risk score-related genes. After risk score correlation analysis with all genes, GO and KEGG functional enrichment analysis was performed by GSEA.

References

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

2. Zi H, He SH, Leng XY, Xu XF, Huang Q, Weng H, et al. Global, regional, and national burden of kidney, bladder, and prostate cancers and their attributable risk factors, 1990-2019. *Mil Med Res* (2021) 8(1):60. doi: 10.1186/s40779-021-00354-z

3. Rawla P. Epidemiology of prostate cancer. *World J Oncol* (2019) 10(2):63–89. doi: 10.14740/wjon1191

4. Cha HR, Lee JH, Ponnazhagan S. Revisiting immunotherapy: A focus on prostate cancer. *Cancer Res* (2020) 80(8):1615–23. doi: 10.1158/0008-5472.CAN-19-2948

5. Kruger S, Ilmer M, Kobold S, Cadilha BL, Endres S, Ormanns S, et al. Advances in cancer immunotherapy 2019 - latest trends. *J Exp Clin Cancer Res* (2019) 38(1):268. doi: 10.1186/s13046-019-1266-0

6. Emens LA, Ascierto PA, Darcy PK, Demaria S, Eggermont AMM, Redmond WL, et al. Cancer immunotherapy: Opportunities and challenges in the rapidly evolving clinical landscape. *Eur J Cancer* (2017) 81:116–29. doi: 10.1016/j.ejca.2017.01.035

7. Ma X, Du T, Zhu D, Chen X, Lai Y, Wu W, et al. High levels of glioma tumor suppressor candidate region gene 1 predicts a poor prognosis for prostate cancer. *Oncol Lett* (2018) 16(5):6749–55. doi: 10.3892/ol.2018.9490

8. Gulley JL, Madan RA. Developing immunotherapy strategies in the treatment of prostate cancer. *Asian J Urol* (2016) 3(4):278–85. doi: 10.1016/j.ajur.2016.08.008

9. Terada N, Akamatsu S, Kobayashi T, Inoue T, Ogawa O, Antonarakis ES. Prognostic and predictive biomarkers in prostate cancer: Latest evidence and clinical implications. *Ther Adv Med Oncol* (2017) 9(8):565–73. doi: 10.1177/1758834017719215

10. Kong HY, Byun J. Emerging roles of human prostatic acid phosphatase. *Biomol Ther (Seoul)* (2013) 21(1):10–20. doi: 10.4062/biomolther.2012.095

11. Deluce JE, Cardenas L, Lalani AK, Maleki Vareki S, Fernandes R. Emerging biomarker-guided therapies in prostate cancer. *Curr Oncol* (2022) 29(7):5054–76. doi: 10.3390/curroncol29070400

12. Shi J, Gao W, Shao F. Pyroptosis: Gasdermin-mediated programmed necrotic cell death. *Trends Biochem Sci* (2017) 42(4):245–54. doi: 10.1016/j.tibs.2016.10.004

13. Ding J, Wang K, Liu W, She Y, Sun Q, Shi J, et al. Pore-forming activity and structural autoinhibition of the gasdermin family. *Nature* (2016) 535(7610):111–6. doi: 10.1038/nature18590

14. Liu X, Zhang Z, Ruan J, Pan Y, Magupalli VG, Wu H, et al. Inflammasome-activated gasdermin d causes pyroptosis by forming membrane pores. *Nature* (2016) 535(7610):153–8. doi: 10.1038/nature18629

15. Fang Y, Tian S, Pan Y, Li W, Wang Q, Tang Y, et al. Pyroptosis: A new frontier in cancer. *BioMed Pharmacother* (2020) 121:109595. doi: 10.1016/j.biopha.2019.109595

16. Xia X, Wang X, Cheng Z, Qin W, Lei L, Jiang J, et al. The role of pyroptosis in cancer: Pro-cancer or pro-"host"? *Cell Death Dis* (2019) 10(9):650. doi: 10.1038/s41419-019-1883-8

17. Nagarajan K, Soundarapandian K, Thorne RF, Li D, Li D. Activation of pyroptotic cell death pathways in cancer: An alternative therapeutic approach. *Transl Oncol* (2019) 12(7):925–31. doi: 10.1016/j.tranon.2019.04.010

18. Gong Z, Li Q, Yang J, Zhang P, Sun W, Ren Q, et al. Identification of a pyroptosis-related gene signature for predicting the immune status and prognosis in lung adenocarcinoma. *Front Bioeng Biotechnol* (2022) 10:852734. doi: 10.3389/fbioe.2022.852734

19. Wang Z, Bao A, Liu S, Dai F, Gong Y, Cheng Y. A pyroptosis-related gene signature predicts prognosis and immune microenvironment for breast cancer based on computational biology techniques. *Front Genet* (2022) 13:801056. doi: 10.3389/fgene.2022.801056

20. Ulitsky I, Bartel DP. lincRNAs: Genomics, evolution, and mechanisms. *Cell* (2013) 154(1):26–46. doi: 10.1016/j.cell.2013.06.020

21. Peng WX, Koirala P, Mo YY. LncRNA-mediated regulation of cell signaling in cancer. *Oncogene* (2017) 36(41):5661–7. doi: 10.1038/onc.2017.184

22. Tan YT, Lin JF, Li T, Li JJ, Xu RH, Ju HQ. LncRNA-mediated posttranslational modifications and reprogramming of energy metabolism in cancer. *Cancer Commun (Lond)* (2021) 41(2):109–20. doi: 10.1002/cac2.12108

23. Lingadahalli S, Jadhao S, Sung YY, Chen M, Hu L, Chen X, et al. Novel lncRNA LINC00844 regulates prostate cancer cell migration and invasion through AR signaling. *Mol Cancer Res* (2018) 16(12):1865–78. doi: 10.1158/1541-7786.MCR-18-0087

24. Barth DA, Juracek J, Slaby O, Pichler M, Calin GA. lncRNA and mechanisms of drug resistance in cancers of the genitourinary system. *Cancers (Basel)* (2020) 12(8):2148. doi: 10.3390/cancers12082148

25. Yan Y, Liu J, Xu Z, Ye M, Li J. lncRNA PCAT14 is a diagnostic marker for prostate cancer and is associated with immune cell infiltration. *Dis Markers* (2021) 2021:9494619. doi: 10.1155/2021/9494619

26. Shukla S, Zhang X, Niknafs YS, Xiao L, Mehra R, Cieślik M, et al. Identification and validation of PCAT14 as prognostic biomarker in prostate cancer. *Neoplasia* (2016) 18(8):489–99. doi: 10.1016/j.neo.2016.07.001

27. Schwalbe N, Wahl B. Artificial intelligence and the future of global health. *Lancet* (2020) 395(10236):1579–86. doi: 10.1016/S0140-6736(20)30226-9

28. Chen S, Jian T, Chi C, Liang Y, Liang X, Yu Y, et al. Machine learning-based models enhance the prediction of prostate cancer. *Front Oncol* (2022) 12:941349. doi: 10.3389/fonc.2022.941349

29. Wu J, Lu L, Wang C, Jiang F. Machine learning-based integration develops a pyroptosis-related lncRNA model to enhance the predicted value of low-grade glioma patients. *J Oncol* (2022) 2022:8164756. doi: 10.1155/2022/8164756

30. Ye Y, Dai Q, Qi H. A novel defined pyroptosis-related gene signature for predicting the prognosis of ovarian cancer. *Cell Death Discovery* (2021) 7(1):71. doi: 10.1038/s41420-021-00451-x

31. Ko DC, Gamazon ER, Shukla KP, Pfuetzner RA, Whittington D, Holden TD, et al. Functional genetic screen of human diversity reveals that a methionine salvage enzyme regulates inflammatory cell death. *Proc Natl Acad Sci U S A* (2012) 109(35):E2343–52. doi: 10.1073/pnas.1206701109

32. Liang C, Fan J, Liang C, Guo J. Identification and validation of a pyroptosis-related prognostic model for gastric cancer. *Front Genet* (2021) 12:699503. doi: 10.3389/fgene.2021.699503

33. Liu T, Zhou YT, Wang LQ, Li LY, Bao Q, Tian S, et al. NOD-like receptor family, pyrin domain containing 3 (NLRP3) contributes to inflammation, pyroptosis, and mucin production in human airway epithelium on rhinovirus infection. *J Allergy Clin Immunol* (2019) 144(3):777–87.e9. doi: 10.1016/j.jaci.2019.05.006

34. Zheng J, Zhou Z, Qiu Y, Wang M, Yu H, Wu Z, et al. A pyroptosis-related gene prognostic index correlated with survival and immune microenvironment in glioma. *J Inflammation Res* (2022) 15:17–32. doi: 10.2147/JIR.S341774

35. Yap JKY, Moriyama M, Iwasaki A. Inflammasomes and pyroptosis as therapeutic targets for COVID-19. *J Immunol* (2020) 205(2):307–12. doi: 10.4049/jimmunol.2000513

36. Hara H, Tsuchiya K, Kawamura I, Fang R, Hernandez-Cuellar E, Shen Y, et al. Phosphorylation of the adaptor ASC acts as a molecular switch that controls the formation of speck-like aggregates and inflammasome activity. *Nat Immunol* (2013) 14(12):1247–55. doi: 10.1038/ni.2749

37. Yang Z, Chen Z, Wang Y, Wang Z, Zhang D, Yue X, et al. A novel defined pyroptosis-related gene signature for predicting prognosis and treatment of glioma. *Front Oncol* (2022) 12:717926. doi: 10.3389/fonc.2022.717926

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

39. Wang Q, Liang D, Shen P, Yu Y, Yan Y, You W. Hsa_circ_0092276 promotes doxorubicin resistance in breast cancer cells by regulating autophagy *via* miR-348/ATG7 axis. *Transl Oncol* (2021) 14(8):101045. doi: 10.1016/j.tranon.2021.101045

40. Zhou J, Jiang YY, Chen H, Wu YC, Zhang L. Tanshinone I attenuates the malignant biological properties of ovarian cancer by inducing apoptosis and autophagy *via* the inactivation of PI3K/AKT/mTOR pathway. *Cell Prolif* (2020) 53(2):e12739. doi: 10.1111/cpr.12739

41. Krause W. Resistance to prostate cancer treatments. *IUBMB Life* (2022). doi: 10.1002/iub.2665

42. Guo J, Zhao C, Zhang X, Wan Z, Chen T, Miao J, et al. A novel 8-gene panel for prediction of early biochemical recurrence in patients with prostate cancer after radical prostatectomy. *Am J Cancer Res* (2022) 12(7):3318–32.

43. Zhao K, Li X, Shi Y, Lu Y, Qiu P, Deng Z, et al. A comprehensive analysis of pyroptosis-related lncRNAs signature associated with prognosis and tumor immune microenvironment of pancreatic adenocarcinoma. *Front Genet* (2022) 13:899496. doi: 10.3389/fgene.2022.899496

44. Xu T, Gu H, Zhang C, Zhang W, Liang X, Cheng X. A novel risk model identified based on pyroptosis-related lncRNA predicts overall survival and associates with the immune landscape of GC patients. *Front Genet* (2022) 13:843538. doi: 10.3389/fgene.2022.843538

45. Hu D, Cao Q, Tong M, Ji C, Li Z, Huang W, et al. A novel defined risk signature based on pyroptosis-related genes can predict the prognosis of prostate cancer. *BMC Med Genomics* (2022) 15(1):24. doi: 10.1186/s12920-022-01172-5

46. Zhang G, Luo Y, Dong W, Zhong W. Characterization of a pyroptosis-related signature for prognosis prediction and immune microenvironment infiltration in prostate cancer. *Comput Math Methods Med* (2022) 2022:8233840. doi: 10.1155/2022/8233840

47. Wang M, Xia H, Yan Q, Liu W, Liu M, Wang X. Identification of pyroptosis-related gene signatures and construction of the risk model to predict BCR in prostate cancer. *Front Mol Biosci* (2022) 9:850758. doi: 10.3389/fmolb.2022.850758

48. Wang Q, Wang Y, Ding J, Wang C, Zhou X, Gao W, et al. A bioorthogonal system reveals antitumour immune function of pyroptosis. *Nature* (2020) 579(7799):421–6. doi: 10.1038/s41586-020-2079-1

49. Zhang Z, Zhang Y, Xia S, Kong Q, Li S, Liu X, et al. Gasdermin e suppresses tumour growth by activating anti-tumour immunity. *Nature* (2020) 579(7799):415–20. doi: 10.1038/s41586-020-2071-9

50. Yang Y, Attwood K, Bshara W, Mohler JL, Guru K, Xu B, et al. High intratumoral CD8(+) T-cell infiltration is associated with improved survival in prostate cancer patients undergoing radical prostatectomy. *Prostate* (2021) 81(1):20–8. doi: 10.1002/pros.24068

51. Davidsson S, Ohlson AL, Andersson SO, Fall K, Meisner A, Fiorentino M, et al. CD4 helper T cells, CD8 cytotoxic T cells, and FOXP3(+) regulatory T cells with respect to lethal prostate cancer. *Mod Pathol* (2013) 26(3):448–55. doi: 10.1038/modpathol.2012.164

52. Weiner AB, Vidotto T, Liu Y, Mendes AA, Salles DC, Faisal FA, et al. Plasma cells are enriched in localized prostate cancer in black men and are associated with improved outcomes. *Nat Commun* (2021) 12(1):935. doi: 10.1038/s41467-021-21245-w

53. Qiu T, Wang L, Liu XH, Weng XD, Kuang YL, Chen ZY, et al. Over-expressing transporters associated with antigen processing increases antitumor immunity response in prostate cancer. *Cell Immunol* (2012) 279(2):167–73. doi: 10.1016/j.cellimm.2012.10.004

54. Sakiyama MJ, Espinoza I, Reddy A, de Carlo F, Kumar A, Levenson AS, et al. Race-associated expression of MHC class I polypeptide-related sequence a (MICA) in prostate cancer. *Exp Mol Pathol* (2019) 108:173–82. doi: 10.1016/j.yexmp.2019.04.010

55. Lauer S, Gresham D. An evolving view of copy number variants. *Curr Genet* (2019) 65(6):1287–95. doi: 10.1007/s00294-019-00980-0

56. Sirohi D, Devine P, Grenert JP, van Ziffle J, Simko JP, Stohr BA. TP53 structural variants in metastatic prostatic carcinoma. *PloS One* (2019) 14(6): e0218618. doi: 10.1371/journal.pone.0218618

57. Al Zoubi MS, Otoum R, Alorjani MS, Al Bashir S, Al Trad B, Abualrja MI, et al. TP53, SPOP and PIK3CA genes status in prostate cancer. *Asian Pac J Cancer Prev* (2020) 21(11):3365–71. doi: 10.31557/APJCP.2020.21.11.3365

58. Nizialek E, Lim SJ, Wang H, Isaacsson Velho P, Yegnasubramanian S, Antonarakis ES. Genomic profiles and clinical outcomes in primary versus secondary metastatic hormone-sensitive prostate cancer. *Prostate* (2021) 81 (9):572–9. doi: 10.1002/pros.24135

59. Yang J, Guo Y, Lu C, Zhang R, Wang Y, Luo L, et al. Inhibition of karyopherin beta 1 suppresses prostate cancer growth. *Oncogene* (2019) 38 (24):4700–14. doi: 10.1038/s41388-019-0745-2

60. Han JH, Lee J, Jeon SJ, Choi ES, Cho SD, Kim BY, et al. *In vitro* and *in vivo* growth inhibition of prostate cancer by the small molecule imiquimod. *Int J Oncol* (2013) 42(6):2087–93. doi: 10.3892/ijo.2013.1898

61. Zha Z, Zhang P, Li D, Liu G, Lu L. Identification and construction of a long noncoding RNA prognostic risk model for stomach adenocarcinoma patients. *Dis Markers* (2021) 2021:8895723. doi: 10.1155/2021/8895723

62. Zhou W, Zhang S, Li HB, Cai Z, Tang S, Chen LX, et al. Development of prognostic indicator based on autophagy-related lncRNA analysis in colon adenocarcinoma. *BioMed Res Int* (2020) 2020:9807918. doi: 10.1155/2020/9807918

63. Pan J, Tang Y, Liu S, Li L, Yu B, Lu Y, et al. LIMD1-AS1 suppressed non-small cell lung cancer progression through stabilizing LIMD1 mRNA *via* hnRNP U. *Cancer Med* (2020) 9(11):3829–39. doi: 10.1002/cam4.2898

64. Xu L, Zhou J, Che J, Wang H, Yang W, Zhou W, et al. Mitochondrial DNA enables AIM2 inflammasome activation and hepatocyte pyroptosis in nonalcoholic fatty liver disease. *Am J Physiol Gastrointest Liver Physiol* (2021) 320(6):G1034–g44. doi: 10.1152/ajpgi.00431.2020

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 04 October 2022
DOI 10.3389/fonc.2022.1007514

OPEN ACCESS

EDITED BY
Mingyue Tan,
Shanghai University of Traditional Chinese Medicine, China

REVIEWED BY
Hua Jiang,
Wannan Medical College, China
Zhiri Tang,
City University of Hong Kong, Hong Kong SAR, China

*CORRESPONDENCE
Roland S. Croner
roland.croner@med.ovgu.de

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 30 July 2022
ACCEPTED 06 September 2022
PUBLISHED 04 October 2022

CITATION
Kahlert UD, Shi W, Strecker M, Scherpinski LA, Wartmann T, Dölling M, Perrakis A, Relja B, Mengoni M, Braun A and Croner RS (2022) COL10A1 allows stratification of invasiveness of colon cancer and associates to extracellular matrix and immune cell enrichment in the tumor parenchyma. *Front. Oncol.* 12:1007514. doi: 10.3389/fonc.2022.1007514

COL10A1 allows stratification of invasiveness of colon cancer and associates to extracellular matrix and immune cell enrichment in the tumor parenchyma

Ulf D. Kahlert[1†], Wenjie Shi[1,2†], Marco Strecker[1†], Lorenz A. Scherpinski[1†], Thomas Wartmann[1], Maximilian Dölling[1], Aristotelis Perrakis[1], Borna Relja[3], Miriam Mengoni[4], Andreas Braun[4] and Roland S. Croner[1*]

[1]University Clinic for General, Visceral, Vascular and Transplantation Surgery, Faculty of Medicine, Otto-von-Guericke-University, Magdeburg, Germany, [2]University Hospital for Gynecology, Pius-Hospital, University Medicine Oldenburg, Oldenburg, Germany, [3]Experimental Radiology, University Clinic of Radiology and Nuclear Medicine, Faculty of Medicine, Otto-von-Guericke-University, Magdeburg, Germany, [4]University Clinic for Dermatology, Faculty of Medicine, Otto-von-Guericke-University, Magdeburg, Germany

Background: Treatment options for metastatic colorectal cancer (CRC) are mostly ineffective. We present new evidence that tumor tissue collagen type X alpha 1 (COL10A1) is a relevant candidate biomarker to improve this dilemma.

Methods: Several public databases had been screened to observe COL10A1 expression in transcriptome levels with cell lines and tissues. Protein interactions and alignment to changes in clinical parameters and immune cell invasion were performed, too. We also used algorithms to build a novel COL10A1-related immunomodulator signature. Various wet-lab experiments were conducted to quantify COL10A1 protein and transcript expression levels in disease and control cell models.

Results: COL10A1 mRNA levels in tumor material is clinical and molecular prognostic, featuring upregulation compared to non-cancer tissue, increase with histomorphological malignancy grading of the tumor, elevation in tumors that invade perineural areas, or lymph node invasion. Transcriptomic alignment noted a strong positive correlation of COL10A1 with transcriptomic signature of cancer-associated fibroblasts (CAFs) and populations of the immune compartment, namely, B cells and macrophages. We verified those findings in

functional assays showing that COL10A1 are decreased in CRC cells compared to fibroblasts, with strongest signal in the cell supernatant of the cells.

Conclusion: COL10A1 abundance in CRC tissue predicts metastatic and immunogenic properties of the disease. COL10A1 transcription may mediate tumor cell interaction with its stromal microenvironment.

KEYWORDS

colon cancer, biomarker, tumor microenvironment, collagen type 10, prognosis

Introduction

Colorectal cancer remains to be one of the most malignant and deadliest cancers worldwide, with over 935 thousand deaths and more than 1.9 million new cases in 2020 (1), despite progressive scientific efforts. Peritoneal and hepatic distant metastases barely provide a median survival rate of 5–9 months upon diagnostic detection (2, 3). Current clinical diagnostics to appreciate the tumor location and spread involve digital rectal examination (DRE), total colposcopy with biopsy, abdominal sonography, thoracal X-ray, carcinoembryonic antigen (CEA) blood levels, and abdominal/thoracal computer tomography, and for rectum carcinoma, rigid rectoscopy, pelvic MR/CT, and rectal end sonography in case of locally limited tumors. One of the major deficiencies in the staging diagnosis of CRC is the detection of malignant lymph nodes and stratification of tumor cases with elevated metastatic risk in low and medium malignancy staging (II/III) (4) (5). On the one hand, this is due to inconsistent cutoff limits and on the other hand to poor sensitivity and specificity of conventional CT (70%, 78%) (6) (71%, 67%) (7) or CT colonography (CTC) (<70%) (8) detecting not only metastatic enlarged but also micro-metastatic lymph nodes with normal size, urging the need for improved diagnostics such as MDCT. Another way is to improve the diagnostic criteria for CT-diagnosed lymph node changes. As a result, based on recent consensus data, the largest short diameter of the suspicious tissue and internal heterogeneity have been identified as the best criteria for CT-assisted malignancy detection (9). Poor diagnosis is particularly problematic, as lymph node status determines whether adjuvant chemotherapy is indicated or not. Colon carcinomas are treated with adjuvant chemotherapy from the Union for International Cancer Control (UICC) stage II/III and rectum carcinoma depending on locality from UICC II in the middle and lower rectum with neoadjuvant regime (upper rectum with adjuvant chemotherapy) (10). Currently, neoadjuvant chemotherapy is not considered the standard of care for CRC patients; however, recent data indicate significant advantages when applying pre-surgical chemotherapy over conventional adjuvant chemotherapy in terms of OS and DSS (11). Contrary to adjuvant chemotherapy, which is started after the pathological evaluation of the resected lymph nodes, initiation and monitoring of neoadjuvant chemotherapy have so far mostly relied on imaging parameters of the tumor area alone. Oncologist and radiologist are frequently confronted with the dilemma of the inability to unequivocally discriminate false positivity of cancer metastasis from actual metastasis, meaning that in patients diagnosed with metastatic disease—and subsequently exposed to adverse-effect-evoking chemotherapy—the lymphatic system was in fact solely reactive to the tumor defense but does not represent lymph nodes with manifested metastasis. Biomarkers that can identify tumor malignancy such as predicting any possible elevated risk for the patient's tumor to enter late stages of metastatic cascades are needed. Our results enforce a previously described collagen isoform to possess the potential to do so, meanwhile also opening a discussion to serve as a direct potential therapeutic target of colon cancer tumor microenvironment.

COL10A1 is a short-chain protein and member of the collagen family of proteins, which are major components of the interstitial extracellular matrix. In addition to the general structural functions of collagen, COL10A1 has also long been attributed to cell–cell interaction. Elevated expression levels have been observed in several malignant tumor types and correlate with tumor progression, invasion, metastasis, and vascularization (12). However, its role in CRC, particular in predicting tumor progression and tumor sub-stratification into cases that would benefit from neoadjuvant therapy, is insufficiently understood. Moreover, little information probing

Abbreviations: CRC, colorectal cancer; COL10A1, collagen type X alpha 1; COLX, Collagen X; CAF, cancer-associated fibroblast; DRE, digital rectal examination; CEA, carcinoembryonic antigen; OS, overall survival; DFS, disease-free survival; DSS, disease-specific survival; TCGA, The Cancer Genome Atlas; TPM, transcripts per million; CCLE, Cell Line Encyclopedia; DFI, disease-free interval; GAPDH, glycerinealdehyde-3-phosphate dehydrogenase; CME, complete mesocolic excision; MACC1, metastasis-associated in colon cancer 1; TRPC, transient receptor potential channel; CTC, CT colonography; DRE, digital rectal examination; IOBR, immune-oncology biological research.

COL10A1 to serve as a micro-environmental niche factor that supports the progression of CRC is available.

Material and methods

Data obtain and preprocessing

Gene expression information and clinical factors of colon cancer were resourced from The Cancer Genome Atlas (TCGA) database. The count data needed to be transferred to transcripts per million (TPM) data format for the next step of the analysis. GSE14297, including 7 normal colon epithelium samples and 18 primary colorectal cancer tissues, were used to validate the gene expression difference between normal and tumor tissues. The cell line expression data was obtained from the Cancer Cell Line Encyclopedia (CCLE) database. Immune cell score data for each sample, according to gene expression, were conducted by ESTIMATE and immune-oncology biological research (IOBR) packages.

Validation COL10A1 mRNA expression in tissues and cell lines

Colon cancer RNA-seq data from TCGA was used to conduct difference gene expression analysis to identify COL10A1 expression differences between normal and tumor tissues. Paired sample validation for COL10A1 was conducted by the TCGA data. In addition, COL10A1 expression difference was validated by the external GSE14297 dataset.

Protein to protein interaction network calculation

Protein to protein interaction network is always used to identify a novel gene's potential function and related network at the protein level. Here, we used the STRING database to show the interaction network of COL10A1 with STRING default setting. Cytoscape was used to visualize the final results.

Association of COL10A1 activation with clinical variables

Patients' clinical characteristics were extracted from the TCGA database. Two groups were formed, namely, one for baseline characteristics and another for tumor invasion factors, according to variable names. The next step was to analyze COL10A1 expression differences in different clinical features.

Association of COL10A1 with consensus transcriptional markers defining tumor microenvironment

We calculated stromal, immune, and estimate scores for each patient based on COL10A1 expression, which was performed by ESTIMATE package. In addition, we also evaluated B cells, cancer-associated fibroblasts (CAFs), CD4 T cells, CD8 T cells, endothelial cells, macrophages, NK cells, and other cells infiltration scores for each sample using the IOBR package. In addition, we also explored this gene expression in single-cell level by an online tool (http://tisch.comp-genomics.org/home/). The detailed correlation between COL10A1 and immune cell markers was calculated by the Spearman test. Considering that immune checkpoints are important for tumor progression, exploring the relationship between COL10A1 and famous immune checkpoints (PD1, CD86, PDL1, CTLA4, LAG3, and TIM3) seemed to be necessary.

Retrieval of COL10A1-related immunomodulators

TISIDB database (13) integrates the interaction between multiple immune genes and tumors. By entering the COL10A1 gene on the website and selecting samples of colon cancer, immunostimulatory factors and immune inhibitors significantly associated with COL10A1 expression can be calculated.

Construction and validation of clinical prognosis signature

Immunostimulatory factors and immune inhibitors significantly associated with COL10A1 were selected from the original expression matrix. Then, we conducted univariate and multivariate Cox regression models to select candidate genes, which were combined with coefficient to construct a prognosis signature. The disease-free interval (DFI) was set as the outcome endpoint. Forty percent of the samples were randomly selected as a test dataset to validate the robustness of the model. This signature also was applied to test OS, PFS, and DSS.

Protein extraction from cell cultures

Cells were lysed and harvested at >80% confluence in the culture flask using Cell Signaling Technology® lysis buffer. The buffer was prepared, and 200 µl was added to a T-25 flask. Cells were scraped with cell scrapers and transferred to a Falcon tube. Subsequently, the cells were treated with ultrasound to ensure

complete disruption. Insoluble cellular components in the lysate were separated by centrifugation (10 min at 14,000×*g*) in a 4°C tempered centrifuge. Supernatants were stored in aliquots at −80°C for further analysis.

Protein extraction of CRC tissue samples

For suspension of the cell pellet, it was diluted 1:10 with radioimmunoprecipitation (RIPA) buffer. The buffer solution contained 10 ml RIPA buffer mixed with 50 µl phenylmethylsulfonyl fluoride (PMSF) and 100 µl protease inhibitor.

Cell lysis was performed using the FastPrep-24TM5G homogenizer. For this purpose, samples were transferred to 2-ml tubes containing Lysing Matrix E and homogenized three times for 30 s each at 8 m/s. For foam regression and final lysis, the samples were incubated on ice for additional 5 min. To separate the samples from the glass beads, a hole was pierced on the bottom of the matrix tube using a cannula. The tube was then placed in another 1.5-ml reaction tube and centrifuged at 3,000×*g* (3 min). The remaining insoluble components were removed by a second centrifugation step at 12,000×*g* for 5 min. The supernatant obtained was aliquoted and stored at −80°C.

RNA isolation

Using the ReliaPrepTM miRNA Cell and Tissue Miniprep System, RNA was obtained directly from the culture flasks. At 90% confluence, the culture supernatant was removed; cells were washed with PBS and lysed using the kit's lysis buffer and processed according to the manufacturer's instructions. Subsequently, measurement of the RNA concentration and first quality control by photometric measurement with the NanoQuant PlateTM (Tecan) were performed.

Isolation of recombinant COLX from overexpressing HEK2973-T

To obtain a positive control of ColX, the culture supernatant of cell line p52 (overexpressing recombinant COLX) was used. The p52 cells were inoculated into T-75 flasks. In these, the cells grew to a confluence of 80%. Then, media was changed from 5% to 0% fecal calf serum (FCS). After 72 h, the culture supernatant was removed.

The remaining cells were removed by centrifugation at 350×*g* for 5 min and transferred to a dialysis tube. Dialysis was performed for 24 h at 0.2 mM Tris, pH 7.5 with solution change after 8 h. Cells were removed by centrifugation at 350×*g* for 5 min. Meanwhile, water was changed twice. The dialysate was then transferred to glass flasks, frozen, and subsequently dried by lyophilization. The finished lyophilizate was then dissolved in water to achieve a 200-fold concentration of the culture supernatant. Since a yield of 50 µg/ml is expected, after lyophilization, an approximate final concentration of 10,000 µg/ml COLX is expected.

Reverse transcription and qPCR

For reverse transcription, LunaScriptTM RT SuperMix Kit was used. One microgram of RNA was transcribed into cDNA for each sample, and a non-reverse transcriptase control was included for each sample for possible non-specific quantitative PCR (qPCR) reaction as caused by contamination with genomic DNA. After RT reaction, samples were diluted at 1:10.

The qPCR was also performed using the LunaScriptTM RT SuperMix kit, and samples were pipetted accordingly. The primers are shown in Table 1. All samples were plotted as triplets and analyzed as mean values. A non-reverse transcription control was also included from each sample to check for contamination with genomic DNA.

Protein extraction, SDS-PAGE, and Western blotting

The protein concentration was determined using the Bradford assay with Bradford solution from Advanced Protein Assay Reagent (Cytoskeleton) kit. The absorbance was measured at 590 nm. A standard curve was generated using bovine serum albumin (BSA).

An adapted sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was performed to detect proteins of interest. The gel was loaded equally with 20 µg protein per gel well by measured protein concentration with Bradford assay for every sample guaranteeing a normalized and comparable standard for every sample. Proteins were denatured in advance with 4× Laemmli buffer + 8% mercaptoethanol at 95°C for 5 min. SDS-PAGE was run overnight in the refrigerator at 6.5 mA per gel.

The proteins separated by SDS-PAGE were transferred to a polyvinylidene difluoride (PVDV) membrane activated with

TABLE 1 Primers for qPCR of COL10A1 and GAPDH.

Gene	Orientation	Sequence from 5′ to 3′
COL10A1	F	AAA GGC CCA CTA CCC AAC AC
	R	ACC TTG CTC TCC TCT TAC TGC
GAPDH	F	CCT GTT CGA CAG TCA GCC GCA T
	R	GAC TCC GAC CTT CAC CTT CCC C

methanol. The blotting chamber was filled with Towbin buffer, and blotting was performed for 90 min at 300 mA. After that, the blot was washed in 1× TBS buffer + 0.1% Tween 20 (TBS/T) for 15 min and then blocked in 5% milk powder (dissolved in TBS/T) for 60 min. After another 5-min wash step with TBS/T, incubation with the primary antibody (dissolved in 5% milk powder—TBS/T) was performed overnight. The next day, four washing steps with TBS/T followed once for 15 min and three times for 5 min. This was followed by incubation with the secondary antibody (dissolved in 5% milk powder—TBS/T) for 60 min. Subsequently, it was washed again four times with TBS/T (1× 15 min, 3× 5min). Finally, the ECL substrate was added to the blots. After 5-min incubation, the images were taken.

Results

COL10A1 transcripts are accumulated in CRC tissue samples but only in a subset of widely applied *in vitro* disease models

The differential expression results show that a total of 1,636 different genes, including 797 downregulated genes and 839 upregulated genes, were screened between normal and tumor tissues (Table 2). Figure 1A indicates that COL10A1 is significantly upregulated in CRC tissues, supported by paired-sample expression validation (Figure 1B) and external dataset GES14297 (Figure 1C).

To verify if widely distributed classical human *in vitro* models of CRC recapitulate physiological relevant levels of the gene, thereby enforcing their use in translational relevant research, we assessed datasets that retrieve expression data of standardized maintained cell lines. We found interesting differences between the cells models, as MDST8, SNU1040, SNUC2A, SW48, and HCT15 showed a significant upregulation of COL10A1. whereas HT115, CL40, RCM1, HCC56, and SW1417 had a very low-level expression (Figure 1D). Additionally, the STRING database shows that COL10A1 closely interacts with several proteins, of which some are famously described as potent promoters of cancer stem cells and mesenchymal transformation, such as MMP13, SOX9, and RUNX2. The computed interactome can be seen in Figure 1E. All the data resource has been shown in Figure 1F.

Diversity-associated variances of COL10A1 expression

Although baseline characteristics are known to be determinants for clinical outcome, these characteristics such as sex, race, and body mass index (BMI) did not differ in their COL10A1 expression levels significantly. Nonetheless, Asians had a minimal higher expression level compared to other ethnicity (Figures 2A–C). To our knowledge, this is the hitherto first comparative assessment of COL10A1 appreciating different diversity setups.

COL10A1 is increased in tumor with high invasion properties

To receive further insights in the clinical translation relevance of COL10A1 in CRC, we performed another screening of COL10A1 activation level with parameters typical describing advanced tumor cell invasion. A significant increase in COL10A1 expression was observed not only in stage T3 and T4 but also in N1 and N2 cases of CRC as compared to low-stage counterparts. Significant stratification did not occur in low-stage comparisons between T1, T2, and N0 (Figures 2D, E) nor was there a difference between M1 and M0 stages (Figure 2F). Elevated COL10A1 expression is associated with advanced tumor stages based on histopathological and image-based tumor staging. Moreover, tumors featuring perineural invasion—an established marker for predicting increased metastatic condition in CRC—have significant elevated levels of COL10A1 expression (Figures 2G–I).

High COL10A1 levels are associated with elevated immune cell infiltration and extracellular matrix score

As a possible mechanism of how malignant cancers enforce their invasive and metastasis properties, the ability of cancer cells to modulate interactions with the immune microenvironment are discussed. In this line, we analyzed COL10A1 transcript in association with expression signals associated with the existence of immune cells and other parameters of immune cell infiltration. In addition to a clear positive correlation of increased COL10A1

TABLE 2 Multivariate Cox regression of COL10A1 related immunomodulators.

Gene symbol	coef	HR	95%CI (Low)	95%CI (High)	*p*-Value
CD244	−1.840308381	0.158768457	0.011655813	2.162648294	0.167248274
CD96	1.802432322	6.064380062	0.875891189	41.98775602	0.067889854
HHLA2	−0.459849416	0.631378714	0.447451581	0.890909984	0.008858794
PDCD1LG2	−1.574446309	0.207122203	0.043256938	0.991739331	0.048798872
TMIGD2	−1.242452254	0.288675444	0.055016077	1.514711989	0.141827337

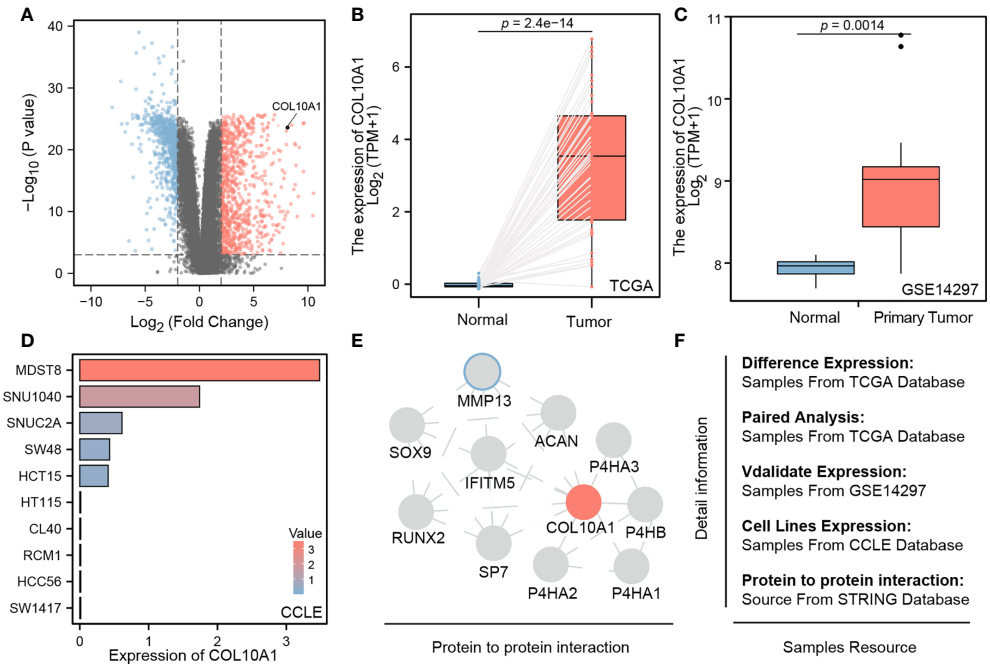


FIGURE 1
(A) COL10A1 upregulated mRNA levels in CRC tissues in comparison with all present genes; **(B, C)** paired sample expression validation of COL10A1 from TCGA and GSE14279; **(D)** different COL10A1 expression levels in common CRC cell lines; **(E)** microenvironmental protein–protein interaction of COL10A1; **(F)** data resources.

expression with stromal immunity, we identified that the extracellular matrix score is upregulated in those cases (r=0.84, r=0.53, r=07, respectively; p<0.001, Figures 3A–C). Moreover, we found the COL10A1 cases are enriched of expression signals describing infiltration of B cells, CAFs, and macrophages (r=0.19, r=0.89, and r=0.66, respectively; p<0.001, Figure 3D). Further expression analysis confirmed these results, as we reveal a correlation of COL10A1 activation with respective consensus markers describing pools of cells such as B-cell markers (CD19, r=0.135, p=0.004; CD79A, r=0.221, p<0. 001), CAFs markers (FAP, PDPN, THY1, ACTA2, COL1A1, PDGFRA, and PDGFRB; p<0.001), and M2 macrophages markers (CD163, r=0.601, p<0.001; VSIG4 r=0.576, p<0.001) (Figures 3E–I). Moreover, the single-cell analysis results also demonstrate that COL10A1 could be expressed in CAF cells (Supplementary Figure S1). Of particular interest was also the significant correlation with immune checkpoint surface proteins such as PD1, CD86, PDL1, CTLA4, LAG3, and TIM3 (p<0.001) (Figures 4A–F), indicating a possible mechanism of how COL10A1-rich CRC facilitates invasion.

COL10A1 related immunomodulators and construction of five gene risk signature to stratify patients' survival probability

A total of 18 immunoinhibitors and 32 immunostimulators of COL10A1-expression-related immunomodulators were identified. Predictive model was built by Cox regression based on the above genes. Fourteen genes were demonstrated to affect patients' outcome (Supplementary Table S1) by univariate results, and five genes are the main body of the model, which are inferred from multivariate results (Table 2). According to the median value of risk score, a high risk score means a poor outcome, while low-risk patients have a contrary prognosis. The area under the curve (AUC) was 0.781 (Figures 5A–C), and the internal validation results also support the above conclusions, while the AUC was 0.750 in the validation dataset (Figures 5D–F). In addition, we applied the model to predict OS, PFS, and DSS, but the model did not accurately predict these new endpoints (Supplementary Figure S2).

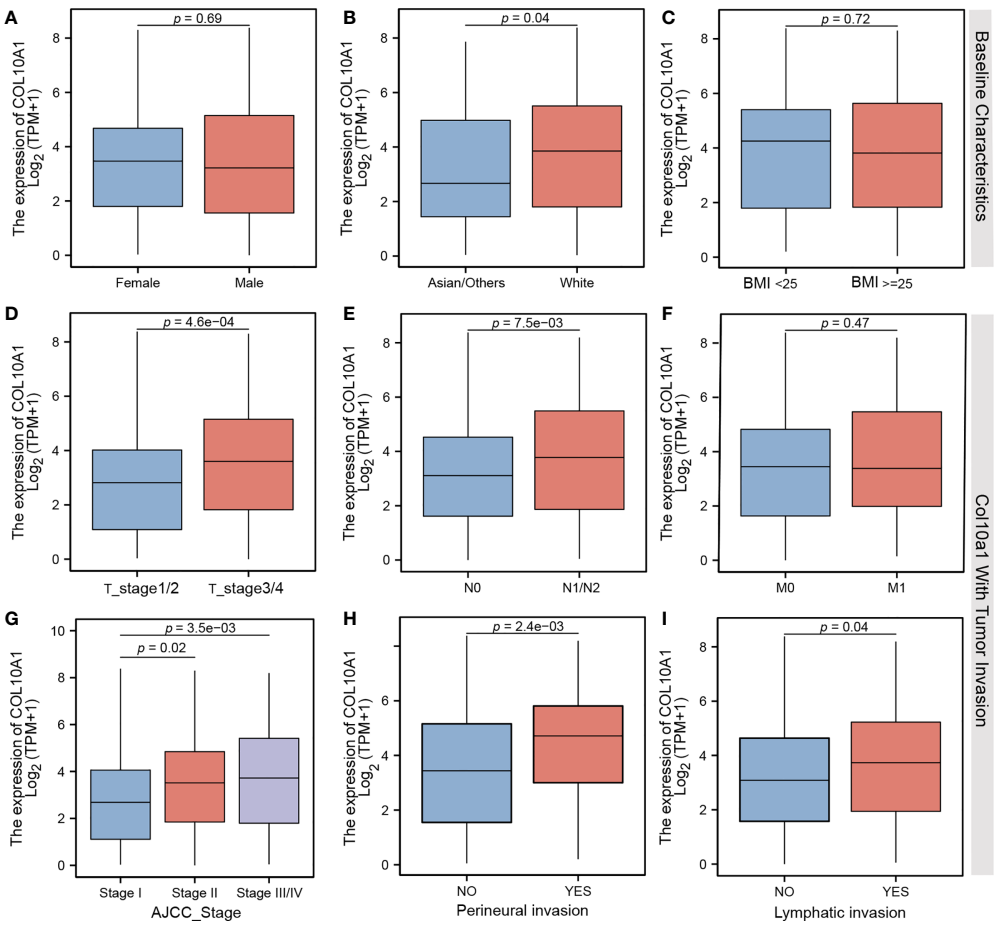


FIGURE 2
Comparison of different mRNA expression levels on **(A)** gender, **(B)** ethnicity, **(C)** body mass index, **(D)** TMN T stages, **(E)** TMN N stages, **(F)** TMN M stages, **(G)** AJCC Classification status, **(H)** perineural invasion, and **(I)** lymphatic invasion of CRC patients.

In vivo and *in vitro* analysis of COL10A1 expression levels in CRC

Microarray analyses by Croner et al. (14) revealed significantly increased expression levels of COL10A1 in tumor tissue samples compared with that in normal tissues in CRC patients. Chapman et al. (15) successfully reproduced these data, but interestingly, high expression levels could not be found in *in vitro* cultured CRC cell lines. To address the question of whether tumor cells themselves exhibit increased expression of COL10A1, CT values were compared between tumor, normal tissue, fibroblasts, and CRC cell lines. This was accomplished by first normalizing the CT values of COL10A1 to GAPDH and second by comparing those RNEs (Figures 6A, B). The tumor tissue significantly shows the highest expression followed by the adjacent normal tissue, which is still higher than any expression of other cell populations, driving the hypothesis of COL10A1 overexpression in tumor stroma, triggered by lateral information transfer between tumor and stromal cells. Our hypothesis based on the bioinformatics analysis that CAFS/fibroblasts is the source of COL10A1 overexpression in CRC is strongly supported due to these results.

COLX protein expression in several cancer and fibroblast cell lines

For PA-597603, the monomeric (~75 kDa) and multimeric forms (~140 kDa) were detected in our recombinant COLX from HEK2973 T cells. Bands at the level of the multimeric form were also weakly found in all CRC cell lines and in fibroblasts and skin sample. Particularly strong bands were detected in the range of ~45 and ~50 kDa and occasionally at ~20 kDa.

The C-terminal antibody MA5-32504 was also able to detect the monomeric form of the recombinant protein at ~75 kDa, which was absent in all other cell lines. In contrast, bands in the ~66-kDa range and at ~6.5 kDa (except LoVo and DLD-1) were detected here.

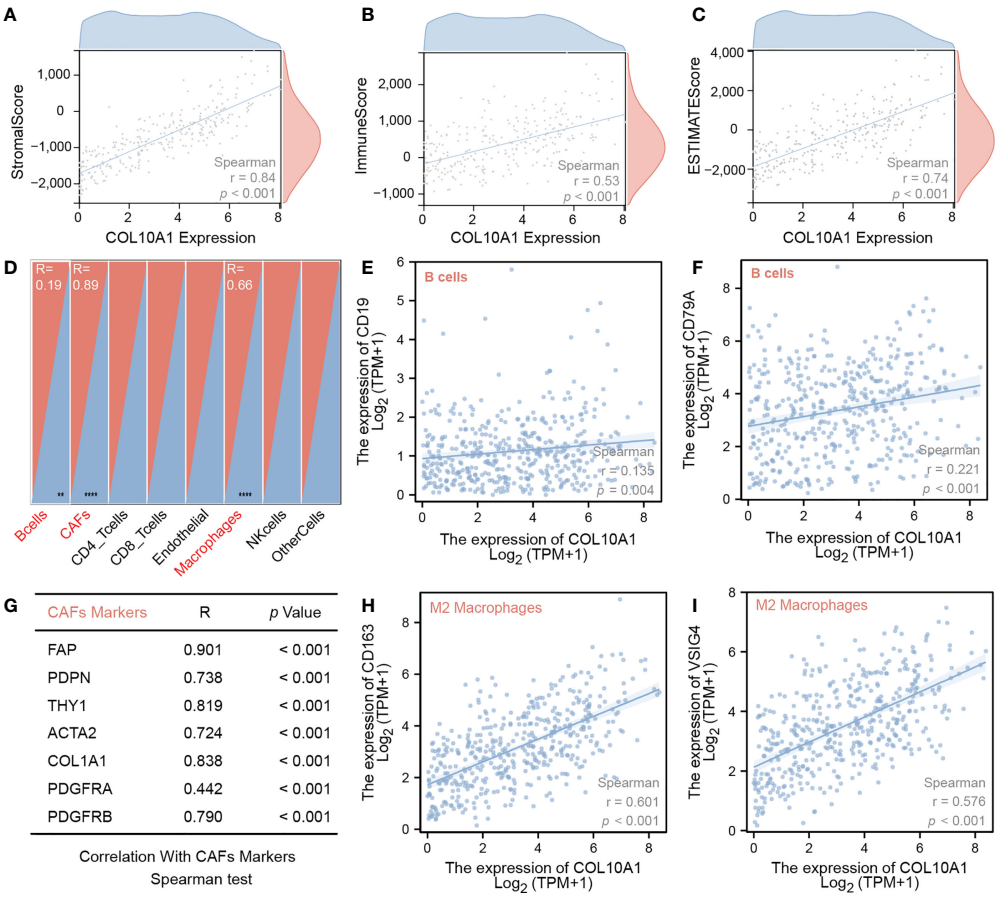


FIGURE 3
Spearman correlation test of COL10A1 mRNA levels and **(A)** stromal score, **(B)** immune score, **(C)** ESTIMATE score, **(E)** CD19, **(F)** CD79A, **(G)** different CAF cell markers, **(H)** CD163, and **(I)** VSIG4; **(D)** overview of correlation analysis of COL10A1 expression an several cell types.

The X53 antibody detected the monomeric form at ~75 kDa and the multimeric form at ~140 kDa in recombinant COLX. In cell lines, the multimeric form was detected at ~170 kDa. The antibody PA5-49198 paralleled with PA5-97603, especially in all CRC cell lines, and in fibroblasts, the ~50-kDa band was the most intense. Other bands were noted at ~45, ~32, and ~20 kDa (Figures 7A–F).

Discussion

Health issues associated with CRC is significant factor of the oncology-related health burden on the society. As prevalence directly relates to the socio-economic development of a country, besides the hotspots Western Europe, Australia, and North America, an increasing incidence can be observed in South America and Eastern Europe, mainly due to lifestyle changes, making CRCs as 1 out of 10 cancer cases in 2020 (CRC total, 1.9×10^6, 935,000 deaths) (1), expected to reach 3×10^6 incidences in 2040 (16). As prevalence rises, drug market size follows with an estimated size of 10.9 billion dollars in 2022 (17) in five most prevalent countries, mainly attributed to adjuvant and neoadjuvant chemotherapy of surgically resectable CRC cancers.

Surgery is the main type of treatment with the strongest positive clinical prognostic consequences for all CRC combined, featuring a rising cost of surgery in recent decades (30,000 cases in 2015 in Germany, CRCs cost of illness ratio reaching roughly 50,000 EUR per patient). The surgical treatment procedure for confirmed CRC depends significantly on whether existing metastasis locally in surrounding lymph nodes exists or not, defining applied resection technique and resection size and aggressive adjuvant therapy. Thus, an accurate diagnosis of this status has an immense impact on the operation procedure and the patient's rehabilitation after surgery. The procedure depends mainly on the localization of the arterial transection, which determines the size of resected colon/rectum segments. Thus, a lymphogenic metastasis follows the regular pattern of an initially longitudinal paracolic (maximum of 10 cm) and then truncal metastasis, which then decides the intestinal

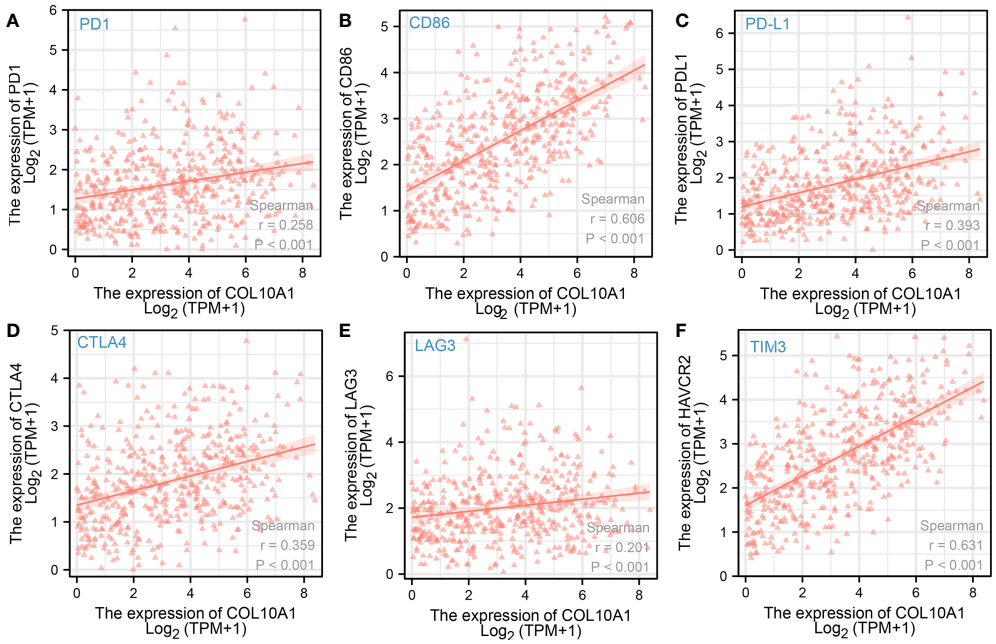


FIGURE 4
Spearman correlation test of COL10A1 expression levels and expression of immune checkpoint molecules **(A)** PD1, **(B)** CD86, **(C)** PD-L1, **(D)** CTLA-4, **(E)** LAG3, and **(F)** TIM3.

resection size with possible partial resection of infiltrated neighboring structures. In colon carcinoma, the spectrum ranges from segmental or hemicolectomies up to complete mesocolic excision (CME) depending on infiltration depth and N status. This will also determine the level of lymphatic resection, i.e., D2—paracolic and intermediate lymph nodes or D3—main, paracolic and intermediate lymph nodes, which is comparable to western CME (18). Therefore, it is becoming increasingly important to improve diagnostic accuracy. A biomarker in the primary tumor that trustfully predicts actual tumor spread would impact both types of the treatment decision. Our project associates to this initiative.

Previous works of others have identified some promising biomarkers; however, hitherto, there is no consensus marker established to support clinical decision making in the before-mentioned manner. Given the economic and clinical importance, the field of research is relatively large and allows only an insufficient

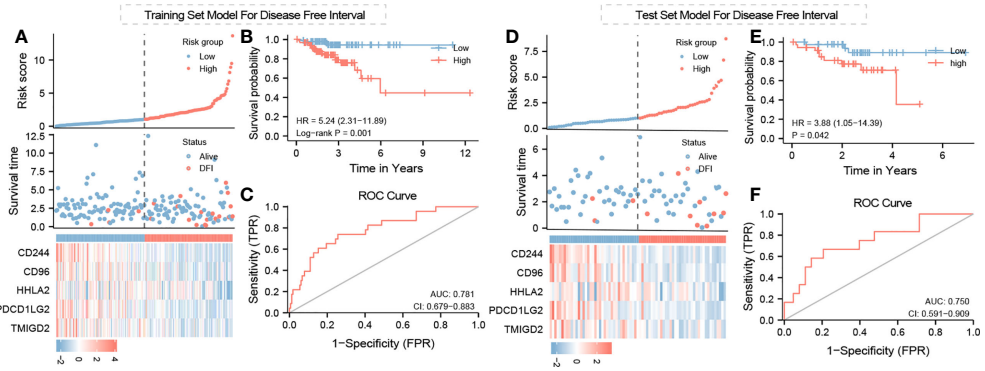


FIGURE 5
Training and test set model for DSS, PFS, and OS: risk score survival time and immunomodulator expression profiles regarding Col10A1 expression and risk group relation for **(A)** training set model and **(D)** test set model; **(B, E)** Kaplan–Meier curve for training set and test set model; **(C, F)** ROC curve for test model validation and optimization of training and test set model.

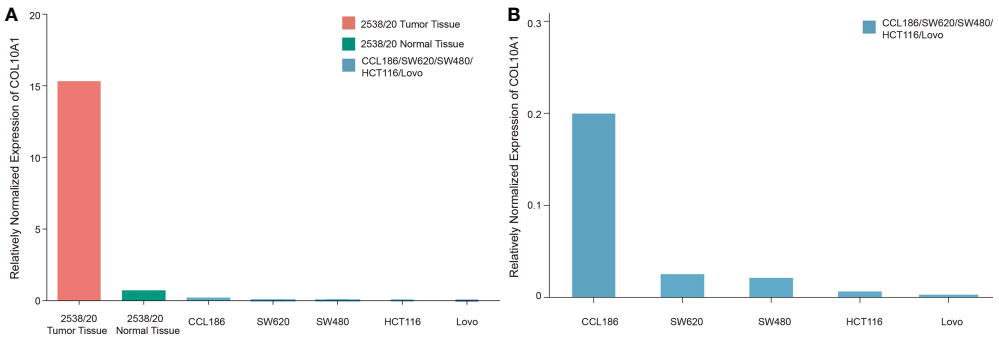


FIGURE 6
Relatively normalized expression (RNE) of COL10A1 in solid tumor, adjacent normal colon tissue, and CRC and fibroblast cell lines. For RNE, the sample with the lowest CT value is set as reference (value: 1) for other tissues and cell lines. CT values were normalized on GAPDH. **(A)** Illustration of multiple increased expression levels in CRC tumor stroma compared with *in vitro* cultured CRC cell lines and fibroblasts. **(B)** Enlarged breakdown of COL10A1 expression levels of common fibroblast and CRC cell lines. Fibroblasts (CCL186) were shown to have an eightfold increased expression compared to the strongest expressing tumor cell SW620.

discussion of the development. The most prominent example has been initially discovered about a decade ago: Stein et al. (19) identified the gene metastasis-associated in colon cancer 1 (MACC1) to possess strong predictive potential to distinguish CRC metastasis risk, and the abundance of the related DNA string in the blood of patients can help to identify cancer recurrence and therapy response. The diagnostic value has been described in independent clinical cohorts, however mostly fundament on retrospective trials (20). To our knowledge, no association to surgical decision making or omission of neoadjuvant treatment in respective MACC1 low-expressing patients in a prospective manner has been conducted. It would be interesting to study the correlation of MACC1 and COL10A1 activation in the analysis of bulk tumor specimen and in

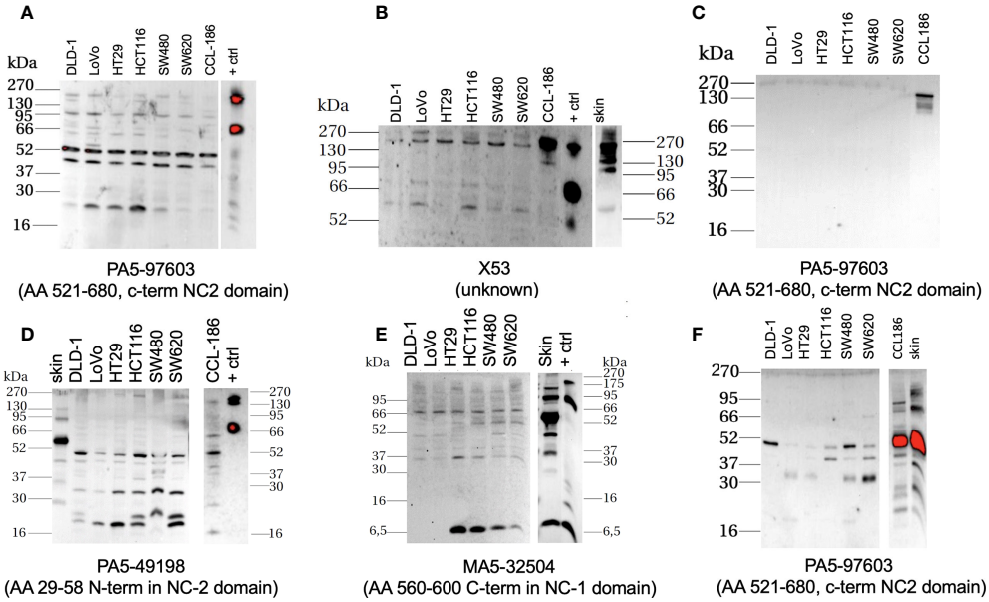


FIGURE 7
Western-blot-based assessment of COL10A1 with four different antibodies (antibody with its putative binding target labeled below the individual membranes, all Thermo Fisher) in extracts from classical CRC cell lines, fibroblast cell line, and skin. Little to low expression of the protein in cancer cells, elevated levels can be detected in the protein lysates and supernatants retrieved from fibroblasts/skin; positive control (+ ctrl) is purified recombinant protein from COL10A1 overexpressing HEK2973 T cells, housekeeping protein: GAPDH, antibody MA5-15738 (Thermo Fisher). **(A, B, D, E)** Cell lysates; **(C)** supernatants 20× concentrated by ultrafiltration, and **(F)** supernatants 70× concentrated by lyophilization.

functional studies applying genetic COL10A loss of function models. From the similar historic time span, in the early 2010, Smith et al. constructed a 34-gene signature that predicts the metastatic spread of CRC based on the experimental model of tumor metastasis (21). Although their signature was validated in prospective clinical trials, Smith signature has not been established in clinical routine, probably due to the necessity of conducting at least 34 multiplex analytics hindering simple and rapid dissemination as POCT. Interestingly, a group at Fudan University analyzed parts of the same datasets that we assessed in this project and identified a five-gene signature that predicts metastasis spread (22). COL10A1 did not come up as their top candidate suggestion, urging that interrogation of the Fudan signature in the context of COL10A1-rich CRC is needed. Very recently, Liu et al. proposed a hub gene signature comprised of four candidates, and the corresponding protein accumulation in the tumor material was verified in independent prospective cohort of patients (23). Importantly, as the new research field of cancer neuroscience currently emerges, indication from the experimental field emerges that nerve growth factors mediate liver metastatic potency of CRC cells (24). Further studies to decipher the composition and tumor-relevant roles of the neural microenvironment at the primary tumor site or in the metastasis site of intestinal tumors are needed and surely will reveal new insights in the diseases. As such, our group recently identified the hitherto unrecognized clinical prognostic role activation levels of sensory nerve chancel of the transient receptor potential channel (TRPC) class in pancreatic cancer (25), a discovery relevant also for developing new therapies for this deadly disease, as members of TRPC are druggable targets with clinical applied inhibitors.

We acknowledge that our work is of a descriptive nature only, and our assumptions are made based on correlative findings. We consider our results to be relevant for the field: our study relies on reusing various publically available, high-quality molecular data from larger patient population retrieved from different datasets that all have been quality approved by the scientific community. It reflects and discriminates regarding gender and ethnic diversity. It is based on current sequencing technologies and molecular tumor diagnostic data, and in our data analysis, we appreciate the importance of the emerging field of intra-tumor spatial heterogeneity to instruct the biological behavior of the disease. As COL10A1-enriched tumors feature increased immune cell infiltration and extracellular matrix components, we assume that CRC COL10A1 activation might either modulate the tumor microenvironment, or vice versa; its expression is a downstream signal of altered immune and stromal environmental interactions. Of particular interest is the very strong correlation of elevated COL10A1 transcription with CAFs, as those pool of cell populations are emerging as modulators of establishing a pro-invasive tumor microenvironment. Functional studies to address this questions, particularly using human model systems that recapitulate cellular and spatial heterogeneity as achieved in patient-derived organoids (26), are underway in our lab. Of note, all of our lab-tested classical CRC cell models show a low abundance of COL10A1 protein expression in Figure 7, which in part reflects the results of the cell line transcription data. However, in our view, it urges the assessment of COL10A1 in clinically more relevant 3D model systems featuring the stroma microenvironment (27). It will be interesting to compare the mRNA/protein levels of COL10A1 in the primary tissue with matching personalized 3D *in vitro* models and study effects of COL10a1 modulation in such conditions. Of note, although not retrieved from orthotopic condition and also resembling high *in vitro* passage model, the tested fibroblasts are high in COL10A protein. Confirmatory studies in patient-matched tumor cell/CAFs co-culture systems are needed to analyze COL10A1 protein/DNA as component of the lateral information system between tumor cells and stroma environment.

The described results further established COL10A1 as a diagnostic marker for predicting progression of colon carcinogenesis, extending previous reports on this protein in the context of colon cancer. The first mentioning of COL10A1 to be specifically upregulated in CRC as compared to normal mucosa related back to Croner et al. in the year 2005 (14). After that, the notable report by Huang et al. described the upregulation of COL10A1 compared to the control tissue in 30 patients (28). Moreover, using protein-based quantification of COL10A1 in tumor specimens based on histological staining and semi-quantitative signal quantification in 197 CRC patients, they identified the significant clinical negative prognostic value of an elevation of the biomarker. Furthermore, a historic study has already proposed COL10A1 serum protein levels to be a minimally invasive and indicative marker for colon cancer detection as compared to its absence in healthy patients (12). It would be interesting to investigate if blood serum levels of COL10A1 protein share a similar prognostic value regarding the metastatic spread of the primary disease as compared to its mRNA abundance in tumor specimens, and to perform a confirmatory study on tumor detection like that reported by Solé et al. In addition, using machine learning algorithm and advance materials to discover the potential value of these gene is also a promising research topic (29, 30). Our data support the initiation of a relevant prospective clinical study to assess COL10A1 expression in tumors aiming to improve the management of colon cancer patients with enlarged lymph node, either by stratifying patient cohorts who do not need to receive neoadjuvant chemotherapy or minimizing the number of patients that require more comprehensive surgical attempt of D3 lymph node resection.

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

RC, AP, and UK: Conceptualization, data curation, formal analysis, roles/writing—original draft, writing—review and editing. RC, UK, WS, MS, LS, TW, MD, MM, and AB: Roles/writing—original draft. RC, AP, and UK: Funding acquisition, methodology, project administration, resources, supervision. All authors contributed to the article and approved the submitted version.

Acknowledgments

UK thanks M Stuerzl, University of Erlangen for stimulating discussion on the phone.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1007514/full#supplementary-material

SUPPLEMENTARY FIGURE 1
COL10A1 expression levels. COL10A1 expression is high in CAF cell cluster **(A, B)** as well as in fibroblasts **(C)**.

SUPPLEMENTARY FIGURE 2
Kaplan-Meier curves of COL10A1 expression levels in groups with „high" and „low" expression on OS **(A)**, PFI**(C)**, DSS(E), ROC curve for validation of Kaplan-Meier curves for OS **(B)**, PFI **(D)**, DSS **(F)**.

References

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

2. Zhou ZX, Zhao LY, Lin T, Liu H, Deng HJ, Zhu HL, et al. Long-term oncologic outcomes of laparoscopic vs open surgery for stages II and III rectal cancer: A retrospective cohort study. *World J Gastroenterol* (2015) 21(18):5505–12. doi: 10.3748/wjg.v21.i18.5505

3. Sato H, Kotake K, Sugihara K, Takahashi H, Maeda K, Uyama I, et al. Clinicopathological factors associated with recurrence and prognosis after R0 resection for stage IV colorectal cancer with peritoneal metastasis. *Dig Surg* (2016) 33(5):382–91. doi: 10.1159/000444097

4. Bipat S, Glas AS, Slors FJM, Zwinderman AH, Bossuyt PMM, Stoker J. Rectal cancer: Local staging and assessment of lymph node involvement with endoluminal US, CT, and MR imaging–a meta-analysis. *Radiology* (2004) 232(3):773–83. doi: 10.1148/radiol.2323031368

5. Puli SR, Reddy JBK, Bechtold ML, Choudhary A, Antillon MR, Brugge WR. Accuracy of endoscopic ultrasound to diagnose nodal invasion by rectal cancers: A meta-analysis and systematic review. *Ann Surg Oncol* (2009) 16(5):1255–65. doi: 10.1245/s10434-009-0337-4

6. Dighe S, Purkayastha S, Swift I, Tekkis PP, Darzi A, A'Hern R, et al. Diagnostic precision of CT in local staging of colon cancers: A meta-analysis. *Clin Radiol* (2010) 65(9):708–19. doi: 10.1016/j.crad.2010.01.024

7. Nerad E, Lahaye MJ, Maas M, Nelemans P, Bakers FCH, Beets GL, et al. Diagnostic accuracy of CT for local staging of colon cancer: A systematic review and meta-analysis. *AJR Am J Roentgenol* (2016) 207(5):984–95. doi: 10.2214/AJR.15.15785

8. Horvat N, Raj A, Liu S, Matkowskyj KA, Knezevic A, Capanu M, et al. CT colonography in preoperative staging of colon cancer: Evaluation of FOxTROT inclusion criteria for neoadjuvant therapy. *AJR Am J Roentgenol* (2019) 212(1):94–102. doi: 10.2214/AJR.18.19928

9. Hong EK, Landolfi F, Castagnoli F, Park SJ, Boot J, Van den Berg J, et al. CT for lymph node staging of colon cancer: Not only size but also location and number of lymph node count. *Abdom Radiol N Y* (2021) 46(9):4096–105. doi: 10.1007/s00261-021-03057-0

10. *Leitlinienprogramm onkologie (Deutsche krebsgesellschaft, deutsche krebshilfe, AWMF): S3-leitlinie kolorektales karzinom, langversion 2.1, 2019, AWMF registrierungsnummer: 021/007OL* (2019). Available at: http://www.leitlinienprogramm-onkologie.de/leitlinien/kolorektales-karzinom/.

11. Cheong CK, Nistala KRY, Ng CH, Syn N, Chang HSY, Sundar R, et al. Neoadjuvant therapy in locally advanced colon cancer: A meta-analysis and systematic review. *J Gastrointest Oncol* (2020) 11(5):847–57. doi: 10.21037/jgo-20-220

12. Solé X, Crous-Bou M, Cordero D, Olivares D, Guinó E, Sanz-Pamplona R, et al. Discovery and validation of new potential biomarkers for early detection of colon cancer. *PloS One* (2014) 9(9):e106748. doi: 10.1371/journal.pone.0106748

13. Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC, et al. TISIDB: an integrated repository portal for tumor–immune system interactions. *Wren J editor Bioinf* (2019) 35(20):4200–2. doi: 10.1093/bioinformatics/btz210

14. Croner RS, Foertsch T, Brueckl WM, Guenther K, Siebenhaar R, Stremmel C, et al. Common denominator genes that distinguish colorectal carcinoma from normal mucosa. *Int J Colorectal Dis* (2005) 20(4):353–62. doi: 10.1007/s00384-004-0664-7

15. Chapman KB, Prendes MJ, Sternberg H, Kidd JL, Funk WD, Wagner J, et al. COL10A1 expression is elevated in diverse solid tumor types and is associated with tumor vasculature. *Future Oncol Lond Engl* (2012) 8(8):1031–40. doi: 10.2217/fon.12.79

16. Gunter MJ, Alhomoud S, Arnold M, Brenner H, Burn J, Casey G, et al. Meeting report from the joint IARC-NCI international cancer seminar series: A focus on colorectal cancer. *Ann Oncol Off J Eur Soc Med Oncol* (2019) 30(4):510–9. doi: 10.1093/annonc/mdz044

17. ltd R and M. *Colorectal cancer (CRC) drugs/ therapeutic market size, share & trends analysis report by drug class (Chemotherapy, immunotherapy), by country (U.S., U.K., Germany, Spain, Italy, France, Japan), and segment forecasts, 2016 -*

2022 (2022). Available at: https://www.researchandmarkets.com/reports/4599566/colorectal-cancer-crc-drugs-therapeutic-market.

18. Włodarczyk M, Włodarczyk J, Trzciński R, Mik M, Dziki Ł, Dziki A. *D3 lymphadenectomy for right colon cancer - włodarczyk - annals of laparoscopic and endoscopic surgery* (2022). Available at: https://ales.amegroups.com/article/view/5424/html.

19. Stein U, Walther W, Arlt F, Schwabe H, Smith J, Fichtner I, et al. MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis. *Nat Med* (2009) 15(1):59–67. doi: 10.1038/nm.1889

20. Radhakrishnan H, Walther W, Zincke F, Kobelt D, Imbastari F, Erdem M, et al. MACC1-the first decade of a key metastasis molecule from gene discovery to clinical translation. *Cancer Metastasis Rev* (2018) 37(4):805–20. doi: 10.1007/s10555-018-9771-8

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

22. Zhou Y, Zang Y, Yang Y, Xiang J, Chen Z. Candidate genes involved in metastasis of colon cancer identified by integrated analysis. *Cancer Med* (2019) 8 (5):2338–47. doi: 10.1002/cam4.2071

23. Liu S, Zhang Y, Zhang S, Qiu L, Zhang B, Han J. Identification of hub genes related to liver metastasis of colorectal cancer by integrative analysis. *Front Oncol* (2021) 11:714866. doi: 10.3389/fonc.2021.714866

24. Lei Y, He X, Huang H, He Y, Lan J, Yang J, et al. Nerve growth factor orchestrates NGAL and matrix metalloproteinases activity to promote colorectal cancer metastasis. *Clin Transl Oncol Off Publ Fed Span Oncol Soc Natl Cancer Inst Mex* (2022) 24(1):34–47. doi: 10.1007/s12094-021-02666-x

25. Shi W, Li C, Wartmann T, Kahlert C, Du R, Perrakis A, et al. Sensory ion channel candidates inform on the clinical course of pancreatic cancer and present potential targets for repurposing of FDA-approved agents. *J Pers Med* (2022) 12 (3):478. doi: 10.3390/jpm12030478

26. Sasaki N, Clevers H. Studying cellular heterogeneity and drug sensitivity in colorectal cancer using organoid technology. *Curr Opin Genet Dev* (2018) 52:117–22. doi: 10.1016/j.gde.2018.09.001

27. Reidy E, Leonard NA, Treacy O, Ryan AE. A 3D view of colorectal cancer models in predicting therapeutic responses and resistance. *Cancers* (2021) 13(2): E227. doi: 10.3390/cancers13020227

28. Huang H, Li T, Ye G, Zhao L, Zhang Z, Mo D, et al. High expression of COL10A1 is associated with poor prognosis in colorectal cancer. *OncoTargets Ther* (2018) 11:1571–81. doi: 10.2147/OTT.S160196

29. Liu L, Chen X, Petinrin OO, Zhang W, Rahaman S, Tang ZR, et al. Machine learning protocols in early cancer detection based on liquid biopsy: A survey. *Life* (2021) 11(7):638. doi: 10.3390/life11070638

30. Xiang Y, Peng X, Kong X, Tang Z, Quan H. Biocompatible AuPd@ PVP core-shell nanoparticles for enhancement of radiosensitivity and photothermal cancer therapy. *Colloids Surfaces A: Physicochemical Eng Aspects* (2020) 594:124652. doi: 10.1016/j.colsurfa.2020.124652

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 12 October 2022
DOI 10.3389/fonc.2022.999654

OPEN ACCESS

EDITED BY
Mingyue Tan,
Shanghai University of Traditional Chinese Medicine, China

REVIEWED BY
Xiaoping Yang,
Hunan Normal University, China
Mei-hua Bao,
Changsha Medical University, China

*CORRESPONDENCE
Jing Wang
wangjing0081@hnca.org.cn
Nayiyuan Wu
wunayiyuan@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 21 July 2022
ACCEPTED 27 September 2022
PUBLISHED 12 October 2022

CITATION
Li H, Liu Z-Y, Chen Y-C, Zhang X-Y, Wu N and Wang J (2022) Identification and validation of an immune-related lncRNAs signature to predict the overall survival of ovarian cancer. *Front. Oncol.* 12:999654. doi: 10.3389/fonc.2022.999654

Identification and validation of an immune-related lncRNAs signature to predict the overall survival of ovarian cancer

He Li[1,2†], Zhao-Yi Liu[2†], Yong-Chang Chen[2], Xiao-Ye Zhang[2], Nayiyuan Wu[2*] and Jing Wang[2,3*]

[1]The Animal Laboratory Center, Hunan Cancer Hospital and The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China, [2]The Central Laboratory, Hunan Cancer Hospital and The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China, [3]Department of Gynecologic Cancer, Hunan Cancer Hospital and The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China

Ovarian cancer (OC) is the most lethal gynecological cancer in women. Studies had reported that immune-related lncRNAs signatures were valuable in predicting the survival and prognosis of patients with various cancers. In our study, the prognostic value of immune-related lncRNAs was investigated in OC patients from TCGA-RNA-seq cohort (n=378) and HG-U133_Plus_2 cohort (n=590), respectively. Pearson correlation analysis was implemented to screen the immune-related lncRNA and then univariate Cox regression analysis was performed to explore their prognostic value in OC patients. Five prognostic immune-related lncRNAs were identified as prognostic lncRNAs. Besides, they were inputted into a LASSO Cox regression to establish and validate an immune-related lncRNA prognostic signature in TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort, respectively. Based on the best cut-off value of risk score, patients were divided into high- and low-risk groups. Survival analysis suggested that patients in the high-risk group had a worse overall survival (OS) than those in the low-risk group in both cohorts. The association between clinicopathological feathers and risk score was then evaluated by using stratification analysis. Moreover, we constructed a nomogram based on risk score, age and stage, which had a strong ability to forecast the OS of the OC patients. The influence of risk score on immune infiltration and immunotherapy response were assessed and the results suggested that patients with high-risk score might recruit multiple immune cells and stromal cells, leading to facilitating immune surveillance evasive. Ultimately, we demonstrated that the risk model was associated with chemotherapy response of multiple antitumor drugs, especially for paclitaxel, metformin and veliparib, which are commonly used in treating OC patients. In conclusion, we constructed a novel immune-related lncRNA signature, which had a potential prognostic value for OC patients and might facilitate personalized counselling for immunotherapy and chemotherapy.

KEYWORDS

ovarian cancer, immune-related lncRNA, prognostic signature, chemotherapy, risk score

Introduction

Ovarian cancer (OC) is the most lethal gynecological cancer among women in worldwide, with 313,959 estimated new cases and 207,252 new deaths in 2020 (1). Due to the ambiguity of early symptoms and the lack of reliable screening strategies, more than 60% OC patients are diagnosed with later-stage. Complete cytoreductive surgery followed by platinum-based chemotherapy is known as the standard first-line treatment protocol for OC patients. However, a high proportion patient will relapse within 2 years of diagnosis (2). Therefore, there is an urgent need to identify prognostic biomarkers to predict the outcome of OC patients.

It is being increasingly recognized that immune system plays vital roles during cancer initiation and progression (3). Moreover, it is suggested that tumor progression and invasion is dependent on intratumoural adaptive immunity and the immunological type, density, and location of immune cells within the tumor samples are superior to TNM staging in predicting the natural history of primary cancers (4, 5). It has been reported that patients whose tumors with more tumor-infiltrating lymphocytes (TILs) predicted longer survival in OC. Besides, recruitment of T-regulatory (Treg) cells in OC can foster immune privilege and predict reduced OS (6, 7). All the evidence convincingly indicated that OC was an immunogenic tumor (8). Therefore, the immune-related prognostic signature might be a potential tool to predict outcome of OC patients.

Long non-coding RNAs (lncRNAs) are a group of RNA molecules whose transcripts are greater than 200nt but not translated into proteins. They participate in various biological progress, such as epigenetic regulation, genetic imprinting, chromatin organization and protein modification (9, 10). Moreover, they participate in immune response including antigen presentation, antigen release, immune cell differentiation and T cells infiltration (11, 12). Lnc-EGFR stimulates Treg cells differentiation and promotes immune invasion in hepatocellular carcinoma (13). Lnc-DC, which is a specific marker of dendritic cells (DCs), promotes the ability of DCs to active T cells (14). LincR-Ccr2-5'AS increases the migration ability of Treg cells (15).

In OC, a new lncRNA small nucleolar RNA host gene 12 (SNHG12) was proved to promote immune escape of OC cells through their crosstalk with M2 macrophages (16). Moreover, lncRNA HOTTIP was suggested to promote the secretion of IL-6 and up-regulate the expression of PD-L1 in neutrophils, leading to the inhibition of T cells activity and acceleration immune escape of OC cells (17). Recently, it was demonstrated that lncRNA XIST could affect the cell proliferation and migration *via* mediating macrophage polarization in both breast cancer and OC (18). In addition, FOXP4-AS1 and MEG8 were revealed to be associated with immune infiltration in OC (19, 20). All these evidences indicated that immune-related lncRNAs played important roles in OC.

Recently, multiple immune-related lncRNA signatures have been identified to predict the OS in various cancers, including breast cancer (21–24), hepatocellular cancer (25), lung cancer (26), cervical cancer (22, 27), colon cancer (28), glioma (29–31), and bladder cancer (32, 33). However, the immune-related prognostic lncRNA signature for predicting the prognosis of OC patients has not been developed. In our study, we aimed to explore the prognostic value of the immune-related lncRNAs in OC and validate an immune-related prognostic lncRNA signature for patients with OC.

Materials and methods

Data acquisition and preprocessing

For TCGA-RNA-Seq training set, mRNA gene expression profiles and corresponding clinical information were downloaded from the TCGA data source (https://xena.ucsc.edu). To increase the statistical power and overcome the systematic errors caused by small sample size, we combined the datasets (GSE26193, GSE30161, GSE63885, GSE9891, GSE18520 and GSE19829) with the HG-U133_Plus_2 platform as the HG-U133_Plus_2 validation set (34–39). All clinical information and microarray data were captured from GEO repository (https://www.ncbi.nlm.nih.gov/geo/). Ultimately, we obtained a TCGA-RNA-seq training cohort with 378 patients and a HG-U133_Plus_2 validation cohort with 590 patients.

Identification of immune-related lncRNAs

The lncRNA annotation file was acquired from the GENCODE website for annotation of the lncRNAs. Consequently, 14826

lncRNAs and 2448 lncRNAs were identified from TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort, respectively (40). The immune-related genes were obtained from the ImmPort database (http://www.immport.org) (41). Pearson correlation analysis was utilized to screen immune-related lncRNAs. Those lncRNAs with r>0.3 and $p<0.001$ were considered as immune-related lncRNAs (25). To assess the prognostic value of immune-related lncRNAs, we further conducted univariate Cox regression analysis by using the "survival" package, and the hazard ratios (HR) with 95% confidence intervals (CIs) were examined. $p < 0.05$ was considered that immune-related lncRNAs were significantly correlated with OS) and served as prognostic immune-related lncRNAs.

OS analysis

OS was defined as the time from randomization to death from any cause. The survival curves were calculated and illustrated by the KM plot with the long-rank test.

Construction of immune-related prognostic lncRNA signature

Based on the prognostic immune-related lncRNAs, a risk signature was constructed by using the "glmnet" package (42). Through 1000 cross-validation, a panel of genes and their LASSO coefficients were obtained. The risk scores for the signature were calculated using the following formula: Risk score=β1X1+β2X2+⋯+βnXn (β, LASSO coefficient; X, the expression of each prognostic immune-related lncRNA in each sample). Based on the best cut-off value of risk score, patients were divided into high-risk and low-risk groups. Kaplan–Meier method with the long-rank test were performed to reveal the OS of the high-risk and low-risk groups by using the "survival" package. Besides, time-dependent relative operating characteristic (ROC curve) and area under the curve (AUC) were applied to assess the prediction ability of the signature. All the time-dependent ROC curves were calculated and drew by "SurvivalROC" and "ggplot2" package, respectively.

Decision tree and prognostic nomogram construction

Decision tree and nomogram model were applied to define significant clinical predictors. Firstly, univariate and multivariate COX regression were performed to select important explanatory variables. Based on the multivariate cox regression results, stage, age and risk score were identified as predictor variables. After then, the "rpart" Package (https://cran.r-project.org/web/packages/tree/index.html) was used to construct decision tree and split patients as different from each other as possible. It was implemented to decide which of these variables to split and the splitting value in each step of the tree's construction (43). Moreover, a nomogram model, which is an individualized risk prediction model to predict the 1, 3, 5-year survival probability, was constructed using the "RMS" package. The calibration curves were used to assess the concordance of the observed and predicted rates of 1, 3, 5-year OS (44).

Estimation of tumor-infiltration, immunotherapy and chemotherapy response

Firstly, the ESTIMATE algorithm (https://bioinformatics.mdanderson.org/public-software/estimate/), which can be applied for assessment of the presence of stromal cells and the infiltration of immune cells in tumor samples using gene expression data, was used to calculate the Estimate score, Immune score, Purity score and stromal score (45). Briefly, we defined ssGSEA based on the signatures related to stromal tissue and immune cell infiltration as Stromal score and Immune score, respectively, and combined the stromal and immune scores as the ESTIMATE score. Purity score was calculated as followed: Purity score= cos (0.6049872018 + 0.0001467884*ESTIMATE score). The correlation of risk score and Estimate score, Immune score, Purity score and stromal score were analyzed by using Pearson correlation analysis. The infiltration of 22 subtypes of tumor-infiltrating immune cells (TIICs) was acquired from CIBERSORT algorithm (http://cibersort.stanford.edu/) (46). Tumor Immune Disfunction and Exclusion (TIDE) algorithm (http://tide.dfci.harvard.edu/), which is a method to accurately predict the outcome of patients treating with immune checkpoint blockade (ICB), were employed to evaluate the immunotherapy response (47, 48). The chemotherapy response was evaluated by using the Genomics of Drug Sensitivity in Cancer database (GDSC, https://www.cancerrxgene.org). The half-maximal inhibitory concentration (IC50) of all drugs commonly used to treat tumors were calculated and represented the drug response. The R package 'pRRopheticRredic' was used with 10fold cross-validation and other parameters by default (49).

Exploration of immune-related lncRNA function

To further explore the function of the five immune-related lncRNA, we firstly assessed the association between the five immune-related lncRNA and immune-related mRNA by using Pearson correlation analysis. Then, the results were converted visually and the co-expression network was identified with Cytoscape software (50). Based on gene expression or the risk score, patients were divided into two groups. GSEA assay was utilized to explore whether a series of priori defined pathways were

enriched in the gene bank derived from DEGs between the two groups (51, 52). FDR<0.05 was identified as enriched. Moreover, the absolute immune scores from gene expression datasets were obtained by LM22 (22 immune cell types) gene signatures of CIBERSORT algorithm (46). The molecular immune cell subtypes related to the five lncRNAs expression were captured by using Spearman correlation analysis (53). Only $p<0.05$ was considered significant.

Cell culture, RNA extraction and real-time quantitative PCR

OC cell lines, SKOV3, A2780, OVCAR8 and OVCAR3, were obtained from Institute of clinical pharmacology, Central South University. All the cell lines were cultured in RPMI-1640 medium with 10% FBS. All cell lines were cultivated at 37°C and 5% CO_2. Total RNAs were extracted from OC cell lines by using Trizol reagent (Takara). After extraction, total RNAs were reverse-transcribed into cDNA using PrimerScript™ RT reagent Kit (RR047A, Takara). Real-time quantitative PCR was performed using the SYBER Premix Ex Taq kit (RR420a, Takara) in Roche-LightCycler 480 system (Roche,USA). Finally, the relative expression of lncRNAs were calculated based in the internal reference GAPDH. The primers of lncRNAs and GAPDH are listed in supplementary Table 1.

Lentivirus infection

The packaged lentivirus vectors of UBXN10-AS1 overexpression (LV-BUXN10-AS1) and empty lentivirus vectors (LV-NC) were purchased from GenePharm (Shanghai, China). For UBXN10-AS1 overexpression, the LV-UBXN10-AS1 or LV-NC were introduced in SKOV3 and A2780 cells at an MOI OF 50-100. After 72h post-infection, the infection efficiency was measured by using RT-qPCR.

Cell proliferation assays

Cell proliferation was assessed by using CCK8 kit (MCE, China). Briefly, cells (1-2 $*10^4$ cells/well) infected with LV-UBXN10-AS1 or LV-NC were seeded into 96-well plates and cultured in a CO_2 incubator for 24,36, 48,72 and 96h. Subsequently, 10μl of CCK8 reagent was added into the wells and the plate was incubated for 1h. Finally, the OD value was measured at 450nm using the microplate reader (54).

Cell migration assay

To detect the cell migration, wounding healing assay was performed. Firstly, cell infected with LV-UBXN10-AS1 or LV-NC were seeded in 24-well plates. A 200ul pipette tip was used to scratch the cell layer, when cells reached 70-80% confluence. Cells were grown for an additional 48h. Microscope images were captured at 0h and 48h (55).

Annexin V/PI apoptosis assay

Cells were plated in 6-well plate with $1*10^5$ cells/well. After 12 hours, cells were infected with lentivirus vectors. After 48 hours incubation, cells were harvested, washed with PBS and incubated with Annexin V and PI, using the Annexin V-APC apoptosis detection kit (KGA1022, KeyGen, China). The flow cytometry analyses were performed with CytoFLEX instrument (Beckman Coulter, USA).

Statistical analysis

The two-tailed Students' t-test was utilized to analyze the significant differences between groups, whereas quantitative differences among groups were analyzed by using the one-way ANOVA. Kaplan–Meier curves and log-rank test were implemented to calculate the OS rate. All statistical analyses were performed using R software (version 3.6.2). * means $p<0.05$, ** means $p<0.01$, ***means $p<0.001$. $p<0.05$ was considered statistically significant.

Results

Identification of immune-related prognostic lncRNAs in OC patients

As shown in Figure 1, we firstly identified 14826 lncRNAs in the TCGA-RNA-seq dataset and 2448 lncRNAs in the HG-U133_Plus_2 dataset, based on the lncRNA annotation file from GENCODE website. Then, the immune-related genes were download from the ImmPort database. Pearson correlation analysis was performed to screen the immune-related lncRNAs. The immune-related lncRNAs were identified as that the expression of lncRNAs were correlated with one or more of the immune-related genes ($|\text{cor}| > 0.3$ and $p < 0.001$). Finally, we obtained 1637 immune-related lncRNAs in TCGA-RNA-seq dataset and 1814 immune-related lncRNAs in the HG-U133_Plus_2 dataset, respectively (Supplement Table 2). To screen immune-related prognostic lncRNAs, the univariate Cox regression was implemented. The forest plot showed that 5 lncRNAs (UBXN10-AS1, TOPORS-AS1, HIPK1-AS1, CELSR3-AS1 and CECR5-AS1) were significantly correlated to prognosis of patients with ovarian cancer. All the lncRNAs were protective factors with hazard ratio (HR) <1 in both datasets (Figure 2A). The Kaplan–Meier curves confirmed that higher

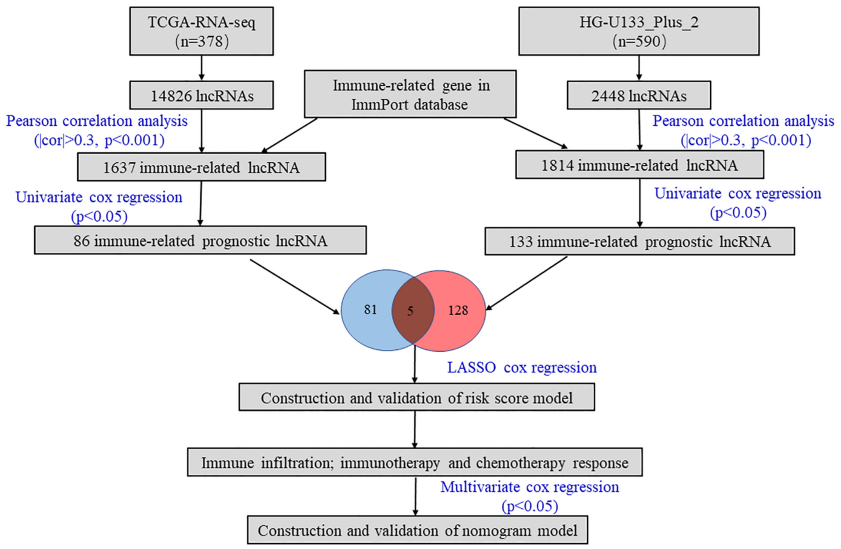


FIGURE 1
The workflow of this study.

expression of all the five lncRNAs were associated with better OS in both cohorts (Figure 2B).

Construction and validation of the immune-related lncRNA signature

In addition, we defined the TCGA-RNA-Seq dataset as discovery cohort and constructed an immune-related lncRNA signature. The risk score for each patient was calculated based on the coefficient for each lncRNA (Supplement Figure 1).

Subsequently, patients were divided into two subgroups dependent on the best cut-off value of risk score. The distributions of the risk score and survival status were listed in the Figure 3A. The heatmap showed that the expression of all the lncRNAs were higher in the low-risk group than in the high-risk group (Figure 3B). Kaplan-Meier survival curves indicated that patients with higher risk score had worse survival rate ($p<0.001$, Figure 3C). Furthermore, we validated the prognostic value of the immune-related lncRNA signature in the HG-U133_Plus_2 cohort. The results were consistent with the findings in the TCGA-RNA-Seq cohort. It's suggested that the higher risk score was associated with shorter OS time and worse survival status (Figures 3D-F). The ROC curves demonstrated that the immune-related prognostic lncNRA signature harbored a promising ability to predict 5-year OS in the TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort (Figures 3G, H). All these demonstrated that the immune-related prognostic lncRNA signature might stably predict the survival outcome of patients with OC.

Association between the prognostic signature and clinicopathological feathers

We attempted to analyze the association between risk score and the clinicopathological feathers. It was suggested that patients with higher age and advanced FIGO stage had higher risk score, while the risk score was not associated with grade in both cohort (Figures 4A, B). Besides, we assessed the prognostic ability of the immune-related prognostic signature by performing a stratification analysis. Compared to patients with lower risk, patients with higher risk had worse OS in younger (<50y), older (≥50y), advanced FIGO stage (III+IV), early grade (G1+G2) and advanced grade (G3+G4) subgroups in the TCGA-RNA-Seq cohort (Figure 4C). Likewise, these results were validated in the HG-U133_Plus_2 cohort (Figure 4D).

Due to the small samples of the early FIGO stage (I+II) subgroup in TCGA-RNA-Seq cohort, there was no significant difference in OS between higher risk patients and lower risk patients ($p=0.17$, Figure 4C). However, we confirmed that the signature retained the ability to predict OS for patients with early stage in HG-U133_Plus_2 cohort ($p=0.0054$, Figure 4D). All these results revealed that it could be served as a potential predictor for OC patients.

Modeling the prognostic nomogram

Firstly, the independent prognostic factors were identified by using the univariate and multivariant cox regression in the

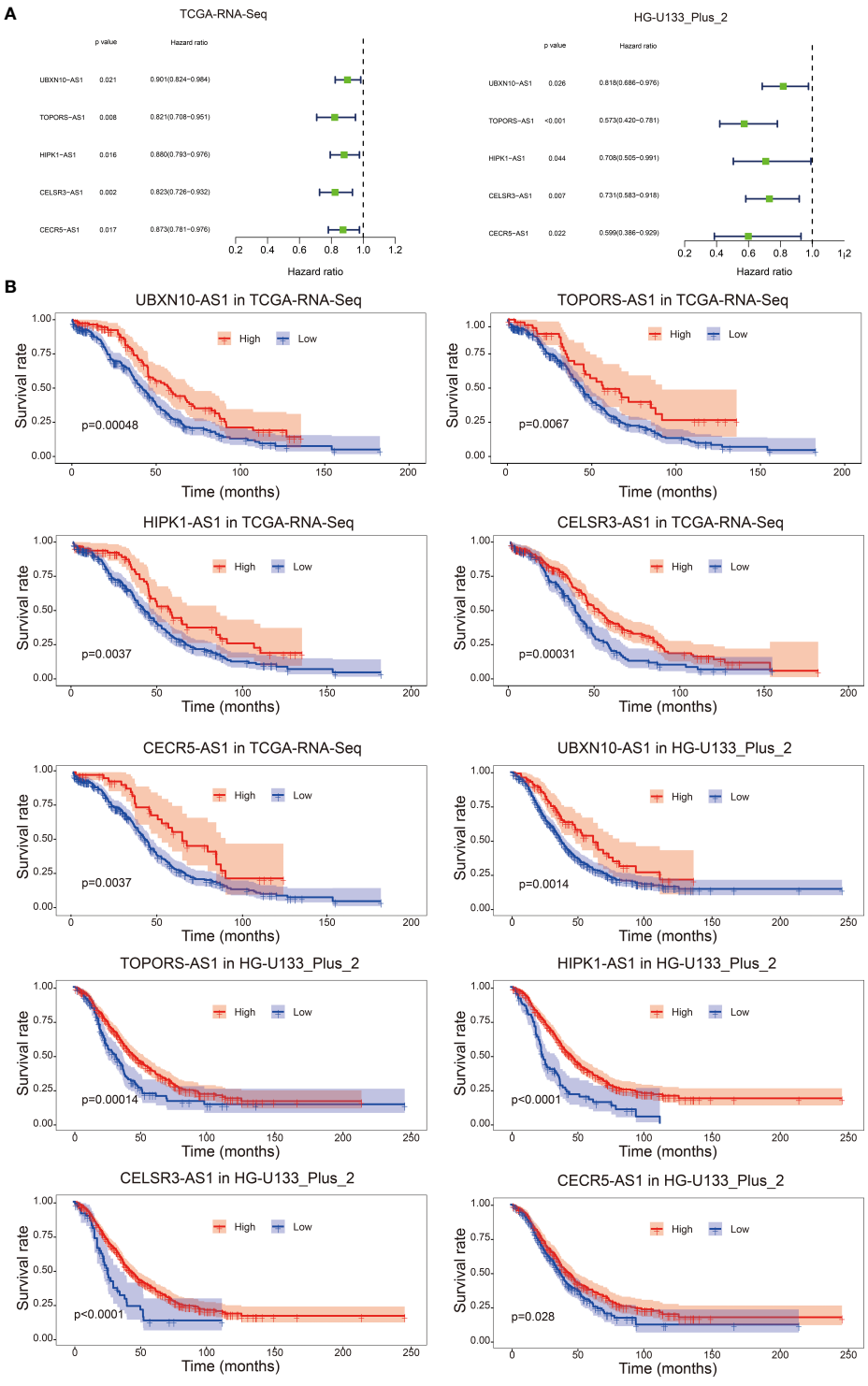


FIGURE 2
Forest plot of the prognostic ability of the five immune-related lncRNAs in TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort **(A)**; Kaplan–Meier curves suggested that expression of the five immune-related lncRNAs were associated with the OS in both TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort **(B)**.

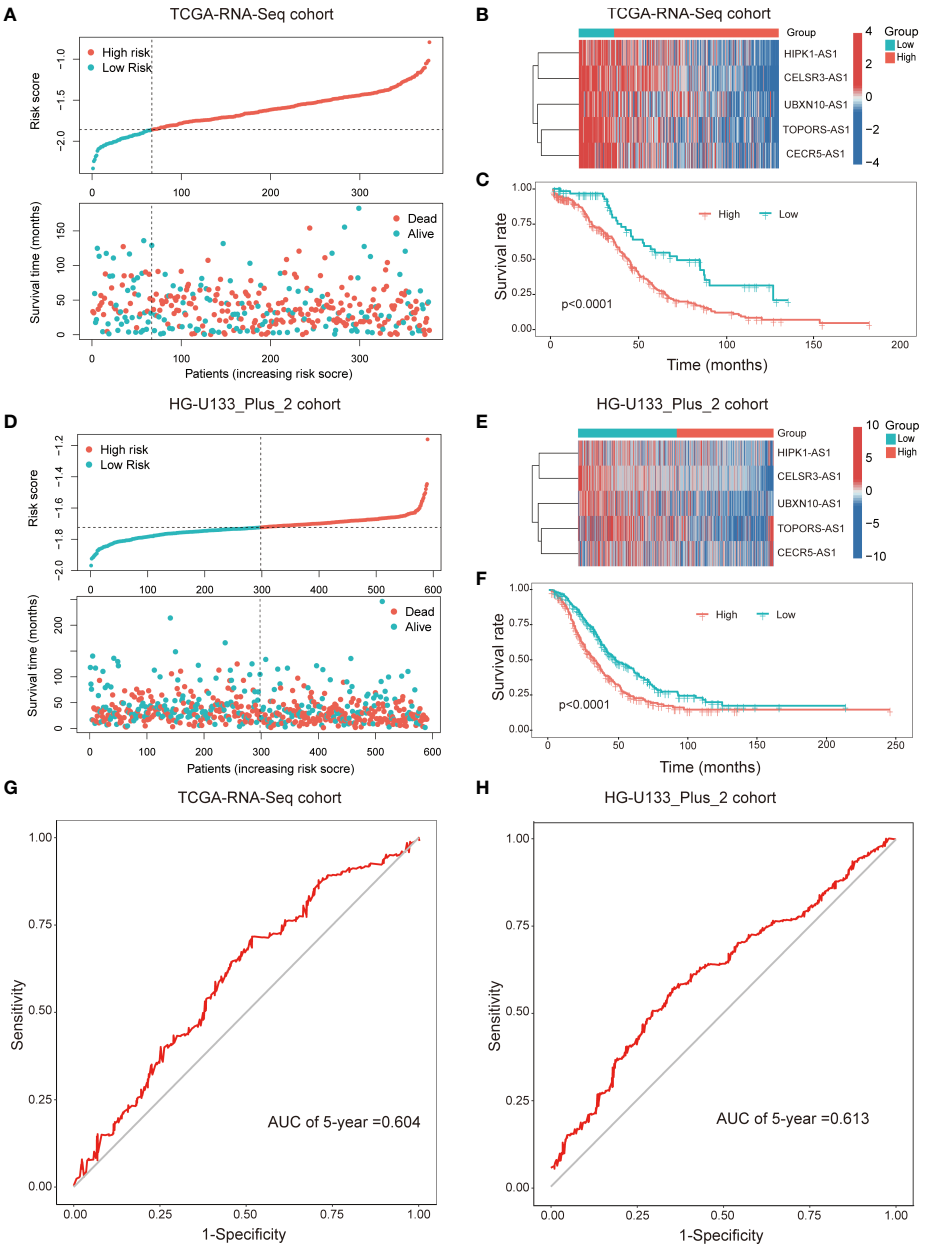


FIGURE 3
The immune-related prognostic signature was established and validated in TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort, respectively. Distributions of risk scores and survival status of OC patients in the TCGA-RNA-Seq cohort **(A)** and HG-U133_Plus_2 cohort **(D)**; Heat map analysis showed the association between risk score and the expression of the five lncRNAs in TCGA-RNA-Seq cohort **(B)** and HG-U133_Plus_2 cohort **(E)**; Kaplan–Meier curves showed that the high-risk subgroup had worse OS than the low-risk subgroup in TCGA-RNA-Seq cohort **(C)** and HG-U133_Plus_2 cohort **(F)**. ROC curves of the immune-related lncRNAs for predicting 5-year survival in TCGA-RNA-Seq cohort **(G)** and HG-U133_Plus_2 cohort **(H)**.

TCGA-RNA-Seq cohort. The univariate cox regression analysis indicated that risk score (HR: 2.971; 95% CI: 1.718-5.136; *p<0.001*), age (HR: 1.022; 95% CI: 1.010-1.035; *p<0.001*), stage (HR: 1.380; 95% CI: 1.032-1.847; *p=0.030*) but not grade (HR: 1.226; 95% CI: 0.828-1.815; *p=0.308*) were associated with OS of patients (Figure 5A). Multivariate cox analysis further proved that risk score (HR: 2.537; 95% CI: 1.443-4.461; *p=0.001*), age (HR: 1.019; 95% CI: 1.007-1.032; *p=0.003*) and stage (HR: 1.377; 95% CI: 1.026-1.849; *p=0.033*) were independent prognostic factors for OC patients (Figure 5A). Therefore, age, FIGO stage and risk score were applied to build a decision tree with five different risk subgroups (Figure 5B). The split at the top of

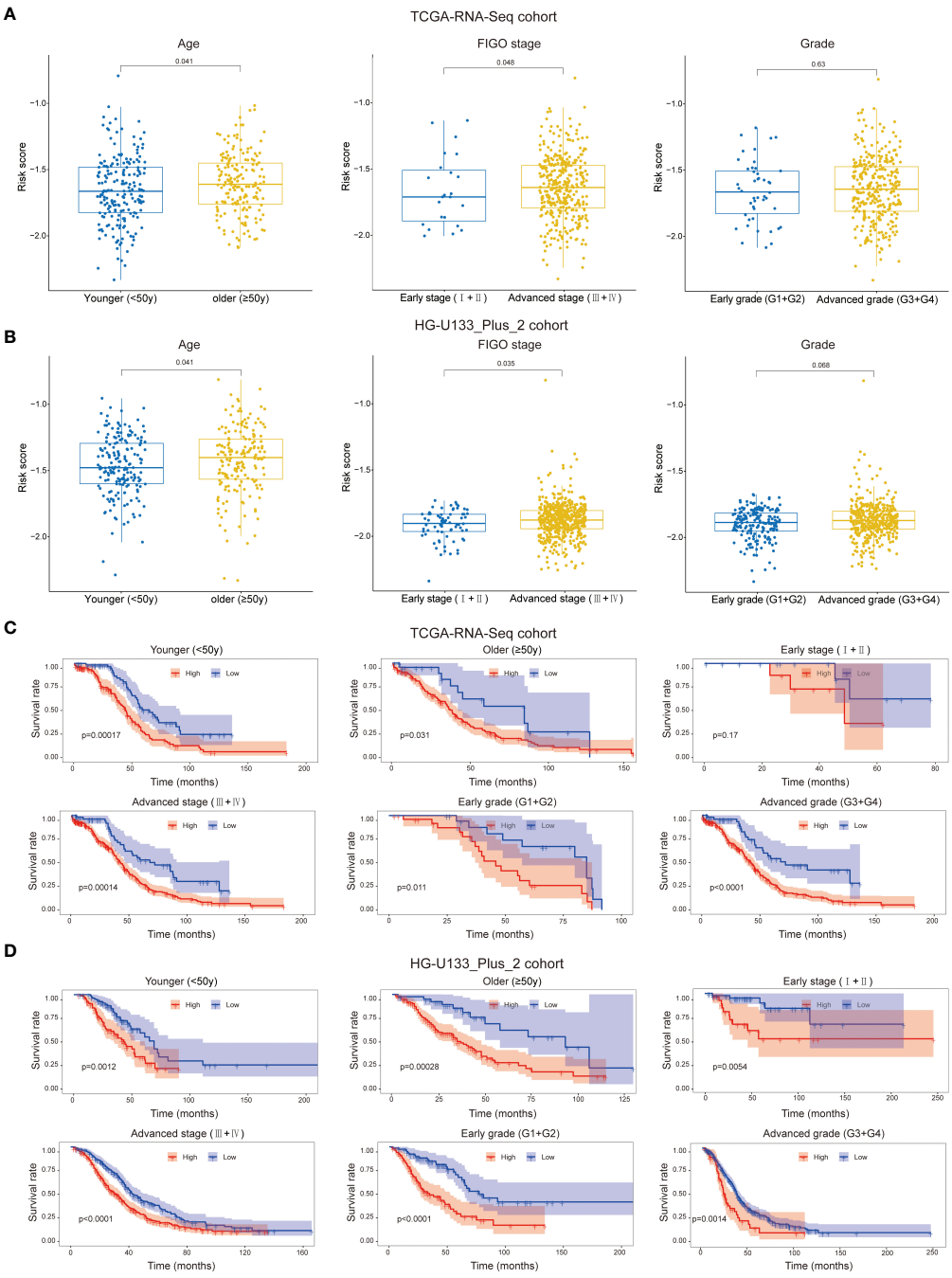


FIGURE 4
Patients with different clinicopathological features (including age, FIGO stage and Grade) had different levels of risk scores in TCGA-RNA-Seq cohort **(A)** and HG-U133_Plus_2 cohort **(B)**. Stratification analysis suggested that the immune-related lncRNAs signature retained its prognostic value in multiple subgroups in TCGA-RNA-Seq cohort **(C)** and HG-U133_Plus_2 cohort **(D)**. The younger and older group were divided based on 50y; FIGO I+II were identified as early stage and FIGO III+IV were identified as advanced stage; G1+G2 were identified as early grade and G3+G4 were identified as advanced grade.

the tree resulted in two large branches: the left-hand branch included patients with early stage; the right-hand branch corresponded to patients with advanced stage. The right branch is further subdivided by age, stage and risk score. Overall, the tree had five terminal nodes, leading to partitioning OC patients in five subgroups. It worth mentioning that compared to patients with younger age (<50y), stage III and high-risk score (31% of overall samples), patients with younger age (<50y), stage III and low-risk score (9% of overall samples) showed higher alive probability (44% *vs* 59%). In order to make the signature more applicable in clinic, a nomogram based on the predictors (including risk score, age and FIGO stage) was established in the TCGA-RNA-Seq cohort (Figure 5C). Calibration plots showed that the observed *vs* predicted rates of 1-, 3- and 5-year OS showed perfect concordance (Figures 5D-F). Moreover, the predictive performance of the nomogram was evaluated by the ROC curve. Compared to other predictors (including age and FIGO stage), the model's 5-year AUC values were higher in both TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort (Supplement Figure 2). KM survival plot analysis showed that patients with high-risk had a worse OS than patients with low-risk subgroup in both TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort ($p<0.001$, $p<0.001$, respectively, Figures 5G, H). These data confirmed that the nomogram had a robust and stable ability to predict the OS for OC patients.

Association between the prognostic signature and immune infiltration and immunotherapy response

To explore the influence of risk score on immune infiltration and immunotherapy response, the ESTIMATEscore, ImmuneScore, PurityScore and StromalScore were calculated to explain immune cell and stromal cell infiltration situation. The correlation analysis results indicated that the risk score was positively correlated with the ESTIMATEscore, ImmuneScore and StromalScore, but negatively correlated PurityScore in TCGA-RNA-Seq cohort (Figures 6A-D). The similar results were validated in the HG-U133_Plus_2 cohort (Supplement Figures 3A-D). After that, the distribution proportion of 22 immune cells in high-risk group and low-risk group were analyzed. In TCGA-RNA-Seq cohort, the distribution proportion of Macrophages cells was higher in high-risk group than low-risk group, whereas the distribution proportion of activated dendritic cells were significantly lower (Figure 6E). In the HG-U133_Plus_2 cohort, not only Macrophages cells and activated dendritic cells but also memory B cells, plasma cells, $CD4^+$ T cells, Treg cells, NK cells, activated mast cells and neutrophils were differently distributed in high-risk group and low-risk group. (Supplement Figure 3E). Besides, the potential response to immunotherapy for each patient was assessed by using the TIDE algorithm. The results suggested that patients with low-risk score were more sensitive to immunotherapy than those with high-risk score in TCGA-RNA-Seq cohort ($p<0.001$, Figure 6F). Taken together, these results indicated that patients with high-risk score might recruit multiple immune cells and stromal cells and facilitate OC immune surveillance evasive.

Analysis the correlation between the risk model and chemotherapy response

Until now, chemotherapy is the main treatment method for OC patients. Therefore, we tried to identify the association between the risk score and chemotherapy response in both TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort (Supplement Table 3). We revealed that a higher risk score was associated with a lower IC50 of chemotherapeutics such as paclitaxel ($p<0.01$), metformin ($p<0.001$) and veliparib ($p<0.001$) in TCGA-RNA-Seq cohort (Figures 6G-I). In HG-U133_Plus_2 cohort, the risk score was also confirmed to be negatively associated with IC50 of paclitaxel ($p<0.05$), metformin ($p<0.001$), and veliparib ($p<0.05$), whereas it was positively associated with the IC50 of cisplatin (Supplement Figures 3F-I), which indicated that the model acted as a potential predictor for chemosensitivity.

Exploration of the five immune-related lncRNA function

To further understand the function of the five immune-related lncRNA, we constructed the co-expression network between the five immune-related lncRNA and immune-related mRNA. As shown in Figure 7A, CELSR3-AS1 and HIPK1-AS1 showed most connections with immune-related mRNAs. Besides, GSEA analysis was performed to further explore and interpret the enrichment results. The annotated top20 pathways were listed in Figures 7B-F. As shown in the bubble charts, all the five lncRNAs, especially TOPORS-AS1, were significantly associated with immune-related pathways. UBXN10-AS1, TOPORS-AS1, CELSR3-AS1 and CECR5-AS1 were significantly associated with chemokine signaling pathway. Except that, TOPORS-AS1, CECR5-AS1 and HIPK1-AS1 participate in antigen processing and presentation. In addition, the associations between lncRNA expression and individual immune cell subtypes were computed by Spearman correlation in TCGA-RNA-Seq cohort and HG-U133_ Plus_2 cohort (Supplement Figures 4A, B). Moreover, there is a significant difference in actin binding, adaptive of immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains, antigen receptor mediated signaling pathway and B cell

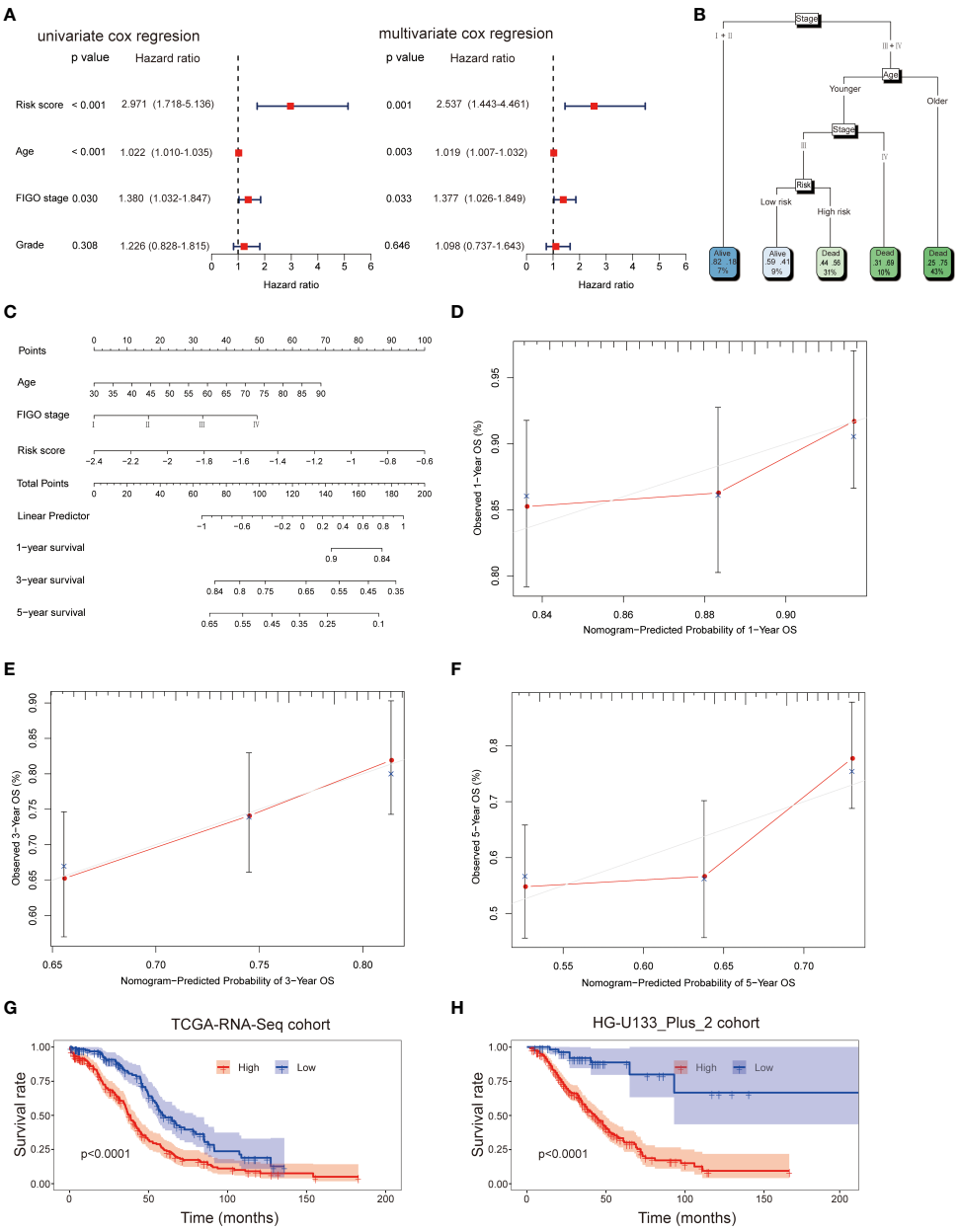


FIGURE 5
Univariate and multivariate analyses revealed that risk score was an independent prognostic predictor in the TCGA-RNA-Seq cohort **(A)**. Construction of decision tree based on risk score, age and stage. The younger and older subgroup were divided based on the median value of age **(B)**. Construction of nomogram based on risk score, age and stage **(C)**. Calibration plots of the nomogram for predicting the probability of OS at 1, 3, and 5-years in the TCGA-RNA-Seq cohort **(D–F)**; KM survival plot analysis showed that patients with high-risk had a worse OS than patients with low-risk subgroup in both TCGA-RNA-Seq cohort and HG-U133_Plus_2 cohort **(G, H)**.

activation between high-risk group and low-risk group in both TCGA-RNA- seq cohort and the HG-U133_Plus_2 cohort (Supplement Figures 4C, D). All these results indicated that the five lncRNAs might affect immune infiltration and facilitate ovarian cancer immune surveillance evasive by regulating immune-related pathways in OC.

Overexpression of UBXN10-AS1 suppressed cell proliferation and migration in OC cell lines

To figure out the function of LncRNAs in OC, the expression of lncRNAs in OC cell lines were detected. Due to the low

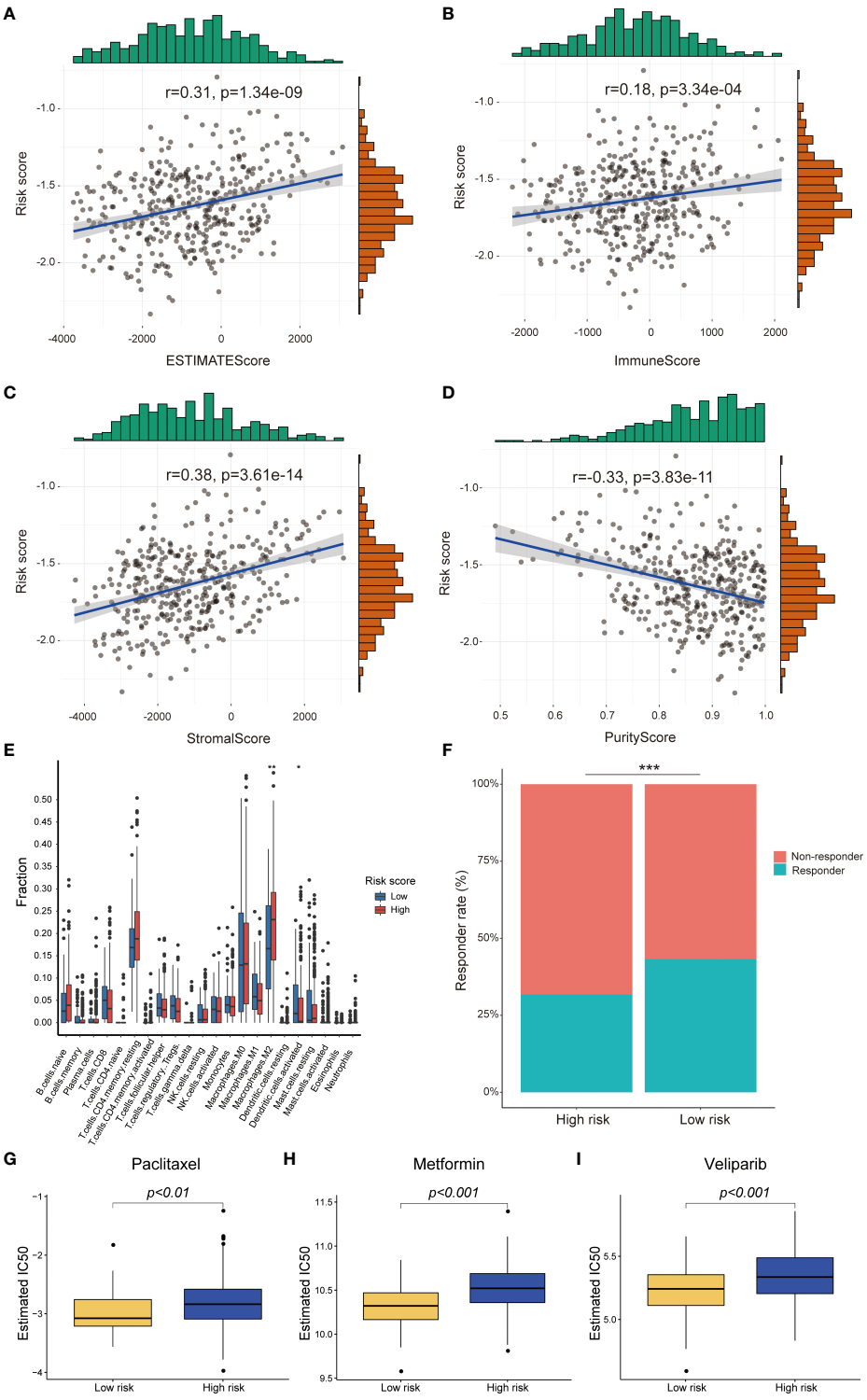


FIGURE 6
Difference between high-risk score group and low-risk score group in immune infiltration, immunotherapy and chemotherapy response prediction in TCGA-RNA-Seq cohort. The risk score was positively correlated with EstimateScore, ImmuneScore, StromalScore and negatively correlated with PurityScore **(A-D)**; The different infiltrated fraction of 22 immune cells between high-risk group and low-risk group **(E)**; The immunotherapy response of patients with OC in high- and low-risk subgroups **(F)**; Estimated IC50 values indicated the chemotherapy response of paclitaxel, metformin and veliparib in TCGA-RNA-Seq cohort **(G-I)**. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

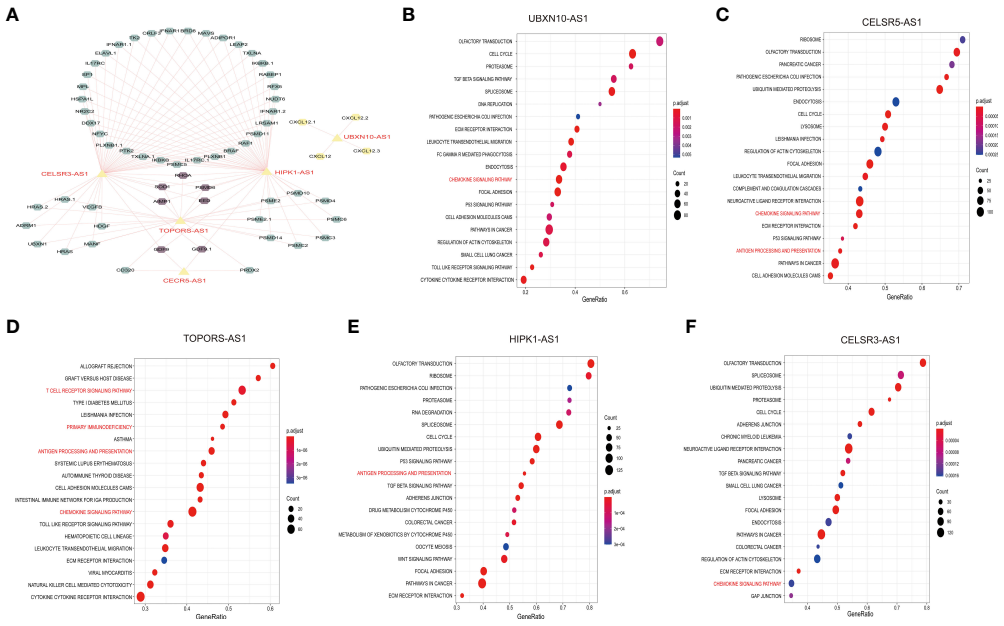


FIGURE 7
Exploration of the five immune-related lncNRAs function. Construction of the co-expression network of the five immune-related lncRNAs and immune-related mRNA **(A)**. GSEA assay to explore the pathways associated with the five immune-related lncRNAs **(B–F)**.

abundance of CELSR3-AS1, CECR5-AS1 and HIPK1-AS1, they were not detected in A2780, SKOV3, OVCAR8 and OVCAR8 cell lines. UBXN10-AS1 were more highly expressed in SKOV3 and A2780 cell lines (Figure 8A). Thus, the function of UBXN10-AS1, as the candidate gene, were further studies in A2780 and SKOV3 cell lines (Figures 8B, C). CCK8 assay revealed that overexpression of UBXN10-AS1 significantly suppressed cell proliferation (Figure 8D). Besides, it could also inhibit the cell migration of SKOV3 and A2780 (Figure 8E). However, UBXN10-AS1 overexpression had no influence on cell apoptosis (Figure 8F). All these results indicated that UBXN10-AS1 might serve as a tumor suppressor in OC.

Discussion

Due to the heterogeneity of OC, it is difficult to blame it on a single specific issue (56). Recently, gene signatures developed by the combination of high-throughput sequencing technology and bioinformatics have been widely used in individualized therapy and prognosis evaluation, which have the better prediction ability than a single biomarker (57). Multiple evidence demonstrated that immune systems made an important contribution to cancer initiation, development, metastasis, and immune escape (58–60). Furthermore, more and more immune-related lncRNAs signatures had been successfully developed and had a perfect prediction accuracy for survival and prognosis in various tumors (61, 62). However, the prediction value of immune-related lncNRAs signature in OC has not been explored.

In our study, we firstly screened immune-related lncRNAs in OC patients from the TCGA-RNA- seq dataset (n = 378) and the HG-U133_Plus_2 dataset (n = 590) by using Pearson correlation analysis. Afterwards, the prognostic significance of immune-related lncRNAs were identified by using univariate cox regression analysis. Finally, five immune-related lncRNAs (including UBXN10-AS1, TOPORS-AS1, HIPK1-AS1, CELSR3-AS1 and CECR5-AS1) were demonstrated to serve as prognostic biomarkers in both TCGA-RNA-seq dataset and the HG-U133_Plus_2 dataset. Recently, it was reported that overexpression of TOPORS-AS1 suppressed cell proliferation and inhibited aggressive cell behaviors, including migration, invasion, and colony formation *via* inhibiting the Wnt/β-catenin pathway in ovarian cancer cells. Moreover, OC patients with high TOPORS-AS1 expression had favorable OS compared to low expression, which was consistent with our study (63). In gastric cancer, it was also proved that the expression of TOPORS-AS1 and its associated gene, NDUFB6 in gastric cancer tissues were significantly lower than that in adjacent tissues (64). All the evidence indicated that TOPORS-AS1 might play important roles in carcinogenesis. Unfortunately, the function of UBXN10-AS1, HIPK1-AS1, CELSR3-AS1 and CECR5-AS1 in OC have not been reported. In colon adenocarcinoma, UBXN10-AS1 was expressed with low

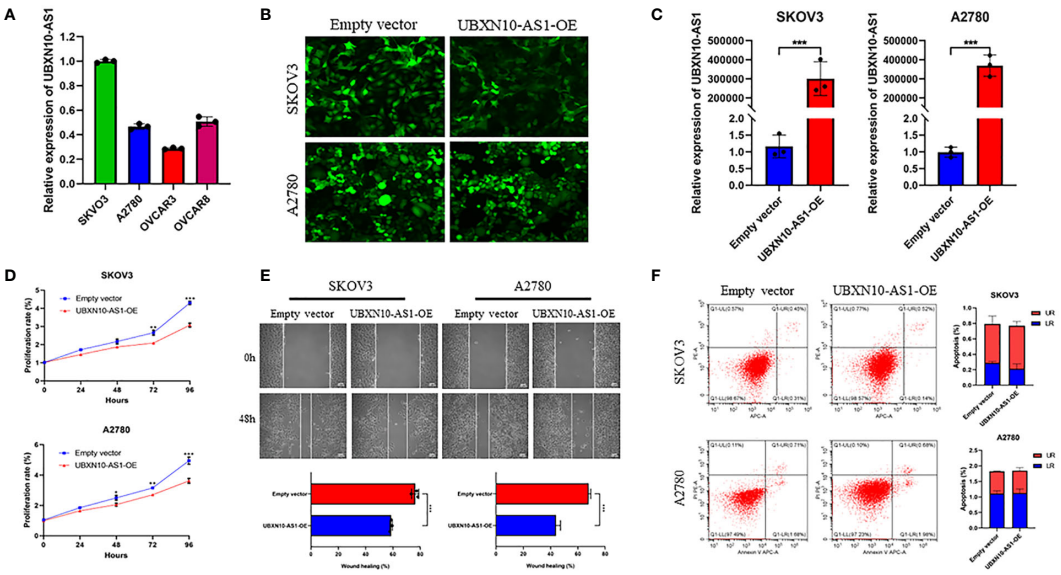


FIGURE 8
UBXN10-AS1 significantly suppressed cell proliferation and inhibited cell migration in SKOV3 and A2780 cell lines. Expression of UBXN10-AS1 in SKOV3, A2780, OVCAR3 and OVCAR8 cell lines **(A)**; Overexpression of UBXN10-AS1 in SKOV3 and A2780 cell lines **(B, C)**; Overexpression of UBXN10-AS1 significantly suppressed cell proliferation in SKOV3 and A2780 cell lines **(D)**; Overexpression of UBXN10-AS1 significantly inhibited cell migration in SKOV3 and A2780 cell lines **(E)**; Overexpression of UBXN10-AS1 had no influence on apoptosis in SKOV3 and A2780 cell lines **(F)**. *p<0.05; **p<0.01; ***p<0.001.

level and overexpression of UBXN10-AS1 suppressed tumor growth *in vivo* and *in vitro* (65). The function of UBXN10-AS1 in OC has not been reported. Therefore, we explored the function of UBXN10-AS1 in cell proliferation and migration in SKOV3 and A2780 cell lines. The results indicated that UBXN10-AS1 could significantly reduce cell proliferation and migration in OC.

Furthermore, we constructed an immune-related lncRNA prognostic signature to predict the OS. Based on the best cutoff value of risk score, all patients were divided into high- and low-risk groups. There was significantly different in OS between both high-risk group and low-risk group. Stratified analysis results revealed that the risk score was associated with age and FIGO stage. By using multivariate cox regression, we demonstrated that risk score was an independent prognostic factor for OC patients. In order to make the signature more applicable in clinic, a nomogram was established. Besides, the potential role of the immune-related signature in immune infiltration and immunotherapy response were investigated. The results indicated that various immune cells, especially tumor associated macrophages (TAMs), were differently distributed in high-risk group and low-risk group. Previous study reported that M2-like TAMs accelerated tumor growth, promoted tumor cell invasion and metastasis, and inhibited immune killing to promote tumor progression, which was consistent with our study (66). Accumulating evidence demonstrated that immune systems make a crucial contribution to the antitumor effects of conventional chemotherapy-based and radiotherapy-based cancer treatments (67, 68). Furthermore, the association between risk model and chemotherapy response were investigated. Our results suggested that the risk model might serve as potential predictor for chemosensitivity of various antitumor drugs, especially for paclitaxel, metformin, and veliparib, which are commonly used in treating OC patients.

In our study, both TCGA-RNA- seq datasets and HG-U133_Plus_2 datasets were included. The sample size is much larger than the studies before, which makes it more robust and reliable. However, there are some limitations. Due to different platforms, gene expression values are subject to sampling bias. Additionally, the roles of the lncRNAs and their interactions with immune-related genes are not confirmed using *in vitro* and *in vivo* experiments.

In summary, we have constructed a novel immune-related lncRNA signature, which have a potential prognostic value for ovarian cancer patients and might facilitate personalized counselling for immunotherapy and chemotherapy. Prospective studies are needed to further validate its predictive accuracy for estimating prognoses of ovarian cancer.

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

HL wrote the manuscript. Z-YL contributed to the data collection and analysis. NW and JW designed the idea and design the study. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Natural Science Foundation of China (NO.81972836), National Key R&D Program (2016YFC1303703) and the Science and Technology Innovation Program of Hunan Province (2020RC2065), the Youth Natural Science Foundation of Hunan Province (2021JJ40321).

Acknowledgments

We would like to thank Dr. Quan Cheng for the bioinformatics technology support.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.999654/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Least absolute shrinkage and selection operator (LASSO) regression was performed, calculating the minimum criteria **(A, B)** and coefficients **(C).**

SUPPLEMENTARY FIGURE 2
Time-dependent receiver operating characteristic (ROC) curves to predict the 5-year OS for the risk score, age and stage in TCGA-RNA-Seq cohort **(A)** and HG-U133_Plus_2 cohort **(B)**.

SUPPLEMENTARY FIGURE 3
Difference between high-risk score group and low-risk score group in immune infiltration, immunotherapy and chemotherapy response prediction in HG-U133_Plus_2 cohort. The risk score was positively correlated with EstimateScore, ImmuneScore, StromalScore and negatively correlated with PurityScore **(A-D)**; The Association of the signature and the distribution of 22 immune cells **(E)**; Estimated IC50 values indicated the chemotherapy response of paclitaxel, metformin, veliparib and Cisplatin in TCGA-RNA-Seq cohort **(F–I)**. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

SUPPLEMENTARY FIGURE 4
Molecular immune cell subtypes related to the five immune-related lncRNAs in HG-U133_Plus_2 cohort **(A)** and TCGA-RNA-Seq cohort **(B)**, respectively. GSEA identify the different pathways between high-risk group and low-risk group in HG-U133_Plus_2 cohort **(C)** and TCGA-RNA-Seq cohort **(D)**, respectively.

SUPPLEMENTARY TABLE 1
The primers in the study.

SUPPLEMENTARY TABLE 2
The immune-related lncRNAs identified in TCGA-RNA-seq dataset and HG-U133_Plus_2 dataset.

SUPPLEMENTARY TABLE 3
The chemotherapy response to anticancer drugs commonly used to treating cancers in TCGA-RNA-seq dataset and HG-U133_Plus_2 dataset.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2021) 71(3):209–249. doi: 10.3322/caac.21660

2. Holmes D. The problem with platinum. *Nature* (2015) 527(7579):S218–9. doi: 10.1038/527S218a

3. Bruni D, Angell HK, Galon J. The immune contexture and immunoscore in cancer prognosis and therapeutic efficacy. *Nat Rev Cancer* (2020) 20(11):662–80. doi: 10.1038/s41568-020-0285-7

4. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Sci (New York NY)* (2006) 313(5795):1960–4. doi: 10.1126/science.1129139

5. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med* (2015) 21(8):938–45. doi: 10.1038/nm.3909

6. Zhang L, Conejo-Garcia JR, Katsaros D, Gimotty PA, Massobrio M, Regnani G, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *New Engl J Med* (2003) 348(3):203–13. doi: 10.1056/NEJMoa020177

7. Curiel TJ, Coukos G, Zou L, Alvarez X, Cheng P, Mottram P, et al. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nat Med* (2004) 10(9):942–9. doi: 10.1038/nm1093

8. Kandalaft LE, Powell DJJr., Singh N, Coukos G. Immunotherapy for ovarian cancer: what's next? *J Clin Oncol* (2011) 29(7):925–33. doi: 10.1200/jco.2009.27.2369

9. Peng WX, Koirala P, Mo YY. LncRNA-mediated regulation of cell signaling in cancer. *Oncogene* (2017) 36(41):5661–7. doi: 10.1038/onc.2017.184

10. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* (2021) 22(2):96–118. doi: 10.1038/s41580-020-00315-9

11. Jiang MC, Ni JJ, Cui WY, Wang BY, Zhuo W. Emerging roles of lncRNA in cancer and therapeutic opportunities. *Am J Cancer Res* (2019) 9(7):1354–66.

12. Yu WD, Wang H, He QF, Xu Y, Wang XC. Long noncoding RNAs in cancer-immunity cycle. *J Cell Physiol* (2018) 233(9):6518–23. doi: 10.1002/jcp.26568

13. Jiang R, Tang J, Chen Y, Deng L, Ji J, Xie Y, et al. The long noncoding RNA lnc-EGFR stimulates T-regulatory cells differentiation thus promoting hepatocellular carcinoma immune evasion. *Nat Commun* (2017) 8:15129. doi: 10.1038/ncomms15129

14. Wang P, Xue Y, Han Y, Lin L, Wu C, Xu S, et al. The STAT3-binding long noncoding RNA lnc-DC controls human dendritic cell differentiation. *Sci (New York NY)* (2014) 344(6181):310–3. doi: 10.1126/science.1251456

15. Hu G, Tang Q, Sharma S, Yu F, Escobar TM, Muljo SA, et al. Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nat Immunol* (2013) 14(11):1190–8. doi: 10.1038/ni.2712

16. Qian M, Ling W, Ruan Z. Long non-coding RNA SNHG12 promotes immune escape of ovarian cancer cells through their crosstalk with M2 macrophages. *Aging (Albany NY)* (2020) 12(17):17122–36. doi: 10.18632/aging.103653

17. Shang A, Wang W, Gu C, Chen C, Zeng B, Yang Y, et al. Long non-coding RNA HOTTIP enhances IL-6 expression to potentiate immune escape of ovarian cancer cells by upregulating the expression of PD-L1 in neutrophils. *J Exp Clin Cancer Res* (2019) 38(1):411. doi: 10.1186/s13046-019-1394-6

18. Zhao Y, Yu Z, Ma R, Zhang Y, Zhao L, Yan Y, et al. lncRNA-Xist/miR-101-3p/KLF6/C/EBPα axis promotes TAM polarization to regulate cancer cell proliferation and migration. *Mol Ther Nucleic Acids* (2021) 23:536–51. doi: 10.1016/j.omtn.2020.12.005

19. Liao C, Wang A, Ma Y, Liu H. Long non-coding RNA FOXP4-AS1 is a prognostic biomarker and associated with immune infiltrates in ovarian serous cystadenocarcinoma. *Med (Baltimore)* (2021) 100(40):e27473. doi: 10.1097/md.0000000000027473

20. Lei J, He ZY, Wang J, Hu M, Zhou P, Lian CL, et al. Identification of MEG8/miR-378d/SOBP axis as a novel regulatory network and associated with immune infiltrates in ovarian carcinoma by integrated bioinformatics analysis. *Cancer Med* (2021) 10(8):2924–39. doi: 10.1002/cam4.3854

21. Shen Y, Peng X, Shen C. Identification and validation of immune-related lncRNA prognostic signature for breast cancer. *Genomics* (2020) 112(3):2640–6. doi: 10.1016/j.ygeno.2020.02.015

22. Chen Q, Hu L, Huang D, Chen K, Qiu X, Qiu B. Six-lncRNA immune prognostic signature for cervical cancer. *Front Genet* (2020) 11:533628. doi: 10.3389/fgene.2020.533628

23. Ma W, Zhao F, Yu X, Guan S, Suo H, Tao Z, et al. Immune-related lncRNAs as predictors of survival in breast cancer: a prognostic signature. *J Trans Med* (2020) 18(1):442. doi: 10.1186/s12967-020-02522-6

24. Li Z, Li Y, Wang X, Yang Q. Identification of a six-Immune-Related long non-coding RNA signature for predicting survival and immune infiltrating status in breast cancer. *Front Genet* (2020) 11:680. doi: 10.3389/fgene.2020.00680

25. Hong W, Liang L, Gu Y, Qi Z, Qiu H, Yang X, et al. Immune-related lncRNA to construct novel signature and predict the immune landscape of human hepatocellular carcinoma. *Mol Ther Nucleic Acids* (2020) 22:937–47. doi: 10.1016/j.omtn.2020.10.002

26. Jin D, Song Y, Chen Y, Zhang P. Identification of a seven-lncRNA immune risk signature and construction of a predictive nomogram for lung adenocarcinoma. *BioMed Res Int* (2020) 2020:7929132. doi: 10.1155/2020/7929132

27. Zheng J, Cao B, Zhang X, Niu Z, Tong J. Immune-related four-lncRNA signature for patients with cervical cancer. *BioMed Res Int* (2020) 2020:3641231. doi: 10.1155/2020/3641231

28. Lin Y, Pan X, Chen Z, Lin S, Chen S. Identification of an immune-related nine-lncRNA signature predictive of overall survival in colon cancer. *Front Genet* (2020) 11:318. doi: 10.3389/fgene.2020.00318

29. Li X, Meng Y. Survival analysis of immune-related lncRNA in low-grade glioma. *BMC Cancer* (2019) 19(1):813. doi: 10.1186/s12885-019-6032-3

30. Zhou M, Zhang Z, Zhao H, Bao S, Cheng L, Sun J. An immune-related six-lncRNA signature to improve prognosis prediction of glioblastoma multiforme. *Mol Neurobiol* (2018) 55(5):3684–97. doi: 10.1007/s12035-017-0572-9

31. Xia P, Li Q, Wu G, Huang Y. An immune-related lncRNA signature to predict survival in glioma patients. *Cell Mol Neurobiol* (2021) 41(2):365–75. doi: 10.1007/s10571-020-00857-8

32. Wang J, Shen C, Dong D, Zhong X, Wang Y, Yang X. Identification and verification of an immune-related lncRNA signature for predicting the prognosis of patients with bladder cancer. *Int Immunopharmacol* (2021) 90:107146. doi: 10.1016/j.intimp.2020.107146

33. Wu Y, Zhang L, He S, Guan B, He A, Yang K, et al. Identification of immune-related LncRNA for predicting prognosis and immunotherapeutic response in bladder cancer. *Aging* (2020) 12(22):23306–25. doi: 10.18632/aging.104115

34. Mateescu B, Batista L, Cardon M, Gruosso T, de Feraudy Y, Mariani O, et al. miR-141 and miR-200a act on ovarian tumorigenesis by controlling oxidative stress response. *Nat Med* (2011) 17(12):1627–35. doi: 10.1038/nm.2512

35. Ferriss JS, Kim Y, Duska L, Birrer M, Levine DA, Moskaluk C, et al. Multi-gene expression predictors of single drug responses to adjuvant chemotherapy in ovarian carcinoma: predicting platinum resistance. *PloS One* (2012) 7(2):e30550. doi: 10.1371/journal.pone.0030550

36. Lisowska KM, Olbryt M, Dudaladava V, Pamuła-Piłat J, Kujawa K, Grzybowska E, et al. Gene expression analysis in ovarian cancer - faults and hints from DNA microarray study. *Front Oncol* (2014) 4:6. doi: 10.3389/fonc.2014.00006

37. Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin Cancer Res* (2008) 14(16):5198–208. doi: 10.1158/1078-0432.Ccr-08-0196

38. Mok SC, Bonome T, Vathipadiekal V, Bell A, Johnson ME, Wong KK, et al. A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor: microfibril-associated glycoprotein 2. *Cancer Cell* (2009) 16(6):521–32. doi: 10.1016/j.ccr.2009.10.018

39. Konstantinopoulos PA, Spentzos D, Karlan BY, Taniguchi T, Fountzilas E, Francoeur N, et al. Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. *J Clin Oncol* (2010) 28(22):3555–61. doi: 10.1200/jco.2009.27.5719

40. Zhang X, Sun S, Pu JK, Tsang AC, Lee D, Man VO, et al. Long non-coding RNA expression profiles predict clinical phenotypes in glioma. *Neurobiol Dis* (2012) 48(1):1–8. doi: 10.1016/j.nbd.2012.06.004

41. Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, et al. ImmPort: disseminating data to the public for the future of immunology. *Immunol Res* (2014) 58(2-3):234–9. doi: 10.1007/s12026-014-8516-1

42. Goeman JJ. L1 penalized estimation in the cox proportional hazards model. *Biometrical J Biometrische Zeitschrift* (2010) 52(1):70–84. doi: 10.1002/bimj.200900028

43. De Felice F, Humbert-Vidan L, Lei M, King A, Guerrero Urbano T. Analyzing oropharyngeal cancer survival outcomes: a decision tree approach. *Br J Radiol* (2020) 93(1111):20190464. doi: 10.1259/bjr.20190464

44. Li H, Wu N, Liu ZY, Chen YC, Cheng Q, Wang J. Development of a novel transcription factors-related prognostic signature for serous ovarian cancer. *Sci Rep* (2021) 11(1):7207. doi: 10.1038/s41598-021-86294-z

45. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* (2013) 4:2612. doi: 10.1038/ncomms3612

46. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337

47. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* (2018) 24(10):1550–8. doi: 10.1038/s41591-018-0136-1

48. Kang Y, Huang J, Liu Y, Zhang N, Cheng Q, Zhang Y. Integrated analysis of immune infiltration features for cervical carcinoma and their associated immunotherapeutic responses. *Front Cell Dev Biol* (2021) 9:573497. doi: 10.3389/fcell.2021.573497

49. Nick TG, Hardin JM. Regression modeling strategies: an illustrative case study from medical rehabilitation outcomes research. *Am J Occup Ther* (1999) 53(5):459–70. doi: 10.5014/ajot.53.5.459

50. Li Z, Wang D, Yin H. A seven immune-related lncRNA signature predicts the survival of patients with colon adenocarcinoma. *Am J Trans Res* (2020) 12(11):7060–78.

51. Xu BF, Liu R, Huang CX, He BS, Li GY, Sun HS, et al. Identification of key genes in ruptured atherosclerotic plaques by weighted gene correlation network analysis. *Sci Rep* (2020) 10(1):10847. doi: 10.1038/s41598-020-67114-2

52. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* (2005) 102(43):15545–50. doi: 10.1073/pnas.0506580102

53. Feng B, Shen Y, Pastor Hostench X, Bieg M, Plath M, Ishaque N, et al. Integrative analysis of multi-omics data identified EGFR and PTGS2 as key nodes in

a gene regulatory network related to immune phenotypes in head and neck cancer. *Clin Cancer Res* (2020) 26(14):3616–28. doi: 10.1158/1078-0432.Ccr-19-3997

54. Gao Y, Liu Y, Liu Y, Peng Y, Yuan B, Fu Y, et al. UHRF1 promotes androgen receptor-regulated CDC6 transcription and anti-androgen receptor drug resistance in prostate cancer through KDM4C-mediated chromatin modifications. *Cancer Lett* (2021) 520:172–83. doi: 10.1016/j.canlet.2021.07.012

55. Xu H, Wang H, Zhao W, Fu S, Li Y, Ni W, et al. SUMO1 modification of methyltransferase-like 3 promotes tumor progression *via* regulating snail mRNA homeostasis in hepatocellular carcinoma. *Theranostics* (2020) 10(13):5671–86. doi: 10.7150/thno.42539

56. Kossaï M, Leary A, Scoazec JY, Genestie C. Ovarian cancer: A heterogeneous disease. *Pathobiology* (2018) 85(1-2):41–9. doi: 10.1159/000479006

57. Yu Y, Feng X, Cang S. A two-microRNA signature as a diagnostic and prognostic marker of pancreatic adenocarcinoma. *Cancer Manage Res* (2018) 10:1507–15. doi: 10.2147/cmar.s158712

58. Leone RD, Powell JD. Metabolism of immune cells in cancer. *Nat Rev Cancer* (2020) 20(9):516–31. doi: 10.1038/s41568-020-0273-y

59. Wei C, Yang C, Wang S, Shi D, Zhang C, Lin X, et al. Crosstalk between cancer cells and tumor associated macrophages is required for mesenchymal circulating tumor cell-mediated colorectal cancer metastasis. *Mol Cancer* (2019) 18(1):64. doi: 10.1186/s12943-019-0976-4

60. Yu H, Kortylewski M, Pardoll D. Crosstalk between cancer and immune cells: role of STAT3 in the tumour microenvironment. *Nat Rev Immunol* (2007) 7 (1):41–51. doi: 10.1038/nri1995

61. Wei C, Liang Q, Li X, Li H, Liu Y, Huang X, et al. Bioinformatics profiling utilized a nine immune-related long noncoding RNA signature as a prognostic target for pancreatic cancer. *J Cell Biochem* (2019) 120(9):14916–27. doi: 10.1002/jcb.28754

62. Yuan M, Wang Y, Sun Q, Liu S, Xian S, Dai F, et al. Identification of a nine immune-related lncRNA signature as a novel diagnostic biomarker for hepatocellular carcinoma. *BioMed Res Int* (2021) 2021:9798231. doi: 10.1155/2021/9798231

63. Fu Y, Katsaros D, Biglia N, Wang Z, Pagano I, Tius M, et al. Vitamin d receptor upregulates lncRNA TOPORS-AS1 which inhibits the wnt/β-catenin pathway and associates with favorable prognosis of ovarian cancer. *Sci Rep* (2021) 11(1):7484. doi: 10.1038/s41598-021-86923-7

64. Mo X, Li T, Xie Y, Zhu L, Xiao B, Liao Q, et al. Identification and functional annotation of metabolism-associated lncRNAs and their related protein-coding genes in gastric cancer. *Mol Genet Genomic Med* (2018) 6(5):728–38. doi: 10.1002/mgg3.427

65. Tang Y, Cai J, Lv B. LncRNA ubiquitin-binding protein domain protein 10 antisense RNA 1 inhibits colon adenocarcinoma progression *via* the miR-515-5p/slit guidance ligand 3 axis. *Bioengineered* (2022) 13(2):2308–20. doi: 10.1080/21655979.2021.2024396

66. Ge Z, Ding S. The crosstalk between tumor-associated macrophages (TAMs) and tumor cells and the corresponding targeted therapy. *Front Oncol* (2020) 10:590941. doi: 10.3389/fonc.2020.590941

67. Zitvogel L, Apetoh L, Ghiringhelli F, Kroemer G. Immunological aspects of cancer chemotherapy. *Nat Rev Immunol* (2008) 8(1):59–73. doi: 10.1038/nri2216

68. Ogino S, Galon J, Fuchs CS, Dranoff G. Cancer immunology–analysis of host and tumor factors for personalized medicine. *Nat Rev Clin Oncol* (2011) 8 (12):711–9. doi: 10.1038/nrclinonc.2011.122

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 13 October 2022
DOI 10.3389/fonc.2022.1008753

OPEN ACCESS

EDITED BY
Fu Wang,
Xi'an Jiaotong University, China

REVIEWED BY
Jian Sun,
Shandong Provincial Hospital, China
Weijia Sun,
Xiangya Hospital, Central South University, China

*CORRESPONDENCE
Zichuan Liu
13719063021@126.com
Haibo Zhang
haibozh@aliyun.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 01 August 2022
ACCEPTED 07 September 2022
PUBLISHED 13 October 2022

CITATION
Xu Z, Liu J, Liu Z and Zhang H (2022) MARCH1 as a novel immune-related prognostic biomarker that shapes an inflamed tumor microenvironment in lung adenocarcinoma.
Front. Oncol. 12:1008753.
doi: 10.3389/fonc.2022.1008753

MARCH1 as a novel immune-related prognostic biomarker that shapes an inflamed tumor microenvironment in lung adenocarcinoma

Zhiyong Xu[1,2†], Jun Liu[3†], Zichuan Liu[4*] and Haibo Zhang[1*]

[1]Department of Oncology, the Second Clinical Medical School, Guangzhou University of Chinese Medicine, Guangzhou, China, [2]Department of Radiotherapy, Southern Theater General Hospital, Guangzhou, China, [3]Department of Pulmonary and Critical Care Medicine, The Second Affiliated Hospital of South China University of Technology, Guangzhou, China, [4]Internal Medicine Section2, Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, China

E3 ubiquitin ligases (E3s), the second most common cancer-related functional protein family, play vital roles in multiple tumors. However, their importance in prognosis and immunotherapy of lung adenocarcinoma (LUAD) is not clear. First, utilizing the data from The Cancer Genome Atlas (TCGA), we comprehensively assessed the expression profile and immunological association of 13 E3s in LUAD patients. Consequently, MARCH1 was considered a candidate for further study. Second, several algorithms were applied to assess the correlation between MARCH1 and immunological characteristics in the LUAD tumor microenvironment. Third, an immune risk score (IRS) was developed to predict the prognosis. Finally, the immunological relationship of MARCH1 in pan-cancer was also estimated. We found that E3s were disordered in LUAD. Among them, MARCH1 was positively correlated with most immunological characteristics, indicating that MARCH1 designed an inflamed TME in LUAD. Coincidently, LUAD with low MARCH1 expression had a poor prognosis and was not sensitive to immune checkpoint blockers. In addition, the IRS could accurately predict the prognosis. In pan-cancer, MARCH1 was also positively correlated with most immunological characteristics. In conclusion, MARCH1 could be a novel and promising biomarker for immune status and effectiveness of immunotherapy for LUAD patients.

KEYWORDS

MARCH1, LUAD, E3s, immunotherapy, pan-cancer

Introduction

Lung adenocarcinoma (LUAD) is the dominating pathological subtype of lung cancer, which is the deadliest and second most prevalent cancer. Its incidence is still increasing worldwide (1). Moreover, the therapeutic outcome is far from satisfactory due to delayed diagnosis and limitation of traditional treatments.

Tumor cells could be identified as abnormal substances and eliminated by immune cells. Meanwhile, they have special mechanisms to evade host immune surveillance (2). Immunotherapy, with fewer off-target effects and longer-lasting responses, could restore the patient's immune system to kill tumor cells through natural mechanisms and is rapidly becoming a focus of oncology research (3). Recent cancer treatment applications of immunotherapy include chimeric antigen receptor T cells, vaccine therapy, and immune checkpoint blockers (ICBs) targeting programmed cell death-ligand 1/programmed death protein 1 (PD-L1/PD-1) and cytotoxic T lymphocyte-associated protein 4 (CTLA-4) (4). In clinical application, the US Food and Drug Administration (FDA) has approved several ICBs to treat non-small cell lung cancer (NSCLC), melanoma, and some other malignant tumors (5). Despite these encouraging results, immunotherapy is only effective for a minority. Accumulated evidence revealed that sensitivity to ICBs was strongly related to tumor immune phenotypes, which were classified as inflamed/infiltrated, immune-excluded, and immune-desert phenotypes based on the T cells' spatial distribution in the tumor microenvironment (TME) (6). An inflamed TME always made immunotherapy more effective than the other two phenotypes. It was characterized by a high PD-L1 and PD-1 expression and a high prevalence of tumor-infiltrating immune cells (TIICs) (7). Consequently, the amount of TIILs and factors regulating the immune cell infiltration, such as cytokines, chemokines, and other components, is crucial for immunotherapy. Meanwhile, elements of inflamed tumors included microsatellite instability (MSI) and tumor mutational burden (TMB) (6, 8). Taken together, these immunologic characteristics within the TME were vital to immunotherapy. Therefore, a biomarker indicating the status of the TME could predict the immunotherapy response.

Ubiquitination, one of the posttranslational modifications, is a cascade that regulates protein degradation by ligating ubiquitin to the target protein. Ubiquitin is activated by binding to ubiquitin-activating enzymes (E1s), subsequently transmitted to ubiquitin-conjugating enzymes (E2s), and finally covalently ligated to a target protein regulated by ubiquitin ligases (E3s) (9). Ubiquitination is an essential system that regulates the stability of numerous pivotal regulatory factors and cellular processes, covering cell cycle, proliferation, apoptosis, and neurotransmission (10). It has been observed to be dysregulated in many cancers (11).

E3s, of which there are about 1,000 members in *Homo sapiens*, can be divided into four categories according to their functional domains: HECT domain-containing type, PHD-finger type, U-box type, and RING-finger type proteins (12). Because of their specificity for substrates, E3s are key regulators in the ubiquitination process. Several immune processes have been linked to their regulation, including immune evasion and antigen presentation, T cell-mediated tolerance, and lymphocyte activation and differentiation (13). Furthermore, ubiquitination of PD-1/PD-L1 *via* E3s seriously alters the protein stabilization and dynamics of PD-1/PD-L1 in cancer immunotherapy (14). However, the relationship between E3s and immunologic signatures in the TME as well as their predictive value in prognosis and immunotherapy efficacy in LUAD remains unknown.

Herein, we obtained 13 E3s, of which the significance in immunity has been uncovered, and demonstrated the relationship between the 13 E3s and immunologic characteristics in the TME. Of interest, MARCH1 was found to have a strong association with the TME. To gain sufficient insight into the role of MARCH1 in LUAD and pan-cancer, we conducted a comprehensive analysis on multiple levels containing mRNA expression, immune signature, patient survival, and chemical compounds. We also established a risk model to predict prognosis and immunotherapy response. Collectively, our systematic analysis provides a comprehensive insight on the biology of MARCH1, which has greater potential value on immunotherapy targets than other E3s.

Materials and methods

Data acquisition

All data, including the pan-cancer RNA sequencing data, somatic mutation data, and detailed clinical data, were acquired from The Cancer Genome Atlas (TCGA) database using UCSC Xena. TMB was calculated with somatic mutation data. MSI data were collected from the study of Bonneville et al. (15).

Expression profiles of E3 ligases

First, the expression profiles of the 13 E3s in tumor tissues and paracarcinoma tissues from LUAD patients were analyzed using the RNA sequencing data. Then, in pan-cancer, differences in MARCH1 level between tumor and paracarcinoma tissues were computed *via* "limma" R package (false discovery rate <0.05, $|\log_2 FC| \geq 1$).

Correlation between MARCH1 and the immunological characteristics in the TME

The characteristics contain the expression level of immunomodulators (16), the expression of immune checkpoints, the infiltration level of TIICs, and the cancer immunity cycle's

activity. The activities of these cancer immunity cycle steps were evaluated by single sample gene set enrichment analysis (17). The association between MARCH1 and immune checkpoints, mismatch repair (MMR) protein was analyzed *via* the Spearman correlation coefficients pan-cancer.

Association between MARCH1 and therapeutic signatures

We summarized the therapeutic signatures from previous studies. Then, their enrichment scores (ESs) were computed *via* the gene set variation analysis R package. The LUAD-linked drug-target genes were filtered out in the DrugBank database. Their levels were compared between low- and high-MARCH1 group.

Screening of immune-related differentially expressed RNAs

Considering the median of MARCH1 mRNA expression, immune score, and stromal score, the latter two computed *via* the ESTIMATE R package, LUAD cohorts were parted into corresponding low and high groups. Differentially expressed RNAs (DERs) were identified *via* the limma R package. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were calculated by ClusterProfiler R package.

Establishment of an immune risk score

With a ratio of 7:3, TCGA-LUAD patients were separated into training and validation sets. Univariate Cox analysis was executed in the training set to identify the correlation between DERs and survival. Then, the immune risk score (IRS) was developed *via* least absolute shrinkage and selector operation (LASSO)-multivariate Cox regression (IRS = $\sum \beta i * RNAi$ βi: the coefficient of the 'i'th IRS RNA expression profile). Referring to the median IRS, patients fell into low and high groups, and their overall survival (OS) was compared by the Kaplan–Meier method and the log-rank test. Furthermore, the IRS was validated in the validation set.

Survival analysis in pan-cancer

To demonstrate the links between MARCH1 expression and OS, survival analysis was carried out in TCGA using the "survival" package in R (18).

Statistical analysis

All statistical analyses were executed utilizing the R software v4.0.3. Correlation between certain variables was gauged using Pearson coefficients. Statistical significance was computed by the log-rank test and defined as $p < 0.05$.

Results

Landscape, prognostic value, and immunological correlation of E3s in LUAD

We obtained the expression of the 13 E3s in LUAD from TCGA database. After a comprehensive analysis, we found that the expressions of CBLB, FBXW7, HUWE1, ITCH, SIAH2, STUB1, SYVN1, TRM2B, and UBR5 were significantly upregulated; MARCH1, RNF128, and TRAF6 were significantly downregulated; and ASB2 had no obvious difference between tumor and paracarcinoma (Figure 1A).

Our goal was to determine the immunological roles of E3s in LUAD. The results uncovered that E3s had a negative or positive correlation with most immunomodulators and TIICs. Among them, MARCH1 and ASB2 were positively correlated with most immunomodulators and all of the TIICs in this analysis (Figures 1B, C). ASB2 expression was correlated with PD-L1, PD-1, CTLA-4, and LAG-3. Simultaneously, MARCH1 expression was correlated with PD-L1, PD-1, and CTLA-4 (Figures 1D–G).

These factors, which are crucial for immunotherapy, were positively linked to MARCH1 expression and were more potent than those of other E3s. Moreover, tumor tissue showed a downregulation of MARCH1. We concluded that the downregulation pattern of MARCH1 may be TME specific, indicating the potential of MARCH1 to be a target to improve LUAD immunotherapy. Hence, MARCH1 was regarded as a candidate gene for further study based on its significance in determining prognosis and immune response.

MARCH1 shapes an inflamed TME in LUAD

As shown in Figure 2A, MARCH1 was positively related to plenty of immunomodulators. Specifically, many major histocompatibility complex molecules (MHCs) were repressed in the low-MARCH1 group. C-X-C motif chemokine ligand (CXCL)9 and CXCL10, two key chemokines promoting the infiltration of CD8+ T, were downregulated in the low-

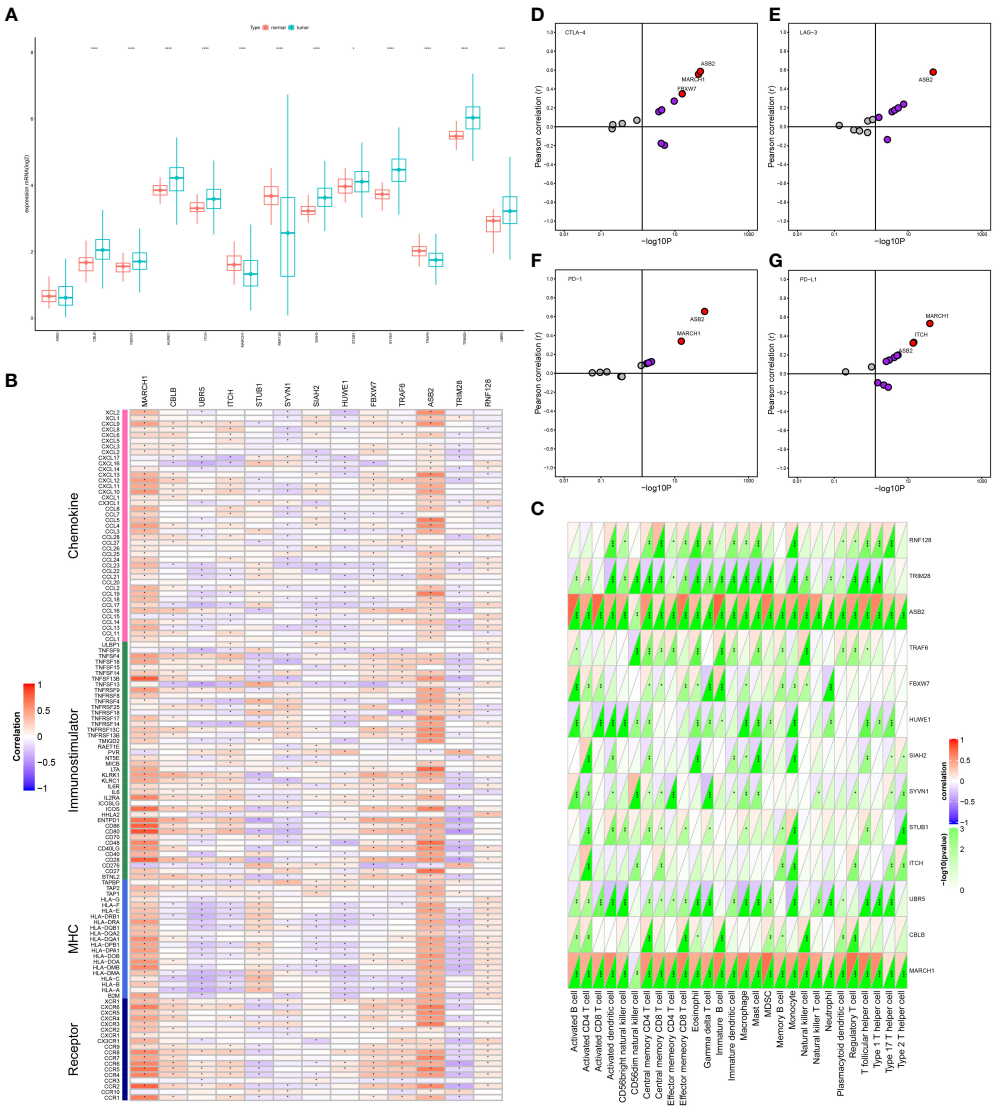


FIGURE 1
Correlation between E3 ligases (E3s) expression and immunological status in lung adenocarcinoma (LUAD). **(A)** The expression pattern of E3s in tumor and paracarcinoma tissues from TCGA database. **(B)** Relation between E3s and immunomodulators. **(C)** Relation between E3s and tumor-infiltrating immune cells (TIICs). **(D–G)** Relation between E3s and four immune checkpoints. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

MARCH1 group. In addition, chemokines, such as C-C motif chemokine ligand (CCL)2–5, CCL19, CXCL11, and their corresponding receptors were positively related to MARCH1. In the low-MARCH1 group, activities of most of the steps (Steps 1–5) of the cycle were significantly decreased, indicating a reduced level of TIICs. Of interest, the activities of Steps 6 and 7 were downregulated in the high-MARCH1 group (Figure 2B). Furthermore, the infiltration level of TIICs was assessed. As anticipated, MARCH1 had a positive correlation with the effector genes of T helper 1 cells, natural killer cells, macrophages, dendritic cells, and CD8+ T cells (Figure 2C). The results also showed that MARCH1 was positively related to numerous immune checkpoints (Figure 2D).

In the IMvigor210 cohort, MARCH1 expression was gradually increased from the desert, excluded, to inflamed tumor immune phenotypes. Moreover, in the groups classified based on PD-L1 (TC0, TC1, TC2) or PD-1 expression (IC0, IC1, IC2), MARCH1 expression was highest in the groups with the

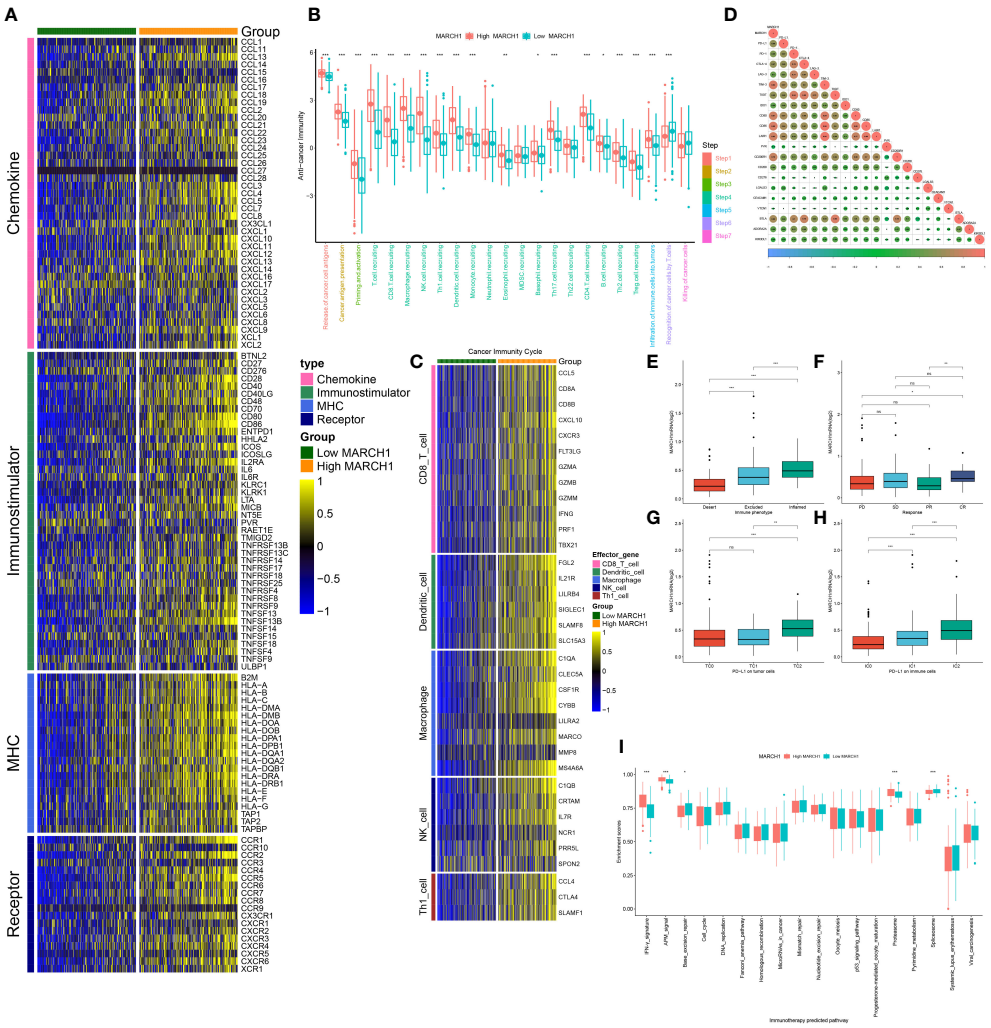


FIGURE 2
MARCH1 created an infiltrated tumor microenvironment (TME) in LUAD. **(A)** Expression of four types of immunomodulators in low- and high-MARCH1 groups. **(B)** Activity of cancer immunity cycle in low- and high-MARCH1 groups. **(C)** Expression of TIICs' effector genes in low- and high-MARCH1 groups. **(D)** Relevance between the expression of MARCH1 and immune checkpoints. **(E)** MARCH1 expression in the three immune phenotypes. **(F)** Expression of MARCH1 in four types of clinical outcome of immunotherapy. **(G, H)** Expression of MARCH1 in different cohorts grouped by PD-L1 or PD-1 expression. **(I)** The ESs of pathways for immunotherapy prediction in low- and high-MARCH1 groups. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$. ns, no significance.

highest PD-L1/PD-1 expression (TC2 and IC2, respectively) (Figures 2E, G, H). Taken together, MARCH1 was strongly linked with the immune phenotype of the TME.

MARCH1 predicts the clinical response to ICB and other therapeutic options in LUAD

From the results above, MARCH1 shaped an inflamed TME in LUAD patients, so patients with higher MARCH1 expression ought to be more sensitive to ICBs. Therefore, we further compared the outcome of LUAD patients with distinct MARCH1 expressions. The result showed that MARCH1 expression was significantly higher in patients with complete response to immunotherapy compared to those patients with progressive and stable disease (Figure 2F). Positive correlation also existed between MARCH1 and the ESs of three immunotherapy-positive gene signatures: IFN-γ signature, APM signal, and proteasome signal (Figure 2I). In addition, MARCH1 had a positive correlation with most individual genes of the T cell inflamed signature (Figure 3A). However, there was no discernible difference in the ESs of the therapeutic targets between low- and high-MARCH1 groups, except for the

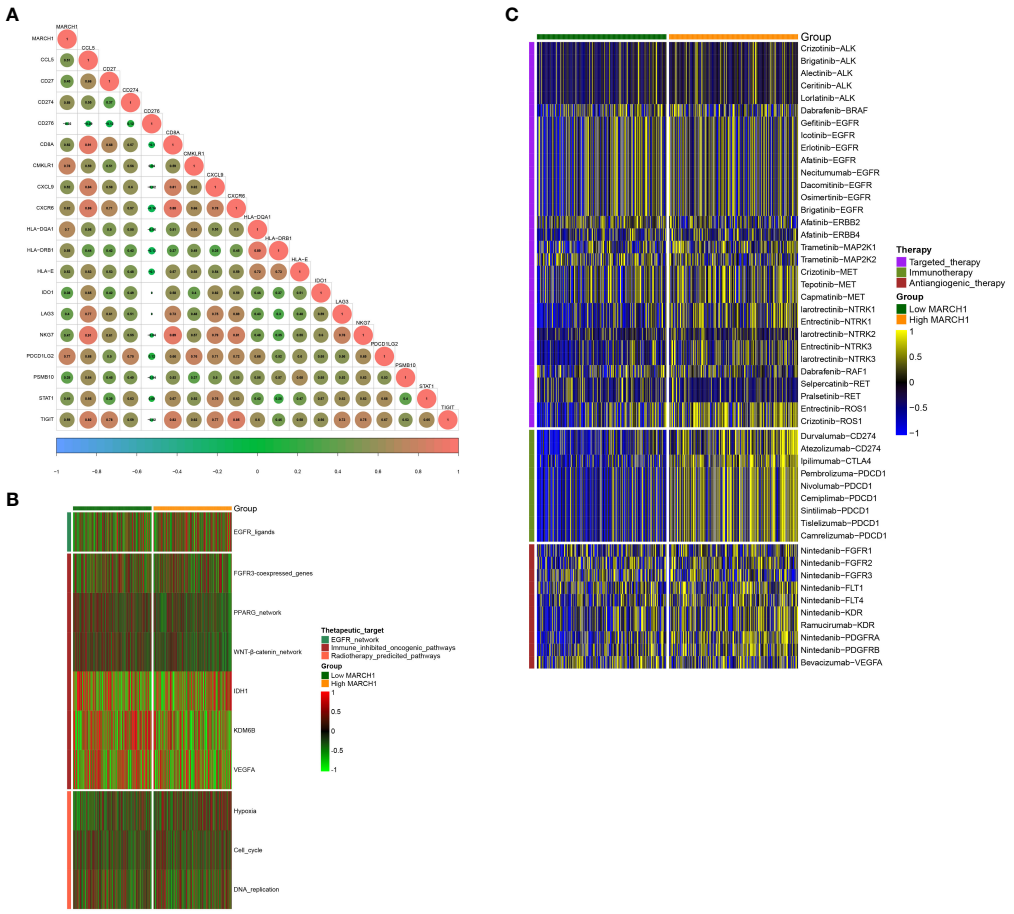


FIGURE 3
MARCH1 predicts the therapeutic sensitivity in LUAD. **(A)** Relevance between MARCH1 and the T cell inflamed gene signature. **(B)** Relevance between MARCH1 and therapeutic targets. **(C)** Relevance between MARCH1 and LUAD-related drug-target genes.

peroxisome proliferator-activated receptor gamma (PPARG) network and WNT-β-catenin network, which were both higher in the former group (Figure 3B). Analysis of the association between MARCH1 and drug-targeted genes unveiled an obviously higher sensitivity to specific targeted therapies and immunotherapies in the high-MARCH1 group (Figure 3C). In a word, ICB could apply to LUAD patients with a high MARCH1 level but not those with a low MARCH1 level.

Immune-related DER identification

In total, 246 common DERs with prognostic significance were screened out (Figure 4A). Notably, there was no overlap among downregulated DERs in the low-MARCH1, high-stromal score, and immune score group. Likewise, no intersection was found among downregulated DERs in the high-MARCH1, low-stromal score, and immune score group (Figures 4D, E). It indicated that MARCH1 expression positively related to stromal score and immune scores in the LUAD TME. Furthermore, GO and KEGG analyses revealed that these DERs were involved in immune-related processes (Figure 4).

IRS establishment and validation

According to univariate Cox analysis, 102 DERs had prognostic values. Among them, seven DERs with minimal λ (0.04141) were considered as the best candidates *via* the LASSO algorithm (Figures 5A–C). Then, a multivariate Cox regression analysis was performed to develop an IRS according to the seven DERs. Considering the IRS median, 350 patients from TCGA training set were sorted into low- (n = 175) and high-IRS groups (n = 175). The result showed that patients from the low-IRS group had remarkably longer OS than those from the high-IRS group. At 1, 3, and 5 years, the AUCs of the IRS were all more than 0.6 (Figure 5D). Furthermore, verification of the prediction accuracy in TCGA validation set displayed that the AUCs of the

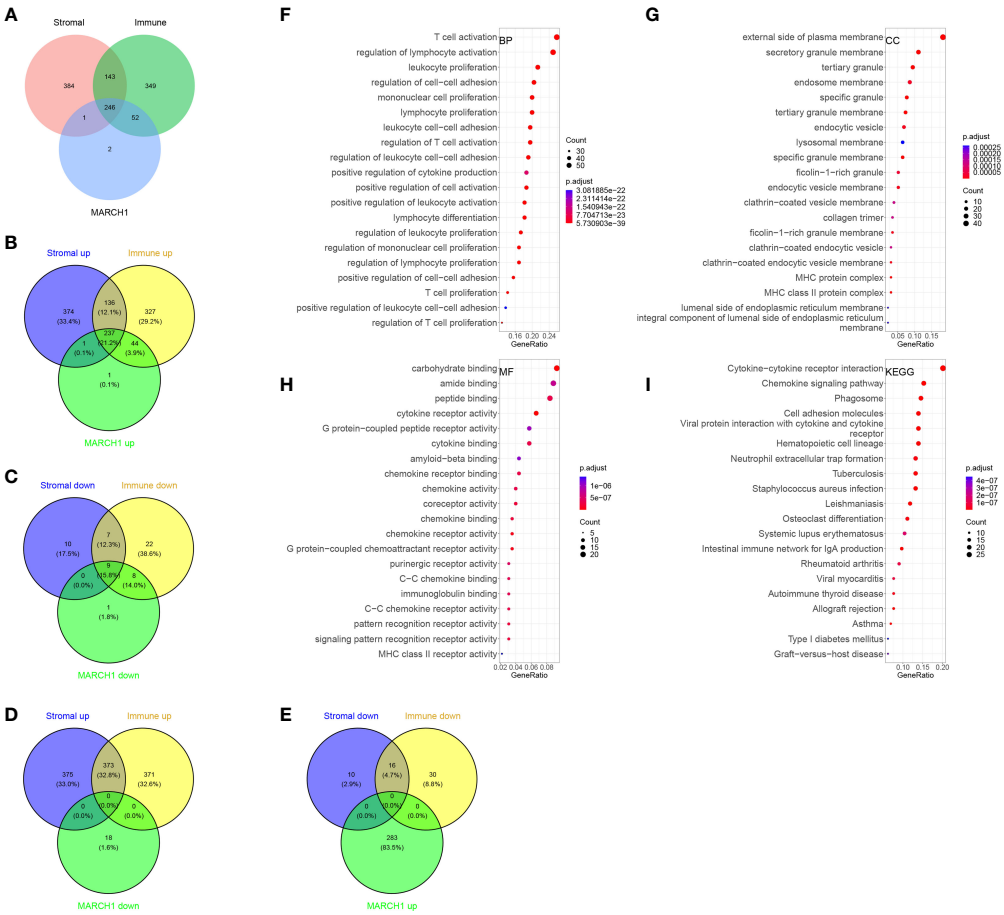


FIGURE 4
Immune-related differentially expressed RNAs (DERs). **(A–E)** Intersection between DERs in different immune/stromal score groups and different MARCH1 groups. **(F–I)** GO and KEGG analyses of the DERs.

IRS in the validation and training sets were very similar (Figure 5E). Taken together, this model could steadily predict the prognosis.

MARCH1 expression profiles and the correlation with prognosis in pan-cancers

To clarify the expression profile of MARCH1 in pan-cancer, MARCH1 levels between tumor and paracarcinoma tissue were compared in 33 cancers. MARCH1 expression was significantly upregulated in breast cancer (BRCA), cervical squamous cell carcinomas (CESC), cholangiocarcinoma (CHOL), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), and stomach adenocarcinoma (STAD). Meanwhile, MARCH1 expression was significantly decreased in colon adenocarcinoma (COAD), LUAD, lung squamous cell carcinoma (LUSC), prostate adenocarcinoma (PAAD), and rectum adenocarcinoma (READ) (Figure 6A).

In pan-cancer, the significance of MARCH1 in prognosis was analyzed. The result revealed that a high expression of MARCH1 was always linked with a better OS in lower grade glioma (LGG), LUAD, and skin cutaneous melanoma (SKCM) (Figure 6B). OS curves in different cancers showing significant differences between high- and low-MARCH1 groups are exhibited in Figures 6C–E.

Genome-wide relation of MARCH1 expression in pan-cancer

The association between MARCH1 and genomic signatures (DNA methylation, somatic copy number, somatic mutation, protein level) was explored *via* the Regulome Explorer web tool.

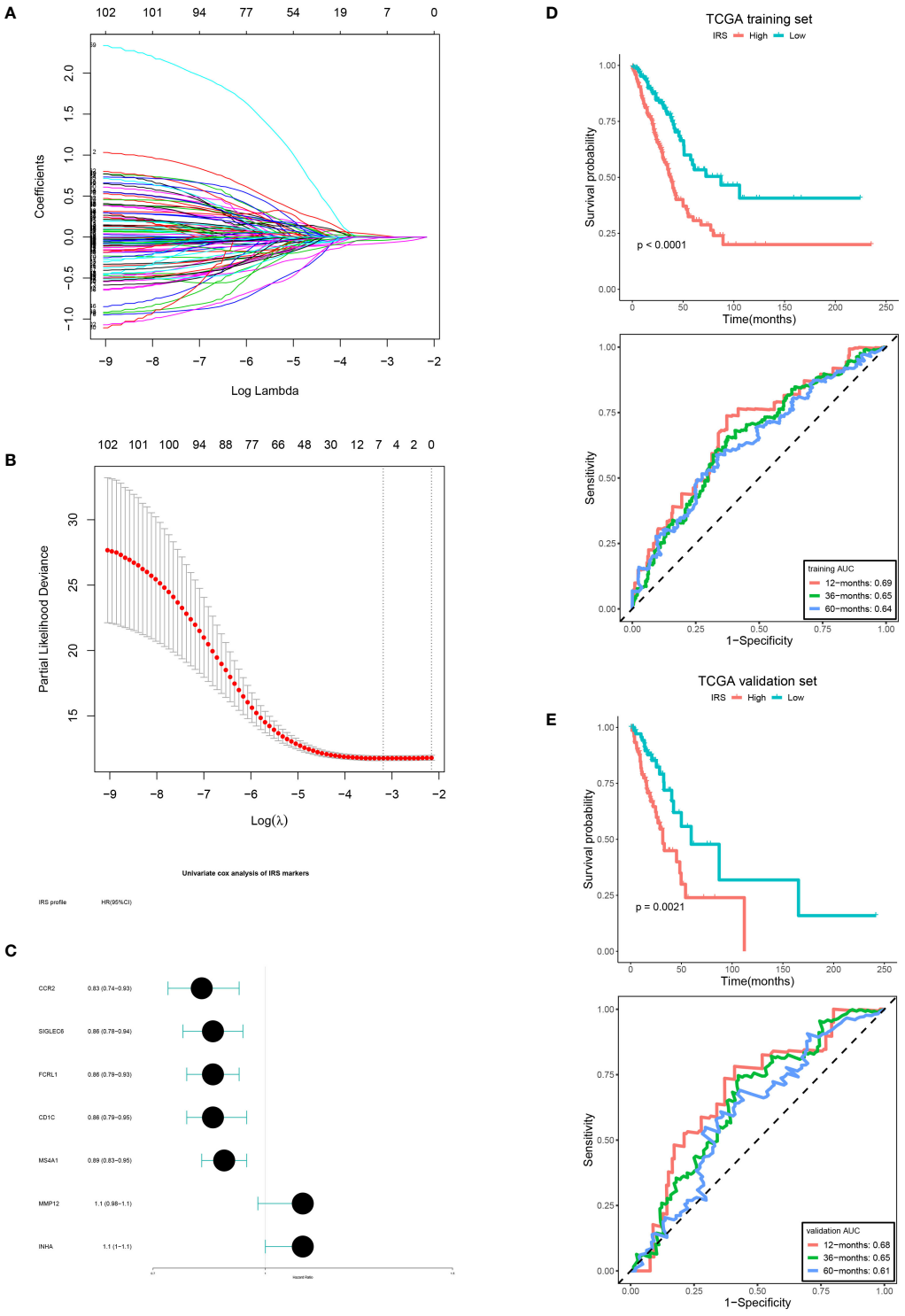


FIGURE 5
IRS development and verification. **(A, B)** LASSO coefficient of the DERs with prognostic value. **(B)** Cross-validation for turning parameter selection. **(C)** IRS markers displayed as forest plot. OS curves of the low and high IRS and AUCs in the training set **(D, E)** validation set.

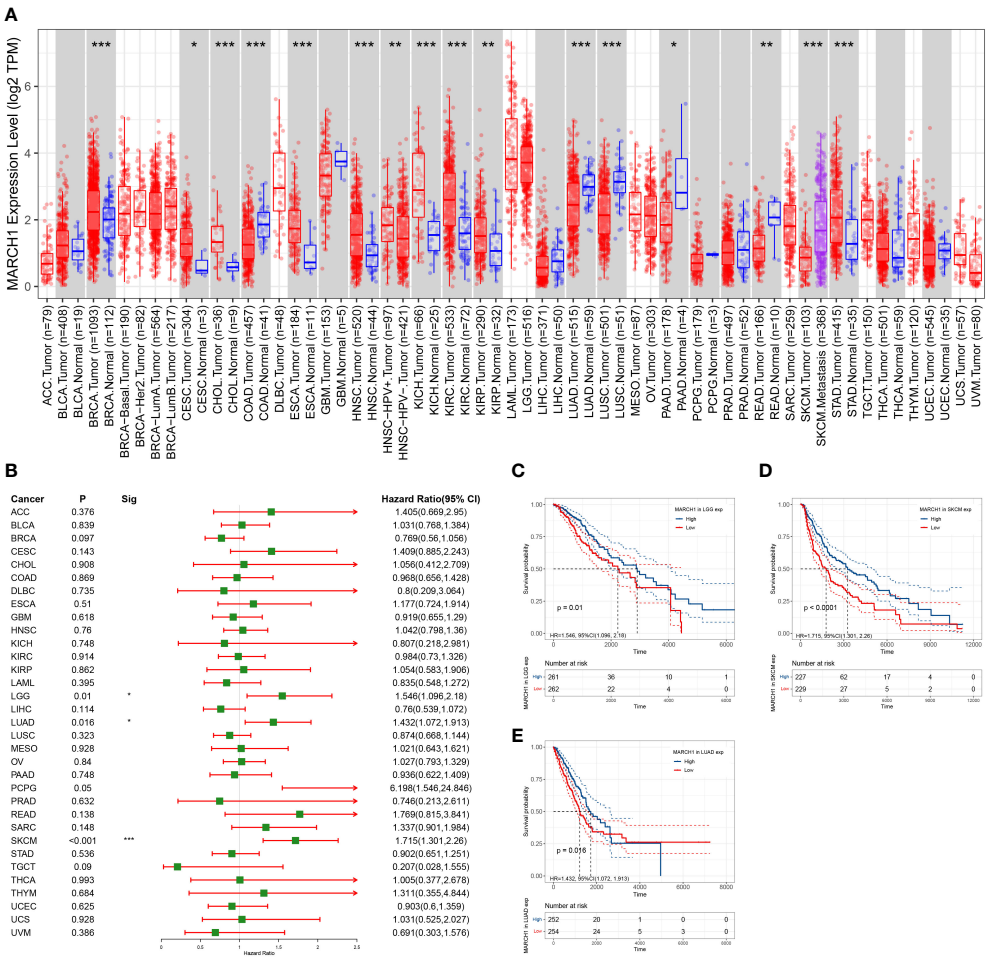


FIGURE 6
Expression profile and prognosis in pan-cancer. **(A)** MARCH1 expression levels in different types of cancer. **(B)** Relation between MARCH1 expression and prognosis in pan-cancers. **(C–E)** OS curves with significance in three types of cancer (LGG, LUAD, and SKCM). *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Circus plots illustrated that genome-wide correlations existed in many cancers. Figure 7 displays the particulars.

Correlation between MARCH1 and immunological characteristics in pan-cancer

We analyzed the associations between MARCH1 expression and immunomodulators, the abundance of TIICs in pan-cancer. The result displayed that MARCH1 had a positive correlation with most immunomodulators and TIICs in pan-cancer, except in KICH and LGG (Figures 8A, D). Furthermore, we found correlations between MARCH1 expression and confirmed immune checkpoints. MARCH1 was also discovered to have a significant positive correlation with large numbers of immune checkpoints in pan-cancer excluding KICH and LGG (Figure 8B). The correlations between MARCH1 expression and five vital MMR signatures, EPCAM, MLH1, MSH2, MutS MSH6, and PMS2, were also detected. The result revealed that MMR signatures, except EPCAM, were positively associated with MARCH1 expression. MARCH1 expression and TMB also had a significant positive association in COAD and ovarian serous cystadenocarcinoma (OV) and a significant negative association in CHOL (Figure 8E). MARCH1 expression was non-significantly correlated with MSI in most types of cancer. However, in COAD, acute myeloid leukemia (LAML), and READ, a higher level of MARCH1 meant significantly higher MSI, while in diffuse large B-cell lymphoma (DLBC), KIRP, LUAD, LUSC, SKCM, and testicular germ cell tumors (TGCT), the opposite trend was observed (Figure 8F).

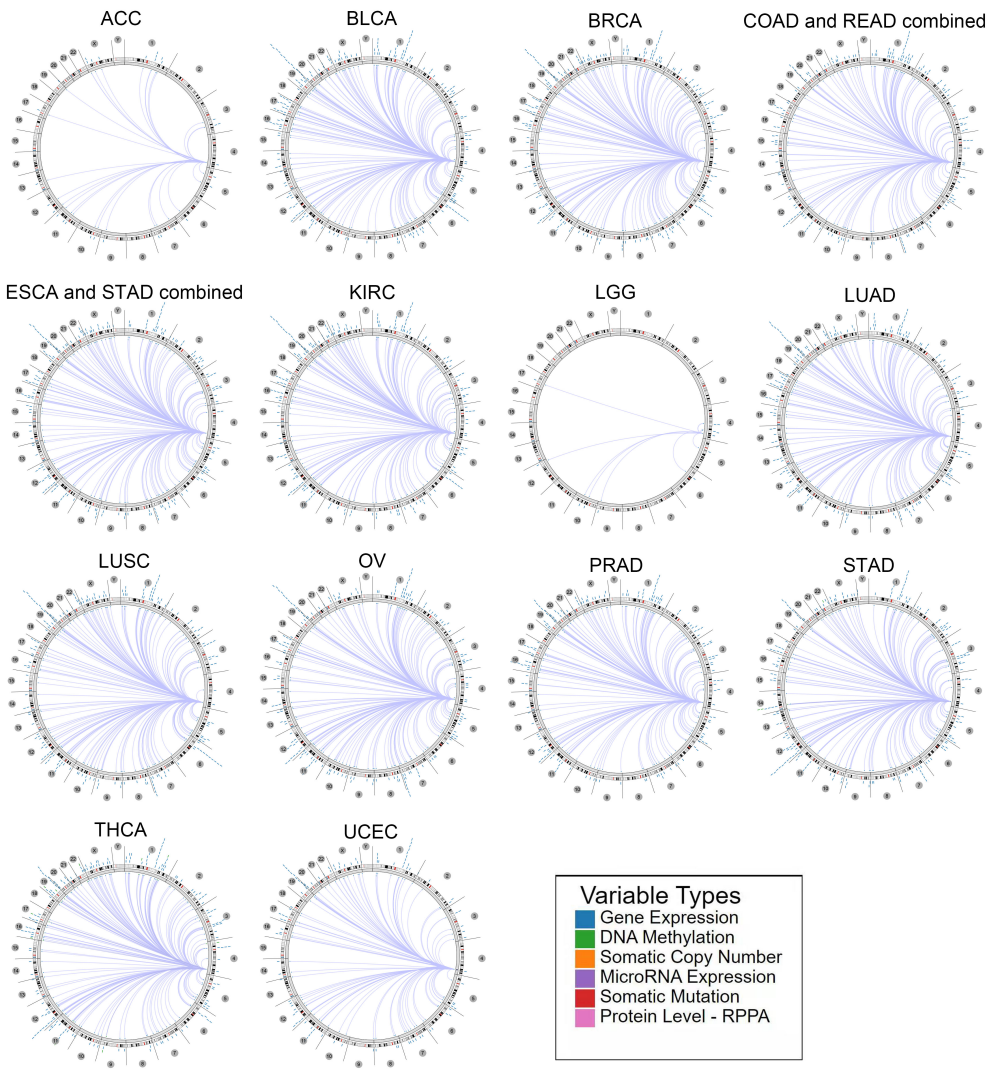


FIGURE 7
Circus plots displayed the relationship between MARCH1 and other genomic signatures.

Discussion

Accumulating evidence shows that E3s are strongly related to cancer immunity (19–22). Nonetheless, their value in prognosis and immunological prediction remains unclear. Hence, we selected 13 E3s that have been reported to be related to the immune system to identify a novel and robust marker that could predict the immunotherapy response.

Dysregulation of E3s is frequently observed in numerous cancers and aids tumor cells evading the immune system (13, 22). Consistently, in this study, 12 out of the 13 E3s were significantly upregulated or downregulated in tumor tissue from LUAD patients (Figure 1A). This situation further implied that E3s play crucial roles in cancer. Considering their function in cancer immunity, the associations between E3s and immunomodulators, immune cells, and immune checkpoints in the TME were analyzed. The results revealed correlations between E3s and most of these immune-related factors. Among those E3s, MARCH1 expression had a positive correlation with most of the immunomodulators, immune cells, and checkpoints in LUAD and many other cancers (Figures 1B, C; 8A, B, D). Therefore, MARCH1 was regarded as a candidate gene for immunotherapy response prediction.

Accurate prediction of the immunotherapy sensitivity could guide the clinical treatment of cancer. Some recognized biomarkers, such as TMB and MMR defects, have been used

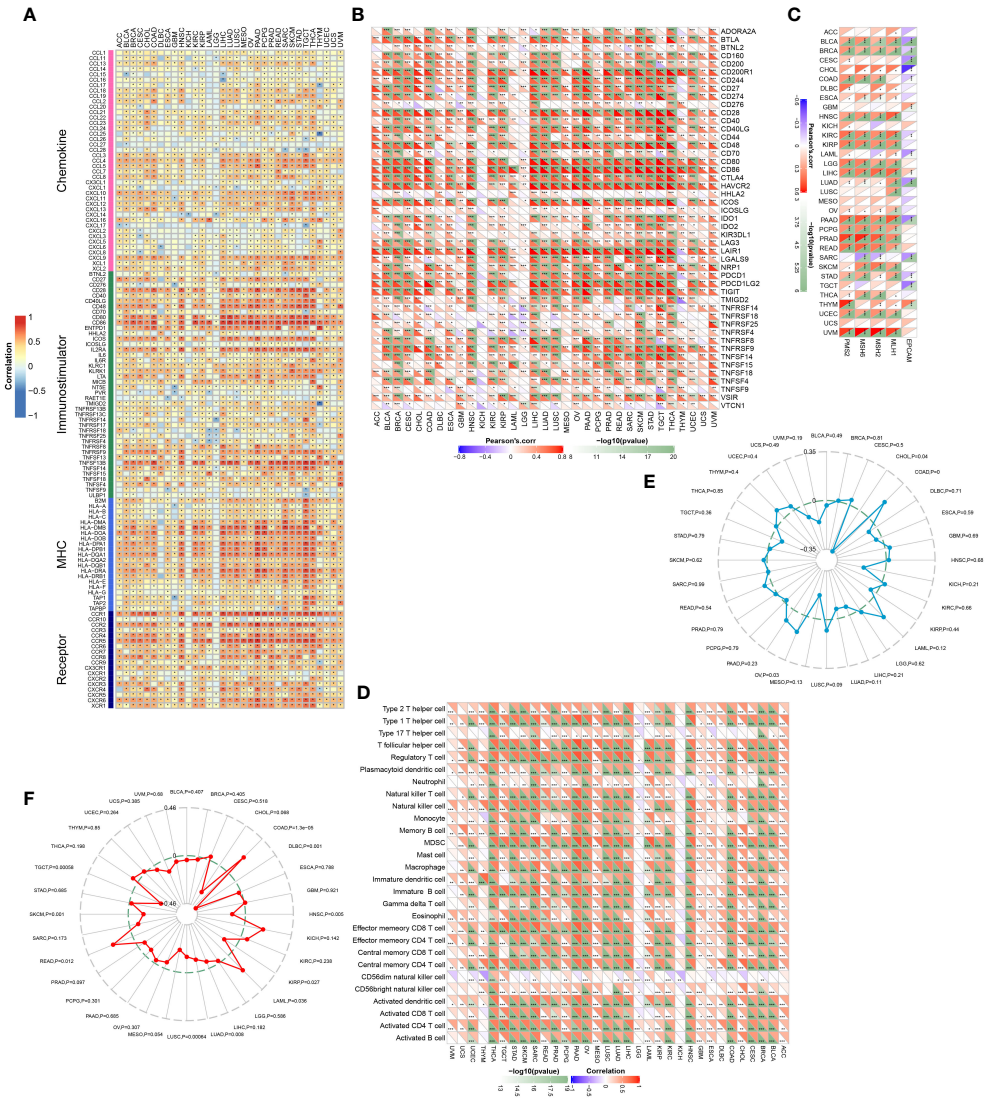


FIGURE 8
Associations between MARCH1 and immunological status in pan-cancer. Correlation between MARCH1 and **(A)** immunomodulators, **(B)** immune checkpoints, **(C)** MMR signatures, **(D)** TIICs, **(E)** TMB, and **(F)** MSI. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

to predict immunotherapy sensitivity (23, 24). Previous studies have discovered that MMR is essential for identifying and repairing mismatched bases during DNA replication (25). Therefore, DNA MMR deficiency typically generated high TMB (26) and MSI (27). They contribute to tumor initiation and are independent predictors of ICB efficacy (25). Recent studies concentrated on discovering more precise, convenient, and economical molecular techniques for clinical applications through the development of personalized medicine in a variety of solid tumors. Consequently, there is an urgent need to search for additional biomarkers that can aid clinical immunotherapy. However, there are still no signatures to accurately predict immunotherapy sensitivity.

In this study, we separated the LUAD patients into two groups depending on their level of MARCH1 expression: low- and high-MARCH1 groups. Interestingly, the expression of most immunomodulators, activities of cancer immunity cycle steps, and ESs of some predictable pathways were elevated in the high-MARCH1 group (Figures 2A, B; 3D). The cancer immunity cycle is the procedure of the immune response to tumor cells. The activities of these steps comprehensively determine the antitumor effect of the complicated immunomodulatory

interplays in the TME (16). In this study, we discovered that MARCH1 was positively associated with nearly all steps (except killing of cancer cells) of the cancer immunity cycle. Upregulation of immune checkpoints, including PD-L1/PD-1, is also an important characteristic of the inflamed TME, which is triggered by preexisting TIICs in the TME (28). ICBs that target these immune checkpoints have provided LUAD patients with the potential for therapeutic effect and survival. Interestingly, we found that the MARCH1 expression level was positively correlated with the expression of immune checkpoints and TIICs. Moreover, there is relevance between MARCH1 expression level and tumor immunotype. The MARCH1 mRNA level ranged from low to high in desert, excluded, and inflamed immune phenotypes (Figure 2E). We also established an IRS for prognosis prediction on the basis of immune-related DERs. Moreover, the IRS model was validated well in the internal validation cohort. In summary, both MARCH1 and IRS may serve as prognostic biomarkers, which robustly illustrate the importance of MARCH1 in prognosis. Furthermore, MARCH1 can also predict the ICB response and define an inflamed TME. High MARCH1 expression always meant that LUAD patients were sensitive to ICBs. However, MARCH1 was negatively correlated with TMB and MSI in LUAD (Figure 8C). This contradictory relationship may interpret why TMB and MSI could not always predict the response to ICBs properly. Therefore, we reckoned that the combination of several signatures to predict the sensitivity to ICBs might be a more accurate way. MARCH1 has displayed its powerful modulation in the immune system *via* controlling stability and transforming of some key immunoreceptors, such as the antigen presenting molecule MHC II and costimulatory molecule CD86 (29). Researchers are currently focusing little on the cancer biology of MARCH1 in certain cancers. MARCH1 could inhibit tumor cell growth *in vivo* and *in vitro* in bladder cancer. Meanwhile, ciRs-6 could increase the expression of MARCH1 *via* sponging miR-653 (30). However, Ying Meng et al. discovered that tumor tissue overexpressed MARCH1 relative to paracarcinoma tissues in ovarian cancer (31). Furthermore, the silencing of MARCH1 could restrain the proliferation, migration, and invasion of tumor cells *via* Wnt/β-catenin and nuclear factor-κB pathways (31). Xie L et al. (32, 33) declared that MARCH1 could also provoke tumor progression in hepatocellular carcinoma *via* PI3K-AKT pathway. Collectively, MARCH1 functions differently depending on the type of cancer. This study revealed that MARCH1 was upregulated in some types of cancer and downregulated in others. In most cancers, excluding LGG and KICH, MARCH1 expression was positively associated with most immunomodulators, checkpoints, and infiltrating immune cells (Figures 8A, B, D). Therefore, the role of MARCH1 in pan-cancer requires further investigation. In LUAD, we found that MARCH1 expression was positively correlated with the abundance of different kinds of TIICs, including activated CD4+ and CD8+ cells. Activated CD4+ and CD8+ T cells could kill tumor cells. In addition, CXCL9 and CXCL10, two key chemokines, could recruit CD8+ T cells into the TME (16) and were upregulated in the high-MARCH1 group (Figure 2A). Collectively, we speculated that MARCH1 may regulate CD8+ T cell recruitment to shape an inflamed TME.

As demonstrated previously, MARCH1 expression is essential for immunotherapy responses. However, MARCH1 is suppressed in LUAD tumor tissue, while the factors that regulate MARCH1 transcription are unknown (34). Therefore, the mechanism by which MARCH1 affects cancer immunity and regulation of MARCH1 expression merit additional research. In addition, LUAD with low MARCH1 expression was insensitive to ICBs. Therefore, it is imperative to seek superior treatment options for LUAD patients expressing low levels of MARCH1.

The research on MARCH1 in cancer immunity is poor. This study firstly demonstrated the role of MARCH1 in prognosis and TME shaping. It also revealed the overall correlation between MARCH1 and immunological characteristics and filled up the gap in this field. MARCH1 is a novelty and robust biomarker to predict the response to immunotherapy and some targeted therapy. It provides a theoretical basis for combined therapy. Additionally, MARCH1 may promote infiltration of CD8+ T cells to shape an inflamed TME and further affect immunotherapy sensitivity. It provides a direction for future research.

There were also limitations in this study. Firstly, clinical and animal studies are necessary to validate the expression profiles of MARCH1 and the correlation between MARCH1 and immunological characteristics. Secondly, the optimal cutoff value for grouping the MARCH1 expression must be determined. Thirdly, more cohorts should be used to validate the results to reduce the batch effects.

Conclusions

This study demonstrated that MARCH1 could shape an inflamed TME and predict the prognosis and immunotherapy sensitivity in LUAD. Therapies that target its regulator to upregulate the expression of MARCH1 may be an efficient means of improving 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.

Author contributions

Conceptualization, ZX and HZ. Data collection, JL. Data analysis, ZX and ZL. Writing, ZX and JL. Revising, ZL and HZ. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Medical Scientific Research Foundation of Guangdong Province, China [No. A2021279].

Conflict of interest

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

Publisher's note

References

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

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

3. Fang Y, Yu A, Ye L, Zhai G. Research progress in tumor targeted immunotherapy. *Expert Opin Drug Deliv* (2021) 18(8):1067–90. doi: 10.1080/17425247.2021.1882992

4. Spella M, Stathopoulos GT. Immune resistance in lung adenocarcinoma. *Cancers (Basel)* (2021) 13(3):384–400. doi: 10.3390/cancers13030384

5. Wei SC, Duffy CR. Fundamental mechanisms of immune checkpoint blockade therapy. *Cancer Discov* (2018) 8(9):1069–86. doi: 10.1158/2159-8290.CD-18-0367

6. Hegde PS, Chen DS. Top 10 challenges in cancer immunotherapy. *Immunity* (2020) 52(1):17–35. doi: 10.1016/j.immuni.2019.12.011

7. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discovery* (2019) 18(3):197–218. doi: 10.1038/s41573-018-0007-y

8. Hegde PS, Karanikas V, Evers S. The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin Cancer Res* (2016) 22(8):1865–74. doi: 10.1158/1078-0432.CCR-15-1507

9. Hershko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem* (1998) 67:425–79. doi: 10.1146/annurev.biochem.67.1.425

10. Antao AM, Tyagi A, Kim KS, Ramakrishna S. Advances in deubiquitinating enzyme inhibition and applications in cancer therapeutics. *Cancers (Basel)* (2020) 12(6):15. doi: 10.3390/cancers12061579

11. Cabana VC, Lussier MP. "From drosophila to human: Biological function of E3 ligase Godzilla and its role in disease," (in eng). *Cells* (2022) 11(3):23. doi: 10.3390/cells11030380

12. Nakayama KI, Nakayama K. Ubiquitin ligases: cell-cycle control and cancer. *Nat Rev Cancer* (2006) 6(5):369–81. doi: 10.1038/nrc1881

13. Senft D, Qi J, Ronai ZA. Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer* (2018) 18(2):69–88. doi: 10.1038/nrc.2017.105

14. Hu X, Wang J, Chu M, Liu Y, Wang ZW, Zhu X. Emerging role of ubiquitination in the regulation of PD-1/PD-L1 in cancer immunotherapy. *Mol Ther* (2021) 29(3):908–19. doi: 10.1016/j.ymthe.2020.12.032

15. Bonneville R, Krook M, Kautto E, Jharna M, Wing M, Chen H, et al. Landscape of microsatellite instability across 39 cancer types. *JCO Precis Oncol* (2017) 2017(3):257–71. doi: 10.1200/PO.17.00073

16. Hu J, Yu A, Othmane B, Qiu D, Li H, Li C, et al. Siglec15 shapes a non-inflamed tumor microenvironment and predicts the molecular subtype in bladder cancer. *Theranostics* (2021) 11(7):3089–108. doi: 10.7150/thno.53649

17. L. Xu, Deng C, Pang B, Zhang X, Liu W, Liao G, et al. TIP: A web server for resolving tumor immunophenotype profiling. *Cancer Res* (2018) 78(23):6575–80. doi: 10.1158/0008-5472.CAN-18-0689

18. Ye W, Luo C, Liu F, Liu Z, Chen F. CD96 correlates with immune infiltration and impacts patient prognosis: A pan-cancer analysis. *Front Oncol* (2021) 11:634617. doi: 10.3389/fonc.2021.634617

19. Liu H, Mintern JD, Villadangos JA. MARCH ligases in immunity. *Curr Opin Immunol* (2019) 58:38–43. doi: 10.1016/j.coi.2019.03.001

20. Song M, Wang C, Wang H, Zhang T. Targeting ubiquitin protein ligase E3 component n-recognin 5 in cancer cells induces a CD8+ T cell mediated immune response. *Oncoimmunology* (2020) 9(1):1746148. doi: 10.1080/2162402X.2020.1746148

21. Paolino M, Choidas A, Wallner S, Pranjic B, Uribesalgo I, Loeser S, et al. The E3 ligase cbl-b and TAM receptors regulate cancer metastasis *via* natural killer cells. *Nature* (2014) 507(7493):508–12. doi: 10.1038/nature12998

22. Fujita Y, Tinoco R, Li Y, Senft D, Ronai ZA. Ubiquitin ligases in cancer Immunotherapy - balancing antitumor and autoimmunity. *Trends Mol Med* (2019) 25(5):428–43. doi: 10.1016/j.molmed.2019.02.002

23. Hodges TR, Ott M, Xiu J, Gatalica Z, Swensen J, Zhou S, et al. Mutational burden, immune checkpoint expression, and mismatch repair in glioma: implications for immune checkpoint immunotherapy. *Neuro Oncol* (2017) 19(8):1047–57. doi: 10.1093/neuonc/nox026

24. Voutsadakis IA. High tumor mutation burden and other immunotherapy response predictors in breast cancers: Associations and therapeutic opportunities. *Target Oncol* (2020) 15(1):127–38. doi: 10.1007/s11523-019-00689-7

25. Ruiz-Bañobre J, Goel A. DNA Mismatch repair deficiency and immune checkpoint inhibitors in gastrointestinal cancers. *Gastroenterology* (2019) 156(4):890–903. doi: 10.1053/j.gastro.2018.11.071

26. McGrail D, Garnett J, Yin J, Dai H, Shih D, Lam T, et al. Proteome instability is a therapeutic vulnerability in mismatch repair-deficient cancer. *Cancer Cell* (2020) 37(3):371–86.e12. doi: 10.1016/j.ccell.2020.01.011

27. Li S, Martin A. Mismatch repair and colon cancer: Mechanisms and therapies explored. *Trends Mol Med* (2016) 22(4):274–89. doi: 10.1016/j.molmed.2016.02.003

28. Gajewski TF, Corrales L, Williams J, Horton B, Sivan A, Spranger S. Cancer immunotherapy targets based on understanding the T cell-inflamed versus non-T cell-inflamed tumor microenvironment. *Adv Exp Med Biol* (2017) 1036:19–31. doi: 10.1007/978-3-319-67577-0_2

29. Baravalle G, Park H, McSweeney M, Ohmura-Hoshino M, Matsuki Y, Ishido S, et al. Ubiquitination of CD86 is a key mechanism in regulating antigen presentation by dendritic cells. *J Immunol* (2011) 187:2966–73:no. doi: 10.4049/jimmunol.1101643

30. Su Y, Feng W, Zhong G, Ya Y, Du Z, Shi J, et al. ciRs-6 upregulates March1 to suppress bladder cancer growth by sponging miR-653. *Aging (Albany NY)* (2019) 11(23):11202–23. doi: 10.18632/aging.102525

31. Meng Y, Hu J, Chen Y, Yu T, Hu L. Silencing MARCH1 suppresses proliferation, migration and invasion of ovarian cancer SKOV3 cells via downregulation of NF-κB and wnt/β-catenin pathways. *Oncol Rep* (2016) 36 (5):2463–70. doi: 10.3892/or.2016.5076

32. Xie L, Dai H, Li M, Yang W, Yu G, Wang X, et al. MARCH1 encourages tumour progression of hepatocellular carcinoma via regulation of PI3K-AKT-β-catenin pathways. *J Cell Mol Med* (2019) 23(5):3386–401. doi: 10.1111/jcmm.14235

33. Xie L, Li M, Liu D, Wang X, Wang P, Dai H, et al. Secalonic acid-f, a novel mycotoxin, represses the progression of hepatocellular carcinoma via MARCH1 regulation of the PI3K/AKT/β-catenin signaling pathway. *Molecules* (2019) 24 (3):393–12. doi: 10.3390/molecules24030393

34. Kaul S, Mittal SK, Roche PA. A major isoform of the E3 ubiquitin ligase march-I in antigen-presenting cells has regulatory sequences within its gene," (in eng). *J Biol Chem* (2018) 293(12):4478–85. doi: 10.1074/jbc.RA118.001775

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 17 October 2022
DOI 10.3389/fonc.2022.1031728

OPEN ACCESS

EDITED BY
Mingyue Tan,
Shanghai University of Traditional Chinese Medicine, China

REVIEWED BY
Wei Shenyu,
Zhejiang Chinese Medical University, China
Jiahang Song,
The First Affiliated Hospital of Nanjing Medical University, China

*CORRESPONDENCE
Haitao Ma
soochowmht@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 30 August 2022
ACCEPTED 28 September 2022
PUBLISHED 17 October 2022

CITATION
Shi Q, Han S, Liu X, Wang S and Ma H (2022) Integrated single-cell and transcriptome sequencing analyses determines a chromatin regulator-based signature for evaluating prognosis in lung adenocarcinoma. *Front. Oncol.* 12:1031728. doi: 10.3389/fonc.2022.1031728

Integrated single-cell and transcriptome sequencing analyses determines a chromatin regulator-based signature for evaluating prognosis in lung adenocarcinoma

Qingtong Shi[1,2†], Song Han[3†], Xiong Liu[2,4], Saijian Wang[2,4] and Haitao Ma[1*]

[1]Department of Thoracic Surgery, The First Affiliated Hospital of Soochow University, Suzhou, China, [2]Department of Thoracic Surgery, The Affiliated Hospital of Yangzhou University, Yangzhou, China, [3]Department of Thoracic Surgery, Suzhou Science and Technology Town Hospital, Suzhou, China, [4]Graduate School of Dalian Medical University, Dalian, China

Background: Accumulating evidence has highlighted the significance of chromatin regulator (CR) in pathogenesis and progression of cancer. However, the prognostic role of CRs in LUAD remains obscure. We aim to detect the prognostic value of CRs in LUAD and create favorable signature for assessing prognosis and clinical value of LUAD patients.

Methods: The mRNA sequencing data and clinical information were obtained from TCGA and GEO databases. Gene consensus clustering analysis was utilized to determine the molecular subtype of LUAD. Cox regression methods were employed to set up the CRs-based signature (CRBS) for evaluating survival rate in LUAD. Biological function and signaling pathways were identified by KEGG and GSEA analyses. In addition, we calculated the infiltration level of immunocyte by CIBERSORT algorithm. The expressions of model hub genes were detected in LUAD cell lines by real-time polymerase chain reaction (PCR).

Results: KEGG analysis suggested the CRs were mainly involved in histone modification, nuclear division and DNA modification. Consensus clustering analysis identified a novel CRs-associated subtype which divided the combined LUAD cohort into two clusters (C1 = 217 and C2 = 296). We noticed that a remarkable discrepancy in survival rate among two clusters. Then, a total of 120 differentially expressed CRs were enrolled into stepwise Cox analyses. Four hub CRs (CBX7, HMGA2, NPAS2 and PRC1) were selected to create a risk signature which could accurately forecast patient outcomes and differentiate patient risk. GSEA unearthed that mTORC1 pathway, PI3K/Akt/mTOR and p53 pathway were greatly enriched in CRBS-high cohort. Moreover, the infiltration percentages of macrophage M0, macrophage M2, resting NK cells, memory B cells, dendritic cells and mast cells were statistically significantly different in

the two groups. PCR assay confirmed the differential expression of four model biomarkers.

Conclusions: Altogether, our project developed a robust risk signature based on CRs and offered novel insights into individualized treatment for LUAD cases.

KEYWORDS

lung adenocarcinoma, chromatin regulator, prognosis, risk signature, immune microenvironment

Introduction

Lung cancer (LC) is the major cause of death for men and women with tumor, representing approximately 18% of all cancer deaths worldwide (1). Up to 90% of LC cases are non-small cell lung cancer (NSCLC), including both lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) histological subtypes, with LUAD occurring most frequently (2). Despite recent advances in clinical treatment, the prognosis for LUAD remains dismal, with a 5-year survival rate of only 19%. With the advent of aging and air pollution in developing countries, the incidence of LUAD remains high and early diagnosis of LUAD becomes essential (3). Unfortunately, we still have limited availability of accurate biomarkers for early diagnosis and individualized treatment of LUAD.

Tumor microenvironment (TME) is the internal environment for tumor cell production and survival, and its cellular components include resident stromal cells and recruited immunocytes in addition to tumor cells (4). TME plays an important role in the tumor growth, metastasis, angiogenesis and treatment resistance and has also a crucial impact on prognosis (5). Therefore, systematic exploration of TME is helpful to clarify the mechanism of tumor occurrence and individualized treatment.

Epigenetic modification is a reversible and heritable process of gene expression in the absence of DNA sequence changes. It is one of the critical regulatory mechanisms at the post-transcriptional level of genes by chromatin regulators (CRs), mainly including DNA methylation, histone modifications, chromatin remodeling and RNA regulation (6). CRs-mediate epigenetic modification regulates the activation of heterozygous promoters or the activity of repressors and trigger changes in gene transcription levels, resulting in cell differentiation, abnormal proliferation and tumorigenesis (7).

Numerous studies have demonstrated that CRs are tightly bound up to the patient outcomes of LC HMGA1, a chromatin remodeler, has been reported to be involved in DNA transcription, replication and repair. Saed and his colleagues have observed that HMGA1 presented higher expression in lung cancer specimens and overexpressed HMGA1 lead to dismal prognosis of LUAD (8). Moreover, HMGA1 was proved to facilitate LUAD cell proliferation and migration through GRP75-induced JNK pathway (9). EZH2, belonging to the polycomb-group (PcG) family, has been reported to be greatly overexpressed in lung specimens, and upregulation of EZH2 predicts dismal survival of NSCLC (10). Geng and his colleagues indicated EZH2 enhances the growth and metastasis of lung cell by Akt pathway (10). RAD51 is well known for its important role in homologous recombination. RAD51 has shown to be upregulated in KRAS mutant lung cancer and could regulate cell survival by enhancing DNA damage repair (11). However, the expression patterns and prognostic value of CRs in LUAD remain largely unknown.

In this academic research, we determined CRs with powerful prognostic values in LUAD and created a risk signature for clinical outcome assessment and immune status prediction of LUAD cases.

Methods

Data collection and processing

We obtained the RNA sequencing (RNA-seq) data of 535 LUAD patients and 59 normal controls and their corresponding clinical features from TCGA database (https://portal.gdc.cancer.gov/) to construct the prognostic signature. The transcription profiling data was downloaded from GEO dataset and was utilized as the validate set.

Determination of differentially expressed CRs

A total of 870 CRs were retrieved from previous research (6). The gene information of all CRs summarized in Supplementary Table S1. The differentially expressed genes (DEGs) between normal and LUAD tumor tissues were determined using the limma R package with a criteria P value<0.05 (12). The generated DEGs and CRs gene sets were subsequently intersected to obtain differentially expressed CRs (DECRs).

Function and pathways enrichment analyses

GO and KEGG enrichment analysis was conducted to obtain the insight into the biological functions and potential pathways of DECRs. Terms with p< 0.05 were listed and visualized using the "clusterProfiler" R packages (13).

Integration of protein–protein interaction (PPI) network

A protein–protein interaction network (PPI) was developed and visualized using the STRING online database (https://cn.string-db.org/) and the Cytoscape (https://cytoscape.org/), respectively (14, 15). Further, the cubHubba plugin in Cytoscape software was used to filter hub genes of the PPI.

Gene consensus cluster analysis

The consensus cluster analysis was conducted using the "ConsensusClusterPlus" R package, based on the combined LUAD cohort (16). To identify the optimal cluster value, we calculated the Delta area and the cumulative distribution function (CDF). Survival analysis was carried out to compare clinical prognoses between different subtypes using "survival" R package.

Construction of the risk signature

Subsequently, Cox regression analyses were performed to obtain candidate CR with remarkable prognostic value. The formula was set up: Risk score = $\sum_{i=1}^{n}(\text{coef} \times \text{Exp}_i)$. "Coef" was defined as the corresponding regression coefficient value, and "Exp" was the expression level of genes in the prognostic model. All patients were divided into low- and high-risk groups according to the median score.

Gene set enrichment analysis (GSEA)

We performed GSEA analysis, including GO and KEGG analysis based on CR related DEGs to identify the potential biological and functional differences of different hierarchical clustering (17). A function term with an adjusted p-value<0.05 and a false discovery rate (FDR)<0.25 was considered enriched.

Estimate of immune infiltrating status

CIBERSORT tool (https://cibersortx.stanford.edu/index.php) was employed to quantify the infiltration status of 22 types of immunocyte fractions in the two LUAD subgroups. P< 0.05 was defined as statistically significant.

Single-cell sequencing analysis

We utilized the Seurat clustering to analyze the single-cell data acquired from the GEO databases. The UMAP dimensional reduction and the t-Distributed Stochastic Neighbor Embedding (tSNE) method, were employed to visualize the gene expression and distribution in dataset GSE131907. Next, the cells were re-clustered with the "SingleR" packages to demonstrate the feature genes of different cell types.

Validation of the model CRs

To detect the expression pattern of a model gene at the mRNA level, GEPIA2 tool was applied. Human Protein Atlas (HPA, https://www.proteinatlas.org/) database was utilized to confirm the protein level of our model genes between LUAD and normal control (18).

Somatic mutation and stem cell characteristics analyses

The somatic mutation data were obtained from TCGA Portal and processed to compare the tumor mutation burden (TMB) in two groups. The mRNAsi is a quantitative index reflecting cancer cells calculated based on gene profiles; The mRNAsi and TMB differences in two subgroups were compared using the independent-samples t-test.

Cell culture

Two human LUAD cell lines (A549 and NCI-H460) and a normal human lung epithelial cell line (BEAS-2B) were purchased from American Type Culture Collection. All cell lines were cultured in RPMI 1640 medium (Sigma) containing 10% fetal bovine serum (Gibco) and 1% antibiotics (100 U/ml penicillin G and 100mg/ml streptomycin) at 37°C in a humidified chamber containing 5% CO2.

RNA extraction and quantitative real-Time PCR

Total cell RNA was extracted by RNA isolation reagent (Takara), then reversed into cDNA by PrimeScript Mix reagent (Takara). SYBR Green Premix (Vazyme biotech) was utilized for PCR reaction system. The value of individual genes

was standardized to GAPDH expression level. Supplementary Table S2 displays primer sequences of all genes.

Statistical analysis

All statistical data in the present project was analyzed by R version 4.0.5 and GraphPad Prism 9. The Kaplan-Meier (KM) analysis was employed to assess the prognostic value of the signature. Moreover, we plotted the receiver operating characteristic (ROC) curve over time to evaluate the prognostic efficacy of the signature.

Results

Characterization of chromatin regulators in LUAD

We first collected 4846 DEGs between LUAD samples and normal cases. A total of 120 DECRs were obtained by taking the intersection of CRs and DEGs gene lists (Figure 1A). Then, GO analysis was employed to detect the underlying function of DECRs. The result disclosed that these genes were mainly enriched in histone modification and DNA modification (Figure 1B). Next, we generated a PPI network to explore the protein interaction among 120 DECRs (Figure 1C). Based on the MCC algorithm, the top ten hub genes were selected to set up a hub network, including CHEK1, CDK1, TOP2A, CDC6, UHRF1, AURKB, PBK, BUB1, TTK and RAD54L (Figure 1D).

Chromatin regulators-based consensus cluster analysis

TCGA-LUAD and GSE14520 were combined into one LUAD cohort (n = 609). We applied consensus cluster analysis to develop a CR-related molecular subtype of LUAD.

The result suggested the entire dataset could be well divided into two subtypes based on the 120 DECRs when k = 2 by increasing the clustering variable (k) from 2 to 9 (Figures 2A–C). PCA analysis shows that DECRs can clearly distinguish two subgroups for clustering analysis (Figure 2D). There were remarkable discrepancies in survival rates among the two clusters (Figure 2E). To evaluate the TME status of two clusters, ESTIMATE algorithm was conducted. As suggested in

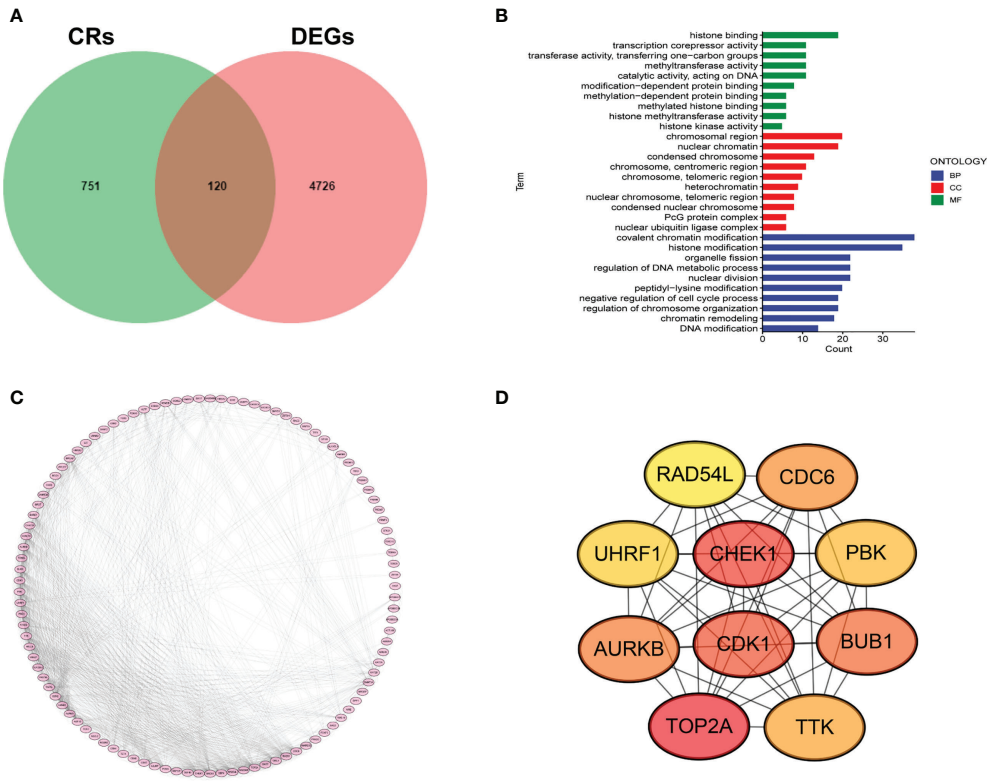


FIGURE 1
Determination of DECRs in LUAD. **(A)** Venn diagram of DECRs. **(B)** KEGG analysis of DECRs. **(C)** PPI network for 120 DECRs. **(D)** The top ten hub genes of DECRs-based PPI network.

Figure 2F, cluster A presented higher stromal score and immune score than that cluster B. In addition, we observed that B cells, T cells, NK cells, dendritic cells and Macrophages showed the most notable difference among the two clusters (Figure 2G).

Construction of the CRBS

To develop an optimal prognostic signature, TCGA-LUAD cohort was selected as the training set. Univariate Cox regression was first employed to determine possible CRs with significant prognostic values (Figure 3A). Subsequently, 12 candidate genes were enrolled into multivariate Cox analysis to create a CRBS that included four risk CRs (Figure 3B). The risk formula was shown as follows: (0.1082 × HMGA2) + (0.3525 × NPAS2) + (0.1909 × PRC1) + (-0.2416 × CBX7). Survival curves illustrated that CBX7 was a potential favorable indicator, and HMGA2, NPAS2 and PRC1 were risky candidate indicators (Figure 3C). Then, we detect the expression differences of four CRs according to TCGA-LUAD dataset. All four CRs were greatly dysregulated between LUAD cases and control samples (Figure 3D). Furthermore, we validated the expression patterns of four

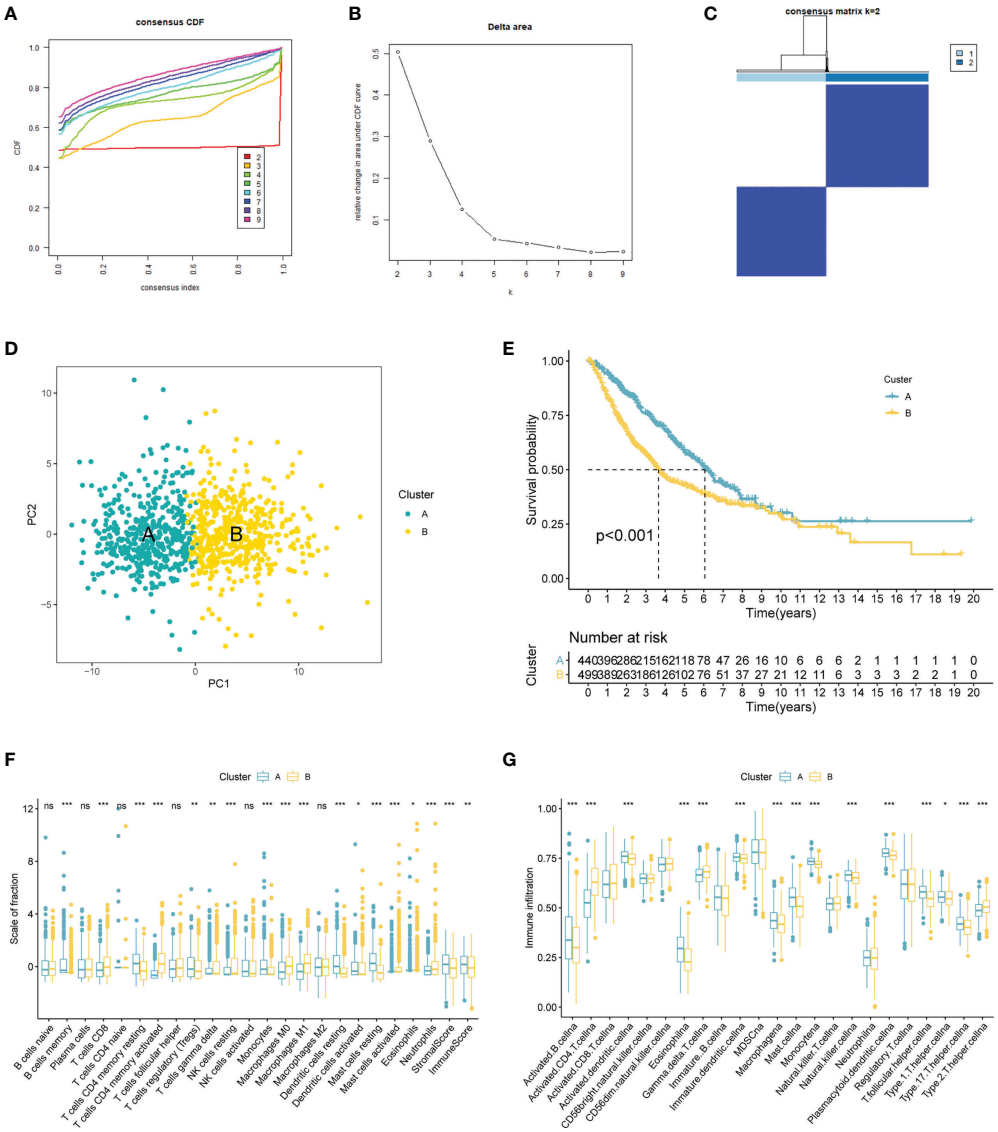


FIGURE 2
CRs-associated clustering analysis. **(A, B)** The CDF value of consensus index. **(C)** Consensus matrix for k=2. **(D)** Principal component analysis of the entire LUAD set. **(E)** The Kaplan–Meier survival analysis. **(F)** Comparison of immunocyte proportions of the CRs-based clusters. **(G)** The differences of TME score 22 the two clusters. ns, no significance; 0.05; *p < 0.05; **p < 0.01; ***p < 0.001.

model genes by qRT-PCR in cell lines. Consistent with the above bioinformatics analysis results, we noticed that CBX7 was downregulated in LUAD cell lines (A549 and HCI-H1975), and HMGA2, NPAS2 and PRC1 were overexpressed in LUAD cell lines compared to BEAS-2B (Figure 3E). Consistent with the above results, we detected the expression patterns of four CRs at IHC level based on HPA database (Figure 3F).

Verification of the CRBS

Figure 4A demonstrated that survival rates are lower in CRBS-high group compared to CRBS-low group in the training set. The AUC (area under the curve) values of 1-, 3-, and 5-year survival rates assessed by the CRBS were 0.729, 0.662, and 0.634, respectively (Figure 4B). Figure 4C summarizes the positive correlation between surviving cases and risk score. Moreover, we observed a similar trend of results in the test set, suggesting the favorable prediction ability of the CRBS (Figures 4D–F). To further unearth the independence of our model, univariate and multivariate Cox regression analyses were employed. Univariate analysis indicated that the risk score was an independent indicator for prognosis in both two datasets (Figures 4G, I). The multivariate method disclosed that risk score was independently associated with the dismal outcome of LUAD cases (Figures 4H, J). At the same time, we explore the performance of the CRBS based on a diversity of clinical subgroups. The results revealed that low risk score was correlated with favorable outcomes in different ages, genders, T stage and N stage cohorts (Figures 5A–D). Similarly, the good prediction capability of the CRBS was confirmed in the T stage and N stage subgroups (Figures 5E, F).

Single-cell sequencing analysis

To decipher the single-cell transcriptome dataset GSE131907, Seurat package was performed. The UMAP analysis suggested the distribution of the 22 LUAD samples (N = 11 and T = 11) with no remarkable batch effects (Figure 6A). All the cells were divided into 12 clusters the through k- Nearest Neighbor (KNN) clustering algorithm (Figure 6B). After performing cell annotation by different cell surface markers, we obtained eight cell subtypes, including B lymphocytes, endothelial cells, epithelial cells, fibroblasts, mast cells, myeloid cells, NK cells and T lymphocytes (Figure 6C). Next, we investigate the location of four CRs at single-cell transcriptome level. In Figure 6D,

HMGA2 and NPAS2 are mainly located in endothelial cells, and PRC1 and CBX7 are mainly located in NK cells and T lymphocytes. In addition, we noticed that the expression of CBX7 was negatively correlated with endothelial cells, whereas NPAS2 was positively correlated with endothelial cells (Figure 6E).

GSEA determines CRBS-associated pathways

In Figure 7A, the top six cancer hallmarks were remarkably enriched in the CRBS-high group, including glycolysis, hypoxia, mTORC1 pathway, MYC target, PI3K/Akt/mTOR and unfolded protein response. In terms of the KEGG analysis, we observed that CRBS-high group was involved in the cell cycle, p53 pathway and ubiquitination response (Figure 7B).

Immune environment analysis

To depict the immune landscape of LUAD, we evaluated the immunocyte infiltration of each case. Figure 8A summarizes the correlation between the 22 immunocyte types. As suggested in Figure 8B, CBX7 was greatly positively associated with memory B cells and resting mast cells. PRC1 was positively associated with activated memory T cells and negatively correlated with resting mast cells. Moreover, macrophage M0, macrophage M2 and resting NK cells were enriched in the CRBS-high group. Cases in CRBS-low group had greatly higher proportions of memory B cells, dendritic cells and mast cells (Figures 8C–H).

Additionally, some immune functions displayed differences between the two groups, including APC co-stimulation, checkpoint, HLA, MHC class I, T cell co-stimulation, and type II IFN response (Figure 9A). Also, we observed that four immune responses (checkpoint, HLA, MHC class I and type II IFN response) had significant differences in the outcome of patients with LUAD (Figures 9B–E).

Clinical potency analysis of the CRBS

TMB has been demonstrated to be useful as an indicator of the efficacy of immunotherapy. We calculated the TMB of each LUAD sample and found that CRBS-high group had a higher TMB than the CRBS-low group (Figures 10A). Moreover, CRBS-high group presented a high level of mRNAsi (Figure 10B). In Figure 10C, most of the immune checkpoint markers were upregulated in the CRBS-high group. The comparison in the expression of m6A markers between the two groups indicated that the expression of ALKBH5, FTO, METTL14, HNRNPC, YTHDF1, YTHDF2, METTL3, RBM15 and WTAP were significant (Figure 10D).

Discussion

LUAD is the most common pathological subtype of lung cancer, which is composed of approximately 40% of lung cancer cases (19). Despite the various efforts in improve, the five-year

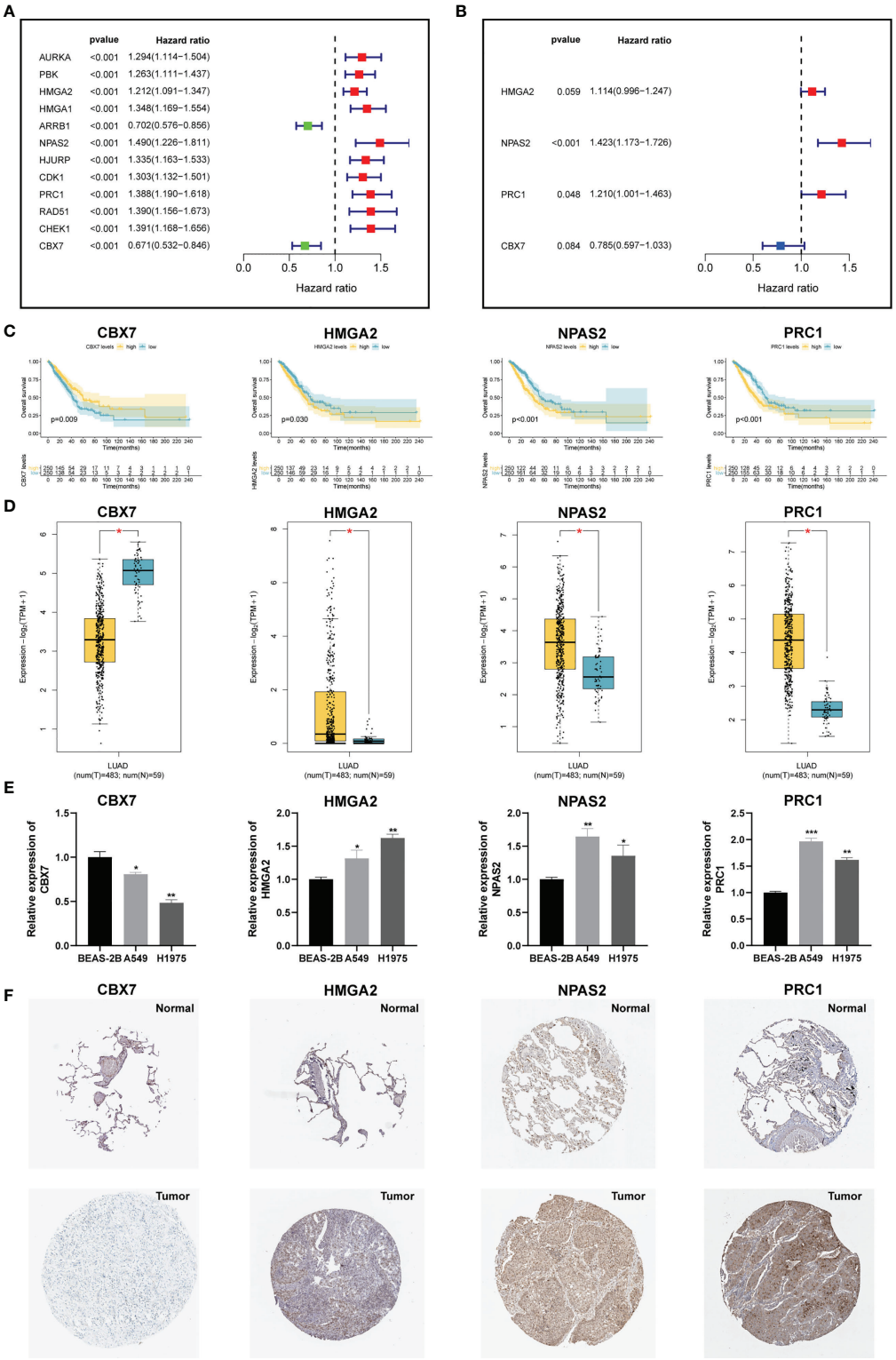


FIGURE 3
Construction of a CRs-based signature (CRBS). **(A, B)** Univariate and multivariate Cox analyses for signature establishment. **(C)** Survival analysis of four model CRs. **(D)** Comparison of differential expression of four model CRs based on GEPIA2 online portal. **(E)** The expression of four model CRs in BEAS-2B, A549 and HCI-H1975 cells line. **(F)** Immunohistochemistry of the CGs according to the HPA database. *p < 0.05; **p < 0.01; ***p < 0.001.

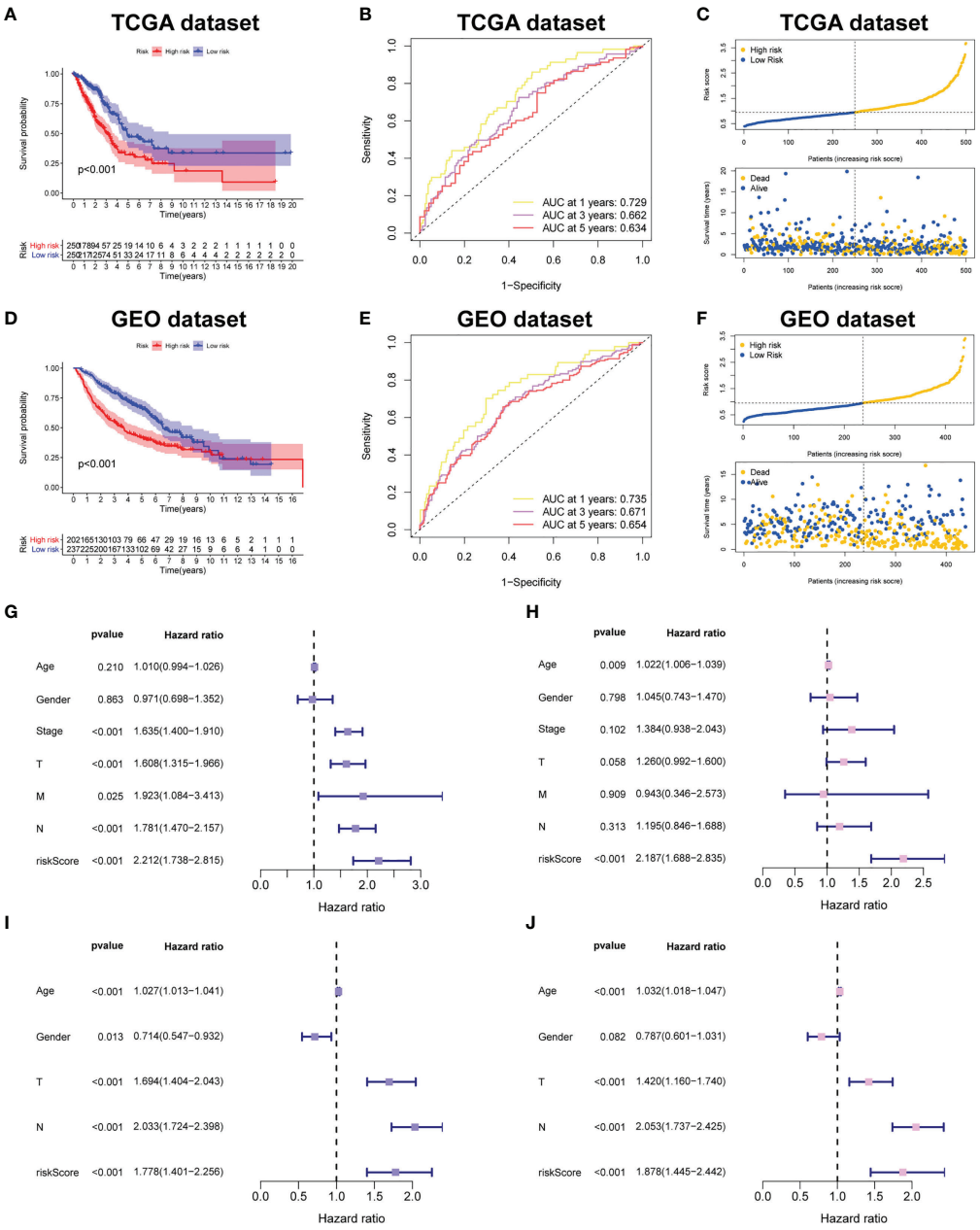


FIGURE 4
Prognostic powerful of the CRBS. **(A, D)** Survival analysis in the TCGA-LUAD and the GSE68465 cohorts. **(B, E)** ROC curves of the CRBS. **(C, F)** The risk distribution plots in two datasets. **(G–J)** Cox relevant regression assessing the independence of the CRBS.

survival rate for LUAD patients remains shabby. Recent studies have suggested that patients with the same histology and TNM stages may have very distinct clinical outcomes, mainly due to their genetic heterogeneities (20). With the rapid development of the next-generation sequencing, a growing number of prognostic signatures based on transcriptome data were established to depict the individual differences, and to forecast the prognosis in various cancers (21–23). Therefore, a more reliable prognostic model based on genetic alterations is urgently needed to provide early detection and personalized treatment for LUAD patients.

It is well known that epigenetic alterations play a considerable role in mediating the tumor progression (24). As indispensable regulatory elements of epigenetics, CRs are involved in the onset and development of various cancer types including multiple myeloma, prostate cancer, hepatocellular carcinoma, and LUAD (25–28). In our current work, a total of

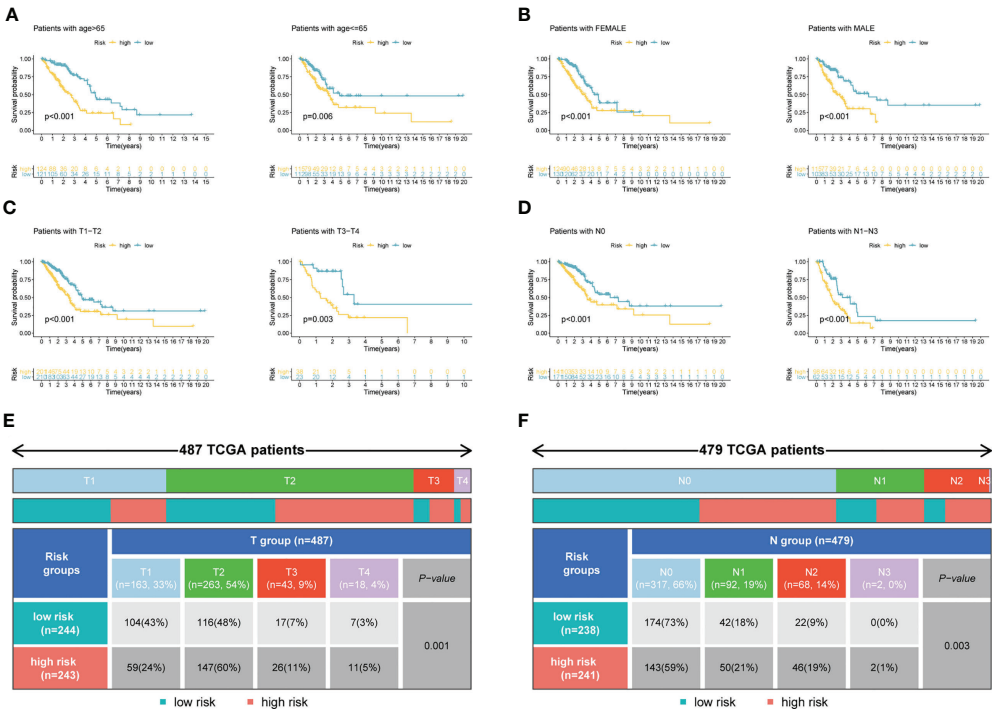


FIGURE 5
Subgroup survival analysis. **(A)** Age subgroup. **(B)** Gender subgroup. **(C)** T stage subgroup. **(D)** N stage subgroup. **(E, F)** Table presenting the distribution of T and N stage subgroups between two risk groups.

four CRs including HMGA2, NPAS2, PRC1, and CBX7, were identified as effective prognostic biomarkers for predicting the prognosis of LUAD. Survival analysis indicated that HMGA2, NPAS2 and PRC1 are potential risky genes since their high expressions are correlated with dismal outcomes of LUAD samples, whereas CBX7 is candidate protective factor given that its high expression is associated with favorable outcomes of LUAD samples. The pro-tumor role of HMGA2 has been

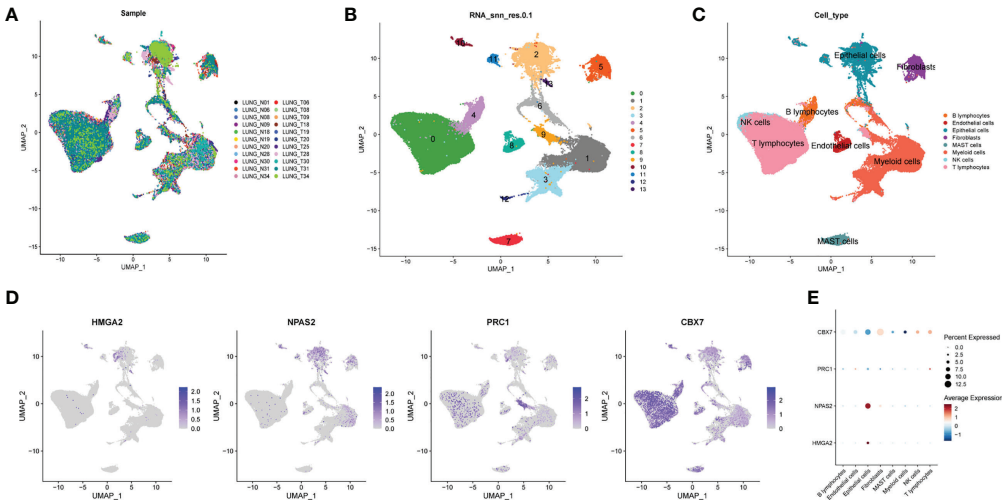


FIGURE 6
Single cell sequencing analysis. **(A)** The integration effect of 22 samples is favorable. **(B)** All cells in 22 samples were divided into 13 subgroups. **(C)** The cells were divided into 8 types of cell subgroups, namely B lymphocytes, endothelial cells, epithelial cells, fibroblasts, mast cells, myeloid cells, NK cells and T lymphocytes. **(D)** Cell location of four model CRs. **(E)** Correlation analysis of four model CRs and 8 types of cell subgroups.

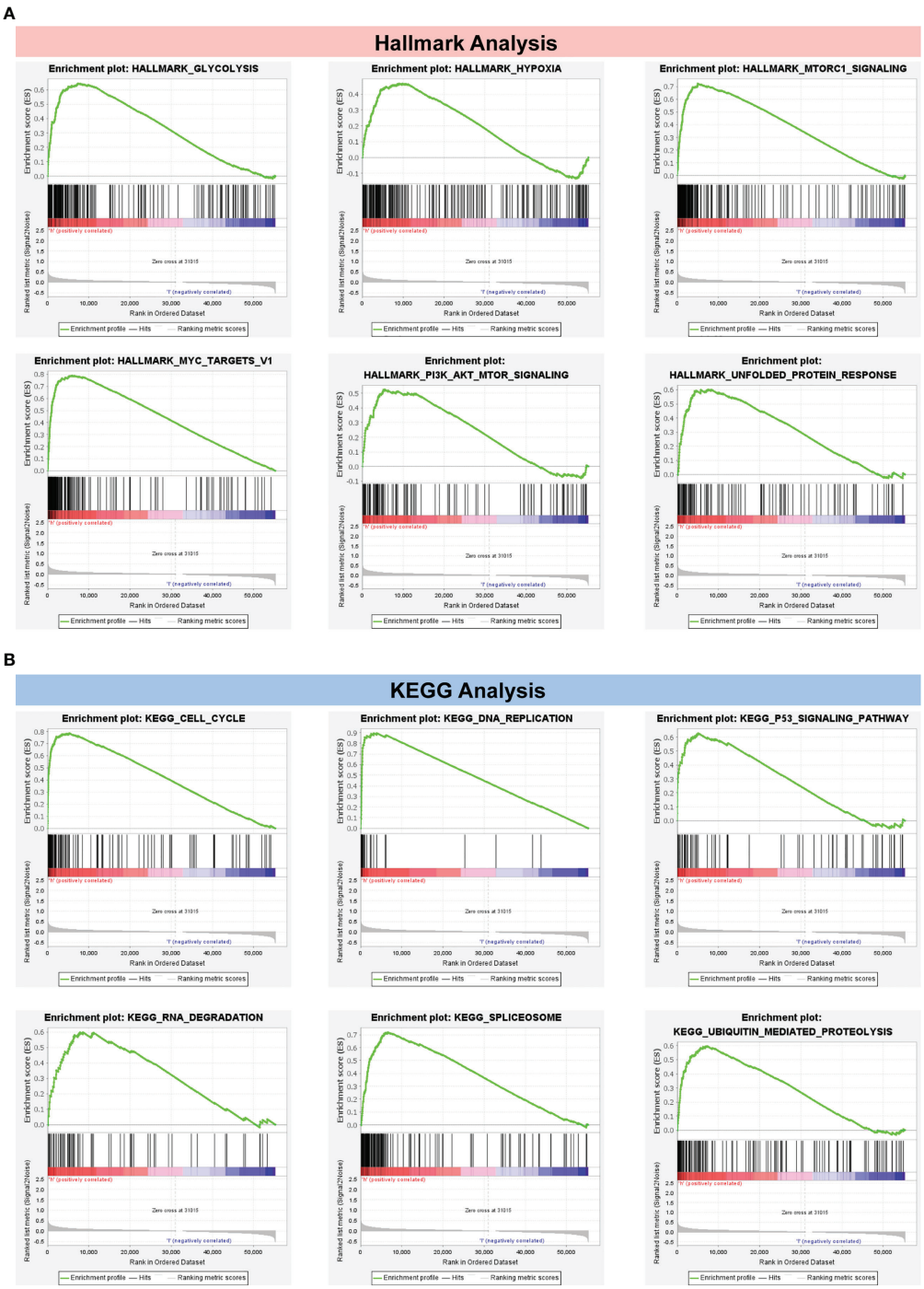


FIGURE 7
Gene Set Enrichment Analysis. **(A)** Hallmark analysis of the CRBS. **(B)** KEGG analysis of the CRBS.

widely reported in multiple cancers including LUAD (29). HMGA2 was found highly expressed in the LUAD tissues compared with normal lung tissues, and HMGA2 silencing notably reduced the growth and metastasis of LUAD cell lines (30). In addition, a mechanistic study revealed that HMGA2 could induce epithelial-mesenchymal transition by activating MAPK/extracellular receptor kinase signaling in LUAD (31). Npas2 has been identified in peripheral tissues, possibly as a modulator of circadian rhythms (32). Qiu et al. once reported that in LUAD, the elevated expression level of NPAS2 is

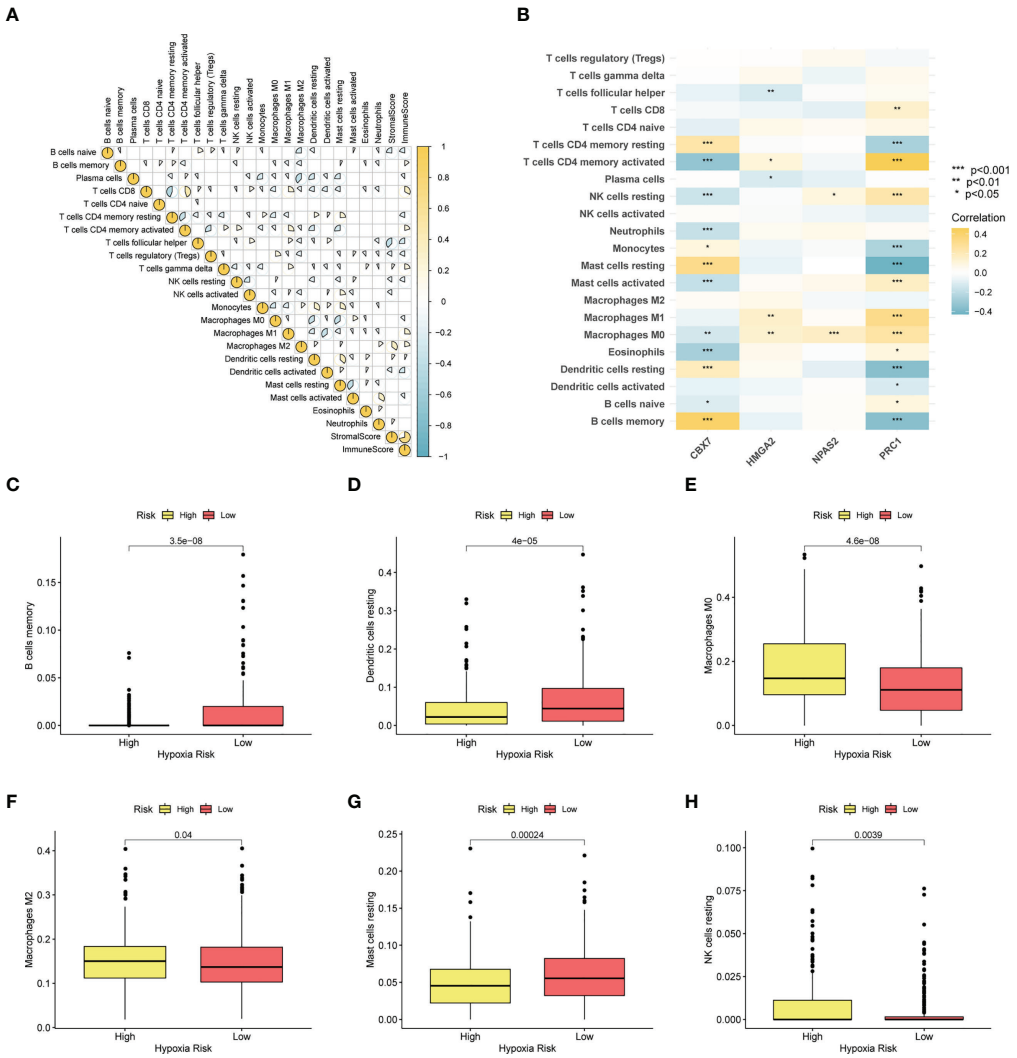


FIGURE 8
Immune infiltration analysis. **(A)** Correlation analysis of 22 immunocyte types. **(B)** Heatmap showing the relationship between four model CRs and immune cells. **(C–H)** The infiltration level differences of memory B cells, dendritic cells, macrophage M0, macrophage M2, mast cells and NK cells among two groups (*p< 0.05; **p< 0.01; ***p< 0.001).

significantly related to poor prognosis (33). Conversely, it has been indicated that LUAD cases with low NPAS2 expression displayed a favorable clinical outcome by another team (34). Therefore, more basic researches are needed to elucidate the exact role of NPAS2 in LUAD. PRC1 has received widespread attention considering its diverse regulatory roles in a number of diseases, especially tumorigenesis (35). It has been suggested that overexpression of PRC1 triggers the onset of various cancers yet its potential roles in LUAD have not been fully understood (36). An ever-growing series of reports has demonstrated the aberrant expression of CBX7 in a variety of tumors (37). Mechanically, CBX7 may exert its tumor suppressor role by inhibiting the Wnt pathway and subsequently restrain the malignant character in LUAD (38).

GSEA unearthed that CRBS-high group were involved in glycolysis signaling, PI3K/AKT/mTOR signaling, and p53 signaling pathway using GSEA. Suppressed oxidative phosphorylation along with enhanced glycolysis, which is called the Warburg phenotype, is considered as metabolic marker of cancers (39). Vaupel et al. once reported that enhanced glycolysis accelerates lactic acid accumulation to impair the immune functions in TME and finally promote malignant progression (40). The PI3K/AKT/mTOR pathway plays a crucial role in diverse biological behaviors including cell growth, migration, metabolism, and death (41). In LUAD, the aberrant activation in this signaling has been indicated to induce uncontrolled growth, drug resistance, sustained angiogenesis, as well as distant metastasis (42). P53 protein

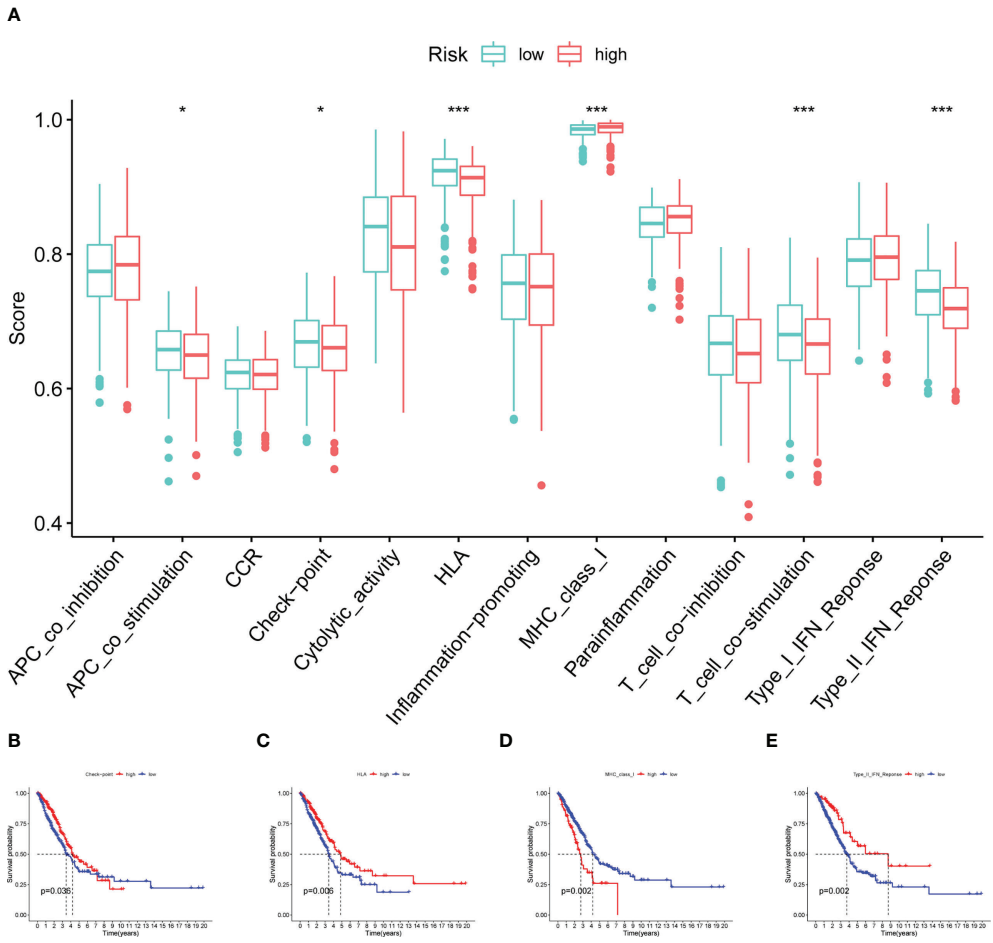


FIGURE 9
Immune function analysis. **(A)** Boxplot showing the relationship between four model CRs and 13 immune functions. **(B–E)** Survival analysis for checkpoint, HLA, MHC class I and type II IFN response (*p< 0.05; ***p< 0.001).

is a transcription factor known as the "guardian gene" because of its significant role in preserving genomic integrity. The mutation of the p53 gene can be detected in a wide spectrum of human malignancies, including the breast, cervical, lung, and prostate cancer (43). More recently, Vokes and his colleagues provided the evidence that p53 alterations were involved in faster resistance evolution and may cooperate with other genomic events to gain resistance to EGFR tyrosine kinase inhibitors (44).

Immunotherapy that emerged recently has achieved promising results in the treatment of LUAD (45). In our work, a comprehensive analysis of tumor-infiltrating immune cells was further conducted to help to clarify the immune infiltration status between the two different risk groups. As a result, the infiltration level of HLA as well as the type 2 IFN was found downregulated remarkably in CRBS-high group. Also, the expression level of the immune checkpoint markers was validated to be correlated with the risk score. CD273, also named B3-H7, is overexpressed in various solid malignancies which serve as a potential therapeutic target (46, 47). Yu and his colleagues disclosed CD273 was upregulated in LUAD, and was correlated with lymph node metastasis (48). Likewise, accumulating studies have indicated the close association between the efficacy of immunotherapy and the CD274 expression (49). VTCN1, also named B7-H4, belongs to the co-stimulatory B7 family molecules and is associated with a poor prognosis in multiple cancer types (50–52). As revealed by a recent study, the elevation of VTCN1 expression is associated with LUAD with EGFR-activating mutations, which can ultimately cause resistance to immunotherapy in LUAD patients (53).

In view of the essential effect of m6A methylation modification in LUAD progression, we unearthed the

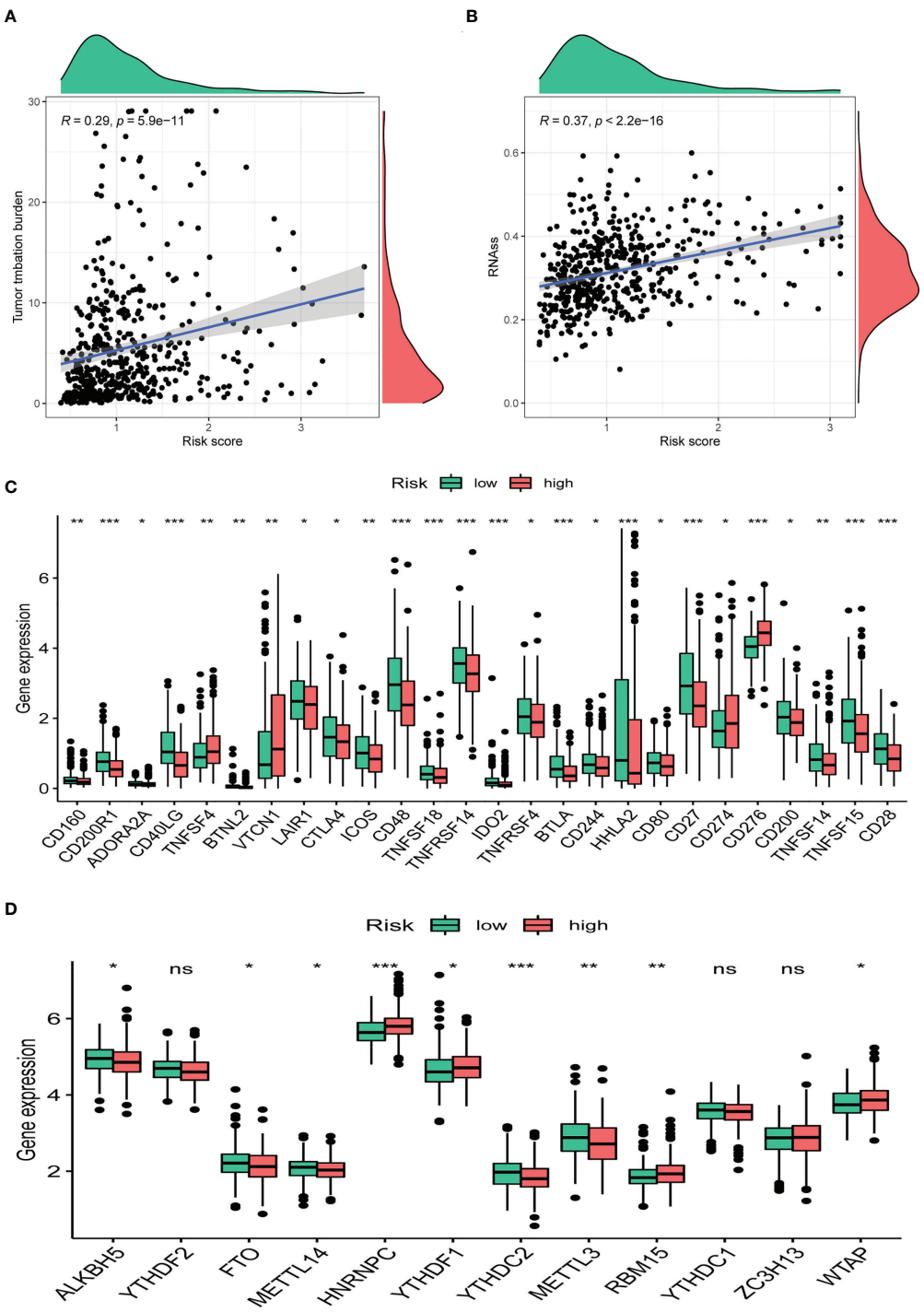


FIGURE 10
Clinical potency analysis. **(A)** TMB analysis of the CRBS. **(B)** Cancer stem cells index analysis the CRBS. Comparison of differential expression of **(C)** immune checkpoints and **(D)** m6A markers (*p< 0.05; **p< 0.01; ***p< 0.001). ns, no significance.

expression patterns of m6A regulators between two risk groups. The results indicated that HNRNPC, YTHDF1, RBM15 and WTAP were enriched in the high-risk group. Lou and his colleagues demonstrated that YTHDF1 could facilitate LUAD growth and survival by enhancing Cyclin B1 translation (54). In addition, YTHDF1 has also been confirmed to have carcinogenic effects in many digestive system tumors including gastric cancer, hepatocellular carcinoma and colorectal cancer (55–57). Cheng

et al. found that overexpression of WTAP correlate with dismal outcome of LC cases. In NSCLC, PCGEM1 could boost cancer cells proliferation by improving WTAP expression (58, 59).

There are still several limitations of the present study that need to be considered. Only expression data in gene level was analyzed to construct the prognostic model, and large-sample clinical data are still needed, as an external cohort, to evaluate the predictive value of our model. Additionally, although we have proven the reliable prognostic capacity of the four CR related genes, fundamental experiments are still needed to validate their precise functions in mediating LUAD progression.

Conclusion

Taken together, our data may help provide opportunities for the development of new therapeutic strategies and elucidate the mechanism of tumor immune escape in LUAD. Our proposed model may usher in novel approaches to predicting prognosis of patients with LUAD.

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

QS and HM visualized the study and took part in the study design. QS, SH, XL, and SW performed the manuscript writing and bioinformatics analysis. All authors read and approved the final manuscript.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1031728/full#supplementary-material

References

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

2. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc* (2008) 83:584–94. doi: 10.4065/83.5.584

3. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* (2012) 150:1107–20. doi: 10.1016/j.cell.2012.08.029

4. Bejarano L, Jordão MJC, Joyce JA. Therapeutic targeting of the tumor microenvironment. *Cancer Discov* (2021) 11:933–59. doi: 10.1158/2159-8290.CD-20-1808

5. Jin M-Z, Jin W-L. The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct Target Ther* (2020) 5:166. doi: 10.1038/s41392-020-00280-x

6. Lu J, Xu J, Li J, Pan T, Bai J, Wang L, et al. FACER: Comprehensive molecular and functional characterization of epigenetic chromatin regulators. *Nucleic Acids Res* (2018) 46:10019–33. doi: 10.1093/nar/gky679

7. Plass C, Pfister SM, Lindroth AM, Bogatyrova O, Claus R, Lichter P. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat Rev Genet* (2013) 14:765–80. doi: 10.1038/nrg3554

8. Saed L, Jeleń A, Mirowski M, Sałagacka-Kubiak A. Prognostic significance of HMGA1 expression in lung cancer based on bioinformatics analysis. *Int J Mol Sci* (2022) 23:6933. doi: 10.3390/ijms23136933

9. Qiao G-B, Wang R-T, Wang S-N, Tao S-L, Tan Q-Y, Jin H. GRP75-mediated upregulation of HMGA1 stimulates stage I lung adenocarcinoma progression by activating JNK/c-JUN signaling. *Thorac Cancer* (2021) 12:1558–69. doi: 10.1111/1759-7714.13944

10. Wang X, Zhao H, Lv L, Bao L, Wang X, Han S. Prognostic significance of EZH2 expression in non-small cell lung cancer: A meta-analysis. *Sci Rep* (2016) 6:19239. doi: 10.1038/srep19239

11. Hu J, Zhang Z, Zhao L, Li L, Zuo W, Han L. High expression of RAD51 promotes DNA damage repair and survival in KRAS-mutant lung cancer cells. *BMB Rep* (2019) 52:151–6. doi: 10.5483/BMBRep.2019.52.2.213

12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43:e47–7. doi: 10.1093/nar/gkv007

13. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: An r package for comparing biological themes among gene clusters. *Omics J Integr Biol* (2012) 16:284–7. doi: 10.1089/omi.2011.0118

14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* (2003) 13:2498–504. doi: 10.1101/gr.1239303

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

16. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26:1572–3. doi: 10.1093/bioinformatics/btq170

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

18. Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome: The human protein atlas. *Protein Sci* (2018) 27:233–44. doi: 10.1002/pro.3307

19. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* (2018) 68:7–30. doi: 10.3322/caac.21442

20. Benusiglio PR, Fallet V, Sanchis-Borja M, Coulet F, Cadranel J. Lung cancer is also a hereditary disease. *Eur Respir Rev Off J Eur Respir Soc* (2021) 30:210045. doi: 10.1183/16000617.0045-2021

21. Dai D, Li Q, Zhou P, Huang J, Zhuang H, Wu H, et al. Analysis of omics data reveals nucleotide excision repair-related genes signature in highly-grade serous ovarian cancer to predict prognosis. *Front Cell Dev Biol* (2022) 10:874588. doi: 10.3389/fcell.2022.874588

22. Song J, Sun Y, Cao H, Liu Z, Xi L, Dong C, et al. A novel pyroptosis-related lncRNA signature for prognostic prediction in patients with lung adenocarcinoma. *Bioengineered* (2021) 12:5932–49. doi: 10.1080/21655979.2021.1972078

23. Buttarelli M, Ciucci A, Palluzzi F, Raspaglio G, Marchetti C, Perrone E, et al. Identification of a novel gene signature predicting response to first-line chemotherapy in BRCA wild-type high-grade serous ovarian cancer patients. *J Exp Clin Cancer Res CR* (2022) 41:50. doi: 10.1186/s13046-022-02265-w

24. Wang H, Wang Q, Cai G, Duan Z, Nugent Z, Huang J, et al. Nuclear TIGAR mediates an epigenetic and metabolic autoregulatory loop *via* NRF2 in cancer therapeutic resistance. *Acta Pharm Sin B* (2022) 12:1871–84. doi: 10.1016/j.apsb.2021.10.015

25. Zhang J, Zhang H, Ding X, Hu J, Li Y, Zhang J, et al. Crosstalk between macrophage-derived PGE2 and tumor UHRF1 drives hepatocellular carcinoma progression. *Theranostics* (2022) 12:3776–93. doi: 10.7150/thno.69494

26. Tirtakusuma R, Szoltysek K, Milne P, Grinev V, Ptasinska A, Chin PS, et al. Epigenetic regulator genes direct lineage switching in MLL/AF4 leukaemia. *Blood* (2022). doi: 10.1182/blood.2021015036

27. Zhang L, Zhang R, Wang J, Chen Y, Qiao C, Shi Q, et al. Identification of clinical implications and potential prognostic models of chromatin regulator mutations in multiple myeloma. *Clin Epigenet* (2022) 14:93. doi: 10.1186/s13148-022-01314-7

28. Wu L, Kou F, Ji Z, Li B, Zhang B, Guo Y, et al. SMYD2 promotes tumorigenesis and metastasis of lung adenocarcinoma through RPS7. *Cell Death Dis* (2021) 12:439. doi: 10.1038/s41419-021-03720-w

29. Li Z, Wu X, Li J, Yu S, Ke X, Yan T, et al. HMGA2-Snai2 axis regulates tumorigenicity and stemness of head and neck squamous cell carcinoma. *Exp Cell Res* (2022) 418:113271. doi: 10.1016/j.yexcr.2022.113271

30. Wang M, Ma M, Yang Y, Li C, Wang Y, Sun X, et al. Overexpression of hsa_circ_0008274 inhibited the progression of lung adenocarcinoma by regulating HMGA2 *via* sponging miR-578. *Thorac Cancer* (2021) 12:2258–64. doi: 10.1111/1759-7714.14059

31. Wang B, Pan L-Y, Kang N, Shen X-Y. PP4R1 interacts with HMGA2 to promote non-small-cell lung cancer migration and metastasis *via* activating MAPK/ERK-induced epithelial-mesenchymal transition. *Mol Carcinog* (2020) 59:467–77. doi: 10.1002/mc.23168

32. Zhou YD, Barnard M, Tian H, Li X, Ring HZ, Francke U, et al. Molecular characterization of two mammalian bHLH-PAS domain proteins selectively expressed in the central nervous system. *Proc Natl Acad Sci USA* (1997) 94:713–8. doi: 10.1073/pnas.94.2.713

33. Qiu M, Chen Y-B, Jin S, Fang X-F, He X-X, Xiong Z-F, et al. Research on circadian clock genes in non-small-cell lung carcinoma. *Chronobiol Int* (2019) 36:739–50. doi: 10.1080/07420528.2018.1509080

34. Gao L-W, Wang G-L. Comprehensive bioinformatics analysis identifies several potential diagnostic markers and potential roles of cyclin family members in lung adenocarcinoma. *OncoTargets Ther* (2018) 11:7407–15. doi: 10.2147/OTT.S171705

35. Angrand P-O. Structure and function of the polycomb repressive complexes PRC1 and PRC2. *Int J Mol Sci* (2022) 23:5971. doi: 10.3390/ijms23115971

36. Pallante P, Forzati F, Federico A, Arra C, Fusco A. Polycomb protein family member CBX7 plays a critical role in cancer progression. *Am J Cancer Res* (2015) 5:1594–601.

37. Li R, Yan Q, Tian P, Wang Y, Wang J, Tao N, et al. CBX7 inhibits cell growth and motility and induces apoptosis in cervical cancer cells. *Mol Ther Oncolytics* (2019) 15:108–16. doi: 10.1016/j.omto.2019.09.002

38. Pei Y-F, He Y, Hu L-Z, Zhou B, Xu H-Y, Liu X-Q. The crosstalk between lncRNA-SNHG7/miRNA-181/cbx7 modulates malignant character in lung adenocarcinoma. *Am J Pathol* (2020) 190:1343–54. doi: 10.1016/j.ajpath.2020.02.011

39. Siddiqui A, Ceppi P. A non-proliferative role of pyrimidine metabolism in cancer. *Mol Metab* (2020) 35:100962. doi: 10.1016/j.molmet.2020.02.005

40. Vaupel P, Schmidberger H, Mayer A. The warburg effect: Essential part of metabolic reprogramming and central contributor to cancer progression. *Int J Radiat Biol* (2019) 95:912–9. doi: 10.1080/09553002.2019.1589653

41. Alzahrani AS. PI3K/Akt/mTOR inhibitors in cancer: At the bench and bedside. *Semin Cancer Biol* (2019) 59:125–32. doi: 10.1016/j.semcancer.2019.07.009

42. Sanaei M-J, Razi S, Pourbagheri-Sigaroodi A, Bashash D. The PI3K/Akt/mTOR pathway in lung cancer; oncogenic alterations, therapeutic opportunities, challenges, and a glance at the application of nanoparticles. *Transl Oncol* (2022) 18:101364. doi: 10.1016/j.tranon.2022.101364

43. Azulay EE, Cooks T, Elkabets M. Potential oncogenic roles of mutant-p53-derived exosomes in the tumor-host interaction of head and neck cancers. *Cancer Immunol Immunother CII* (2020) 69:285–92. doi: 10.1007/s00262-019-02450-5

44. Vokes NI, Chambers E, Nguyen T, Coolidge A, Lydon CA, Le X, et al. Concurrent TP53 mutations facilitate resistance evolution in EGFR-mutant lung adenocarcinoma. *J Thorac Oncol Off Publ Int Assoc Study Lung Cancer* (2022) 17:779–92. doi: 10.1016/j.jtho.2022.02.011

45. Kaur J, Elms J, Munn AL, Good D, Wei MQ. Immunotherapy for non-small cell lung cancer (NSCLC), as a stand-alone and in combination therapy. *Crit Rev Oncol Hematol* (2021) 164:103417. doi: 10.1016/j.critrevonc.2021.103417

46. Zhou W-T, Jin W-L. B7-H3/CD276: An emerging cancer immunotherapy. *Front Immunol* (2021) 12:701006. doi: 10.3389/fimmu.2021.701006

47. Wang S, Zhang X, Ning H, Dong S, Wang G, Sun R. B7 homolog 3 induces lung metastasis of breast cancer through Raf/MEK/ERK axis. *Breast Cancer Res Treat* (2022) 193:405–16. doi: 10.1007/s10549-022-06520-8

48. Yu T-T, Zhang T, Lu X, Wang R-Z. B7-H3 promotes metastasis, proliferation, and epithelial-mesenchymal transition in lung adenocarcinoma. *OncoTargets Ther* (2018) 11:4693–700. doi: 10.2147/OTT.S169811

49. Oitabén A, Fonseca P, Villanueva MJ, García-Benito C, López-López A, Garrido-Fernández A, et al. Emerging blood-based biomarkers for predicting immunotherapy response in NSCLC. *Cancers* (2022) 14:2626. doi: 10.3390/cancers14112626

50. Iizuka A, Nonomura C, Ashizawa T, Kondou R, Ohshima K, Sugino T, et al. A T-cell-engaging B7-H4/CD3-bispecific fab-scFv antibody targets human breast cancer. *Clin Cancer Res Off J Am Assoc Cancer Res* (2019) 25:2925–34. doi: 10.1158/1078-0432.CCR-17-3123

51. Liu Z, Jin K, Zeng H, Shao F, Chang Y, Wang Y, et al. B7-H4 correlates with clinical outcome and immunotherapeutic benefit in muscle-invasive bladder cancer. *Eur J Cancer Oxf Engl 1990* (2022) 171:133–42. doi: 10.1016/j.ejca.2022.05.022

52. Schalper KA, Carvajal-Hausdorf D, McLaughlin J, Altan M, Velcheti V, Gaule P, et al. Differential expression and significance of PD-L1, IDO-1, and B7-H4 in human lung cancer. *Clin Cancer Res Off J Am Assoc Cancer Res* (2017) 23:370–8. doi: 10.1158/1078-0432.CCR-16-0150

53. Lu Y, Wu F, Cao Q, Sun Y, Huang M, Xiao J, et al. B7-H4 is increased in lung adenocarcinoma harboring EGFR-activating mutations and contributes to immunosuppression. *Oncogene* (2022) 41:704–17. doi: 10.1038/s41388-021-02124-6

54. Lou X, Ning J, Liu W, Li K, Qian B, Xu D, et al. YTHDF1 promotes cyclin B1 translation through m6A modulation and contributes to the poor prognosis of lung adenocarcinoma with KRAS/TP53 Co-mutation. *Cells* (2021) 10:1669. doi: 10.3390/cells10071669

55. Wang S, Gao S, Zeng Y, Zhu L, Mo Y, Wong CC, et al. N6-methyladenosine reader YTHDF1 promotes ARHGEF2 translation and RhoA signaling in colorectal cancer. *Gastroenterology* (2022) 162:1183–96. doi: 10.1053/j.gastro.2021.12.269

56. Zhao X, Chen Y, Mao Q, Jiang X, Jiang W, Chen J, et al. Overexpression of YTHDF1 is associated with poor prognosis in patients with hepatocellular carcinoma. *Cancer Biomark Sect Dis Markers* (2018) 21:859–68. doi: 10.3233/CBM-170791

57. Chen X-Y, Liang R, Yi Y-C, Fan H-N, Chen M, Zhang J, et al. The m6A reader YTHDF1 facilitates the tumorigenesis and metastasis of gastric cancer *via* USP14 translation in an m6A-dependent manner. *Front Cell Dev Biol* (2021) 9:647702. doi: 10.3389/fcell.2021.647702

58. Cheng H, Wang S-J, Li Z, Ma Y, Song Y-R. ING2-WTAP is a potential therapeutic target in non-small cell lung cancer. *Biochem Biophys Res Commun* (2022) 605:31–8. doi: 10.1016/j.bbrc.2022.02.041

59. Weng L, Qiu K, Gao W, Shi C, Shu F. LncRNA PCGEM1 accelerates non-small cell lung cancer progression *via* sponging miR-433-3p to upregulate WTAP. *BMC Pulm Med* (2020) 20:213. doi: 10.1186/s12890-020-01240-5

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 03 November 2022
DOI 10.3389/fonc.2022.1003222

OPEN ACCESS

EDITED BY
Fu Wang,
Xi'an Jiaotong University, China

REVIEWED BY
Lincan Duan,
Third Affiliated Hospital of Kunming Medical University, China
Qiqiao Du,
The First Affiliated Hospital of Sun Yat-sen University, China
Chen Li,
Free University of Berlin, Germany

*CORRESPONDENCE
Zengding Zhou
xueshengz@qq.com
Jiajing Cheng
chengjiajing1963@gmail.com
Jinlong Qin
qinjinlong121074@163.com

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 26 July 2022
ACCEPTED 29 September 2022
PUBLISHED 03 November 2022

CITATION
Zeng Q, Jiang H, Lu F, Fu M, Bi Y, Zhou Z, Cheng J and Qin J (2022) Prediction of the immunological and prognostic value of five signatures related to fatty acid metabolism in patients with cervical cancer. *Front. Oncol.* 12:1003222. doi: 10.3389/fonc.2022.1003222

Prediction of the immunological and prognostic value of five signatures related to fatty acid metabolism in patients with cervical cancer

Qiongjing Zeng[1†], Huici Jiang[1†], Fang Lu[1], Mingxu Fu[1], Yingying Bi[1], Zengding Zhou[2*], Jiajing Cheng[1*] and Jinlong Qin[1*]

[1]Department of Obstetrics and Gynecology, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, Shanghai, China, [2]Department of Burn Surgery, Ruijin Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

A growing attention has been attached to the role of fatty acid metabolism (FAM) in the development of cancer, and cervical cancer (CC) is still the primary cause of cancer-associated death in women worldwide. Therefore, it is imperative to explore the possible prognostic significance of FAM in CC. In this study, CC samples and corresponding normal samples were acquired from the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx). Single sample gene set enrichment analysis (ssGSEA) was conducted for calculating FAM-related scores (FAMRs) to screen FAM-related genes (FAMRGs). Two subtypes related to FAM were identified by consistent clustering. Among them, subtype C2 had a poor prognosis, and C1 had a high level of immune cell infiltration, while C2 had a high possibility of immune escape and was insensitive to chemotherapy drugs. Based on the differentially expressed genes (DEGs) in the two subtypes, a 5-gene signature (PLCB4, FBLN5, TSPAN8, CST6, and SERPINB7) was generated by the least absolute shrinkage and selection operator (LASSO) and Akaike information criterion (AIC). The model demonstrated a high prognostic accuracy (area under the curve (AUC)>0.7) in multiple cohorts and was one independent prognostic factor for CC patients. Accordingly, FAMRGs can be adopted as a biomarker for the prediction of CC patients' prognosis and help guide the immunotherapy of CC.

KEYWORDS

prediction, immunological and prognostic value, five signatures, acid metabolism, cervical cancer

Introduction

Since the World Health Organization (WHO) called for the worldwide elimination of cervical cancer (CC) in 2018, various preventive measures for CC have emerged one after another, among which human papillomavirus (HPV) vaccine and cervical screening are the two most effective interventions (1). However, such prevention and treatment schemes are extremely limited by resources and basic health facilities, and the coverage of them in low-and middle-income countries is less than one tenth of that in developed countries, so CC is still the primary cause of cancer-associated death in poor countries over the globe (2, 3). At present, there is still a need to develop a brand-new screening technology that can identify the symptoms in the incubation period and early stage of CC, and is affordable in most regions, thus reducing the difference in the incidence of CC worldwide due to the gap in resources and infrastructure by greatly lowering the incidence of CC in developing countries (4, 5). Therefore, a faster and more cost-effective screening method for CC is still wanted worldwide (6).

As the next-generation sequencing technology and the accumulation of CC sequencing data develop, it becomes clear and cost-effective to find biomarkers for prognosis assessment and treatment of CC through genome-wide analysis (7). Valuable decision-making guidance can be provided for clinicians by unbiased synthesis of various data, screening of molecular characteristics of cancer-causing subgroups of CC, and re-classification of them, so that more medical resources can be concentrated on high-risk CC patients who really have disease progression, and the economic and psychological burden caused by HPV vaccination and cervical screening can be greatly reduced (7–9).

Compared with normal cells, tumor cells often have different cell metabolic phenotypes to meet the energy needs of rapid cell proliferation and growth (10, 11). Recently, a growing number of studies have found that lipid metabolism disorder often occurs in the development of various human malignant tumors including prostate cancer (12) and colon cancer (13), and the change of FAM has greatly promoted the energy conversion of cancer cells (14). All the activities of tumor cells are inseparable from the intake and synthesis of fatty acids (15). The gradually accumulated fatty acids seem to be bound up with the disease recurrence and unfavorable prognosis of patients, and the metabolic characteristics of fatty acids may become a new target of anti-cancer therapy (12, 14, 15).

Zhang et al. (16) have found that fatty acid-binding protein 5 induces lymphatic metastasis of CC through metabolic reprogramming, but the clinical significance of FAM-related characteristics in CC is still under investigation, and it is still a challenge to identify stable fatty acid-related signature. This study comprehensively analyzed the expression, immune and prognostic characteristics of fatty acid metabolism-related genes (FAMRGs) in CC, identified two different CC subtypes associated with FAM and their immune characteristics, and verified the FAM-associated prognosis model by multiple cohorts, which provided a theoretical basis for forecasting the survival risk of CC patients.

Methods

Variation analysis acquisition and pre-processing of data sets

From the Cancer Genome Atlas (TCGA, https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga) (17) and Genotype-Tissue Expression (GTEx) (https://commonfund.nih.gov/gtex) (18), the data about expression profile of CC tissues and normal cervical tissues were downloaded. Their batch effect was eliminated by the remove batch effect function of limma in the R package, and two data sets were corrected by the normalize between arrays function. Principal Component Analysis (PCA) was used for evaluating the degree of batch effect removal. Totally 300 data of CC expression profiles were downloaded from the GSE44001 dataset of Gene Expression Omnibus (GEO) as a verification set (19), and files of the probe platform were downloaded. The probe ID numbers were annotated to gene symbols. Probes corresponding to multiple genes meantime were removed, and the value of probes with the same gene expression was averaged.

In addition, the 272 tumour samples in TCGA were assigned to a training set and a verification set in the random manner based on the proportion of 1:1 after 100 times of random grouping with replacement to facilitate the subsequent model construction.

Limma in the R package was adopted in the variation analysis of different groups, and the differentially expressed genes (DEGs) were screened with the absolute value of $\log_2$ (fold change) > $\log_2$ (1.2) and FDR< 0.05.

Single-sample gene-set enrichment analysis

The fatty acid metabolism-related scores (FAMRs) were calculated through ssGSEA and R package GSVA after downloading the FAMRGs sets in the molecular signature database (MSigDb, c2.cp.kegg. v7.4.symbols) (20). The rcorr function in Hmisc in the R package was adopted for determining the correlation of FAMRs with DEGs. The correlation with FAMRGs was found with cor > 0.2 and FDR<0.05 as the filter condition.

Survival analysis

Univariate Cox analysis was carried out by the coxph function of Survival in the R package to screen the genes associated with CC patients' prognosis, with $p<0.05$ as the filter condition. The log rank test was adopted for analyzing the survival differences between groups and corresponding Kaplan-Meier (K-M) curves were drawn.

Construction of FAM-related subtypes

272 CC samples were consistently clustered using ConsensusClusterPlus in the R package, and 500 times of bootstraps were performed by the pam algorithm and "Pearson" as the measurement distance. Each bootstrap process covered 80% of patients in the training set. With the number of clusters set to 2 to 10, the consistency matrix and consistency cumulative distribution function were calculated for determining the optimal classification.

Analysis of immune escape characteristics

According to the previous research (21, 22), the molecular characterization of aneuploid score, nosilent mutation rate, fraction altered, number of segments, and homologous recombination defects were collected to evaluate tumour immunogenicity among different subtypes, and maftools in the R package was used for visually analyzing the mutation data of the top 10 genes with significant differences in expression.

Calculation of the difference in immune microenvironment among different subtypes

The CIBERSORT algorithm in IOBR of the R package was adopted for calculating the relative abundance of 22 kinds of immune cells in CC (23), and the ESTIMATE algorithm was adopted for calculating the matrix score and immune score of each sample of CC (24).

Prediction of clinical efficacy

With the Tumour Immune Dysfunction and Exclusion (TIDE) algorithm developed by Jiang et al. (25), TIDE, IFNG, Dysfunction, Exclusion, and TAM.M2 scores were downloaded from TIDE (http://tide.dfci.harvard.edu) for predicting the clinical treatment response of different subtypes, and the Wilcox.test was used for comparing the scores among different subtypes. Additionally, the half-maximum inhibitory concentration (IC50) of traditional drugs was downloaded from Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/) (26), and pRRophetic in the R packet was used for predicting the chemotherapy response of CC samples.

Construction of prognosis-related signature based on FAMRGs

The glmnet in the R package was used for further feature selection by the least absolute shrinkage and selection operator (LASSO), and a risk model was built by 10-fold cross-validation. According to Akaike information criterion (AIC), the complexity of the model was evaluated, and the number of parameters was gradually deleted to acquire the optimal model. The RiskScore of patients with different subtypes was calculated ($\text{RiskScore} = \sum_{i=1}^{n} Coef(i) * Exp(i)$), and Coef was taken as the characteristic coefficient of each signature. Exp presented the expression of each signature in CC samples. The samples of RiskScore with Z score and RiskScore > 0 were assigned to a high-risk group and those with scores<0 to a low-risk group, and the timerROC in the R package was used to evaluate the prediction accuracy of different risk levels. The rms in the R package was adopted for establishing nomograms to predict the1-year, 3-year and 5-year overall survival rates and calculate the prognosis risk of individual patients. The Decision Curve Analysis (DCA) curve was drawn by ggDCA in the R-packet for evaluating the clinical predictive performance of the model.

Clinical sample collection and qPCR validation

100 cases of cervical cancer tissues and 100 cases of adjacent tissues in our hospital were collected, and qPCR verification of PLCB4, FBLN5, TSPAN8, CST6, and SERPINB7 genes was performed. The tissue samples were fully ground with liquid nitrogen, 1 ml of Trizol (Invitrogen) solution was added, mixed well, and placed at room temperature for 5 minutes to fully lyse; (the sample name should be marked on the tube cover and tube wall) qPCR verification was carried out according to the specific operation steps of qPCR. Primers: PLCB4, F:ACAG ATACGAGGAGGAATCC, R: TCCATGTCAGAAAGAAGCC; FBLN5, F: CATCAATACTGAAGGCGGG, R: TCATCAAT GTCTAAGCACTGG; TSPAN8, F: CAAGAAGAGTTTAA ATGCTGCG, R: AGGCACATAATTCAGGATAGTG; CST6, F: TACTACTTCCGAGACACGC, R: AGGAAGTACTTG ATGCCGG; SERPINB7, F: TCCCACAAGGATTATGATC TCAG, R: CTCAATGTAGTCCTTATGAAAGCC. The relative expression levels of PLCB4, FBLN5, TSPAN8, CST6 and SERPINB7 genes were calculated by $2^{-\triangle\triangle CT}$.

Results

Screening of FAMRGs

The working route of this study is shown in Figure 1. According to PCA, two data sets were clustered together mainly according to their sources (Figure 2A), but after the integration of these data sets, the samples in the two data sets were mixed, and the batch effect between the data sets was eliminated (Figure 2B).

As shown in Figure 3A, 487 DEGs were selected from tumor samples of CC and corresponding normal samples, of which 120 DEGs were up-regulated and 367 DEGs were down-regulated. Furthermore, ssGSEA results revealed notably fewer FAMRGs in tumor samples than those in normal samples (Figure 3B) and also revealed differences in FAM between CC tissues and normal tissues. Among them, 48 DEGs were greatly associated with FAM (Figure 3C). Univariate Cox analysis showed that 7 FAMRGs including S100A11 were bound up with the prognosis of CC patients (Figure 3D; Supplementary Table 1).

Identification of two different FAM-related subtypes based on FAMRG

Based on the cumulative distribution function (CDF) and CDF Delta area curve, the optimal number of clusters (Figures 4A, B) was determined. When k=2, there was a comparatively stable clustering result, and two subtypes (C1, and C2) were obtained (Figure 4C). Further analysis of the prognosis of these two CC subtypes revealed a notably lower survival rate in patients from the C2 group that that in patients from the C1 group at the same time ($p<0.05$, Figure 4D). Similarly, the same difference in GSE44001 was found. The same method was adopted for processing the CC samples from GSE44001. Patients in Group C2 still had an unfavorable prognosis (Figure 4E), which was similar to the results of the data set from TCGA. The findings indicate that the two subtypes based on FAMRG can be transplanted in different research cohorts. The Chi-square test was used for comparing the distribution of different clinicopathological features between the two subtypes, and the results revealed notable differences in the living conditions of CC patients in the TCGA cohort between the two groups (Figure 5, $p<0.05$).

Immune characteristics of FAM-related subtypes

The results revealed a notably higher fraction altered in C1 than that in C2 (Figure 6A), and also showed that genes with significant differences in CC such as TTN and PIK3CA had higher mutation frequency in C1 (Figure 6B). The potential function of FAM-related subtypes in CC was further analyzed, and the proportion of 22 immune cell types between the two subtypes was evaluated by CIBERSORT. Compared with C2, the proportion of B cells navie, Plasma cells, T cells memory resting and T cells regulatory (Tregs) in C1 was significantly lower, while T cells memory activated, Macrophages M1 and Dendritic cells activated were significantly enriched (Figure 7A). C1 got higher immune score and estimate score than C2, and C1 had a higher level of immune cell infiltration (Figure 7B). As shown in Figure 7D, the TIDE score of subtype C2 in the TCGA cohort was higher than that of C1, suggesting that subtype C2 was more likely to escape and less likely to benefit from immunotherapy. The IC50 of 6 traditional chemotherapeutic drugs in C1 was significantly lower than that in C2, and these drugs were more effective in C1 patients (Figure 7E).

Construction and verification of prognosis-related model of FAMRGs

Totally 558 DEGs of the two-fatty acid-related subtypes were screened by variation analysis (Figure 8A), and 58 DEGs related to prognosis were further filtered by univariate Cox analysis in the training set (Supplementary Table 2). When lambda= 0.0385, the model reached the optimal state (Figures 8B, C), and the parameters were further compressed to obtain a model composed of five genes: Riskscores = 0.48 × PLCB4 + 0.49 × FBLN5 + 0.15 × TSPAN8 + 0.38 × CST6 + 0.30 × SERPINB7 (See Supplementary Table 3 for detailed descriptions of the genes.)

The Risk Score of each sample was calculated. As shown in Figures 8D and E, the training set and validation set of TCGA both revealed a shorter survival time in CC patients from the high-risk group than that from the low-risk one ($p<0.05$). Moreover, this model had high accuracy in the prediction and classification of CC in one year, three years and five years (area under the curve (AUC)>0.7). For further verifying the generalization ability of the model, all TCGA data and the independent data set GSE44001 were verified. The results, as shown in Figures 8F and G, were in agreement with those of the training set of TCGA. FAMRGs prognosis-associated model was a prognosis scoring system with high precision (AUC>0.7), and the high-risk group had an unfavorable prognosis.

Association of RiskScore with other clinicopathological features and its prognostic value

The associations of RiskScore with subtype, T. Stage, N. Stage, M. Stage, Stage, age, Event and Grade were tested. As

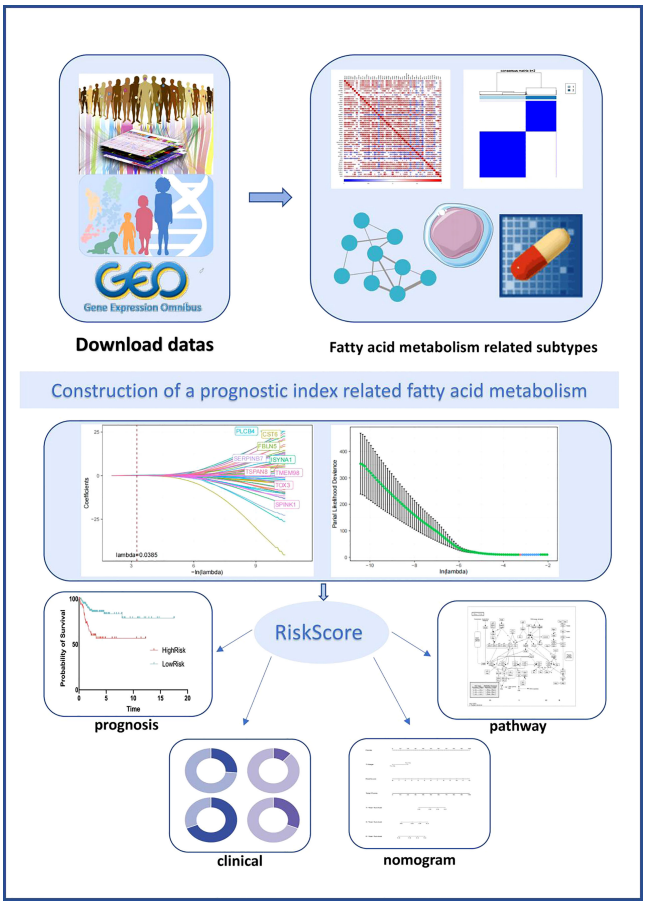


FIGURE 1
Graphical abstract of the construction of a prognostic index associated with fatty acid metabolism in cervical cancer.

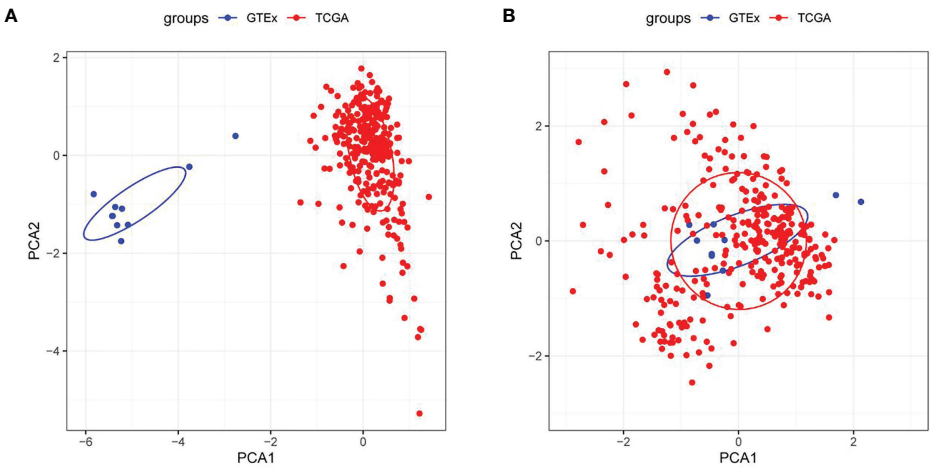


FIGURE 2
Evaluation of sample clustering by the principal component analysis (PCA) **(A)** PCA diagram between two data sets before the batch effect was removed; **(B)** PCA diagram between two data sets after the batch effect was removed. TCGA, The Cancer Genome Atlas; GTEx, Genotype-Tissue Expression.

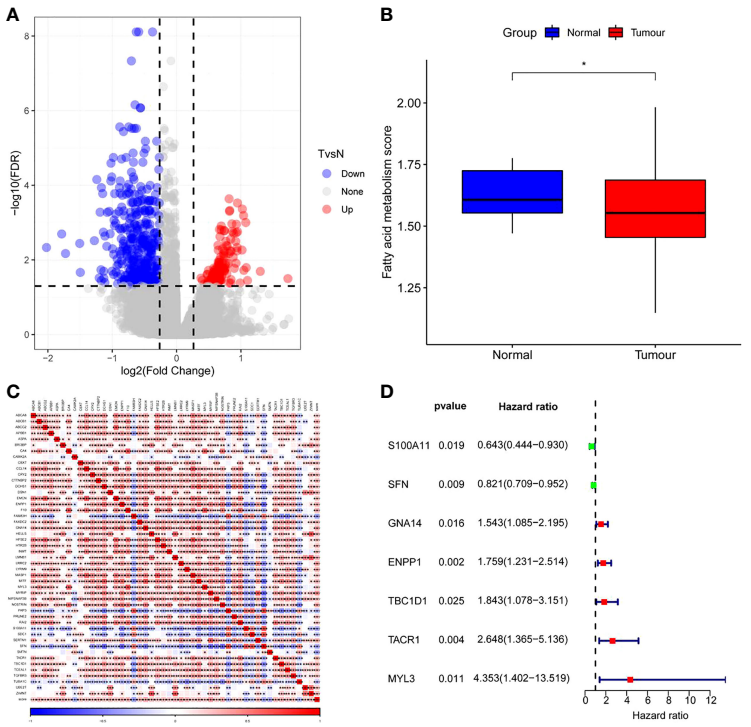


FIGURE 3
Screening of fatty acid metabolism-related genes (FAMRGs) **(A)**Volcano map of variation analysis between CC samples and normal samples; **(B)** Comparison of fatty acid metabolism-related scores (FAMRs) between CC samples and corresponding normal samples; **(C)** FAMRGs-related Heat map, **(D)** Forest map of prognosis-related FAMRGs. * vs p<0.05. *p<0.05,**p<0.01,***p<0.001.

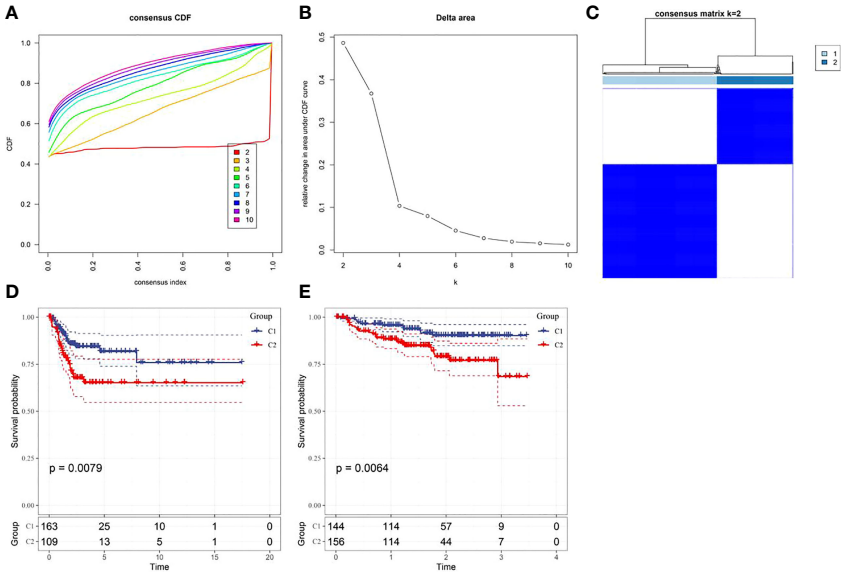


FIGURE 4
Construction of fatty acid metabolism-related subtypes and prognosis **(A)** Consensus clustering samples between each category number k in the TCGA cohort; **(B)** Delta area curve of cumulative distribution function (CDF) of TCGA cohort sample; **(C)** Heat map of sample clustering when k=2; **(D)** Kaplan-Meier curve of two subtypes in the TCGA cohort; **(E)** The prognostic Kaplan-Meier curves of the two subtypes in the GSE44001 cohort.

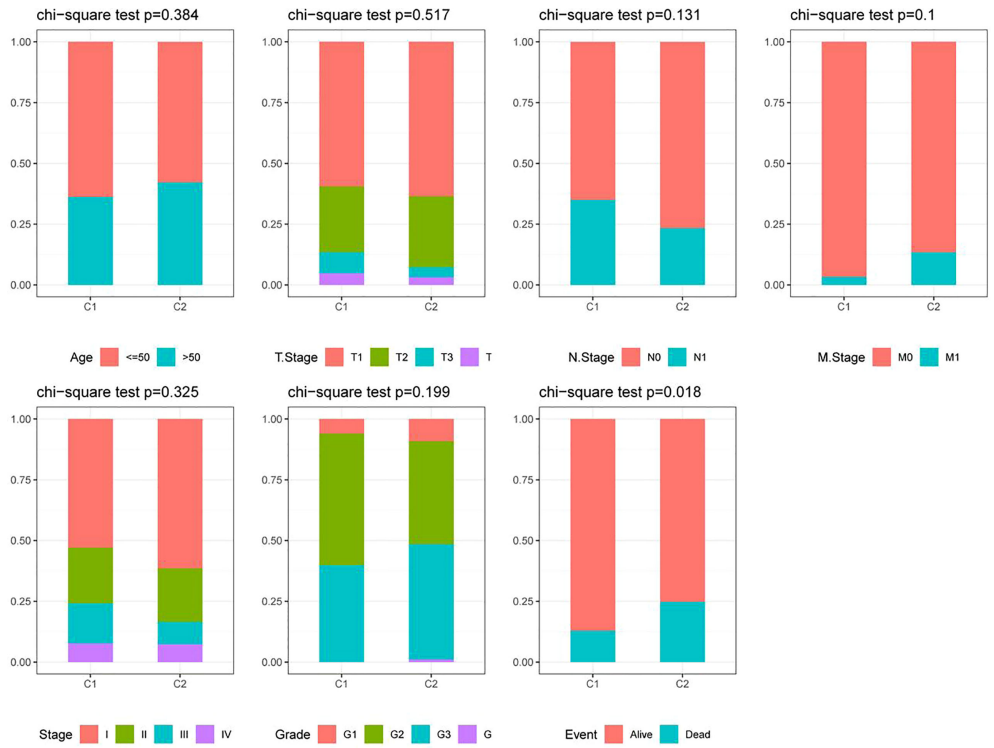


FIGURE 5
Distribution of clinicopathological features between two fatty acid metabolism-related subtypes in the TCGA cohort.

shown in Figure 9A, the proportion of subtype C2 and dead population in the high-risk group was higher ($p<0.05$). Further comparison of the difference in RiskScore among people with different clinicopathological features revealed higher RiskScore in people with subtype C2, age ≤ 50 and death (Figure 9B).

Univariate and multivariate Cox regression analysis was used for evaluating the prognostic value of RiskScore and other clinicopathological characteristics in CC. As shown in Figures 10A, B, T. Stage and RiskScore were independent prognostic factors of CC patients, and RiskScore was the most significant prognostic factor. Then, a nomogram composed of T. Stage and RiskScore was constructed. According to Figure 10C, RiskScore made the greatest contribution to the survival prediction of CC patients. The nomogram correction map and DCA curve showed that RiskScore had higher predictive performance than other clinicopathological features.

Biological pathway of potential regulation of FAMRGs prognosis-related model

For better studying the potential function of the FAMRGs prognosis-related model, the score of each KEGG pathway in CC patients was calculated by GSVA package, and 90 significant pathways were calculated in the high-and low-risk groups ($p< 0.05$, Supplement Table 3), as shown in Figure 11A. Among them, there were 53 significant pathways in high and low risk groups ($p< 0.001$). The association of enrichment score with RiskScore was analyzed (Figure 11B; Supplementary Table 4). The FAMRGs prognosis-related model was significantly bound up with signals including O GLYCAN BIOSYNTHESIS, CELL CYCLE, BASAL TRANSCRIPTION FACTORS, and P53_SIGNALING_PATHWAY, which was similar to our previous research results (Figure 7C). There were significant differences in FAM-related subtypes among 10 classic oncogenic pathways (27), and FAMRGs prognosis-related model was strongly bound up with these signals.

Clinical cohort qPCR validation

100 cervical cancer tissues and 100 paracancerous tissues were collected from our hospital for qPCR verification of PLCB4, FBLN5, TSPAN8, CST6, and SERPINB7. The results showed that PLCB4, FBLN5, TSPAN8, CST6, and SERPINB7 were highly expressed in cervical cancer tissues (Figure 12, $p<0.05$).

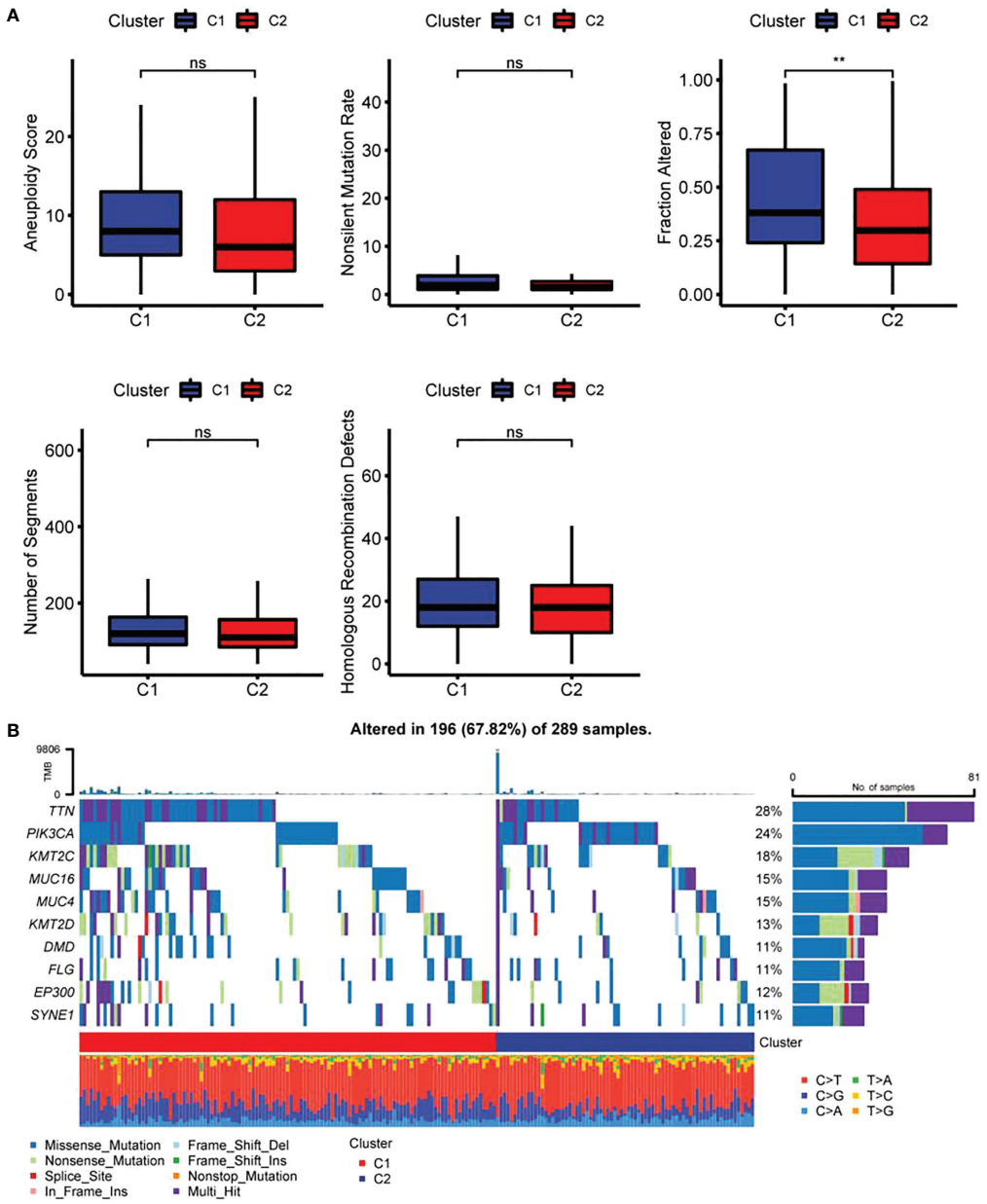


FIGURE 6
Differences in genome mutation between two fatty acid-related subtypes **(A)** Differences of Homologous Recombination Defects, Aneuploidy Score, Fraction Altered, Number of Segments and Tumor mutation burden in molecular subtypes in the TCGA cohort; **(B)** Somatic mutation landscape in two molecular subtypes. ** p<0.01, ns, P>0.05.

Discussion

Compared with sugar metabolism and amino acid metabolism, FAM has received less attention, but the importance of fatty acids in the development of cancer is increasingly recognized (28, 29). As a crucial component of the membrane matrix, the fatty acid is a crucial messenger and fuel source for energy production (30). Compared with normal cells, tumour cells are more likely to rely on *de novo* synthesis to synthesize fatty acids for energy metabolism and membrane formation for the maintenance of the rapid growth and proliferation of cells (31). It is worth noting that there are many fatty acids and metabolic by-products of them, each of which has different feedback mechanisms and regulation nodes and affects

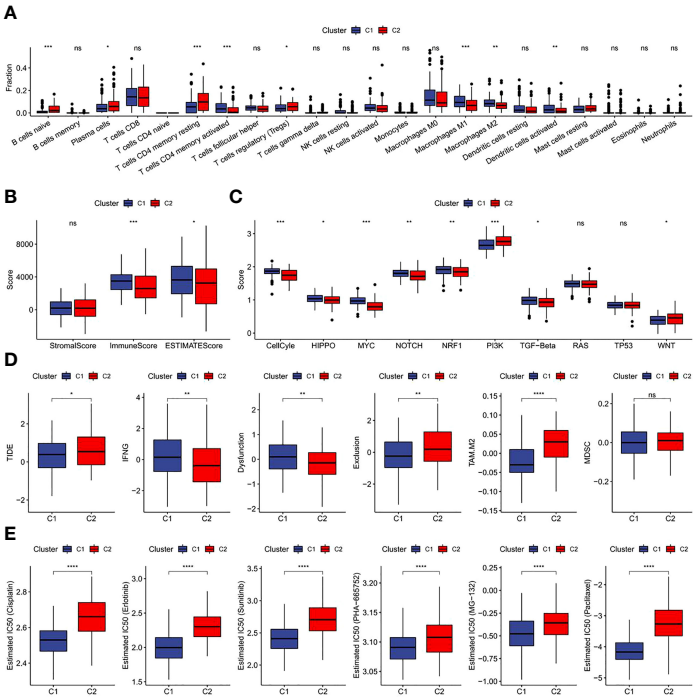


FIGURE 7
Immunological characteristics, immunotherapy, chemotherapy and target therapy responses of two fatty acid-related subtypes **(A)** Difference of 22 immune cell scores between different molecular subtypes in the TCGA cohort; **(B)** Difference of ESTIMATE immune infiltration between different molecular subtypes in the TCGA cohort; **(C)** Difference in the score of 10 pathways related to tumor abnormality between different subgroups in the TCGA cohort; **(D)** Difference of TIDE analysis results between different groups in the TCGA cohort; **(E)** The box plots of the estimated IC50 for drug in TCGA-CECS. *p<0.05,**p<0.01,***p<0.001, ****p<0.0001, ns, P>0.05.

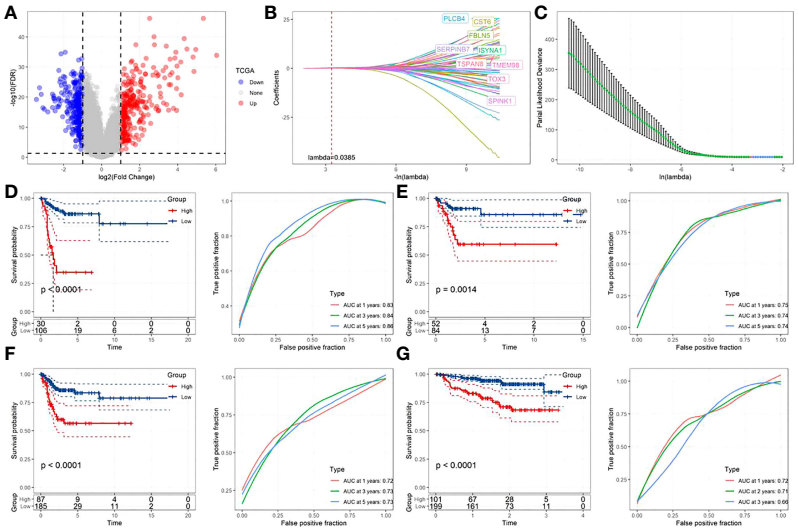


FIGURE 8
Construction and verification of prognostic correlation model of FamRGS **(A)**Volcanic map of the variation analysis of molecular subtypes; **(B)** The changing trajectory of each independent variable, with the horizontal axis representing the log value of the independent variable lambda, and the vertical axis representing the coefficient of the independent variable; **(C)** Confidence interval under each lambda; **(D)** AUC curve and KM curve of the risk model of the training set data from TCGA; **(E)** AUC curve and KM curve of the risk model of the verification set data from TCGA; **(F)** AUC curve and KM curve of the risk model of all data sets from TCGA; **(G)** AUC curve and KM curve of the risk model of data set from GSE44001.

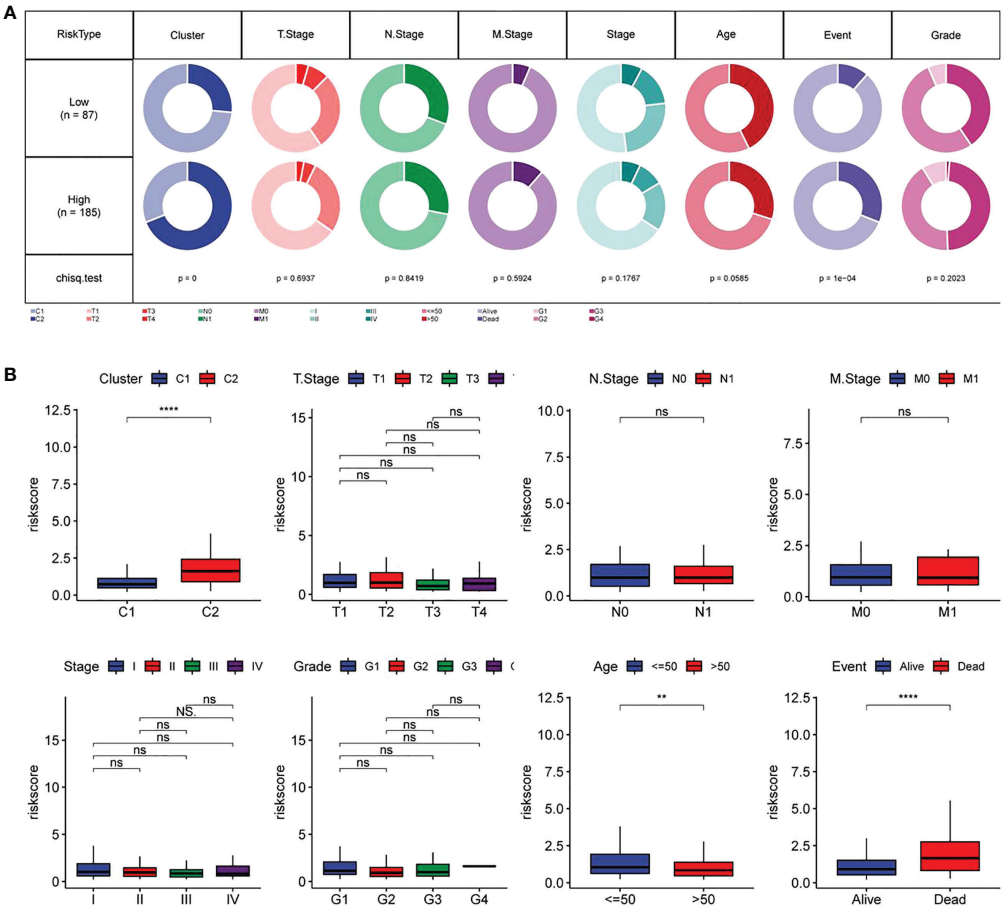


FIGURE 9
Relationship between RiskScore and different clinicopathological features of CC patients **(A)** Comparison in terms of the distribution of different clinical characteristics between high-and low-risk groups in TCGA data sets; **(B)** Differences in RiskScore of different clinical characteristics in TCGA data sets. **p<0.01, ****p<0.001, ns, P>0.05.

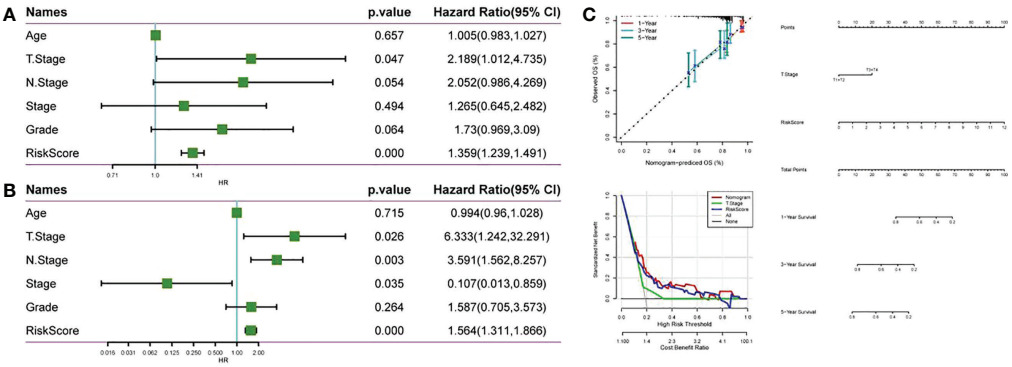


FIGURE 10
Compared with other clinicopathological features, RiskScore has higher prognostic value. **(A)** Forest map of clinical phenotype in the TCGA cohort and univariate Cox analysis of RiskScore; **(B)** Forest map of clinical phenotype in the TCGA cohort and multivariate Cox analysis of RiskScore; **(C)** Anomogram, nomogram alignment map and DCA curve constructed by TCGA data set.

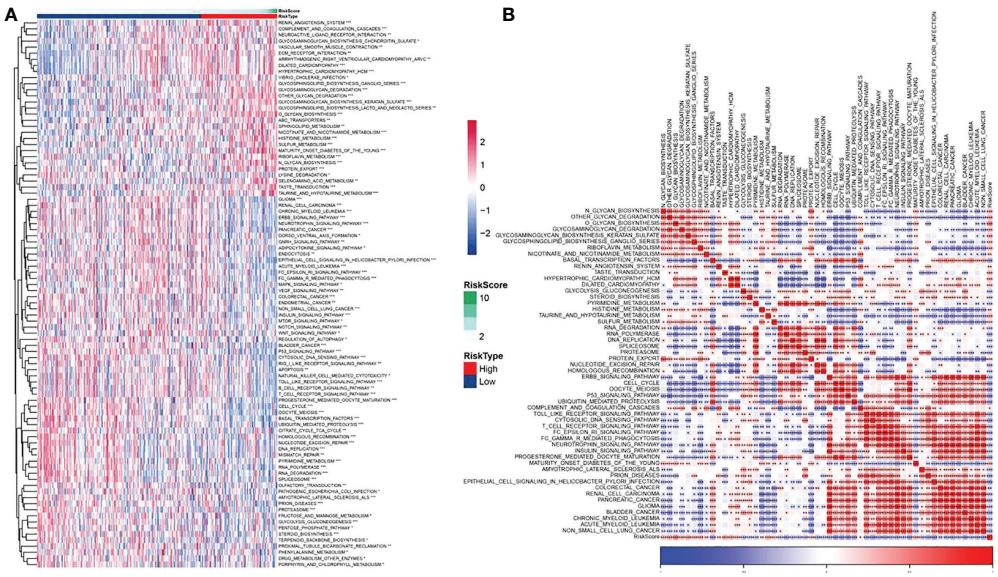


FIGURE 11
Biological pathway for potential regulation of FAMRGs prognosis-related model **(A)** Heat map in which GSVA analysis of TCGA dataset pathway showed significant enrichment scores of related pathways in high and low risk groups, and **(B)** Heat map of correlation analysis of pathways with significant differences in TCGA data sets and RiskScores (*p<0.05,**p<0.01,***p<0.001).

many ways of disease behavior through extremely complex regulatory networks (32). Different tumor subtypes may drive specific lipid phenotypes (28, 31). Therefore, identifying a potential subtype related to FAM helps predict the CC patients' prognosis.

By integrating cervical sample data from TCGA and GTEx, DEGs related to FAMRs were searched and molecular subtypes related to FAM (C1 and C2) were constructed. C2 showed poor prognosis in both TCGA cohort and GSE44001 cohort, independent verification set, and the proportion of deaths in subtype C2 was notably higher than that in subtype C1. In addition, a high Fraction of Altered was found in C1, and TTN and PIK3CA, common drivers in CC, have a high mutation frequency (14). C1 shows a higher proportion of T cells, macrophages and dendritic cells activated, and immune score and estimate score than C2, which means a lower level of tumor purity (24). The proportion of immune cells is high in the samples with lower tumor purity (33), and the inflammatory reaction caused by immune cells will increase the cell mutation rate and activate stronger anti-tumor characteristics and faster reaction speed (34, 35). The samples with higher tumor mutation load often show better immunotherapy effects (34, 36). These results were in agreement with our research results. The IC50 of traditional chemotherapeutic drugs in subtype C1 was notably lower than that in subtype C2, and subtype C1 was less likely to escape from immune surveillance than C2 and was more sensitive to immunotherapy.

Then, a prognostic model related to FAM was constructed for CC based on the two subtypes of DEGs, and the generalization and prediction accuracy of the model was repeatedly verified by multiple cohorts. Patients with subtype C2 and dead ones accounted for a higher proportion in the high-risk group, and these patients had a higher RiskScore. Consistent with our expectations, RiskScore can serve as one independent prognostic factor to predict the CC patients' prognosis and contributes greatly to the prediction of the survival of CC. The accuracy of the model prediction has been further confirmed.

As we described above, the up regulation of FAM contributes to cell membrane production and signal transmission, including activation signals (4, 37). The enrichment of multiple signals is significantly different in different risk levels. Research has pointed out that one of the key mechanisms of signal transduction in CC cells is the glycosylation of proteins (38). As a glycoprotein on the cell surface, N-Glycon directly affects cell signal transduction and is the diagnostic target of malignant transformation in the early stage of CC (39–41). O-glycan can be used as a biological marker of proliferation, senescence and metastasis of CC cells by

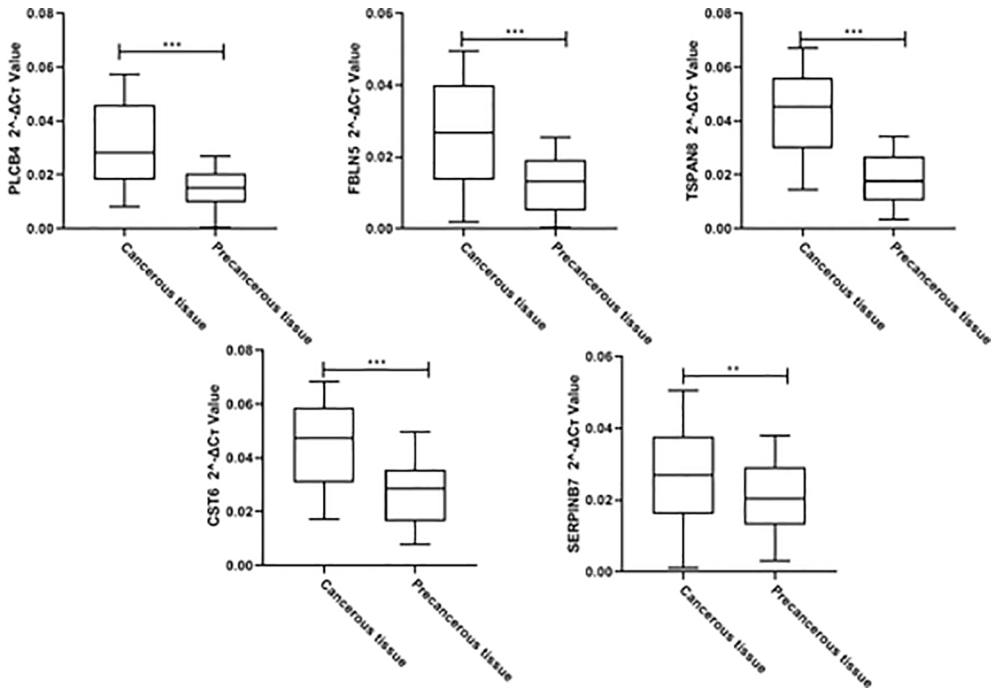


FIGURE 12
Clinical cohort qPCR validationClinical cohort qPCR validation for PLCB4, FBLN5, TSPAN8, CST6, and SERPINB7 (**p<0.01,*** p<0.001).

regulating immune response and controlling cell metabolism (42, 43). In addition, in the G2 phase, *de novo* synthesis is enhanced to synthesize lipids, which ensures the membrane material needed for mitosis and promotes cell proliferation (44). Subtype C1 had a higher score on cell cycle than subtype C2, and RiskScore and cell cycle enrichment showed a significantly negative correlation. The high-risk group and patients with subtype C2 may escape from the control of the cell cycle and fails, leading to continuous cell division and promoting cancer progress (45, 46).

Although some studies have explored biomarkers related to FAM in clear cell renal cell carcinoma (47) and bladder cancer (48), this study has revealed molecular subtypes related to FAM in CC for the first time and constructed a FAM-related prognostic model with strong predictive ability. It provides some new insights for accurate screening of CC, which is helpful to guide clinical treatment and prognosis prediction.

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 Shanghai Fourth People's Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

QZ, HJ, and FL are the mainly responsible for the writing of the article. YB and MF are mainly responsible for research design. ZZ, JC, and JQ are co-corresponding authors. They are responsible for ensuring that the descriptions are accurate and agreed by all authors. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Shanghai Science and Technology Commission/General Project of Natural Science Foundation of Shanghai in 2021(21ZR1440600).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1003222/full#supplementary-material

References

1. Canfell K. Towards the global elimination of cervical cancer. *Papillomavirus Res* (2019) 8:100170. doi: 10.1016/j.pvr.2019.100170

2. Bruni L, Diaz M, Barrionuevo-Rosas L, Herrero R, Bray F, Bosch FX, et al. Global estimates of human papillomavirus vaccination coverage by region and income level: a pooled analysis. *Lancet Global Health* (2016) 4(7):e453–e63. doi: 10.1016/S2214-109X(16)30099-7

3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492

4. Bedell SL, Goldstein LS, Goldstein AR, Goldstein AT. Cervical cancer screening: past, present, and future. *Sexual Med Rev* (2020) 8(1):28–37. doi: 10.1016/j.sxmr.2019.09.005

5. Basu P, Mittal S, Vale DB, Kharaji YC. Secondary prevention of cervical cancer. *Best Pract Res Clin Obstetrics Gynaecology.* (2018) 47:73–85. doi: 10.1016/j.bpobgyn.2017.08.012

6. Vu M, Yu J, Awolude OA, Chuang L. Cervical cancer worldwide. *Curr problems cancer.* (2018) 42(5):457–65. doi: 10.1016/j.currproblcancer.2018.06.003

7. Gao G, Johnson SH, Kasperbauer JL, Eckloff BW, Tombers NM, Vasmatzis G, et al. Mate pair sequencing of oropharyngeal squamous cell carcinomas reveals that HPV integration occurs much less frequently than in cervical cancer. *J Clin Virology.* (2014) 59(3):195–200. doi: 10.1016/j.jcv.2013.12.006

8. Hu Z, Ma D. The precision prevention and therapy of HPV-related cervical cancer: new concepts and clinical implications. *Cancer Med* (2018) 7(10):5217–36. doi: 10.1002/cam4.1501

9. Tuna M, Amos CI. Next generation sequencing and its applications in HPVassociated cancers. *Oncotarget* (2017) 8(5):8877. doi: 10.18632/oncotarget.12830

10. DeBerardinis RJ, Lum JJ, Hatzivassiliou G, Thompson CB. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab* (2008) 7(1):11–20. doi: 10.1016/j.cmet.2007.10.002

11. Vaupel P, Schmidberger H, Mayer A. The warburg effect: essential part of metabolic reprogramming and central contributor to cancer progression. *Int J Radiat Biol* (2019) 95(7):912–9. doi: 10.1080/09553002.2019.1589653

12. Dang Q, Chen Y-A, Hsieh J-T. The dysfunctional lipids in prostate cancer. *Am J Clin Exp Urology.* (2019) 7(4):273. doi: 10.3390/cancers13092000

13. Piccinin E, Cariello M, Moschetta A. Lipid metabolism in colon cancer: role of liver X receptor (LXR) and stearoyl-CoA desaturase 1 (SCD1). *Mol Aspects Med* (2021) 78:100933. doi: 10.1016/j.mam.2020.100933

14. Balaban S, Lee LS, Varney B, Aishah A, Gao Q, Shearer RF, et al. Heterogeneity of fatty acid metabolism in breast cancer cells underlies differential sensitivity to palmitate-induced apoptosis. *Mol Oncol* (2018) 12(9):1623–38. doi: 10.1002/1878-0261.12368

15. Chen M, Huang J. The expanded role of fatty acid metabolism in cancer: new aspects and targets. *Precis Clin Med* (2019) 2(3):183–91. doi: 10.1093/pcmedi/pbz017

16. Zhang C, Liao Y, Liu P, Du Q, Liang Y, Ooi S, et al. FABP5 promotes lymph node metastasis in cervical cancer by reprogramming fatty acid metabolism. *Theranostics* (2020) 10(15):6561. doi: 10.7150/thno.44868

17. Hutter C, Zenklusen JC. The cancer genome atlas: creating lasting value beyond its data. *Cell* (2018) 173(2):283–5. doi: 10.1016/j.cell.2018.03.042

18. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The genotype-tissue expression (GTEx) project. *Nat Genet* (2013) 45(6):580–5. doi: 10.1038/ng.2653

19. Clough E, Barrett T. The gene expression omnibus database. *Stat genomics: Springer;* (2016) p:93–110. doi: 10.1007/978-1-4939-3578-9_5

20. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* (2011) 27(12):1739–40. doi: 10.1093/bioinformatics/btr260

21. Li H, Li J, Zhang C, Zhang C, Wang H. TERT mutations correlate with higher TMB value and unique tumor microenvironment and may be a potential biomarker for anti-CTLA4 treatment. *Cancer Med* (2020) 9(19):7151–60. doi: 10.1002/cam4.3376

22. De Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* (2014) 346(6206):251–6. doi: 10.1126/science.1253462

23. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337

24. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* (2013) 4(1):1–11. doi: 10.1038/ncomms3612

25. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* (2018) 24(10):1550–8. doi: 10.1038/s41591-018-0136-1

26. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* (2012) 41(D1):D955–D61. doi: 10.1093/nar/gks1111

27. Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in the cancer genome atlas. *Cell* (2018) 173(2):321–37.e10. doi: 10.1016/j.cell.2018.03.035

28. Koundouros N, Poulogiannis G. Reprogramming of fatty acid metabolism in cancer. *Br J cancer.* (2020) 122(1):4–22. doi: 10.1038/s41416-019-0650-z

29. Currie E, Schulze A, Zechner R, Walther TC, Farese RVJr. Cellular fatty acid metabolism and cancer. *Cell Metab* (2013) 18(2):153–61. doi: 10.1016/j.cmet.2013.05.017

30. Kimura I, Ichimura A, Ohue-Kitano R, Igarashi M. Free fatty acid receptors in health and disease. *Physiol Rev* (2020) 100(1):171–210. doi: 10.1152/physrev.00041.2018

31. Li Z, Zhang H. Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cell Mol Life Sci* (2016) 73(2):377–92. doi: 10.1007/s00018-015-2070-4

32. Nagarajan SR, Butler LM, Hoy AJ. The diversity and breadth of cancer cell fatty acid metabolism. *Cancer Metab* (2021) 9(1):1–28. doi: 10.1186/s40170-020-00237-2

33. Mao Y, Feng Q, Zheng P, Yang L, Liu T, Xu Y, et al. Low tumor purity is associated with poor prognosis, heavy mutation burden, and intense immune phenotype in colon cancer. *Cancer Manage Res* (2018) 10:3569. doi: 10.2147/CMAR.S171855

34. Liu R, Yang F, Yin J-Y, Liu Y-Z, Zhang W, Zhou H-H. Influence of tumor immune infiltration on immune checkpoint inhibitor therapeutic efficacy: A

computational retrospective study. *Front Immunol* (2021) 12:2397. doi: 10.3389/fimmu.2021.685370

35. Robbins PF, Lu Y-C, El-Gamil M, Li YF, Gross C, Gartner J, et al. Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nat Med* (2013) 19(6):747. doi: 10.1038/nm.3161

36. Ma W, Li W, Xu L, Liu L, Xia Y, Yang L, et al. Identification of a gene prognostic model of gastric cancer based on analysis of tumor mutation burden. *Pathol Oncol Res* (2021) 27:1609852. doi: 10.3389/pore.2021.1609852

37. Mullen PJ, Yu R, Longo J, Archer MC, Penn LZ. The interplay between cell signalling and the mevalonate pathway in cancer. *Nat Rev Cancer* (2016) 16 (11):718–31. doi: 10.1038/nrc.2016.76

38. Xu Z, Zhang Y, Ocansey DK, Wang B, Mao F. Glycosylation in cervical cancer: new insights and clinical implications. *Front Oncol* (2021) 11. doi: 10.3389/fonc.2021.706862

39. Jin Y, Kim SC, Kim HJ, Ju W, Kim YH, Kim H-J. Increased sialylation and reduced fucosylation of exfoliated cervical cells are potential markers of carcinogenesis in the cervix. *Clin Chem Lab Med (CCLM).* (2016) 54(11):1811–9. doi: 10.1515/cclm-2015-1014

40. Hua S, Saunders M, Dimapasoc LM, Jeong SH, Kim BJ, Kim S, et al. Differentiation of cancer cell origin and molecular subtype by plasma membrane n-glycan profiling. *J Proteome Res* (2014) 13(2):961–8. doi: 10.1021/pr400987f

41. Martinez-Morales P, Cruz IM, Roa-de la Cruz L, Maycotte P, Salinas JSR, Zamora VJV, et al. Hallmarks of glycogene expression and glycosylation pathways in squamous and adenocarcinoma cervical cancer. *PeerJ* (2021) 9:e12081. doi: 10.7717/peerj.12081

42. Solórzano C, Mayoral MÁ, de los Angeles Carlos M, Berumen J, Guevara J, Chávez FR, et al. Overexpression of glycosylated proteins in cervical cancer recognized by the machaerocereus eruca agglutinin. *Folia Histochemica Cytobiologica.* (2012) 50(3):398–406. doi: 10.5603/FHC.2012.0054

43. Kim M, Kim YS, Kim H, Kang MY, Park J, Lee DH, et al. O-Linked n-acetylglucosamine transferase promotes cervical cancer tumorigenesis through human papillomaviruses E6 and E7 oncogenes. *Oncotarget* (2016) 7(28):44596. doi: 10.18632/oncotarget.10112

44. Icard P, Fournel L, Wu Z, Alifano M, Lincet H. Interconnection between metabolism and cell cycle in cancer. *Trends Biochem Sci* (2019) 44(6):490–50. doi: 10.1016/j.tibs.2018.12.007

45. Matthews HK, Bertoli C, de Bruin RA. Cell cycle control in cancer. *Nat Rev Mol Cell Biol* (2022) 23(1):74–88. doi: 10.1038/s41580-021-00404-3

46. Lundberg A, Yi JJJ, Lindström LS, Tobin NP. Reclassifying cancer: Defining tumour cell cycle activity in terms of its tissue of origin in over 13,000 samples. *NPJ Precis Oncol* (2022) 6(1):59. doi: 10.1038/s41698-022-00302-7

47. Zhao Z, Liu Y, Liu Q, Wu F, Liu X, Qu H, et al. The mRNA expression signature and prognostic analysis of multiple fatty acid metabolic enzymes in clear cell renal cell carcinoma. *J Cancer.* (2019) 10(26):6599. doi: 10.7150/jca.33024

48. Vantaku V, Dong J, Ambati CR, Perera D, Donepudi SR, Amara CS, et al. Multi-omics integration analysis robustly predicts high-grade patient survival and identifies CPT1B effect on fatty acid metabolism in bladder cancer. *Clin Cancer Res* (2019) 25(12):3689–701. doi: 10.1158/1078-0432.CCR-18-1515

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 08 December 2022
DOI 10.3389/fonc.2022.1054564

OPEN ACCESS

EDITED BY
Mingyue Tan,
Shanghai University of Traditional Chinese Medicine, China

REVIEWED BY
Zhijie Xu,
Central South University, China
Kai Fang,
Shanghai Pudong Hospital, China

*CORRESPONDENCE
Juan Tang
tangjuan@ntu.edu.cn
Jinfeng Ji
jijfeng@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 04 October 2022
ACCEPTED 28 October 2022
PUBLISHED 08 December 2022

CITATION
Gao L, Wang X, Wang X, Wang F, Tang J and Ji J (2022) A prognostic model and immune regulation analysis of uterine corpus endometrial carcinoma based on cellular senescence. *Front. Oncol.* 12:1054564. doi: 10.3389/fonc.2022.1054564

A prognostic model and immune regulation analysis of uterine corpus endometrial carcinoma based on cellular senescence

Lulu Gao[1†], Xiangdong Wang[2,3†], Xuehai Wang[3], Fengxu Wang[3], Juan Tang[3*] and Jinfeng Ji[2*]

[1]Department of Obstetrics and Gynecology, Nantong Maternal and Child Health Hospital Affiliated to Nantong University, Nantong, China, [2]Department of Integrated Traditional Chinese and Western Internal Medicine, Nantong Tumor Hospital, Affiliated Tumor Hospital of Nantong University, Nantong, China, [3]Department of Occupational Medicine and Environmental Toxicology, Nantong Key Laboratory of Environmental Toxicology, School of Public Health, Nantong University, Nantong, China

Background: This study aimed to explore the clinical significance of cellular senescence in uterine corpus endometrial carcinoma (UCEC).

Methods: Cluster analysis was performed on GEO data and TCGA data based on cellular senescence related genes, and then performed subtype analysis on differentially expressed genes between subtypes. The prognostic model was constructed using Lasso regression. Survival analysis, microenvironment analysis, immune analysis, mutation analysis, and drug susceptibility analysis were performed to evaluate the practical relevance. Ultimately, a clinical nomogram was constructed and cellular senescence-related genes expression was investigated by qRT-PCR.

Results: We ultimately identified two subtypes. The prognostic model divides patients into high-risk and low-risk groups. There were notable discrepancies in prognosis, tumor microenvironment, immunity, and mutation between the two subtypes and groups. There was a notable connection between drug-sensitive and risk scores. The nomogram has good calibration with AUC values between 0.75-0.8. In addition, cellular senescence-related genes expression was investigated qRT-PCR.

Conclusion: Our model and nomogram may effectively forecast patient prognosis and serve as a reference for patient management.

KEYWORDS

UCEC, cellular senescence, bioinformatics, prognosis, multi-omics analysis

Introduction

Uterine corpus endometrial carcinoma (UCEC) is one of the three major gynecological malignancies, second only to cervical cancer in incidence (1, 2). Hypertension, diabetes, obesity, infertility, and family history are risk factors for UCEC (3). However, because of the scarcity of effective timely detection of UCEC, many patients have progressed to advanced stages by the time they are diagnosed (4). At the same time, the poor prognosis for patients who develop metastases despite treatment is now a pressing issue (5). Treatment options other than first-line chemotherapy drugs remain limited (6). Studies show that the treatment and prognosis of patients can be assessed through predictive models and biomarkers (7). However, there are no credible biomarkers to assess the outcome for UCEC.

Cellular senescence is the central process of aging, bringing the cell cycle to a permanent standstill (8). Cellular senescence can promote repair and prevent tumorigenesis. Meanwhile, some degenerative diseases and cancers are associated with abnormal accumulation of senescent cells (9, 10). Senescent tumor cells can modulate the tumor microenvironment (TME), transform surrounding unsenescent cells into senescent cells, and recruit and activate immune cells to produce anti-tumor and pro-tumor effects (8, 9). Cellular senescence is capable of limiting tumor growth progression and is considered a potential therapeutic target (11). Adriamycin and bleomycin can induce senescence and thus exert anti-tumor effects. Therefore, studying the effects of cellular senescence in tumors can help develop new approaches to tumor therapy (12). However, the role of cellular senescence in UCEC and the relationship with UCEC prognosis remains unclear.

Materials and methods

Data collection

From TCGA and GEO databases, the gene expression and clinical data of UCEC were downloaded. The GEO cohort GSE119041 and TCGA cohort were acquired (13). Among them, patients in the integrated cohort of the TCGA cohort and the GEO cohort were randomly divided into training cohort and testing cohort at the ratio of 1:1, the integrated cohort was also defined as validation cohort. We normalized the expression of the genes by using "ComBat" algorithm from the "sva" package (14). Patients with inadequate clinical data and survival information were eliminated.

The clustering analysis

We collected 307 cellular senescence related genes from the previous study (15). Full details of these genes were shown in Table S1. The "ConsensusClusterPlus" package was used to perform consistent unsupervised cluster analysis to classify patients into different subtypes. We screened out clusters with high intra-type correlation and low inter-type correlation for subsequent analysis (16).

Multi-omics analysis of UCEC subtypes based on senescence genes

First, to validate the categorization of patient subtypes, we used principal component analysis (PCA). We investigated the link with the subtypes and patient clinical characteristics. We then performed a survival analysis using the "survival" package to draw Kaplan–Meier curves to assess differences in survival between subtypes. Next, we explored the differences in the TME between different subtypes. Violin plots were used to show the distribution of TME scores for each sample across subtypes. a score of 22 immune cells was obtained by the CIBERSORT method (17). To measure the amount of immune cell infiltration, the single sample gene set enrichment analysis (ssGSEA) technique was utilized (18). Finally, we explored differences in PD-L1 and PD-L2 expression among different subtypes.

Enrichment analysis

Using the "clusterProfiler" software package, we performed Gene Ontology (GO) analysis to identify functions for these genes, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis to identify enriched pathways for these genes (19). We retained analysis results with p-values less than 0.05 and displayed them in bar graphs.

Difference analysis

Based on gene expression between the two subtypes, we screened for genes that differed between the two subtypes (20). In addition, we analyzed the pathways that differed between the two subtypes by means of KEGG enrichment analysis.

The differential genes clustering analysis and multi-omics analysis

First, we used the same method as above for cluster analysis. Then, we explored the association of this subtype with clinical factors and performed survival analysis. Besides, we performed TMB analysis and checkpoint analysis of PD-L1 and PD-L2.

Model construction and evaluation

In the training cohort, we performed the least absolute and selection operator (LASSO) regression analysis to select cellular

senescence related genes to connect to the prognosis. The model's predictive performance was tested using test and validation cohorts. Based on the median risk score, we classified the patients into two groups: high-risk and low-risk. Between the two groups, we investigated variations in clinical features and patient outcomes. The time-dependent receiver operating characteristic (ROC) curve was utilized to assess the model's accuracy. Besides, univariate and multivariate cox analyses were also performed (21).

Multi-omics analysis for the model

First, the link between risk scores and clinical factors was investigated. We then explored the TME based on the model. One-class logistic regression (OCLR) machine-learning algorithm was used to quantify the stemness of tumor samples by calculating cancer stem cell indices (22). Pearson analysis was used to reveal the correlation of risk score and RNAss. Between the two groups, the GSEA analysis was carried out to evaluate variations in enriched pathways. Besides, we also performed immune microenvironment (IME) analysis. We immunotyped the patients and investigated the association with both risk score and immunotyping to learn more about the based on risk score and immunity.

Studies showed that tumor mutational burden (TMB) correlates with IME (23). Therefore, we calculated TMB for each sample by somatic mutation profiles and investigated the link between risk score and TMB. Based on the median TMB, we separated patients into high-TMB and low-TMB groups and performed survival analysis. In addition, we combined TMB with risk scores for survival analysis. Besides, we analyzed the relationship among riskscores and microsatellite instability (MSI) and immunophenoscore (IPS).

The "PRROPHOPIC" pack includes hundreds of medicines (24). From it, we calculated the half inhibitory concentration (IC50) value of the drug and screened out the drugs with significant differences in the two risk groups.

Nomogram construction and evaluation

We created a nomogram using the riskscores and clinical data. The nomogram's accuracy was assessed using the C-index, ROC curve, and calibration curve.

Quantitative RT-PCR

A total of 12 UCEC tissues from patients in the Nantong Maternal and Child Health Hospital Affiliated to Nantong University were paired with normal tissues. The Ethics Committee of the Nantong Maternal and Child Health Hospital Affiliated to Nantong University approved the study. All patients signed the informed consent form. Use TRIZOL reagent (Thermo Fisher Scientific, USA) to separate total RNA from the sample, then use Revert Aid first strand cDNA synthesis kit (Thermo Fisher Scientific, USA) to reverse transcribe it into cDNA, and use SYBR Green PCR kit (Takara, Tokyo, Japan) for real-time quantitative PCR (qRT-PCR) analysis. GAPDH was used to regulate the relative expression of genes. The sequence is listed in Supplementary Table S3.

Results

Establishment and assessment of senescence subtypes

We included 593 patients from both TCGA and GEO cohorts in our study for further analysis. Based on cellular senescence related gene expression, we classified patients using a consensus clustering approach(Figure S1). The results of the analysis show that k=2 is the optimal number of groups (Figure 1A). We then divided them into subtype A and subtype B based on the above results. PCA analysis indicated that subtypes A and B successfully distinguished patients (Figure 1B). Survival analysis incidated that our subtype successfully stratified the survival of patients, and the survival time of subtype A was longer (Figure 1C). However, after comparing the clinical factors of the patients, we found no difference in the expression of pyroptotic genes with age, stage, grade, survival status, and histological type (Figure 1D).

Multi-omics analysis of different senescence subtypes

TME plays a key role in tumorigenesis and progression. Therefore, we first analyzed the TME. Violin plots showed significant differences in stromal, immune, and ESTIMATE scores between the two subtypes (Figure 2A). We further analyzed the immune-related functions and infiltration of immune cells of two subtypes based on the above results. A subtype had higher infiltration levels of NK cells activated, T cells regulatory, and T cells CD8, while B cells naive, T cells follicular helper, and Macrophages M1 had greater levels of infiltration in the B subtype (Figure 2B). ssGSEA analysis further confirmed that immune cell infiltration levels differed significantly between the two subtypes (Figure 2C). Besides, the expression of HLA-A, HLA-DMA, and HLA-F was higher in subtype A, while the expression of HLA-DMB and HLA-DOA in subtype B was higher (Figure 2D). The results of the checkpoint analysis indicated B subtype showed greater levels of PD-L1 and PD-L2 expression (Figures 2E, F).

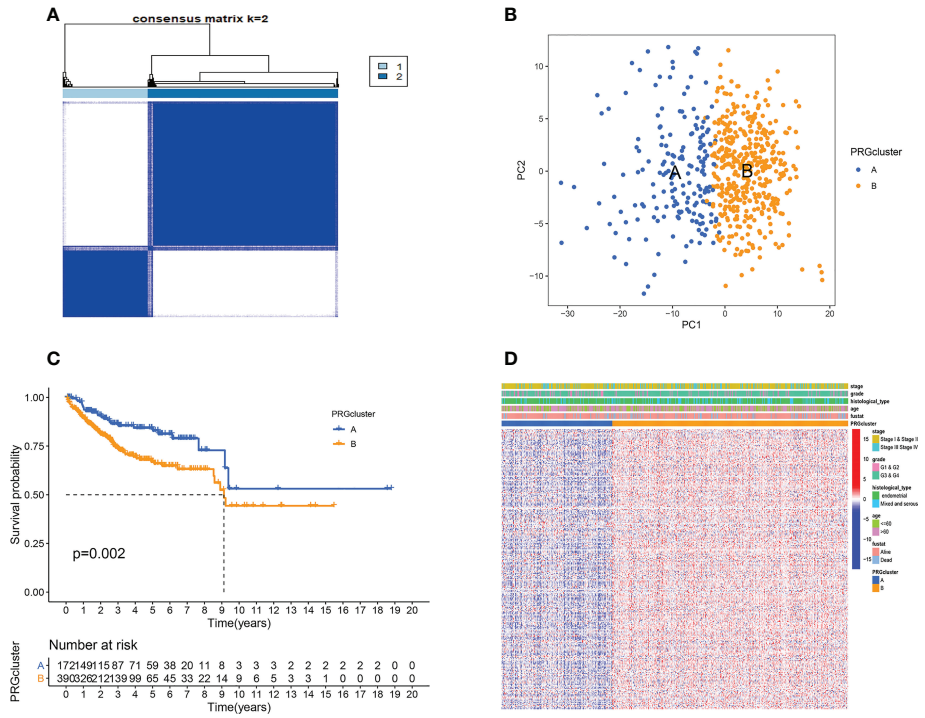


FIGURE 1
Cellular senescence subtypes and clinical assessment. **(A)** Two subtypes and their associated regions. **(B)** PCA analysis. There are significant differences between the two subtypes. **(C)** Survival analysis. Subtype B has a poorer prognosis. **(D)** There were no differences in clinical factors between the two subtypes.

We also analyzed gene function and enriched pathways. GO enrichment analysis revealed these genes were primarily associated with cell mitosis, metabolism of genetic material, and ATP metabolism (Figure 2G). KEGG enrichment analysis revealed these genes were primarily associated with cell cycle, protein processing, transport, and DNA replication (Figure 2H). Besides, it also revealed subtype A was substantially more concentrated in lipid metabolism, and subtype B had considerable cell cycle, cell division, and tumor enrichment (Figure 2I).

Differential genes subtypes

Through differential analysis, we identified 1219 differential genes. Based on these genes, we used the same cohort and method to further subtype the patients (Figure S2A). We found dividing patients into two subtypes (A and B) was optimal (Figure S2B). Besides, the survival time of the two subtypes was significantly different (Figure S2C). However, the heatmap showed no differences in clinical factors between the two subtypes (Figure S2D).

Then, we performed TME analysis. The results showed that subtype A had higher stromalscore, immunescore, and estimatescore, while subtype B had higher tumorpurity (Figures 3A–D). In addition, the A subtype of NK cells activated, T cells regulatory (Tregs), and T cells CD8 have a higher degree of infiltration, and the B subtype of Macrophages M1, T cells follicular helper, and B cells naive have a higher degree of infiltration (Figure 3E). The results of ssGSEA analysis further confirmed that immune cell infiltration differed significantly between the two subtypes (Figure 3F). At the same time, the PD-L1 and PD-L2 genes of subtype B are highly expressed (Figures 3G, H). Figure 3I showed that the expression of HLA-related genes of the two subtypes was significantly different. This is basically consistent with the analysis of cellular senescence subtypes.

Model construction and evaluation

After LASSO analysis, a total of 4 genes were screened (Figures 4A, B). The model's calculating formula was as follows: riskscore = BZW2*0.44481118 - NRIP1*0.38695576 + ARHGAP29*0.22408622 + SIX1*0.18719355. Based on the median risksocre in the training cohorts, patients in the three cohorts were separated into high- and low-risk groups. Figure 4C showed the distribution of patients grouped by two cellular senescence subtypes, two differential gene subtypes, high

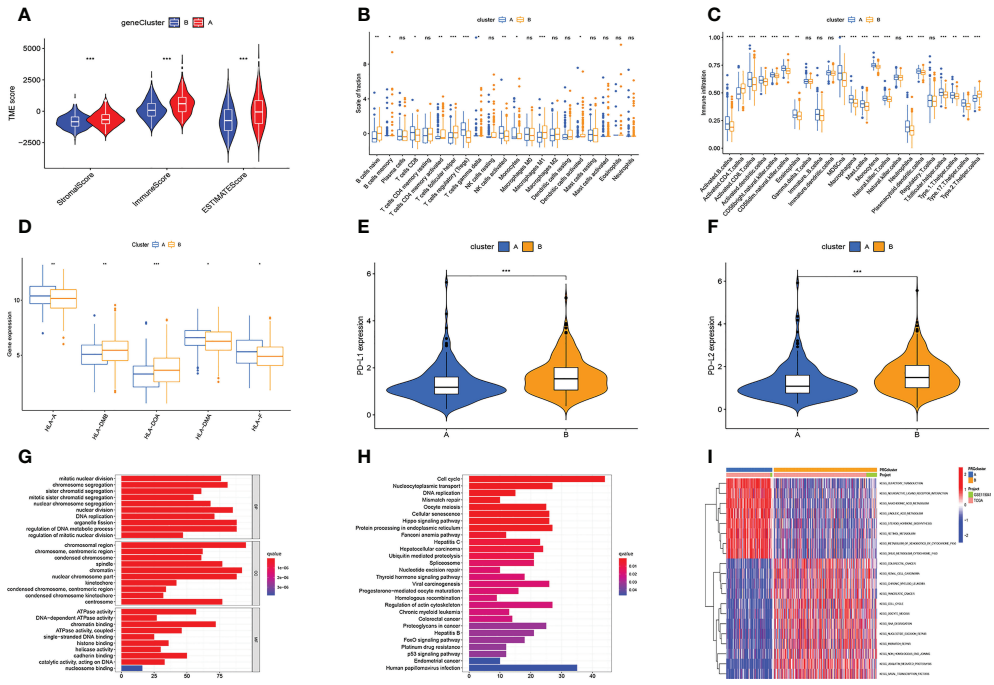


FIGURE 2
Multi-omics analysis based on senescence cluster. **(A)** TME analysis. Subtype A has a higher TME score. **(B, C)** Differences in immune cell infiltration levels. **(D)** Differences in HLA-related gene expression levels. **(E, F)** The PD-L1 and PD-L2 genes of subtype B are highly expressed. **(G, H)** The GO **(G)** and KEGG **(H)** enrichment analysis. **(I)** Differential KEGG enriched pathways between the two subtypes. Adjusted p-values were shown as ns, not significant; *p < 0.05; **p < 0.01; ***p < 0.001.

and low-risk groups, and survival status. We also observed that both the cellular senescence subtype and the differential gene subtype had a higher risk score for the B subtype (Figures 4D, E). Figure 4F shows that RNAss values are positively correlated with risk scores. Furthermore, the risk score was linked to patient's clinical factors. The higher risk score, the more advanced and poorly differentiated tumors, and the greater the likelihood of death (Figures S3A–D). We also found a lower risk score for tumors originating from endometrial tissue and a higher risk for mixed and serous tissue (Figure S3E).

We then analyzed the relationship of the model to patient survival. Patients were separated into high-risk and low-risk groups based on the median (Figures S4A–C). At the same time, the number of patient deaths was proportional to the risk score (Figures S4D–F). Furthermore, in the high-risk group, BZW2, ARHGAP29, and SIX1 were overexpressed, whereas NRIP1 was overexpressed in the low-risk group (Figures S4G–I).

Then, we evaluated the accuracy of the model. The high-risk group had the worst prognosis among the three groups (Figures 5A–C). Figures S4A–H showed the results of survival analysis for clinical factors. The AUC of the training cohort at 1, 3, and 5 years was 0.652, 0.722, and 0.771, respectively (Figure 5D). The AUC of the test cohort at 1, 3, and 5 years was 0.621, 0.619, and 0.645, respectively (Figure 5E). The AUC of the validation cohort at 1, 3, and 5 years was 0.644, 0.671, and 0.697, respectively (Figure 5F).

Independent prognostic analysis

For independent prognostic analysis, univariate and multivariate COX regression models were utilized. The results of the univariate COX analysis are as follows (Table S1). In the training cohort, histological type, stage, riskscore were independent prognostic factors. The grade was also an independent prognostic factor in the testing cohort and validation cohort. In the three cohorts, multivariate COX analysis demonstrated riskscore and stage were independent predictive variables (Table S2).

The model's multi-omics analysis

First, GSEA analysis revealed the high-risk group was mostly associated with cardio-renal diseases (Figure 6A). The low-risk group was mostly associated with immunity and rejection (Figure 6B). Then, we analyzed the relationship between the TME and the model. StromalScore, ImmuneScore, and

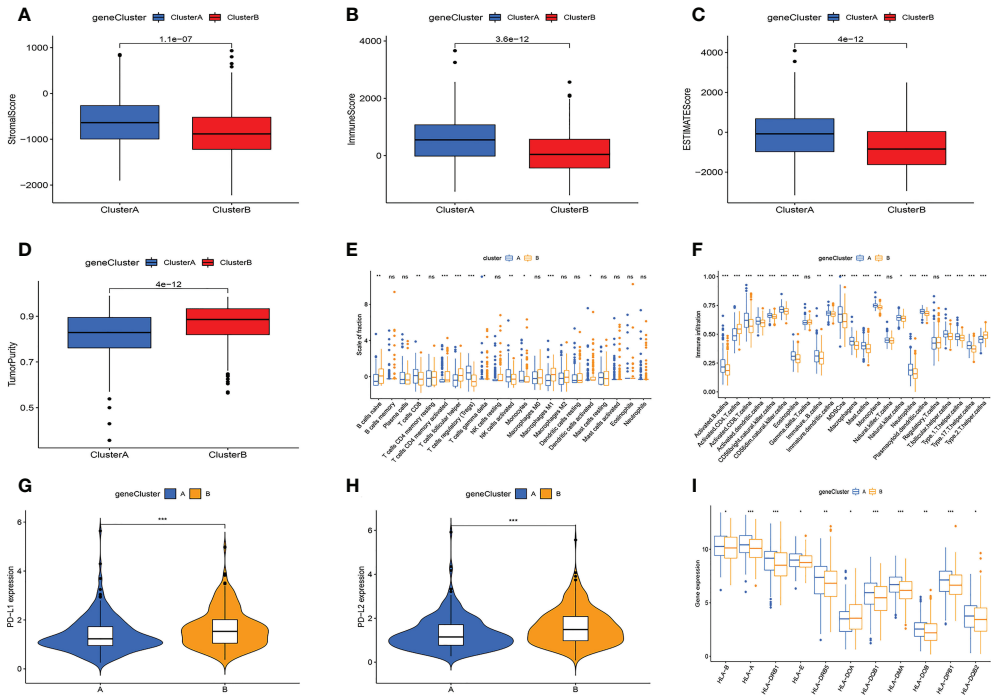


FIGURE 3
Multi-omics analysis based on differential genes cluster. **(A-D)** TME analysis. Subtype A has higher TME scores and subtype B has higher tumor purity. **(E, F)** The amount of immune cell infiltration differed significantly. **(G)** PD-L1, PD-L2 expression levels are higher in subtype **(B-H)** The A subtype has higher expression levels of HLA-related genes. Adjusted p-values were shown as ns, not significant; *p < 0.05; **p < 0.01; ***p < 0.001.

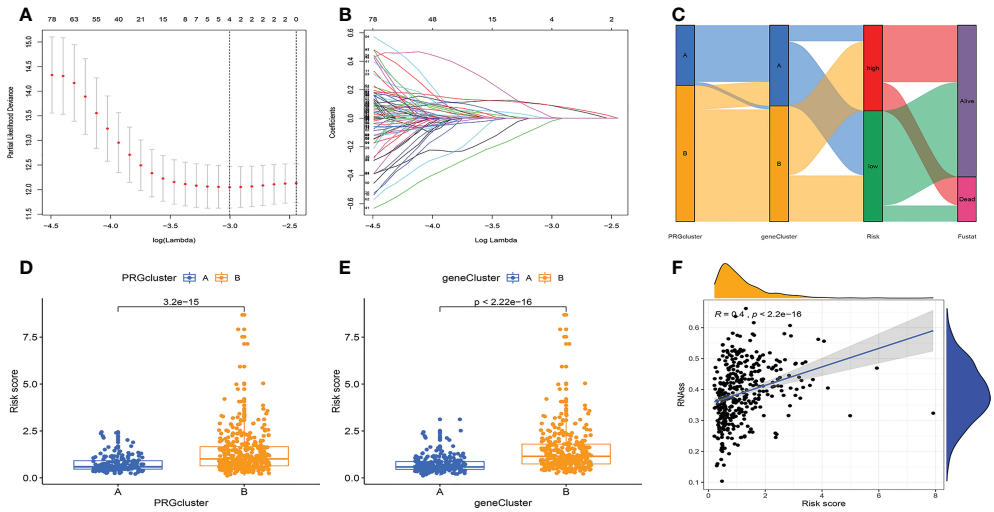


FIGURE 4
Prognostic model construction. **(A, B)** LASSO regression analysis. 4 genes were screened to build a prognostic model. **(C)** Distribution of different subtypes, risk groups, and survival outcomes. **(D, E)** Distribution of risk scores for different subtypes. **(F)** RNAss values are positively correlated with risk scores.

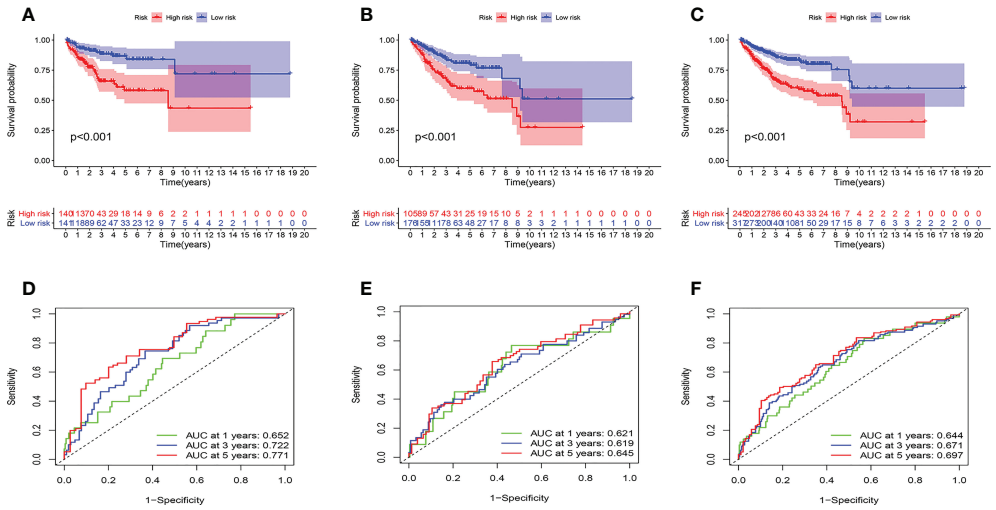


FIGURE 5
Prognostic model evaluation. **(A-C)** survival analysis. the high-risk group had a worse prognosis in training **(A)**, test **(B)**, and validation **(C)** cohorts. **(D-F)** ROC curves. The AUC value of the model is basically between 0.6 and 0.7 in three cohorts.

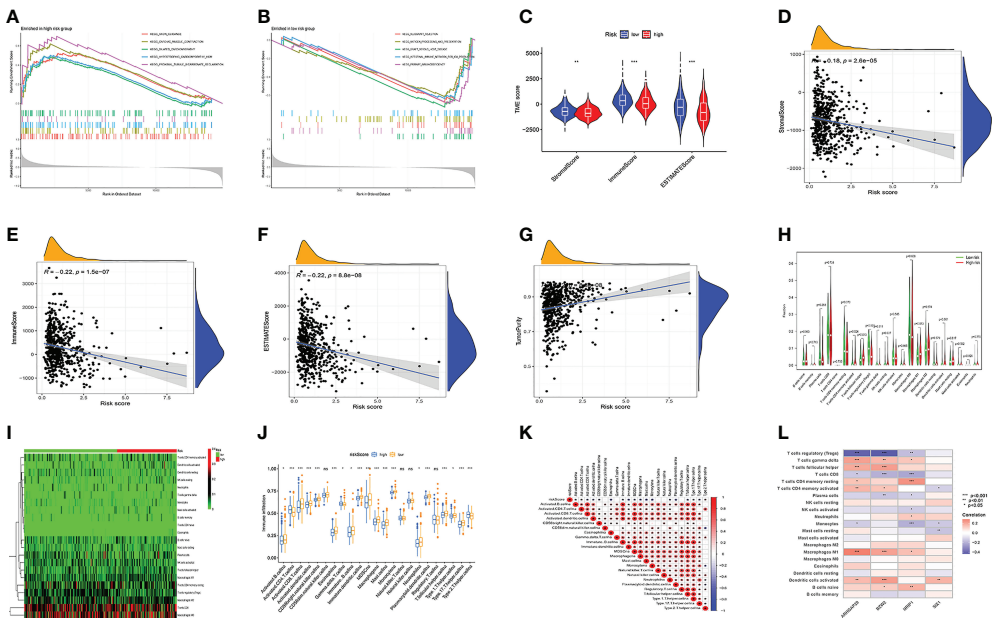


FIGURE 6
Model multi-omics analysis. **(A, B)** GSEA enrichment analysis. **(C-F)** TME analysis. the low-risk group had higher StromalScore, ImmuneScore, and ESTIMATEScore. Risk Score is inversely proportional to StromalScore, ImmuneScore, and ESTIMATEScore, and proportional to TumorPurity. **(G-I)** The amount of immune cell infiltration differed significantly. **(J)** Immune cell infiltration and risk score were linked. **(K)** immune cells and Risk score were linked. **(L)** immune cells and model genes were linked. Adjusted p-values were shown as ns, not significant; *p < 0.05; **p < 0.01; ***p < 0.001.

ESTIMATEScore were greater in low-risk group (Figure 6C). In addition, riskScore is inversely proportional to StromalScore, ImmuneScore, and ESTIMATEScore, and proportional to TumorPurity (Figures 6D–G). Figure S5A illustrated the distribution of immune cell in two groups per patient. We then investigated the model's connection to immune cell infiltration. Besides, T cells CD4 memory activated, T cells follicular helper, T cells regulatory, NK cells resting,

Macrophages M1, and Dendritic cells activated were distinct in the two groups (Figures 6H, I). SsGSEA analysis also confirmed that in the high-risk group, most immune cells had higher infiltration levels (Figure 6J). The risk score was significantly associated with immune cells, model genes (Figures 6K, L). We found that T cell regulatory were negatively correlated with riskscore, and all the rest of cells had a positive correlation to risk score. (Figures S5B–F). Figure 6K showed the relationship between model genes and immune cells. Then, we divided patients into four subtypes based on their immunity (Figure S5G). Different types of immune infiltration correspond to tumor promotion and tumor inhibition, including C1 (wound healing), C2 (INF-g dominance), C3 (inflammation) and C4 (lymphocyte depleted) (25). The risk score for the C2 subtype was the greatest, while the risk score for the C3 subtype was the lowest (Figure S5H). In addition, significant variations between the two groups were also seen in the expression of immunological checkpoint genes (Figure S5I). Among them, CTLA4, PDCD1LG2, and PDCD1 were most associated with risk scores (Figure S5J). Risk scores were inversely correlated with PDCD1LG2, CTLA4, and PDCD1, and favorably correlated with PDCD1LG2 (Figures S5K–M).

Studies have demonstrated that TMB can serve as an important component of composite predictors to guide tumor immunotherapy (26). We found that the three genes with the greatest mutation probability in the high-risk group were TP53, PIK3CA, and PTEN, while the three genes with the highest mutation probability in the low-risk group were PTEN, ARID1A, and PIK3CA (Figures 7A, B). We then performed survival analysis. The prognosis of patients with high-TMB scores and high risk score was greater (Figures 7C, D). The research by Ganesh et al. illustrated MSI is closely related to the sensitivity to immunotherapy (27). The low MSI accounted for the least, and the high MSI group had the lowest risk score (Figures 7E, F). To further guide the patient's treatment, we performed a drug sensitivity analysis. First, we screened out drugs related to model genes, including Tamoxifen, Dasatinib, Panobinostat, etc (Figure 8A). Next, we further screened drugs sensitive to the high-risk group, including Gemcitabine, Doxorubicin, Docetaxel, Cisplatin, Vinorelbine, Paclitaxel, Vinblastine (Figures 8B–H).

Nomogram construction and validation

We built a nomogram by combining riskscore and clinical factors. According to the nomogram, the 1-, 3-, and 5-year mortality rates for the patients were 0.0104, 0.0445, and 0.0644, respectively (Figure 9A). The calibration curve showed the nomogram had an excellent calibration (Figure 9B). The C-index showed that the nomogram performed better than the risk score and clinical factors (Figure 9C). The same conclusion was drawn from the ROC curve, with the AUC of 0.751, 0766, and 0.786 in the nomogram at years 1, 3, and 5, respectively (Figures 9D–F).

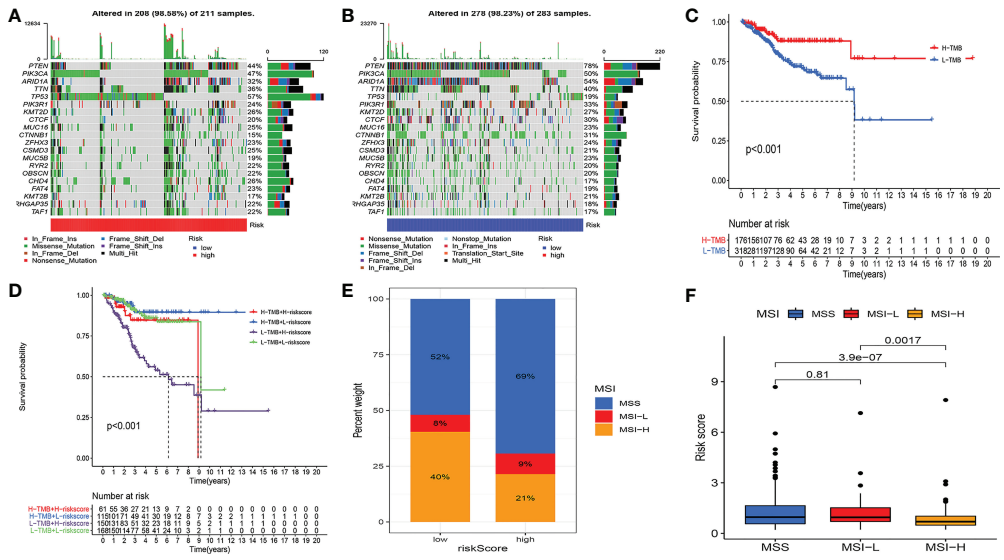


FIGURE 7
Multi-omics analysis of the model. **(A, B)** Gene mutation frequencies in both groups. **(C, D)** Survival analysis. H-TMB has a better prognosis. **(E, F)** MSI analysis. The low MSI accounted for the least, and the high MSI group had the lowest risk score.

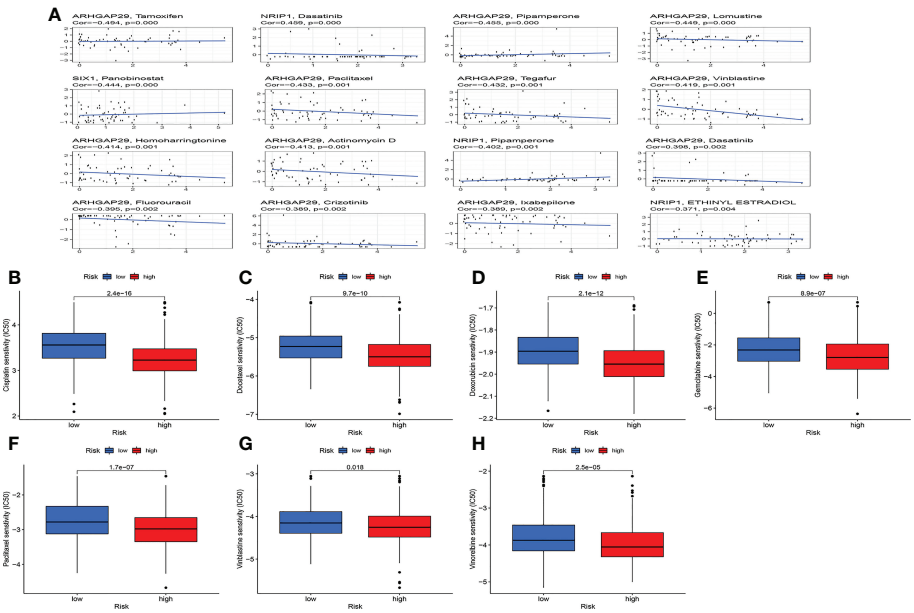


FIGURE 8
Drug sensitivity analysis. **(A)** Relationship between model genes and sensitive drugs. **(B-H)** Sensitive drugs in high-risk group.

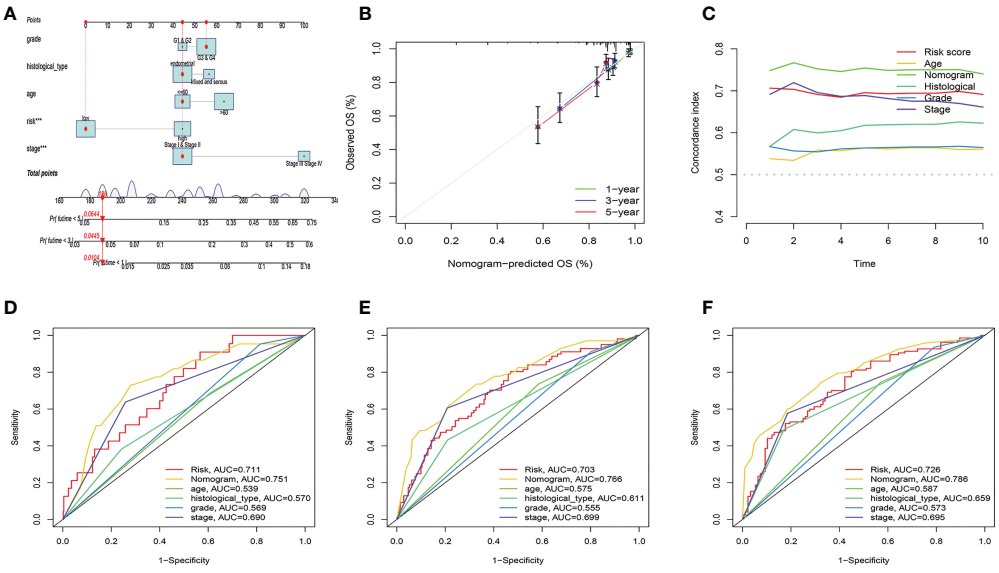


FIGURE 9
Nomogram construction and evaluation. **(A)** According to the nomogram, the 1-, 3-, and 5-year mortality rates for the patients were 0.0104, 0.0445, and 0.0644, respectively. **(B)** Calibration curve for nomogram. **(C)** C-index curve. **(D-F)** ROC curves.

Validating gene expression level of cellular senescence-related genes in UCEC samples

To validate the expression levels of cellular senescence-related genes, we used qRT-PCR to detect the expression levels of six cellular senescence-related genes in 12 UCEC samples and 12 normal tissues. The results indicated that ARHGAP29 expression was significantly higher in tumor samples, while GNLY and NRIP1expression was significantly lower in UCEC samples. There was no significant difference in BATF, BZW2 and SIX1 expression (Figure S6).

Discussion

In this study, to evaluate the involvement of senescence genes in UCEC, we did a complete bioinformatics analysis. Based on the senescence gene, we began by categorizing the patients into two groups. Subtype B has a worse prognosis. TME, immune checkpoint gene expression, and immune function also differed significantly between the two subtypes. We further discovered two gene subtypes based on the differential genes. The results of correlation analysis showed that different genes can be used as indicators of patient prognosis and TME. Therefore, the prognostic model was built using differential genes. The model's predictive ability was proven using survival analysis and ROC curves. Furthermore, this prognostic model was significantly associated with clinical factors, TME, immune-related markers, TMB, MSI, and drug sensitivity. Finally, we built a nomogram by combining riskscore and clinical factors. The results showed that the nomogram was successful in stratifying patients and guiding them in prognostic assessment and treatment selection.

In this study, we verified the expression levels of cell senescence-related genes in tumor tissues and normal tissues. Perhaps due to the small sample size, there was no difference in BATF, BZW2 and SIX1 expression between tumor and normal tissues. It is necessary to expand the sample size to further verify this result. Senescence is a steady state that removes sick cells and stabilizes the collective internal environment (11). It is also thought to prevent tumor development (28). However, recent studies have found that tumor progression can also be caused by cellular senescence (29). Senescent cells secrete signaling molecules that affect tumor proliferation, invasion and metastasis, and angiogenesis (30). In addition, the senescence of some tumor cells is reversible and they can escape cellular senescence and re-enter the cell cycle, which is an important cause of tumor recurrence and progression (31). As a result, it is critical to thoroughly investigate the clinical importance of cellular senescence in malignancies. However, there are currently no studies on the role of cellular senescence in UCEC.

Four genes have been identified as being involved in illness development and progression. BZW2 is a protein that has a role in cell adhesion (32). Huang et al. showed that BZW2 promoted colorectal cancer progression (33). NRIP1 is a nuclear receptor protein, and Its high expression is linked to a bad prognosis of gastric cancer (34). ARHGAP29 is a GTPase that stimulates prostate cancer development and metastasis (35). SIX1 is a transcription factor with an important role in tumorigenesis (36, 37). Our prognostic model combines these four genes, which will give us a better understanding for cancer cells.

The function of programmed cell death in tumor therapy and TME are receiving increasing attention (38, 39). Tumor growth must evade tumor immunity, which is also considered an important marker of tumor progression (40, 41). Despite breakthroughs in the treatment of aggressive malignancies with immunotherapy, a large minority of patients still have no impact on treatment (42, 43). The immune microenvironment of UCEC can predict patient survival (44). In this study, GSEA analysis revealed that the low-risk group was mostly associated to immunity. In addition, our study also found the riskscore was inversely related to the patient's stromalscore, immunescore, estimatescore and proportional to tumorpurity. At the same time, we also found that major immune checkpoint genes were up-regulated in the low-risk group. This means that patients with low-risk scores are more immunogenic and may benefit from immunotherapy. Therefore, our study may guide the immunotherapy of UCEC patients.

Studies have shown that immunotherapy is more effective in people with a high TMB (45). Tissue TMB can also predict patient response to immune checkpoint therapy (46). TP53 mutation is an independent marker of poor prognosis (47). There is also evidence that human carcinogens can induce TP53 mutations (48). Our study also reached similar conclusions. The mutation rate of TP53 is substantially greater in the high-risk group than in the low-risk group. This helps us explore the causes of tumorigenesis and the choice of treatment options for patients. Besides, drug resistance of tumors has always been one of the challenges of UCEC treatment (49). It is also difficult to effectively treat advanced cases (50). To this end, our study screened drug candidates for relevanche to prognostic models.

Our study has some limitations. First, our studies are all from public databases. Due to the limited access to public data sets and the limited amount of data, the clinicopathological parameters analyzed in this study were not comprehensive, and there were errors or biases. In the future, we will conduct basic experiments *in vivo* or *in vitro* to confirm our findings. Second, our study was a retrospective study. Future prospective clinical validation is needed.

This is the first prognostic model of UCEC based on cellular senescence genes to our knowledge. Our analyses reveal a broad range of regulatory regulatory mechanisms that facilitate individualized treatment and prognosis prediction in patients.

Conclusion

We constructed a UCEC prognostic model based on cellular senescence genes and combined with clinical factors to construct nomograms, which showed good predictive performance. Using this model, the prognosis and TME of UCEC patients can be accurately estimated. Furthermore, our findings may lead to new approaches for UCEC treatment.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the Nantong Maternal and Child Health Hospital Affiliated to Nantong University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JT and JJ conceived the study and participated in the study design and performance. LG and XDW conducted the bioinformatics analysis and manuscript writing. XHW and FW revised the manuscript. All authors contributed and approved the submitted version.

Acknowledgments

We would like to extend our gratitude to the researchers and study patients for their contributions.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.1054564/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Cluster analysis. Cluster analysis heatmap of cellular senescence genes (k = 3-9).

SUPPLEMENTARY FIGURE 2
Cluster analysis. **(A)** Cluster analysis heatmap of differential genes (k = 3-9). **(B)** Dividing patients into two subtypes was optimal. **(C)** Survival analysis. Subtype B has a poorer prognosis. **(D)** There were no differences in clinical factors between the two subtypes.

SUPPLEMENTARY FIGURE 3
Association of clinical factors with risk score. (A-E) Risk score in patients with different age, grade, fustat, and histological_type.

SUPPLEMENTARY FIGURE 4
Model evaluation. **(A-C)** Patients were divided into high- and low-risk groups based on the median risk score. **(D-F)** As the risk value increased, the proportion of UCEC patients who died increased. **(G-I)** in the high-risk group, BZW2, ARHGAP29, and SIX1 were highly expressed, whereas, in the low-risk group, NRIP1 was highly expressed.

SUPPLEMENTARY FIGURE 5
Immunoassay of the model. **(A)** The distribution of immune cells for each sample in the two groups. **(B-F)** The relationship between risk scores and immune cells. T cell regulatory (Tregs) were negatively correlated with risk score, and the remaining cells were all positively correlated with risk score. **(G, H)** The relationship between immunophenotyping and risk score. **(I, J)** Immune checkpoint analysis. There were also significant differences in the expression of immune checkpoint genes between the two groups. **(K-M)** PDCD1LG2 was positively associated with risk scores, CTLA4, and PDCD1 were negatively associated with risk scores.

SUPPLEMENTARY FIGURE 6
The cellular senescence-related genes expression was investigated by qPT-PCR.

References

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

2. Sponholtz TR, Palmer JR, Rosenberg L, Hatch EE, Adams-Campbell LL, Wise LA. Reproductive factors and incidence of endometrial cancer in U.S. black women. *Cancer causes control: CCC* (2017) 28(6):579–88. doi: 10.1007/s10552-017-0880-4

3. Busch EL, Crous-Bou M, Prescott J, Chen MM, Downing MJ, Rosner BA, et al. Endometrial cancer risk factors, hormone receptors, and mortality prediction. *Cancer epidemiol Biomarkers Prev* (2017) 26(5):727–35. doi: 10.1158/1055-9965.EPI-16-0821

4. Suryo Rahmanto Y, Shen W, Shi X, Chen X, Yu Y, Yu ZC, et al. Inactivation of Arid1a in the endometrium is associated with endometrioid tumorigenesis through transcriptional reprogramming. *Nat Commun* (2020) 11(1):2717. doi: 10.1038/s41467-020-16416-0

5. Stubert J, Gerber B. Current issues in the diagnosis and treatment of endometrial carcinoma. *Geburtshilfe und Frauenheilkunde* (2016) 76(2):170–5. doi: 10.1055/s-0035-1558230

6. Lee YC, Lheureux S, Oza AM. Treatment strategies for endometrial cancer: current practice and perspective. *Curr Opin obstetrics gynecol* (2017) 29(1):47–58. doi: 10.1097/GCO.0000000000000338

7. Argentiero A, Solimando AG, Krebs M, Leone P, Susca N, Brunetti O, et al. Anti-angiogenesis and immunotherapy: Novel paradigms to envision tailored approaches in renal cell-carcinoma. *J Clin Med* (2020) 9(5):1594. doi: 10.3390/jcm9051594

8. Calcinotto A, Kohli J, Zagato E, Pellegrini L, Demaria M, Alimonti A. Cellular senescence: Aging, cancer, and injury. *Physiol Rev* (2019) 99(2):1047–78. doi: 10.1152/physrev.00020.2018

9. Hernandez-Segura A, Nehme J, Demaria M. Hallmarks of cellular senescence. *Trends Cell Biol* (2018) 28(6):436–53. doi: 10.1016/j.tcb.2018.02.001

10. Sharpless NE, Sherr CJ. Forging a signature of *in vivo* senescence. *Nat Rev Cancer* (2015) 15(7):397–408. doi: 10.1038/nrc3960

11. Sun JX, Liu CQ, Xu JZ, An Y, Xu MY, Zhong XY, et al. A four-Cell-Senescence-Regulator-Gene prognostic index verified by genome-wide CRISPR can depict the tumor microenvironment and guide clinical treatment of bladder cancer. *Front Immunol* (2022) 13:908068. doi: 10.3389/fimmu.2022.908068

12. Zhou L, Niu Z, Wang Y, Zheng Y, Zhu Y, Wang C, et al. Senescence as a dictator of patient outcomes and therapeutic efficacies in human gastric cancer. *Cell Death Discovery* (2022) 8(1):13. doi: 10.1038/s41420-021-00769-6

13. Gultekin O, Gonzalez-Molina J, Hardell E, Moyano-Galceran L, Mitsios N, Mulder J, et al. FOXP3+ T cells in uterine sarcomas are associated with favorable prognosis, low extracellular matrix expression and reduced YAP activation. *NPJ Precis Oncol* (2021) 5(1):97. doi: 10.1038/s41698-021-00236-6

14. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinf (Oxford England)* (2012) 28(6):882–3. doi: 10.1093/bioinformatics/bts034

15. Lin W, Wang X, Wang Z, Shao F, Yang Y, Cao Z, et al. Comprehensive analysis uncovers prognostic and immunogenic characteristics of cellular senescence for lung adenocarcinoma. *Front Cell Dev Biol* (2021) 9:780461. doi: 10.3389/fcell.2021.780461

16. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinf (Oxford England)* (2010) 26 (12):1572–3. doi: 10.1093/bioinformatics/btq170

17. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* (2015) 12 (5):453–7. doi: 10.1038/nmeth.3337

18. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* (2015) 160(1-2):48–61. doi: 10.1016/j.cell.2014.12.033

19. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* (2017) 45 (D1):D353–d361. doi: 10.1093/nar/gkw1092

20. Liu J, Geng R, Ni S, Cai L, Yang S, Shao F, et al. Pyroptosis-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with UCEC. *Mol Ther Nucleic Acids* (2022) 27:1036–55. doi: 10.1016/j.omtn.2022.01.018

21. Ni X, Chen C, Cui G, Ding W, Liu J. Crosstalk of RNA adenosine modification-related subtypes, establishment of a prognostic model, and immune infiltration characteristics in ovarian cancer. *Front Immunol* (2022) 13:932876. doi: 10.3389/fimmu.2022.932876

22. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* (2018) 173(2):338–354.e315. doi: 10.1016/j.cell.2018.03.034

23. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* (2019) 51(2):202–6. doi: 10.1038/s41588-018-0312-8

24. Geeleher P, Cox N, Huang RS. pRRophetic: an r package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS One* (2014) 9(9):e107468. doi: 10.1371/journal.pone.0107468

25. Tamborero D, Rubio-Perez C, Muiños F, Sabarinathan R, Piulats JM, Muntasell A, et al. A pan-cancer landscape of interactions between solid tumors and infiltrating immune cell populations. *Clin Cancer res: An Off J Am Assoc Cancer Res* (2018) 24(15):3717–28. doi: 10.1158/1078-0432.CCR-17-3509

26. Jardim DL, Goodman A, de Melo Gagliato D, Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell* (2021) 39 (2):154–73. doi: 10.1016/j.ccell.2020.10.001

27. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol* (2019) 16(6):361–75. doi: 10.1038/s41575-019-0126-x

28. Lee S, Schmitt CA. The dynamic nature of senescence in cancer. *Nat Cell Biol* (2019) 21(1):94–101. doi: 10.1038/s41556-018-0249-2

29. Hanahan D. Hallmarks of cancer: New dimensions. *Cancer Discovery* (2022) 12(1):31–46. doi: 10.1158/2159-8290.CD-21-1059

30. Faget DV, Ren Q, Stewart SA. Unmasking senescence: context-dependent effects of SASP in cancer. *Nat Rev Cancer* (2019) 19(8):439–53. doi: 10.1038/s41568-019-0156-2

31. De Blander H, Morel AP, Senaratne AP, Ouzounova M, Puisieux A. Cellular plasticity: A route to senescence exit and tumorigenesis. *Cancers* (2021) 13 (18):4561. doi: 10.3390/cancers13184561

32. Guo Z, Neilson LJ, Zhong H, Murray PS, Zanivan S, Zaidel-Bar R. E-cadherin interactome complexity and robustness resolved by quantitative proteomics. *Sci Signaling* (2014) 7(354):rs7. doi: 10.1126/scisignal.2005473

33. Huang L, Chen S, Fan H, Ai F, Sheng W. BZW2 promotes the malignant progression of colorectal cancer *via* activating the ERK/MAPK pathway. *J Cell Physiol* (2020) 235(5):4834–42. doi: 10.1002/jcp.29361

34. Fang D, Lu G. Expression and role of nuclear receptor-interacting protein 1 (NRIP1) in stomach adenocarcinoma. *Ann Trans Med* (2020) 8(20):1293. doi: 10.21037/atm-20-6197

35. Shimizu K, Matsumoto H, Hirata H, Ueno K, Samoto M, Mori J, et al. ARHGAP29 expression may be a novel prognostic factor of cell proliferation and invasion in prostate cancer. *Oncol Rep* (2020) 44(6):2735–45. doi: 10.3892/or.2020.7811

36. Rafiq A, Aashaq S, Jan I, Beigh MA. SIX1 transcription factor: A review of cellular functions and regulatory dynamics. *Int J Biol macromol* (2021) 193(Pt B):1151–64. doi: 10.1016/j.ijbiomac.2021.10.133

37. Wu W, Ren Z, Li P, Yu D, Chen J, Huang R, et al. Six1: a critical transcription factor in tumorigenesis. *Int J Cancer* (2015) 136(6):1245–53. doi: 10.1002/ijc.28755

38. Wang H, Lin D, Yu Q, Li Z, Lenahan C, Dong Y, et al. A promising future of ferroptosis in tumor therapy. *Front Cell Dev Biol* (2021) 9:629150. doi: 10.3389/fcell.2021.629150

39. Niu X, Chen L, Li Y, Hu Z, He F. Ferroptosis, necroptosis, and pyroptosis in the tumor microenvironment: Perspectives for immunotherapy of SCLC. *Semin Cancer Biol* (2022) 86(3):273–85. doi: 10.1016/j.semcancer.2022.03.009

40. Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. *Nat Immunol* (2013) 14(10):1014–22. doi: 10.1038/ni.2703

41. Batlle E, Massagué J. Transforming growth factor-β signaling in immunity and cancer. *Immunity* (2019) 50(4):924–40. doi: 10.1016/j.immuni.2019.03.024

42. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* (2012) 12(4):252–64. doi: 10.1038/nrc3239

43. Di Giacomo AM, Calabrò L, Danielli R, Fonsatti E, Bertocci E, Pesce I, et al. Long-term survival and immunological parameters in metastatic melanoma patients who responded to ipilimumab 10 mg/kg within an expanded access

programme. *Cancer immunol immunother: CII* (2013) 62(6):1021–8. doi: 10.1007/s00262-013-1418-6

44. Li BL, Wan XP. Prognostic significance of immune landscape in tumour microenvironment of endometrial cancer. *J Cell Mol Med* (2020) 24(14):7767–77. doi: 10.1111/jcmm.15408

45. Fancello L, Gandini S, Pelicci PG, Mazzarella L. Tumor mutational burden quantification from targeted gene panels: major advancements and challenges. *J immunother Cancer* (2019) 7(1):183. doi: 10.1186/s40425-019-0647-4

46. Marabelle A, Fakih M, Lopez J, Shah M, Shapira-Frommer R, Nakagawa K, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol* (2020) 21(10):1353–65. doi: 10.1016/S1470-2045(20)30445-9

47. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harbor Perspect Biol* (2010) 2 (1):a001008. doi: 10.1101/cshperspect.a001008

48. Pfeifer GP, Besaratinia A. Mutational spectra of human cancer. *Hum Genet* (2009) 125(5-6):493–506. doi: 10.1007/s00439-009-0657-2

49. Yu C, Qi H, Zhang Y, Zhao W, Wu G. Elevated expression of gamma-glutamyl hydrolase is associated with poor prognosis and altered immune signature in uterine corpus endometrial carcinoma. *Front Genet* (2021) 12:764194. doi: 10.3389/fgene.2021.764194

50. Yang J, Li H, Hu S, Zhou Y. ACE2 correlated with immune infiltration serves as a prognostic biomarker in endometrial carcinoma and renal papillary cell carcinoma: Implication for COVID-19. *Aging* (2020) 12(8):6518–35. doi: 10.18632/aging.103100

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 23 January 2023
DOI 10.3389/fonc.2023.1100045

OPEN ACCESS

EDITED BY
Fu Wang,
Xi'an Jiaotong University, China

REVIEWED BY
Ning Wu,
Institute of Oceanology, Chinese Academy of Sciences (CAS), China
Yutong Zhao,
The Ohio State University, United States
Zhe Liu,
Tianjin Medical University, China

*CORRESPONDENCE
Wen Ning
ningwen108@nankai.edu.cn
Lian Li
lilian523@nankai.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 16 November 2022
ACCEPTED 04 January 2023
PUBLISHED 23 January 2023

CITATION
Zhao C, Chen Z, Zhu L, Miao Y, Guo J, Yuan Z, Wang P, Li L and Ning W (2023) The BMP inhibitor follistatin-like 1 (FSTL1) suppresses cervical carcinogenesis. *Front. Oncol.* 13:1100045. doi: 10.3389/fonc.2023.1100045

The BMP inhibitor follistatin-like 1 (FSTL1) suppresses cervical carcinogenesis

Chenjing Zhao[1†], Zhongjie Chen[2†], Li Zhu[2], Yunheng Miao[3], Jiasen Guo[1], Zhiyong Yuan[2], Ping Wang[2], Lian Li[1*] and Wen Ning[1*]

[1]State Key Laboratory of Medical Chemical Biology, Tianjin Key Laboratory of Protein Sciences, College of Life Sciences, Nankai University, Tianjin, China, [2]Department of Radiation Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin's Clinical Research Center for Cancer, Tianjin, China, [3]Institute of Entomology, College of Life Sciences, Nankai University, Tianjin, China

Follistatin-like 1 (FSTL1) is a cancer-related matricellular secretory protein with contradictory organ-specific roles. Its contribution to the pathogenesis of cervical carcinoma is still not clear. Meanwhile, it is necessary to identify novel candidate genes to understand cervical carcinoma's pathogenesis further and find potential therapeutic targets. We collected cervical carcinoma samples and matched adjacent tissues from patients with the locally-advanced disease and used cervical carcinoma cell lines HeLa and C33A to evaluate the effects of FSTL1 on CC cells. The mRNA transcription and protein expression of FSTL1 in cervical carcinoma tumor biopsy tissues were lower than those of matched adjacent tissues. Patients with a lower ratio of *FSTL1* mRNA between the tumor and its matched adjacent tissues showed a correlation with the advanced cervical carcinoma FIGO stages. High expression of FSTL1 markedly inhibited the proliferation, motility, and invasion of HeLa and C33A. Regarding mechanism, FSTL1 plays its role by negatively regulating the BMP4/Smad1/5/9 signaling. Our study has demonstrated the tumor suppressor effect of FSTL1, and these findings suggested a potential therapeutic target and biomarker for cervical carcinoma.

KEYWORDS

FSTL1, tumor suppressor, cervical cancer, FIGO stages, BMP4/Smad1/5/9 signaling.

Introduction

Cervical carcinoma (CC) is the fourth most pervasive female malignancy in the world, with over 500,000 diagnosed cases and over 300,000 deaths each year (1). During the past decade, the incidence of CC has effectively reduced profit from the introduction of organized screening programs and human papilloma virus (HPV) vaccination programs (2). However, about 90% of CC deaths occur in non-developed countries, where morbidity and disease-specific mortality continue to increase (1). China has the most significant number of CC patients, with about 110,000 new cases and 60,000 mortality in the single year of 2020, which

is equivalent to 18.2% of newly diagnosed CC cases and 17.3% of deaths worldwide (3). After chemoradiotherapy, patients with failing or recurring metastatic CC still suffer a poor prognosis, even incorporating the anti-VEGF medication bevacizumab and novel immunotherapeutic approaches (4). Therefore, it is essential to identify novel target genes for in-depth understanding the pathogenesis of CC and predicting the prognosis of CC.

Follistatin-like 1 (FSTL1), a matricellular protein which initially discovered as a TGF-β1-inducible protein (5), belongs to the Fst-SPARC family (6). As a protein widely present in mammalian tissues, FSTL1 plays significant roles in the extracellular matrix and regulates cellular proliferation, survival, differentiation, and migration associated with development and disease, including cardiovascular diseases, arthritis, and organ fibrosis (7). The carcinogenesis of FSTL1 (previously named TSC-36) was first discovered when researchers found that FSTL1 was reduced and even undetectable in various v-myc/v-ras-transformed cells and human cancer cells (8). Recently, more and more works have identified the potential of FSTL1 as a tumor suppressor because of its ability to negatively regulate the motility and invasion of ovarian (9), renal (10), lung (11), and nasopharyngeal cancer cells (12). However, controversial data have reported that FSTL1 is riched in astrocytic brain tumors with high expression (13) and enhances the metastasis of cancer cells *via* activating diverse signaling pathway in breast (integrin β3/Wnt) (14), esophageal (NFkB–BMP) (15), hepatocellular (TGF-β1) (16), gastric (AKT) (17), and colorectal cancers (FAK) (18).

FSTL1 plays a role in development and disease to a large extent by regulating the TGF-β/BMP4 signaling (8, 19). Our previous studies on lung development also showed that FSTL1 interferes with alveolar differentiation mediated by the BMP4-Smad1/5/8 signaling (20). BMP4 is associated with many aspects of carcinogenesis but has different effects on different cancer types (11, 21). Recently, researchers have reported that FSTL1 up-regulates the BMP4-Smad signaling in lung adenocarcinoma (11), while in glioblastoma, FSTL1 down-regulates the same signaling (21). Therefore, the effect and mechanism of FSTL1 in cancer progression remain to be explored to a large extent.

The clinical significance of FSTL1 in CC is rarely reported, and the signaling of FSTL1 driving cervical carcinogenesis is not elucidated. In this study, the FSTL1 expression was found to be reduced, whereas BMP4/Smad signaling was more activated in biopsies of CC tumors than in matched adjacent tissues. The low ratio of *FSTL1* mRNA expression between the tumor and its matched adjacent tissue was associated with the poor prognosis in CC. The characterization of the function of FSTL1 in cervical carcinogenesis was also carried out in cultured human CC cells (HeLa and C33A). Our data demonstrated the tumor suppressor effect of FSTL1, suggesting its potential role as a therapeutic target and a prognostic marker for CC.

Materials and methods

Subjects

We collected the samples of CC tumors and matched adjacent tissues (2 cm from the tumor) from 15 patients with locally-advanced disease from 2018 to 2021 at Tianjin Medical University Cancer Institute and Hospital (TMUCIH). The pathological diagnose of each patient was assigned using the established criterion (22). The levels of FSTL1 in patients' cervical tissues were detected separately using qRT-PCR (No.1-11), western blot (No.12-15), and immunohistochemistry (No.15). We also collected the blood samples of eight patients and eight normal control individuals with matched age, sex, and weight. ELISA was used to detect FSTL1 level of peripheral blood samples The clinical information of CC patients is presented in Table 1, and the characteristics of normal control individuals are summarized in Table 2.

This study followed the principles of the Declaration of Helsinki. Approval was authorized by the Ethics Committee of TMUCIH (approval number: Ek2018137; date of approval: 20 November 2018). All sample donors in the study confirmed and signed the informed consent to publish this article.

Methods of data analysis and manu-experiments

The detailed methods were described in Supplementary Materials. The judgment of significance of all experimental results followed by consistent standard: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$ and ns: $p > 0.05$. Data were expressed as mean ± SE.

Results

The expression of FSTL1 is lower in CC patients

To confirm the broad significance of FSTL1 in CC, we first analyzed *FSTL1* mRNA expression in a CESC cohort (CC, n = 306; adjacent, n = 3) that was collected from the TCGA database. We observed a 3.2-fold reduction in *FSTL1* mRNA level ($\Delta\log_2 = -1.67$) in CC tumors compared with adjacent tissues (Figure 1A). We also obtained tumors and the matched adjacent tissues from an independent cohort of patients with locally-advanced CC and measured similar declines in *FSTL1* expression. As shown in Figure 1B, among the 11 pairs of biopsies (Table 1, patient No. 1-11) examined, the *FSTL1* mRNA transcription level in each tumor was markedly lower than that in the matched adjacent tissue using qRT-PCR. The reduction of FSTL1 protein expression was further detected through densitometric analysis of western blot (Table 1, patient No. 12-15, Figure 1C) and immunohistochemistry staining (Table 1, patient No. 15, Figure 1D). These data indicated the reduced FSTL1 expression in CC. Unfortunately, the circulating levels of FSTL1 CC patients' serum were comparable to those of healthy controls (Tables 1, 2, patient No. 13-20, Figure 1E).

Low FSTL1 expression ratio in CC is connected to advanced FIGO stage and poor prognosis

We further investigated the association between decreased FSTL1 levels and poor prognosis of CC. In the CESC cohort, CC patients with low *FSTL1* mRNA expression tended to have a poor prognosis

TABLE 1 Clinical characteristics of CC patients.

Pat.	Age	Pathology	FIGO stage	Follow-up	Ratio (Tumor/Adjacent)
1	45	SCC	IB2	CR	-1.45
2	56	SCC	IIB	CR	-2.88
3	60	SCC	IIB	CR	-2.08
4*	67	SCC	IIA1	CR	-1.99
5	55	SCC	IIIC2	PR	-2.46
6	45	SCC	IIIC1	PR	-4.81
7	51	SCC	IIB	CR	-2.13
8	58	SCC	IIA1	CR	-1.93
9	48	SCC	IIIC1	PD	-2.35
10	29	SCCC	IIIC1	UT	-2.65
11	59	SCCC	IVB	UT	-3.52
12	69	SCC	IIIB	/	/
13	49	SCC	IIIB	/	/
14	55	SCC	IIIB	/	/
15	48	SCC	IIB	/	/
16	61	SCC	IIB	/	/
17	57	SCC	IIIB	/	/
18	67	SCC	IIB	/	/
19	41	SCC	IIIC1	/	/
20	63	SCC	IIB	/	/

Patient and treatment characteristics. The detailed descriptions of all abbreviations in the column "FIGO stage" are listed in (22). In the column "Follow-up", CR, Complete response; PR, Partial response; PD, Progressive disease; UT, Under-treatment. *Discontinued therapy due to myocardial infarction.

(FIGO stage: 0, Log_2FSTL1 = 9.11 ± 0.46; I/II, Log_2FSTL1 = 7.14 ± 0.09; III-IV, Log_2FSTL1 = 6.84 ± 0.15), but the difference was not significant (Figure 2A). Similarly, no prognostic significance of the *FSTL1* mRNA expression was observed in our patient cohort (Table 1, patient No. 1-11; Figure 2B). Besides, no significant difference between the expression level of FSTL1 and the survival probability of patients shared on the TCGA database (Figure S1). However, interestingly, when calculating the *FSTL1* mRNA expression ratio between the tumor and matched adjacent tissues (Table 1), the reduced ratio was positively correlated with FIGO stage (I-II, Δ= -2.08 ± 0.19; III-IV, Δ= -3.16 ± 0.46, Figure 2C), which suggested that CC patients with relatively lower *FSTL1* mRNA transcription might tend to have a poor prognosis, including pelvic or retroperitoneal lymph node metastasis. We further followed up with these 11 CC patients who received standard chemoradiotherapy. As expected, six FIGO stage I-II patients showed complete response (CR) and were in stable condition. In contrast, two FIGO stage III patients (No. 5, IIIC2r; No. 6, IIIC1r) showed partial response (PR) and one patient (No. 9) with IIIC1r developed progressive disease (PD) at a median of 22 months (range, 18-31 months) follow-up. More seriously, patient

TABLE 2 Characteristics of normal control individuals.

Normal control	Age (year)	Weight (kg)
1	49	66
2	55	61
3	48	65
4	61	63
5	57	58
6	67	75
7	41	78
8	63	60

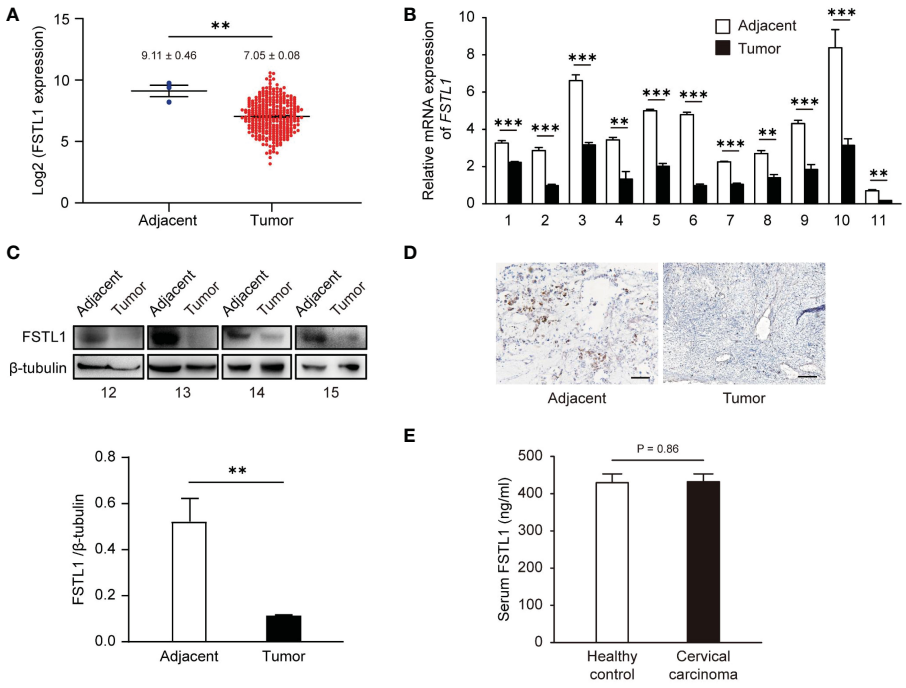


FIGURE 1
The level of FSTL1 frequently decreased in CC, which was related to the poor prognosis. **(A)** *FSTL1* mRNA transcription level of was analyzed using TCGA database containing 306 CC samples and 3 adjacent tissue samples. The mRNA **(B)** and protein **(C)** expression levels in CC tissues were lower than those in adjacent tissues. **(D)** The FSTL1 IHC staining in CC tissues was weaker than adjacent tissues. Scale bar, 150 μm. **(E)** ELISA showed no significant differences in serum FSTL1 levels between cervical cancer patients and healthy controls (n = 8). **p < 0.01. ***p < 0.001.

No.9 developed multiple distant metastases, including liver, bone and lymph nodes metastasis, and died eight months after standard chemoradiotherapy, bevacizumab and checkpoint inhibitors. The newly recruited patients with FIGO stage III-IV (No. 10, IIIC1; No. 11, IVB) are still under treatment. In conclusion, clinical data suggested that a low FSTL1 expression ratio could predict advanced CC stages to a certain extent.

FSTL1 inhibits CC cell proliferation

Before evaluating the regulatory effect of FSTL1 on CC cell proliferation, we first tested the levels of mRNA transcription and protein expression of FSTL1 in two CC cell lines, HeLa and C33A. And both levels in HeLa were significantly lower than those in C33A and the normal cervical epithelial cell line (H8) (Figures 3A, B). Overexpression of FSTL1 with transient transfection of pcFstl1 into HeLa cells significantly increased FSTL1 protein expression when compared with HeLa cells transfected with the empty vector (pcDNA3.1) (Figure 3C). Parallelly, the knockdown of FSTL1 by siRNA in C33A significantly decreased FSTL1 protein expression (Figure 3D). The overexpression of FSTL1 inhibited HeLa cell proliferation, as determined by cell number counting (Figure 3E), MTT assay (Figure 3F), and EdU staining (Figure 3G). Moreover, the deficiency of FSTL1 enhanced the proliferation of C33A (Figures 3E–G). In summary, the results indicated that FSTL1 might play the role as a tumor suppressor, and its high expression in CC cells can inhibit cell growth *in vitro*.

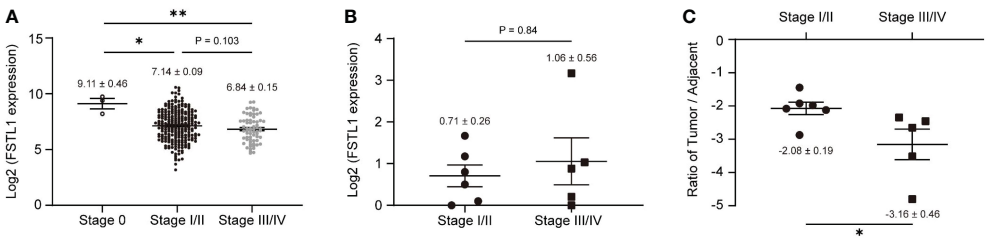


FIGURE 2
Low expression of FSTL1 in CC was connected to the poor prognosis. **(A)** Based on the cases from the TCGA database, *FSTL1* transcription decreased with the development of a poor prognosis. **(B)** *FSTL1* mRNA transcription of 11 CC tumor tissues showed no significant correlation between *FSTL1* and the FIGO stage. **(C)** With the development of a poor prognosis, the ratio of *FSTL1* mRNA transcription between the tumors and their matched adjacent tissues was reduced. *p < 0.05. **p < 0.01.

10.3389/fonc.2023.1100045

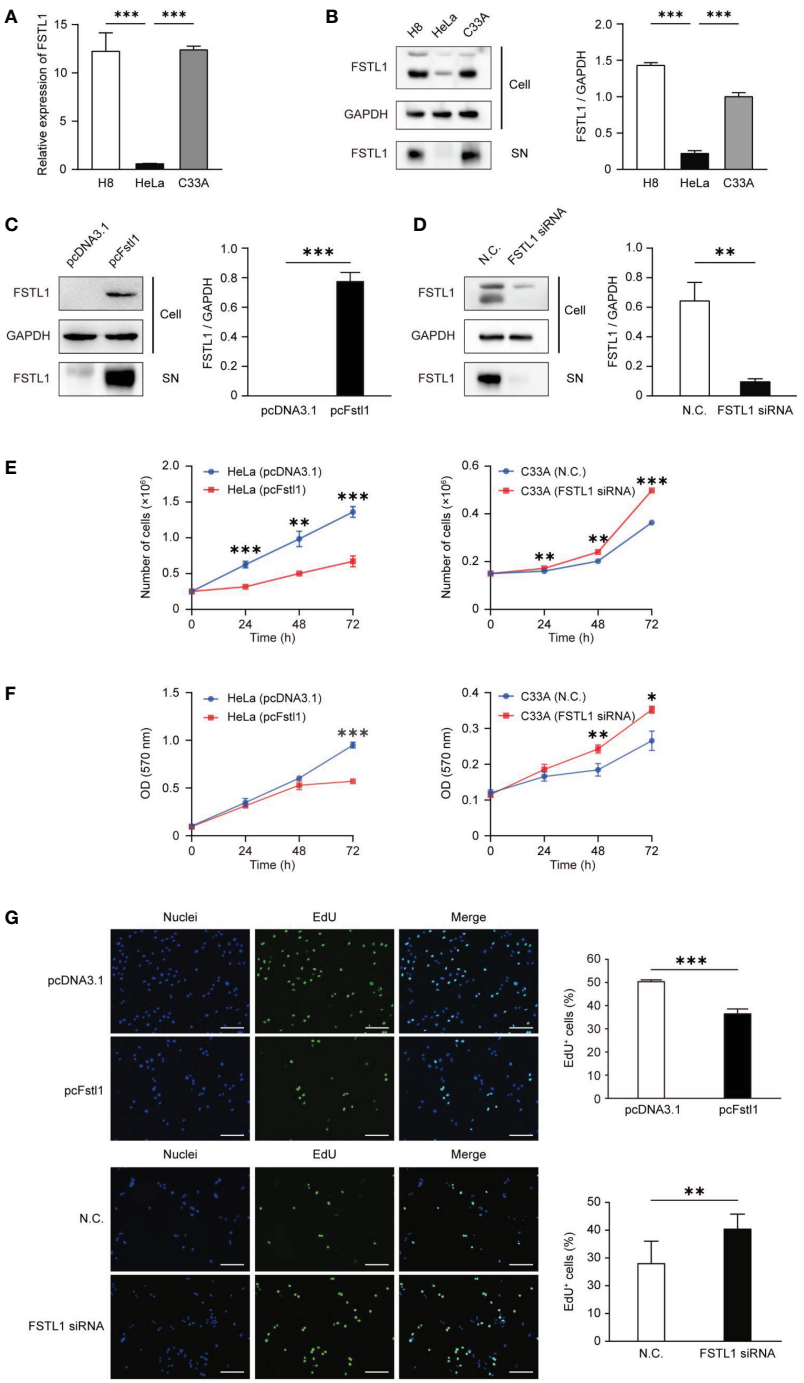


FIGURE 3
FSTL1 inhibited CC cell proliferation. **(A)** The comparison of *FSTL1* mRNA transcription level in cell lines HeLa, C33A and H8. **(B)** FSTL1 protein expression in cell extracts (labeled as "Cell") of HeLa, C33A and H8 cells and in the medium (labeled as "supernatant [SN]"). **(C)** FSTL1 protein expression in HeLa undergone the transfection of pcDNA3.1 or pcFstl1. **(D)** FSTL1 protein expression in C33A undergone the interference of siRNA. **(E)** The number of proliferating cells after overexpression of FSTL1 in HeLa (left) and knockdown in C33A (right). **(F)** The formazan production in HeLa and C33A cells for diverse time duration. **(G)** The nucleuses of HeLa and C33A cells were stained in blue by Hoechst, which represented the total number of cells. The cells in active proliferation were stained in green by EdU. Then the percentage of cell proliferation was calculated. Scale bar, 50 μm. *p < 0.05. **p < 0.01. ***p < 0.001.

FSTL1 has little effect on CC cell apoptosis

The impact of FSTL1 on the survival of HeLa and C33A cells was also tested. Overexpression of FSTL1 in HeLa or knockdown FSTL1 in C33A slightly changed the level of the cleaved form of Caspase-3, the marker of Caspase-3 activation in apoptotic signaling. Besides, the overexpression or deficiency of FSTL1 had little effect on the expression of Bcl-2, an anti-apoptotic protein that is often used as a marker showing apoptotic activity (Figure 4A). Consistently, FACS analysis also showed slightly changing but insignificant proportions of apoptotic cells in HeLa cells with FSTL1 overexpression and C33A cells with FSTL1 deficiency (Figure 4B). The data above suggested that FSTL1 slightly affected CC cells' apoptosis *in vitro*, but not significantly.

FSTL1 suppresses the motility and invasion of CC cell

The motility and invasion *in vitro* of CC cells with the overexpression and knockdown of FSTL1 were also detected. Compared with HeLa cells transfected with pcDNA3.1, the transwell migration assay showed significant reductions of migratory cells with the high expression level of FSTL1 (H8 and HeLa transfected with pcFstl1) was identified (Figure 5A). Consistently, the deficiency of FSTL1 in C33A cells resulted in a significant increase of migratory cells (Figure 5B). Moreover, the invasion of HeLa and C33A cells through Matrigel was significantly inhibited when FSTL1 was high-expressed (Figures 5C, D). Meanwhile, the overexpression of FSTL1 also caused a decrease in MMP2 expression which is related to tumor metastasis. And unsurprisingly, MMP2 expressed higher when FSTL1 was knocked down in C33A cells (Figure 5E). These data further demonstrated that FSTL1 is a tumor suppressor, and its high expression can significantly inhibit the motility and invasion of CC cells *in vitro*.

FSTL1 inhibits BMP4-Smad signaling in CC

The FSTL1-BMP4-Smad signaling has been reported in lung adenocarcinoma (11) and glioblastoma (21), but the role of FSTL1 in BMP4-Smad signaling remains controversial. To determine the molecular basis of the anti-tumor activity of FSTL1 in CC, we first examined the Smad-mediated BMP4 signaling. Compared with the matched adjacent tissues, the phosphorylation level of Smad1/5/9 in tumor biopsy tissues was higher in patients with the locally-advanced

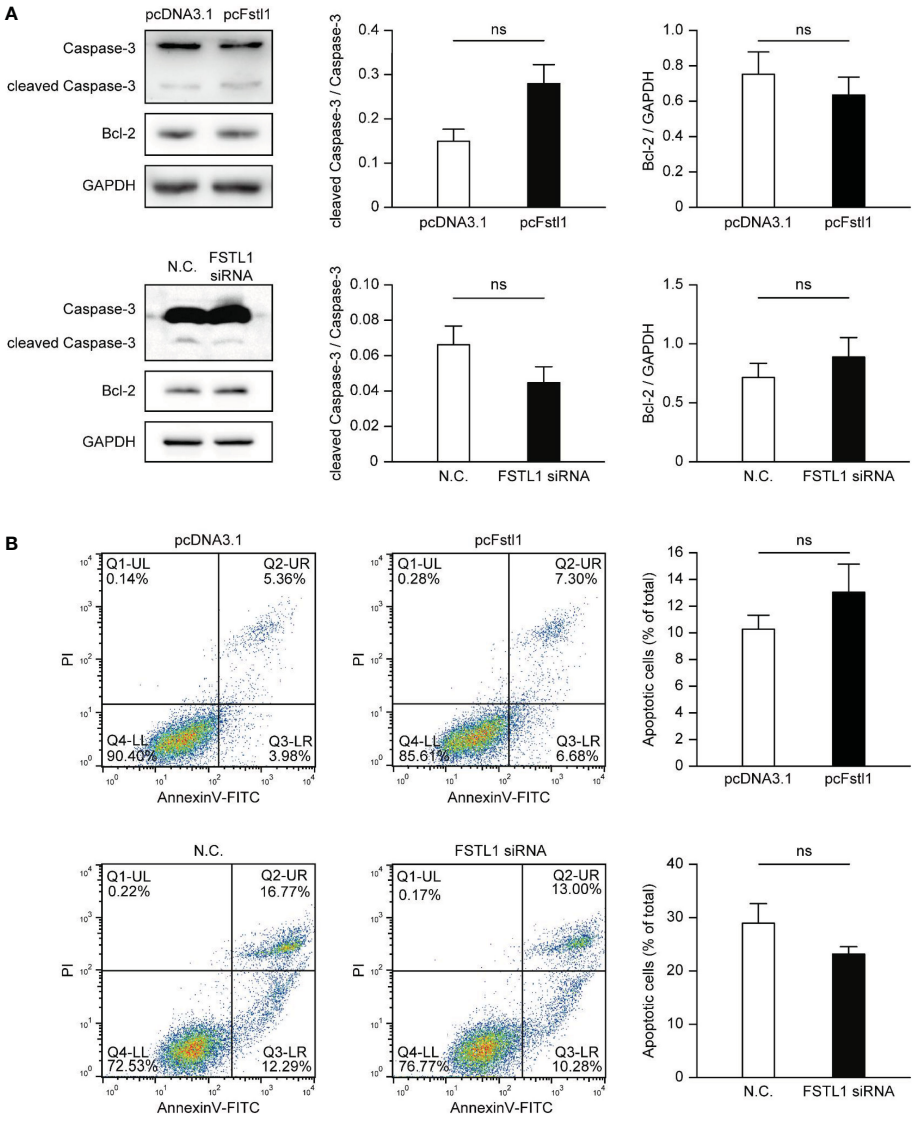


FIGURE 4
FSTL1 had little effect on the apoptosis of CC cells. **(A)** The expression of four proteins, including total Caspase-3, cleaved Caspase-3, Bcl-2, and GAPDH. GAPDH was selected as a loading control. **(B)** FACS showed HeLa cells with overexpression of FSTL1 and C33A cells with knockdown of FSTL1. The graph illustrates the induction of apoptosis in HeLa cells (up) and C33A cells (down). Percentages for each quadrant were pooled together and each column showed the average of three independent experiments. The Q4-LL represented the normal live cells. The Q1-UL represented necrotic cells. The Q2-UR represented the cells undergone late apoptosis. And the Q3-LR represented the cells undergone early apoptosis. ns: $p > 0.05$.

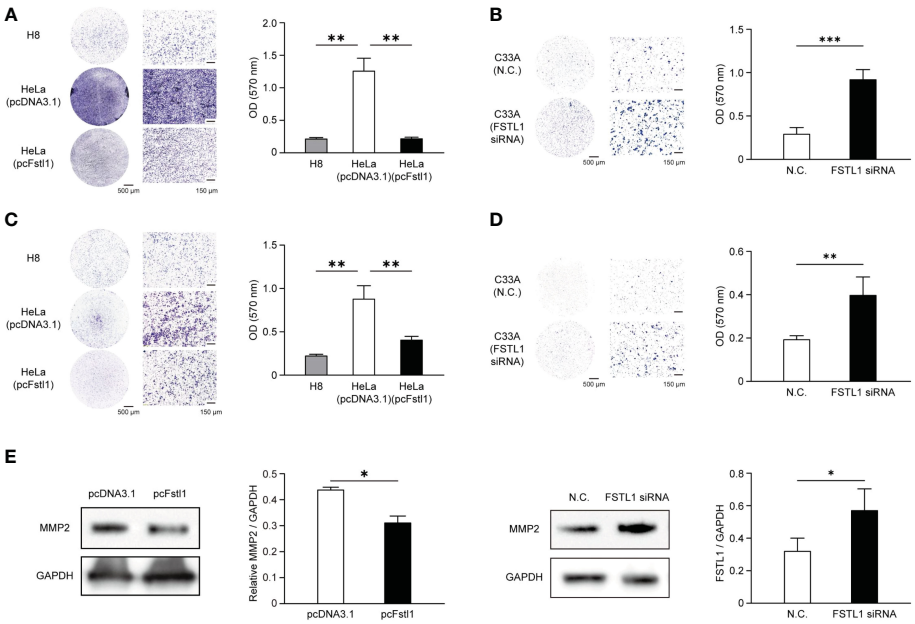


FIGURE 5
FSTL1 inhibited the mobility and invasion of CC cells. Representative images of crystal violet-stained HeLa migratory cells transfected with an empty plasmid or a pcFstl1 plasmid after the motility **(A)** and invasion **(C)** assay compared with H8 cells. The same experiments of cell mobility **(B)** and invasion **(D)** were also performed using C33A cells. Quantification of the migratory **(A, B)** and invading **(C, D)** cells by solubilizing the crystal violet and spectrophotometric reading at OD 570 nm. **(E)** The expression of the invasion-related protein MMP2 in HeLa cells transfected with either an empty control vector or pcFstl1 (left) and in C33A cells undergone the knockdown of FSTL1 (right). *p < 0.05. **p < 0.01 ***p < 0.001.

disease (Patient No. 12; Figure 6A). This corresponded to a lower level of FSTL1 protein in the tumor than in its matched adjacent tissues (Figures 1C and 6A). These findings implied that FSTL1 might function in the negative control of BMP4-Smad signaling in CC.

To further examine the inhibiting effect of FSTL1 on BMP4-Smad signaling in CC, we overexpressed FSTL1 in HeLa and knocked down FSTL1 in C33A. As shown in Figure 6B, BMP4-induced activation of Smad1/5/9 signaling was suppressed by the high expression of FSTL1 in CC cells. Moreover, FSTL1 high-expression inhibited the BMP4-induced CC cell proliferation as detected by EdU assay (Figure 6C) and CC cell metastases as detected by MMP2 expression (Figure 6D). Therefore, the results supported the involvement of the FSTL1-BMP4-Smad signaling in CC and demonstrated the negative effect of FSTL1 on regulating the BMP4-Smad signaling in CC progression.

Discussion

The worldwide gynecological malignancy, cervical carcinoma (CC), is a primary cause of female tumor-related deaths in non-developed countries (23). Despite advances in treatments, little progress has been made in treating patients with progressing CC, and the prognosis is poor. One of the hurdles to improving the effectiveness of treatment and developing precise treatment strategies is short of an in-depth study of the molecular mechanisms of cervical carcinogenesis. Here we provided new insights into the pathogenesis of CC and demonstrated the tumor suppressor effect of FSTL1 in CC. We analyzed the clinical samples as well as conducted *in vitro* experiments to validate that FSTL1 holds the potential to be a promising therapeutic target and possible biomarker for CC prognosis prediction.

We found evidence to prove the tumor suppressor function of FSTL1 in cervical carcinogenesis. FSTL1 expression decreased in the CC tumor tissues compared with its matched adjacent tissues. High expression of FSTL1 suppressed the proliferation, motility, and invasion, but affected little on HeLa and C33A cells' apoptosis. However, the same experiments demonstrated that a normal cervical epithelial cell line (like H8) could not be affected by changing the FSTL1 expression level (Figure S2). In summary, FSTL1 in CC shows a similar carcinogenesis suppressor function as in ovarian (9), renal (10), lung (11), and nasopharyngeal cancers (12).

Recent studies have reported that the tumor suppressor function of FSTL1 can further predict the prognosis of patients. For example, the IHC analysis and survival analysis of the public data both reveal a positive correlation between FSTL1 level and overall survival in lung adenocarcinoma patients (11). Liu et al. further found an SNP (rs1259293) in the genomic coding region of FSTL1, which is connected with a rising risk and poor postoperative prognosis of renal cell carcinoma (24). Here, we found that the decreased *FSTL1* mRNA expression ratio between the tumor and its matched adjacent tissues, instead of the expression of *FSTL1* mRNA itself, is correlated with the FIGO stage. Our data suggest a novel calculation method to highlighting the prognostic value of FSTL1 in CC.

The critical role of BMP4 in cancer pathogenesis has been reported (25). The expression level of BMP4 is usually varied in diverse types of tumors, and BMP4 inhibits cancer growth and metastasis in most types of tumors, although contradictory or

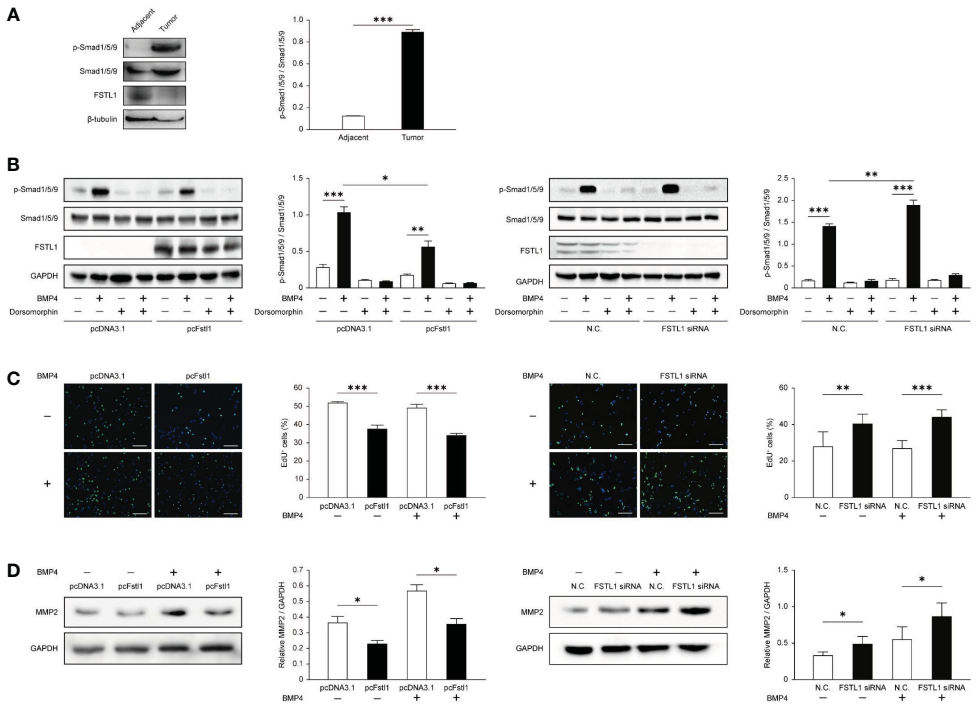


FIGURE 6
FSTL1 negatively regulated the BMP4/Smad1/5/9 signaling in CC. **(A)** p-Smad1/5/9, Smad1/5/9, FSTL1, and β-tubulin expression in CC tissues and matched adjacent tissues. **(B)** p-Smad1/5/9, Smad1/5/9, FSTL1, and GAPDH expression in HeLa transfected with pcDNA3.1 or pcFstl1 (left), and in C33A with knockdown of FSTL1 (right). Both cell lines were treated with BMP4 and/or dorsomorphin to prove that the BMP4-Smad signaling can be normally activated or blocked in the CC cells. **(C)** The proliferation of HeLa cells undergone 16 h BMP4 treatment after the Fstl1-transfected for 24 h (left) and of C33A cells undergone the same BMP4 treatment after FSTL1 knockdown for 48 h (right). The cell nucleuses were stained in blue by Hoechst, and the cells in active proliferation were stained in green by EdU. Then the proportion of cell proliferation was calculated. Scale bar, 50 μm. **(D)** The expression of protein MMP2 after 16 h BMP4 treatment in HeLa cells transfected with a plasmid pcFstl1 (left) and C33A undergone FSTL1 knockdown (right). *p < 0.05. **p < 0.01. ***p < 0.001.

conflicting results have been reported as well (21, 26). Jin and colleagues reported the high expression of FSTL1 in high-grade gliomas, and it facilitates glioma growth by negatively regulating the BMP4-Smad signaling (21). Chiou and colleagues showed low expression of FSTL1 and BMP4 in lung adenocarcinoma (11). They found that FSTL1 prevents the nicotine-induced proliferation of lung cancer cell lines. Different from the above studies, we observed low FSTL1 expression and high BMP4-Smad1/5/9 signaling activity in CC and found that FSTL1 high expression may attenuate the BMP4-promoted migration of CC cells. The precise mechanisms by which the FSTL1-BMP4-Smad axis plays its role in the pathogenesis of CC need further study.

Conclusions

In summary, our study has demonstrated that FSTL1 has a tumor suppressor effect in CC. The low expression of FSTL1 calculated based on the mRNA expression ratio between the tumor and its matched adjacent tissues can predict the poor prognosis of CC to a certain extent. High expression of FSTL1 suppressed the proliferation, motility, and invasion of CC cells *in vitro*. The mechanism of this action was through the negative control of the BMP4/Smad1/5/9 signaling. This study puts forward novel insights into the molecular mechanisms of FSTL1 in CC and suggests that FSTL1 is a potential therapeutic target and possible biomarker for CC.

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 the Ethics Committee of TMUCIH. 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

Conceptualization: WN. Methodology: CZ. Software: YM. Validation: JG. Formal analysis: CZ. Investigation: LZ, ZY and PW. Resources: ZC. Data curation: CZ and YM. Writing—original draft preparation: CZ, ZC, LL and WN. Writing—review and editing: CZ, ZC, LL and WN. Visualization: CZ. Supervision: WN and LL. Project administration: WN and LL. Funding acquisition: WN and LL. All

authors have read and agreed to the published version of 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 (NSFC) grants 82030001, 31471373, 31871460 (to WN), 82070077 (to LL), and the Fundamental Research Funds for the Central Universities of Nankai University 63171410 (to WN).

Conflict of interest

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

The reviewer ZL declared a shared parent affiliation with the authors ZC, LZ, ZY, PW to the handling editor at the time of review.

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1100045/full#supplementary-material

References

1. Arbyn M, Weiderpass E, Bruni L, de Sanjose S, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: A worldwide analysis. *Lancet Glob Health* (2020) 8(2):e191–203. doi: 10.1016/S2214-109X(19)30482-6

2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492

3. Bruni L, Albero G, Serrano B MM, Gómez D, Muñoz J, Bosch FX, et al. ICO/IARC information centre on HPV and cancer (HPV information centre). Human papillomavirus and related diseases in China. *Summary Rep* (2021) 22.

4. Kang Y, Huang J, Liu Y, Zhang N, Cheng Q, Zhang Y. Integrated analysis of immune infiltration features for cervical carcinoma and their associated immunotherapeutic responses. *Front Cell Dev Biol* (2021) 9:573497. doi: 10.3389/fcell.2021.573497

5. Shibanuma M, Mashimo J, Mita A, Kuroki T, Nose K. Cloning from a mouse osteoblastic cell line of a set of transforming-growth-factor-beta 1-regulated genes, one of which seems to encode a follistatin-related polypeptide. *Eur J Biochem* (1993) 217(1):13–9. doi: 10.1111/j.1432-1033.1993.tb18212.x

6. Hambrock HO, Kaufmann B, Muller S, Hanisch FG, Nose K, Paulsson M, et al. Structural characterization of TSC-36/Flik: Analysis of two charge isoforms. *J Biol Chem* (2004) 279(12):11727–35. doi: 10.1074/jbc.M309318200

7. Mattiotti A, Prakash S, Barnett P, van den Hoff MJB. Follistatin-like 1 in development and human diseases. *Cell Mol Life Sci* (2018) 75(13):2339–54. doi: 10.1007/s00018-018-2805-0

8. Mashimo J, Maniwa R, Sugino H, Nose K. Decrease in the expression of a novel TGF beta1-inducible and ras-recision gene, TSC-36, in human cancer cells. *Cancer Lett* (1997) 113(1-2):213–9. doi: 10.1016/S0304-3835(97)04700-9

9. Chan QK, Ngan HY, Ip PP, Liu VW, Xue WC, Cheung AN. Tumor suppressor effect of follistatin-like 1 in ovarian and endometrial carcinogenesis: A differential expression and functional analysis. *Carcinogenesis* (2009) 30(1):114–21. doi: 10.1093/carcin/bgn215

10. Tan X, Zhai Y, Chang W, Hou J, He S, Lin L, et al. Global analysis of metastasis-associated gene expression in primary cultures from clinical specimens of clear-cell renal-cell carcinoma. *Int J Cancer* (2008) 123(5):1080–8. doi: 10.1002/ijc.23637

11. Chiou J, Su CY, Jan YH, Yang CJ, Huang MS, Yu YL, et al. Decrease of FSTL1-BMP4-Smad signaling predicts poor prognosis in lung adenocarcinoma but not in squamous cell carcinoma. *Sci Rep* (2017) 7(1):9830. doi: 10.1038/s41598-017-10366-2

12. Wang H, Wu S, Huang S, Yin S, Zou G, Huang K, et al. Follistatin-like protein 1 contributes to dendritic cell and T-lymphocyte activation in nasopharyngeal carcinoma patients by altering nuclear factor kappab and jun n-terminal kinase expression. *Cell Biochem Funct* (2016) 34(8):554–62. doi: 10.1002/cbf.3227

13. Reddy SP, Britto R, Vinnakota K, Aparna H, Sreepathi HK, Thota B, et al. Novel glioblastoma markers with diagnostic and prognostic value identified through transcriptome analysis. *Clin Cancer Res* (2008) 14(10):2978–87. doi: 10.1158/1078-0432.CCR-07-4821

14. Cheng S, Huang Y, Lou C, He Y, Zhang Y, Zhang Q. FSTL1 enhances chemoresistance and maintains stemness in breast cancer cells *via* integrin beta3/Wnt signaling under miR-137 regulation. *Cancer Biol Ther* (2019) 20(3):328–37. doi: 10.1080/15384047.2018.1529101

15. Lau MC, Ng KY, Wong TL, Tong M, Lee TK, Ming XY, et al. FSTL1 promotes metastasis and chemoresistance in esophageal squamous cell carcinoma through NFkappaB-BMP signaling cross-talk. *Cancer Res* (2017) 77(21):5886–99. doi: 10.1158/0008-5472.CAN-17-1411

16. Loh JJ, Li TW, Zhou L, Wong TL, Liu X, Ma VWS, et al. FSTL1 secreted by activated fibroblasts promotes hepatocellular carcinoma metastasis and stemness. *Cancer Res* (2021) 81(22):5692–705. doi: 10.1158/0008-5472.CAN-20-4226

17. Wu M, Ding Y, Wu N, Jiang J, Huang Y, Zhang F, et al. FSTL1 promotes growth and metastasis in gastric cancer by activating AKT related pathway and predicts poor survival. *Am J Cancer Res* (2021) 11(3):712–28.

18. Gu C, Wang X, Long T, Wang X, Zhong Y, Ma Y, et al. FSTL1 interacts with VIM and promotes colorectal cancer metastasis *via* activating the focal adhesion signalling pathway. *Cell Death Dis* (2018) 9(6):654. doi: 10.1038/s41419-018-0695-6

19. Rider CC, Mulloy B. Bone morphogenetic protein and growth differentiation factor cytokine families and their protein antagonists. *Biochem J* (2010) 429(1):1–12. doi: 10.1042/BJ20100305

20. Geng Y, Dong Y, Yu M, Zhang L, Yan X, Sun J, et al. Follistatin-like 1 (Fstl1) is a bone morphogenetic protein (BMP) 4 signaling antagonist in controlling mouse lung development. *Proc Natl Acad Sci U S A* (2011) 108(17):7058–63. doi: 10.1073/pnas.1007293108

21. Jin X, Nie E, Zhou X, Zeng A, Yu T, Zhi T, et al. Fstl1 promotes glioma growth through the BMP4/Smad1/5/8 signaling pathway. *Cell Physiol Biochem* (2017) 44(4):1616–28. doi: 10.1159/000485759

22. Bhatla N, Berek JS, Cuello Fredes M, Denny LA, Grenman S, Karunaratne K, et al. Revised FIGO staging for carcinoma of the cervix uteri. *Int J Gynaecol Obstet* (2019) 145(1):129–35. doi: 10.1002/ijgo.12749

23. Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical cancer. *Lancet* (2019) 393(10167):169–82. doi: 10.1016/S0140-6736(18)32470-X

24. Liu Y, Han X, Yu Y, Ding Y, Ni C, Liu W, et al. A genetic polymorphism affects the risk and prognosis of renal cell carcinoma: association with follistatin-like protein 1 expression. *Sci Rep* (2016) 6:26689. doi: 10.1038/srep26689

25. Kallioniemi A. Bone morphogenetic protein 4-a fascinating regulator of cancer cell behavior. *Cancer Genet* (2012) 205(6):267–77. doi: 10.1016/j.cancergen.2012.05.009

26. Chen L, Yi X, Goswami S, Ahn YH, Roybal JD, Yang Y, et al. Growth and metastasis of lung adenocarcinoma is potentiated by BMP4-mediated immunosuppression. *Oncoimmunol* (2016) 5(11):e1234570. doi: 10.1080/2162402X.2016.1234570

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 31 January 2023
DOI 10.3389/fonc.2023.1104262

OPEN ACCESS

EDITED BY
Fu Wang,
Xi'an Jiaotong University, China

REVIEWED BY
Yang Xiao,
Southeast University, China
Liao Yuanmei,
Andalusian Center of Molecular Biology and Regenerative Medicine (CABIMER), Spain

*CORRESPONDENCE
Cuifang Han
hancuifang@gdmu.edu.cn
Hongbing Yu
hbyu@gdmu.edu.cn
Zhiwei He
hezhiwei@gdmu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Oncology, a section of the journal Frontiers in Oncology

RECEIVED 21 November 2022
ACCEPTED 02 January 2023
PUBLISHED 31 January 2023

CITATION
Han C, Chen J, Huang J, Zhu R, Zeng J, Yu H and He Z (2023) Single-cell transcriptome analysis reveals the metabolic changes and the prognostic value of malignant hepatocyte subpopulations and predict new therapeutic agents for hepatocellular carcinoma.
Front. Oncol. 13:1104262.
doi: 10.3389/fonc.2023.1104262

Single-cell transcriptome analysis reveals the metabolic changes and the prognostic value of malignant hepatocyte subpopulations and predict new therapeutic agents for hepatocellular carcinoma

Cuifang Han[1†*], Jiaru Chen[1,2†], Jing Huang[1], Riting Zhu[1,2], Jincheng Zeng[3], Hongbing Yu[1*] and Zhiwei He[1*]

[1]Guangdong Provincial Key Laboratory of Medical Molecular Diagnostics, The First Dongguan Affiliated Hospital, Guangdong Medical University, Dongguan, China, [2]School of Pharmacy, Guangdong Medical University, Dongguan, China, [3]Dongguan Key Laboratory of Medical Bioactive Molecular Developmental and Translational Research, Guangdong Provincial Key Laboratory of Medical Molecular Diagnostics, Guangdong Medical University, Dongguan, China

Background: The development of HCC is often associated with extensive metabolic disturbances. Single cell RNA sequencing (scRNA-seq) provides a better understanding of cellular behavior in the context of complex tumor microenvironments by analyzing individual cell populations.

Methods: The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) data was employed to investigate the metabolic pathways in HCC. Principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) analysis were applied to identify six cell subpopulations, namely, T/NK cells, hepatocytes, macrophages, endothelial cells, fibroblasts, and B cells. The gene set enrichment analysis (GSEA) was performed to explore the existence of pathway heterogeneity across different cell subpopulations. Univariate Cox analysis was used to screen genes differentially related to The Overall Survival in TCGA-LIHC patients based on scRNA-seq and bulk RNA-seq datasets, and LASSO analysis was used to select significant predictors for incorporation into multivariate Cox regression. Connectivity Map (CMap) was applied to analysis drug sensitivity of risk models and targeting of potential compounds in high risk groups.

Results: Analysis of TCGA-LIHC survival data revealed the molecular markers associated with HCC prognosis, including MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3, and CYP2C9. The RNA expression of 11 prognosis-related differentially expressed genes (DEGs) in normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 and HepG2 were

compared by qPCR. Higher KPNA2, LAGE3, SF3B4, CCT3 and GTPBP4 protein expression and lower CYP2C9 and PON1 protein expression in HCC tissues from Gene Expression Profiling Interactive Analysis (GEPIA) and Human Protein Atlas (HPA) databases. The results of target compound screening of risk model showed that mercaptopurine is a potential anti-HCC drug.

Conclusion: The prognostic genes associated with glucose and lipid metabolic changes in a hepatocyte subpopulation and comparison of liver malignancy cells to normal liver cells may provide insight into the metabolic characteristics of HCC and the potential prognostic biomarkers of tumor-related genes and contribute to developing new treatment strategies for individuals.

KEYWORDS

cancer metabolism, hepatocellular carcinoma, malignant hepatocytes, prognostic biomarker, single-cell RNA sequencing

1 Introduction

The mortality rate for liver cancer is the third highest among all cancers, and it is the sixth most frequent cancer overall (1). Hepatocellular carcinoma (HCC) is a tumour of hepatocellular origin. HCC is the predominant pathological type of primary liver cancer (PLC), as it represents 75-85% of all instances of PLC (2). A vast majority of HCCs are caused by chronic disease, and most of these cases reportedly evolve from chronic liver disease. This is primarily because of viral infections, including hepatitis B virus (HBV) and hepatitis C virus (HCV), and alcohol misuse (3). It is recommended that patients diagnosed with HCC in the early stages receive surgical resection, liver transplantation, and local resection (radiofrequency ablation) according to the Barcelona Clinic Liver Cancer (BCLC) staging system. Those in the intermediate stage are widely treated with trans-arterial chemoembolization (TACE), whereas systemic therapies are mainly considered for advanced-stage patients. Advanced-stage patients are often symptomatic, although they exhibit some degree of impaired liver function (4, 5). Notably, few or no treatments are available to improve survival rates for patients in advanced stages.

The development of treatment modalities for advanced HCC has dramatically expanded recently. To date, the FDA has approved several oral tyrosine kinase inhibitors (lenvatinib, regorafenib and cabozantinib), immune checkpoint inhibitors (nivolumab and pembrolizumab) and immunotherapies, such as monoclonal antibodies (6–8). These therapies have steadily improved the overall survival (OS) of HCC patients. However, the prognosis for HCC patients continues to be poor because of recurrence and elevated metastasis rates (9). HCC features have been attributed to a small subpopulation of tumour cells that carry more aggressive genetic or phenotypic alterations that allow them to escape conventional detection methods (10).

Although conventional bulk RNA sequencing (bulk RNA-seq) can provide sufficient gene expression profiles of large blocks of tissue, it does not effectively distinguish between different cell lineages and cellular interactions (11). Recently, the emergence of single-cell sequencing technology has bridged the gap between traditional high-throughput sequencing technologies and microarray data to provide genomic, transcriptomic, and epigenetic information from individual cells (12). Tumours consist of three major cell types, namely, malignant, immune and stromal cells, whose spatiotemporal interactions constitute a complex ecosystem (13). Unravelling the interactions between these types involves understanding tumour development and prognosis and therapeutic options. Since the advent of single-cell sequencing, various researchers have produced a relatively complete picture of human cell atlas, which has subsequently provided a great reference for understanding the complex composition of the organs of the body (14). Additionally, single-cell sequencing has been extensively employed to reveal the molecular mechanisms underlying HCC. For instance, studies have mapped the single-cell landscape of the early recurrent HCC ecosystem by relying on the high recurrence and low survival rates of HCC patients to advance the immunotherapy guidelines for HCC (13). Numerous studies have utilized single-cell sequencing techniques to elucidate the heterogeneity of malignant tumour cells, stromal cells, and immune cells. The large scale single-cell omics study targeting tumor-associated T cells published by Zhang et al. sketched the tumor immune landscape and laid the groundwork for a multifaceted understanding of T-cell characteristics associated with liver cancer (15). Single-cell technology can also identify rare subpopulations that were previously undetected by bulk RNA sequencing techniques, and these cell types are pivotal in determining tumor characteristics, including stemness-associated malignant cells and cancer-associated fibroblasts (16–18).

Abbreviations: HCC, hepatocellular carcinoma; PCA, Principal component analysis; UMAP, Uniform Manifold Approximation and Projection; KEGG, Kyoto Encyclopedia of Genes and Genomes; PON1, paraoxonase 1; TNM, tumor-node-metastasis; GSEA, gene set enrichment analysis; CNV, Copy number variation; DEDs, Differentially Expressed Genes; OS, Overall Survival; LASSO, least absolute shrinkage and selection operator; MARCKSL1, MARCKS Like 1; SPP1, Secreted Phosphoprotein 1; BSG, Basigin, also called CD147 or EMMPRIN; CCT3, chaperonin containing TCP1 subunit 3; LAGE3, L antigen family member 3; KPNA2, karyopherin subunit alpha 2; SF3B4, Splicing Factor 3b Subunit 4; GTPBP4, GTP Binding Protein 4; PON1, Paraoxonase 1; CFHR3, Complement factor H-related 3; CYP2C9, cytochrome P450 family 2 subfamily C member 9.

The reprogramming of energy metabolism characterizes tumour cells and causes rapid cell growth and proliferation. Thus, it is one of the hallmarks of cancer. Tumour cells actively take up glucose through the uncommon process of anaerobic glycolysis (Warburg effect). Studies have shown that this process provides energy to tumour cells, permitting intermediates to enter the anabolic bypass to maintain the *de novo* synthesis of nucleotides, lipids, and amino acids needed for cell proliferation (19). HCC is closely linked to metabolic abnormalities, as the liver is the primary metabolic organ. Most previous studies concerned with liver cancer have focused on sequencing at the tissue level to reveal the overall metabolic alterations. Single-cell sequencing technology can compensate for the shortcomings of bulk sequencing, thereby allowing one to pinpoint the cell groups most significantly associated with metabolic alterations from a large number of cell types. This also allows researchers to comprehensively describe the overall changes in gene expression patterns and reveal changes across specific cell groups. Therefore, scRNA-seq and bulk RNA-seq integration are important techniques for studying tumour development and heterogeneity. We analysed published single-cell transcriptome sequencing data to identify metabolically relevant HCC subpopulations, namely, hepatic epithelial cells. We then used the identified differentially expressed genes to designate a prognostic model for HCC patients.

2 Materials and methods

2.1 Data collection

The scRNA-seq data for HCC patients were acquired from GEO (https://www.ncbi.nlm.nih.gov/geo/, accession number GSE149614) and TCGA (https://portal.gdc.cancer.gov/) databases, respectively. TCGA-LIHC samples with complete clinical information were utilized as the model training set, and HCC samples from the GEO database (GSE76427) were utilized as the external validation set.

We first constructed a human liver cell atlas by performing cell classification and marker gene identification relying on Seurat. There were 17 samples in total from 10 HCC patients. These included 8 tumour samples (PT), 8 normal paraneoplastic samples (NTL), and one metastatic lymph node sample (MLN). The data for these samples were obtained from the GSE149614 project.

2.2 Identification of HCC cell subtypes

The scRNA-seq data were assessed by the Seurat package implemented in R software (4.1.1), with the exclusion of samples with more than 30% mitochondrial genes. The data were normalized using the Normalize Data function, and 2,000 genes with high intercellular coefficients of variation were subsequently extracted. Principal component analysis (PCA) was then performed, with 15 PCs selected for subsequent uniform manifold approximation and projection (UMAP) analysis. Cell types within the obtained clusters were annotated by the reported cell marker genes, and the expression matrix was generated for further analysis.

2.3 Analysis of intercellular communications

To investigate the potential interactions between tumor and paracancerous normal HCC samples, we employed the CellChat (1.5.0) package to analyse intercellular communication. We performed CellChat analysis of the annotated cellular gene expression profile data according to the official workflow. This package mimics intercellular communication by assessing the binding ligands and receptors along with their cofactors (20). Depending on receptor expression in one cell type and ligand expression in the other, enriched receptor–ligand interactions between the two cell types were inferred. Signaling pathways were visualized using the "netVisual_aggregate" function, where ligands were defined as efferent signals and receptors were defined as afferent signals.

2.4 Identification of important metabolic pathways at the single-cell level

Next, we employed the 'scMetabolism' package (0.2.1) to calculate the metabolic state between different cell types in the HCC dataset. This package combines published gene sets from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and the Reactome database to easily quantify single-cell metabolic activity. (21). Here, we used the authors' integrated list of metabolism-related gene sets from the Reactome database to explore metabolic pathway changes among six cell subpopulations and further looked at metabolic changes in epithelial cell subpopulations between tumor and paracancerous normal HCC samples.

2.5 Copy number variation analysis

To identify malignant cells in HCC patients, we compared patterns of chromosomal gene expression across cancer cells to those of their putative noncancerous counterparts using the infercnv package (version 1.12.0). First, we downloaded the human genome annotation file from the gencode database (https://www.gencodegenes.org/human/), converting it into a genomic location file. We used paracancerous epithelial tissue expression profiles from HCC patients as a reference group. Because our data were 10x scRNA-seq data, we set 0.1 as the cut-off value, and the denoise = T. Referring to the two indicators used by Itay Tirosh et al. to determine benign versus malignant cells, here we used the overall copy number variant (CNV) and the correlation with the average CNV of the top 5% of cells from the same tumor to estimate the malignancy or non-malignancy of the cells (22). The following correlation reference thresholds for determining the malignancy or not of cells were given: malignant cells: overall CNV > 0.2 & CNV correlation of the top 5% of tumors > 0.2; non-malignant cells: overall CNV < 0.2 & CNV correlation of the top 5% of tumors < 0.2.

2.6 Identification of significantly related pathways across different epithelial cell types

After scoring individual cells using a variety of enrichment methods, we derived multiple gene set enrichment score matrixes using the 'irGSEA' package (https://github.com/chuiqin/irGSEA/). Next, we calculated the differentially expressed gene sets for every single cell subpopulation within the enrichment score matrix for every gene set using the Wilcoxon test. Employing heat maps, certain specific enrichment pathways were labelled and then visualized.

2.7 Generation and validation of prognostic features

Univariate Cox analysis was used to screen genes associated with OS in TCGA-LIHC patients based on scRNA-seq and bulk RNA-seq datasets, and then, LASSO analysis was used to select significant predictors for incorporation into multivariate Cox regression. Next, we selected and used prognostic characteristics to generate polygenic risk scores and stratify TCGA-LIHC samples into either low- or high-risk groups. We also generated time-dependent receiver operating characteristic (ROC) curves to assess the predictive power of the prognostic features. The GSE76427 dataset was used to validate the prognostic value of the prognostic features. The entire analysis and visualization processes were performed by the survival, survminer, rms, and time ROC packages in R.

2.8 Gene expression of prognostic genes

Total RNA from cells was extracted with TRIzol reagent (Thermo Fisher Scientific, 15596026) following the manufacturer's instructions. Complementary DNA (cDNA) was synthesized and PCRs with cDNA as template were performed using a real-time detector (The Applied Biosystems QuantStudio 5 Real-Time PCR System) using Hieff qPCR SYBR Green Master Mix. The primer sequences are shown in Supplementary Table S1. Transcript levels were normalized against beta-actin levels as an internal reference and were evaluated using the 2- Δ ΔCt method. All experiments were repeated three times.

The Human Protein Atlas (HPA) tool was used to visually display the protein expression of prognostic genes in the form of immunohistochemical staining. The Gene Expression Profiling Interaction Analysis (GEPIA) database was applied to further demonstrate the credibility of the results.

2.9 Cell culture and western blot

An immortalized nontumorigenic normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 and HepG2 were purchased from the Fenghui Biotech Co., Ltd. (Hunan, China) with STR report. The MIHA cells were cultured in RPMI-1640 and HCC-LM3 and HepG2 were cultured in Dulbecco's modified Eagle medium (Gibco, Gaithersburg, MD, USA) with 10% fetal bovine serum (FBS, Sigma), 100 μg/mL penicillin and 100 μg/mL streptomycin (Solarbio, Shanghai, China) at 37°CC and 5% CO^2.

Total protein was extracted by using Takara kit. The Protein concentration was detected by BCA assay. The primary antibodies used in this study were anti-CYP2C9 (1:1000, Abcam), anti-PON1 (1:1000, Abcam) anti-beta-Actin (1:1000, Cell Signaling Technology).

2.10 Drug sensitivity analysis

Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/) is the largest pharmacogenomic database that is freely accessible for predicting responses to anticancer drugs. GDSC comprises 2 databases, namely, GDSC1, which contains 958 cell lines and 367 drugs, and GDSC2, which contains 805 cell lines and 198 drugs (23). To explore the differences in drug treatment effects among HCC patients, drug inferred sensitivity scores were assessed in GDSC2 by the 'oncoPredict' package.

2.11 Connective map analysis

The Connective Map (CMap) database stores a large-scale resource of expression profile data of cell lines under different drug treatments, which allows rapid targeting of drug candidates for the treatment of target diseases based on aberrant transcriptomic features in tumor cells (24). These drugs have an inverse relationship with tumor-promoting factors and may regulate aberrantly expressed genes in the opposite direction.

Recently, Yang et al. used the Library of Integrated Network-based Cellular Signatures (LINCS) database to demonstrate that using the eXtreme Sum (XSum) algorithm is most likely to yield optimal results in matching compounds and disease features, demonstrating better drug retrieval performance than the other five available methods, and obtaining practical targets with desirable results in liver cancer (25). In addition, the parameters for achieving the best prediction performance in this study were set at a number of disease molecular features of 100. Considering the significant difference in dimensionality between CMap data and LINCS, we incorporate more query signatures using top300 genes for XSum analysis for potential drug prediction.

2.12 Statistical analysis

All statistical analyses were carried out using packages implemented in R version 4.2.0 (https://www.r-project.org/). Student's t test was used to perform comparisons of continuous variables between two groups, and the Wilcoxon rank sum test was used to compare more than two groups. Kaplan–Meier curves with log-rank statistics were used to compare differences in OS between the two groups. Statistical significance was represented by $p < 0.05$.

3 Results

3.1 Single-cell gene expression profiles reveal six major cell types in the TME of primary HCC tumours

We performed descending and unsupervised cell clustering to recognize cell types based on their expression profiles. The raw dataset was read using the Seurat package. Then, an initial screening of genes and cells was performed using the following criteria: a gene had to be expressed in at least 3 cells, and at least 200 genes were measured in this cell. This was followed by further quality control to extract cells with >200 and <8000 expressed genes and <30% of mitochondrial genes. Next, the data were normalized to obtain 2000 highly variable genes for subsequent downscaling. Removal of the cell cycle effect resulted in an expression matrix comprising 58,475 cells and 24,746 genes. Next, we employed known marker genes to define broad cell categories and obtained the following six major cell subpopulations: T/NK cells, hepatocytes, macrophages, endothelial cells, fibroblasts, and B cells (Figures 1A, B). Cells from tumours and normal paracancerous tissues from different patients were classified into six categories (Figure 1C). Because proliferation is a hallmark of tumour cells, we employed the cell cycle scoring method to analyse the cell cycle. This image shows the results indicated that most of the cells were in the G1 phase, and a small number of cells were in the G2/M and S phases (Figure 1D).

3.2 Genes associated with the glucose and lipid metabolic pathway are upregulated in hepatocytes

To explore the existence of pathway heterogeneity across different cell subpopulations, we performed pathway activity and GSEA using signature genomes. Numerous pathways associated with cancer were upregulated in the hepatocyte subpopulation; these pathways included oxidative phosphorylation, glycolysis, and the metabolism of fatty acids, bile acids, and xenobiotics (Figure 2A). Next, we used the scMetabolism package to calculate scores for each metabolic pathway in each cell. We found that the epithelial cell subpopulation was enriched in most metabolic pathways, mainly those regulating pyruvate metabolism, the citrate tricarboxylic acid cycle, and the metabolism of triglycerides, pyruvates, lipids, carbohydrates, amino acids and their derivatives, ketone bodies, glucose, and fatty acids, and FoxO-mediated oxidative stress (Figure 2B). The genes of glucose metabolism and lipid metabolism pathways were also upregulated in epithelial cells (Figures 2C, D). To determine the differences in metabolic pathways of hepatic epithelial cells between tumor and paracancerous tissues, we extracted a separate subpopulation of hepatocytes and analysed the enrichment of metabolic pathways. Strikingly, we found an opposite trend between the glucose metabolism and lipid metabolism pathways in tumour and paracancerous cells (Figure 2E). Consequently, we subjected this cell subpopulation to more in-depth analysis.

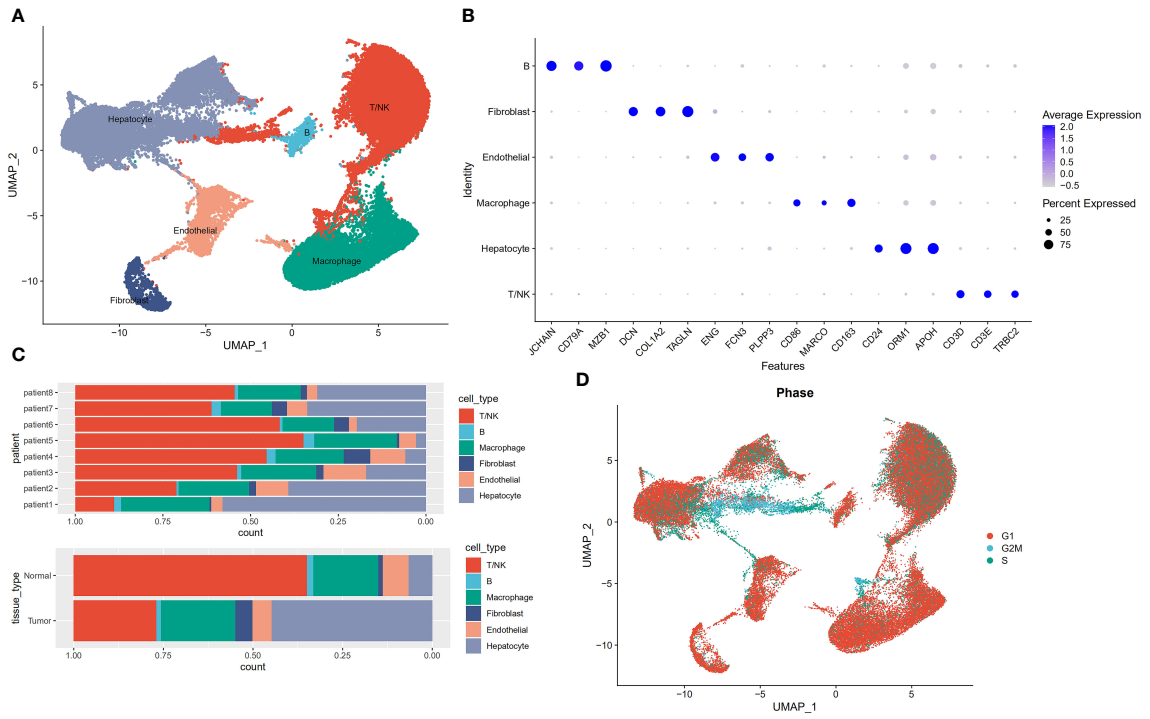


FIGURE 1
Profiles of single cells isolated from 8 primary liver cancer lesions with matching adjacent samples. **(A)** Uniform manifold approximation and projection (UMAP) plot of the analysed single cells. Each colour reflects one cell type. **(B)** Expression of marker genes for Hepatocytes, Macrophage cells, Endothelial cells, Fibroblasts, Mast cells, B cells, and T/NK cells. **(C)** Distribution of cells derived from either different patients or different sample origins. **(D)** UMAP clustering of 58,475 cells. Every colour represents a distinct cell cycle stage.

3.3 Pattern of intercellular communication between tumour and normal paracancerous tissues

We constructed a communication network between tumour samples and normal paracancerous samples to characterize alterations in signalling pathways (Figure 3A). A total of 642 and 499 significant ligand–receptor (LR) interactions were identified between the cell types present in tumour and normal paracancerous tissues, respectively (Supplementary Table S2). Differences between the number of communications among all cell populations between tumour and normal samples are illustrated in Figure 3B. In summary, tumour samples exhibited more cellular interactions than their normal counterparts, a phenomenon that was even more pronounced in the overall signalling pattern (Supplementary Figure S1). Next, we investigated the potential efferent and afferent signals among these six cell types and the specific molecular pairs. We found that the tumour samples consistently had more signal pairs than normal samples regardless of efferent or afferent signalling. The potential signalling pathways specific to tumour samples included SPP1, VTN, OCLN, CD46, GDF, EPHA, AGRN, PERIOSTIN, and HSPG. In normal samples, endothelial cells and T/NK cells were the main signalling providers and receptors, respectively, whereas in tumour samples, fibroblasts and macrophages represented the main signalling providers and receptors, respectively (Figures 3C, D). The overall communication probabilities of cells from tumour samples and normal sample sources were significantly different. Among the ligand receptors for intercellular communication in the normal sample sources, multiple pathways take part in inflammatory and immune responses, including pathways involving MHC-I, MHC-II, CXCL, complement, CCL, and TNF. In tumour samples, the intercellular interactions were mainly active in signalling pathways, including pathways involving SPP1, VTN, NOTCH, THY1, and CD46 (Figure 3E). To further elucidate the relationship between hepatocytes and other cell subpopulations, we generated a network plot of differences in number and strength. We found that hepatocytes had significantly higher interactions with endothelial cells and fibroblasts but a weaker association with immune cells (Figure 3F-G).

Differential analysis of all ligand–receptor pairs in hepatocytes and other cell types revealed significantly different patterns between tumour and adjacent normal tissues (Figures 3H–I). Studies have shown that CD74 promotes tumour cell growth by interacting with MIF (26). Remarkably, MIF-(CD74$^+$CD44) signalling between hepatocytes and T/NK and macrophages, which mediates immunosuppressive effects that have previously been illustrated for promoting cancer progression (27). Blocking MIF-CD74 signalling not only inhibits the proliferation of HCC cells but also exerts antitumour effects. Therefore, MIF/CD74 axis inhibition could be an effective treatment for HCC (28). SPP1 encodes osteopontin (OPN), a phosphorylated glycoprotein expressed in various tissues and cells associated with human diseases (29, 30). Notably, OPN is crucial in tumour progression, including HCC metastasis and prognosis, since it drives the evolutionary adaptation of tumour cells in the tumour microenvironment. Strikingly, SPP1-CD44 signaling was present between hepatocytes and T/NK cells,

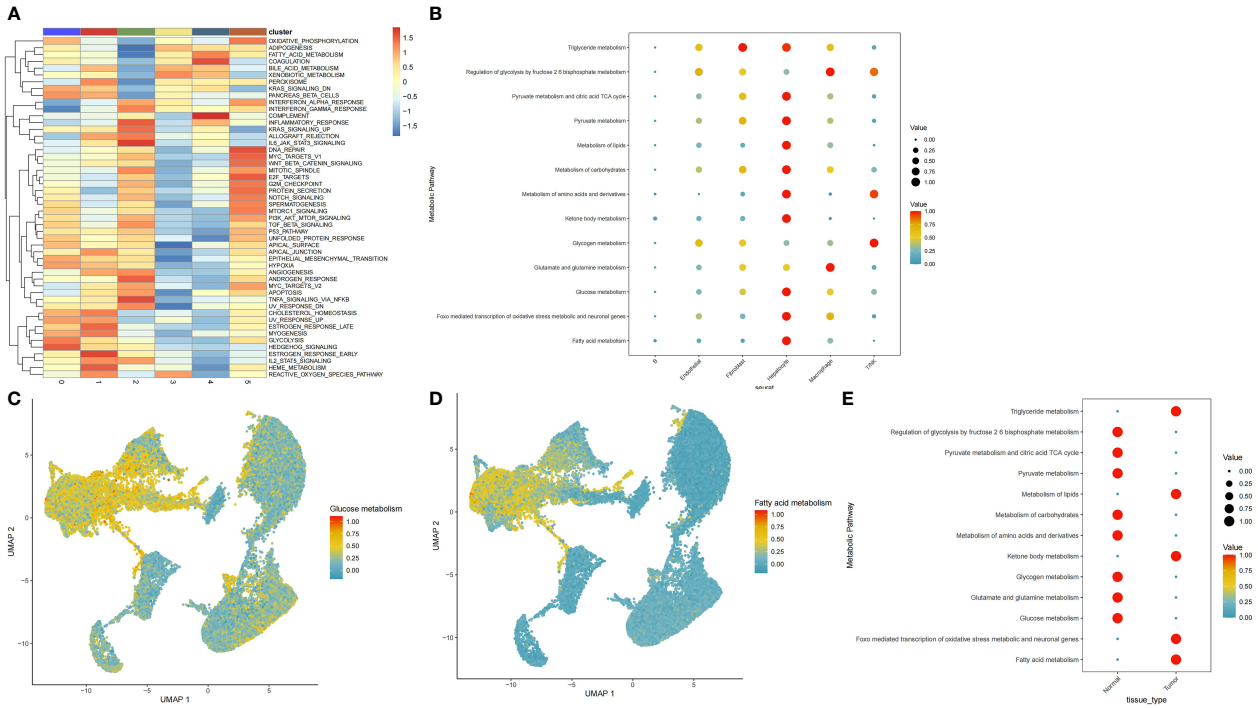


FIGURE 2
Distribution of glucose and lipid metabolic pathways in cellular subpopulations. **(A)** Functional annotation of six cellular subpopulations. **(B)** Dot plots show the specific metabolic pathways that were enriched in each cell subpopulation. **(C, D)** Scatter plots demonstrating highly expressed glucose and lipid metabolic pathways in hepatocytes cells. **(E)** Metabolic pathways comparison in hepatocytes cells from tumour versus paracancerous tissue.

macrophages, and fibroblasts in tumor samples, but not in normal samples adjacent to cancer, further supporting the critical role of SPP1 in the tumor ecosystem.

3.4 Transcriptome heterogeneity of hepatocytes in HCC

Despite previous batch effects, tumour cells continued to show patient-specific expression patterns. This suggests a high degree of heterogeneity, which could possibly be caused by CNVs. Six major cell subpopulations were identified after the entire malignant and normal hepatocytes reclustering (Figure 4A). In addition, UMAP plots revealed distinct clusters of malignant cells that corresponded to the sample origin (Figure 4B). Figure 4C illustrates the marker genes for each cell subpopulation. Next, the irGSEA package was employed to perform scRNA-seq gene set enrichment analysis and found that these subpopulations have unique activation signals. These signals include the Hedgehog signalling pathway (subpopulation 0), the early oestrogen response (subpopulation 1), the IL6/STAT3 and TNF signalling pathways (subpopulation 2), the xenobiotic metabolism and reactive oxygen species signalling pathways (subpopulation 3), and the KRAS signalling pathway (subpopulation 4). Moreover, multiple cell proliferation-related pathways were upregulated in subpopulation 5; these pathways included those involving the MYC targets V1 and V2, G2M checkpoints, E2F targets, WNT signalling, and P53 targets (Figure 4D). Activated KRAS is a major driver of cancer stem cell (CSC) proliferation and tumour metastasis (31). The results of the present study revealed that the KRAS signalling pathway was significantly upregulated in subpopulation 4, and the marker genes for CSCs were also distributed in this subpopulation (Figure 4E).

3.5 Profiles of chromosomal CNV in hepatocytes subpopulations

Next, we determined the chromosomal CNVs in each sample based on transcriptomic data to understand the malignancy of the epithelial cell subpopulation. This image shows the results revealed low and high CNV in adjacent normal epithelial cell subpopulations (control samples) and tumour epithelial cells, respectively. Chromosome amplification primarily occurs within chromosomes 1, 3, 5, 6, 7, 8, 12, 15, 17, 20, 21, and 22, with deletions most commonly observed in chromosomes 4, 9, 10, 11, 13, 14, 16 and 18 (Figure 5A). First, the copy value (CNV value) was calculated based on the sum of squares for all genes in each sample. Next, we ranked the CNV values of the tumour cells, using the top 5% as a reference, and then calculated correlation coefficients between other epithelial and reference cells. The determination of tumour cells was achieved at a threshold CNV >0.2 and a correlation coefficient >0.2. With CNV value as the horizontal coordinate and correlation coefficient as the vertical coordinate, black dots represent tumor cells and blue dots represent normal cells (Figure 5B). Finally, 13,502 tumor cells and 1,718 normal cells were identified and projected on the UMAP map (Figure 5C).

Thereafter, we employed the FindAllMarkers function and set the screening conditions logfc = 0.25 (difference multiplicity), min. pct =

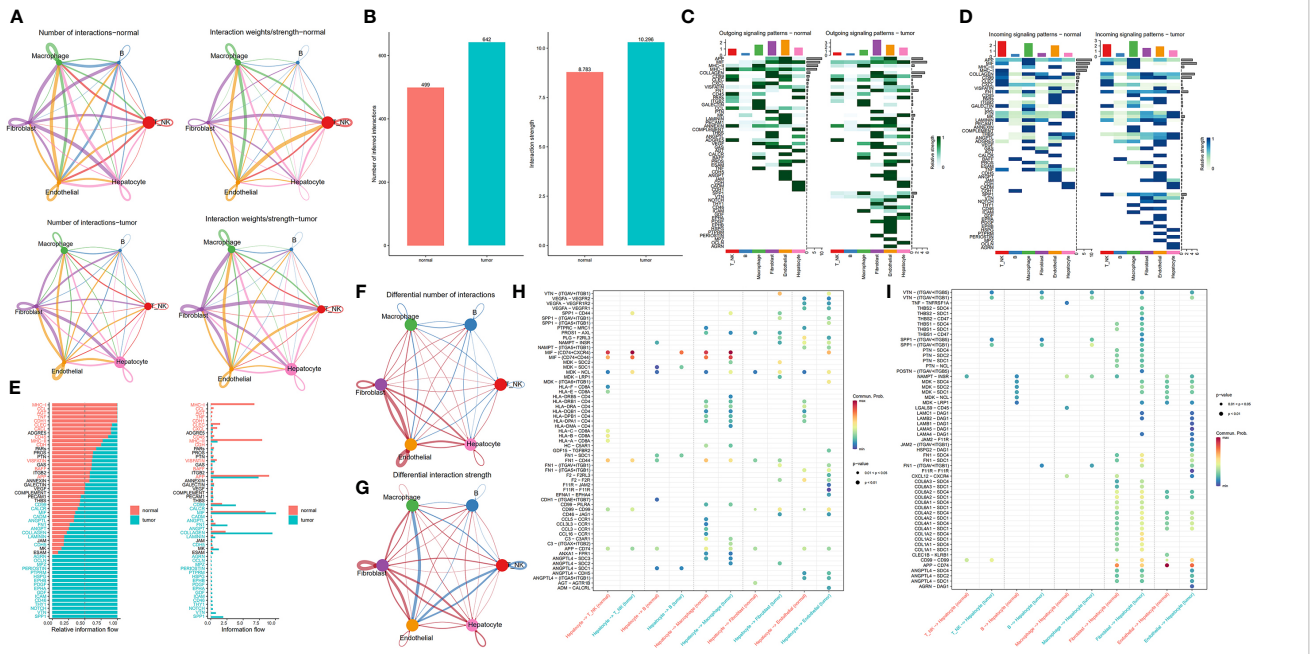


FIGURE 3
Comparison of cellular interactions between samples from tumour and adjacent normal tissues. **(A)** Cellular interaction number and strength. **(B)** Bar graph illustrating the total number (left) and weight (right) of ligand–receptor interactions between samples from tumour and adjacent normal tissues. **(C, D)** Heatmap showing possible afferent or efferent signalling pathways between cells. **(E)** Comparative profiles of pathway signal intensities indicating conserved and specific signalling pathways in tumour and normal tissue samples. **(F, G)** Communication quantity and intensity differences network. Red and blue colours represent upregulated and downregulated pathways, respectively, relative to normal tissues. **(H, I)** Dot plots show the variation in the signalling action of hepatocytes relative to other cell types.

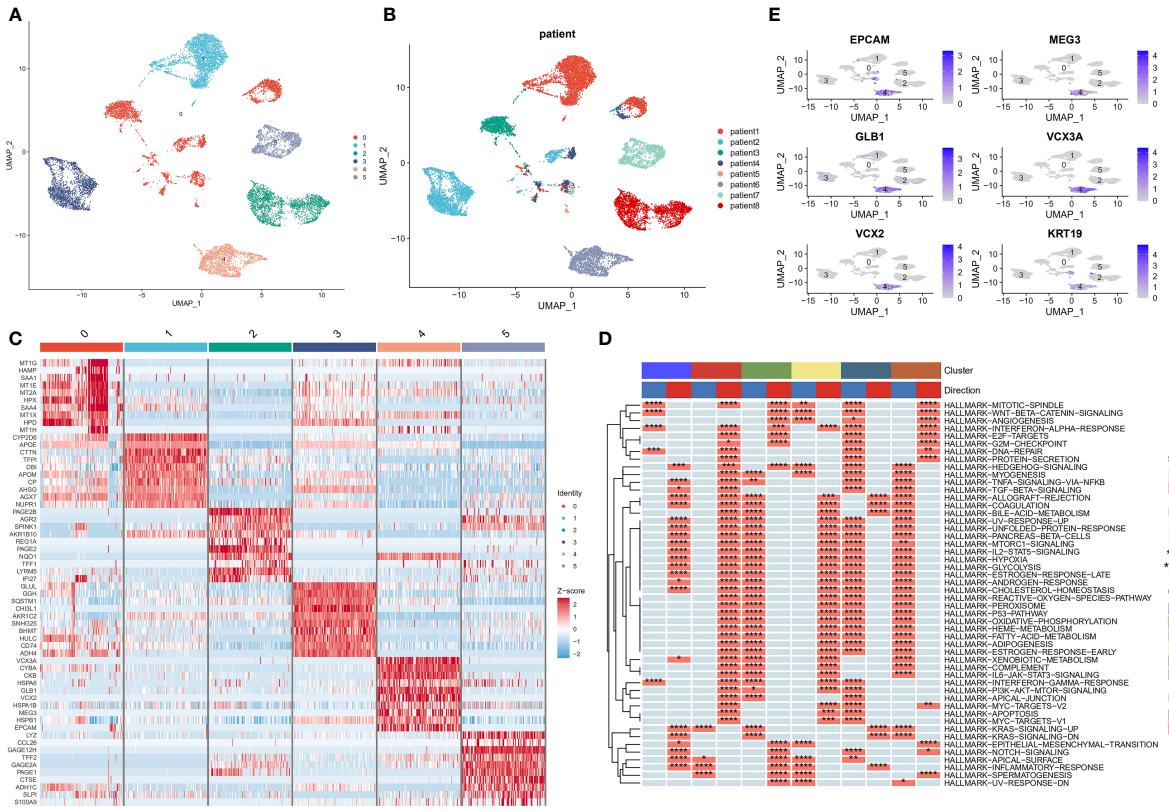


FIGURE 4
Transcriptome heterogeneity of malignant cells in HCC tissues. **(A)** UMAP plots of six different epithelial cell subpopulations. **(B)** UMAP plots demonstrating the heterogeneity among patients. **(C)** Heatmap of the top 10 differentially expressed genes(DEGs) across six epithelial cell clusters. **(D)** Single-cell pathway analysis of six subpopulations. **(E)** Scatter plot showing marker genes for cancer stem cells.

0.25 (minimum differential gene expression ratio) and pct. diff >0.1 (pct.1-pct.2) to identify marker genes in the hepatic malignant and normal epithelial cell subsets. The results revealed a total of 564 marker genes (Supplementary Table S3). We hypothesize that their function in HCC differs from that in normal epithelial cells, although further research exploration is needed.

3.6 Malignant hepatocyte subpopulations are associated with HCC prognosis

Next, we explored the prognostic role of hepatocyte subpopulations in HCC patients. Analysis of the mRNA expression data from HCC samples across the TCGA database yielded 2,900 differentially expressed genes (Figure 6A). Marker genes from malignant and nonmalignant cells of hepatocyte subpopulations intersected with DEGs related to HCC development in the TCGA database. Notably, 2,900 DEGs overlapped with 564 marker genes, resulting in 203 differentially expressed marker genes in HCC. These were subsequently named hepatocyte differential genes (HDGs) (Figure 6B). Univariate Cox regression analysis revealed 101 differentially expressed marker genes that were significantly related to the prognosis of HCC patients. To obtain a more robust prognostic profile, we employed the LASSO regression algorithm at 10-fold cross-validation with a lambda-min of 0.06321515 to designate a prognostic model consisting of 11 genes, namely, MARCKSL1 (MARCKS Like 1), SPP1 (Secreted Phosphoprotein 1), BSG (Basigin, also called CD147 or EMMPRIN), CCT3 (chaperonin containing TCP1 subunit 3), LAGE3 (L antigen family member 3), KPNA2 (karyopherin subunit alpha 2), SF3B4 (Splicing Factor 3b Subunit 4), GTPBP4 (GTP Binding Protein 4), PON1 (Paraoxonase 1), CFHR3 (Complement factor H-related 3) and CYP2C9 (cytochrome P450 family 2 subfamily C member 9) (Figures 6C, D).

Next, the median risk score was used to stratify the patients into high- and low-risk groups. Patients in the low-risk group showed significantly higher OS rates than their counterparts in the high-risk group ($p<0.001$) (Figure 6E). Application of the 11-gene signature in the validation cohort also indicated that patients in the low-risk group had longer OS rates than their counterparts in the high-risk group ($p<0.001$) (Figure 6F). To test the prognostic performance of the 11-gene signature, time-dependent ROC curves were generated targeting TCGA-LIHC samples. The results revealed area under the curve (AUC) values of 0.8, 0.7, and 0.7 for 1-, 3- and 5-year survival, respectively, in the testing cohort (Figure 6G) and 0.8, 0.8, and 0.87, respectively, in the validation cohort (GSE76427) (Figure 6H). These findings suggest that the 11-gene signature had good prognostic value in both cohorts. For the association analysis between the clinicopathological characteristics and the prognostic model, we analysed gender, TNM, stage, and risk scores in the TCGA-LIHC sample. The multivariate Cox regression analysis results revealed that the risk score was a significant independent prognostic factor for patients with LIHC ($p<0.001$) (Figure 6I). Moreover, we generated a

nomogram encompassing gender, stage, age, grade, risk score and 1-, 2- and 3-year survival. Next, we employed a one-sample GSEA approach to calculate scores for each sample across 175 pathways based on the risk score to identify relevant regulatory pathways. Thereafter, the correlation between each pathway and the risk score was no less than 0.3 for the evaluation. The results revealed 39 positive and 50 negative correlations with the sample risk score. Pathways that were positively correlated with the risk score included those related to cancer development, whereas the negatively correlated pathways included those regulating glycolysis/glycogenesis, glycine, and the metabolism of fatty acids, serine, threonine, glyoxylate and dicarboxylate (Figures 6K, L).

3.7 The relative RNA expression level and protein expression level of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9

Based on the initial trend of differentially up- and down-regulated genes (Supplementary Table S4), To further investigate the gene expression characteristics of 11 prognosis-related differentially expressed genes (MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9) in the high-risk and low-risk groups of HCC patients, we performed a correlation analysis between gene expression levels and risk scores. The results showed that all eight genes were positively correlated with risk scores, except for PON1, CFHR3 and CYP2C9, whose mRNA expression levels were significantly negatively correlated with risk scores (Supplementary Figure S3A). Meanwhile, in order to classify the high and low risk genes, we could see from the forest plot of 11 prognostic genes that the hazard ratio of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4 and GTPBP4 were all greater than 1, suggesting that these 8 genes might be poor prognostic factors and belong to high risk genes, while PON1, CFHR3 and CYP2C9 were all less than 1, suggesting that these three genes may be factors with a better prognosis (Supplementary Figure S3B).

The RNA expression of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9 in normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 (high metastatic HCC cells) and HepG2(low metastatic HCC cells) were compared by qPCR. It was found that CYP2C9, PON1 and CFHR3 were low expressed and MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4 were over expressed in human hepatoma cells compared with normal human hepatocyte cells (Unpaired t-test, $p<0.01$) (Figure 7A). Figure 7B shows the results of the protein expression levels of CYP2C9 and PON1 were down regulated in HepG2 and HCC-LM3 compared to MIHA. At the same time, CPTAC database analysis results showed that the protein expression of PON1, CFHR3 and CYP2C9 were low expressed in tumor tissues compared with paracancer normal tissue, while other genes were over expressed (Figure 7C).

Furthermore, immunohistochemical analysis from HPA database confirmed higher KPNA2, LAGE3, SF3B4, CCT3 and GTPBP4 protein expression and lower CYP2C9 and PON1 protein expression in HCC tissues (Figures 7D–J).

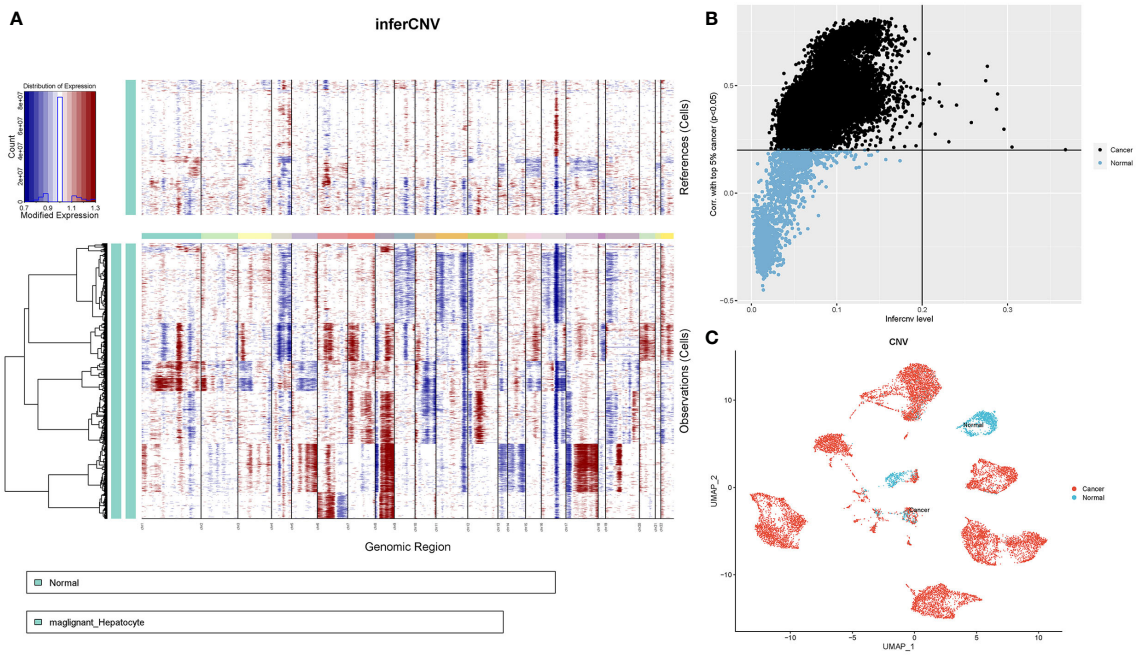


FIGURE 5
CNV analysis of HCC patient epithelial cells. **(A)** Heatmap showing CNV patterns in epithelial cells across 17 HCC samples. **(B)** Epithelial cells were classified as either malignant or nonmalignant. The horizontal coordinate represents the CNV value of the cell, whereas the vertical coordinate denotes the correlation coefficient of the top 5% of CNV values of tumour cells. **(C)** Distribution of tumour versus normal epithelial cells on the UMAP plot based on copy number variation.

3.8 Drug sensitivity analysis of risk models and targeting of potential compounds in high risk groups using connectivity map (CMap)

To determine the impact of risks on clinical practice, we evaluated the IC50 values of several chemotherapeutic agents in the high- and low-risk groups using the "oncoPredict" package. This analysis identified 123 drugs that were statistically significant ($p < 0.01$) (Supplementary Table S5). The results showed that afatinib, dasatinib, 5-fluorouracil, lapatinib, SCH772984, and cediranib had lower IC50 values in the high-risk group than in the low-risk group, suggesting that patients in the high-risk group may benefit more from these drugs. In contrast, JQ1, AT13148, axitinib, AZ960, AZD1208, and irinotecan had lower IC50 values in the low-risk group, suggesting that low-risk patients may benefit more from the above chemotherapeutic agents (Figure 8A).

While single-cell sequencing strategies are powerful tools for constructing disease signatures specific to individual cell types, CMap provides unprecedented convenience for researchers to tightly link the triad of drug, gene and disease in a context where deep understanding is lacking, as this method does not require the detailed mechanism of action or drug target to be provided in advance to predict therapeutic potential. Therefore, by combining a high-resolution single-cell sequencing strategy with CMap, we have been able to directly target effective therapeutic agents based on individual cell-level expression signatures and thus provide a more accurate prediction for screening potential drugs for disease.

We used a computational drug discovery strategy based on "signature reversion" (32) to identify drugs with a high risk of reversion using the large amount of data in the CMap database (Figure 8B). The top 300 genes with the highest fold change in the high- and low-risk groups were extracted for XSum analysis (Supplementary Table S6). The results of the CMap analysis revealed several compounds with gene expression patterns opposite to those specific to the high-risk group, with lower CMap scores indicating a higher perturbation ability. PHA.00816795, mercaptopurine, W.13, NU.1025 and arachidonyltrifluoromethane were the five potentially valuable small molecule drug candidates, as they were ranked as the top 5 candidates (Figure 8C). Among the top three of these candidates is mercaptopurine, which is a common chemotherapeutic drug that produces anticancer effects by interfering

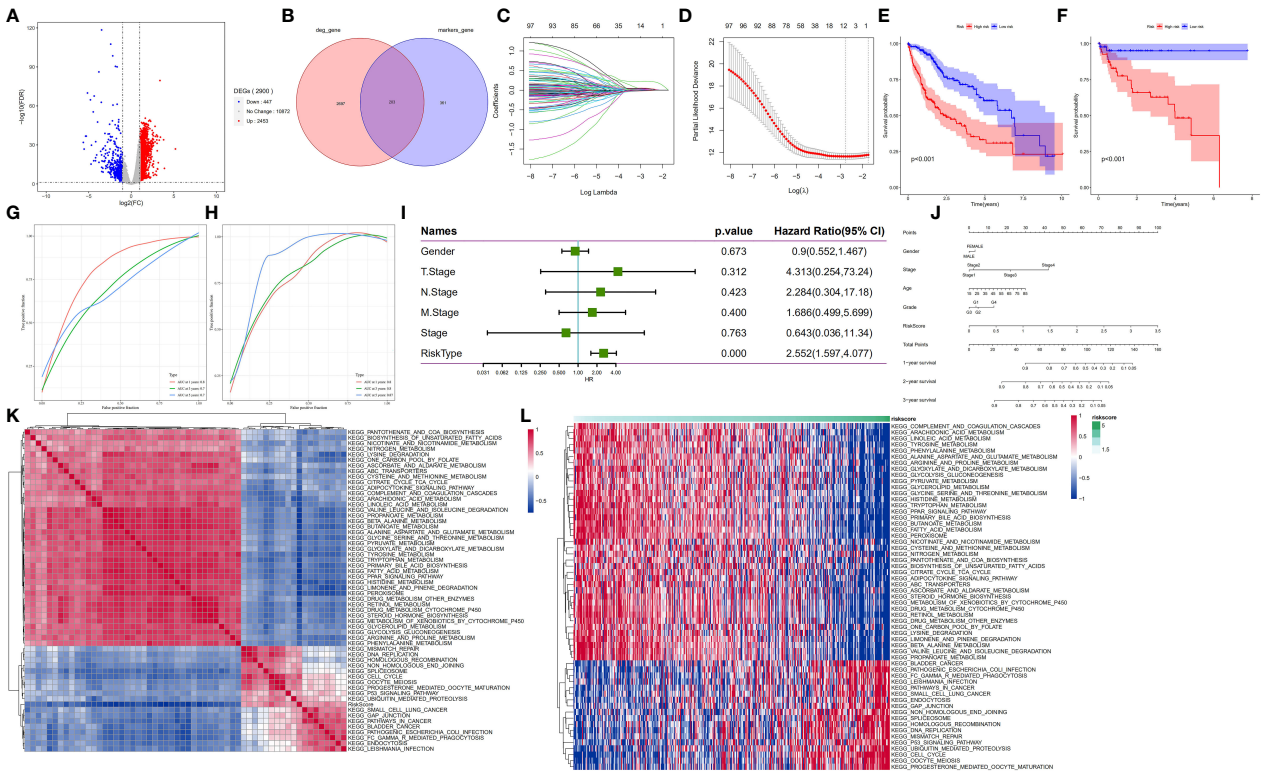


FIGURE 6
HDG identification and validation in the training (TCGA-LIHC) and validation cohorts (GSE76427). **(A)** The volcano plot of DEGs in the TCGA-LIHC dataset. **(B)** The intersection of DEGs of TCGA-LIHC cohort with marker genes of epithelial cell subpopulation of HCC. **(C, D)** Coefficient distribution plots of log(λ) sequences **(C)** and selection of optimal parameters (lambda) in the LASSO model **(D)**. **(E, F)** Kaplan–Meier survival curves illustrate the prognostic value of the 11-gene signature in the training cohort **(E)** and validation cohort **(F)**. **(G, H)** Distribution of the 11-gene signature risk scores and survival status of HCC patients in the training cohort **(G)** and validation cohort **(H)**. ROC curves showing the value of the 11-gene signature in predicting the OS rates of HCC patients at 1, 3, and 5 years in both cohorts. **(I)** Forest plot showing multivariate Cox analysis results. **(J)** Nomogram showing the prediction of OS at 1, 2, and 3 years. **(K, L)** Regulatory pathways potentially related to risk score.

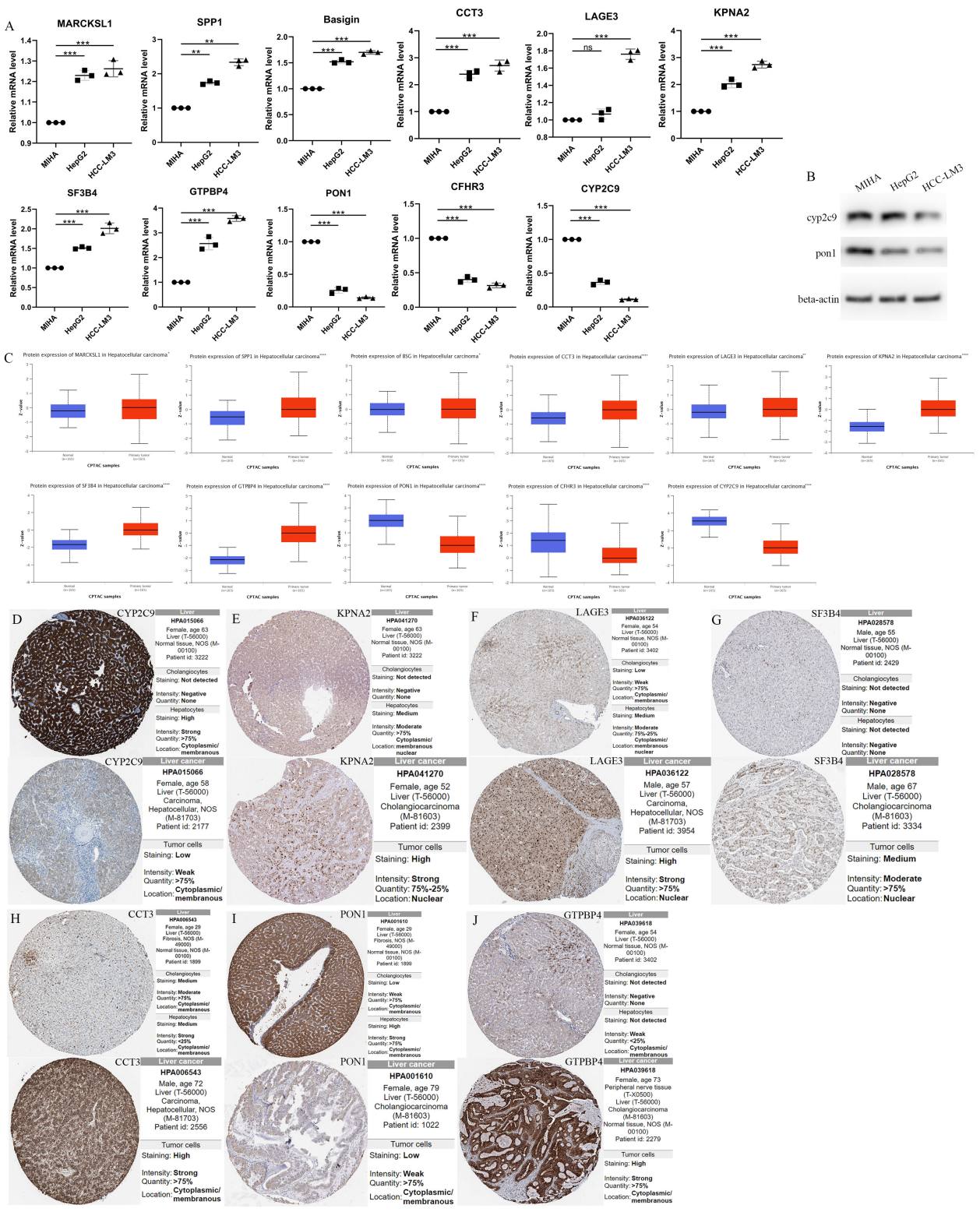


FIGURE 7
The Relative RNA Expression Level and Protein Expression Level of prognosis-related differentially expressed genes. **(A)** The Relative RNA Expression Level of MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3 and CYP2C9. **(B)** Expression of CYP2C9 and PON1 in normal human hepatocyte cell line MIHA and HCC cell lines HCC-LM3 and HepG2 through western blot analysis. **(C)** Box plots showed the differential protein expression of 11 hub genes in the CPTAC dataset in HCC tumor tissue and adjacent normal. **(D-J)** Immunohistochemical analysis of the CYP2C9, KPNA2, LAGE3, SF3B4, CCT3, PON1 and GTPBP4 in HCC and liver tissues from the HPA database. HCC, hepatocellular carcinoma; CPTAC, The National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium. HPA, Human Protein Atlas. (Unpaired t-test, **P < 0.01, ***P < 0.001, ****p < 0.0001 and ns, no significance).

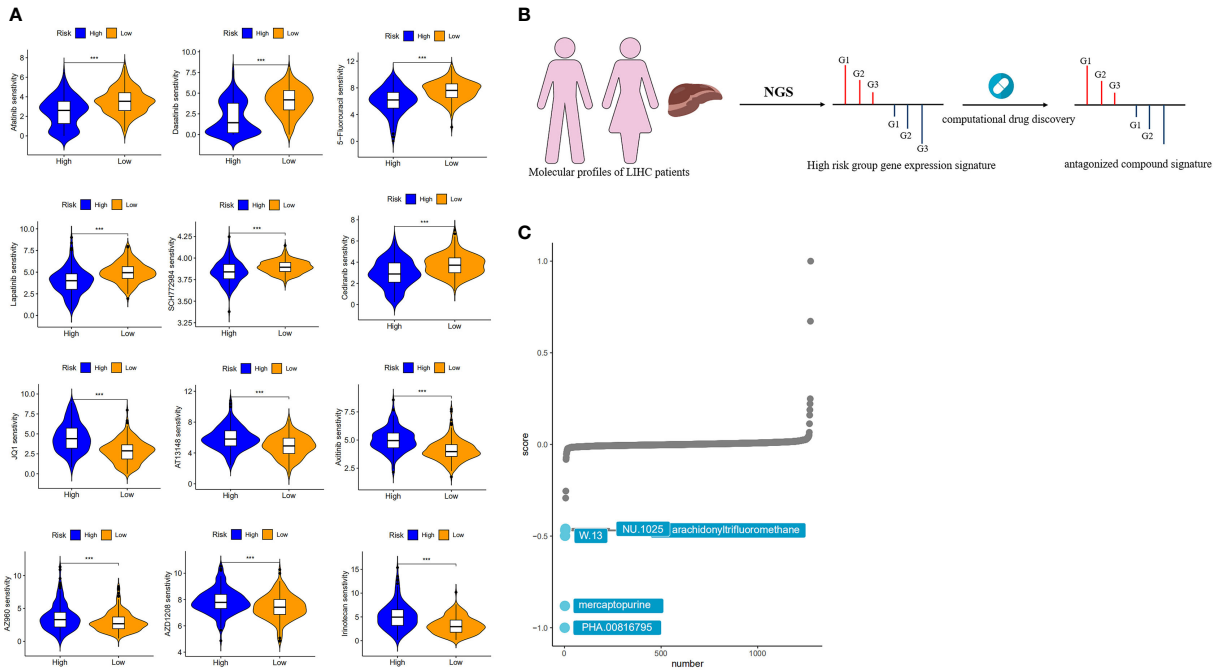


FIGURE 8
Drug sensitivity analysis and target compound screening for risk models. **(A)** Sensitivity analysis of chemotherapeutic agents between different risk groups. **(B)** How the "signature reversion"-based calculation method works. **(C)** The top 5 drugs with the lowest CMap scores.

with cell division or DNA synthesis (33). Yu et al. obtained five drugs associated with HCC by integrating multiple data to define the types of genes, considering the effect of genetic changes on HCC and the positive and negative relationships between drugs and HCC (34). Among these drugs, mercaptopurine is a potential anti-HCC drug.

4 Discussion

The liver is a major site for many metabolic processes, and metabolic dysregulation is vital for HCC progression and development (35). Evidence from numerous studies has shown that HCC originates from adult hepatocytes (36, 37). In this study, we found that HCC occurs in adult hepatocytes. Moreover, there were metabolic changes in the hepatic epithelial cells. While normal hepatocytes produce energy primarily through oxidative phosphorylation, malignant hepatocytes convert glucose into lactate through glycolysis to generate energy, a phenomenon known as the Warburg effect (38). The dysregulation of oxidative phosphorylation is related to elevated HCC tumorigenicity (39, 40). The liver synthesizes lactic acid and can store and breakdown lipids. Therefore, in HCC, aberrant lipid metabolism generates the lipids required for membrane formation and energy production, and posttranslational modifications support tumorigenesis (41). In our study, the gluconeogenic pathway (aerobic gluconeogenesis) was found to be enhanced in normal hepatic epithelial cells adjacent to cancerous cells, whereas the lipid metabolism pathway was enriched in malignant hepatocytes (Figure 2E).

HCC is a heterogeneous disease influenced by multiple factors, which makes it difficult to diagnose and perform individualized treatment. HCC patients are often diagnosed after curative surgical approaches are no longer possible because these patients are at an advanced stage of the disease. Traditional sequencing methods often mask the underlying heterogeneity in phenotypically defined cell subpopulations. In contrast, scRNA-seq allows the in-depth exploration of tumour heterogeneity and the analysis of tumour development, drug resistance, intercellular communication and immune infiltration patterns (12). Thus, this technique was employed to comprehensively analyse the HCC landscape at single-cell resolution.

To understand the interactions among hepatocytes and stromal cells and immune cells, we conducted intercellular communication analysis. This revealed enhanced interactions between hepatic epithelial cells and fibroblasts and reduced contact with immune cells, macrophages and endothelial cells in tumour samples compared to normal adjacent samples (Supplementary Figure S2). Cancer-associated fibroblasts (CAFs) are a major part of the tumour stroma and contribute to HCC progression. Furthermore, CAFs interact with tumour cells, immune cells, or vascular endothelial cells in the TME through direct intercellular contacts or indirect paracrine interactions to promote HCC (42–44). Similarly, Wang et al. performed a single-cell level analysis of samples from normal and malignant livers and found that in HCC, the most significant alteration was the expansion of ACTA2+ fibroblast populations and malignant cells. This suggests that the transition of hepatocytes from normal to malignant is accompanied by alterations in intercellular contact with other cells in the tumour microenvironment, which produce the complex intra- and intertumoral heterogeneity of HCC (45).

Differentially expressed genes between malignant and normal hepatocytes were identified by analysing copy number variations in single-cell transcriptome data and isolating malignant and nonmalignant cells from hepatocytes. In addition, analysis of

TCGA-LIHC survival data revealed the molecular markers associated with HCC prognosis,including MARCKSL1, SPP1, BSG, CCT3, LAGE3, KPNA2, SF3B4, GTPBP4, PON1, CFHR3, and CYP2C9. The identified prognostic risk factors showed good prediction performance in both HCC cohorts. Based on this, we also constructed a nomogram risk assessment model, which combines risk scores with clinical characteristics to facilitate the clinical application of HCC. It has been reported that CYP2C9 is involved in the metabolism of many carcinogens and drugs, and is down-regulated in HCC (46). Wang et al. used time serial transcriptome to reveal that *Cyp2c29* is a key gene in the development of hepatocellular carcinoma in the mouse model, and its overexpression enhances the production of 14,15-EET and inhibits inflammation induced hepatocellular proliferation by inhibiting the IKK-NF-κB pathway during liver injury (47). Meanwhile, the expression of the human homologous of *Cyp2c29* gene in mice was positively correlated with the survival time of HCC patients, further suggesting that CYP2C epoxygenases may be a potential therapeutic target for liver disease. Chen and others have revealed lncZic2/depletion/MARCKS/MARCKSL1 pathways can eliminate the liver tumor–initiating cells (TICs) (48). The overexpression of myristoylated alanine-rich protein kinase C substrate (MARCKS) and MARCKS like 1(MARCKSL1) can drive the self-renewal of TICs. Yang et al. demonstrated that BSG may be a tumor-promoting factor in HCC (49). The potential diagnostic role of BSG in differentiating HCC specimens from non-tumor specimens was demonstrated by analysis of multiple cohorts. BSG mRNA expression levels were significantly upregulated in both HCC specimens and HCC cell lines, and significantly shorter Overall Survival (OS) (P = 0.0014) and Disease Free Survival (DFS) (P = 0.0097) were observed in patients with high BSG expression relative to those with low BSG expression. Han et al. revealed that CCT3 is a new complementary biomarker for HCC screening and diagnosis (50). Several studies have shown that CCT3 is overexpressed in HCC patients by quantitative RT-PCR and western blotting. CCT3 can influence the progression of HCC by affecting phosphorylation signaling and translocation of STAT3/STAT3 into the nucleus of HCC cells (51, 52). The study of Li et al. showed that LAGE3 has prognostic value in HCC, which may affect the progression path of HCC tumor by promoting the proliferation, survival, migration, invasion and anti-apoptosis of HCC cells through the PI3K/AKT/mTOR and Ras/RAF/MAPK pathways (53). Guo et al. identified KPNA2 as a potential diagnostic and prognostic biomarker for HCC, which may affect HCC cell proliferation and migration by regulating cell cycle and DNA replication (54). Splicing factor 3b subunit 4 (SF3B4) has been revealed to be associated with the diagnosis and prognosis of HCC (55, 56). Liu et al. further demonstrated that SF3B4 drives cell proliferation and metastasis in HCC (57). Deng et al. further studied the mechanism and revealed the interaction between SF3B4 and ENAH in HCC, that is, SF3B4-regulated ENAH promotes the development of HCC by activating Notch signaling (58). It has been reported that Guanosine triphosphate binding protein 4 (GTPBP4) is associated with poor prognosis in HCC patients (59). Additional reports have explored the role of GTPBP4 in metabolic regulation and the potential mechanisms involved in HCC development and metastasis (60). GTPBP4 induces the dimer conformation of PKM2 through the SUMOylation to promote the aerobic glycolysis of HCC, thus promoting the progression and metastasis of HCC (61). Serum Paraoxonase 1 (PON1) has been reported as a biomarker for evaluating microvascular infiltration in hepatocellular carcinoma. Complement factor H related 3 (CFHR3) can be used to predict the prognosis of HCC. Overexpression of CFHR3 can affect the proliferation and apoptosis of hepatocellular carcinoma (62). Recent reports suggest that overexpression of CFHR3 may be a potential strategy for overcoming hypoxia and treating HCC (63). These studies confirm the significance and plausibility of these prognostic signatures.

Currently, liver transplantation and resection are efficient treatment options for early-stage disease; however, these treatments are appropriate for only 20-30% of HCC patients (64). Chemotherapy is another viable treatment option for advanced HCC. Recently, there has been significant progress in the development of molecularly targeted treatments for liver cancer (65). These include sorafenib, levatinib, and regorafenib, which have been approved as first- and second-line treatments for HCC. In this study, the sensitivities of HCC to various treatments were predicted. Low-risk patients showed higher sensitivity to afatinib, dasatinib, 5-fluorouracil, lapatinib, SCH772984, and cediranib than high-risk patients, which may be attributed to their higher metabolic activity. Various drugs were suggested for low-risk group patients, such as JQ1, AT13148, axitinib, AZ960, AZD1208, and irinotecan. Cancerous cells have the potential to evade the immune system (66). Immune escape can be achieved through a variety of mechanisms. Thus, therapeutic strategies that block checkpoint inhibitors of the PD-1/PD-L1 and CTLA-4 pathways can promote tumour-reactive T-cell aggregation, thereby improving the antitumour response (67, 68).

To the best of our knowledge, the 11-gene signature is the first to explore the overall molecular prognostic feature of subpopulations associated with metabolic disorders from single-cell sequencing data. This risk model exhibited excellent ability to predict the prognosis of HCC patients, and the AUC values at year 1, year 3 and year 5 were all greater than 0.7, with the optimal value of 0.8. Meanwhile, a novel XSum algorithm was used to predict potential drugs targeting high-risk groups from the Cmap database, and 5 drugs were finally obtained, including PHA.00816795, mercaptopurine, W.13, NU. 1025 and arachidonyl trifluoromethane. Small molecule drugs, serving as candidates, embrace potential value conducive to providing medication strategies for accurate treatment of HCC patients.

This research also has certain drawbacks. First, more perspective data with larger sample size should be collected to validate the accuracy of our established prognostic model. Second, the characteristics of different fractionated epithelial cells have not been generated and validated. Further in-depth analysis from specific epithelial cell subtypes closely related to metabolic changes will be conducive to obtaining more accurate and valid prognostic characteristics. Nevertheless, scRNA-seq analysis sheds new light on the metabolic characteristics of individual cell subsets in HCC, and anchors the survival and prognosis of relevant cell subsets with the most significant metabolic changes, which is a key step forward in clinical practice.

In conclusion, the present study identified prognostic genes significantly associated with metabolic changes in a hepatocyte

subpopulation at the single-cell level, and explored the heterogeneity of this subpopulation and its interrelationships with other cells in the tumor microenvironment. A prognostic model for OS prediction in HCC patients was established and validated and the results demonstrated favourable predictive ability. Additionally, differences in chemosensitivity between high-risk and low-risk groups were evaluated, and five potential drugs that might reverse the risk score were forecasted. These results provided an in-depth understanding of the metabolic characteristics of HCC. Furthermore, the characteristics of potential prognostic biomarker can be clarified through the comparison of tumor-related genes constructed by liver malignant cells and normal hepatocytes. The above may be conducive to new strategies of individualized therapy.

Data availability statement

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

Ethics statement

Ethical approval was not provided for this study on human participants because TCGA and GEO belong to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevai1t articles. Our study is based on open source data, so there are no ethical issues and other conflicts of interest. The Cancer Genome Atlas (TCGA), a landmark cancer genomics program, molecularly characterized over 20,000 primary cancer and matched normal samples spanning 33 cancer types. This joint effort between NCI and the National Human Genome Research Institut1te began in 2006, bringing together researchers from diverse disciplines and multiple institut1tions. Over the next dozen years, TCGA generated over 2.5 petabytes of genomic, epigenomic, transcriptomic, and proteomic data. The data, which has already led to improvements in our ability to diagnose, treat, and prevent cancer, will remain publicly available for anyone in the research community to use. (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

CH and JC contributed conception and design of the study; ZH and HY refined the research design idea. JC and JH collected the data; CH and JC performed the statistical analysis; JH and RZ accomplished the RT-qPCR assay and Western blot; CH and JZ further supplemented and validated the data and assisted with the interpretation of the results; CH and JC drafted the manuscript; ZH and HY revised the logic of the manuscript and polished the language. All authors contributed to manuscript and approved the submitted version.

Funding

The present study was supported by the National Natural Science Foundation of China (project no. 81902855), Natural Science Foundation of Guangdong Province (project no. 2020A1515010026), Funds for PhD Researchers of Guangdong Medical University in 2019 (project no. B2019019), and Discipline Construction Project of Guangdong Medical University (project no. 4SG22005G).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1104262/full#supplementary-material

References

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

2. Singal AG, Lampertico P, Nahon P. Epidemiology and surveillance for hepatocellular carcinoma: New trends. *J Hepatol* (2020) 72(2):250–61. doi: 10.1016/j.jhep.2019.08.025

3. Li X-Y, Shen Y, Zhang L, Guo X, Wu J. Understanding initiation and progression of hepatocellular carcinoma through single cell sequencing. *Biochim Biophys Acta (BBA)-Reviews Cancer* (2022) p:188720. doi: 10.1016/j.bbcan.2022.188720

4. Llovet J, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers* (2021) 7(1):6. doi: 10.1038/s41572-020-00240-3

5. Balogh J, Victor D 3rd, Asham EH, Burroughs SG, Boktour M, Saharia A, et al. Hepatocellular carcinoma: a review. *J hepatocellular carcinoma* (2016) 3:41. doi: 10.2147/JHC.S61146

6. Zhang T, Merle P, Wang H, Zhao H, Kudo M. Combination therapy for advanced hepatocellular carcinoma: do we see the light at the end of the tunnel? *Hepatobiliary Surg Nutr* (2021) 10(2):180. doi: 10.21037/hbsn-2021-7

7. Todesca P, Marzi L, Critelli RM, Cuffari B, Caporali C, Turco L, et al. Angiopoietin-2/Tie2 inhibition by regorafenib associates with striking response in a patient with aggressive hepatocellular carcinoma. *Hepatology* (2019) 70(2):745–7. doi: 10.1002/hep.30588

8. Gao X, Chen H, Huang X, Li H, Liu Z, Bo X. ARQ-197 enhances the antitumor effect of sorafenib in hepatocellular carcinoma cells via decelerating its intracellular clearance. *OncoTargets Ther* (2019) 12:1629. doi: 10.2147/OTT.S196713

9. Li D, Sedano S, Allen R, Gong J, Cho M, Sharma S, et al. Current treatment landscape for advanced hepatocellular carcinoma: patient outcomes and the impact on quality of life. *Cancers* (2019) 11(6):841. doi: 10.3390/cancers11060841

10. Vande Lune P, Abdel Aal AK, Klimkowski S, Zarzour JG, Gunn AJ. Hepatocellular carcinoma: Diagnosis, treatment algorithms, and imaging appearance after transarterial chemoembolization. *J Clin Transl Hepatol* (2018) 6(2):175–88. doi: 10.14218/JCTH.2017.00045

11. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* (2019) 37(7):773–82. doi: 10.1038/s41587-019-0114-2

12. González-Silva L, Quevedo L, Varela I. Tumor functional heterogeneity unraveled by scRNA-seq technologies. *Trends Cancer* (2020) 6(1):13–9. doi: 10.1016/j.trecan.2019.11.010

13. Sun Y, Wu L, Zhong Y, Zhou K, Hou Y, Wang Z, et al. Single-cell landscape of the ecosystem in early-relapse hepatocellular carcinoma. *Cell* (2021) 184(2):404–421.e16. doi: 10.1016/j.cell.2020.11.041

14. Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, et al. Construction of a human cell landscape at single-cell level. *Nature* (2020) 581(7808):303–9. doi: 10.1038/s41586-020-2157-4

15. Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, et al. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* (2017) 169(7):1342–1356.e16. doi: 10.1016/j.cell.2017.05.035

16. Zheng H, Pomyen Y, Hernandez MO, Li C, Livak F, Tang W, et al. Single-cell analysis reveals cancer stem cell heterogeneity in hepatocellular carcinoma. *Hepatology* (2018) 68(1):127–40. doi: 10.1002/hep.29778

17. Ho D, Tsui YM, Sze KM, Chan LK, Cheung TT, Lee E, et al. Single-cell transcriptomics reveals the landscape of intra-tumoral heterogeneity and stemness-related subpopulations in liver cancer. *Cancer Lett* (2019) 459:176–85. doi: 10.1016/j.canlet.2019.06.002

18. Yu L, Shen N, Shi Y, Shi X, Fu X, Li S, et al. Characterization of cancer-related fibroblasts (CAF) in hepatocellular carcinoma and construction of CAF-based risk signature based on single-cell RNA-seq and bulk RNA-seq data. *Front Immunol* (2022) 13. doi: 10.3389/fimmu.2022.1009789

19. Satriano L, Lewinska M, Rodrigues PM, Banales JM, Andersen JB. Metabolic rearrangements in primary liver cancers: cause and consequences. *Nat Rev Gastroenterol Hepatol* (2019) 16(12):748–66. doi: 10.1038/s41575-019-0217-8

20. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. *Nat Commun* (2021) 12(1):1–20. doi: 10.1038/s41467-021-21246-9

21. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discovery* (2022) 12(1):134–53. doi: 10.1158/2159-8290.CD-21-0316

22. Tirosh I, Izar B, Prakadan SM, Wadsworth MH2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* (2016) 352(6282):189–96. doi: 10.1126/science.aad0501

23. Maeser D, Gruener RF, Huang RS. oncoPredict: an r package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Briefings Bioinf* (2021) 22(6):bbab260. doi: 10.1093/bib/bbab260

24. Al Mahi N, Zhang EY, Sherman S, Yu JJ, Medvedovic M. Connectivity map analysis of a single-cell RNA-sequencing-derived transcriptional signature of mTOR signaling. *Int J Mol Sci* (2021) 22(9):4371. doi: 10.3390/ijms22094371

25. Yang C, Zhang H, Chen M, Wang S, Qian R, Zhang L, et al. A survey of optimal strategy for signature-based drug repositioning and an application to liver cancer. *Elife* (2022) 11:e71880. doi: 10.7554/eLife.71880

26. Tanese K, Hashimoto Y, Berkova Z, Wang Y, Samaniego F, Lee JE, et al. Cell surface CD74-MIF interactions drive melanoma survival in response to interferon-γ. *J Invest Dermatol* (2015) 135(11):2901. doi: 10.1038/jid.2015.259

27. Jeong H, Lee SY, Seo H, Kim BJ. Recombinant mycobacterium smegmatis delivering a fusion protein of human macrophage migration inhibitory factor (MIF) and IL-7 exerts an anticancer effect by inducing an immune response against MIF in a tumor-bearing mouse model. *J Immunother Cancer* (2021) 9(8). doi: 10.1136/jitc-2021-003180

28. Wirtz T, Saal A, Bergmann I, Fischer P, Heinrichs D, Brandt EF, et al. Macrophage migration inhibitory factor exerts pro-proliferative and anti-apoptotic effects via CD74 in murine hepatocellular carcinoma. *Br J Pharmacol* (2021) 178(22):4452–67. doi: 10.1111/bph.15622

29. Ma L, Du W, Chen Z, Xiang C. Single-cell atlas of tumor clonal evolution in liver cancer. *bioRxiv* (2020). doi: 10.1101/2020.08.18.254748

30. Song Z, Chen W, Athavale D, Ge X, Desert R, Das S, et al. Osteopontin takes center stage in chronic liver disease. *Hepatol (Baltimore Md.)* (2021) 73(4):1594–608. doi: 10.1002/hep.31582

31. Moon B, Jeong WJ, Park J, Kim TI, Min Do S, Choi KY. Role of oncogenic K-ras in cancer stem cell activation by aberrant wnt/β-catenin signaling. *J Natl Cancer Institute* (2014) 106(2):djt373. doi: 10.1093/jnci/djt373

32. Li J, Zheng S, Chen B, Butte AJ, Swamidass SJ, Lu Z. A survey of current trends in computational drug repositioning. *Briefings Bioinf* (2016) 17(1):2–12. doi: 10.1093/bib/bbv020

33. Sadia M. *A review on CRISPR-cas system based applications in oncology*. Brac University (2022).

34. Yu L, Xu F, Gao L. Predict new therapeutic drugs for hepatocellular carcinoma based on gene mutation and expression. *Front Bioengineering Biotechnol* (2020) 8:8. doi: 10.3389/fbioe.2020.00008

35. Du D, Liu C, Qin M, Zhang X, Xi T, Yuan S, et al. Metabolic dysregulation and emerging therapeutical targets for hepatocellular carcinoma. *Acta Pharm Sin B* (2021) 12(2):558–80. doi: 10.1016/j.apsb.2021.09.019

36. Font-Burgada J, Shalapour S, Ramaswamy S, Hsueh B, Rossell D, Umemura A, et al. Hybrid periportal hepatocytes regenerate the injured liver without giving rise to cancer. *Cell* (2015) 162(4):766–79. doi: 10.1016/j.cell.2015.07.026

37. Marquardt JU. *Deconvolution of the cellular origin in hepatocellular carcinoma: Hepatocytes take the center stage*. Wiley Online Library (2016) p. 1020–3.

38. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the warburg effect: the metabolic requirements of cell proliferation. *science* (2009) 324(5930):1029–33. doi: 10.1126/science.1160809

39. Lee Y-K, Lim JJ, Jeoun UW, Min S, Lee EB, Kwon SM, et al. Lactate-mediated mitoribosomal defects impair mitochondrial oxidative phosphorylation and promote hepatoma cell invasiveness. *J Biol Chem* (2017) 292(49):20208–17. doi: 10.1074/jbc.M117.809012

40. Soukupova J, Malfettone A, Hyrossova P, Hernandez-Alvarez MI, Penuelas-Haro I, Bertran E, et al. Role of the transforming growth factor-β in regulating hepatocellular carcinoma oxidative metabolism. *Sci Rep* (2017) 7(1):1–15. doi: 10.1038/s41598-017-12837-y

41. Zaidi N, Lupien L, Kuemmerle NB, Kinlaw WB, Swinnen JV, Smans K. Lipogenesis and lipolysis: the pathways exploited by the cancer cells to acquire fatty acids. *Prog Lipid Res* (2013) 52(4):585–9. doi: 10.1016/j.plipres.2013.08.005

42. Peng H, Zhu E, Zhang Y. Advances of cancer-associated fibroblasts in liver cancer. *biomark Res* (2022) 10(1):1–18. doi: 10.1186/s40364-022-00406-z

43. Liu C, Liu L, Chen X, Cheng J, Zhang H, Zhang C, et al. LSD1 stimulates cancer-associated fibroblasts to drive Notch3-dependent self-renewal of liver cancer stem–like CellsLSD1 regulates liver CSC self-renewal via Notch3 signaling. *Cancer Res* (2018) 78(4):938–49. doi: 10.1158/0008-5472.CAN-17-1236

44. Zhang M, Yang H, Wan L, Wang Z, Wang H, Ge C, et al. Single-cell transcriptomic architecture and intercellular crosstalk of human intrahepatic cholangiocarcinoma. *J Hepatol* (2020) 73(5):1118–30. doi: 10.1016/j.jhep.2020.05.039

45. Wang H, Feng C, Lu M, Zhang B, Xu Y, Zeng Q, et al. Integrative single-cell transcriptome analysis reveals a subpopulation of fibroblasts associated with favorable prognosis of liver cancer patients. *Trans Oncol* (2021) 14(1):100981. doi: 10.1016/j.tranon.2020.100981

46. Yu D, Green B, Marrone A, Guo Y, Kadlubar S, Lin D, et al. Suppression of CYP2C9 by microRNA hsa-miR-128-3p in human liver cells and association with hepatocellular carcinoma. *Sci Rep* (2015) 5(1):1–9. doi: 10.1038/srep08534

47. Wang Q, Tang Q, Zhao L, Zhang Q, Wu Y, Hu H, et al. Time serial transcriptome reveals Cyp2c29 as a key gene in hepatocellular carcinoma development. *Cancer Biol Med* (2020) 17(2):401. doi: 10.20892/j.issn.2095-3941.2019.0335

48. Chen Z, Liu Y, Yao L, Guo S, Gao Y, Zhu P. The long noncoding RNA lncZic2 drives the self-renewal of liver tumor–initiating cells via the protein kinase c substrates MARCKS and MARCKSL1. *J Biol Chem* (2018) 293(21):7982–92. doi: 10.1074/jbc.RA117.001321

49. Yang L, Deng WL, Zhao BG, Xu Y, Wang XW, Fang Y, et al. FOXO3-induced lncRNA LOC554202 contributes to hepatocellular carcinoma progression via the miR-485-5p/BSG axis. *Cancer Gene Ther* (2022) 29(3):326–40. doi: 10.1038/s41417-021-00312-w

50. Qian E-N, Han SY, Ding SZ, Lv X. Expression and diagnostic value of CCT3 and IQGAP3 in hepatocellular carcinoma. *Cancer Cell Int* (2016) 16(1):1–8. doi: 10.1186/s12935-016-0332-3

51. Zhang Y, Wang Y, Wei Y, Wu J, Zhang P, Shen S, et al. Molecular chaperone CCT3 supports proper mitotic progression and cell proliferation in hepatocellular carcinoma cells. *Cancer Lett* (2016) 372(1):101–9. doi: 10.1016/j.canlet.2015.12.029

52. Cui X, Hu ZP, Li Z, Gao PJ, Zhu JY. Overexpression of chaperonin containing TCP1, subunit 3 predicts poor prognosis in hepatocellular carcinoma. *World J Gastroenterology: WJG* (2015) 21(28):8588. doi: 10.3748/wjg.v21.i28.8588

53. Li Y, Xiong H. Correlation of LAGE3 with unfavorable prognosis and promoting tumor development in HCC via PI3K/AKT/mTOR and Ras/RAF/MAPK pathways. *BMC Cancer* (2022) 22(1):1–13. doi: 10.1186/s12885-022-09398-3

54. Guo X, Wang Z, Zhang J, Xu Q, Hou G, Yang Y, et al. Upregulated KPNA2 promotes hepatocellular carcinoma progression and indicates prognostic significance across human cancer types. *Acta Biochim Biophys Sin* (2019) 51(3):285–92. doi: 10.1093/abbs/gmz003

55. Shen Q, Nam SW. SF3B4 as an early-stage diagnostic marker and driver of hepatocellular carcinoma. *BMB Rep* (2018) 51(2):57. doi: 10.5483/BMBRep.2018.51.2.021

56. Iguchi T, Komatsu H, Masuda T, Nambara S, Kidogami S, Ogawa Y, et al. Increased copy number of the gene encoding SF3B4 indicates poor prognosis in hepatocellular carcinoma. *Anticancer Res* (2016) 36(5):2139–44.

57. Liu Z, Li W, Pang Y, Zhou Z, Liu S, Cheng K, et al. SF3B4 is regulated by microRNA-133b and promotes cell proliferation and metastasis in hepatocellular carcinoma. *EBioMedicine* (2018) 38:57–68. doi: 10.1016/j.ebiom.2018.10.067

58. Deng G, Luo Y, Zhang Y, Zhang J, He Z. Enabled homolog (ENAH) regulated by RNA binding protein splicing factor 3b subunit 4 (SF3B4) exacerbates the proliferation, invasion and migration of hepatocellular carcinoma cells via notch signaling pathway. *Bioengineered* (2022) 13(2):2194–206. doi: 10.1080/21655979.2021.2023983

59. Liu W-B, Jia WD, Ma JL, Xu GL, Zhou HC, Peng Y, et al. Knockdown of GTPBP4 inhibits cell growth and survival in human hepatocellular carcinoma and its prognostic significance. *Oncotarget* (2017) 8(55):93984. doi: 10.18632/oncotarget.21500

60. Zhou Q, Yin Y, Yu M, Gao D, Sun J, Yang Z, et al. GTPBP4 promotes hepatocellular carcinoma progression and metastasis via the PKM2 dependent glucose metabolism. *Redox Biol* (2022) 56:102458. doi: 10.1016/j.redox.2022.102458

61. Ding G-Y, Zhu XD, Ji Y, Shi GM, Shen YH, Zhou J, et al. Serum PON1 as a biomarker for the estimation of microvascular invasion in hepatocellular carcinoma. *Ann Trans Med* (2020) 8(5):204. doi: 10.21037/atm.2020.01.44

62. Liu H, Zhang L, Wang P. Complement factor h–related 3 overexpression affects hepatocellular carcinoma proliferation and apoptosis. *Mol Med Rep* (2019) 20(3):2694–702. doi: 10.3892/mmr.2019.10514

63. Zeng Z, Lei S, Wang J, Yang Y, Lan J, Tian Q, et al. A novel hypoxia-driven gene signature that can predict the prognosis of hepatocellular carcinoma. *Bioengineered* (2022) 13(5):12193–210. doi: 10.1080/21655979.2022.2073943

64. Wang S, Xie J, Zou X, Pan T, Yu Q, Zhuang Z, et al. Single-cell multiomics reveals heterogeneous cell states linked to metastatic potential in liver cancer cell lines. *Iscience* (2022) 25(3):103857. doi: 10.1016/j.isci.2022.103857

65. Yan T, Yu L, Zhang N, Peng C, Su G, Jing Y, et al. The advanced development of molecular targeted therapy for hepatocellular carcinoma. *Cancer Biol Med* (2022) 19 (6):802. doi: 10.20892/j.issn.2095-3941.2021.0661

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

67. Marin-Acevedo JA, Kimbrough EO, Lou Y. Next generation of immune checkpoint inhibitors and beyond. *J Hematol Oncol* (2021) 14(1):1–29. doi: 10.1186/s13045-021-01056-8

68. Faivre S, Rimassa L, Finn RS. Molecular therapies for HCC: Looking outside the box. *J Hepatol* (2020) 72(2):342–52. doi: 10.1016/j.jhep.2019.09.010

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

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