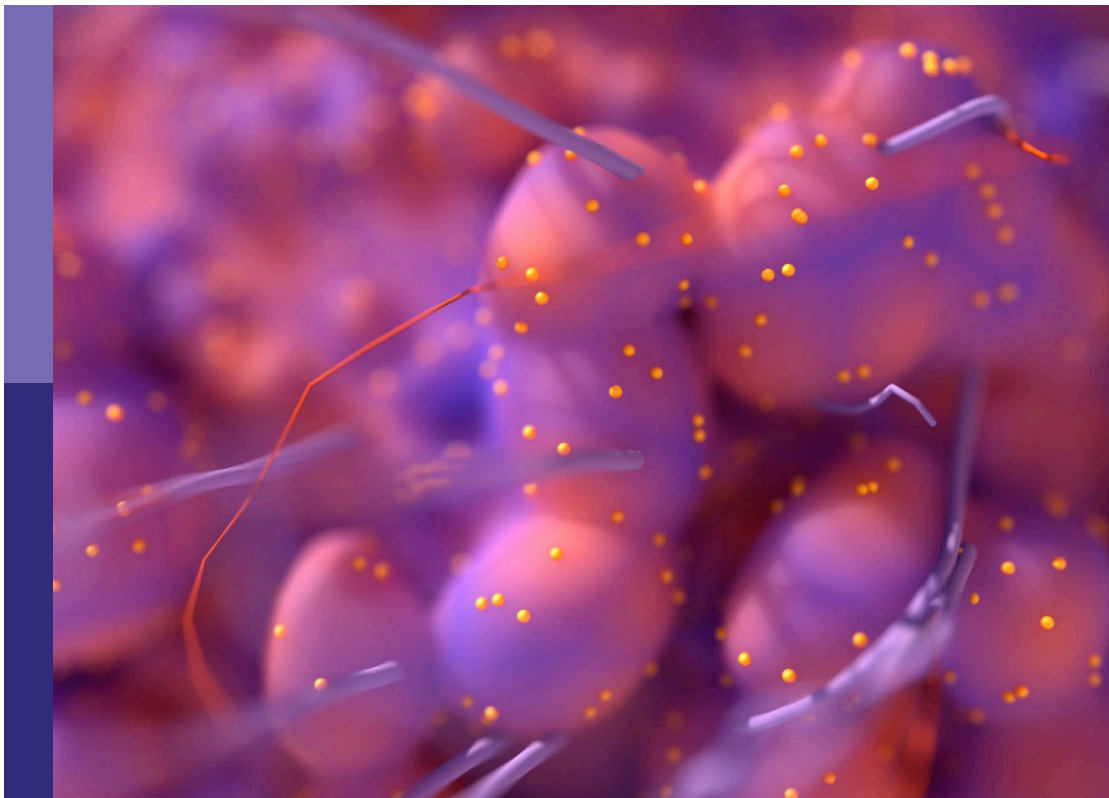
frontiers | Research Topics

# Application and innovation of multiomics technologies in clinical oncology

**Edited by**
Ye Wang, Xiaoming Xing, Xinmin Li and Anton A. Buzdin

**Published in**
Frontiers in Oncology
Frontiers in Immunology



FRONTIERS EBOOK COPYRIGHT STATEMENT

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

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

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

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

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

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

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

ISSN 1664-8714
ISBN 978-2-83252-156-4
DOI 10.3389/978-2-83252-156-4

## About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

# Application and innovation of multiomics technologies in clinical oncology

**Topic editors**

Ye Wang — Qingdao University, The Affiliated Qingdao Central Hospital of Qingdao University, China

Xiaoming Xing — The Affiliated Hospital of Qingdao University, China

Xinmin Li — University of California, Los Angeles, United States

Anton A. Buzdin — European Organisation for Research and Treatment of Cancer, Belgium

**Citation**

Wang, Y., Xing, X., Li, X., Buzdin, A. A., eds. (2023). *Application and innovation of multiomics technologies in clinical oncology*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83252-156-4

# Table of contents

07 Editorial: Application and innovation of multiomics technologies in clinical oncology
Xinmin Li, Anton Budzin and Ye Wang

09 Identification of a 3-miRNA Signature Associated With the Prediction of Prognosis in Nasopharyngeal Carcinoma
Jinhui Zhou, Bo Zhang, Xin Zhang, Chengyu Wang and Yu Xu

20 Comprehensive Analysis Identified ASF1B as an Independent Prognostic Factor for HBV-Infected Hepatocellular Carcinoma
Xianmo Wang, Huawei Yi, Jiancheng Tu, Wen Fan, Jiahao Wu, Li Wang, Xiang Li, Jinrong Yan, Huali Huang and Rong Huang

29 COL5A2 Inhibits the TGF-β and Wnt/β-Catenin Signaling Pathways to Inhibit the Invasion and Metastasis of Osteosarcoma
Yan-Long Han, Dan Luo, Kakeng Habaxi, Julaiti Tayierjiang, Wei Zhao, Wei Wang, Wumaierjiang Aikebaier and Li Wang

38 Construction of a Co-Expression Network for lncRNAs and mRNAs Related to Urothelial Carcinoma of the Bladder Progression
Yeqing Mao, Chao Wen and Zitong Yang

49 The Cell Cycle-Associated Protein CDKN2A May Promotes Colorectal Cancer Cell Metastasis by Inducing Epithelial-Mesenchymal Transition
Wei-Kun Shi, Yun-Hao Li, Xue-Shan Bai and Guo-Le Lin

58 Search for Key Genes and Functional Pathways of Ulcerative Colitis to Colon Cancer Based on Bioinformatics
Shengbao Wang, Lingling Zhen, Xiaoli Li, Xu Fu, Peiwu Li and Dekui Zhang

67 Transcriptomic Portraits and Molecular Pathway Activation Features of Adult Spinal Intramedullary Astrocytomas
Nikolay Konovalov, Stanislav Timonin, Dmitry Asyutin, Mikhail Raevskiy, Maxim Sorokin, Anton Buzdin and Stanislav Kaprovoy

82 Investigating the Prognostic Significance of Pyroptosis-Related Genes in Gastric Cancer and Their Impact on Cells' Biological Functions
Jie Yin, Gang Che, Wankun Wang, Shitu Chen and Jian Liu

90 The Differential Expression of CD47 may be Related to the Pathogenesis From Myelodysplastic Syndromes to Acute Myeloid Leukemia
Xiao Yan, Binbin Lai, Xuyan Zhou, Shujun Yang, Qunfang Ge, Miao Zhou, Cong Shi, Zhijuan Xu and Guifang Ouyang

98 miR-139-5p Was Identified as Biomarker of Different Molecular Subtypes of Breast Carcinoma
Haohang Sun, Ji Dai, Mengze Chen, Qi Chen, Qiong Xie, Weijun Zhang, Guoqing Li and Meidi Yan

109 LncRNA-AC02278.4 Is a Novel Prognostic Biomarker That Promotes Tumor Growth and Metastasis in Lung Adenocarcinoma
Xin Chen, Fan Zhou, Wenjun Ren, Jishu Guo, Xiaobin Huang, Jun Pu, Xiaoqun Niu and Xiulin Jiang

120 Cyclin-Dependent Kinase 6 Identified as the Target Protein in the Antitumor Activity of *Tetrastigma hemsleyanum*
Chaoguang Wei, Yuxiang Zhao, Tao Ji, Yong Sun, Xudong Cai and Xin Peng

134 Identification of a 12-Gene Signature and Hub Genes Involved in Kidney Wilms Tumor *via* Integrated Bioinformatics Analysis
Guoping Huang and Jianhua Mao

145 A Prognostic Gene Signature for Hepatocellular Carcinoma
Rong Chen, Meng Zhao, Yanli An, Dongfang Liu, Qiusha Tang and Gaojun Teng

156 Identification by Bioinformatics Analysis of Potential Key Genes Related to the Progression and Prognosis of Gastric Cancer
Wencang Gao and Min Yang

165 Integrative Analysis Identifies Cell-Type-Specific Genes Within Tumor Microenvironment as Prognostic Indicators in Hepatocellular Carcinoma
Zi-Li Huang, Bin Xu, Ting-Ting Li, Yong-Hua Xu, Xin-Yu Huang and Xiu-Yan Huang

177 Exploring the Potential of Pyroptosis-Related Genes in Predicting Prognosis and Immunological Characteristics of Pancreatic Cancer From the Perspective of Genome and Transcriptome
Jing Zhang, Xiaomin You, Dong Kang and Guoxiong Zhou

190 The Function, Role and Process of DDX58 in Heart Failure and Human Cancers
Ping Yu, Peng Liang, Shifeng Pang, Wenjian Yuan, Yuxiang Zhao and Qiaojuan Huang

202 Colorectal Cancer Chemotherapy Drug Bevacizumab May Induce Muscle Atrophy Through CDKN1A and TIMP4
Qun Xu, Jinyou Li, Yue Wu, Wenjing Zhou and Zherong Xu

212 Effects of Tumor-Derived DNA on CXCL12-CXCR4 and CCL21-CCR7 Axes of Hepatocellular Carcinoma Cells and the Regulation of Sinomenine Hydrochloride
Conghuan Shen, Jianhua Li, Ruidong Li, Zhenyu Ma, Yifeng Tao, Quanbao Zhang and Zhengxin Wang

227 Comprehensive Analysis of TRP Channel–Related Genes in Patients With Triple-Negative Breast Cancer for Guiding Prognostic Prediction
Haojie Zhang, Xiangsheng Zhang, Xiaohong Wang, Hongguang Sun, Changran Hou, Yue Yu, Song Wang, Fangxu Yin and Zhenlin Yang

237 Genomic Profiling Reveals Novel Predictive Biomarkers for Chemo-Radiotherapy Efficacy and Thoracic Toxicity in Non-Small-Cell Lung Cancer
Kewen He, Shaotong Zhang, Jiaohui Pang, Jiani C. Yin, Dianbin Mu, Jun Wang, Hong Ge, Jie Ma, Zhe Yang, Xiaoli Zheng, Lihua Dong, Junli Zhang, Pengyu Chang, Li Li, Shanshan Tang, Hua Bao, Xue Wu, Xiaonan Wang, Yang Shao, Jinming Yu and Shuanghu Yuan

252 MiR-520a-5p/PPP5C regulation pattern is identified as the key to gemcitabine resistance in pancreatic cancer
Ruibiao Fu, Qian Shao, Bin Yang, Yan Chen, Qinghuang Ye, Xi Chen and Jinhui Zhu

261 Construction of lncRNA and mRNA co-expression network associated with nasopharyngeal carcinoma progression
Xu Lu, Xing Chen, Xinke Wang, Jing Qing, Ji Li and Yunyun Pan

269 Bioinformatics analysis and experimental verification of the prognostic and biological significance mediated by fatty acid metabolism related genes for hepatocellular carcinoma
Xiao-Ren Zhu, Jia-Qi Zhu, Yu-Fei Chen, Yuan-Yuan Liu, Jing-Jing Lu, Jun Sun, Shi-Qing Peng, Min-Bin Chen and Yi-Ping Du

288 The global change of gene expression pattern caused by PTEN mutation affects the prognosis of glioblastoma
Shengjun Zhou, Haifeng Wang, Yi Huang, Yiwen Wu and Zhiqing Lin

296 The prognostic value of immune-related genes *AZGP1, SLCO5A1,* and *CTF1* in Uveal melanoma
Wanpeng Wang and Sha Wang

307 Comprehensive prognostic and immunological analysis of CCT2 in pan-cancer
Wenming Lv, Lin Shi, Jiebing Pan and Shengbao Wang

319 HOXA1 is a radioresistance marker in multiple cancer types
Lu He, Min Liang, Weisheng Guo, Jinquan Liu and Yi Yu

330 Single-cell transcriptome analysis reveals T-cell exhaustion in denosumab-treated giant cell tumor of bone
Meiling Yang, Fen Wang, Guohao Lu, Mingzhe Cheng, Wei Zhao and Changye Zou

344 lncRNA NEAT1 promotes the proliferation and metastasis of hepatocellular carcinoma by regulating the FOXP3/PKM2 axis
Junping Pan, Yingzhe Hu, Chenlu Yuan, Yafu Wu and Xinhua Zhu

356 DIRAS3, GPR171 and RAC2 were identified as the key molecular patterns associated with brain metastasis of breast cancer
Ji Dai, Qi Chen, Guoqing Li, Mengze Chen, Haohang Sun and Meidi Yan

365 Prognostic signature construction of energy metabolism-related genes in pancreatic cancer
Hao Liu, Jianhua Zhang, Chaoguang Wei, Zhao Liu, Wei Zhou, Pan Yang, Yifu Gong and Yuxiang Zhao

376 MIR4435-2HG in exosomes promotes gastric carcinogenesis by inducing M2 polarization in macrophages
Chaofeng Li, Zhengju Chen, Jinli Gao, Tao Tang, Lei Zhou, Guochao Zhang, Dongdong Zhang, Chao Shen, Lei Guo and Tao Fu

386 PGAM5: A necroptosis gene associated with poor tumor prognosis that promotes cutaneous melanoma progression
Jianzhong Peng, Tao Wang, Chao Yue, Xianyan Luo and Peng Xiao

395 Distinct gene mutation profiles among multiple and single primary lung adenocarcinoma
Yadong Wang, Guanghui Wang, Haotian Zheng, Jichang Liu, Guoyuan Ma, Gemu Huang, Qingtao Song and Jiajun Du

410 Analysis of the regulatory mechanisms of prognostic immune factors in thyroid cancer
Yin Tian, Tao Xie and Xue Sun

422 Identification of prognosis-related gene features in low-grade glioma based on ssGSEA
Yuanzhi He, Zhangping Lin and Sanyang Tan

434 LASSO-based screening for potential prognostic biomarkers associated with glioblastoma
Yin Tian, Li'e Chen and Yun Jiang

443 Wnt signaling pathway-related gene PRICKLE1 is a prognostic biomarker for esophageal squamous cell carcinoma
Jinxian He, Gaofeng Liang, Hui Tian, Yiqing Wang, Li Yu, Wang Lv, Jian Hu and Weiyu Shen

451 The bioinformatic approach identifies *PARM1* as a new potential prognostic factor in osteosarcoma
Haijun Feng, Liping Wang, Jie Liu and Shengbao Wang

462 The role of physics in multiomics and cancer evolution
Lucie E. Gourmet and Simon Walker-Samuel

471 Heterogeneous pattern of gene expression driven by *TTN* mutation is involved in the construction of a prognosis model of lung squamous cell carcinoma
Zhao Liu, Xiaowen Zhao, Ruihong Wang, Xiangyue Tang, Yuxiang Zhao, Guanghui Zhong, Xin Peng and Chunlin Zhang

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 28 March 2023
DOI 10.3389/fonc.2023.1179829

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Heather Cunliffe,
University of Otago, New Zealand

*CORRESPONDENCE
Ye Wang
yewang@qdu.edu.cn

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

RECEIVED 05 March 2023
ACCEPTED 16 March 2023
PUBLISHED 28 March 2023

CITATION
Li X, Budzin A and Wang Y (2023) Editorial: Application and innovation of multiomics technologies in clinical oncology. *Front. Oncol.* 13:1179829. doi: 10.3389/fonc.2023.1179829

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

# Editorial: Application and innovation of multiomics technologies in clinical oncology

Xinmin Li[1], Anton Budzin[2,3,4] and Ye Wang[5]*

[1]Department of Pathology & Laboratory Medicine, Technology Center for Genomics & Bioinformatics, University of California, Los Angeles (UCLA), Los Angeles, CA, United States, [2]Laboratory for Translational Genomic Bioinformatics, Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russia, [3]World-Class Research Center, I.M. Sechenov First Moscow State Medical University, Moscow, Russia, [4]World-Class Research Center "Digital Biodesign And Personalized Healthcare", Sechenov First Moscow State Medical University, Moscow, Russia, [5]Clinical Laboratory, Qingdao Central Hospital, The Second Affiliated Hospital of Medical College of Qingdao University, Qingdao, China

KEYWORDS

**clinical oncology, multiomics technologies, genomics, transcriptomics, biomarkers**

Editorial on the Research Topic
Application and innovation of multiomics technologies in clinical oncology

Cancer is a complex disease that involves multiple layers of genetic, epigenetic, and metabolic alterations and interactions to orchestrate tumor growth and progression. Multiomics technologies empower researchers to generate and integrate these different types of data to provide coherent and complete picture of cancer biology, and to effectively identify key molecular, genetic alterations, pathways, and networks that drive cancer development and progression. These findings can lead to the discovery of new biomarkers, new diagnostic tools, new strategies to overcome resistance and treat cancer patients.

This special issue entitled "Application and Innovation of Multiomics Technologies in Clinical Oncology" compiled 43 articles, highlighting different applications and findings using genomics, transcriptomics and multiomics in a wide range of cancer types. One of the powerful and common omics'applications is its ability to identify molecular signatures and biomarkers for cancer classification, metastasis, prognosis and treatment. This special issue provides ample examples of such, including miR-139-5p and a 3-miRNA signature as potential biomarkers for subtyping breast carcinoma (Yan et al.) and predicting survival in nasopharyngeal carcinoma (Zhou et al.), LncRNA-AC02278.4 and ASF1B as Prognostic Biomarkers for Tumor Growth and Metastasis in Lung Adenocarcinoma (Chen et al.), and poor outcome in the HBV-infected hepatocellular carcinoma (Wang et al.), immune-related and energy metablolism-related gene signatures with prognostic potential for pancreatic cancer (Chen et al.) and hepatocellular carcinoma (Liu et al.).

Of particular interest, Huang et al. used univariable Cox regression analyses and differential gene expression analysis with bulk RNAseq data to identify eight genes that are closely associated with the recurrence-free survivals (RFS) of HCC patients in both training and three validation cohorts. They subsequently built a multivariable Cox model based on the expression of these genes to predict the risk of HCC recurrence. The results showed that the Eight-Gene Cox (EGC) model outperformed other published models and could serve as independent predictor in predicting HCC recurrence. Furthermore, the authors

investigated the cell-type-specific expression patterns and their biological functions of these eight marker genes in tumor microenvironment using single-cell RNA sequencing data. Interestingly, these genes had different expression patterns in different cell types. For example, PLCB1 and SLC22A7 were predominantly expressed in malignant cells while FASLG and IL2RB were specifically expressed in T cells, suggesting their functional roles in these given cell types.

The results of this study are not only significant in terms of providing a novel model for HCC recurrence prediction but also in shedding light on the underlying mechanisms behind HCC recurrence. It demonstrates the importance of the tumor microenvironment and highlights the need for a more comprehensive understanding of the interplay between tumor cells and their surrounding microenvironment in cancer development and progression. The insights gained into the functional roles of the HCC recurrence marker genes provide clues for further research and the development of new therapeutic strategies for the treatment of HCC.

In addition to gene expression signatures and biomarkers, the DNA markers are equally informative for predicting treatment efficacy, toxicity and clinical outcome. He et al. identified genetic determinants that can explain individual variations in radiosensitivity and radiotherapy-associated toxicity in 122 patients with unresectable stage III non-small-cell lung cancer (NSCLC). The research confirmed previously reported genetic association with clinical outcomes and also identified the *FGFR* family genes, *MET*, *PTEN*, and *NOTCH2* as potential novel and independent risk factors of poor post-CRT survival by using a 474-cancer and radiotherapy-related gene panel. These novel biomarkers may offer new therapeutic targets for patients with NSCLC. The identification of genetic markers in relation to radiotherapy-associated thoracic toxicity is critical for the design of better treatment strategies to reduce or eliminate the risk of toxicities associated with CRT.

This study has several important clinical implications. It demonstrated the clinical utility of focused NGS panels in identifying predictive biomarkers for response to CRT, reinforced the importance of pre-treatment genetic information to better inform CRT outcomes and clinical actions in stage III unresectable NSCLCs, and stressed that integrated analysis of multiple alterations can lead to improved stratification of the risk populations.

The utilization of a multi-omics approach has also facilitated the comprehension of molecular pathways, mechanisms, and mechanistic models for tumor development by compiling various types of data and constructing new data-driven concepts. In this Research Topic, such efforts were documented and published for several common and rare cancers, including colon (Wang et al.), gastric (Gao and Yang, 2022), bladder (Mao et al.), nasopharyngeal carcinoma (Lu et al.), and adult spinal intramedullary astrocytomas (Konovalov et al.), the latter paper is supplemented with the first-of-its-kind collection of experimental RNAseq profiles for this orphan cancer.

As the field of omics continues to evolve and expand, a multiomics approach is becoming increasingly important for advancing our understanding of complex diseases such as cancer. By embracing a multiomics approach that combine and integrate multiple layers of molecular information, we can accelerate the pace of discovery and translation of omics research into clinical practice in cancer.

## Author contributions

XL, AB, and YW: writing original draft and editing. All authors contributed to the article and approved the submitted version.

## Funding

AB research was supported by the ERA-NET RUS Plus Program (GOTICS Project), and the Russian Foundation for Basic Research grant 20-515-76007. YW research was supported by the the National Natural Science Foundation of China (No. 81670822).

## Acknowledgments

We would like to thank all the authors, reviewers, and technical editorial stuff for their contributions to this issue.

## 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.

frontiers in Oncology

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fonc.2021.823603

# Identification of a 3-miRNA Signature Associated With the Prediction of Prognosis in Nasopharyngeal Carcinoma

*Jinhui Zhou*[1], *Bo Zhang*[2], *Xin Zhang*[1], *Chengyu Wang*[1] *and Yu Xu*[1]*

[1] Department of Otolaryngology Head and Neck Surgery, The Affiliated Huai'an No.1 People's Hospital of Nanjing Medical University, Huai'an, China, [2] Teaching and Research Section of Otolaryngology, Hubei University of Science and Technology, Xianning, China

Nasopharyngeal carcinoma (NPC) is a malignant tumor caused by an infection of the epithelial cells of the nasopharynx, which is highly metastatic and aggressive. Due to the deep anatomical site and atypical early symptoms, the majority of NPC patients are diagnosed at terminal stages. There is growing evidence that microRNAs offer options for early detection, accurate diagnosis, and prediction of malignancy treatment response. Therefore, the purpose of this article was to identify microRNAs that predict the prognosis of patients with NPC by integrating biological information analysis. In this study, we utilized the GSE36682 dataset rooted in the Gene Expression Omnibus (GEO) data bank, including 62 cases of NPC tissues and six cases of non-cancerous tissues. The miRNAs were subjected to weighted gene co-expression network analysis, and hub miRNAs were screened for differentially upregulated miRNAs from modules highly correlated with tumor progression. We took a lot of time to calculate the risk scores of miRNA markers for 62 NPC patients, and incidentally combined the clinical survival information of patients to finally identify the three key miRNAs, and then divided the patients into low- and high-risk groups. Kaplan-Meier curve analysis revealed that the overall survival of patients in the high-risk group was obviously shorter than that of the low-risk group. Subsequently, the target genes of the three miRNAs were predicted and analyzed for functional enrichment. In summary, a prognostic predictive risk model based on three miRNA profiles may increase prognostic predictive value and provide reference information for the precise treatment of nasopharyngeal carcinoma.

**Keywords: nasopharyngeal carcinoma, WGCNA, miRNA-mNRA, lncRNA-mRNA co-expression network, prognostic marker**

OPEN ACCESS

***Edited by:***
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

***Reviewed by:***
*Gong Yifu,*
*Ningbo University, China*
*Lili Zhou,*
*First Affiliated Hospital of Jilin University, China*

***Correspondence:**
*Yu Xu*
*xuyu7h5j@163.com*

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

***Received:*** *27 November 2021*
***Accepted:*** *31 December 2021*
***Published:*** *27 January 2022*

***Citation:***
*Zhou J, Zhang B, Zhang X, Wang C and Xu Y (2022) Identification of a 3-miRNA Signature Associated With the Prediction of Prognosis in Nasopharyngeal Carcinoma. Front. Oncol. 11:823603. doi: 10.3389/fonc.2021.823603*

## INTRODUCTION

Nasopharyngeal carcinoma (NPC) is one of the frequently occurring malignancies of the head and neck, with a unique geographic distribution and a relatively high incidence in Southeast Asia, including in southeastern China (as well as Guangdong and Hong Kong), India, and Thailand (1–3), though it is fairly rare in Western countries (4). It is generally accepted that the development of

nasopharyngeal carcinoma is closely associated with the Epstein-Barr virus (EBV), cumulative inheritance in the context of susceptibility genetics and environmental factors, and genome-wide epigenetic modifications of tumor-associated genes (5). For the past few years, the mortality and incidence of nasopharyngeal carcinoma have gradually decreased due to effective screening and treatment strategies such as radiotherapy and combination chemotherapy, as well as the advent of immunotherapy (6–8). However, local recurrence and distant metastases are very common in patients with advanced NPC. Statistically, more than 30% of stage III and IV patients develop local recurrence or distant metastases within five years of receiving combination therapy (9, 10). Therefore, there is a need to clearly identify some sensitive biomarkers and new therapeutic targets for pre-NPC diagnosis and provide effective therapeutic measures.

The development of NPC is a complex, multistep, multifactorial, and aberrantly regulated process of signaling pathways involving aberrant activation of certain inactivation and oncogenes of tumor suppressor genes (11). Non-coding RNAs are usually dysregulated in tumor pathogenesis. MicroRNAs (miRNAs), a class of functionally short non-coding RNAs, are related to all markers of carcinoma development, metastasis, and progression, and have become one of the hot spots in carcinoma research (12). Elicited from microarray-based transcriptome profiling, an increasing number of studies have shown that many miRNAs are closely associated with the development and progression of nasopharyngeal carcinoma. For example, Jiang et al. (13) found that EBV-encoded miR-BART2-5p showed an elevated pattern in the serum of preclinical NPC patients and could promote NPC metastasis by inhibiting Rnd3. Zhou et al. (14) analyzed microarray data from the plasma of nasopharyngeal carcinoma patients and highlighted miR-548 and miR-940 as potential diagnostic biomarkers for nasopharyngeal carcinoma with high sensitivity and specificity. In addition, there have also been many studies conducted to find miRNA signatures associated with the prognosis of nasopharyngeal carcinoma *via* high-throughput miRNA expression profiling. Bruce et al. (15) identified a 4-miRNA signature (miR-140, miR-449b, miR-154, and miR-34c) associated with the risk of distant metastasis. Zhang et al. successfully identified a 4-miRNA signature by integrating biological information analysis for predicting the prognosis of patients with nasopharyngeal carcinoma. However, although many studies have analyzed miRNA expression profiles in nasopharyngeal carcinoma, unfortunately the detailed molecular mechanisms of NPC have not been fully elucidated.

In this research, we identified miRNA expression based on 62 nasopharyngeal carcinoma patient samples in the GSE36682 dataset, subjected them to weighted gene co-expression network analysis (WGCNA), and identified miRNAs associated with nasopharyngeal carcinoma progression. A prognostic prediction risk model was constructed based on the 3-miRNA signature to provide prognosis prediction, diagnosis, or precise treatment for potential candidate biomarkers of nasopharyngeal carcinoma.

# METHODS

## Data Sources

We obtained research from the Gene Expression Omnibus database using the keywords "nasopharyngeal carcinoma" and "miRNA," and then manually reviewed and selected cohorts containing miRNA expression and clinical survival information. The training GSE36682 dataset (platform: GPL15311, Human miRNA 1K) was obtained from Sun Yat-sen University, China and contained 62 nasopharyngeal cancer tissues and six non-cancerous tissues; the validation GSE32960 (16) dataset contained 312 nasopharyngeal cancer tissues and 18 normal tissues, with clinical survival data from the Zhang et al. (17) study obtained from the original investigators.

## Differential Expression Analysis

Differential expression of miRNAs was studied with the limma R software package; "adjusted P value<0.05 and |logFC|≥1" was defined as the threshold miRNA differential expression screen, and heatmap and volcano maps were plotted using the pheatmap and ggplot2 packages, respectively.

## Weighted Gene Co-Expression Network Analysis (WGCNA)

Using the WGCNA online resource, Pearson correlation coefficients between genes were calculated by using miRNAs from the WGCNA dataset and appropriate soft thresholds were selected β As a backup. A one-step method was used to construct a gene network, transforming the adjacency matrix into a topological overlap matrix (TOM) and generating a hierarchical clustering tree of genes. The DynamicTreeCut method was used to identify highly correlated co-expressed gene modules with the threshold set to cutHeight = 0.2 and minSize = 20. The Pearson correlation test was used to analyze the relationship between module signature genes (MEs) and clinical features.

## Lasso Regression Analysis

First, we performed the lasso regression algorithm, and 10-fold cross-validation was used to determine the parameters to obtain a suitable model. The multivariate Cox regression analysis was then performed on the genes obtained from the lasso regression, and the multivariate regression coefficient was calculated for each miRNA. Finally, the risk score equation could be constructed. Patients were divided into high- and low-risk groups according to the optimal risk score cutoff values. We first compared the overall survival time of the two groups, which would use to Kaplan Meier survival curve analysis, and subsequently evaluated the predictive value of the miRNA markers *via* time-dependent ROC.

## Cell Culture

Human nasopharyngeal epithelial cells NP69 and human nasopharyngeal carcinoma cell line SUNE-1 cells were rooted in the Shanghai Cell Bank, Chinese Academy of Sciences. SUNE-1 cells were cultured in RPMI1640 medium containing 10% fetal bovine serum (FBS) and NP69 cells were cultured in

keratinocyte/serum-free (K-S) medium containing growth factors (Gibco, Grand Island, NY, USA). The serum-free (K-SFM) medium and all cells were cultured in a 37°C, 5% CO2 incubator.

## Fluorescence Quantitative PCR

Total RNA was obtained from NP69 and SUNE-1 cells according to the instructions of TRIzol (Invitrogen, USA). 1 μg of RNA was reverse transcribed into cDNA according to the instructions of the PrimeScriptTM RT reagent kit (TaKaRa, Japan) and 1 μg of RNA was reverse transcribed into cDNA according to the instructions of the SYBR Green reagent (TaKaRa, Japan). PCR amplification was performed according to the instructions of the SYBR Green reagent (TaKaRa, Japan). The PCR reaction system was as follows: 1 μL cDNA, 0.5 μL each upstream and downstream primers, 10μL SYBR Premix Ex Taq, 8μL ddH2O. The reaction conditions were: pre-denaturation at 95°C for 7 mins, followed by pre-denaturation at 95°C (5s), 60°C (30s), and 72°C (3 mins). The reaction was performed for 35 cycles at 95°C (5s), 60°C (30s), and 72°C (3 mins). The $2^{-\Delta\Delta CT}$ method calculated the relative expression of target genes. The experiments were repeated three times and the mean values were taken. The primers used in this study were listed in **Table 1**.

## Prediction of miRNA Target Genes

Three miRNAs (ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141) were analyzed using the online tool tarBase v8.0 for potential target genes. The target gene selection threshold was prediction score ≥ 8.0.

## Functional and Pathway Enrichment Analysis

Kyoto Encyclopedia and Gene Ontology (GO) of Genes and Genomes (KEGG) pathway enrichment analyses were performed on target genes using Data Bank for Integrated Discovery (DAVID) (v6.8) Visualization and Annotation.

## Statistical Analysis

Similar experiments were repeated at least three times in this study and expressed using the mean ± standard deviation. The data were subjected to an independent samples t-test and multi-factor ANOVA using Graphpad Prism 6.0.

**TABLE 1** | The sequence of the primers.

| Subject | Primer sequence |
|---|---|
| ebv-miR-BART19-3p | Forward: 5'-CGCGTTTTGTTTGCTTGGG-3'<br>Reverse:5'-AGTGCAGGGTCCGAGGTATT-3' |
| hsa-miR-135b | Forward:5'-CGCGTATGGCTTTTCATTCCT-3'<br>Reverse:5'-GCGCGTAACACTGTCTGGTAA-3' |
| hsa-miR-141 | Forward: 5'-AGTGCAGGGTCCGAGGTATT-3'<br>Reverse: 5'-AGTGCAGGGTCCGAGGTATT-3' |
| U6 | Forward:5'-GACTGCGCAAGGATGACAC -3'<br>Reverse:5'-CAGTGCGTGTCGTGGAGTC-3' |

# RESULTS

## Identification of Differentially Expressed miRNAs

The GEO database was used to obtain the NPC-related miRNA expression dataset GSE36682, which included 62 NPC specimens and six non-cancerous tissue specimens. Next, a total of 121 DEmiRNAs were obtained using $|logFC| \geq 1$ and adjust $P < 0.05$ as screening thresholds (**Figure 1A**). These included 63 expressing upregulated miRNAs and 58 expressing downregulated miRNAs (**Figure 1B**).

## Identification of Gene Co-Expression Modules

To explore the miRNA expression patterns in NPC, we subjected miRNAs in GSE36682 to WGCNA analysis. To ensure a scale-free network, we chose a soft threshold of $\beta = 8$ (**Figure 2A**), generated a hierarchical clustering tree using the WGCNA package as the soft threshold power, and then identified a total of six modules (**Figures 2B–D**). Then, we constructed a co-expression network of associations between clinical features and these modules (**Figure 2E**). Notably, the blue modules were significantly positively correlated with tumor progression. Therefore, the blue module most associated with tumor development status was defined as the SUR module. The scatter plot of correlation between gene module affiliation and gene salience in the blue module is shown in **Figure 2F**.

## Hub miRNAs and the Survival Prognosis Prediction Model

A total of 98 miRNAs were rooted in the blue module and overlapped with 63 variedly upregulated miRNAs, so that a total of 33 hub miRNAs could be obtained (as in **Figure 3A**). Next, we constructed a prognostic signature by using the lasso Cox regression model to analyze the expression levels of hub miRNAs. A predictive 3-miRNA signature model (**Figures 3B, C**) was constructed based on the minimal criterion (Lambda.min = 0.0614) selecting ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141, whose predictive risk scores consisted mainly of the following: Riskscore=(-0.05623)*ebv-miR-BART19-3p+(0.8671)*hsa-miR-135b+(0.1991)*hsa-miR-141.

The sample was divided into high- and low-risk groups according to the risk score ranking, using the best cut-off risk score as the threshold. The results of Kaplan-Meier analysis showed that patients in the high-risk group had a significantly worse prognosis than those in the low-risk group (**Figure 3D**). The model for predicting the overall survival (OS) period of patients was verified by applying ROC curves, and we found that the AUC values of this risk model for predicting prognosis at 1, 3, and 5 years were 0.69, 0.76, and 0.87, respectively, indicating that the model has high accuracy in predicting prognosis survival of patients with nasopharyngeal carcinoma (**Figure 3E**). In addition, we also validated the model using the validation set GSE32960, and the results showed that the prognosis of high-risk patients was significantly worse than that of low-risk patients in the validation set (**Figures 3F, G**).

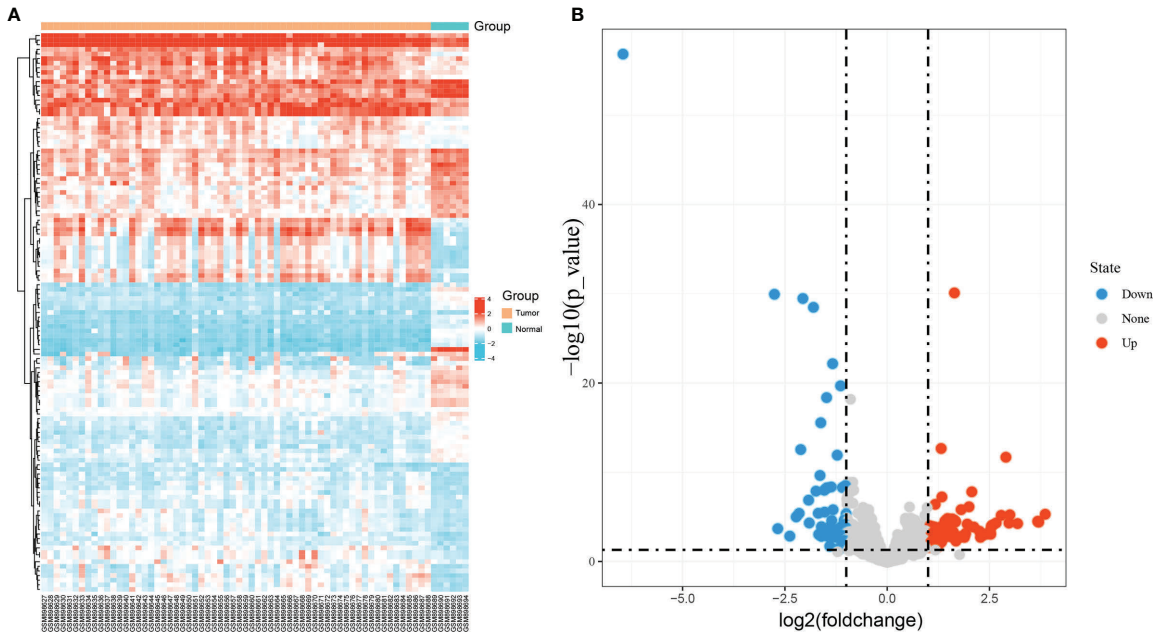



**FIGURE 1** | Identification of differentially expressed miRNAs in GSE36682. **(A)** heat map showing differentially expressed miRNAs; **(B)** volcano map showing differentially expressed miRNAs.

## Hub miRNA Expression and Survival Prognosis

By analyzing the expression of three miRNAs in patients with nasopharyngeal carcinoma, the results showed that ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141 were upgraded in nasopharyngeal carcinoma tissues compared to normal tissues (**Figures 4A–C**). The patients were split into high- and low-expression groups, and it was found that the overall survival of those with high expression of hsa-miR-135b and hsa-miR-141 was significantly shorter than that of patients in the low expression group (**Figures 4E, F**), and the difference was significant ($P < 0.05$); however, the expression level of ebv-miR-BART19-3p was not associated with overall survival, with the difference not being significant ($P > 0.05$, **Figure 4D**).

## qRT-PCR

We detected the expression of ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141 in nasopharyngeal carcinoma cell lines *via* real-time fluorescent quantitative PCR, and the results showed that compared to human nasopharyngeal epithelial cells NP69, ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141 expression levels were significantly higher in human nasopharyngeal cancer SUNE-1 cells ($P < 0.05$, **Figure 5**).

## Predicted Target mRNAs

The target genes of hsa-miR-135b, ebv-miR-BART19-3p, and hsa-miR-141 were predicted by 335, 28, and tarBase v8.0, respectively, with 151 target genes obtained. The targeting relationships are shown in **Figure 6**. Among them, ZFAND5 could be observed as a common target gene for all three.

## Target Gene Function and Pathway Enrichment Analysis

The predicted target genes of ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141 were analyzed through the online analysis website Metascape to obtain PPI protein network interactions and visualize the gene information and network construction, as shown in **Figure 7A**. The KEGG and GO enrichment analyses of the predicted target genes are shown in **Figures 7B, C**. The results of the KEGG analysis revealed that the predicted target genes were mainly enriched in the Hippo signaling pathway, cell cycle, and MicroRNAs in cancer; the results of the GO analysis showed that the predicted target genes were mainly enriched in cellular metabolic processes.

# DISCUSSION

In recent years, targeted therapies, less toxic and more effective chemotherapy modalities, and new technologies have provided new ideas for the treatment of nasopharyngeal carcinoma (18). Among them, gene therapy has received widespread attention for its increasing potential as a novel therapeutic approach (19). Beyond affecting all biological processes, including apoptosis and proliferation, miRNAs are unquestionably the best candidates for therapeutic agents. One of the most important advantages of miRNAs over conventional methods that target a single gene is their ability to target multiple molecules. This makes them very effective in regulating different biological processes related to normal and malignant cell biology (20). In turn, the most critical and significant challenge in the development of miRNA-based

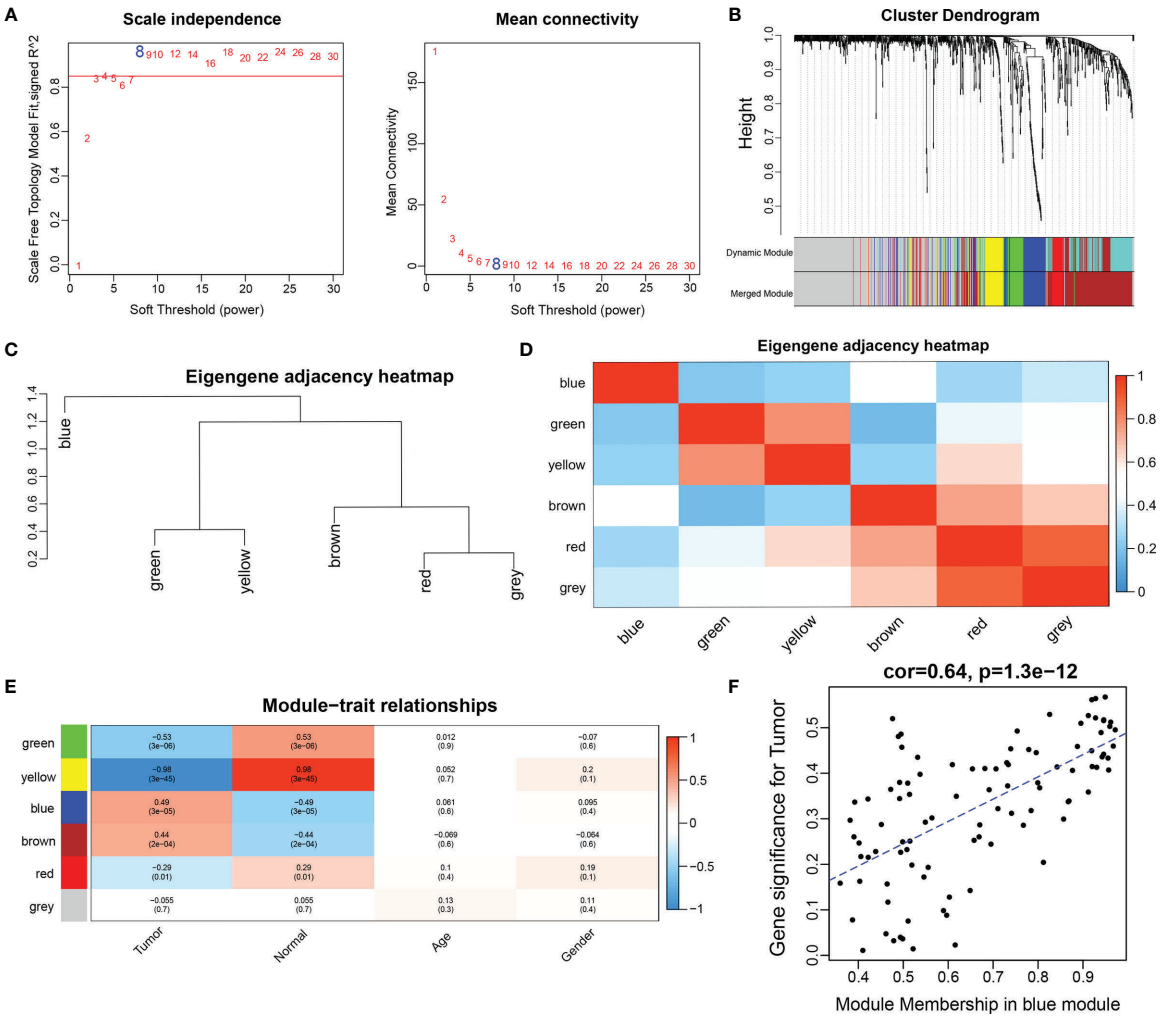



**FIGURE 2** | Construction of weighted gene co-expression network. **(A)** network topographies with different soft threshold powers; **(B)** gene tree diagram obtained by clustering based on the same topological overlap with the corresponding color rows indicating the module colors; **(C)** clustering analysis of each module; **(D)** correlation analysis of each module; **(E)** correlation between each module and clinical information; **(D)** scatter plot of correlation between gene module affiliation and gene significance in blue modules). **(F)** scatter plot of correlation between gene module affiliation and gene significance in blue modules.

therapeutic modalities lies in identifying the best miRNA candidate or miRNA target for each disease. Therefore, the search for new key prognostic biomarkers is an urgent need if we wish to improve the prognosis and clinical treatment of patients with nasopharyngeal carcinoma.

In this research, the GSE36682 dataset was used to construct a co-expression network of clinical features and intermodular associations, i.e., to select modules related to NPC occurrence and development, and ultimately to identify 33 differentially upregulated hub miRNAs *via* the common method of differential expression analysis. Then, a prognostic prediction risk model based on the three miRNA signatures was constructed using lasso regression analysis, and the prognosis prediction of NPC patients at 1, 3, and 5 years was very precise. The three miRNAs, ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141, were highly expressed in nasopharyngeal cancer. hsa-miR-141 and hsa-miR-135b patients with high expression had a significantly worse prognosis than those with low expression, but ebv-miR-BART19-3p expression levels were not significantly correlated with prognosis.

The Epstein-Barr virus (EBV) is one of eight human herpes viruses that infect more than 95% of the world's population (21). EBV persists as a latent infection in the B-lymphatic system throughout life and maintains a well-balanced relationship with humans. Once the fragile EBV-host equilibrium is disrupted, EBV can show its pathogenic potential (22). It is now well known that EBV belongs to one of the mechanisms of nasopharyngeal carcinogenesis, and some EBV genomic products, such as viral proteins, RNAs, and miRNAs, may be involved in nasopharyngeal carcinogenesis (23). EBV also expresses 44 mature microRNAs, including the BamHI fragment H rightward open reading frame 1 (BHRF1) miRNA and the

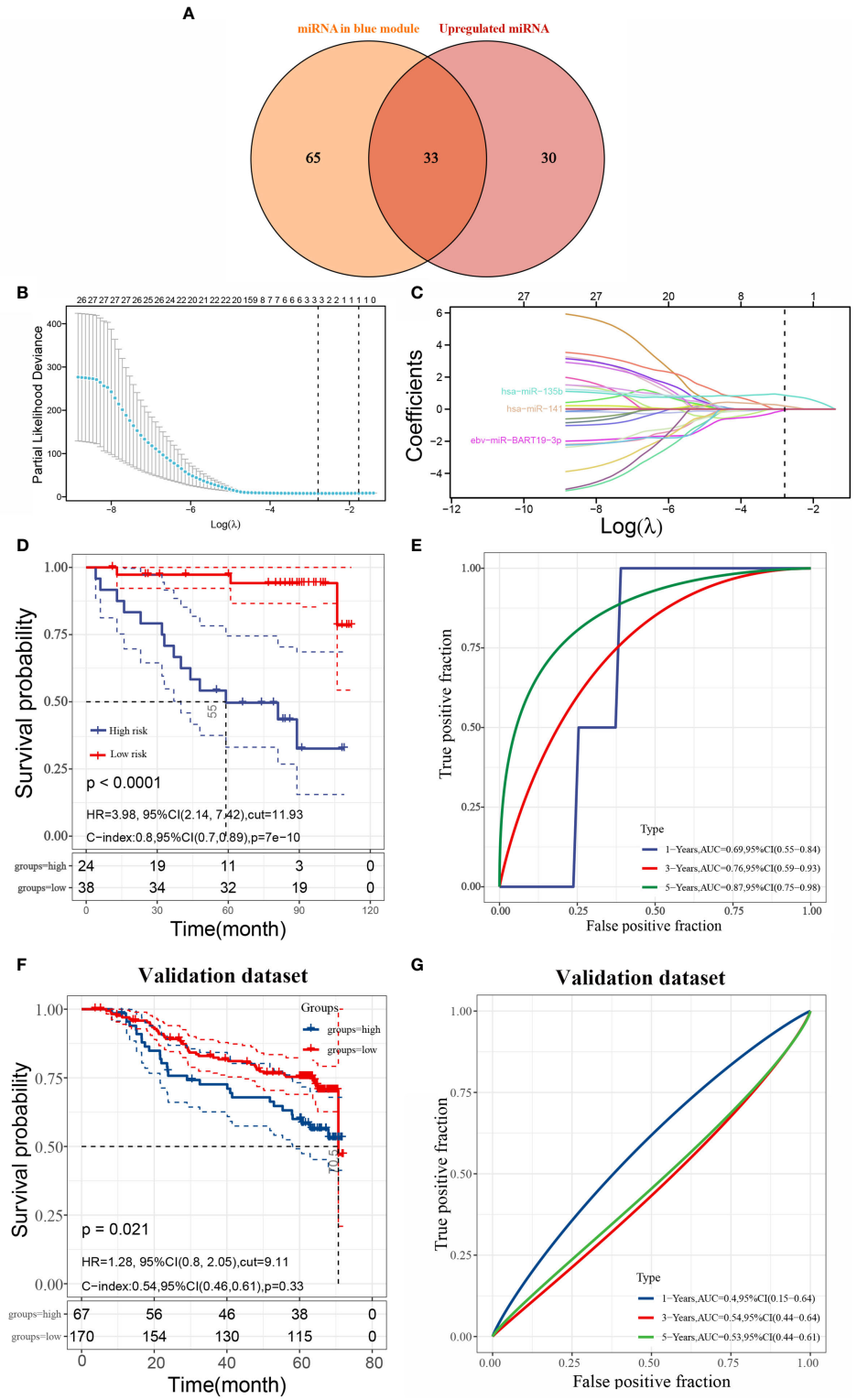



**FIGURE 3** | Identification of hub miRNAs and construction of survival prognosis prediction model **(A)** intersection of differentially upregulated miRNAs and miRNAs from the blue module; **(B)** lasso coefficient screening; **(C)** trajectory plot of lasso variables; **(D)** Kaplan-Meier curves of overall survival (OS) derived from risk scores of three miRNA features in the training set; **(E)** comparison of risk scores of three miRNA markers in the training set for predicting 1-, 3-, and 5-year ROC curves for survival; **(F)** Kaplan-Meier curves for overall survival (OS) derived from the risk scores of the three miRNA signatures in the validation set; **(E)** ROC curves comparing risk scores of three miRNA markers in the validation set for predicting 1-, 3-, and 5-year survival). **(G)** ROC curves comparing risk scores of three miRNA markers in the validation set for predicting 1-, 3-, and 5-year survival.

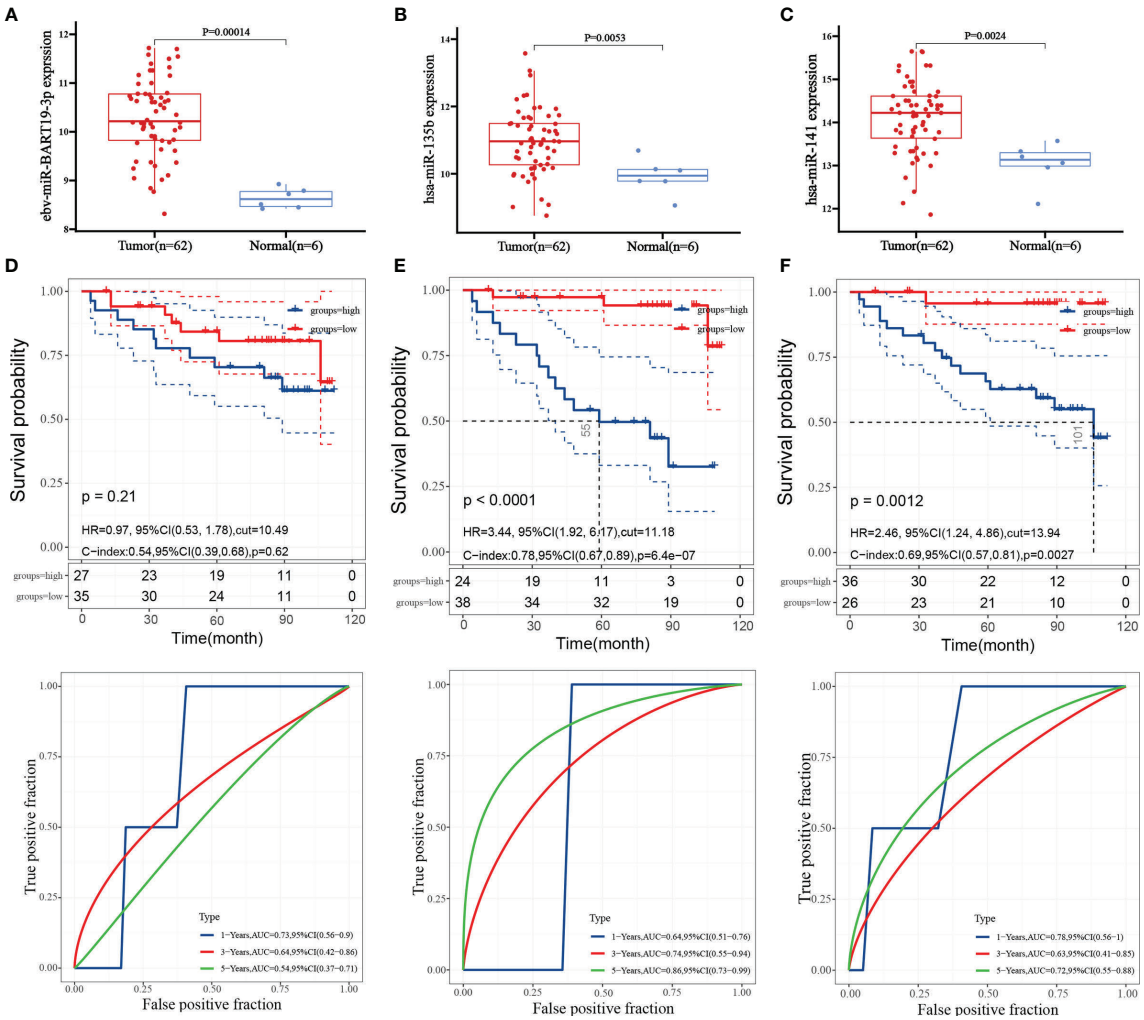



**FIGURE 4** | Hub miRNA expression and survival prognosis. **(A)** expression level of ebv-miR-BART19-3p; **(B)** expression level of hsa-miR-135b; **(C)** expression level of hsa-miR-141; **(D)** Kaplan-Meier survival curve and ROC curve of ebv-miR-BART19-3p; **(E)** Kaplan-Meier survival curve and ROC curve of hsa-miR-135b; **(F)** Kaplan-Meier survival curves and ROC curves for hsa-miR-141.

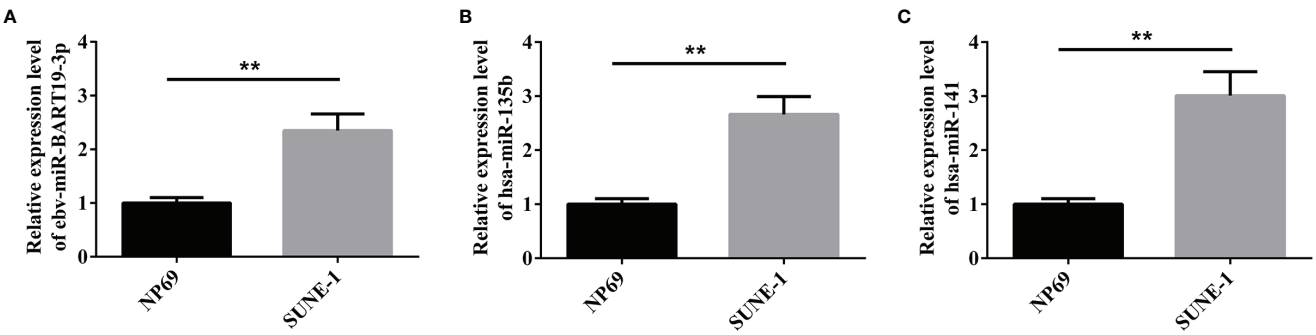



**FIGURE 5** | qPCR to verify the expression of the three miRNAs. **(A)** expression level of ebv-miR-BART19-3p; **(B)** expression level of hsa-miR-135b; **(C)** expression level of hsa-miR-141; **P < 0.01.

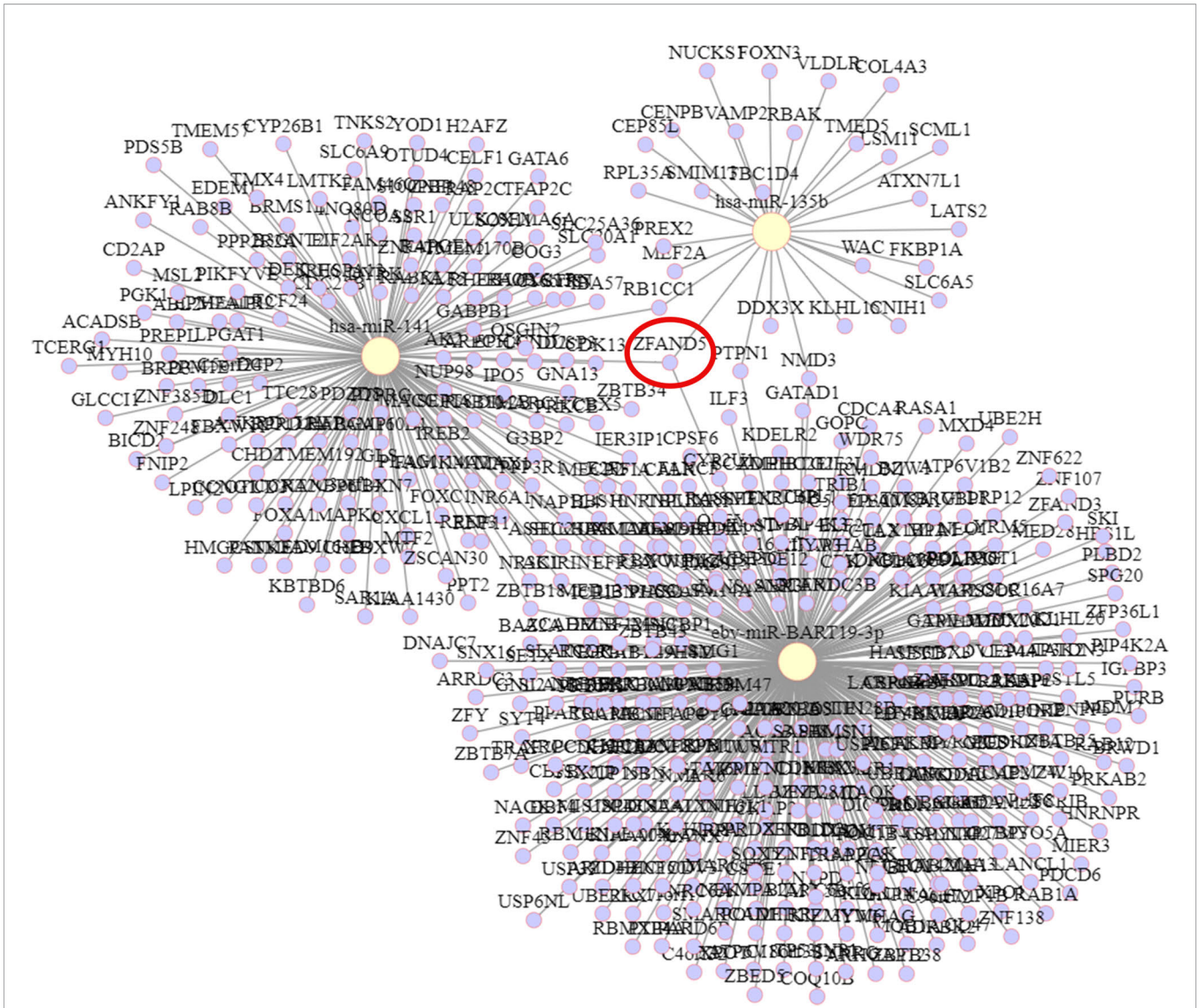



**FIGURE 6** | Predicted target mRNA of miRNA.

BamHI A rightward transcript (BART) miRNAs (24). Among them, BART miRNA is highly expressed in nasopharyngeal carcinoma and EBV-associated gastric cancer and regulates the expression of multiple genes at the post-transcriptional level (25, 26). During NPC development, a minimum of 105 host genes are regulated by EBV miRNAs, which affect five important signaling pathways, including Wnt signaling (27). Zhang et al. (28) constructed the EBV-encoded miRNAs in chronic active EBV infection (CAEBV), EBV-associated phagocytic syndrome (EBV-HLH), and nasopharyngeal carcinoma (NPC). A comprehensive expression profile of EBV-encoded miRNAs in patients showed that miR-BART19-3p was upregulated in all these diseases and suggested that miR-BART19-3p is involved in the tumorigenesis of EBV-related diseases and may be a potential therapeutic target. In the present study, we demonstrated high expression of ebv-miR-BART19-3p in nasopharyngeal carcinoma *via* dataset analysis and fluorescent quantitative PCR analysis. has-miR-135b was first discovered to play a role in somatic stem cell differentiation (29, 30). hsa-miR-135b has been described in the literature as an oncogenic factor in most tumor tissues, such as colon cancer (31), oral cancer (32), and breast cancer (33), while hsa-miR-135b expression was upregulated in metastatic tissues of nasopharyngeal carcinoma (34). Similarly, the results of the present study demonstrated that hsa-miR-135b was highly expressed in nasopharyngeal carcinoma cells. miR-141 is a member of the miR-200 family (35). Dysregulation of miR-200 family members in cancer is associated with growth, apoptotic response, and metastasis regulation (36). However, numerous studies have yielded conflicting results regarding the role of miR-141 in the progression of different tumor types (37, 38), though some previous studies (39) have confirmed that miR-141 is upregulated in nasopharyngeal carcinoma tissues and that

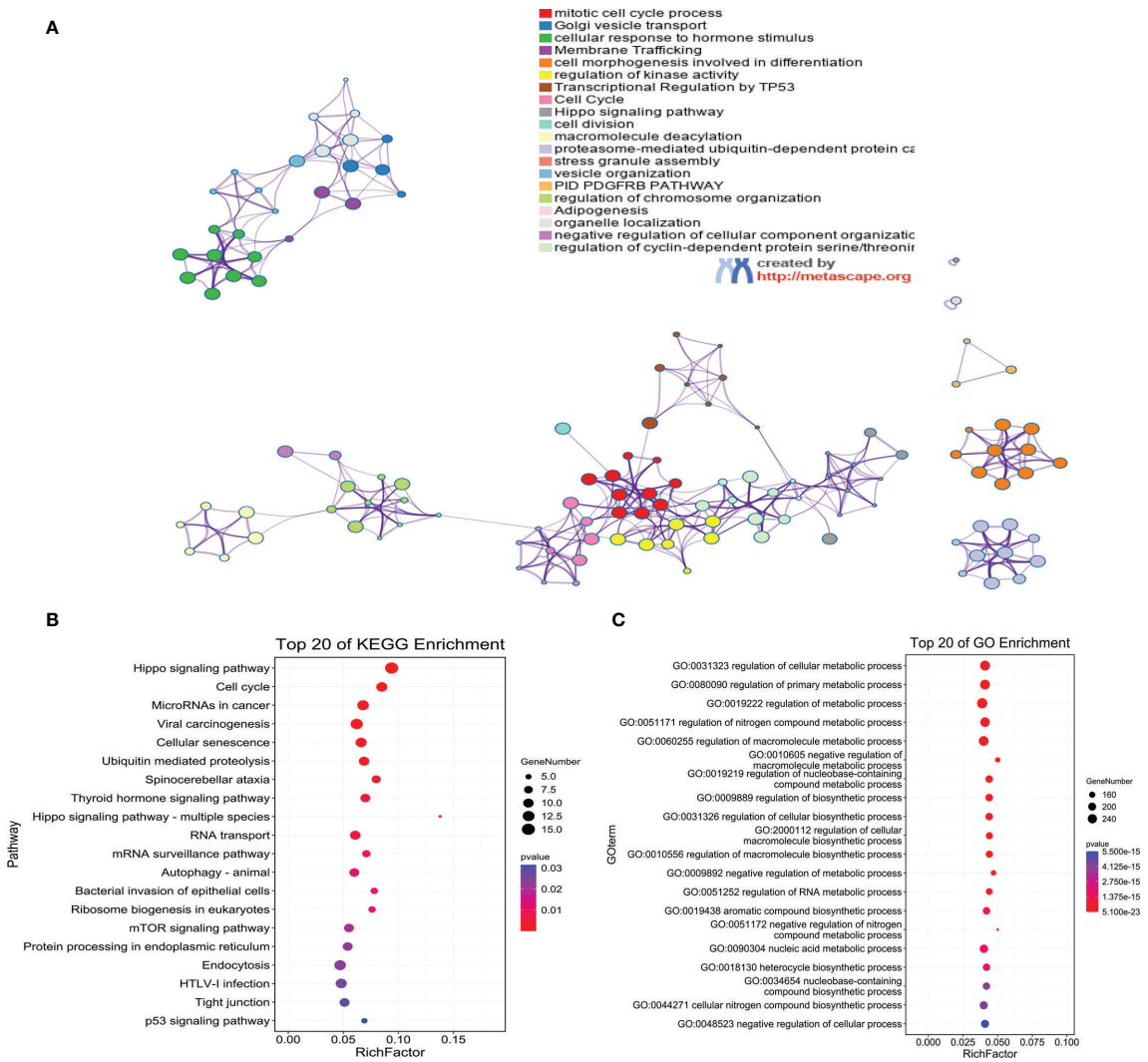



**FIGURE 7** | Functional annotation of target genes. **(A)** PPI network diagram of target genes; **(B)** KEGG enrichment analysis; **(C)** GO enrichment analysis.

high miR-141 levels are negatively correlated with overall survival in nasopharyngeal carcinoma patients. In the present study, we also obtained the same results, where miR-141 was highly expressed in nasopharyngeal carcinoma cells and analysis of the dataset samples revealed that high miR-141 expression was significantly associated with poor prognosis in nasopharyngeal carcinoma patients.

In addition, miRNAs, as important regulators of biological pathways, can regulate target gene expression through translational repression or mRNA degradation and play a key role in tumorigenesis (40, 41). We subsequently predicted the target genes of ebv-miR-BART19-3p, hsa-miR-135b, and hsa-miR-141 and identified ZFAND5 as a common target gene for all three. ZFAND5, a member of the ZFAND family (42), enhances protein degradation by activating the ubiquitin-proteasome system (43). Finally, we performed KEGG and GO functional enrichment analyses of all target genes, and the results showed that genes were mainly enriched in the Hippo signaling pathway, cell cycle, and cellular metabolic processes. Initially, studies on the Hippo signaling pathway mainly focused on the regulation of organ size. Numerous studies have shown that the Hippo signaling pathway inhibits cell growth. In mammals, the upstream membrane protein receptors of the Hippo signaling pathway act as receptors for extracellular growth inhibitory signals and, once sensed, activate a series of kinase cascade phosphorylation reactions that ultimately phosphorylate the downstream effectors' YAP and TAZ. Aberrant regulation of the Hippo pathway has been reported in several cases, including nasopharyngeal carcinoma (44, 45).

Of course, this study is not a perfect one, it still has some limitations. The sample size selected for the training dataset is too small, and we need a larger cohort to test the constructed miRNA signature to prove its robustness. Furthermore, while the prediction results and functional analysis for the target genes suggest that they

are associated with some key signaling pathways, the functional analysis results need further experimental validation.

In summation, our work constructed a prognostic predictive risk model for patients with nasopharyngeal carcinoma based on three miRNA signatures (ebv-miR-BART19-3p, hsa-miR-135b, hsa-miR-141), which can be used to predict the overall survival of patients with nasopharyngeal carcinoma. In addition, the risk model combines human miRNA and EBV. This will provide new ideas for the development of novel therapeutic strategies for the diagnosis and treatment of nasopharyngeal carcinoma. However, experimental validation and detailed biological information analysis are needed for the miRNA downstream mechanism.

## 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

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

## REFERENCES


1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global Cancer Statistics, 2012. *CA: Cancer J Clin* (2015) 65(2):87–108. doi: 10.3322/caac.21262
2. Chen Y-P, Chan AT, Le Q-T, Blanchard P, Sun Y, Ma J. Nasopharyngeal Carcinoma. *Lancet* (2019) 394(10192):64–80. doi: 10.1016/S0140-6736(19)30956-0
3. Lee HM, Okuda KS, González FE, Patel V. Current Perspectives on Nasopharyngeal Carcinoma. *Hum Cell Transform* (2019) 1164:11–34. doi: 10.1007/978-3-030-22254-3_2
4. Mahdavifar N, Towhidi F, Makhsosi BR, Pakzad R, Moini A, Ahmadi A, et al. Incidence and Mortality of Nasopharynx Cancer and Its Relationship With Human Development Index in the World in 2012. *World J Oncol* (2016) 7(5-6):109. doi: 10.14740/wjon980w
5. Hildesheim A, Wang C-P. Genetic Predisposition Factors and Nasopharyngeal Carcinoma Risk: A Review of Epidemiological Association Studies, 2000–2011: Rosetta Stone for NPC: Genetics, Viral Infection, and Other Environmental Factors. *Semin Cancer Biol* (2012) 22(2):107-16. doi: 10.1016/j.semcancer.2012.01.007
6. Cao S-M, Simons MJ, Qian C-N. The Prevalence and Prevention of Nasopharyngeal Carcinoma in China. *Chin J Cancer* (2011) 30(2):114. doi: 10.5732/cjc.010.10377
7. Fang W, Zhang J, Hong S, Zhan J, Chen N, Qin T, et al. EBV-Driven LMP1 and IFN-γ Up-Regulate PD-L1 in Nasopharyngeal Carcinoma: Implications for Oncotargeted Therapy. *Oncotarget* (2014) 5(23):12189. doi: 10.18632/oncotarget.2608
8. Perri F, Scarpati GDV, Caponigro F, Ionna F, Longo F, Buonopane S, et al. Management of Recurrent Nasopharyngeal Carcinoma: Current Perspectives. *OncoTargets Ther* (2019) 12:1583. doi: 10.2147/OTT.S188148
9. Lee AW, Ng WT, Chan JY, Corry J, Mäkitie A, Mendenhall WM, et al. Management of Locally Recurrent Nasopharyngeal Carcinoma. *Cancer Treat Rev* (2019) 79:101890. doi: 10.1016/j.ctrv.2019.101890
10. Lee A, Ma B, Ng WT, Chan A. Management of Nasopharyngeal Carcinoma: Current Practice and Future Perspective. *J Clin Oncol* (2015) 33(29):3356–64. doi: 10.1200/JCO.2015.60.9347
11. Yang K-B, Xu C, Chen Y-P, Vermorken JB, O'Sullivan B, Ma J. Advances in the Pathogenesis and Therapeutic Strategies for Nasopharyngeal Carcinoma. *Front Oncol* (2021) 11:654. doi: 10.3389/fonc.2021.647809
12. Jansson MD, Lund AH. MicroRNA and Cancer. *Mol Oncol* (2012) 6(6):590–610. doi: 10.1016/j.molonc.2012.09.006
13. Jiang C, Li L, Xiang Y-Q, Lung ML, Zeng T, Lu J, et al. Epstein–Barr Virus miRNA BART2-5p Promotes Metastasis of Nasopharyngeal Carcinoma by Suppressing Rnd3. *Cancer Res* (2020) 80(10):1957–69. doi: 10.1158/0008-5472.CAN-19-0334
14. Zhuo X, Zhou W, Li D, Chang A, Wang Y, Wu Y, et al. Plasma microRNA Expression Signature Involving miR-548q, miR-630 and miR-940 as Biomarkers for Nasopharyngeal Carcinoma Detection. *Cancer Biomarkers* (2018) 23(4):579–87. doi: 10.3233/CBM-181852
15. Bruce JP, Hui AB, Shi W, Perez-Ordonez B, Weinreb I, Xu W, et al. Identification of a microRNA Signature Associated With Risk of Distant Metastasis in Nasopharyngeal Carcinoma. *Oncotarget* (2015) 6(6):4537. doi: 10.18632/oncotarget.3005
16. Liu N, Chen N-Y, Cui R-X, Li W-F, Li Y, Wei R-R, et al. Prognostic Value of a microRNA Signature in Nasopharyngeal Carcinoma: A microRNA Expression Analysis. *Lancet Oncol* (2012) 13(6):633–41. doi: 10.1016/S1470-2045(12)70102-X
17. Zhang S, Yue W, Xie Y, Liu L, Li S, Dang W, et al. The Four−microRNA Signature Identified by Bioinformatics Analysis Predicts the Prognosis of Nasopharyngeal Carcinoma Patients. *Oncol Rep* (2019) 42(5):1767–80. doi: 10.3892/or.2019.7316
18. Kang Y, He W, Ren C, Qiao J, Guo Q, Hu J, et al. Advances in Targeted Therapy Mainly Based on Signal Pathways for Nasopharyngeal Carcinoma. *Signal Transduction Targeted Ther* (2020) 5(1):1–20. doi: 10.1038/s41392-020-00394-2
19. Liu T, Li J, Wu X, Zhang S, Lu Z, Li G, et al. Transferrin-Targeting Redox Hyperbranched Poly (Amido Amine)-Functionalized Graphene Oxide for Sensitized Chemotherapy Combined With Gene Therapy to Nasopharyngeal Carcinoma. *Drug Deliv.* (2019) 26(1):744–55. doi: 10.1080/10717544.2019.1642421
20. Ji W, Sun B, Su C. Targeting microRNAs in Cancer Gene Therapy. *Genes* (2017) 8(1):21. doi: 10.3390/genes8010021
21. Wang Y, Guo Z, Shu Y, Zhou H, Wang H, Zhang W. BART miRNAs: An Unimaginable Force in the Development of Nasopharyngeal Carcinoma. *Eur J Cancer Prev* (2017) 26(2):144. doi: 10.1097/CEJ.0000000000000221
22. Taylor GS, Long HM, Brooks JM, Rickinson AB, Hislop AD. The Immunology of Epstein-Barr Virus–Induced Disease. *Annu Rev Immunol* (2015) 33:787–821. doi: 10.1146/annurev-immunol-032414-112326
23. Gu A-D, Zeng M-S, Qian C-N. The Criteria to Confirm the Role of Epstein-Barr Virus in Nasopharyngeal Carcinoma Initiation. *Int J Mol Sci* (2012) 13(10):13737–47. doi: 10.3390/ijms131013737
24. Pfeffer S, Zavolan M, Grässer FA, Chien M, Russo JJ, Ju J, et al. Identification of Virus-Encoded microRNAs. *Science* (2004) 304(5671):734–6. doi: 10.1126/science.1096781
25. Lo AKF, Dawson CW, Jin DY, Lo KW. The Pathological Roles of BART miRNAs in Nasopharyngeal Carcinoma. *J Pathol* (2012) 227(4):392–403. doi: 10.1002/path.4025
26. Marquitz AR, Raab-Traub N. The Role of miRNAs and EBV BARTs in NPC. *Semin Cancer Biol* (2012) 22(2):166-72. doi: 10.1016/j.semcancer.2011.12.001
27. Zeng Z, Huang H, Huang L, Sun M, Yan Q, Song Y, et al. Regulation Network and Expression Profiles of Epstein-Barr Virus-Encoded microRNAs and Their Potential Target Host Genes in Nasopharyngeal Carcinomas. *Sci China Life Sci* (2014) 57(3):315–26. doi: 10.1007/s11427-013-4577-y
28. Zhang Q, Luo D, Xie Z, He H, Duan Z. The Oncogenic Role of miR-BART19-3p in Epstein-Barr Virus-Associated Diseases. *BioMed Res Int* (2020) 2020. doi: 10.1155/2020/5217039
29. Schaap-Oziemlak AM, Raymakers RA, Bergevoet SM, Gilissen C, Jansen BJ, Adema GJ, et al. MicroRNA hsa-miR-135b Regulates Mineralization in

Osteogenic Differentiation of Human Unrestricted Somatic Stem Cells. *Stem Cells Dev* (2010) 19(6):877–85. doi: 10.1089/scd.2009.0112
30. Si Y-J, Ren Q-H, Bi L. miR-135b-5p Regulates Human Mesenchymal Stem Cell Osteogenic Differentiation by Facilitating the Hippo Signaling Pathway. *Int J Clin Exp Pathol* (2017) 10(7):7767.
31. Valeri N, Braconi C, Gasparini P, Murgia C, Lampis A, Paulus-Hock V, et al. MicroRNA-135b Promotes Cancer Progression by Acting as a Downstream Effector of Oncogenic Pathways in Colon Cancer. *Cancer Cell* (2014) 25 (4):469–83. doi: 10.1016/j.ccr.2014.03.006
32. Lopes CB, Magalhães LL, Teófilo CR, Alves APN, Montenegro RC, Negrini M, et al. Differential Expression of hsa-miR-221, hsa-miR-21, hsa-miR-135b, and hsa-miR-29c Suggests a Field Effect in Oral Cancer. *BMC Cancer* (2018) 18 (1):1–10. doi: 10.1186/s12885-018-4631-z
33. Arigoni M, Barutello G, Riccardo F, Ercole E, Cantarella D, Orso F, et al. miR-135b Coordinates Progression of ErbB2-Driven Mammary Carcinomas Through Suppression of MID1 and MTCH2. *Am J Pathol* (2013) 182 (6):2058–70. doi: 10.1016/j.ajpath.2013.02.046
34. Mi J-L, Xu M, Liu C, Wang R-S. Identification of Novel Biomarkers and Small-Molecule Compounds for Nasopharyngeal Carcinoma With Metastasis. *Medi* (2020) 99(32). doi: 10.1097/MD.0000000000021505
35. Chen X, Wang X, Ruan A, Han W, Zhao Y, Lu X, et al. miR-141 Is a Key Regulator of Renal Cell Carcinoma Proliferation and Metastasis by Controlling EphA2 Expression. *Clin Cancer Res* (2014) 20(10):2617–30. doi: 10.1158/1078-0432.CCR-13-3224
36. Zhao G, Wang B, Liu Y, Zhang J-G, Deng S-C, Qin Q, et al. miRNA-141, Downregulated in Pancreatic Cancer, Inhibits Cell Proliferation and Invasion by Directly Targeting MAP4K4. *Mol Cancer Ther* (2013) 12(11):2569–80. doi: 10.1158/1535-7163.MCT-13-0296
37. Van Jaarsveld M, Helleman J, Boersma A, Van Kuijk P, Van Ijcken W, Despierre E, et al. miR-141 Regulates KEAP1 and Modulates Cisplatin Sensitivity in Ovarian Cancer Cells. *Oncogene* (2013) 32(36):4284–93. doi: 10.1038/onc.2012.433
38. 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
39. Liu Y, Zhao R, Wang H, Luo Y, Wang X, Niu W, et al. miR-141 is Involved in BRD7-Mediated Cell Proliferation and Tumor Formation Through Suppression of the PTEN/AKT Pathway in Nasopharyngeal Carcinoma. *Cell Death Dis* (2016) 7(3):e2156–e. doi: 10.1038/cddis.2016.64
40. Zheng Z, Qu J-Q, Yi H-M, Ye X, Huang W, Xiao T, et al. MiR-125b Regulates Proliferation and Apoptosis of Nasopharyngeal Carcinoma by Targeting A20/NF-κb Signaling Pathway. *Cell Death Dis* (2017) 8(6):e2855–e. doi: 10.1038/cddis.2017.211
41. Zuo L, Zhang J, Liu L, Zhou Q, Du S, Xin S, et al. Cadherin 6 is Activated by Epstein–Barr Virus LMP1 to Mediate EMT and Metastasis as an Interplay Node of Multiple Pathways in Nasopharyngeal Carcinoma. *Oncogenesis* (2017) 6(12):1–16. doi: 10.1038/s41389-017-0005-7
42. Dai M, Chen S, Wei X, Zhu X, Lan F, Dai S, et al. Diagnosis, Prognosis and Bioinformatics Analysis of lncRNAs in Hepatocellular Carcinoma. *Oncotarget* (2017) 8(56):95799. doi: 10.18632/oncotarget.21329
43. Lee D, Takayama S, Goldberg AL. ZFAND5/ZNF216 is an Activator of the 26S Proteasome That Stimulates Overall Protein Degradation. *Proc Natl Acad Sci* (2018) 115(41):E9550–E9. doi: 10.1073/pnas.1809934115
44. Zhang Y, Wang M, Xu X, Liu Y, Xiao C. Matrine Promotes Apoptosis in SW480 Colorectal Cancer Cells *via* Elevating MIEF1-Related Mitochondrial Division in a Manner Dependent on LATS2-Hippo Pathway. *J Cell Physiol* (2019) 234(12):22731–41. doi: 10.1002/jcp.28838
45. Zhu Y, He D, Bo H, Liu Z, Xiao M, Xiang L, et al. The MRVI1-AS1/ATF3 Signaling Loop Sensitizes Nasopharyngeal Cancer Cells to Paclitaxel by Regulating the Hippo–TAZ Pathway. *Oncogene* (2019) 38(32):6065–81. doi: 10.1038/s41388-019-0858-7

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 24 February 2022
doi: 10.3389/fonc.2022.838845

# Comprehensive Analysis Identified ASF1B as an Independent Prognostic Factor for HBV-Infected Hepatocellular Carcinoma

*Xianmo Wang[1†], Huawei Yi[1†], Jiancheng Tu[2], Wen Fan[1], Jiahao Wu[1], Li Wang[1], Xiang Li[1], Jinrong Yan[1], Huali Huang[3,4*] and Rong Huang[1*]*

*[1] Clinical Laboratory, The First People's Hospital of Jingzhou, The First Affiliated Hospital of Yangtze University, Jingzhou, China, [2] Clinical Laboratory, The Second Clinical College of Wuhan University, Wuhan, China, [3] Department of Neurology, The Fifth Affiliated Hospital of Guangxi Medical University, Nanning, China, [4] Department of Neurology, The First People's Hospital of Nanning, Nanning, China*

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Han Wu,*
*Eastern Hepatobiliary Surgery Hospital, China*
*Haiqing Ma,*
*Guangdong Provincial People's Hospital, China*

***Correspondence:**
*Huali Huang*
*hualihuang999@Outlook*
*Rong Huang*
*HRZ8890@outlook.com*

*[†]These authors share first authorship*

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

**Received:** *18 December 2021*
**Accepted:** *24 January 2022*
**Published:** *24 February 2022*

**Citation:**
*Wang X, Yi H, Tu J, Fan W, Wu J, Wang L, Li X, Yan J, Huang H and Huang R (2022) Comprehensive Analysis Identified ASF1B as an Independent Prognostic Factor for HBV-Infected Hepatocellular Carcinoma. Front. Oncol. 12:838845. doi: 10.3389/fonc.2022.838845*

**Purpose:** Hepatitis B (HBV)-infected hepatocellular carcinoma is one of the most common cancers, and it has high incidence and mortality rates worldwide. The incidence of hepatocellular carcinoma has been increasing in recent years, and existing treatment modalities do not significantly improve prognosis. Therefore, it is important to find a biomarker that can accurately predict prognosis.

**Methods:** This study was analyzed using the The Cancer Genome Atlas (TCGA) database and validated by the International Cancer Genome Consortium (ICGC) database. The STRING database was used to construct a gene co-expression network and visualize its functional clustering using Cytoscape. A prognostic signature model was constructed to observe high and low risk with prognosis, and independent prognostic factors for HBV-infected hepatocellular carcinoma were identified by Cox regression analysis. The independent prognostic factors were then analyzed for expression and survival, and their pathway enrichment was analyzed using gene set enrichment analysis (GSEA).

**Results:** 805 differentially expressed genes (DEGs) were obtained by differential analysis. Protein–protein interaction (PPI) showed that DEGs were mostly clustered in functional modules, such as cellular matrix response, cell differentiation, and tissue development. Prognostic characterization models showed that the high-risk group was associated with poor prognosis, while Cox regression analysis identified ASF1B as the only independent prognostic factor. As verified by expression and prognosis, ASF1B was highly expressed

in HBV-infected hepatocellular carcinoma and led to a poor prognosis. GSEA showed that high ASF1B expression was involved in cell cycle-related signaling pathways.

**Conclusion:** Bioinformatic analysis identified ASF1B as an independent prognostic factor in HBV-infected hepatocellular carcinoma, and its high expression led to a poor prognosis. Furthermore, it may promote hepatocellular carcinoma progression by affecting cell cycle-related signaling pathways.

**Keywords: lung cancer, HBV, ASF1B, prognosis, ASF1B overexpression**

## INTRODUCTION

Liver cancer is one of the most common cancers worldwide and is the leading cause of cancer-related deaths. The incidence of liver cancer has continued to increase in recent years and has been concentrated to countries with high sociodemographic indices (1). There were more than 900,000 new cases of liver cancer in 2020, with the location of occurrence concentrated to the eastern, southeastern, and northern regions of Asia (2). In addition, there are gender and ethnic differences in the incidence of liver cancer (3). Currently, liver cancer is still very difficult to treat. Although it can be treated by surgical resection, organ transplantation, and chemotherapy, patients' prognoses have not improved greatly in recent years (4). The prognosis prediction of hepatocellular carcinoma relies mostly on clinical phenotype and biochemical features, but their prognostic value is limited. Therefore, the search for a new biomarker to improve diagnostic accuracy and prognostic prediction is particularly important.

Many risk factors are involved in the development of hepatocellular carcinoma. Significant among these risk factors are viral infections and poor lifestyles (5). The vast majority of liver cancer cases can be attributed to chronic viral infections, with hepatitis B virus (HBV) being one of the major types (6). HBV is a small hepatophilic DNA virus that replicates by selectively infecting hepatocytes and undergoing retrotransposition (7). Upon the infection of liver tissue, HBV double-stranded DNA integrates with the human genome and promotes hepatocellular carcinoma progression through oncogenic activation or inactivation (8). The nature of this behavior is to induce genetic instability, affect tumor-associated signaling pathways, and alter the immune response, thereby disrupting normal hepatocyte programming. Chronic infection *via* the HBV virus usually leads to acute or chronic hepatitis and cirrhosis, followed by progressive evolution to hepatocellular carcinoma (9). The HBV vaccination is currently the main way to prevent HBV infection, and it has been effective in reducing the incidence of infection in endemic areas (10). However, the total number of liver cancer deaths due to HBV infection is still increasing (11). This may be because HBV may mediate genetic mutations and tumor heterogeneity in hepatocellular carcinoma tissues, resulting in a different therapeutic sensitivity to chemotherapy than in HBV-negative hepatocellular carcinoma patients (12).

In recent years, research has focused on identifying factors associated with the prognosis of HBV-infected hepatocellular carcinoma. A clinical study showed that a family history of hepatocellular carcinoma was associated with reduced overall prognostic survival (13), several miRNAs were found to predict survival in patients with hepatitis B-related hepatocellular carcinoma (5), and 17 pivotal genes were identified as potential prognostic markers for HBV-infected hepatocellular carcinoma (14). In this study, we cut through patients with liver cancer due to HBV infection to find new biomarkers associated with liver cancer prognosis.

## DATA AND METHODS

### Data Sources

Gene expression matrix and clinical information were obtained from the TCGA database, which yielded 371 cases of liver cancer tissue samples and 50 cases of normal tissue samples following a search. The samples were divided into HBV-positive (145 cases) and HBV-negative (226 cases) groups based on whether the patients were infected with HBV. In addition, the gene expression profiles of hepatocellular carcinoma were downloaded from the International Cancer Genome Consortium (ICGC) database, which included 243 cases of hepatocellular carcinoma tissue samples and 202 cases of paraneoplastic tissues.

### Differential Analysis

The limma package in R was used to study the differential expression of mRNA. The liver cancer samples were divided into HBV infection-positive and HBV-negative groups, and genes with differential expression in the two groups were sought. $|\log2FC| < 0.3785$, $P<0.05$ was defined as the screening threshold for differentially expressed genes (DEGs).

### Protein-Protein Interaction Networks (PPI)

DEGs were imported into the STRING database (https://string-db.org/) to obtain gene-to-gene interaction network relationships. Cytoscape was used to visualize gene interaction network relationship maps and to perform module screening of PPI networks using the MCODE plugin. Metascape (https://metascape.org/gp/index.html#/main/step1) was used to resolve the biological pathway enrichment and gene function annotation of the interaction networks.

### Minimum Absolute Contraction and Selection Operator Analysis

Dimensionality reduction and the construction of prognostic feature models were achieved based on the least absolute shrinkage and selection operator (LASSO). LASSO regression was performed using the R package glmnet to screen potential

prognostic risk features. A risk score formula was obtained from the model, which was used to calculate the risk score for each sample. All patients were divided into high- and low-risk groups based on median values. The R packages survival and survminer were used to analyze survival differences between the two groups, and the R package timeROC was used to analyze the receiver operating characteristic (ROC) analysis performed at 1, 3, and 5 years. Cox risk proportional regression analysis was used to determine the independent prognostic factors in the prognostic model.

## Expression and Survival Analysis

The Wilcoxon rank-sum test was used to analyze the gene expression distribution in the two groups of samples, and mapping was achieved with the R package ggplot2. Survival analysis was performed using the R package surv, and log rank was used to test Kaplan–Meier (KM) survival analysis and to compare the survival differences between the two groups. $P < 0.05$ was considered statistically significant.

## Single Gene Set Enrichment Analysis

Single gene set enrichment analysis (GSEA) was used on the TCGA dataset to identify the biological pathways between the high and low ASF1B expression groups of HBV-infected hepatocellular carcinoma. |NES|>1, NOM p-val<0.05, and FDR q-val<0.25 were defined as screening thresholds based on which the Kyoto Encyclopedia of Genes and Gene (KEGG), and Hallmark signaling pathways were obtained. The top biological processes that were altered were selected according to the normalized enrichment score (NES) ranking.

# RESULTS

## Identification of Differentially Expressed Genes Between HBV-Positive and HBV-Negative Samples

Hepatocellular carcinoma tissue samples from the TCGA database were divided into two groups, HBV infection positive and negative, and the clinical information was compared (**Table 1**). The prognosis of hepatocellular carcinoma associated with positive HBV infection was found to be worse by survival analysis (**Figure 1A**). Therefore, this was used for differential analysis to seek DEGs, and a total of 805 genes were found to be significantly differentially expressed. Of

**TABLE 1** | Comparison of clinical information.

| characteristics | C1 | C2 | P_value |
|---|---|---|---|
| OS | 145 | 226 | 0.002 |
| Gender | 145 | 226 | 0.962 |
| Race | 238 | 373 | 0.57 |
| T_stage | 144 | 225 | 0 |
| N_stage | 144 | 226 | 0.173 |
| M_stage | 145 | 226 | 0.857 |
| TNM_stage | 140 | 207 | 0 |
| Grade | 144 | 222 | 0.081 |
| New_tumor_event_type | 74 | 99 | 0.005 |

*C1 is HBV-positive, C2 is HBV-negative.*

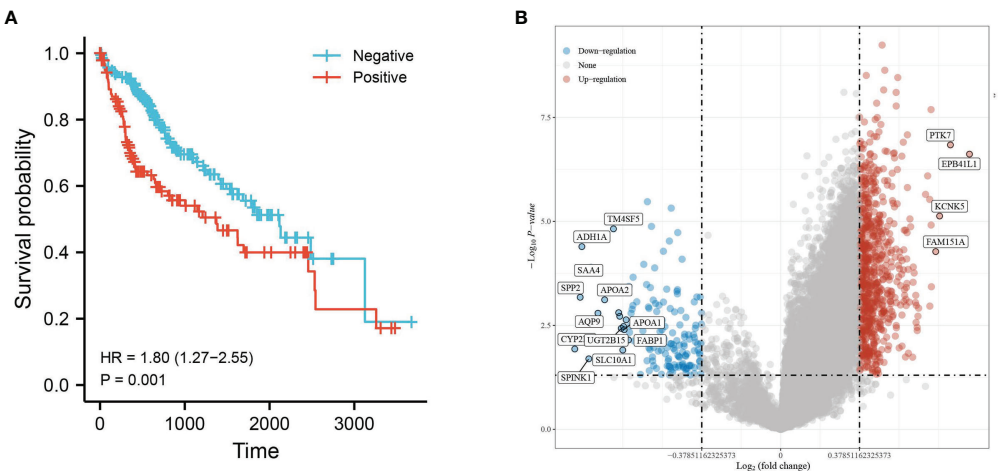



**FIGURE 1** | Differential gene expression analysis of HBV infection-positive and negative hepatocellular carcinomas. **(A)** KM survival curves based on groupings of HBV infections or not and **(B)** volcano plot with upregulated genes in red and downregulated genes in blue.

these, 703 were upregulated genes, and 147 were downregulated genes (**Figure 1B**).

## Identification of Key Functional Gene Modules

The STRING database was used to analyze the gene–gene interactions, Metascape was used to show the enrichment of gene functional terms, and Cytoscape was used to visualize this result. All DEGs have functional linkages between them, and they are classified into different modules by functional clustering (**Figure 2A**). The most representative functional terms or pathways in the top 20 functional clusters are shown in **Table 2**, and these include cell–matrix response, cell differentiation, and tissue development. The gene interaction network map was subdivided into 30 gene modules using the MCODE plugin. The module with the largest score (score: 37.231), which contains 40 nodes and 726 edges (**Figure 2B**). Each node represents a gene, and the higher the score of the module, the more critical and typical the genes are. The genes in this module were selected for subsequent analysis.

## High-Risk Score Is Associated With Poorer Prognosis

The prognostic characteristic risk model was used to study the prognostic effects of 40 genes in key modules on HBV infection related liver cancer samples. The model is a riskscore formula containing multiple genes, and each gene has weight. The analysis showed that the model contained 8 genes. The

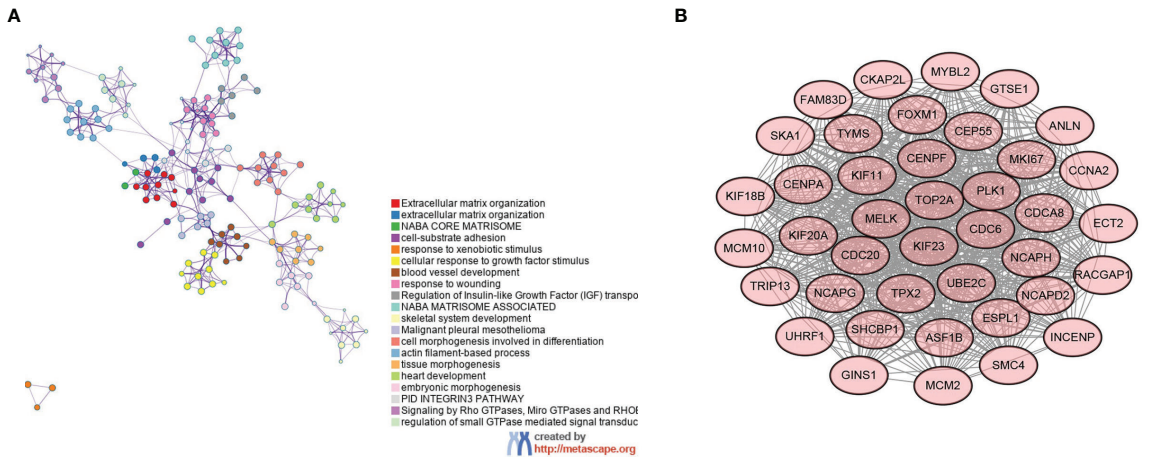



**FIGURE 2** | Gene interaction network diagram and key gene modules. **(A)** Functional interaction network relationship diagram of genes—different colors represent different functional clusters—and **(B)** key gene modules.

**TABLE 2** | Top 20 functional terms.

| Category | Description | P |
|---|---|---|
| Reactome Gene Sets | Extracellular matrix organization | 7.41E-27 |
| GO Biological Processes | extracellular matrix organization | 1.23E-26 |
| Canonical Pathways | NABA CORE MATRISOME | 3.72E-25 |
| GO Biological Processes | cell-substrate adhesion | 1.26E-23 |
| GO Biological Processes | response to xenobiotic stimulus | 2.57E-22 |
| GO Biological Processes | cellular response to growth factor stimulus | 1.02E-21 |
| GO Biological Processes | blood vessel development | 1.29E-21 |
| GO Biological Processes | response to wounding | 1.74E-20 |
| Reactome Gene Sets | Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs) | 9.55E-20 |
| Canonical Pathways | NABA MATRISOME ASSOCIATED | 4.57E-19 |
| GO Biological Processes | skeletal system development | 2.82E-18 |
| WikiPathways | Malignant pleural mesothelioma | 4.27E-18 |
| GO Biological Processes | cell morphogenesis involved in differentiation | 3.47E-16 |
| GO Biological Processes | actin filament-based process | 5.89E-16 |
| GO Biological Processes | tissue morphogenesis | 9.33E-16 |
| GO Biological Processes | heart development | 2.45E-15 |
| GO Biological Processes | embryonic morphogenesis | 2.63E-15 |
| Canonical Pathways | PID INTEGRIN3 PATHWAY | 1.74E-14 |
| Reactome Gene Sets | Signaling by Rho GTPases, Miro GTPases and RHOBTB3 | 9.77E-14 |
| GO Biological Processes | regulation of small GTPase mediated signal transduction | 3.80E-13 |

risk score was calculated using Riskscore = (0.0103)*KIF20 A + (-0.0493)*TYMS + (-0.1866)*ASF1B + (0.1975)*CDCA8 + (0.0657)*CDC20 + (0.1702)*CENPA + (0.0192)*FAM83 D + (0.1364)* TRIP13, which was used for each sample. Using the TCGA dataset as the test set and the ICGC dataset as the validation set, the samples were divided into high- and low-risk groups based on the median cut-off value (cut = 1.2), and **Figure 3A** shows their distribution. The survival of different risk groups for HBV-infected hepatocellular carcinoma is shown in **Figure 3B**, with the high-risk group pointing to a worse prognosis. The predictive power of this prognostic risk model was demonstrated by ROC curves, with more accurate predictions at 1 and 3 years (**Figure 3C**). The model was thereafter validated using the ICGC dataset, and the results were consistent with the test set, showing that the high-risk group was associated with a poorer prognosis (**Figures 3D–F**).

## ASF1B Is an Independent Prognostic Factor for HBV-Infected Hepatocellular Carcinoma

Multi-factor Cox regression analysis was performed using risk score values and each clinical phenotype. Among them, risk score (P = 0), age (P = 0.035), and TNM stage (P = 0.024) were significant, indicating that the above prognostic risk score model is credible (**Figure 4A**). The mountain range plot demonstrates the distribution of the eight prognostic features in the risk score formula for HBV-infected hepatocellular carcinoma, with the dashed line showing the median of each feature (**Figure 4B**). Subsequent univariate and multifactorial Cox regression analyses were performed, and ASF1B was found to be an independent prognostic factor for HBV-infected hepatocellular carcinoma (**Figures 4C, D**).

## High ASF1B Expression Is Associated With a Poor Prognosis in HBV-Infected Hepatocellular Carcinoma

The analysis of samples from the TCGA database revealed that ASF1B was highly expressed in hepatocellular carcinoma tissues, especially HBV-infected hepatocellular carcinoma samples (**Figures 5A, B**). Similarly, samples from the ICGC database yielded the same conclusion, and their high expression was associated with an overall poor prognosis (**Figures 5C, D**). To further understand the prognostic impact of ASF1B, we analyzed the OS, DFS, DSS and PFS based on the TCGA database. The results showed that high ASF1B expression led to a poorer prognosis (**Figures 5E–H**). Meanwhile, ROC plots were performed for each prognosis type to predict prognoses at 1, 3, and 5 years, and the diagnostic accuracy of ASF1B was good in all cases (**Figures 5I–L**).

## ASF1B Is Associated With the Cell Cycle Signaling Pathway

In HBV-positive liver cancer tissues, ASF1B was mainly enriched in signaling pathways related to cell cycle progression

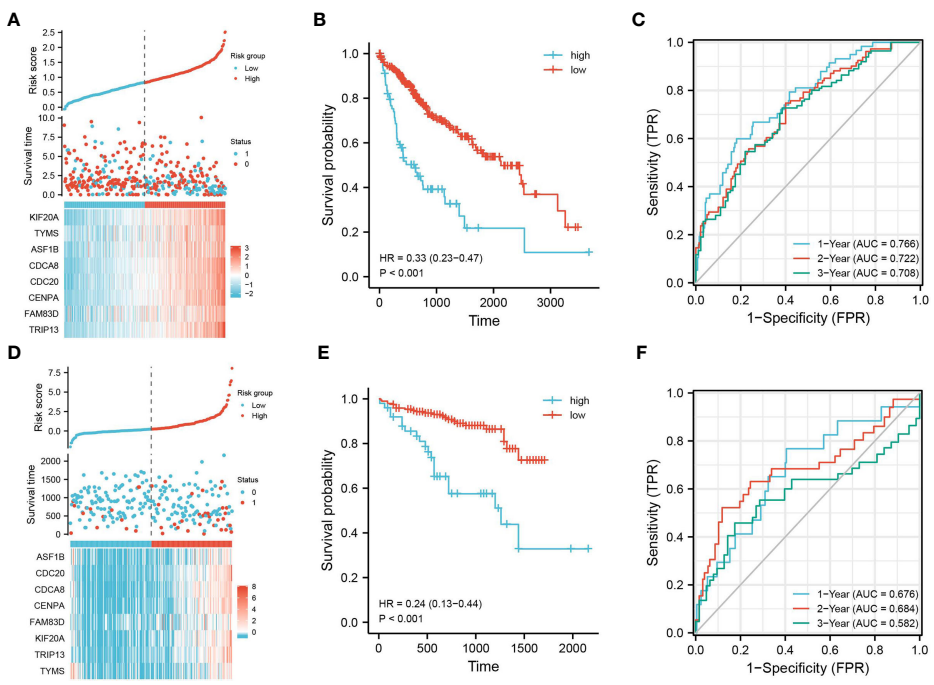



**FIGURE 3** | Construction of prognostic feature model. Using the TCGA dataset as the test set, **(A)** the distribution of high- and low-risk samples; **(B)** KM survival curve; **(C)** time-dependent ROC curve; using the ICGC dataset as the validation set, **(D)** distribution of high- and low-risk samples; **(E)** KM survival curve; and **(F)** time-dependent ROC curve.

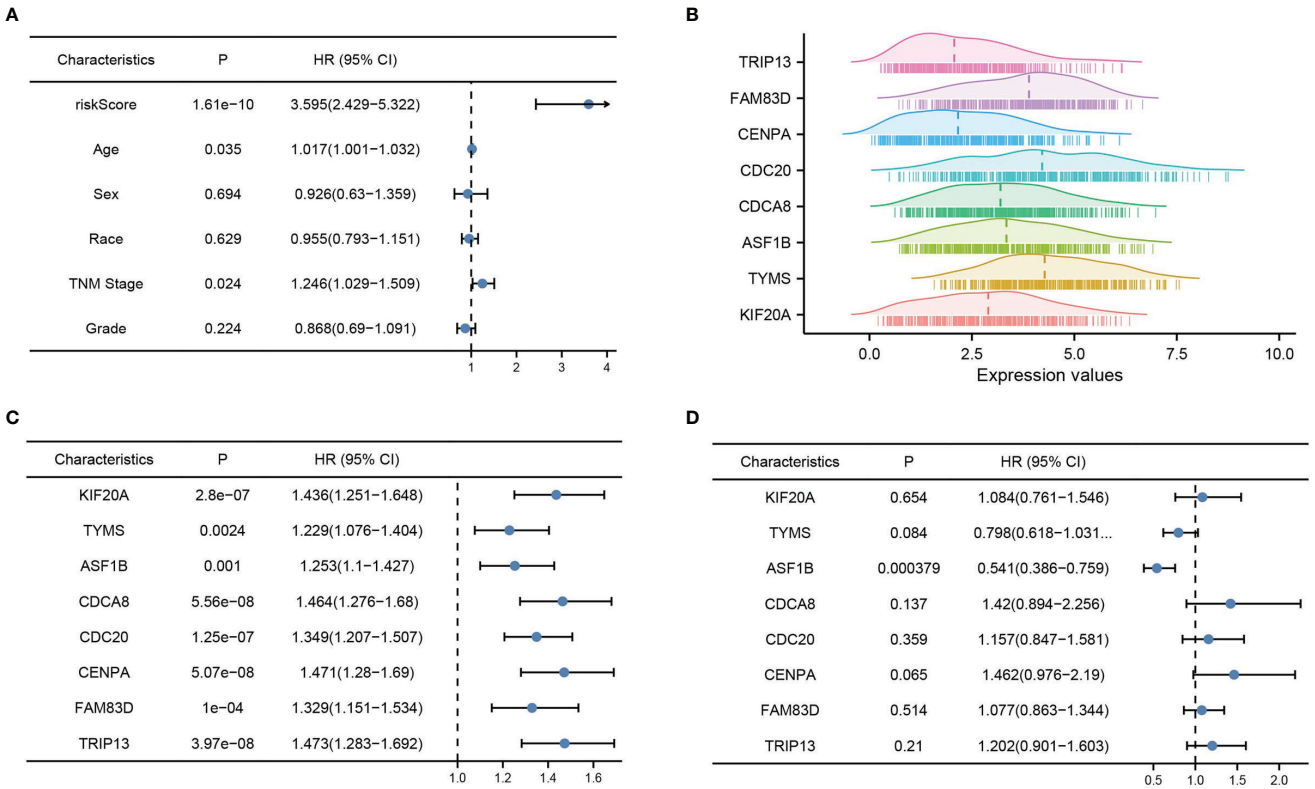



**FIGURE 4** | Risk model validation and independent prognostic factor analysis. **(A)** Multi-factor Cox analysis of risk score and each clinical factor, **(B)** mountain range plot of each factor in the risk score formula, **(C)** one-way Cox analysis of each factor in the risk score formula, and **(D)** multi-way Cox analysis of each factor in the risk score formula. $p < 0.05$ was considered statistically significant.

(**Figures 6A, B**). These pathways involve E2F targets, G2M checkpoints, RNA spliceosome assembly, DNA replication, and other pathways, and their activation is an accelerated cell cycle process that facilitates cancer progression. This suggests that ASF1B may promote the progression of HBV-infected hepatocellular carcinoma by affecting cell cycle-related signaling pathways.

## DISCUSSION

With advances in technology and clinical management, there have been remarkable achievements in our understanding of the pathogenesis of liver cancer and in advances in treatment modalities. However, liver cancer's high morbidity and mortality rates are still of great concern. It is estimated that approximately 240 million people worldwide have been infected with HBV (15). Of these, 70–90% of HBV-infected patients develop hepatocellular carcinoma (16). A high HBV viral load is also an important risk factor for the recurrence of infection after surgery in patients with advanced hepatocellular carcinoma (17). We obtained 850 DEGs by grouping hepatocellular carcinoma patients according to whether they were infected with HBV, and these DEGs were mostly associated with cell differentiation and tissue development. Based on the modeling of prognostic characteristics, we grouped HBV-infected liver cancer samples based on risk scores. The high-risk group in the model was associated with a poor prognosis, and the model had good prognostic predictive power (**Figure 3**). The model included seven pivotal genes associated with prognostic risk and a prognostic independent factor for hepatocellular carcinoma. ASF1B was derived by univariate and multifactorial Cox regression analyses.

Cancer development and progression are associated with the dysregulation of chromatin regulators, including histone variant proteins and histone chaperone proteins (18). ASF1B is an important member of the H3/H4 family of histone chaperone proteins and is mainly involved in cell proliferation (19). ASF1B has a role in regulating the nucleosome structure of chromatin, and its cooperation with chromatin assembly factor 1 (CAF-1) promotes replication-dependent chromatin assembly (20). ASF1B has been identified early in the disease progression of breast cancer. Corpet A et al. noted that overexpression of ASF1B correlated with clinical data and disease outcomes in breast cancer, meaning that ASF1B has diagnostic and prognostic value (21). In recent years, the role of ASF1B in the

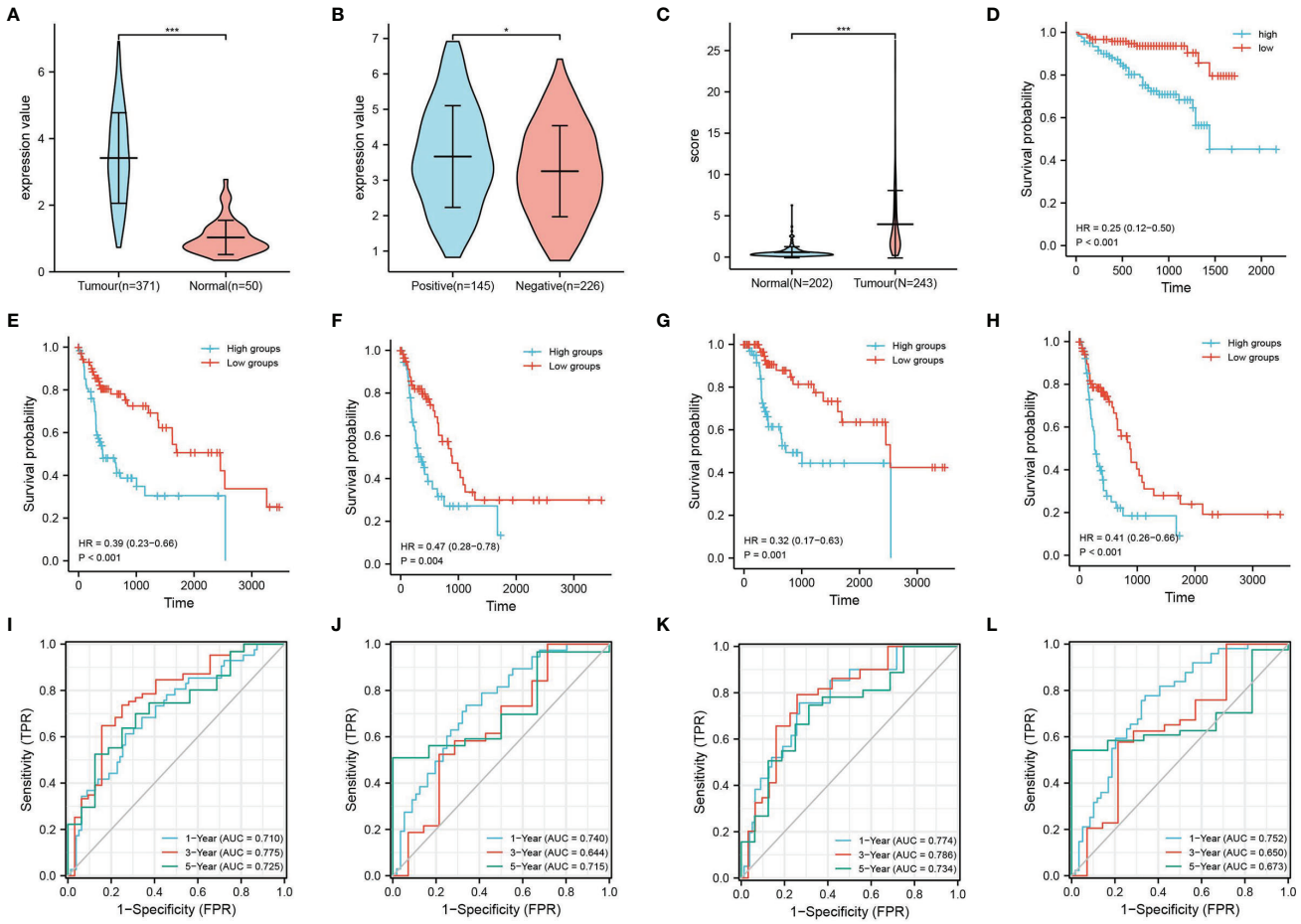



**FIGURE 5** | Expression and survival analysis of ASF1B. TCGA database: **(A)** ASF1B expression in liver cancer tissues and normal tissue samples and **(B)** ASF1B expression in HBV-positive and HBV-negative liver cancer samples. ICGC database: **(C)** ASF1B expression in liver cancer tissues and normal tissue samples; **(D)** survival analysis of ASF1B in primary liver cancer; **(E–H)** are the KM survival curves for overall survival (OS), progression-free survival (PFS), disease-free survival (DFS), and disease-specific survival (DSS) of the KM survival curves; and **(I–L)** ROC curves of OS, DFS, DSS, and PFS, respectively. *P < 0.05, ***P < 0.001.

progression of other cancers has been explored; high ASF1B expression has pointed to a poor prognosis for lung adenocarcinoma and has been associated with advanced tumor stage and tumor progression (22). The expression of ASF1B can also promote multiple myeloma progression (23). In addition, ASF1B expression has been associated with cancer-related pathways. In prostate cancer, ASF1B promotes cancer progression by affecting the PI3K/AKT signaling pathway (24). In cervical cancer, ASF1B promotes cell invasion and affects prognosis by activating the Wnt/β-Catenin signaling pathway (25). However, its role in HBV-infected hepatocellular carcinoma has not been effectively validated. In this study, we analyzed the prognostic significance of ASF1B expression on OS, DFS, DSS, and PFS in HBV-infected hepatocellular carcinoma patients. The results showed that high ASF1B expression led to poorer overall survival and the potential to shorten the interval for tumor progression, tumor recurrence, and death. GSEA showed that high ASF1B expression exhibited great significance regarding pathways associated with cell cycle progression.

Cancer is essentially a group of diseases with persistent excessive cell division tightly regulated by cell cycle control mechanisms (26). The G1 and S phases are two very important periods in the cell cycle, and their changes affect cell proliferation (27). It has been demonstrated that ASF1B is a G1- and S-phase regulator, and the knockdown of ASF1B expression induces G1-phase cell cycle arrest (24). Interestingly, HBV can dysregulate cellular signaling pathways, such as Wnt/FZD/β-catenin, PI3K/Akt/mTOR, and Ras/Raf/MAPK, associated with liver cancer development (28). In this study, we found that the high expression of ASF1B in HBV-infected hepatocellular carcinoma patients was mostly associated with cell cycle-related pathways. However, we mostly focused on functions such as DNA replication, mitosis,

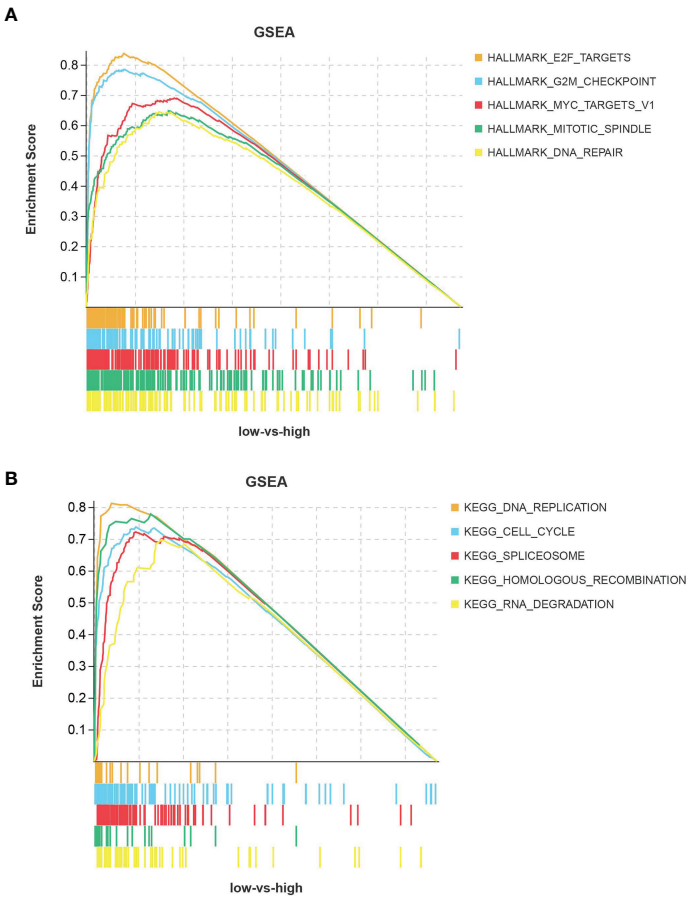



**FIGURE 6** | Pathway enrichment analysis of ASF1B. **(A)** Hallmark pathway and **(B)** KEGG pathway.

etc. ASF1B was previously identified in human islet β-cell replication, and it significantly increased islet β-cell proliferation (29). In addition, ASF1B overexpression promotes melanoma cell growth and adhesion and inhibits apoptosis (30). Furthermore, ASF1B has been shown to promote the proliferation and invasion of clear cell renal cell carcinoma and gastric cancer cells through the activation of the AKT and Bax/Bcl-2-p53 pathways (31, 32).

In conclusion, the present study is the first to analyze ASF1B based on HBV-infected liver cancer tissues. The results showed that ASF1B was highly expressed in HBV-infected compared to non-HBV-infected hepatocellular carcinoma tissues. An association between high ASF1B expression and clinical prognosis was also revealed. Our study showed that ASF1B is an independent prognostic factor in HBV-infected hepatocellular carcinoma and that its high expression leads to a poor prognosis in multiple survival types. In addition, GSEA showed that ASF1B is associated with cell cycle pathways in the progression of HBV-infected hepatocellular carcinoma. These studies contribute to the understanding of the protumor role of ASF1B, but its specific mechanisms require further validation.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## AUTHOR CONTRIBUTIONS

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

## FUNDING

This work was supported by 2015 Jingzhou Science and Technology Plan (Medical and Health) under grant 2015-04 and Central Funds Guiding the Local Science and Technology Development of Hubei Province under grant 2019ZYYD066.

## REFERENCES


1. Liu Z, Jiang Y, Yuan H, Fang Q, Cai N, Suo C, et al. The Trends in Incidence of Primary Liver Cancer Caused by Specific Etiologies: Results From the Global Burden of Disease Study 2016 and Implications for Liver Cancer Prevention. *J Hepatol* (2019) 70(4):674–83. doi: 10.1016/j.jhep.2018.12.001
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
3. Petrick JL, McGlynn KA. The Changing Epidemiology of Primary Liver Cancer. *Curr Epidemiol Rep* (2019) 6(2):104–11. doi: 10.1007/s40471-019-00188-3
4. Jiao Y, Li Y, Jiang P, Han W, Liu Y. PGM5: A Novel Diagnostic and Prognostic Biomarker for Liver Cancer. *PeerJ* (2019) 7:e7070. doi: 10.7717/peerj.7070
5. Zhen Y, Xinghui Z, Chao W, Yi Z, Jinwen C, Ruifang G, et al. Several microRNAs Could Predict Survival in Patients With Hepatitis B-Related Liver Cancer. *Sci Rep* (2017) 7(1):1–8. doi: 10.1038/srep45195
6. de Martel C, Maucort-Boulch D, Plummer M, Franceschi S. World-Wide Relative Contribution of Hepatitis B and C Viruses in Hepatocellular Carcinoma. *Hepatology* (2015) 62(4):1190–200. doi: 10.1002/hep.27969
7. Feng J, Yang G, Liu Y, Gao Y, Zhao M, Bu Y, et al. LncRNA PCNAP1 Modulates Hepatitis B Virus Replication and Enhances Tumor Growth of Liver Cancer. *Theranostics* (2019) 9(18):5227. doi: 10.7150/thno.34273
8. Furuta M, Tanaka H, Shiraishi Y, Unida T, Imamura M, Fujimoto A, et al. Characterization of HBV Integration Patterns and Timing in Liver Cancer and HBV-Infected Livers. *Oncotarget* (2018) 9(38):25075. doi: 10.18632/oncotarget.25308
9. Schweitzer A, Horn J, Mikolajczyk RT, Krause G, Ott JJ. Estimations of Worldwide Prevalence of Chronic Hepatitis B Virus Infection: A Systematic Review of Data Published Between 1965 and 2013. *Lancet* (2015) 386 (10003):1546–55. doi: 10.1016/S0140-6736(15)61412-X
10. Nelson NP, Easterbrook PJ, McMahon BJ. Epidemiology of Hepatitis B Virus Infection and Impact of Vaccination on Disease. *Clinics Liver Dis* (2016) 20 (4):607–28. doi: 10.1016/j.cld.2016.06.006
11. Choi J, Han S, Kim N, Lim YS. Increasing Burden of Liver Cancer Despite Extensive Use of Antiviral Agents in a Hepatitis B Virus-Endemic Population. *Hepatology* (2017) 66(5):1454–63. doi: 10.1002/hep.29321
12. Liu DX, Li PP, Guo JP, Li LL, Guo B, Jiao HB, et al. Exosomes Derived From HBV–Associated Liver Cancer Promote Chemoresistance by Upregulating Chaperone–Mediated Autophagy. *Oncol Lett* (2019) 17(1):323–31. doi: 10.3892/ol.2018.9584
13. Li Z-L, Han J, Liu K, Xing H, Wu H, Lau WY, et al. Association of Family History With Long-Term Prognosis in Patients Undergoing Liver Resection of HBV-Related Hepatocellular Carcinoma. *Hepatobil Surg Nutr* (2019) 8(2):88. doi: 10.21037/hbsn.2018.11.20
14. Tang Y, Zhang Y, Hu X. Identification of Potential Hub Genes Related to Diagnosis and Prognosis of Hepatitis B Virus-Related Hepatocellular Carcinoma *via* Integrated Bioinformatics Analysis. *BioMed Res Int* (2020) 2020. doi: 10.1155/2020/4251761
15. Balogh J, Victor IIID, Asham EH, Burroughs SG, Boktour M, Saharia A, et al. Hepatocellular Carcinoma: A Review. *J Hepatocell Carcinoma* (2016) 3:41. doi: 10.2147/JHC.S61146
16. Shiani A, Narayanan S, Pena L, Friedman M. The Role of Diagnosis and Treatment of Underlying Liver Disease for the Prognosis of Primary Liver Cancer. *Cancer Control* (2017) 24(3):1073274817729240. doi: 10.1177/1073274817729240
17. Lee I-C, Lei H-J, Chau G-Y, Yeh Y-C, Wu C-J, Su C-W, et al. Predictors of Long-Term Recurrence and Survival After Resection of HBV-Related Hepatocellular Carcinoma: The Role of HBsAg. *Am J Cancer Res* (2021) 11 (7):3711.
18. Gurard-Levin ZA, Quivy J-P, Almouzni G. Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics. *Annu Rev Biochem* (2014) 83:487–517. doi: 10.1146/annurev-biochem-060713-035536
19. Paul PK, Rabaglia ME, Wang C-Y, Stapleton DS, Leng N, Kendziorski C, et al. Histone Chaperone ASF1B Promotes Human β-Cell Proliferation *via* Recruitment of Histone H3. 3. *Cell Cycle* (2016) 15(23):3191–202. doi: 10.1080/15384101.2016.1241914
20. Papadopoulos P, Kafasi A, De Cuyper IM, Barroca V, Lewandowski D, Kadri Z, et al. Mild Dyserythropoiesis and β-Like Globin Gene Expression Imbalance Due to the Loss of Histone Chaperone ASF1B. *Hum Genomics* (2020) 14(1):1–12. doi: 10.1186/s40246-020-00283-3
21. Corpet A, De Koning L, Toedling J, Savignoni A, Berger F, Lemaître C, et al. Asf1b, the Necessary Asf1 Isoform for Proliferation, Is Predictive of Outcome in Breast Cancer. *EMBO J* (2011) 30(3):480–93. doi: 10.1038/emboj.2010.335
22. Feng Z, Zhang J, Zheng Y, Wang Q, Min X, Tian T. Elevated Expression of ASF1B Correlates With Poor Prognosis in Human Lung Adenocarcinoma. *Personalized Med* (2021) 18(2):115–27. doi: 10.2217/pme-2020-0112
23. Wang C, Li M, Wang S, Jiang Z, Liu Y. LINC00665 Promotes the Progression of Multiple Myeloma by Adsorbing miR-214-3p and Positively Regulating the Expression of PSMD10 and ASF1B. *OncoTargets Ther* (2020) 13:6511. doi: 10.2147/OTT.S241627
24. Han G, Zhang X, Liu P, Yu Q, Li Z, Yu Q, et al. Knockdown of Anti-Silencing Function 1B Histone Chaperone Induces Cell Apoptosis *via* Repressing PI3K/Akt Pathway in Prostate Cancer. *Int J Oncol* (2018) 53(5):2056–66. doi: 10.3892/ijo.2018.4526
25. Zhang W, Li H, Sun X, Shi Y, Yang Y, Xie C, et al. Anti-Silencing Function 1b Overexpression Affects Prognosis and Promotes Invasion in Cervical Carcinoma *via* Activation of Wnt/β-Catenin Signaling Pathway. (2021). doi: 10.21203/rs.3.rs-345176/v1
26. Matthews HK, Bertoli C, de Bruin RA. Cell Cycle Control in Cancer. *Nat Rev Mol Cell Biol* (2021), 1–15. doi: 10.1038/s41580-021-00404-3
27. Bai J, Li Y, Zhang G. Cell Cycle Regulation and Anticancer Drug Discovery. *Cancer Biol Med* (2017) 14(4):348. doi: 10.20892/j.issn.2095-3941.2017.0033
28. Torresi J, Tran BM, Christiansen D, Earnest-Silveira L, Schwab RHM, Vincan E. HBV-Related Hepatocarcinogenesis: The Role of Signalling Pathways and Innovative *Ex Vivo* Research Models. *BMC Cancer* (2019) 19(1):1–14. doi: 10.1186/s12885-019-5916-6
29. Banerjee RR, Bhatnagar S. *ASF1B Chaperones Histone 3.3 to the β-Cell Cycle Dance*. Taylor & Francis (2017). doi: 10.1080/15384101.2016.1260989
30. Shi X, Xu X, Shi N, Chen Y, Fu M. MicroRNA-520d-3p Suppresses Melanoma Cells Proliferation by Inhibiting the Anti-Silencing Function 1B Histone Chaperone. *Bioengineered* (2021) 12(2):10703–15. doi: 10.1080/21655979.2021.2001914
31. Jiangqiao Z, Tao Q, Zhongbao C, Xiaoxiong M, Long Z, Jilin Z, et al. Anti-Silencing Function 1B Histone Chaperone Promotes Cell Proliferation and Migration *via* Activation of the AKT Pathway in Clear Cell Renal Cell Carcinoma. *Biochem Biophys Res Commun* (2019) 511(1):165–72. doi: 10.1016/j.bbrc.2019.02.060
32. Li M, Huang L, Xiao Y, Wang C, Li Q, Qin C. Elevated ASF1B Promotes Apoptosis in Gastric Tumour Cells Through the Bax/Bcl-2-P53 Pathway and Is Related to a Good Prognosis. (2021). doi: 10.21203/rs.3.rs-929676/v1


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 24 February 2022
doi: 10.3389/fonc.2022.813809

# COL5A2 Inhibits the TGF-β and Wnt/β-Catenin Signaling Pathways to Inhibit the Invasion and Metastasis of Osteosarcoma

*Yan-Long Han[1], Dan Luo[2], Kakeng Habaxi[1], Julaiti Tayierjiang[1], Wei Zhao[1], Wei Wang[1], Wumaierjiang Aikebaier[1] and Li Wang[1]**

[1] Department of Bone and Joint Surgery Center, People's Hospital of Xinjiang Uygur Autonomous Region, Urumqi, China,
[2] Department of Respiratory and Critical Care Medicine, People's Hospital of Xinjiang Uygur Autonomous Region, Urumqi, China

OPEN ACCESS

**Edited by:**
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

**Reviewed by:**
Mu He,
Michigan State University, United States
Yucheng Zeng,
University of Cincinnati, United States

***Correspondence:**
Li Wang
hanyanlong777@163.com

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

**Received:** 12 November 2021
**Accepted:** 18 January 2022
**Published:** 24 February 2022

**Citation:**
Han Y-L, Luo D, Habaxi K, Tayierjiang J, Zhao W, Wang W, Aikebaier W and Wang L (2022) COL5A2 Inhibits the TGF-β and Wnt/β-Catenin Signaling Pathways to Inhibit the Invasion and Metastasis of Osteosarcoma. Front. Oncol. 12:813809. doi: 10.3389/fonc.2022.813809

Osteosarcoma is the most common skeletal malignancy and is the second leading cause of cancer death in adolescents. Its highly aggressive nature and high propensity to metastasize lead to an extremely poor prognosis for patients with osteosarcoma. Therefore, finding a suitable treatment has become a matter of urgency. In this study, we first divided the samples into metastatic and non-metastatic groups using the Target database and obtained 1136 differentially expressed genes (DEGs) using differential analysis. A PPI network was constructed to analyze the network of action relationships among DEGs, and the top 10 genes were derived using the MCC algorithm in Cytoscape software. A risk scoring system for 10 key genes was constructed using the LASSO-COX prognostic risk model, and genes associated with osteosarcoma prognosis were screened based on multifactorial COX. COL5A2 gene was highly expressed in metastatic osteosarcoma and led to a poor prognosis. Furthermore, qRT-PCR and immunofluorescence assays confirmed the high expression of COL5A2 in human osteosarcoma cells. CCK-8 assay and scratch WB was used to determine whether the downregulation of COL5A2 expression inhibits the TGF-β signaling and Wnt/β-Catenin signaling pathways. In this study, we screened COL5A2 for prognostic relevance to osteosarcoma through bioinformatics analysis and demonstrated that COL5A2 inhibited osteosarcoma invasion and metastasis by suppressing the TGF-β signaling and Wnt/β-Catenin signaling pathways.

**Keywords: osteosarcoma, COL5A2, TGF-β, Wnt/β-catenin, tumor metastasis**

## INTRODUCTION

Osteosarcoma (OS) is the most common primary skeletal malignancy (1). Both adolescents and older adults have a peak incidence, which results in a bimodal distribution of osteosarcoma incidence (2). Various factors contribute to the increased incidence of osteosarcomas, such as race and gender, but younger age is the most common risk factor (3). Osteosarcoma has been reported to

account for over 50% of malignant bone tumors in children and is the second leading cause of cancer death in adolescents (4). The commonly used treatment in the past was surgical resection, but this approach did not improve the survival of patients with metastatic osteosarcoma (5). Modern combination treatments such as intensive chemotherapy have now led to a 35%–50% increase in five-year survival in patients with localized disease (6, 7). However, the five-year survival rate remains low for the patient population with metastatic lesions (8). The main reasons for the low long-term survival of patients with osteosarcoma are the highly aggressive and early systemic metastatic nature of the osteosarcoma (9). Approximately 20% of patients are identified as metastatic at diagnosis, and less than 15% of patients without metastatic disease are treated for five-year survival (10). In contrast, pulmonary metastases are the most prominent type of distant metastasis in osteosarcoma, with a five-year survival rate of no more than 30% (2).

Distant metastasis is a predictor of poor tumor prognosis, and specific metabolic pathways can promote secondary tumor formation through the induction of signaling pathways (11). During the treatment of cancer patients, tumor progression was found to be accelerated when treatment was stopped and the epithelial-mesenchymal transition (EMT) process was induced (12). EMT is a plastic cellular process that is associated with cancer cell invasion and metastasis (13). Moreover, EMT is a hallmark of metastatic cancer; there are multiple important signaling pathways involved in accelerating the EMT process and tumor metastasis, including the transforming growth factor-β1 (TGF-β) signaling and Wnt/β-Catenin signaling pathways (14). TGF-β is a secreted cytokine that plays an important role in EMT, cell biological progression, etc. (15). TGF-β acts as a tumor suppressor in the early stage of tumorigenesis, while the subsequent TGF-β signaling pathway can effectively induce the EMT process, which makes the cells migratory and invasive and, thus, promote cancer metastasis (16). Similarly, the Wnt/β-Catenin signaling pathway drives tumor initiation and progression and is required for biological processes, such as cell proliferation, migration, invasion, and angiogenesis (17). Furthermore, β-Catenin translocation from the cytoplasm to the nucleus activates the target genes MMP7 and C-MYC, thereby playing an active role in tumor metastasis (18). Moreover, the Wnt/β-Catenin signaling pathway has been shown to promote tumor metastasis by regulating the EMT process in thyroid and pancreatic cancers (19, 20).

Genes have been reported to promote osteosarcoma progression through the TGF-β signaling pathway (21) or Wnt/β-Catenin signaling pathway (22) and are associated with the prognosis of osteosarcoma patients. This study aimed to identify new prognostic key factors associated with tumor metastasis for osteosarcoma. The samples were divided into metastatic and non-metastatic groups, and genes associated with osteosarcoma prognosis were identified through differential analysis, PPI, and LASSO-COX analysis. Subsequently, *in vitro* experiments were used to validate the expression and explore the biological functions of the key and further investigate the mechanisms of key genes on osteosarcoma metastasis.

# MATERIALS AND METHODS

## Data Source and Processing

Clinical samples and RNA-seq data of 89 cases of osteosarcoma were obtained from the Target database (https://ocg.cancer.gov/programs/target). Human osteosarcoma MG-63, Saos-2, and U-2OS cell lines—which were purchased from the American Typical Culture Collection (ATCC, USA)—were used for *in vitro* experiments. Data were statistically analyzed using SPSS 17.0, and results are expressed as $x \pm s$. Component data were compared using a t-test, and $P < 0.05$ was considered as a statistically significant difference in data.

## Analysis of Variance With Protein-Protein Interaction Network (PPI)

The samples were divided into transferred and non-transferred groups and analysis of variance (ANOVA) was performed using the R package "limma," with $|\log2\ FC| > 0.3785$ and $P < 0.05$ as the screening condition. The STRING database (https://string-db.org/) was used to obtain the reciprocal network relationships among differentially expressed genes (DEGs) and then imported into Cytoscape software to eliminate isolated nodes and use the MCC algorithm to derive the top-10 ranked genes.

## Construction of a Prognostic Model

The LASSO-COX prognostic model was constructed by the R package "glmnet" to obtain a risk score formula containing multiple genes, and the samples were divided into two groups—high and low risk—in accordance with the calculation results. In addition, the R package "survivor" was used to perform survival analysis, aLog-rank was used to test the survival difference between high and low-risk groups, and the prediction accuracy of the model was analyzed using time ROC. The genes included in the model were subsequently analyzed using multifactorial COX regression, and genes with $P < 0.05$ were used as the key genes. The R package "ggplot2" was used to obtain the expression distribution of the key genes in the transferred and non-transferred groups, and the differences between the two groups were analyzed using the Wilcoxon rank-sum test.

## Cell Culture and Transfection

The human osteosarcoma MG-63, Saos-2, and U-2OS cell lines were cultured in a DMEM medium containing 10% FBS, and the cells were cultured and passaged at 37°C in an atmosphere of 5% CO2. In addition, MG-63 cells in the logarithmic growth phase were taken and divided into control, OE-COL5A2, and si-COL5A2 groups for transfection; the cells were collected for subsequent experiments 48h after transfection.

## Quantitative Real-Time PCR (qRT-PCR) for the Detection of Key Gene Expression

The transfected MG-63 cells were treated with Trizol reagent and RNA was extracted. cDNA was synthesized using a reverse transcription kit according to the manufacturer's instructions. qRT-PCR was performed to amplify the reaction products, using GAPDH as the internal reference and the primer sequences were GAPDH, forward primer: CTGGGCTACACTGAGCACC, and

the results were expressed as 2-ΔΔCt. The reaction products were synthesized by qRT-PCR with GAPDH as the internal reference and the primer sequences were GAPDH, forward primer: CTGGGCTACACTGAGCACC, reverse primer: AAGTGGTCGTTGAGGGCAATG; and COL5A2, forward primer: GACTGTGCCGACCCTGTAAC, reverse primer: CCTGGACGACCACGTATGC.

## Immunofluorescence (IF) Analysis

The transfected MG-63, Saos-2, and U-2OS cells were washed with PBS, fixed with 4% paraformaldehyde for 20 min, and then the cells were incubated in 0.3% Triton X-100 for 20 min. The closure was performed at room temperature, followed by immunofluorescence staining with primary antibody COL5A2 (ab7046,abcam). Cells were washed with PBS and incubated again with fluorescently labeled secondary antibodies and, finally, images were obtained using fluorescence microscopy.

## Western Blotting (WB)

MG-63 cells at the logarithmic growth stage were taken, cells were lysed with lysis buffer, and total proteins were extracted. The total protein concentration was determined using the BCA kit. Thereafter, protein samples were subjected to SDS-PAGE and transferred to PVDF membranes using anti-COL5A2 (ab7046, abcam), β-catenin (ab32572, abcam), Vimentin (ab92547, abcam), E-cadherin (ab233611, abcam), and N-cadherin (ab254512, abcam); these samples were incubated overnight at 4°C and then the membranes were immunoblotted with horseradish peroxidase-coupled secondary antibodies. β-actin (ab8226,abcam) was used as a control and the protein signal was detected using a gel imaging system.

## CCK-8 Detection oCell Proliferation

The transfected MG-63 cells were digested with trypsin and inoculated in 96-well plates, added with DMEM medium, and incubated at 37°C in an incubator with 5% CO2. According to the manufacturer's instructions, CCK-8 solution was added after 0h, 6h, 12h, 24h, and 48h of incubation, respectively, and the incubation was continued for 3h. The absorbance value (OD) was measured at a wavelength of 450 nm; three experiments were performed in parallel, and the average value was taken for the calculation.

## Scratch Assay to Detect Cell Migration

The transfected human osteosarcoma MG-63 cells were inoculated in a six-well plate. When the cell density reached 90%, a 20 μL pipette tip was used to gently scrape the bottom of the well. The cells were washed using PBS solution, the scratched-down cells were removed, serum-free medium was added and incubated at 37°C in a 5% CO2 incubator, and the degree of healing of the scratches was observed at 0h and 24h, respectively. The cell migration rate (%) = (initial scratch width value - corresponding time point scratch width value)/initial scratch width value × 100%.

## Transwell Assay for Cell Invasion

Transfected human osteosarcoma MG-63 cells were taken and inoculated into the upper chamber of transwell coated with Matrigel matrix gel, and serum-free medium was added to maintain cell survival. Medium containing 10% FBS was added to the lower chamber of the transwell as a chemical elicitor, and the cells remaining on the membrane surface were removed after incubation for 48h. The membranes were fixed with 100% ethanol for 5 min and then stained with 5% crystalline violet for 5 min. Photographs were taken using a light microscope, and the cell numbers were counted and averaged.

# RESULTS

## Preliminary Gene Screening

The osteosarcoma samples in the Target database were divided into metastatic and non-metastatic groups for differential gene expression analysis, and 1100 up-regulated genes and 36 down-regulated genes were obtained (**Figure 1A**). PPI networks were constructed using DEGs, which were used to observe the protein-protein interactions (**Figure 1B**). The top 10 key genes (COL5A2, BGN, ITGB4, COL5A1, ITGB3, TLN1, FLNA, CTNNB1, FN1, LOX) were subsequently obtained by the MCC algorithm (**Figure 1C**).

## COL5A2: A Key Gene Related to the Prognosis of Osteosarcoma

A LASSO-COX prognostic risk model was constructed based on the top-10 ranked key genes, and a risk scoring system was obtained when the minimum λ was 0.0645 (**Figure 2A, B**). The corresponding risk score values were calculated based on risk score = (0.1982)*COL5A2 + (-0.043) × ITGB3 + (-0.0562) × TLN1 + (-0.2716) × FLNA, and the samples were divided into high-risk and low-risk groups. KM survival curves reveal that the high-risk group was associated with a poor prognosis of osteosarcoma (P = 0.00148, HR = 3.181) (**Figure 2C**). Furthermore, the ROC curve revealed that the risk model had a good predictive ability (**Figure 2D**). Multi-factorial COX regression analysis of four genes in the risk model related to the prognosis of osteosarcoma led to the identification of COL5A2 as the key gene associated with osteosarcoma prognosis (**Figure 2E**). COL5A2 was significantly expressed in the osteosarcoma metastasis group (P = 0.0089) and its high expression led to a worse prognosis (P = 0.018, HR = 2.2) (**Figures 2F, G**).

## High Expression of COL5A2 in Osteosarcoma Cells

To determine the relationship between COL5A2 and osteosarcoma progression, this study first verified the expression of COL5A2 using qRT-PCR and immunofluorescence techniques. As shown in the figure, the expression of col5a2 in human osteosarcoma MG-63 cells was analyzed by immunofluorescence technique (**Figure 3A**). Subsequently, human osteosarcoma MG-63 cells were divided into control, COL5A2-enhanced, and COL5A2-knockdown groups and the different expression levels of COL5A2 in MG-63 cells were verified using qRT-PCR with GAPDH as the control gene (**Figure 3B**). In addition, the results of this experiment demonstrated that the cells were successfully transfected.

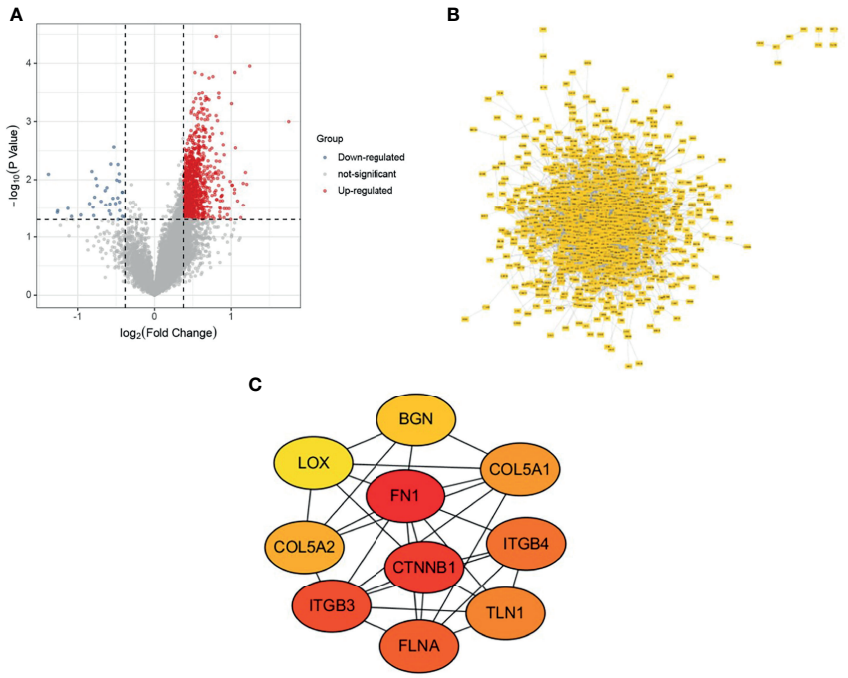



**FIGURE 1** | Variance analysis and PPI construction. **(A)** Volcano map; **(B)** PPI network map; **(C)** Top 10 genes.

## The Pro-Tumor Effect of COL5A2

In order to investigate the biological role of COL5A2 in osteosarcoma, OERNA with COL5A2-enhancing effect was transfected, and its expression was stably upregulated in MG-63 cells, which is a finding verified in **Figure 3B**. The proliferation rate was significantly faster in the experimental group compared to the control group and revealed a significant difference from 6h onward (**Figure 4A**). Similarly, the migration ability of osteosarcoma cells was enhanced after transfection with OERNA of COL5A2 (**Figure 4B**). In addition, the Transwell assay revealed that the upregulation of COL5A2 promoted the invasive effect of osteosarcoma cells (**Figure 4C**). All these experiments demonstrated that the upregulation of COL5A2 expression contributes to the promotion of osteosarcoma progression.

## COL5A2 Can Affect Osteosarcoma Progression Through the Wnt/β-Catenin and TGF-β Signaling Pathways

TGF-β signaling and Wnt/β-catenin signaling can promote the EMT process by increasing the expression of N-cadherin and Vimentin or decreasing the expression of E-cadherin, thereby promoting tumor invasion and metastasis (15, 23). In this study, we first examined the expression of Wnt/β-catenin and TGF-β pathway proteins by WB; the results revealed that the knockdown of COL5A2 decreased the expressions of β-catenin and TGF-β1 (**Figures 5A**). Subsequently, the experimental subjects were divided into two groups treated with or without TGF-β1, each comprising a control group and a si-COL5A2 group. WB analysis revealed that N-cadherin and Vimentin expression was downregulated by the knockdown of COL5A2, while E-cadherin expression was upregulated with the knockdown of COL5A2 (**Figures 5B**). Using the same method to detect the effect of COL5A2 knockdown on the Wnt/β-catenin signaling pathway again, the same conclusion can be drawn (**Figures 5C**). These results suggest that the downregulation of COL5A2 can promote osteosarcoma metastasis by affecting either the Wnt/β-catenin signaling pathway or the TGF-β signaling pathway.

# DISCUSSION

Osteosarcoma is an important threat to the health of children and adolescents. Currently, surgical resection is often used to control local tumors at the primary tumor site, while neoadjuvant multidrug chemotherapy is mostly used for unresectable cases, like distant metastases (24, 25). However, the high propensity to metastasize leads to a high mortality rate and often a poor prognosis for patients with osteosarcoma (26). Studies have revealed that several proteins or RNAs are tumor-promoting factors in osteosarcoma and they play a regulatory or inducing role in tumor progression (27, 28). This study aimed to identify new prognostic factors associated with osteosarcoma metastasis and to briefly describe the mechanisms by which they affect cancer cell progression.

Osteosarcoma samples were grouped based on whether or not they metastasized and screened by bioinformatics analysis to obtain COL5A2. Type V α2 collagen (COL5A2) is mainly found in the basement membrane of the skin and provides a template for the composition of type V collagen (29). It was revealed that the upregulation of the collagen family gene (COL) can be used as a diagnostic marker for osteosarcoma (30). In contrast,

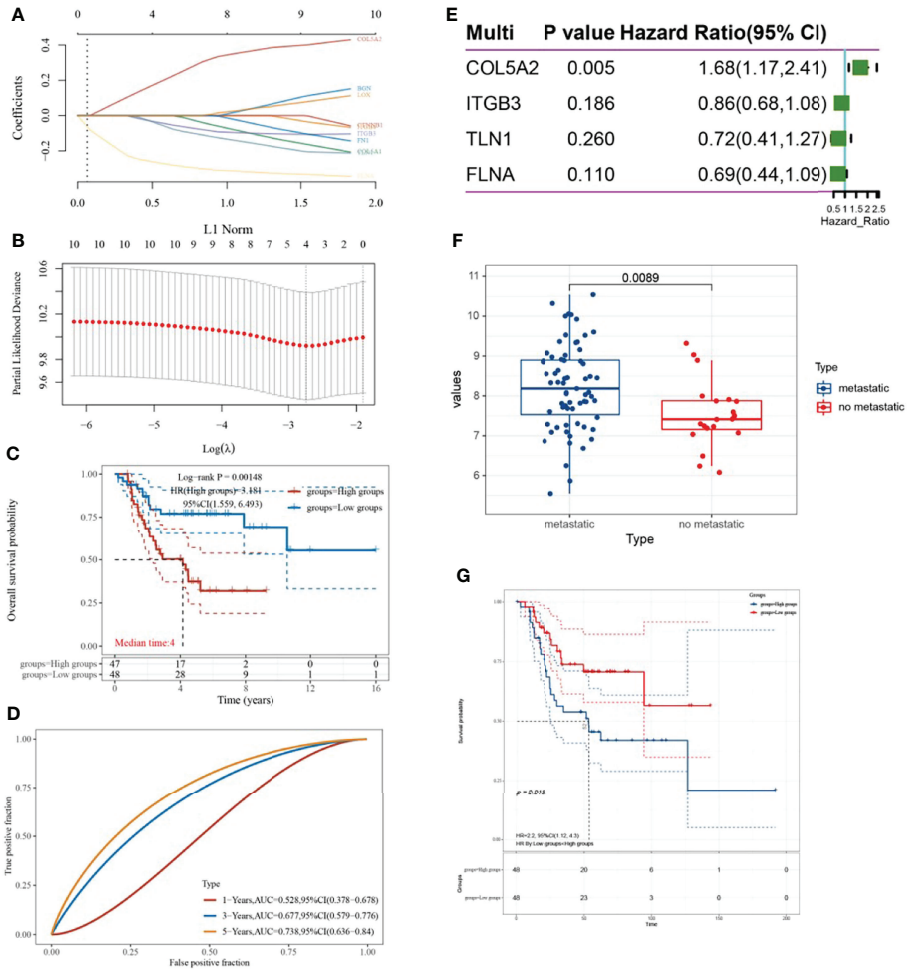



**FIGURE 2** | Establishment of the LASSO-COX prognostic model and screening of key genes. **(A, B)** λ parameters for selected features and partial likelihood deviation plotted against log(λ) using the LASSO-COX model; **(C)** KM survival curve based on risk grouping; **(D)** time ROC curve; **(E)** forest plot; **(F)** box line plot with COL5A2 expression distribution based on the presence or absence of metastatic grouping; **(G)** relationship between high and low COL5A2 expression and prognosis of osteosarcoma.

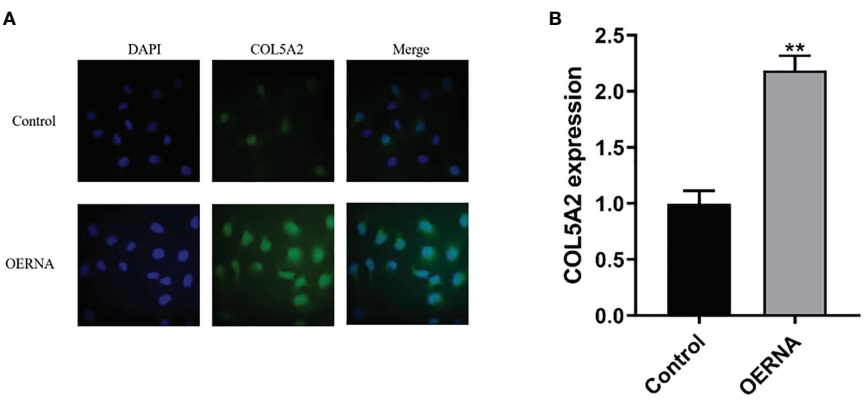



**FIGURE 3** | COL5A2 expression validation. **(A)** Immunofluorescence to verify the differential expression of COL5A2 in different types of osteosarcoma cell lines up to; **(B)** qRT-PCR to verify the high and low expression levels of COL5A2 in osteosarcoma **P<0.01.

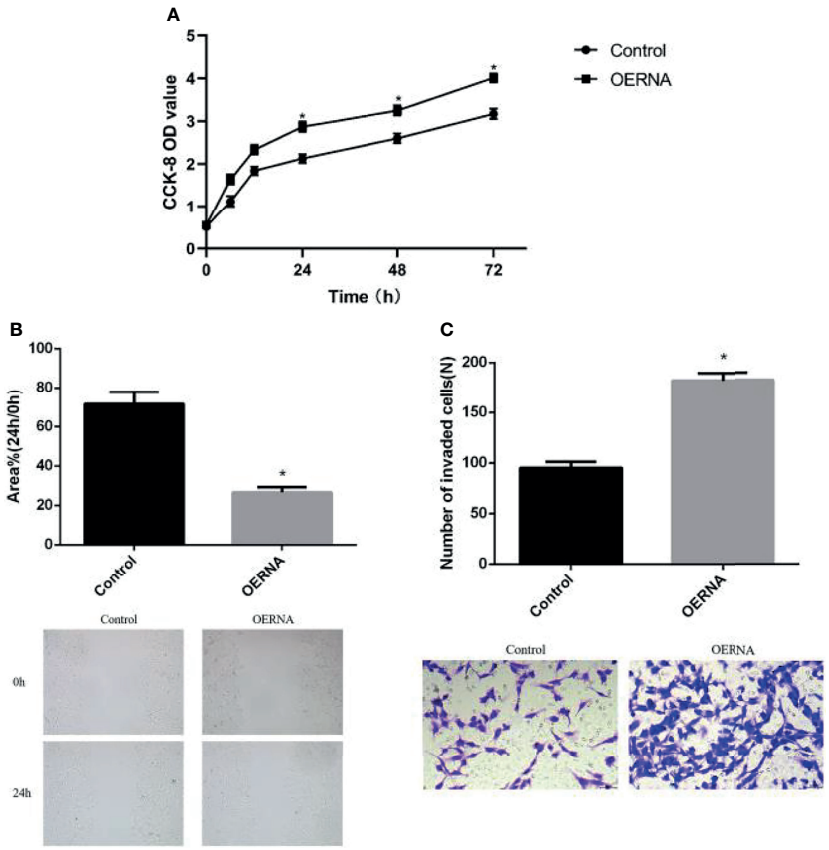



**FIGURE 4** | High expression of COL5A2 promotes the progression of osteosarcoma. **(A)** Enhanced promotion of osteosarcoma cell proliferation by COL5A2 (CCK-8 assay) ($P < 0.05$); **(B)** scratch assay to observe the migration ability of osteosarcoma cells after the upregulation of COL5A2 expression *P<0.05; **(C)** transwell assay was used to assess the invasiveness of osteosarcoma cells after the upregulation of COL5A2 expression *P<0.05.

COL5A2, a member of the collagen family, is involved in immune system regulation, angiogenesis, and tumor metastasis, and plays a role in promoting tumor progression (31, 32). It has been demonstrated that high COL5A2 expression promotes the proliferation and invasion of prostate cancer cells (33), is associated with poor prognosis in patients with gastric cancer (34), and promotes the proliferation of colon cancer cells by activating the Wnt/β-catenin signaling pathway (35). In this study, the analysis of bioinformatics revealed that COL5A2 indicated a high expression status in metastatic osteosarcoma and led to a poor prognosis. To explore the role and functions played by COL5A2 in osteosarcoma progression, a series of *in vitro* experiments were performed in this study. qRT-PCR and immunofluorescence assays confirmed the high expression of COL5A2 in human osteosarcoma cells. Subsequent CCK-8 assays, scratch assays, and transwell assays confirmed that enhanced COL5A2 expression contributes to the proliferation, invasion, and migration of osteosarcoma cells. These results suggest that high COL5A2 expression may be associated with the high metastatic properties of osteosarcoma.

In the current study, EMT is considered to be a key factor involved in tumor cell invasion and metastasis, and it has been reported that metastasis can be promoted by regulating the EMT process in colorectal cancer (23), breast cancer (36), and head and neck squamous cell carcinoma (37). Moreover, in osteosarcoma, the EMT process is also associated with tumor cell metastasis (38). The tGF-β signaling and Wnt/β-catenin signaling pathways are two important pathways that activate the EMT process and are both involved in tumor invasion and metastasis (39, 40). tGF-β is an important factor that regulates the immune response *in vivo*, and tumor cells can manipulate the regulatory process of TGF-β. TGF-β can lead to cell cycle arrest and apoptosis in the early stages of cancer, but cancer cells usually transform TGF-β signaling into a tumor-promoting function by mutating key components of the TGF-β signaling pathway (41). Dysregulation of the TGF-β signaling pathway is often a signal for tumorigenesis. In patients with osteosarcoma, the upregulation of TGF-β expression can lead to increased chemoresistance and subsequently poor prognosis (42). Similarly, there are numerous studies that confirm that the Wnt/β-catenin signaling pathway is one of the major oncogenic pathways involved in the progression of human osteosarcoma (43). β-catenin is a key mediator of Wnt signaling and contributes to stable cell-to-cell contacts (44). An aberrant Wnt/β-catenin signaling pathway promotes cancer stem cell renewal, cell proliferation, and

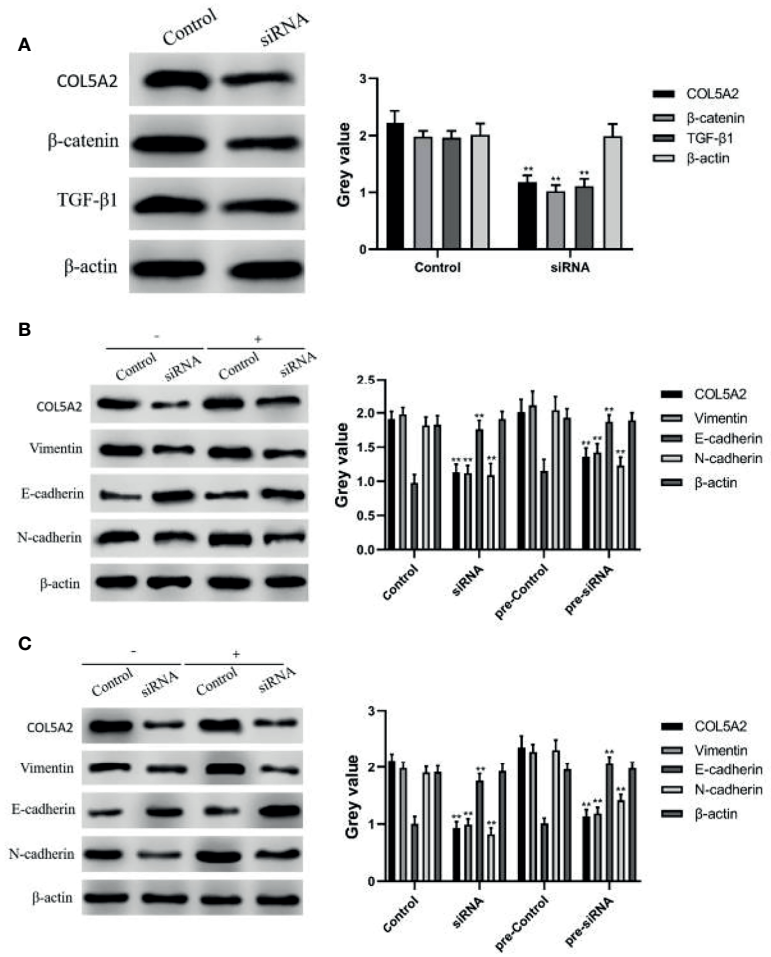



**FIGURE 5** | WB to verify the relationship between COL5A2 and signaling pathways. **(A)** Expression of β-catenin, TGFF-β1 was detected by WB after reducing the expression of COL5A2 **P<0.01. **(B)** protein expression levels of COL5A2, E-cadherin, N-cadherin, and Vimentin were assessed by WB after treatment with or without TGF-β1 **P<0.01. **(C)** protein expression levels of COL5A2, E-cadherin, N-cadherin, and Vimentin were assessed by WB after treatment with or without Wnt/β-catenin treatment and the protein expression levels of COL5A2, E-cadherin, N-cadherin, and Vimentin **P<0.01.

differentiation as well as plays an important role in tumorigenesis (45). The invasive and metastatic process of several cancers can be promoted by activating the Wnt/β-catenin signaling pathway (46, 47). Further, different genes have been shown to inhibit the growth of osteosarcoma by suppressing the TGF-β signaling pathway or the Wnt/β-catenin signaling pathway (48, 49). In this study, we further determined that the knockdown of COL5A2 expression inhibited the TGF-β signaling and Wnt/β-catenin signaling pathways, thereby suppressing the invasion and migration of osteosarcoma.

In the present study, COL5A2 was the only gene screened for prognostic relevance in patients with osteosarcoma based on whether they were grouped metastatically. COL5A2 is highly expressed in patients with metastatic osteosarcoma and indicates a poor prognostic survival. In the study by Chen et al., COL5A2 was investigated as a gene regulatory mediator and the authors simply demonstrated that the knockdown of COL5A2 inhibited metastasis and proliferation of osteosarcoma (50). This is consistent with our findings that reveal that enhanced COL5A2 expression accelerated the proliferation, invasion, and migration rate of osteosarcoma cells. Furthermore, WB experiments confirmed that COL5A2 could achieve tumor suppression by inhibiting the TGF-β signaling or Wnt/β-catenin signaling pathways.

## 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

Y-LH, DL, KH, WA, and LW: Conceptualization, Methodology, Software. Y-LH, DL and LW: Data curation, Writing- Original draft preparation. WW: Visualization, Investigation. JT, WZ, WW and WA: Software, Validation. Y-LH and LW: Writing- Reviewing and Editing.

## REFERENCES


1. Gianferante DM, Mirabello L, Savage SA. Germline and Somatic Genetics of Osteosarcoma—Connecting Aetiology, Biology and Therapy. *Nat Rev Endocrinol* (2017) 13(8):480–91. doi: 10.1038/nrendo.2017.16
2. Huang X, Zhao J, Bai J, Shen H, Zhang B, Deng L, et al. Risk and Clinicopathological Features of Osteosarcoma Metastasis to the Lung: A Population-Based Study. *J Bone Oncol* (2019) 16:100230. doi: 10.1016/j.jbo.2019.100230
3. Sadykova LR, Ntekim AI, Muyangwa-Semenova M, Rutland CS, Jeyapalan JN, Blatt N, et al. Epidemiology and Risk Factors of Osteosarcoma. *Cancer Invest* (2020) 38(5):259–69. doi: 10.1080/07357907.2020.1768401
4. Sobhan MR, Mostafavi S, Mazaheri M, Neamatzadeh H. Genetics of Pediatric Bone Tumors: A Systematic Review. *Iranian J Pediatr Hematol Oncol* (2017) 7 (2):117–29. doi: 10.1016/s46005-010-0040-1
5. Song K, Song J, Lin K, Chen F, Ma X, Jiang J, et al. Survival Analysis of Patients With Metastatic Osteosarcoma: A Surveillance, Epidemiology, and End Results Population-Based Study. *Int orthopaedics* (2019) 43(8):1983–91. doi: 10.1007/s00264-019-04348-4
6. Lindsey BA, Markel JE, Kleinerman ES. Osteosarcoma Overview. *Rheumatol Ther* (2017) 4(1):25–43. doi: 10.1007/s40744-016-0050-2
7. Shin SH, Jeong HJ, Han I, Cho HS, Kim H-S. Osteosarcoma and Chondrosarcoma of the Shoulder: Site-Specific Comparative Analysis. *Orthopedics* (2013) 36(2):e179–e85. doi: 10.3928/01477447-20130122-20
8. Guo J, Reddick WE, Glass JO, Ji Q, Billups CA, Wu J, et al. Dynamic Contrast-Enhanced Magnetic Resonance Imaging as a Prognostic Factor in Predicting Event-Free and Overall Survival in Pediatric Patients With Osteosarcoma. *Cancer* (2012) 118(15):3776–85. doi: 10.1002/cncr.26701
9. Sheen H, Kim W, Byun BH, Kong C-B, Song WS, Cho WH, et al. Metastasis Risk Prediction Model in Osteosarcoma Using Metabolic Imaging Phenotypes: A Multivariable Radiomics Model. *PloS One* (2019) 14(11): e0225242. doi: 10.1371/journal.pone.0225242
10. Song Z, Pearce MC, Jiang Y, Yang L, Goodall C, Miranda CL, et al. Delineation of Hypoxia-Induced Proteome Shifts in Osteosarcoma Cells With Different Metastatic Propensities. *Sci Rep* (2020) 10(1):1–17. doi: 10.1038/s41598-019-56878-x
11. Elia I, Doglioni G, Fendt S-M. Metabolic Hallmarks of Metastasis Formation. *Trends Cell Biol* (2018) 28(8):673–84. doi: 10.1016/j.tcb.2018.04.002
12. Redfern AD, Spalding LJ, Thompson EW. The Kraken Wakes: Induced EMT as a Driver of Tumour Aggression and Poor Outcome. *Clin Exp Metastasis* (2018) 35(4):285–308. doi: 10.1007/s10585-018-9906-x
13. Boulding T, McCuaig R, Tan A, Hardy K, Wu F, Dunn J, et al. LSD1 Activation Promotes Inducible EMT Programs and Modulates the Tumour Microenvironment in Breast Cancer. *Sci Rep* (2018) 8(1):1–18. doi: 10.1038/s41598-017-17913-x
14. Yue B, Liu C, Sun H, Liu M, Song C, Cui R, et al. A Positive Feed-Forward Loop Between LncRNA-CYTOR and Wnt/β-Catenin Signaling Promotes Metastasis of Colon Cancer. *Mol Ther* (2018) 26(5):1287–98. doi: 10.1016/j.ymthe.2018.02.024
15. Xie F, Ling L, van Dam H, Zhou F, Zhang L. TGF-β Signaling in Cancer Metastasis. *Acta Biochim Biophys Sin* (2018) 50(1):121–32. doi: 10.1093/abbs/gmx123
16. Wu Y, Shen Z, Wang K, Ha Y, Lei H, Jia Y, et al. High FMNL3 Expression Promotes Nasopharyngeal Carcinoma Cell Metastasis: Role in TGF-β1-Induced Epithelia-to-Mesenchymal Transition. *Sci Rep* (2017) 7(1):1–14. doi: 10.1038/srep42507
17. Tammela T, Sanchez-Rivera FJ, Cetinbas NM, Wu K, Joshi NS, Helenius K, et al. A Wnt-Producing Niche Drives Proliferative Potential and Progression in Lung Adenocarcinoma. *Nature* (2017) 545(7654):355–9. doi: 10.1038/nature22334
18. Zhang M, Du H, Wang L, Yue Y, Zhang P, Huang Z, et al. Thymoquinone Suppresses Invasion and Metastasis in Bladder Cancer Cells by Reversing EMT Through the Wnt/β-Catenin Signaling Pathway. *Chemico-Biol Interact* (2020) 320:109022. doi: 10.1016/j.cbi.2020.109022
19. Liu L, Wu B, Cai H, Li D, Ma Y, Zhu X, et al. Tiam1 Promotes Thyroid Carcinoma Metastasis by Modulating EMT *via* Wnt/β-Catenin Signaling. *Exp Cell Res* (2018) 362(2):532–40. doi: 10.1016/j.yexcr.2017.12.019
20. Liu F, Xia Z, Zhang M, Ding J, Feng Y, Wu J, et al. SMARCAD1 Promotes Pancreatic Cancer Cell Growth and Metastasis Through Wnt/β-Catenin-Mediated EMT. *Int J Biol Sci* (2019) 15(3):636. doi: 10.7150/ijbs.29562
21. Zhao H, Peng C, Lu X, Guo M, Yang T, Zhou J, et al. PDCD5 Inhibits Osteosarcoma Cell Metastasis *via* Targeting TGF-β1/Smad Signaling Pathway and is Associated With Good Prognosis. *Am J Trans Res* (2019) 11(2):1116.
22. Mao X, Jin Y, Feng T, Wang H, Liu D, Zhou Z, et al. Ginsenoside Rg3 Inhibits the Growth of Osteosarcoma and Attenuates Metastasis Through the Wnt/β-Catenin and EMT Signaling Pathway. Evidence-Based Complementary and Alternative Medicine. *Evid Based Complement Alternat Med* (2020) 2020 (17):232–9. doi: 10.1155/2020/6065124
23. Li Q, Lai Q, He C, Fang Y, Yan Q, Zhang Y, et al. RUNX1 Promotes Tumour Metastasis by Activating the Wnt/β-Catenin Signalling Pathway and EMT in Colorectal Cancer. *J Exp Clin Cancer Res* (2019) 38(1):1–13. doi: 10.1186/s13046-019-1330-9
24. Eaton BR, Schwarz R, Vatner R, Yeh B, Claude L, Indelicato DJ, et al. Osteosarcoma. *Pediatr Blood Cancer* (2021) 68:e28352. doi: 10.1002/pbc.28352
25. Miwa S, Shirai T, Yamamoto N, Hayashi K, Takeuchi A, Igarashi K, et al. Current and Emerging Targets in Immunotherapy for Osteosarcoma. *J Oncol* (2019) 102:1023–31. doi: 10.1155/2019/7035045
26. Yi Z, Pu Y, Gou R, Chen Y, Ren X, Liu W, et al. Silencing of RIPK4 Inhibits Epithelial–Mesenchymal Transition by Inactivating the Wnt/β–Catenin Signaling Pathway in Osteosarcoma. *Mol Med Rep* (2020) 21(3):1154–62. doi: 10.3892/mmr.2020.10939
27. Wang J-Y, Yang Y, Ma Y, Wang F, Xue A, Zhu J, et al. Potential Regulatory Role of lncRNA-miRNA-mRNA Axis in Osteosarcoma. *Biomed Pharmacother* (2020) 121:109627. doi: 10.1016/j.biopha.2019.109627
28. Chen S, Li Y, Zhi S, Ding Z, Wang W, Peng Y, et al. WTAP Promotes Osteosarcoma Tumorigenesis by Repressing HMBOX1 Expression in an M 6 A-Dependent Manner. *Cell Death Dis* (2020) 11(8):1–14. doi: 10.1038/s41419-020-02847-6
29. Ilieș RF, Aioanei CS, Cătană A, Halmagyi S-R, Lukacs I, Tokes R-E, et al. Involvement of COL5A2 and TGF−β1 in Pathological Scarring. *Exp Ther Med* (2021) 22(4):1–4. doi: 10.3892/etm.2021.10501
30. Wu D, Chen K, Bai Y, Zhu X, Chen Z, Wang C, et al. Screening of Diagnostic Markers for Osteosarcoma. *Mol Med Rep* (2014) 10(5):2415–20. doi: 10.3892/mmr.2014.2546 doi: 10.3892/mmr.2014.2546
31. Ding Y-L, Sun S-F, Zhao G-L. COL5A2 as a Potential Clinical Biomarker for Gastric Cancer and Renal Metastasis. *Med* (2021) 100(7):1–8. doi: 10.1097/MD.0000000000024561
32. Meng X-Y, Shi M-J, Zeng Z-H, Chen C, Liu T-Z, Wu Q-J, et al. The Role of COL5A2 in Patients With Muscle-Invasive Bladder Cancer: A Bioinformatics Analysis of Public Datasets Involving 787 Subjects and 29 Cell Lines. *Front Oncol* (2019) 8:659. doi: 10.3389/fonc.2018.00659
33. Ren X, Chen X, Fang K, Zhang X, Wei X, Zhang T, et al. COL5A2 Promotes Proliferation and Invasion in Prostate Cancer and Is One of Seven Gleason-Related Genes That Predict Recurrence-Free Survival. *Front Oncol* (2021) 11:426. doi: 10.3389/fonc.2021.583083
34. Tan Y, Chen Q, Xing Y, Zhang C, Pan S, An W, et al. High Expression of COL5A2, a Member of COL5 Family, Indicates the Poor Survival and Facilitates Cell Migration in Gastric Cancer. *Biosci Rep* (2021) 41(4): BSR20204293. doi: 10.1042/BSR20204293
35. Wang J, Jiang Y-H, Yang P-Y, Liu F. Increased Collagen Type V α2 (COL5A2) in Colorectal Cancer Is Associated With Poor Prognosis and Tumor Progression. *OncoTargets Ther* (2021) 14:2991. doi: 10.2147/OTT.S288422
36. Neelakantan D, Zhou H, Oliphant MU, Zhang X, Simon LM, Henke DM, et al. EMT Cells Increase Breast Cancer Metastasis *via* Paracrine GLI Activation in Neighbouring Tumour Cells. *Nat Commun* (2017) 8(1):1–14. doi: 10.1038/ncomms15773
37. Chen L, Li YC, Wu L, Yu GT, Zhang WF, Huang CF, et al. TRAF 6 Regulates Tumour Metastasis Through EMT and CSC Phenotypes in Head and Neck Squamous Cell Carcinoma. *J Cell Mol Med* (2018) 22(2):1337–49. doi: 10.1111/jcmm.13439
38. Jiang X, Zhang Z, Song C, Deng H, Yang R, Zhou L, et al. Glaucocalyxin A Reverses EMT and TGF-β1-Induced EMT by Inhibiting TGF-β1/Smad2/3 Signaling Pathway in Osteosarcoma. *Chemico-Biol Interact* (2019) 307:158–66. doi: 10.1016/j.cbi.2019.05.005

39. Tripathi V, Shin J-H, Stuelten CH, Zhang YE. TGF-β-Induced Alternative Splicing of TAK1 Promotes EMT and Drug Resistance. *Oncogene* (2019) 38 (17):3185–200. doi: 10.1038/s41388-018-0655-8
40. Qu T, Zhao Y, Chen Y, Jin S, Fang Y, Jin X, et al. Down-Regulated MAC30 Expression Inhibits Breast Cancer Cell Invasion and EMT by Suppressing Wnt/β-Catenin and PI3K/Akt Signaling Pathways. *Int J Clin Exp Pathol* (2019) 12(5):1888.
41. Korkut A, Zaidi S, Kanchi RS, Rao S, Gough NR, Schultz A, et al. A Pan-Cancer Analysis Reveals High-Frequency Genetic Alterations in Mediators of Signaling by the TGF-β Superfamily. *Cell Syst* (2018) 7(4):422–37.e7. doi: 10.1016/j.cels.2018.08.010
42. Zhou L, Tang J, Hu F, Liao Y, Li R, Zhou Y, et al. Effects of Different Levels of TGF-β Expression and Tumor Cell Necrosis Rates in Osteosarcoma on the Chemotherapy Resistance of Osteosarcoma. *J Bone Oncol* (2020) 23:100299. doi: 10.1016/j.jbo.2020.100299
43. Chang J, Li Y, Wang X, Hu S, Wang H, Shi Q, et al. Polyphyllin I Suppresses Human Osteosarcoma Growth by Inactivation of Wnt/β-Catenin Pathway *In Vitro* and In Vivo. *Sci Rep* (2017) 7(1):1–12. doi: 10.1038/s41598-017-07194-9
44. Cheng X, Xu X, Chen D, Zhao F, Wang W. Therapeutic Potential of Targeting the Wnt/β-Catenin Signaling Pathway in Colorectal Cancer. *Biomed Pharmacother* (2019) 110:473–81. doi: 10.1016/j.biopha.2018.11.082
45. Zhang Y, Wang X. Targeting the Wnt/β-Catenin Signaling Pathway in Cancer. *J Hematol Oncol* (2020) 13(1):1–16. doi: 10.1186/s13045-020-00990-3
46. Zhou P, Li Y, Li B, Zhang M, Liu Y, Yao Y, et al. NMIIA Promotes Tumor Growth and Metastasis by Activating the Wnt/β-Catenin Signaling Pathway and EMT in Pancreatic Cancer. *Oncogene* (2019) 38(27):5500–15. doi: 10.1038/s41388-019-0806-6
47. Yang S, Liu Y, Li M-Y, Ng CS, S-l Y, Wang S, et al. FOXP3 Promotes Tumor Growth and Metastasis by Activating Wnt/β-Catenin Signaling Pathway and EMT in Non-Small Cell Lung Cancer. *Mol Cancer* (2017) 16(1):1–12. doi: 10.1186/s12943-017-0700-1
48. Tan T, Chen J, Hu Y, Wang N, Chen Y, Yu T, et al. Dihydrotanshinone I Inhibits the Growth of Osteosarcoma Through the Wnt/β-Catenin Signaling Pathway. *OncoTargets Ther* (2019) 12:5111. doi: 10.2147/OTT.S204574
49. Sun Y, Jiang X, Lu Y, Zhu J, Yu L, Ma B, et al. Oridonin Prevents Epithelial-Mesenchymal Transition and TGF-β1-Induced Epithelial-Mesenchymal Transition by Inhibiting TGF-β1/Smad2/3 in Osteosarcoma. *Chemico-Biol Interact* (2018) 296:57–64. doi: 10.1016/j.cbi.2018.09.013
50. Chen H, Liu W, Zhong L, Liao D, Zhang R, Kang T, et al. NKX2-2 Suppresses Osteosarcoma Metastasis and Proliferation by Downregulating Multiple Target Genes. *J Cancer* (2018) 9(17):3067. doi: 10.7150/jca.26382


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 25 February 2022
doi: 10.3389/fonc.2022.835074

# Construction of a Co-Expression Network for lncRNAs and mRNAs Related to Urothelial Carcinoma of the Bladder Progression

*Yeqing Mao*[1]*, *Chao Wen*[2] and *Zitong Yang*[2]

[1] Urology Department, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, China, [2] Medical College, Zhejiang University, Hangzhou, China

Carcinoma of urinary bladder is the most familiar cancer of the urinary tract, with the highest incidence in men. However, its prognosis and treatment have not improved significantly in the last 30 years. The main reason for this may be related to the alteration and regulation of genes. These alterations in genes that play a crucial role in cell cycle regulation may result in high-grade tumors and may alter drug sensitivity. Notably, the role of lncRNA in bladder cancer, especially the lncRNA-mRNA regulatory network, has not been fully elucidated. In this manuscript, we compared RNA sequencing (RNA-seq) data from 19 normal bladder tissues and 411 primary bladder tumor tissues using The Cancer Genome Atlas (TCGA) data bank, subjected differentially expressed mRNAs and lncRNAs to weighted gene co-expression network analysis, and screened out modules highly correlated with tumor progression. Subsequently, a lncRNA-mRNA co-expression network was built, and two key mRNAs were identified *via* COX regression analysis. Kaplan-Meier curve analysis revealed that the overall survival of sick people in the high-risk section was significantly shorter than those in the low-risk section. Therefore, this lncRNA-mRNA-based co-expression pattern may be used clinically to predict the prognosis of carcinoma of urinary bladder people. Our study not only provides a genetic target for carcinoma of urinary bladder therapy but also provides new ideas for people in the medical profession to discover the treatment of various tumors.

**Keywords: urothelial carcinoma of the bladder, WGCNA, COX analysis, lncRNA-mRNA co-expression network, prognostic marker**

OPEN ACCESS

***Edited by:***
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

***Reviewed by:***
*Ning Li,*
*Fourth Affiliated Hospital of China Medical University, China*
*Wenhu Zhou,*
*Central South University, China*

***Correspondence:**
*Yeqing Mao*
*1313019@zju.edu.cn*

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

***Received:*** *14 December 2021*
***Accepted:*** *24 January 2022*
***Published:*** *25 February 2022*

***Citation:***
*Mao Y, Wen C and Yang Z (2022) Construction of a Co-Expression Network for lncRNAs and mRNAs Related to Urothelial Carcinoma of the Bladder Progression. Front. Oncol. 12:835074. doi: 10.3389/fonc.2022.835074*

## INTRODUCTION

Carcinoma of urinary bladder is one of the most familiar malignancies worldwide (1), dominated by urothelial carcinoma of the bladder (BLCA), which accounts for approximately 90% of bladder cancers (2). It is predicted that there were 573,278 new cases of BCa and 212,536 deaths in 185 countries worldwide in 2020 (3). The traditional cure for carcinoma of urinary bladder mainly include surgical excision and chemotherapy. Hence, the recurrence rate is not low, with an overall 5-year survival rate of 15-20% (4, 5). Additionally, surgical treatment and chemotherapy are quite

limited for advanced bladder cancer (6, 7). This high recurrence rate may be partly explained by the poorly understood pathogenesis of BCa (8). Therefore, exploring the pathogenesis of BCa and identifying accurate and effective biomarkers based on its clinical spectrum is vital for early diagnosis without obvious clinical symptoms, assessment of prognosis, and the development of effective treatment strategies.

For the past few years, high-throughput transcriptome sequencing has become extremely usual, revealing that up to 70% of the human genome has been transcribed (9). Hence, Most of the transcribed genes detected by high-throughput sequencing are non-coding genes, which may be associated with noncoding RNA (lncrnas) longer than 200 nucleotides (10). LncRNAs may be the most critical regulators of gene expression, cell growth, cell differentiation, cell development and chromatin dynamics (11). Thousands of lncRNAs have been shown to be mutated or aberrantly expressed in all kinds of cancers. For example, in bladder cancer, MEG3 overexpression promotes apoptosis and inhibits cell proliferation in BCa cells (12). LncRNA MALAT1, an oncogene in lung cancer, is expressed in association with metastasis and survival in lung cancer (13). However, because of our limited understanding of lncrnas, it is still a difficult task to identify lncrnas associated with cancer. One of the essential molecular mechanisms of lncRNAs is as competing endogenous RNAs (ceRNAs), which act as sponges for microRNAs (miRNAs). Aberrant expression of miRNAs has also been found in many types of cancer (14, 15). Additionally, lncRNAs may affect post-transcriptional modifications of mRNAs, suggesting that their mechanism of action is also related to the targeting of mRNAs. Thus, lncRNA-induced disruption of target-gene mRNA transcription is an effective strategy to identify cancer-critical functional lncRNAs (16, 17). Hence, because the lack of simultaneous analysis of lncRNA and mRNA expression levels in bladder cancer, few studies have reported the existence of lncRNA-mRNA regulatory networks associated with bladder cancer progression.

This study identified mRNAs and lncRNAs differentially expressed during bladder cancer progression based on 411 BLCA patient samples from The Cancer Genome Atlas (TCGA) database, subjected them to weighted gene co-expression network analysis (WGCNA), and confirmed mRNAs and lncRNAs associated with bladder cancer progression. A multi-step approach was used to construct a bladder cancer progression-associated lncRNA-mRNA co-expression network to reveal the potential roles of bladder cancer-associated mRNAs and lncRNAs. This study provides useful information for exploring potential candidate biomarkers for diagnosis, prognosis and drug targets in bladder cancer.

# METHODS

## Data Source Center

RNA-Seq gene expression profiles of urothelial carcinoma of the bladder (BLCA) patients were rooted in The Cancer Genome Atlas data bank, including FPKM and count formats. Clinical information, such as survival time and survival status, was downloaded from the TCGA portal. Later, we use R software to perform data extraction and sorting, and then use the coding/non coding classification provided by gencode/Ensembl, including the classification that only produces "antisense" and "lncrna", "lncRNA". Subsequently, the lncRNA and mRNA expression matrices and clinical data were got.

## Differential Expression Analysis

Differential expression of mRNA and lncRNA was studied using the edgeR package of R software. Adjusted P-values were analyzed in TCGA to correct for false-positive results. "P-value < 0.05 and $|\log FC| \geq 1$" was defined as the threshold for lncRNA and mRNA differential expression screens, and volcano plots were painted using the R package ggplot2.

## Weighted Gene Co-Expression Network Analysis

Using the WGCNA data bank of R software, Pearson correlation coefficients between genes were got from differentially expressed mRNAs and lncRNAs used for WGCNA; later, the appropriate soft threshold $\beta$ was chosen to ensure no network scalability. A one-step method was used to construct a gene network, transforming the adjacency matrix into a topological overlap matrix (TOM) and generating a hierarchical clustering tree of genes using hierarchical clustering. The DynamicTreeCut method was used to identify highly correlated co-expressed gene modules, with the threshold set to cutHeight = 0.99 and minSize = 20. The Pearson correlation test analyzed the relationship between module eigengene (ME) and clinical features.

## Functional and Pathway Enrichment Analyses

Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene ontology (GO) pathway enrichment analyses were figured out that for all genes in the blue module using Database for Annotation, Visualisation and Integrated Discovery (v6.8).

## Protein-Protein Interaction Network Construction (PPI) and Hub Gene Screening

The STRING data bank was used to identify known and predicted protein-protein interactions. STRING was also used to analyze all genes in the blue module, construct the PPI network and apply MCODE in the Cytoscape (v3.8.2) software to screen the hub genes in the PPI network.

## lncRNA-mRNA co-Expression Network Establishment

In order to clarify co-expressed lncRNA-mRNA pairs, we calculated Pearson correlation coefficients from the expression values between mRNA pairs and each differentially expressed lncRNA. The threshold of the Pearson correlation coefficient was set to > 0.5, and the corresponding FDR was set to < 0.01. In total, 248 lncRNAs and 1,308 mRNAs were identified from 57,634 co-expression relationships.

## Survival Analysis

The optimal cut-off point for risk stratification was determined using X-tile (version 3.6.1; Yale University, New Haven, CT, USA). Gene expression in BLCA was divided into high and low expression groups according to the optimal cut-off value. Kaplan-Meier survival analysis was performed on both parts, and the log-rank test difference was statistically significant ($P < 0.05$). All analyses were performed using R 3.1.0. Multivariate and univariate COX regression analyses were used to assess the relationship between survival and gene expression levels.

## Gene Set Enrichment Analysis

Usually, we use GSEA to divide the sample into two plates (low expression and high expression), so that we can determine the effect of gene expression on tumor. The screening conditions were FDR < 0.25 and $P < 0.05$.

## Cell Culture

People bladder carcinoma cell line HT-1376 and human bladder epithelial cell line HCV-29 were obtained from Cell Bank of Type Culture Collection, Chinese Academy of Sciences. The cells were cultured in RPMI-1640 medium (Invitrogen, California, USA), then added with 10% fetal bovine serum (FBS, GIBCO, California, USA), and cultured in humidified air including 5% $CO_2$ at 37°C.

## qRT-PCR

Total RNA was rooted in cells using TRIzol reagent (Invitrogen, CA, USA). RNA was reverse-transcribed to cDNA by PrimeScript RT Master Mix (Takara, Japan). RT-PCR analyses were performed using the 7500 Fast Real-Time PCR System (Applied Biosystems, Forster City, CA,USA) with the primers in **Table 1**. Data were explained by $2^{-\Delta\Delta Ct}$ method and GAPDH was the internal reference of lncRNA DEPDC-AS1, CCNB1 and CDC20.

# RESULTS

## Identification of Differentially Expressed mRNAs and lncRNAs

In all 411 cancer tissue samples and 19 paracancer tissue samples were got from the TCGA-BLCA data bank. From the results of the screening based on lncrna expression in each sample shown by volcano plots, we can see 2168 upregulated lncrnas and 1226 downregulated lncrnas (**Figure 1A**); **Figure 1B** response screens 925 upregulated mRNAs and 1454 downregulated mRNAs.

**TABLE 1** | Primer sequences.

| Genes | Primer sequences |
|---|---|
| lncRNA DEPDC-AS1 | Forward: 5'-TGGCCCCAACTCCTGGTGACT-3' |
| | Reverse: 5'-CCTGTTTTAAAAATCATTGCGGAAGCTCA-3' |
| CCNB1 | Forward: 5'-AATAAGGCGAAGATCAACATGGC-3' |
| | Reverse: 5'-TTTGTTACCAATGTCCCCAAGAG-3' |
| CDC20 | Forward: 5'-GCACAGTTCGCGTTCGAGA-3' |
| | Reverse: 5'-CTGGATTTGCCAGGAGTTCGG-3' |
| GAPDH | Forward: 5'-CTGGGCTACACTGAGCACC-3' |
| | Reverse: 5'-AAGTGGTCGTTGAGGGCAATG-3' |

## Identification of Gene Co-Expression Plate

In order to probe the co-expression patterns of mRNAs and lncRNAs in BLCA, we screened 3,394 differentially expressed lncRNAs and 2,379 differentially expressed mRNAs obtained from the TCGA database for WGCNA analysis. A power of $\beta=3$ was set as the soft threshold of the scale-free network (**Figure 2A**). As explained in **Figure 2B**, the clustering dendrogram contained five co-expression modules, denoted by turquoise, blue, brown, yellow and grey. We explored the module functions by generating an intrinsic gene neighbor-joining heat map and found that the blue module strongly correlated with other modules. Moreover, by analyzing the correlation between each module and its clinical features, we confirmed that the blue module was a significant module highly related to BLCA tumor progression (**Figures 2C, D**).

## Enrichment Analysis of Genes in the Blue Plate

In order to know the biological functions related to tumor progression in the blue plate, the co-expressed mRNAs were annotated using KEGG and GO. KEGG analysis showed that the mRNAs associated with tumor progression were significantly enriched in the cell cycle and DNA repair (**Figure 3A**); GO analysis revealed that the mRNAs associated with tumor progression were significantly enriched in the cell cycle and cell cycle progression (**Figure 3B**). The cell cycle is considered the most central function among the KEGG- and GO-enriched pathways because exchanges with other pathways strongly depend on its presence (**Figure 3C**).

## Construction of lncRNA-mRNA Co-Expression Network and PPI Network

By analyzing the lncRNA-mRNA co-expression patterns in the blue plate, 1,558 co-expression relationships, including those between 22 lncRNAs and 158 mRNAs, were obtained. The key lncRNA DEPDC1-AS1 and 152 co-expressed mRNAs were obtained by screening based on degree (**Figure 4A**). The 152 mRNAs were analyzed using the online analysis website STRING to obtain PPI protein network interactions and further visualize the gene information and network construction (**Figure 4B**). The properties of each node in the network graph were identified and visualized using the MCODE plugin in Cytoscape (**Figure 4C**) and the Top 3 MCODE; the functional enrichment is shown in **Table 2**. Thirty-three mRNAs in the largest sub-network were selected as key mRNAs. The co-expression network of lncRNA DEPDC1-AS1 and 33 hub mRNAs is shown in **Figure 4D**.

## Prognostic Analysis of 33 Co-Expressed mRNAs in Bladder Cancer

The 33 co-expressed mRNAs were subjected to univariate and multivariate analyses (**Table 3**). The consequences of the univariate COX analysis showed that none of the 33 mRNAs

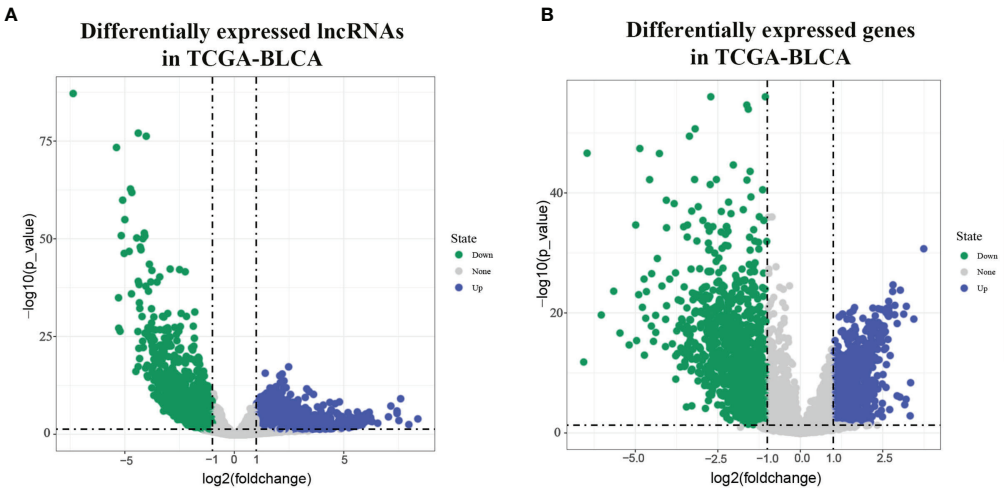



**FIGURE 1** | Differentially expressed lncRNAs and mRNAs in TCGA-BLCA. [**(A)** volcano diagram showing BLCA differentially expressed lncRNA in TCGA; **(B)** volcano diagram showing BLCA differentially expressed mRNA in TCGA].

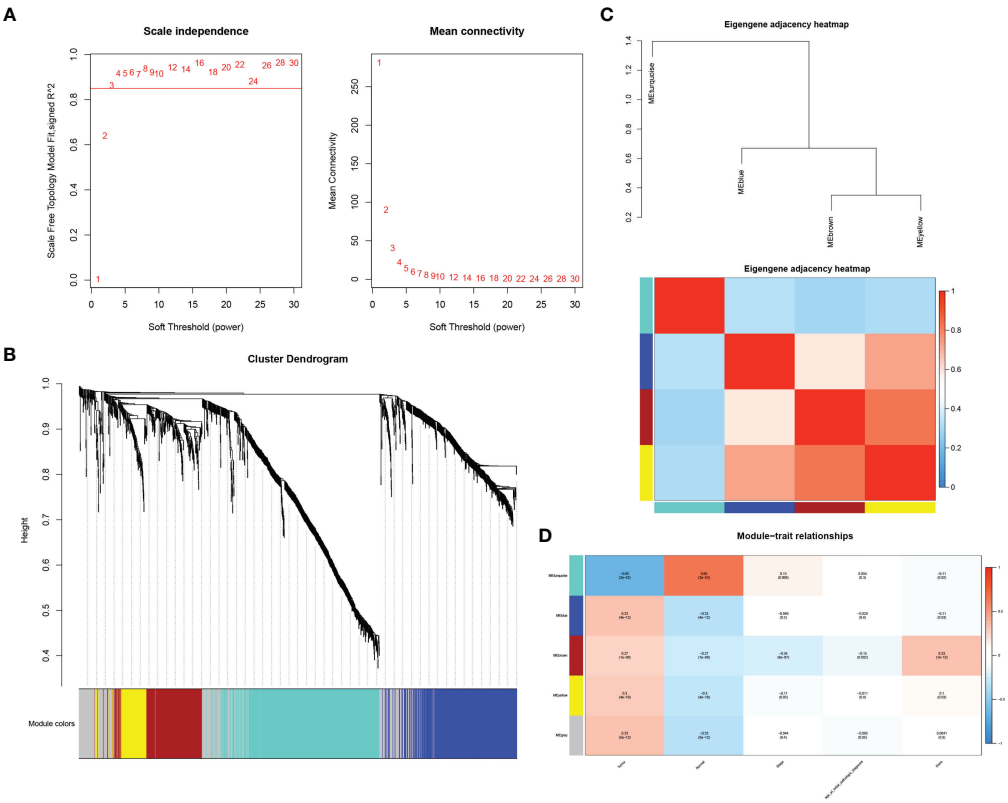



**FIGURE 2** | Construction of weighted gene co-expression network. [**(A)** selection of soft threshold β; **(B)** correlation heat map of neighboring modules in WGCNA; **(C)** clustering dendrogram of gene modules; **(D)** in the module feature heat map, each column corresponds to clinical parameters and each row corresponds to a module feature gene].

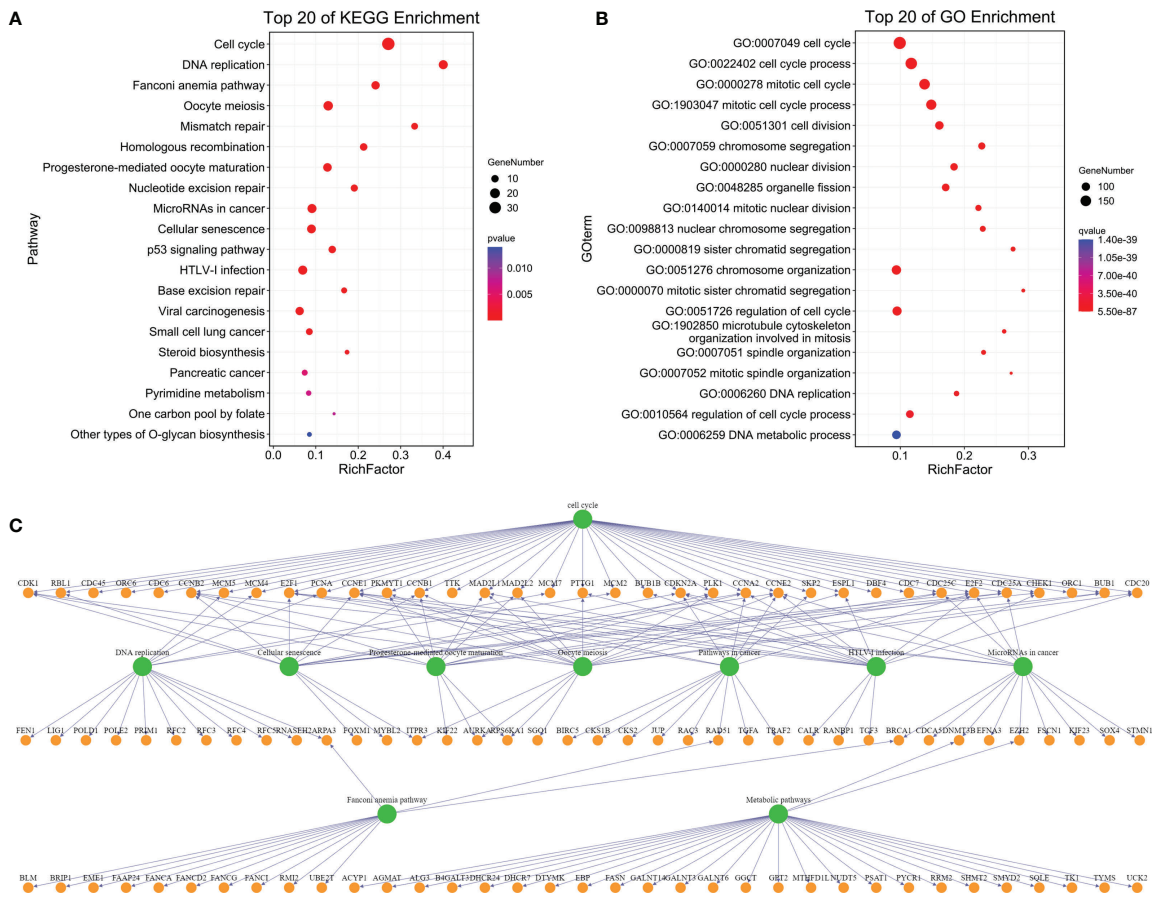



**FIGURE 3** | Enrichment analysis of genes in the blue module. [**(A)** Top 20 enrichment of KEGG pathway of the mRNA in a blue module; **(B)** Top 20 enrichment of GO pathway of the mRNA in the blue module; **(C)** Interaction and overlapping (of the top 10 pathways)].

was statistically significant. However, in the univariate analysis, due to the correlation between the independent variables, the effect of the independent variable on the dependent variable reflected not only its effect but also a comprehensive result after including the effect of the variable itself and the confounding effect of other variables. In the multivariate analysis, the regression plate was constructed to adjust for the effects of other confounding factors so that the factor's actual effect on the dependent variable was revealed. Therefore, the results of the multifactorial analysis show that cell cycle protein B1 (CCNB1) and cell division cycle 20 (CDC20) are risk factors for the prognosis of carcinoma of urinary bladder patients.

## Prognostic Analysis of lncRNA DEPDC-AS1, CCNB1 and CDC20 in BLCA

As shown in **Figures 5A–C**, lncRNA DEPDC-AS1 as well as CCNB1 and CDC20 had significantly higher expressions in carcinoma of urinary bladder tissues compared to normal paracancerous tissues. Additionally, the expression of DEPDC-AS1 was significantly and positively correlated with the expression of CCNB1 and CDC20 in BLCA (**Figures 5D, E**). DEPDC-AS1, CCNB1 and CDC20 expressions were classified into low and high expression groups according to the optimal cut-off values; the survival analysis results revealed that the high expression of DEPDC-AS1, CCNB1 and CDC20 was significantly related to poor prognosis in BLCA patients, ROC results showed that the AUC for 1 year of DEPDC-AS1, CCNB1 and CDC20 was 0.56, 0.58 and 0.63, 3 years of DEPDC-AS1, CCNB1 and CDC20 was 0.55, 0.53 and 0.65, of DEPDC-AS1, CCNB1 and CDC20 was 0.57, 0.55 and 0.56 (**Figures 5F–H**). Despite the results of AUC value was not good enough, but it might be of some value when in combination.

## Gene Set Enrichment Analysis of CDC20 and CCNB1

We wanted to find out all the functions of ccnb1 and Cdc20, using GSEA and classifying the top 5 enriched pathways (FDR Q-value < 0.250, NOM p-value < 0.050) based on the normalized enrichment score (NES). As shown in **Figure 6**, the Reactome Database enrichment results showed that both CCNB1 and CDC20 were mainly enriched in cell-cycle and DNA-replication pathways. The Hallmark enrichment results showed

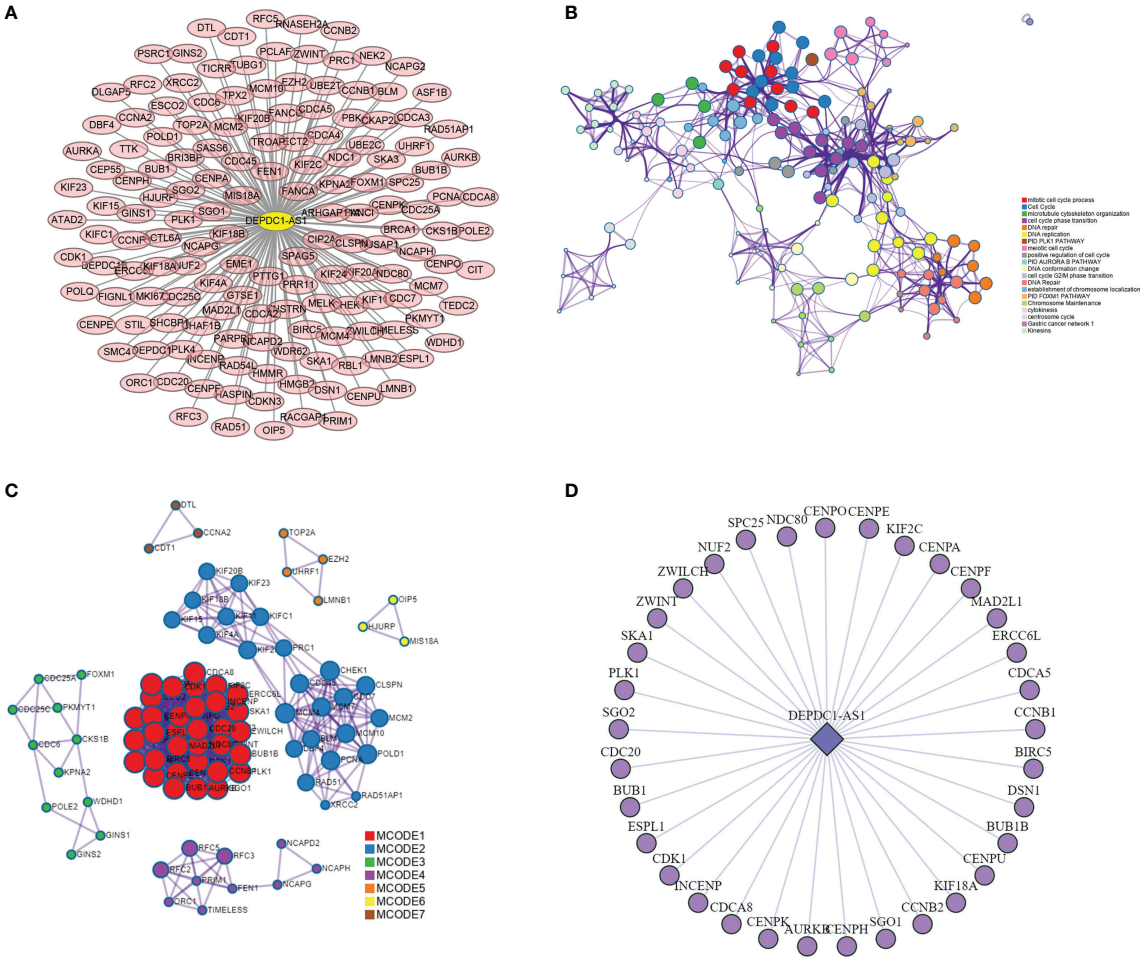



**FIGURE 4** | lncRNA-mRNA co-expression network and PPI network construction. [**(A)** lncRNA-mRNA co-expression network in the blue module; **(B)** protein-protein interaction network of co-expressed mRNAs in co-expression network; **(C)** MCODE plug-in screening out the highest-scoring sub-network; **(D)** co-expression network of lncRNA DEPDC1-AS1 and mRNAs in MCODE1].

that both CCNB1 and CDC20 were positive for G2M checkpoint and E2F targets.

## The Expression of CCNB1, lncRNA DEPDC-AS1 and CDC20 in BC Cells

The expression of lncRNA DEPDC-AS1, CCNB1 and CDC20 in carcinoma of urinary bladder cell line was determined by qRT-PCR analysis. The results figured out that the expression of lncRNA DEPDC-AS1, CCNB1 and CDC20 were significantly increased in HT-1376 cells compared with that in HCV-29 (**Figures 7A–C**). *In vitro* experiments results reflected the same results of the biology information analysis.

# DISCUSSION

Lncrnas are emerging as regulators of a wide range of biological functions. These newly characterized regulators play important and broad roles in cancer development and progression (18). However, the biological functions of most lncRNAs involved in epigenetic regulation and their role in risk stratification and prognosis have not been investigated. To date, some lncRNAs, such as UCA1, HHOTAIR and H19, have been detected in BCa cells. LncRNA overexpression can promote chemoresistance by regulating the Wnt signalling pathway and may serve as a potential diagnostic biomarker for BCa (19). However, there is a lack of comprehensive database providing resources for experimental validation of lncRNA functions. Experimental validation of the roles of the numerous lncRNAs is complex, laborious and very expensive. Biology information analysis is an approach increasingly used for target gene and protein studies. WGCNA is a systems biology approach for characterizing correlation patterns among genes in microarray samples, allowing the identification of modules of highly correlated genes for the study of potential functions (20). Recent researches have shown that WGCNA has been widely used for

**TABLE 2 |** Pathway and process enrichment analysis on Top 3 MCODE.

| MCODE | GO | Description | $Log_{10}(P)$ |
|---|---|---|---|
| MCODE_1 | R-HSA-2500257 | Resolution of Sister Chromatid Cohesion | -75.5 |
| MCODE_1 | R-HSA-68882 | Mitotic Anaphase | -69.5 |
| MCODE_1 | R-HSA-2555396 | Mitotic Metaphase and Anaphase | -69.5 |
| MCODE_2 | GO:1903047 | mitotic cell cycle process | -25.7 |
| MCODE_2 | R-HSA-176187 | Activation of ATR in response to replication stress | -20.3 |
| MCODE_2 | GO:0033260 | Nuclear DNA replication | -19.8 |
| MCODE_3 | R-HSA-69278 | Cell Cycle, Mitotic | -13.6 |
| MCODE_3 | R-HSA-1640170 | Cell Cycle | -12.8 |
| MCODE_3 | R-HSA-69242 | S Phase | -10.8 |

**TABLE 3 |** Univariate and multivariate Cox regression analysis of the 33 co-expressed mRNAs in BLCA patients.

| Variable | Univariate analysis | | | Multivariate analysis | | |
|---|---|---|---|---|---|---|
| | β | P-value | Exp (95% CI for exp) | β | P-value | Exp (95% CI for exp) |
| CENPU | 0.077 | 0.355 | 1.080 (0.918-1.271) | 0.146 | 0.326 | 1.157 (0.865-1.549) |
| CCNB1 | 0.113 | 0.168 | 1.120 (0.953-1.316) | 0.509 | 0.026 | 1.663 (1.063-2.602) |
| KIF18A | -0.004 | 0.957 | 0.996 (0.853-1.162) | -0.445 | 0.029 | 0.641 (0.429-0.956) |
| BUB1B | 0.121 | 0.145 | 1.128 (0.959-1.327) | 0.539 | 0.040 | 1.714 (1.025-2.868) |
| ERCC6L | -0.083 | 0.314 | 0.921 (0.784-1.081) | -0.328 | 0.013 | 0.720 (0.555-0.934) |
| CDC20 | 0.111 | 0.123 | 1.117 (0.970-1.287) | 0.420 | 0.046 | 1.522 (1.007-2.300) |
| BUB1 | -0.001 | 0.988 | 0.999 (0.855-1.167) | -0.361 | 0.110 | 0.697 (0.448-1.086) |
| CCNB2 | 0.049 | 0.537 | 1.051 (0.898-1.229) | -0.332 | 0.186 | 0.718 (0.439-1.174) |
| AURKB | 0.005 | 0.946 | 1.005 (0.866-1.166) | -0.222 | 0.238 | 0.801 (0.554-1.158) |
| BIRC5 | 0.087 | 0.273 | 1.090 (0.934-1.273) | 0.159 | 0.479 | 1.172 (0.755-1.820) |
| CDCA8 | 0.058 | 0.462 | 1.060 (0.908-1.237) | -0.034 | 0.889 | 0.967 (0.601-1.554) |
| CENPA | 0.057 | 0.433 | 1.059 (0.918-1.222) | 0.152 | 0.509 | 1.164 (0.742-1.828) |
| CENPE | 0.042 | 0.616 | 1.043 (0.885-1.229) | -0.171 | 0.439 | 0.843 (0.547-1.299) |
| CDK1 | 0.015 | 0.858 | 1.015 (0.860-1.198) | -0.174 | 0.423 | 0.841 (0.550-1.286) |
| CDCA5 | 0.071 | 0.359 | 1.074 (0.922-1.250) | -0.140 | 0.546 | 0.870 (0.552-1.369) |
| CENPF | 0.054 | 0.468 | 1.055 (0.913-1.219) | 0.030 | 0.887 | 1.030 (0.684-1.552) |
| CENPO | 0.050 | 0.653 | 1.051 (0.845-1.307) | -0.085 | 0.716 | 0.918 (0.580-1.453) |
| CENPK | 0.005 | 0.955 | 1.005 (0.846-1.194) | -0.290 | 0.100 | 0.748 (0.530-1.057) |
| CENPH | 0.015 | 0.879 | 1.015 (0.837-1.232) | -0.020 | 0.908 | 0.980 (0.697-1.377) |
| DSN1 | 0.060 | 0.590 | 1.061 (0.855-1.318) | 0.033 | 0.864 | 1.033 (0.711-1.502) |
| KIF2C | 0.063 | 0.440 | 1.065 (0.908-1.249) | -0.193 | 0.510 | 0.824 (0.465-1.463) |
| MAD2L1 | 0.027 | 0.744 | 1.027 (0.873-1.209) | 0.002 | 0.992 | 1.002 (0.685-1.465) |
| INCENP | 0.128 | 0.159 | 1.137 (0.951-1.358) | 0.437 | 0.058 | 1.547 (0.986-2.429) |
| ESPL1 | 0.087 | 0.276 | 1.091 (0.933-1.277) | 0.127 | 0.496 | 1.136 (0.787-1.638) |
| NDC80 | 0.029 | 0.696 | 1.030 (0.890-1.191) | 0.005 | 0.978 | 1.005 (0.707-1.428) |
| NUF2 | -0.022 | 0.746 | 0.978 (0.856-1.118) | -0.240 | 0.133 | 0.786 (0.575-1.076) |
| SGO2 | 0.077 | 0.432 | 1.080 (0.891-1.309) | 0.078 | 0.758 | 1.082 (0.657-1.781) |
| SGO1 | -0.044 | 0.570 | 0.957 (0.821-1.115) | -0.189 | 0.292 | 0.828 (0.582-1.177) |
| PLK1 | 0.099 | 0.177 | 1.104 (0.956-1.276) | 0.156 | 0.370 | 1.169 (0.831-1.644) |
| SKA1 | 0.097 | 0.232 | 1.102 (0.940-1.292) | 0.392 | 0.052 | 1.480 (0.996-2.199) |
| SPC25 | 0.014 | 0.860 | 1.015 (0.864-1.191) | -0.103 | 0.623 | 0.902 (0.598-1.360) |
| ZWINT | 0.045 | 0.583 | 1.046 (0.890-1.230) | 0.016 | 0.933 | 1.016 (0.702-1.470) |
| ZWILCH | 0.083 | 0.493 | 1.086 (0.857-1.377) | 0.187 | 0.415 | 1.205 (0.770-1.886) |

the screening and identification of disease susceptibility genes and candidate targets (21, 22). Another method is that we can also integrate the expression profiles of protein coding genes and lncrna into the co-expression model, so as to study the characteristics of lncrna in different biological processes and cancers.

Numerous researched have figured out that the occurrence of BCa is highly related to the abnormal expression of non-coding RNAs and protein-coding genes. Wang et al. combined the miRNA mRNA regulatory network, lncrna miRNA regulatory network and lncrna mRNA co expression network to obtain a three-layer network, and then calculated the topological characteristics of each node in the network, including degree, compactness and intermediation. They also identified mirna-93 and mirna-195 as controllers of three-layer networks related to BCA and regulators of many target genes, the imbalance of these target genes may be closely related to the pathogenesis of BCa (23). However, few studies have reported the existence of a lncRNA-mRNA co-expression network associated with BCa progression. LncRNAs are more likely to be co-expressed with

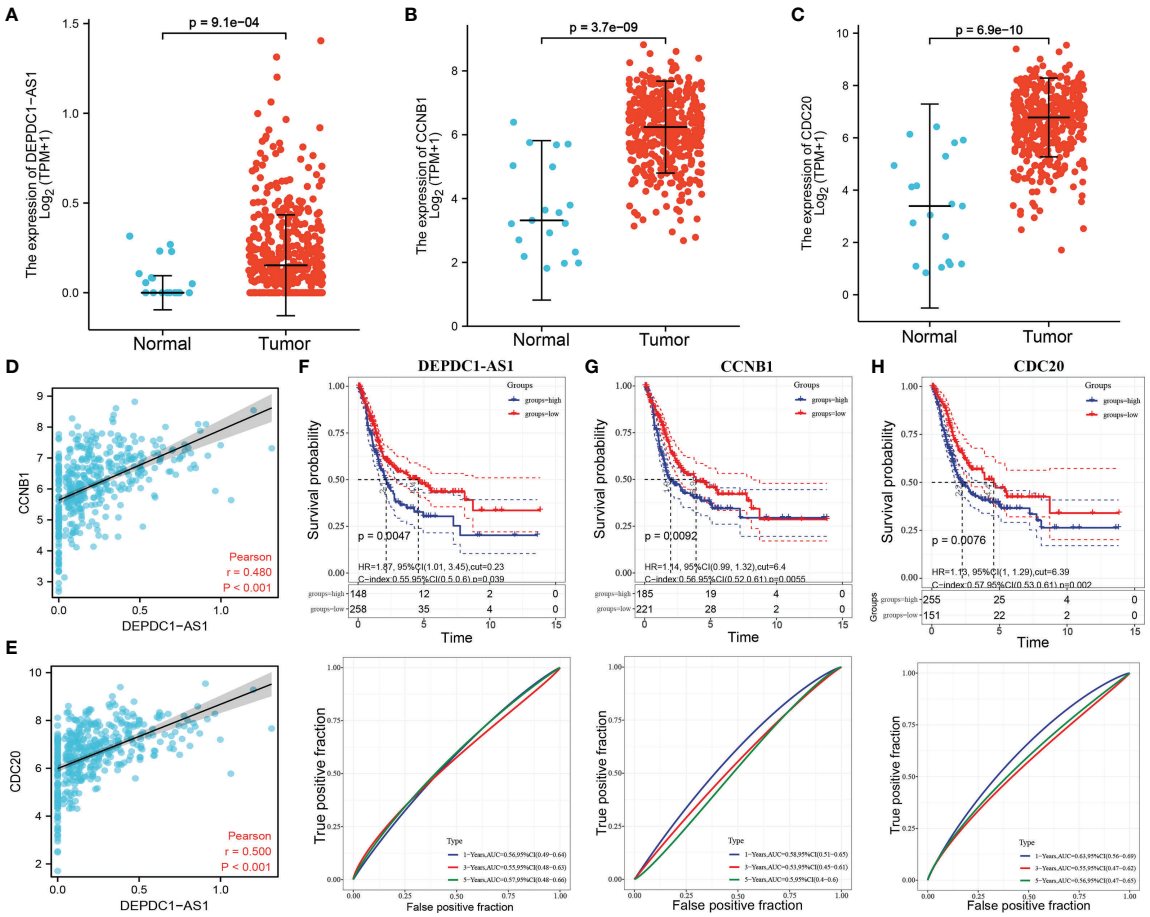



**FIGURE 5** | Expression and prognostic analysis of lncRNA and mRNA in BLCA. [**(A)** DEPDC1-AS1 expression in BLCA; **(B)** CCNB1 expression in BLCA; **(C)** CDC20 expression in BLCA; **(D)** correlation between DEPDC1-AS1 and CCNB1 expression in BLCA; **(E)** correlation between DEPDC1-AS1 and CDC20 expression in BLCA; **(F)** KM curve of DEPDC1- AS1 expression with KM curve and time-dependent ROC curve of overall survival; **(G)** KM curve and time-dependent ROC curve of CCNB1 expression with overall survival; **(H)** KM curve and time-dependent ROC curve of CCNB1 expression with overall survival].

neighboring coding genes through a cis-regulatory mechanism (24). Previous studies have shown that some lncRNAs can be co-expressed with the corresponding coding genes based on the ceRNA theory (17). Based on this theory, we analyzed and characterized the lncRNA-mRNA co-expression network in bladder cancer.

This manuscript used differentially expressed mRNAs and lncRNAs in TCGA-BLCA for WGCNA to build a functional lncRNA-mRNA co-expression network related to carcinoma of urinary bladder progression. A key lncRNA (DEPDC1-AS1) and two mRNAs (CCNB1 & CDC20) were then identified *via* univariate and multivariate COX analyses. Studies targeting DEPDC1-AS1 alone in cancer have not been reported. However, LuLu et al. (25) first demonstrated that DEPDC1-AS1 was associated with the prognosis of lung adenocarcinoma, and Yuan C et al. (26) constructed a new triple-length non-coding RNA risk scoring system for predicting the prognosis of triple-negative breast cancer, which included DEPDC1-AS1. While the present study demonstrated for the first time that DEPDC1-AS1 was associated with the prognosis of BLCA, its relevant mechanism of action in BLCA needs to be further explored. CCNB1 is a monitoring protein that initiates the process from the G2 phase to mitosis. Numerous studies have reported that CCNB1 is overexpressed in many tumors—hepatocellular carcinoma (27), colon cancer (28), and pancreatic cancer (29)—and promotes tumor cell proliferation. A recent study identified CCNB1 as a potential target and a new key biomarker for the prognosis of human BCa (30); this study's results confirm this. In cell cycle progression, CDC20 is an important control factor of cell cycle checkpoints. Its most important function is to bind to the late promoter complex (APC/C), which in turn regulates the degradation of the isolate inhibitor protein (31, 32). Thus, dysregulation of CDC20 may have a considerable impact on cell growth and tumorigenesis. Furthermore, many studies have figured out that CDC20 is a carcinogen that promotes cancer development (33, 34).

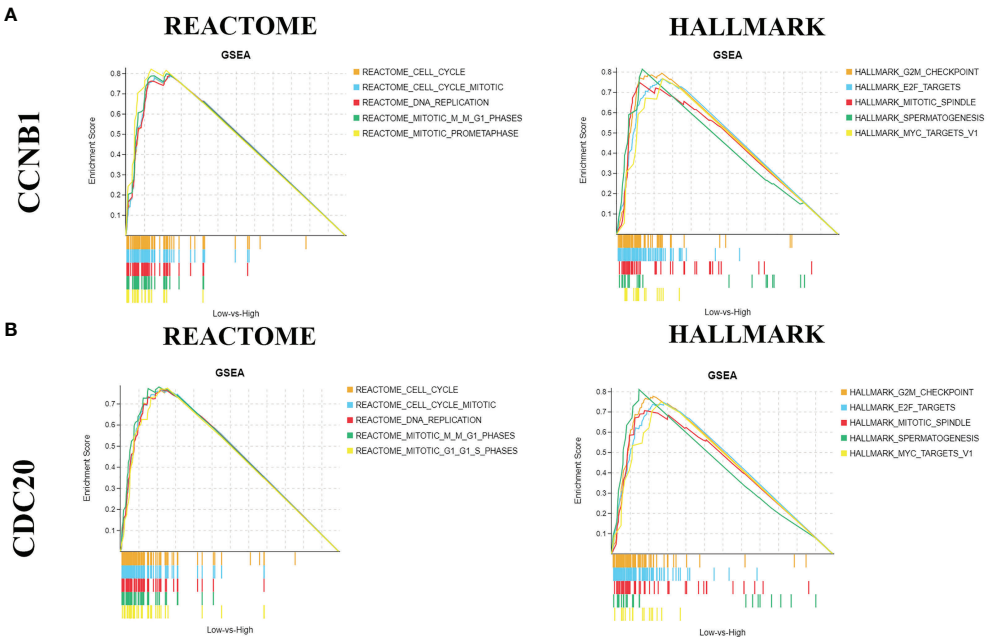



**FIGURE 6** | Gene set enrichment analysis. [**(A)** REACTOME and HALLMARK pathway analysis of CCNB1; **(B)** BREACTOME and HALLMARK pathway analysis of CDC20].

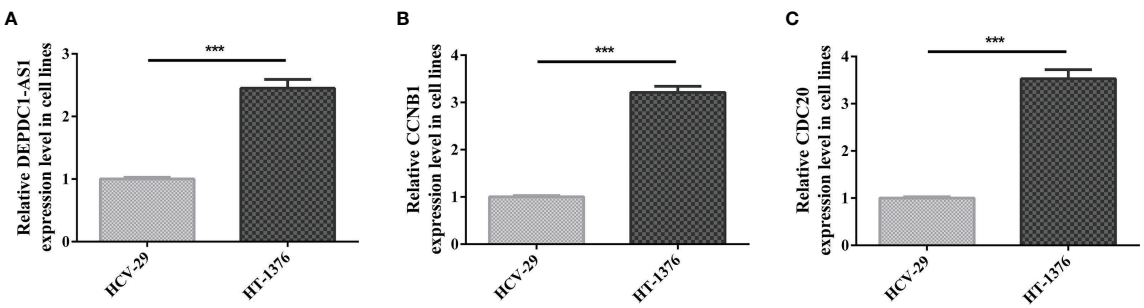



**FIGURE 7** | Relative expression of genes quantified by qRT-PCR. [**(A)** Relative expression level of lncRNA DEPDC-AS1; **(B)** Relative expression level of CCNB1; **(C)** Relative expression level of CDC20; ***P < 0.001].

Choi et al. (35) found elevated CDC20 expression in patients with uroepithelial bladder cancer (UBC) and that high CDC20 expression in UBC patients was related to short recurrence-free survival and poor overall survival. The results of this study confirmed these findings. Additionally, the gene set enrichment analysis results showed that the CCNB1 and CDC20 enrichment pathways are similar in that both act on the regulatory mechanisms of BCa development by regulating the progression of the cell cycle.

In addition, lncRNA can play multiple roles in the regulation of its target genes. Firstly, lncRNA can play a role through cis or trans regulation of its target genes (36, 37). A single lncRNA may regulate multiple target genes through different mechanisms. Secondly, lncRNA can also be used as a guide to promote DNA protein interaction. It should be noted that lncRNA can play a role as enhancer RNA as well when recruiting protein complexes to induce DNA cyclization and target gene transcriptional activation (38). Thirdly, lncRNAs can also be used as bait to bind to proteins involved in transcription to prevent them from binding to DNA target proteins, or as competitive endogenous RNA (ceRNAs) to bind to miRNAs to prevent negative regulation of target genes (39). The binding of lncRNA to mRNA also leads to intron retention, which promotes the selective splicing of mRNAs (40). Chen et al. (41), found that

LncRNA PVT1 accelerates malignant phenotypes of bladder cancer cells by modulating miR-194-5p/BCLAF1 axis as a ceRNA. However, there were no reports on the mechanism of how DEPDC1-AS1 works in the regulation of the cancer progression. Luckily, this provide us a novel insights into and a new research direction of the DEPDC1-AS1 in cancers.

In conclusion, our analysis by synthesis provided new insights into the lncRNA-mRNA co-expression network in bladder cancer progression. It indicates that lncRNA plays an important role in bladder cancer. LncRNA-DEPDC1-AS1 may serve as a biomarker to predict survival in BCa patients. Moreover, it may mediate the BCa cell cycle through a co-expression network involving CCNB1 and CDC20, thus affecting BCa progression. In addition, the biological function of lncRNA-DEPDC1-AS1 requires further laboratory characterization and study. Therefore, the results and conclusions obtained in this study may provide an essential theoretical basis for future experimental studies on lncRNA's role in bladder cancer.

## DATA AVAILABILITY STATEMENT

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

## ETHICS STATEMENT

Ethical approval/written informed consent was not required for the study of animals/human participants in accordance with the local legislation and institutional requirements.

## AUTHOR CONTRIBUTIONS

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

## REFERENCES


1. Kulkarni GS, Black PC, Sridhar SS, Kapoor A, Zlotta AR, Shayegan B, et al. Canadian Urological Association Guideline: Muscle-Invasive Bladder Cancer. *Can Urol Assoc J* (2019) 13(8):230. doi: 10.5489/cuaj.5902
2. Antoni S, Ferlay J, Soerjomataram I, Znaor A, Jemal A, Bray F. Bladder Cancer Incidence and Mortality: A Global Overview and Recent Trends. *Eur Urol* (2017) 71(1):96–108. doi: 10.1016/j.eururo.2016.06.010
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
4. Guidance N. Bladder Cancer: Diagnosis and Management of Bladder Cancer. *BJU Int* (2017) 120(6):755–65.
5. Aggen DH, Drake CG. Biomarkers for Immunotherapy in Bladder Cancer: A Moving Target. *J Immunother Cancer* (2017) 5(1):1–13. doi: 10.1186/s40425-017-0299-1
6. Chen C-L, Liu C-Y, Cha T-L, Hsu C-Y, Chou Y-C, Wu S-T, et al. Does Radical Cystectomy Outperform Other Bladder Preservative Treatments in Elderly Patients With Advanced Bladder Cancer? *J Chin Med Assoc* (2015) 78(8):469–74. doi: 10.1016/j.jcma.2015.05.003
7. Jordan EJ, Iyer G. Targeted Therapy in Advanced Bladder Cancer: What Have We Learned? *Urol Clinics* (2015) 42(2):253–62. doi: 10.1016/j.ucl.2015.01.006
8. Zhang C, Berndt-Paetz M, Neuhaus J. Identification of Key Biomarkers in Bladder Cancer: Evidence From a Bioinformatics Analysis. *Diagnostics* (2020) 10(2):66. doi: 10.3390/diagnostics10020066
9. Li L, Peng Q, Gong M, Ling L, Xu Y, Liu Q. Using lncRNA Sequencing to Reveal a Putative lncRNA-mRNA Correlation Network and the Potential Role of PCBP1-AS1 in the Pathogenesis of Cervical Cancer. *Front Oncol* (2021) 11. doi: 10.3389/fonc.2021.634732
10. Chen C-T, Wang P-P, Mo W-J, Zhang Y-P, Zhou W, Deng T-F, et al. Expression Profile Analysis of Prognostic Long Non-Coding RNA in Adult Acute Myeloid Leukemia by Weighted Gene Co-Expression Network Analysis (WGCNA). *J Cancer* (2019) 10(19):4707. doi: 10.7150/jca.31234
11. Bhan A, Mandal SS. LncRNA HOTAIR: A Master Regulator of Chromatin Dynamics and Cancer. *Biochim Biophys Acta (BBA) Rev Cancer* (2015) 1856 (1):151–64. doi: 10.1016/j.bbcan.2015.07.001
12. Zhang X, Gejman R, Mahta A, Zhong Y, Rice KA, Zhou Y, et al. Maternally Expressed Gene 3, an Imprinted Noncoding RNA Gene, Is Associated With Meningioma Pathogenesis and Progression. *Cancer Res* (2010) 70(6):2350–8. doi: 10.1158/0008-5472.CAN-09-3885
13. Gutschner T, Hämmerle M, Eißmann M, Hsu J, Kim Y, Hung G, et al. The Noncoding RNA MALAT1 Is a Critical Regulator of the Metastasis Phenotype of Lung Cancer Cells. *Cancer Res* (2013) 73(3):1180–9. doi: 10.1158/0008-5472.CAN-12-2850
14. Ruff SM, Ayabe RI, Malekzadeh P, Good ML, Wach MM, Gonzales MK, et al. MicroRNA-210 May Be a Preoperative Biomarker of Malignant Pheochromocytomas and Paragangliomas. *J Surg Res* (2019) 243:1–7. doi: 10.1016/j.jss.2019.04.086
15. Jiang J, Bi Y, Liu XP, Yu D, Yan X, Yao J, et al. To Construct a ceRNA Regulatory Network as Prognostic Biomarkers for Bladder Cancer. *J Cell Mol Med* (2020) 24(9):5375–86. doi: 10.1111/jcmm.15193
16. Liu Y, Zhang R, Qiu F, Li K, Zhou Y, Shang D, et al. Construction of a lncRNA–PCG Bipartite Network and Identification of Cancer-Related lncRNAs: A Case Study in Prostate Cancer. *Mol Biosyst* (2015) 11(2):384–93. doi: 10.1039/C4MB00439F
17. Wang P, Ning S, Zhang Y, Li R, Ye J, Zhao Z, et al. Identification of lncRNA-Associated Competing Triplets Reveals Global Patterns and Prognostic Markers for Cancer. *Nucleic Acids Res* (2015) 43(7):3478–89. doi: 10.1093/nar/gkv233
18. Fan Y, Shen B, Tan M, Mu X, Qin Y, Zhang F, et al. Long Non-Coding RNA UCA 1 Increases Chemoresistance of Bladder Cancer Cells by Regulating Wnt Signaling. *FEBS J* (2014) 281(7):1750–8. doi: 10.1111/febs.12737
19. Guo X, Gao L, Liao Q, Xiao H, Ma X, Yang X, et al. Long Non-Coding RNAs Function Annotation: A Global Prediction Method Based on Bi-Colored Networks. *Nucleic Acids Res* (2013) 41(2):e35–e. doi: 10.1093/nar/gks967
20. Langfelder P, Horvath S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf* (2008) 9(1):1–13. doi: 10.1186/1471-2105-9-559
21. Liu Z, Li M, Hua Q, Li Y, Wang G. Identification of an Eight-lncRNA Prognostic Model for Breast Cancer Using WGCNA Network Analysis and a Cox–Proportional Hazards Model Based on L1-Penalized Estimation. *Int J Mol Med* (2019) 44(4):1333–43. doi: 10.3892/ijmm.2019.4303
22. Ren Z-H, Shang G-P, Wu K, Hu C-Y, Ji T. WGCNA Co-Expression Network Analysis Reveals ILF3-AS1 Functions as a CeRNA to Regulate PTBP1 Expression by Sponging miR-29a in Gastric Cancer. *Front Genet* (2020) 11:39. doi: 10.3389/fgene.2020.00039
23. Wang X, Ding Y, Wang J, Wu Y. Identification of the Key Factors Related to Bladder Cancer by lncRNA-miRNA-mRNA Three-Layer Network. *Front Genet* (2020) 10:1398. doi: 10.3389/fgene.2019.01398
24. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative Annotation of Human Large Intergenic Noncoding RNAs Reveals Global Properties and Specific Subclasses. *Genes Dev* (2011) 25(18):1915–27. doi: 10.1101/gad.17446611
25. Lu L, Liu L-P, Zhao Q-Q, Gui R, Zhao Q-Y. Identification of a Ferroptosis-Related LncRNA Signature as a Novel Prognosis Model for Lung Adenocarcinoma. *Front Oncol* (2021) 11:2443. doi: 10.3389/fonc.2021.675545
26. Yuan C, Yuan H, Chen L, Sheng M, Tang W. A Novel Three-Long Noncoding RNA Risk Score System for the Prognostic Prediction of Triple-Negative Breast Cancer. *Biomarkers Med* (2020) 15(1):43–55. doi: 10.2217/bmm-2020-0505
27. Zou Y, Ruan S, Jin L, Chen Z, Han H, Zhang Y, et al. CDK1, CCNB1, and CCNB2 Are Prognostic Biomarkers and Correlated With Immune Infiltration

in Hepatocellular Carcinoma. *Med Sci Monit: Int Med J Exp Clin Res* (2020) 26:e925289–1. doi: 10.12659/MSM.925289
28. Lei X, Jing J, Zhang M, Guan B, Dong Z, Wang C. Bioinformatic Identification of Hub Genes and Analysis of Prognostic Values in Colorectal Cancer. *Nutr Cancer* (2021) 73(11-12), 2568–78. doi: 10.1080/01635581.2020.1841249
29. Zhang H, Zhang X, Li X, Meng WB, Bai ZT, Rui SZ, et al. Effect of CCNB1 Silencing on Cell Cycle, Senescence, and Apoptosis Through the P53 Signaling Pathway in Pancreatic Cancer. *J Cell Physiol* (2019) 234(1):619–31. doi: 10.1002/jcp.26816
30. Fang H, Huang Z, Zeng X, Wan J, Wu J, Di J. SMC4, CCNB1 and CKS1B as Potential Targets and New Critical Biomarkers for the Prognosis of Human Bladder Cancer. (2021). doi: 10.21203/rs.3.rs-595148/v1
31. Wu F, Lin Y, Cui P, Li H, Zhang L, Sun Z, et al. Cdc20/p55 Mediates the Resistance to Docetaxel in Castration-Resistant Prostate Cancer in a Bim-Dependent Manner. *Cancer Chemother Pharmacol* (2018) 81(6):999–1006. doi: 10.1007/s00280-018-3578-8
32. Wang L, Zhang J, Wan L, Zhou X, Wang Z, Wei W. Targeting Cdc20 as a Novel Cancer Therapeutic Strategy. *Pharmacol Ther* (2015) 151:141–51. doi: 10.1016/j.pharmthera.2015.04.002
33. Tang J, Lu M, Cui Q, Zhang D, Kong D, Liao X, et al. Overexpression of ASPM, CDC20, and TTK Confer a Poorer Prognosis in Breast Cancer Identified by Gene Co-Expression Network Analysis. *Front Oncol* (2019) 9:310. doi: 10.3389/fonc.2019.00310
34. Zhuang L, Yang Z, Meng Z. Upregulation of BUB1B, CCNB1, CDC7, CDC20, and MCM3 in Tumor Tissues Predicted Worse Overall Survival and Disease-Free Survival in Hepatocellular Carcinoma Patients. *BioMed Res Int* (2018) 2018. doi: 10.1155/2018/7897346
35. Choi J-W, Kim Y, Lee J-H, Kim Y-S. High Expression of Spindle Assembly Checkpoint Proteins CDC20 and MAD2 Is Associated With Poor Prognosis in Urothelial Bladder Cancer. *Virchows Archiv* (2013) 463(5):681–7. doi: 10.1007/s00428-013-1473-6
36. Guttman M, Rinn JL. Modular Regulatory Principles of Large Non-Coding RNAs. *Nature* (2012) 482(7385):339–46. doi: 10.1038/nature10887
37. Geisler S, Coller J. RNA in Unexpected Places: Long Non-Coding RNA Functions in Diverse Cellular Contexts. *Nat Rev Mol Cell Biol* (2013) 14 (11):699–712. doi: 10.1038/nrm3679
38. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long Noncoding RNAs With Enhancer-Like Function in Human Cells. *Cell* (2010) 143(1):46–58. doi: 10.1016/j.cell.2010.09.001
39. Chiu H-S, Somvanshi S, Patel E, Chen T-W, Singh VP, Zorman B, et al. Pan-Cancer Analysis of lncRNA Regulation Supports Their Targeting of Cancer Genes in Each Tumor Context. *Cell Rep* (2018) 23(1):297–312. e12. doi: 10.1016/j.celrep.2018.03.064
40. Romero-Barrios N, Legascue MF, Benhamed M, Ariel F, Crespi M. Splicing Regulation by Long Noncoding RNAs. *Nucleic Acids Res* (2018) 46(5):2169–84. doi: 10.1093/nar/gky095
41. Chen M, Zhang R, Le Lu JD, Chen C, Ding K, Wei X, et al. LncRNA PVT1 Accelerates Malignant Phenotypes of Bladder Cancer Cells by Modulating miR-194-5p/BCLAF1 Axis as a ceRNA. *Aging (Albany NY)* (2020) 12 (21):22291. doi: 10.18632/aging.202203

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 03 March 2022
doi: 10.3389/fonc.2022.834235

# The Cell Cycle-Associated Protein CDKN2A May Promotes Colorectal Cancer Cell Metastasis by Inducing Epithelial-Mesenchymal Transition

*Wei-Kun Shi*[1†], *Yun-Hao Li*[1†], *Xue-Shan Bai*[2] *and Guo-Le Lin*[1*]

[1] Department of General Surgery, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China, [2] Plastic Surgery Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Colorectal cancer (CRC) is a common gastrointestinal malignancy, and recurrence and metastasis contribute considerably to its high mortality. It is well known that the epithelial-mesenchymal transition (EMT) accelerates the rate of cancer cell dissemination and migration, thus promoting cancer metastasis. Targeted therapy is a common modality for cancer treatment, and it can play a role in inhibiting cancer progression. In this study, bioinformatics was used to search for genes associated with the prognosis of CRC. First, differential analysis was performed on colon and rectal cancer samples to obtain 2,840 and 3,177 differentially expressed genes (DEGs), respectively. A Venn diagram was then used to identify 262 overlapping genes from the two groups of DEGs and EMT-related genes. The overlapping genes were subjected to batch survival analysis and batch expression analysis successively, and nine genes were obtained whose high expression in CRC led to a poor prognosis. The least absolute shrinkage and selection operator (LASSO) prognostic model was then constructed to obtain the risk score formula. A nomogram was constructed to seek prognostic independent factors to obtain CDKN2A. Finally, CCK-8 assay, flow cytometry and western blotting assays were performed to analyze the cellular biological function of CDKN2A. The results showed that knockdown of CDKN2A expression inhibited HT-29 cell proliferation, promoted apoptosis and cell cycle progression, and affected the EMT process in CRC.

**Keywords: CDKN2A, colorectal cancer, epithelial-mesenchymal transition, cancer metastasis, cell proliferation and migration**

OPEN ACCESS

***Edited by:***
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

***Reviewed by:***
*Liu He,*
*Wuhan University, China*
*Peng Xiao,*
*Zhejiang University, China*

***Correspondence:**
*Guo-Le Lin*
*linguole@126.com*

*[†]These authors have contributed equally to this work*

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

***Received:*** *13 December 2021*
***Accepted:*** *31 January 2022*
***Published:*** *03 March 2022*

***Citation:***
*Shi W-K, Li Y-H, Bai X-S and Lin G-L (2022) The Cell Cycle-Associated Protein CDKN2A May Promotes Colorectal Cancer Cell Metastasis by Inducing Epithelial-Mesenchymal Transition. Front. Oncol. 12:834235. doi: 10.3389/fonc.2022.834235*

## INTRODUCTION

Colorectal cancer (CRC) is a common gastrointestinal malignancy and the leading cause of cancer deaths. The 2020 global cancer statistics show that there were nearly 1.9 million new cases of CRC (9.8% of cases) and more than 900,000 deaths (9.2% of cases) (1). The five-year survival rate for CRC is highly stage-dependent: early-stage survival rates exceed 90%, while advanced-stage rates reach only 10% (2). This is because early-stage CRC is usually not easily detectable and is not diagnosed until it has spread substantially. Patients with CRC often exhibit symptoms such as

dyspareunia, colorectal bleeding, and changes in gastrointestinal motility (3). The causes of the disease are mostly related to poor lifestyle habits, for example, smoking, an unhealthy diet, alcohol abuse, and a lack of physical activity (4). Fortunately, the risk of CRC can be reduced through effective preventive measures, such as cancer screening and a healthy lifestyle (5). However, even though many screening tools can be used to detect and reduce the incidence of CRC, nearly a quarter of cases are still diagnosed with advanced metastases (6). Statistically, the five-year survival rate of patients with metastatic CRC is less than 15%, which is a significant concern (7).

The key characteristic of epithelial-mesenchymal transition (EMT) is the acquisition of a mesenchymal phenotype by epithelial cells (8). In a normal physiological context, EMT promotes developmental processes and wound healing (9). However, in tumor progression, EMT is a dynamic process. Cancer cells adjust their metabolism to meet their growth demands, and certain metabolic pathways involved can directly contribute to EMT (10). In this context, EMT promotes cancer cell dissemination and migration by enhancing cell mobility, which leads to cancer metastasis (11). The hallmark alterations observed in cells undergoing EMT are reduced E-cadherin expression and enhanced N-cadherin and vimentin expression, which usually occur prior to tumor invasion (12). These modifications are mediated by multiple transcription factors (TFs) that directly repress E-cadherin expression and promote the change of epithelial cells to a mesenchymal state (13). In a variety of cancers, including CRC, EMT confers metastatic and stem cell properties on cancer cells (14). There is also evidence that EMT is associated with drug resistance to multiple drugs (15). Therefore, it is important to develop EMT inhibitors for cancer therapy.

Targeted therapies play a unique role in inhibiting cancer progression by directly inhibiting cell proliferation, differentiation, and migration (16). The activity of the TFs involved in EMT can be regulated by a variety of kinases, and their signaling pathways can also serve as effective therapeutic targets. Previous studies have shown that the activation of certain signaling pathways can promote EMT in cancer cells (17). In addition, EMT can be induced by proteins and selectively trigger gene expression programs in cancer stem cells (18).

Therefore, in this study, we focused on identifying EMT-related genes that are potentially associated with CRC prognosis and conducted *in vitro* experiments to validate the biological functions of these potential genes.

# MATERIALS AND METHODS

## Data Sources

Gene expression matrices and RNA-seq data of colon and rectal cancers were obtained from the TCGA database, as well as mRNA expression data of the corresponding normal tissue samples. The data contained 455 colon cancer samples and 165 rectal cancer samples, and they corresponded to 41 and 10 normal tissue samples, respectively. EMT-related gene sets were downloaded from the Molecular Signatures Database (MsigDB, https://www.gsea-msigdb.org/gsea/msigdb/) and contained 1,263 genes.

## Differential Analysis

The R package Limma was used to study the differential expression of mRNAs, and $P < 0.05$ and $|\log2FC| > 1$ were defined as the screening threshold for differential expression. The intersection of differentially expressed genes (DEGs) and EMT-related genes was determined using a Venn diagram constructed to obtain overlapping genes associated with colon cancer, rectal cancer, and EMT.

## Construction of the LASSO Prognostic Model

The prognostic feature model was constructed using the R package glmnet to investigate the relationship between genes and CRC prognosis. The least absolute shrinkage and selection operator (LASSO) regression algorithm was used to select features using 10-fold cross-validation. The LASSO prognostic model is a risk score formula that includes multiple genes. The included samples were divided into two groups: high risk and low risk. The log-rank test was used for the survival analysis and to compare the survival difference between the two groups, and timeROC analysis was used to compare the predictive accuracy and risk scores of the pivotal genes.

## Survival Analysis

Gene expression levels in cancer versus para cancer cells were compared using the t-test, and the R package ggplot2 was used for picture plotting. The R package survival was used to analyze survival differences between high- and low-expression groups, and the results are presented as KM survival curves, with P values and hazard ratios (HR) with 95% confidence intervals (95% CI) obtained by log-rank test and univariate Cox proportional hazards regression. Univariate and multivariate Cox analyses were used to identify prognostic independent factors. The R package forest plot was used to draw forest plots, and the R package "rms" was used to create column line plots (nomograms). P value < 0.05 was considered statistically significant.

## Cell Culture

Human HT-29 CRC cells (HTB-38$^{TM}$) were purchased from the American Type Culture Collection (ATCC). The HT-29 cell line was cultured in McCoy's 5A medium containing 10% FBS in a humidified incubator with 5% $CO_2$ at 37°C.

## SiRNA Transfection

siRNA specifically targeting CDKN2A were designed and synthesis by GenePharma Corporation (Shanghai, China), the sequence of sense is GGGUCCCAGUCUGCAGUUATT, the sequence of antisense is UAACUGCAGACUGGGACCCTT. Cells were added to 6-well plates and cultured for 24 h. siRNA was transfected into cells using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. The medium was replaced after 6 h with fresh medium containing 10% FBS, and the cells were collected after 24 h for

the following assays. The negative control group consisted of si-NC transfection-treated HT-29 cells, which were cultured in parallel, as described above.

## Cell Proliferation Assay

Cells were added to 96-well plates and incubated overnight for cell transfection. They continued to be incubated at 37°C in a humid atmosphere of 5% $CO_2$. CCK-8 reagent was added to the wells 0, 6, 12, 24, 48, or 72 h post-transfection, and the cells were incubated for 2 h. The absorbance values (optical density, OD) were measured at 450 nm using an enzyme marker.

## Western Blotting

Cells were lysed using RIPA buffer, and the total protein concentration was determined using the BCA assay. Total proteins were separated using 10% SDS-PAGE and transferred to polyvinylidene difluoride membranes. The membranes were blocked with 5% skim milk diluted in TBST for 1 h and then incubated with primary antibodies overnight at 4°C. The primary antibodies were anti-CDKN2A (ab185620, abcam), anti-N-cadherin (ab245117, abcam), anti-E-cadherin (ab40772, abcam), anti-vimentin (ab92547, abcam), and anti-β-actin (ab8226, abcam). After washing, the secondary antibody coupled to horseradish peroxide was added, and the membranes were incubated for 2 h at room temperature. Protein bands were detected using an enhanced chemiluminescence kit.

## Flow Cytometry Assay

The effect of CDKN2A on the apoptosis and cell cycle processes of HT-29 cells was analyzed using flow cytometry. For apoptosis, the transfected cells were collected and treated with trypsin, and the supernatant was removed after centrifugation at 3,000 × g at room temperature. Cells were suspended in the binding buffer according to the supplier's instructions and stained with membrane-linked protein V-FITC/Propidium iodide (PI) in the dark, and then the apoptosis rate was determined using flow cytometry. Similarly, transfected HT-29 cells were collected, fixed overnight in 75% ethanol at 4°C, stained with PI according to the manufacturer's instructions, and subsequently subjected to a cell cycle assay using flow cytometry.

# RESULTS

## Differential Analysis

Differential analysis was performed for colon cancer (T = 455, Nor = 41) and rectal cancer (T = 165, Nor = 10), and the results showed that there were 1,401 upregulated genes and 1,439 downregulated genes in colon cancer (**Figure 1A**), and 1,356 upregulated genes and 1,821 downregulated genes in rectal cancer (**Figure 1B**). The intersection results showed that there were 262 overlapping genes among the differential genes and EMT-related genes (**Figure 1C**). Survival analysis was performed on the 262 overlapping genes, of which 24 genes were associated with prognosis (**Figure 2A**). The genes with a risk ratio (HR) > 1 were subsequently subjected to bulk expression analysis, and nine genes were highly expressed in CRC tissues and six genes were highly expressed in paracancer tissues (**Figure 2B**).

## Construction of the Prognostic Risk Model

The nine EMT-related genes that were found to be highly expressed in CRC cells were subjected to LASSO regression analysis to construct characteristic prognostic models (**Figures 3A, B**). When the minimum lambda value was 0.0114, the following risk score formula was obtained: Riskscore = (0.1879)*PLOD3+(0.0427)*STC2+(0.1019)*CDKN2A+(0.0578)*SNAI1+(0.3033)*TIMP1. According to the calculation results of the formula, the samples are divided into high-risk group and low-risk group (**Figure 3C**). The KM survival curves demonstrated the difference in survival between the high-risk and low-risk groups, with the low-risk group found to have better survival compared to the high-risk group (**Figure 3D**).

## CDKN2A Is an Independent Prognostic Factor for Colorectal Cancer (CRC)

Genes in the risk score formula were defined as risk genes, and a nomogram was constructed to analyze the prognostic value of the five risk genes. Combining the results of the univariate and multifactorial Cox analyses showed CDKN2A to be an independent prognostic factor for CRC, and that age and M stage are also significant (**Figures 4A, B**). The nomogram model had good predictive power and demonstrated the prognostic predictive ability of each factor for CRC (**Figures 4C, D**).

## Comparison of CDKN2A Expression With Survival

To further understand the prognostic significance of CDKN2A in CRC, we analyzed the prognosis in terms of the three survival types. The results for overall survival (OS) (**Figure 5A**), progression-free survival (PFS) (**Figure 5B**), and disease-specific survival (DSS) (**Figure 5C**) showed that high CDKN2A expression indicated a worse prognosis. The median OS time was 28.8 months longer for patients with low CDKN2A expression compared to those with high CDKN2A expression. The ROC curves of OS (**Figure 5D**), PFS (**Figure 5E**) and DSS (**Figure 5F**) show the area under the curve (AUC) for this risk score is around 0.6.

## Knockdown of CDKN2A Expression Inhibits HT-29 Cell Proliferation

The cell proliferative ability was then determined *via* a CCK-8 assay. The results showed that knockdown of CDKN2A expression inhibited cell proliferation (**Figure 6A**).

## CDKN2A Promotes CRC Progression *Via* Epithelial-Mesenchymal Transition (EMT)

To verify the relationship between CDKN2A and EMT in CRC cells, we determined the protein levels of EMT signaling pathway-related molecules using WB. The results showed that knockdown of CDKN2A expression in HT-29 cells led to enhanced of E-cadherin expression and suppression expression of N-cadherin and vimentin at the protein levels compared with the negative control cells (**Figure 6B**). Thus, it is evident that knockdown of CDKN2A expression inhibits EMT in CRC cells.

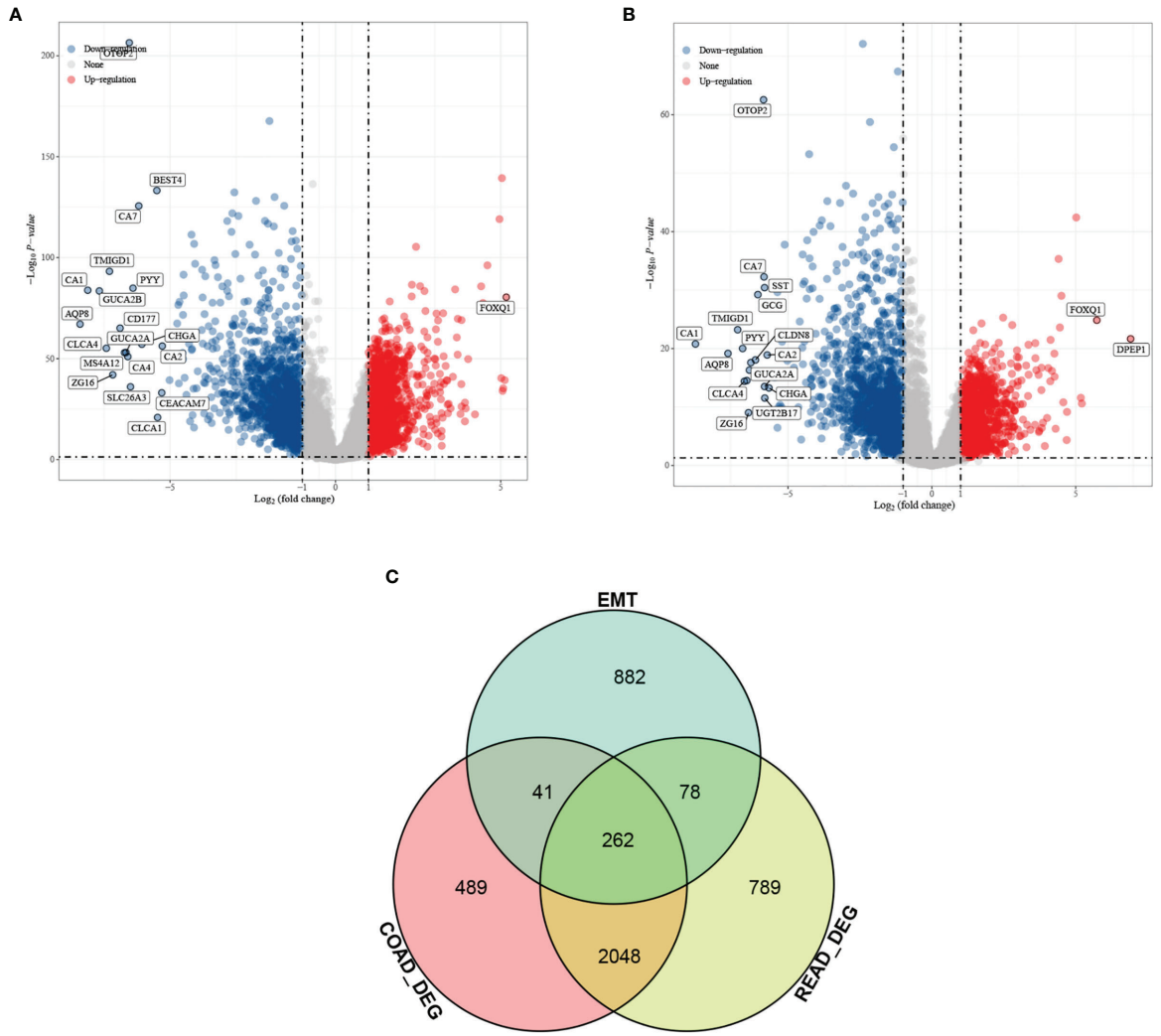



**FIGURE 1** | Differential analysis results from screening of EMT-related genes associated with prognosis of colorectal cancer (CRC). Volcano plot. Red indicates upregulated genes, and blue indicates downregulated genes. **(A)** Differential analysis of colon cancer and para cancer; **(B)** differential analysis of rectal cancer and para cancer; and **(C)** intersection results of differentially expressed genes (DEGs) in colon cancer, DEGs in rectal cancer, and EMT-related genes.

## CDKN2A Is Involved in the Regulation of Cell Cycle and Apoptosis of HT-29 Cells

Flow cytometry was utilized to understand the effect of CDKN2A on the cell cycle and apoptosis. Compared with the negative control cell samples, the percentage of G1 phase cells in the treated HT-29 cell samples increased significantly (**Figure 6C**). Also, the percentage of apoptotic HT-29 cells was significantly higher in the cell samples with knocked-down CDKN2A than in the control cell samples (**Figure 6D**).

## DISCUSSION

CRC is a highly lethal cancer, and its mortality is mainly attributed to recurrence and distant metastasis (19). There is a strong association between EMT and tumor metastasis, and EMT has been identified as a major cause of CRC metastasis (20). In this study, bioinformatics analysis showed that CDKN2A (p14) was an independent prognostic factor of colorectal cancer, and its high expression could induce EMT and mediate adverse clinical outcomes in CRC patients.

CDKN2A mainly encodes two proteins, p14 and p16 (21). p16 protein is a cyclin dependent kinase inhibitor, which can bind with CDK4 and CDK6 and prevent the phosphorylation of retinoblastoma protein, so as to prevent the process of cell cycle (22). p14 protein is a splice variant of CDKN2A. It has no amino acid homology with p16 due to frameshift, and can regulate the activity of p53 (23). As a tumor suppressor, CDKN2A is involved in B-cell differentiation, cell survival, and cell cycle progression (24). CDKN2A is also associated with cancer prognosis. Deletion of CDKN2A is associated with a poor prognosis in soft tissue sarcomas and is an independent prognostic factor in HPV-negative head and

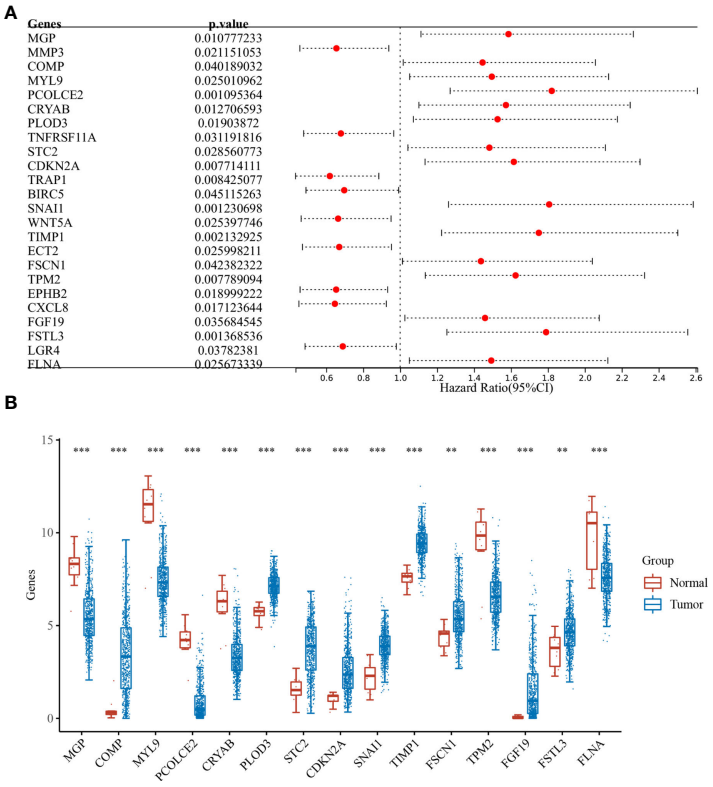



**FIGURE 2** | Batch survival and expression analysis. **(A)** Results of batch survival analysis of overlapping genes, HR > 1 indicates that the gene is a protective factor for cancer; and **(B)** expression analysis of genes with survival significance. **P < 0.01, ***P < 0.001.

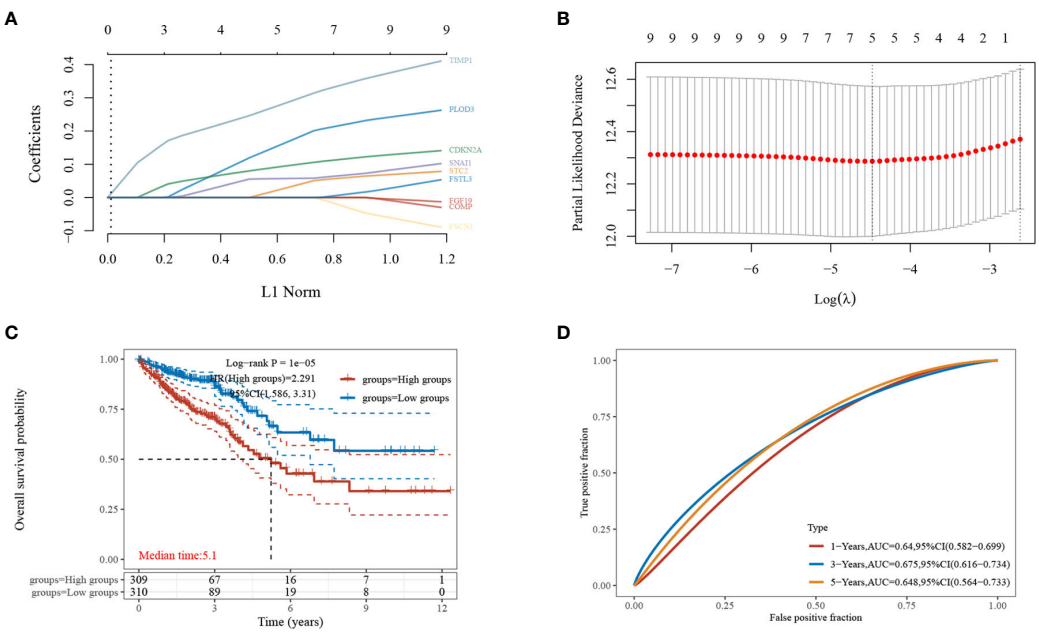



**FIGURE 3** | Establishment of the least absolute shrinkage and selection operator (LASSO)-Cox prognostic model. **(A, B)** Selected characteristic coefficients are shown by lambda parameters, and the number of prognostic factors was determined using LASSO regression analysis; **(C)** distribution of high and low risk samples; **(D)** KM survival analysis of the high- and low-risk samples for CRC in the TCGA dataset.

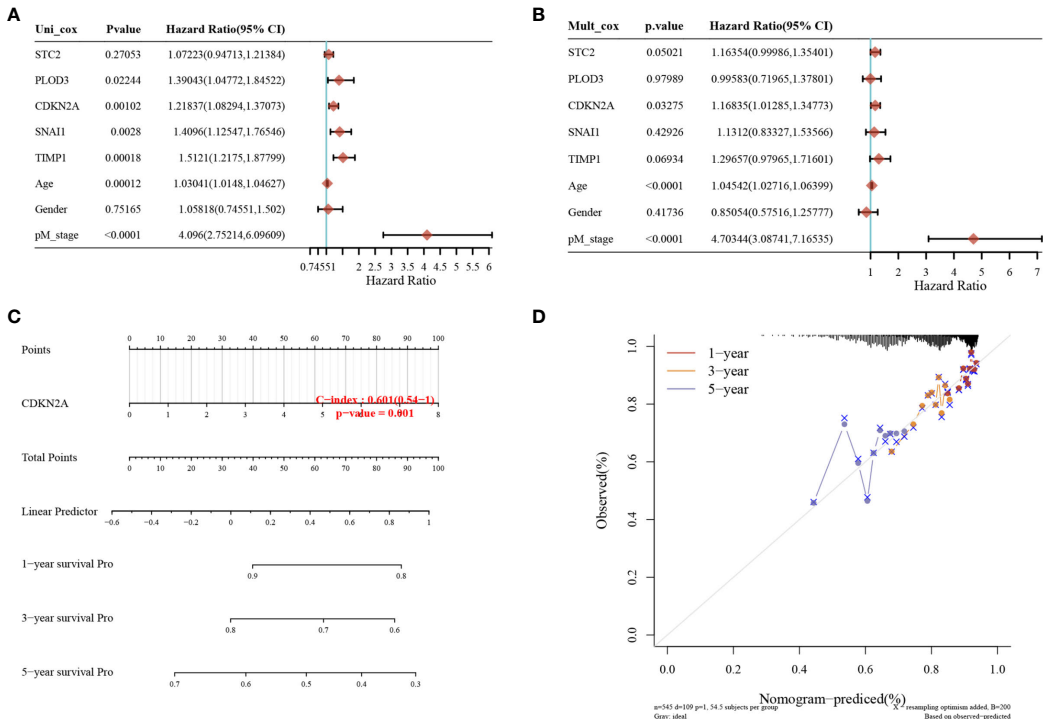



**FIGURE 4** | Identification of independent prognostic factors for CRC and single-gene prognostic analysis. **(A)** Single-factor Cox regression analysis; **(B)** multi-factor Cox regression analysis; and **(C, D)** the nomogram model used to predict the OS of patients with CRC at one, three, and five years. $p < 0.05$ was considered statistically significant.

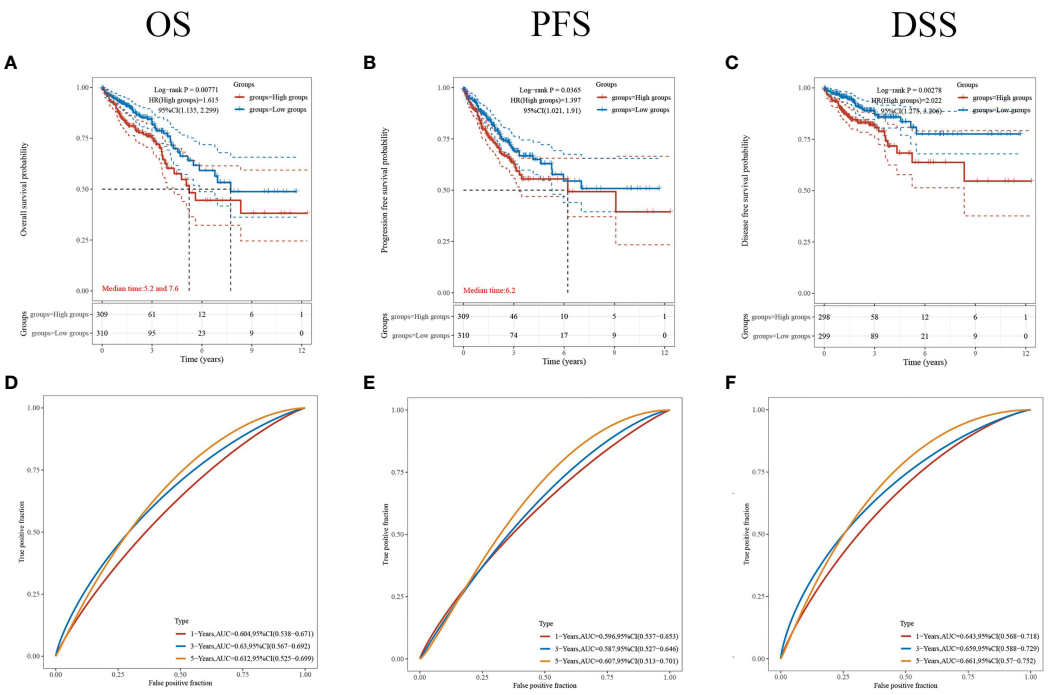



**FIGURE 5** | Comparison of different survival types. **(A–C)** The OS, PFS and DSS of patients with CRC who expressed high levels of CDKN2A. **(D–F)** ROC curves of OS, PFS and DSS. $p < 0.05$ was considered statistically significant.

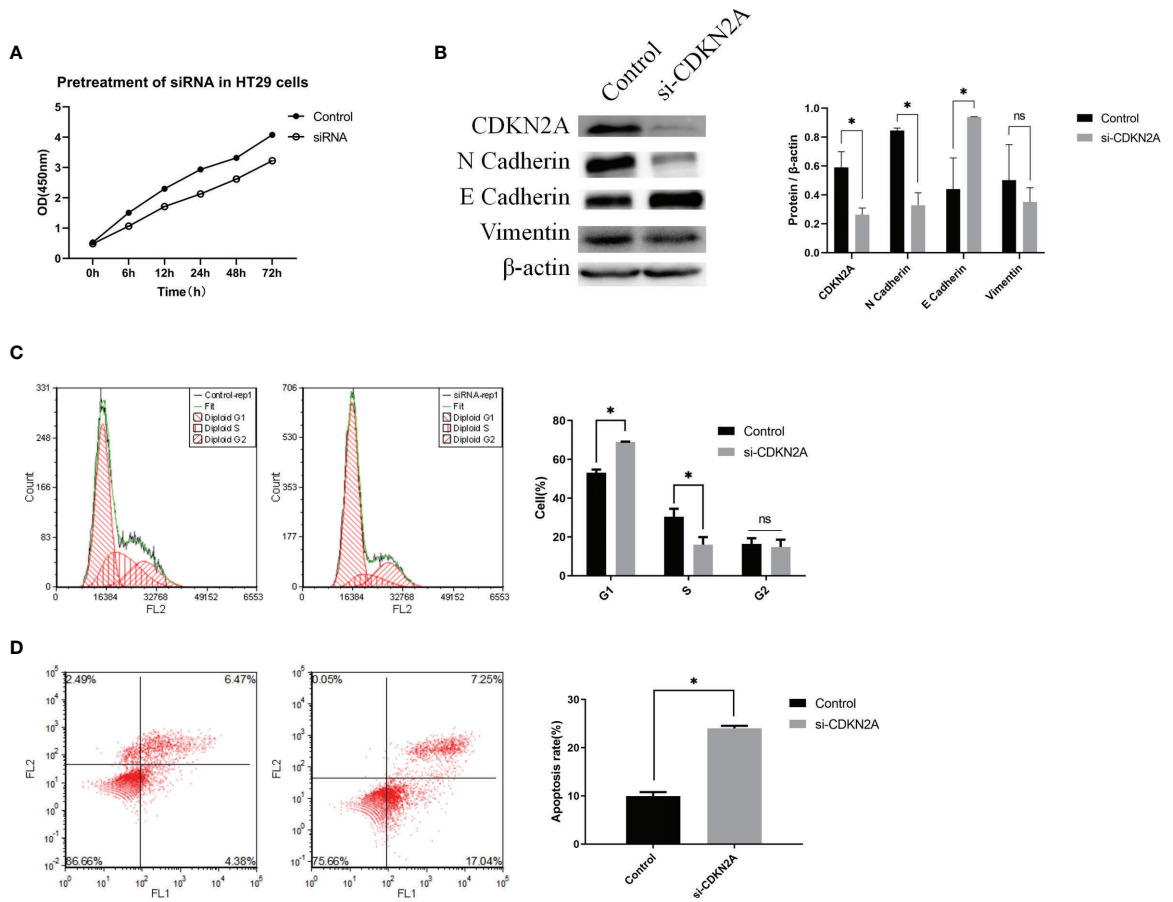



**FIGURE 6** | *In vitro* experiments validating the biological role of CDKN2A. **(A)** CCK-8 assay; **(B)** Protein expression of CDKN2A, E-cadherin, N-cadherin, and vimentin. **(C)** Flow cytometric analysis of the cell cycle and statistical results of the cell cycle; **(D)** Flow cytometric analysis of apoptosis and statistical results of apoptosis; The results are shown as the mean ± SD of the data from three sets of replicate experiments, *P < 0.05, ns, non significant.

neck squamous cell carcinomas (21, 25). Some reports suggest that hypermethylated CDKN2A is a predictor of poor prognosis of colorectal cancer (26). Interestingly, we also found that high CDKN2A expression is an independent prognostic factor for CRC and is associated with a poor prognosis. The reason for the contradiction between the two conclusions may be related to the heterogeneity between samples. In addition, there are regional differences in the risk of poor prognosis of colon cancer caused by CDKN2A methylation. In addition, the mechanism by which CDKN2A promotes CRC progression has been investigated. It may promote the proliferation and metastasis of tumor cells through ILF3-AS1/EZH2/H3K27me3/CDKN2A axis (27). Different from that study, in this study we analyzed the relationship between CDKN2A and colorectal cancer metastasis from EMT.

EMT plays an important role in tumor metastasis. The onset of the EMT process was often accompanied by decreased expression of E-cadherin and enhanced expression of N-cadherin and vimentin. E-cadherin is a transmembrane protein, and its downregulation is usually associated with the invasion of early tumor cells. N-cadherin is commonly found in non-epithelial cells, and its upregulation induces EMT and cancer stem cell properties (28). Vimentin is an intermediate filament protein that regulates cellular traits and migratory capacity during cell metastasis to support the EMT phenomenon (29). It is reported that EMT related prognostic features can not only be used as a prediction tool for high recurrence risk of cancer (30), but also promote invasion and metastasis (31). In the present study, WB results showed that knocking down CDKN2A expression in HT-29 cells was followed by enhanced E-cadherin expression and suppression of N-cadherin and vimentin expression. This implies that CDKN2A can induce EMT. Coincidentally, the occurrence of cancer invasion and metastasis is often related to EMT. In pancreatic cancer and melanoma, CDKN2A is often noticed due to mutations (22, 32). It is also reported that CDKN2A is lost in locally advanced or metastatic tumor tissues (33). Studies have shown that the mutation or defect of CDKN2A in cancer cells can induce EMT, so as to promote the invasion and metastasis of cancer cells (34, 35). The inactivation of p14 protein can worsen colorectal cancer and lead to tumor size (36).

In conclusion, this study combined bioinformatics analysis with *in vitro* assays to identify CDKN2A as an independent

prognostic factor in CRC and to confirm its biological role. The results show that CDKN2A (p14) has a protumor effect, and the increase of its expression can induce the occurrence of EMT.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.gsea-msigdb.org/gsea/msigdb/.

## AUTHOR CONTRIBUTIONS

WS and YL contributed equally to this work. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by Beijing Major Science and Technology Projects under grant No. D171100002617003.

## 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. Van der Jeught K, Xu H-C, Li Y-J, Lu X-B, Ji G. Drug Resistance and New Therapies in Colorectal Cancer. *World J Gastroenterol* (2018) 24(34):3834. doi: 10.3748/wjg.v24.i34.3834
3. Mattiuzzi C, Sanchis-Gomar F, Lippi G. Concise Update on Colorectal Cancer Epidemiology. *Ann Trans Med* (2019) 7(21):609. doi: 10.21037/atm.2019.07.91
4. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal Cancer Statistics, 2020. *CA Cancer J Clin* (2020) 70(3):145–64. doi: 10.3322/caac.21601
5. Hampel H, de la Chapelle A. The Search for Unaffected Individuals With Lynch Syndrome: Do the Ends Justify the Means? *Cancer Prev Res* (2011) 4 (1):1–5. doi: 10.1158/1940-6207.CAPR-10-0345
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. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2019. *CA Cancer J Clin* (2019) 69(1):7–34. doi: 10.3322/caac.21551
8. Pastushenko I, Blanpain C. EMT Transition States During Tumor Progression and Metastasis. *Trends Cell Biol* (2019) 29(3):212–26. doi: 10.1016/j.tcb.2018.12.001
9. Derynck R, Weinberg RA. EMT and Cancer: More Than Meets the Eye. *Dev Cell* (2019) 49(3):313–6. doi: 10.1016/j.devcel.2019.04.026
10. Georgakopoulos-Soares I, Chartoumpekis DV, Kyriazopoulou V, Zaravinos A. EMT Factors and Metabolic Pathways in Cancer. *Front Oncol* (2020) 10:499. doi: 10.3389/fonc.2020.00499
11. Aiello NM, Kang Y. Context-Dependent EMT Programs in Cancer Metastasis. *J Exp Med* (2019) 216(5):1016–26. doi: 10.1084/jem.20181827
12. Zeng Q, Zhang Y, Zhang W, Guo Q. Baicalein Suppresses the Proliferation and Invasiveness of Colorectal Cancer Cells by Inhibiting Snail–induced Epithelial–Mesenchymal Transition. *Mol Med Rep* (2020) 21(6):2544–52. doi: 10.3892/mmr.2020.11051
13. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in Cancer. *Nat Rev Cancer* (2018) 18(2):128–34. doi: 10.1038/nrc.2017.118
14. Lim SH, Becker TM, Chua W, Ng W, de Souza P, Spring KJ. Circulating Tumour Cells and the Epithelial Mesenchymal Transition in Colorectal Cancer. *J Clin Pathol* (2014) 67(10):848–53. doi: 10.1136/jclinpath-2014-202499
15. Shintani Y, Okimura A, Sato K, Nakagiri T, Kadota Y, Inoue M, et al. Epithelial to Mesenchymal Transition is a Determinant of Sensitivity to Chemoradiotherapy in Non-Small Cell Lung Cancer. *Ann Thorac Surg* (2011) 92(5):1794–804. doi: 10.1016/j.athoracsur.2011.07.032
16. Xie Y-H, Chen Y-X, Fang J-Y. Comprehensive Review of Targeted Therapy for Colorectal Cancer. *Signal Transduct Target Ther* (2020) 5(1):1–30. doi: 10.1038/s41392-020-0116-z
17. Boulding T, McCuaig R, Tan A, Hardy K, Wu F, Dunn J, et al. LSD1 Activation Promotes Inducible EMT Programs and Modulates the Tumour Microenvironment in Breast Cancer. *Sci Rep* (2018) 8(1):1–18. doi: 10.1038/s41598-017-17913-x
18. Li Q, Lai Q, He C, Fang Y, Yan Q, Zhang Y, et al. RUNX1 Promotes Tumour Metastasis by Activating the Wnt/β-Catenin Signalling Pathway and EMT in Colorectal Cancer. *J Exp Clin Cancer Res* (2019) 38(1):1–13. doi: 10.1186/s13046-019-1330-9
19. Li T, Hu P-S, Zuo Z, Lin J-F, Li X, Wu Q-N, et al. METTL3 Facilitates Tumor Progression *via* an M 6 A-IGF2BP2-Dependent Mechanism in Colorectal Carcinoma. *Mol Cancer* (2019) 18(1):1–15. doi: 10.1186/s12943-019-1038-7
20. Lin J, Shi Z, Yu Z, He Z. LncRNA HIF1A-AS2 Positively Affects the Progression and EMT Formation of Colorectal Cancer Through Regulating miR-129-5p and DNMT3A. *Biomed Pharmacother* (2018) 98:433–9. doi: 10.1016/j.biopha.2017.12.058
21. Bui NQ, Przybyl J, Trabucco SE, Frampton G, Hastie T, van de Rijn M, et al. A Clinico-Genomic Analysis of Soft Tissue Sarcoma Patients Reveals CDKN2A Deletion as a Biomarker for Poor Prognosis. *Clin Sarcoma Res* (2019) 9(1):1–11. doi: 10.1186/s13569-019-0122-5
22. Zeng H, Jorapur A, Shain AH, Lang UE, Torres R, Zhang Y, et al. Bi-Allelic Loss of CDKN2A Initiates Melanoma Invasion *via* BRN2 Activation. *Cancer Cell* (2018) 34(1):56–68. e9. doi: 10.1016/j.ccell.2018.05.014
23. Inoue K, Fry EA. Aberrant Expression of P14arf in Human Cancers: A New Biomarker? *Tumor Microenviron* (2018) 1(2):37. doi: 10.4103/tme.tme_24_17
24. Pfeifer H, Raum K, Markovic S, Nowak V, Fey S, Obländer J, et al. Genomic CDKN2A/2B Deletions in Adult Ph+ ALL are Adverse Despite Allogeneic Stem Cell Transplantation. *Blood* (2018) 131(13):1464–75. doi: 10.1182/blood-2017-07-796862
25. Chen WS, Bindra RS, Mo A, Hayman T, Husain Z, Contessa JN, et al. CDKN2A Copy Number Loss is an Independent Prognostic Factor in HPV-Negative Head and Neck Squamous Cell Carcinoma. *Front Oncol* (2018) 8:95. doi: 10.3389/fonc.2018.00095
26. Xing X, Cai W, Shi H, Wang Y, Li M, Jiao J, et al. The Prognostic Value of CDKN2A Hypermethylation in Colorectal Cancer: A Meta-Analysis. *Br J Cancer* (2013) 108(12):2542–8. doi: 10.1038/bjc.2013.251
27. Hong S, Li S, Bi M, Yu H, Yan Z, Liu T, et al. lncRNA ILF3-AS1 Promotes Proliferation and Metastasis of Colorectal Cancer Cells by Recruiting Histone Methylase EZH2. *Mol Ther Nucleic Acids* (2021) 24:1012–23. doi: 10.1016/j.omtn.2021.04.007
28. Sommariva M, Gagliano N. E-Cadherin in Pancreatic Ductal Adenocarcinoma: A Multifaceted Actor During Emt. *Cells* (2020) 9 (4):1040. doi: 10.3390/cells9041040
29. Loh C-Y, Chai JY, Tang TF, Wong WF, Sethi G, Shanmugam MK, et al. The E-Cadherin and N-Cadherin Switch in Epithelial-to-Mesenchymal Transition: Signaling, Therapeutic Implications, and Challenges. *Cells* (2019) 8(10):1118. doi: 10.3390/cells8101118
30. Wang K, Song K, Ma Z, Yao Y, Liu C, Yang J, et al. Identification of EMT-Related High-Risk Stage II Colorectal Cancer and Characterisation of Metastasis-Related Genes. *Br J Cancer* (2020) 123(3):410–7. doi: 10.1038/s41416-020-0902-y
31. Wang X, Lai Q, He J, Li Q, Ding J, Lan Z, et al. LncRNA SNHG6 Promotes Proliferation, Invasion and Migration in Colorectal Cancer Cells by Activating TGF-β/Smad Signaling Pathway *via* Targeting UPF1 and Inducing EMT *via* Regulation of ZEB1. *Int J Med Sci* (2019) 16(1):51. doi: 10.7150/ijms.27359
32. Tu Q, Hao J, Zhou X, Yan L, Dai H, Sun B, et al. CDKN2B Deletion is Essential for Pancreatic Cancer Development Instead of Unmeaningful Co-Deletion Due to Juxtaposition to CDKN2A. *Oncogene* (2018) 37(1):128–38. doi: 10.1038/onc.2017.316

33. Yang C, Cimera RS, Aryeequaye R, Jayakumaran G, Sarungbam J, Al-Ahmadie HA, et al. Adverse Histology, Homozygous Loss of CDKN2A/B, and Complex Genomic Alterations in Locally Advanced/Metastatic Renal Mucinous Tubular and Spindle Cell Carcinoma. *Mod Pathol* (2021) 34 (2):445–56. doi: 10.1038/s41379-020-00667-9
34. Merkle J, Breunig M, Schmid M, Allgöwer C, Krüger J, Melzer MK, et al. CDKN2A-Mutated Pancreatic Ductal Organoids From Induced Pluripotent Stem Cells to Model a Cancer Predisposition Syndrome. *Cancers* (2021) 13 (20):5139. doi: 10.3390/cancers13205139
35. Jeong E-H, Lee T-G, Ko YJ, Kim SY, Kim H-R, Kim H, et al. Anti-Tumor Effect of CDK Inhibitors on CDKN2A-Defective Squamous Cell Lung Cancer Cells. *Cell Oncol* (2018) 41(6):663–75. doi: 10.1007/s13402-018-0404-6
36. Hibi K, Nakayama H, Koike M, Kasai Y, Ito K, Akiyama S, et al. Colorectal Cancers With Both P16 and P14 Methylation Show Invasive Characteristics. *Jpn J Cancer Res* (2002) 93(8):883–7. doi: 10.1111/j.1349-7006.2002.tb01333.x

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 15 March 2022
doi: 10.3389/fonc.2022.857148

# Search for Key Genes and Functional Pathways of Ulcerative Colitis to Colon Cancer Based on Bioinformatics

*Shengbao Wang[1†], Lingling Zhen[2†], Xiaoli Li[3], Xu Fu[4], Peiwu Li[1*] and Dekui Zhang[5*]*

*[1] Emergency Center, Gansu Emergency Medicine Clinical Research Center, Lanzhou University Second Hospital, Lanzhou, China, [2] Infectious Department, Lanzhou University Second Hospital, Lanzhou, China, [3] Digestive Department, Lanzhou University Second Hospital, Lanzhou, China, [4] Key Laboratory of Emergency Medicine, Lanzhou University Second Hospital, Lanzhou, China, [5] Department of Gastroenterology, Key Laboratory of Digestive Diseases, Lanzhou University Second Hospital, Lanzhou, China*

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Zhengdong Zhang,*
*First Affiliated Hospital of Chengdu Medical College, China*
*Jun Ma,*
*First Affiliated Hospital of Xinjiang Medical University, China*

***Correspondence:**
*Dekui Zhang*
*zhangdk8616@126.com*
*Peiwu Li*
*lipeiw@lzu.edu.cn*

*†These authors share first authorship*

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

**Received:** *18 January 2022*
**Accepted:** *16 February 2022*
**Published:** *15 March 2022*

**Citation:**
*Wang S, Zhen L, Li X, Fu X, Li P and Zhang D (2022) Search for Key Genes and Functional Pathways of Ulcerative Colitis to Colon Cancer Based on Bioinformatics. Front. Oncol. 12:857148. doi: 10.3389/fonc.2022.857148*

Ulcerative colitis (UC) is a persistent and diffuse inflammatory disease of the intestine. It is widely prevalent in developed countries. Approximately 30% of patients with UC suffer from widespread and aggressive colitis and are at increased risk of colon cancer. In this study, the genetic features and potential molecular mechanisms shared between UC and colorectal cancer were investigated. The datasets from GEO and TCGA were analyzed to obtain differentially expressed genes, of which there were 116 overlapping genes. A module containing 15 genes was obtained using String and Cytoscape to analyze the module and identify hub genes. Weighted gene co-expression network analysis (WGCNA) was used to identify co-expression modules associated with UC and colon cancer, with 52 overlapping genes. Functional clustering of the two gene cohorts was performed using the Metascape online tool, with three significant functions or pathways associated with both gene cohorts. A total of 19 key genes were included, and CCT2 was identified after expression and survival analyses. CCT2 is highly expressed in colon cancer and lowly expressed in UC, and its low expression is associated with a poor prognostic ratio. This study reveals, for the first time, that CCT2 may be a promoter of UC transformation into colon cancer and identifies new gene candidates that could be used as biomarkers or potential therapeutic targets.

**Keywords: ulcerative colitis, colon cancer, key genes, diffuse inflammatory disease, cancer**

## INTRODUCTION

Ulcerative colitis (UC), a type of inflammatory bowel disease (IBD), is a persistent and diffuse inflammatory disease with typical symptoms, such as bloody diarrhea, abdominal pain, and stool urgency (1). UC is a chronic immune-mediated disease that is widely prevalent in developed countries and is rapidly emerging in newly industrialized countries (2). Some data have shown that its prevalence is about 0.286% in the United States, 5.05% in Norway, and 0.00667% in Malaysia (3). UC is common in all age groups but highly prevalent in adults over 6050 years of age (4). Family genetic susceptibility and previous smoking history are among the most important risk factors associated with UC. UC is clinically heterogeneous, and approximately 30% of patients suffer from a widespread and aggressive form of colitis

(5). Since there is no cure, patients usually seek pharmacological treatment. However, the inability to achieve durable remission leads approximately 15% of patients to opt for surgical removal of most or even all of the colon within 20 years of diagnosis (6).

Persistent inflammatory irritation of the colon puts patients with UC at a significantly increased risk of developing colon cancer, and the leading cause of death in patients with UC is colitis-associated cancer (CACC) (7). Although several studies have been conducted in this area, the causes and mechanisms responsible for this transformation are currently not well understood. This may be due to the glycosylation of the oncoprotein MUC1, which undergoes abnormal changes in UC, and this abnormally glycosylated MUC1 promotes the development and progression of cancer (8). For example, a study by Michael Kvorjak et al. proposed that immune cells may affect the abnormal glycosylation of MUC1, which induces cancer. In addition, Yao et al. started with inflammation and elaborated that validation and inflammatory cytokines may promote the progression of UC and colitis-related cancers (9). These studies attempt to find the factors that jointly promote the progression of UC and colon cancer, and hopefully, this will induce a therapeutic entry point.

Microarray and bioinformatics approaches have been widely used to screen for genetic changes at the genomic level (10). In this study, mRNA microarray datasets from the gene expression omnibus (GEO) and the cancer genome atlas (TCGA) were analyzed to obtain differentially expressed genes in UC and colon cancer tissues. Protein–protein interaction (PPI) network analysis was then performed to explore the relationship between the different genes. The weighted gene co-expression network was also used to obtain key genes associated with UC and colon cancer. The functional clustering of key genes was analyzed jointly with the Metascape online tool to find the functional pathways linking UC and colon cancer progression. Finally, the key genes associated with the progression of both diseases were obtained by analyzing the gene expression levels and survival analysis.

# METHODOLOGY

## Data Sources

Microarrays for UC were obtained from the GEO (https://www.ncbi.nlm.nih.gov/geo/) database. GSE37283 included four pure UC samples, five normal samples, and 11 UC samples with tumor formation. GSE38713 included 13 normal samples and 30 UC samples. GSE74604 included 30 normal samples and 30 colorectal cancer samples. Gene expression matrix and clinical information data of colon cancer and paracancer were obtained from the TCGA database, including 41 normal samples and 456 tumor samples. GSE37283, GSE38713, and GSE74604 were integrated into a new dataset containing normal, UC, and COAD tissues named Merge_1.

## Variance Analysis

The samples were divided into normal, UC, and tumor tissues. The datasets were analyzed separately for variance analysis using the R package limma package. $|\log2FC| > 1$ and $P < 0.05$ were defined as screening thresholds. Upset plots are a visualization tool for showing the common or unique differential genes between GSE37283, GSE38713, and TCGA.

## PPI Network Construction and Module Analysis

PPI networks were constructed using the String online database to analyze the functional interactions between proteins. The MCODE plug-in in Cytoscape was used to cluster the generated networks, defining tightly connected regions into a module. The selection criteria for the MCODE plug-in were Degree Cutoff = 2, Node Score Cutoff = 0.2, K-Core = 2, Max. Depth = 100. 0.2, K-Core = 2, and Max. Depth = 100.

## Weighted Gene Co-Expression Network Analysis (WGCNA)

WGCNA can be used to identify highly synergistic gene sets, and identify candidate biomarker genes or therapeutic targets according to the interconnection of gene sets and the association between gene sets and phenotypes. WGCNA was used to find biologically significant co-expressed gene modules and to explore the relationship between gene networks and diseases. In this study, WGCNA was used to obtain gene modules associated with UC and colon cancer. First, a suitable soft threshold β was selected according to the criteria of a scale-free network. Then, a hierarchical clustering dendrogram was constructed, and similarly expressed genes were classified into different modules. Finally, appropriate gene modules were selected by calculating the correlation between module feature genes (ME) and clinical features.

## Functional Clustering of Gene Modules

Metascape is a web-based portal that provides comprehensive gene list annotation and analysis resources (11). Metascape was used to simultaneously analyze key genes from PPI and WGCNA for pathway enrichment and functional annotation of genes to deeply understand the association between genes and diseases. Thus, the common or unique functions and pathways of the genes in both modules were annotated. The key gene sets were enriched and analyzed to understand their Kyoto gene Encyclopedia (KEGG) pathway and Gene Ontology (GO).

## Expression of Key Genes

The GSE37283, GSE38713, and GSE74604 datasets were merged using the R package named GM1. The merged datasets contained normal, UC, and colon cancer tissue samples, and differences in the expression of key genes were observed using a one-way ANOVA. Survival analysis of key genes was performed using the R package SURVIVAL, and a log-rank test was used to analyze the survival differences between the high and low expression of genes.

# RESULTS

## Identification and Analysis of Differentially Expressed Genes (DEGs)

DEGs were identified in the GSE37283, GSE38713, and TCGA datasets, respectively. In GSE37283, there were 134 downregulated genes and 887 upregulated genes (**Figure 1A**). In GSE38713, 372

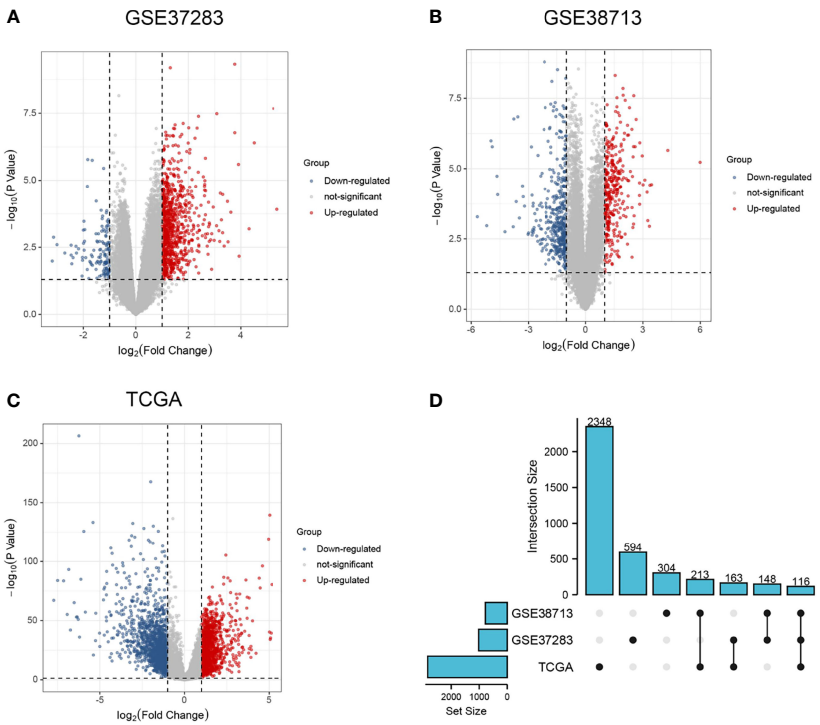



**FIGURE 1** | Differential expression analysis and Upset plots. **(A–C)** are the differential analysis from GSE37283, GSE38713, and TCGA datasets, respectively, and the results are shown as volcano plots. Blue indicates downregulation, and red indicates upregulation. **(D)** Upset plots showing the intersection or concatenation parts of the three differentially expressed genes.

downregulated genes and 330 upregulated genes were found (**Figure 1B**). There were 1439 downregulated genes and 1401 upregulated genes in the TCGA–COAD genes (**Figure 1C**). As shown in the Upset plot, the common and unique parts of the three groups of DEGs were shown, in which a total of 116 genes overlapped (**Figure 1D**).

## Analysis of Key Modules of DEGs

The screened DEGs were imported into the STRING database for PPI network construction and visualized using Cytoscape (**Figure 2A**). Networks with rich terms in DEGs were visualized, and the top 20 most significant terms included NABA MATRISOME ASSOCIATED, the IL-17 signaling pathway, and other functions or pathways (**Figure 2B**). Six gene modules were obtained using the MCODE plug-in in Cytoscape, and the most tightly connected module was selected. It included 15 nodes and 69 edges (**Figure 2C**).

## Selection of Co-Expression Modules for UC and COAD

Sixteen and 10 modules were identified by WGCNA in GSE37283 and GSE74604, respectively, with each color representing a different module. Heat maps of module–phenotype correlations were drawn based on Spearman correlation coefficients to assess the relationship between modules and features (**Figures 3A, C**). In GSE37283, darkseagreen4, floral white, and light cyan were highly correlated with UC, and darkseagreen4 (R = 0.79, P = 3.8e−5) was selected as the key module (**Figure 3B**). In GSE74604, blue salmon was highly associated with COAD, and blue (R = 0.94, P = 1.1e−28) was selected as the key module (**Figure 3D**). The intersection of the genes in the two key modules was taken using a Venn diagram, and 52 overlapping genes were obtained (**Figure 3E**).

## Functional Clustering

This study was conducted using Metascape to understand the functions and pathways of key genes obtained using different analysis methods. The key gene modules obtained using PPI and WGCNA analyses were named PPI-KG and WGCNA-KG, respectively. The overlap and shared term connections between them are shown as circle plots (**Figure 4A**). The two sets of gene lists were merged for PPI network analysis, and the functional terms between them were colored separately according to the counts (**Figure 4B**). A subset of enriched terms was selected and presented as a network graph to further capture the relationships between terms, where terms with a similarity > 0.3 were connected by edges (**Figure 4C**). The top 20 clusters and their representative enriched terms are presented in the table, with the regulation of neutrophil degranulation, ossification, and protein hydrolysis being the pathways significantly enriched in both gene lists, containing a total of 19 genes (**Table 1**).

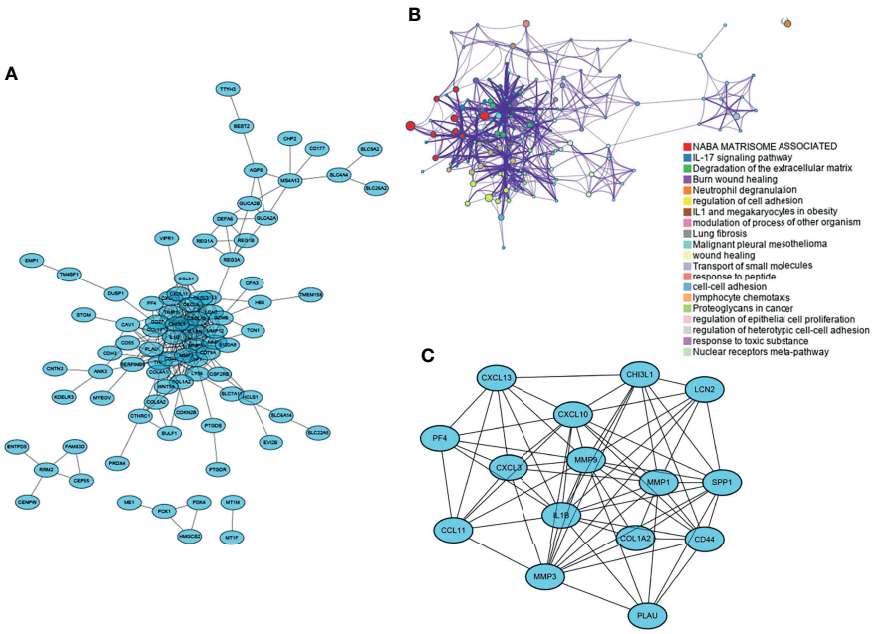



**FIGURE 2** | Functional interoperability network of DEGs and selection of key modules. **(A)** PPI network of DEGs; **(B)** rich terminology network in DEGs, each color represents one most prominent term; **(C)** most important gene modules in the PPI network.

## Enrichment Analysis

Based on the above results, we enriched and analyzed 19 genes in an attempt to understand their biological pathways or functions. The results showed that 19 genes were most associated with the KEGG pathway of IL − 17 signaling pathway (**Figure 5A**); In biological processes, it is related to neutrophil activation, degranulation, and its mediated immune response (**Figure 5B**).

## Expression and Survival Analysis

Nineteen genes related to the regulatory pathways of neutrophil degranulation, ossification, and protein hydrolysis were extracted and analyzed for their expression levels in different tissues in Merge_1. Most genes were significantly differentially expressed in normal and COAD tissues or in normal and UC tissues (**Figure 6**). Only HSP90AB1, RIPK2, DDX21, UCHL5, GTPBP4, and PCID2 were significantly different in UC and COAD tissues, and they all showed overexpression in colon cancer tissues compared to UC tissues (**Figures 6A–H**). A univariate survival analysis of 19 genes revealed that only CCT2 was associated with colon cancer prognosis, and its low expression was associated with poor prognosis (**Figures 7A, B**).

# DISCUSSION

In recent years, the cellular and molecular mechanisms by which chronic inflammation is associated with tumorigenesis have been progressively explored. In comparison, patients with a long course of UC have a significantly higher chance of developing colitis-associated cancer (12). This may be due to the chronic inflammation caused by persistent irritation or infection that can induce repeated epithelial cell stimulation and infiltration of immune cells and soluble mediators, thus providing a beneficial environment for tumor development (13, 14). It has also been shown that increased expression of certain proteins or activation of pathways may play an important role in transforming UC into colon cancer (9). In this study, we conducted bioinformatics analysis based on the latter to seek key genes and mechanisms that may promote the transition from UC to colon cancer. Three independent gene microarray datasets and one TCGA-COAD dataset were bioinformatically analyzed. gse37283, gse38713, and TCGA-COAD were used to perform differential analysis, and 116 overlapping DEGs were identified. A key module containing 15 genes was obtained using Cytoscape's MCODE plug-in. Key modules associated with UC and colon cancer were also identified using WGCNA analysis, with 52 genes. Functional clustering was performed using Metascape for both gene cohorts, which shared functions or pathways for regulating neutrophil degranulation, ossification, and protein hydrolysis.

Neutrophils are the first responders to inflammation and infection, and their elevated ratio to lymphocytes is a prognostic indicator of overall survival in some cancers (15). As an important component of cancer-associated inflammation, neutrophils are recruited in the tumor microenvironment to promote tumor progression (16). While neutrophil degranulation allows the antimicrobial agents contained in the granules to kill bacteria, excessive degranulation can also damage the host tissue (17). Ossification is defined as the formation of bone or bone material,

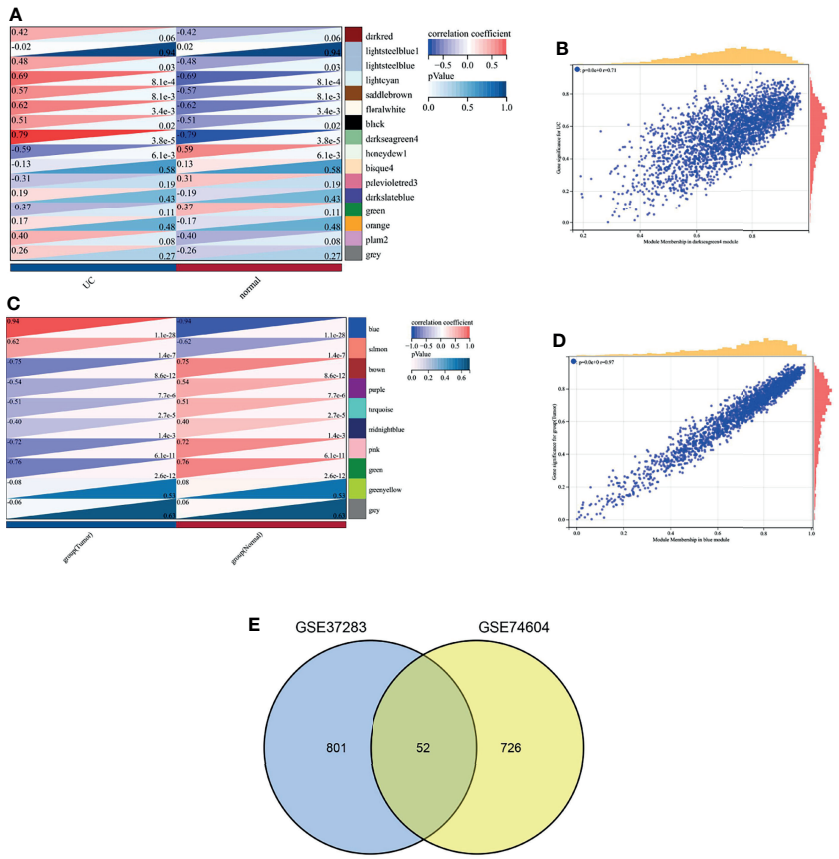



**FIGURE 3** | Weighted gene co-expression network analysis. **(A)** Heat map of correlation between gene modules and phenotypes in GSEA37283; **(B)** Scatter plot of correlation between key module genes and phenotypes in GSEA37283; **(C)** Heat map of correlation between gene modules and phenotypes in GSE74604; **(D)** Scatter plot of correlation between key module genes and phenotypes in GSEA74604; **(E)** Venn diagram of two sets of key module genes.

or the process of converting fibrous tissue or cartilage into bone or bone material (18). In contrast, regulating protein hydrolysis is an essential process for all cells, and it is particularly important when cells adapt to a new environment or coordinate development (19). Genes related to the regulation of neutrophil degranulation, ossification, and protein hydrolysis were extracted, and their expression in different tissues and their prognoses concerning colon cancer were observed. Significant differences were obtained in the expression of CCT2 in UC and colon cancer tissues. CCT2 is a subunit of CCT, which is a molecular chaperone protein that contributes to the correct folding of proteins, thus ensuring the stability of dynamic cellular homeostasis (20). CCT expression has been reported to be upregulated in rapidly proliferating tumor cells to promote tumor growth through efficient protein production (21). CCT expression has been shown to drive cell invasion and proliferation (22), and elevated mRNA levels of most of its subunits are associated with poor tumor prognosis (23). As one of the important subunits of CCT, CCT also plays an important role in tumors. Studies have shown that CCT2 expression negatively correlates with survival in breast cancer patients, which supports our results (24). CCT2 can also regulate protein folding functions associated with hypoxia in colorectal cancer (25). In this study, CCT2 expression showed an increasing trend in normal, UC, and colon cancer tissues. This expression difference also responds to the possibility that CCT2 may be closely associated with the development of colonic inflammation or cancer.

However, it has not been reported whether CCT directly relates to UC transformation into colorectal cancer. The available studies have suggested that CCT may contribute to malignant transformation in an inflammatory setting. The NF-κB pathway has long been recognized as a classic proinflammatory signaling pathway (26). Increased NF-κB activity indirectly promotes the formation of neutrophil extracellular traps (NET), which is one of the antimicrobial mechanisms of neutrophils (27). As with degranulation, the occurrence of this mechanism may allow neutrophils to rapidly enter into suicidal cell death, inducing chronic inflammation and cancer (28). However, CCT may promote NF-κB activity. There is evidence that the binding of LOX-1, a novel CCT substrate protein with an important link between chronic inflammation and cancer (29), to CCT may

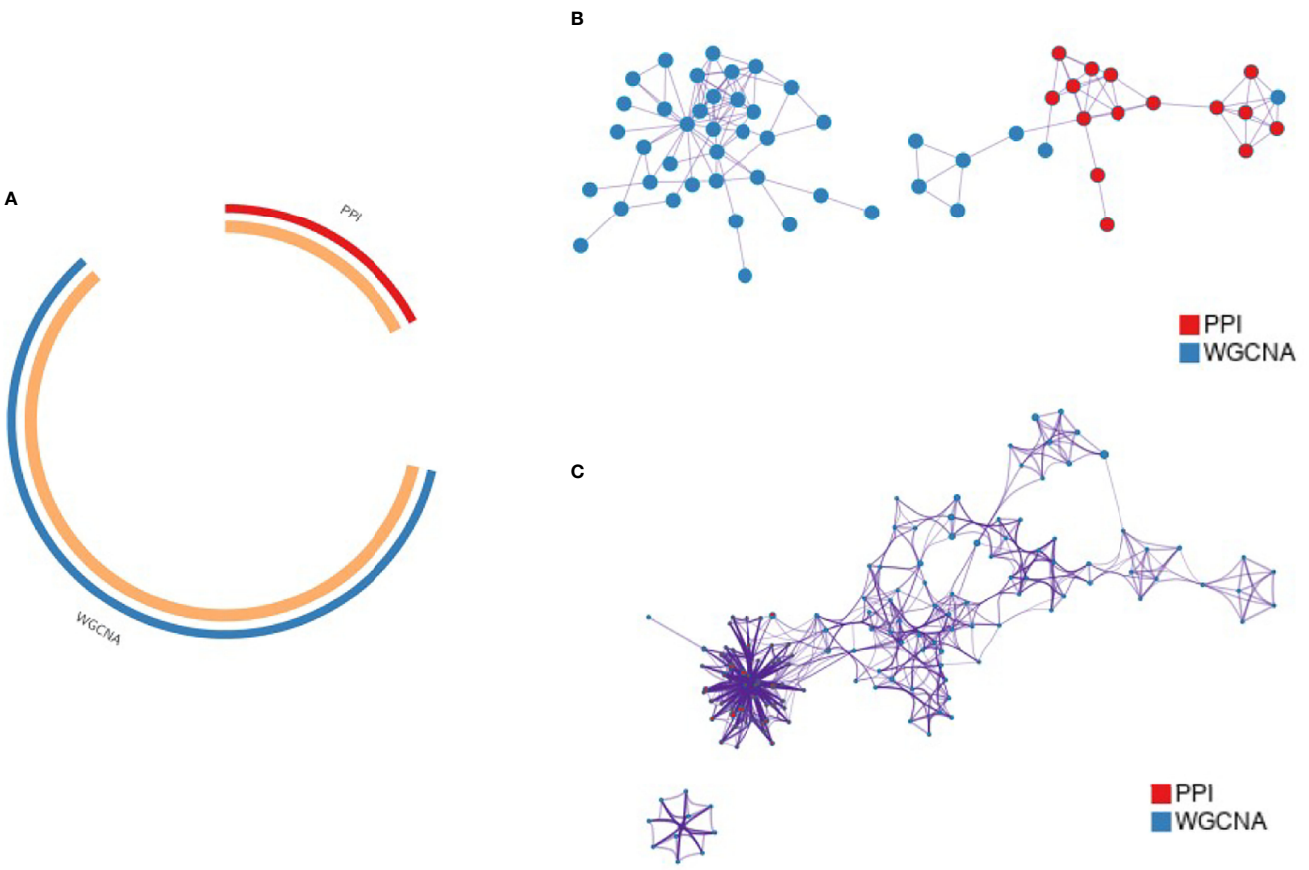



**FIGURE 4** | Functional clustering of key genes by Metascape. **(A)** Circos plot showing the overlap of key gene modules from PPI and WGCNA; **(B)** PPI network of genes in the merged cohort; **(C)** Pie chart showing the enriched terminology network of genes, with different colors indicating different cohorts.

**TABLE 1** | Top 20 clusters and their representative enrichment terms.

| _PATTERN_ | Category | Description | Count | Log10(P) |
|---|---|---|---|---|
| PPI | WikiPathways | Overview of proinflammatory and profibrotic mediators | 10 | –20.38 |
| PPI | KEGG Pathway | IL-17 signaling pathway | 8 | –16.39 |
| WGCNA | GO Biological Processes | ribonucleoprotein complex biogenesis | 12 | –10.97 |
| PPI | WikiPathways | Lung fibrosis | 5 | –10 |
| PPI | Canonical Pathways | PID AVB3 OPN PATHWAY | 4 | –8.91 |
| WGCNA | WikiPathways | Nucleotide metabolism | 4 | –7.53 |
| PPI | Reactome Gene Sets | Cell surface interactions at the vascular wall | 4 | –6.27 |
| PPI WGCNA | Reactome Gene Sets | Neutrophil degranulation | 9 | –5.95 |
| PPI WGCNA | GO Biological Processes | ossification | 7 | –5.62 |
| WGCNA | GO Biological Processes | nucleobase-containing compound biosynthetic process | 9 | –5.47 |
| WGCNA | GO Biological Processes | protein import into mitochondrial matrix | 3 | –5.27 |
| PPI | KEGG Pathway | Amoebiasis | 3 | –4.78 |
| WGCNA | GO Biological Processes | positive regulation of phosphatase activity | 3 | –4.48 |
| WGCNA | Reactome Gene Sets | ESR-mediated signaling | 5 | –3.86 |
| WGCNA | GO Biological Processes | regulation of translation | 6 | –3.85 |
| WGCNA | Reactome Gene Sets | Translation | 5 | –3.84 |
| WGCNA | GO Biological Processes | translational initiation | 3 | –3.75 |
| PPI WGCNA | GO Biological Processes | regulation of proteolysis | 8 | –3.63 |
| PPI | GO Biological Processes | regulation of peptidyl-serine phosphorylation | 4 | –3.52 |
| PPI | GO Biological Processes | regulation of chromosome organization | 4 | –3.3 |

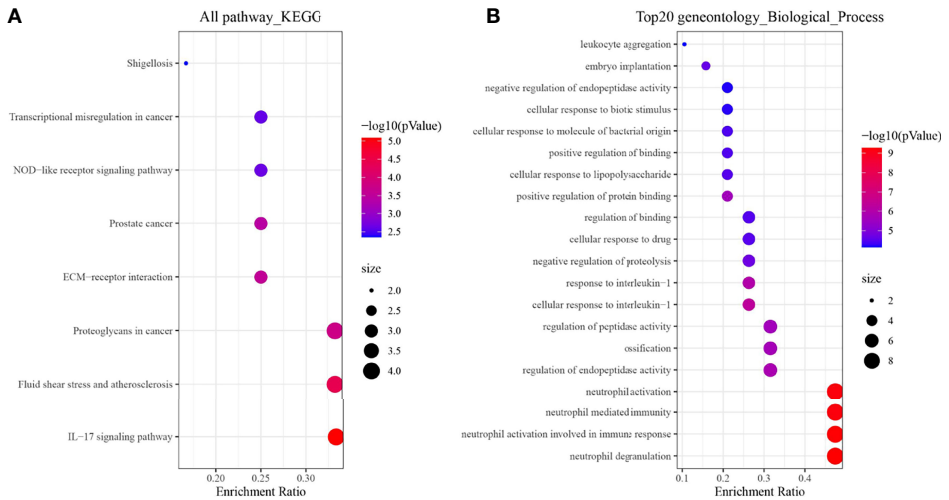



**FIGURE 5** | Enrichment analysis results of 19 genes. **(A)** KEGG; **(B)** Biological process terms of the top 20 GO.

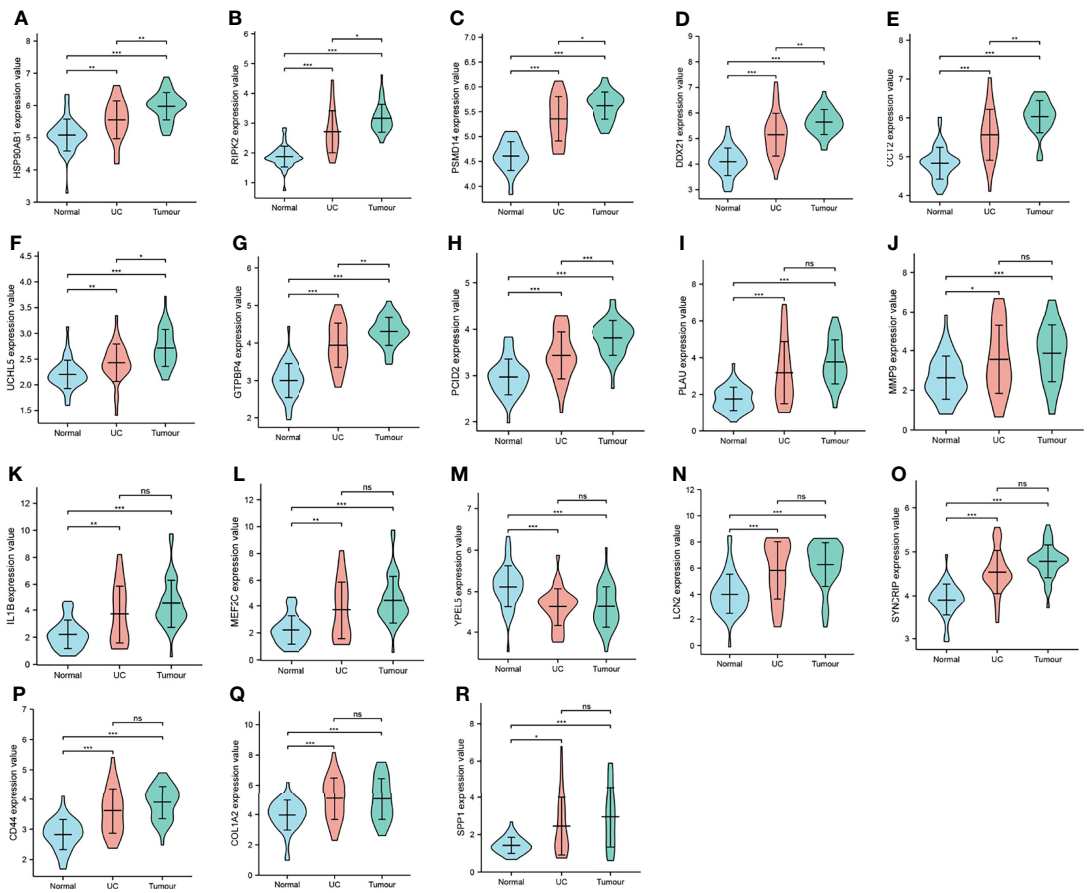



**FIGURE 6** | Expression distribution of key genes. **(A–R)** Expression level distribution of 18 key genes in normal, UC, and COAD tissues. *P < 0.05, **P < 0.01, ***P < 0.001, ns, no significant. CHIL1 was not significantly different in all the three tissues, so it is not shown.

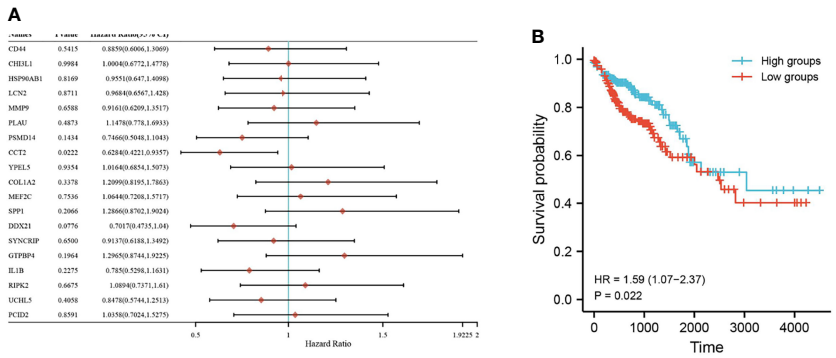



**FIGURE 7** | Survival of key genes in colorectal cancer samples. **(A)** Forest plot with one-way Cox analysis for 19 genes; **(B)** Survival curve plot for CCT2. P < 0.05 is considered statistically significant.

activate the inflammatory transcription factor NF-κB, thereby inducing malignant transformation (30). Altering the expression level or activity of CCT may also lead to increased NF-κB activity and carcinogenesis (31). This shows that CCT has a proinflammatory role and may also indirectly promote the transformation of inflammation into cancer.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## AUTHOR CONTRIBUTIONS

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

## FUNDING

This work was supported by The Natural Science Foundation of Gansu Province(20JR10RA724), The Second Hospital of Lanzhou University Cuiying Science and Technology Innovation Program Project (CY2020-MS06) and Science and Technology Plan Project of Lanzhou City (2020-ZD-91).

## REFERENCES

1. Tripathi K, Feuerstein JD. New Developments in Ulcerative Colitis: Latest Evidence on Management, Treatment, and Maintenance. *Drugs Context* (2019) 8:212572. doi: 10.7573/dic.212572
2. Jones G-R, Lyons M, Plevris N, Jenkinson PW, Bisset C, Burgess C, et al. IBD Prevalence in Lothian, Scotland, Derived by Capture–Recapture Methodology. *Gut* (2019) 68(11):1953–60. doi: 10.1136/gutjnl-2019-318936
3. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide Incidence and Prevalence of Inflammatory Bowel Disease in the 21st Century: A Systematic Review of Population-Based Studies. *Lancet* (2017) 390(10114):2769–78. doi: 10.1016/S0140-6736(17)32448-0
4. Rönnblom A, Karlbom U. Acute Severe Attacks of Ulcerative Colitis in a Population-Based Cohort: Epidemiology, Treatment and Outcome. *Scand J Gastroenterol* (2020) 55(5):555–9. doi: 10.1080/00365521.2020.1757143
5. Solberg IC, Lygren I, Jahnsen J, Aadland E, Høie O, Cvancarova M, et al. Clinical Course During the First 10 Years of Ulcerative Colitis: Results From a Population-Based Inception Cohort (IBSEN Study). *Scand J Gastroenterol* (2009) 44(4):431–40. doi: 10.1080/00365520802600961
6. Eisenstein M. Ulcerative Colitis: Towards Remission. *Nature* (2018) 563 (7730):S33–S. doi: 10.1038/d41586-018-07276-2
7. Zhang Q, Sha S, Xu B, Liang S, Wu K. Prevalence of Colorectal Cancer in Patients With Ulcerative Colitis: A Retrospective, Monocenter Study in China. *J Cancer Res Ther* (2015) 11(4):899. doi: 10.4103/0973-1482.143345
8. Kvorjak M, Ahmed Y, Miller ML, Sriram R, Coronnello C, Hashash JG, et al. Crosstalk Between Colon Cells and Macrophages Increases ST6GALNAC1 and MUC1-sTn Expression in Ulcerative Colitis and Colitis-Associated Colon Cancer. *Cancer Immunol Res* (2019) 8(2):167–78. doi: 10.1158/2326-6066.CIR-19-0514
9. Yao D, Dong M, Dai C, Wu S. Inflammation and Inflammatory Cytokine Contribute to the Initiation and Development of Ulcerative Colitis and its Associated Cancer. *Inflamm Bowel Dis* (2019) 25(10):1595–602. doi: 10.1093/ibd/izz149
10. Jiang CH, Yuan X, Li JF, Xie YF, Zhang AZ, Wang XL, et al. Bioinformatics-Based Screening of Key Genes for Transformation of Liver Cirrhosis to Hepatocellular Carcinoma. *J Trans Med* (2020) 18(1):1–11. doi: 10.1186/s12967-020-02229-8
11. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10(1):1–10. doi: 10.1038/s41467-019-09234-6
12. Low END, Mokhtar NM, Wong Z, Raja Ali RA. Colonic Mucosal Transcriptomic Changes in Patients With Long-Duration Ulcerative Colitis Revealed Colitis-Associated Cancer Pathways. *J Crohn's Colitis* (2019) 13 (6):755–63. doi: 10.1093/ecco-jcc/jjz002
13. Zhang Y, Kong W, Jiang J. Prevention and Treatment of Cancer Targeting Chronic Inflammation: Research Progress, Potential Agents, Clinical Studies and Mechanisms. *Sci China Life Sci* (2017) 60(6):601–16. doi: 10.1007/s11427-017-9047-4
14. Galdiero MR, Marone G, Mantovani A. Cancer Inflammation and Cytokines. *Cold Spring Harbor Perspect Biol* (2018) 10(8):a028662. doi: 10.1101/cshperspect.a028662
15. Mollinedo F. Neutrophil Degranulation, Plasticity, and Cancer Metastasis. *Trends Immunol* (2019) 40(3):228–42. doi: 10.1016/j.it.2019.01.006

16. Galdiero MR, Varricchi G, Loffredo S, Mantovani A, Marone G. Roles of Neutrophils in Cancer Growth and Progression. *J Leukocyte Biol* (2018) 103 (3):457–64. doi: 10.1002/JLB.3MR0717-292R
17. Eichelberger KR, Goldman WE. Manipulating Neutrophil Degranulation as a Bacterial Virulence Strategy. *PloS Pathog* (2020) 16(12):e1009054. doi: 10.1371/journal.ppat.1009054
18. Yang Y-Q, Tan Y-Y, Wong R, Wenden A, Zhang L-K, Rabie ABM. The Role of Vascular Endothelial Growth Factor in Ossification. *Int J Oral Sci* (2012) 4 (2):64–8. doi: 10.1038/ijos.2012.33
19. Raivio TL. Regulation of Proteolysis in the Gram-Negative Bacterial Envelope. *J Bacteriol* (2018) 200(3):e00639–17. doi: 10.1128/JB.00639-17
20. Dekker C, Stirling PC, McCormack EA, Filmore H, Paul A, Brost RL, et al. The Interaction Network of the Chaperonin CCT. *EMBO J* (2008) 27(13):1827–39. doi: 10.1038/emboj.2008.108
21. Yokota S-i, Yamamoto Y, Shimizu K, Momoi H, Kamikawa T, Yamaoka Y, et al. Increased Expression of Cytosolic Chaperonin CCT in Human Hepatocellular and Colonic Carcinoma. *Cell Stress Chaperones* (2001) 6 (4):345. doi: 10.1379/1466-1268(2001)006<0345:IEOCCC>2.0.CO;2
22. Showalter AE, Martini AC, Nierenberg D, Hosang K, Fahmi NA, Gopalan P, et al. Investigating Chaperonin-Containing TCP-1 Subunit 2 as an Essential Component of the Chaperonin Complex for Tumorigenesis. *Sci Rep* (2020) 10 (1):1–14. doi: 10.1038/s41598-020-57602-w
23. Macario AJ, de Macario EC eds. Chaperonins in Cancer: Expression, Function, and Migration in Extracellular Vesicles. In: *Seminars in Cancer Biology*. Elsevier. doi: 10.1016/j.semcancer.2021.05.029
24. Ghozlan H, Showalter A, Lee E, Zhu X, Khaled AR. Chaperonin-Containing TCP1 Complex (CCT) Promotes Breast Cancer Growth Through Correlations With Key Cell Cycle Regulators. *Front Oncol* (2021) 11:663877. doi: 10.3389/fonc.2021.663877
25. Park SH, Jeong S, Kim BR, Jeong YA, Kim JL, Na YJ, et al. Activating CCT2 Triggers Gli-1 Activation During Hypoxic Condition in Colorectal Cancer. *Oncogene* (2020) 39(1):136–50. doi: 10.1038/s41388-019-0972-6
26. Lawrence T. The Nuclear Factor NF-κb Pathway in Inflammation. *Cold Spring Harbor Perspect Biol* (2009) 1(6):a001651. doi: 10.1101/cshperspect.a001651
27. Fonseca Z, Díaz-Godínez C, Mora N, Alemán OR, Uribe-Querol E, Carrero JC, et al. Entamoeba Histolytica Induce Signaling *via* Raf/MEK/ERK for Neutrophil Extracellular Trap (NET) Formation. *Front Cell Infect Microbiol* (2018) 8:226. doi: 10.3389/fcimb.2018.00226
28. Brostjan C, Oehler R. The Role of Neutrophil Death in Chronic Inflammation and Cancer. *Cell Death Discov* (2020) 6(1):1–8. doi: 10.1038/s41420-020-0255-6
29. Khaidakov M, Mitra S, Kang B-Y, Wang X, Kadlubar S, Novelli G, et al. Oxidized LDL Receptor 1 (OLR1) as a Possible Link Between Obesity, Dyslipidemia and Cancer. *PloS One* (2011) 6(5):e20277. doi: 10.1371/journal.pone.0020277
30. Hirsch HA, Iliopoulos D, Joshi A, Zhang Y, Jaeger SA, Bulyk M, et al. A Transcriptional Signature and Common Gene Networks Link Cancer With Lipid Metabolism and Diverse Human Diseases. *Cancer Cell* (2010) 17 (4):348–61. doi: 10.1016/j.ccr.2010.01.022
31. Roh S-H, Kasembeli M, Bakthavatsalam D, Chiu W, Tweardy DJ. Contribution of the Type II Chaperonin, TRiC/CCT, to Oncogenesis. *Int J Mol Sci* (2015) 16(11):26706–20. doi: 10.3390/ijms161125975

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 21 March 2022
doi: 10.3389/fonc.2022.837570

# Transcriptomic Portraits and Molecular Pathway Activation Features of Adult Spinal Intramedullary Astrocytomas

*Nikolay Konovalov*[1], *Stanislav Timonin*[1], *Dmitry Asyutin*[1], *Mikhail Raevskiy*[2,3,4,5], *Maxim Sorokin*[3,4,5], *Anton Buzdin*[2,3,4,5,6*] *and Stanislav Kaprovoy*[1]

[1] Burdenko Neurosurgical Center, Moscow, Russia, [2] Omicsway Corp., Walnut, CA, United States, [3] Moscow Institute of Physics and Technology, Moscow, Russia, [4] Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia, [5] I.M. Sechenov First Moscow State Medical University, Moscow, Russia, [6] Oncobox Ltd., Moscow, Russia

OPEN ACCESS

**Edited by:**
Tzu Pin Lu,
National Taiwan University, Taiwan

**Reviewed by:**
Giorgio Carrabba,
University of Milano-Bicocca, Italy
Mirza Pojskic,
University Hospital of Giessen and Marburg, Germany
Nir Shimony,
Geisinger Health System, United States

***Correspondence:**
Anton Buzdin
buzdin@oncobox.com

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

**Received:** 16 December 2021
**Accepted:** 21 February 2022
**Published:** 21 March 2022

**Citation:**
Konovalov N, Timonin S, Asyutin D, Raevskiy M, Sorokin M, Buzdin A and Kaprovoy S (2022) Transcriptomic Portraits and Molecular Pathway Activation Features of Adult Spinal Intramedullary Astrocytomas. Front. Oncol. 12:837570. doi: 10.3389/fonc.2022.837570

In this study, we report 31 spinal intramedullary astrocytoma (SIA) RNA sequencing (RNA-seq) profiles for 25 adult patients with documented clinical annotations. To our knowledge, this is the first clinically annotated RNA-seq dataset of spinal astrocytomas derived from the intradural intramedullary compartment. We compared these tumor profiles with the previous healthy central nervous system (CNS) RNA-seq data for spinal cord and brain and identified SIA-specific gene sets and molecular pathways. Our findings suggest a trend for SIA-upregulated pathways governing interactions with the immune cells and downregulated pathways for the neuronal functioning in the context of normal CNS activity. In two patient tumor biosamples, we identified diagnostic *KIAA1549-BRAF* fusion oncogenes, and we also found 16 new SIA-associated fusion transcripts. In addition, we bioinformatically simulated activities of targeted cancer drugs in SIA samples and predicted that several tyrosine kinase inhibitory drugs and thalidomide analogs could be potentially effective as second-line treatment agents to aid in the prevention of SIA recurrence and progression.

**Keywords: spinal intramedullary astrocytoma, glioblastoma, transcriptomic (RNA-Seq), RNA sequencing (RNA-seq), molecular pathway activation, gene expression, spinal cord neoplasms**

## INTRODUCTION

Spinal intramedullary astrocytoma (SIA) is a rare subtype of glioma comprising about 2%–4% of all primary central nervous system (CNS) neoplasms and approximately 6%–8% of tumors occurring in the spinal cord. SIAs are mainly observed as low-grade tumors (WHO I and II) (1). Five-year overall survival rate of patients with low-grade SIA is 70%–80% and declines to 14%–28% for grades III–IV (2). However, clinical data on prognostic biomarkers and tumor molecular data associated with treatment outcomes are needed for patients with spinal astrocytoma due to a particularly low frequency of these tumors and lack of successful therapeutic regimens. In addition, diagnosis and treatment of these neoplasms is often challenging given their ambiguous manifestations such as

back pain, limb weakness, paresthesia, and bowel and bladder dysfunction (1, 3). Spinal cord tumors are more frequently diagnosed in children (3, 4) but also occur in adults (4).

Surgical resection remains the main primary treatment for intramedullary astrocytomas of the spinal cord (3, 5). In turn, second-line treatments usually include radiation therapy and chemotherapy. It was also reported that the use of adjuvant radiation therapy can result in an increase in the overall survival of patients, especially for the lower-grade tumors (6). However, the optimal regimen for an adjunctive therapy including chemotherapy settings has not yet been precisely determined (1, 3, 5). Treating spinal cord astrocytomas remains problematic to date, and morbidity and mortality depend on various factors. In order to better understand these relevant factors linked with the outcomes of spinal cord astrocytomas, several studies were conducted. Due to the complications associated with clinical diagnosis, indecision about optimal surgical treatment, and second-line treatment failures, most reports on intramedullary astrocytomas represent either small cohort retrospective analyses and case studies or data capturing current changes in treatment options (1, 3, 5). Due to the low incidence of these tumors, prospective clinical investigations are problematic to perform, and alternative chemo- and targeted therapeutic agents and regimens were poorly explored for SIA (7–9).

On the other hand, RNA expression profiles may serve as potent predictors of tumor sensitivity to targeted therapeutics, as shown in clinical investigations for microarray (10) and RNA sequencing (RNA-seq) (11) data. Furthermore, molecular pathway activation levels can be calculated using high-throughput gene expression profiles (12, 13) and translated into next-generation biomarkers (14–16) for algorithmic scoring of cancer drug efficiencies (17, 18). Moreover, aggregation of expression data of single gene products into pathways or signatures results in significantly more robust expression-level biomarkers, as deduced theoretically (19) and shown on real cancer molecular data (20, 21). Thus, RNA-seq profiles can be used for finding effective cancer prognostic or predictive biomarkers and assist in finding better clinical treatment regimens (22–24). However, there is a dearth of clinically annotated molecular profiles of SIA that could be used for such a purpose.

In this study, we report 31 new SIA RNA-seq profiles for 25 patients with documented clinical annotations. As far as we know, the current study presents the first clinically annotated RNA-seq dataset of spinal astrocytomas derived from the intradural intramedullary compartment.

We compared these tumor profiles with the previous healthy CNS RNA-seq data of spinal cord and brain samples (24, 25). We identified differentially expressed gene (DEG) sets in SIA and molecular pathways and analyzed the occurrence of known diagnostic and new fusion transcripts. In addition, we calculated prognostic balanced efficiency scores for known targeted drugs and identified a fraction of them that could be potentially helpful as second-line treatment agents to aid in the prevention of SIA recurrence and progression.

# RESULTS

SIAs were not previously characterized on transcriptome-wide level, and in this study, we aimed to analyze RNA-seq profiles of SIA samples in comparison with healthy brain and spinal cord samples obtained from Genotype-Tissue Expression (GTEx) Portal (25) and Atlas of RNA sequencing profiles for normal human tissues (ANTE) database (24).

## Spinal Intramedullary Astrocytoma Diagnosis and Biosamples

Overall, 31 tumor tissue samples were taken from 14 male and 11 female donors who were diagnosed between 2003 and 2018 with SIA (23 pilocytic astrocytomas, 4 glioblastomas, 2 anaplastic astrocytomas, and 2 astrocytomas with uncategorized histological subtype; **Figure 1**). The mean age was 32.73 years (range 18–69 years) and 30.00 years (18–56 years), respectively. The biosamples were formalin-fixed paraffin-embedded (FFPE) histologically characterized tumor samples with at least 70% cancer cells. Clinical annotations of tumor tissue specimens investigated in this study and their patient origin are summarized in **Supplementary Table S1**. Kaplan–Meier plots for progression-free survival and overall survival are shown in **Supplementary Figure S1**.

## Characteristics of Sequencing and Mapping

A total of 10,509.3 million reads were obtained for 31 independent libraries of SIA tissues (median 55.3 million reads per sample). Most reads reached Phred-like quality scores (Q-scores) at the Q30 level, indicating that the probability of an incorrect base call is 0.001%. The average coverage of sequencing depth reached approximately 53.45× of the human transcriptome. After alignment, 98.57% to 99.12% uniquely aligned reads were mapped to the reference human genome.

## Primary Comparison of RNA Sequencing Profiles of Spinal Intramedullary Astrocytomas and Healthy Central Nervous System Tissues

To further characterize SIA transcriptomic data, we compared using principal component analysis (PCA) distributions of RNA-seq profiles among the SIA samples (n = 31) and publicly available datasets of normal spinal cord (n = 159) and brain (n = 9) tissues from GTEx and ANTE databases, respectively. PCA was performed to investigate cross-dataset compatibility in order to select proper reference group(s) for SIA comparison. The profiles from ANTE database were chosen because they were obtained using the same reagents, equipment, and protocols as for the current experimental SIA sampling (24). The GTEx reference group of samples was selected because this is, to our knowledge, currently the biggest publicly available collection of healthy spinal cord RNA-seq profiles [30]. Performing two-step expression analysis allowed us to select the DEGs between SIA and normal neural tissue explored using the same RNA-seq platform (brain samples from ANTE) and using a different

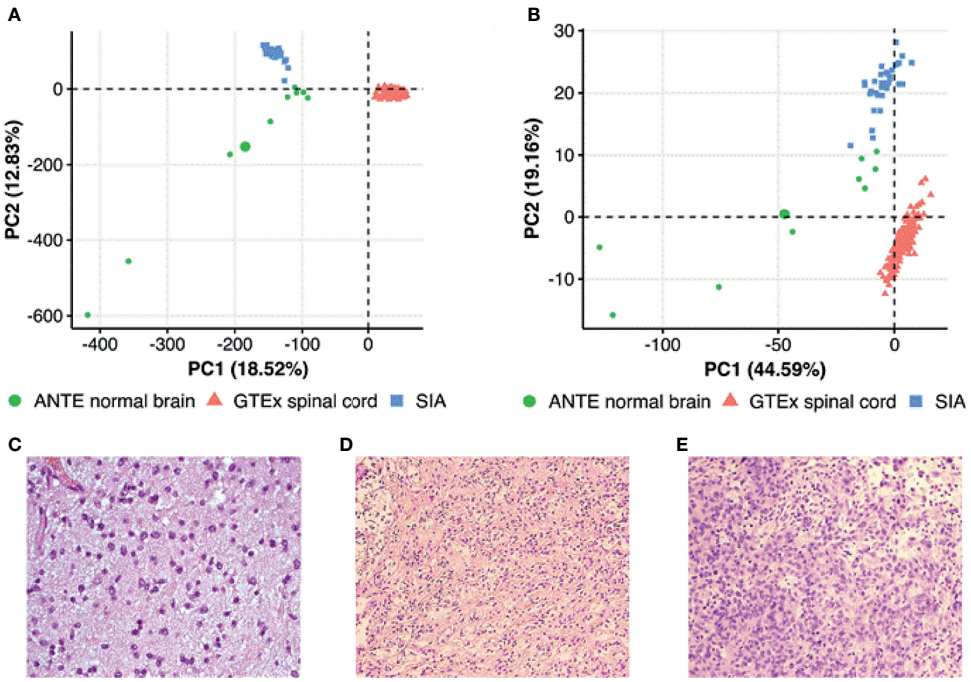



**FIGURE 1** | Principal component analysis (PCA) of **(A)** gene expression and **(B)** pathway activation levels (PALs) for spinal intramedullary astrocytoma (SIA) tissue samples and publicly available spinal cord and brain normal samples from The Genotype-Tissue Expression (GTEx) Portal and Atlas of Normal Tissue Expression (ANTE). PALs were calculated according to Borisov et al. (12) with the 168 healthy CNS tissue samples taken as the controls for SIA. Histological examples of **(C)** a diffuse, **(D)** a pilocytic, and **(E)** an anaplastic astrocytoma.

platform but for the same tissue type (GTEx spinal cord samples). Unfortunately, normal spinal cord samples were not available in the ANTE database. We hope that this approach allowed to establish differential gene expression profiles without the influence of platform-specific batch effect.

PCA was performed first in the space of $\log_{10}$ transformed quantile normalized gene counts. We observed tissue-specific sample clustering corresponding to the biological nature of the datasets under analysis, where SIA samples formed a separate cluster (**Figure 1A**) . In addition, we performed PCA for brain and spinal cord from GTEx, SIA, and ANTE normal brain data normalized using quantile normalization, DESeq2, and harmonized/batch corrected using XPN (26), CuBlock (27), and Shambhala (28). It appeared that GTEx brain profiles clustered with ANTE normal brain even in case of DESeq2 normalization, indicating that further batch correction was not necessary (**Supplementary Figure S2**).

Then, we performed PCA based on pathway activation levels (PALs) of 1,611 molecular pathways (29) calculated using the same transcriptomic data for each sample under study (**Figure 1B**). In case of pathway upregulation or downregulation, PALs can take positive or negative values, respectively, thus quantitatively reflecting the extent of a pathway activation or inhibition relatively to the control group of samples. Zero PAL values suggest unaffected activity of a molecular pathway. Thus, PAL values can be used as the quantitative functional characteristic of the interactome under analysis (14). We calculated PALs according to Borisov et al. (12), with the 168 healthy CNS tissue samples taken as the controls.

On the PAL-based PCA plot, we observed similar clustering as for the gene expression-based PCA (**Figure 1B**), thus strongly suggesting that SIA samples should be independently compared to each of the above healthy CNS tissue datasets.

## Differential Gene Expression Analysis

Subsequently, we performed paired differential gene expression analysis between SIA samples relatively to each dataset of healthy CNS tissues (**Supplementary Table S2**). Overall, 1,949 genes, 1,766 (90.61%) upregulated and 183 (9.39%) downregulated, were found to be statistically significantly differentially expressed [$|\log_2 FC|>5$, false discovery rate (FDR)-adjusted $p$-value <0.05] between SIA and GTEx spinal cord samples (**Figure 2A**). In turn, 382 DEGs, 102 (26.70%) upregulated and 280 (73.30%) downregulated, were found for the comparison between SIA and ANTE healthy brain samples (**Figure 2B**).

These DEG sets were then intersected with respect to $\log_2 FC$ sign (**Figures 3A, B**). In order to test whether an observed number of common differential genes can support random or non-random intersection hypothesis, we performed perturbation test for randomness according to Sorokin et al. (30) with 1,000 random gene sets. The percentile of the observed case precedent in the distribution of random intersections was considered as a measure of statistical significance.

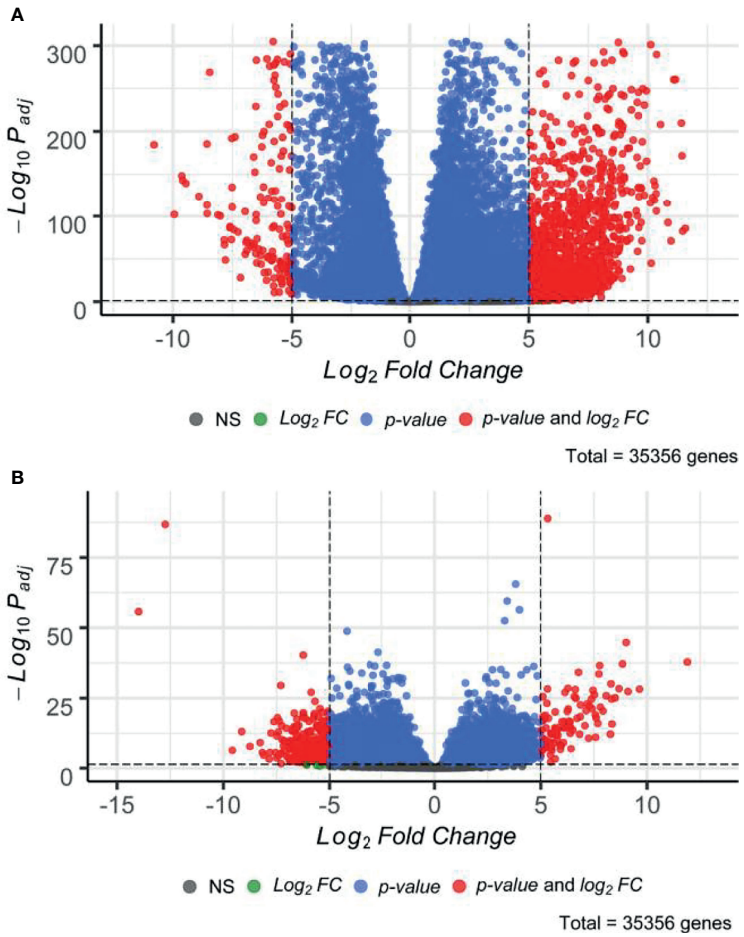



**FIGURE 2** | Distribution of differentially expressed genes between spinal intramedullary astrocytomas (SIAs) relative to **(A)** GTEx healthy spinal cord and **(B)** ANTE healthy brain samples.

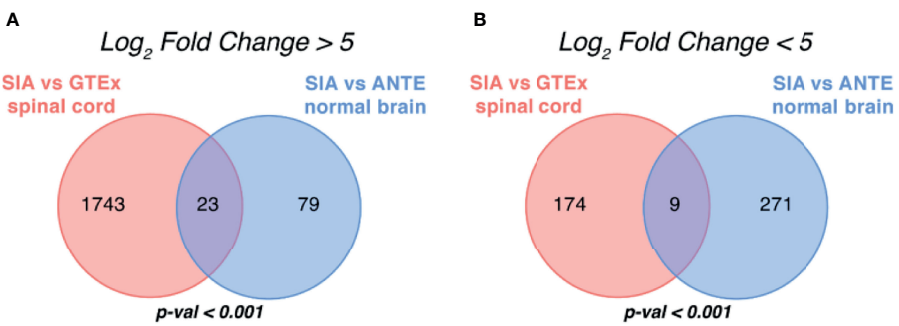



**FIGURE 3** | Intersection of differentially expressed gene sets between spinal intramedullary astrocytomas (SIAs), GTEx healthy spinal cord, and ANTE healthy brain samples. Intersections of **(A)** upregulated and **(B)** downregulated differentially expressed gene sets between SIA–GTEx spinal cord and SIA–ANTE normal brain samples are shown; *p*-values for intersection significance obtained in perturbation test are highlighted in bold.

In total, 32 genes, 23 (71.88%) upregulated and 9 (28.12%) downregulated, were commonly differentially expressed in SIA samples according to both comparisons, which supported the hypothesis that the intersections between the DEGs were non-random ($p < 0.01$).

## Gene Ontology Enrichment Analysis

To evaluate potential functional similarities of the above 32 SIA-specific differential genes and the underlying molecular and cellular processes, we then performed Gene Ontology (GO) terms enrichment analysis (**Figures 4A, B**). We identified significantly enriched 563 functional GO terms, where 340 terms (60.39%) were for upregulated and 223 (39.61%) were for downregulated DEGs. For statistical estimates, we used Benjamini–Hochberg method for FDR correction (31) and *p*-value threshold 0.05 (32).

Interestingly, most of the enriched terms for upregulated DEGs were related to the regulation of an innate and adaptive immune response, thus supporting a concept that immune microenvironment may play a crucial role in the development of spinal astrocytomas (33). In contrast, for downregulated DEGs, the most strongly enriched terms were linked with complex neuronal processes, such as cognition, learning ability, and regulation of neurotransmitter secretion and transport. The latter supports specific functional impairments occurring in astrocytomas in comparison with healthy CNS tissues (34).

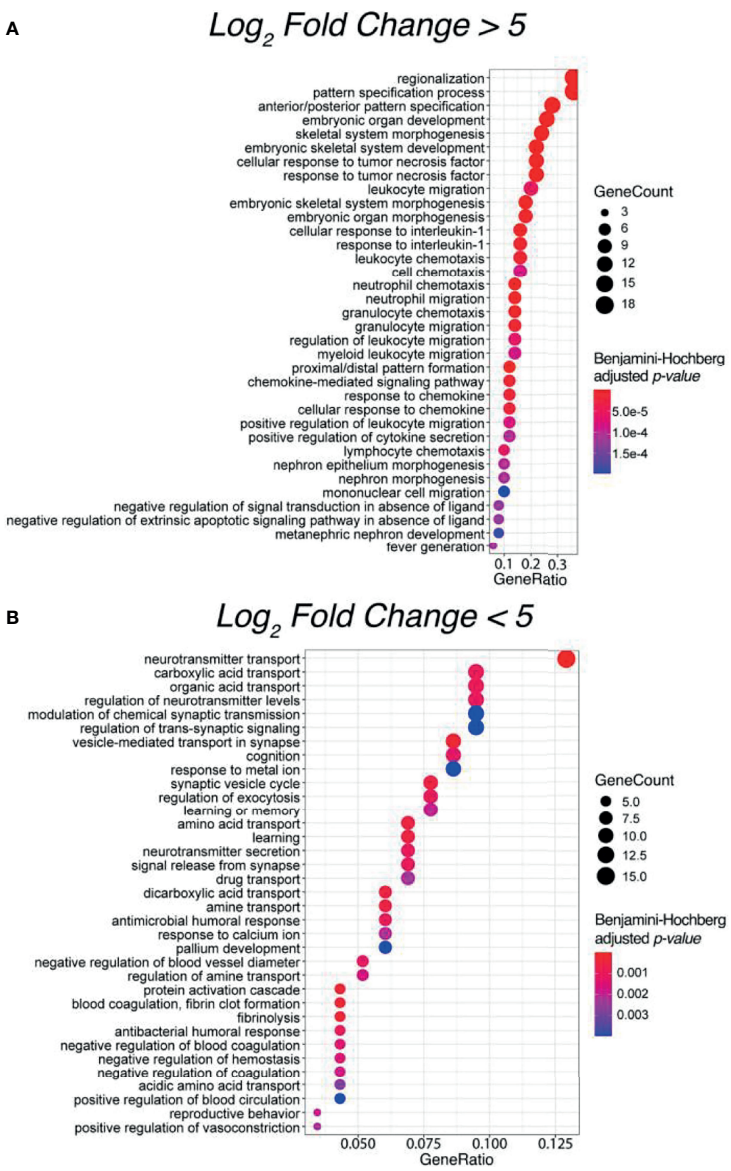



**FIGURE 4** | Top 35 enriched Gene Ontology (GO) terms for significantly **(A)** upregulated and **(B)** downregulated differentially expressed genes between spinal intramedullary astrocytomas (SIAs) and healthy CNS tissues: GTEx spinal cord and ANTE normal brain samples.

## Differential Pathway Activation Analysis

We then performed differential PAL analysis for SIA samples relative to healthy CNS tissues (**Figure 5**). In total, 1,611 molecular pathways including 10 and more gene products were interrogated from Oncobox pathway databank (29).

When comparing SIA and GTEx healthy spinal cord samples, we identified 468 differentially regulated pathways, 272 (58.12%) of them were activated and 196 (41.88%) were suppressed (**Figure 5**). In turn, for the comparison between SIA and ANTE normal brain samples, 464 differential pathways were identified; among them, 254 (54.74%) were activated and 210 (45.26%) were suppressed (**Figure 5**).

The intersections between these pathway sets returned 397 common differential molecular pathways [222 (55.92%) activated and 175 (44.08%) inhibited; **Supplementary Tables S3**, **S4**]. This supported non-random intersection between the differential pathways in both comparisons, $p < 0.001$ (**Figure 5**).

In both comparisons, top activated pathways deal with intracellular signal transduction and with the immune response, whereas top downregulated pathways are responsible for translational regulation and neurotransmitter activities (**Figure 6**). Except for the new feature of translational regulation, this trend was in line with the results obtained previously for the GO terms enrichment in the SIA differential genes (**Figure 4**).

## Simulated Activities of Anticancer Targeted Therapeutics

A chemotherapeutic treatment of SIA remains a challenging and poorly investigated field, and we performed computational simulation whether anticancer targeted drugs (ATDs) that are currently in use for other CNS tumors could be repurposed as second-line treatment options for SIA. To this end, we utilized Oncobox method for predicting efficiencies of ATDs based on gene expression and molecular pathway activation data (35). This returns for every drug a tumor sample-specific value of balanced drug efficiency score [drug score (DS)]. DS reflects an expected responsiveness of a tumor to a specific drug, where higher values mean higher expected efficacy of an ATD. Furthermore, drugs with positive DS are predicted to be potentially beneficial, and drugs with negative DS—potentially harmful (35). This method was shown to be clinically beneficial in a prospective clinical investigation on high-grade human solid tumors [16] and was effective for individual selection of experimental/off-label chemo- and targeted therapeutic settings, e.g., Buzdin et al. (36) and Moisseev et al. (37). In glioblastoma, Oncobox method could effectively predict tumor response on temozolomide, a DNA-alkylating agent whose activity is antagonized by *MGMT* gene products (38).

By using Oncobox algorithm, we identified 85 ATDs with positive DS in the SIA–GTEx spinal cord comparison, and 70 ATDs with positive DS in the SIA–ANTE healthy brain comparison (**Figures 7**, **8**). In these lists, there were 66 common drugs, thus evidencing non-random intersection between the two comparison results (**Figure 8** and **Supplementary Tables S5**, **S6**).

We then assessed available clinical trial reports for the top 20 DS-ranked drugs among these common 66 ATDs with the biggest DS values for different CNS tumors (**Table 1** and **Figure 7**). Interestingly, the top predicted drugs mostly represented the classes of tyrosine kinase inhibitors (i.e., regorafenib, lenvatinib, nintedanib, sorafenib, dovitinib, sunitinib, tivozanib, pazopanib, imatinib, foretinib, dasatinib, erdafitinib) and thalidomide analogs (thalidomide and pomalidomide). Many of these drugs were previously investigated for CNS tumors and related cancers like neuroblastoma (**Table 1**).

On the other hand, drugs with the predicted negative drug scores that were, therefore, algorithmically not recommended belonged mainly to cyclin-dependent kinase inhibitors and androgenic and anabolic steroid (AAS) classes.

## Spinal Intramedullary Astrocytoma Fusion Gene Transcripts

Chromosomal rearrangements resulting in fusion genes and abnormal transcripts in some cases may become clinically actionable targets of specific cancer therapeutics (71). Fusion transcripts combine exons of 2 or more genes and may serve as the oncogenic drivers in many cancers including CNS tumors (72, 73).

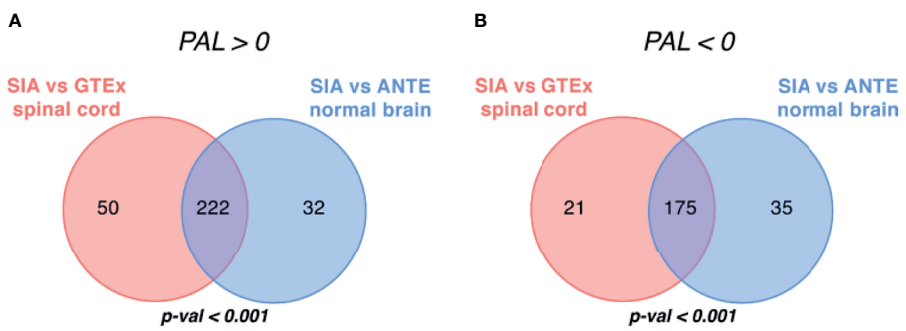



**FIGURE 5** | Intersection of differentially regulated molecular pathways between spinal intramedullary astrocytomas (SIAs), GTEx healthy spinal cord, and ANTE healthy brain samples. Intersections of significantly **(A)** upregulated (PAL >0) and **(B)** downregulated (PAL <0) molecular pathways between SIA–GTEx spinal cord and SIA–ANTE normal brain samples are shown; *p*-values for intersection significance obtained in perturbation test are highlighted in bold.

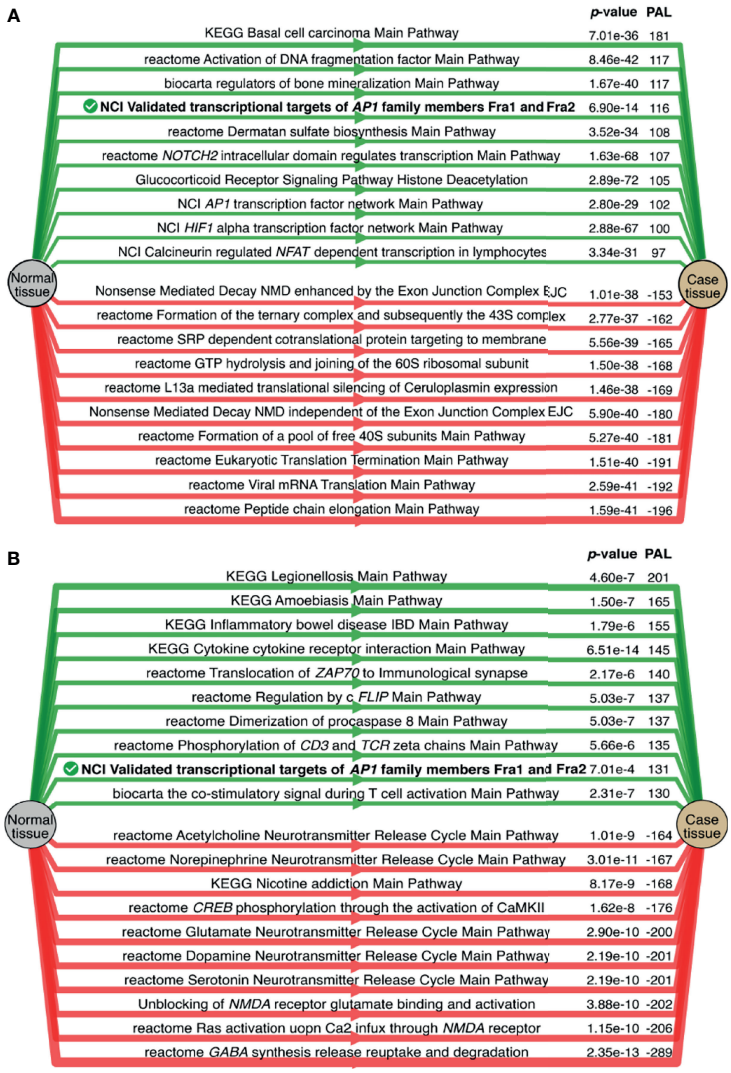



**FIGURE 6** | Top 10 activated (green) and suppressed (red) molecular pathways for the comparisons of spinal intramedullary astrocytomas (SIAs) with **(A)** GTEx healthy spinal cord and **(B)** ANTE normal brain samples. Common top differential pathways are shown in bold.

We used RNA-seq profiles for SIA patients to detect fusion transcripts presenting in spinal astrocytoma and focused on the fusions where at least on partner gene was a serine/threonine or tyrosine kinase. This allowed to select potentially druggable fusion genes. We found, in total, 16 different fusion transcripts identified by *STAR-Fusion* software (74) (**Figure 9** and **Supplementary Figure S3**). One of them, *KIAA1549-BRAF*, was found in two SIA patients and preserved BRAF kinase domain (**Figures 10A, B**). Interestingly, this fusion transcript was previously reported to confer a clinically less aggressive phenotype in pediatric low-grade astrocytoma (75). It was also found less abundant in the adult compared to pediatric patients with pilocytic astrocytoma (76). Other fusions, to our knowledge, were not reported previously and were not found in ChimerDB fusion database.

## DISCUSSION

SIAs are rare tumors comprising 6%–8% of all spinal cord tumors, and finding effective lines of treatment for SIA is a challenging task (1, 3). Currently, second-line treatments after surgical resection may include radiation and chemotherapy, where the regimen for adjunctive therapy was not optimally defined. Distance of tumor extension, type of surgery, and adjuvant therapy were significantly associated with SIA patients' survival in a previous study (77). The limited number of SIA clinical cases results in an absence of prospective studies (7–9). H3-K27 mutation previously showed diagnostic relevance and defined phenotypically and molecularly a distinct set of tumors (78). However, these mutations rarely occur in SIA with just several cases described in the literature (79).

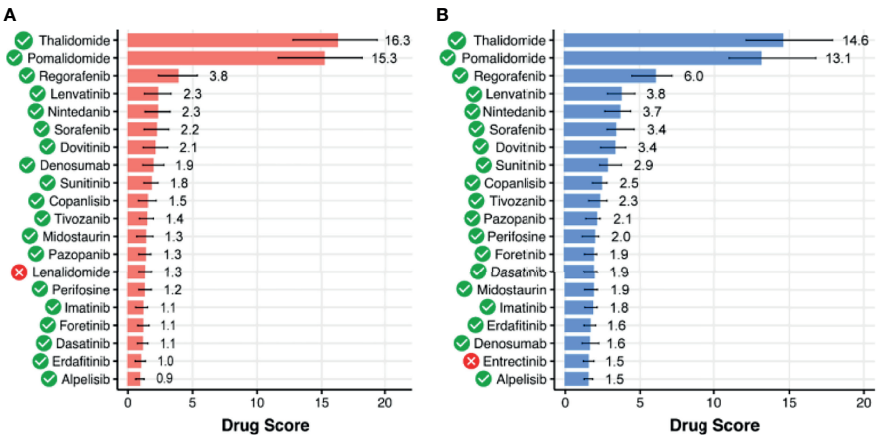



**FIGURE 7** | Top 20 targeted therapeutics ranked by drug score for spinal intramedullary astrocytomas (SIAs) separately normalized on **(A)** healthy GTEx spinal cord and **(B)** ANTE normal brain samples. Targeted therapeutics that are common between the two top-20 lists are shown with green marks.

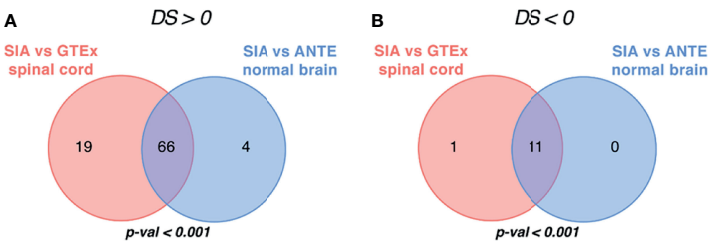



**FIGURE 8** | Intersection of targeted therapeutics assessed by Oncobox algorithm with **(A)** positive and **(B)** negative drug score (DS) predicted for spinal intramedullary astrocytomas (SIAs) separately normalized on GTEx healthy spinal cord and on ANTE normal brain samples; *p*-values for intersection significance obtained in perturbation test are highlighted in bold.

Here, we report the first RNA-seq molecular dataset with documented clinical annotations for 31 samples of 25 SIA patients. While there were studies by Biczok et al. (80) and Zhang et al. (81) on SIA molecular profiling, the current study is, to our knowledge, the only one with publicly available sequencing data. In addition, Biczok et al. (80) performed only targeted RNA-seq for 55 genes to detect fusion genes, while we investigated gene expression profiling using total RNA sequencing.

By comparing the experimental data obtained with the healthy brain and spinal cord CNS tissues, we analyzed SIA-specific DEGs, enrichment of GO terms, activation of molecular pathways, presence of fusion transcripts, and simulated efficacies of anticancer targeted drugs. Our study has certain limitations such as retrospective design and a small number of patients with mixed pathology of low-grade and high-grade spinal cord astrocytomas, although taking into account the rarity of this disease, this is expected.

Our results provide clues on possible molecular mechanisms of spinal astrocytoma and on its biomarkers. Indeed, a group of 23 differential SIA-upregulated genes found in the study was significantly enriched by GO terms mostly linked with regulation of innate and adaptive immune response. This strongly supports a role of the immune microenvironment in SIA development and progression. At the same time, a group of nine differentially downregulated genes was enriched by the terms dealing with neuronal and cognitive functions, thus reflecting their impairment in the cancer tissue.

Interestingly, a secreted extracellular matrix protein periostin was among the 23 SIA upregulated genes. Periostin was previously associated with prognosis and performance status in gliomas (82). Moreover, Mikheev et al. (83) showed that periostin knockdown impaired the survival of xenografted glioma stem cells and thus concluded that targeting periostin may be a promising strategy. Our study supports these findings and points to a potential role of periostin also in spinal astrocytomas.

Furthermore, our algorithmic simulation predicted that 66 targeted therapeutics can be potentially beneficial for SIA treatment. Some of them were already tested for CNS tumors and passed Phases I or II of clinical trials. We speculate that they could be repurposed from being used in other CNS tumors, and related tumors such as neuroblastoma, to improving the second-line treatment of SIA. Interestingly, the most highly ranked drugs

**TABLE 1 |** Overview of existing clinical trials conducted across Central Nervous System (CNS)-related tumors for target drugs with the highest drug score predicted for spinal intramedullary astrocytoma (SIA) samples.

| Drug | Nosology | Clinical Trials | Reference |
|---|---|---|---|
| Thalidomide | Advanced secondary glioblastoma | – | (39) |
| | Neuroblastoma | – | (40) |
| | Recurrent high-grade gliomas (anaplastic mixed glioma, anaplastic astrocytoma, or glioblastoma multiforme) | Phase II | (41) |
| | Astrocytoma | Phase II-Thalidomide, Temozolomide Tamoxifen | (42) |
| Pomalidomide | Pediatric recurrent, progressive, and refractory CNS tumors | Phase I (Phase II failed) NCT02415153 | (43) |
| | Glioblastoma multiforme | – | (44) |
| Regorafenib | Glioblastoma multiforme | Phase II NCT02926222 | (45) |
| | Recurrent high-grade astrocytoma | - (very poor performance) | (46) |
| | Neuroblastoma | – | (47) |
| Lenvatinib | Recurrent glioblastoma multiforme | Phase II NCT01433991 | (48) |
| | Pediatric solid tumors, including CNS tumors | Phase I/II NCT03245151 | (49) |
| Nintedanib (BIBF 1120) | Recurrent glioblastoma multiforme | Phase II (clinically non-relevant antitumor activity) NCT01251484 | (50) |
| | Glioblastomas/Anaplastic oligoastrocytoma/Gliosarcoma/Anaplastic Astrocytoma (AA)/Anaplastic oligodendroglioma (AO) | Phase II NCT01380782 | (51) |
| Sorafenib | Recurrent or progressive low-grade astrocytomas | Phase II NCT01338857 | (52) |
| | Progressive high-grade glioma | – | (53) |
| | Neuroblastoma | – | (54, 55) |
| Dovitinib | Recurrent glioblastoma | Phase II NCT01753713 | (56) |
| Sunitinib | Recurrent glioblastoma | Phase II/III NCT03025893 | (57) |
| | Recurrent glioblastoma and anaplastic astrocytoma | Phase II NCT00606008 | (58) |
| | Neuroblastoma | – | (59) |
| Copanlisib | – | – | – |
| Tivozanib | Recurrent glioblastoma | Phase II NCT01846871 | (60) |
| Pazopanib | Recurrent glioblastoma | Phase II NCT00459381 | (61) |
| Perifosine | Recurrent glioblastoma | Phase I - Temsirolimus, Perifosine NCT01051557 | (62) |
| | Recurrent or refractory pediatric CNS and solid tumors | Phase I | (63) |
| Foretenib | Glioblastoma multiforme | – | (64, 65) |
| Dasatinib | Glioblastoma | - (lack of activity against recurrent glioblastoma) | (66) |
| Midostaurin | – | – | – |
| Imatinib | Relapsed/Refractory neuroblastoma | Phase II NCT00030667 | (67, 68) |
| | Recurrent glioblastoma | Phase II NCT00010049 | (69) |
| | Metastatic pilocytic astrocytoma | – | (70) |
| Denosumab | – | – | – |
| Alpelisib | – | – | – |

(thalidomide and its derivatives) also reflect the top GO terms enriched in the SIA-upregulated gene set, i.e., the cellular response on tumor necrosis factor (**Figure 4A**), whose pathway is a primary molecular target for these drugs (84). Many tyrosine kinase inhibitory drugs were also predicted to be beneficial in SIA treatment.

In contrast, there are some cancer therapeutics we predicted to be potentially harmful for treating SIA, which mainly related to cyclin-dependent kinase inhibitors (-*Ciclibs*) and AAS hormones.

At the level of molecular pathway analysis, we could identify several specific molecular features of SIA. For example, the top upregulated pathways were associated with transcriptional targets of AP1 (Activator Protein-1) family member transcription factors *FOSL1* and *FOSL2*. Interestingly, AP1 transcription factors FOS and Fra1 were found to be upregulated in pilocytic astrocytomas (85). Moreover, Fra1 was shown to control architecture and migratory nature of glioblastoma cells (86). This protein is also linked with promotion of glioma aggressiveness through epithelial–

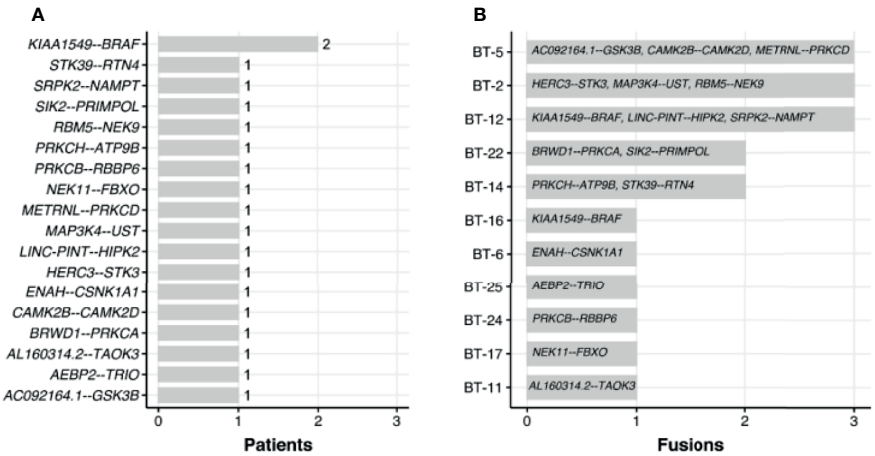



**FIGURE 9** | Occurrence of fusions found across spinal intramedullary astrocytoma (SIA) samples by number of patients **(A)** and fusions **(B)**.

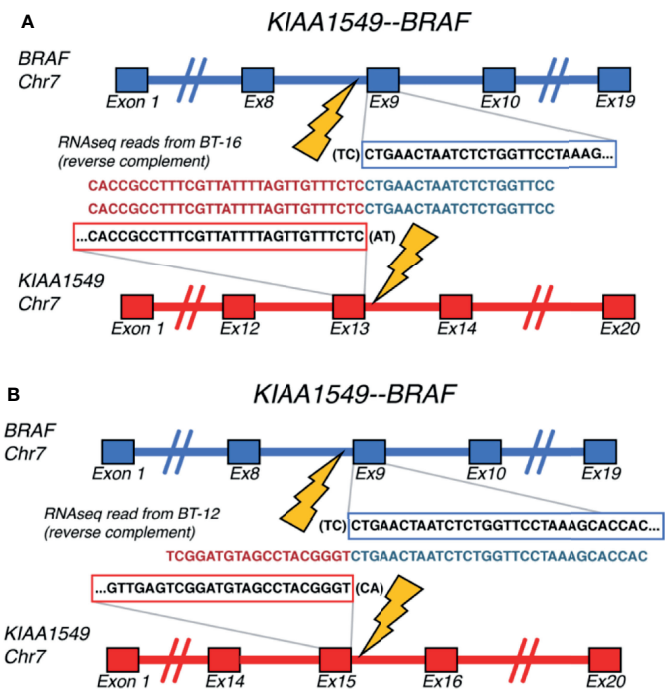



**FIGURE 10** | Schematic representation of the *KIAA1549-BRAF* fusion transcripts identified for **(A)** BT-16 and **(B)** BT-12 samples of spinal intramedullary astrocytoma (SIA).

mesenchymal transition (87) and overall glioblastoma invasion (88). In turn, downregulation of Fra1 enhances drug sensitivity in breast cancer cells (89). Also, experimental Fra1 inhibitors significantly suppressed tumor growth and lymph node metastasis of head and neck cancers in a patient-derived xenograft model (90). Thus, our results suggest that Fra1 could be investigated as a potential drug target in rare CNS tumors, such as SIA.

Finally, 16 different fusion transcripts identified in this study suggest the occurrence of chromosomal rearrangements resulting in fusion oncogenes and abnormal transcripts in SIA. Moreover, *KIAA1549-BRAF* fusion detected for two adult SIA patients was previously relatively frequently found in pediatric pilocytic astrocytomas [47] and was reported to confer a clinically less aggressive phenotype in pediatric low-grade astrocytoma (75). Thus, fusion transcripts found can be

potentially clinically relevant for SIAs and should be further tested, since many gene fusions were reported as oncogenic drivers in CNS tumors (72, 73).

# MATERIALS AND METHODS

## Experimental Clinical Biosamples

This study was performed in agreement with the ethical principles of Declaration of Helsinki. Retrospective biosamples were obtained from patients diagnosed with SIAs who had undergone surgery at the spinal department of Burdenko Neurosurgical Center, Moscow. From all the patients involved or from their legal representatives, informed written consents to participate in this study were collected. The study design and consent collection procedure were approved by the local ethical committee of the Burdenko Neurosurgical Center. For all patients enrolled and for their biosamples, the consent was obtained for publication of age, sex, histological tumor type, diagnosis, and molecular data including RNA-seq profiles but not including whole-genome and/or whole-exome sequencing data.

Biosamples were FFPE tumor tissue blocks that were evaluated and confirmed by a pathologist who estimated a proportion of tumor cells and determined the histological type of a tumor. In this study, only FFPE blocks with at least 70% tumor cells were analyzed. In total, 31 samples for 25 SIAs meeting the above criteria were obtained for further molecular screenings (**Supplementary Table S5**).

## Preparation of Libraries and RNA Sequencing

RNA libraries were generated and sequenced according to Suntsova et al. (24). RNA was extracted using RecoverAll™ Total Nucleic Acid Isolation Kit (Invitrogen). RNA concentrations were measured with Qubit RNA Assay Kit, and Agilent 2100 bioanalyzer was used to measure RNA Integrity Number (RIN). Depletion of ribosomal RNA was performed using RNA Hyper Kit (Roche), and then library concentrations and fragment length distributions were measured with Qubit (Life Technologies) and Agilent Tapestation (Agilent), respectively. The RNA-seq was performed using Illumina NextSeq 550 engine for 50-bp single-end reads and approximately 30 million raw reads per sample using standard protocol. Single-end sequencing was used because SIA samples were FFPE tissue blocks that typically have a strong degree of RNA degradation. Primary sequencing data quality control was performed with Illumina SAV, and demultiplexing was made according to Suntsova et al. (24) with Illumina Bcl2fastq2 v 2.17 software.

## RNA Sequencing Data Processing

SIA profiles were processed according to Suntsova et al. (24). STAR aligner (91) was used to process FASTQ files from RNA-seq in "GeneCounts" mode for Ensembl human transcriptome annotation GRCh38.89. The gene names for 36,596 annotated genes were converted to HGNC (HUGO Gene Nomenclature Committee) gene symbols from Ensembl IDs according to Complete HGNC dataset, version of August 17, 2021 (https://www.genenames.org). Further quality control metrics for RNA-seq data were obtained with NCBI MAGIC software (92–94). RNA-seq profiles were preprocessed by quantile normalization method (95), and then differential expression analysis was performed using *DESeq2* (96), and visualized with R package *EnhancedVolcano* (97). Genes that were considered significantly differentially expressed had to pass a threshold of FDR-adjusted *p*-values <0.05 (98). GO enrichment analysis was conducted using *clusterProfile* (v.4.2.1) and *org.Hs.eg.db* (v.3.8.2) R packages. Fusion transcripts were detected with *STAR-Fusion* tool (74), and PCA and visualization were done for $\log_{10}$ transformed counts of all genes using *pca2d* R (v.3.6.0) and *prcomp* software. Code for data analysis is available at: https://github.com/raevskymichail/SIA_analysis.

## Healthy Tissue Transcriptomic Data

We used healthy tissue transcriptomic profiles obtained for normal human spinal cord biosamples from GTEx project portal (25) and for normal brain samples from the ANTE database (24). In total, 9 ANTE and 159 GTEx normal CNS samples were analyzed. Raw count quantification in GENCODEv26 annotation was obtained from GTEx portal.

## Calculation of Pathway Activation Levels

Algorithmically annotated molecular pathway graphs were taken from our previously published database (29). PALs were calculated with the Oncobox bioinformatic platform. It allows quantitative assessment of PALs using RNA-seq data and functionally annotated collection of molecular pathways (12, 29). We used a set of 1,611 pathways with 10 or more gene products included because of previously reported poor theoretically estimated data aggregation effect for smaller pathways (15).

This method of calculating PAL showed a strong potential to suppress batch effects (15, 16, 21) and to minimize the artifacts introduced by the methods of experimental transcriptome analysis (13, 99). An absolute value of PAL reflects the strength of the pathway up/downregulation, while a positive or negative sign indicates its activation or suppression, accordingly (12). To calculate PAL, each sample RNA-seq profile was normalized on mean geometric levels of gene expression in the relevant control dataset.

## Targeted Drug Efficiency Simulation

Drug score [Balanced Efficiency Score (BES)] for cancer-targeted drugs was calculated according to Tkachev et al. (35), whose method is based on the analysis of targeted molecular pathway activation and relative expression levels of drug target genes.

## Testing of Intersection Significance

To test whether an observed number of overlapping differential genes or pathways between the two intersecting datasets is significant, for every comparison, we performed 1,000 random intersections according to Sorokin et al. (30). In every case, two

random samples from the corresponding gene sets under comparison were taken. Then, these random samples were intersected for 1,000 iterations, and numbers of randomly obtained common genes were registered. Then, *p*-value of intersection significance was calculated as an expected fraction of random intersects that is equal to or higher than the experimentally observed number of overlapping genes.

## CONCLUSIONS

We report here the first clinically annotated RNA-seq dataset for 31 tumor tissue samples of 25 patients with SIA, a rare CNS tumor. Bioinformatic analysis revealed the presence of characteristic *KIAA1549-BRAF* fusion transcripts in two samples and 16 new fusions each present in one SIA patient. For the first time, differential gene and molecular pathway analysis showed that the top SIA-upregulated pathways govern interactions with the immune cells, whereas the top inhibited pathways deal with normal neuronal activities. In addition, we found SIA-specific activation of molecular targets for cancer drugs: several tyrosine kinase inhibitors and thalidomide analogs. While this is a theoretical prediction, we propose that they could be further investigated as second-line treatment agents to aid in the prevention of SIA recurrence and progression.

## DATA AVAILABILITY STATEMENT

Transcriptomic data used across the analyses was deposited in the Sequencing Read Archive under BioProject accession: PRJNA763174. Normalized gene counts are presented in **Supplementary Dataset 1**. Description of clinically relevant (age, sex, diagnosis, tumor histotype) information is given in **Supplementary Table S1**.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethical Committee of the Burdenko Neurosurgical Center. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

Software: MR. Visualization: MR. Validation: MR. Writing—original draft: MR, MS, and AB. Writing—review and editing: MR, MS, and AB. Investigation: MR. Data curation: MR. Conceptualization: MS and AB. Methodology: MS. Supervision: MS. Project administration: AB. Patient treatment: NK, DA, SK, and ST. Tissue sample collection: NK, DA, SK, and ST. Writing—review and editing: SK and ST. All authors have read and agreed to the published version of the article.

## FUNDING

This research was funded by The Russian Foundation for Basic Research, grant number 18-29-01042.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.837570/full#supplementary-material

**Supplementary Figure 1 |** Kaplan-Meier plots for progression-free and overall survival.

**Supplementary Figure 2 |** PCA plots for log-transformed counts processed using **(A)** quantile normalization; **(B)** DESeq2 normalization; **(C)** XPN; **(D)** CuBlock; **(E, F)** Shambhala.

**Supplementary Figure 3 |** Schematic representation of 16 different fusion transcripts identified by STAR-Fusion for spinal intramedullary astocytoma (SIA) transcriptomic samples

**Supplementary Data Sheet 1 |** Normalized gene counts.

**Supplementary Table 1 |** Clinical annotations of spinal intramedullary astrocytomas tissue specimens analysed in this study.

**Supplementary Table 2 |** Gene lists for the intersection of upregulated and downregulated differentially expressed gene sets between spinal intramedullary astrocytomas (SIA), GTEx healthy spinal cord, and ANTE healthy brain samples.

**Supplementary Table 3 |** Pathway-activation-levels (PAL) from the differential pathway activation analysis estimated for 1611 molecular pathways from Oncobox pathway databank between spinal intramedullary astrocytomas (SIA) and GTEx healthy spinal cord samples.

**Supplementary Table 4 |** Pathway-activation-levels (PAL) from the differential pathway activation analysis calculated across 1611 molecular pathways from Oncobox pathway databank between spinal intramedullary astrocytomas (SIA) and ANTE normal brain samples.

**Supplementary Table 5 |** Simulated activities of anticancer targeted therapeutics (Drug Scores), calculated between spinal intramedullary astrocytomas (SIA) and GTEx healthy spinal cord samples.

**Supplementary Table 6 |** Simulated activities of anticancer targeted therapeutics (Drugs Scores), estimated between spinal intramedullary astrocytomas (SIA) and ANTE normal brain samples.

## REFERENCES

1. Das J, Hoang S, Mesfin F. Intramedullary Spinal Cord Tumors. *Neurocirugia* (2021) 2:213–8. doi: 10.1016/S1130-1473(91)71148-2

2. Milano MT, Johnson MD, Sul J, Mohile NA, Korones DN, Okunieff P, et al. Primary Spinal Cord Glioma: A Surveillance, Epidemiology, and End Results Database Study. *J Neurooncol* (2010) 98:83–92. doi: 10.1007/s11060-009-0054-7

3. Ogunlade J, Wiginton JG IV, Elia C, Odell T, Rao SC. Primary Spinal Astrocytomas: A Literature Review. *Cureus* (2019) 11:e5247. doi: 10.7759/CUREUS.5247
4. Townsend N, Handler M, Fleitz J, Foreman N. Intramedullary Spinal Cord Astrocytomas in Children. *Pediatr Blood Cancer* (2004) 43:629–32. doi: 10.1002/PBC.20082
5. Hussain I, Parker WE, Barzilai O, Bilsky MH. Surgical Management of Intramedullary Spinal Cord Tumors. *Neurosurg Clin N Am* (2020) 31:237–49. doi: 10.1016/J.NEC.2019.12.004
6. Helal A, Alvi MA, Everson M, Mikula AL, Cohen Cohen S, Bydon M, et al. Prognostic Factors Independently Associated With Improved Progression-Free Survival After Surgical Resection in Patients With Spinal Cord Astrocytomas: An Institutional Case Series. *Oper Neurosurg* (2022) 22:106–14. doi: 10.1227/ONS.0000000000000084
7. Raco A, Piccirilli M, Landi A, Lenzi J, Delfini R, Cantore G. High-Grade Intramedullary Astrocytomas: 30 Years' Experience at the Neurosurgery Department of the University of Rome "Sapienza". *J Neurosurg Spine* (2010) 12:144–53. doi: 10.3171/2009.6.SPINE08910
8. Raco A, Esposito V, Lenzi J, Piccirilli M, Delfini R, Cantore G. Long-Term Follow-Up of Intramedullary Spinal Cord Tumors: A Series of 202 Cases. *Neurosurgery* (2005) 56:972–81. doi: 10.1227/01.NEU.0000158318.66568.CC
9. Minehan K, Brown P, Scheithauer B, Krauss W, Wright M. Prognosis and Treatment of Spinal Cord Astrocytoma. *Int J Radiat Oncol Biol Phys* (2009) 73:727–33. doi: 10.1016/J.IJROBP.2008.04.060
10. Rodon J, Soria JC, Berger R, Miller WH, Rubin E, Kugel A, et al. Genomic and Transcriptomic Profiling Expands Precision Cancer Medicine: The WINTHER Trial. *Nat Med* (2019) 25:751–8. doi: 10.1038/s41591-019-0424-4
11. Poddubskaya E, Sorokin M, Garazha A, Glusker A, Moisseev A, Sekacheva M, et al. Clinical Use of RNA Sequencing and Oncobox Analytics to Predict Personalized Targeted Therapeutic Efficacy. *J Clin Oncol* (2020) 38:e13676–6. doi: 10.1200/JCO.2020.38.15_suppl.e13676
12. Borisov N, Sorokin M, Garazha A, Buzdin A. Quantitation of Molecular Pathway Activation Using RNA Sequencing Data. *Methods Mol Biol* (2020) 2063:189–206. doi: 10.1007/978-1-0716-0138-9_15
13. Ozerov IV, Lezhnina KV, Izumchenko E, Artemov AV, Medintsev S, Vanhaelen Q, et al. In Silico Pathway Activation Network Decomposition Analysis (iPANDA) as a Method for Biomarker Development. *Nat Commun* (2016) 7:13427. doi: 10.1038/ncomms13427
14. Buzdin A, Sorokin M, Garazha A, Sekacheva M, Kim E, Zhukov N, et al. Molecular Pathway Activation – New Type of Biomarkers for Tumor Morphology and Personalized Selection of Target Drugs. *Semin Cancer Biol* (2018) 53:110–24. doi: 10.1016/j.semcancer.2018.06.003
15. Borisov N, Suntsova M, Sorokin M, Garazha A, Kovalchuk O, Aliper A, et al. Data Aggregation at the Level of Molecular Pathways Improves Stability of Experimental Transcriptomic and Proteomic Data. *Cell Cycle* (2017) 16:1810–23. doi: 10.1080/15384101.2017.1361068
16. Sorokin M, Raevskiy M, Zottel A, Šamec N, Vidmar MS, Matjašič A, et al. Large-Scale Transcriptomics-Driven Approach Revealed Overexpression of CRNDE as a Poor Survival Prognosis Biomarker in Glioblastoma. *Cancers* (2021) 13:3419. doi: 10.3390/CANCERS13143419. 2021, Vol 13, Page 3419.
17. Artemov A, Aliper A, Korzinkin M, Lezhnina K, Jellen L, Zhukov N, et al. A Method for Predicting Target Drug Efficiency in Cancer Based on the Analysis of Signaling Pathway Activation. *Oncotarget* (2015) 6:29347–56. doi: 10.18632/oncotarget.5119
18. Zhu Q, Izumchenko E, Aliper AM, Makarev E, Paz K, Buzdin AA, et al. Pathway Activation Strength is a Novel Independent Prognostic Biomarker for Cetuximab Sensitivity in Colorectal Cancer Patients. *Hum Genome Var* (2015) 2:15009. doi: 10.1038/hgv.2015.9
19. Crispo F, Notarangelo T, Pietrafesa M, Lettini G, Storto G, Sgambato A, et al. Braf Inhibitors in Thyroid Cancer: Clinical Impact, Mechanisms of Resistance and Future Perspectives. *Cancers (Basel)* (2019) 11:1388. doi: 10.3390/cancers11091388
20. Borisov NM, Terekhanova NV, Aliper AM, Venkova LS, Smirnov PY, Roumiantsev S, et al. Signaling Pathways Activation Profiles Make Better Markers of Cancer Than Expression of Individual Genes. *Oncotarget* (2014) 5:10198–205. doi: 10.18632/oncotarget.2548
21. Kamashev D, Sorokin M, Kochergina I, Drobyshev A, Vladimirova U, Zolotovskaia M, et al. Human Blood Serum can Donor-Specifically Antagonize Effects of EGFR-Targeted Drugs on Squamous Carcinoma Cell Growth. *Heliyon* (2021) 7:E06394. doi: 10.1016/j.heliyon.2021.e06394
22. Davis C, Hitz B, Sloan C, Chan E, Davidson J, Gabdank I, et al. The Encyclopedia of DNA Elements (ENCODE): Data Portal Update. *Nucleic Acids Res* (2018) 46:D794–801. doi: 10.1093/NAR/GKX1081
23. Huang X, Stern D, Zhao H. Transcriptional Profiles From Paired Normal Samples Offer Complementary Information on Cancer Patient Survival–Evidence From TCGA Pan-Cancer Data. *Sci Rep* (2016) 6:20567. doi: 10.1038/SREP20567
24. Suntsova M, Gaifullin N, Allina D, Reshetun A, Li X, Mendeleeva L, et al. Atlas of RNA Sequencing Profiles for Normal Human Tissues. *Sci Data* (2019) 6:36. doi: 10.1038/s41597-019-0043-4
25. Consortium TGte. The GTEx Consortium Atlas of Genetic Regulatory Effects Across Human Tissues. *Science (80- )* (2020) 369:1318–30. doi: 10.1126/SCIENCE.AAZ1776
26. Shabalin AA, Tjelmeland H, Fan C, Perou CM, Nobel AB. Merging Two Gene-Expression Studies *via* Cross-Platform Normalization. *Bioinformatics* (2008) 24:1154–60. doi: 10.1093/bioinformatics/btn083
27. Junet V, Farrés J, Mas JM, Daura X. CuBlock: A Cross-Platform Normalization Method for Gene-Expression Microarrays. *Bioinformatics* (2021) 37:2365–73. doi: 10.1093/bioinformatics/btab105
28. Borisov N, Shabalina I, Tkachev V, Sorokin M, Garazha A, Pulin A, et al. Shambhala: A Platform-Agnostic Data Harmonizer for Gene Expression Data 06 Biological Sciences 0604 Genetics. *BMC Bioinf* (2019) 20:66. doi: 10.1186/s12859-019-2641-8
29. Sorokin M, Borisov N, Kuzmin D, Gudkov A, Zolotovskaia M, Garazha A, et al. Algorithmic Annotation of Functional Roles for Components of 3,044 Human Molecular Pathways. *Front Genet* (2021) 12:617059. doi: 10.3389/fgene.2021.617059
30. Sorokin M, Ignatev K, Barbara V, Vladimirova U, Muraveva A, Suntsova M, et al. Molecular Pathway Activation Markers Are Associated With Efficacy of Trastuzumab Therapy in Metastatic HER2-Positive Breast Cancer Better Than Individual Gene Expression Levels. *Biochem* (2020) 85:758–72. doi: 10.1134/S0006297920070044
31. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the False Discovery Rate in Behavior Genetics Research. *Behav Brain Res* (2001) 125:279–84. doi: 10.1016/S0166-4328(01)00297-2
32. Guan Q, Chen R, Yan H, Cai H, Guo Y, Li M, et al. Differential Expression Analysis for Individual Cancer Samples Based on Robust Within-Sample Relative Gene Expression Orderings Across Multiple Profiling Platforms. *Oncotarget* (2016) 7:68909–20. doi: 10.18632/ONCOTARGET.11996
33. Hewedi IH, Radwan NA, Shash LS, Elserry TH. Perspectives on the Immunologic Microenvironment of Astrocytomas. *Cancer Manag Res* (2013) 5:293. doi: 10.2147/CMAR.S48942
34. Pletschko T, Felnhofer A, Lamplmair D, Dorfer C, Czech T, Chocholous M, et al. Cerebellar Pilocytic Astrocytoma in Childhood: Investigating the Long-Term Impact of Surgery on Cognitive Performance and Functional Outcome. *Dev Neurorehabil* (2018) 21:415. doi: 10.1080/17518423.2017.1370502
35. Tkachev V, Sorokin M, Garazha A, Borisov N, Buzdin A. "Oncobox Method for Scoring Efficiencies of Anticancer Drugs Based on Gene Expression Data". In: K Astakhova, SA Bukhari, editors. *Methods Mol Biol*. New York, NY: Springer US (2020). p. 235–55. doi: 10.1007/978-1-0716-0138-9_17
36. Buzdin A, Sorokin M, Garazha A, Glusker A, Aleshin A, Poddubskaya E, et al. RNA Sequencing for Research and Diagnostics in Clinical Oncology. *Semin Cancer Biol* (2019) 60:311–23. doi: 10.1016/j.semcancer.2019.07.010
37. Moisseev A, Albert E, Lubarsky D, Schroeder D, Clark J. Transcriptomic and Genomic Testing to Guide Individualized Treatment in Chemoresistant Gastric Cancer Case. *Biomedicines* (2020) 8:67. doi: 10.3390/BIOMEDICINES8030067
38. Kim EL, Sorokin M, Kantelhardt SR, Kalasauskas D, Sprang B, Fauss J, et al. Intratumoral Heterogeneity and Longitudinal Changes in Gene Expression Predict Differential Drug Sensitivity in Newly Diagnosed and Recurrent Glioblastoma. *Cancers (Basel)* (2020) 12:520. doi: 10.3390/cancers12020520
39. Hassler MR, Sax C, Flechl B, Ackerl M, Preusser M, Hainfellner JA, et al. Thalidomide as Palliative Treatment in Patients With Advanced Secondary Glioblastoma. *Oncology* (2015) 88:173–9. doi: 10.1159/000368903
40. Gesundheit B, Moser A, Or R, Klement G. Successful Antiangiogenic Therapy for Neuroblastoma With Thalidomide. *J Clin Oncol* (2007) 25:5321–4. doi: 10.1200/JCO.2007.12.3331

41. Fine HA, Figg WD, Jaeckle K, Wen PY, Kyritsis AP, Loeffler JS, et al. Phase II Trial of the Antiangiogenic Agent Thalidomide in Patients With Recurrent High-Grade Gliomas. *J Clin Oncol* (2000) 18:708–15. doi: 10.1200/JCO.2000.18.4.708
42. Rabbani G, Benzil D, Wallam MN, Chen B, Hoang A, Kancherla R, et al. Combination Therapy With Thalidomide, Temozolomide and Tamoxifen Improves Quality of Life in Patients With Malignant Astrocytomas. *Anticancer Res* (2007) 27:2729–36.
43. Fangusaro J, Mitchell DA, Kocak M, Robinson GW, Baxter PA, Hwang EI, et al. Phase 1 Study of Pomalidomide in Children With Recurrent, Refractory, and Progressive Central Nervous System Tumors: A Pediatric Brain Tumor Consortium Trial. *Pediatr Blood Cancer* (2021) 68:e28756. doi: 10.1002/PBC.28756
44. Luo Z, Wang B, Liu H, Shi L. TNF Inhibitor Pomalidomide Sensitizes Glioblastoma Cells to EGFR Inhibition. *Ann Clin Lab Sci* (2020) 50:474–80.
45. Detti B, Scoccianti S, Lucidi S, Maragna V, Teriaca MA, Ganovelli M, et al. Regorafenib in Glioblastoma Recurrence: A Case Report. *Cancer Treat Res Commun* (2021) 26:100263. doi: 10.1016/J.CTARC.2020.100263
46. Kebir S, Rauschenbach L, Radbruch A, Lazaridis L, Schmidt T, Stoppek AK, et al. Regorafenib in Patients With Recurrent High-Grade Astrocytoma. *J Cancer Res Clin Oncol* (2019) 145:1037–42. doi: 10.1007/S00432-019-02868-5
47. Subramonian D, Phanthilath N, Rinehardt H, Flynn S, Huo Y, Zhang J, et al. Regorafenib is Effective Against Neuroblastoma *In Vitro* and *In Vivo* and Inhibits the RAS/MAPK, PI3K/Akt/mTOR and Fos/Jun Pathways. *Br J Cancer* (2020) 123:568–79. doi: 10.1038/S41416-020-0905-8
48. Reardon DA, Pan E, Fan J, Mink J, Barboriak DP, Vredenburgh JJ, et al. A Phase 2 Trial of the Multitargeted Kinase Inhibitor Lenvatinib (E7080) in Patients (PTS) With Recurrent Glioblastoma (GBM) And Disease Progression Following Prior Bevacizumab Treatment. *Ann Oncol* (2012) 23:ix146. doi: 10.1016/S0923-7534(20)33001-5
49. dela Cruz FS, Fox E, Muscal JA, Kirov I, Geller JI, DuBois SG, et al. A Phase I/II Study of Lenvatinib (LEN) Plus Everolimus (EVE) in Recurrent and Refractory Pediatric Solid Tumors, Including CNS Tumors. *J Clin Oncol* (2020) 38:10527–7. doi: 10.1200/JCO.2020.38.15_SUPPL.10527
50. Muhic A, Poulsen HS, Sorensen M, Grunnet K, Lassen U. Phase II Open-Label Study of Nintedanib in Patients With Recurrent Glioblastoma Multiforme. *J Neurooncol* (2013) 111:205–12. doi: 10.1007/S11060-012-1009-Y
51. Norden AD, Schiff D, Ahluwalia MS, Lesser GJ, Nayak L, Lee EQ, et al. Phase II Trial of Triple Tyrosine Kinase Receptor Inhibitor Nintedanib in Recurrent High-Grade Gliomas. *J Neurooncol* (2015) 121:297–302. doi: 10.1007/S11060-014-1631-Y
52. Karajannis MA, Legault G, Fisher MJ, Milla SS, Cohen KJ, Wisoff JH, et al. Phase II Study of Sorafenib in Children With Recurrent or Progressive Low-Grade Astrocytomas. *Neuro Oncol* (2014) 16:1408–16. doi: 10.1093/NEUONC/NOU059
53. Hassler MR, Ackerl M, Flechl B, Sax C, Wöhrer A, Widhalm G, et al. Sorafenib for Patients With Pretreated Recurrent or Progressive High-Grade Glioma: A Retrospective, Single-Institution Study. *Anticancer Drugs* (2014) 25:723–8. doi: 10.1097/CAD.0000000000000077
54. Kakodkar NC, Peddinti RR, Tian Y, Guerrero LJ, Chlenski A, Yang Q, et al. Sorafenib Inhibits Neuroblastoma Cell Proliferation and Signaling, Blocks Angiogenesis, and Impairs Tumor Growth. *Pediatr Blood Cancer* (2012) 59:642–7. doi: 10.1002/PBC.24004
55. Okada K, Nakano Y, Yamasaki K, Nitani C, Fujisaki H, Hara J. Sorafenib Treatment in Children With Relapsed and Refractory Neuroblastoma: An Experience of Four Cases. *Cancer Med* (2016) 5:1947–9. doi: 10.1002/CAM4.784
56. Sharma M, Schilero C, Peereboom DM, Hobbs BP, Elson P, Stevens GHJ, et al. Phase II Study of Dovitinib in Recurrent Glioblastoma. *J Neurooncol* (2019) 144:359–68. doi: 10.1007/S11060-019-03236-6
57. Brahm CG, Van Linde ME, Labots M, Kouwenhoven MCM, Sanchez Aliaga E, Enting RH, et al. 436tip A Phase II/III Trial of High-Dose, Intermittent Sunitinib in Patients With Recurrent Glioblastoma: The STELLAR Study. *Ann Oncol* (2019) 30:v143–4. doi: 10.1093/annonc/mdz243
58. Pan E, Yu D, Yue B, Potthast L, Chowdhary S, Smith P, et al. A Prospective Phase II Single-Institution Trial of Sunitinib for Recurrent Malignant Glioma. *J Neurooncol* (2012) 110:111–8. doi: 10.1007/S11060-012-0943-Z
59. Calero R, Morchon E, Johnsen JI, Serrano R. Sunitinib Suppress Neuroblastoma Growth Through Degradation of MYCN and Inhibition of Angiogenesis. *PloS One* (2014) 9:e95628. doi: 10.1371/JOURNAL.PONE.0095628
60. Kalpathy-Cramer J, Chandra V, Da X, Ou Y, Emblem KE, Muzikansky A, et al. Phase II Study of Tivozanib, an Oral VEGFR Inhibitor, in Patients With Recurrent Glioblastoma. *J Neurooncol* (2017) 131:603–10. doi: 10.1007/S11060-016-2332-5
61. Iwamoto FM, Lamborn KR, Robins HI, Mehta MP, Chang SM, Butowski NA, et al. Phase II Trial of Pazopanib (GW786034), an Oral Multi-Targeted Angiogenesis Inhibitor, for Adults With Recurrent Glioblastoma (North American Brain Tumor Consortium Study 06-02). *Neuro Oncol* (2010) 12:855. doi: 10.1093/NEUONC/NOQ025
62. Kaley TJ, Panageas KS, Pentsova EI, Mellinghoff IK, Nolan C, Gavrilovic I, et al. Phase I Clinical Trial of Temsirolimus and Perifosine for Recurrent Glioblastoma. *Ann Clin Transl Neurol* (2020) 7:429–36. doi: 10.1002/ACN3.51009
63. Becher OJ, Millard NE, Modak S, Kushner BH, Haque S, Spasojevic I, et al. A Phase I Study of Single-Agent Perifosine for Recurrent or Refractory Pediatric CNS and Solid Tumors. *PloS One* (2017) 12:e0178593. doi: 10.1371/JOURNAL.PONE.0178593
64. Gortany NK, Panahi G, Ghafari H, Shekari M, Ghazi-Khansari M. Foretinib Induces G2/M Cell Cycle Arrest, Apoptosis, and Invasion in Human Glioblastoma Cells Through C-MET Inhibition. *Cancer Chemother Pharmacol* (2021) 87:827–42. doi: 10.1007/S00280-021-04242-0
65. Fotoohi M, Hadi N, Namazi F. Effect of Foretinib on Matrix Metalloproteinase-2 (MMP2) Expression in Glioblastoma. *Res Mol Med* (2019) 6:20–7. doi: 10.18502/RMM.V6I4.4800
66. Schiff D, Sarkaria J. Dasatinib in Recurrent Glioblastoma: Failure as a Teacher. *Neuro Oncol* (2015) 17:910–1. doi: 10.1093/NEUONC/NOV086
67. Morandi F, Amoroso L, Dondero A, Castriconi R, Parodi S, Luksch R, et al. Updated Clinical and Biological Information From the Two-Stage Phase II Study of Imatinib Mesylate in Subjects With Relapsed/Refractory Neuroblastoma. *Oncoimmunology* (2018) 7:e1468953. doi: 10.1080/2162402X.2018.1468953/SUPPL_FILE/KONI_A_1468953_SM4972.ZIP
68. Beppu K, Jaboine J, Merchant MS, Mackall CL, Thiele CJ. Effect of Imatinib Mesylate on Neuroblastoma Tumorigenesis and Vascular Endothelial Growth Factor Expression. *J Natl Cancer Inst* (2004) 96:46–55. doi: 10.1093/JNCI/DJH004
69. Raymond E, Brandes AA, Dittrich C, Fumoleau P, Coudert B, Clement PMJ, et al. Phase II Study of Imatinib in Patients With Recurrent Gliomas of Various Histologies: A European Organisation for Research and Treatment of Cancer Brain Tumor Group Study. *J Clin Oncol* (2008) 26:4659–65. doi: 10.1200/JCO.2008.16.9235
70. McLaughlin ME, Robson CD, Kieran MW, Jacks T, Pomeroy SL, Cameron S. Marked Regression of Metastatic Pilocytic Astrocytoma During Treatment With Imatinib Mesylate (STI-571, Gleevec): A Case Report and Laboratory Investigation. *J Pediatr Hematol Oncol* (2003) 25:644–8. doi: 10.1097/00043426-200308000-00012
71. Ferguson SD, Zhou S, Huse JT, de Groot JF, Xiu J, Subramaniam DS, et al. Targetable Gene Fusions Associate With the IDH Wild-Type Astrocytic Lineage in Adult Gliomas. *J Neuropathol Exp Neurol* (2018) 77:437. doi: 10.1093/JNEN/NLY022
72. Parker B, Zhang W. Fusion Genes in Solid Tumors: An Emerging Target for Cancer Diagnosis and Treatment. *Chin J Cancer* (2013) 32:594–603. doi: 10.5732/CJC.013.10178
73. Samii A, Sorokin M, Kar S, Makovskaia L, Garazha A, Hartmann C, et al. Case of Multifocal Glioblastoma With Four Fusion Transcripts of ALK, FGFR2, NTRK2, and NTRK3 Genes Stresses the Need for Tumor Tissue Multisampling for Transcriptomic Analysis. *Cold Spring Harb Mol Case Stud* (2021) 7:a006100. doi: 10.1101/MCS.A006100
74. Haas BJ, Dobin A, Li B, Stransky N, Pochet N, Regev A. Accuracy Assessment of Fusion Transcript Detection *via* Read-Mapping and *De Novo* Fusion Transcript Assembly-Based Methods. *Genome Biol* (2019) 20:213. doi: 10.1186/s13059-019-1842-9
75. Hawkins C, Walker E, Mohamed N, Zhang C, Jacob K, Shirinian M, et al. BRAF-KIAA1549 Fusion Predicts Better Clinical Outcome in Pediatric Low-

Grade Astrocytoma. *Clin Cancer Res* (2011) 17:4790–8. doi: 10.1158/1078-0432.CCR-11-0034
76. Hasselblatt M, Riesmeier B, Lechtape B, Brentrup A, Stummer W, Albert F, et al. BRAF-KIAA1549 Fusion Transcripts are Less Frequent in Pilocytic Astrocytomas Diagnosed in Adults. *Neuropathol Appl Neurobiol* (2011) 37:803–6. doi: 10.1111/J.1365-2990.2011.01193.X
77. Yuan C, Yao Q, Cheng L, Zhang C, Ma L, Guan J, et al. Prognostic Factors and Nomogram Prediction of Survival Probability in Primary Spinal Cord Astrocytoma Patients. *J Neurosurg Spine* (2021) 35:651–62. doi: 10.3171/2021.1.SPINE202017
78. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: A Summary. *Acta Neuropathol* (2016) 131:803–20. doi: 10.1007/s00401-016-1545-1
79. Palpan Flores A, Rodríguez Domínguez V, Esteban Rodriguez I, Román de Aragón M, Zamarrón Pérez Á. H3K27M-Mutant Glioma in Thoracic Spinal Cord and Conus Medullaris With Pilocytic Astrocytoma Morphology: Case Report and Review of the Literature. *Br J Neurosurg* (2021), 1–7. doi: 10.1080/02688697.2021.1988054
80. Biczok A, Strübing FL, Eder JM, Egensperger R, Schnell O, Zausinger S, et al. Molecular Diagnostics Helps to Identify Distinct Subgroups of Spinal Astrocytomas. *Acta Neuropathol Commun* (2021) 9:119. doi: 10.1186/s40478-021-01222-6
81. Zhang M, Iyer RR, Azad TD, Wang Q, Garzon-Muvdi T, Wang J, et al. Genomic Landscape of Intramedullary Spinal Cord Gliomas. *Sci Rep* (2019) 9:18722. doi: 10.1038/s41598-019-54286-9
82. Tian B, Zhang Y, Zhang J. Periostin is a New Potential Prognostic Biomarker for Glioma. *Tumor Biol* (2014) 35:5877–83. doi: 10.1007/s13277-014-1778-3
83. Mikheev AM, Mikheeva SA, Trister AD, Tokita MJ, Emerson SN, Parada CA, et al. Periostin is a Novel Therapeutic Target That Predicts and Regulates Glioma Malignancy. *Neuro Oncol* (2015) 17:372–82. doi: 10.1093/neuonc/nou161
84. Lin C, Lecca D, Yang L, Luo W, Scerba M, Tweedie D, et al. 3,6'-Dithiopomalidomide Reduces Neural Loss, Inflammation, Behavioral Deficits in Brain Injury and Microglial Activation. *Elife* (2020) 9:1–78. doi: 10.7554/ELIFE.54726
85. Jeyapalan J, Doctor G, Jones T, Alberman S, Tep A, Haria C, et al. DNA Methylation Analysis of Paediatric Low-Grade Astrocytomas Identifies a Tumour-Specific Hypomethylation Signature in Pilocytic Astrocytomas. *Acta Neuropathol Commun* (2016) 4:54. doi: 10.1186/S40478-016-0323-6
86. Debinski W, Gibo D. Fos-Related Antigen 1 (Fra-1) Pairing With and Transactivation of JunB in GBM Cells. *Cancer Biol Ther* (2011) 11:254–62. doi: 10.4161/CBT.11.2.13953
87. Zhang L, Liu H, Mu X, Cui J, Peng Z. Dysregulation of Fra1 Expression by Wnt/β-Catenin Signalling Promotes Glioma Aggressiveness Through Epithelial-Mesenchymal Transition. *Biosci Rep* (2017) 37:BSR20160643. doi: 10.1042/BSR20160643
88. Amos S, Redpath G, Dipierro C, Carpenter J, Hussaini I. Epidermal Growth Factor Receptor-Mediated Regulation of Urokinase Plasminogen Activator Expression and Glioblastoma Invasion *via* C-SRC/MAPK/AP-1 Signaling Pathways. *J Neuropathol Exp Neurol* (2010) 69:582–92. doi: 10.1097/NEN.0B013E3181E008FE
89. Shi R, Peng H, Yuan X, Zhang X, Zhang Y, Fan D, et al. Down-Regulation of C-Fos by shRNA Sensitizes Adriamycin-Resistant MCF-7/ADR Cells to Chemotherapeutic Agents *via* P-Glycoprotein Inhibition and Apoptosis Augmentation. *J Cell Biochem* (2013) 114:1890–900. doi: 10.1002/JCB.24533
90. Zhang M, Hoyle R, Ma Z, Sun B, Cai W, Cai H, et al. FOSL1 Promotes Metastasis of Head and Neck Squamous Cell Carcinoma Through Super-Enhancer-Driven Transcription Program. *Mol Ther* (2021) 29:2583–600. doi: 10.1016/J.YMTHE.2021.03.024
91. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* (2013) 29:15–21. doi: 10.1093/bioinformatics/bts635
92. Bartlett P, Shawe-taylor J. Generalization Performance of Support Vector Machines and Other Pattern Classifiers, in: *Adv Kernel Methods Support Vector Learn* (1999). Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.42.6950 (Accessed June 26, 2019).
93. Borisov N, Buzdin A. New Paradigm of Machine Learning (ML) in Personalized Oncology: Data Trimming for Squeezing More Biomarkers From Clinical Datasets. *Front Oncol* (2019) 9:658. doi: 10.3389/fonc.2019.00658
94. Borisov N, Sorokin M, Tkachev V, Garazha A, Buzdin A. Cancer Gene Expression Profiles Associated With Clinical Outcomes to Chemotherapy Treatments. *BMC Med Genomics* (2020) 13:111. doi: 10.1186/s12920-020-00759-0
95. Zhao Y, Wong L, Goh Bin WW. How to do Quantile Normalization Correctly for Gene Expression Data Analyses. *Sci Rep* (2020) 10:1–11. doi: 10.1038/s41598-020-72664-6
96. Love MI, Huber W, Anders S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data With Deseq2. *Genome Biol* (2014) 15:550. doi: 10.1186/s13059-014-0550-8
97. Blighe K, Rana S, Lewis M. *EnhancedVolcano: Publication-Ready Volcano Plots With Enhanced Colouring and Labeling* (2018). Available at: https://github.com/kevinblighe/EnhancedVolcano.
98. Feng J, Meyer CA, Wang Q, Liu JS, Shirley Liu X, Zhang Y. GFOLD: A Generalized Fold Change for Ranking Differentially Expressed Genes From RNA-Seq Data. *Bioinformatics* (2012) 28:2782–8. doi: 10.1093/BIOINFORMATICS/BTS515
99. Buzdin AA, Zhavoronkov AA, Korzinkin MB, Venkova LS, Zenin AA, Smirnov PY, et al. Oncofinder, A New Method for the Analysis of Intracellular Signaling Pathway Activation Using Transcriptomic Data. *Front Genet* (2014) 5:55. doi: 10.3389/fgene.2014.00055

**Conflict of Interest:** Author MR was employed by Omicsway Corp and AB was employed by Omicsway Corp. and Oncobox Ltd.

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 28 March 2022
doi: 10.3389/fonc.2022.861284

# Investigating the Prognostic Significance of Pyroptosis-Related Genes in Gastric Cancer and Their Impact on Cells' Biological Functions

*Jie Yin*[1†], *Gang Che*[2†], *Wankun Wang*[2], *Shitu Chen*[2*†] *and Jian Liu*[2*†]

[1] *Department of Radiation Oncology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China,*
[2] *Department of Surgical Oncology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China*

**Objective:** To probe into the role of pyroptosis-related genes in gastric cancer.

**Methods:** To establish pyroptosis-related genes, observe their expression in gastric cancer, and analyze the prognosis of pyroptosis-related genes in gastric cancer by single-factor COX, which showed that only GSDME had prognostic significance in gastric cancer. The mRNA expression profiles and lncRNA expression profiles of gastric cancer downloaded from the Cancer Genome Atlas were combined for weighted gene regulatory network analysis, after which the lncRNA nodes of the module to which GSDME belongs were extracted to obtain the lncRNAs–GSDME interactions, which were visualized with Cytoscape network plots. Finally, the effects of GSDME on the proliferation, migration, and apoptosis of gastric cancer cells were observed with CCK8, and flow cytometry.

**Results:** Our results show that only GSDME has prognostic significance in gastric cancer, and show that it has an important role in a variety of tumors. In addition, our results show that 16 lncRNAs have a significant interaction with GSDME. Finally, the experimental analysis showed that knocking down the expression level of GSDME could affect the growth as well as apoptosis of gastric cancer cells.

**Conclusion:** The significant prognostic significance of GSDME in gastric cancer and the fact that affecting GSDME expression inhibits gastric cancer cell growth suggest that GSDME can be used as a predictive biomarker.

**Keywords: gastric cancer, pyroptosis, GSDME, cell proliferation, pyroptosis-related genes**

OPEN ACCESS

***Edited by:***
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

***Reviewed by:***
*Xuefeng Xia,*
*Nanjing Drum Tower Hospital, China*
*Fei Liu,*
*The Fifth Affiliated Hospital of Sun Yat-Sen University, China*

***Correspondence:**
*Jian Liu*
*liujianzy@zju.edu.cn*
*Shitu Chen*
*11618339@zju.edu.cn*

*[†]These authors have contributed equally to this work*

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

***Received:*** *24 January 2022*
***Accepted:*** *01 March 2022*
***Published:*** *28 March 2022*

***Citation:***
*Yin J, Che G, Wang W, Chen S and Liu J (2022) Investigating the Prognostic Significance of Pyroptosis-Related Genes in Gastric Cancer and Their Impact on Cells' Biological Functions. Front. Oncol. 12:861284. doi: 10.3389/fonc.2022.861284*

## INTRODUCTION

Stomach cancer has caused many deaths. Gastric cancer causes more than 720,000 deaths per year (1, 2). More than 90% of gastric cancers are adenocarcinomas, which can be subdivided into cardia and non-cardia tumors, respectively, according to the site of the tumor (3). Cardia cancer occurs in the region adjacent to the esophagogastric junction and therefore has the same epidemiological features as esophageal adenocarcinomas. Non-cardia carcinomas, also known as distal gastric carcinomas, often occur in the lower part of the stomach (4). Due to the lack of typical early signs, 40% of them have metastatic disease at the time of diagnosis (5). When patients are found to have metastases, they are already in the middle to late stages of gastric cancer, and Patients with advanced gastric cancer are bound to die, ranging from approximately 10% to 30% at five years (6).

Pyroptosis amplifies or maintains inflammation by releasing pro-inflammatory cellular contents that tightly control the inflammatory response and orchestrate antimicrobial host defenses. It is an innate immune effector mechanism for fighting intracellular pathogens (7, 8). Heat illnesses are linked to many diseases, with different tissue and genetic backgrounds having different effects on cancer. On the one hand, pyroptosis can inhibit tumor development; on the other hand, as pro-inflammatory death, pyroptosis can form a microenvironment suitable for tumor cell growth, thus promoting tumor growth (9, 10). With the progress of research, the role of pyroptosis in tumors has become more prominent, as it may affect all level of carcinogenesis. Therefore, this study investigated the potential impact of genes related to pyrogenesis in gastric cancer.

# METHODS

## Data Acquisition

There were 375 gastric cancer tissue samples and 32 corresponding normal tissue samples. We also downloaded gastric cancer mRNA expression profiles and lncRNA expression profiles from the TCGA dataset for subsequent analysis.

## Pyroptosis-Related Gene Acquisition

The human genes with BP_PYROPTOSIS in their annotation results were NLRC4, NLRP1, AIM2, TREM2, GSDMC, GSDMB, NLRP9, NLRP6, ELANE, NAIP, CASP8, CASP4 GZBP1, APIP, GZMB, GZMA, CASP1, DHX9, GSDMA, GSDME, and GSDMD. Thus, the set of functional genes related to focal death was constructed.

## Expression Analysis

Statistical significance was considered achieved when $p < 0.05$.

## Pan-Cancer Analysis

The correlation between MSI and GSDME expression levels in pan-cancer was analyzed using the R language with the fmsb package. It was considered statistically significant when $p < 0.05$. **Table 1** displays 33 tumor abbreviations in the TCGA database.

## GSEA Analysis

To understand the impact of various biological functional gene sets of GSDME in gastric cancer, GSEA was used for enrichment analysis. GSDME was classified into high and low expression parts. Subsequently, GSEA V3.0 software was used to analyze the enrichment results of genes. A nominal $P < 0.05$ and a false discovery rate (FDR) < 25% were selected as cut-off criteria. We selected the top five-ranked analysis results.

## Weighted Gene Regulatory Network Analysis

A soft threshold for network construction is first selected. Second, the adjacency matrix is transformed into a topology matrix. The genes are clustered based on the topology matrix using the average chain hierarchy clustering method. We set the minimum base number for each gene network module at 30. After determining the gene modules by the dynamic cut method, the feature vectors of each module are calculated in turn. Next, the modules are clustered, and the closer modules are merged into new modules.

## Sample Collection

The gastric cancer tissues and the corresponding adjacent cancer tissues were collected from 5 GC patients, and the liquid nitrogen was preserved. Two cases were male and three cases were female. Sample collection excluded patients with major diseases. The study was informed by the patient and obtained by the Ethics committee of our hospital (ethical batch number: IIT20210716A).

## Cell Source and Culture

Human gastric cancer cell lines (HS-746T, HT-X2557, MKN74, HTX2401C) and human gastric mucosal cells (GES-1, HT-X1964) were got from Shenzhen Haodi Huatuo Biotechnology Co. These cells were cultured with Dulbecco's modified Eagle medium. Cells were supplemented with 10% fetal bovine serum and penicillin-streptomycin solution at 37°C and 5% CO2. After reaching 80–90% confluence with wall growth, cells were digested with 25% trypsin for passaging.

## Cell Transfection

Cells were transferred to six-well plates at 2 x 10 five cells/well and incubated overnight at 37°C in an incubator. Si-NC and Si-GSDME were transfected with HS-746T and MKN74 cells using a Lipofectamine 2000 kit according to the kit instructions into DMEM containing 10% FBS and incubated for 6 h after transfection.

## Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR) Assay

Grind the gastric cancer tissue to prepare a cell suspension. Total RNA was extracted from cells using the TRIzol Kit (Invitrogen, USA).Then, $1.7 \leq OD260/OD280 \leq 2.1$ RNA samples could be used for analysis. Subsequently, the total RNA was reverse transcribed using a reverse transcription kit. The volume was adjusted by 20 μL of total reaction volume and 10 μL of SYBR Premix Ex Taq II (2X), 2 μL of cDNA, 0.8 μL of upstream and downstream primers, and the addition of sterile purified water. The amplification conditions were: 95°C for 30 s, 95°C for 5 s, 60°C for 30 s, and 40 cycles. GAPDH was used as an internal reference for GSDME. All primers were purchased from Shanghai Gene Chemistry Co. The forward primer for GSDME was 5′-TGCC TACGGTGTCATTGAGTT-3′; its reverse primer was TCTGGCATGTCTATGAATGCAAA-3′. The forward primer for GAPDH was 5′-GAGTCAACGGATTTGGTCGT-3′, and its reverse primer was 5′-GACAAGCTTCCCGTTCTCAG-3′. Now, $2^{-\Delta\Delta}$Ct was used to analyze the data, and the experiment was repeated three times.

## Cell Counting Kit-8 (CCK8) Assay

Cell proliferation assays were performed using CCK-8 (Beyotime Biotechnology, China) according to the kit's instructions: Cells were transfected for 24 h and incubated in 96-well plates at $2.5\times10^{3}$ cells/well. After 24, 48, and 72 h of incubation, 10 μL of CCK-8 solution was added to each well and incubated at room temperature for 2 h. Subsequently, the optical density of each

**TABLE 1** | TCGA database tumor abbreviations.

| Abbr | Full name |
|---|---|
| ACC | Adrenocortical carcinoma |
| BLCA | Bladder urothelial carcinoma |
| BRCA | Breast invasive carcinoma |
| CESC | Cervical squamous cell carcinoma and endocervical adenocarcinoma |
| CHOL | Cholangiocarcinoma |
| COAD | Colon adenocarcinoma |
| COADREAD | Colon adenocarcinoma/Rectum adenocarcinoma/Esophageal carcinoma |
| DLBC | Lymphoid neoplasm diffuse large B-cell lymphoma |
| ESCA | Esophageal carcinoma |
| FPPP | FFPE Pilot Phase II |
| GBM | Glioblastoma multiforme |
| GBMLGG | Glioma |
| HNSC | Head and neck squamous cell carcinoma |
| KICH | Kidney chromophobe |
| KIPAN | Pan-kidney cohort (KICH+KIRC+KIRP) |
| KIRC | Kidney renal clear cell carcinoma |
| KIRP | Kidney renal papillary cell carcinoma |
| LAML | Acute myeloid leukemia |
| LGG | Brain lower grade glioma |
| LIHC | Liver hepatocellular carcinoma |
| LUAD | Lung adenocarcinoma |
| LUSC | Lung squamous cell carcinoma |
| MESO | Mesothelioma |
| OV | Ovarian serous cystadenocarcinoma |
| PAAD | Pancreatic adenocarcinoma |
| PCPG | Pheochromocytoma and paraganglioma |
| PRAD | Prostate adenocarcinoma |
| READ | Rectum adenocarcinoma |
| SARC | Sarcoma |
| SKCM | Skin cutaneous melanoma |
| STAD | Stomach adenocarcinoma |
| STES | Stomach and esophageal carcinoma |
| TGCT | Testicular germ cell tumors |
| THCA | Thyroid carcinoma |
| THYM | Thymoma |
| UCEC | Uterine corpus endometrial carcinoma |
| UCS | Uterine carcinosarcoma |
| UVM | Uveal melanoma |

well at 450 nm was detected using an enzyme marker. The experiment was repeated three times.

## Flow Cytometry

The Annexin V-FITC Apoptosis Detection Kit (Beyotime, China) measured the apoptosis in gastric cancer cells treated as indicated. The experiments were repeated three times.

## Statistical Analysis

The data obtained were plotted and analyzed using GraphPad 8. Comparisons between groups were made using independent t-tests, and comparisons between multiple parts were made using one-way ANOVA.

# RESULTS

## Expression Levels of Pyroptosis-Related Genes in Gastric Cancer

We performed bulk expression analysis obtained and observed the expression of pyroptosis-related genes in gastric cancer (n = 375) and paracancerous tissue (n = 32). Most of the pyroptosis-related genes were upregulated in gastric cancer. See **Figure 1**.

## Prognostic Analysis of Pyroptosis-Related Genes

The results showed that only GSDME had prognostic significance in gastric cancer. The KM survival curve analysis also showed that the high-level survival of GSDME was lower than the low-level survival of GSDME. See **Figure 2**.

## Role of GSDME in Multiple Tumors

To continue to probe the role of GSDME in tumors, we also observed the expression of GSDME in multiple tumors. GSDME was highly expressed in CHOL, ESCA, GBM and so on, and low expressed in BLCA, BRCA, KICH, PRAD, THCA, and UCEC. A univariate cox analysis revealed the prognostic significance of GSDME in BLCA, KIRP, LIHC, KIRC, HNSC, LGG, KICH, and ACC. An immune checkpoint correlation analysis demonstrated that GSDME was closely associated with immune checkpoints TNFSF9, TNFSF15, TNFSF18, TNFSF4, TNFRSF25, TNFRSF4, TNFRSF8, LGALS9, NRP1, CD276, CD40, and CD200 in gastric

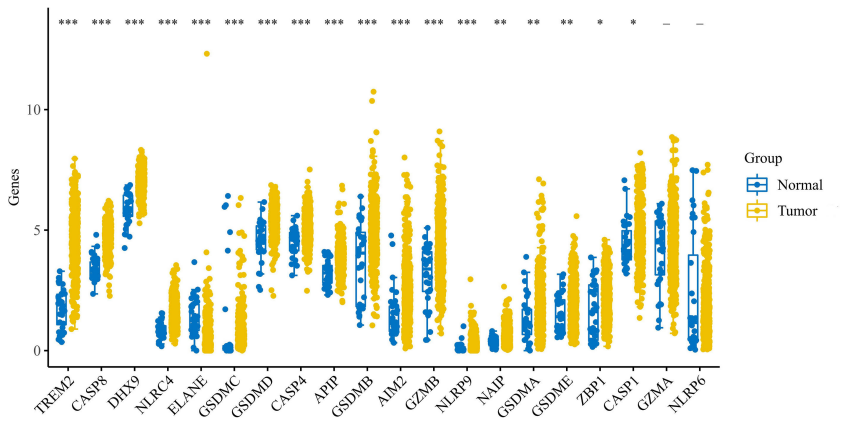



**FIGURE 1** | Expression of pyroptosis-related genes in gastric cancer. Expression of pyroptosis-related genes in gastric cancer. * indicates $p < 0.05$ for comparison between gastric cancer and paraneoplasia; ** indicates $p < 0.01$ for comparison between gastric cancer and paraneoplasia; *** indicates $p < 0.001$ for comparison between gastric cancer and paraneoplasia.

cancer. A neoantigen analysis depicted that GSDME was only correlated in KIRP. An immunomutational load correlation indicated that GSDME was significantly correlated with immunomutational load in STAD, MESO, PAAD, TGCT, BLCA, and LUSC. Microsatellite correlation analysis showed that GSDME was significantly correlated with microsatellite instability in STAD and SARC. See **Figure 3**.

## GSEA Analysis

We divided the samples into 2parts according to the median expression levels of GSDME in gastric cancer and then analyzed the biological processes or signaling pathways involved in high GSDME expression in gastric cancer. We only screened the top five-ranked signaling pathways. The analysis revealed that the high expression of GSDME in gastric cancer was found in ECM_RECEPTOR_INTERACTION, HYPERTROPHIC_CARDIOMYOPATHY_HCM, DILATED_CARDIOMYOPATHY, GLYCOSAMINOGLYCAN_BIOSYNTHESIS_CHONDROITIN_SULFATE, and FOCAL_ADHESION, which have important roles. See **Figure 4**.

## GSDME and the Stemness Score of Gastric Cancer Cells

We then divided the group into high and low expression groups based on the median GSDME expression. The mRNAsi scores between the two groups and the normal group were observed. The results showed that the mRNAsi score was significantly higher in the GSDME high expression group compared to the normal group. Surprisingly, the mRNAsi score was again significantly higher in the GSDME low expression group compared to the high expression. This may be a shift in the progression of gastric cancer, but the exact reason for this remains to be explored. See **Figure 5**.

## WGCNA Analysis

We combined the mRNA expression profiles obtained from the TCGA database with lncRNA expression profiles and performed

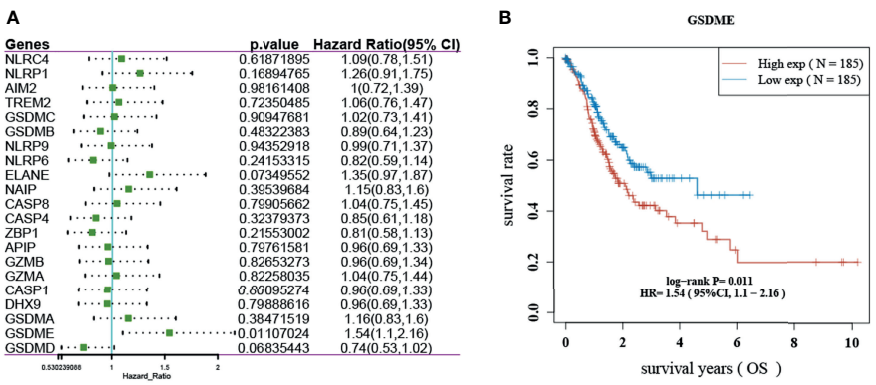



**FIGURE 2** | Prognosis of pyroptosis-related genes in gastric cancer. **(A)** Forest plot showing the OS prognosis of pyroptosis-related genes in gastric cancer. **(B)** KM survival curve of GSDME in gastric cancer.

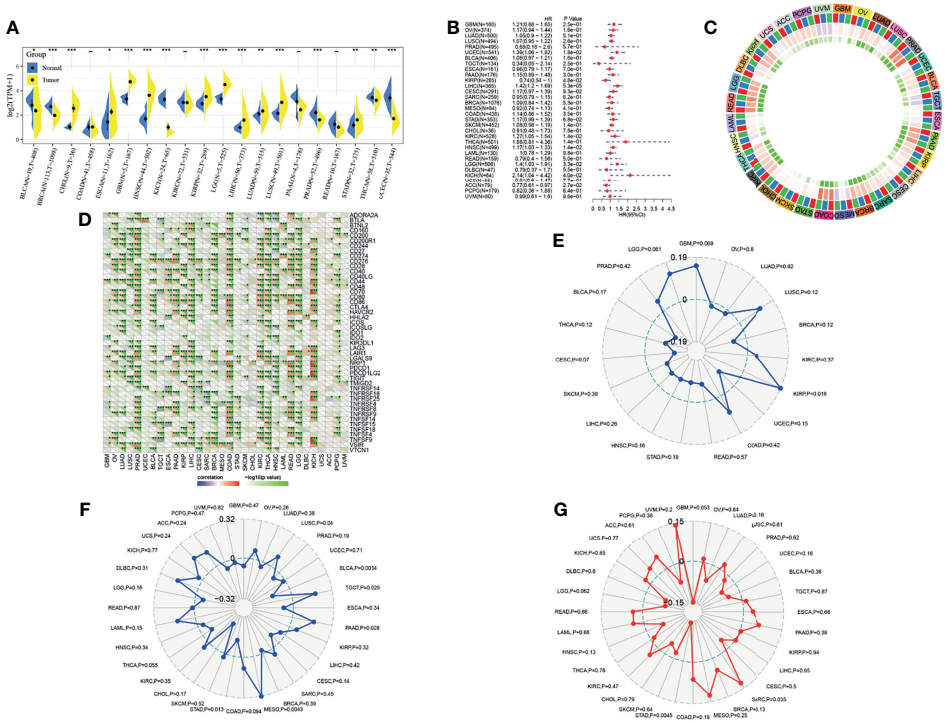



**FIGURE 3** | Analysis of GSDME in multiple tumors. **(A)** Expression of GSDME in multiple tumors (* indicates $p < 0.05$ for comparison between cancer and paracancer; ** indicates $p < 0.01$ for comparison between cancer and paracancer; *** indicates $p < 0.001$ for comparison between cancer and paracancer). **(B)** Prognostic analysis of GSDME in multiple tumors. **(C)** Correlation of GSDME with immune microenvironment scores in multiple tumors (blue, minimum color; red, maximum value color; green, minimum p-value color). **(D)** Correlation of GSDME with immune checkpoints in multiple tumors. **(E)** Correlation of GSDME with immune neoantigens in multiple tumors. **(F)** Correlation of GSDME with immune mutational load in multiple tumors. **(G)** correlation of GSDME with microsatellites in multiple tumors.

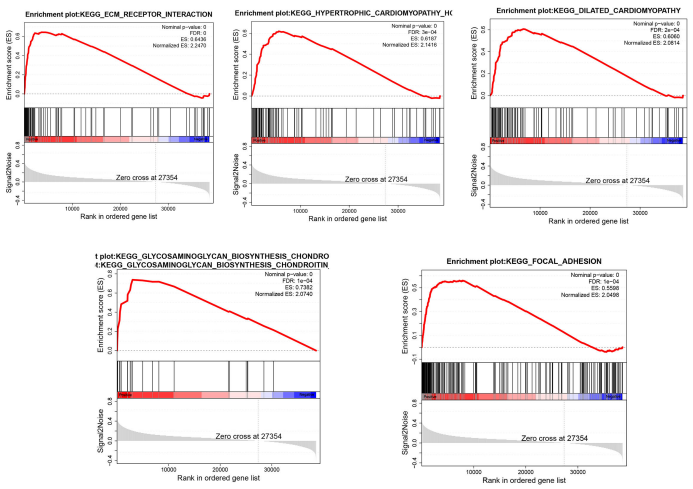



**FIGURE 4** | Biological processes involved in the high expression of GSDME in gastric cancer.

WGCNA analysis. We found the module darkorange2 to which GSDME belongs. Next, we extracted the lncRNA nodes belonging to this module and obtained 16 lncRNAs with interactions with GSDME, namely, LINC01106, LINC01547, LINC00265, LINC00910, LINC01560, LINC01123, LINC00654, LINC01521, LINC00998, LINC00680, LINC00339, LINC00630, LINC00205, LINC01057, LINC00539, and LINC00863. Finally, we used Cytoscape (https://cytoscape.org/) for visualization. See **Figure 6**.

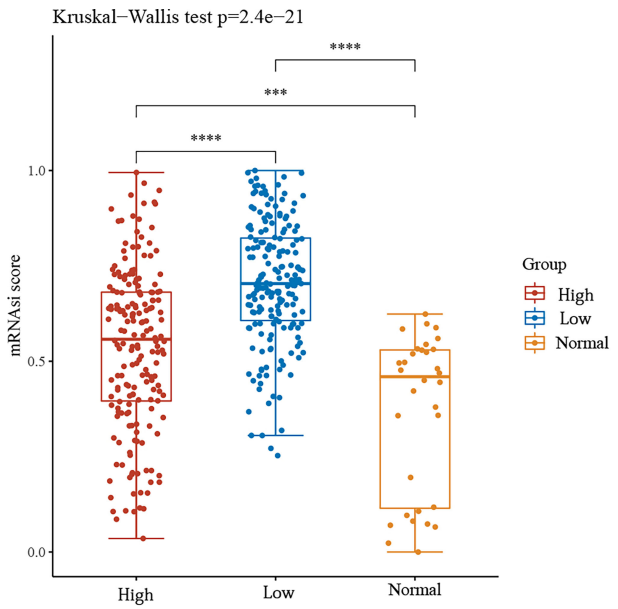



**FIGURE 5** | mRNAsi scores between high and low GSDME expression groups and normal groups. ***P < 0.001, ****P < 0.0001.

## Experimental Observation of the Effect of GSDME on Gastric Cancer Cells

Finally, we performed experiments to observe whether GSDME influences the biological behavior of gastric cancer cells. We first observed the expression of GSDME in gastric cancer tissue, gastric cancer cell lines HS-746T and MKN74 and gastric mucosal cells GES-1. The expression level of GSDME was significantly higher in gastric cancer tissue, HS-746T and MKN74 than in GES-1. We then further observed the effect of GSDME on gastric cancer cells by knocking down the expression level of GSDME to observe its effect on gastric cancer cells. The expression level of GSDME was reduced, the growth ability of gastric cancer cells were significantly slowed, and the apoptosis of cells was markedly increased. See **Figure 7**.

## DISCUSSION

GSDME related to the induction of secondary necrosis/scarring death. GSDME expression may also control the transition, such that in response to apoptotic stimuli, cells lacking GSDME expression undergo apoptosis without progressing. GSDME also plays distinct roles in a variety of tumors. In the present study, the expression of GSDME also differed in assorted tumors. Moreover, GSDME was closely associated with immune checkpoints in most tumors, including gastric cancer. GSDME expression enhanced the phagocytosis of tumor cells by tumor-associated macrophages and enhanced the number and function of tumor-infiltrating natural killer cells and CD8+ T cells, It can be cleaved by activated caspase-3 to produce its N-terminal fragment (GSDME-NT) (11, 12). It has also been indicated that targeted drugs induce GSDME-mediated cellular scorching in melanoma, linking the tumor immune microenvironment to

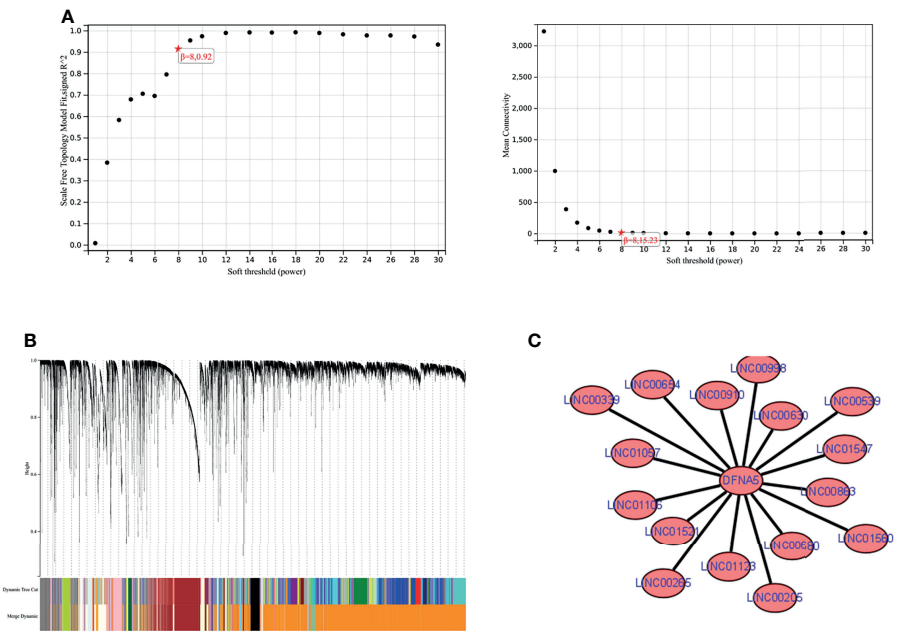



**FIGURE 6** | WGCNA analysis. **(A)** soft threshold determination. **(B)** gene clustering map. **(C)** lncRNAs-GSDME interaction network map.

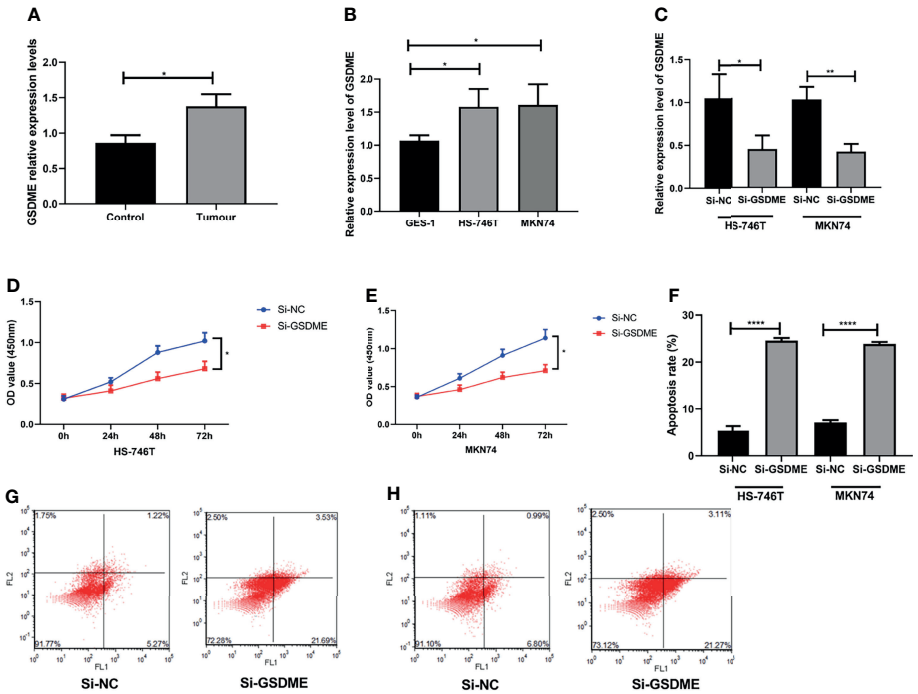



**FIGURE 7** | Effect of the knockdown of GSDME expression on gastric cancer cells. **(A)**. Relative expression levels of GSDME in gastric cancer and adjacent to paracancerous tissues. **(B)** Expression of GSDME in human gastric cancer cell lines HS-746T and MKN74 and gastric mucosal cells GES-1. **(C)** Transfection efficiency. **(D)** Effect of reduced GSDME levels on the proliferation of HS-746T cells. **(E)** Effect of reduced GSDME levels on the proliferation of MKN74 cells. **(F)** Promotion of apoptosis after GSDME levels were reduced. **(G)** Flow apoptosis diagram based on HS-746T cells. **(H)** Flow apoptosis diagram based on MKN74 cells. *P < 0.05, **P < 0.01, ****P < 0.0001.

T cell-mediated anti-tumor immunity (13). It is suggested that GSDME may related to cancer therapy and antitumor immunity, but the potential effect of GSDME on immune cells in gastric cancer needs to be investigated in more depth.

In rectal and esophageal squamous cell carcinoma also significantly higher than normal tissue and promote tumor progression (14, 15). In the present study, only GSDME among several pyroptosis genes had prognostic significance in gastric cancer. The KM curve revealed that the high expression of GSDME was associated with a poor prognosis for patients. The stemness score also showed a significantly higher mRNAsi score in the GSDME high expression group compared to the normal group. Interestingly, the mRNAsi score was again significantly higher in the GSDME low expression group compared to the high expression. However, the exact mechanism of development remains to be investigated.

Tumor development is inseparable from the proliferation and migration of cancer cells. Conversely, the inhibition of cancer cell proliferation may stunt tumor growth. In breast cancer, cell survival is strongly correlated with GSDMB expression. In addition, reports have demonstrated that GSDMB overexpression reduces cell viability (16). Gene deletion of GSDME promoted drug resistance (17). Another study reported that the knockdown of GSDME shifted loplatin-induced cell death from cell scorching to apoptosis, but did not affect the growth and tumor formation of colon cancer cells treated with loplatin (18). Although GSDME has been less studied in gastric cancer, one study has shown that GSDME converts chemotherapy-induced caspase-3-dependent apoptosis to cellular scorching in gastric cancer cells (19).Our *in vitro* experiments also show that changes in GSDME expression may affect the biological behavior of gastric cancer cells, which may provide new ideas for the treatment of gastric cancer.

LncRNA is a transcript with more than 200 nucleotides and no protein-coding potential. However, it is frequently dysregulated in cancer (20, 21). In therapeutic agents, LncRNAs can exert clinical therapeutic effects on tumors by inhibiting the transcription of mRNAs and blocking their function in combination with proteins (22, 23). In the present study, we found that 16 lncRNAs had reciprocal relationships with GSDME by WGCNA analysis, namely, LINC01106, LINC01547, LINC00265, LINC00910, LINC01560, LINC01123, LINC00654, LINC01521 LINC00998, LINC00680, LINC00339, LINC00630, LINC00205, LINC01057, LINC00539, and LINC00863. However, no clinical studies indicate that these lncRNAs have a regulatory relationship with GSDME. Correspondingly, more in-depth studies are needed.

In conclusion, GSDME is highly expressed in gastric cancer. The knockdown of GSDME expression can inhibit the

growth of gastric cancer cells. GSDME, then, has prognostic significance in gastric cancer and can be used as a predictive biomarker.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## AUTHOR CONTRIBUTIONS

These authors contributed equally: JY and GC. All authors contributed to the article and approved the submitted version.

## FUNDING

The Key Research and Development Project of Zhejiang Province (2019C03071).

## REFERENCES

1. Noto JM, Peek RM Jr. The Gastric Microbiome, its Interaction With Helicobacter Pylori, and Its Potential Role in the Progression to Stomach Cancer. *PLoS Pathog* (2017) 13. doi: 10.1371/journal.ppat.1006573
2. Poorolajal J, Moradi L, Mohammadi Y, Cheraghi Z, Gohari-Ensaf F. Risk Factors for Stomach Cancer: A Systematic Review and Meta-Analysis. *Epidemiol Health* (2020) 42. doi: 10.4178/epih.e2020004
3. Etemadi A, Safiri S, Sepanlou SG, Ikuta K, Bisignano C, Shakeri R, et al. The Global, Regional, and National Burden of Stomach Cancer in 195 Countries, 1990–2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* (2020) 5:42–54. doi: 10.1016/S2468-1253(19)30328-0
4. Rawla P, Barsouk A. Epidemiology of Gastric Cancer: Global Trends, Risk Factors and Prevention. *Prz Gastroenterol* (2019) 14:26. doi: 10.5114/pg.2018.80001
5. Koemans W, van der Kaaij R, Boot H, Buffart T, Veenhof A, Hartemink K, et al. Cytoreductive Surgery and Hyperthermic Intraperitoneal Chemotherapy Versus Palliative Systemic Chemotherapy in Stomach Cancer Patients With Peritoneal Dissemination, the Study Protocol of a Multicentre Randomised Controlled Trial (PERISCOPE Ii). *BMC Cancer* (2019) 19:1–8. doi: 10.1186/s12885-019-5640-2
6. Cao F, Hu Y, Chen Z, Han W, Lu W, Xu J, et al. Circulating Long Noncoding RNAs as Potential Biomarkers for Stomach Cancer: A Systematic Review and Meta-Analysis. *World J Surg Oncol* (2021) 19:1–13. doi: 10.1186/s12957-021-02194-6
7. Wang Y-Y, Liu X-L, Zhao R. Induction of Pyroptosis and its Implications in Cancer Management. *Front Oncol* (2019) 9:971. doi: 10.3389/fonc.2019.00971
8. Zeng C, Wang R, Tan H. Role of Pyroptosis in Cardiovascular Diseases and its Therapeutic Implications. *Int J Biol Sci* (2019) 15:1345. doi: 10.7150/ijbs.33568
9. 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
10. 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) 10:1–13. doi: 10.1038/s41419-019-1883-8
11. Zhang C-C, Li C-G, Wang Y-F, Xu L-H, He X-H, Zeng Q-Z, et al. Chemotherapeutic Paclitaxel and Cisplatin Differentially Induce Pyroptosis in A549 Lung Cancer Cells *via* Caspase-3/GSDME Activation. *Apoptosis* (2019) 24:312–25. doi: 10.1007/s10495-019-01515-1
12. 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) 579:415–20. doi: 10.1038/s41586-020-2071-9
13. Erkes DA, Cai W, Sanchez IM, Purwin TJ, Rogers C, Field CO, et al. Mutant BRAF and MEK Inhibitors Regulate the Tumor Immune Microenvironment *via* Pyroptosis. *Cancer Discov* (2020) 10:254–69. doi: 10.1158/2159-8290.CD-19-0672
14. Tan G, Huang C, Chen J, Zhi FJ. HMGB1 Released From GSDME-Mediated Pyroptotic Epithelial Cells Participates in the Tumorigenesis of Colitis-Associated Colorectal Cancer Through the ERK1/2 Pathway. *J Hematol Oncol Cancer Discov* (2020) 13:1–11. doi: 10.1186/s13045-020-00985-0
15. Wu M, Wang Y, Yang D, Gong Y, Rao F, Liu R, et al. A PLK1 Kinase Inhibitor Enhances the Chemosensitivity of Cisplatin by Inducing Pyroptosis in Oesophageal Squamous Cell Carcinoma. *EBioMedicine* (2019) 41:244–55. doi: 10.1016/j.ebiom.2019.02.012
16. Zhang Z, Zhang H, Li D, Zhou X, Qin Q, Zhang Q. Caspase-3-Mediated GSDME Induced Pyroptosis in Breast Cancer Cells Through the ROS/JNK Signalling Pathway. *J Cell Mol Med* (2021) 25:8159–68. doi: 10.1111/jcmm.16574
17. Lu H, Zhang S, Wu J, Chen M, Cai M-C, Fu Y, et al. Molecular Targeted Therapies Elicit Concurrent Apoptotic and GSDME-Dependent Pyroptotic Tumor Cell Death. *Clin Cancer Res* (2018) 24:6066–77. doi: 10.1158/1078-0432.CCR-18-1478
18. Yu J, Li S, Qi J, Chen Z, Wu Y, Guo J, et al. Cleavage GSDME by caspase-3 Determines Lobaplatin-induced Pyroptosis Colon Cancer Cells. *Cell Death Dis* (2019) 10:1–20. doi: 10.1038/s41419-019-1441-4
19. Wang Y, Yin B, Li D, Wang G, Han X, Sun XJB. GSDME Mediates Caspase-3-Dependent Pyroptosis in Gastric Cancer. *Clin Cancer Res* (2018) 495:1418–25. doi: 10.1016/j.bbrc.2017.11.156
20. Liang Y, Song X, Li Y, Chen B, Zhao W, Wang L, et al. LncRNA BCRT1 Promotes Breast Cancer Progression by Targeting miR-1303/PTBP3 Axis. *Mol Cancer* (2020) 19(1):85. doi: 10.1186/s12943-020-01206-5
21. Wu P, Mo Y, Peng M, Tang T, Zhong Y, Deng X, et al. Emerging Role of Tumor-Related Functional Peptides Encoded by lncRNA and circRNA. *Mol Cancer* (2020) 19:1–14. doi: 10.1186/s12943-020-1147-3
22. Jiang M-C, Ni J-J, Cui W-Y, Wang B-Y, Zhuo W. Emerging Roles of lncRNA in Cancer and Therapeutic Opportunities. *Am J Cancer Res* (2019) 9:1354.
23. He S, Yang S, Zhang Y, Li X, Gao D, Zhong Y, et al. LncRNA ODIR1 Inhibits Osteogenic Differentiation of hUC-MSCs Through the FBXO25/H2BK120ub/H3K4me3/OSX Axis. *Cell Death Dis* (2019) 10:1–16. doi: 10.1038/s41419-019-2148-2
24. Mai FY, He P, Ye JZ, Xu LH, Ouyang DY, Li CG, et al. Caspase-3-Mediated GSDME Activation Contributes to Cisplatin-and Doxorubicin-Induced Secondary Necrosis in Mouse Macrophages. (2019) 52:e12663. doi: 10.1111/cpr.12663

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 31 March 2022
doi: 10.3389/fonc.2022.872999

# The Differential Expression of CD47 may be Related to the Pathogenesis From Myelodysplastic Syndromes to Acute Myeloid Leukemia

*Xiao Yan[1], Binbin Lai[1], Xuyan Zhou[2], Shujun Yang[3], Qunfang Ge[1], Miao Zhou[1], Cong Shi[4], Zhijuan Xu[1] and Guifang Ouyang[1]**

[1] Haematology Department of Ningbo First Hospital, Ningbo Clinical Research Center for Hematologic Malignancies, Ningbo, China, [2] Haematology Department, The Affiliated Hospital of Medical School of Ningbo University, Ningbo, China, [3] Medical Research Center of Ningbo First Hospital, Ningbo, China, [4] Stem Cell Transplantation Laboratory of Ningbo First Hospital, Institute of Hematology of Ningbo First Hospital, Ningbo, China

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Zhijian Xiao,*
*Chinese Academy of Medical Sciences and Peking Union Medical College, China*
*Na Li,*
*Peter MacCallum Cancer Centre, Australia*
*Di Yang,*
*Huazhong University of Science and Technology, China*

***Correspondence:**
*Guifang Ouyang*
*oygf@nbdyyy.com*

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

**Received:** *10 February 2022*
**Accepted:** *07 March 2022*
**Published:** *31 March 2022*

**Citation:**
*Yan X, Lai B, Zhou X, Yang S, Ge Q, Zhou M, Shi C, Xu Z and Ouyang G (2022) The Differential Expression of CD47 may be Related to the Pathogenesis From Myelodysplastic Syndromes to Acute Myeloid Leukemia. Front. Oncol. 12:872999. doi: 10.3389/fonc.2022.872999*

Myelodysplastic syndrome (MDS) can lead to the development of peripheral blood cytopenia and abnormal cell morphology. MDS has the potential to evolve into AML and can lead to reduced survival. CD47, a member of the immunoglobulin family, is one molecule that is overexpressed in a variety of cancer cells and is associated with clinical features and poor prognosis in a variety of malignancies. In this study, we analyzed the expression and function of CD47 in MDS and AML, and further analyzed its role in other tumors. Our analysis revealed significantly low CD47 expression in MDS and significantly high expression in AML. Further analysis of the function or pathway of CD47 from different perspectives identified a relationship to the immune response, cell growth, and other related functions or pathways. The relationship between CD47 and other tumors was analyzed from four aspects: DNA methyltransferase, TMB, MSI, and tumor cell stemness. Changes in gene expression levels have a known association with aberrant DNA methylation, and this methylation is the main mechanism of tumor suppressor gene silencing and clonal variation during the evolution of MDS to AML. Taken together, our findings support the hypothesis that the differential expression of CD47 might be related to the transformation of MDS to AML.

**Keywords: mds, AML, cd47, DNA methylation, expression**

## INTRODUCTION

Myelodysplastic syndromes (MDS) belong to clonal hematopoietic stem cell disorders that present with ineffective hematopoiesis, including peripheral cytopenia and abnormal cell morphologies (1). The World Health Organization (WHO) subdivides MDS patients' risk levels from low to high according to the molecular and morphological characteristics, the number of primitive cells in the bone marrow, and the degree of peripheral cytopenia (2). Age differences are evident in the prevalence of MDS, as MDS is uncommon in children or adolescents. By contrast, a progressive increase occurs in its incidence in people between the ages of 40 and 80 years (3). Another hematopoietic disease involving myeloid or multipotent progenitor cells is acute myeloid leukemia

(AML), a malignant clonal disease with molecular heterogeneity. The various clonal disorders associated with AML are caused by a failure in differentiation and an uncontrolled proliferation of hematopoietic progenitor cells, leading to the accumulation of many different cytogenetic disorders (4).

Both MDS and AML are clinically genetically heterogeneous myeloid stem cell disorders, and patients with MDS are at a high risk of progression to AML (5). Patients whose MDS progresses to AML have a lower response to standard therapy compared to patients with new-onset AML (6). In recent years, many new drugs have emerged to treat MDS, and allogeneic HSCT provides possibility to cure MDS (7). However, the stringent conditions for transplantation and the conversion to AML experienced by some patients have led to poor overall survival (OS) in MDS (8). This has prompted a search for the genetic basis of this conversion. One potential molecule that may be involved is Cluster of differentiation 47 (CD47).

CD47 is a cell surface glycoprotein molecule belonging to the immunoglobulin superfamily. It is a widely expressed transmembrane protein in human cells and is overexpressed on the surface of many cancer cells. Its binding to signal-regulated protein α (SIRPα) signals cancer cells to escape from macrophage-mediated phagocytosis (9), thereby promoting tumor progression. CD47 expression has also been associated with the clinical features and prognosis of a variety of malignancies; for example, its expression is associated with poor prognosis and the pathological features of colorectal cancer (10) and it affects the progression-free survival (PFS) of patients with non-small cell lung cancer (11). CD47 also appears to have an important role in several hematological malignancies, such as acute lymphoblastic leukemia (ALL) and AML (12). Clinical studies of CD47 antibodies or targeted drugs for MDS or AML have been carried out, but most have been based on blocking CD47 expression to restore the phagocytosis of foreign cells by macrophages.

In the present study, we examined the role of CD47 from the perspective of the transformation of MDS to AML. In recent years, the development of genetic testing technology has greatly assisted the study of the genetic landscapes of diseases and their relationship to the pathogenesis and progression of diseases such as MDS (13). Comprehensive genomic analysis of MDS and AML has enabled the detection and differentiation of drivers and subclonal mutations, while informing risk prediction and defining targeted therapies (14). Here, we used high-throughput microarray technology and comprehensive bioinformatics analysis to explore the potential relationships between CD47 expression and functional pathways in MDS and AML.

# DATA AND METHODS

## Data Sources

Data microarrays related to MDS and AML were obtained from the Gene Expression Omnibus (GEO) database. The gse30029 dataset included 90 AML samples and 31 normal control samples, the gse24395 dataset included 12 AML samples and 5 normal control samples, the gse10695 dataset included 20 MDS samples and 10 normal control samples, and the gse30195 dataset included 15 MDS samples and 4 normal control samples. Bone marrow RNA-seq data for AML were obtained from the Target database for a total of 240 cases. RNA-seq data of tumors were obtained from The Cancer Genome Atlas (TCGA) database and paraneoplastic data from the GTEx dataset. Tumor cell sequencing data were obtained from CCLE for 27 tissues.

## Gene Expression Analysis

The distribution of CD47 expression in the two groups of samples was analyzed using the Wilcoxon rank sum test, and violin plots were drawn with the R package ggplot2. Spearman's correlation analysis was used to describe the correlation between quantification and variables without normal distribution, and the correlation maps of genes were presented using the R package pheatmap. A value of $p<0.05$ was considered statistically significant, with an absolute value of the correlation coefficient closer to 1, indicating a stronger correlation.

## Analysis of Differences

Microarrays for MDS or AML were obtained from the GEO database, and samples were categorized into disease or control groups based on clinical information. Differential expression of mRNA was analyzed using the R package limma package with $P<0.05$, and $|\log2FC|>1$ was defined as the screening threshold for differentially expressed genes (DEG). The results of the differential analysis for each data set were presented using volcano plots. The overlap of the two DEG sets was then used in subsequent analysis.

## Protein-Protein Interaction Network

We analyzed the functional pathways common to MDS and AML by PPI analysis of the overlapping genes. The STRING database (https://string-db.org/) was used to obtain the network relationship map of the overlapping genes, and the key functional gene modules were obtained by the MCODE plug-in in Cytoscape software. We used GeneMANIA (https://genemania.org/) to construct a PPI network centered on CD47, which included association data for protein and genetic interactions, pathways, co-expression, co-localization, and protein structural domain similarity.

## Functional and Pathway Enrichment Analysis

Metascape (https://metascape.org/gp/index.html#/main/step1), a website for analyzing gene or protein lists, was used to analyze the functional clustering of gene sets. The R package ClusterProfiler package was used to analyze gene sets for gene ontology (GO) and the Kyoto Encyclopedia of Genes (KEGG), and $P<0.05$ was considered significant. Gene set enrichment analysis (GSEA) was performed to investigate the biological signaling pathways between high and low CD47 expression. $|NES|>1$,NOM p-val<0.05,FDR q-val<0.25 were the pathway screening thresholds.

# RESULTS

## Expression Distribution of CD47

Further analysis of the CD47 expression levels in MDS and AML using GEO microarrays revealed expression in MDS tissues (**Figures 1A, B**) and high expression in AML tissues (**Figures 1C, D**) compared to control tissues. Integrate the data of these four chips and eliminate outlier samples to observe the expression of CD47. It was found that the expression level of CD47 in AML was significantly higher than that in MDS (**Figure 1E**). The waterfall diagram shows the mutations of 10 genes in AML. Compared with other genes, the mutation frequency of CD47 is not high, only 1% (**Figure 1F**). However, we found that mutations in CD47 involved only one variant classification of missense mutations (**Figure 1G**). This mutation usually causes protein abnormalities.

## Functional Analysis of CD47-Related Genes

A PPI network of 21 genes centered on CD47 was constructed using GeneMANIA (**Figure 2A**). GO functional enrichment and KEGG pathway analysis were performed for these 21 genes. Significantly enriched GO terms included leukocyte migration, cell adhesion and activation, and integrin-mediated signaling pathways (**Figure 2B**), while significantly enriched KEGG pathways included EMC receptor interactions, adherent spots, human papillomavirus infection, and PI3K-Akt signaling pathway (**Figure 2C**).

## Identification and Analysis of DEGs

We used differential analysis to discern the common functions and pathways in MDS and AML disease progression. Overall, 693 downregulated genes, and 924 upregulated genes were identified in GSE30195 (**Figure 3A**) and 1327 downregulated genes and 1162 upregulated genes in GSE30029 (**Figure 3B**). The Venn diagram yielded 248 overlapping genes that were differentially expressed in both MDS and AML (**Figure 3C**).

## Function of Key Gene Modules

The functional relationships of 248 overlapping genes were presented as PPI network diagrams (**Figure 4A**). Subsequently, four modules were obtained with the MCODE plug-in, and module 1 was selected as the key gene module (**Figure 4B**). The module was scored as 4.5 and contained 5 nodes and 9 edges, each node representing a gene. Enrichment analysis of these 5 genes was performed using Metascape, with each color indicating a functional cluster (**Figure 4C**). The significantly enriched functional terms include cytokinesis, nuclear division, and meiosis I phase. The relationship between CD47 and these five genes was analyzed using the Spearman correlation, and CD47 showed a varying degree of correlation with each gene; the greatest correlation was with CKAP2 (**Figure 3D**).

## Functional Pathways of CD47 in MDS and AML

The GSEA results showed that the CD47 high expression group in MDS tissues was mainly enriched in Hallmark pathways, such as heme metabolism, G2M checkpoint, and E2F target (**Figure 5A**), and in KEGG pathways, such as pressin-regulated water reabsorption, cell cycle, and spliceosome (**Figure 5B**). By contrast, CD47 in AML tissues was mainly enriched in Hallmark pathways, such as protein secretion, TGF-β signaling, and NOTCH signaling (**Figure 5C**), and KEGG pathways, such as ribosomes, ubiquitin-mediated protein hydrolysis, and amyl-TRNA biosynthesis (**Figure 5D**).

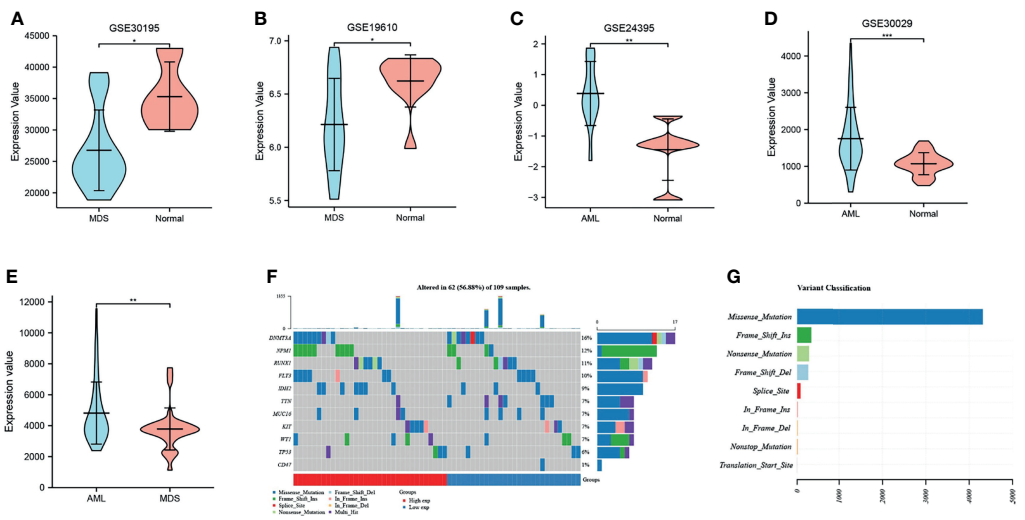



**FIGURE 1** | The expression and mutation of CD47. **(A, B)** CD47 expression analysis in MDS and normal control samples with datasets GSE30196 and GSE19610; **(C, D)** CD47 expression in AML and normal control samples using the GSE24395 and GSE30029 data sets. *P < 0.05, **P < 0.01, ***P < 0.001. **(E)** expression level of CD47 in AML and MDS. **(F)** Oncoplot shows the somatic landscape of acute myeloid leukemia. **(G)** Variant classification of CD47 mutations.

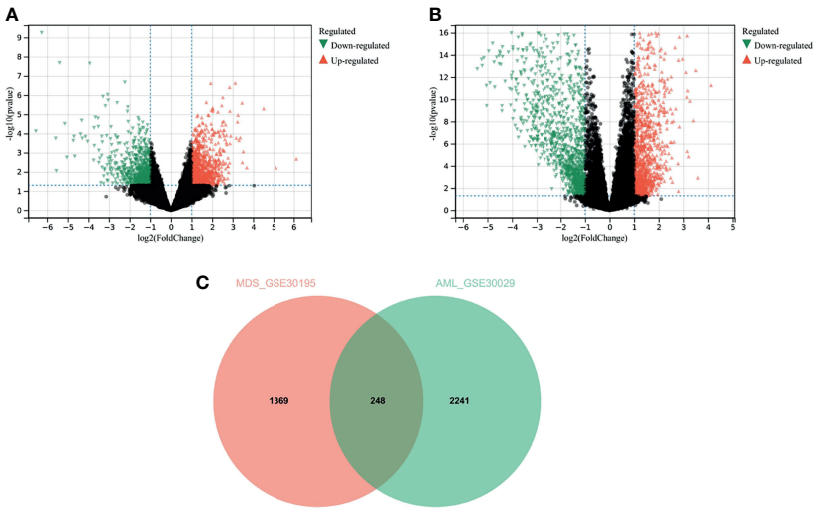



**FIGURE 2** | Results of analysis of variance. **(A)** analysis of variance for the MDS dataset GSE30195; **(B)** analysis of variance for the AML dataset GSE30029; **(C)** Venn diagram for the overlapping genes of the two sets of DEGs.

## DISCUSSION

In recent years, immunotherapy has shown a therapeutic effect against a variety of malignancies. One of the key cell types in the innate immune response is the macrophage, and cells expressing CD47 have the ability to evade the clearance by macrophages and other phagocytes. Consequently, CD47 is considered to be a main macrophage checkpoint (15). Given this relationship between CD47 and macrophages, many studies are now analyzing the roles of CD47 in tumor immunotherapy.

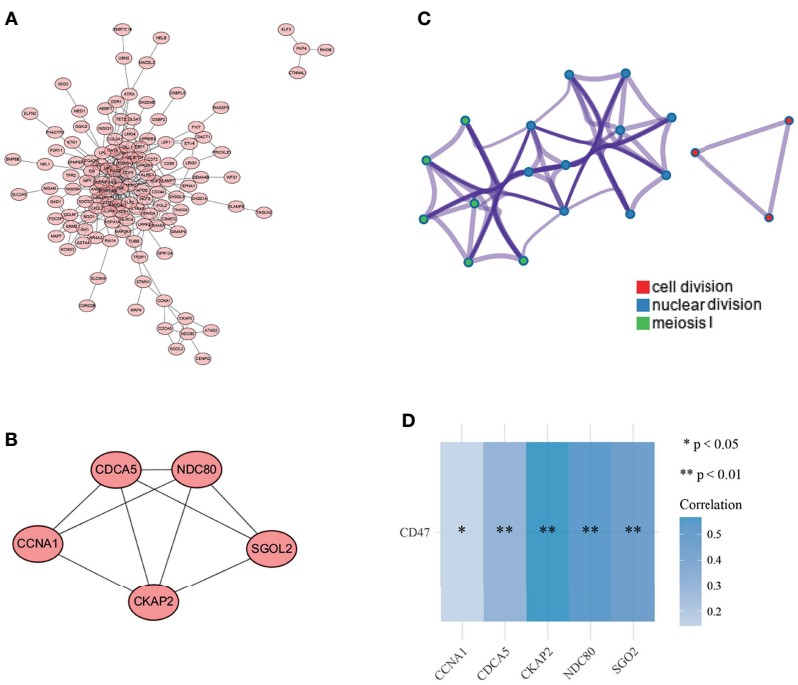



**FIGURE 3** | Selection of key gene modules. **(A)** PPI network of overlapping genes; **(B)** gene modules with the highest tightness were analyzed by MCODE plug-in **(C)**; functional clustering network map of overlapping genes; **(D)** correlation analysis of CD47 with 5 key genes in the target database (sgol2 is also called sgO2). *P < 0.05, **P < 0.01.

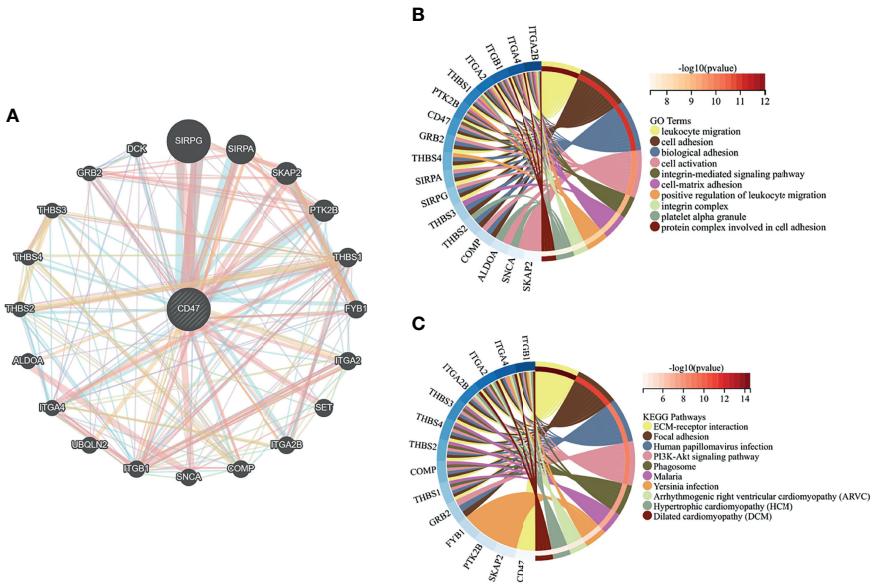



**FIGURE 4** | Gene interaction network and enrichment analysis. **(A)** Construction of a PPI network with 21 genes centered on CD47; **(B)** Top 10 GO functional terms; **(C)** Top 10 KEGG pathways.

CD47 has shown potent anticancer potentials in a variety of hematologic malignancies, including MDS and AML, and a number of CD47-related antibodies or target drugs have emerged. For example, the humanized anti-CD47 monoclonal antibody CC-90002 enables tumor cells killed by macrophages by blocking the CD47/SIRPα interaction (16). Similarly, a SIRPα-αCD123 antibody, reported by Siret Tahk et al, intervenes against tumors by blocking local CD47 and binding to a single molecule on specific leukemic stem cells (17). These studies have revealed the mechanisms of actions of CD47

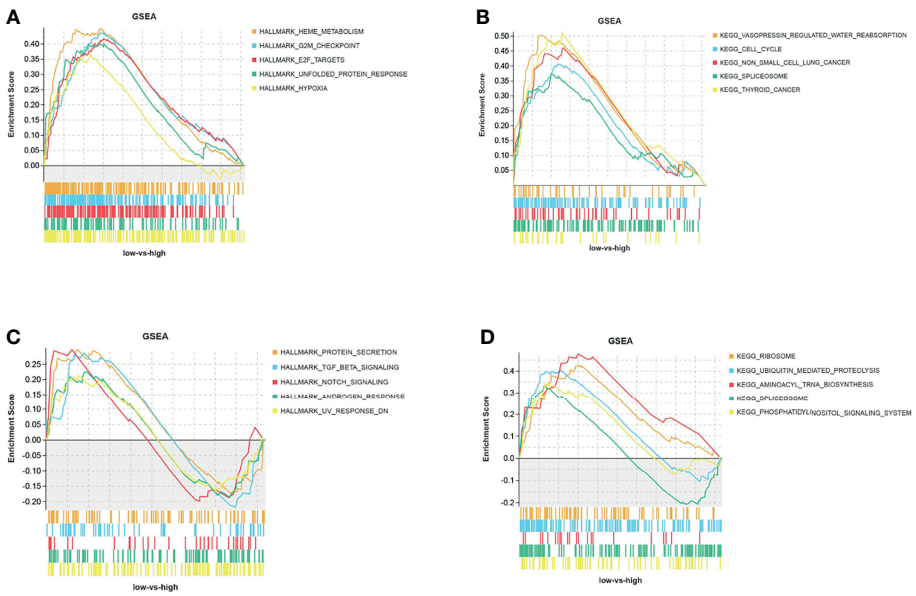



**FIGURE 5** | GSEA analysis of CD47 Hallmark and KEGG pathways. **(A)** Hallmark and **(B)** KEGG pathways of CD47 in MDS tissues from the GSE30195 dataset, and **(C)** hallmark and **(D)** KEGG pathways of CD47 in AML tissues from the GSE30029 dataset.

in MDS or AML from a horizontal perspective, whereas the present study explored the effect and association of CD47 expression on the transformation of MDS to AML from a longitudinal perspective.

MDS is a set of diseases associated with ineffective hematopoiesis, which is characterized by increased apoptosis of early and mature hematopoietic cells (18). In clinical practice, MDS is considered to transition to AML when the number of bone marrow blasts exceeds20% (19). In MDS, DNA methylation or DNA repair, chromatin modification, RNA splicing, signal transcription, and mucin regulation are the main mutational targets, and these mutational modalities share a common clonal origin with AML (20). This is because mutations in MDS and AML allow a well-organized expansion of the initiating clone, while the functional interactions that exist between mutations also determine disease progression.

Many studies now consider mutational events involving active signaling, myeloid transcription, or tumor suppressors as necessary for the progression of MDS to AML (21). A clear example is the difference in the mutation frequency of certain genes in MDS and AML. For example, some genes, such as receptor tyrosine kinases (FLT3 and KIT) and RAS pathway genes, have a higher proportion of mutations in AML, whereas mutations in splicing factors (SF) and epigenetic regulators, among others, are more prevalent in MDS (20). Another major cause of conversion of MDS to AML is DNA methylation (22), an epigenetic modification that regulates gene expression and is a key event in tumorigenesis (23). DNA methylation involves the covalent bonding of a methyl group at the cytosine 5' carbon position of the genomic CpG dinucleotide in the presence of DNA methyltransferase. Many malignancies, including AML and MDS, typically exhibit aberrant DNA methylation and altered histone modifications that result in altered gene expression (24). For example, a study by Wen Jing Ding et al. demonstrated that upregulation of the expression of RAP1GTPase activating protein 1 (Rap1GAP), a gene involved in hematopoietic regulation, was associated with a lower methylation status of the promoter region of this gene in MDS patients (25).

CD47 has been reported to show high expression in AML tissues, and higher CD47 mRNA expression is an independent factor for poor prognosis in AML patients (26). One study showed that CD47 expression in MDS gradually increased with the evolution of risk scores in the International Prognostic Scoring System (IPSS-R), suggesting that CD47 expression levels may contribute to the progression from MDS to AML (27). In this study, CD47 was found to be significantly expressed in AML but poorly expressed in MDS. Meanwhile, the expression of CD47 correlated with DNA methyltransferase in AML. Therefore, we hypothesized that this differential expression of CD47 was associated with an aberrant DNA methylation status.

In general, DNA hypermethylation induces transcriptional repression of oncogenes, while hypomethylation induces activation of oncogenes (28). However, increased expression of promoter methylation-regulated genes occasionally occurs (29). Demethylation drugs have been used clinically to treat patients with MDS, and the methylation of certain oncogenes is confirmed to lead to the development or progression of MDS (30). Aberrant DNA methylation is also the main mechanism of tumor suppressor gene silencing and clonal variation during the evolution of MDS to AML (31). Thus, aberrant DNA methylation may induce the transformation of MDS to AML by altering the expression of CD47, although this speculation still needs confirmation by further studies.

In addition, the functional pathways of key genes were analyzed. The results showed that the genes associated with AML and MDS were significantly enriched in cell cycle related signal pathways including cell division. Coincidentally, GSEA results showed that CD47 was also related to cell growth, cell cycle and other related signal pathways in AML and MDS. The activation of CD47 may induce the growth of tumor cells and accelerate the proliferation and transformation (32). It is reported that the increase of cell proliferation is conducive to cell mutation and leukemia transformation (33). Misreplication during cell division can lead to the increase of mutation load and the loss and accumulation of methylation, which will also coordinate the regulation of cell cycle during tissue formation (34, 35). These evidences suggest that abnormal cell cycle is closely related to the cell mutation and DNA methylation. However, there is no final conclusion whether the expression of CD47 promotes the transformation of acute myeloid leukemia by affecting cell cycle related pathways. This study puts forward this assumption here, but the specific mechanism still needs to be verified by in further experiments.

In this study, we analyzed the expression of CD47 in MDS and AML and conducted a functional pathway analysis of CD47. We found that CD47 was differentially expressed in MDS and AML, and the difference in CD47 expression may reflect an abnormal DNA methylation status, which may be associated with the conversion of MDS to AML.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

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

## FUNDING

Zhejiang Provincial Health Science and Technology Program 2022KY316; 2021KY997; Zhejiang Provincial Natural Science Foundation LY20H080001.

## REFERENCES


1. Bewersdorf JP, Carraway H, Prebet T. Emerging Treatment Options for Patients With High-Risk Myelodysplastic Syndrome. *Ther Adv Hematol* (2020) 11:2040620720955006. doi: 10.1177/2040620720955006
2. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 Revision to the World Health Organization Classification of Myeloid Neoplasms and Acute Leukemia. *Blood J Am Soc Hematol* (2016) 127(20):2391–405. doi: 10.1182/blood-2016-03-643544
3. Greenberg PL, Stone RM, Al-Kali A, Barta SK, Bejar R, Bennett JM, et al. Myelodysplastic Syndromes, Version 2.2017, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Cancer Netw* (2017) 15(1):60–87. doi: 10.6004/jnccn.2017.0007
4. Yu J, Li Y, Li T, Li Y, Xing H, Sun H, et al. Gene Mutational Analysis by NGS and its Clinical Significance in Patients With Myelodysplastic Syndrome and Acute Myeloid Leukemia. *Exp Hematol Oncol* (2020) 9(1):1–11. doi: 10.1186/s40164-019-0158-5
5. Cortes JE, Heidel FH, Hellmann A, Fiedler W, Smith BD, Robak T, et al. Randomized Comparison of Low Dose Cytarabine With or Without Glasdegib in Patients With Newly Diagnosed Acute Myeloid Leukemia or High-Risk Myelodysplastic Syndrome. *Leukemia* (2019) 33(2):379–89. doi: 10.1038/s41375-018-0312-9
6. Greenberg PL, Young NS, Gattermann N. Myelodysplastic Syndromes. *ASH Educ Program Book* (2002) 2002(1):136–61. doi: 10.1182/asheducation-2002.1.136
7. Duncavage EJ, Jacoby MA, Chang GS, Miller CA, Edwin N, Shao J, et al. Mutation Clearance After Transplantation for Myelodysplastic Syndrome. *New Engl J Med* (2018) 379(11):1028–41. doi: 10.1056/NEJMoa1804714
8. OK Weinberg, RP Hasserjian eds. The Current Approach to the Diagnosis of Myelodysplastic Syndromes☆. In: *Seminars in Hematology*. Elsevier. doi: 10.1053/j.seminhematol.2018.05.015
9. Zhang W, Huang Q, Xiao W, Zhao Y, Pi J, Xu H, et al. Advances in Anti-Tumor Treatments Targeting the CD47/Sirpα Axis. *Front Immunol* (2020) 11:18. doi: 10.3389/fimmu.2020.00018
10. Arrieta O, Aviles-Salas A, Orozco-Morales M, Hernández-Pedro N, Cardona AF, Cabrera-Miranda L, et al. Association Between CD47 Expression, Clinical Characteristics and Prognosis in Patients With Advanced non-Small Cell Lung Cancer. *Cancer Med* (2020) 9(7):2390–402. doi: 10.1002/cam4.2882
11. Kim H, Jee S, Kim Y, Sim J, Bang S, Son HK, et al. Correlation of CD47 Expression With Adverse Clinicopathologic Features and an Unfavorable Prognosis in Colorectal Adenocarcinoma. *Diagnostics* (2021) 11(4):668. doi: 10.3390/diagnostics11040668
12. Chao MP, Alizadeh AA, Tang C, Jan M, Weissman-Tsukamoto R, Zhao F, et al. Therapeutic Antibody Targeting of CD47 Eliminates Human Acute Lymphoblastic Leukemia. *Cancer Res* (2011) 71(4):1374–84. doi: 10.1158/0008-5472.CAN-10-2238
13. Maciejewski JP, Balasubramanian SK. Clinical Implications of Somatic Mutations in Aplastic Anemia and Myelodysplastic Syndrome in Genomic Age. *Hematol 2014 Am Soc Hematol Educ Program Book* (2017) 2017(1):66–72. doi: 10.1182/asheducation-2017.1.66
14. Tawana K, Drazer MW, Churpek JE. Universal Genetic Testing for Inherited Susceptibility in Children and Adults With Myelodysplastic Syndrome and Acute Myeloid Leukemia: Are We There Yet? *Leukemia* (2018) 32(7):1482–92. doi: 10.1038/s41375-018-0051-y
15. Chao MP, Weissman IL, Majeti R. The Cd47–Sirpα Pathway in Cancer Immune Evasion and Potential Therapeutic Implications. *Curr Opin Immunol* (2012) 24(2):225–32. doi: 10.1016/j.coi.2012.01.010
16. Russ A, Hua AB, Montfort WR, Rahman B, Riaz IB, Khalid MU, et al. Blocking Don't Eat Me" Signal of CD47-SIRPα in Hematological Malignancies, an In-Depth Review. *Blood Rev* (2018) 32(6):480–9. doi: 10.1016/j.blre.2018.04.005
17. Tahk S, Vick B, Hiller B, Schmitt S, Marcinek A, Perini ED, et al. Sirpα-αcd123 Fusion Antibodies Targeting CD123 in Conjunction With CD47 Blockade Enhance the Clearance of AML-Initiating Cells. *J Hematol Oncol* (2021) 14(1):1–17. doi: 10.1186/s13045-021-01163-6
18. Albitar M, Manshouri T, Shen Y, Liu D, Beran M, Kantarjian HM, et al. Myelodysplastic Syndrome is Not Merely "Preleukemia". *Blood J Am Soc Hematol* (2002) 100(3):791–8. doi: 10.1182/blood.V100.3.791
19. Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, et al. The World Health Organization Classification of Neoplastic Diseases of the Hematopoietic and Lymphoid Tissues: Report of the Clinical Advisory Committee Meeting, Airlie House, Virginia, November, 1997. *Ann Oncol* (1999) 10(12):1419–32. doi: 10.1023/A:1008375931236
20. Ogawa S. Genetics of MDS. *Blood J Am Soc Hematol* (2019) 133(10):1049–59. doi: 10.1182/blood-2018-10-844621
21. Li X, Xu F, Wu L-Y, Zhao Y-S, Guo J, He Q, et al. A Genetic Development Route Analysis on MDS Subset Carrying Initial Epigenetic Gene Mutations. *Sci Rep* (2020) 10(1):1–8. doi: 10.1038/s41598-019-55540-w
22. Shih L, Huang C, Wang P, Wu J, Lin T, Dunn P, et al. Acquisition of FLT3 or N-Ras Mutations is Frequently Associated With Progression of Myelodysplastic Syndrome to Acute Myeloid Leukemia. *Leukemia* (2004) 18(3):466–75. doi: 10.1038/sj.leu.2403274
23. Dexheimer GM, Alves J, Reckziegel L, Lazzaretti G, Abujamra AL. DNA Methylation Events as Markers for Diagnosis and Management of Acute Myeloid Leukemia and Myelodysplastic Syndrome. *Dis Markers* (2017) 2017. doi: 10.1155/2017/5472893
24. Larsson CA, Cote G, Quintás-Cardama A. The Changing Mutational Landscape of Acute Myeloid Leukemia and Myelodysplastic Syndrome. *Mol Cancer Res* (2013) 11(8):815–27. doi: 10.1158/1541-7786.MCR-12-0695
25. Ding WJ, Yang Y, Chen ZX, Wang YY, Dong WL, Cen JN, et al. Methylation Level of Rap1GAP and the Clinical Significance in MDS. *Oncol Lett* (2018) 16 (6):7287–94. doi: 10.3892/ol.2018.9503
26. Majeti R, Chao MP, Alizadeh AA, Pang WW, Jaiswal S, Gibbs KDJr., et al. CD47 is an Adverse Prognostic Factor and Therapeutic Antibody Target on Human Acute Myeloid Leukemia Stem Cells. *Cell* (2009) 138(2):286–99. doi: 10.1016/j.cell.2009.05.045
27. Boasman K, Simmonds MJ, Rinaldi CR. Expression of CD47 and Calr in Myeloproliferative Neoplasms and Myelodysplastic Syndrome: Potential New Therapeutical Targets. *Am Soc Hematol Washington DC* (2019) 134:5031. doi: 10.1182/blood-2019-128422
28. Franco R, Schoneveld O, Georgakilas AG, Panayiotidis MI. Oxidative Stress, DNA Methylation and Carcinogenesis. *Cancer Lett* (2008) 266(1):6–11. doi: 10.1016/j.canlet.2008.02.026
29. Wang L-H, Choi Y-L, Hua X-Y, Shin Y-K, Song Y-J, Youn S-J, et al. Increased Expression of Sonic Hedgehog and Altered Methylation of its Promoter Region in Gastric Cancer and its Related Lesions. *Modern Pathol* (2006) 19 (5):675–83. doi: 10.1038/modpathol.3800573
30. Tsujioka T, Yokoi A, Uesugi M, Kishimoto M, Tochigi A, Suemori S, et al. Effects of DNA Methyltransferase Inhibitors (DNMTIs) on MDS-Derived Cell Lines. *Exp Hematol* (2013) 41(2):189–97. doi: 10.1016/j.exphem.2012.10.006
31. Jiang Y, Dunbar A, Gondek LP, Mohan S, Rataul M, O'Keefe C, et al. Aberrant DNA Methylation is a Dominant Mechanism in MDS Progression to AML. *Blood J Am Soc Hematol* (2009) 113(6):1315–25. doi: 10.1182/blood-2008-06-163246
32. Boukhari A, Alhosin M, Bronner C, Sagini K, Truchot C, Sick E, et al. CD47 Activation-Induced UHRF1 Over-Expression is Associated With Silencing of Tumor Suppressor Gene P16ink4a in Glioblastoma Cells. *Anticancer Res* (2015) 35(1):149–57. doi: 10.1186/s13046-016-0453-5
33. Schnerch D, Yalcintepe J, Schmidts A, Becker H, Follo M, Engelhardt M, et al. Cell Cycle Control in Acute Myeloid Leukemia. *Am J Cancer Res* (2012) 2 (5):508.
34. Jung H, Kim HS, Kim JY, Sun J-M, Ahn JS, Ahn M-J, et al. DNA Methylation Loss Promotes Immune Evasion of Tumours With High Mutation and Copy Number Load. *Nat Commun* (2019) 10(1):1–12. doi: 10.1038/s41467-019-12159-9
35. Shim S, Lee HG, Park O-S, Shin H, Lee K, Lee H, et al. Dynamic Changes in DNA Methylation Occur in TE Regions and Affect Cell Proliferation During Leaf-to-Callus Transition in Arabidopsis. *Epigenetics* (2022) 17(1):41–58. doi: 10.1080/15592294.2021.1872927


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers in Oncology

ORIGINAL RESEARCH
published: 31 March 2022
doi: 10.3389/fonc.2022.857714

# miR-139-5p Was Identified as Biomarker of Different Molecular Subtypes of Breast Carcinoma

*Haohang Sun[1], Ji Dai[2], Mengze Chen[1], Qi Chen[1], Qiong Xie[1], Weijun Zhang[1], Guoqing Li[1] and Meidi Yan[1]**

*[1] General Surgery I (Thyroid, Breast, Vascular, Hernia Surgery), General Hospital of Zhenhai District People's Hospital Medical Group, Ningbo, China, [2] Department of General Surgery, Zhenhai District People's Hospital, Ningbo, China*

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Yuqing Hang,*
*Salk Institute for Biological Studies, United States*
*Hailun Zhou,*
*Guangxi Medical University, China*

***Correspondence:**
*Meidi Yan*
*2510700535@qq.com*

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

**Received:** *19 January 2022*
**Accepted:** *24 February 2022*
**Published:** *31 March 2022*

**Citation:**
*Sun H, Dai J, Chen M, Chen Q, Xie Q, Zhang W, Li G and Yan M (2022) miR-139-5p Was Identified as Biomarker of Different Molecular Subtypes of Breast Carcinoma. Front. Oncol. 12:857714. doi: 10.3389/fonc.2022.857714*

Located on chromosome 11q13.4, miR-139-5p has been confirmed by several studies as a possible attractive biomarker for cancer, including breast cancer, but its mechanism of correlation in different molecular subtypes of breast cancer has not been reported. In this study, comprehensive bioinformatics analysis was used to evaluate the expression of miR-139-5p in different molecular subtypes of breast cancer (luminal A, luminal B, HER2-enriched, and basal-like). The target genes of miR-139-5p were predicted by using an online database TargetScan and miRDB, and three key genes, FBN2, MEX3A, and TPD52, were screened in combination with differentially expressed genes in different molecular subtypes of breast cancer. The expression of the three genes was verified separately, and the genes were analyzed for pathway and functional enrichment. Bone marrow mesenchymal stem cells (BMSC) are another kind of highly plastic cell population existing in bone marrow besides hematopoietic stem cells. BMSC can affect the proliferation and migration of cancer cells, promote the metastasis and development of cancer, and regulate the tumor microenvironment by secreting exosome mirnas, thus affecting the malignant biological behavior of tumor cells. Finally, human bone marrow mesenchymal stem cells exosomes were obtained by ultracentrifugation, and the morphology of exosomes was observed by transmission electron microscopy. The expression of miR-139-5p in normal breast cells MCF-10A, human breast cancer cell line MDA-MB-231 cells, and BMSCs-derived exosomes were compared; the exosomes and MDA-MB-231 cells were co-cultured to observe their effects on the proliferation of the MDA-MB-231 cells. Human bone marrow mesenchymal stem cell-derived exosomes inhibited the growth of breast cancer cells and promoted the expression of FBN2, MEX3A, and TPD52 by transporting miR-139-5p.

**Keywords: miR-139-5p, mesenchymal stem cells, exosomes, molecular isoforms, breast carcinoma**

## INTRODUCTION

Breast cancer is a global female health problem (1). According to recent Global Cancer Observatory (GLOBOCAN) 2020 data, the incidence of breast cancer in women has surpassed lung cancer as the leading cause of cancer incidence worldwide, with an estimated 2.3 million new cases, accounting for 11.7% of all cancer cases (2). In addition, the incidence of breast cancer is rising rapidly in South America, Africa, and Asian countries in transition, as well as in high-income Asian countries (Japan, China, and Korea) (3, 4). Breast cancer is a complex heterogeneous tumor that can be classified according to histological features as hormone receptor (ER+, PR+/-), human epidermal growth factor receptor 2+ (HER2+), and triple-negative breast cancer (TNBC), also known as basal cell type, which are key factors affecting patient prognosis (5, 6). Although breast cancer treatment and early diagnosis have continued to improve and the 5-year survival rate of breast cancer patients has increased, breast cancer is still considered the second leading cause of cancer-related death (7).

Pathology plays a key role in understanding complex diseases such as cancer. However, in our country, data on key epidemiological findings are lacking, and studies have been conducted to show that the proportion of breast cancer subtypes varies in different populations (8). However, these traditional classifications do not reflect the diversity of breast cancers. For example, patients with HER2-negative or ER-negative tumors do not respond to neoadjuvant therapy targeting HER2. Women who are HER2-positive or ER-positive tend to show recurrent responses to such targeted therapy (9). To better predict patient prognosis, the overall classification of breast cancer has changed in the last decade or so. Microarray-based gene expression profiling helps to identify histopathological types and molecular subtypes of breast cancer (10). The Cancer Genome Atlas (TCGA), extensive analysis of protein levels, microRNA, and DNA based on the detection of ER, PR, HER2, and low expression proliferating cell nuclear antigen-67 (Ki-67) indicators have helped to establish a more refined subtype of breast cancer. Molecular subtypes include luminal A, luminal B, HER-enriched, and basal-like, each of which has changed the way breast cancer is treated (11). Each subtype is associated with a different incidence, prognosis, preferential metastatic organs, treatment response, recurrence, or disease-free survival outcomes (12–14).

MicroRNAs are single-stranded short non-coding RNAs, and their genes are usually found in clusters, distributed on all chromosomes except the Y chromosome (15). An important regulator, miR-139-5p is closely associated with the proliferation, invasion, and metastasis of a variety of tumors (16). Its mechanism of action in breast cancer has also been discovered by many studies. However, its mechanism of action in different molecular subtypes of breast cancer has not been overly reported. Besides, miRNAs loaded in exosomes have also been reported, demonstrating that exosomes may act as a miRNA transport to regulate intercellular communication (17) and exosomes secreted by the BMSCs have been reported to have therapeutic potential (18). Exosomes are small intraluminal vesicles that are secreted by various cells and can deliver intracellular contents, such as microRNAs (miRNAs), messenger RNAs (mRNAs) and proteins (19, 20). It is known that multiple target genes can exist simultaneously in the same miRNA (21). Therefore, this study investigated the expression of miR-139-5p and its target genes in different molecular subtypes of breast cancer, aiming to provide potential targets for the treatment of different molecular subtypes of breast cancer.

## METHODS

### Data Sources

The RNA-Seq, miRNA-seq, and clinical information of patients with breast invasive carcinoma (BRCA) were downloaded from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). If multiple probes were detecting the same miRNA expression during the analysis, the average of the miRNA expression was taken as the expression value of that miRNA. For the analysis of patient clinical information, the clinical information of patients with unknown survival time and survival time of 0 was deleted.

### Variance Analysis

Differentially expressed genes were screened using the edgeR package in R v4.0.3, with a $|\log FC| \geqq 1$ and an adjust $P$ value < 0.05. Volcano plots of differentially expressed genes were plotted using ggplot2.

### Kaplan-Meier Survival Analysis

The survival analysis was performed using Survival in the R package. The $p$-values and hazard ratios (HR) with 95% confidence intervals (CI) in the Kaplan-Meier curves were derived by log-rank test and univariate Cox proportional hazards regression.

### Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Enrichment Analysis

The GO analysis and KEGG analysis of genes were performed using the DAVID 6.8 database (https://david.ncifcrf.gov/). Enrichment results with $p < 0.05$ or false discovery rate (FDR) < 0.05 were statistically significant.

### Gene Set Enrichment Analysis

The samples were divided into high and low expression groups according to the median gene expression, and RNA-seq profiles were loaded to GSEA to investigate key gene-related signaling pathways in the high and low-risk groups. FDR < 0.25, $P < 0.05$ were considered to be significantly enriched.

### Cell Source and Culture

The human normal breast cells MCF-10A, human breast cancer cell line MDA-MB-231 cells, and human bone marrow mesenchymal stem cells (BMSCs) were purchased from the American Type Culture Collection (ATCC) cell bank (Manassas, VA, USA). The MCF-10A cells were cultured in

DMEM/F12 medium (Gibco) containing 2.5 mM glutamine, 20 ng/ml epidermal growth factor, 0.01 mg/ml insulin, 500 ng/ml hydrocortisone, and 5% horse serum in DMEM/F12 medium (Gibco). The MDA-MB-231 cells were cultured in RPMI 1640 medium (Gibco) containing 10% fetal bovine serum (FBS) and 0.5% penicillin-streptomycin. The BMSCs were cultured in DMEM medium (Gibco) containing 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin in DMEM medium (Gibco). All cells were routinely cultured at 37°C in a 5% $CO_2$ incubator.

## Isolation and Identification of Exosomes

The BMSCs were grown and fused to 80%, rinsed in phosphate buffered saline (PBS), and cultured in DMEM medium without exosomal fetal bovine serum for 2 d. The exosomes were extracted by ultracentrifugation: the cell culture medium was collected, centrifuged at 500r/min for 12 min to remove the cells, and then centrifuged at 5000r/min for 12 min to remove the cell fragments. Collect supernatant filtered by 0.22 μM membrane, low temperature ultracentrifugation for 2 h, take the precipitation and dissolve it in PBS for subsequent experiments. The morphology of exosomes was observed by transmission electron microscope. The exosome specific protein markers CD63, CD9 and HSP70 were detected by western blot.

## Cell Co-Culture with Exosome

The MDA-MB-231 cells were divided into a control and an exosome group. The cells were inoculated in 6-well plates 24 h before treatment, and when the growth fusion reached 70%, 200 μL of exosomes were added to each well of the exosome group and co-cultured with the cells for 48 h. and PBS was added to the control group, and the cells were collected for subsequent experiments.

## Quantitative Real-Time PCR

The total RNA was extracted from the cells using Trizol reagent (Invitrogen), and NanoDrop (Thermo Fisher Scientific) was applied to detect the concentration and purity of the RNA. The cDNA was reverse transcribed into cDNA according to the instructions of the PrimeScript RT Reagent Kit (Takara), and this cDNA was used as a template for PCR reactions, using the ABI Step One Real-time PCR System (Thermo Fisher Scientific) according to the instructions of the SYBR GREEN kit (TaKaRa). The specific primers used for qRT-PCR were synthesized by Shanghai Bioengineering Co (**Table 1**). The mean values were taken.

## CCK-8 Assay for Cell Proliferation Viability

The MDA-MB-231 cells in the logarithmic growth phase were inoculated with $1\times10^4$ cells/well in a 96-well plate and incubated at 37°C for 24, 48, 72, and 96 h. Next, 10 μL of CCK-8 solution was added to each well and incubated continuously for 4 d at 37°C in a constant temperature incubator. Finally, the absorbance value of each well was measured at 450 nn, using an enzyme marker, and the absorbance value was used to represent the cell proliferation level.

## Statistical Methods

Statistical software (SPSS 21.0, IBM Corp. Armonk, USA) was used for data analysis, and Prism 8.0 (Graphpad, USA) was used for graphical presentation. Measures were expressed as mean standard deviation (Mean ± SD), and each experiment was repeated at least three times. Comparisons between two groups for measures that obeyed a normal distribution were performed using the independent sample t-test. Comparisons between multiple groups were made using a one-way ANOVA followed by Tukey's *post hoc* test. A two-tailed $P < 0.05$ was considered statistically significant.

**TABLE 1** | Primer sequences.

| Gene | Primer sequence (5'-3') |
|---|---|
| miR-139-5p | Forward: CGCGTCTACAGTGCACGTGTC |
| | Reverse: AGTGCAGGGTCCGAGGTATT |
| FBN2 | Forward: TGGATTTTGTTCCCGTCCTAAC |
| | Reverse: CAACGTCCACCATTCTGACAT |
| MEX3A | Forward: TGGAGAACTAGGATGTTTCGGG |
| | Reverse: GAGGCAGAGTTGATCGAGAGC |
| TPD52 | Forward: AGCATCTAGCAGAGATCAAGCG |
| | Reverse: AGCCAACAGACGAAAAAGCAG |
| U6 | Forward: CTCGCTTCGGCAGCACA |
| | Reverse: AACGCTTCACGAATTTGCGT |
| GAPDH | Forward: CTGGGCTACACTGAGCACC |
| | Reverse: AAGTGGTCGTTGAGGGCAATG |

# RESULTS

## Expression of miR-139-5p in Breast Cancer and Prognosis

The expression of miR-139-5p in luminal A, luminal B, HER2-enriched, and basal-like breast cancers in the TCGA database was observed; the results showed that compared with normal samples, miR-139-5p expression was reduced in luminal A, luminal B, HER2-enriched, and basal-like breast cancers (**Figure 1A**). The differences pf the expression level of miR-139-5p between basal-like breast cancer and HER2-enriched breast cancer were statistically significant ($P < 0.05$). The prognosis of the four different types of breast carcinoma (**Figure 1B**) showed that there was a difference in prognosis between the four types of breast carcinoma ($P < 0.05$).

## Screening of Differentially Expressed Genes in Different Types of Breast Cancer

Using $|\log FC|>1$ and adjusted $P < 0.05$ as screening conditions, differentially expressed genes were screened from different types of breast cancer in the TCGA database and visualized, using volcano plots. A total of 1674 downregulated genes and 905 upregulated genes were obtained from the screening of luminal A breast cancer (**Figure 2A**); from the screening of luminal B breast cancer, 2270 downregulated genes and 1186 upregulated genes were obtained (**Figure 2B**); and from the screening of HER2-enriched breast cancer, 2513 downregulated genes and 974 upregulated genes were obtained (**Figure 2C**). A total of 2445 downregulated genes and 1092 upregulated genes (**Figure 2D**) were obtained in the basal-like breast cancer screen.

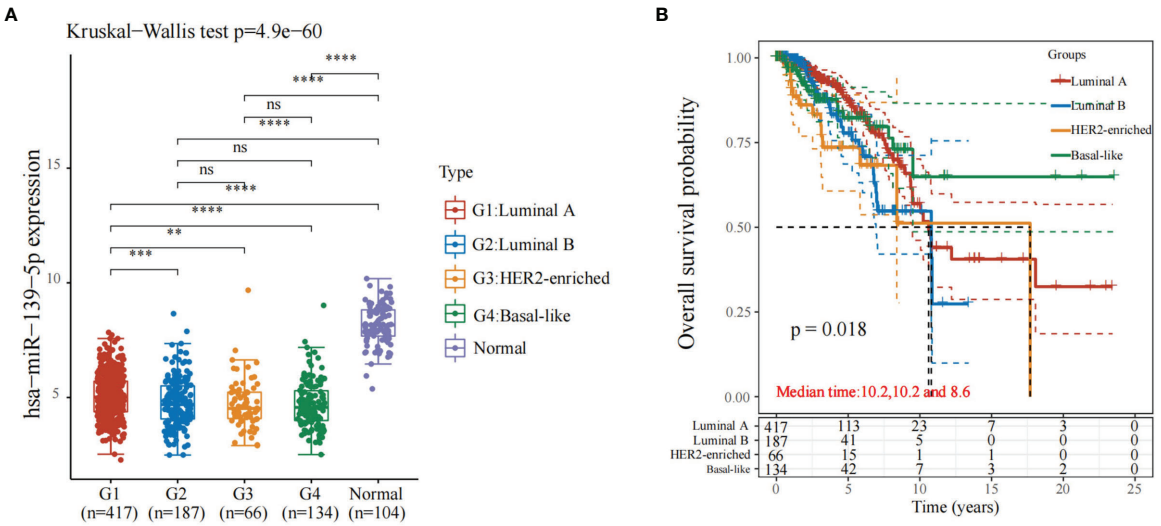



**FIGURE 1** | Expression and prognosis of miR-139-5p in different types of breast cancer. **(A)** Expression of miR-139-5p; **(B)** Prognosis of different types of breast cancer. **P < 0.01, ***P < 0.001, ****P < 0.0001. ns, no significance.

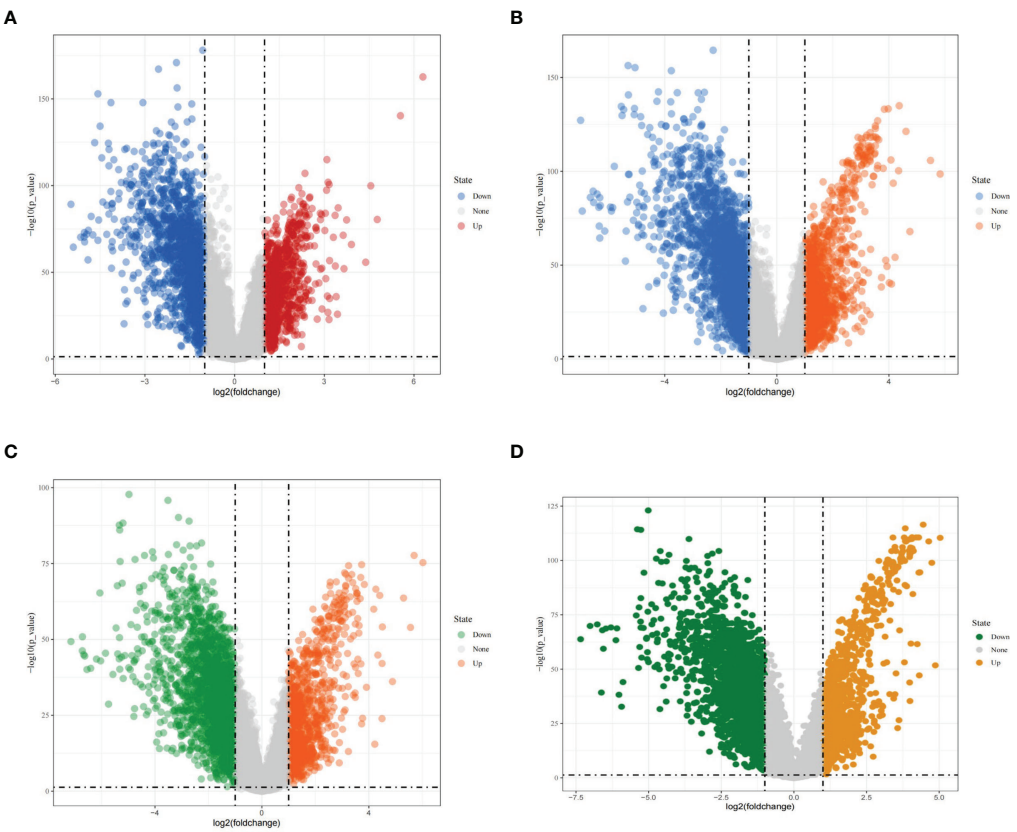



**FIGURE 2** | Differentially expressed genes in different types of breast cancer. **(A)** Volcano diagram showing differentially expressed genes in luminal A breast cancer; **(B)** Volcano diagram showing differentially expressed genes in luminal B breast cancer; **(C)** Volcano diagram showing differentially expressed genes in HER2-enriched breast cancer; **(D)** Volcano diagram showing differentially expressed genes in basal-like breast cancer.

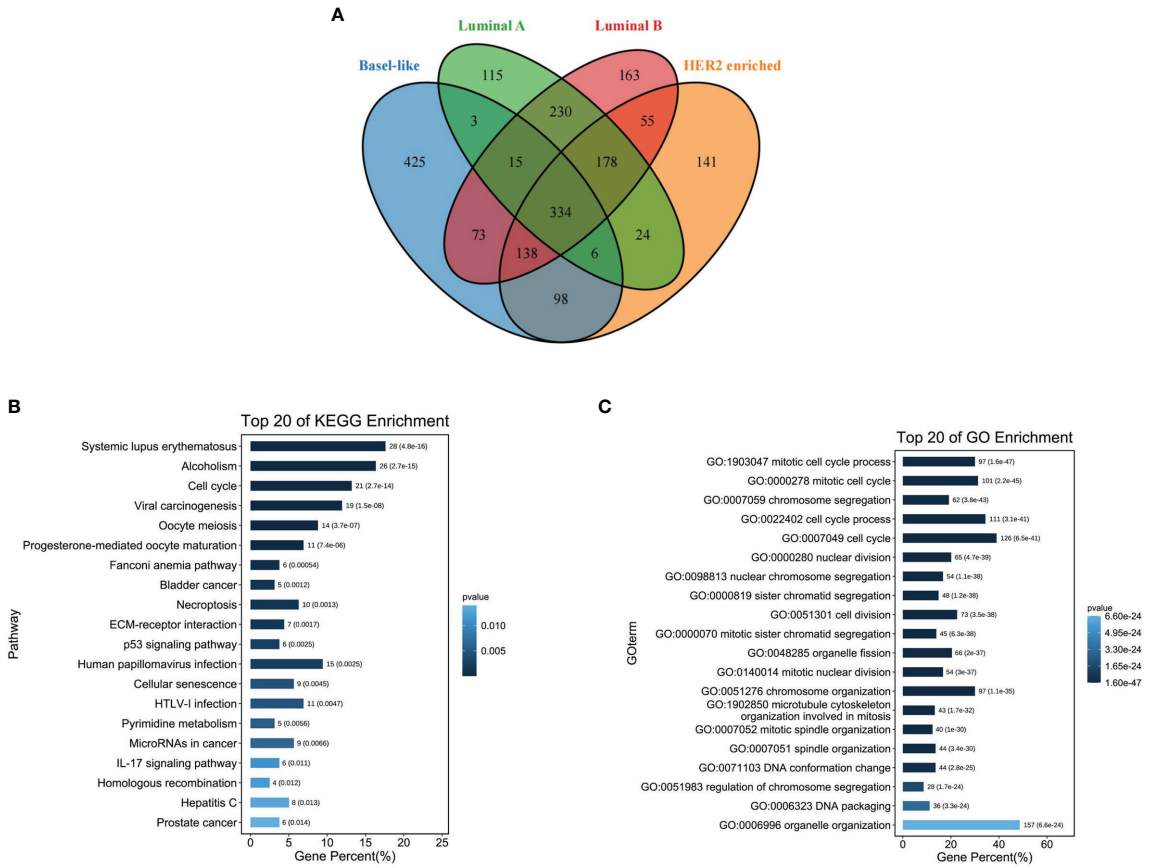



**FIGURE 3** | Functional enrichment analysis of differentially expressed genes in four types of breast cancer. **(A)** Venn diagram showing four types of breast cancer differentially up-regulated overlapping genes; **(B)** Overlapping gene KEGG pathway enrichment analysis; **(C)** Overlapping gene GO function enrichment analysis.

## Functional Enrichment Analysis of Differentially Expressed Genes in Four Types of Breast Cancer

The differentially up-regulated genes in the four types of breast cancer overlapped, and a total of 334 overlapping genes were obtained (**Figure 3A**). Subsequently, the KEGG pathway analysis was performed on the 334 differentially up-regulated genes, which involved a total of 20 pathways, mainly enriched in the systemic lupus erythematosus pathway, alcohol moderate, cell cycle, and other pathways (**Figure 3B**). The GO enrichment analysis showed that 334 differentially upregulated genes were mainly enriched in the cell cycle (**Figure 3C**).

## Target Gene Prediction of miR-139-5p

The target genes of miR-139-5p were predicted by miRDB and TargetScan 8.0, respectively. Six hundred and twenty target genes were obtained in miRDB, 432 target genes in TargetScan 8.0, and 240 target genes were obtained after overlapping (**Figure 4A**); the KEGG and GO functional enrichment analyses were performed on the 240 target genes. The results showed that the 240 target genes were mainly enriched in cancer-related pathways (**Figure 4B**) and multicellular biological functions (**Figure 4C**).

## Expression and Prognosis of predicted Target Genes and Differentially Upregulated Genes

A total of three key genes were obtained after overlapping the predicted target genes with the differentially upregulated genes in the four types of breast cancer: FBN2, MEX3A, and TPD52 (**Figure 5A**). Their expression and prognosis in the four types of breast cancer were observed separately, and the results showed that FBN2, MEX3A, and TPD52 were more expressed in luminal A, HER2, and TPD52 than in normal samples. The results showed that FBN2, MEX3A, and TPD52 were highly expressed in luminal A, luminal B, HER2-enriched, and basal-like breast cancers (**Figures 5B–D**) compared with normal samples. Besides, the overall survival time of different types of breast cancer were showed significant differences (**Figure 5E**).

## Functional Annotation of FBN2, MEX3A, and TPD52

As shown in **Figure 6**, the three KEGG pathways and hallmark pathways most significantly associated with high expression of FBN2, MEX3A, and TPD52 are given. Among them, the high expression of FBN2 was mainly enriched in extracellular matrix

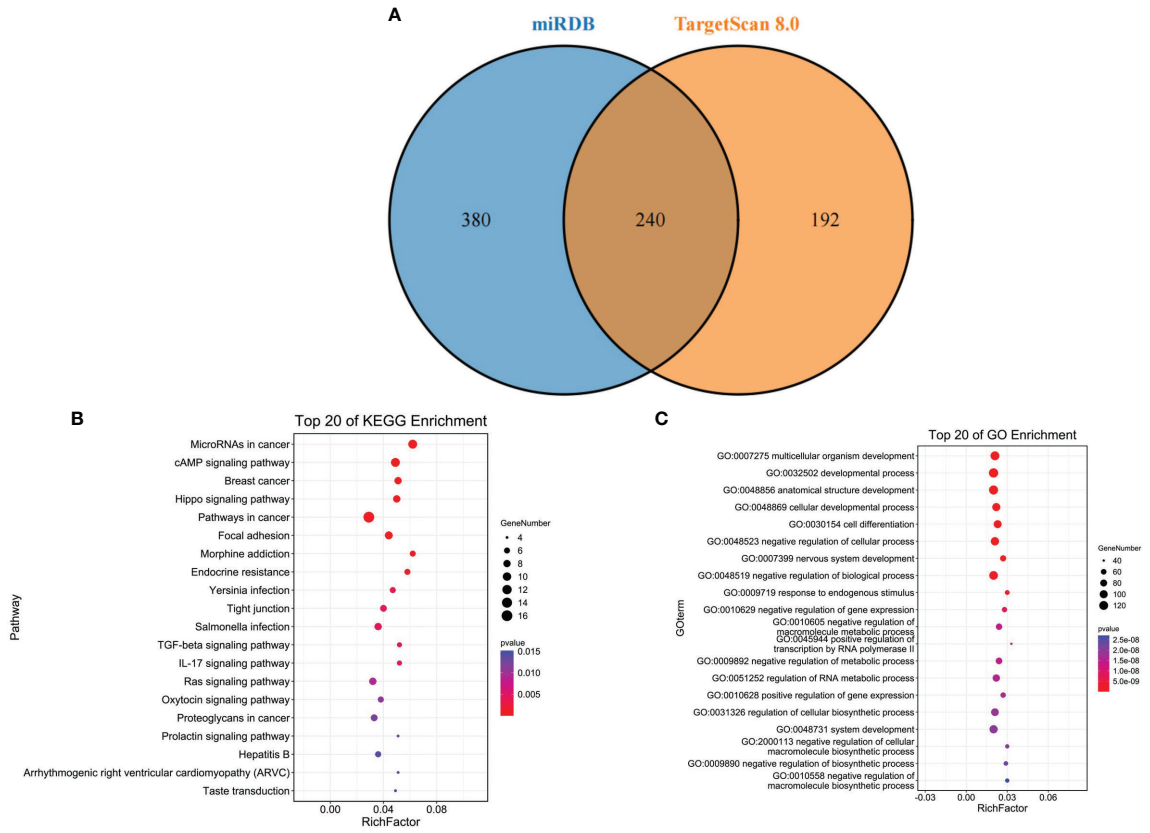



**FIGURE 4** | Target gene prediction of miR-139-5p. **(A)** miR-139-5p in the miRDB database with target gene overlap genes in the TargetScan 8.0 database; **(B)** Overlap gene KEGG pathway enrichment analysis; **(C)** Overlap gene GO function enrichment analysis.

(ECM) receptor interaction and epithelial-mesenchymal transition (EMT) pathway (**Figure 6A**); the high expression of MEX3A was mainly enriched in cell cycle and G2M checkpoint (**Figure 6B**); and the high expression of TPD52 was mainly enriched in cell cycle and E2F targets (**Figure 6C**).

## Exosome Identification of BMSCs

By using the exosome miRNA database (http://bioinfo.life.hust.edu.cn/EVmiRNA#!/), the search revealed that miR-139-5p can be specifically expressed in MSCs (**Figure 7A**). Transmission electron microscopy showed that the exosomes were vesicle-shaped with a diameter of 100–200 nm (**Figure 7B**). The western blot showed that the exosomes could express specific proteins such as CD9, CD63, and HSP70 (**Figure 7C**).

## Expression of miR-139-5p and Its Target Genes in Cells and Effect on Proliferation

In order to further analyze the effect of miR-139-5p on MDAMB-231 cells and its target genes, we examined the expression of miR-139-5p in cells and exocrine and its effect on the proliferation of MDA-MB-231 cells. The qPCR results showed (**Figures 8A–D**) that miR-139-5p expression was decreased in the MDA-MB-231 cells, and FBN2, MEX3A, and TPD52 expression was increased compared with the MCF-10A cells, and the differences were all statistically significant ($P < 0.05$). Further detection of miR-139-5p expression in exosomes showed that miR-139-5p expression was significantly higher in BMSCs exosomes than in MDA-MB-231 cells. In addition, FBN2,MEX3A, and TPD52 expression were significantly lower in MDA-MB-231 cells co-cultured with exosomes compared with MDA-MB-231 cells ($P < 0.05$). The CCK-8 results showed (**Figure 8E**) that after 48 h co-culture of exosomes with MDAMB-231 cells, compared with MDA-MB-231 cells, cell viability was significantly reduced in the exosome group compared with the MDA-MB-231 cells ($P < 0.05$).

# DISCUSSION

Breast cancer remains one of the leading causes of death in women worldwide. Attributed to clinical, pathological, and biological factors, it is defined as a heterogeneous disease and, as such, presents differently in different populations, and treatment modalities should be individualized (22). The St. Gallen Consensus has led the development of personalized treatment for clinical and biological subtypes of breast cancer over the years. The Consensus can also be used to make informed adjuvant treatment decisions (23). However, the prevalence of

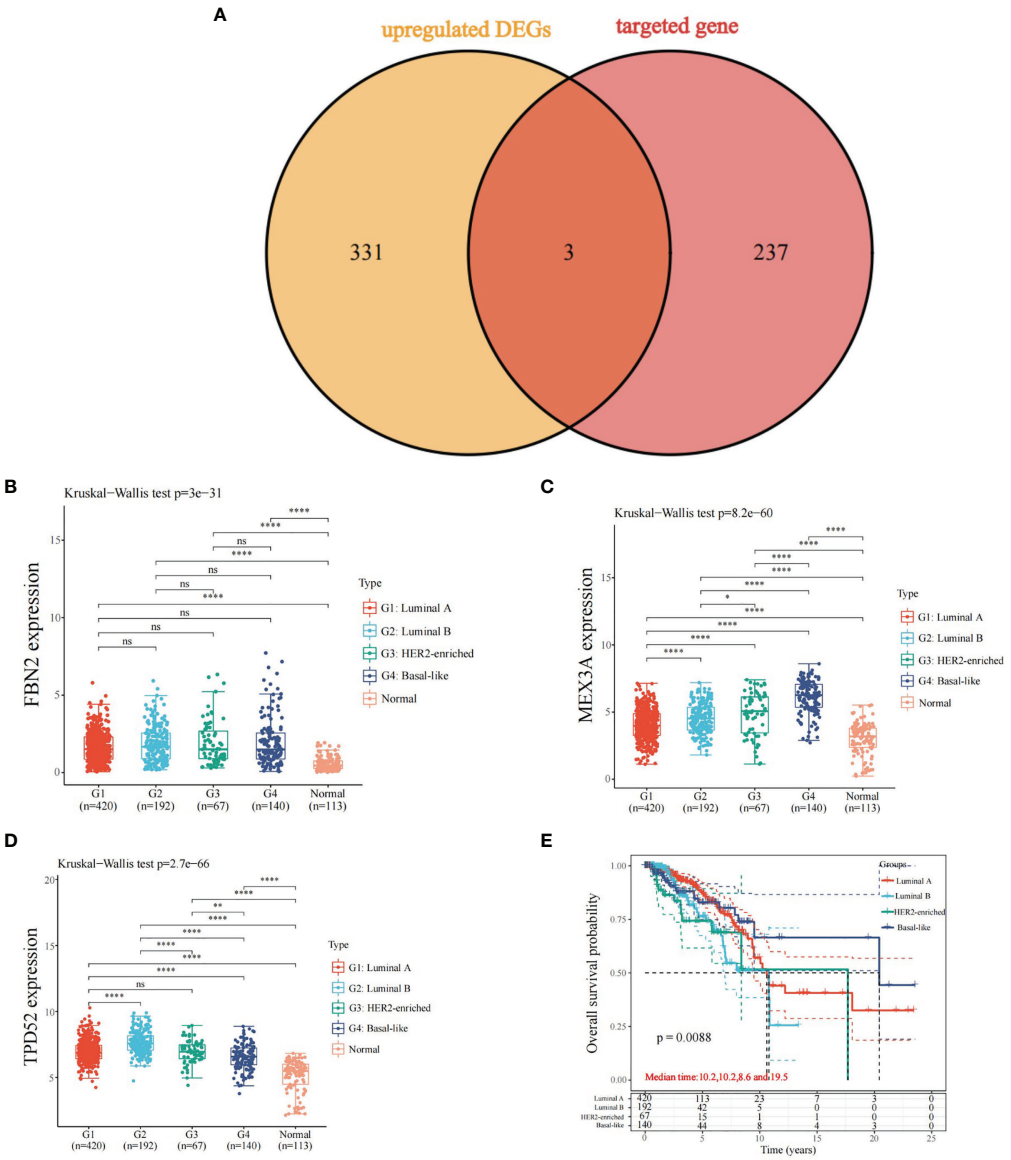



**FIGURE 5** | Expression and prognosis of predicted target genes and differentially up-regulated genes. **(A)** Overlapping genes of predicted target genes and differentially up-regulated genes; **(B)** Expression of FBN2 in different types of breast cancer; **(C)** Expression of MEX3A in different types of breast cancer; **(D)** Expression of TPD52 in different types of breast cancer; **(E)** Overall survival time of different types of breast cancer; *P < 0.05, ****P < 0.0001. ns, no significance.

molecular subtypes of breast cancer has not been extensively studied in developing countries. Over the past 20 years, scientists have studied the role of microRNAs in cancer development. An attractive advantage of microRNAs as cancer biomarkers is their stability in circulating fluid compared with other RNA species, allowing for noninvasive detection and tumor surveillance (24, 25). Numerous studies of differentially expressed genes and mRNAs have shown that miR-139-5p is one of the most important miRNAs in tumorigenesis, and an additional advantage is the ability of a single miRNA to simultaneously regulate many downstream signaling pathways (26).

In this study, we first analyzed the expression of miR-139-5p in different molecular subtypes of breast cancer and found that miR-139-5p was lowly expressed in luminal A, luminal B, HER2-enriched, and basal-like breast cancers, and the expression was significantly different in different molecular subtypes. Some studies have confirmed that miR-139-5p is lowly expressed in breast cancers (27), and the results of the present study also confirmed this finding and found that expression differences existed in different molecular subtypes. Subsequently, we identified three key genes, FBN2, MEX3A, and TPD52, by predicting miR-139-5p target genes and overlaying them with

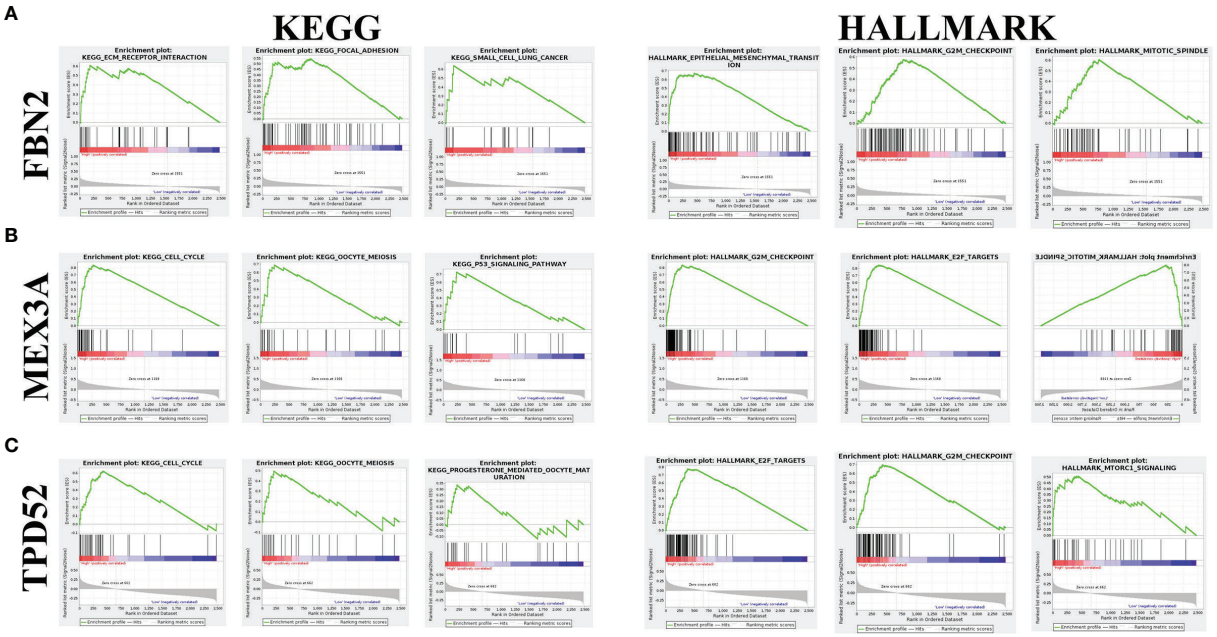



**FIGURE 6** | Enrichment analysis of FBN2, MEX3A, and TPD52 in KEGG and hallmark datasets. **(A)** Top three most relevant pathways of FBN2 in KEGG and hallmark databases; **(B)** Top three most relevant pathways of MEX3A in KEGG and hallmark databases; **(C)** Top three most relevant pathways of TPD52 in KEGG and hallmark databases.

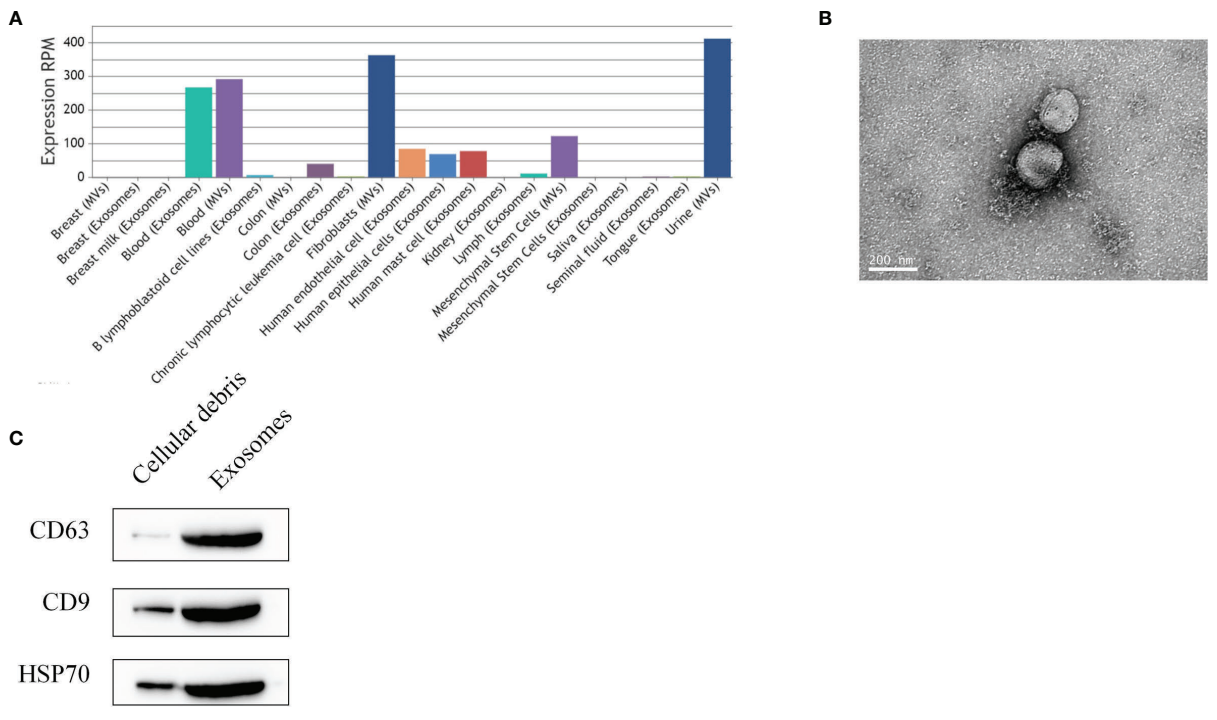



**FIGURE 7** | Identification of exosomes. **(A)** hsa-miR-139-5p expression of extracellular vesicles; **(B)** Transmission electron microscopy observation of exosomes; **(C)** Western blot detection of exosomal surface marker proteins of BMSCs.

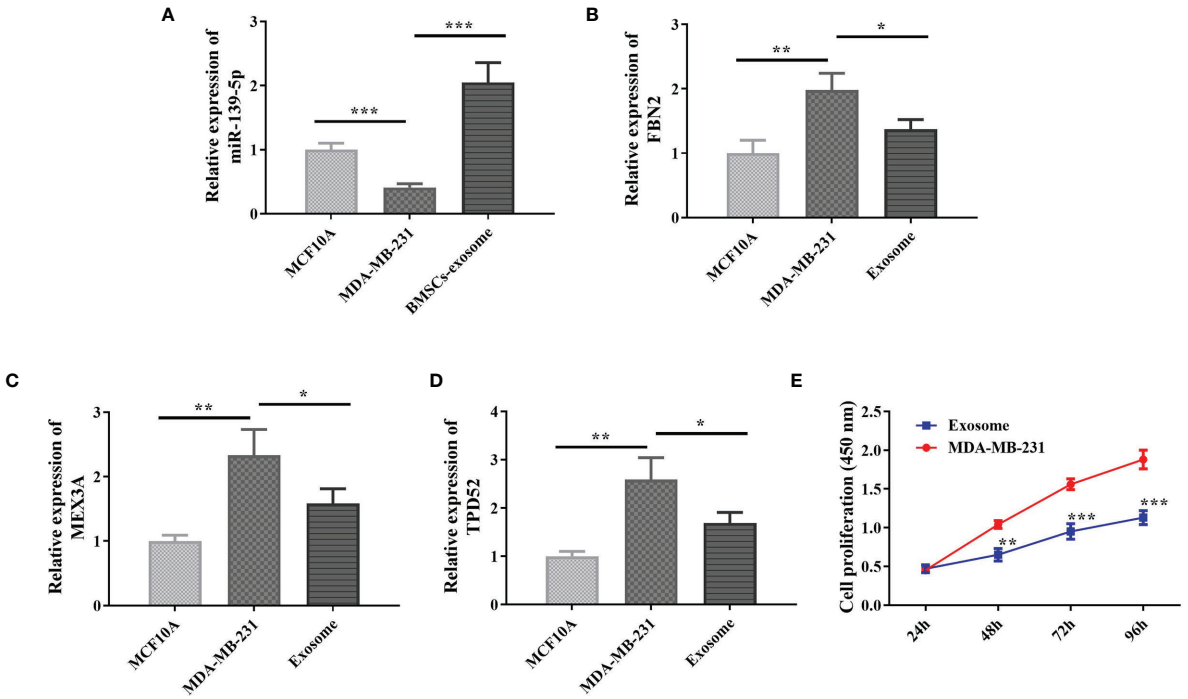



**FIGURE 8** | Expression of miR-139-5p and its target genes in cells and effect on proliferation. **(A)** Expression of miR-139-5p in cells; **(B)** Expression of FBN2 in cells; **(C)** Expression of MEX3A in cells; **(D)** Expression of TPD52 in cells; **(E)** Effect of BMSCs exosomes on cell proliferation, compared with control, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$).

differentially expressed genes in different molecular subtypes, and selected differentially upregulated genes. Fibrillin-2 (FBN2), first expressed at the site of contact between epithelial and mesenchymal cells, is an extracellular calcium-binding microfibril involved in multiple biological pathways, including bone mineralization, osteoblast maturation, and calcium-binding (28). It has been identified as a diagnostic biomarker for smooth muscle sarcoma and rhabdomyosarcoma (29, 30). However, studies on FBN2 have focused on its methylation, and aberrant methylation of FBN2 has been found in breast cancer, non-small cell lung cancer, and esophageal squamous cell carcinoma (31–33), and the promoter region of the FBN2 gene has been repeatedly reported to be hypermethylated in several types of cancer. However, methylation may actually lead to downregulation of FBN2 in primary tumors (34). In contrast, FBN2 was found to be upregulated in different molecular subtypes of breast cancer in the present study, and this result, although different from the findings of previous studies, gives rise to the idea to study the role of FBN2 in a deeper way. MEX3A belongs to a family of RNA-binding proteins, consisting of four members (MEX3A-D). Many studies have pointed out that MEX3A is involved in mRNA regulation and influences the development of various diseases, especially human cancers, such as gastric cancer, bladder cancer, colorectal cancer, liver cancer, and glioblastoma. Jiang et al. (35) found that MEX3 was highly expressed in triple-negative breast cancer and promoted the proliferation and migration of triple-negative breast cancer through a PI3K/AKT signaling pathway. Shi et al. (36) found that MEX3A promotes the development of breast cancer by regulating PIK3CA. In the present study, not only was MEX3A highly expressed in different molecular subtypes of breast cancer, but also the expression was significantly different among the four different molecular subtypes, indicating the specificity of MEX3A expression. Therefore, we suggest that MEX3A may be a potential target for personalized treatment for different molecular subtypes of breast cancer. The tumor protein D52 (TPD52) is amplified from the human chromosome 8q21 amplification region to an oncogene and is highly expressed in many cancers, such as ovarian cancer and prostate cancer (37, 38). In addition, TPD52 is also overexpressed in breast cancer (39). There is increasing evidence that TPD52 is involved in cell transformation, proliferation, apoptosis, and metastasis (40, 41). Zhang et al. (39) found that TPD52 expression was significantly increased in breast cancer tissues and cells, and miR-449 deletion promoted proliferation and metastasis of breast cancer cells by regulating TPD52. In contrast, in the present study, whether TPD52, a target gene of miR-139-5p, plays a role in the proliferation and metastasis of breast cancer cells needs to be further investigated.

To further demonstrate the regulatory role of miR-139-5p in breast cancer, we found that miR-139-5p was expressed in the extracellular bodies of mesenchymal stem cells. In addition, bone MSCs have been shown to contribute to tumorigenic processes, including proliferation, metastasis, and drug resistance in a

variety of cancers, mainly through the secretion of paracrine factors or cell-cell interactions (42, 43). Therefore, we used human BMSCs to isolate and obtain their exosomes to overexpress miR-139-5p as a transporter of miR-139-5p into breast cancer cells. The isolated products were confirmed to be exosomes and expressed miR-139-5p. Subsequently, they were co-cultured with MDA-MB-231 cells and found to inhibit the proliferation ability of MDA-MB-231 cells. In previous studies, miR-139-5p has been identified as a tumor suppressor in breast (44), colorectal (45) and esophageal squamous cell carcinoma (17), where it has been shown to inhibit migration, invasion, and metastasis. In this study, we found that miR-139-5p not only can be a potential biomarker in breast cancer, but also in different molecular subtypes of breast cancer. miR-139-5p may be a good candidate for further research in advanced preclinical studies as a therapeutic target, not only for breast cancer, but also for solid tumors.

In summary, this study revealed by comprehensive biological information analysis that miR-139-5p was lowly expressed in various kinds of breast cancer, and miR-139-5p was highly expressed in exosomes rooted in BMSCs. Bone marrow mesenchymal stem cells-derived exosomal miR-139-5p inhibited the proliferation of MDA-MB-231 cells, laying an experimental foundation for the clinical application of exosomal miR-139-5p gene in breast cancer treatment. In addition, the key target genes of miR-139-5p, FBN2, MEX3A, and TPD52 also provide potential targets for personalized treatment of different molecular subtypes of breast cancer.

## DATA AVAILABILITY STATEMENT

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

## AUTHORS CONTRIBUTION

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

## FUNDING

This work was supported by Ningbo Medical Science and Technology Program Project (2018A29), Ningbo Natural Science Foundation Project (2019A610302), Zhejiang Province Medical and Health Clinical Research Application Project (2022KY1173).

## REFERENCES

1. Barzaman K, Karami J, Zarei Z, Hosseinzadeh A, Kazemi MH, Moradi-Kalbolandi S, et al. Breast Cancer: Biology, Biomarkers, and Treatments. *Int Immunopharmacol* (2020) 84:1065355. doi: 10.1016/j.intimp.2020.106535
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: Globocan Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: Cancer J Clin* (2021) 71:209–495. doi: 10.3322/caac.21660
3. Heer E, Harper A, Escandor N, Sung H, McCormack V, Fidler-Benaoudia MM. Global Burden and Trends in Premenopausal and Postmenopausal Breast Cancer: A Population-Based Study. *Lancet Global Health* (2020) 8: e1027–375. doi: 10.1016/S2214-109X(20)30215-1
4. Joko-Fru WY, Jedy-Agba E, Korir A, Ogunbiyi O, Dzamalala CP, Chokunonga E, et al. The Evolving Epidemic of Breast Cancer in Sub-Saharan Africa: Results From the African Cancer Registry Network. *Int J Cancer* (2020) 147:2131–415. doi: 10.1002/ijc.33014
5. Yeo SK, Guan J-L. Breast Cancer: Multiple Subtypes Within a Tumor? *Trends Cancer* (2017) 3:753–605. doi: 10.1016/j.trecan.2017.09.001
6. Holm J, Eriksson L, Ploner A, Eriksson M, Rantalainen M, Li J, et al. Assessment of Breast Cancer Risk Factors Reveals Subtype Heterogeneity. *Cancer Res* (2017) 77:3708–175. doi: 10.1158/0008-5472.CAN-16-2574
7. Waks AG, Winer EP. Breast Cancer Treatment. *Jama* (2019) 321:316–165. doi: 10.1001/jama.2018.20751
8. Rangarajan B, Shet T, Wadasadawala T, Nair NS, Sairam RM, Hingmire SS, et al. Breast Cancer: An Overview of Published Indian Data. *South Asian J Cancer* (2016) 5:086–925. doi: 10.4103/2278-330X.187561
9. Ng CK, Schultheis AM, Bidard F-C, Weigelt B, Reis-Filho JS. Breast Cancer Genomics From Microarrays to Massively Parallel Sequencing: Paradigms and New Insights. *JNCI: J Natl Cancer Institute* (2015) 107(5):djv015. doi: 10.1093/jnci/djv015
10. Jönsson G, Staaf J, Vallon-Christersson J, Ringnér M, Holm K, Hegardt C, et al. Genomic Subtypes of Breast Cancer Identified by Array-Comparative Genomic Hybridization Display Distinct Molecular and Clinical Characteristics. *Breast Cancer Res* (2010) 12:1–145. doi: 10.1186/bcr2596
11. Ma CX, Ellis MJ. The Cancer Genome Atlas: Clinical Applications for Breast Cancer. *Oncology* (2013) 27:1263–635.
12. Soliman NA, Yussif SM. Ki-67 as a Prognostic Marker According to Breast Cancer Molecular Subtype. *Cancer Biol Med* (2016) 13:4965. doi: 10.20892/j.issn.2095-3941.2016.0066
13. Kim Y-J, Kim J-S, Kim IA. Molecular Subtype Predicts Incidence and Prognosis of Brain Metastasis From Breast Cancer in Seer Database. *J Cancer Res Clin Oncol* (2018) 144:1803–165. doi: 10.1007/s00432-018-2697-2
14. Abubakar M, Sung H, Bcr D, Guida J, Tang TS, Pfeiffer RM, et al. Breast Cancer Risk Factors, Survival and Recurrence, and Tumor Molecular Subtype: Analysis of 3012 Women From an Indigenous Asian Population. *Breast Cancer Res* (2018) 20:1–145. doi: 10.1186/s13058-018-1033-8
15. Lu TX, Rothenberg ME. Microrna. *J Allergy Clin Immunol* (2018) 141:1202–075. doi: 10.1016/j.jaci.2017.08.034
16. Zhang H-D, Jiang L-H, Sun D-W, Li J, Tang J-H. Mir-139-5p: Promising Biomarker for Cancer. *Tumor Biol* (2015) 36:1355–65. doi: 10.1007/s13277-015-3199-3
17. Liu R, Yang M, Meng Y, Liao J, Sheng J, Pu Y, et al. Tumor-Suppressive Function of Mir-139-5p in Esophageal Squamous Cell Carcinoma. *PloS One* (2013) 8:e770685. doi: 10.1371/journal.pone.0077068
18. Chen L, Lu F-B, Chen D-Z, Wu J-L, Xu L-M, Zheng M-H, et al. Bmscs-Derived Mir-223-Containing Exosomes Contribute to Liver Protection in Experimental Autoimmune Hepatitis. *Mol Immunol* (2018) 93:38–46. doi: 10.1016/j.molimm.2017.11.008
19. Sundararajan V, Sarkar FH, Ramasamy TS. The Versatile Role of Exosomes in Cancer Progression: Diagnostic and Therapeutic Implications. *Cell Oncol* (2018) 41:223–525. doi: 10.1007/s13402-018-0378-4
20. Lan F, Qing Q, Pan Q, Hu M, Yu H, Yue X. Serum Exosomal Mir-301a as a Potential Diagnostic and Prognostic Biomarker for Human Glioma. *Cell Oncol* (2018) 41:25–335. doi: 10.1007/s13402-017-0355-3
21. Ritchie W. microRNA Target Prediction. *Methods Mol Biol* (2017) 1513:193–200. doi: 10.1007/978-1-4939-6539-7_13
22. Januškevičienė I, Petrikaitė V. Heterogeneity of Breast Cancer: The Importance of Interaction Between Different Tumor Cell Populations. *Life Sci* (2019) 239:1170095. doi: 10.1016/j.lfs.2019.117009

23. Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists Guideline Recommendations for Immunohistochemical Testing of Estrogen and Progesterone Receptors in Breast Cancer (Unabridged Version). *Arch Pathol Lab Med* (2010) 134:e48–725. doi: 10.5858/134.7.e48
24. Austin TBM. First Microrna Mimic Enters Clinic. *Nat Biotechnol* (2013) 31:577. doi: 10.1038/nbt0713-577
25. Saliminejad K, Khorram Khorshid HR, Soleymani Fard S, Ghaffari SH. An Overview of Micrornas: Biology, Functions, Therapeutics, and Analysis Methods. *J Cell Physiol* (2019) 234:5451–655. doi: 10.1002/jcp.27486
26. Jia S, Zhai H, Zhao M. Micrornas Regulate Immune System *via* Multiple Targets. *Discovery Med* (2014) 18:237–475.
27. Sun D-W, Mao L, Zhang J, Jiang L-H, Li J, Wu Y, et al. Mir-139-5p Inhibits the Biological Function of Breast Cancer Cells by Targeting Notch1 and Mediates Chemosensitivity to Docetaxel. *Biochem Biophys Res Commun* (2015) 465:702–13. doi: 10.1016/j.bbrc.2015.08.053
28. Nistala H, Lee-Arteaga S, Smaldone S, Siciliano G, Carta L, Ono RN, et al. Fibrillin-1 and-2 Differentially Modulate Endogenous TGF-β and BMP Bioavailability During Bone Formation. *J Cell Biol* (2010) 190(6):1107–21. doi: 10.1083/jcb.201003089
29. Gu H-Y, Zhang C, Guo J, Yang M, Zhong H-C, Jin W, et al. Risk Score Based on Expression of Five Novel Genes Predicts Survival in Soft Tissue Sarcoma. *Aging (Albany NY)* (2020) 12:38075. doi: 10.18632/aging.102847
30. Lian X, Bond JS, Bharathy N, Boudko SP, Pokidysheva E, Shern JF, et al. Defining the Extracellular Matrix of Rhabdomyosarcoma. *Front Oncol* (2021) 11:385. doi: 10.3389/fonc.2021.601957
31. Chen H, Suzuki M, Nakamura Y, Ohira M, Ando S, Iida T, et al. Aberrant Methylation of Fbn2 in Human non-Small Cell Lung Cancer. *Lung Cancer* (2005) 50:43–495. doi: 10.1016/j.lungcan.2005.04.013
32. Tsunoda S, Smith E, De Young NJ, Wang X, Tian Z-Q, Liu J-F, et al. Methylation of Cldn6, Fbn2, Rbp1, Rbp4, Tfpi2, and Tmeff2 in Esophageal Squamous Cell Carcinoma. *Oncol Rep* (2009) 21:1067–735. doi: 10.3892/or_00000325
33. Kikuyama M, Takeshima H, Kinoshita T, Okochi-Takada E, Wakabayashi M, Akashi-Tanaka S, et al. Development of a Novel Approach, the Epigenome-Based Outlier Approach, to Identify Tumor-Suppressor Genes Silenced by Aberrant DNA Methylation. *Cancer Lett* (2012) 322:204–125. doi: 10.1016/j.canlet.2012.03.016
34. Yi JM, Dhir M, Guzzetta AA, Iacobuzio-Donahue CA, Heo K, Yang KM, et al. DNA Methylation Biomarker Candidates for Early Detection of Colon Cancer. *Tumor Biol* (2012) 33:363–725. doi: 10.1007/s13277-011-0302-2
35. Jiang S, Meng L, Chen X, Liu H, Zhang J, Chen F, et al. Mex3a Promotes Triple Negative Breast Cancer Proliferation and Migration *via* the Pi3k/Akt Signaling Pathway. *Exp Cell Res* (2020) 395:1121915. doi: 10.1016/j.yexcr.2020.112191
36. Shi X, Sun Y, Zhang Y, Wang W, Xu J, Guan Y, et al. Mex3a Promotes Development and Progression of Breast Cancer Through Regulation of Pik3ca. *Exp Cell Res* (2021) 404:1125805. doi: 10.1016/j.yexcr.2021.112580
37. Byrne JA, Maleki S, Hardy JR, Gloss BS, Murali R, Scurry JP, et al. Mal2 and Tumor Protein D52 (Tpd52) are Frequently Overexpressed in Ovarian Carcinoma, But Differentially Associated With Histological Subtype and Patient Outcome. *BMC Cancer* (2010) 10:1–115. doi: 10.1186/1471-2407-10-497
38. Ge J, Mao L, Xu W, Fang W, Wang N, Ye D, et al. Mir-103a-3p Suppresses Cell Proliferation and Invasion by Targeting Tumor Protein D52 in Prostate Cancer. *J Invest Surg* (2021) 34:984–25. doi: 10.1080/08941939.2020.1738602
39. Zhang Z, Wang J, Gao R, Yang X, Zhang Y, Li J, et al. Downregulation of Microrna-449 Promotes Migration and Invasion of Breast Cancer Cells by Targeting Tumor Protein D52 (Tpd52). *Oncol Res* (2017) 25:7535. doi: 10.3727/096504016X14772342320617
40. Li J, Li Y, Liu H, Liu Y, Cui B. The Four-Transmembrane Protein Mal2 and Tumor Protein D52 (Tpd52) are Highly Expressed in Colorectal Cancer and Correlated With Poor Prognosis. *PloS One* (2017) 12:e01785155. doi: 10.1371/journal.pone.0178515
41. Dasari C, Reddy KRK, Natani S, Murthy T, Bhukya S, Ummanni R. Tumor Protein D52 (Isoform 3) Interacts With and Promotes Peroxidase Activity of Peroxiredoxin 1 in Prostate Cancer Cells Implicated in Cell Growth and Migration. *Biochim Biophys Acta (BBA) Mol Cell Res* (1866) 2019):1298–309. doi: 10.1016/j.bbamcr.2019.04.007
42. Ridge SM, Sullivan FJ, Glynn SA. Mesenchymal Stem Cells: Key Players in Cancer Progression. *Mol Cancer* (2017) 16:1–105. doi: 10.1186/s12943-017-0597-8
43. Shi Y, Du L, Lin L, Wang Y. Tumour-Associated Mesenchymal Stem/Stromal Cells: Emerging Therapeutic Targets. *Nat Rev Drug Discov* (2017) 16:35–525. doi: 10.1038/nrd.2016.193
44. Gu S, Luo J, Yao W. The Regulation of Mir-139-5p on the Biological Characteristics of Breast Cancer Cells by Targeting Col11a1. *Math Biosci Eng* (2020) 17:1428–415. doi: 10.3934/mbe.2020073
45. Song M, Yin Y, Zhang J, Zhang B, Bian Z, Quan C, et al. Mir-139-5p Inhibits Migration and Invasion of Colorectal Cancer by Downregulating Amfr and Notch1. *Protein Cell* (2014) 5:851–615. doi: 10.1007/s13238-014-0093-5


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 06 April 2022
doi: 10.3389/fonc.2022.860961

# LncRNA-AC02278.4 Is a Novel Prognostic Biomarker That Promotes Tumor Growth and Metastasis in Lung Adenocarcinoma

*Xin Chen[1†], Fan Zhou[2†], Wenjun Ren[3†], Jishu Guo[4†], Xiaobin Huang[1], Jun Pu[1*], Xiaoqun Niu[5*] and Xiulin Jiang[3*]*

[1] Department of Neurosurgery, The second Affiliated Hospital of Kunming Medical University, Kunming, China, [2] Hematology and Rheumatology Department, The Pu'er People's Hospital, Pu'er, China, [3] Department of Cardiovascular Surgery, The First People's Hospital of Yunnan Province, Kunming, China, [4] Institute for Ecological Research and Pollution Control of Plateau Lakes, School of Ecology and Environmental Sciences, Yunnan University, Kunming, China, [5] Department of Respiratory Medicine, Second Hospital of Kunming Medical University, Kunming, China, [6] Kunming College of Life Science, University of Chinese Academy of Sciences, Beijing, China

OPEN ACCESS

**Edited by:**
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

**Reviewed by:**
Dulari Jayarathna, Queensland University of Technology, Australia
Jennifer H. Gunter, Queensland University of Technology, Australia

***Correspondence:**
Xiulin Jiang
jiangxiulin@mail.kiz.ac.cn
Xiaoqun Niu
673409256@qq.com
Jun Pu
pujun303@aliyun.com

[†]These authors have contributed equally to this work

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

**Received:** 24 January 2022
**Accepted:** 08 March 2022
**Published:** 06 April 2022

**Citation:**
Chen X, Zhou F, Ren W, Guo J, Huang X, Pu J, Niu X and Jiang X (2022) LncRNA-AC02278.4 Is a Novel Prognostic Biomarker That Promotes Tumor Growth and Metastasis in Lung Adenocarcinoma. Front. Oncol. 12:860961. doi: 10.3389/fonc.2022.860961

LncRNA-AC02278.4 (ENSG00000248538) is a long non-coding RNA (lncRNA) found to be highly expressed in multiple human cancers including lung adenocarcinoma (LUAD). However, the underlying biological function and potential mechanisms of AC02278.4 driving the progression of LUAD remain unclear. In this study, we investigated the role of AC02278.4 in LUAD and found that AC02278.4 expression was significantly increased in datasets extracted from The Cancer Genome Atlas. Increased expression of lncRNA-AC02278.4 was correlated with advanced clinical parameters. Receiver operating characteristic (ROC) curve analysis revealed the significant diagnostic ability of AC02278.4 [area under the ROC curve (AUC) = 0.882]. In addition, gene set enrichment analysis (GSEA) enrichment showed that AC02278.4 expression was correlated with immune response-related signaling pathways. Finally, we determined that AC02278.4 regulated cell proliferation and migration of LUAD *in vitro*. Our clinical sample results also confirmed that AC02278.4 was highly expressed in LUAD and correlated with adverse clinical outcomes. In conclusion, our data demonstrated that AC02278.4 was correlated with progression and immune infiltration and could serve as a prognostic biomarker for LUAD.

**Keywords: lncRNA-AC02278.4, lung adenocarcinoma, prognosis, biomarker, cell proliferation, cell migration**

## INTRODUCTION

Lung cancer is the leading cause of cancer-related death worldwide, according to cancer statistics 2020. The incidence rate of lung cancer ranks second, while the death rate of lung cancer ranks first (1). Lung cancer includes small cell lung carcinoma (SCLC) and non-SCLC (NSCLC). NSCLC includes lung adenocarcinoma (ADC) (LUAD), lung squamous cell carcinoma (SCC), and large cell lung carcinoma. NSCLC accounts for approximately 85% of all cases (2). Despite various treatments

being applied during diagnostic and therapeutic procedures for lung cancer, the 5-year survival rate of lung cancer still remains poor (3). Therefore, elucidating the molecular mechanisms of lung oncogenesis and identifying new therapeutic targets or biomarkers are essential for effectively preventing the development of lung cancer.

The initiation and progression of lung cancer are very complicated processes that involve genetic mutations, tumor microenvironment, and the abnormal activation of epigenetic modification (4–6). Simultaneously, epigenetic changes in lung cancer such as histone modifications (7), DNA methylation (8), and non-coding RNAs (ncRNAs) (9) have been far and wide studied. Long ncRNAs (lncRNAs) are a class of RNA molecules longer than 200 nucleotides in length with considerable potential to be involved in cancer development (10, 11). Emerging evidence has confirmed that lncRNAs participated in modulating gene expression *via* cis or trans manner (12).

Various regulatory mechanisms for an lncRNA have been well established. LncRNAs can 1) affect downstream gene expression *via* inhibiting the activity of RNA polymerase II or affect chromatin remodeling and histone modification state; 2) modulate the mRNA splicing process *via* complementary binding with pre-mRNAs; 3) interact with protein and regulate protein activity; 4) function as scaffolds to promote RNA–protein complexes form; 5) modulate the subcellular localization of proteins; and 6) function as miRNA sponge to regulate gene expression (13). Moreover, accumulating evidence has confirmed that lncRNAs correlate with the pathogenesis of lung cancers through regulating the proliferation and invasion, cell cycle, cell autophagy, cell apoptosis, stemness of lung cancer stem cell, chemotherapy resistance, and tumor microenvironment (14). For example, Gong et al. found that lncRNA JPX was highly expressed in lung cancer and correlated with the tumor size and an advanced stage. Forced expression of JPX facilitated lung cancer cell proliferation *in vitro* and facilitated lung tumor growth *in vivo* (15). Lu et al. reported that lnc-IGFBP4-1 was overexpressed in lung cancer tissues and that its higher expression was associated with TNM stage and lymph node metastasis. Depletion of lnc-IGFBP4-1 significantly inhibited cell proliferation and induced apoptosis. Further research showed that lnc-IGFBP4-1 *via* affecting the expression of HK2, PDK1, and LDHA led to enhancing the ATP production level and is involved in lung cancer progression (16). Furthermore, lncRNA AFAP1-AS1 was found to modulated NSCLC cell proliferation *via* interacting with EZH2 and recruiting EZH2 to the promoter regions of p21, thus inhibiting p21 expression (17). However, the clinical value of lncRNA-AC02278.4 in LUAD has not been explored. Hence, this study aimed to investigate the role of lncRNA-AC02278.4 in the progression of LUAD.

In this study, we compared the expression of AC02278.4 between LUAD tissues and normal samples and explored the correlation between AC02278.4 expression and the clinical significance of LUAD. Furthermore, we explored the prognostic and diagnostic value of AC02278.4 in LUAD. Meanwhile, the correlation between AC02278.4 expression and immune infiltration was analyzed to explore the potential mechanisms involved in AC02278.4 modulation in the progression of LUAD. Finally, the biological role of AC02278.4 was identified in LUAD. In summary, we demonstrated the potential role of AC02278.4 in regulating tumor progression and its potential application in the diagnosis and prognostic evaluation in LUAD.

# MATERIALS AND METHODS

## Data Collection

TCGA-LUAD cohort data and corresponding clinical information of 535 LUAD patients were downloaded from The Cancer Genome Atlas (TCGA) website (https://portal.gdc.cancer.gov/repository). LUAD patients were classified into low- and high-AC02278.4 expression groups according to the median AC02278.4 expression value. AC02278.4 expression data from GSE81089 datasets were downloaded from the Gene Expression Omnibus (GEO) database and validated for expression analyses. The gene expression profiles were normalized using the scale method provided in the "limma" R package. Data analysis was performed with the R (version 3.6.3) and ggplot2 [3.3.3] packages. The expression data were normalized to transcripts per kilobase million (TPM) values before further analysis. Besides, the receiver operating characteristic (ROC) curve was used to evaluate the diagnostic value of AC02278.4 using the pROC R package and ggplot2 R package.

## Nomogram Construction and Evaluation

Based on the multivariate Cox analysis results, a nomogram was established to predict the prognosis of LUAD patients. According to the prognosis model, each patient's risk score was calculated as the total score of each parameter, which could predict the prognosis of LUAD patients (18).

## Gene Set Enrichment Analysis

The gene set enrichment analysis (GSEA) software was utilized to analyze the potential signaling pathway and molecular function in LUAD (19, 20). A customized Perl script was used to perform GSEA between high-AC02278.4 and low-AC02278.4 groups. According to the default statistical methods, an adjusted $p$-value <0.05 was considered significant.

## Immune Infiltration Analysis by Single-Sample Gene Set Enrichment Analysis

A GSVA R package was used to examine the LUAD immune infiltration of 24 tumor-infiltrating immune cells in tumor samples through single-sample gene set enrichment analysis (ssGSEA) (21, 22). The correlation between AC02278.4 and infiltration levels of immune cells was analyzed by Spearman's correlation, and these immune cells with the different expression groups of AC02278.4 were analyzed by rank-sum test.

## Cell Culture

BEAS-2B cell line was purchased from Cell Bank of Kunming Institute of Zoology and cultured in BEGM media (Lonza, Basel,

Switzerland; CC-3170). Lung cancer cell lines, including A549, H1299, and SPC-A1, were purchased from Cobioer (Nanjing, China) with short tandem repeat (STR) documents, and were cultured in RPMI-1640 medium (Corning, Manassas, VA, USA) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin.

## Constructs, Lenti-Viral Preparation, and Establishment of Different Cell Lines

For shRNA knockdown experiments, independent shRNAs targeting a different region of AC02278.4 RNA were constructed using a pLKO.1 vector (Addgene, Cambridge, MA, USA), and the oligo sequences were provided as follows. Lentiviruses were generated according to the manufacturer's protocol as previously documented (5), and indicated cells were infected by viruses twice with 48- and 72-h viral supernatants containing 4 μg/ml of polybrene, and stable cell lines were established by appropriate puromycin selection. The two independent AC02278.47 targeting sequences are as follows: shRNA#1, 5′-GGCACTTCGTGGCTGAACCGA-3′; shRNA#2, 5′-GGGGAACAATGGCTTCAGCAG-3′.

## Cell Proliferation Assays

For BrdU incorporation assay, indicated cells were cultured in 8-well chamber slides for 24 h, pretreated with or without SC79 (Beyotime, Shanghai, China; SF2730) for another 24 h and then treated with 10 μM of BrdU (Abcam, Cambridge, UK; ab142567) for 20 min. Subsequently, indicated cells were fixed with 4% paraformaldehyde (PFA) at room temperature for 20 min and then incubated with BrdU primary antibody (Abcam, ab6326) followed by secondary antibody detection. The cell nuclei were stained with DAPI as previously documented (5).

## Cell Migration Assays

Cell migration assay was performed as previously described (15). Briefly, indicated cells were seeded into 6-well plates ($9 \times 10^5$/cell) and incubated for 1 day, and then a straight line was scraped with pipette tips. Detached cells were removed. Photographs were taken at the indicated time, and the relative traveled distance was measured. For the trans-well migration assay, $2.5 \times 10^4$ cells/well in 100 μl of serum-free medium were plated in a 24-well plate chamber insert, and the lower chamber was filled with 10% FBS. After incubation for 24 h, cells were fixed with 4% PFA, washed, and then stained with 0.5% crystal violet for further imaging.

## Real-time RT-PCR assay

In the real-time RT-PCR assay, cells were lysed by RNAiso Plus (Takara Bio, Beijing, China; Cat. 108-95-2). Total RNAs were extracted according to the manufacturer's protocol and then reverse transcribed by using RT reagent Kit (Takara Bio, Beijing, China). The primer used in this study is as follows: β-actin-F: AAG TGTGACGTGGACATCCGC, β-actin-R: CCGGACTCGTCA TACTCCTGCT, AC02278.4-F: TCAAGGCATCATGTGTCATT, AC02278.4-R: ACCTGCTGAAGCCATTGTTCC.

## Statistical Analysis

For the datasets from TCGA database, statistical analyses were performed using R. The Wilcoxon rank-sum test and chi-square test were used to estimate the association between AC02278.4 and clinical pathologic characteristics. The Kaplan–Meier method was used to calculated LUAD patient survival rates. Univariate and multivariate Cox analyses were performed to assess the correlation between clinical features and overall survival (OS), disease-free survival, and progression-free survival (PFS). For the data regarding the function of AC02278.4, Graph Pad Prism 7.0 was used for statistical analyses. Student's *t*-test evaluated the statistical significance between groups. The significance of the data between the two experimental groups was determined by Student's *t*-test, and multiple group comparisons were analyzed by one-way ANOVA. $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***), were significant.

# RESULTS

## AC02278.4 Was Overexpressed in Lung Adenocarcinoma

We simultaneously analyzed the expression profiles of AC02278.4 (ENSG00000248538) based on TCGA database. The results confirmed that AC02278.4 was highly expressed in multiple cancer types compared to the normal samples (**Figure 1A**). To examine the expression of AC02278.4 in LUAD and normal samples, we analyzed the expression of AC02278.4 in 535 tumor tissues and 59 normal prostate tissues of TCGA data, and we uncovered that AC02278.4 was upregulated in LUAD tissues (**Figure 1B**). There were 59 pairs of LUAD cancer samples and matched adjacent normal samples in TCGA data. The expression level of AC02278.4 was also higher in LUAD samples than matched adjacent normal samples (**Figure 1C**). Moreover, we found that AC02278.4 increased in lung cancer tissues by analyzing the GEO dataset (**Figure 1D**). To validate the expression of AC02278.4 in LUAD, we first employed qRT-PCR assay to detect the expression of AC02278.4 in 20 pairs of LUAD tissues and matched adjacent normal tissues. It was identified that AC02278.4 was significantly upregulated in LUAD tissue than adjacent normal tissue (**Figure 1E**).

## Overexpression of AC02278.4 Was Associated With Adverse Clinical Parameters in Lung Adenocarcinoma

To examine the clinical relevance of AC02278.4 in LUAD, 535 LUAD patients with clinical parameters were classified into two subgroups based on the mean value of relative AC02278.4 expression. We then explored the correlations between AC02278.4 expression and clinical parameters, including pathologic stage, TNM stage, and residual tumor. Regarding the tumor pathological stage, a significant increase in AC02278.4 expression was observed in LUAD patients in stages 1, 2, 3, and 4 (**Figure 2A**). Based on the cancer stage, AC02278.4 expression was higher in patients with LUAD classified as T stage (T1, T2,

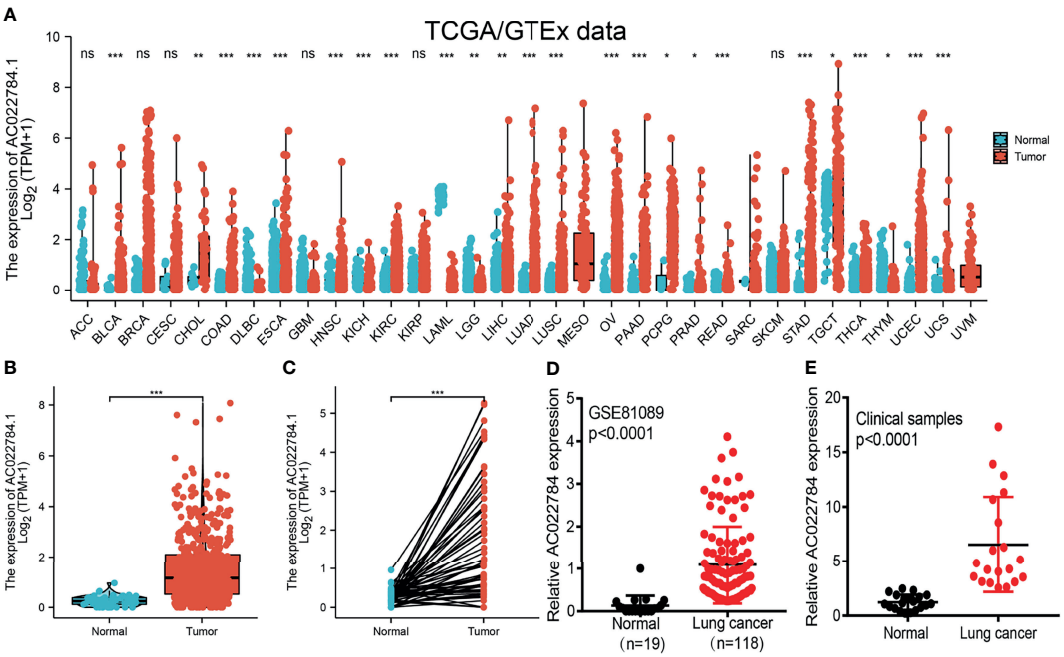



**FIGURE 1** | lncRNA-AC02278.4 was highly expressed in lung adenocarcinoma. **(A)** Expression of lncRNA-AC02278.4 in diverse human cancer based on the TCGA and GTEx datasets. **(B)** Expression of lncRNA-AC02278.4 in lung cancer based on TCGA dataset. **(C)** Expression levels of lncRNA-AC02278.4 in 59 paired adjacent normal tissues and paired samples. **(D)** Expression levels of lncRNA-AC02278.4 in lung cancer based on the GEO dataset. **(E)** Relative lncRNA-AC02278.4 expression detected by RT-qPCR in 20 paired lung cancer and non-cancerous tissues. ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower-grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumors; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma; TCGA, The Cancer Genome Atlas; GTEx, Genotype-Tissue Expression; GEO, Gene Expression Omnibus. NS, $p > 0.05$,*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

and T3), N stage (N0, N1, and N2), and M stage (M0 and M1) (**Figures 2B–D**). We also found that AC02278.4 expression was significantly correlated with the residual tumor stage (**Figure 2E**). Furthermore, ROC analysis showed that the AC02278.4 could be used to differentiate LUAD patients from normal control with a specificity (AUC = 0.882) (**Figure 2F**). Additionally, Kaplan–Meier analysis showed that LUAD patients with higher AC02278.4 expression were associated with poor OS, disease-free survival, and PFS (**Figures 2G–I**). To further validate the correlation between AC02278.4 expression and OS, we examined the prognosis of AC02278.4 in lung cancer by clinical samples from Yunnan Cancer Hospital (N = 181). The results also showed that high AC02278.4 expression had a worse OS than low AC02278.4 expression (**Figure 2J**).

## Univariate and Multivariate Cox Regression Analyses of Different Parameters on Overall Survival

We conducted a univariate Cox regression analysis in the TCGA-LAUD cohort to determine whether AC02278.4 expression level and pathologic stage might be valuable prognostic biomarkers. In the univariate Cox regression analysis, high expression of AC02278.4, pathologic stage, and TNM stage were associated with OS in LUAD patients. To ascertain whether AC02278.4 expression level could be an independent prognostic factor for patients with LUAD, a multivariate Cox regression analysis was performed. We confirmed that increased AC02278.4 expression was a significant independent prognostic factor in the TCGA-LAUD cohort that directly correlated with adverse clinical outcomes, along with T stage (**Table 1**).

## Construction and Validation of AC02278.4-Based Nomogram

The multivariate analysis result confirmed that AC02278.4 is an independent prognostic factor in LUAD. We then constructed a prediction model for OS, disease-free survival, and PFS by integrating AC02278.4 expression and pathologic stage. We established a nomogram to integrate AC02278.4 as a LUAD biomarker; higher total points on the nomogram for OS, progression-free interval (PFI), and disease-specific survival (DSS) indicated a worse prognosis (**Figures 3A–E**). A higher point on the nomogram represented a worse prognostic factor.

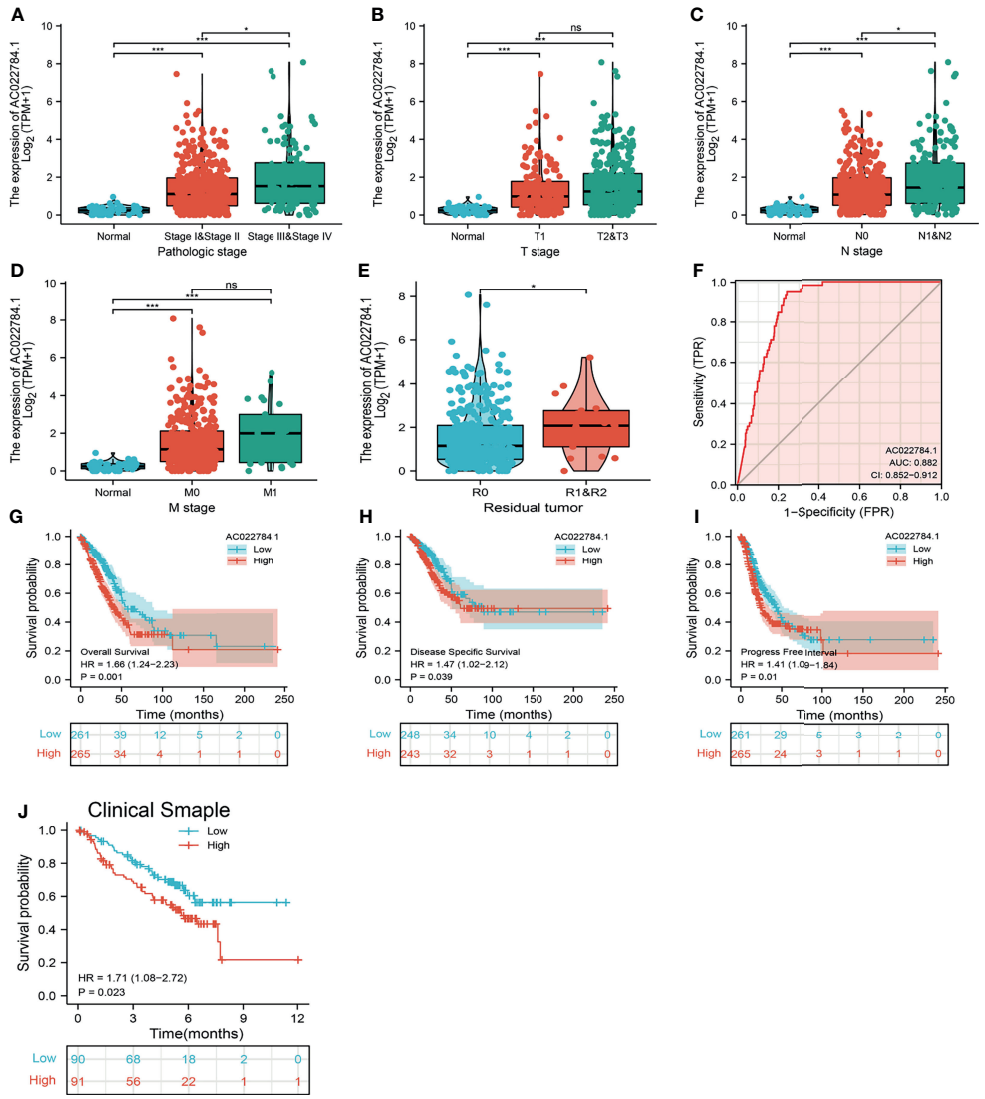



**FIGURE 2** | Clinical significance of lncRNA-AC02278.4 in lung adenocarcinoma. Correlation between lncRNA-AC02278.4 expression and clinical parameters, including **(A)** pathological stage, **(B–D)** TNM stage, and **(E)** residual tumor. **(F)** ROC curves were used to determine the diagnostic value of lncRNA-AC02278.4 in lung adenocarcinoma. **(G–J)** Kaplan–Meier survival curves showed that lung adenocarcinoma patients with high lncRNA-AC02278.4 expression exhibited poor overall survival, disease-specific survival, and progression-free survival of AC02278.4 in LUAD as determined by TCGA-LUAD dataset and clinical samples. NS, $p > 0.05$, $^{*}p < 0.05$, $^{***}p < 0.001$. ROC, receiver operating characteristic; LUAD, lung adenocarcinoma.

Based on the calibration curve of nomograms for OS, DSS, and PFS, the predictions conformed well to observations in all patients, and the test showed no deviation from the perfect fit. The nomogram had a C-index of 0.657 and contained 1,000 bootstrap replicates [95% CI: (0.634–0.679)] for OS. We also found DSS [C-index: 0.594, CI: (0.573–0.615)] and PFS (C-index: 0.645, CI: 0.617–0.673). It was found that the bias-corrected line in the calibration plot was close to the ideal curve, indicating a strong correlation between predicted values and observed values (**Figures 3D–F**). In summary, these results indicated that the nomogram can well predict short- or long-term survival of LUAD patients. We also used the GEO dataset to validate the above result, as is shown in **Figures 3G, H**; using the GSE31552 dataset, we found that this nomogram could well predict the OS of LUAD patients.

## AC02278.4-Related Signaling Pathways Based on Gene Set Enrichment Analysis

GSEA was used to identify potential signaling pathways that are activated by high AC02278.4 expression. As shown in **Figures 4A, B**, there were eight significant signaling pathways related to the high AC02278.4 expression phenotype, including the focal adhesion, epithelial cell signaling in *Helicobacter pylori* infection, Wnt signaling pathway, VEGF signaling pathway, toll-

**TABLE 1** | Univariate and multivariate Cox regression analyses of different parameters on overall survival in lung adenocarcinoma.

| Characteristics | Total (N) | Univariate analysis | | Multivariate analysis | |
|---|---|---|---|---|---|
| | | Hazard ratio (95% CI) | *p*-Value | Hazard ratio (95% CI) | *p*-Value |
| T stage | 523 | | | | |
| T1 and T2 | 457 | | | | |
| T3 and T4 | 66 | 2.317 (1.591–3.375) | <0.001 | 1.638 (1.018–2.635) | 0.042 |
| N stage | 510 | | | | |
| N0 and N1 | 437 | | | | |
| N3 and N2 | 73 | 2.321 (1.631–3.303) | <0.001 | 1.293 (0.626–2.674) | 0.488 |
| Pathologic stage | 518 | | | | |
| Stage II and stage I | 411 | | | | |
| Stage IV and stage III | 107 | 2.664 (1.960–3.621) | <0.001 | 1.802 (0.839–3.871) | 0.131 |
| M stage | 377 | | | | |
| M0 | 352 | | | | |
| M1 | 25 | 2.136 (1.248–3.653) | 0.006 | 1.192 (0.541–2.626) | 0.664 |
| AC022784 1 | 526 | 1.251 (1.141–1.372) | <0.001 | 1.168 (1.053–1.296) | 0.003 |

like receptor signaling pathway, T-cell receptor signaling pathway, B-cell receptor signaling pathway, and cytokine–cytokine receptor interaction (**Figures 4A, B**).

## Correlation Between AC02278.4 Expression and Immune Infiltration

Using the ssGSEA method, we explored the correlation between AC02278.4 expression and infiltrating immune cells in LUAD. These results confirmed that AC02278.4 expression was negatively associated with infiltration levels of Th17 cells, B cells, CD8 T cells, eosinophils, cytotoxic cells, Tem, Th1 cells, macrophages, T cells, mast cells, Tcm, DC, TFH, T helper cells, and iDC ($p < 0.001$) and was positively correlated with that of Th2 cells, Tgd, and NK CD56bright cells (**Figure 5A**). Furthermore, we found that patients with AC02278.4 high-expression group showed a reduction in the numbers of infiltrating aDC, B cells, CD8 T cells, cytotoxic cells, DC, eosinophils, iDC, macrophages, T helper cells, T cells, mast cells, Tcm, Tem, TFH, and Th1 cells (**Figure 5B**).

## AC02278.4 Promotes Proliferation, Migration, and Invasion of Lung Adenocarcinoma Cells *In Vitro*

The above studies indicated that AC02278.4 expression was distinctly upregulated in LUAD tissues, and AC02278.4 might influence the progression in LUAD. To further investigate the biological role of AC02278.4 in LUAD, we first confirmed that the expression of AC02278.4 was significantly upregulated in H1299, A549, and SPC-A1 lung cancer cell lines compared to the human bronchial epithelial cells (BEAS2B) (**Figure 6A**). Furthermore, specific siRNA for AC02278.4 was used to construct A549 and SPC-A1 cells with stable knockdown of AC02278.4 expression. The knockdown efficiencies in transformed cell lines were detected by qRT-PCR analysis (**Figures 6B, C**). It was confirmed that knockdown of AC02278.4 reduced the proliferative capacity of A549 and SPC-A1 cells (**Figures 6D, E**) upon BrdU assays. Moreover, transwell assay and wound healing revealed that the migration and invasion abilities of A549 and SPC-A1 cells were significantly inhibited through downregulating AC02278.4 expression level (**Figures 6F–I**). Using various bioinformatics tools, we also constructed an mRNA–miRNA–lncRNA network; further study should be conducted to verify this result (**Figure 6J**). Briefly, it was suggested that AC02278.4 plays the role of an oncogene in lung cancer cells.

# DISCUSSION

Recent studies consistently report that lncRNAs can modulate various hallmarks of lung cancer, including cancer cell proliferation, metabolic reprogramming, angiogenesis, cancer cell invasion and metastasis, and immunosurveillance (23, 24). It has been shown that lncRNAs may be effective and specific molecular markers for lung cancer diagnosis. For example, Tan et al. found that Linc00152 has an AUC value of 0.742 and may serve as a diagnostic marker distinguishing NSCLC (25). Furthermore, univariate and multivariate analyses as standard and reliable statistical methods are utilized to confirm whether an lncRNA can be regarded as an independent tumor marker for predicting the prognosis of lung cancer patients. By univariate and multivariate analyses, some lncRNAs have also been identified as independent prognostic markers in lung cancer. For instance, EGFR−AS1 was shown to be increased in NSCLC and correlated with a poor prognosis. With forced EGFR−AS1 expression promoting NSCLC cell proliferation and chemoresistance *via* regulating miR-223/IGF1R axis, univariate and multivariate analyses reported EGFR−AS1 as an independent prognostic marker in lung cancer (26).

Combining bioinformatics analyses, we found that AC02278.4 was highly expressed in LUAD, which was also correlated with OS, disease-free survival, and PFS in the LUAD patients of TCGA data. Our clinical sample analysis demonstrated that AC02278.4 was overexpressed in lung cancer. We also verified that increased AC02278.4 expression was a significant independent prognostic factor in the TCGA-LUAD cohort that directly correlated with adverse clinical outcomes. We also established a nomogram to integrate AC02278.4 as a LUAD biomarker; higher total points on the nomogram for OS, PFI, and DSS indicated a worse prognosis.

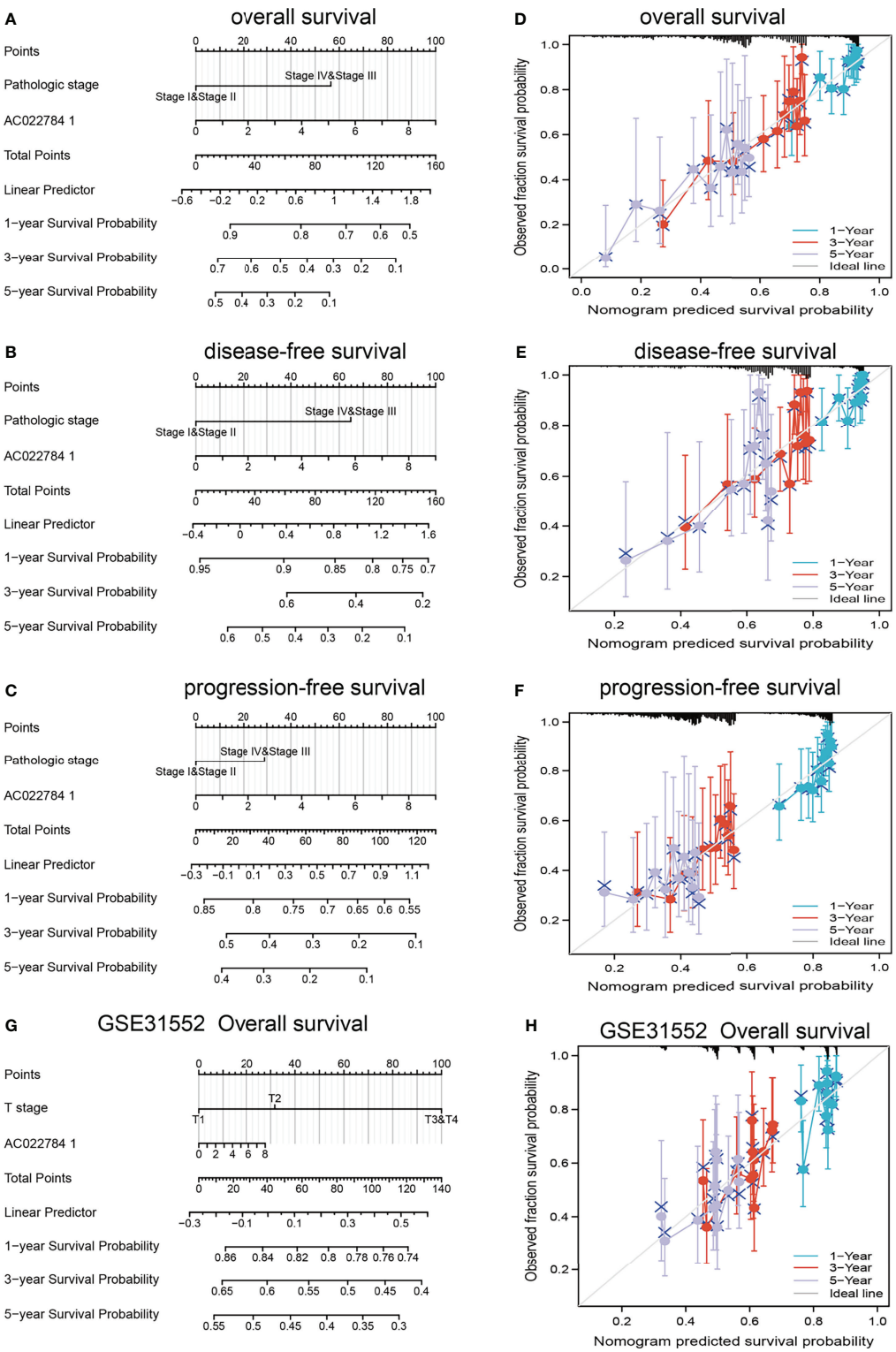



**FIGURE 3** | Construction and performance validation of the lncRNA-AC02278.4-based nomogram for lung adenocarcinoma patients. Nomogram to predict the **(A)** overall survival, **(B)** disease-specific survival, and **(C)** progression-free survival for lung cancer patients. The calibration curve and Hosmer–Lemeshow test of nomograms in the TCGA-lung adenocarcinoma cohort for **(D)** overall survival, **(E)** disease-specific survival, and **(F)** progression-free survival. **(G, H)** Using GEO dataset to validation of the lncRNA-AC02278.4-based nomogram for lung adenocarcinoma patients. GEO, Gene Expression Omnibus.

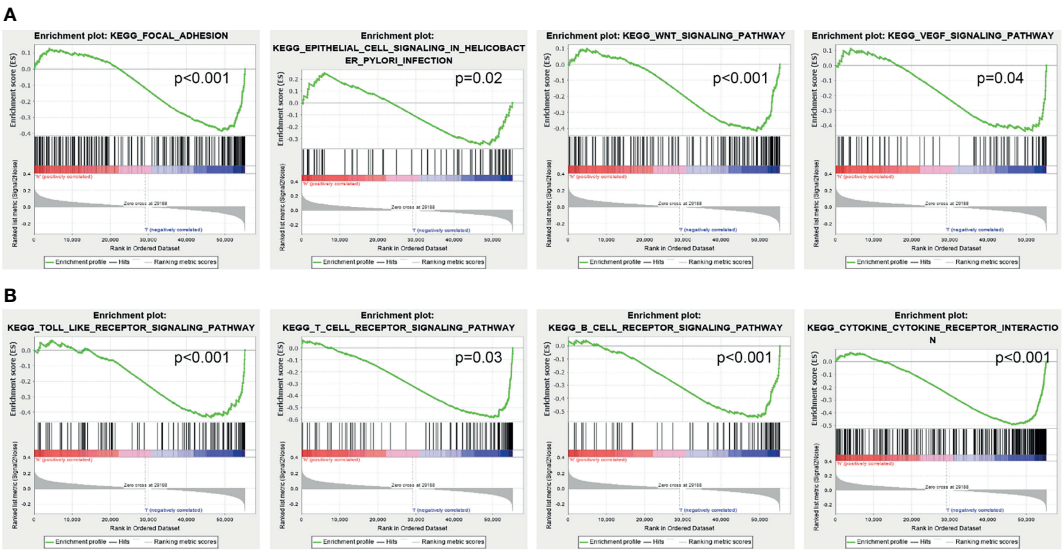



**FIGURE 4** | Identification of lncRNA-AC02278.4-related signaling pathways in lung adenocarcinoma. **(A, B)** Identification of lncRNA-AC02278.4-related signaling pathways by GSEA software. GSEA, gene set enrichment analysis.

In this study, we investigated the underlying mechanisms through which AC02278.4 affected the progression of LUAD. GSEA enrichment confirmed that AC02278.4 was significantly associated with focal adhesion, epithelial cell signaling in *H. pylori* infection, Wnt signaling pathway, VEGF signaling pathway, toll-like receptor signaling pathway, T-cell receptor signaling pathway, B-cell receptor signaling pathway, and cytokine–cytokine receptor interaction, which indicated that AC02278.4 might have a crucial role in immune-response regulation and cell proliferation. Furthermore, we also explored the correlation between AC02278.4 expression and diverse immune infiltration levels in LUAD.

In this finding, we found that AC02278.4 expression was negatively associated with infiltration levels of Th17 cells, B cells, CD8 T cells, eosinophils, cytotoxic cells, Tem, Th1 cells, macrophages, T cells, mast cells, Tcm, DC, TFH, T helper cells, and iDC ($p < 0.001$) and was positively correlated with those of Th2 cells, Tgd, and NK CD56bright cells and thus plays an

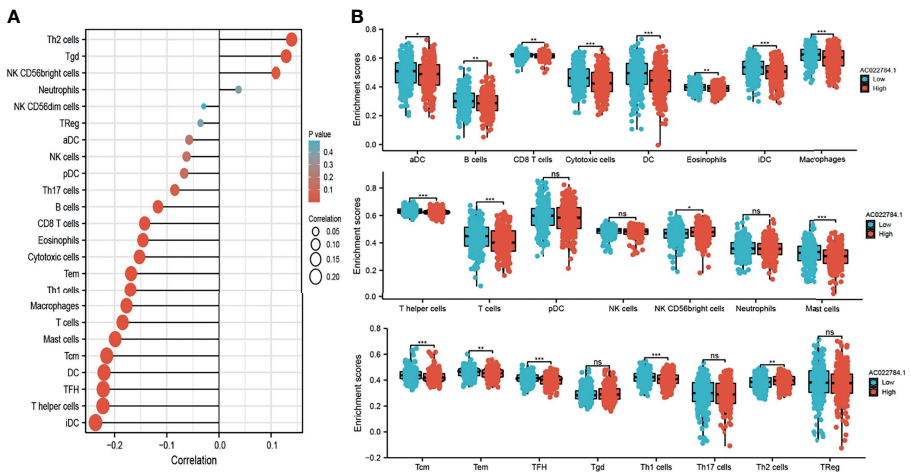



**FIGURE 5** | Correlation analysis of lncRNA-AC02278.4 expression and infiltration levels of immune cells in LUAD tissues. **(A)** Correlation between the relative abundances of 24 immune cells and lncRNA-AC02278.4 expression level. **(B)** Box plots of the correlations between lncRNA-AC02278.4 or molecular model expression and infiltration levels of immune cells. NS, $p > 0.05$,*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. LUAD, lung adenocarcinoma.

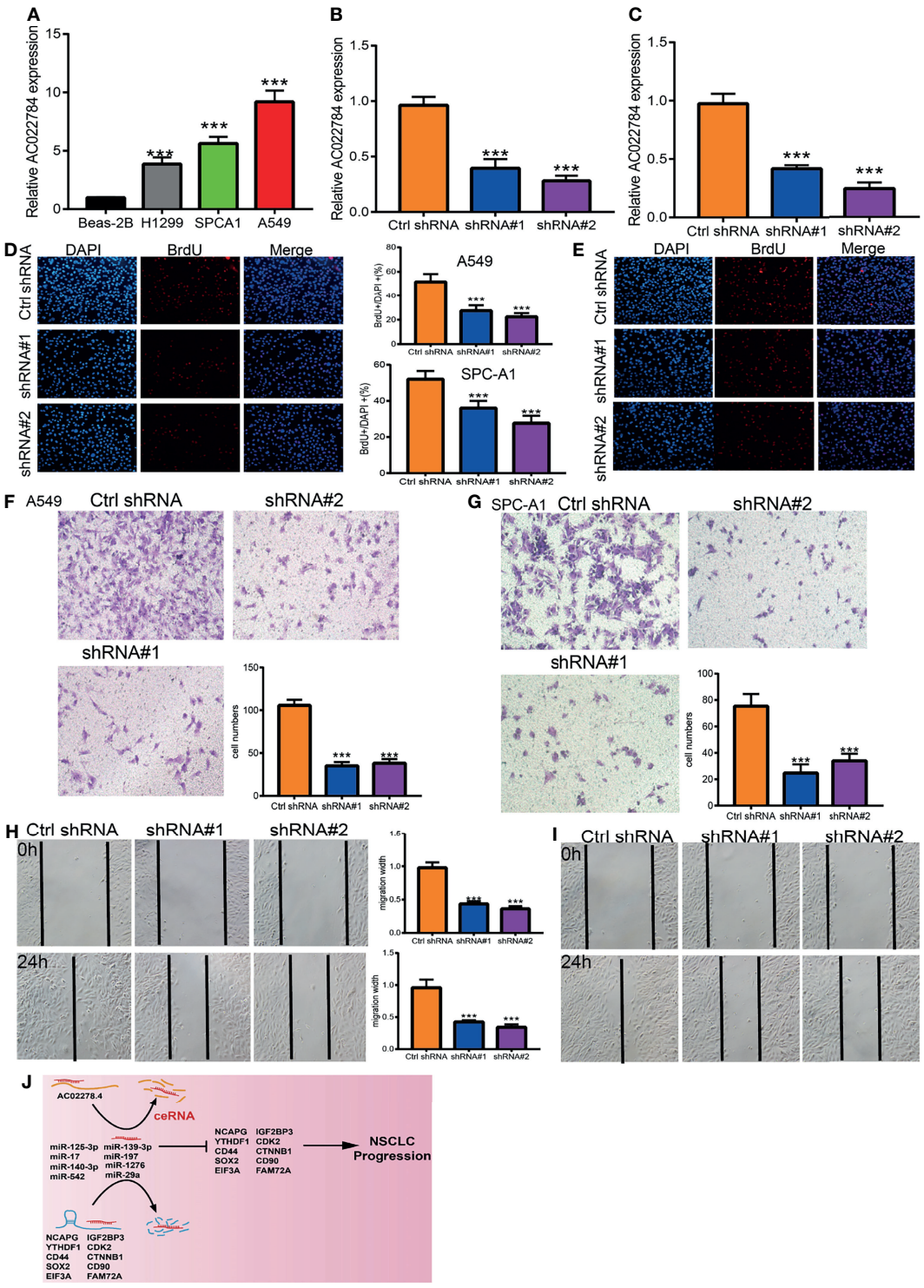



**FIGURE 6** | LncRNA-AC02278.4 promotes LUAD cell proliferation, migration, and invasion *in vitro*. **(A)** The relative expression level of lncRNA-AC02278.4 in lung adenocarcinoma cancerous cell lines, including H1299, SPCA1, and A549 examined by real-time RT-PCR, compared to normal human bronchial epithelial cell line: BEAS-2B. **(B, C)** Establishment of lncRNA-AC02278.4 knockdown cell lines in A549 and H1299 verified by real-time RT-PCR. **(D, E)** Knockdown of lncRNA-AC02278.4 significantly inhibits cell proliferation in A549 and H1299 cells, as measured with a BrdU Alexa Fluor Imaging Kit. Scale bar: 50 μm. **(F–I)** Knockdown of lncRNA-AC02278.4 dramatically inhibits A549 and H1299 cell migration's ability as examined by transwell and wound healing assays. **(J)** The network of lncRNA–miRNA–mRNA. ***$p < 0.001$. LUAD, lung adenocarcinoma.

oncogenic role in LUAD. *In vitro*, depletion of AC02278.4 in A549 and SPC-A1 cells inhibited cell proliferation and migration. Based on the above findings, we proposed that AC02278.4 exerts an essential function in regulating the pathologic progression of LUAD.

This study improves our understanding of the correlation between lncRNA-AC02278.4 and LUAD, but some limitations still exist. First, although we explored the correlation between AC02278.4 and immune infiltration in LUAD patients, there is a lack of experiments to validate the function of AC02278.4 in the tumor microenvironment regulation of LUAD. Second, we uncover that depletion of AC02278.4 inhibits cell proliferation and cell migration of LUAD cells. However, the potential molecular mechanisms of AC02278.4 in cancer progression need to be explored in further studies. Third, we did not conduct the *in vivo* experiments to validate the function of AC02278.4 in the tumor metastasis and tumor microenvironment regulation of LUAD. In the future, we will pay more attention to the function of AC02278.4 in tumor metastasis and tumor microenvironment regulation of LUAD. Furthermore, we will perform more *in vivo* and *in vitro* experiments to explore the function and the potential molecular mechanisms of AC02278.4 in tumor metastasis and tumor microenvironment regulation of LUAD.

Overall, our results confirmed that AC02278.4 could serve as a potential novel prognostic biomarker for LUAD. Moreover, we explored the underlying evidence indicating that AC02278.4 regulates immune cell infiltration in the tumor microenvironment in LUAD patients. Therefore, these findings are potentially valuable in advancing our current understanding of not only the role of AC02278.4 but also its translational use in LUAD prognosis and immunotherapy.

## CONCLUSION

In summary, our findings provide the first evidence that AC02278.4 has significantly increased expression in LUAD and may serve as a promising diagnostic and prognostic biomarker for 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 authors.

## AUTHOR CONTRIBUTIONS

XC, FZ, WR and JG designed this work and performed related assay. XH analyzed the data. JP, XN and XJ supervised and wrote the manuscript. All authors have read and approved the final version of the manuscript.

## FUNDING

This work was supported by the Applied Basic Research Project of Yunnan Provincial Science and Technology Department and Kunming Medical University, No.2020001AY070001-117. The Open Project of The First People's Hospital of Yunnan Province Clinical Medicine Center (2021LCZXXF-XZ03).

## REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2020. *CA: A Cancer J Clin* (2020) 70(1):7–30. doi: 10.3322/caac.21590
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(5):584–94. doi: 10.1016/S0025-6196(11)60735-0
3. Reck M, Rabe KF. Precision Diagnosis and Treatment for Advanced Non-Small-Cell Lung Cancer. *N Engl J Med* (2017) 377(9):849–61. doi: 10.1056/NEJMra1703413
4. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, et al. Genomic Landscape of Non-Small Cell Lung Cancer in Smokers and Never-Smokers. *Cell* (2012) 150(6):1121–34. doi: 10.1016/j.cell.2012.08.024
5. Shi Y, Fan S, Wu M, Zuo Z, Li X, Jiang L, et al. YTHDF1 Links Hypoxia Adaptation and non-Small Cell Lung Cancer Progression. *Nat Commun* (2019) 10(1):4892. doi: 10.1038/s41467-019-08460-2
6. Duruisseaux M, Esteller M. Lung Cancer Epigenetics: From Knowledge to Applications. *Semin Cancer Biol* (2018) 51:116–28. doi: 10.1016/j.semcancer.2017.09.005
7. Chen Y, Liu X, Li Y, Quan C, Zheng L, Huang K. Lung Cancer Therapy Targeting Histone Methylation: Opportunities and Challenges. *Comput Struct Biotechnol J* (2018) 16:211–23. doi: 10.1016/j.csbj.2018.06.001
8. Hulbert A, Jusue-Torres I, Stark A, Chen C, Rodgers K, Lee B, et al. Early Detection of Lung Cancer Using DNA Promoter Hypermethylation in Plasma and Sputum. *Clin Cancer Res* (2017) 23(8):1998–2005. doi: 10.1158/1078-0432.CCR-16-1371
9. Li Y, Yin Z, Fan J, Zhang S, Yang W. The Roles of Exosomal miRNAs and lncRNAs in Lung Diseases. *Signal Transduct Target Ther* (2019) 4:47. doi: 10.1038/s41392-019-0080-7
10. Ginn L, Shi L, Montagna M, Garofalo M. LncRNAs in Non-Small-Cell Lung Cancer. *Noncoding RNA* (2020) 6(3):25. doi: 10.3390/ncrna6030025
11. Osielska MA, Jagodziński PP. Long non-Coding RNA as Potential Biomarkers in Non-Small-Cell Lung Cancer: What do We Know So Far? *BioMed Pharmacother* (2018) 101:322–33. doi: 10.1016/j.biopha.2018.02.099
12. Harries LW. Long non-Coding RNAs and Human Disease. *Biochem Soc Trans* (2012) 40(4):902–6. doi: 10.1042/BST20120020
13. Wilusz JE, Sunwoo H, Spector DL. Long Noncoding RNAs: Functional Surprises From the RNA World. *Genes Dev* (2009) 23(13):1494–504. doi: 10.1101/gad.1800909
14. Chen Z, Lei T, Chen X, Gu J, Huang J, Lu B, et al. Long non-Coding RNA in Lung Cancer. *Clin Chim Acta* (2020) 504:190–200. doi: 10.1016/j.cca.2019.11.031
15. Pan J, Fang S, Tian H, Zhou C, Zhao X, Tian H, et al. lncRNA JPX/miR-33a-5p/Twist1 Axis Regulates Tumorigenesis and Metastasis of Lung Cancer by Activating Wnt/β-Catenin Signaling. *Mol Cancer* (2020) 19(1):9. doi: 10.1186/s12943-020-1133-9
16. Yang B, Zhang L, Cao Y, Chen S, Cao J, Wu D, et al. Overexpression of lncRNA IGFBP4-1 Reprograms Energy Metabolism to Promote Lung Cancer Progression. *Mol Cancer* (2017) 16(1):154. doi: 10.1186/s12943-017-0722-8
17. Yin D, Lu X, Su J, He X, De W, Yang J, et al. Long Noncoding RNA AFAP1-AS1 Predicts a Poor Prognosis and Regulates Non-Small Cell Lung Cancer Cell Proliferation by Epigenetically Repressing P21 Expression. *Mol Cancer* (2018) 17(1):92. doi: 10.1186/s12943-018-0836-7
18. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinf* (2011) 12:77. doi: 10.1186/1471-2105-12-77

19. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *Omics* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
20. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc Natl Acad Sci USA* (2005) 102(43):15545–50. doi: 10.1073/pnas.0506580102
21. 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
22. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer. *Immunity* (2013) 39(4):782–95. doi: 10.1016/j.immuni.2013.10.003
23. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
24. Zhang X, Xie K, Zhou H, Wu Y, Li C, Liu Y, et al. Role of non-Coding RNAs and RNA Modifiers in Cancer Therapy Resistance. *Mol Cancer* (2020) 19 (1):47. doi: 10.1186/s12943-020-01171-z
25. Li N, Feng XB, Tan Q, Luo P, Jing W, Zhu M, et al. Identification of Circulating Long Noncoding RNA Linc00152 as a Novel Biomarker for Diagnosis and Monitoring of Non-Small-Cell Lung Cancer. *Dis Markers* (2017) 2017:7439698. doi: 10.1155/2017/7439698
26. Xu YH, Tu JR, Zhao TT, Xie SG, Tang SB. Overexpression of lncRNA EGFR −AS1 is Associated With a Poor Prognosis and Promotes Chemotherapy Resistance in Non−Small Cell Lung Cancer. *Int J Oncol* (2019) 54(1):295–305. doi: 10.3892/ijo.2018.4629

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 11 April 2022
doi: 10.3389/fonc.2022.865409

# Cyclin-Dependent Kinase 6 Identified as the Target Protein in the Antitumor Activity of *Tetrastigma hemsleyanum*

*Chaoguang Wei*[1,2], *Yuxiang Zhao*[1,2], *Tao Ji*[1], *Yong Sun*[3], *Xudong Cai*[1] *and Xin Peng*[1*]

[1] Ningbo Municipal Hospital of TCM, Affiliated Hospital of Zhejiang Chinese Medical University, Ningbo, China, [2] United New Drug Research and Development Center, Biotrans Technology Co., Ltd, Ningbo, China, [3] State Key Laboratory of Food Science and Technology, Nanchang University, Nanchang, China

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Vikas Malik,*
*Columbia University, United States*
*Luis E. Arias-Romero,*
*National Autonomous University of Mexico, Mexico*

***Correspondence:**
*Xin Peng*
*pengx@nit.zju.edu.cn*

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

**Received:** *29 January 2022*
**Accepted:** *29 March 2022*
**Published:** *11 April 2022*

**Citation:**
*Wei C, Zhao Y, Ji T, Sun Y, Cai X and Peng X (2022) Cyclin-Dependent Kinase 6 Identified as the Target Protein in the Antitumor Activity of Tetrastigma hemsleyanum. Front. Oncol. 12:865409. doi: 10.3389/fonc.2022.865409*

**Background:** *Tetrastigma hemsleyanum* (*T. hemsleyanum*) is widely used as an adjuvant drug for tumor therapy but its antitumor therapeutic targets and molecular mechanisms have remained unclear. The prediction and analysis of natural products has previously used only network pharmacology methods to identify potential target proteins from public databases. In this study, we use comprehensive bioinformatics analysis and experimental verification to determine the antitumor mechanism of *T. hemsleyanum*.

**Methods:** Network pharmacology analysis was used to predict the potential *in vivo* target proteins of *T. hemsleyanum*. The expression matrix and clinical data to perform an analysis of hub genes were collected from the TCGA and GTEx databases, specifically the analysis of expression, prognosis, tumor immune cell infiltration analysis, immune checkpoint genes, microsatellite instability, tumor mutational burden, tumor neoantigen, and immune microenvironment, which identify the roles and biological functions of the hub genes in pan-cancer. Finally, gene set enrichment analysis was used to verify the biological processes and signaling pathways involved in the pan-cancer expression profile.

**Results:** We found 124 potential *in vivo* target proteins of *T. hemsleyanum* through network pharmacological analysis, and five hub genes (AKR1C1, MET, PTK2, PIK3R1, and CDK6) were then screened by protein–protein interaction (PPI) network analysis and molecular complex detection analysis (MCODE). Experimental intervention with an aqueous extract of *T. hemsleyanum* verified that these hub genes are the target proteins involved in the regulation of *T. hemsleyanum* in cells. A pan-cancer analysis then confirmed that CDK6 and MET are potential targets upon which *T. hemsleyanum* may exert antitumor action, especially in ACC, CESC, LGG, and PAAD. The CDK6 protein targeted by *T. hemsleyanum* is also involved in the immune and mutation process of pan-cancer, especially in the regulation of immune cell infiltration, immune checkpoint gene

expression, microsatellite instability, tumor mutation burdens, and tumor neoantigens. Together, these analyses show that *T. hemsleyanum* affects tumor immune regulation and genomic stability. Finally, a gene set enrichment analysis confirmed that *T. hemsleyanum* regulates the cell cycle checkpoint.

**Conclusions:** We found that *T. hemsleyanum* can behave as an antitumor agent by acting as a potential cell cycle checkpoint inhibitor in CDK6-driven tumors, such as ACC, CESC, LGG, and PAAD, and that it acts as a tyrosine kinase receptor inhibitor that inhibits the expression of the proto-oncogene MET. Combined with an analysis of immune and mutation correlations in pan-cancer, we determined that *T. hemsleyanum* may function biologically as an immune regulator and interfere with the stability of the tumor genome, which is worthy of further study.

**Keywords: CDK6, MET, antitumor, pan-cancer analysis, *Tetrastigma hemsleyanum***

## INTRODUCTION

*Tetrastigma hemsleyanum* Diels et Gilg (*T. hemsleyanum*) is a valuable medicinal plant mainly found in the eastern and southern regions of China and Southeast Asia (1). The plant is traditionally used as an herbal medicine for the treatment of hepatitis and rheumatic arthritis and for regulating immune function (2, 3). Previous research has found that flavonoids, lipid soluble substances, and polysaccharides extracted from *T. hemsleyanum* have strong antiproliferative effects on a variety of cancers (4), including leukemia and carcinomas of the lung (5), liver (6), uterus (7), and intestines (8). To date, more than one hundred compounds have been isolated and identified from *T. hemsleyanum*, including flavonoids and their glycosides, phenolic acids, and alkaloids (9), but the antitumor therapeutic targets and molecular mechanisms of the agents extracted from *T. hemsleyanum* have remained poorly understood.

The cell cycle in normal cells is finely regulated and cell cycle checkpoints prevent uncontrolled proliferation and repair any damaged DNA. Growth factors may activate cell cycle progression and are regulated by two groups of proteins: cyclins and cyclin-dependent kinases (CDK) (10). CDKs are a family of serine/threonine protein kinases in complex with cyclins that regulate various critical cellular processes, including cell cycle progression, transcription, neurogenesis, and apoptosis. Deregulation of CDKs is directly related to tumorigenesis, and the important roles of CDKs and cyclins in promoting cancer genesis and progression makes them attractive targets for drug inhibition. To date, at least 20 CDKs and 30 cyclins have been reported, and of these, emerging evidence suggests that CDK6 has a vital transcriptional role in tumor angiogenesis and thus represents a promising target for antitumor treatment (11).

In recent years, with the development of next-generation sequencing techniques and omics and the establishment of The Cancer Genome Atlas (TCGA) database, a new pipeline called pan-cancer analysis has developed to study the correlations between the expression of single genes or gene sets of at least 33 tumors and their phenotypes, such as prognosis, tumor immune cell infiltration, immune checkpoint gene expression, microsatellite instability, tumor mutational burden (TMB), and the immune microenvironment. The *in vivo* action targets of natural products can often be obtained through network pharmacological analysis and combined with the technical methods of network pharmacology and pan-cancer analysis, we can thus explore the different functions and related phenotypes of natural product targets in various tumors, which offer a new approach to screening the possible uses of natural products by mining their pharmacological mechanisms. In this study we examine, through a series of analyses, whether CDK6 is a potential target of *T. hemsleyanum* and whether *T. hemsleyanum* might have an antitumor function as a cell cycle checkpoint inhibitor.

## METHODS

### Data Collection

The gene expression profile data and clinical data for pan-cancer analysis were obtained from the TCGA database. Due to the small number of normal samples in the database and in order to improve the feasibility of the study, we added normal samples from the Genotype-Tissue Expression (GTEx) database. The abbreviations for the tumors in the pan-cancer analysis are shown in **Table 1**.

### Network Pharmacology Analysis

In previous published research, we searched for and analyzed the components of *T. hemsleyanum* (12). The Swiss Target Prediction database (http://www.swisstargetprediction.ch/) and UniProt database (https://www.uniprot.org/) were used to identify compound-related targets with probability values of 0.5 or greater and the probability value derived from a cross-validation analysis was used to rank the targets and to estimate the accuracy of the predictions.

### Protein–Protein Interaction (PPI) Network Analysis

The compound-related targets were imported into the STRING (13), BioGrid (14), OmniPath (15), and InWeb_IM (16)

**TABLE 1** | The abbreviations of tumors in pan-cancer analysis.

| Tumor Name | Abbreviations |
|---|---|
| Adrenocortical carcinoma | ACC |
| Bladder Urothelial Carcinoma | BLCA |
| Breast invasive carcinoma | BRCA |
| Cervical squamous cell carcinoma and endocervical adenocarcinoma | CESC |
| Cholangiocarcinoma | CHOL |
| Colon adenocarcinoma | COAD |
| Rectum adenocarcinoma | READ |
| Lymphoid Neoplasm Diffuse Large B-cell Lymphoma | DLBC |
| Esophageal carcinoma | ESCA |
| Glioblastoma multiforme | GBM |
| Head and Neck squamous cell carcinoma | HNSC |
| Kidney Chromophobe | KICH |
| Kidney renal clear cell carcinoma | KIRC |
| Kidney renal papillary cell carcinoma | KIRP |
| Acute Myeloid Leukemia | LAML |
| Low grade glioma | LGG |
| Liver hepatocellular carcinoma | LIHC |
| Lung adenocarcinoma | LUAD |
| Lung squamous cell carcinoma | LUSC |
| Mesothelioma | MESO |
| Ovarian serous cystadenocarcinoma | OV |
| Pancreatic adenocarcinoma | PAAD |
| Pheochromocytoma and Paraganglioma | PCPG |
| Prostate adenocarcinoma | PRAD |
| Rectum adenocarcinoma | READ |
| Sarcoma | SARC |
| Skin Cutaneous Melanoma | SKCM |
| Stomach adenocarcinoma | STAD |
| Stomach and Esophageal carcinoma | STES |
| Testicular Germ Cell Tumors | TGCT |
| Thyroid carcinoma | THCA |
| Thymoma | THYM |
| Uterine Corpus Endometrial Carcinoma | UCEC |
| Uveal Melanoma | UVM |

databases, and the files of interactions between the network nodes were obtained. Cytoscape was then used to visualize the PPI network and molecular complex detection (MCODE) was used to screen the hub nodes.

## Gene Function Enrichment Analysis

The Gene Ontology (GO) biological process project, the Kyoto Encyclopedia of Genes and Genomes (KEGG), and the Hallmark Gene Sets served as biological functioning or signaling pathway sources for conducting gene function enrichment analysis. The compound-related targets were imported into Metascape (http://metascape.org/) (17) for functional enrichment analysis and terms with a minimum count of 3, p-values below.01, and enrichment factors greater than 1.5, were used as the threshold for the enrichment analysis; the terms at the top of the selected enrichment indexes were visualized by bar graphs.

## Cell Culture

HEK293 cells (CRL-1573; American Type Culture Collection, VA, USA) were cultured in MEM medium (11090073; Invitrogen, MA, USA) with 10% FBS at 37°C in 5% carbon dioxide according to the manufacturer's instructions. Before the experiment, the cells underwent and passed cell identification and mycoplasma detection for quality control.

## Quantitative PCR Assay

The gene expression levels of AKR1C1, MET, PTK2, PIK3R1, and CDK6 were determined by quantitative PCR. HEK293 cells were sub-packed into 12-well plates and cultured overnight; the method of preparing an aqueous extract of *T. hemsleyanum* is given in previous studies (2). The HEK293 cells were treated with 0, 150, or 300 ug/ml aqueous extract of *T. hemsleyanum* then lysed with RNA-isolating Total RNA Extraction Reagent (R401-01; Vazyme, Nanjing, China) to prepare high-purity RNA, according to the manufacturer's instructions. HiScript III 1st Strand cDNA Synthesis (Kit R312-01; Vazyme, Nanjing, China) was used to prepare the cDNA library for reverse transcription. The $-2^{\Delta\Delta CT}$ method was used to quantify the gene expression levels and the experiment was repeated three times independently. Student's t-test was used for comparative analysis, with a p-value below.05 considered statistically significant. Primer information is shown in **Table 2**.

## Gene Expression Analysis

The gene expression matrix was obtained from the TCGA and GTEx databases. The expression distributions of the genes in tumor tissues and normal tissues were analyzed using a Wilcoxon test. A p-value below.05 was considered statistically significant. The results were visualized by a violin plot prepared using ggplot2.

## Gene Prognostic Analysis

Clinical data were obtained from the TCGA database and a univariate Cox regression analysis was performed on the screened hub genes. A forest map was used to display the p-value, HR, and 95% CI of each variable, prepared using the forestplot R package. Statistical analysis was performed using R software. A Wilcoxon test was used to compare the two groups of data, with a p-value below.05 considered to be statistically significant. The gene prognosis relationship, screened using the univariate Cox analysis, was verified by Kaplan–Meier survival analysis and the receiver operating characteristic (ROC) curve. The R packages survival and survminer were used to perform the Kaplan–Meier curve, logrank test, and univariate Cox proportional hazards regression analyses and the ROC curve was prepared with the R package timeROC.

## Correlation Analysis of Tumor Immune Cell Infiltration

The CIBERSORT and quanTIseq methods were used to cross-verify the correlation between the gene expression and tumor immune cell infiltration, which was performed with the immunedeconv package. A Wilcoxon test was used to compare the two groups of data, and a p-value below.05 was considered statistically significant. The correlations between selected gene expressions and tumor immune cell infiltration levels in pan-cancer were visualized by a circle plot.

**TABLE 2** | Primer information.

| Primer Name | Sequence (5'→3') | Tm (°C) |
|---|---|---|
| h-AKR1C1-F | TCCAGTGTCTGTAAAGCCAGG | 58 |
| h-AKR1C1-R | CCAGCAGTTTTCTCTGGTTGAA | 58 |
| h-MET-F | AGCAATGGGGAGTGTAAAGAGG | 58 |
| h-MET-R | CCCAGTCTTGTACTCAGCAAC | 58 |
| h-PTK2-F | GCTTACCTTGACCCCAACTTG | 58 |
| h-PTK2-R | ACGTTCCATACCAGTACCCAG | 58 |
| h-PIK3R1-F | ACCACTACCGGAATGAATCTCT | 58 |
| h-PIK3R1-R | GGGATGTGCGGGTATATTCTTC | 58 |
| h-CDK6-F | CCAGATGGCTCTAACCTCAGT | 58 |
| h-CDK6-R | AACTTCCACGAAAAAGAGGCTT | 58 |
| β-actin-F | CTCCATCCTGGCCTCGCTGT | 58 |
| β-actin-R | GCTGTCACCTTCACCTTCC | 58 |

## Correlation Analysis of Immune Checkpoint Gene Expression

Wilcoxon analysis was used to compare the correlations between the expression levels of the selected genes and the gene expression levels of immune checkpoints such as SIGLEC15, IDO1, CD274 (PD-L1), HAVCR2, PDCD1 (PD-1), CTLA4, LAG3, and PDCD1LG2 (PD-1), and pheatmap software was used to visualize the results. A p-value below.05 was considered statistically significant.

## Correlation Analysis of Microsatellite Instability (MSI)

Spearman's correlation coefficients between the selected gene expression levels and MSI levels were calculated (18). The relationships between the selected genes and the MSI levels of four tumors (ACC, CESC, LGG, and PAAD) were visualized by a lollipop plot and the correlations between the selected gene expressions and MSI levels in pan-cancer were visualized by a radar plot. A p-value below.05 was considered statistically significant.

## Correlation Analysis of TMB

Spearman's correlation coefficients between the selected gene expression levels and TMB levels were calculated (19). The relationships between the selected genes and the TMB levels of four tumors (ACC, CESC, LGG, and PAAD) were visualized by a lollipop plot and the correlations between the selected gene expressions and TMB levels in pan-cancer were visualized by a radar plot. A p-value below.05 was considered statistically significant.

## Correlation Analysis of Tumor Neoantigens in Pan-Cancer

The correlations between the selected gene expressions and neoantigen levels in pan-cancer were visualized by a radar plot. A p-value below.05 was considered statistically significant.

## Correlation Analysis of the Immune Microenvironment in Pan-Cancer

The immune microenvironment analysis of tumor tissue consisted of three parts: an ESTIMATEScore, an ImmuneScore, and a StromalScore. The correlations between the selected gene expressions and immune microenvironment levels in pan-cancer were visualized by a circle plot. A p-value below.05 was considered statistically significant.

## Gene Set Enrichment Analysis (GSEA) in Pan-Cancer

To observe the biological functions of the selected genes in the tumor, we divided the pan-cancer samples into high and low groups according to the gene expression levels. GSEA was used to enrich the KEGG pathways in the high expression group and the Hallmark pathways in the low expression group. The threshold was | NES | > 1, p <.05, FDR <.25.

# RESULTS

## Identification of Potential Target Proteins of *T. hemsleyanum*

As shown in **Table S1**, a total of 124 compound-related targets were found after eliminating duplicates. Enrichment analysis and PPI network analysis were performed by Metascape to explore the biological functions of the 124 targets. The complete PPI network is shown in **Figure 1A**. Compared with proteins at the edge of the network, those at the center have more connections, indicating that they may play an important biological function. All information regarding PPI members is given in **Table S2**.

PPI members were enriched and analyzed at the functional level using the GO biological process project, KEGG, and Hallmark Gene Sets, as shown in **Figure 1B**. From the GO biological processes, 455 pathways were found; ranking these enrichment pathways by $\log_{10}$ (P), the top 10 pathways were positive regulation of transferase activity, protein phosphorylation, cellular response to hormone stimulus, one-carbon metabolic process, cellular response to lipid, icosanoid metabolic process, cellular response to chemical stress, daunorubicin metabolic process, organic hydroxy compound metabolic process, and peptidyl-serine modification—the top 20 pathways are shown in **Table 3**. From KEGG, 49 pathways were enriched and the top 10 were nitrogen metabolism, pathways in cancer, PI3K-Akt signaling pathway, ovarian steroidogenesis, progesterone-mediated oocyte maturation, insulin resistance, arachidonic acid metabolism, microRNAs in cancer, phospholipase D signaling

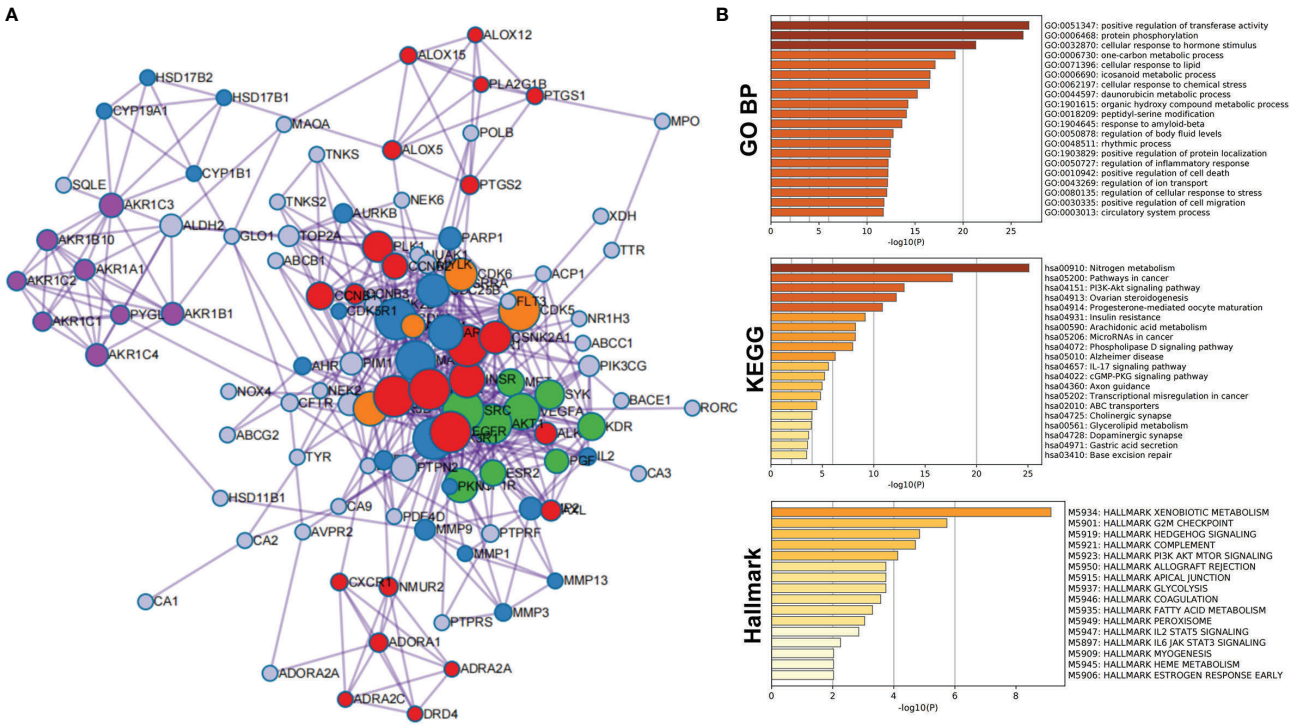



**FIGURE 1** | Protein–protein interaction network and enrichment analysis of network members of potential target proteins in *T. hemsleyanum*. **(A)** PPI network. **(B)** Enrichment analysis of network members of potential target proteins by GO, KEGG, and Hallmark.

pathway, and Alzheimer's disease; the top 20 enriched terms are presented in **Table 4**. From the Hallmark Gene Sets, 16 pathways were enriched and the top 10 were xenobiotic metabolism, G2M checkpoint, hedgehog signaling, complement, PI3K-Akt-mTOR signaling, apical junction, glycolysis, allograft rejection, coagulation, and fatty acid metabolism; details of all enriched terms are presented in **Table 5**.

The MCODE algorithm in Cytoscape was then used to identify neighborhoods in which proteins were densely connected. Each MCODE network was assigned a unique color, as shown in **Figure 2A**. Five MCODEs were found in the PPI Network: MCODE1 (composed of 24 genes), MCODE2 (21 genes), MCODE3 (9 genes), MCODE4 (8 genes), and MCODE5 (4 genes). GO enrichment analysis was applied to each MCODE network to assign "meanings" to the network components; the three terms with the lowest p-values were retained and are shown in **Table 6**. Hub targets were the key nodes for evaluating the essence of the whole network and were identified by MCODE and MCC algorithms. AKR1C1, MET, PTK2, PIK3R1, and CDK6 were identified as the crucial hub targets in the network.

To verify the accuracy of the network pharmacology and PPI analysis, we used experimental means to determine whether an aqueous extract of *T. hemsleyanum* could change the expression levels of the hub genes. As shown in **Figure 2B**, Q-PCR analysis demonstrated, for all concentrations of the aqueous extract of *T. hemsleyanum*, that the expression levels of each hub gene decreased significantly compared with the control group and that they decreased significantly along a gradient with increased concentration. These findings corroborate the *in silico* analysis and indicate that *T. hemsleyanum* significantly inhibits the expression of the target proteins.

## Gene Expression and Prognosis Analysis of Potential Target Proteins for *T. hemsleyanum* in Pan-Cancer

We used R software to calculate the difference in AKR1C1 expression between normal samples and tumor samples for each tumor and used non-paired Wilcoxon rank sum and signed rank tests to determine significance, as shown in **Figure 3A**. Significant up-regulation was observed in three tumors—HNSC, KIRP, and LUSC—and significant down-regulation was observed in 23 tumors—ACC, BLCA, BRCA, CESC, COAD, ESCA, GBM, KICH, KIRC, LAML, LGG, LIHC, LUAD, LUSC, OV, PAAD, PRAD, READ, SKCM, STAD, TGCT, THCA, and UCS. However, the expression of AKR1C1 in a univariate Cox prognostic analysis was inconsistent with its expression in tumors and normal tissues, as shown in **Figure 4A**. Generally speaking, proto-oncogenes are highly expressed in tumor tissues, which predicts poor prognosis and vice versa for tumor suppressor genes. In the PTK2 and PIK3R1 genes, the results were the reverse (**Figures 3C, D**, **4C, D**), so in the following analysis, these contradictory results are excluded.

**TABLE 3 |** List of tops enriched GO terms.

| GO | Category | Description | Count | % | $Log_{10}(P)$ |
|---|---|---|---|---|---|
| GO:0051347 | GO Biological Processes | positive regulation of transferase activity | 32 | 25.81 | -26.85 |
| GO:0006468 | GO Biological Processes | protein phosphorylation | 34 | 27.42 | -26.25 |
| GO:0032870 | GO Biological Processes | cellular response to hormone stimulus | 26 | 20.97 | -21.32 |
| GO:0006730 | GO Biological Processes | one-carbon metabolic process | 12 | 9.68 | -19.17 |
| GO:0071396 | GO Biological Processes | cellular response to lipid | 23 | 18.55 | -17.08 |
| GO:0006690 | GO Biological Processes | icosanoid metabolic process | 14 | 11.29 | -16.56 |
| GO:0062197 | GO Biological Processes | cellular response to chemical stress | 18 | 14.52 | -16.5 |
| GO:0044597 | GO Biological Processes | daunorubicin metabolic process | 7 | 5.65 | -15.23 |
| GO:1901615 | GO Biological Processes | organic hydroxy compound metabolic process | 20 | 16.13 | -14.26 |
| GO:0018209 | GO Biological Processes | peptidyl-serine modification | 15 | 12.1 | -14.1 |
| GO:1904645 | GO Biological Processes | response to amyloid-beta | 10 | 8.06 | -13.63 |
| GO:0050878 | GO Biological Processes | regulation of body fluid levels | 17 | 13.71 | -12.72 |
| GO:0048511 | GO Biological Processes | rhythmic process | 15 | 12.1 | -12.45 |
| GO:1903829 | GO Biological Processes | positive regulation of protein localization | 18 | 14.52 | -12.41 |
| GO:0050727 | GO Biological Processes | regulation of inflammatory response | 17 | 13.71 | -12.2 |
| GO:0010942 | GO Biological Processes | positive regulation of cell death | 20 | 16.13 | -12.18 |
| GO:0043269 | GO Biological Processes | regulation of ion transport | 21 | 16.94 | -12.14 |
| GO:0080135 | GO Biological Processes | regulation of cellular response to stress | 21 | 16.94 | -12.06 |
| GO:0030335 | GO Biological Processes | positive regulation of cell migration | 19 | 15.32 | -11.76 |
| GO:0003013 | GO Biological Processes | circulatory system process | 18 | 14.52 | -11.7 |

After screening all the data corresponding to expression and prognosis, it was determined that CDK6 and MET were potential target proteins for *T. hemsleyanum* and were significantly up-regulated in ACC (**Figure 3B**), whose high expression leads to poor prognosis (**Figure 4B**); CDK6 also showed similar results in CESC, LGG, and PAAD (**Figures 3E**, **4E**). We therefore chose CDK6 as the focus of downstream analysis. To verify the cancer-promoting effect of CDK6 in four tumors (ACC, CESC, LGG, and PAAD), Kaplan–Meier survival analysis was used to determine that CDK6 is not beneficial for patient prognosis and the ROC curve was used to characterize the relationship between the expression level of CDK6 and the prognosis—that is, the accuracy of the CDK6 model used to evaluate the survival of tumor patients. As shown in **Figure 5**, survival analysis for ACC showed that the high expression of CDK6 led to a poor prognosis for patients (HR = 1.77, p <.001), and the area under the curve (AUC) at 1, 3, and 5 years was greater than 0.7, indicating that the expression of CDK6 can be confidently used to predict adverse survival outcomes of ACC patients. Similarly, for CESC, Kaplan–Meier survival analysis also showed that the high expression of CDK6 led to poor prognosis for patients (HR = 1.24, p <.01), but the AUC at 1, 3, and 5 years was less than 0.7, indicating that the model of CDK6 expression used to predict clinical outcomes for patients in CESC has limitations. For LGG (HR = 1.75, p <.0001) and PAAD (HR = 1.87, p <.001), CDK6 led to a poor prognosis for patients, and the AUC was

**TABLE 4 |** List of tops enriched KEGG terms.

| GO | Category | Description | Count | % | $Log_{10}(P)$ |
|---|---|---|---|---|---|
| hsa00910 | KEGG Pathway | Nitrogen metabolism | 12 | 9.68 | -25.09 |
| hsa05200 | KEGG Pathway | Pathways in cancer | 24 | 19.35 | -17.64 |
| hsa04151 | KEGG Pathway | PI3K-Akt signaling pathway | 17 | 13.71 | -12.95 |
| hsa04913 | KEGG Pathway | Ovarian steroidogenesis | 9 | 7.26 | -12.19 |
| hsa04914 | KEGG Pathway | Progesterone-mediated oocyte maturation | 10 | 8.06 | -10.84 |
| hsa04931 | KEGG Pathway | Insulin resistance | 9 | 7.26 | -9.16 |
| hsa00590 | KEGG Pathway | Arachidonic acid metabolism | 7 | 5.65 | -8.22 |
| hsa05206 | KEGG Pathway | MicroRNAs in cancer | 12 | 9.68 | -8.2 |
| hsa04072 | KEGG Pathway | Phospholipase D signaling pathway | 9 | 7.26 | -7.95 |
| hsa05010 | KEGG Pathway | Alzheimer disease | 11 | 8.87 | -6.25 |
| hsa04657 | KEGG Pathway | IL-17 signaling pathway | 6 | 4.84 | -5.59 |
| hsa04022 | KEGG Pathway | cGMP-PKG signaling pathway | 7 | 5.65 | -5.22 |
| hsa04360 | KEGG Pathway | Axon guidance | 7 | 5.65 | -4.97 |
| hsa05202 | KEGG Pathway | Transcriptional misregulation in cancer | 7 | 5.65 | -4.82 |
| hsa02010 | KEGG Pathway | ABC transporters | 4 | 3.23 | -4.45 |
| hsa04725 | KEGG Pathway | Cholinergic synapse | 5 | 4.03 | -3.98 |
| hsa00561 | KEGG Pathway | Glycerolipid metabolism | 4 | 3.23 | -3.93 |
| hsa04728 | KEGG Pathway | Dopaminergic synapse | 5 | 4.03 | -3.66 |
| hsa04971 | KEGG Pathway | Gastric acid secretion | 4 | 3.23 | -3.56 |
| hsa03410 | KEGG Pathway | Base excision repair | 3 | 2.42 | -3.47 |

**TABLE 5** | List of tops enriched Hallmark terms.

| GO | Category | Description | Count | % | $Log_{10}(P)$ |
|---|---|---|---|---|---|
| M5934 | Hallmark Gene Sets | HALLMARK XENOBIOTIC METABOLISM | 11 | 8.87 | -9.15 |
| M5901 | Hallmark Gene Sets | HALLMARK G2M CHECKPOINT | 8 | 6.45 | -5.74 |
| M5919 | Hallmark Gene Sets | HALLMARK HEDGEHOG SIGNALING | 4 | 3.23 | -4.84 |
| M5921 | Hallmark Gene Sets | HALLMARK COMPLEMENT | 7 | 5.65 | -4.71 |
| M5923 | Hallmark Gene Sets | HALLMARK PI3K AKT MTOR SIGNALING | 5 | 4.03 | -4.13 |
| M5915 | Hallmark Gene Sets | HALLMARK APICAL JUNCTION | 6 | 4.84 | -3.74 |
| M5937 | Hallmark Gene Sets | HALLMARK GLYCOLYSIS | 6 | 4.84 | -3.74 |
| M5950 | Hallmark Gene Sets | HALLMARK ALLOGRAFT REJECTION | 6 | 4.84 | -3.74 |
| M5946 | Hallmark Gene Sets | HALLMARK COAGULATION | 5 | 4.03 | -3.57 |
| M5935 | Hallmark Gene Sets | HALLMARK FATTY ACID METABOLISM | 5 | 4.03 | -3.3 |
| M5949 | Hallmark Gene Sets | HALLMARK PEROXISOME | 4 | 3.23 | -3.04 |
| M5947 | Hallmark Gene Sets | HALLMARK IL2 STAT5 SIGNALING | 5 | 4.03 | -2.85 |
| M5897 | Hallmark Gene Sets | HALLMARK IL6 JAK STAT3 SIGNALING | 3 | 2.42 | -2.25 |
| M5906 | Hallmark Gene Sets | HALLMARK ESTROGEN RESPONSE EARLY | 4 | 3.23 | -2.03 |
| M5945 | Hallmark Gene Sets | HALLMARK HEME METABOLISM | 4 | 3.23 | -2.03 |
| M5909 | Hallmark Gene Sets | HALLMARK MYOGENESIS | 4 | 3.23 | -2.03 |

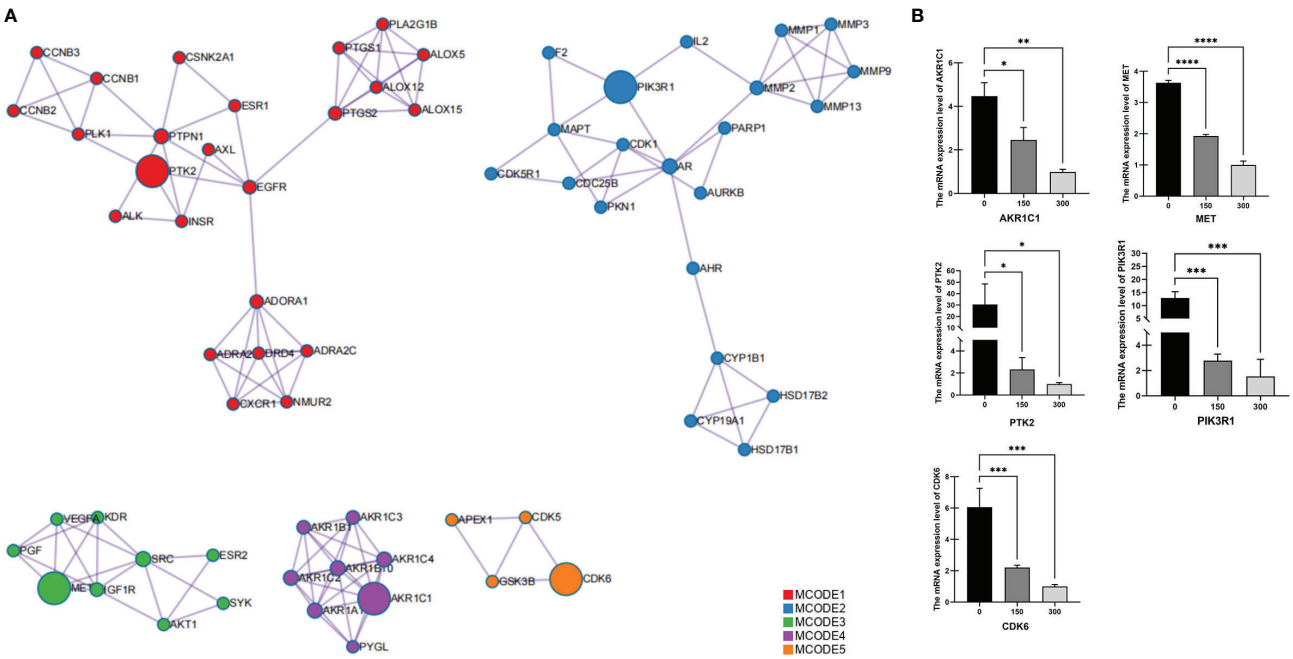



**FIGURE 2** | Identification of protein–protein interaction network hub genes for potential target proteins of *T. hemsleyanum*. **(A)** Hub gene screened by MCODE. **(B)** Quantitative PCR analysis confirmed that the aqueous extract of *T. hemsleyanum* was involved in the gene expression regulation of hub genes AKR1C1, MET, PTK2, PIK3R1, and CDK6. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001

around 0.7, indicating that it has potential as a biomarker for the clinical prognosis of these tumors.

## Pan-Cancer Analysis of the Relationship Between the *T. hemsleyanum*-Targeted Protein CDK6 and Immunity

CDK6-related tumor immune infiltration analysis using the CIBERSORT and quanTIseq methods produced different results. To ensure the reliability of the analysis, this study only discusses tumors and immune cells for which similar results were obtained from both methods. For ACC, the infiltration of B cells (B cell plasm) and dendritic cells (myeloid dendritic cell resting) was positively correlated with the expression of CDK6, while the infiltration of M2 macrophages was negatively correlated (**Figures 6A, B**, **8E**). At the pan-cancer level, the effect of CDK6 on tumor immune cell infiltration in different tumors is very heterogeneous. For example, in LUSC, ESCA, STAD, CHOL, DLBC, and PCPG tumors, the expression of CDK6 is not related to the infiltration

**TABLE 6 |** Top three enrichment terms of each MCODE components in PPI network.

| Network | GO | Annotation | $Log_{10}(P)$ |
|---|---|---|---|
| MCODE1 | GO:0045859 | regulation of protein kinase activity | -15.5 |
| | GO:0051347 | positive regulation of transferase activity | -14.8 |
| | GO:0033674 | positive regulation of kinase activity | -13.7 |
| MCODE2 | GO:0034644 | cellular response to UV | -12.6 |
| | GO:0071482 | cellular response to light stimulus | -11.3 |
| | GO:0009411 | response to UV | -11.0 |
| MCODE3 | hsa04510 | Focal adhesion | -13.6 |
| | ko04510 | Focal adhesion | -13.6 |
| | WP306 | Focal Adhesion | -13.5 |
| MCODE4 | GO:0044597 | daunorubicin metabolic process | -25.0 |
| | GO:0044598 | doxorubicin metabolic process | -25.0 |
| | GO:0030638 | polyketide metabolic process | -25.0 |
| MCODE5 | GO:0021766 | hippocampus development | -7.2 |
| | GO:0021761 | limbic system development | -6.8 |
| | GO:0021543 | pallium development | -6.2 |

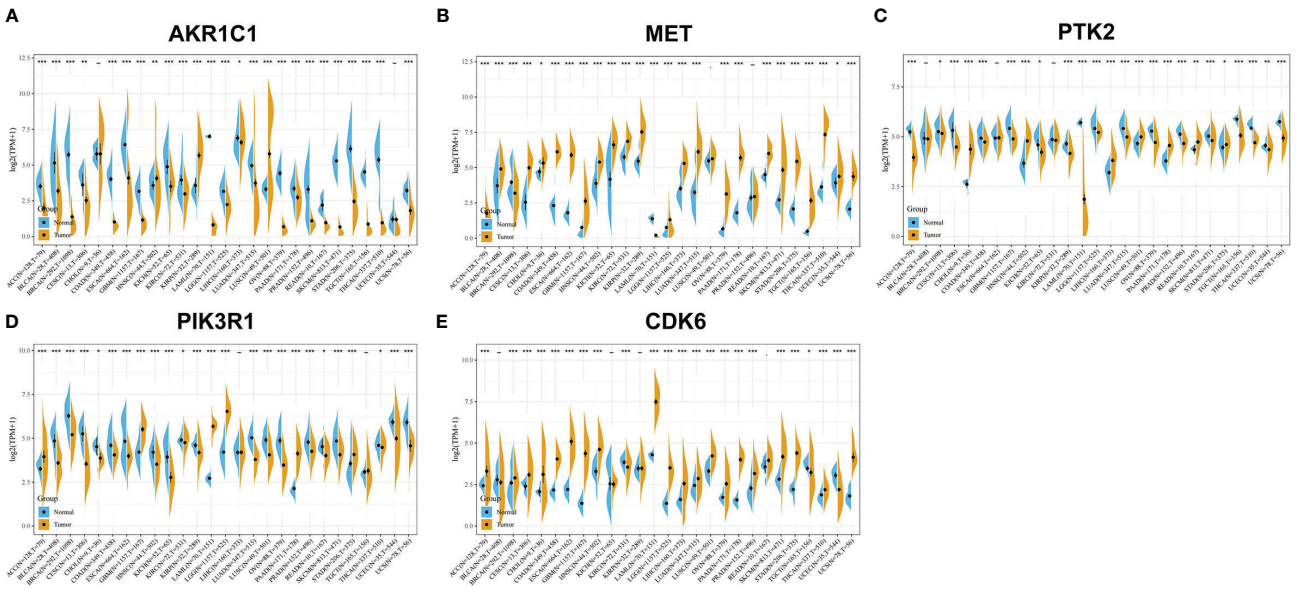



**FIGURE 3** | Gene expression analysis of potential target proteins of *T. hemsleyanum* in pan-cancer. **(A)** AKR1C1. **(B)** MET. **(C)** PTK2. **(D)** PIK3R1. **(E)** CDK6. *p < 0.05, **p < 0.01, ***p < 0.001.

and distribution of most immune cells, but in BLCA, TGCT, BRCA, KIRC, LAML, and ACC tumors, the relationship is significant (**Figure 7A**).

An analysis of the correlations between the expression of CDK6 and immune checkpoints such as SIGLEC15, IDO1, CD274 (PD-L1), HAVCR2, PDCD1 (PD-1), CTLA4, LAG3, and PDCD1LG2 (PD-1) shows that the expression of CDK6 in ACC is negatively correlated with the transcripts of most of the immune checkpoint genes, while it is positively correlated with most such genes in LGG. For PAAD and CESC, the expression of CDK6 in ACC was not related to the transcript expression of most immune checkpoint genes (**Figure 6C**). A comprehensive evaluation of the correlations between CDK6 and tumor immune checkpoint genes at the pan-cancer level shows that CDK6 is similar in its infiltration of tumor immune cells. CDK6 is not related to the expression levels of immune checkpoint genes in LUSC, ESCA, STAD, and CHOL, but it is related to tumors such as BLCA, TGCT, BRCA, and KIRC, indicating that, from the perspective of immunity, the two analyses are synergistic (**Figure 7B**).

At the level of the immune microenvironment, the top three tumors are, according to the StromalScore scoring rules, BRCA (R = .369, p <.001), BLCA (R = .36, p <.001), and KIRC (R = .297, p <.001); according to ImmuneScore, they are BLCA (R = .36, p <.001), BRCA (R = .369, p <.001), and LAML (R = -0.326, p <.001); and according to ESTIMATEScore, they are BLCA (R = .36, p <.001), BRCA (R = .369, p <.001), and KIRC (R = .297, p <.001) (**Figure 8D**).

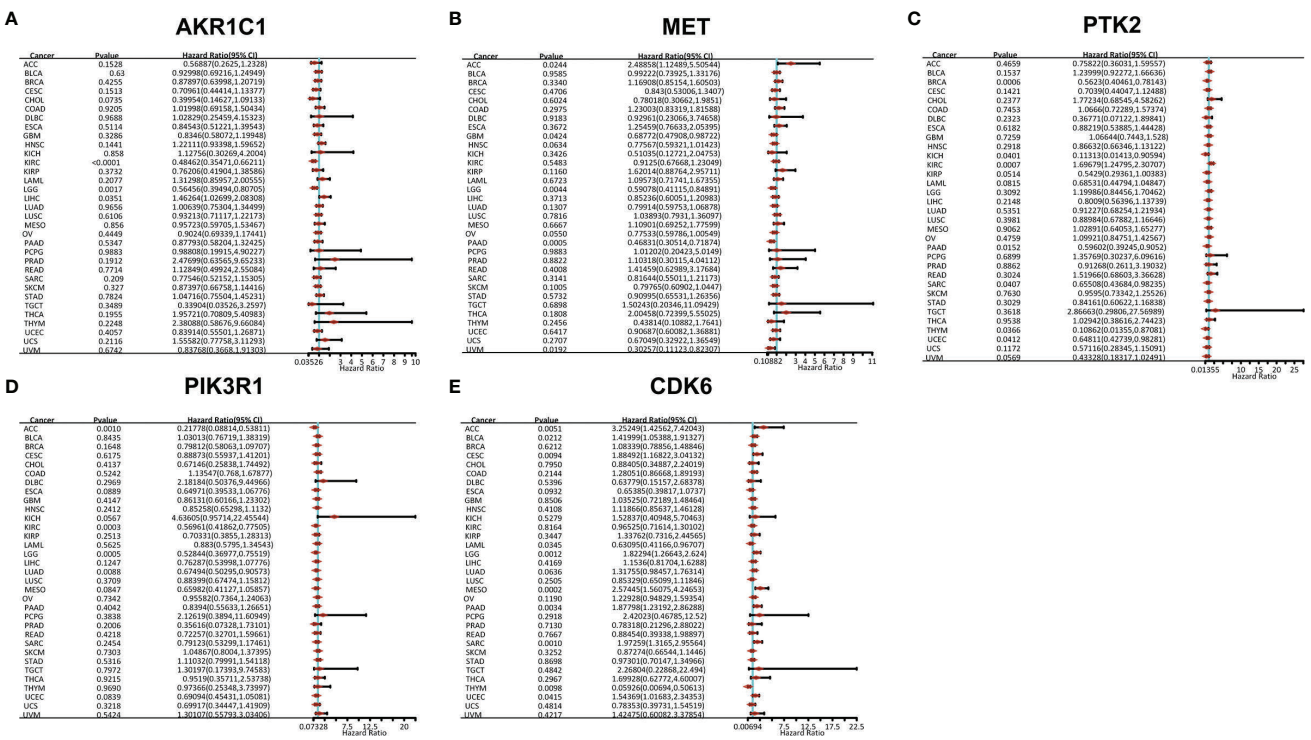



**FIGURE 4** | Gene prognostic analysis of potential target proteins of *T. hemsleyanum* in pan-cancer. **(A)** AKR1C1. **(B)** MET. **(C)** PTK2. **(D)** PIK3R1. **(E)** CDK6.

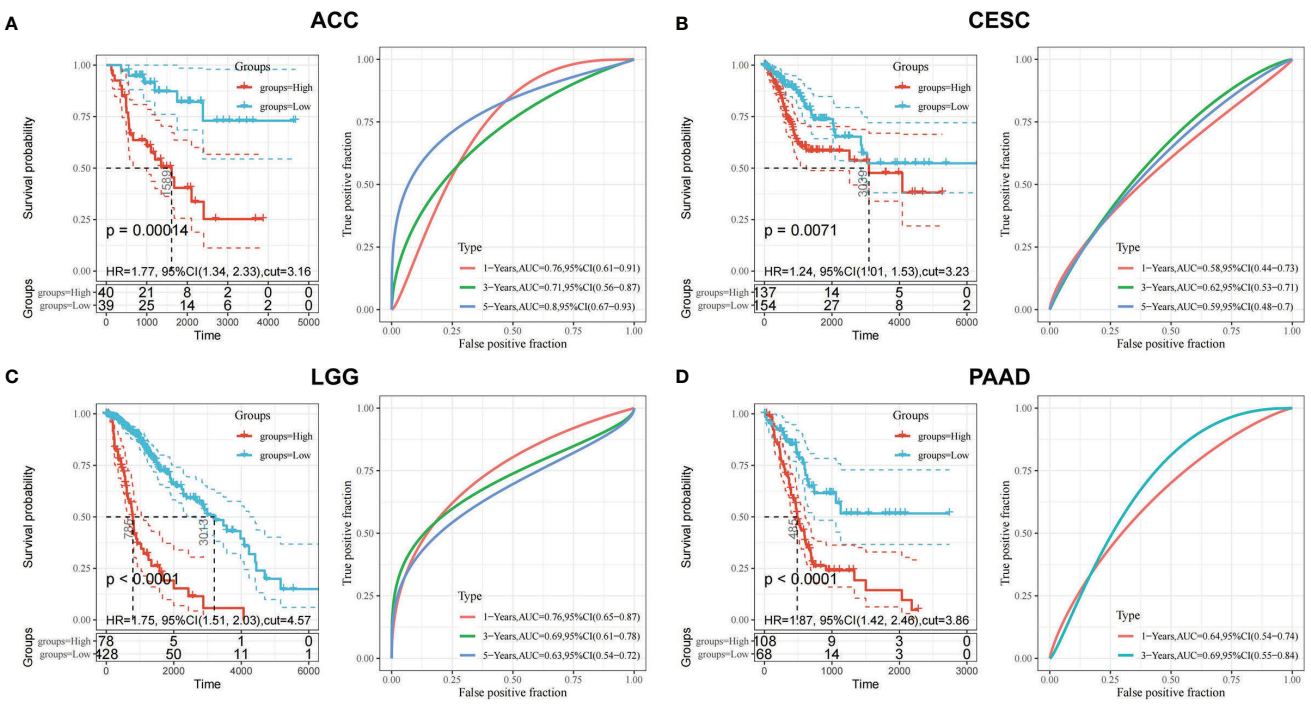



**FIGURE 5** | Prognostic validation of *T. hemsleyanum* potential target protein CDK6 in four tumors. **(A)** Kaplan–Meier survival analysis of CDK6 in ACC and ROC curve. **(B)** CESC. **(C)** LGG. **(D)** PAAD.

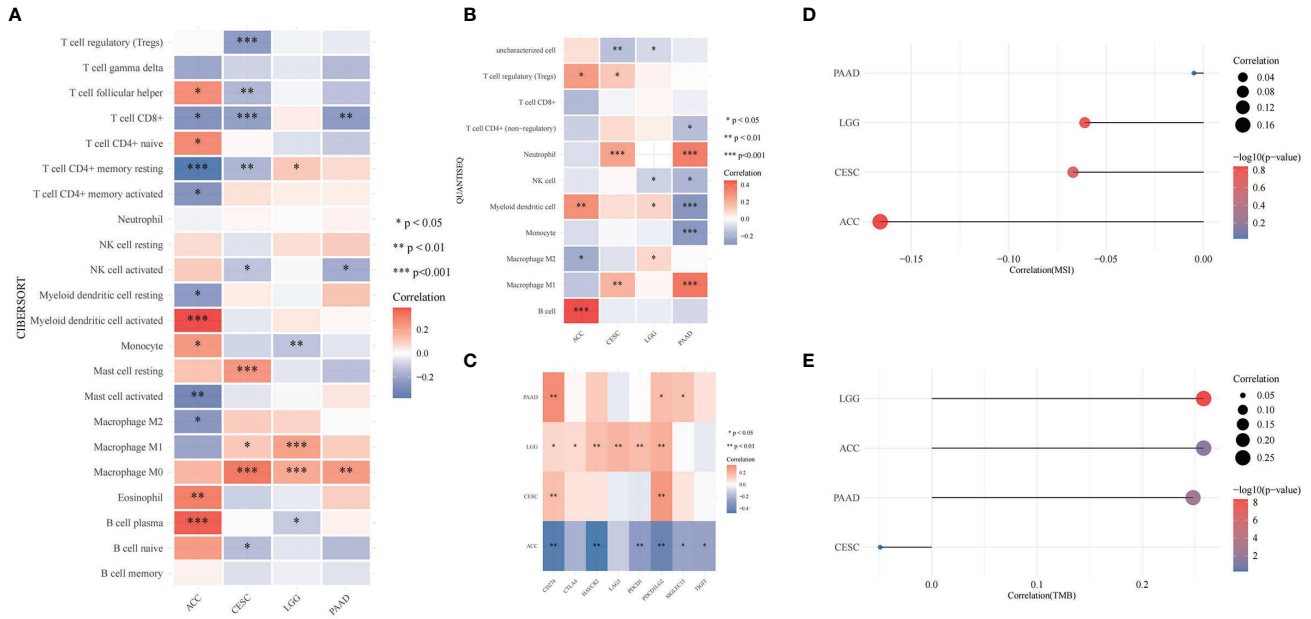



**FIGURE 6** | Correlation of *T. hemsleyanum*-targeted protein CDK6 with immunity and mutation of four tumors. **(A)** Analysis of tumor immune cell infiltration by CIBERSORT. **(B)** Analysis of tumor immune cell infiltration by quanTIseq. **(C)** Correlation analysis of immune checkpoints. **(D)** Correlation analysis of microsatellite instability. **(E)** Correlation analysis of TMB. *p < 0.05, **p < 0.01, ***p < 0.001.

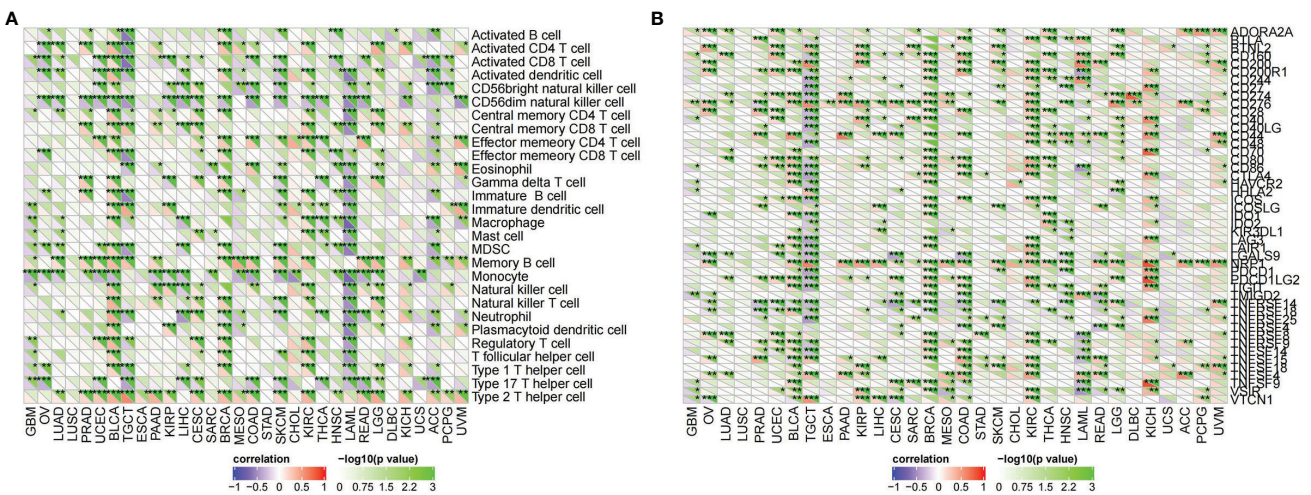



**FIGURE 7** | Correlation of *T. hemsleyanum*-targeted protein CDK6 with immunity in pan-cancer. **(A)** Analysis of tumor immune cell infiltration. **(B)** Correlation analysis of immune checkpoints. *p < 0.05, **p < 0.01, ***p < 0.001.

## Pan-Cancer Analysis of the Relationship Between *T. hemsleyanum*-Targeted Protein CDK6 and Mutation

In the microsatellite instability analysis, it was found that all four tumors for which CDK6 indicated a poor prognosis were negatively correlated with MSI; that is, CDK6 promotes the stability of the cell chromatin genome. Therefore, inhibiting the activity of CDK6 cells may affect the survival of tumor cells by inducing instability in the cell genome. The four tumors, in descending order of the magnitudes of their correlations with MSI are ACC, CESC, LGG, and PAAD (**Figure 6D**). At the pan-cancer level, the tumors with positive correlations between

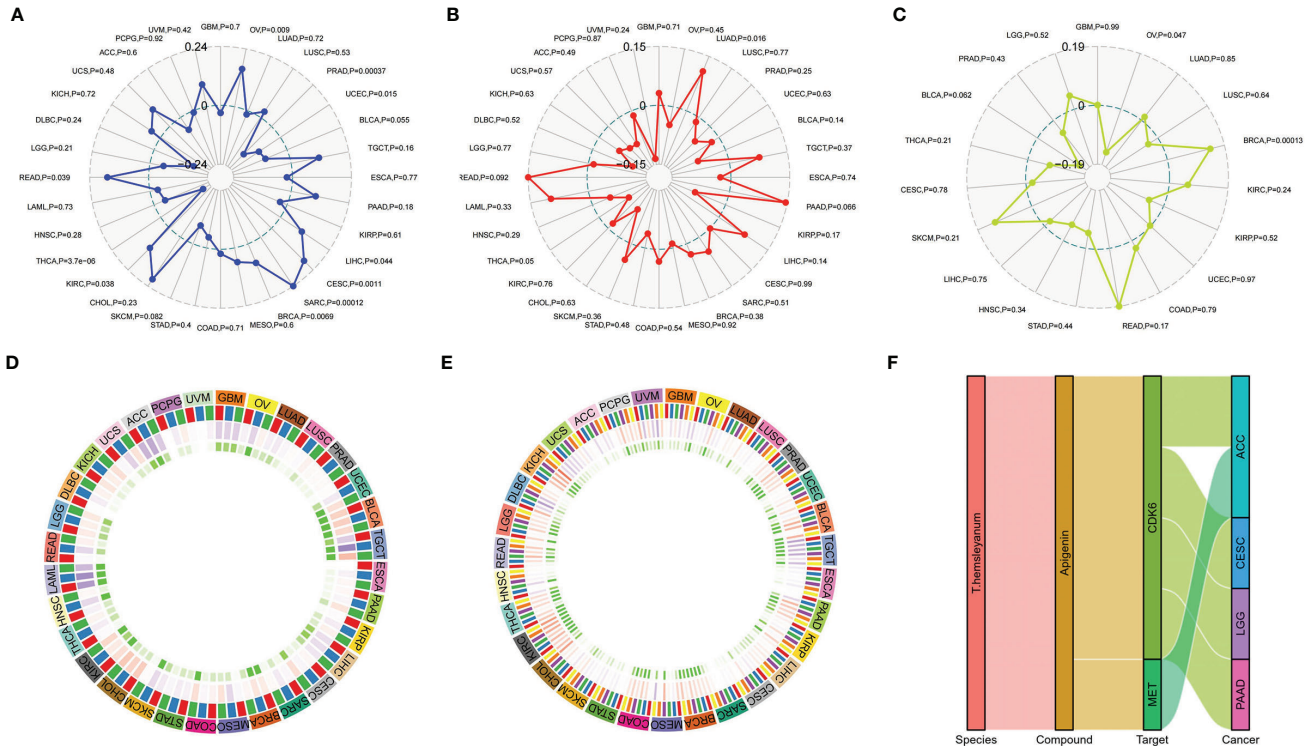



**FIGURE 8** | Correlation of *T. hemsleyanum*-targeted protein CDK6 with immunity and mutation in pan-cancer. **(A)** Correlation analysis of microsatellite instability. **(B)** Correlation analysis of TMB. **(C)** Correlation analysis of tumor neoantigens. **(D)** Correlation analysis of immune microenvironment. **(E)** Correlation analysis of tumor immune cell infiltration in pan-cancer. **(F)** Sanky plot of antitumor activity of *T. hemsleyanum*.

CDK6 expression and MSI are READ, KIRC, BRCA, SARC, CESC, LIHC, and OV and the tumors with negative correlations are THCA, UCEC, and PRAD (**Figure 8A**).

In terms of TMB, the expression of CDK6 was related to LGG, ACC, and PAAD, while CESC had little relationship with TMB (**Figure 6E**). At the pan-cancer level, LUAD was positively correlated with TMB, while THCA was negatively correlated. TMB is related to tumor neoantigen (**Figure 8B**); and generally speaking if TMB increases, patients can often benefit from targeted therapy, which is related to the generation of tumor neoantigen. At the pan-cancer level, in BRCA, the expression level of CDK6 was positively correlated with tumor neoantigen and negatively correlated with OV, showing that the roles and functions of CDK6 vary in different tumor mutations (**Figure 8C**). In general, we confirmed the potential pattern of anti-tumor activity of Tetrastigma hemsleyanum through a series of bioinformatics analysis, as shown in **Figure 8F**.

## Biological Function Identification of *T. hemsleyanum*-Targeted Protein CDK6 in Pan-Cancer

To study the function of CDK6 expression in pan-cancer, we divided the human pan-cancer samples into high and low expression groups according to the expression levels of CDK6 and we analyzed the enrichment of signal pathways in KEGG and Hallmark in both groups using GSEA. The three signal pathways most significantly enriched in the two databases in both groups are shown in **Figure 9A**. For the KEGG high expression group, focal adhesion (NES = -2.1, p <.001, FDR = .012), regulation of actin cytoskeleton (NES = -2.1, p <.001, FDR = .01), and glioma (NES = -2, p <.001, FDR = .024) were significantly enriched. For the KEGG low expression group, ribosome was the only significant enrichment pathway (NES = 1.9, p <.002, FDR = .11) (**Figure 9B**). For the Hallmark high expression group, mitotic spindle (NES = -2.2, p <.001, FDR <.001) and UV response (NES = -2, p = .002, FDR = .024) are related to the biological function of the G2M checkpoint (NES = -1.9, p = .011, FDR = .056) (**Figure 9C**), while for the Hallmark low expression group, no significant biological processes or signal pathways were enriched (**Figure 9D**).

## DISCUSSION

CDK6 has been shown to play an important role in regulating cell cycle progression and the up-regulation of CDK6 activity is closely related to the occurrence and development of multiple types of cancer, including breast cancer, hematological malignancies, and several solid tumors. CDK6 is overexpressed in cancer cells but only detected at low levels in non-cancer cells and CDK6-null mice have been found to develop normally (11),

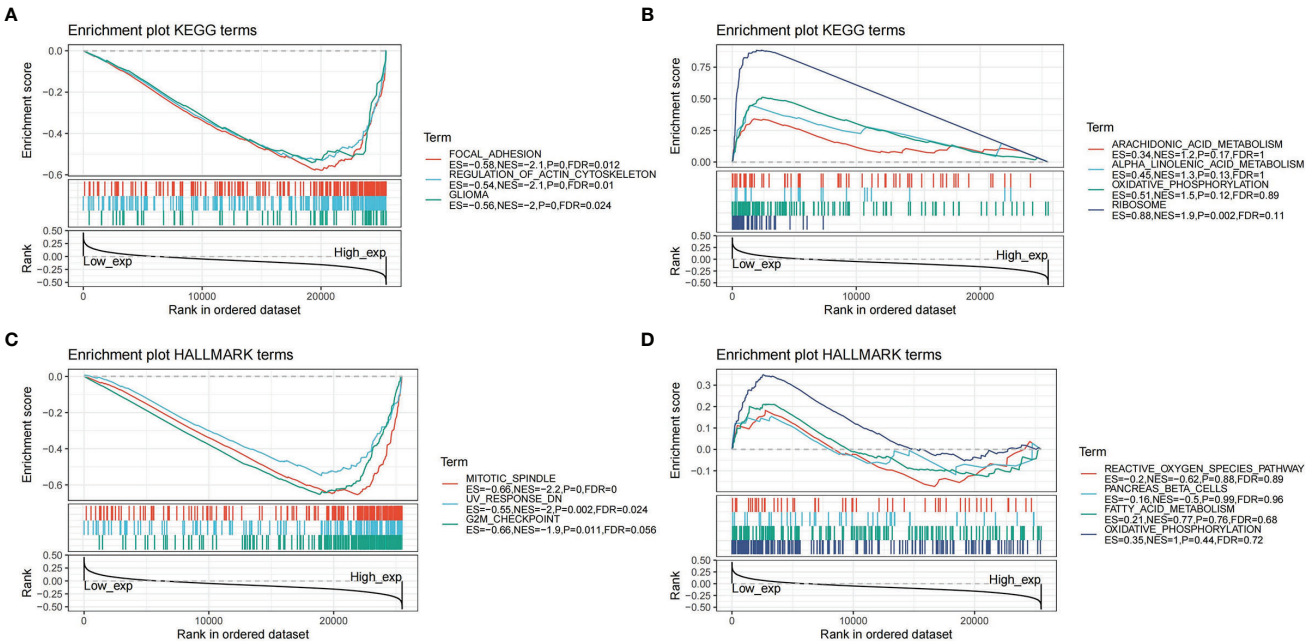



**FIGURE 9** | Gene set enrichment analysis in pan-cancer. **(A)** Enrichment plot of KEGG terms in high expression of CDK6. **(B)** Enrichment plot of KEGG terms in low expression of CDK6. **(C)** Enrichment plot of Hallmark terms in high expression of CDK6. **(D)** Enrichment plot of Hallmark terms in low expression of CDK6.

suggesting a potential low-toxicity therapeutic strategy through the inhibition of the expression of CDK6. Selective small-molecule CDK6 inhibitors therefore have tremendous potential as antitumor drugs and we confirmed herein the inhibitory effects of *T. hemsleyanum* extract on the expression of CDK6.

In general, CDK6 participates in the protection of cell genome stability and prevents cell mutations, which could lead to tumorigenesis (20) and similar results were found in the present study: *T. hemsleyanum* specifically inhibits the expression levels of CDK6, making it easy to cause genomic damage and to kill tumor cells, especially adrenocortical carcinoma, cervical cancer, low grade gliomas, and pancreatic cancer. The TMB results underline the biological function of CDK6 in maintaining cell stability. In previous research, overexpression of DNA damage and cell cycle–dependent proteins were observed to be associated with poor survival in 79 adrenocortical carcinoma patients and were accompanied by the significant up-regulation of genes involved in DNA damage and the regulation of cell cycle pathways; indeed, greater expression of CDK6 was associated with worse survival irrespective of age or sex (21).

Notably, tumor suppressor p16INK4A is indispensable for the survival of cervical carcinoma cell lines and oncogenic p16INK4A activity depends on the inhibition level of CDK6 (22). Similarly, miR-145 overexpression has been found to inhibit the proliferative ability of human cervical carcinoma cells by downregulating the expression of CDK6 (23), and abnormally expressed microRNA is closely associated with pancreatic cancer, with miR-3613-5p having been found to increase the metastasis of pancreatic cancer by targeting CDK6 (24). Conversely, sequential treatment with CDK6 inhibitors following DNA-damaging chemotherapy was shown to improve the therapeutic effect in pancreatic cancer, which was due to the action of CDK6 inhibitors on the homologous recombinant proteins responsible for repairing chromosome damage (25).

Mesenchymal–epithelial transition tyrosine kinase receptor (MET or c-MET) is a high-affinity proto-oncogene receptor tyrosine kinase encoded by the MET proto-oncogene. Hepatocyte growth factor (HGF) is the native peptide ligand of the MET receptor and aberrant HGF/MET activation drives oncogenic pathways involved in the development and progression of several human cancers, including renal, gastrointestinal, lung, and breast carcinomas, as well as glioblastoma multiforme (GBM) (26, 27). MET mutation and/or amplification is often found in genitourinary malignancies and has been shown in phase I clinical trials to lead to worse prognoses (28). Recent studies have indicated that many tumors display MET/HGF pathway abnormalities, which can be seen from c-MET overexpression and from MET mutation and/or amplification. Drugs that target c-MET may offer a new strategy for the control of aggressive cancers and the combination of HGF/MET inhibitors with conventional chemotherapy in patients affected by different cancer types has shown promising results. In studies of liver cancer cell lines and mice with orthotopic tumors, MET inhibitors were shown to promote liver tumor evasion of the immune response by stabilizing PDL1 (29). The MET receptor has thus emerged as a druggable target across several human cancers and agents targeting MET and HGF, including small molecules such as

crizotinib, tivantinib, and cabozantinib, have shown therapeutic effects in different tumors (30). In the present study, we demonstrated the inhibitory effects of *T. hemsleyanum* extract on the expression of MET. The active constituents of *T. hemsleyanum* are therefore promising candidates for improving the clinical activity of MET inhibitors and have the potential to improve survival rates and enhance therapeutic efficacy in human cancers.

Natural products generally have a variety of action targets and their biological processes and signal pathways are similarly multiple. In contrast, many widely used targeted drugs, such as immune checkpoint inhibitors or metabolic checkpoint inhibitors, may induce drug resistance in tumor cells as the target-sensitive main clone created by tumor heterogeneity gives way to non-sensitive subclones, gradually inducing drug resistance under the effect of mutation-screening caused by the targeted drugs. The multi-target effect of natural products therefore has the potential to overcome some of the drug resistance of targeted drugs and is likely to either trigger synergistic death in the process of drug use or to sensitize the targeted drugs in combination with the primary chemotherapy regimen, thus reducing drug toxicity and improving pharmacokinetics. Through the integrated network pharmacology pan-cancer analysis of the effective components of *T. hemsleyanum*, this study has shown that *T. hemsleyanum* acts as a cell cycle checkpoint inhibitor and a tyrosine kinase receptor inhibitor and has demonstrated the multi-target nature of *T. hemsleyanum* as a natural product, which was verified by bioinformatics analysis combined with experiments. The pivotal genes CDK6 and MET of the *T. hemsleyanum*-inhibited tumor were confirmed, providing a new theoretical basis for the clinical application of *T. hemsleyanum*.

In fact, many studies have already reported the antitumor effects of *T. hemsleyanum*, both *in vitro* and *in vivo*, and it has been widely used in clinics as an adjuvant drug to chemotherapy treatment. In this study we explored, for the first time, the antitumor mechanisms of *T. hemsleyanum* through network pharmacology combined with pan-cancer analysis. As a common screening method for natural product action targets, network pharmacological analysis relies on an existing database for text mining. Combined with pan-cancer analysis—as a means of comprehensively describing biological functions—and with phenotypic correlation analysis of molecules in different tumors, it represents a new strategy for the study of the antitumor activity of natural products.

## 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

XP designed the project and experiment, CGW, YXZ, XDC, TJ and YS have completed the analysis, CGW, TJ and YS completed the experiment, XP, YXZ and CGW wrote the draft of the manuscript, and all the authors have read the manuscript and approved the publication of the manuscript.

## FUNDING

This research was supported by the Major Science and Technology Projects of Breeding New Varieties of Agriculture in Zhejiang Province (2021C02074) and by the Ningbo City Science and Technology Innovation 2025 Major Research Project (2019B10008).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.865409/full#supplementary-material.

## REFERENCES

1. Peng X, Li MJ, Hao WU, Chen HJ, Zhang ZY. Co-Regulation Role of Endogenous Hormones and Transcriptomics Profiling Under Cold Stress in Tetrastigma Hemsleyanum. *J Plant Growth Regul* (2021) 40:1992–2006. doi: 10.1007/s00344-020-10246-6
2. Lou TL, Ji T, Peng X, Ji WW, Yuan LX, Wang J, et al. Extract From Tetrastigma Hemsleyanum Leaf Alleviates Pseudomonas Aeruginosa Lung Infection: Network Pharmacology Analysis and Experimental Evidence. *Front Pharmacol* (2021) 12:587850. doi: 10.3389/fphar.2021.587850
3. Ji W, Peng X, Lou T, Wang J, Qiu W. Total Flavonoids From Tetrastigma Hemsleyanum Ameliorates Inflammatory Stress in Concanavalin A–induced Autoimmune Hepatitis Mice by Regulating Treg/Th17 Immune Homeostasis. *Inflammopharmacol* (2019) 27:1297–307. doi: 10.1007/s10787-019-00599-0
4. Ru Y, Chen X, Wang J, Guo L, Lin Z, Peng X, et al. Polysaccharides From Tetrastigma Hemsleyanum Diels et Gilg: Extraction Optimization, Structural Characterizations, Antioxidant and Antihyperlipidemic Activities in Hyperlipidemic Mice. *Int J Biol Macromol* (2019) 125:1033–41. doi: 10.1016/j.ijbiomac.2018.11.236
5. Li YL, Zhang YR, Wang YX, Yu X, Yu T, Zheng XD, et al. Tetrastigma Hemsleyanum Leaf Flavones Have Anti-NSCLC Ability by Triggering Apoptosis Using the Akt-mTOR Pathway. *Food Biosci* (2021) 41(5):100914. doi: 10.1016/j.fbio.2021.100914
6. Peng X, Zhang YY, Wang J, Ji QY. Ethylacetate Extract From Tetrastigma Hemsleyanum Induces Apoptosis *via* the Mitochondrial Caspase-Dependent Intrinsic Pathway in HepG2 Cells. *Tumor Biol* (2016) 37(1):865–76. doi: 10.1007/s13277-015-3579-8
7. Xiong Y, Wu XW, Rao LQ. Tetrastigma Hemsleyanum (Sanyeqing) Root Tuber Extracts Induces Apoptosis in Human Cervical Carcinoma HeLa Cells. *J Ethnopharmacol* (2015) 165:46–53. doi: 10.1016/j.jep.2015.02.030
8. Chu Q, Jia RY, Chen W, Liu YY, Li YL, Ye X, et al. Purified Tetrastigma Hemsleyanum Vines Polysaccharide Attenuates EC-Induced Toxicity in Caco-2 Cells and Caenorhabditis Elegans *via* DAF-16/FOXO Pathway. *Int J Biol Macromol* (2020) 150:1192–202. doi: 10.1016/j.ijbiomac.2019.10.128
9. Ji T, Ji WW, Wang J, Chen HJ, Peng X, Cheng KJ, et al. A Comprehensive Review on Traditional Uses, Chemical Compositions, Pharmacology

Properties and Toxicology of Tetrastigma Hemsleyanum. *J Ethnopharmacol* (2021) 264:113247. doi: 10.1016/j.jep.2020.113247
10. Nardone V, Barbarino M, Angrisani A, Correale P, Pastina P, Cappabianca S, et al. CDK4, CDK6/cyclin-D1 Complex Inhibition and Radiotherapy for Cancer Control: A Role for Autophagy. *Int J Mol Sci* (2021) 22:8391. doi: 10.3390/ijms22168391
11. Tadesse S, Yu M, Kumarasiri M, Le BT, Wang SD. Targeting CDK6 in Cancer: State of the Art and New Insights. *Cell Cycle* (2015) 14(20):3220–30. doi: 10.1080/15384101.2015.1084445
12. Ji T, Ji WW, Wang J, Peng X, Xu Z, Cao W, et al. Tetrastigma Hemsleyanum Alleviates Sarcoidosis Through Metabolomic Regulation and Th17/Treg Immune Homeostasis. *J Funct Foods* (2022) 88:104910. doi: 10.1016/j.jff.2021.104910
13. 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:D607–613. doi: 10.1093/nar/gky1131
14. Stark C, Breitkreutz B-J, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: A General Repository for Interaction Datasets. *Nucleic Acids Res* (2006) 34:D535–539. doi: 10.1093/nar/gkj109
15. Turei D, Korcsmáros T, Saez-Rodriguez J. OmniPath: Guidelines and Gateway for Literature-Curated Signaling Pathway Resources. *Nat Methods* (2016) 13:966–7. doi: 10.1038/nmeth.4077
16. Li T, Wernersson R, Hansen RB, Horn H, Mercer J, Slodkowicz G, et al. A Scored Human Protein–Protein Interaction Network to Catalyze Genomic Interpretation. *Nat Methods* (2017) 14:61–4. doi: 10.1038/nmeth.4083
17. Zhou YY, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6
18. Bonneville R, Krook MA, Kautto EA, Miya J, Wing MR, Chen HZ, et al. Landscape of Microsatellite Instability Across 39 Cancer Types. *JCO Precis Oncol* (2017) 2017:PO.17.00073. doi: 10.1200/PO.17.00073
19. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. *Immunity* (2018) 48(4):812–830.e14. doi: 10.1016/j.immuni.2018.03.023
20. Qie S, Diehl JA. Cyclin D1, Cancer Progression and Opportunities in Cancer Treatment. *J Mol Med (Berl)* (2016) 94(12):1313–26. doi: 10.1007/s00109-016-1475-3
21. Subramanian C, Cohen MS. Over Expression of DNA Damage and Cell Cycle Dependent Proteins are Associated With Poor Survival in Patients With Adrenocortical Carcinoma. *Surgery* (2019) 165(1):202–10. doi: 10.1016/j.surg.2018.04.080
22. McLaughlin-Drubin ME, Park D, Munger K. Tumor Suppressor P16ink4ais Necessary for Survival of Cervical Carcinoma Cell Lines. *PNAS* (2013) 110 (40):16175–80. doi: 10.1073/pnas.1310432110
23. Zhang J, Wang L, Li BL, Huo MP, Mu MT, Liu JJ, et al. miR–145 Downregulates the Expression of Cyclin–Dependent Kinase 6 in Human Cervical Carcinoma Cells. *Exp Ther Med* (2014) 8:591–4. doi: 10.3892/etm.2014.1765
24. Cao R, Wang K, Long M, Guo M, Sheng L, Zhan M, et al. miR-3613-5p Enhances the Metastasis of Pancreatic Cancer by Targeting CDK6. *Cell Cycle* (2020) 19(22):3086–95. doi: 10.1080/15384101.2020.1831254
25. Salvador-Barbero B, Alvarez-Fernandez M, Zapatero-Solana E, Bakkali AE, Menendez M, Lopez-Casas P, et al. CDK4/6 Inhibitors Impair Recovery From Cytotoxic Chemotherapy in Pancreatic Adenocarcinoma. *Cancer Cell* (2020) 37:340–53. doi: 10.1016/j.ccell.202.01.007
26. Sabari JK, Leonardi GC, Shu CA, Umeton R, Montecalvo J, Ni A, et al. PD-L1 Expression, Tumor Mutational Burden, and Response to Immunotherapy in Patients With MET Exon 14 Altered Lung Cancers. *Ann Oncol* (2018) 29:2085–91. doi: 10.1093/annonc/mdy334
27. Comoglio PM, Trusolino L, Boccaccio C. Known and Novel Roles of the MET Oncogene in Cancer: A Coherent Approach to Targeted Therapy. *Nat Rev Cancer* (2018) 18:341–58. doi: 10.1038/s41568-018-0002-y
28. Jardim D, Gagliato D, Falchook G, Zinner R, Wheler J, Janku F, et al. MET Abnormalities in Patients With Genitourinary Malignancies and Outcomes With C-MET Inhibitors. *Clin Genitourin Cancer* (2015) 13(1):e19–26. doi: 10.1016/j.clgc.2014.06.017
29. Li H, Li CW, Li XQ, Ding QQ, Guo L, Liu S, et al. MET Inhibitors Promote Liver Tumor Evasion of the Immune Response by Stabilizing PDL1. *Gastroenterology* (2019) 156:1849–61. doi: 10.1053/j.gastro.2019.01.252
30. Moosavi F, Giovannetti E, Peters G, Firuzi O. Combination of HGF/MET-Targeting Agents and Other Therapeutic Strategies in Cancer. *Crit Rev Oncol Hemat* (2021) 160:103234. doi: 10.1016/j.critrevonc.2021.103234

**Conflict of Interest:** CGW and YXZ were employed by Biotrans Technology Co., LTD.

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 11 April 2022
doi: 10.3389/fonc.2022.877796

# Identification of a 12-Gene Signature and Hub Genes Involved in Kidney Wilms Tumor *via* Integrated Bioinformatics Analysis

*Guoping Huang* * *and Jianhua Mao*

*Department of Nephrology, The Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China*

Wilms tumor (WT), also known as nephroblastoma, is a rare primary malignancy in all kinds of tumor. With the development of second-generation sequencing, the discovery of new tumor markers and potential therapeutic targets has become easier. This study aimed to explore new WT prognostic biomarkers. In this study, WT-miRNA datasets GSE57370 and GSE73209 were selected for expression profiling to identify differentially expressed genes. The key gene miRNA, namely hsa-miR-30c-5p, was identified by overlapping, and the target gene of candidate hsa-miR-30c-5p was predicted using an online database. Furthermore, 384 genes were obtained by intersecting them with differentially expressed genes in the TARGET-WT database, and the genes were analyzed for pathway and functional enrichment. Kaplan–Meier survival analysis of the 384 genes yielded a total of 25 key genes associated with WT prognosis. Subsequently, a prediction model with 12 gene signatures (BCL6, CCNA1, CTHRC1, DGKD, EPB41L4B, ERRFI1, LRRC40, NCEH1, NEBL, PDSS1, ROR1, and RTKN2) was developed. The model had good predictive power for the WT prognosis at 1, 3, and 5 years (AUC: 0.684, 0.762, and 0.774). Finally, ERRFI1 (hazard ratios [HR] = 1.858, 95% confidence intervals [CI]: 1.298–2.660) and ROR1 (HR = 0.780, 95% CI: 0.609–0.998) were obtained as independent predictors of prognosis in WT patients by single, multifactorial Cox analysis.

**Keywords: miRNA, prognosis, Wilms tumor, target genes, risk model**

OPEN ACCESS

***Edited by:***
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

***Reviewed by:***
*Dezhong Zhou,*
*Xi'an Jiaotong University, China*
*Wei Ge,*
*Shandong University, China*

***Correspondence:**
*Guoping Huang*
*6510018@zju.edu.cn*

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

***Received:*** *17 February 2022*
***Accepted:*** *07 March 2022*
***Published:*** *11 April 2022*

***Citation:***
*Huang G and Mao J (2022) Identification of a 12-Gene Signature and Hub Genes Involved in Kidney Wilms Tumor via Integrated Bioinformatics Analysis. Front. Oncol. 12:877796. doi: 10.3389/fonc.2022.877796*

## INTRODUCTION

Wilms tumor (WT), also known as nephroblastoma, is the second most common intra-abdominal tumor and the most common primary renal tumor in children (1). Approximately 75% of children with WT develop the condition between the ages of 1 and 5 years, most commonly at the age of (2). After years of clinical exploration, the five-year survival rate for WT has improved from less than 30% to 85%–90%. However, the recurrence rate remains at 15%–50% (3). The treatment and prognosis of nephroblastoma in children are related to histological staging (4), and five-year survival rates have reached 90% for children with good histological types of WT after moderate treatment (5, 6).

The treatment of nephroblastoma continues to be based on surgical resection supplemented by a combination of chemotherapy and radiotherapy (7, 8). However, treatment-related complications

remain a problem for many children with WT, with treatment often triggering nausea and vomiting, loss of appetite, anemia, alopecia, and neutropenia, which subsequently affect patients' psychological well-being (9). At the same time, surgical removal of the diseased kidney has limitations. Therefore, current treatments are not entirely appropriate for some populations, especially infants and children and patients with bilateral tumors (10). Therefore, the key to improving patient prognosis is to improve treatment based on clinical and biological risk factors, and further stratification of current treatment options based on the prognostic value of tumor biology would be an important approach to improving WT prognosis (11, 12). Along with the development of CRISPR/Cas9 gene-editing technology, artificially modified chimeric antigen receptor T cell immunotherapy, and aptamer technology, precise genetic and biological therapies could be a new option for the treatment of nephroblastoma (13–15).

With the development of second-generation sequencing, the discovery of new tumor markers and potential therapeutic targets has become more accessible. Advances in RNA sequencing technologies have revealed the complexity of the human genome. The study of the RNA transcriptome is one of the most important challenges facing biology today, as RNAs represent new potential biomarkers and drug targets (16, 17). Currently, a growing number of studies on WT have identified many key mRNAs that are closely associated with the prognosis of this tumor (18, 19). It is well known that inter-individual heterogeneity usually constitutes only a more traditional prognostic system. For example, risk stratification based on the TNM staging system alone is not sufficient, nor is it sufficient to provide an accurate prediction of survival outcomes. A study by Lin et al. (20) identified a 5 mRNA signature as a new potential prognostic biomarker for WT, beyond which models have not been over reported. Therefore, additional prognostic models are needed to predict survival outcomes in pediatric WT patients.

This study was a comprehensive study to analyze differential genes in WT samples through multiple datasets and to develop validated gene signatures for predicting prognosis in WT patients, as well as to further screen key genes to provide a research basis for future biological treatment and clinical diagnosis of WT.

# METHODOLOGY

## Data sources

RNA-seq data and clinical information for TARGET-WT were downloaded from the UCSC Xena platform (http://xena.ucsc.edu/), which included 126 cancer tissue samples and six paracancer tissue samples. The TCGA expression matrix was obtained by data fusion and ID transformation of raw TCGA counts data. Searches were performed in Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) using the keyword "Wilms tumor" followed by manual review and selection of cohorts containing miRNA, mRNA expression, GSE57370 (21, 22) (Platform: GPL16770), containing 62 WT cancer tissues and four non-cancerous tissues; GSE73209 (23) (Platform: GPL10558) containing 32 WT cancer tissues and 6 noncancerous tissues. If more than one probe detected the same miRNA expression during the analysis, the average of that miRNA expression was taken as the expression value of that miRNA. For the analysis of patient clinical information, the clinical information of patients with unknown survival times and those equal to zero were deleted.

## Differential Expression Analysis

We applied the limma package of R software (v4.0.3) to perform normalization and base-2 logarithm conversion for the matrix data for each GEO dataset. "Adjusted P value < 0.05 and $|\log FC| \geq 1$" were defined as the thresholds for differentially expressed gene screening, and overlapping genes were analyzed by Venn plot using the ggplot2 package to plot heat maps and volcano maps, respectively.

## Target Gene Prediction

The target genes of key miRNAs were predicted using the miRDB online database (http://mirdb.org/).

## Functional Enrichment Analysis

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analyses were performed on genes using the DAVID 6.8 database (https://david.ncifcrf.gov/). Enrichment results with $P < 0.05$ or FDR < 0.05 were considered significant.

## Kaplan–Meier Survival Analysis

Survival analysis was performed using Survival in the R package. P-values and HR with 95% CI in Kaplan–Meier curves were derived using log-rank tests and univariate Cox proportional hazards regression.

## LASSO Regression Model Construction

The expression matrix integrating the initial genes of the model with patient survival status and survival time was constructed. Furthermore, the LASSO regression algorithm was used for feature selection, and 10-fold cross-validation was used to determine the parameters among which the key genes associated with the patient survival cycle were screened. The genes obtained from the LASSO regression were then subjected to multifactorial Cox regression analysis, and the multifactorial regression coefficients of each gene were calculated to construct a risk score equation. Based on the median risk score, the patients were divided into high-risk and low-risk groups. Kaplan–Meier survival curve analysis was used to compare the overall survival time of the two groups, and the predictive value of the genetic markers was evaluated using time-related ROC.

## Single-Gene Enrichment Analysis (GSEA)

We obtained the GSEA software (version 3.0) from the GSEA website, divided the samples into high and low expression groups based on the median value of gene expression levels, and downloaded the c2.cp.kegg.v7.4.symbols.gmt and h.all.v7.4.symbols.gmt from the Molecular Signatures Database (symbols.gmt subsets) to evaluate relevant pathways and molecular mechanisms based on gene expression profiles and phenotypic groupings, setting a minimum gene set of five and a maximum gene

set of 5000, with 1000 resamplings and a screening condition of FDR < 0.25 and P < 0.05.

## Statistical Analysis

R software (v4.0.3) was used for data analysis, and the Wilcoxon rank sum test was used between gene and miRNA expression groups in the data samples. Cox regression analysis was performed using SPSS 25.0, and P < 0.05 was considered statistically significant.

# RESULTS

## Screening of Differentially Expressed Genes in Wilms Tumor

The GEO database was used to obtain the WT-related miRNA expression dataset GSE57370, which included 62 WT tissue samples and four normal kidney tissue samples, and the mRNA expression dataset GSE73209, which included 32 WT tissue samples and six normal kidney tissue samples. Using |logFC|≥1 and adjusted P < 0.05 as screening thresholds, five upregulated miRNAs and 45 downregulated miRNAs were obtained in GSE57370 (**Figure 1A**), and a total of 58 upregulated genes and 459 downregulated genes were obtained in GSE73209 (**Figure 1B**). Lastly, one intersection gene, hsa-miR-30c-5p (**Figure 1C**), was obtained for both sets of differentially expressed genes.

## Target Gene Prediction of hsa-miR-30c-5p

RNA-seq data of TARGET-WT were downloaded from the UCSC Xena platform (http://xena.ucsc.edu/), in which a total of 126 cancer tissue samples and six para-carcinoma tissue samples were obtained, with |logFC|≥1 and adjusted P < 0.05 as screening thresholds, and 2217 and 2059 upregulated and downregulated genes, respectively (**Figure 2A**). The miRDB database was used to predict the target genes for hsa-miR-30c-5p, and 1545 target genes were obtained. The differentially expressed genes and target genes overlapped separately, and 384 genes were obtained (**Figure 2B**). Lastly, hsa-miR-30c-5p was related to 384 genes, as shown in **Figure 2C**.

## Functional Enrichment Analysis of 384 Genes

Subsequently, 384 genes were analyzed for the KEGG pathway and GO functional enrichment. KEGG involved a total of 20 pathways, mainly enriched in microRNAs in cancer, other types of O-glycan biosynthesis, and Ubiquitin mediated proteolysis (**Figure 3A**). GO enrichment analysis showed that 384 genes were mainly enriched in the nucleoplasm, regulation of the cellular metabolic process, etc. (**Figures 3B–D**).

## Kaplan–Meier Survival Analysis of 384 Genes

Kaplan–Meier survival analysis of 384 genes was used to analyze the relationship between gene expression and the overall survival of WTs. The results showed that 25 genes were significantly associated with the overall survival of Wilms tumor, namely ADRA2A, BCL6, CA10, CCNA1, CTHRC1, DGKD, EPB41L4B, ERRFI1, GALNT3, JAM2, LRRC40, MTF2, NCEH1 NEBL,

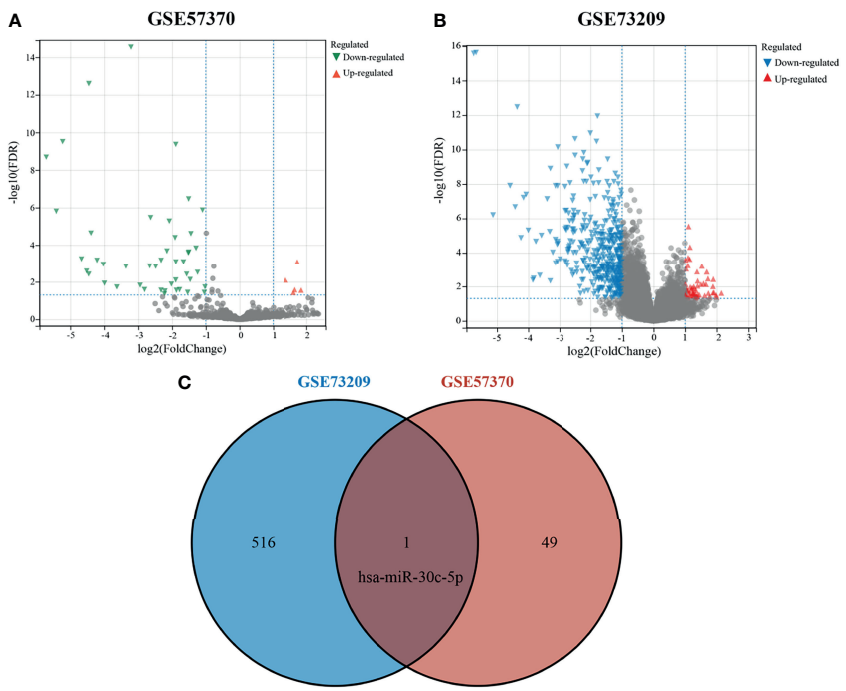



**FIGURE 1** | Differentially expressed genes in Wilms tumor. **(A)** Volcano plot showing differentially expressed miRNAs in GSE57370; **(B)** Volcano plot showing differentially expressed genes in GSE73209; **(C)** Venn plot showing intersecting genes.

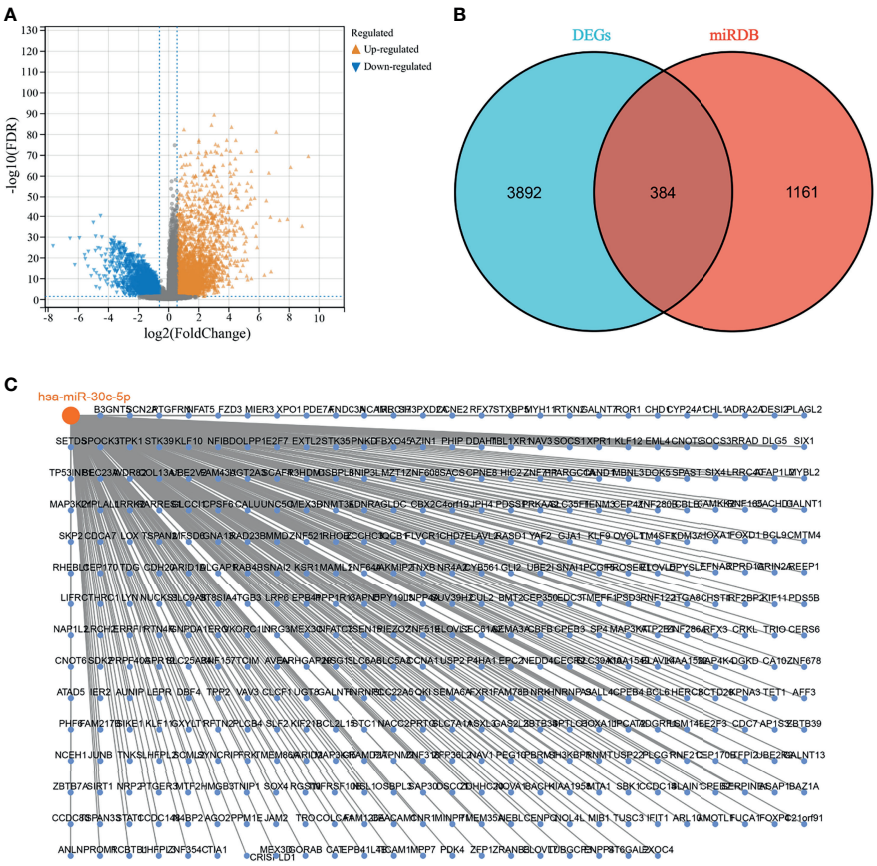



**FIGURE 2** | hsa-miR-30c-5p target genes. **(A)** Wilms tumor differentially expressed genes in the Target database; **(B)** Venn diagram showing the intersection of differentially expressed genes and target genes; **(C)** hsa-miR-30c-5p with its target genes.

OSBPL3, PDS5B, PDSS1, RFX3, ROR1, RTKN2, SLAIN1, SPTLC3, TICAM1, TUBGCP3, and ZFP36L2 (see **Table 1**).

## Construction of a Prognostic Risk Model

The LASSO regression model screened 25 genes to identify key genes affecting WT prognosis, and the model was optimal when the number of genetic variables included in the model was 12 (lambda.min = 0.0238, **Figures 4A, B**), which were the key genes associated with the prognosis of WT patients, namely BCL6, CCNA1, CTHRC1, DGKD EPB41L4B, ERRFI1, LRRC40, NCEH1, NEBL, PDSS1, ROR1, and RTKN2. Moreover, a prediction model based on the 12 gene signatures was constructed (**Figure 4C**) whose predicted risk scores consisted mainly of the following:

$$\begin{aligned}\text{Risk score} = {} & (-0.2011) * \text{BCL6} + (0.1837) * \text{CCNA1} + \\ & (-0.1889) * \text{CTHRC1} + (0.1227) * \text{DGKD} \\ & + (-0.0891) * \text{EPB41L4B} + (0.5903) * \text{ERRFI1} + \\ & (0.3992) * \text{LRRC40} + (-0.1926) * \text{NCEH1} \\ & + (-0.2065) * \text{NEBL} + (0.4799) * \text{PDSS1} + \\ & (-0.1969) * \text{ROR1} * (-0.2957) * \text{RTKN2}\end{aligned}$$

Risk scores were calculated according to the formula, and the median risk score was used as the threshold to divide the sample into high-risk and low-risk groups. The results for the Kaplan–Meier survival analysis showed that patients in the high-risk group had a significantly worse prognosis than those in the low-risk group (**Figure 4D**). In addition, the accuracy of the model in predicting patients' OS period was verified using subject working curves, and we found that the risk model predicted AUC values of 0.684, 0.762, and 0.774 for the prognosis of WT patients at 1, 3, and 5 years, respectively. These results indicate that the model has some accuracy in predicting the prognosis survival of WT patients (**Figure 4E**).

## Univariate and Multifactor Cox Regression Analysis of Risk Score Grouping and Clinicopathological Indicators

Single-factor Cox analysis was used to analyze the relationship between clinicopathological indicators and prognosis, and the results demonstrated that risk score grouping and tumor stage were significantly associated with patient prognosis (**Figure 5A**). To adjust the interaction between variables and to understand the independent prognostic value of variables, indicators with

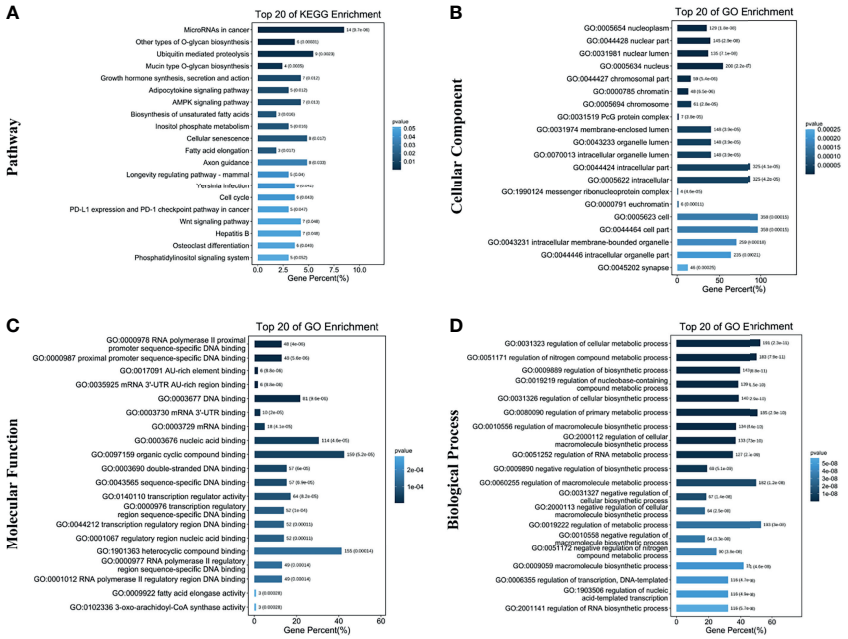



**FIGURE 3** | Functional enrichment analysis of target genes. **(A)** Top 20 terms in the KEGG pathway enrichment; **(B)** Top 20 terms in the cellular component enrichment; **(C)** Top 20 terms in the molecular function enrichment; **(D)** Top 20 terms in the biological process enrichment.

**TABLE 1** | The 25 genes related to overall survival according to Kaplan–Meier survival analysis.

| Gene | P-value | HR | 95% CI |
|---|---|---|---|
| ADRA2A | 0.028 | 0.544 | 0.315-0.937 |
| BCL6 | 0.002 | 0.409 | 0.233-0.719 |
| CA10 | 0.004 | 0.445 | 0.255-0.778 |
| CCNA1 | 0.017 | 1.951 | 1.126-3.380 |
| CTHRC1 | 0.011 | 0.489 | 0.282-0.848 |
| DGKD | 0.048 | 1.731 | 1.004-2.984 |
| EPB41L4B | 0.018 | 0.519 | 0.302-0.892 |
| ERRFI1 | 0.016 | 1.961 | 1.132-3.398 |
| GALNT3 | 0.020 | 0.521 | 0.301-0.903 |
| JAM2 | 0.044 | 0.572 | 0.332-0.986 |
| LRRC40 | 0.046 | 1.732 | 1.009-2.973 |
| MTF2 | 0.029 | 1.833 | 1.063-3.161 |
| NCEH1 | 0.047 | 0.580 | 0.339-0.993 |
| NEBL | 0.047 | 0.578 | 0.336-0.992 |
| OSBPL3 | 0.008 | 0.472 | 0.272-0.818 |
| PDS5B | 0.025 | 1.863 | 1.080-3.215 |
| PDSS1 | 0.044 | 1.743 | 1.016-2.991 |
| RFX3 | 0.031 | 1.832 | 1.057-3.174 |
| ROR1 | 0.042 | 0.571 | 0.332-0.980 |
| RTKN2 | 0.025 | 1.864 | 1.081-3.213 |
| SLAIN1 | 0.040 | 1.759 | 1.025-3.018 |
| SPTLC3 | 0.007 | 0.468 | 0.268-0.816 |
| TICAM1 | 0.019 | 0.521 | 0.302-0.898 |
| TUBGCP3 | 0.026 | 1.856 | 1.076-3.200 |
| ZFP36L2 | 0.016 | 0.510 | 0.294-0.884 |

significant single-factor analysis were introduced into the model for multifactor regression analysis, and the results of Jie showed that risk score grouping and tumor stage could be used as independent predictors of patient prognosis for WT (**Figure 5B**).

## Single-Factor and Multifactor Cox Analysis for 12 Genes

Univariate and multifactor Cox analyses were performed for the 12 genes in the model. The results of the univariate analysis

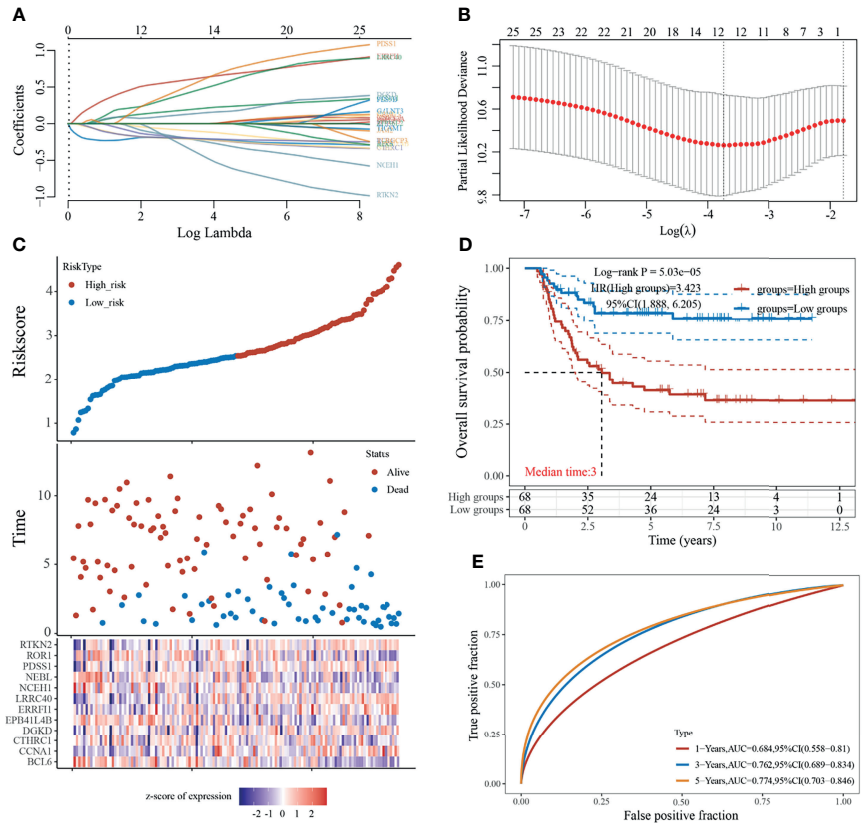



**FIGURE 4** | Construction of prognostic risk model. **(A)** Coefficients of selected features are shown by lambda parameter; **(B)** Partial likelihood deviance versus log (λ) was drawn using the LASSO Cox regression model; **(C)** Risk score, survival time, and survival status in dataset; **(D)** Kaplan–Meier survival analysis of the gene signature; **(E)** Time-dependent ROC analysis of the gene signature.

showed that BCL6, CCNA1, EPB41L4B, ERRFI1, LRRC40, NEBL, and ROR1 were significantly associated with patient prognosis (**Table 2**). Consequently, they were included in the multifactor analysis, and the results showed that ERRFI1 and ROR1 could be used as independent predictors of patient prognosis (**Table 3**).

## Expression of ERRFI1 and ROR1 and Prognosis

The expression levels of ERRFI1 and ROR1 were analyzed by integrating 126 cases of WT tissue samples and six cases of paracancerous tissue samples from the TARGET database. The results showed that ERRFI1 was significantly downregulated in the WT (**Figure 6A**), and ROR1 was significantly upregulated (**Figure 6C**). KM survival analysis showed that high expression of ERRFI1 and low expression of ROR1 were significantly associated with poor patient prognosis (**Figures 6B, D**).

## Single-Gene Functional Enrichment Analysis

The GSEA results showed that the three KEGG pathways and the HALLMARK pathway were most significantly associated with ERRFI1 high expression. Among them, high ERRFI1 expression was mainly enriched in the complement and coagulation cascade, the p53 signaling pathway, and epithelial cell signaling in *H. pylori* infection-related cells (**Figure 7A**) High expression of ERRFI1 was positive for TNF-α signaling pathway *via* NK-κB, epithelial-mesenchymal transition, and hypoxia (**Figure 7B**) signaling pathway, PPAR signaling pathway, and endocytosis (**Figure 7C**). High ROR1 expression was positive for epithelial-mesenchymal transition, PI3K-AKT-mTOR signaling, and UV response (**Figure 7D**).

# DISCUSSION

Numerous studies have reported that mRNA plays a crucial role in the tumorigenesis development of WT (24). However, with the development of detection techniques, a single mRNA expression pattern is no longer sufficient to accurately predict the prognosis of WT. In addition, the role of miRNAs in altered gene expression should not be neglected. Zhu et al. (25) reported that miR-92a-3p inhibited the proliferation, migration, and invasion of WT cells by regulating the NOTCH1 signaling pathway. Therefore, the identification of differentially expressed miRNAs represents a promising strategy. However, heterogeneous results are primarily generated by studies with relatively limited sample sizes, several

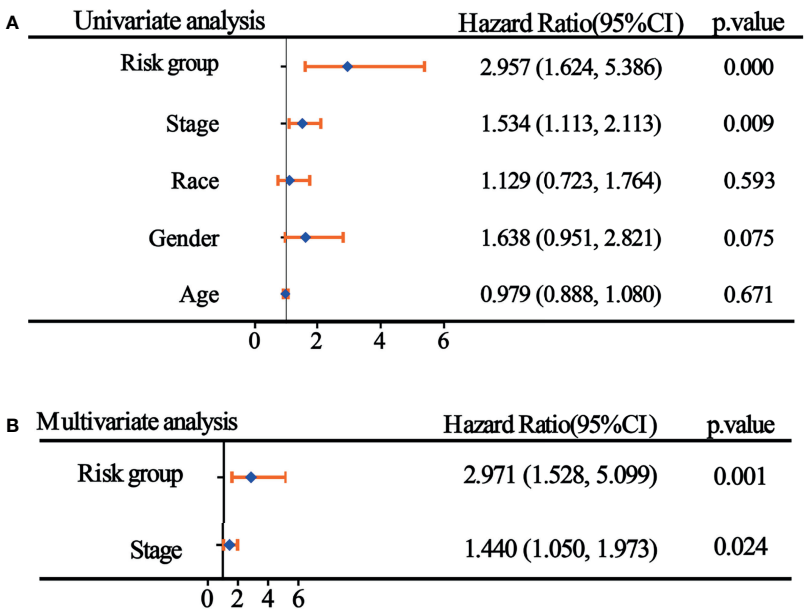



**FIGURE 5** | Univariate and multifactor Cox regression analyses of risk score groupings and clinicopathological indicators. **(A)** Results of single-factor Cox analysis; **(B)** Results of multifactor Cox regression analysis.

**TABLE 2** | Univariate cox analysis of 12 genes.

| Gene | B | SE | P-value | HR 95%CI |
|---|---|---|---|---|
| BCL6 | -0.489 | 0.158 | 0.002 | 0.613 (0.450-0.835) |
| CCNA1 | 0.227 | 0.109 | 0.037 | 1.255 (1.013-1.555) |
| CTHRC1 | -0.091 | 0.088 | 0.297 | 0.913 (0.769-1.084) |
| DGKD | 0.360 | 0.191 | 0.059 | 1.434 (0.986-2.085) |
| EPB41L4B | -0.361 | 0.183 | 0.049 | 0.697 (0.487-0.998) |
| ERRFI1 | 0.378 | 0.161 | 0.019 | 1.459 (1.064-2.001) |
| LRRC40 | 0.624 | 0.264 | 0.018 | 1.866 (1.112-3.133) |
| NCEH1 | -0.405 | 0.255 | 0.112 | 0.667 (0.404-1.099) |
| NEBL | -0.275 | 0.113 | 0.015 | 0.760 (0.609-0.948) |
| PDSS1 | 0.364 | 0.276 | 0.188 | 1.439 (0.837-2.471) |
| ROR1 | -0.246 | 0.109 | 0.025 | 0.782 (0.631-0.969) |
| RTKN2 | 0.216 | 0.178 | 0.225 | 1.240 (0.876-1.758) |

candidate miRNAs, or a lack of experimental validation. In this study, we identified one hub miRNA—miR-30c-5p—by integrating the dataset and constructed a 12-gene marker-based prognostic prediction model for WT based on the prediction of miR-30c-5p target genes and the results of WT differentially expressed gene analysis.

MiR-30 c-5p (Previous ID: miR-30c) was first identified in 2002 by Lagos-Quintana et al. (26) from mouse heart and brain tissues, and its sequence is highly conserved across species. Several studies have reported that miR-30c-5p is aberrantly expressed in different tumor tissues, sera, and cell lines and is associated with clinical features and prognostic factors in a variety of cancers, including lung cancer (27), breast cancer (28), and colorectal cancer (29). In 2010, Heinzelmann et al. (30) screened 12 miRNAs, including miR-30c-5p, which were lowly expressed in highly invasive renal clear cell carcinoma with early metastasis. In an in-depth study, it was found that miR-30c-5p expression levels were significantly lower in primary tumors with metastasis compared to normal kidney tissue and primary renal clear cell carcinoma without metastasis, and miR-30c-5p expression levels were significantly correlated with the 5-year progression-free survival of patients (31). Thus, miR-30c-5p may be a useful indicator for the early prediction of renal cancer metastasis, and different expression levels may be associated with specific distant metastasis. This study focused on the role of the target genes of miR-30c-5p in WT. We predicted the target genes of miR-30c-5p and obtained a total of 384 genes in combination with differentially expressed genes in TARGET-WT. Furthermore, KM survival analysis of these 384 genes was performed to obtain 25 genes associated with cervical cancer prognosis, and LASSO Cox regression was used as a machine-learning algorithm to construct a prognostic risk model with 12 gene signatures (i.e.,

**TABLE 3** | Multivariate cox analysis of genes.

| Gene | B | SE | P-value | HR 95%CI |
|---|---|---|---|---|
| BCL6 | -0.284 | 0.194 | 0.144 | 0.753 (0.515-1.102) |
| CCNA1 | 0.207 | 0.129 | 0.109 | 1.230 (0.955-1.585) |
| EPB41L4B | -0.005 | 0.230 | 0.983 | 0.995 (0.634-1.561) |
| ERRFI1 | 0.620 | 0.183 | 0.001 | 1.858 (1.298-2.660) |
| LRRC40 | 0.375 | 0.329 | 0.254 | 1.455 (0.763-2.773) |
| NEBL | -0.235 | 0.140 | 0.094 | 0.790 (0.600-1.041) |
| ROR1 | -0.249 | 0.126 | 0.048 | 0.780 (0.609-0.998) |

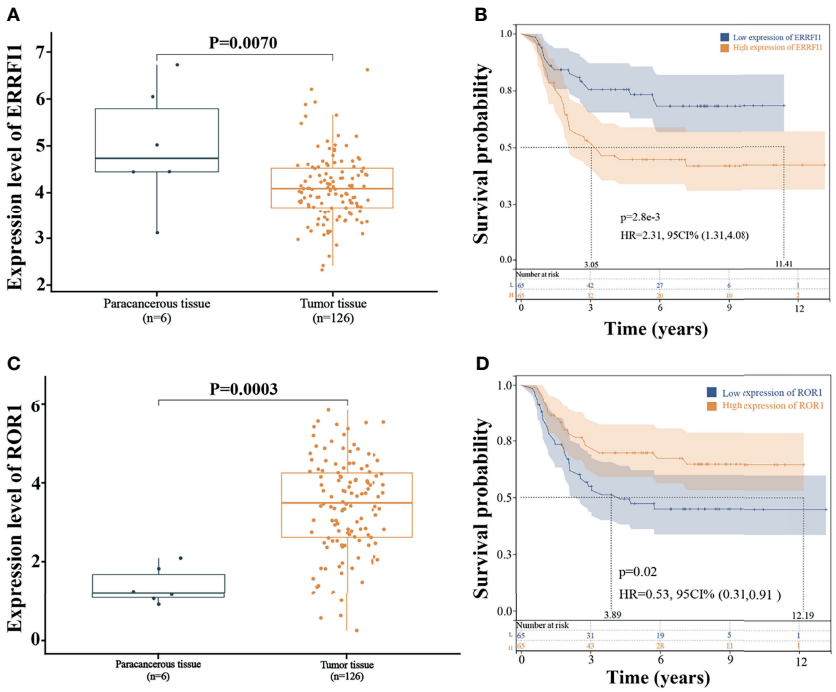



**FIGURE 6** | Expression and prognosis of ERRFI1 and ROR1. **(A)** Expression level of ERRFI1 in the TARGET-WT database; **(B)** Prognosis analysis of ERRFI1 in the TARGET-WT database; **(C)** Expression level of ROR1 in the TARGET-WT database; **(D)** Prognosis analysis of ROR1 in the TARGET-WT database.

BCL6, CCNA1, CTHRC1, DGKD, EPB41L4B, ERRFI1, LRRC40, NCEH1, NEBL, PDSS1, ROR1, and RTKN2), and the patients were divided into high-risk and low-risk groups according to risk scores. The prognosis of patients in the high-risk group was significantly worse than that in the low-risk group, and the working curves of subjects were used to verify that the model had good predictive ability for 1, 3, and 5-year prognoses of WT patients, and the risk group could be used as an independent predictor of WT prognosis in patients. Previously, gene and miRNA signatures were developed for WT prognosis prediction. However, the number of models is far from enough.

Finally, we performed univariate and multivariate Cox analyses for each of the 12 genes in the model and found that ERRFI1 and ROR1 served as independent predictors. ERBB receptor Feedback Inhibitor 1 (ERRFI1), the product of mitogen-inducible gene 6, through its ERBB-binding region, docks with the EGFR kinase structural domain docking, inhibiting EGFR activation and downstream signaling (32). We found low expression of ERRFI1 in WT cancer tissues after comparing its expression in WT, and it has been reported that ERRFI1 is also frequently mutated or downregulated in breast cancer (33), lung cancer (34), and glioblastoma (35). However, we found that high ERRFI1 expression was significantly associated with a poor prognosis in WT patients.

ERRFI1 is considered a tumor suppressor that directly inhibits epidermal growth factor receptors. However, new studies have found that the role of ERRFI1 depends on EGFR levels; therefore, in a low EGFR setting, downregulation of ERRFI1 leads to higher migration rates and promotes cell growth (36). In the Jäger K et al. (37) human study, ERRFI1 upregulation was found to be significantly associated with poor prognosis in metastatic melanoma, and in the present study, ERRFI1 upregulation was similarly found to be associated with poor prognosis in WT. Receptor tyrosine kinase-like orphan

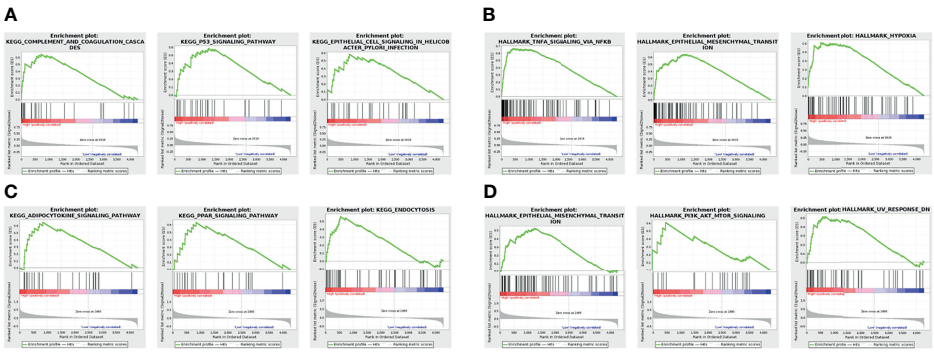



**FIGURE 7** | Single-gene functional enrichment analysis. **(A)** Results of ERRFI1 enrichment analysis in the KEGG pathway; **(B)** Results of ERRFI1 enrichment analysis in the HALLMARK pathway; **(C)** Results of ROR1 enrichment analysis in the KEGG pathway; **(D)** Results of ROR1 enrichment analysis in the HALLMARK pathway.

receptor 1 (ROR1) is a member of the ROR family of type I transmembrane receptors with ligand-bound extracellular and intracellular tyrosine kinase domains.

During embryogenesis, ROR1 plays a physiological role in neural, auditory, skeletal, and vascular organogenesis, but studies have shown that ROR1 is absent or expressed at low levels in most adult tissues (38, 39). However, as an oncoprotein, ROR1 can reemerge in hematological and solid tumors, especially in histologically advanced tumors, where ROR1 may promote tumor cell migration through Wnt5a signaling or interaction with other receptors (40, 41). In the present study, we found high expression of ROR1 in WT cancer tissues. Hodjattallah Rabbani et al. (42) verified that detecting a high level of ROR1 expression in blood cells may help in the early detection of renal malignancies. Notably, Kaplan–Meier survival analysis found that low ROR1 expression was significantly associated with poor patient prognosis, a result that is contrary to most studies that have partially demonstrated poor prognosis in patients with high ROR1 expression, such as ovarian cancer (43), colorectal cancer (44), and so on. Therefore, the prognostic predictive role of ROR1 expression in WT remains to be explored in depth.

In summary, this study constructed a 12-gene signature prognostic risk model based on the target genes of miR-30c-5p and determined that ERRFI1 and ROR1 could be used as independent predictors of WT prognosis in patients. However, this study has several shortcomings. These include the lack of biological behavior studies for the identified ERRFI1 and ROR1 and the fact that there are few studies on ERRFI1 and ROR1 in WT. However, it also has potential research value. In addition, WT is a relatively rare primary malignancy, and more clinical samples and survival information are needed to validate the results of this study. Although the risk model constructed in this study displayed better performance, the LASSO Cox algorithm used will adjust the parameters appropriately during the calculation to avoid degradation of the algorithm's performance (45). Therefore, more accurate machine-learning methods, a larger number of clinical samples, and *in vitro* experiments need to be further developed in future studies.

## 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

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

## REFERENCES

1. Aldrink JH, Heaton TE, Dasgupta R, Lautz TB, Malek MM, Abdessalam SF, et al. Update on Wilms Tumor. *J Pediatr Surg* (2019) 54(3):390–7. doi: 10.1016/j.jpedsurg.2018.09.005
2. Pater L, Melchior P, Rübe C, Cooper BT, McAleer MF, Kalapurakal JA, et al. Wilms Tumor. *Pediatr Blood Cancer* (2021) 68:e28257. doi: 10.1002/pbc.28257
3. Spreafico F, Pritchard Jones K, Malogolowkin MH, Bergeron C, Hale J, De Kraker J, et al. Treatment of Relapsed Wilms Tumors: Lessons Learned. *Expert Rev Anticancer Ther* (2009) 9(12):1807–15. doi: 10.1586/era.09.159
4. Trah J, Arand J, Oh J, Pagerols-Raluy L, Trochimiuk M, Appl B, et al. Lithocholic Bile Acid Induces Apoptosis in Human Nephroblastoma Cells: A non-Selective Treatment Option. *Sci Rep* (2020) 10(1):1–8. doi: 10.1038/s41598-020-77436-w
5. S Irtan, PF Ehrlich, K Pritchard-Jones eds. Wilms Tumor:"State-of-the-Art" Update, 2016. In: *Seminars in Pediatric Surgery*. Elsevier. doi: 10.1053/j.sempedsurg.2016.09.003
6. Gratias EJ, Dome JS, Jennings LJ, Chi Y-Y, Tian J, Anderson J, et al. Association of Chromosome 1q Gain With Inferior Survival in Favorable-Histology Wilms Tumor: A Report From the Children's Oncology Group. *J Clin Oncol* (2016) 34(26):3189. doi: 10.1200/JCO.2015.66.1140

7. Dome JS, Graf N, Geller JI, Fernandez CV, Mullen EA, Spreafico F, et al. Advances in Wilms Tumor Treatment and Biology: Progress Through International Collaboration. *J Clin Oncol* (2015) 33(27):2999. doi: 10.1200/JCO.2015.62.1888
8. Lopyan NM, Ehrlich PF. Surgical Management of Wilms Tumor (Nephroblastoma) and Renal Cell Carcinoma in Children and Young Adults. *Surg Oncol Clinics* (2021) 30(2):305–23. doi: 10.1016/j.soc.2020.11.002
9. Ross A, Gomez O, Wang X, Lu Z, Abdelhafeez H, Davidoff A, et al. Timing of Adjuvant Chemotherapy After Laparotomy for Wilms Tumor and Neuroblastoma. *Pediatr Surg Int* (2021) 37(11):1585–92. doi: 10.1007/s00383-021-04968-1
10. Schmidt A, Warmann SW, Urla C, Schaefer J, Fideler F, Fuchs J. Patient Selection and Technical Aspects for Laparoscopic Nephrectomy in Wilms Tumor. *Surg Oncol* (2019) 29:14–9. doi: 10.1016/j.suronc.2019.02.007
11. Treger TD, Chowdhury T, Pritchard-Jones K, Behjati S. The Genetic Changes of Wilms Tumour. *Nat Rev Nephrol* (2019) 15(4):240–51. doi: 10.1038/s41581-019-0112-0
12. Zhou N, Yan B, Ma J, Jiang H, Li L, Tang H, et al. Expression of TCF3 in Wilms' Tumor and Its Regulatory Role in Kidney Tumor Cell Viability, Migration and Apoptosis In Vitro. *Mol Med Rep* (2021) 24(3):1–8. doi: 10.3892/mmr.2021.12281
13. Jiang D, Chen J, Fan Z, Tan D, Zhao J, Shi H, et al. CRISPR/Cas9-Induced Disruption of Wt1a and Wt1b Reveals Their Different Roles in Kidney and Gonad Development in Nile Tilapia. *Dev Biol* (2017) 428(1):63–73. doi: 10.1016/j.ydbio.2017.05.017
14. Hong B, Dong R. Research Advances in the Targeted Therapy and Immunotherapy of Wilms Tumor: A Narrative Review. *Trans Cancer Res* (2021) 10(3):1559–67. doi: 10.21037/tcr-20-3302
15. Mardanpour K, Rahbar M, Mardanpour S, Mardanpour N, Rezaei M. CD8+ T-Cell Lymphocytes Infiltration Predict Clinical Outcomes in Wilms' Tumor. *Tumor Biol* (2020) 42(12):1010428320975976. doi: 10.1177/1010428320975976
16. Yuan N, Zhang G, Bie F, Ma M, Ma Y, Jiang X, et al. Integrative Analysis of Lncrnas and Mirnas With Coding Rnas Associated With Cerna Crosstalk Network in Triple Negative Breast Cancer. *OncoTargets Ther* (2017) 10:5883. doi: 10.2147/OTT.S149308
17. Huang CG, Li FX, Pan S, Xu CB, Dai JQ, Zhao XH. Identification of Genes Associated With Castration–Resistant Prostate Cancer by Gene Expression Profile Analysis. *Mol Med Rep* (2017) 16(5):6803–13. doi: 10.3892/mmr.2017.7488
18. Charlton J, Pavasovic V, Pritchard-Jones K. Biomarkers to Detect Wilms Tumors in Pediatric Patients: Where Are We Now? *Future Oncol* (2015) 11(15):2221–34. doi: 10.2217/fon.15.136
19. Cone EB, Dalton SS, Van Noord M, Tracy ET, Rice HE, Routh JC. Biomarkers for Wilms Tumor: A Systematic Review. *J Urol* (2016) 196(5):1530–5. doi: 10.1016/j.juro.2016.05.100
20. Lin XD, Wu YP, Chen SH, Sun XL, Ke ZB, Chen DN, et al. Identification of a Five-Mrna Signature as a Novel Potential Prognostic Biomarker in Pediatric Wilms Tumor. *Mol Genet genomic Med* (2020) 8(1):e1032. doi: 10.1002/mgg3.1032
21. Wegert J, Ishaque N, Vardapour R, Geörg C, Gu Z, Bieg M, et al. Mutations in the SIX1/2 Pathway and the DROSHA/DGCR8 Mirna Microprocessor Complex Underlie High-Risk Blastemal Type Wilms Tumors. *Cancer Cell* (2015) 27(2):298–311. doi: 10.1016/j.ccell.2015.01.002
22. Ludwig N, Werner TV, Backes C, Trampert P, Gessler M, Keller A, et al. Combining Mirna and Mrna Expression Profiles in Wilms Tumor Subtypes. *Int J Mol Sci* (2016) 17(4):475. doi: 10.3390/ijms17040475
23. Karlsson J, Valind A, Jansson C, O'Sullivan MJ, Mengelbier LH, Gisselsson D. Aberrant Epigenetic Regulation in Clear Cell Sarcoma of the Kidney Featuring Distinct DNA Hypermethylation and EZH2 Overexpression. *Oncotarget* (2016) 7(10):11127. doi: 10.18632/oncotarget.7152
24. Martins AG, Pinto AT, Domingues R, Cavaco BM. Identification of a Novel CTR9 Germline Mutation in a Family With Wilms Tumor. *Eur J Med Genet* (2018) 61(5):294–9. doi: 10.1016/j.ejmg.2017.12.010
25. Zhu S, Zhang L, Zhao Z, Fu W, Fu K, Liu G, et al. Microrna–92a–3p Inhibits the Cell Proliferation, Migration and Invasion of Wilms Tumor by Targeting NOTCH1. *Oncol Rep* (2018) 40(2):571–8. doi: 10.3892/or.2018.6458
26. Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T. Identification of Tissue-Specific Microrna From Mouse. *Curr Biol* (2002) 12(9):735–9. doi: 10.1016/S0960-9822(02)00809-6
27. Zhou Y, Shi H, Du Y, Zhao G, Wang X, Li Q, et al. Lncrna DLEU2 Modulates Cell Proliferation and Invasion of Non-Small Cell Lung Cancer by Regulating Mir-30c-5p/SOX9 Axis. *Aging (Albany NY)* (2019) 11(18):7386. doi: 10.18632/aging.102226
28. Pei B, Li T, Qian Q, Fan W, He X, Zhu Y, et al. Downregulation of Microrna-30c-5p was Responsible for Cell Migration and Tumor Metastasis *via* COTL1-Mediated Microfilament Arrangement in Breast Cancer. *Gland Surg* (2020) 9(3):747. doi: 10.21037/gs-20-472
29. Guo Y, Guo Y, Chen C, Fan D, Wu X, Zhao L, et al. Circ3823 Contributes to Growth, Metastasis and Angiogenesis of Colorectal Cancer: Involvement of Mir-30c-5p/TCF7 Axis. *Mol Cancer* (2021) 20(1):1–21. doi: 10.1186/s12943-021-01372-0
30. Heinzelmann J, Henning B, Sanjmyatav J, Posorski N, Steiner T, Wunderlich H, et al. Specific Mirna Signatures Are Associated With Metastasis and Poor Prognosis in Clear Cell Renal Cell Carcinoma. *World J Urol* (2011) 29(3):367–73. doi: 10.1007/s00345-010-0633-4
31. Heinzelmann J, Unrein A, Wickmann U, Baumgart S, Stapf M, Szendroi A, et al. Micrornas With Prognostic Potential for Metastasis in Clear Cell Renal Cell Carcinoma: A Comparison of Primary Tumors and Distant Metastases. *Ann Surg Oncol* (2014) 21(3):1046–54. doi: 10.1245/s10434-013-3361-3
32. Cairns J, Fridley BL, Jenkins GD, Zhuang Y, Yu J, Wang L. Differential Roles of ERRFI1 in EGFR and AKT Pathway Regulation Affect Cancer Proliferation. *EMBO Rep* (2018) 19(3):e44767. doi: 10.15252/embr.201744767
33. Mojica CAR, Ybañez WS, Olarte KCV, Poblete ABC, Bagamasbad PD. Differential Glucocorticoid-Dependent Regulation and Function of the ERRFI1 Gene in Triple-Negative Breast Cancer. *Endocrinology* (2020) 161(7):bqaa082. doi: 10.1210/endocr/bqaa082
34. Wu K, Chen X, Feng J, Zhang S, Xu Y, Zhang J, et al. Capilliposide C From Lysimachia Capillipes Restores Radiosensitivity in Ionizing Radiation-Resistant Lung Cancer Cells Through Regulation of ERRFI1/EGFR/STAT3 Signaling Pathway. *Front Oncol* (2021) 11:1017. doi: 10.3389/fonc.2021.644117
35. Duncan CG, Killela PJ, Payne CA, Lampson B, Chen WC, Liu J, et al. Integrated Genomic Analyses Identify ERRFI1 and TACC3 as Glioblastoma-Targeted Genes. *Oncotarget* (2010) 1(4):265. doi: 10.18632/oncotarget.137
36. Vu HL, Rosenbaum S, Capparelli C, Purwin TJ, Davies MA, Berger AC, et al. MIG6 is MEK Regulated and Affects EGF-Induced Migration in Mutant NRAS Melanoma. *J Invest Dermatol* (2016) 136(2):453–63. doi: 10.1016/j.jid.2015.11.012
37. Jäger K, Larribère L, Wu H, Weiss C, Gebhardt C, Utikal J. Expression of Neural Crest Markers GLDC and ERRFI1 Is Correlated With Melanoma Prognosis. *Cancers* (2019) 11(1):76. doi: 10.3390/cancers11010076
38. Dave H, Anver MR, Butcher DO, Brown P, Khan J, Wayne AS, et al. Restricted Cell Surface Expression of Receptor Tyrosine Kinase ROR1 in Pediatric B-Lineage Acute Lymphoblastic Leukemia Suggests Targetability With Therapeutic Monoclonal Antibodies. *PloS One* (2012) 7(12):e52655. doi: 10.1371/journal.pone.0052655
39. Fukuda T, Chen L, Endo T, Tang L, Lu D, Castro JE, et al. Antisera Induced by Infusions of Autologous Ad-CD154-Leukemia B Cells Identify ROR1 as an Oncofetal Antigen and Receptor for Wnt5a. *Proc Natl Acad Sci* (2008) 105(8):3047–52. doi: 10.1073/pnas.0712148105
40. Yu J, Chen L, Cui B, Wu C, Choi MY, Chen Y, et al. Cirmtuzumab Inhibits Wnt5a-Induced Rac1 Activation in Chronic Lymphocytic Leukemia Treated With Ibrutinib. *Leukemia* (2017) 31(6):1333–9. doi: 10.1038/leu.2016.368
41. Li C, Wang S, Xing Z, Lin A, Liang K, Song J, et al. A ROR1–HER3–Lncrna Signalling Axis Modulates the Hippo–YAP Pathway to Regulate Bone Metastasis. *Nat Cell Biol* (2017) 19(2):106–19. doi: 10.1038/ncb3464
42. Rabbani H, Ostadkarampour M, Manesh AHD, Basiri A, Tehrani MJ, Forouzesh F. Expression of ROR1 in Patients With Renal Cancer-a Potential Diagnostic Marker. *Iranian Biomed J* (2010) 14(3):77.
43. Zhang H, Qiu J, Ye C, Yang D, Gao L, Su Y, et al. ROR1 Expression Correlated With Poor Clinical Outcome in Human Ovarian Cancer. *Sci Rep* (2014) 4(1):1–7. doi: 10.1038/srep05811
44. Zhou J-K, Zheng Y-Z, Liu X-S, Gou Q, Ma R, Guo C-L, et al. ROR1 Expression as a Biomarker for Predicting Prognosis in Patients With Colorectal Cancer. *Oncotarget* (2017) 8(20):32864. doi: 10.18632/oncotarget.15860

45. Scott IA, Cook D, Coiera EW, Richards B. Machine Learning in Clinical Practice: Prospects and Pitfalls. *Med J Aust* (2019) 211(5):203–5. doi: 10.5694/mja2.50294

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 27 April 2022
doi: 10.3389/fonc.2022.841530

# A Prognostic Gene Signature for Hepatocellular Carcinoma

*Rong Chen[1], Meng Zhao[2], Yanli An[3,4], Dongfang Liu[3], Qiusha Tang[3] and Gaojun Teng[4]**

[1] Department of Oncology, Zhongda Hospital, Nanjing, China, [2] School of Basic Medicine, Zhengzhou University, Zhengzhou, China, [3] Medical School of Southeast University, Nanjing, China, [4] Department of Radiology, Medical School of Southeast University, Nanjing, China

Hepatocellular carcinoma is the third most common cause of cancer-related deaths in China and immune-based therapy can improve patient outcomes. In this study, we investigated the relationship between immunity-associated genes and hepatocellular carcinoma from the prognostic perspective. The data downloaded from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) and the Gene Expression Omnibus (GEO) was screened for gene mutation frequency using the maftools package. Immunity-associated eight-gene signature with strong prognostic ability was constructed and proved as an independent predictor of the patient outcome in LIHC. Seven genes in the immune-related eight-gene signature were strongly associated with the infiltration of M0 macrophages, resting mast cells, and regulatory T cells. Our research may provide clinicians with a quantitative method to predict the prognosis of patients with liver cancer, which can assist in the selection of the optimal treatment plan.

**Keywords: hepatocellular carcinoma, bioinformatics analysis, immune-related prognostic genes, immune cells, gene signature**

OPEN ACCESS

***Edited by:***
*Anton A. Buzdin, European Organisation for Research and Treatment of Cancer, Belgium*

***Reviewed by:***
*Nikolay Mikhaylovich Borisov, Moscow Institute of Physics and Technology, Russia*
*Qiaoyou Weng, Peking University, China*

***Correspondence:**
*Gaojun Teng*
*101008344@seu.edu.cn*

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

***Received:*** *11 January 2022*
***Accepted:*** *28 March 2022*
***Published:*** *27 April 2022*

***Citation:***
*Chen R, Zhao M, An Y, Liu D, Tang Q and Teng G (2022) A Prognostic Gene Signature for Hepatocellular Carcinoma. Front. Oncol. 12:841530. doi: 10.3389/fonc.2022.841530*

## INTRODUCTION

Hepatocellular carcinoma (HCC) is the fifth most common cancer in the world (1, 2). China accounts for 55% of new HCC cases and HCC-related deaths annually (3). Standard treatment approaches for HCC include surgery, liver transplantation, targeted therapy, radiotherapy, immunotherapy, and chemotherapy; however, the therapeutic effect is still not satisfactory (4). Thus, in China a 5-year survival rate for patients with HCC is only 14.1% (5). The overall poor outcome can be attributed to the fact that patients are already at the advanced stage when diagnosed and only less than 30% of them can be operated (6). Therefore, in order to improve patients' survival, it is important to explore new diagnostic and therapeutic targets, including disease-specific biomarkers and prognostic molecular models.

Mutations in the TP53 gene encoding an important tumour suppressor protein are commonly found in diverse human cancers (7–9). Wild-type TP53 can activate apoptosis-related pathways to induce cancer cell death and prevent tumour growth (10), whereas loss-of-function mutations in the TP53 gene can induce uncontrolled tumour cell proliferation (11, 12), as shown for oesophageal cancer (13, 14). In breast cancer, the frequency of TP53 mutations is as high as 80%, which exceeds even that of BRCA1 mutations (15) and which can account for shorter lifespan of patients with mutated TP53 (16). In high-grade ovarian cancer, the TP53 mutation rate is approximately 97% (17). Adavosertib can increase the chemotherapeutic drug sensitivity of cancer cells harbouring

TP53 mutations (18, 19), and it was reported that in patients with platinum-sensitive ovarian cancer, adavosertib combined with paclitaxel and carboplatin can improve progression-free survival (PFS).

TP53 mutations are also very common in HCC and have been detected in 13–48% of patients (20–24). Patients with HCC and mutated TP53 had shorter overall and relapse-free survival (25). Previous studies indicate that the TP53 mutation status is associated with distinct immune reactions (26, 27); thus, in HCC TP53 genetic alterations resulted in decreased immune response (28). It has been reported that mutations in both low-density lipoprotein receptor-related protein 1B (LRP1B) and TP53 may be a prognostic biomarker predicting a better effect of immunotherapy in patients with HCC (29). The predictive value of the TP53 mutation status was also shown in the treatment and prognosis of other cancers. Thus, in squamous cell carcinoma of the head and neck (HNSCC), mutations in methylguanine-DNA methyltransferase and TP53 were related to a poorer prognosis (30). Recent studies indicate that the cooperative effect of poly (ADP-ribose) polymerase (PARP) inhibitors and ionic radiation or chemotherapy depends on the TP53 function (31, 32). Moreover, reactivation of mutant TP53 combined with olaparib resulted in more efficient inhibition of tumour growth in the preclinical model of triple negative breast cancer with a TP53 mutation (33). PARP inhibitors showed activity in a subset of colorectal cancer cell lines and preservation of the TP53 function may increase the likelihood of a favourable response (34).

As TP53 mutations play a significant role in many types of cancers including HCC, previous studies have been limited to the investigation of this particular gene (22, 23, 28). However, it is established that cancer is a heterogeneous multi-stage disease caused by the interaction of numerous gene products and signalling factors. Therefore, an integrative functional genomics approach should help in deciphering the molecular features of liver cancer. In previous studies, either the sample size was insufficient, which undermined the reliability of conclusions (35), or the records on baseline clinical features and therapeutic regimens and even the information included in the dataset were incomplete (36). Therefore, comprehensive testing and analysis are required to identify more reliable diagnostic biomarkers and therapeutic targets in HCC.

In this study, we used The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases to obtain and screen highly mutated genes in primary liver cancer, construct an immune-related gene signature, and explore the relationship between immune cells and patient prognosis.

# MATERIALS AND METHODS

## Acquisition of Liver Hepatocellular Carcinoma (LIHC) Data and Screening of Highly Mutated Genes

The gene expression data on 364 cases of LIHC were downloaded from TCGA database (37) using RTCGAToolbox (38) and used as a training set. The LIHC gene chip and clinical survival data of 115 LIHC cases contained in the GSE76427 dataset (39) were downloaded from the GEO database (40) and used as the validation set. Maftools (41) was used to identify the top 20 highly mutated genes and visualize mutations and their frequencies in all samples, which were then divided into two groups according to the presence of mutations in the gene with the highest mutation frequency.

## Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA)

To determine the pathways differentially expressed between patient groups, we performed GSEA, a computational method to detect functionally relevant genes (42, 43) and GSVA, a non-parametric approach to calculate sample-wise gene set enrichment scores for gene expression data (44). For enrichment analysis, we used GSEA 4.0.3 software and file 'c2.all.v.7.1.symbols.gmt' as the reference gene set, and performed 1,000 genome replacements to determine the standardized enrichment score for each analysis; P-values and false discovery rates less than 0.05 were considered to indicate statistical significance. File 'h.all.v7.1.symbols.gmt' was used as the reference gene set for GSVA performed with clusterProfiler (45). P-value less than 0.05 indicated statistical significance.

## Determination of the Immunity-Associated Gene Signature

Univariate Cox regression was used to analyze the association between immunity-related genes and the prognosis of patients with LIHC; forest plots were constructed for visualization. Screening of immune genes correlated with disease prognosis was used as the basis for signature construction; $P < 0.05$ was the selection criterion. We applied machine learning methods and Lasso regression, which is widely used in search of prognostic biomarkers (46), to generate a new gene combination for each iteration; 1,000 Lasso regressions were performed on candidate genes and the best gene signature was determined based on the area under the curve (AUC). Next, we calculated the risk score for each patient according to the gene expression level and divided patients into groups. The optimal prognostic immune-gene signature was verified by Cox regression analysis.

## Gene Ontology (GO) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) Analyses

GO is widely used to annotate gene functions (47), and KEGG is a common method to analyze pathway enrichment (48). For studying the functions of genes associated with LIHC prognosis and the related molecular mechanisms, an R package clusterProfiler (45) was applied to perform GO and KEGG analyses.

## External Validation of the Immunity-Related Gene Signature

The risk score for each sample in the validation set was calculated according to the optimal gene signature and used to assign

patients to high- and low-risk groups. The receiver operating characteristic (ROC) curve at different time points was used to analyze the prognostic potential of the optimal gene signature in the validation set.

## Clinical Subgroup Analysis and Nomogram Construction

We grouped patients with LIHC according to clinical characteristics (gender, age, pathological stage, TNM stage, and metastasis) and performed Kaplan-Meier analysis on the samples in each group. Next, we constructed a nomogram including the predictive information on clinical features and gene signature.

## Analysis of Correlation Between Immune Cell Infiltration and the Optimal Gene Signature

An online CIBERSORT tool (49) was used to analyze the distribution and infiltration of 22 types of immune cells in the high- and low-risk groups. Principal Component Analysis (PCA) was applied to the data to determine the difference in immune cell infiltration between the two groups. We also evaluated the inter-group differences in the composition, interaction, and infiltration of the 22 immune cell types. Further, the association between immune cell infiltration and LIHC prognosis was explored using Kaplan-Meier analysis.

## q-RCR

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 (Analytik Jena AG, qTower 3.2G; Jena, Germany) using BeyoFast SYBR Green qPCR Mix (Bio-Rad, 1708882AP, Shanghai, China). The primer sequences were as follows: GAPDH Forward: 5'-AC AGCCTCAAGATCATCAGC-3'; GAPDH Reverse: 5'-GGT CATGAGTCCTTCCACGAT-3'; CCR3 Forward: 5'- CACA AGCCAGGGAGAAGTGA-3'; CCR3 Reverse: 5'- TTTTCA CAGAGCAGGCCCAC -3'; CHGA Forward: 5'- CAGCGGT TTTGAAGATGAACTC -3'; CHGA Reverse: 5'- ACTTTT CTCTGCCTCCTTGGAA -3'; EPO Forward: 5'- GCTGCAT GTGGATAAAGCCG -3'; EPO Reverse: 5'- TGATT GTTCGGAGTGGAGCA -3'; LECT2 Forward: 5'- CTGCTCA AAGAAGTCAGAGGC -3'; LECT2 Reverse: 5'- GCGTACA CAGTAGATCCAGCA -3'; NROB1 Forward: 5'- AGGGGGTA AAGAGGCGCTA -3'; NROB1 Reverse: 5'- CTTGA TTTGTGCTCGTGGGC -3'; S100A9 Forward: 5'- GGAACG CAACATAGAGACCA -3'; S100A9 Reverse: 5'- GATCTTTT CGCACCAGCTCTT -3'; SEMA4F Forward: 5'- CCTGC CTCCCACACACTTTA -3'; SEMA4F Reverse: 5'- ACC ATCCAGTCAATCCTGCG -3'; SPP1 Forward: 5'- CAAA TACCCAGATGCTGTGGC -3'; SPP1 Reverse: 5'- TGG TCATGGCTTTCGTTGGA -3'. Transcript levels were normalized against GAPDH levels as an internal reference and were evaluated using the 2-ΔΔCt method. All experiments were repeated three times.

## Statistical Analysis

All statistical analyses were performed in R package. Cox regression analysis was applied to verify the association of patient survival with the gene signature and the expression of each signature gene. Kaplan-Meier analysis was used to evaluate the survival of patients in the high- and low-risk groups. Pearson correlation analysis was performed to determine the correlation between the prognostic gene signature and infiltration of prognosis-related immune cells. P-values and false discovery rates less than 0.05 were considered to indicate statistical significance.

# RESULTS

## Mutant Genes in LIHC

A flowchart of this study is shown in **Figure 1**. Using the maftools package (41), we identified the top 20 highly mutated genes in LIHC: SPTA1, CACNA1E, HMCN1, ARID1A, XIRP2, AXIN1, OBSCN, LRP1B, FLG, CSMD3, APOB, BCA13, RYR2, MUC4, PCLO, ALB, MUC16, CTNNB1, TTN, and TP53; among them TP53 had the highest mutation frequency (**Figure 2**).

## GSEA and GSVA Results

According to the TP53 mutation status, a total of 364 samples were divided into the TP53 and NO_TP53 groups. GSEA showed that four immune-related pathways: Hoffmann-large-to-small-

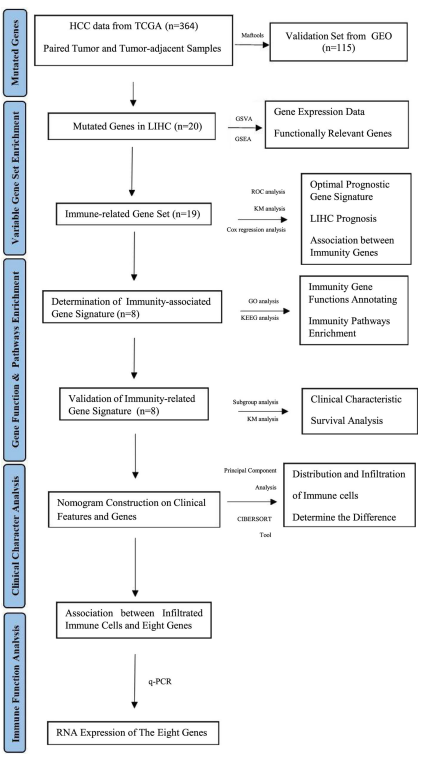



**FIGURE 1** | Bioinformatics Algorithm of prognostic gene set for hepatocellular carcinoma.

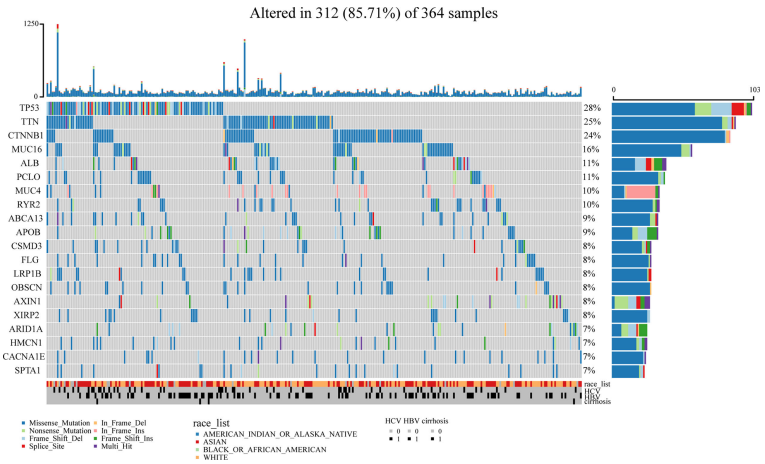



**FIGURE 2** | Identification of mutated genes. The maftools package was used and we identified the top 20 highly mutated genes in LIHC. TP53 had the highest mutation frequency among them.

pre-bil-lymphocyte-up, croonquist-IL6-deprivation-dn, mori-large-pre-bil-lymphocyte-up and lee-early-t-lymphocyte-up were enriched in the TP53 group (**Supplementary Figure 1A**). GSVA confirmed that many immune-related KEGG pathways including myc-targets-v1, orc1-signalling, ical-junction, folded-protein-response, apoptotic-spindle, f-targets, 3-pathway, m-checkpoint, response-up, c-targets-v2, glycolysis, apoptosis, 2-stat5-signalling, 3k-akt-mtor-signalling, and complement were activated were enriched in the TP53 group, which further indicating that the activation of TP53 might participate in the process of immune process (**Supplementary Figure 1B**).

## Identification of Immune-Related Prognostic Genes and Signature Construction

A single-factor Cox regression model showed that 19 immunity-related genes: BIRC5, CALCR, CCR3, CHGA, COLEC12, CXCL8, EPO, FABP6, FGF9, IKBKE, MAPT, NR0B1, S100A11, S100A2, S100A9, SEMA4F, SPP1, STC2, and TNFRSF11B were associated with LIHC prognosis (**Supplementary Figure 2**). For the accuracy of predicting the optimal gene signature for LIHC, we performed iterative Lasso Cox regression analysis, which identified a prognostic signature comprising eight genes: LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 (**Figure 3A**). ROC analysis showed that the eight-gene signature had a good predictive ability (**Figure 3B**), whereas Kaplan-Meier analysis revealed that the overall survival of patients in the low-risk group was significantly better than that in the high-risk group ($P < 0.001$; **Figure 3C**). **Figure 3D** shows the survival status, risk score distribution, and expression of the signature genes. The mortality rate was significantly higher in the high-risk than in the low-risk group and each signature gene was differentially expressed in the two groups. Cumulatively, these results indicated that the signature comprising eight immunity-associated genes could be a significant prognostic indicator in LIHC.

## Verification of the Optimal Immune-Associated Gene Signature in the External Validation Set

Analysis of the survival status, risk score distribution, and gene expression of eight-gene signature in the validation set (**Supplementary Figure 3A**) confirmed that the prognosis of patients in the low-risk group was significantly better than that in the high-risk group, thus verifying the prognostic ability of the signature. Kaplan-Meier survival analysis showed that the eight-gene signature could predict the prognosis for patients with LIHC in the external verification set ($P = 0.0017$) (**Supplementary Figure 3B**). ROC analysis of survival prognosis indicated that the eight-gene signature had a strong ability to predict 3-year (AUC = 0.71), 5-year (AUC = 0.78), and 7-year (AUC = 0.68) survival of patients with LIHC (**Supplementary Figure 3C**). Comparison of the immune-related eight-gene signature with the established LIHC prognostic biomarkers showed that the prediction based on the gene signature was significantly more reliable (**Supplementary Figure 3D**).

## Independent Predicting Ability of the Eight-Gene and Construction of a Prognostic Nomogram

Patients were regrouped and Kaplan-Meier survival analysis was performed based on clinicopathological characteristic. The results indicated that even if the clinical features were regrouped, the survival in the high-risk group was always poor ($P < 0.05$ for all; **Figure 4**). In addition, we combined clinicopathological characteristics and the immune-related eight-gene signature and constructed a prognostic nomogram (**Figure 5**), which could aid in the clinical decision regarding the treatment plan.

## Immune Cell Infiltration in the Two Risk Groups

PCA revealed the difference in immune cell infiltration between the two risk groups (**Supplementary Figure 4A**). The results of

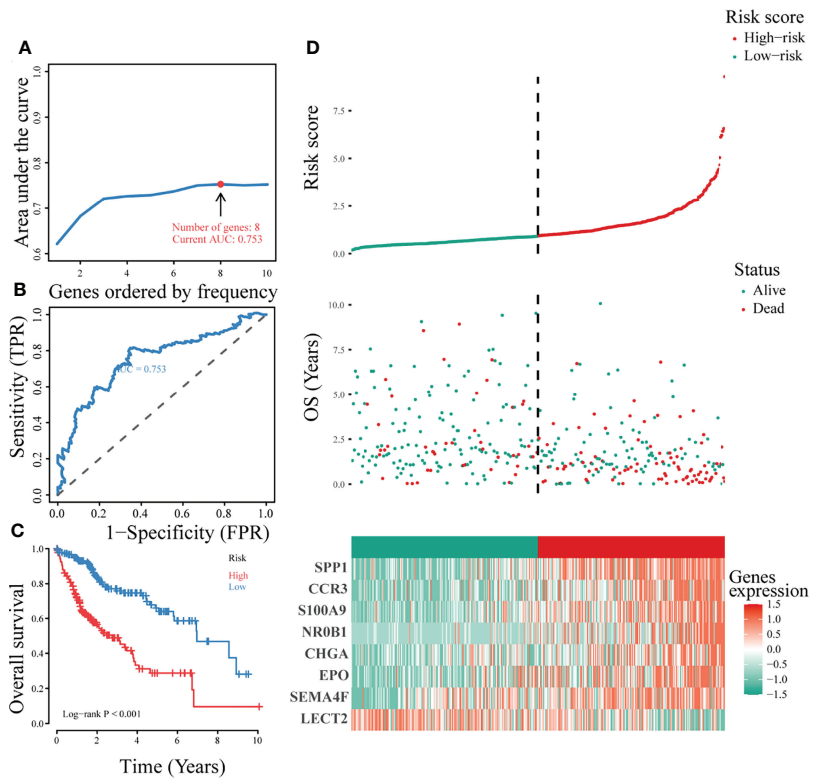



**FIGURE 3** | Construction of the optimal immune gene signature associated with LIHC prognosis. **(A)** Iterative Lasso Cox regression analysis used to construct the immune gene signature based on the size of the AUC. **(B)** ROC analysis of the optimal immunity-associated gene signature. **(C)** Kaplan-Meier curves of different risk groups. **(D)** The risk factor association diagram showing risk score distribution, survival status, and expression of the signature genes in the two risk groups.

correlation analysis showed that the infiltration of CD8+ T cells was positively correlated with that of regulatory T cells (Tregs), M1 macrophages, and follicular helper T cells (**Supplementary Figure 4B**). However, the infiltration of naïve B cells was negatively correlated with that of CD8+ T cells, macrophages (M0, M1, and M2), monocytes, resting dendritic and NK cells. Immune cell interaction network revealed that M0 macrophages, activated NK cells, naïve B cells, and resting CD4+ memory T cells had the strongest, whereas activated dendritic and mast cells, naïve CD4+ T cells, and resting dendritic cells – the weakest association with other immune cells (**Supplementary Figure 4C**). Immune cell composition analysis revealed that activated NK cells had the highest infiltration rate and activated dendritic cells – the lowest infiltration rate (**Supplementary Figure 4D**).

## Association of Immune Cell Infiltration With LIHC Prognosis and the Eight-Gene Signature

Analysis of the correlation between immune cell infiltration and prognosis showed that the infiltration of gamma delta T cells, eosinophils, and M0 and M2 macrophages indicated a poorer prognosis, whereas that of CD8+ T cells, M1 macrophages, and NK cells suggested a better prognosis for patients with LIHC (**Figure 6**). The results of the constructed correlation heat map for the signature genes revealed that M0 macrophages, resting mast cells, and Tregs showed negative correlation with CCR3, EPO, NR0B1, S100A9, SEMA4F, and SPP1, and positive correlation with LECT2 (**Figure 7**).

## The Relative RNA Expression Level of LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1

The RNA expression of LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 in normal human hepatic epithelial cells HL-02 and human hepatoma cells BEL_7402 were compared by qPCR. It was found that LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 were low expressed in human hepatoma cells compared with normal human hepatic epithelial cells (Unpaired t-test, $p<0.01$) (**Figure 8**).

# DISCUSSION

The current clinical problems of liver cancer are mainly associated with untimely diagnosis and treatment, which can be attributed to a special double blood supply structure of the liver allowing the formation of a microenvironment providing

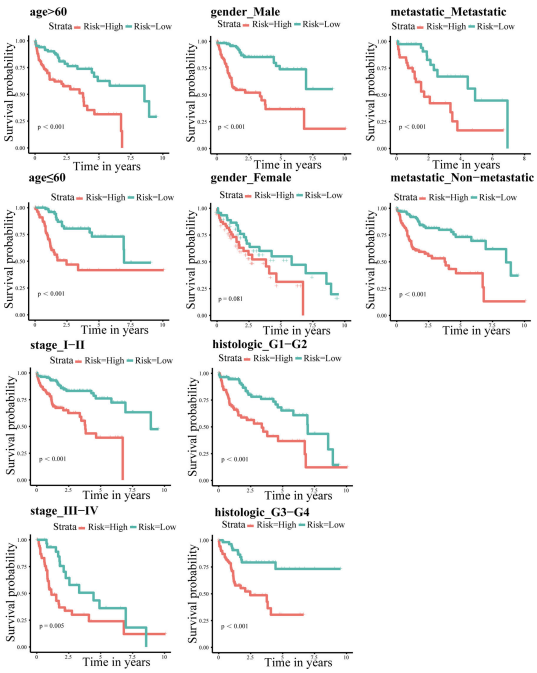



**FIGURE 4** | Kaplan-Meier survival analysis according to individual clinicopathological characteristics (age, gender, metastasis, and TNM and pathological staging). Red and green indicate high- and low-risk groups, respectively.

autoimmune tolerance (50). This phenomenon, together with the immune escape of liver cancer cells, indicates that targeted immune-therapeutics should be an effective treatment for HCC (51). In recent years, the research on the mechanism of liver cancer immunotherapy has made great progress (52), but there are still many challenges. Our study identified an immunity-related eight-gene signature that can be used as an independent prognostic indicator of the LIHC outcome, offering a quantitative clinical method to predict patient's survival. Analysis of the association between LIHC prognosis and immune cell infiltration showed that each gene in the immune-related eight-gene signature was strongly related to M0 macrophage infiltration.

Our analysis of TCGA database indicated that the mutation frequency of the TP53 gene was the highest among the genes mutated in LIHC, and GSEA and GSVA revealed important pathways enriched in patients harbouring TP53 mutations. We also determined an immunity-related eight-gene prognostic signature and performed analysis of its association with immune cell infiltration, which showed that the infiltration of M1

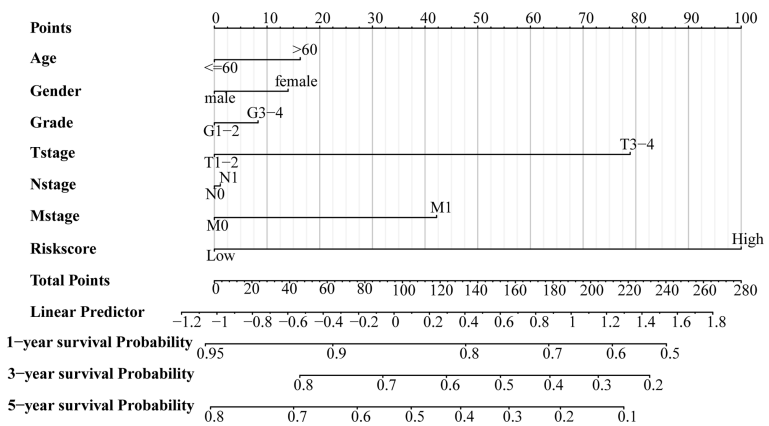



**FIGURE 5** | A prognostic nomogram for the overall survival of patients with LIHC.

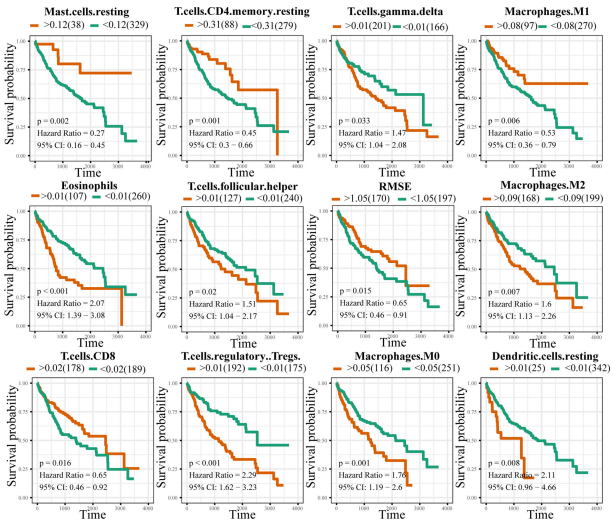



**FIGURE 6** | Correlation between immune cell infiltration and LIHC prognosis. Orange and green colours indicate the high- and low-risk groups, respectively.

macrophages, resting CD4+ memory T cells, activated NK cells, and CD8+ T cell suggested a better prognosis. Previous reports indicate that liver cancer tissues are characterized with a high expression level of PD-L1, CTLA4, lymphocyte activation gene 3, and other immunosuppressive molecules, which is negatively associated with the tumour infiltration of IFNγ+ T lymphocytes. Antibody treatment could increase the rate of CD8+ tumour-infiltrating T lymphocytes and the production of cytokines in liver cancer tissues (53). The infiltration of T cells before and after immunotherapy could be used to evaluate the effect of drugs enhancing the response to immune checkpoint blockers and to determine whether T cell infiltration by itself could predict the outcome of immunotherapy (54). It was reported that ependymin related protein 1 (EPDR1) and BRCA1 are correlated with immune cell infiltration and prognosis in HCC (55, 56). Recently, a nine immune-related gene model with an independent prognostic capability for HCC has been developed and shown to be associated with immune cell infiltration (57).

Using univariate Cox regression, we analysed the relationship among immune-associated genes and the prognosis of patients with LIHC, which were then divided into groups according to risk scores, and the optimal prognostic signature containing eight genes was established using iterative Lasso Cox regression analysis. ROC analysis performed in the external validation set revealed that the immune-associated eight-gene signature had a significant ability to predict 3-, 5-, and 7-year prognosis for patients with LIHC. Furthermore, compared with common prognostic biomarkers of LIHC, the eight-gene signature showed a superior predictive power and was proved to be an independent prognostic predictor of patient survival. We constructed a nomogram combining the clinicopathological characteristics and the immune-related eight-gene signature to offer clinicians a

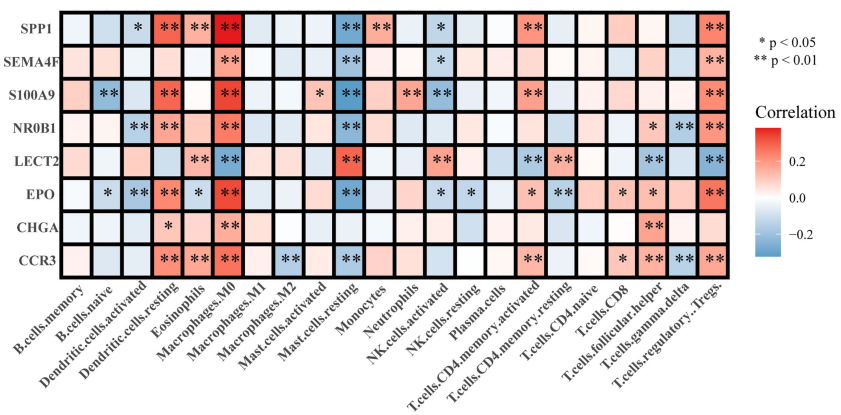



**FIGURE 7** | Correlation between the signature genes and immune cell infiltration.

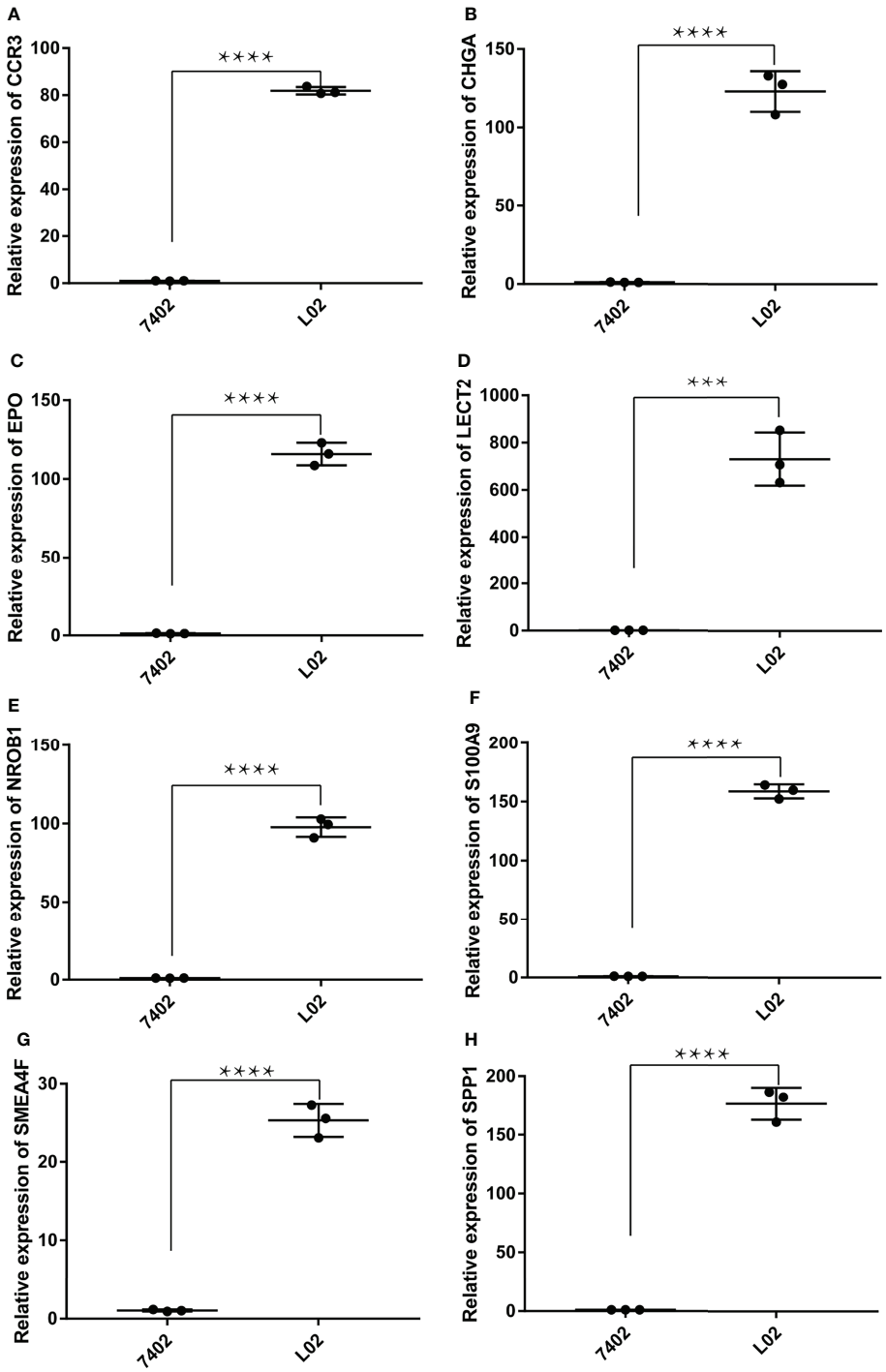



**FIGURE 8** | The relative RNA expression level of LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1. The RNA expression of CCR3 **(A)**, CHGA **(B)**, EPO **(C)**, LECT2 **(D)**, NR0B1 **(E)**, S100A9 **(F)**, SEMA4F **(G)** and SPP1 **(H)** were low expressed in human hepatoma cells compared with normal human hepatic epithelial cells (Unpaired t-test, ***$P < 0.001$, ****$P < 0.0001$).

quantitative method for predicting the LICH outcome, which should aid in the selection of optimal treatment approaches. Analysis of immune cell infiltration revealed that Tregs, activated NK cells, and M0 macrophages had the highest, whereas activated dendritic cells – the lowest infiltration rate in the high-risk group.

The immunity-related genes composing our eight gene prognostic signature have been previously shown to be involved

in oncogenesis. Thus, SEMA4F encoding semaphorin 4F plays a role in axonal growth cone guidance (58) and induction of neurogenesis in prostate cancer (59). The expression of the erythropoietin-encoding EPO gene is related to apoptosis, survival, and proliferation in the early stages of clear cell renal cell carcinoma (60) and has been identified as a distinct prognostic factor for overall and metastases-free survival and locoregional control in locally advanced HNSCC (61). The overexpression of S100A9 encoding calgranulin B has been suggested to play a vital role in the progress of oral squamous cell carcinoma (OSCC) and may serve as a diagnostic and prognostic biomarker for OSCC (62) and nonsmall-cell lung carcinoma (63).

In order to determine the association between LIHC prognosis and immune cell infiltration, we constructed a correlation heat map, which showed that seven genes in the immune-related eight-gene signature were negatively associated with the infiltration of M0 macrophages, resting mast cells, and Tregs. Previous studies have shown that S100A9 plays a significant role in the regulation of immune response and inflammation in most tumours and that it promotes cancer metastasis by accelerating tumour cell proliferation and invasion (64–66), which is consistent with the role of tumour-associated inflammation in supporting metastasis and cancer progression (66–68). Leukocyte cell-derived chemotaxin-2, a 16-kDa secreted protein encoded by the LECT2 gene (69), is involved in the regulation of the tumour microenvironment (70) and plays a critical role in hepatic oncogenesis. Thus, LECT2 deletion modifies the tumour microenvironment and alters cancer phenotypes, suggesting that it is a promising immunotherapeutic target in liver cancer (71). Another study has found that LECT2 expression in HCC is strongly correlated with tumour angiogenesis (72).

It was shown that the expression of C motif chemokine receptor 3 (CCR3) was correlated with malignancy of tumour cells (73). CCR3 ligand may be up-regulated by tumour-related inflammation and involved in the progress of renal cell carcinoma (74), whereas the CCR3/eotaxin-1 loop could induce malignant cell growth in T-cell lymphomas (75, 76). Secreted phosphoprotein 1 (SPP1), also known as osteopontin, is a multifunctional protein first characterized as a biomarker in epithelial cell transformation (77) and suggested to function as an enhancer of HCC growth targeted by miR-181c, thus representing a potential candidate biomarker for HCC diagnosis and therapy (78). We also found that LECT2, SEMA4F, EPO, CHGA, NR0B1, S100A9, CCR3, and SPP1 were low expressed in hepatocellular carcinoma, which may be new cancer therapeutic targets.

## 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 analyzed the data and wrote the manuscript. MZ designed the framework of the paper. YA, DL, and QT interpreted data. GT revised the manuscript. The final version of the manuscript has been approved by all authors.

## FUNDING

We are grateful for the support of National Natural Science Foundation of China (81827805), National Key R&D Program of China (2018YFA0704100, 2018YFA0704100), the Jiangsu Provincial Medical Youth Talent (QNRC2016816), the Project of Jiangsu Provincial Health and Family Planning Commission (H2018090).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.841530/full#supplementary-material

**Supplementary Figure 1 |** Comparison of the TP53 and NO_TP53 groups using GSEA and GSVA. **(A)** File 'c2.all.v7.1.symbols.gmt' was used as the reference gene set for GSEA. **(B)** File 'h.all.v7.1.symbols.gmt' was used as the reference gene set for GSVA.

**Supplementary Figure 2 |** Single-factor Cox regression model.

**Supplementary Figure 3 |** Confirmation of the optimal gene signature using an external validation set. **(A)** A centralized risk factor correlation diagram for the validation set. **(B)** Kaplan-Meier analysis of the correlation between the eight-gene signature and patient survival. **(C)** ROC analysis of 3-, 5-, and 7-year prognosis for patients with LIHC. **(D)** Time-conditioned ROC curve of the eight-gene signature and common prognostic biomarkers of LIHC.

**Supplementary Figure 4 |** Immune cell infiltration analysis. **(A)** Correlations in immune cell infiltration. **(B)** Immune cell interaction network; circle size (large to small) is proportional to the intensity of interaction (strong to weak). **(C)** Differences in immune cell infiltration between high- and low-risk groups. **(D)** Immune cell composition analysis.

## REFERENCES

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer Statistics, 2007. *CA Cancer J Clin* (2007) 57:43–66. doi: 10.3322/canjclin.57.1.43
2. Bruix J, Reig M, Sherman M. Evidence-Based Diagnosis, Staging, and Treatment of Patients With Hepatocellular Carcinoma. *Gastroenterology* (2016) 150:835–53. doi: 10.1053/j.gastro.2015.12.041
3. Wang Q, Zhang P, Li Z, Feng X, Lv C, Zhang H, et al. Evaluation of Polymer Nanoformulations in Hepatoma Therapy by Established Rodent Models. *Theranostics* (2019) 9:1426–52. doi: 10.7150/thno.31683
4. Bruix J, Qin S, Merle P, Granito A, Huang YH, Bodoky G, et al. Regorafenib for Patients With Hepatocellular Carcinoma Who Progressed on Sorafenib Treatment (RESORCE): A Randomised, Double-Blind, Placebo-Controlled, Phase 3 Trial. *Lancet* (2017) 389:56–66. doi: 10.1016/S0140-6736(16)32453-9
5. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Niksic M, et al. Global Surveillance of Trends in Cancer Survival 2000-14 (Concord-3): Analysis of Individual Records for 37 513 025 Patients Diagnosed With One of 18 Cancers From 322 Population-Based Registries in 71 Countries. *Lancet* (2018) 391:1023–75. doi: 10.1016/S0140-6736(17)33326-3

6. Forner A, Llovet JM, Bruix J. Hepatocellular Carcinoma. *Lancet* (2012) 379:1245–55. doi: 10.1016/S0140-6736(11)61347-0
7. Stratton MR. Exploring the Genomes of Cancer Cells: Progress and Promise. *Science* (2011) 331:1553–8. doi: 10.1126/science.1204040
8. Kandoth C, McLellan MD, Vandin F, Ye K, Niu B, Lu C, et al. Mutational Landscape and Significance Across 12 Major Cancer Types. *Nature* (2013) 502:333–9. doi: 10.1038/nature12634
9. Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 Mutations in Human Cancers. *Science* (1991) 253:49–53. doi: 10.1126/science.1905840
10. da Silva GN, Evangelista AF, Magalhaes DA, Macedo C, Bufalo MC, Sakamoto-Hojo ET, et al. Expression of Genes Related to Apoptosis, Cell Cycle and Signaling Pathways Are Independent of TP53 Status in Urinary Bladder Cancer Cells. *Mol Biol Rep* (2011) 38:4159–70. doi: 10.1007/s11033-010-0536-x
11. Cardin R, Piciocchi M, Tieppo C, Maddalo G, Zaninotto G, Mescoli C, et al. Oxidative DNA Damage in Barrett Mucosa: Correlation With Telomeric Dysfunction and p53 Mutation. *Ann Surg Oncol* (2013) 20 Suppl 3:S583–9. doi: 10.1245/s10434-013-3043-1
12. Di Agostino S, Strano S, Blandino G. Gender, Mutant p53 and PML: A Growing "Affaire" in Tumor Suppression and Oncogenesis. *Cell Cycle* (2013) 12:1824–5. doi: 10.4161/cc.25174
13. Greenblatt MS, Bennett WP, Hollstein M, Harris CC. Mutations in the p53 Tumor Suppressor Gene: Clues to Cancer Etiology and Molecular Pathogenesis. *Cancer Res* (1994) 54:4855–78.
14. Song Y, Li L, Ou Y, Gao Z, Li E, Li X, et al. Identification of Genomic Alterations in Oesophageal Squamous Cell Cancer. *Nature* (2014) 509:91–5. doi: 10.1038/nature13176
15. Lalloo F, Evans DG. Familial Breast Cancer. *Clin Genet* (2012) 82:105–14. doi: 10.1111/j.1399-0004.2012.01859.x
16. Olivier M, Langerod A, Carrieri P, Bergh J, Klaar S, Eyfjord J, et al. The Clinical Value of Somatic TP53 Gene Mutations in 1,794 Patients With Breast Cancer. *Clin Cancer Res* (2006) 12:1157–67. doi: 10.1158/1078-0432.CCR-05-1029
17. Ahmed AA, Etemadmoghadam D, Temple J, Lynch AG, Riad M, Sharma R, et al. Driver Mutations in TP53 Are Ubiquitous in High Grade Serous Carcinoma of the Ovary. *J Pathol* (2010) 221:49–56. doi: 10.1002/path.2696
18. Rajeshkumar NV, De Oliveira E, Ottenhof N, Watters J, Brooks D, Demuth T, et al. Mk-1775, a Potent Wee1 Inhibitor, Synergizes With Gemcitabine to Achieve Tumor Regressions, Selectively in p53-deficient Pancreatic Cancer Xenografts. *Clin Cancer Res* (2011) 17:2799–806. doi: 10.1158/1078-0432.CCR-10-2580
19. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated Analysis of Somatic Mutations and Focal Copy-Number Changes Identifies Key Genes and Pathways in Hepatocellular Carcinoma. *Nat Genet* (2012) 44:694–8. doi: 10.1038/ng.2256
20. Cancer Genome Atlas Research Network. Electronic address and N. Cancer Genome Atlas Research. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* (2017) 169:1327–41.e23. doi: 10.1016/j .cell.2017.05.046
21. Zhou X, Hao Q, Lu H. Mutant p53 in Cancer Therapy-the Barrier or the Path. *J Mol Cell Biol* (2019) 11:293–305. doi: 10.1093/jmcb/mjy072
22. Parrales A, Iwakuma T. Targeting Oncogenic Mutant p53 for Cancer Therapy. *Front Oncol* (2015) 5:288. doi: 10.3389/fonc.2015.00288
23. Takai A, Dang HT, Wang XW. Identification of Drivers From Cancer Genome Diversity in Hepatocellular Carcinoma. *Int J Mol Sci* (2014) 15:11142–60. doi: 10.3390/ijms150611142
24. Ahn SM, Jang SJ, Shim JH, Kim D, Hong SM, Sung CO, et al. Genomic Portrait of Resectable Hepatocellular Carcinomas: Implications of RB1 and FGF19 Aberrations for Patient Stratification. *Hepatology* (2014) 60:1972–82. doi: 10.1002/hep.27198
25. Liu J, Ma Q, Zhang M, Wang X, Zhang D, Li W, et al. Alterations of TP53 are Associated With a Poor Outcome for Patients With Hepatocellular Carcinoma: Evidence From a Systematic Review and Meta-Analysis. *Eur J Cancer* (2012) 48:2328–38. doi: 10.1016/j.ejca.2012.03.001
26. Dong ZY, Zhong WZ, Zhang XC, Su J, Xie Z, Liu SY, et al. Potential Predictive Value of TP53 and KRAS Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. *Clin Cancer Res* (2017) 23:3012–24. doi: 10.1158/1078-0432.CCR-16-2554
27. Biton J, Mansuet-Lupo A, Pecuchet N, Alifano M, Ouakrim H, Arrondeau J, et al. Tp53, STK11, and EGFR Mutations Predict Tumor Immune Profile and the Response to Anti-PD-1 in Lung Adenocarcinoma. *Clin Cancer Res* (2018) 24:5710–23. doi: 10.1158/1078-0432.CCR-18-0163
28. Long J, Wang A, Bai Y, Lin J, Yang X, Wang D, et al. Development and Validation of a TP53-Associated Immune Prognostic Model for Hepatocellular Carcinoma. *EBioMedicine* (2019) 42:363–74. doi: 10.1016/j.ebiom.2019.03.022
29. Tan LL, Jiang XL, Xu LX, Li G, Feng CX, Ding X, et al. TP53-Induced Glycolysis and Apoptosis Regulator Alleviates Hypoxia/Ischemia-Induced Microglial Pyroptosis and Ischemic Brain Damage. *Neural Regen Res* (2021) 16:1037–43. doi: 10.4103/1673-5374.300453
30. Matsuda S, Mafune A, Kohda N, Hama T, Urashima M. Associations Among Smoking, MGMT Hypermethylation, TP53-Mutations, and Relapse in Head and Neck Squamous Cell Carcinoma. *PloS One* (2020) 15:e0231932. doi: 10.1371/journal.pone.0231932
31. Sizemore ST, Mohammad R, Sizemore GM, Nowsheen S, Yu H, Ostrowski MC, et al. Synthetic Lethality of PARP Inhibition and Ionizing Radiation Is P53-Dependent. *Mol Cancer Res* (2018) 16:1092–102. doi: 10.1158/1541-7786.MCR-18-0106
32. Augustine T, Maitra R, Zhang J, Nayak J, Goel S. Sensitization of Colorectal Cancer to Irinotecan Therapy by PARP Inhibitor Rucaparib. *Invest New Drugs* (2019) 37:948–60. doi: 10.1007/s10637-018-00717-9
33. Na B, Yu X, Withers T, Gilleran J, Yao M, Foo TK, et al. Therapeutic Targeting of BRCA1 and TP53 Mutant Breast Cancer Through Mutant p53 Reactivation. *NPJ Breast Cancer* (2019) 5:14. doi: 10.1038/s41523-019-0110-1
34. Smeby J, Kryeziu K, Berg KCG, Eilertsen IA, Eide PW, Johannessen B, et al. Molecular Correlates of Sensitivity to PARP Inhibition Beyond Homologous Recombination Deficiency in Pre-Clinical Models of Colorectal Cancer Point to Wild-Type TP53 Activity. *EBioMedicine* (2020) 59:102923. doi: 10.1016/j.ebiom.2020.102923
35. Niyaz M, Ainiwaer J, Abudureheman A, Zhang L, Sheyhidin I, Turhong A, et al. Association Between TP53 Gene Deletion and Protein Expression in Esophageal Squamous Cell Carcinoma and Its Prognostic Significance. *Oncol Lett* (2020) 20:1855–65. doi: 10.3892/ol.2020.11709
36. Wang L, Yan K, He X, Zhu H, Song J, Chen S, et al. LRP1B or TP53 Mutations Are Associated With Higher Tumor Mutational Burden and Worse Survival in Hepatocellular Carcinoma. *J Cancer* (2021) 12:217–23. doi: 10.7150/jca.48983
37. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nat Genet* (2013) 45:1113–20. doi: 10.1038/ng.2764
38. Samur MK. RTCGAToolbox: A New Tool for Exporting TCGA Firehose Data. *PloS One* (2014) 9:e106397. doi: 10.1371/journal.pone.0106397
39. Grinchuk OV, Yenamandra SP, Iyer R, Singh M, Lee HK, Lim KH, et al. Tumor-Adjacent Tissue Co-Expression Profile Analysis Reveals Pro-Oncogenic Ribosomal Gene Signature for Prognosis of Resectable Hepatocellular Carcinoma. *Mol Oncol* (2018) 12:89–113. doi: 10.1002/1878-0261.12153
40. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for Functional Genomics Data Sets–Update. *Nucleic Acids Res* (2013) 41:D991–5. doi: 10.1093/nar/gks1193
41. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res* (2018) 28:1747–56. doi: 10.1101/gr.239244.118
42. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proc Natl Acad Sci USA* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
43. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. Gsea-P: A Desktop Application for Gene Set Enrichment Analysis. *Bioinformatics* (2007) 23:3251–3. doi: 10.1093/bioinformatics/btm369
44. Hanzelmann S, Castelo R, Guinney J. GSVA: Gene Set Variation Analysis for Microarray and RNA-seq Data. *BMC Bioinform* (2013) 14:7. doi: 10.1186/1471-2105-14-7
45. Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
46. Tibshirani R. The Lasso Method for Variable Selection in the Cox Model. *Stat Med* (1997) 16:385–95. doi: 10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3

47. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: Tool for the Unification of Biology. The Gene Ontology Consortium. *Nat Genet* (2000) 25:25–9. doi: 10.1038/75556
48. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* (2000) 28:27–30. doi: 10.1093/nar/28.1.27
49. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust Enumeration of Cell Subsets From Tissue Expression Profiles. *Nat Methods* (2015) 12:453–7. doi: 10.1038/nmeth.3337
50. Jenne CN, Kubes P. Immune Surveillance by the Liver. *Nat Immunol* (2013) 14:996–1006. doi: 10.1038/ni.2691
51. Wu ZJL, Li K, Zhang K, Gong WD. [Research Progress of Immunotherapy Alone and in Combination for Liver Cancer]. *Zhonghua Gan Zang Bing Za Zhi* (2020) 28:471–4. doi: 10.3760/cma.j.cn501113-20200520-00262
52. Kroemer G, Galluzzi L, Kepp O, Zitvogel L. Immunogenic Cell Death in Cancer Therapy. *Annu Rev Immunol* (2013) 31:51–72. doi: 10.1146/annurev-immunol-032712-100008
53. Chew V, Lai L, Pan L, Lim CJ, Li J, Ong R, et al. Delineation of an Immunosuppressive Gradient in Hepatocellular Carcinoma Using High-Dimensional Proteomic and Transcriptomic Analyses. *Proc Natl Acad Sci USA* (2017) 114:E5900–9. doi: 10.1073/pnas.1706559114
54. Lu LC, Lee YH, Chang CJ, Shun CT, Fang CY, Shao YY, et al. Increased Expression of Programmed Death-Ligand 1 in Infiltrating Immune Cells in Hepatocellular Carcinoma Tissues After Sorafenib Treatment. *Liver Cancer* (2019) 8:110–20. doi: 10.1159/000489021
55. Chen R, Zhang Y. EPDR1 Correlates With Immune Cell Infiltration in Hepatocellular Carcinoma and Can Be Used as a Prognostic Biomarker. *J Cell Mol Med* (2020) 24:12107–18. doi: 10.1111/jcmm.15852
56. Mei J, Wang R, Xia D, Yang X, Zhou W, Wang H, et al. BRCA1 Is a Novel Prognostic Indicator and Associates With Immune Cell Infiltration in Hepatocellular Carcinoma. *DNA Cell Biol* (2020) 39:1838–49. doi: 10.1089/dna.2020.5644
57. Wang D, Liu J, Liu S, Li W. Identification of Crucial Genes Associated With Immune Cell Infiltration in Hepatocellular Carcinoma by Weighted Gene Co-expression Network Analysis. *Front Genet* (2020) 11:342. doi: 10.3389/fgene.2020.00342
58. Francks C, Fisher SE, Olson RK, Pennington BF, Smith SD, DeFries JC, et al. Fine Mapping of the Chromosome 2p12-16 Dyslexia Susceptibility Locus: Quantitative Association Analysis and Positional Candidate Genes SEMA4F and OTX1. *Psychiatr Genet* (2002) 12:35–41. doi: 10.1097/00041444-200203000-00005
59. Ayala GE, Dai H, Powell M, Li R, Ding Y, Wheeler TM, et al. Cancer-Related Axonogenesis and Neurogenesis in Prostate Cancer. *Clin Cancer Res* (2008) 14:7593–603. doi: 10.1158/1078-0432.CCR-08-1164
60. Stoyanoff TR, Rodriguez JP, Todaro JS, Espada JD, Colavita JP, Brandan NC, et al. Tumor Biology of Non-Metastatic Stages of Clear Cell Renal Cell Carcinoma; Overexpression of Stearoyl Desaturase-1, EPO/EPO-R System and Hypoxia-Related Proteins. *Tumour Biol* (2016) 37:13581–93. doi: 10.1007/s13277-016-5279-4
61. Seibold ND, Schild SE, Gebhard MP, Noack F, Schroder U, Rades D. Prognosis of Patients With Locally Advanced Squamous Cell Carcinoma of the Head and Neck. Impact of Tumor Cell Expression of EPO and EPO-R. *Strahlenther Onkol* (2013) 189:559–65. doi: 10.1007/s00066-013-0320-7
62. Qu ZF, Ma H, Duan XF, Wu R, Zou Y. [The Expression and Significance of S100A9 in Oral Squamous Cell Carcinoma]. *Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* (2017) 31:219–22. doi: 10.13201/j.issn.1001-1781.2017.03.013
63. Koh HM, An HJ, Ko GH, Lee JH, Lee JS, Kim DC, et al. Prognostic Role of S100A8 and S100A9 Protein Expressions in Non-Small Cell Carcinoma of the Lung. *J Pathol Transl Med* (2019) 53:13–22. doi: 10.4132/jptm.2018.11.12
64. Lang B, Shang C, Meng L. Targeted Silencing of S100A8 Gene by miR-24 to Increase Chemotherapy Sensitivity of Endometrial Carcinoma Cells to Paclitaxel. *Med Sci Monit* (2016) 22:1953–8. doi: 10.12659/MSM.899179
65. Srikrishna G. S100A8 and S100A9: New Insights Into Their Roles in Malignancy. *J Innate Immun* (2012) 4:31–40. doi: 10.1159/000330095
66. Lim SY, Yuzhalin AE, Gordon-Weeks AN, Muschel RJ. Tumor-Infiltrating Monocytes/Macrophages Promote Tumor Invasion and Migration by Upregulating S100A8 and S100A9 Expression in Cancer Cells. *Oncogene* (2016) 35:5735–45. doi: 10.1038/onc.2016.107
67. Zhang X, Ai F, Li X, She X, Li N, Tang A, et al. Inflammation-Induced S100A8 Activates Id3 and Promotes Colorectal Tumorigenesis. *Int J Cancer* (2015) 137:2803–14. doi: 10.1002/ijc.29671
68. Funk S, Mark R, Bayo P, Flechtenmacher C, Grabe N, Angel P, et al. High S100A8 and S100A12 Protein Expression Is a Favorable Prognostic Factor for Survival of Oropharyngeal Squamous Cell Carcinoma. *Int J Cancer* (2015) 136:2037–46. doi: 10.1002/ijc.29262
69. Ito M, Nagata K, Kato Y, Oda Y, Yamagoe S, Suzuki K, et al. Expression, Oxidative Refolding, and Characterization of Six-Histidine-Tagged Recombinant Human LECT2, a 16-kDa Chemotactic Protein With Three Disulfide Bonds. *Protein Expr Purif* (2003) 27:272–8. doi: 10.1016/S1046-5928(02)00634-4
70. Hwang HJ, Jung TW, Hong HC, Seo JA, Kim SG, Kim NH, et al. LECT2 Induces Atherosclerotic Inflammatory Reaction *Via* CD209 Receptor-Mediated JNK Phosphorylation in Human Endothelial Cells. *Metabolism* (2015) 64:1175–82. doi: 10.1016/j.metabol.2015.06.001
71. L'Hermitte A, Pham S, Cadoux M, Couchy G, Caruso S, Anson M, et al. Lect2 Controls Inflammatory Monocytes to Constrain the Growth and Progression of Hepatocellular Carcinoma. *Hepatology* (2019) 69:160–78. doi: 10.1002/hep.30140
72. Chen CK, Yu WH, Cheng TY, Chen MW, Su CY, Yang YC, et al. Inhibition of VEGF165/VEGFR2-Dependent Signaling by LECT2 Suppresses Hepatocellular Carcinoma Angiogenesis. *Sci Rep* (2016) 6:31398. doi: 10.1038/srep31398
73. Jin L, Liu WR, Tian MX, Jiang XF, Wang H, Zhou PY, et al. CCL24 Contributes to HCC Malignancy *Via* RhoB- Vegfa-VEGFR2 Angiogenesis Pathway and Indicates Poor Prognosis. *Oncotarget* (2017) 8:5135–48. doi: 10.18632/oncotarget.14095
74. Johrer K, Zelle-Rieser C, Perathoner A, Moser P, Hager M, Ramoner R, et al. Up-Regulation of Functional Chemokine Receptor CCR3 in Human Renal Cell Carcinoma. *Clin Cancer Res* (2005) 11:2459–65. doi: 10.1158/1078-0432.CCR-04-0405
75. Kleinhans M, Tun-Kyi A, Gilliet M, Kadin ME, Dummer R, Burg G, et al. Functional Expression of the Eotaxin Receptor CCR3 in CD30+ Cutaneous T-Cell Lymphoma. *Blood* (2003) 101:1487–93. doi: 10.1182/blood-2002-02-0475
76. Wang C, Wang Y, Hong T, Cheng B, Gan S, Chen L, et al. Blocking the Autocrine Regulatory Loop of Gankyrin/STAT3/CCL24/CCR3 Impairs the Progression and Pazopanib Resistance of Clear Cell Renal Cell Carcinoma. *Cell Death Dis* (2020) 11:117. doi: 10.1038/s41419-020-2306-6
77. Han X, Wang W, He J, Jiang L, Li X. Osteopontin as a Biomarker for Osteosarcoma Therapy and Prognosis. *Oncol Lett* (2019) 17:2592–8. doi: 10.3892/ol.2019.9905
78. Wang J, Hao F, Fei X, Chen Y. SPP1 Functions as an Enhancer of Cell Growth in Hepatocellular Carcinoma Targeted by miR-181c. *Am J Transl Res* (2019) 11:6924–37.


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 04 May 2022
doi: 10.3389/fonc.2022.881015

# Identification by Bioinformatics Analysis of Potential Key Genes Related to the Progression and Prognosis of Gastric Cancer

*Wencang Gao*[*†] *and Min Yang*[†]

*Department of Oncology, The Second Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China*

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Juanjuan Zhu,*
*Capital Medical University, China*
*Dhananjay Shukla,*
*Guru Ghasidas Vishwavidyalaya, India*
*Feng Jiang,*
*Fudan University, China*

***Correspondence:**
*Wencang Gao*
*20164580@zcmu.edu.cn*

[†]*These authors share first authorship*

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

**Received:** *22 February 2022*
**Accepted:** *25 March 2022*
**Published:** *04 May 2022*

**Citation:**
*Gao W and Yang M (2022) Identification by Bioinformatics Analysis of Potential Key Genes Related to the Progression and Prognosis of Gastric Cancer. Front. Oncol. 12:881015. doi: 10.3389/fonc.2022.881015*

**Objective:** Despite increasingly sophisticated medical technology, the prognosis of patients with advanced gastric cancer is still not objectively certain. Therefore, it is urgent to identify new diagnostic and prognostic biomarkers. To identify potential critical genes related to gastric cancer's staging mechanism and to the prognosis of gastric cancer.

**Methods:** Dynamic trend analysis was conducted to find genes with similar trends in gastric cancer staging in order to explore the differentially expressed genes in gastric cancer and identify the intersection of the results of the dynamic trend analysis. Functional predictive analysis were performed on the obtained genes to observe the expression of prognostic genes in gastric cancer and in gastric cancer stages as well as the correlation with tumor immune cell infiltration. Gastric cancer samples were collected and sequenced for follow-up analysis based on the results of the Cancer Genome Atlas (TCGA) database analysis.

**Results:** The expression of genes enriched in module 0 had a similar trend in gastric cancer staging. 3213 differential genes were screened. A total of 50 intersection genes were obtained among genes with similar trends, of which only 10 genes have prognostic significance in gastric cancer. These 10 genes were correlated with macrophage infiltration in varying degrees. In addition, we found that AGT was significantly abnormally expressed in the results of sample sequencing. AGT was related to the occurrence of gastric cancer and interacted with brd9, golph3, nom1, klhl25, and psmd11.

**Conclusion:** AGT has prominent abnormal expression in gastric cancer and may promote gastric cancer progression. This study provides a new direction for further exploring potential biomarkers and molecular targeted gastric cancer therapy.

**Keywords: gastric cancer, TNM stage, bioinformatics analysis, RNA-seq, tumor immune cell infiltration**

## INTRODUCTION

Gastric cancer is a common upper gastrointestinal cancer (1). It often occurs among the middle-aged and elderly and is not common in young people (under the age of 45), who represent no more than 10% of patients suffering from the disease (2). With improvements in treatment technology, the five-year survival rate of early gastric cancer can reach more than 95%, but the symptoms of gastric cancer are rare and nonspecific in the early stages of the disease (3). This multifactorial disease is associated with environmental and genetic factors and is usually diagnosed in its late stages, with a median survival period of less than 12 months (4, 5). Data show that the prognosis of patients is related to the tumor stage (6). Therefore, more and more attention has been paid to clinical staging in recent years. Obtaining accurate, reliable clinical staging is a challenging but crucial problem.

The Union for international cancer control (UICC) and the American Joint Committee on cancer (AJCC) maintain the TNM cancer control staging system, a tool used by physicians to stage different types of cancer (T), existence in the lymph nodes (N), and metastasis (M). At present, the TNM staging system is used to evaluate most malignant tumors worldwide, including gastric cancer (7). The accurate staging of gastric cancer is a crucial part of patient management, as tumor staging is the most important predictor of survival at the initial diagnosis. In addition, the treatment scheme for the tumor will vary according to different disease stages (8). Tumor prognosis prediction and treatment recommendations for gastric cancer are mainly guided by the TNM staging system (9). Preventable factors that play a role in gastric cancer development are also important. This study aimed to find biomarkers that may be involved in the progression and staging of gastric cancer to stratify patients and better target treatment.

## METHODS AND ANALYSIS

### Data Collection

We extracted data from the Cancer Genome Atlas (TCGA) database, obtaining the mRNA information of gastric cancer samples, including 375 primary cancers and 32 non-tumor tissues. The original count data and relevant clinical information of patients were also downloaded and extracted. Gene expression information was matched to the clinical stage information, and unknown or incomplete clinical information was removed.

### Sample Collection

The gastric cancer tissues and adjacent tissues of five patients with gastric cancer in our hospital (two males and three females) were collected and preserved in liquid nitrogen. This study was approved by the Ethics Committee of the First Affiliated Hospital of Zhejiang University of Traditional Chinese Medicine. The duties, composition, operating procedures, and records of this ethics review committee follow the Measures for Ethical Review of Biomedical Research Involving Human Beings, the International Ethical Guidelines for Health-Related Research Involving Human Beings, the Declaration of Helsinki, international ethical guidelines such as GCP and ICH-GCP, and relevant domestic laws and regulations (2022-LW-015-01).

### Dynamic Trend Analysis

Trend analysis can classify genes with similar change characteristic patterns within a changing trend. We used the OmicShare online tool (https://www.omicshare.com/tools/) for analysis. $P$ <.05 was considered statistically significant. The number of trends is chosen to be 20.

### Analysis of Differentially Expressed Genes

We used the limma package to identify differentially expressed genes between tumor tissues and normal tissues, followed by a test to analyze differentially expressed genes. Adj.$P$ <.05 and $|\log2\ (FC)| > 1.0$ were used as cutoff values.

### Analysis of Tumor Immune Cell Infiltration

The correlation between gene expression and immune cells (including B cells, CD4 + T cells, CD8 + T cells, macrophages, neutrophils, and dendritic cells) was based on the TIMER database. $P$ <.05 was considered statistically significant. For the spearman test, there is generally a significant correlation for $|r|>0.95$; a high correlation for $|r|\geq0.8$; a moderate correlation for $0.5\leq|r|<0.8$; and a low correlation for $0.3\leq|r|<0.5$.

### Survival Analysis

Univariate Cox analysis was used to analyze the genes with prognostic characteristics in gastric cancer and to display the $p$-value, risk coefficient HR, and confidence interval. The KM curve was drawn with the "gene and survival" tool, which the R software package implements. $P$ <.05 was considered statistically significant.

### Weighted Gene Coexpression Network Analysis

WGCNA has been widely used to identify disease-related gene modules and to extract potential therapeutic targets. First, the integral function in the WGCNA software package in R was used to select the appropriate soft threshold power β. We constructed a topological overlap matrix and hierarchical clustering tree between the module detected genes. Analysis and biological function analysis then further screened the key co expression modules. For visualization, the genes coexpressed with AGT were analyzed by Cytoscape software (https://cytoscape.org/).

## RESULTS

### Dynamic Trend Analysis

First, we performed trend analysis based on TCGA to observe the dynamic expression of genes in different stages. Our analysis results show that the genes enriched in modules 0, 12, 16, 17, and 18 were statistically significant, and the genes enriched in module

0 show a downward linear trend. In module 0, 2,234 genes were enriched (see **Figure 1** for details).

## Screening of Differentially Expressed Genes

We then analyzed the differential expression between gastric cancer and normal tissues based on TCGA. Taking $P$ <.05 and | log2 (FC) | > 1 as the screening thresholds, a total of 3,213 differential genes were displayed, of which 2,702 were up-regulated and 511 were down-regulated (see **Figure 2** for details).

## Intersecting Genes

Next, we intersected the genes of module 0 and the differential genes in the trend analysis results, yielding a total of 50 intersecting genes. (see **Figure 3** for details).

## Survival Analysis

In addition, we analyzed the prognosis of the 50 intersecting genes and observed the effect of their expression on the prognosis of patients with gastric cancer. Ten of the 50 genes (RXRG, AGT, BCHE, UBE2QL1, PLCXD3, ADCYAP1R1, NRCAM, MAMDC2, CDH19, and GAMT) had prognostic significance in gastric cancer and were represented by KM curves (see **Figure 4**).

## Gene Expression in Gastric Cancer Tissues and Diverse Stages

We then observed the expression of those 10 genes in gastric cancer and normal tissues. The results show that, compared with normal tissues, the expression of AGT and NrCAM in gastric cancer was significantly up-regulated while the expression of the other eight genes was down-regulated in gastric cancer (see **Figure 5** for details). In addition, we observed the expression of the 10 genes in various stages of gastric cancer, as shown in **Figure 6**.

## Relationship Between Gene Expression and Immune Cell Infiltration

We further explored the relationship between those 10 genes and immune cells. The results of the immune cell correlation analysis show that the expression level of the genes was most closely related to macrophages. The expression levels of BCHE、MAMDC2 and PLCXD3 were moderately positively correlated with the degree of macrophage infiltration while the expression levels of AGT、GAMT、CDH19、ADCYAP1R1、RXRG and UBE2QL1 were weakly positively correlated with the degree of macrophage infiltration (see **Figure 7** for details).

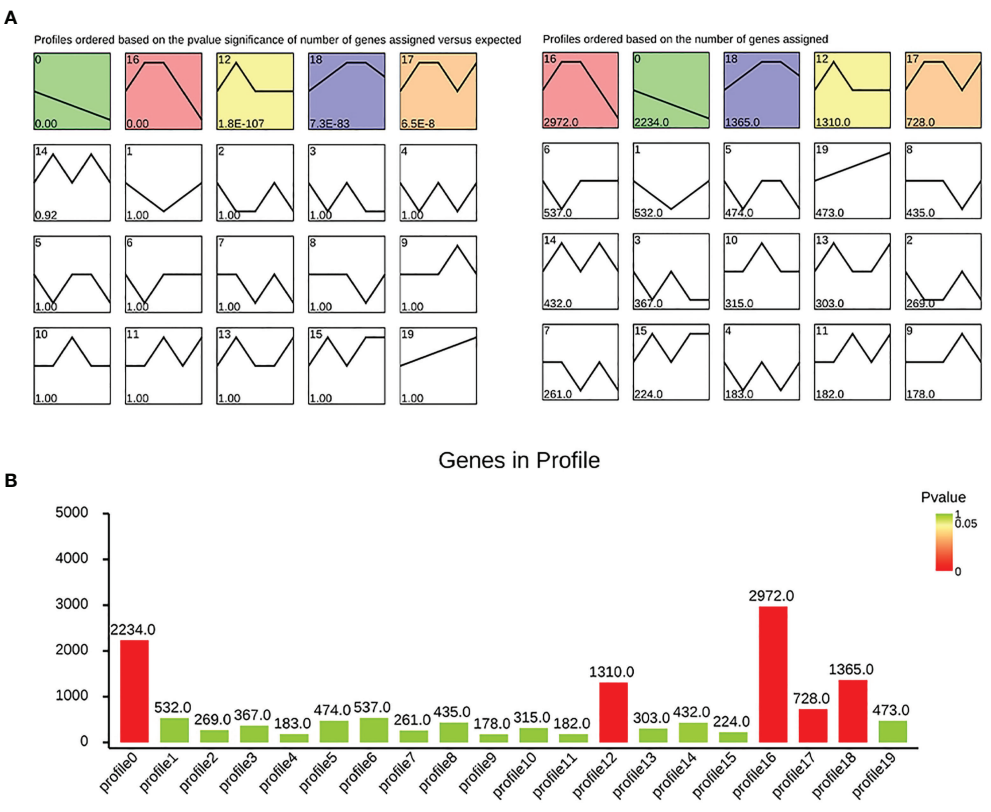



**FIGURE 1** | Dynamic trend analysis: **(A)** the significance and gene number of each module; **(B)** histogram showing the situation of each module.

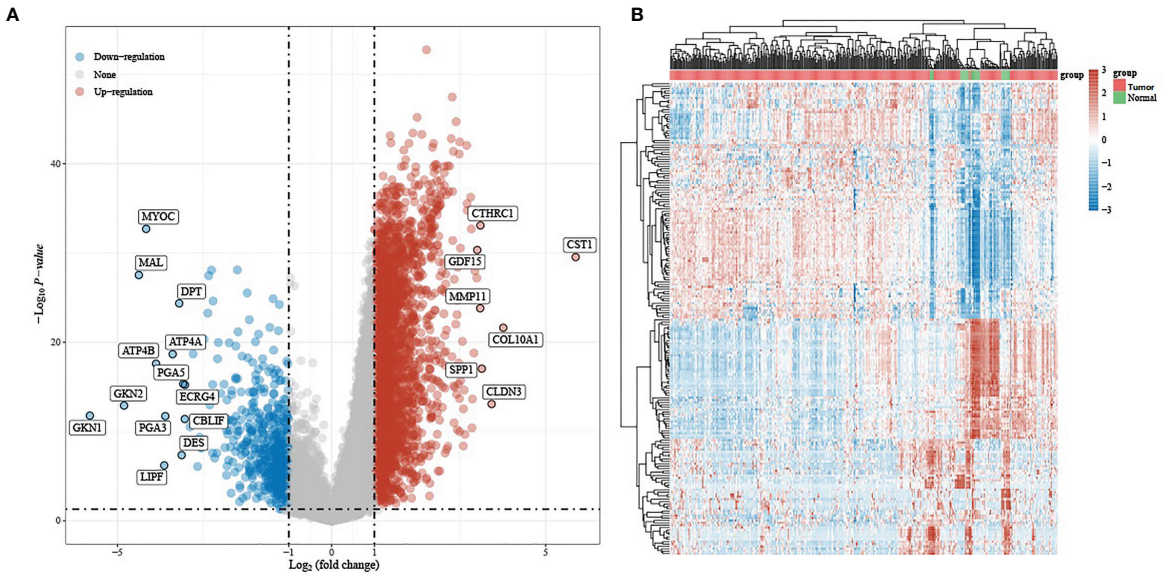



**FIGURE 2** | Differential expression analysis: **(A)** volcano plot of differentially expressed genes, with red indicating up-regulated genes and blue indicating down-regulated genes; **(B)** heat map of differential gene expression.

## Correlation Analysis of Sample Sequencing

After expression analysis and the observation of the expression in the stages of gastric cancer, we found that AGT and NrCAM were up-regulated in gastric cancer and that their appearance could increase with the increase of the stage. Moreover, the predictive results of these two genes in gastric cancer show that the higher their expression, the worse the prognosis. Therefore, we collected gastric cancer tissue samples for transcriptome analysis to further observe the expression of AGT and NRCAM in gastric cancer. The results reveal that only AGT had significant expression differences in gastric cancer tissues, and the up-regulated expression was consistent with the above analysis results based on TCGA. Next, we performed WGCNA analysis based on the expression matrix obtained by sequencing, identified the module where AGT is located (antiquewhite1), and

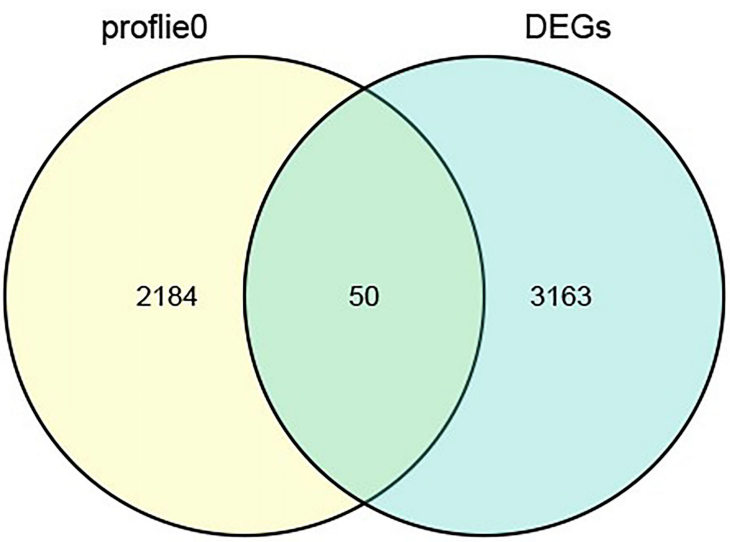



**FIGURE 3** | Venn diagram of the intersection.

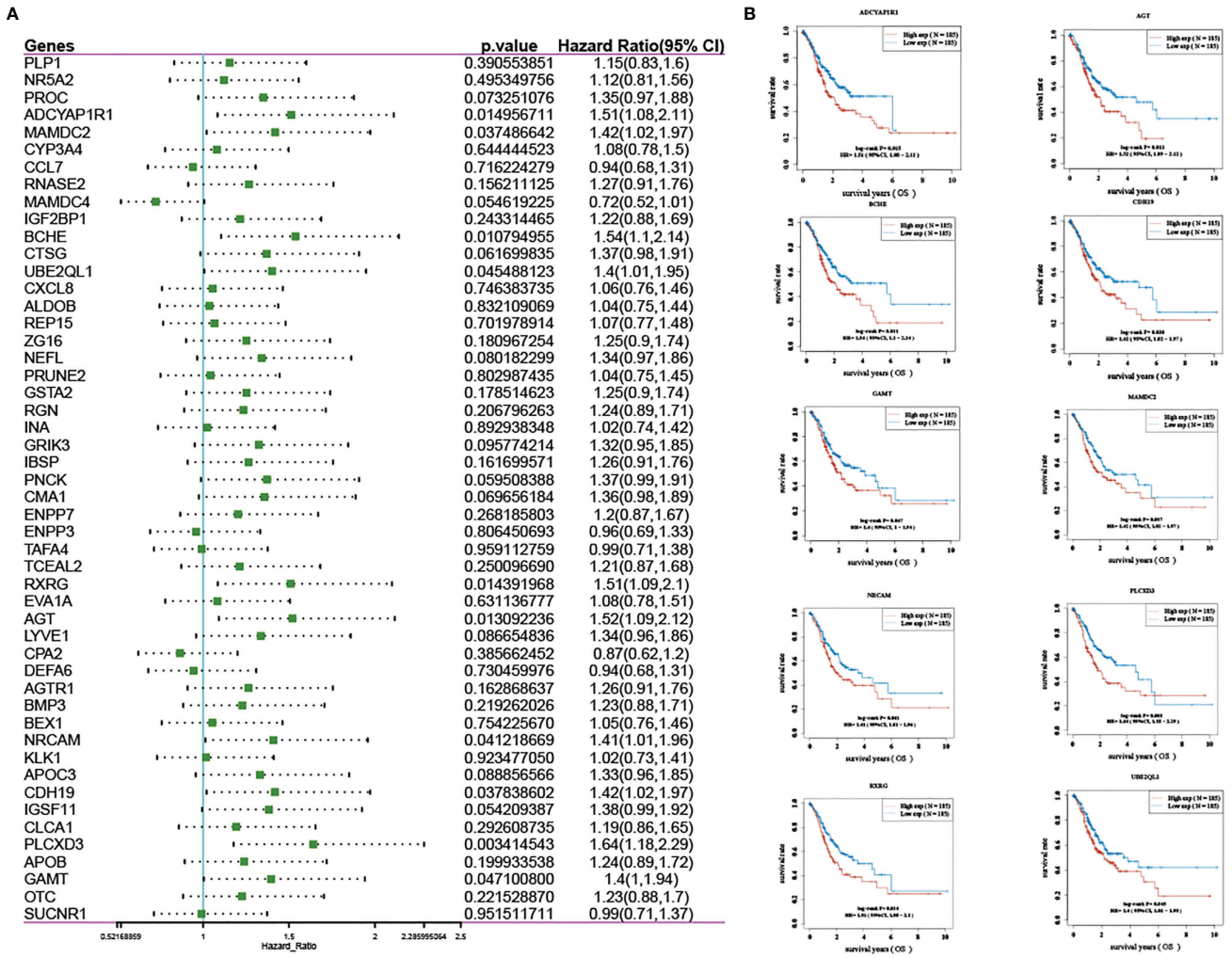



**FIGURE 4** | Prognosis analysis: **(A)** forest map of screening prognostic genes; **(B)** KM curve of genes with prognostic significance.

explored the genes that have a coexpression relationship with AGT. The results show that BRD9,GOLPH3, NOM1, KLHL25, and PSMD11 interacted with AGT (see **Figure 8** for details).

## DISCUSSION

Like other solid tumors, gastric cancer comprises several molecular subtypes with diverse biological characteristics. The increasing understanding of molecular pathways provides a basis for innovative therapies (10). Molecular targeted therapy can be used to identify specific oncogenic targets of tumor cells, killing tumor cells (11). In addition, the therapy has been shown to regulate the immune response, as some molecularly targeted drugs can increase tumor antigen expression and promote the antigen presentation of antigen-presenting cells to induce a more robust antitumor immune response (12). The expression of mRNA, a transient intermediate between genes and proteins, has shown therapeutic potential in various applications, including viral vaccines, protein replacement therapy, cancer immunotherapy, cell reprogramming, and genome editing (13). Therefore, mRNA has become a new therapeutic agent for preventing and treating various diseases (14). Using dynamic trend analysis, differential analysis, and predictive analysis, this study screened 10 genes with prognostic significance in gastric cancer (RXRG, AGT, BCHE, UBE2QL1, PLCXD3, ADCYAP1R1, NRCAM, MAMDC2, CDH19, and GAMT). Expression analysis and stage expression analysis also showed that they had high or low expression in gastric cancer and changed dynamically with the evolution of the gastric cancer stage.

The immune microenvironment plays an essential role in the development of digestive system cancer. Tumor infiltrating immune cells play a crucial role in suppressing or supporting tumor growth and development and can be effectively targeted by drugs, associated with patient survival time (15). Other studies have shown that macrophages induce PD-L1 expression and help gastric cancer cells escape cytotoxic T cell death. In addition, they can promote the

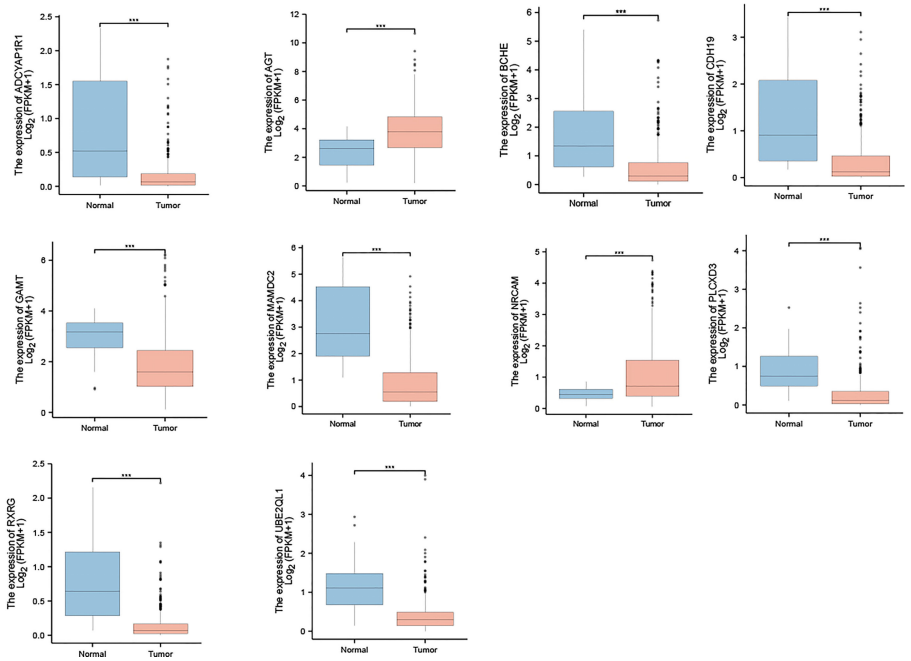



**FIGURE 5** | Expression of 10 genes in gastric cancer and normal tissues. ***p < 0.001.

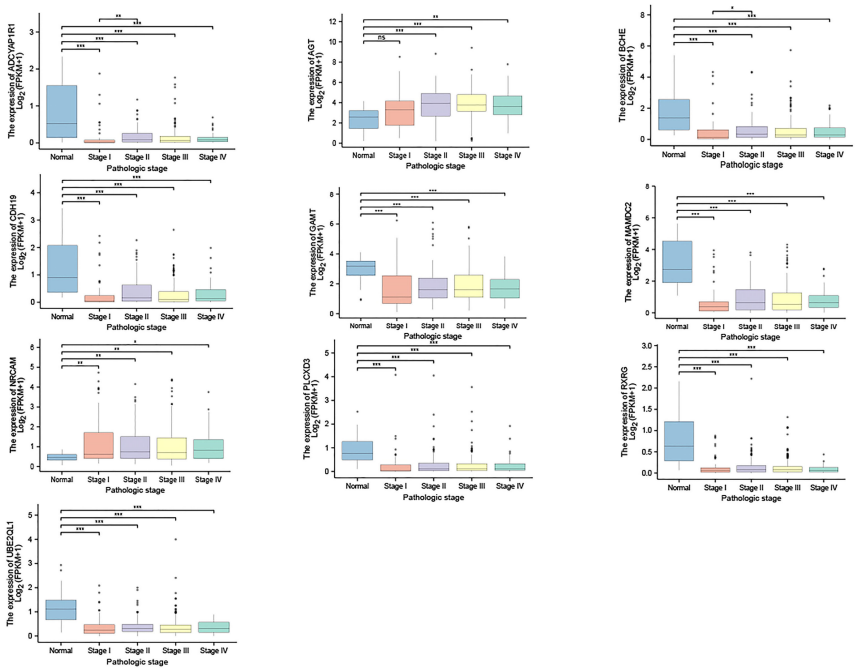



**FIGURE 6** | Expression of 10 genes in diverse stages of gastric cancer. *p < 0.05; **p < 0.01; ***p < 0.001; ns, non significant.

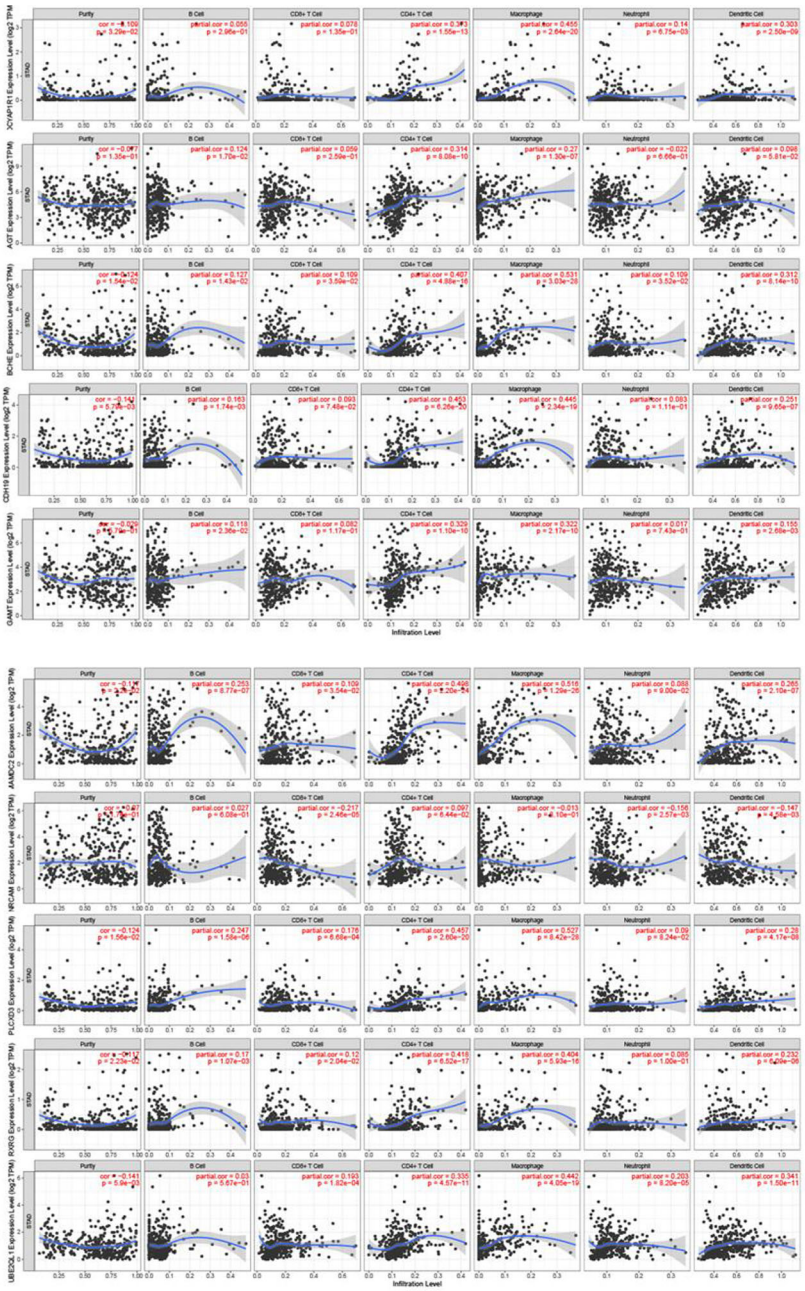



**FIGURE 7** | 10 Relationship between gene expression and immune cells.

proliferation of gastric cancer cells by regulating the expression of PD-L1 (16).

Moreover, macrophage infiltration is significantly correlated with the prognosis of patients with gastric cancer (17). ADCYAP1R1 has anti-inflammatory, neuroprotective and regenerative properties, and its expression may also be influenced by immune/inflammatory stimuli (18). It has been shown that deletion of ADCYAP1R1 in naïve mice results in the absence of retinal ganglion neurons and their dendrites, while increased axonal pathology and secondary increases in microglia/macrophages are also evident in the optic nerve (19). CDH19 is thought to be a member of calmodulin and establishes and maintains intercellular junctions (20). Some studies have shown that CDH19 may be a potential candidate as an immunotherapeutic target for breast cancer patients (21). MAMDC2 is a proteoglycan of completely unknown function, containing five MAM structural domains. They mediate interactions and stability and can act as an adhesion structural domain for surface receptors (22). Some studies have

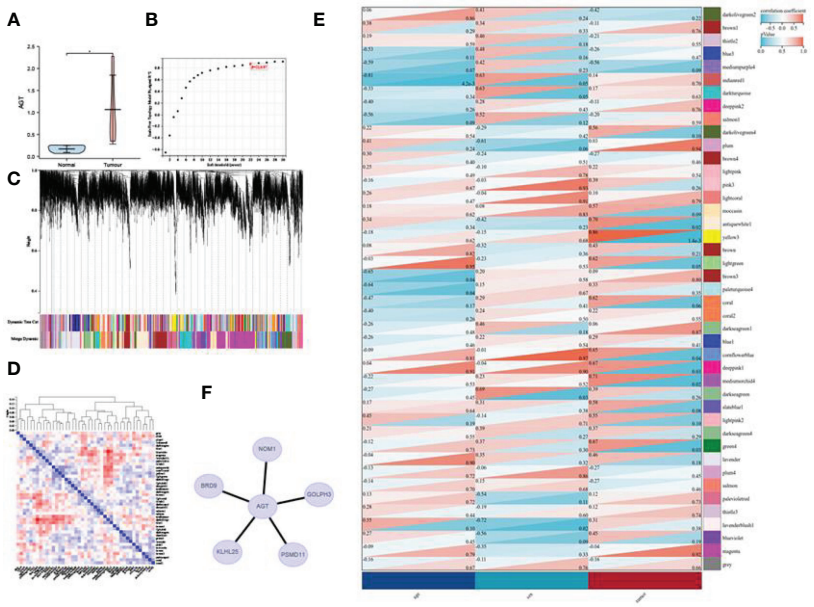



**FIGURE 8** | WGCNA analysis: **(A)** Expression of AGT in gastric cancer based on the sequencing expression matrix; **(B)** soft threshold; **(C)** gene clustering; **(D)** module eigenvector clustering; **(E)** module and phenotype correlation heat map; **(F)** AGT coexpression network. *p < 0.05.

also shown that MAMDC2 plays a key role in the development of invasive ductal carcinoma of the breast (23). Although no clear studies have shown their direct role with immune cells in tumors, they also show that they play an important role in tumors. We found that the expression levels of BCHE, MAMDC2 and PLCXD3 were moderately positively correlated with the degree of macrophage infiltration while the expression levels of AGT, GAMT, CDH19, ADCYAP1R1, RXRG and UBE2QL1 were weakly positively correlated with the degree of macrophage infiltration. These results show that these genes can potentially regulate the recruitment and activation of immune cells in gastric cancer.

This study found that AGT is highly expressed in gastric cancer and up-regulated with the increase of the stage. This continuous high expression may also be related to the poor prognosis of patients. The results based on sample sequencing also show that the expression of AGT was up-regulated in gastric cancer as compared with the standard group. AGT is an inactive precursor of potent vasoactivity and the salt-retention hormone angiotensin II, and an elevated level of it is the primary precursor in the pathogenesis of hypertension (24, 25). AGT also plays a critical role in tumors. Oncogenic effects of high glucose on breast cancer cell growth and metastasis (26). Studies have shown that AGT can also inhibit the invasion and migration of colorectal cancer (27). In this study, we performed a WGCNA of the sequenced expression matrix, which revealed that AGT was located in the antiquewhite1 module, closely related to gastric cancer. Based on the genes of this module, we explored the genes that have a coexpression relationship with AGT. The analysis shows that BRD9, GOLPH3, NOM1, KLHL25, and PSMD11 interact with AGT.

In conclusion, the expression of AGT is up-regulated in gastric cancer and the gastric cancer stages, and its continuous up-regulation is closely related to the poor prognosis of patients with gastric cancer. It is suggested that AGT may contribute to gastric cancer progression and could be used as a potential biomarker or therapeutic target in gastric cancer. However, there are deficiencies in this study, and *in vitro* experiments are needed to determine the critical signal pathways related to AGT.

## 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

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

## FUNDING

This study was supported by the key discipline “minimally invasive oncology of acupuncture and Moxibustion” (2017-XK-A12) of Zhejiang Administration of traditional Chinese medicine and the inheritance studio of Pangde Xiang famous old traditional Chinese medicine experts (2A12012014).

## REFERENCES


1. Yusefi AR, Lankarani KB, Bastani P, Radinmanesh M, Kavosi Z. Risk Factors for Gastric Cancer: A Systematic Review. *Asian Pac J Cancer Prev* (2018) 19 (3):591. doi: 10.22034/APJCP.2018.19.3.591
2. Machlowska J, Baj J, Sitarz M, Maciejewski R, Sitarz R. Gastric Cancer: Epidemiology, Risk Factors, Classification, Genomic Characteristics and Treatment Strategies. *Int J Mol Sci* (2020) 21(11):4012. doi: 10.3390/ijms21114012
3. Wu D, Zhang P, Ma J, Xu J, Yang L, Xu W, et al. Serum Biomarker Panels for the Diagnosis of Gastric Cancer. *Cancer Med* (2019) 8(4):1576–83. doi: 10.1002/cam4.2055
4. Song Z, Wu Y, Yang J, Yang D, Fang XJTB. Progress in the Treatment of Advanced Gastric Cancer. *Tumour Biol* (2017) 39(7):1010428317714626. doi: 10.1177/1010428317714626
5. Zhang X-y, Zhang P-y. Gastric Cancer: Somatic Genetics as a Guide to Therapy. *J Med Genet* (2017) 54(5):305–12. doi: 10.1136/jmedgenet-2016-104171
6. Ratti M, Lampis A, Hahne JC, Passalacqua R, Valeri NJC, Sciences ML. Microsatellite Instability in Gastric Cancer: Molecular Bases, Clinical Perspectives, and New Treatment Approaches. *Cell Mol Life Sci* (2018) 75 (22):4151–62. doi: 10.1007/s00018-018-2906-9
7. Li Z, Wang Y, Shan F, Ying X, Wu Z, Xue K, et al. ypTNM Staging After Neoadjuvant Chemotherapy in the Chinese Gastric Cancer Population: An Evaluation on the Prognostic Value of the AJCC Eighth Edition Cancer Staging System. *Gastric Cancer* (2018) 21(6):977–87. doi: 10.1007/s10120-018-0830-1
8. Lim W, Ridge CA, Nicholson AG, Mirsadraee S. The 8th Lung Cancer TNM Classification and Clinical Staging System: Review of the Changes and Clinical Implications. *Quant Imaging Med Surg* (2018) 8(7):709. doi: 10.21037/qims.2018.08.02
9. Zubarayev M, Min E-K, Son T. Clinical and Molecular Prognostic Markers of Survival After Surgery for Gastric Cancer: Tumor-Node-Metastasis Staging System and Beyond. *Transl Gastroenterol Hepatol* (2019) 4:59. doi: 10.21037/tgh.2019.08.05
10. Baretton G, Aust DJDP. Aktuelle Biomarker Beim Magenkarzinom. *Pathologe* (2017) 38(2):93–7. doi: 10.1007/s00292-017-0271-3
11. Lee YT, Tan YJ, Oon CE. Molecular Targeted Therapy: Treating Cancer With Specificity. *Eur J Pharmacol* (2018) 834:188–96. doi: 10.1016/j.ejphar.2018.07.034
12. Li Z, Liu Y, Fang X, Shu Z. Nanomaterials Enhance the Immunomodulatory Effect of Molecular Targeted Therapy. *Int J Nanomed* (2021) 16:1631. doi: 10.2147/IJN.S290346
13. Kowalski PS, Rudra A, Miao L, Anderson DG. Delivering the Messenger: Advances in Technologies for Therapeutic mRNA Delivery. *Mol Ther* (2019) 27(4):710–28. doi: 10.1016/j.ymthe.2019.02.012
14. Hou X, Zaks T, Langer R, Dong Y. Lipid Nanoparticles for mRNA Delivery. *Nat Rev Mater* (2021) 6(12):1078–94. doi: 10.1038/s41578-021-00358-0
15. Yang S, Liu T, Cheng Y, Bai Y, Liang G. Immune Cell Infiltration as a Biomarker for the Diagnosis and Prognosis of Digestive System Cancer. *Cancer Sci* (2019) 110(12):3639–49. doi: 10.1111/cas.14216
16. Ju X, Zhang H, Zhou Z, Chen M, Wang Q. Tumor-Associated Macrophages Induce PD-L1 Expression in Gastric Cancer Cells Through IL-6 and TNF-α Signaling. *Exp Cell Res* (2020) 396(2):112315. doi: 10.1016/j.yexcr.2020.112315
17. Xiao Z, Hu L, Yang L, Wang S, Gao Y, Zhu Q, et al. Tgfβ2 is a Prognostic-Related Biomarker and Correlated With Immune Infiltrates in Gastric Cancer. *J Cell Mol Med* (2020) 24(13):7151–62. doi: 10.1111/jcmm.15164
18. Tsumuraya T, Ohtaki H, Song D, Sato A, Watanabe J, Hiraizumi Y, et al. Human Mesenchymal Stem/Stromal Cells Suppress Spinal Inflammation in Mice With Contribution of Pituitary Adenylate Cyclase-Activating Polypeptide (PACAP). *J Neuroinflamm* (2015) 12(1):1–13. doi: 10.1186/s12974-015-0252-5
19. Van C, Condro MC, Ko HH, Hoang AQ, Zhu R, Lov K, et al. Targeted Deletion of PAC1 Receptors in Retinal Neurons Enhances Neuron Loss and Axonopathy in a Model of Multiple Sclerosis and Optic Neuritis. *Neurobiol Dis* (2021) 160:105524. doi: 10.1016/j.nbd.2021.105524
20. Shi Y, Han T, Liu C. CircRNA Hsa_Circ_0006220 Acts as a Tumor Suppressor Gene by Regulating miR-197-5p/CDH19 in Triple-Negative Breast Cancer. *Ann Transl Med* (2021) 9(15):1236. doi: 10.21037/atm-21-2934
21. Zhang Z, Li J, He T, Ding J. Bioinformatics Identified 17 Immune Genes as Prognostic Biomarkers for Breast Cancer: Application Study Based on Artificial Intelligence Algorithms. *Front Oncol* (2020) 10:330. doi: 10.3389/fonc.2020.00330
22. Wierer M, Prestel M, Schiller HB, Yan G, Schaab C, Azghandi S, et al. Compartment-Resolved Proteomic Analysis of Mouse Aorta During Atherosclerotic Plaque Formation Reveals Osteoclast-Specific Protein Expression. *Mol Cell Proteomics* (2018) 17(2):321–34. doi: 10.1074/mcp.RA117.000315
23. Meng L, Xu Y, Xu C, Zhang W. Biomarker Discovery to Improve Prediction of Breast Cancer Survival: Using Gene Expression Profiling, Meta-Analysis, and Tissue Validation. *Onco Targets Ther* (2016) 9:6177. doi: 10.2147/OTT.S113855
24. Brand E, Chatelain N, Paillard F, Tiret L, Visvikis S, Lathrop M, et al. Detection of Putative Functional Angiotensinogen (AGT) Gene Variants Controlling Plasma AGT Levels by Combined Segregation-Linkage Analysis. *Eur J Hum Genet* (2002) 10(11):715–23. doi: 10.1038/sj.ejhg.5200874
25. Fajar JK, Pikir BS, Sidarta EP, Saka PNB, Akbar RR, Tamara F, et al. The Genes Polymorphism of Angiotensinogen (AGT) M235T and AGT T174M in Patients With Essential Hypertension: A Meta-Analysis. *Gene Rep* (2019) 16:100421. doi: 10.1016/j.genrep.2019.100421
26. Sun S, Sun Y, Rong X, Bai L. High Glucose Promotes Breast Cancer Proliferation and Metastasis by Impairing Angiotensinogen Expression. *Biosci Rep* (2019) 39(6):BSR20190436. doi: 10.1042/BSR20190436
27. Chen W, Chen Y, Zhang K, Yang W, Li X, Zhao J, et al. AGT Serves as a Potential Biomarker and Drives Tumor Progression in Colorectal Carcinoma. *Int Immunopharmacol* (2021) 101:108225. doi: 10.1016/j.intimp.2021.108225


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 30 May 2022
doi: 10.3389/fonc.2022.878923

# Integrative Analysis Identifies Cell-Type-Specific Genes Within Tumor Microenvironment as Prognostic Indicators in Hepatocellular Carcinoma

*Zi-Li Huang[1,2†], Bin Xu[1,3†], Ting-Ting Li[4], Yong-Hua Xu[2], Xin-Yu Huang[1] and Xiu-Yan Huang[1*]*

[1] Department of General Surgery, Shanghai Jiaotong University Affiliated Sixth People's Hospital, Shanghai, China, [2] Department of Radiology, Xuhui District Central Hospital of Zhongshan Hospital, Fudan University, Shanghai, China, [3] Department of General Surgery, The Tenth People's Hospital of Tongji University, Shanghai, China, [4] Department of Infectious Disease, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China

OPEN ACCESS

**Edited by:**
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

**Reviewed by:**
Takeo Toshima,
Kyushu University, Japan
Yong Yi,
Fudan University, China

***Correspondence:**
Xiu-Yan Huang
xyhuang1119@163.com

[†]These authors have contributed equally to this work

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

**Received:** 18 February 2022
**Accepted:** 13 April 2022
**Published:** 30 May 2022

**Citation:**
Huang Z-L, Xu B, Li T-T, Xu Y-H, Huang X-Y and Huang X-Y (2022) Integrative Analysis Identifies Cell-Type-Specific Genes Within Tumor Microenvironment as Prognostic Indicators in Hepatocellular Carcinoma. Front. Oncol. 12:878923. doi: 10.3389/fonc.2022.878923

**Background:** Hepatocellular carcinoma (HCC) is a leading cause of cancer-related mortality worldwide, but effective early detection and prognostication methods are lacking.

**Methods:** The Cox regression model was built to stratify the HCC patients. The single-cell RNA sequencing data analysis and gene set enrichment analysis were employed to investigate the biological function of identified markers. PLCB1 gain- or loss-of-function experiments were performed, and obtained HCC samples were analyzed using quantitative real-time PCR and immunohistochemistry assay to validate the biological function of identified markers.

**Results:** In this study, we developed a model using optimized markers for HCC recurrence prediction. Specifically, we screened out 8 genes through a series of data analyses, and built a multivariable Cox model based on their expression. The risk stratifications using the Eight-Gene Cox (EGC) model were closely associated with the recurrence-free survivals (RFS) in both training and three validation cohorts. We further demonstrated that this risk stratification could serve as an independent predictor in predicting HCC recurrence, and that the EGC model could outperform other models. Moreover, we also investigated the cell-type-specific expression patterns of the eight recurrence-related genes in tumor microenvironment using single-cell RNA sequencing data, and interpreted their functional roles from correlation and gene set enrichment analyses, *in vitro* and *in vivo* experiments. Particularly, PLCB1 and SLC22A7 were predominantly expressed in malignant cells, and they were predicted to promote angiogenesis and to help maintain normal metabolism in liver, respectively. In contrast, both FASLG and IL2RB were specifically expressed in T cells, and were highly correlated

with T cell marker genes, suggesting that these two genes might assist in maintaining normal function of T cell-mediated immune response in tumor tissues.

**Conclusion:** In conclusion, the EGC model and eight identified marker genes could not only facilitate the accurate prediction of HCC recurrence, but also improve our understanding of the mechanisms behind HCC recurrence.

**Keywords: recurrence-free survival, single-cell RNA sequencing, tumor microenvironment, T cell, immune response**

## INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most common malignancies and a leading cause of cancer-related mortality worldwide (1). Common etiologic factors for HCC are HBV/HCV infection, use of alcohol and tobacco, and cirrhosis (2). As few clinical signs can be observed in early-stage HCC, delayed diagnoses often lead to poor prognoses (3). Although targeted therapy and immunotherapy have improved the prognoses of HCC patients to some extent, the odds of recurrence within five years are still relatively high (4–6).

Targeted therapies and immunotherapies are considered promising strategies for unresectable HCC, and therapeutic breakthrough now heavily relies on the identification of potential molecular markers (7). So far, there are several models for predicting HCC prognosis, using different gene signatures such as HOXD9, SPP1, SPINK1, TXNRD1 and MAGEB6, and the satisfying performances of those models suggested that these genes could serve as possible drug targets (8, 9). Since inflammation and immune response play an essential role in the initiation and development of HCC, molecules participating in these processes are receiving growing attention (10). High density of CD3+ and CD8+ cells, along with PD-L1 expression are considered as useful markers for post-surgical, relapse-free survival (1). The α-fetoprotein (AFP) levels are found to be an indicator of recurrence after direct-acting antiviral agent (DAA) therapy for HCC patients with hepatitis C virus infection (11, 12), while another study has demonstrated that IL-11 is associated with HCC recurrence in patients after surgical resection, and blocking IL-11-STAT3 signaling could help prevent post-surgical recurrence (13). Aside from these findings, there is a study focused on predictive lncRNA markers in HCC, and MSC-AS1, POLR2J4, EIF3J-AS1, SERHL, RMST, and PVT1 were recognized as prognostic indicators, which are mainly associated with TGF-β signaling and cellular apoptosis. Meanwhile, a recent study has investigated DNA methylation-driven genes in the growth and metastasis of HCC, concluding that SPP1 and LCAT are candidate markers for HCC recurrence (14). With the aid of advanced sequencing technologies, it is now possible to explore mechanisms behind HCC tumorigenesis from diverse perspectives. Nonetheless, novel predictive biomarkers of HCC recurrence are still inadequately addressed. Here we attempt to develop a predictive model using optimized markers for postsurgical recurrence of HCC, and to interpret their potential functional significance in HCC recurrence.

## MATERIALS AND METHODS

### Data Download and Preprocessing

The three pre-normalized gene expression datasets with clinical characteristics were downloaded from the Gene Expression Omnibus (GEO) database under accession number GSE14520 (15) and GSE76427 (16). The Clinical Proteomic Tumor Analysis Consortium (CPTAC) HCC data was obtained from NODE (The National Omics Data Encyclopedia) with under accession number OEP000321. The clinical characteristics of the three cohorts were summarized in **Supplementary Table S1**. Gene expressions were logarithmically transformed and then used for downstream data analyses.

### Survival Analysis

The Cox proportional hazard regression analysis was employed to identify the recurrence-related genes, and a multivariable Cox model was constructed using identified markers. Specifically, the differential expression analysis was first conducted to identify differentially expressed genes (DEGs) in HCC (Student *t* test, *P*-value < 0.05). Subsequently, the univariable Cox proportional hazard regression analysis was performed to detect DEGs correlated with recurrence-free survival (RFS) in HCC patients (Log rank test, *P*-value < 0.05). It should be noted that for each sample, the expression status of any given gene was denoted as high or low, using its median expression levels in all samples as a cutoff. Furthermore, the MMPC (17) (Max-Min Parents and Children) algorithm was used to identify the best combination of gene signatures with a *P*-value lower than 0.05, and the multivariable Cox model was built based on the selected gene set. The HCC patients were stratified into two risk groups according to the median value of the risk scores, which were estimated using the Cox model and the expression status of the eight signature genes. The Cox model and MMPC algorithm were implemented in R survival (18) and MXM packages (19).

### The Comparison of Cox Models Built on Different Signatures

The hazard ratios and 95% confidence intervals were calculated using different Cox regression models. The forest plots were plotted using R survcomp (20).

### Single-Cell RNA Sequencing Data

We collected the HCC single-cell RNA sequencing (scRNA-seq) data from GEO database under accession number GSE125449 (21), from which the Set 1 was used in this study. The cell types

were pre-defined in the previous study. The t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis and dimensionality reduction were conducted to visualize the clustering of cell types. The cell-type-specific marker genes were identified from the differential expression analysis. The scRNA-seq data analysis was implemented in R Seurat package (22).

## Gene Set Enrichment Analysis (GSEA)

Prior to GSEA, the Spearman correlation analysis was first conducted in this study. Subsequently, genes were ranked according to Spearman correlation coefficients (SCC). GSEA was then conducted to identify the gene sets enriched by genes correlated with identified markers. This analysis was implemented in R clusterProfiler (23). The hepatocyte specific markers were obtained from a liver single-cell RNA-seq study (24).

## Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Analysis

The core genes in the set of T cell marker genes were subjected to the KEGG pathway analysis. The Fisher's exact test was used to measure the statistical significance. The KEGG pathway analysis was implemented in R clusterProfiler package (23).

## Cell Culture

The HCC cell lines, MHCC97H, MHCC97L, SMMC7721 and Hep3B, were cultured following procedures stated in a previous report (25). Specifically, a highly metastatic HCC cell line, MHCC97, were originated from LCI-D20 tumor, which was a subclone possessing high metastatic potentials (26) (up to 100% pulmonary metastatic rate in MHCC97H was reported using orthotopic inoculation). The human HCC cell line Hep3B with very low invasive potentials (27) was also prepared in this study. Human HCC cell lines Hep3B and MHCC97H were obtained from the Liver Cancer Institute, Shanghai Medical College of Fudan University, and cultured in Dulbecco modified Eagle medium (DMEM) (Gibco-BRL, Gaithersburg, MD, USA) supplemented with 10% fetal bovine serum (HyClone, Logan, UT, USA) in a humidified incubator containing 5% $CO_2$ at 37°C. In all experiments, no antibiotics were used.

## Overexpression and Knockdown of PLCB1 Gene

The PLCB1 cDNA was cloned following a previous study (28), and a scrambled sequence was used as a control. The small interfering RNA (siRNA) sequences were designed to target PLCB1. Cell transfection was performed using the Lipofectamine 2000 reagent (Thermo Fisher Scientific) in antibiotic-free medium. The detailed siRNA sequences were as follows:

(+) 5'-CAGAAGAGUGUCAGAACAATT-3',
(–) 5'-UUGUUCUGACACUCUUCUGTT-3'.

## Clinical Specimens

A total of 32 snap-frozen HCC tissues were obtained from the Hospital Clinic and examined using quantitative real-time PCR (qPCR). Informed consent was obtained from each patient, and the Research Ethics Committee of Hospital approved all aspects of this study. Cases with hepatitis B from 2019 to 2021 were confirmed by experienced pathologists based on WHO criteria. The clinical characteristics of the three cohorts were summarized in **Supplementary Table S1**.

## Quantitative Real-Time PCR

Quantitative PCR assays of cDNA were performed using a CFX96 Real-time PCR system (Bio-Rad). Relative expressions of the target transcripts were quantified and normalized to the expression of the reference gene GAPDH. Specifically, expressions of target genes in each sample were first assessed by qRT-PCR independently. Target cDNAs were amplified using the following probe set:

B3GAT3: forward 5'-CTCGGCCAGCCATGTGAC-3'; reverse 5'-TCCAGCCATATCCACAGGGA-3'.

EEF1D: forward 5'-TTCATCAGTCTTCCCGCGTC-3'; reverse 5'-CTTGTTGACCCAGATCCCCC-3'.

NRP1: forward 5'-GACCTGGGGGAGGAGAAGAT-3'; reverse 5'-GATCCTGAATGGGTCCCGTC-3'.

PLCB1: forward 5'-GGACTGACCCTCAGGGATTTT-3'; reverse 5'-AAGCCACGAGATTCAAATGGG-3'.

FASLG: forward 5'-CTTGGTAGGATTGGGCCTGG-3'; reverse 5'-CTGGCTGGTAGACTCTCGGA-3'.

IL2RB: forward 5'-CATGTCTCAGCCAGGGCTTC-3'; reverse 5'-GTTGCATCTGTGGGTCTCCA-3'.

SLC22A7: forward 5'-CCAGAGTCCAAGGGTCTATGT-3'; reverse 5'-ATCAAGGATGGATGAGCAGAG-3'.

STX11: forward 5'-ACACGTAAGCAGGAAGCAGC-3'; reverse 5'-GTGCAGTGGGCCAAATGATG-3'.

GAPDH: forward 5'-ACCCACTCCTCCACCTTTG-3'; reverse 5'-CTCTTGTGCTCTTGCTGGG-3'.

## Immunohistochemistry Assay

Intratumoral microvessels immunostained for CD31 were counted using the method described by Weidner et al. (29). All slides were independently assessed by two board-certified pathologists that were blinded in this experiment. Any disagreements in the microvessel count were resolved by consensus.

## Cell Invasion Assay

Cell invasion assays were performed in a Transwell apparatus (Millipore), following the manufacturer's instructions. Briefly, after being serum-starved for 24 h, the cells were seeded in the top chamber of the Transwell apparatus (100 μL DMEM, 5 × 104 cells/well), which was coated with collagen IV, and 600 μL conditioned medium was added to the bottom chamber of the apparatus. The cells were subsequently incubated for 48 h at 37°C and then fixed with 0.1% methanol for 10 min before hematoxylin staining. Migrated cells were counted in three independent experiments, and the results were presented as means ± SD.

## MMP9 Activity Assay

MMP9 activity was measured using a human MMP9 Activity ELISA System (Amersham Pharmacia Biotech, Piscataway, NJ), following the manufacturer's instructions. The plate was read at 450 nm in a SPECTRAmax 250 Microplate Spectrophotometer

(Molecular Devices, Sunnyvale, CA). Assays were repeated in triplicate.

### *In Vivo* Study

Male athymic BALB/c nu/nu mice of 18–20 g at 5 weeks' age were handled according to the recommendations of the National Institutes of Health Guidelines for Care and Use of Laboratory Animals. The experimental protocol was approved by the Shanghai Medical Experimental Animal Care Committee. Human HCC tumor models produced by MHCC97H were established in nude mice by orthotopic inoculation. We silenced the expression of PLCB1 in MHCC97H cells using the small interface RNA (siRNA). Then, $1 \times 10^5$ cells were subcutaneously injected to form the tumors, which were inoculated *in situ* in nude mice (KD group), while CON group presented siRNA negative group. Briefly, the left lobe of the liver was exposed under anesthesia, and a part of the liver surface was mechanically injured with scissors. Then, a piece of MHCC97H tumor tissue (size 2 × 2 × 2 mm) was fixed within the liver tissue.

## RESULTS

### Identification of Recurrence-Related Genes in HCCs

To build a predictive model for HCC recurrence, we first conducted univariable Cox regression analyses and a differential gene expression analysis to identify recurrence-related genes, using a gene expression dataset with a long-term follow-up of patients from GEO under accession number GSE14520 (**Figure 1A**). Specifically, we identified 138 recurrence-related genes (**Supplementary Table S2**) and screened out 8 genes to construct a multivariable Cox model, including B3GAT3, EEF1D, NRP1, PLCB1, FASLG, IL2RB, SLC22A7, and STX11, using MMPC algorithm. Notably, expressions of the first four recurrence-related genes and of the remaining ones were negatively and positively correlated with the recurrence-free survival (RFS) in HCC patients, respectively (**Figure 1B**). Consistently, the expressions of upregulated or downregulated genes appeared to be negatively or positively correlated with RFS in HCC, compared with the adjacent normal tissues (**Figure 1C**, Wilcoxon rank sum test, $P < 0.001$). These results indicated that the eight genes were particularly relevant to HCC recurrence.

### Construction and Independent Validation of the Eight-Gene Cox (EGC) Model

Based on expression profiles of the eight selected genes, we built an Eight-Gene Cox (EGC) model, and HCC patients were stratified into high and low risk groups by the median risk scores (see *Materials and Methods*). Significant differences in RFS between risk groups were observed in the training cohort (GSE14520) (**Figure 2A**, log rank test, P < 0.0001). Consistently, in two validation cohorts [CPTAC (Clinical Proteomic Tumor

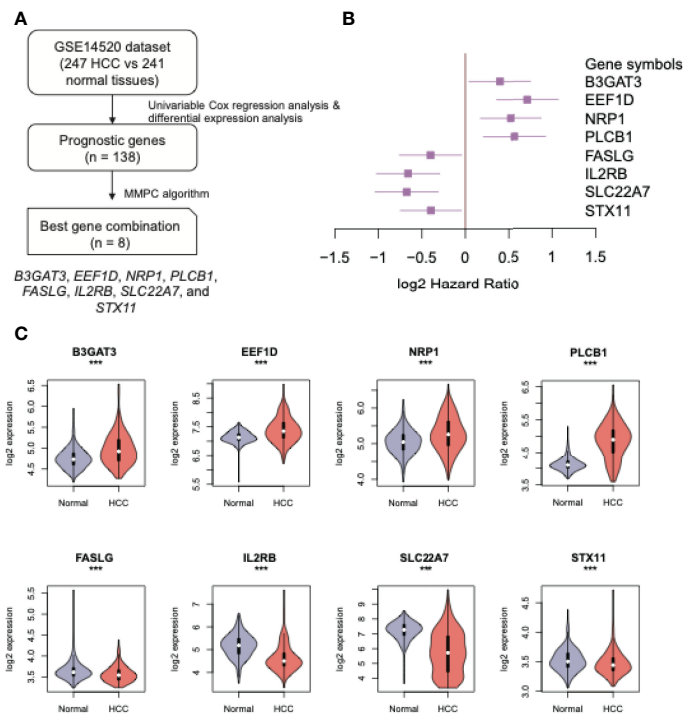



**FIGURE 1** | The eight recurrence-related genes in HCC. **(A)** The workflow for the identification of the eight recurrence-related genes. **(B)** The forest plot displays the log2 (hazard ratios) (the boxes) and its 95% confidence intervals (the two ends) of the eight recurrence-related genes. **(C)** The expression patterns of the eight recurrence-related genes in HCC and adjacent normal tissues. The orange and purple colors represent the HCC and adjacent normal tissues, respectively. ***: P-value < 0.05.

Analysis Consortium) and GSE76427], the RFS in the high-risk group was shorter (**Figures 2B, C**). Furthermore, we collected 32 snap-frozen HCC tissues, and measured RNA abundance of the eight genes using quantitative real-time PCR (qPCR). Consistently, in this cohort, significantly shorter RFS was observed in the high-risk group, compared with the low-risk group (**Figure 2D** and **Supplementary Table S3**, log-rank test, *P*-value < 0.05), further suggesting that the EGC model could ensure robust predictions of HCC recurrence. In addition, we also compared the performance of MMPC with that of stepwise regression and Lasso algorithms, respectively, and found that MMPC achieved the highest C-index and statistical significance compared with stepwise regression and Lasso algorithms (**Supplementary Table S4**). These results suggested that the EGC model provided an enhanced performance and consistency in HCC recurrence prediction.

## The Risk Stratification by EGC Model Is an Independent Predictor of HCC Recurrence

To further demonstrate that the EGC model was an independent predictor of HCC recurrence, we built a multivariable Cox model based on this risk stratification and four other prognostically relevant risk factors, including Barcelona Clinic Liver Cancer (BCLC) stage, Alpha Fetoprotein (AFP), tumor size, and age. Notably, the validation cohort under accession number GSE76427 was excluded in this analysis, due to a lack of clinical characteristics. In the multivariable Cox model, significant p-values for BCLC and the risk stratification were observed in the training and the validation (CPTAC) cohorts (**Tables 1**, **2**). More importantly, the statistical significance of the risk stratification was higher than BCLC. These results indicated that the risk stratification by EGC model was an independent prognostic factor in HCC recurrence prediction.

## The EGC Model Outperforms Previously Published Models

To demonstrate superiority of the EGC model, we compared the EGC model with three previously reported HCC recurrence-predicting models. The samples in two validation datasets were also stratified into two risk groups based on expression profiles of different sets of recurrence-related genes from those studies. The comparison of the EGC model and those three models revealed that the EGC model performance was better than the others (9, 30, 31) (**Figures 3A, B**). It should be noted that the other three models were also relatively less significant in GSE76427 cohort. These results demonstrated that the EGC model outperforms previous published models in risk stratification.

## Cell Type-Specific Expression Patterns of the Eight Recurrence-Related Genes

To elucidate the expression patterns of these recurrence-related genes in the tumor microenvironment, we collected single-cell RNA sequencing (scRNA-seq) data of 12 HCC samples from a previous study. The t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis and dimensionality reduction of the scRNA-seq data revealed that the cells could be divided into eight cell clusters, and seven out of these cell clusters could be annotated using the cell type specific marker genes reported by Ma L. et al. (21), namely B cells, cancer-associated fibroblasts (CAF), hepatic progenitor cell (HPC)-like, malignant cells, T cell, tumor-associated macrophages (TAM), and tumor-associated

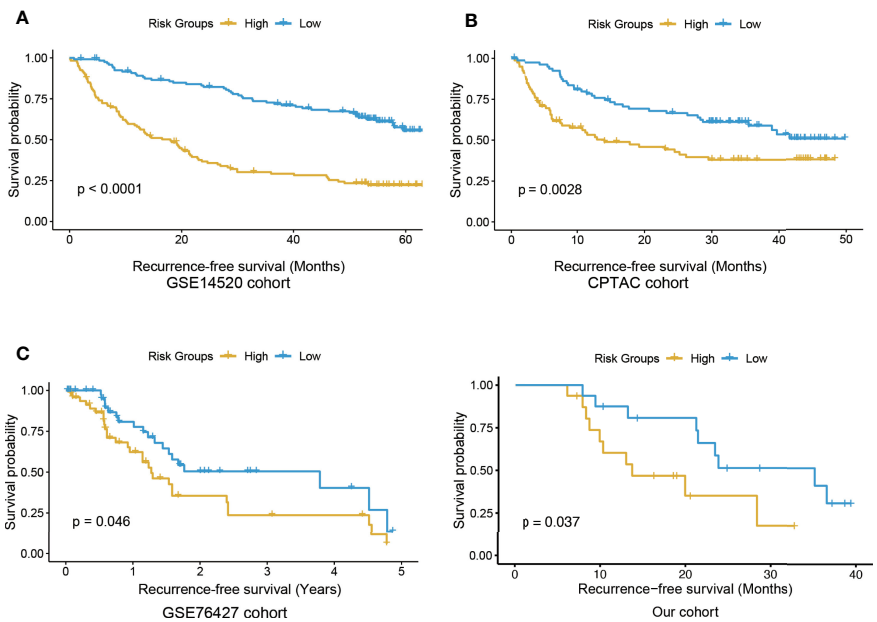



**FIGURE 2** | The Kaplan-Meier curves for the high and low risk groups in the training and three validation cohorts. The KM curves of training and three validation cohorts were displayed in **(A–D)**, respectively. The yellow and blue curves represent the high and low risk groups, respectively.

**TABLE 1 |** The coefficients and statistical significance of the risk stratification and four prognostic factors in the multivariable model (training cohort).

| Factor | | Coef. | Exp (coef.) | Se (coef.) | z | Pr (>\|z\|) | |
|---|---|---|---|---|---|---|---|
| BCLC | | | | | | | |
| | A | 0.88 | 2.42 | 0.43 | 2.07 | 3.89E-02 | * |
| | B | 1.10 | 2.99 | 0.50 | 2.19 | 2.84E-02 | * |
| | C | 1.67 | 5.29 | 0.49 | 3.41 | 6.52E-04 | * |
| AFP | | | | | | | |
| | <300 ng/ml | -0.04 | 0.96 | 0.18 | -0.24 | 8.09E-01 | |
| Tumor size | | | | | | | |
| | <5cm | 0.09 | 1.10 | 0.21 | 0.46 | 6.48E-01 | |
| Age | | | | | | | |
| | <50 yrs. | -0.20 | 0.82 | 0.18 | -1.10 | 2.70E-01 | |
| Risk group | | | | | | | |
| | Low-risk | -1.15 | 0.32 | 0.20 | -5.71 | 1.11E-08 | *** |

**: P-value < 0.05, ***: P-value < 0.001.*

endothelial cells (TEC) (**Figure 4A**). Among the eight recurrence-related genes, PLCB1, FASLG, IL2RB, and SLC22A7 were found to be specifically expressed in only one cell type. Specifically, PLCB1 and SLC22A7 were expressed in malignant cells, while both FASLG and IL2RB were expressed in T cells (**Figure 4B**). In contrast, B3GAT3, EEF1D, NRP1, and STX11 were expressed in multiple cell types. These results indicated that the expression patterns of SLC22A7, PLCB1, FASLG, and IL2RB might be cell type-specific.

## Interpretation of Cell-Type Specific Recurrence-Related Genes in HCC

To investigate how those four cell-type-specific genes participated in HCC recurrence, we conducted correlation analysis and gene set enrichment analysis (GSEA) to interpret their functional roles using the training cohort. SLC22A7 was specifically expressed in malignant cells, however, this gene was downregulated in HCC tissues (**Figure 1C**). The genes whose expressions were positively correlated with SLC22A7 expression were highly enriched in the hepatocyte marker genes (**Figure 5A**), further suggesting that SLC22A7 was a hepatocyte-specific marker gene. As shown in **Figure 5B**, highly similar expression patterns of SLC22A7 and of the core genes in those SLC22A7-related ones were observed. The KEGG pathway enrichment analysis of these genes revealed that they were highly enriched in liver-related metabolism (**Figure 5C**), suggesting that SLC22A7 was involved in liver-related metabolism. PLCB1 was also identified as a malignant cell-specific gene, and was upregulated in HCC. The GSEA revealed that the genes correlated with PLCB1, including HEY1, EZH2, VEGFA, E2F3, and STC1, were enriched in angiogenesis (**Figures 5D, E**), suggesting that the increased expression of PLCB1 might be associated with HCC recurrence *via* tumor-associated angiogenesis.

As FASLG and IL2RB were identified as two candidate T cell specific marker genes, we thus speculated that the downregulation of these two genes might be associated with T cell-mediated anticancer properties. Consistently, expressions of these two genes were positively correlated with expressions of T cell marker genes (**Figures 6A, B**). Notably, CD96, TRAT1, CD6, CD3D, CD3E, CD3G, TRAC, BCL11B, PRKCQ, CD2, TRBC1, SH2D1A, GIMAP5, and LCK were identified as the core genes in T cell marker genes (**Figures 6C, D**). These results indicated that low expression levels of FASLG and IL2RB might be associated with attenuated T cell-mediated anticancer activities, thus resulting in HCC recurrence.

In addition, we also conducted GSEA for the remaining genes, including B3GAT3, EEF1D, NRP1, and STX11, to interpret their biological functions in HCC. Of note, B3GAT3, EEF1D, NRP1, and STX11 were predicted to be involved in gap junction trafficking and regulation, protein translation, collagen formation, and G-protein coupling receptors (GPCRs)

**TABLE 2 |** The coefficients and statistical significance of the risk stratification and four prognostic factors in the multivariable model (validate (CPTAC) cohort).

| Factor | | Coef. | Exp (coef.) | Se (coef.) | z | Pr (>\|z\|) | |
|---|---|---|---|---|---|---|---|
| BCLC | | | | | | | |
| | B | 0.09 | 1.10 | 0.45 | 0.21 | 8.36E-01 | |
| | C | 0.89 | 2.42 | 0.42 | 2.12 | 3.38E-02 | * |
| AFP | | | | | | | |
| | <300 ng/ml | 0.02 | 1.02 | 0.24 | 0.07 | 9.41E-01 | |
| Tumor size | | -0.26 | 0.77 | 0.40 | -0.64 | 5.25E-01 | |
| | <5cm | | | | | | |
| Age | | 0.32 | 1.38 | 0.24 | 1.31 | 1.89E-01 | |
| | <50 yrs. | | | | | | |
| Risk group | | | | | | | |
| | Low-risk | -0.54 | 0.58 | 0.24 | -2.29 | 2.19E-02 | * |

**: P-value < 0.05.*

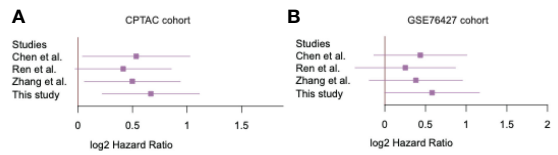



**FIGURE 3** | The significance of the recurrence predictive Cox models. The forest plots display the performance of the four Cox models in two validation cohorts **(A, B)**.

signaling according to GSEA, respectively (**Supplementary Table S5**).

## PLCB1 Silencing Decreases the Microvessel Density (MVD) in HCC

As PLCB1 was upregulated in malignant cells and was predicted to be involved in angiogenesis according to GSEA, we then investigated whether PLCB1 overexpression could lead to increased microvessel density in HCC. Two HCC cell lines, MHCC97H (high metastatic potential) and Hep3B (low metastatic potential), were selected for this analysis, following our previous studies (25, 32). Firstly, we silenced and overexpressed PLCB1 in MHCC97H and Hep3B cells, using the small interface RNA (siRNA) and an PLCB1 expression vector (n = 3), respectively. Nude mice were subcutaneously inoculated with those cells. Then, *in situ* HCC nude mice model was established. As shown in **Figure 7A**, the mRNA expression levels of PLCB1 were significantly reduced after siRNA silencing (KD group), while an increase was observed after PLCB1 overexpression (OE group) (n = 3). Furthermore, we measured the MVD in tumor tissues from the nude mice, and found that the KD group had significantly lower MVD than the controls (CON group) (**Figures 7B, C**, t test, $P$-value < 0.05), indicating that PLCB1 silencing could decrease the MVD in HCC.

## SLC22A7 Inhibits HCC Metastasis by Reducing MMP9 Activity

As SLC227A was specifically expressed in hepatocytes, we then examined its expression patterns in cell lines with varied metastatic potentials. We quantified the expression levels of SLC22A7 in HCC cell lines including Hep3B, SMMC7721, MHCC97L, and MHCC97H. Specifically, we found that SLC22A7 was expressed higher in lowly metastatic Hep3B cells than highly metastatic MHCC97H cells (student-$t$ test, $P < 0.001$, **Figure 8A**). Accordingly, lower expression of SLC22A7 in highly metastatic HCC cells was also observed in the RNA sequencing

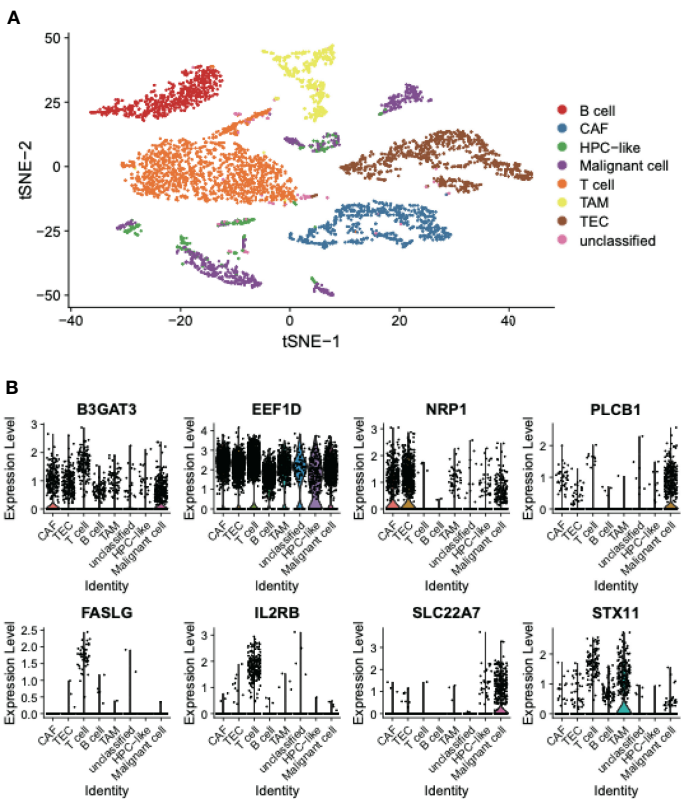



**FIGURE 4** | The cell type specific expression in HCC tissues. **(A)** The t-SNE plot represents the distribution of cell types in HCC tissues. **(B)** The expression patterns of the eight recurrence-related genes in the EGC model. The points represent the expression levels of the cells.

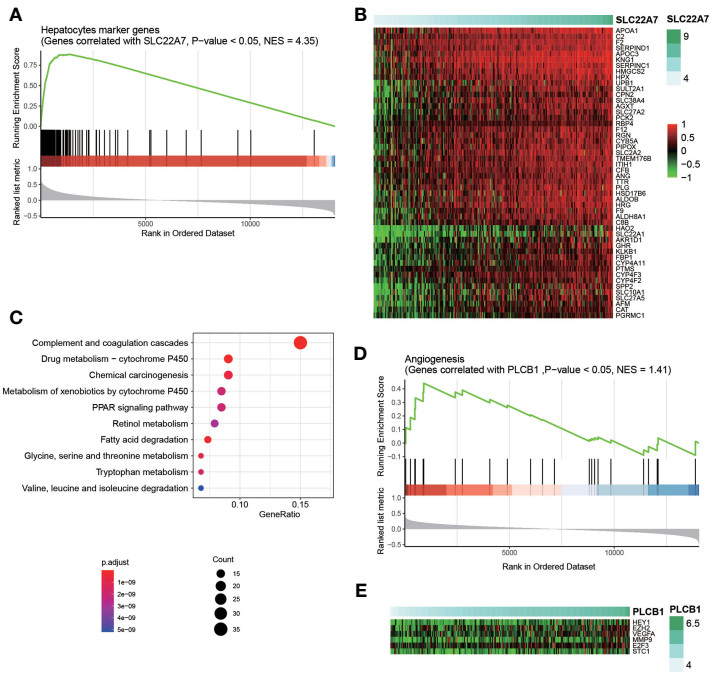



**FIGURE 5** | Interpretation of SLC22A7 and PLCB1 in HCC recurrence. **(A)** The statistical significance of the genes correlated with SLC22A7 enriched in the hepatocytes marker genes. The y axis on the bottom represents the Spearman correlation coefficients. **(B)** The heatmap for the core genes positively correlated SLC22A7 and also specifically expressed in hepatocytes. **(C)** The KEGG pathways enriched by the core genes positively correlated SLC22A7 and also specifically expressed in hepatocytes. **(D)** The statistical significance of the genes correlated with PLCB1 enriched in the angiogenesis-related genes. **(E)** The heatmap for the core genes positively correlated PLCB1 and also involved in angiogenesis.

data from a previous study (33) (**Figure 8B**). These results indicated that low expression of SLC22A7 might be associated with HCC metastasis.

To further explore the functional roles of SLC22A7 in HCC metastasis, we overexpressed SLC22A7 in MHCC97H cells, and used MHCC97H cells transfected with the empty vector as controls. Compared with the control group, SLC22A7 overexpression significantly attenuated the invasive ability of MHCC97H cells in Transwell assay (**Figure 8C**). Moreover, we also evaluated the activity of matrix metalloproteinase-9

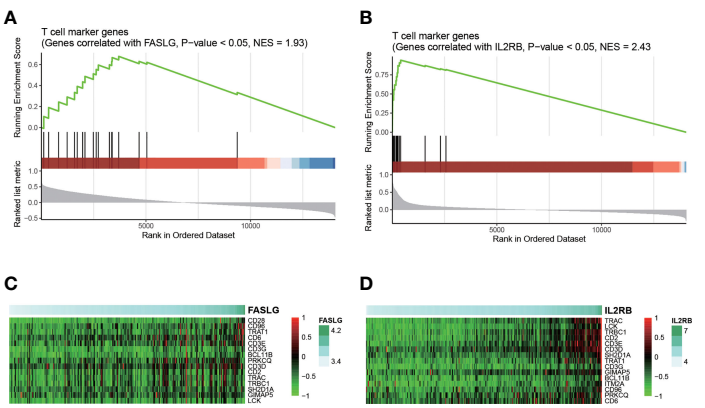



**FIGURE 6** | Interpretation of FASLG and IL2RB in HCC recurrence. The statistical significance of the genes correlated with FASLG and IL2RB enriched in T cell marker genes are displayed in **(A, B)**, respectively. The heatmaps for the core genes positively correlated FASLG and IL2RB and also specifically expressed in T cell are displayed in **(C, D)**, respectively.

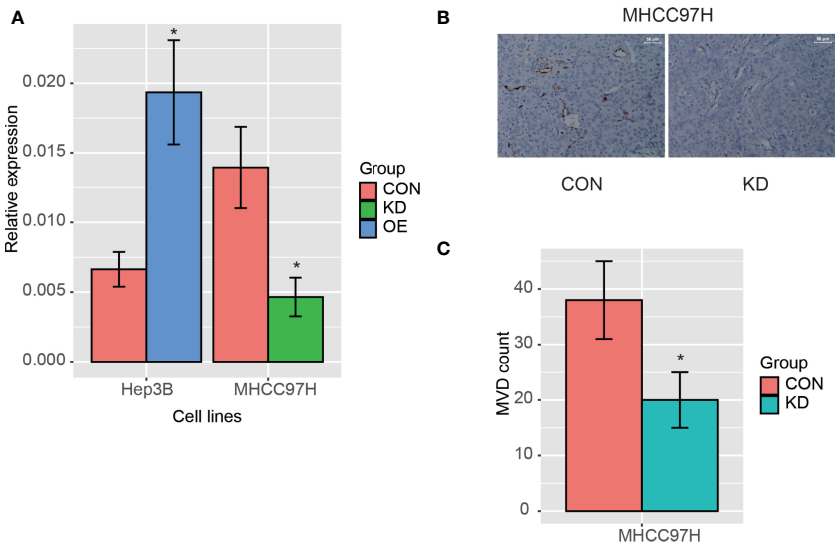



**FIGURE 7** | The PLCB1 expression in HCC cell lines. **(A)** The RNA expression of PLCB1 by the knockdown and overexpression of PLCB1 mRNA. **(B)** The microvessel density (MVD) in HCC tissue with PLCB1 knockdown and the control. **(C)** The MVD count in HCC tissue with PLCB1 knockdown and the control. *: P-value < 0.05.

(MMP9) in MHCC97H using an ELISA kit, and found that MMP9 activity was significantly decreased after SLC22A7 overexpression (**Figure 8D**), suggesting that SLC22A7 might prevent HCC invasion by inhibiting MMP9 activity.

## DISCUSSION

There is still a lack of effective early detection and prognostication methods for HCC, which remains a leading cause of cancer-related mortality worldwide. In the present study, we have attempted to develop a predictive model with optimized markers for HCC recurrence. Specifically, we screened out 8 genes, including B3GAT3, EEF1D, NRP1, PLCB1, FASLG, IL2RB, SLC22A7, and STX11, through a series of data analyses such as differential expression analysis, Cox regression analysis, and MMPC algorithm. The risk stratification based on the Eight-Gene Cox (EGC) model was closely associated with the RFS of HCC patients in both training and three validation cohorts. Of note, the reduced strength of EGC model in the validation cohorts might be resulted from the reduced sample size. We have demonstrated that the EGC model could serve as an independent predictor of HCC recurrence, through a multivariable Cox model constructed with the risk stratification (samples were stratified as either low or high risk) and four prognostically relevant variables. Consistently, the superiority of the EGC model was demonstrated by comparing it with other three models (9, 30, 31). In short, the EGC model could accurately predict the risk of recurrence in HCC.

As single-cell RNA sequencing analysis facilitates the identification of HCC recurrence-related genes and interpretation of their functional roles at a higher resolution, we investigated the expression patterns of selected genes in different cell populations from HCC tissues. Among those eight genes in the EGC model, B3GAT3, EEF1D, NRP1, and STX11 were not expressed in a cell type-specific manner. B3GAT3 is a member of glucuronyltransferase, and is involved in glycosaminoglycans biosynthesis (34). High expression of B3GAT3 has been reported to be associated with poor prognosis in liver cancer (35). Moreover, EEF1D, an eukaryotic translation elongation factor, also exhibits prognostic significance in multiple cancers (36). NRP1, a transmembrane co-receptor for semaphorins and heparin-binding pro-angiogenic cytokines, has been reported to be upregulated in hepatocellular carcinoma, as it contributes to tumor growth and vascular remodeling (37). However, STX11 was found downregulated in HCC tissues, and it has been reported to function as a tumor suppressor gene in peripheral T-cell lymphomas (38). These results suggested that these four genes might also be associated with HCC recurrence in similar manners.

Particularly, PLCB1 and SLC22A7 were predominantly expressed in malignant cells, while both FASLG and IL2RB were specifically expressed in T cells (**Figure 4B**). PLCB1 was identified as a malignant cell-specific genes in HCC, and has been reported to be implicated in breast cancer (39), HCC (40), and ovarian cancer (41). However, its functional role in HCC has not been fully unveiled. In this study, we found that PLCB1 silencing could decrease the MVD in HCC, suggesting that PLCB1 silencing might inhibit the angiogenesis of HCC. SLC22A7, which was downregulated in HCC yet specifically expressed in malignant cells, has been reported to affect

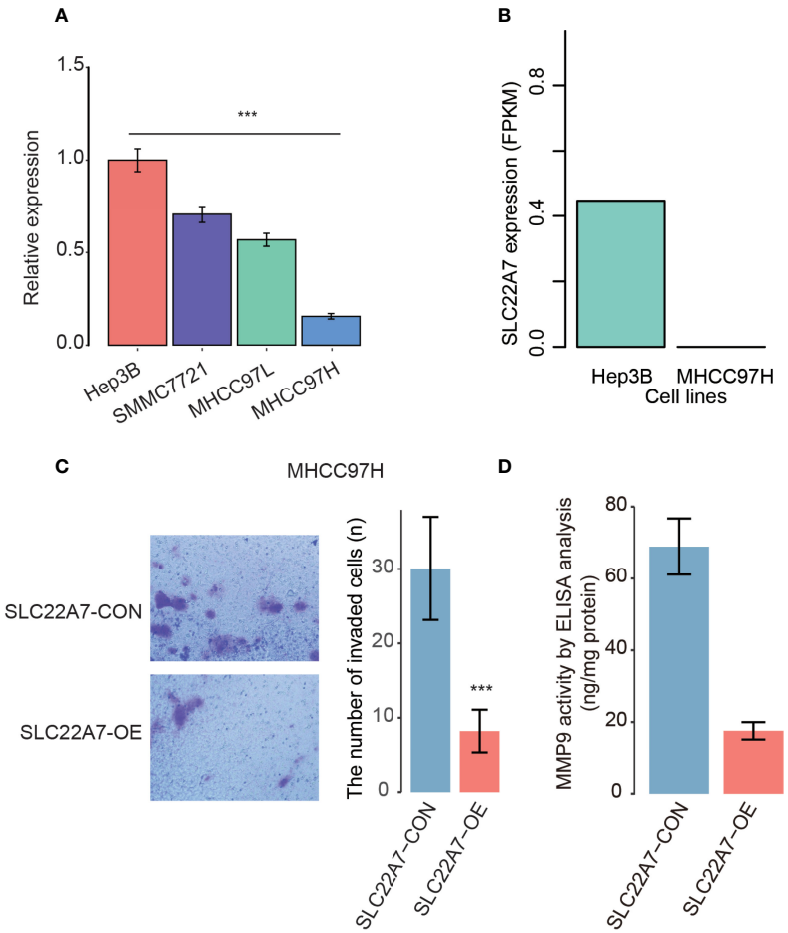



**FIGURE 8** | The SLC22A7 expression in HCC cell lines. **(A)** The SLC22A7 expression levels detected by qRT-PCR. The expression levels were normalized to Hep3B. **(B)** The SLC22A7 expression levels quantified by RNA-seq data. **(C)** The number of invaded cells in SLC22A7-overexpression (SLC22A7-OE) and control (SLC22A7-CON) MHCC97H cell lines. **(D)** MMP9 activity in SLC22A7-OE and SLC22A7-CON groups by ELISA analysis. ***: P-value < 0.001.

mitochondrion and oxidoreductase in noncancerous liver tissues, thereby promoting the occurrence of HCC (42). The KEGG pathway analysis further confirmed that SLC22A7 was associated with normal metabolism in human liver. Kudo et al. reported that SLC22A7 was the best predictor of multicentric occurrence (MO)-free survival (MFS), and SLC22A7 downregulation was associated with mitochondrion and oxidoreductase activity (42). Moreover, reduced SLC22A7 expression in the liver could indicate a significant risk of HCC development in chronic hepatitis C, which was independent of other risk factors as described by Yasui et al. (43). As the hepatocyte-specific markers were abundantly expressed in highly differentiated liver cancer cells, the close association between SLC22A7 and hepatocyte-specific markers suggested that SLC22A7 might participate in tumor differentiation. Moreover, we also observed that SLC22A7 was expressed lower in highly metastatic MHCC97H cells than lowly metastatic Hep3B cells. It is well recognized that histological differentiation grade of HCC is associated with recurrence (44–46). To further explore the functional roles of SLC22A7 in HCC metastasis, we overexpressed SLC22A7 in MHCC97H cells, and found that SLC22A7 overexpression significantly decreased the invasive ability of MHCC97H cells and reduced the activity of matrix metalloproteinase-9 (MMP9), suggesting that SLC22A7 might prevent HCC invasion by inhibiting MMP9 activity. In addition, it is widely reported that FASLG and IL2RB were specifically expressed in T cells (47–50). High infiltration of T cells was associated with low recurrence rate in HCC (51, 52), and cytotoxic T cells were key contributors to anticancer immunity in multiple cancers (53–55). Thus, positive correlation between the expressions of FASLG/IL2RB and T cell markers suggested that FASLG/IL2RB might promote T cell-induced anticancer immunity in HCC.

In addition, the present study still had some limitations. One is the lack of a larger cohort when developing the proposed Cox model. Another limitation is that *in vivo*

experiments in HCC mice models should be proposed to validate the molecular mechanisms. In conclusion, the EGC model and eight identified marker genes could not only facilitate the accurate prediction of HCC recurrence, but also improve our understanding of the mechanisms behind HCC recurrence.

## 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

X-YaH designed and coordinated the study; Z-LH, BX and T-TL performed the experiments, acquired and analyzed data; X-YaH, Z-LH, Y-HX, T-TL, and X-YuH interpreted the data; Z-LH, BX, and X-YaH wrote the manuscript; all authors approved the final version of the article.

## FUNDING

This study was supported by grants from the Project funded by the central government to guide local scientific and technological development (YDZX20213100001001), and the Interdisciplinary Program of Shanghai Jiao Tong University (No. YG2017MS13).

## ACKNOWLEDGMENTS

The authors acknowledge the contribution of all investigators at all participating study sites.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.878923/full#supplementary-material

## REFERENCES

1. Gabrielson A, Wu Y, Wang H, Jiang J, Kallakury B, Gatalica Z, et al. Intratumoral CD3 and CD8 T-Cell Densities Associated With Relapse-Free Survival in HCC. *Cancer Immunol Res* (2016) 4(5):419–30. doi: 10.1158/2326-6066.CIR-15-0110
2. Massarweh NN, El-Serag HB. Epidemiology of Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma. *Cancer Control* (2017) 24 (3):1073274817729245. doi: 10.1177/1073274817729245
3. Vilgrain V, Pereira H, Assenat E, Guiu B, Ilonca AD, Pageaux GP, et al. Efficacy and Safety of Selective Internal Radiotherapy With Yttrium-90 Resin Microspheres Compared With Sorafenib in Locally Advanced and Inoperable Hepatocellular Carcinoma (SARAH): An Open-Label Randomised Controlled Phase 3 Trial. *Lancet Oncol* (2017) 18(12):1624–36. doi: 10.1016/S1470-2045(17)30683-6
4. Rizzo A, Dadduzio V, Ricci AD, Massari F, Di Federico A, Gadaleta-Caldarola G, et al. Lenvatinib Plus Pembrolizumab: The Next Frontier for the Treatment of Hepatocellular Carcinoma? *Expert Opin Investig Drugs* (2021) 2021:1–8. doi: 10.1080/13543784.2021.1948532
5. Rizzo A, Brandi G. Biochemical Predictors of Response to Immune Checkpoint Inhibitors in Unresectable Hepatocellular Carcinoma. *Cancer Treat Res Commun* (2021) 27:100328. doi: 10.1016/j.ctarc.2021.100328
6. Rizzo A, Ricci AD. PD-L1, TMB, and Other Potential Predictors of Response to Immunotherapy for Hepatocellular Carcinoma: How can They Assist Drug Clinical Trials? *Expert Opin Investig Drugs* (2021) 2021:1–9. doi: 10.1080/13543784.2021.1972969
7. Llovet JM, Montal R, Sia D, Finn RS. Molecular Therapies and Precision Medicine for Hepatocellular Carcinoma. *Nat Rev Clin Oncol* (2018) 15 (10):599–616. doi: 10.1038/s41571-018-0073-4
8. Long J, Zhang L, Wan X, Lin J, Bai Y, Xu W, et al. A Four-Gene-Based Prognostic Model Predicts Overall Survival in Patients With Hepatocellular Carcinoma. *J Cell Mol Med* (2018) 22(12):5928–38. doi: 10.1111/jcmm.13863
9. Chen PF, Li QH, Zeng LR, Yang XY, Peng PL, He JH, et al. A 4-Gene Prognostic Signature Predicting Survival in Hepatocellular Carcinoma. *J Cell Biochem* (2019) 120(6):9117–24. doi: 10.1002/jcb.28187
10. Keenan BP, Fong L, Kelley RK. Immunotherapy in Hepatocellular Carcinoma: The Complex Interface Between Inflammation, Fibrosis, and the Immune Response. *J Immunother Cancer* (2019) 7(1):267. doi: 10.1186/s40425-019-0749-z
11. Asahina Y, Tsuchiya K, Nishimura T, Muraoka M, Suzuki Y, Tamaki N, et al. Alpha-Fetoprotein Levels After Interferon Therapy and Risk of Hepatocarcinogenesis in Chronic Hepatitis C. *Hepatology* (2013) 58 (4):1253–62. doi: 10.1002/hep.26442
12. Nakano M, Koga H, Ide T, Kuromatsu R, Hashimoto S, Yatsuhashi H, et al. Predictors of Hepatocellular Carcinoma Recurrence Associated With the Use of Direct-Acting Antiviral Agent Therapy for Hepatitis C Virus After Curative Treatment: A Prospective Multicenter Cohort Study. *Cancer Med* (2019) 8 (5):2646–53. doi: 10.1002/cam4.2061
13. Wang D, Zheng X, Fu B, Nian Z, Qian Y, Sun R, et al. Hepatectomy Promotes Recurrence of Liver Cancer by Enhancing IL-11-STAT3 Signaling. *EBioMedicine* (2019) 46:119–32. doi: 10.1016/j.ebiom.2019.07.058
14. Long J, Chen P, Lin J, Bai Y, Yang X, Bian J, et al. DNA Methylation-Driven Genes for Constructing Diagnostic, Prognostic, and Recurrence Models for Hepatocellular Carcinoma. *Theranostics* (2019) 9(24):7251–67. doi: 10.7150/thno.31155
15. Roessler S, Jia HL, Budhu A, Forgues M, Ye QH, Lee JS, et al. A Unique Metastasis Gene Signature Enables Prediction of Tumor Relapse in Early-Stage Hepatocellular Carcinoma Patients. *Cancer Res* (2010) 70(24):10202–12. doi: 10.1158/0008-5472.CAN-10-2607
16. Grinchuk OV, Yenamandra SP, Iyer R, Singh M, Lee HK, Lim KH, et al. Tumor-Adjacent Tissue Co-Expression Profile Analysis Reveals Pro-Oncogenic Ribosomal Gene Signature for Prognosis of Resectable Hepatocellular Carcinoma. *Mol Oncol* (2018) 12(1):89–113. doi: 10.1002/1878-0261.12153
17. Brown LE, Tsamardinos L, Aliferis CF. A Novel Algorithm for Scalable and Accurate Bayesian Network Learning. *St Heal T* (2004) 107:711–5. doi: 10.3233/978-1-60750-949-3-711
18. Li JCA. Modeling Survival Data: Extending the Cox Model. *Sociol Method Res* (2003) 32(1):117–20. doi: 10.1177/0049124103031004005
19. Lagani V, Athineou G, Farcomeni A, Tsagris M, Tsamardinos I. Feature Selection With the R Package MXM: Discovering Statistically Equivalent Feature Subsets. *J Stat Softw* (2017) 80(7)1–25. doi: 10.18637/jss.v080.i07
20. Schroder MS, Culhane AC, Quackenbush J, Haibe-Kains B. Survcomp: An R/Bioconductor Package for Performance Assessment and Comparison of Survival Models. *Bioinformatics* (2011) 27(22):3206–8. doi: 10.1093/bioinformatics/btr511
21. Ma L, Hernandez MO, Zhao Y, Mehta M, Tran B, Kelly M, et al. Tumor Cell Biodiversity Drives Microenvironmental Reprogramming in Liver Cancer. *Cancer Cell* (2019) 36(4):418–30.e416. doi: 10.1016/j.ccell.2019.08.007
22. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive Integration of Single-Cell Data. *Cell* (2019) 177(7):1888. doi: 10.1016/j.cell.2019.05.031

23. Yu GC, Wang LG, Han YY, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *Omics-a J Integr Biol* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118
24. Aizarani N, Saviano A, Sagar, Mailly L, Durand S, Herman JS, et al. A Human Liver Cell Atlas Reveals Heterogeneity and Epithelial Progenitors. *Nature* (2019) 572(7768):199–204. doi: 10.1038/s41586-019-1373-2
25. Huang XY, Huang ZL, Zhang PB, Huang XY, Huang J, Wang HC, et al. CircRNA-100338 Is Associated With mTOR Signaling Pathway and Poor Prognosis in Hepatocellular Carcinoma. *Front Oncol* (2019) 9:392. doi: 10.3389/fonc.2019.00392
26. Sun FX, Tang ZY, Lui KD, Ye SL, Xue Q, Gao DM, et al. Establishment of a Metastatic Model of Human Hepatocellular Carcinoma in Nude Mice *via* Orthotopic Implantation of Histologically Intact Tissues. *Int J Cancer* (1996) 66 (2):239–43. doi: 10.1002/(SICI)1097-0215(19960410)66:2<239::AID-IJC17>3.0.CO;2-7
27. Giannelli G, Fransvea E, Marinosci F, Bergamini C, Colucci S, Schiraldi O, et al. Transforming Growth Factor-Beta1 Triggers Hepatocellular Carcinoma Invasiveness *via* Alpha3beta1 Integrin. *Am J Pathol* (2002) 161(1):183–93. doi: 10.1016/S0002-9440(10)64170-3
28. Li J, Zhao X, Wang D, He W, Zhang S, Cao W, et al. Up-Regulated Expression of Phospholipase C, Beta1 Is Associated With Tumor Cell Proliferation and Poor Prognosis in Hepatocellular Carcinoma. *OncoTargets Ther* (2016) 9:1697–706. doi: 10.2147/OTT.S97189
29. Weidner N, Semple JP, Welch WR, Folkman J. Tumor Angiogenesis and Metastasis–Correlation in Invasive Breast Carcinoma. *N Engl J Med* (1991) 324(1):1–8. doi: 10.1056/NEJM199101033240101
30. Ren X, Ji Y, Jiang X, Qi X. Downregulation of CYP2A6 and CYP2C8 in Tumor Tissues Is Linked to Worse Overall Survival and Recurrence-Free Survival From Hepatocellular Carcinoma. *BioMed Res Int* (2018) 2018:5859415. doi: 10.1155/2018/5859415
31. Zhang BH, Yang J, Jiang L, Lyu T, Kong LX, Tan YF, et al. Development and Validation of a 14-Gene Signature for Prognosis Prediction in Hepatocellular Carcinoma. *Genomics* (2020) 112(4):2763–71. doi: 10.1016/j.ygeno.2020.03.013
32. Huang XY, Huang ZL, Huang J, Xu B, Huang XY, Xu YH, et al. Exosomal circRNA-100338 Promotes Hepatocellular Carcinoma Metastasis *via* Enhancing Invasiveness and Angiogenesis. *J Exp Clin Cancer Res* (2020) 39 (1):20. doi: 10.1186/s13046-020-1529-9
33. Qiu Z, Li H, Zhang Z, Zhu Z, He S, Wang X, et al. A Pharmacogenomic Landscape in Human Liver Cancers. *Cancer Cell* (2019) 36(2):179–93.e111. doi: 10.1016/j.ccell.2019.07.001
34. Budde BS, Mizumoto S, Kogawa R, Becker C, Altmuller J, Thiele H, et al. Skeletal Dysplasia in a Consanguineous Clan From the Island of Nias/ Indonesia is Caused by a Novel Mutation in B3GAT3. *Hum Genet* (2015) 134(7):691–704. doi: 10.1007/s00439-015-1549-2
35. Zhang YL, Ding C, Sun L. High Expression B3GAT3 Is Related With Poor Prognosis of Liver Cancer. *Open Med (Wars)* (2019) 14:251–8. doi: 10.1515/med-2019-0020
36. Hassan MK, Kumar D, Naik M, Dixit M. The Expression Profile and Prognostic Significance of Eukaryotic Translation Elongation Factors in Different Cancers. *PloS One* (2018) 13(1):e0191377. doi: 10.1371/journal.pone.0191377
37. Berge M, Allanic D, Bonnin P, de Montrion C, Richard J, Suc M, et al. Neuropilin-1 is Upregulated in Hepatocellular Carcinoma and Contributes to Tumour Growth and Vascular Remodelling. *J Hepatol* (2011) 55(4):866–75. doi: 10.1016/j.jhep.2011.01.033
38. Yoshida N, Tsuzuki S, Karube K, Takahara T, Suguro M, Miyoshi H, et al. STX11 Functions as a Novel Tumor Suppressor Gene in Peripheral T-Cell Lymphomas. *Cancer Sci* (2015) 106(10):1455–62. doi: 10.1111/cas.12742
39. Zhou X, Wang S, Wang Z, Feng X, Liu P, Lv XB, et al. Estrogen Regulates Hippo Signaling *via* GPER in Breast Cancer. *J Clin Invest* (2015) 125(5):2123–35. doi: 10.1172/JCI79573
40. Lin D, Fu Z, Yang G, Gao D, Wang T, Liu Z, et al. Exportin-5 SUMOylation Promotes Hepatocellular Carcinoma Progression. *Exp Cell Res* (2020) 395 (2):112219. doi: 10.1016/j.yexcr.2020.112219
41. Fan Y, Wang L, Han XC, Ma HY, Zhang N, Zhe L. LncRNA MIF-AS1 Aggravates the Progression of Ovarian Cancer by Sponging miRNA-31-5p. *Eur Rev Med Pharmacol Sci* (2020) 24(5):2248–55. doi: 10.26355/eurrev_202003_20490
42. Kudo A, Mogushi K, Takayama T, Matsumura S, Ban D, Irie T, et al. Mitochondrial Metabolism in the Noncancerous Liver Determine the Occurrence of Hepatocellular Carcinoma: A Prospective Study. *J Gastroenterol* (2014) 49(3):502–10. doi: 10.1007/s00535-013-0791-4
43. Yasui Y, Kudo A, Kurosaki M, Matsuda S, Muraoka M, Tamaki N, et al. Reduced Organic Anion Transporter Expression is a Risk Factor for Hepatocellular Carcinoma in Chronic Hepatitis C Patients: A Propensity Score Matching Study. *Oncology* (2014) 86(1):53–62. doi: 10.1159/000356643
44. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular Carcinoma. *Nat Rev Dis Primers* (2021) 7(1):6. doi: 10.1038/s41572-020-00240-3
45. Shinkawa H, Tanaka S, Kabata D, Takemura S, Amano R, Kimura K, et al. The Prognostic Impact of Tumor Differentiation on Recurrence and Survival After Resection of Hepatocellular Carcinoma Is Dependent on Tumor Size. *Liver Cancer* (2021) 10(5):461–72. doi: 10.1159/000517992
46. Song J, Zhou H, Gu D, Xu Y. Hepatocellular Carcinoma Differentiation: Research Progress in Mechanism and Treatment. *Front Oncol* (2021) 11:790358. doi: 10.3389/fonc.2021.790358
47. Nakano K, Saito K, Mine S, Matsushita S, Tanaka Y. Engagement of CD44 Up-Regulates Fas Ligand Expression on T Cells Leading to Activation-Induced Cell Death. *Apoptosis* (2007) 12(1):45–54. doi: 10.1007/s10495-006-0488-8
48. Kirkin V, Cahuzac N, Guardiola-Serrano F, Huault S, Luckerath K, Friedmann E, et al. The Fas Ligand Intracellular Domain is Released by ADAM10 and SPPL2a Cleavage in T-Cells. *Cell Death Differ* (2007) 14 (9):1678–87. doi: 10.1038/sj.cdd.4402175
49. Meier P, Dayer E, Ronco P, Blanc E. Dysregulation of IL-2/IL-2R System Alters Proliferation of Early Activated CD4+ T Cell Subset in Patients With End-Stage Renal Failure. *Clin Nephrol* (2005) 63(1):8–21. doi: 10.5414/CNP63008
50. Nomura J, Hosoi T, Okuma Y, Nomura Y. The Presence and Functions of Muscarinic Receptors in Human T Cells: The Involvement in IL-2 and IL-2 Receptor System. *Life Sci* (2003) 72(18-19):2121–6. doi: 10.1016/S0024-3205(03)00071-7
51. Hu ZQ, Xin HY, Luo CB, Li J, Zhou ZJ, Zou JX, et al. Associations Among the Mutational Landscape, Immune Microenvironment, and Prognosis in Chinese Patients With Hepatocellular Carcinoma. *Cancer Immunol Immunother* (2021) 70(2):377–89. doi: 10.1007/s00262-020-02685-7
52. Schoenberg MB, Li X, Li X, Han Y, Hao J, Miksch RC, et al. The Predictive Value of Tumor Infiltrating Leukocytes in Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis. *Eur J Surg Oncol* (2021) 47 (10):2561–70. doi: 10.1016/j.ejso.2021.04.042
53. Raskov H, Orhan A, Christensen JP, Gogenur I. Cytotoxic CD8(+) T Cells in Cancer and Cancer Immunotherapy. *Br J Cancer* (2021) 124(2):359–67. doi: 10.1038/s41416-020-01048-4
54. Drijvers JM, Sharpe AH, Haigis MC. The Effects of Age and Systemic Metabolism on Anti-Tumor T Cell Responses. *Elife* (2020) 9:1–29. doi: 10.7554/eLife.62420
55. Lanitis E, Dangaj D, Irving M, Coukos G. Mechanisms Regulating T-Cell Infiltration and Activity in Solid Tumors. *Ann Oncol* (2017) 28(suppl_12): xii18–32. doi: 10.1093/annonc/mdx238


**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 YY declared a shared parent affiliation with the author YHX 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.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 16 June 2022
doi: 10.3389/fonc.2022.932786

# Exploring the Potential of Pyroptosis-Related Genes in Predicting Prognosis and Immunological Characteristics of Pancreatic Cancer From the Perspective of Genome and Transcriptome

*Jing Zhang*[1†], *Xiaomin You*[1†], *Dong Kang*[2] *and Guoxiong Zhou*[1*]

[1] Department of Gastroenterology, Affiliated Hospital of Nantong University, Nantong University, Nantong, China,
[2] Department of General Surgery, Rugao Hospital of Traditional Chinese Medicine, Rugao, China

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Fangshi Xu,*
*The Second Affiliated Hospital of Xi'an Jiaotong University, China*
*Yisha Zhao,*
*Zhejiang University, China*
*Xiaoqing Li,*
*Chongqing Hospital of Traditional Chinese Medicine, China*

***Correspondence:**
*Guoxiong Zhou*
*zhougx@ntu.edu.cn*

[†]*These authors have contributed equally to this work*

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

**Received:** *30 April 2022*
**Accepted:** *16 May 2022*
**Published:** *16 June 2022*

**Citation:**
*Zhang J, You X, Kang D and Zhou G (2022) Exploring the Potential of Pyroptosis-Related Genes in Predicting Prognosis and Immunological Characteristics of Pancreatic Cancer From the Perspective of Genome and Transcriptome. Front. Oncol. 12:932786. doi: 10.3389/fonc.2022.932786*

**Objective:** To probe into the role of pyroptosis-related genes in pancreatic carcinoma.

**Methods:** Herein, we conducted a comprehensive bioinformatics analysis to evaluate tumor-immune infiltration and tumor mutation burden, the correlations between PRGs, and microsatellite instability and found that 33 PRGS were up- or down-regulated in PC. Then we built the PPI network, which was downloaded from the STRING database. Using TCGA cohort median risk score, PC subjects from the Gene Expression Composite cohort (GEO) data resource were stratified into two risk categories, with the low-PC risk group harboring a higher overall survival (OS) (P = 0.011). We employed the ssGSEA approach to quantify immune cell abundance in separate risk groups separated by risk signature while assessing variations in immune cell invasion. Chemotherapeutic drugs were retrieved from the Genomics of Drug Sensitivity in Cancer (GDSC) data resource.

**Results:** Eight prognostic PRG models (CASP4, GSDMC, IL-18, NLRP1, NLRP2, PLCG1, TIRAP, and TNF) were established *via* LASSO Cox regression to estimate the OS of PC subjects with medium-to-high accuracy.

**Conclusion:** Our study is the first to identify a pyroptotic-related prognostic gene feature for PC, providing more options for the prognostic prediction of PC.

**Keywords: pyroptosis, pancreatic cancer, prognostic prediction, risk stratification, immune microenvironment**

## INTRODUCTION

Pancreatic cancer (PC) is a highly malignant tumor of the digestive tract with increased incidence in the recent years (1) along with a poor overall prognosis (2). PC is known as the "king of cancer" not only for its aggressiveness and rapid progression but also for its frequent discovery at an advanced stage of disease for the first time. At this time, radical surgery is uncommon and is not responsive to radiotherapy or chemotherapy. The overall 5-yearsurvival rate was only 8% (3, 4). However, when

many patients are first diagnosed with pancreatic cancer, it has already advanced, the untreated median survival time is about 6 months, and some patients after positive surgery, or chemotherapy, immunotherapy, and other treatments of advanced pancreatic cancer survival, are still not optimistic.

Although breakthroughs in genetic characterization have increased our knowledge of disease features and heterogeneity, the lack of real progress in PC during the last few decades has caused substantial concern (5). Given the constraints of PC therapy, novel therapeutic targets are necessary to enhance PC clinical outcomes. There are many related studies on pancreatic cancer worldwide, such as the recently identified super-enhancers such as JQ1, IBET, and SLC1A5 which can promote the EMT progression in pancreatic cancer (6–8). Current first-line therapies for advanced pancreatic cancer including gemcitabine combined with albumin-binding paclitaxel, FOLFIRINOX (fluorouracil, leucovorin, irinotecan and oxaliplatin), and modified FOLFIRINOX (9), or PARP inhibitors, such as olaparib, have been approved for germline patients with germline BRCA1 or BRCA2 mutations (10) but have not significantly improved survival. In the surgical specimens of pancreatic cancer, the pancreatic cancer tumor can be found as a solid tumor and containing a large amount of extracellular matrix in the tumor tissue, which significantly reduces the drug penetration into the tumor. It has been shown that hyperbaric oxygen significantly consumes the main components of the extracellular matrix, thus promoting the infiltration of cytotoxic T cells into the tumor parenchyma to facilitate drug penetration (11). At the same time, a prospective study mentioned that preoperative HBO pretreatment of pancreatic surgery can effectively reduce postoperative complications in patients (12). Therefore, HBO and gemcitabine may improve drug resistance and open a new chapter in the treatment of pancreatic cancer, a fatal malignant tumor, which still needs to be further explored. However, very few people can be effectively used in clinical practice. As a result, improved prognostic approaches that are both trustworthy and accurate are urgently needed to make targeted therapy more practical. The pyroptosis-related genes we studied may also provide new targets for the treatment of pancreatic cancer (13).

Pyroptosis, also known as cellular inflammatory necrosis, is a new kind of programmed cell death (14, 15). Pyroptotic cells have cellular swelling as well as many bubble-like protrusions. Under an electron microscope, pyroptotic cells may be seen forming a large number of vesicles. Following the development of these vesicles, openings form on the cell membrane, which ruptures and allows the contents to escape (16). The gasdermin family, which comprises gasdermin-A through to gasdermin-E and pejvakin, is the major executioner of pyroptosis (17). Shearing and multimerization of the gasdermin family proteins result in the cleavage of the N-terminal along with the C-terminal junctional structural domains as well as the release of activated N-terminal domains that dock to membrane lipids, cardiolipin, and phosphatidylinositol and localize into cell membrane pores (18). Cellular gasdermin family proteins have pores in the cell membrane that range from 10 to 20 nm in size, and cell contents are slowly discharged *via* membrane pores, generating increased inflammatory reactions (19, 20). Cells flatten progressively, forming 1–5-m apoptotic vesicle-like protrusions (scorched vesicles), and cells inflate gradually until the plasma membrane ruptures, exhibiting characteristics such as nuclear condensation and chromatin DNA disintegration (21, 22). Pyroptosis was first recognized as a remarkable process for fighting infection, and multiple investigations show that it is also important in tumorigenesis (23). Inflammatory vesicles, gasdermin proteins, and proinflammatory cytokines are linked with carcinogenesis, infiltration, and metastasis as pivotal elements of pyroptosis (24).

Until now, there have been many studies on prognostic models, such as recent ones identifying novel signatures for the prognosis of multiple cancers (19, 25–28). GSDMD and GSDME play significant roles in defense against intracellular pathogens and tumors. A recent study showed that GSDMD-mediated pyroptosis is closely associated with the prognosis of hepatocellular carcinoma, rectal adenocarcinoma, and cutaneous melanoma. The authors of this study highlighted the role in the antitumor immunity of GSDMD and argued for developing future drugs for activating GSDMD (29). L61H10, a thiopyran derivative of curcumin, can induce pyroptosis through the caspase-3/GSDME pathway. It has a good antitumor activity against lung cancer (30). L61H10 can promote the expression of antiapoptotic genes by regulating NF-κB in inducing the transition from apoptosis to pyroptosis. However, curcumin bioavailability *in vivo* is low, and its use for lung cancer treatment needs further research (31). Increasing the induction of tumor-related pyroptosis can reduce the tumor volume. If this move can be further studied and applied to clinical practice, it may be a crucial step in the process of human play against tumors. According to Wang et al., the bio-orthogonal system that is based on Phe-BF3 desilylation is a potent tool for chemical biology, implying that pyroptosis-triggered inflammation generates robust antitumor immunity and may synergize with checkpoint blockade (32). Therefore, the study of pyroptosis-related genes in pancreatic cancer is important for the treatment of pancreatic cancer.

It was found that application of this bio-orthogonal system to gas proteins revealed that less than 15% of tumor cells with pyroptosis were capable of clearing the entire 4T1 breast tumor graft (33). Various danger-linked signaling molecules along with cytokines are stimulated and produced when pyroptosis occurs, coupled with a severe inflammatory response as well as immune system activation (34). Some investigations have documented that pyroptosis' powerful proinflammatory activity may influence tumor immune microenvironment modulation (35).

A considerable reduction in the number as well as activity of CD8+ T cells was linked with aberrant GSDMD expression (36). Additionally, pyroptosis plays an indispensable role in NK cell antitumor activity (37). We confirm that pyroptosis is important in tumorigenesis and antitumor processes based on our present results; nevertheless, its particular activities in PC have received less attention. As a result, we conducted a comprehensive investigation to compare the expression levels of PRGs in

healthy and pancreatic adenocarcinoma tissues, investigate their prognostic worthiness, and assess the relationships of pyroptosis with the tumor immune microenvironment. We established that pyroptosis is strongly linked to tumorigenesis and that pyroptosis may cross talk with metabolites in the TME as well as particular targets on the cell membrane to execute an antitumor impact (38).

# MATERIALS AND METHODS

## Dataset

The Cancer Genome Atlas (TCGA) data resource was adopted to abstract PC-clinical information, PC-RNA sequencing patterns (n = 178), and non-malignant pancreas epithelium RNA sequencing patterns (n = 4) (39). We excluded PC subjects who had no survival time, retaining 177 subjects for further research. Besides, as an external verification data set, we abstracted GSE62452 along with the GSE71729 data sets from the GEO data resource, which included 69 subjects (four cases were excluded) and 125 subjects with primary PC, respectively. Meanwhile, genes were uncovered using the GENCODE data resource (40), GPL6244, and GPL20769 annotation documents. Additionally, 33 PRGs were abstracted on the basis of previous research (41). TCGA data resource was utilized to generate mutation along with CNV data. Furthermore, the maftools program was adopted to generate the mutation frequencies of 33 PRGs in PC subjects (42).

## Construction of the PPI Network

Thirty-three PRGs were loaded into a STRING data resource (confidence = 0.900) to generate PPI networks in order to identify cross-talking genes.

## Sample Classification for Aging Patterns

We adopted the "ConsensusClusterPlus" R tool to conduct two classifications on the basis of the 33 PRGs. The optimal k value was established across all PC subjects by estimating the inflection point of the sum of squared error (SSE) according to the prognostic ARG expression. The decline rate reduced after k = i; thus, k = i was selected. Moreover, in pyroptosis profiles, we executed Kaplan–Meier survival assessment on cluster-1 and -2 groups.

## Construction of Risk Signature and Nomogram

Differential PRGs in malignant and non-malignant tissues were assessed *via* the "limma" package (|logFC| >2 along with $P < 0.05$). Univariate Cox regression was then adopted to filter prognostic PRGs ($P < 0.05$). These prognostic genes were then utilized to determine genes involved in signature creation *via* multivariate Cox coupled with LASSO regression models. To generate the model and control the complexity of LASSO regression, we employed the suitable λ. The following formula was used to determine the risk score: OS score of $\text{OS score of Risk} = \sum_{i=1}^{n} \text{Coef}_i * x_i$ (43). Besides, for verification, we implemented Kaplan–Meier survival assessment, ROC analysis, and a calibration curve. For clinical utility, the comparative value ($2^{-\Delta ct}$) was calculated from qRT-PCR results and used for score calculation, and the score was further standardized and simplified to generate a riskscore (44). The riskscore was calculated as follows: Riskscore (qRT-PCR) = (score-min)/max.

## Functional Enrichment Analysis

The "ggplot2" and "clusterProfiler" R tools were adopted to execute enrichment analysis in 33 PRGs. The data of the "clusterProfiler" tool were utilized to conduct gene ontology (GO) assessment along with the Kyoto Encyclopedia of Genes and Genomes (KEGG) assessment. Moreover, GSEA enrichment analysis was done in numerous risk groups differentiated by risk signature (42).

## Comprehensive Immune, TMB, and MSI Analyses

We adopted the ssGSEA approach to quantify the abundance of immune cells in distinct risk groups separated by risk signature while investigating disparities in immune cell invasion. We also adopted the Timer algorithm to assess the immune cell correlation. More critically, we assessed the levels of TMB along with MSI expressions in each risk signature gene.

## Drug Sensitivity Analysis

Chemotherapeutic drugs were abstracted from the Genomics of Drug Sensitivity in Cancer (GDSC) data resource, and IC50 was computed *via* the "pRRophetic" R tool (42, 43, 45).

## Cell Culture and Real-Time PCR

A total of 14 tumor tissue samples and 10 normal tissue samples were obtained from PC patients who underwent tumor resection. All tissue samples were collected from the Affiliated Hospital of Nantong University with the approval by the Medical Ethics Committee. The National Infrastructure of Cell Line Resource provided a human normal pancreatic duct epithelial cell line (HPDE6-C7) and human pancreatic cancer cells (PANC-1, BXPC-3, CFPAC-1). They were grown in DMEM or RPMI-1640 enriched with 10% FBS, along with 1% penicillin/streptomycin, and incubated at 37°C in a humidified incubator containing 5% $CO_2$. Total RNA was extracted from Hp, PANC, BX, and CF cells with TRIzol (Invitrogen). The complement DNA (cDNA) was prepared with a RevertAid First Strand cDNA Synthesis Kit (TaKaRa Bio, Kusatsu, Japan). All quantitative real-time PCRs were carried out in triplicate with SYBR Master Mix (TaKaRa Bio) on a LightCycler 480 II instrument (Roche, Basel, Switzerland). The primers were used as follows:

RT-IL-18-F:TCTTCATTGACCAAGGAAATCGG,
RT-IL-18-R:TCCGGGGTGCATTATCTCTAC;
RT-TNF-F:ACCCTCACACTCACAAACCA,
RT-TNF-R:ATAGCAAATCGGCTGACGGT;
RT-NLRP1-F:CCCCATCCCTCTGAGCTAC,
RT-NLRP1-R:ACTTAACAGGCCCAATAGGAA;
RT-CASP4-F:TGGCAGAAGGCAACCACAGAA,
RT-CASP4-R:TTTGTTCCACCAAGTTATCC.

# RESULTS

## Differential Expression and Mutation Landscape of PRGs in PC Patients

We explored the mutation landscape of 33 PRGs that could be annotated in TCGA-PAAD cohort. First, we revealed the copy number variation (CNV) of 33 PRGs in PC patients (**Figure 1A**). All PRGs had copy number amplification or deletion, among which GSDMA showed the highest amplification frequency and CASP3 revealed the highest loss frequency. Mutations were found in 7 (4.12%) of 170 PC samples, with most genes mutating at about 1% (**Figure 1B**). In the assessment of the incidence of somatic mutations in 33 PRGs, missense mutation was the most common variation classification. Single-nucleotide polymorphism (SNP) was the most common variation type, and C > T was dominant in the single-nucleotide variant (SNV) (**Figure 1C**). In addition, we identified the corresponding position of 33 PRGs on the chromosomes (**Figure 1D**). Meanwhile, we compared the expression levels of PRGs in TCGA-PAAD cohort, and the heatmap revealed six differentially expressed PRGs, including PLCG1, PRKACA, TNF, NOD2, NLRC4, and NLRP3 (**Figure 1E**). Interestingly, the boxplot demonstrated that the above six genes were downregulated in tumor tissue (**Figure 1F**). To further explore the cross talk between PRGs, we conducted PPI analysis, and the minimum confidence required for analysis was set at 0.9 (**Figure 1G**). Finally, we showed a correlation network of 33 PRGs (red: positive correlation; blue: negative correlation), and the results revealed more red line segments than blue line segments (**Figure 1H**).

## Pyroptosis Patterns in PC Patients

To explore the association between PRG expression and PC occurrence, we performed consensus clustering in PC patients and found that when k = 2, the intra-group correlation was the highest, while the inter-group correlation was low (**Figure 2A**). It indicated that PC patients could be well divided into two groups according to 33 PRGs. According to the above algorithm, all patients were classified into subgroups C1 and C2. In addition, the overall survival (OS) between the two subgroups was also compared; excitingly, the survival time of the C2 subgroup was remarkably shorter than that of the C1 subgroup (P< 0.001), as displayed in **Figure 2B**. Heatmap results demonstrated remarkable differences in age and survival status between the subtypes (P< 0.05), as depicted in **Figure 2C**. To clarify other biological functions of PRGs, except for regulating pyroptosis, we conducted GO and KEGG enrichment analyses. GO enrichment analysis revealed that PRGs were mainly involved in regulating cytokine production, interleukin-1 function, and other biological processes (**Table S1**). In addition, KEGG analysis demonstrated that PRGs were mainly involved in the NOD-like receptor signaling pathway, *Salmonella* infection, and other cascades (**Table S2**).

## A PRG-Based Risk Stratification System Was Developed in PC Patients

Transcriptome data and survival information were matched, and TCGA-PAAD cohort eventually included 177 patients for follow-up analysis. Univariate Cox regression assessment was adopted for preliminary screening of prognostic genes, and 11 genes were identified (**Figure 2D**). Risk stratification systems containing 8-PRGs were constructed based on the optimal λ values by LASSO (**Figures 2E, F**) and multivariate Cox regression assessment (**Figure 2G**). The risk score was computed as follows: risk score = (0.6617*CASP4-exp.) + (0.2903*GSDMC-exp.) + (0.2719*IL18-exp.) + (0.3557*NLRP1-exp.) + (0.0659*NLRP2-exp.) + (-0.5057*PLCG1 exp.) + (-0.2744*TIRAP-exp.) + (-0.3883*TNF exp.). According to the median score computed by the formula, 177 PC patients were divided into low- and high-risk subgroups. PCA revealed that different risk patients were well divided into two groups (**Figure 3A**). In contrast with the low-PC risk group, subjects in the high-PC risk group had shorter survival time and remarkable differences in OS time (**Figure 3B**). Time-dependent receiver operating characteristic (ROC) assessment was adopted to assess the sensitivity along with the specificity of the risk stratification system. We found that the areas under the ROC curve (AUC) were 0.740, 0.707, and 0.741 at 1, 3, and 5 years, respectively (**Figure 3C**). In addition, the risk profile also showed the same results (**Figures 3D, E**). Kaplan–Meier analysis was performed on GEO and TCGA cohorts of each PRG participating in the risk stratification system (**Figure S1**).

## External Validation

An external validation cohort of 190 PC patients was abstracted from the GEO database (GSE62452 and GSE71729). The data of the two datasets were standardized by the "sva" package before further analysis. On the basis of the median risk score in TCGA cohort, 190 patients in the GEO cohort were stratified into the low-risk group (94 patients) and high-risk group (96 patients), as presented in **Figures 3F, G**. Kaplan–Meier analysis demonstrated that the survival difference between the two subgroups was statistically significant (P = 0.011, **Figure 3H**). ROC curve analysis of the GEO cohort indicated that our risk stratification system had a good prediction effect (AUC values of 1, 3, and 5 years were 0.518, 0.560, and 0.562, respectively), as depicted in **Figure 3I**.

## Construction Nomogram Based on Independent Prognostic Factors

Univariate and multivariate Cox regressions were adopted to assess whether the risk score was an independent prognostic factor in PC subjects. Univariate Cox regression data revealed that risk score was a risk factor in predicting poor survival in TCGA (**Figure 4A**) and GEO cohorts (**Figure 4C**) (HR =1.537, HR: 1.261). Multivariate analysis also illustrated that risk score was an independent prognostic factor (HR = 1.517, HR: 1.353) (**Figures 4B, D**). We combined the clinical characteristics commonly employed in clinical work to establish a nomogram to estimate survival probability (**Figure 4E**). Calibration curves illustrated that the prediction of overall survival was relatively linked to the standard curve in the two cohorts (**Figures 4F, G**).

## Comparison of Immunity Between Subgroups

We use the ssGSEA algorithm to further compare the enrichment score of 16 immune cells and 13 immune-linked cascades in low- and high-risk subgroups. Interestingly, there was no correlation

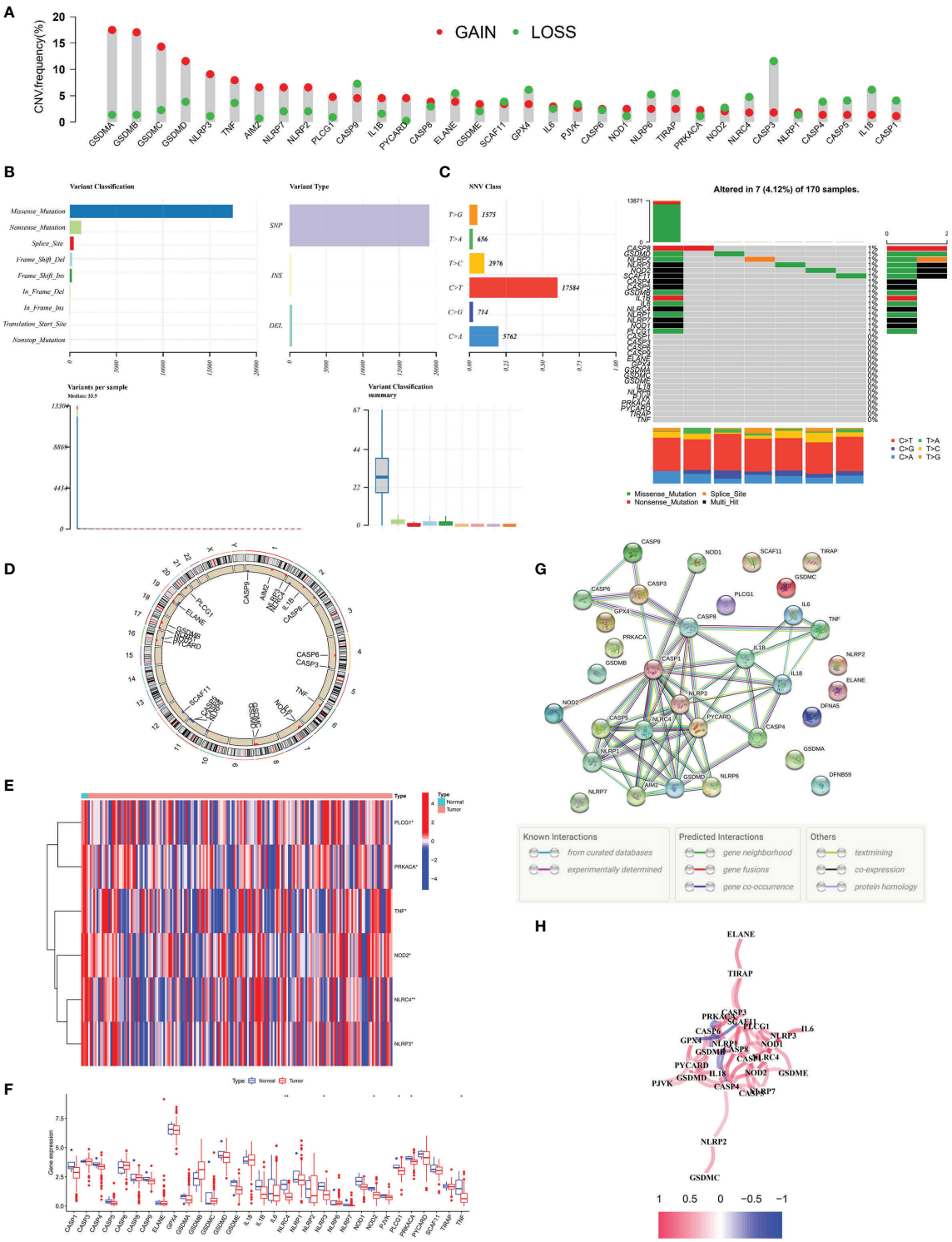



**FIGURE 1** | Landscape of 33 PRGs in the PC cohort. **(A)** The CNV alteration of 33 PRGs in the PC cohort. The height of the column represents the alteration frequency. **(B, C)** The mutation frequency along with the classification of 33 PRGs in PC. PRG: pyroptosis-related gene, PC: pancreatic cancer, SNP: single-nucleotide polymorphism, INS: insertion, and DEL: deletion. **(D)** The position of 33 PRGs on the chromosome. **(E)** Heatmap (blue: low expression level; red: high expression level) of PRGs between normal (N, brilliant blue) and tumor tissues (T, red). **(F)** Comparison of the gene expression of 33 types of PRGs between normal and tumor tissues in TCGA cohort. **(G)** PPI network illustrating the cross talks of PRGS (cross talk score = 0.9). **(H)** The association network of PRGs (red line: positive association; blue line: negative relationship). The depth of colors reflects the strength of relevance). *P < 0.05; **P < 0.01.

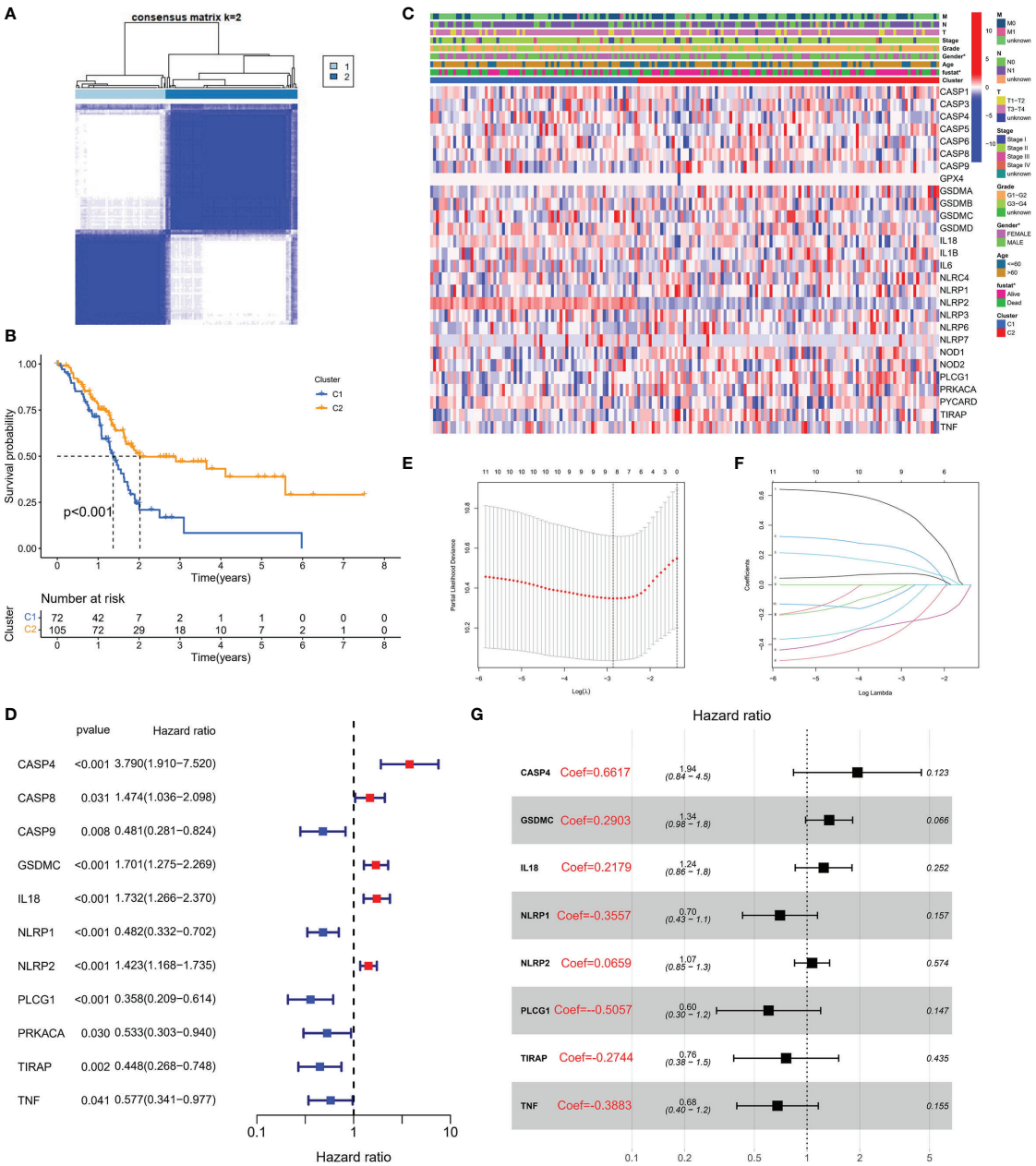



**FIGURE 2** | Molecular subtypes and risk status based on PRG expression. **(A)** According to the consensus clustering matrix (k = 2), 190 PC subjects were stratified into two groups. **(B)** Kaplan–Meier curves for the two clusters. **(C)** The heatmap along with the clinicopathologic characteristics of the two clusters established using these DEGs and clinical variables. (T refers to tumor, N refers to lymph node, and M refers to distant metastasis). T1 means that the tumor is less than 2 cm, T2 is the tumor in 2–4 cm, T3 is the tumor larger than 4 cm, and T4 is the tumor regardless of size but invades the celiac artery, superior mesenteric artery, or common hepatic artery. N refers to lymph nodes, N1 refers to regional metastases in one to three lymph nodes, N2 refers to regional metastases in four or more lymph nodes, and M1 refers to distant metastasis of the tumor. On this basis, the combination of the three TNM indicators was used to draw a specific stage (stage). G1, G2, and G3 are the degree of tumor differentiation (G1: highly differentiated; G2: moderately differentiated; G3: poorly differentiated). **(D)** Univariate Cox regression analysis of PRGs in TCGA cohort. **(E)** Cross-verification was adopted to fine-tune the selection of parameters in LASSO regression. **(F)** LASSO regression of 11 genes linked to PC. **(G)** Multivariate Cox regression analysis of PRGs in TCGA cohort.

about immune cells and immune-related cascades in each group in TCGA cohort (**Figures 5A, D**). However, mast and NK cells were lower in high-risk patients than in low-risk patients in the GSE62452 cohort (**Figure 5B**). TCGA and GSE62452 cohorts also showed no statistical differences in immune function (**Figure 5E**). In the GSE71792 cohort, there were five types of immune cells and six types of immune function that differed remarkably between groups (**Figures 5C, F**). In addition, we elucidated the

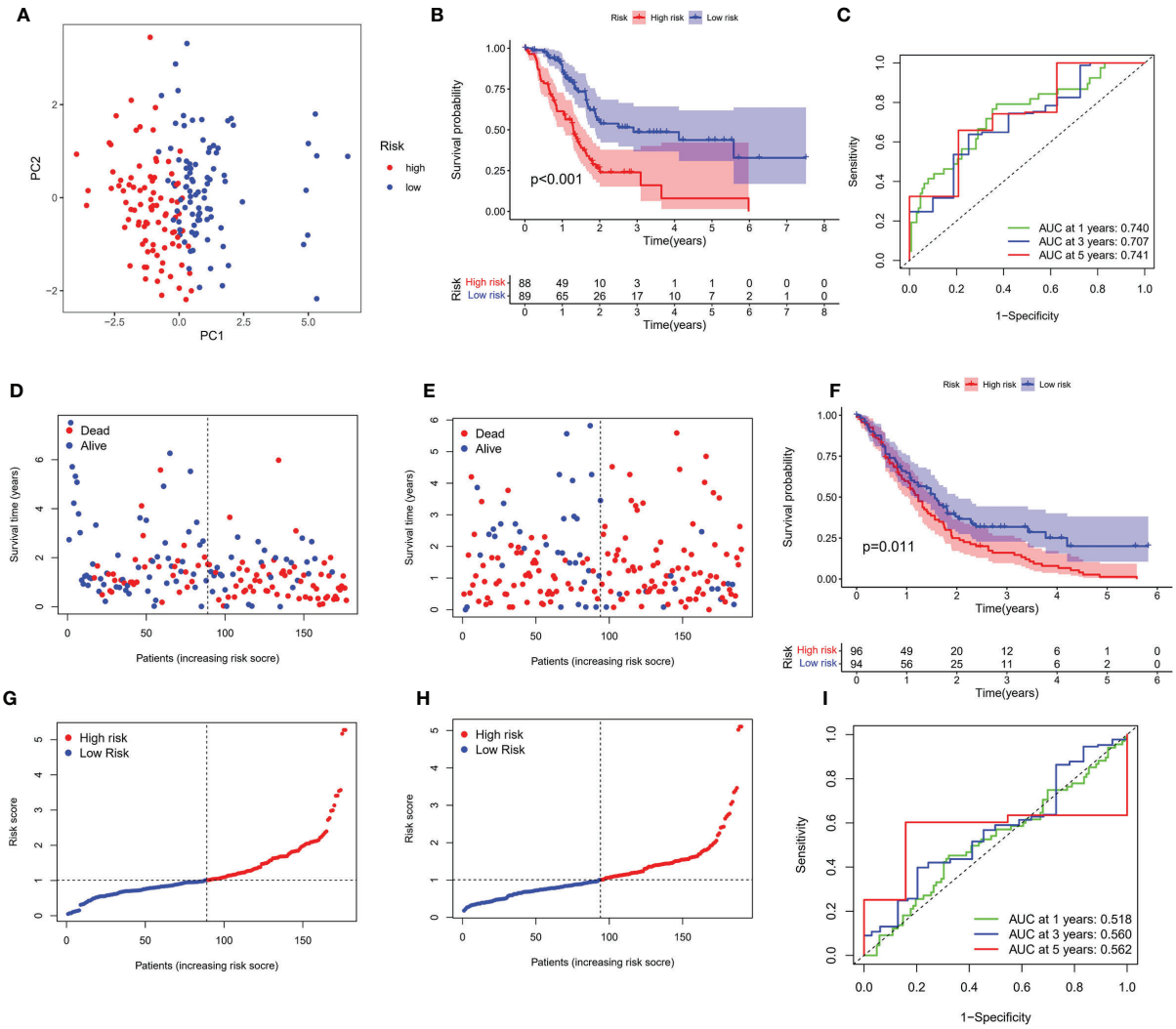



**FIGURE 3** | Verification of risk score. **(A)** PCA map for PC in TCGA cohort. **(B)** Kaplan–Meier curves for comparison of PC risks between low-PC and high-PC risk groups. **(C)** ROC curves illustrating the risk score prediction efficiency. **(D)** Each patient's survival rate is shown in TCGA cohort (low-PC risk cluster: on the left side of the dotted line; high-PC risk class: on the right side of the dotted line). **(E)** The survival rate for each patient in TCGA cohort (low-PC risk group: on the left side of the dotted line; high-PC risk class: on the right side of the dotted line). Patient distribution on the basis of the risk score. **(F)** Each patient's survival rate is shown in GEO cohort. **(G)** The survival rate for each patient in TCGA cohort. **(H)** Kaplan–Meier curves for comparison of PC risks between low-PC and high-PC risk groups. **(I)** ROC curves illustrating the risk score prediction efficiency.

immune infiltration of prognostic PRGs in the TIMER database (**Figure S2**).

## TMB, MSI, and Drug Sensitivity Analyses

TMB can be used as a biomarker to predict the efficacy of immunotherapy for PC, and MSI is also considered a marker for cancer immunotherapy. To determine whether 8-PRGs involved in the risk signature can also be used as biomarkers for drug selection, we analyzed the correlation between 8-PRGs and drug, MSI, and TMB. The results indicated that MSI was negatively linked with IL18 and TNF (**Figure S3**). TMB was negatively linked with NLRP1 and positively with IL18 (**Figure S4**). Herein, drug sensitivity assessment demonstrated that IL18 expression was negatively linked to some or most drugs in the tumor therapeutic response Portal database (**Table S3**).

## Validation of PRGs in Clinical Samples and Cell Lines

For further screening, we calculated the MCC score in the PPI network and identified NLRP1, IL18, TNF, and CASP4 as the hub genes in the protein level (**Figures 6A–C**). Meanwhile, the IHC imaging from the HPA database further confirmed the expression in different tissues (**Figures 6D–G**). Similar to the results in **Figure 1F**, the protein expression in different tissues was also significantly different. RT-qPCR assessment was performed in four cell lines, comprising three tumor cell lines

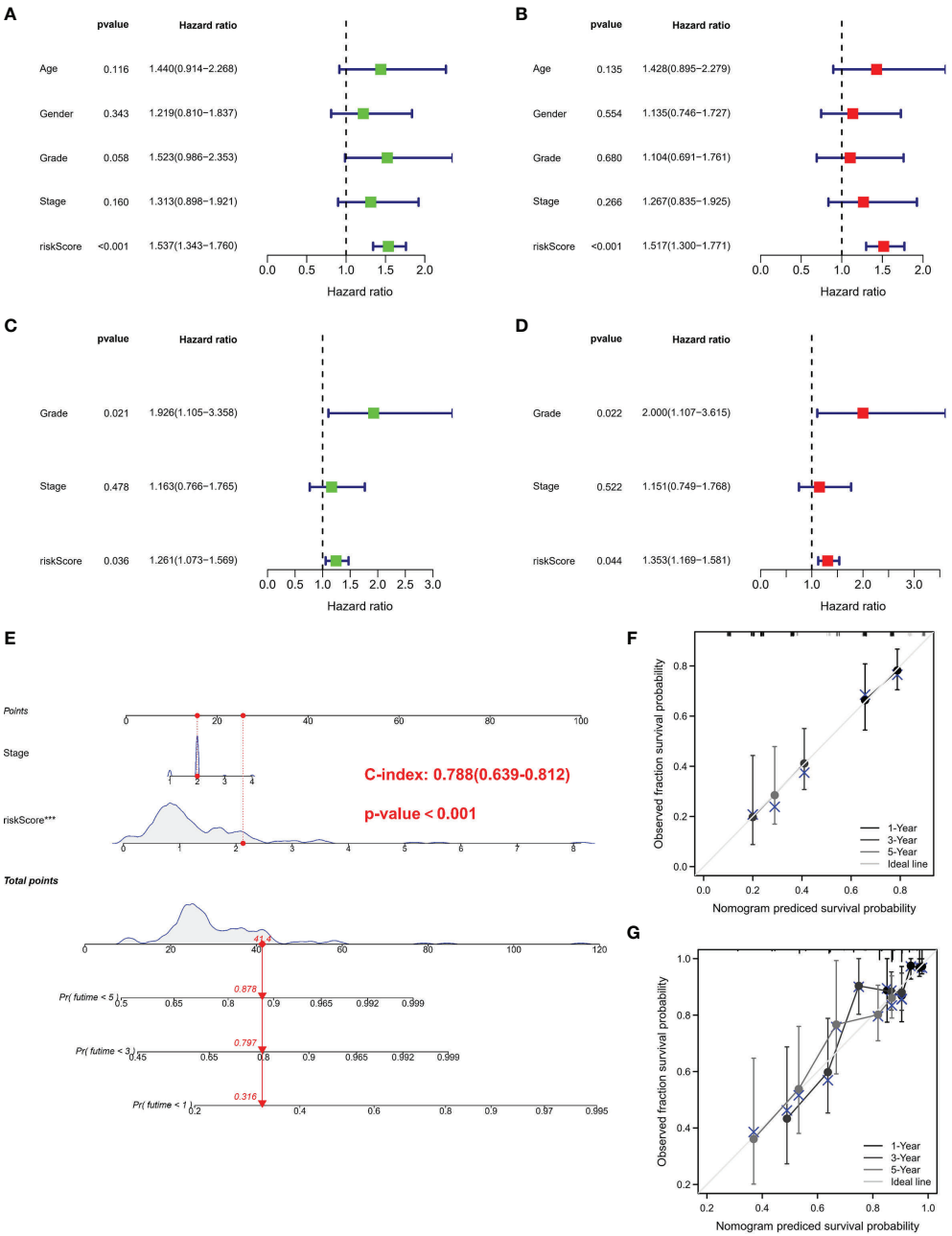



**FIGURE 4** | Independent prognostic validation and construction of nomogram. **(A)** Univariate assessment for TCGA cohort (grade: degree of tumor differentiation, G1 to G3; stage: the size of the primary tumor and the degree to which the cancer has spread in the patient's body, I to IV). **(B)** Multivariate analysis for TCGA cohort. **(C)** Univariate analysis for the GEO cohort. **(D)** Multivariate analysis for the GEO cohort. **(C, F)** Heatmap (green: low expression; red: high expression) for the connections between clinicopathologic features and risk groups. **(E)** Nomogram. **(F, G)** Nomogram to predict 1-, 3-, and 5-year overall survival rates of PC patients. Calibration curve for the overall survival nomogram model in the discovery group. A dashed diagonal line represents the ideal nomogram. PRG, pyroptosis-related gene; PC, pancreatic cancer.

along with a human immortalized normal ductal epithelial cell line (**Figure 6H**). Compared with the human immortalized normal ductal epithelial cell line, the expression of NLRP1 and TNF in tumor cell lines was remarkably higher, the expression of NLRP1 was the highest in BX cell lines, and the expression of TNF was the highest in PANC-1 cell lines. IL18 and CASP4 expressions were not observed in the Bx cell lines. In CF cell lines, the expression of these four genes was remarkably higher. For clinical samples, we used the GTEx database, TCGA database, and qRT-PCR assay to analyze the expression levels

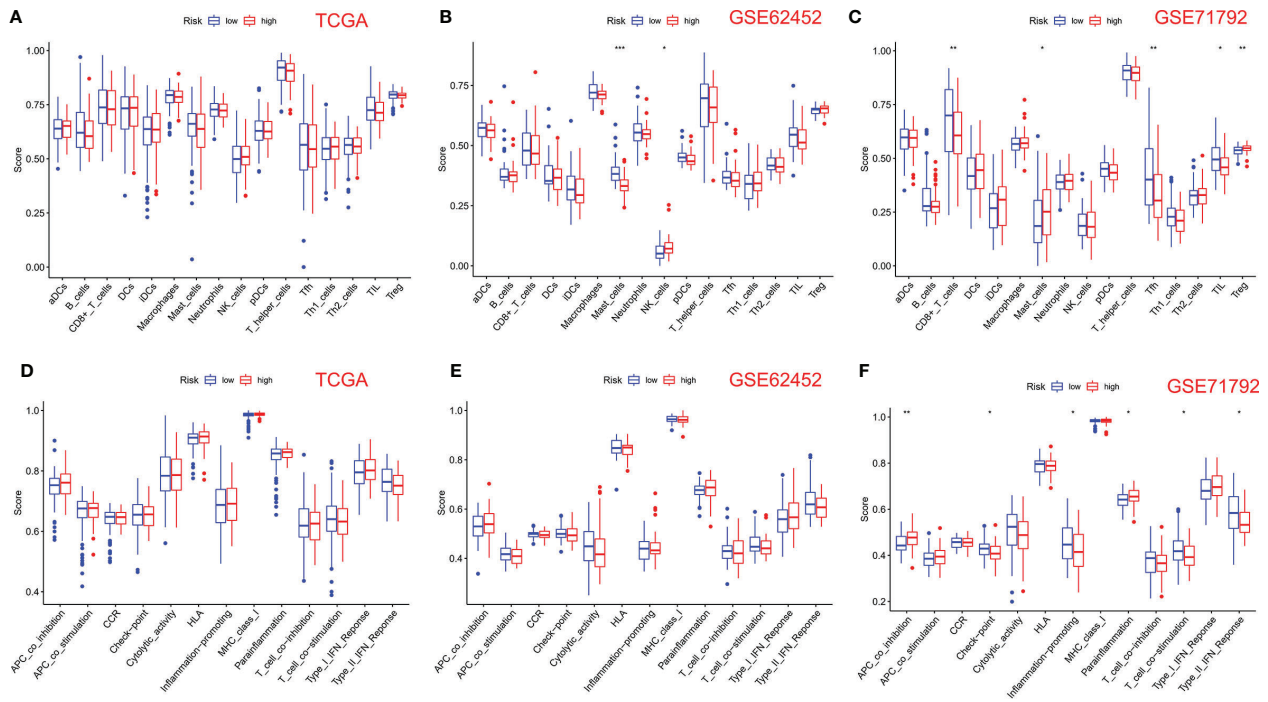



**FIGURE 5** | Comparison of ssGSEA scores for immune cells and cascades. **(A, D)** Comparison of enrichment scores of 16 types of immune cells and 13 immune-related cascades between low- (blue box) and high-risk (red box) groups in TCGA cohort. **(B, E)** Comparison of tumor immunity between low- (blue box) and high-risk (red box) groups in the GSE62452 cohort. **(C, F)** Comparison of tumor immunity between low- (blue box) and high-risk (red box) groups in the GSE71792 cohort. P-values were shown as follows: ns, not significant; *P < 0.05; **P < 0.01; ***P < 0.001.

of four PRGs. As shown in **Figures 7A–E**, tumor tissues showed obviously higher expression levels of NLRP1, IL18, TNF, and CASP4 than did the normal tissues. Importantly, survival analysis also showed that the high-risk group had a shorter survival time in our hospital cohort (P < 0.05) (**Figure 7F**). The experimental findings presented above were congruent with the predictions made by bioinformatics approaches.

## DISCUSSION

Pyroptosis is a recently identified kind of programmed cell death that has a dual role in cancer progress and therapeutic methods. Pyroptosis is known to be associated with tumor suppression. Activation of the pyroptosis pathway can effectively inhibit tumor progression, and some cancer cells escape tumor suppression also by escaping the pyroptosis pathway (46). Nevertheless, the function of PRG in PC remains unclear, and our work was designed to elucidate its role. We focused on defining the expression and prognostic worthiness of PRGs in PC. When comparing PC to normal tissues, we reported an improvement in IL18 expression and a reduction in TNF expression. A prognosis study found a low survival rate in PC patients with elevated NLRP1, NLRP2, IL18, and CASP4 expressions. These findings were consistent with previous reports.

As an inflammasome of NLRP3, IL-18 can accelerate atherosclerosis in mice (47). TNF is one of potential inducers of necroptosis in PDA (48). NLRP1 has been shown to predispose individuals to multiple self-healing palmoplantar carcinoma (MSPC) and familial keratosis lichenoides chronica (FKLC) (49). CASP4 and IL-18 are both involved in the pathogenesis of alcoholic hepatitis (AH) but has opposite effects in the pathogenesis. A cross-analysis has identified CASP4 as a commonly up-regulated gene known to be involved in the non-canonical inflammasome pathway. CASP4 deficiency reduces the severity of AH. Conversely, the deficiency of interleukin-18, the key antimicrobial cytokine, aggravates hepatic bacterial load, GSDMD activation, and AH (50). Additionally, we performed a functional enrichment analysis of PRGs and revealed that these 33 PRGs were mainly involved in the NOD-like receptor signaling pathway, TNF signaling pathway, regulating inflammatory response, pyroptosis, apoptosis, and Toll-like receptor signaling pathway. These findings suggested that these 33 PRGs may also be critical in PC oncogenesis and progression. Nowadays, there have also been many studies on the treatment of pancreatic cancer. A study of nutrient innervation in pancreatic cancer found that TRK-NGF inhibitors could interfere with the axon–nerve axis and reduce the nutrient supply of PDAC to reduce tumor recurrence (38, 51). The macrophage phenotypic switch-related signature could predict metastasis and survival in pancreatic cancer patients (52). In addition, CXC chemokine expression and their biological functions in pancreatic cancer may demonstrate good performance in PC patient prognosis and immunotherapeutic

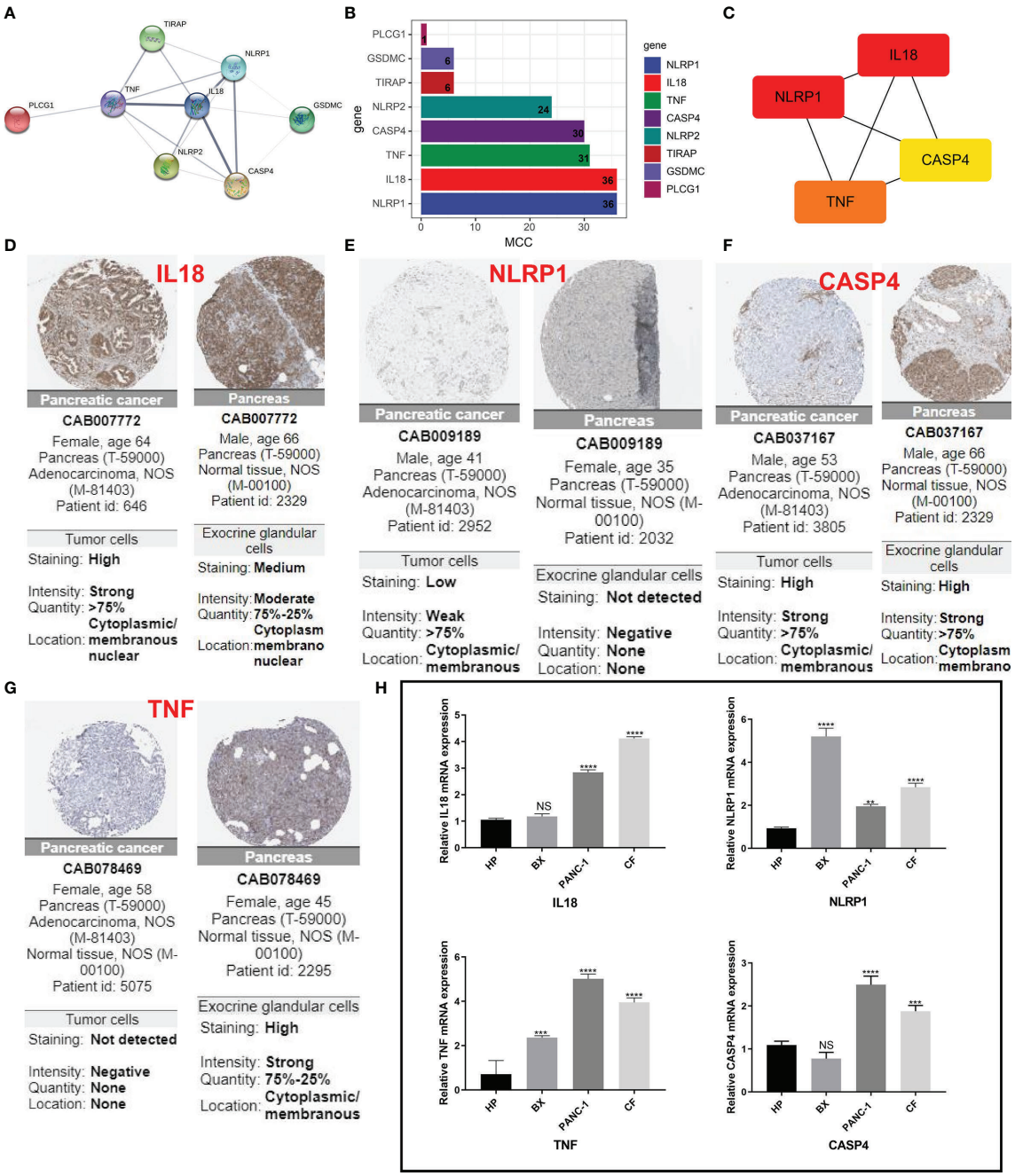



**FIGURE 6** | IHC and immunohistochemical verification. **(A)** PPI network showing the cross talks of eight pyroptosis-related genes. **(B, C)** The MCC algorithm is used to calculate the expression score of these eight genes in pancreatic cancer and links between the four genes with higher expression scores. **(D–G)** HPA data resource was also used to explore the expression of these four markers in pancreatic cancer tissue samples and normal tissues. **(H)** The mRNA levels from different cells as determined by real-time PCR analysis. **P < 0.01; ***P < 0.001; **** < 0.0001. NS means not significant.

target therapy prediction (53). A prognostic gene model based on eight prognostic PRGs (CASP4, GSDMC, IL-18, NLRP1, NLRP2, PLCG1, TIRAP, and TNF) was constructed using LASSO Cox regression analysis and could predict the overall survival of PC patients with medium-to-high accuracy. In comparison to an ideal model, a predictive nomogram revealed that 3- and 5-year overall survival rates in the entire cohort could be predicted relatively well. Our research was the first to identify a pyroptosis-related prognostic gene signature for PC, expanding the number of options for prognostic prediction in the disease.

Our study has some limitations. Due to the extremely low survival rate for pancreatic cancer, our risk stratification system

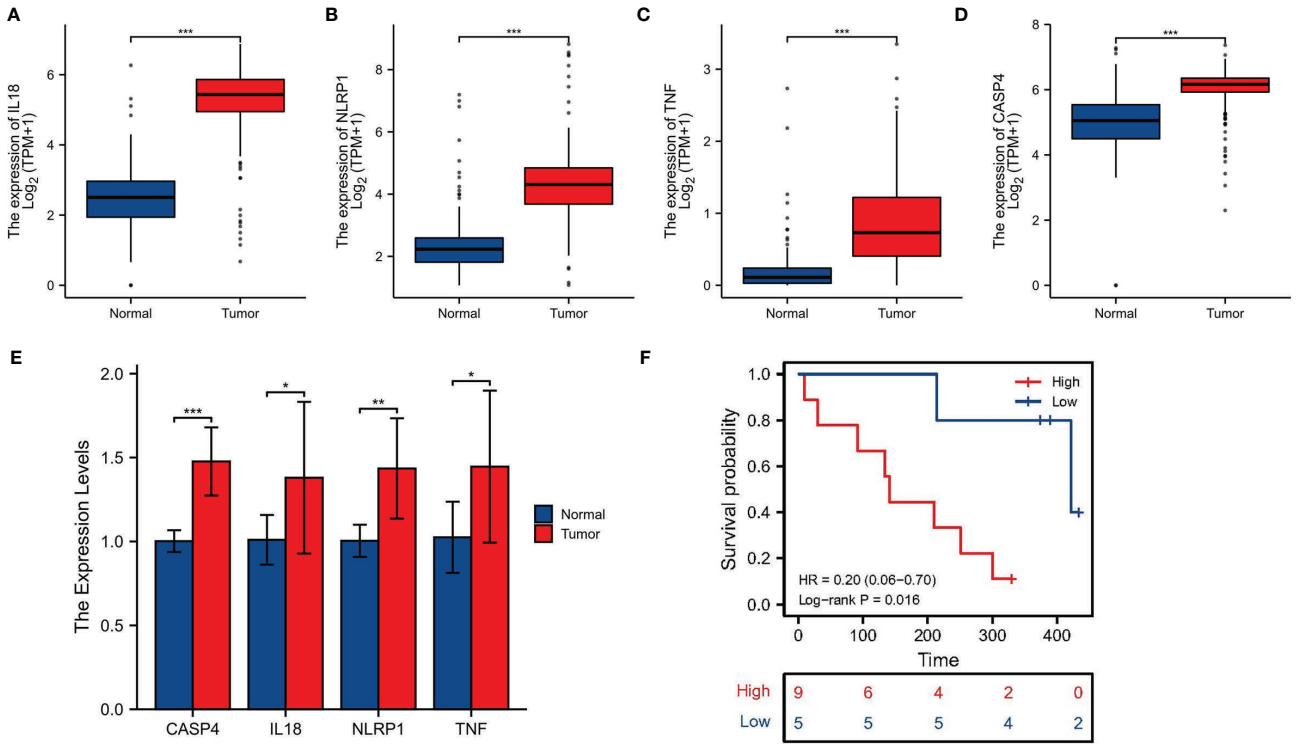



**FIGURE 7** | Clinical sample validation. **(A–D)** The expression of four signature-associated genes between tumor tissues and normal tissues on basis of the GTEx database and TCGA database. **(E)** qRT-PCR of the expression of four signature-associated genes in clinical samples. **(F)** Survival analysis. *p < 0.05; **p < 0.01; ***P < 0.001.

is not predictive in external experiments. Other limitations include the relatively small number of normal samples from TCGA database, and the hypothesis did not use animal models to validate. To be precise, these factors must be verified in further investigations. Therefore, in the future, we will create animal models to test these hypotheses further.

## 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 study was authorized by the Ethical Committee of Affiliated Hospital of Nantong University (2022-L096).

## AUTHOR CONTRIBUTIONS

JZ and XY conceived of and designed the study. DK were responsible for the materials and supplement experiments. JZ drafted the manuscript. GZ critically revised the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was funded by the National Natural Science Foundation of China (81572397).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.932786/full#supplementary-material

**Supplementary Table 1** | Result of GO analysis.

**Supplementary Table 2** | Result of KEGG analysis.

**Supplementary Table 3** | Results of drug correlation analysis.

**Supplementary Figure 1** | Survival analysis of signature-associated genes in different data sets.

**Supplementary Figure 2** | The relationship of eight prognostic PRGs with immune infiltration.

**Supplementary Figure 3** | Correlation analysis between MSI and PRG in PC.

**Supplementary Figure 4** | Correlation analysis between TMB and PRG in PC.

## REFERENCES


1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2020. *CA Cancer J Clin* (2020) 70:7–30. doi: 10.3322/caac.21590
2. Zhang L, Sanagapalli S, Stoita A. Challenges in Diagnosis of Pancreatic Cancer. *World J Gastroenterol* (2018) 24:2047–60. doi: 10.3748/wjg.v24.i19.2047
3. Orth M, Metzger P, Gerum S, Mayerle J, Schneider G, Belka C, et al. Pancreatic Ductal Adenocarcinoma: Biological Hallmarks, Current Status, and Future Perspectives of Combined Modality Treatment Approaches. *Radiat Oncol* (2019) 14:141. doi: 10.1186/s13014-019-1345-6
4. Le N, Sund M, Vinci A, Pancreas Gcgo. Prognostic and Predictive Markers in Pancreatic Adenocarcinoma. *Dig Liv Dis* (2016) 48:223–30. doi: 10.1016/j.dld.2015.11.001
5. Collisson EA, Bailey P, Chang DK, Biankin AV. Molecular Subtypes of Pancreatic Cancer. *Nat Rev Gastroenterol Hepatol* (2019) 16:207–20. doi: 10.1038/s41575-019-0109-y
6. Zhu X, Zhang T, Zhang Y, Chen H, Shen J, Jin X, et al. A Super-Enhancer Controls TGF- Beta Signaling in Pancreatic Cancer Through Downregulation of TGFBR2. *Cell Signal* (2020) 66:109470. doi: 10.1016/j.cellsig.2019.109470
7. Yu W, Ma Y, Roy SK, Srivastava R, Shankar S, Srivastava RK. Ethanol Exposure of Human Pancreatic Normal Ductal Epithelial Cells Induces EMT Phenotype and Enhances Pancreatic Cancer Development in KC (Pdx1-Cre and LSL-Kras(G12D) ) Mice. *J Cell Mol Med* (2022) 26:399–409. doi: 10.1111/jcmm.17092
8. Xu F, Wang H, Pei H, Zhang Z, Liu L, Tang L, et al. SLC1A5 Prefers to Play as an Accomplice Rather Than an Opponent in Pancreatic Adenocarcinoma. *Front Cell Dev Biol* (2022) 10:800925. doi: 10.3389/fcell.2022.800925
9. Liu L, Huang X, Shi F, Song J, Guo C, Yang J, et al. Combination Therapy for Pancreatic Cancer: Anti-PD-(L)1-Based Strategy. *J Exp Clin Cancer Res* (2022) 41:56. doi: 10.1186/s13046-022-02273-w
10. Golan T, Hammel P, Reni M, Van Cutsem E, Macarulla T, Hall MJ, et al. Maintenance Olaparib for Germline BRCA-Mutated Metastatic Pancreatic Cancer. *N Engl J Med* (2019) 381:317–27. doi: 10.1056/NEJMoa1903387
11. Liu X, Ye N, Liu S, Guan J, Deng Q, Zhang Z, et al. Hyperbaric Oxygen Boosts PD-1 Antibody Delivery and T Cell Infiltration for Augmented Immune Responses Against Solid Tumors. *Adv Sci (Weinh)* (2021) 8:e2100233. doi: 10.1002/advs.202100233
12. Bosco G, Casarotto A, Nasole E, Camporesi E, Salvia R, Giovinazzo F, et al. Preconditioning With Hyperbaric Oxygen in Pancreaticoduodenectomy: A Randomized Double-Blind Pilot Study. *Anticancer Res* (2014) 34:2899–906.
13. Liu M, Zhang Y, Yang J, Cui X, Zhou Z, Zhan H, et al. ZIP4 Increases Expression of Transcription Factor ZEB1 to Promote Integrin Alpha3beta1 Signaling and Inhibit Expression of the Gemcitabine Transporter ENT1 in Pancreatic Cancer Cells. *Gastroenterology* (2020) 158:679–92.e1. doi: 10.1053/j.gastro.2019.10.038
14. Kovacs SB, Miao EA. Gasdermins: Effectors of Pyroptosis. *Trends Cell Biol* (2017) 27:673–84. doi: 10.1016/j.tcb.2017.05.005
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. Miao EA, Rajan JV, Aderem A. Caspase-1-Induced Pyroptotic Cell Death. *Immunol Rev* (2011) 243:206–14. doi: 10.1111/j.1600-065X.2011.01044.x
17. Broz P, Pelegrin P, Shao F. The Gasdermins, a Protein Family Executing Cell Death and Inflammation. *Nat Rev Immunol* (2020) 20:143–57. doi: 10.1038/s41577-019-0228-2
18. 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) 535:111–6. doi: 10.1038/nature18590
19. 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) 10:650. doi: 10.1038/s41419-019-1883-8
20. Yu J, Li S, Qi J, Chen Z, Wu Y, Guo J, et al. Cleavage of GSDME by Caspase-3 Determines Lobaplatin-Induced Pyroptosis in Colon Cancer Cells. *Cell Death Dis* (2019) 10:193. doi: 10.1038/s41419-019-1441-4
21. Zhang Y, Chen X, Gueydan C, Han J. Plasma Membrane Changes During Programmed Cell Deaths. *Cell Res* (2018) 28:9–21. doi: 10.1038/cr.2017.133
22. Frank D, Vince JE. Pyroptosis Versus Necroptosis: Similarities, Differences, and Crosstalk. *Cell Death Differ* (2019) 26:99–114. doi: 10.1038/s41418-018-0212-6
23. Tan Y, Chen Q, Li X, Zeng Z, Xiong W, Li G, et al. Pyroptosis: A New Paradigm of Cell Death for Fighting Against Cancer. *J Exp Clin Cancer Res* (2021) 40:153. doi: 10.1186/s13046-021-01959-x
24. Kolb R, Liu GH, Janowski AM, Sutterwala FS, Zhang W. Inflammasomes in Cancer: A Double-Edged Sword. *Protein Cell* (2014) 5:12–20. doi: 10.1007/s13238-013-0001-4
25. Ye Y, Dai Q, Qi H. A Novel Defined Pyroptosis-Related Gene Signature for Predicting the Prognosis of Ovarian Cancer. *Cell Death Discovery* (2021) 7:71. doi: 10.1038/s41420-021-00451-x
26. Lin W, Chen Y, Wu B, Chen Y, Li Z. Identification of the Pyroptosisrelated Prognostic Gene Signature and the Associated Regulation Axis in Lung Adenocarcinoma. *Cell Death Discovery* (2021) 7:161. doi: 10.1038/s41420-021-00557-2
27. Xu F, Zhang Z, Yuan M, Zhao Y, Zhou Y, Pei H, et al. M6A Regulatory Genes Play an Important Role in the Prognosis, Progression and Immune Microenvironment of Pancreatic Adenocarcinoma. *Cancer Invest* (2021) 39:39–54. doi: 10.1080/07357907.2020.1834576
28. Xu F, Zhang Z, Zhao Y, Zhou Y, Pei H, Bai L, et al. Bioinformatic Mining and Validation of the Effects of Ferroptosis Regulators on the Prognosis and Progression of Pancreatic Adenocarcinoma. *Gene* (2021) 795:145804. doi: 10.1016/j.gene.2021.145804
29. Qiu S, Hu Y, Dong S. Pan-Cancer Analysis Reveals the Expression, Genetic Alteration and Prognosis of Pyroptosis Key Gene GSDMD. *Int Immunopharmacol* (2021) 101:108270. doi: 10.1016/j.intimp.2021.108270
30. Chen L, Weng B, Li H, Wang H, Li Q, Wei X, et al. A Thiopyran Derivative With Low Murine Toxicity With Therapeutic Potential on Lung Cancer Acting Through a NF-kappaB Mediated Apoptosis-to-Pyroptosis Switch. *Apoptosis* (2019) 24:74–82. doi: 10.1007/s10495-018-1499-y
31. Zhao C, Liu Z, Liang G. Promising Curcumin-Based Drug Design: Mono-Carbonyl Analogues of Curcumin (MACs). *Curr Pharm Des* (2013) 19:2114–35.
32. 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) 579:421–6. doi: 10.1038/s41586-020-2079-1
33. Dupaul-Chicoine J, Yeretssian G, Doiron K, Bergstrom KS, McIntire CR, LeBlanc PM, et al. Control of Intestinal Homeostasis, Colitis, and Colitis-Associated Colorectal Cancer by the Inflammatory Caspases. *Immunity* (2010) 32:367–78. doi: 10.1016/j.immuni.2010.02.012
34. Tang R, Xu J, Zhang B, Liu J, Liang C, Hua J, et al. Ferroptosis, Necroptosis, and Pyroptosis in Anticancer Immunity. *J Hematol Oncol* (2020) 13:110. doi: 10.1186/s13045-020-00946-7
35. Erkes DA, Cai W, Sanchez IM, Purwin TJ, Rogers C, Field CO, et al. Mutant BRAF and MEK Inhibitors Regulate the Tumor Immune Microenvironment *via* Pyroptosis. *Cancer Discov* (2020) 10:254–69. doi: 10.1158/2159-8290.CD-19-0672
36. Xi G, Gao J, Wan B, Zhan P, Xu W, Lv T, et al. GSDMD Is Required for Effector CD8(+) T Cell Responses to Lung Cancer Cells. *Int Immunopharmacol* (2019) 74:105713. doi: 10.1016/j.intimp.2019.105713
37. 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) 579:415–20. doi: 10.1038/s41586-020-2071-9
38. Zhang JY, Zhou B, Sun RY, Ai YL, Cheng K, Li FN, et al. The Metabolite Alpha-KG Induces GSDMC-Dependent Pyroptosis Through Death Receptor 6-Activated Caspase-8. *Cell Res* (2021) 31:980–97. doi: 10.1038/s41422-021-00506-9
39. Blum A, Wang P, Zenklusen JC. SnapShot: TCGA-Analyzed Tumors. *Cell* (2018) 173:530. doi: 10.1016/j.cell.2018.03.059
40. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: The Reference Human Genome Annotation for The ENCODE Project. *Genome Res* (2012) 22:1760–74. doi: 10.1101/gr.135350.111
41. Liang H, Yu T, Han Y, Jiang H, Wang C, You T, et al. LncRNA PTAR Promotes EMT and Invasion-Metastasis in Serous Ovarian Cancer by Competitively Binding miR-101-3p to Regulate ZEB1 Expression. *Mol Cancer* (2018) 17:119. doi: 10.1186/s12943-018-0870-5

42. Feng S, Xia T, Ge Y, Zhang K, Ji X, Luo S, et al. Computed Tomography Imaging-Based Radiogenomics Analysis Reveals Hypoxia Patterns and Immunological Characteristics in Ovarian Cancer. *Front Immunol* (2022) 13:868067. doi: 10.3389/fimmu.2022.868067
43. Feng S, Yin H, Zhang K, Shan M, Ji X, Luo S, et al. Integrated Clinical Characteristics and Omics Analysis Identifies a Ferroptosis and Iron-Metabolism-Related lncRNA Signature for Predicting Prognosis and Therapeutic Responses in Ovarian Cancer. *J Ovarian Res* (2022) 15:10. doi: 10.1186/s13048-022-00944-y
44. Zheng S, Zou Y, Liang JY, Xiao W, Yang A, Meng T, et al. Identification and Validation of a Combined Hypoxia and Immune Index for Triple-Negative Breast Cancer. *Mol Oncol* (2020) 14:2814–33. doi: 10.1002/1878-0261.12747
45. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of Drug Sensitivity in Cancer (GDSC): A Resource for Therapeutic Biomarker Discovery in Cancer Cells. *Nucleic Acids Res* (2013) 41:D955–61. doi: 10.1093/nar/gks1111
46. Yang F, Bettadapura SN, Smeltzer MS, Zhu H, Wang S. Pyroptosis and Pyroptosis-Inducing Cancer Drugs. *Acta Pharmacol Sin* (2022). doi: 10.1038/s41401-022-00887-6
47. Westerterp M, Fotakis P, Ouimet M, Bochem AE, Zhang H, Molusky MM, et al. Cholesterol Efflux Pathways Suppress Inflammasome Activation, NETosis, and Atherogenesis. *Circulation* (2018) 138:898–912. doi: 10.1161/CIRCULATIONAHA.117.032636
48. Seifert L, Werba G, Tiwari S, Giao Ly NN, Alothman S, Alqunaibit D, et al. The Necrosome Promotes Pancreatic Oncogenesis *via* CXCL1 and Mincle-Induced Immune Suppression. *Nature* (2016) 532:245–9. doi: 10.1038/nature17403
49. Zhong FL, Mamai O, Sborgi L, Boussofara L, Hopkins R, Robinson K, et al. Germline NLRP1 Mutations Cause Skin Inflammatory and Cancer Susceptibility Syndromes *via* Inflammasome Activation. *Cell* (2016) 167:187–202.e17. doi: 10.1016/j.cell.2016.09.001
50. Khanova E, Wu R, Wang W, Yan R, Chen Y, French SW, et al. Pyroptosis by Caspase11/4-Gasdermin-D Pathway in Alcoholic Hepatitis in Mice and Patients. *Hepatology* (2018) 67:1737–53. doi: 10.1002/hep.29645
51. Banh RS, Biancur DE, Yamamoto K, Sohn ASW, Walters B, Kuljanin M, et al. Neurons Release Serine to Support mRNA Translation in Pancreatic Cancer. *Cell* (2020) 183:1202–18.e25. doi: 10.1016/j.cell.2020.10.016
52. Li MX, Wang HY, Yuan CH, Ma ZL, Jiang B, Li L, et al. Establishment of a Macrophage Phenotypic Switch Related Prognostic Signature in Patients With Pancreatic Cancer. *Front Oncol* (2021) 11:619517. doi: 10.3389/fonc.2021.619517
53. Huang J, Chen Z, Ding C, Lin S, Wan D, Ren K. Prognostic Biomarkers and Immunotherapeutic Targets Among CXC Chemokines in Pancreatic Adenocarcinoma. *Front Oncol* (2021) 11:711402. doi: 10.3389/fonc.2021.711402

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 22 June 2022
doi: 10.3389/fonc.2022.911309

# The Function, Role and Process of DDX58 in Heart Failure and Human Cancers

*Ping Yu*[1], *Peng Liang*[1], *Shifeng Pang*[1], *Wenjian Yuan*[1], *Yuxiang Zhao*[2,3]* *and Qiaojuan Huang*[1]*

[1] Department of Cardiology, The Second Affiliated Hospital of Guangxi Medical University, Nanning, China, [2] United New Drug Research and Development Center, Biotrans Technology Co., LTD., Ningbo, China, [3] Institute of Bioengineering, Biotrans Technology Co., LTD., Ningbo, China

**Background:** Heart failure (HF) is the most common outcome of cardiovascular disease, and an increasing number of patients with heart failure die from noncardiac causes, such as cancer. Epidemiological data suggest that ischemic cardiomyopathy–induced HF (ischemic HF) may be associated with an increased incidence of cancer. This study aimed to investigate the possible mechanisms of the association between ischemic HF and cancer, as well as potential therapeutic targets.

**Methods:** Weighted gene co-expression network analysis was performed to analyze the correlations between phenotypes and gene modules using immune cells as phenotypes. Differential analysis was then performed to screen differentially expressed genes (DEGs) in ischemic HF and normal control samples. The macrophage-related Brown module was identified as the key module, and immune-related DEGs were obtained by taking the intersection of the Brown module, DEGs, and immune-related genes using a Venn diagram. DDX58 was identified as the key gene using a protein–protein interaction network and expression analyses and validated using immunohistochemistry. Kaplan–Meier survival analysis was performed to analyze the correlation between DDX58 expression and tumor prognosis. Spearman correlation analysis was performed to assess the correlation between DDX58 expression and immune cell infiltration.

**Results:** DDX58 was identified as a key immune-related gene associated with ischemic HF and was highly expressed in most cancer types. The survival analysis revealed a significant negative correlation between high DDX58 expression and prognosis in multiple tumor types. Moreover, DDX58 expression was significantly associated with immune cell infiltration and immune checkpoint gene expression in many cancer types.

**Conclusion:** DDX58 is a key immune-related gene in ischemic HF and may play a crucial role in the relationship between ischemic HF and cancer. Pan-cancer analysis suggests that DDX58 is a promising clinical prognostic marker for most cancers and may be a therapeutic target for cancer patients and ischemic HF patients at an increased risk of cancer.

**Keywords: heart failure, cancer, immune infiltration, DDX58, ischemic cardiomyopathy**

OPEN ACCESS

***Edited by:***
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

***Reviewed by:***
*Xuewei Chen,*
*Tianjin Academy of Environmental Protection Sciences (TAES), China*
*Chao Zhang,*
*Tongji University, China*

***Correspondence:**
*Yuxiang Zhao*
*yuxiang_zhao@biotranstech.com*
*Qiaojuan Huang*
*huangqiaojuan@gxmu.edu.cn*

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

***Received:*** *02 April 2022*
***Accepted:*** *26 May 2022*
***Published:*** *22 June 2022*

***Citation:***
*Yu P, Liang P, Pang S, Yuan W, Zhao Y and Huang Q (2022) The Function, Role and Process of DDX58 in Heart Failure and Human Cancers. Front. Oncol. 12:911309. doi: 10.3389/fonc.2022.911309*

## INTRODUCTION

Heart failure (HF) is a complex and severe syndrome occurring when the heart is unable to deliver sufficient blood to meet its own needs under normal filling pressures (1). HF currently affects more than 60 million people worldwide, and the number continues to increase (2). The major causes of HF include coronary artery disease, hypertension, diabetes, and obesity, and the risk increases over time (3). The phenotype of HF patients varies across regions. In Africa, the main causes are hypertensive heart disease and dilated cardiomyopathy. In Asia and South America, the main cause is coronary artery disease. In Western countries, ischemic causes are significantly more prominent than in East Asia (4–6).

HF treatment techniques have improved in recent years. However, an increasing number of HF patients die from noncardiac causes, such as cancer (7, 8). Growing evidence suggests a causal relationship between HF and cancer (8). In addition to the increased risk of HF due to cancer itself and its treatments (9–11), patients with ischemic cardiomyopathy–induced HF (ischemic HF) have a 71% higher risk of developing cancer. This trend usually emerges 1.5 years after HF diagnosis and is most pronounced in HF patients with reduced ejection fractions (7, 12, 13). Epidemiological data suggest that HF patients diagnosed with cancer are more prone to in-hospital complications and death (14), and the European Society of Cardiology has called on medical professionals to pay more attention to the incidence of cancer in HF patients (15). However, due to the complexity of the interaction between HF and cancer, the pathophysiological mechanisms involved are unclear.

Immune system dysfunction appears to play a crucial role in myocardial remodeling and cancer development after myocardial infarction (16, 17). The monocyte system may be an important link between HF and cancer (18). The advent of immunotherapy has ushered in a new direction in the treatment of cancer and other autoimmune diseases, such as cardiovascular disease. In clinical practice, immunotherapies targeting various immune checkpoints have shown significant efficacy against various cancer types. However, currently available immune checkpoint inhibitors are expensive and have many cardiotoxic side effects, the most common of which is myocarditis (19). Myocarditis induced by immunodetection site inhibitors has been reported to have morbidity and mortality rates of up to 50% (20). Therefore, the discovery of safer and more effective therapeutic approaches is of great importance for the treatment of cancer and other autoimmune diseases. In addition, the recently proposed therapeutic strategy for improving DDX58-mediated innate immunity can be used as a complementary mechanism to compensate for the lack of immune checkpoint therapy (21, 22). To that end, this study aimed to explore the role of ischemic HF immune-related genes in pan-cancer through a comprehensive bioinformatics analysis, hoping to gain new insights into the association between ischemic HF and cancer and to reveal potential therapeutic targets that can help reduce the incidence of cancer in ischemic HF patients.

## MATERIALS AND METHODS

### Data Sources

Twelve cases of ischemic HF and five normal control samples were downloaded from the GSE42955 dataset of the Gene Expression Omnibus (GEO) database, and 19 cases of HF and five normal control samples were downloaded from the GSE26887 dataset. Expression data for various tumor cell lines were downloaded from the Cancer Cell Line Encyclopedia (CCLE) database (23). Gene expression profiles of 33 cancers and normal tissues were obtained from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) databases (24). A total of 7398 immune-related genes were downloaded from InnateDB (25).

### Single-Sample Gene Set Enrichment Analysis

The Immune Cell Abundance Identifier (ImmuCellAI) tool (http://bioinfo.life.hust.edu.cn/ImmuCellAI#!/ ) was used to estimate GSE42955 microarray data for single-sample Gene Set Enrichment Analysis (ssGSEA) to observe the infiltration abundance of 24 immune cells in ischemic HF and normal control samples, including 18 T cell subtypes and six other immune cells: DC, B cells, monocytes, macrophages, NK, neutrophils, $CD4^+$ T, $CD8^+$ T, NKT, Tgd, CD4 naive, CD8 naive, Tc, Tex, Tr1, nTreg, iTreg, Th1, Th2, Th17, Tfh, Tcm, Tem, and MAIT. Pearson correlation analysis was performed to assess the correlations between gene expression and immune cell infiltration. A value of $P < 0.05$ was considered statistically significant.

### Weighted Gene Co-Expression Network Analysis

Weighted gene co-expression network analysis (WGCNA) was performed using the R package "WGCNA" to identify gene modules associated with key immune cells in GSE42955. A soft threshold of $\beta = 14$ was chosen to achieve a scale-free topology by generating an adjacency matrix between pairs of genes using Pearson correlation analysis. Gene dendrograms and module colors were created using phase dissimilarity. The most significant gene modules (26) were identified by assessing the correlations between phenotypes and module genes.

### Analysis of Variance

The samples in the GSE42955 dataset were divided into ischemic HF and normal control samples, and differential analysis was performed using the R package "limma." The results were expressed as volcano plots with a screening threshold of $|\log2FC| > 0.3785$ for differentially expressed genes (DEGs) at $P < 0.05$. Venn diagrams were used to take intersections of significant DEGs, key WGCNA modules, and immune-related genes to obtain immune-related DEGs (IRDEGs).

### Protein–Protein Interaction Network

Metascape (https://metascape.org/gp/index.html#/main/step1) was used to analyze the biological functions of IRDEGs (27). Additionally, the obtained IRDEGs were imported into the

STRING database (https://string-db.org/) to obtain a network of interactions. The top 10 most important genes were identified using the MCC algorithm in the CytoHubba plugin in Cytoscape software.

## Expression of Key Genes

The samples were divided into two parts based on the clinical characteristics or median expression of macrophages to compare the expression levels of key genes in different groups. Violin plots were used to represent the expression levels of genes or immune cells. Sankey plots were used to visualize the distribution of samples in different subgroups.

## Immunohistochemical Validation

Myocardial tissues from rats with HF after myocardial infarction and rats in a sham-operated group were used for immunohistochemical analysis. Animal experiment ethics has been approved by The Animal Care & Welfare Committee of Guangxi Medical University (No.201904027), and all experimental operations ensure animal ethical norms are followed. Briefly, after antigen repair and endogenous peroxidase elimination, sections were placed in a goat serum blocking solution at 37°C for 10–15 min. They were then incubated overnight at 4°C with a primary antibody against DDX58 (1:100; Affinity). The sections were subsequently rinsed with PBS three times and then incubated with a secondary antibody at 37°C for 20 min to develop color and hematoxylin lines for restaining. Samples were observed and imaged using light microscopy (Olympus, Tokyo, Japan).

## Functional and Pathway Clustering

DEG enrichment analysis was performed to identify significant Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The level of statistical significance was set to $P < 0.05$. Moreover, ssGSEA was used to identify the KEGG and Hallmark pathways of the samples (28). Thresholds of $|NES| > 1$, NOM $P < 0.05$, and FDR $q < 0.25$ were used.

## Expression and Survival Analyses

Kaplan–Meier survival analysis was performed to assess the expression levels of DDX58 in different tumor tissues and analyze the correlations between DDX58 expression and overall survival (OS), disease-specific survival (DSS), and progression-free interval (PFI) in cancer patients. Univariate Cox regression analysis was performed to determine the hazard ratio (HR), with HR > 1 indicating high DDX58 expression as a high-risk factor, and a log-rank test was used to determine the *P*-values, with the level of statistical significance set to $P < 0.05$.

## Immune Infiltration Analysis

The correlation between DDX58 expression levels and immune cell infiltration was analyzed in 33 cancers using the Tumor Immune Evaluation Resource (TIMER) database and the CIBERSORT algorithm. Spearman correlation analysis was performed to determine the *P*-values (29). The immune, stromal, and ESTIMATE scores of different tumors were calculated using the R package "ESTIMATE" to predict tumor purity (30, 31) and to assess the correlations between the scores and DDX58 expression. Furthermore, 47 common immune checkpoint genes were collected, and their correlations with DDX58 expression were assessed using Spearman rank correlation analysis. Values of $P < 0.05$ were considered statistically significant.

# RESULTS

## Immune Infiltration in Ischemic HF

In recent years, the importance of immune cells has been demonstrated in various diseases, including HF (32, 33). In this study, immune cell infiltration into ischemic HF and normal control tissues was first analyzed using ssGSEA. A heat map was constructed to visualize the infiltration of 24 immune cells into HF and normal control tissues (**Figure 1A**). The infiltration abundances of these immune cells were then compared. The results showed that the abundance of Tc and $CD8^+$ T cells was significantly higher in HF than in normal control samples, whereas that of macrophages was significantly lower in HF samples (**Figure 1B**). These results suggest that the inflammatory response generated by these immune cells may be associated with ischemic HF.

## Genetic Features Associated With Immune Cells

WGCNA was performed to further investigate the genetic features associated with immune cell infiltration. **Figure 2A**

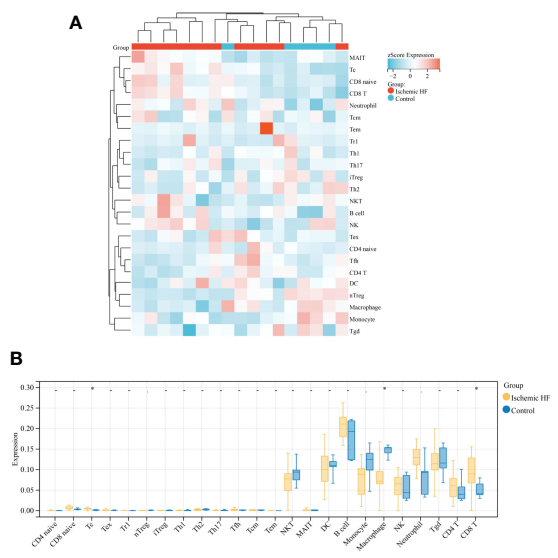



**FIGURE 1** | Comparison of immune infiltration in HF and normal control samples **(A)** Heat map of immune infiltration between Ischemic HF and Control samples based on ssGSEA score analysis; **(B)** Comparison of immune cell infiltration in Ischemic HF and Control samples. *$P < 0.05$.

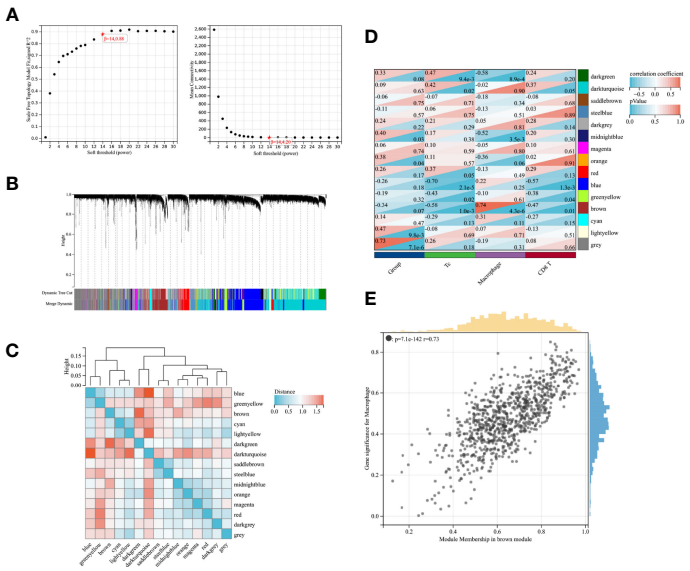



**FIGURE 2** | WGCNA screening of gene modules associated with immune infiltration **(A)** Scale-free fit indices and mean connectivity for soft thresholds; **(B)** Dendrogram of immune-associated genes based on phase dissimilarity metric clustering; **(C)** Module feature vector clustering; **(D)** Heat map of gene modules with feature correlation; **(E)** Module feature genes associated with Macrophage abundance in brown modules.

shows the scale-free fit indices and average connectivity for various soft threshold capabilities. The genes were divided into 15 modules using hierarchical clustering analysis and represented using branches of the clustering tree and different colors (**Figures 2B, C**). Heat maps were created by calculating the correlations between the gene models and phenotypes (**Figure 2D**). All gene modules showed weak correlations with Tc and $CD8^+$ T cells, with correlation values less than 0.5. The Brown module had the strongest correlation with macrophages ($r = 0.73$, $P = 7.1e\text{-}142$; **Figure 2E**).

## Analysis of Variance

A total of 737 DEGs, including 547 downregulated and 190 upregulated genes, were identified in GSE42955 and compared with normal control samples (**Figure 3A**). The biological functions of DEGs were subsequently investigated using GO and KEGG pathway enrichment analyses. Most DEGs were enriched in GO terms related to immune responses, such as immune system processes and innate immune responses (**Figure 3B**). Most of the enriched KEGG pathways were associated with diseases such as influenza A, systemic lupus erythematosus (SLE), and pertussis, as well as with immune-related pathways such as complement and coagulation cascade reactions, antigen processing, and presentation (**Figure 3C**). DEG, in the Brown module gene, and InnateDB immune-related gene overlapping to obtain 172 immune-related DEGs (**Figure 3D**).

## Selection of Key Genes

Functional enrichment analysis showed that 172 IRDEGs were mainly involved in the signaling and regulation of cytokines in the immune system, regulation of the immune system and immune response, interferon signaling, and response to and regulation of bacteria or viruses (**Figure 4A**). A network diagram was drawn by analyzing the functional role relationships between these IRDEGs (**Figure 4B**). Ten genes were obtained using the MCC algorithm: IRF9, MX1, MX2, IFIH1, IFIT1, IFIT2, IFIT3, STAT1, DDX58, and OAS2 (**Figure 4C**).

## Validation of Potential Ischemic HF Biomarkers

The Sankey diagram in **Figure 5A** shows the distribution among sample types, macrophage infiltration abundance, and high and low DDX58 expression in GSE42955. In the high macrophage infiltration group, Tc cells showed lower infiltration levels, while macrophages and $CD8^+$ T cells showed higher infiltration abundances (**Figure 5B**). The infiltration levels of immune cells were also lower in the high macrophage infiltration group (**Figure 5C**). Among the 10 key genes, only DDX58 was differentially expressed in this group (**Figure 5D**). High DDX58 expression was also observed in the HF samples (**Figure 5E**). These results suggest that high DDX58 levels are closely associated with both high macrophage infiltration levels and HF. Immunohistochemistry was performed to verify the expression levels of DDX58 in rat HF samples. The results showed significantly higher DDX58 expression in the post-infarction HF samples than in the sham-operated group (**Figure 5F**).

## Biological Functions of DDX58

GSEA was performed to evaluate the potential biological mechanisms of macrophages. The main KEGG pathways associated with the immune response were the TOLL-like receptor signaling pathway, complement and coagulation

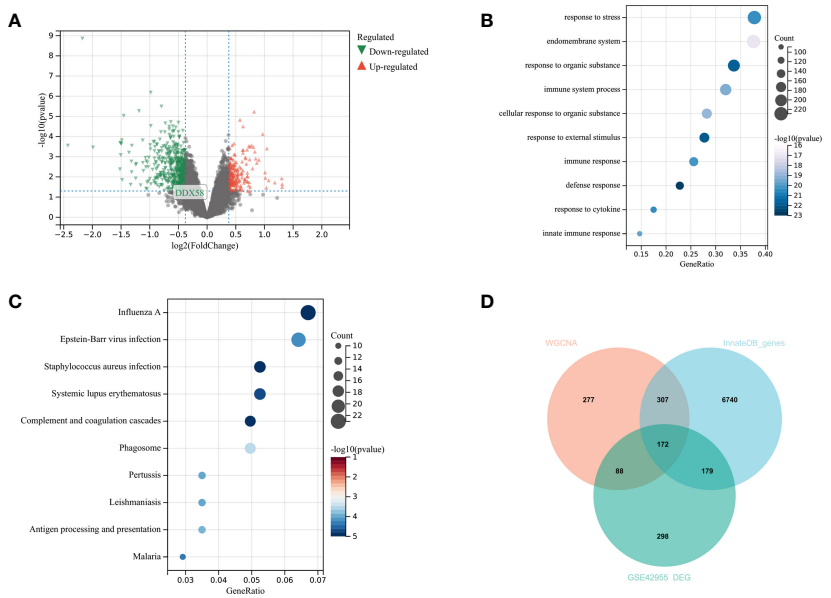



**FIGURE 3** | Analysis of differences in GSE42955 **(A)** Volcano plot of significant DEGs between ischemic heart failure and normal control samples; **(B, C)** top 10 enriched GO terms and KEGG pathways of DEGs, respectively; **(D)** Venn diagram of DEGs, WGCNA key modules and immune gene sets.

cascades, NOD-like receptor signaling pathway, FCγR-mediated phagocytosis, and chemokine signaling pathway (**Figure 6A**). The main Hallmark pathways were interferon γ response, interferon α response, IL6 JAK STAT3 signaling, IL2 STAT5 signaling, and TNFA signaling *via* NFKB (**Figure 6B**). The biological function of DDX58 in HF was then evaluated. Significant enrichment was observed in immune response and inflammation-related diseases such as SLE and FCγR-mediated phagocytosis and the TOLL-like receptor signaling pathway and other KEGG pathways (**Figure 6C**), TNFA signaling *via* NFKB, IL6 JAK STAT3 signaling, interferon γ response, interferon α response, and other Hallmark pathways (**Figure 6D**). Thus, the pathways enriched by macrophages and DDX58 were similar.

## Expression of DDX58 in Different Cancer Types

The expression of DDX58 in various cancer types was analyzed to explore its relationship with pan-cancer. An analysis based on the CCLE database showed that DDX58 was stably expressed in 21 tumor cell lines (**Figure 7A**). An analysis based on the TCGA database identified higher DDX58 expression in Breast invasive carcinoma (BRCA), Cholangiocarcinoma (CHOL), Esophageal carcinoma (ESCA), Glioblastoma multiforme (GBM), Head and Neck squamous cell carcinoma (HNSC), Kidney renal clear cell carcinoma (KIRC), Brain Lower Grade Glioma (LGG), Liver hepatocellular carcinoma (LIHC), and Stomach adenocarcinoma (STAD) than in normal tissues (**Figure 7B**). To compensate for the lack of normal tissues, a combined analysis using the GTEx and TCGA databases showed that DDX58 was highly expressed in most tumor types (**Figure 7C**). These findings suggest that DDX58 expression is closely related to tumorigenesis.

## Prognostic Potential of DDX58 in Pan-Cancer

Survival analysis is a common method for disease prognosis studies and is often used to explore the impact of prognostic factors on disease outcomes. The analysis of the correlation between DDX58 expression and OS in patients with cancer revealed a significant negative correlation between DDX58 expression and OS in Adrenocortical carcinoma (ACC), LGG, Lung adenocarcinoma (LUAD), and Pancreatic adenocarcinoma (PAAD) and a significant positive correlation in KIRC, Mesothelioma (MESO), and skin cutaneous melanoma (SKCM) (**Figure 8**). The DSS analysis showed that DDX58 expression negatively correlated with DSS in Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC), LGG, LUAD, PAAD, and Uterine Corpus Endometrial Carcinoma (UCEC) and positively correlated with DSS in KIRC and SKCM (**Supplementary Figure 1**). The PFI analysis showed that high DDX58 expression significantly attenuated PFI in patients with six tumors: DLBC, LGG, LUAD, PAAD, Prostate adenocarcinoma (PRAD), and UCEC (**Supplementary Figure 2**). These findings indicate that DDX58 expression affects the prognosis of various tumors, suggesting that DDX58 may be a pan-cancer prognostic indicator.

## Correlation Between DDX58 and Cancer Immune Infiltration

Immune cell infiltration plays an important role in the progression and prognostic links of tumors. The analysis of the correlation between DDX58 expression and immune cell infiltration in different tumor tissues showed that DDX58 expression significantly correlated with the abundance of six

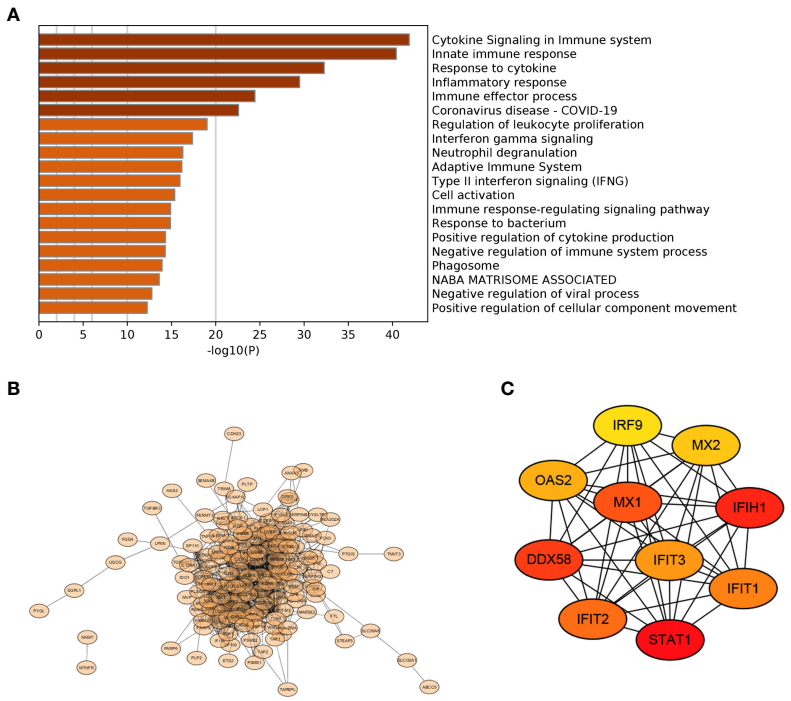



**FIGURE 4** | PPI selection of key genes **(A)** Functional terms of 172 IRDEGs; **(B)** PPI network diagram of 172 IRDEGs; **(C)** Top 10 important genes in the MCC algorithm.

types of immune cell infiltration in most cancers, including BRCA, Colon adenocarcinoma (COAD), KIRC, LGG, LIHC, Lung squamous cell carcinoma (LUSC), Ovarian serous cystadenocarcinoma (OV), PAAD, PRAD, Rectum adenocarcinoma (READ), and SKCM. A significant correlation was not observed only in Uveal Melanoma (UVM). In the immune cell infiltration analysis of 33 cancer types, DDX58 expression significantly correlated with macrophages in 24 types (**Figure 9A**). In the analysis of the correlations between DDX58 expression and 22 immune cells in 33 tumors using CIBERSORT, Macrophage_M1 showed the most prominent expression (**Figure 9B**). The analysis of the correlations between DDX58 expression and the immune, stromal, and ESTIMATE scores revealed significant correlations with all three scores in 20 cancer types. The three cancer types most strongly correlating with the immune score were Bladder Urothelial Carcinoma (BLCA), COAD, and HNSC. The three cancer types most strongly correlating with the stromal score were COAD, LGG, and SKCM. The three cancer types most strongly correlating with the ESTIMATE score were BLCA, COAD, and SKCM (**Figure 9C**). This suggests that DDX58 may be involved in the regulation of tumor immune cell infiltration. To examine whether DDX58 could be a target for tumor immunotherapy, the correlations between its expression levels and immune checkpoint genes were investigated. The results showed that DDX58 expression significantly correlated with the expression of 47 immune checkpoint genes in most cancer types, 44 in COAD, and 43 in Thyroid carcinoma (THCA) (**Figure 9D**). Overall, these results suggest that DDX58 plays an important role in tumor immune infiltration.

## DISCUSSION

HF caused by ischemic heart disease is a common cause of morbidity and mortality in developed countries. Although pharmacological treatments and interventions have improved the OS of HF patients over the past 20 years, the prevalence of HF and hospitalization rates continue to increase (34). Moreover, although the development of therapeutic techniques such as early thrombolysis or coronary revascularization has greatly reduced HF-associated mortality, mortality with noncardiac causes, such as cancer, has increased. A significant proportion of HF patients develop cancer (7, 8), which has drawn attention to the biological HF–cancer intersection mechanisms. Cardiac oncology mostly focuses on the study of cancer therapy–related cardiac dysfunction (CTRCD). However, besides CTRCD, epidemiological and experimental studies suggest that HF may be a cancer susceptibility disease, leading to a new direction in cardiac oncology called reverse cardio-oncology (7, 12, 13, 18, 35). Recent evidence suggests that immune cell reprogramming may be a mechanism of HF-induced tumor growth. However, the associated pathophysiological mechanisms have yet to be elucidated (18). Therefore, in this study, we analyzed ischemic HF for immune cell infiltration and identified DDX58 as an important biomarker using

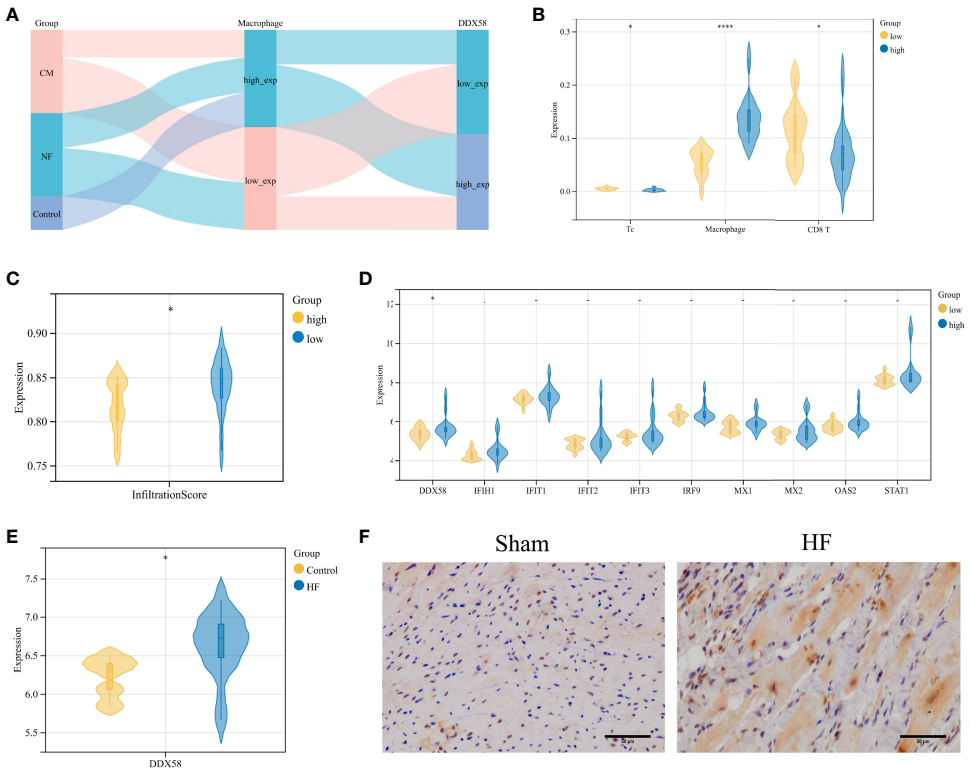



**FIGURE 5** | Validation of potential biomarkers for ischemic heart failure **(A)** Sankey diagram of GSE42955; **(B)** Comparison of Macrophage, Tc, CD8$^+$ T cell infiltration abundance in high and low Macrophage infiltration groups in GSE42955; **(C)** Differences in immune infiltration abundance in high Macrophage infiltration and low Macrophage infiltration groups; **(D)** Differences in expression levels of 10 genes in high Macrophage infiltration and low Macrophage infiltration abundance; **(E)** Expression level of DDX58 in GSE26887; **(F)** IHC staining of DDX58 in post-infarction heart failure group and sham-operated group (×400). *$P < 0.05$, ****$P < 0.0001$.

comprehensive bioinformatics. We also investigated DDX58 expression, prognosis, and immune cell infiltration in various cancer types through pan-cancer analysis, confirming that DDX58 is a potential therapeutic target for ischemic HF patients at an increased risk of developing cancer.

Our enrichment analysis of DEGs between ischemic HF and normal control samples suggests that they are significantly associated with biological processes or pathways related to the immune system or innate immune responses. The innate immune system includes immune cells and cytokines, which mediate the onset and development of inflammation. The innate immune responses are triggered by receptors widely expressed on the surfaces of immune cells or in cytoplasmic lysates (36). Previous studies have shown that immune cells such as macrophages, mast cells, monocytes, neutrophils, B cells, and T cells reside in or infiltrate cardiac tissue (37). We found significant differences in the expression of Tc, CD8$^+$ T cells, and macrophages between ischemic HF and normal tissue samples. T lymphocyte–mediated immune responses play an important role in HF. CD8$^+$ T cells systematically expand in chronic ischemic HF, and their depletion is independently associated with death (38–40). Moreover, in the cardiovascular system, macrophages maintain homeostasis by removing senescent cells and promoting angiogenesis (41).

Our WGCNA suggests that only the gene modules associated with macrophages are related to ischemic HF. Macrophages play an important role in tissue inflammation and wound healing and are thought to be a major contributor to the inflammatory and fibrotic processes in HF (42, 43). They release CC chemokines in myocardial infarction to recruit large numbers of monocytes and induce differentiation into macrophages for phagocytosis, a process that can be mediated by many proteins (44). In line with Wei et al. (45), we screened 10 central genes using protein–protein interaction network analysis and identified the only immune-related gene associated with ischemic HF: DDX58. DExD/H-box helicase 58 (DDX58, also known as RIG-I) is a protein involved in viral double-stranded RNA recognition and type I IFN production and was originally described as a key mediator of antiviral and innate immune responses (46). It has been reported that autophagy mediates the degradation of DDX58 (47). Autophagy is a biological process that employs phagocytosis to prevent pathogen invasion in host cells. Cells such as macrophages and neutrophils are specialized phagocytes (48). Mature phagocytes can be converted into autophagosomes

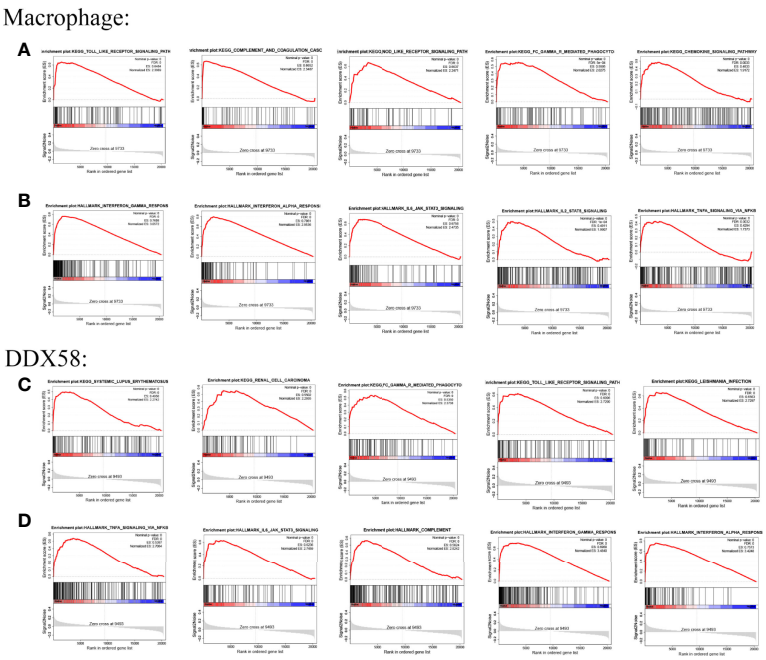



**FIGURE 6** | GSEA results of Macrophage and DDX58 **(A, B)** are the KEGG pathway and Hallmark pathway of Macrophage, respectively; **(C, D)** are the KEGG pathway and Hallmark pathway of DDX58, respectively.

that transport cargo to lysosomes for degradation (49). Autophagy can affect antiviral immune responses by selectively degrading downregulated immune factors—for example, by disrupting the interaction between DDX58 and LRRC25, which leads to the degradation of DDX58, thereby stabilizing the DDX58 protein to positively regulate DDX58-mediated type I IFN signaling (47, 50). Moreover, the RIG-I-like receptor signaling pathway is closely associated with monocyte infiltration, and its activation may be associated with inflammation and cardiomyocyte apoptosis (51, 52). However, there is no direct evidence of a direct relationship between DDX58 and the development of ischemic HF. Our study is the first to identify significantly high DDX58 expression in ischemic HF tissues and a high degree of macrophage infiltration.

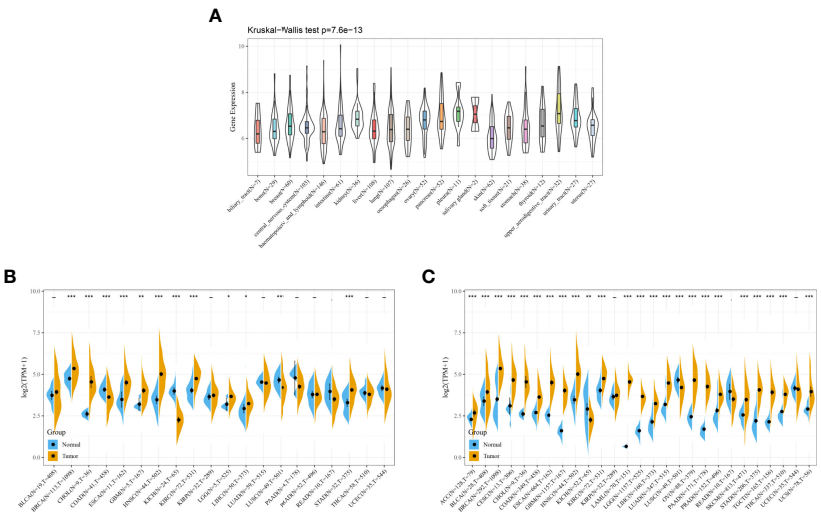



**FIGURE 7** | Expression of DDX58 in different cancer types **(A)** Expression level of DDX58 in different cancer cell lines; **(B)** Expression level of DDX58 in TCGA; **(C)** Expression level of DDX58 in the combination of GTEx database and TCGA. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

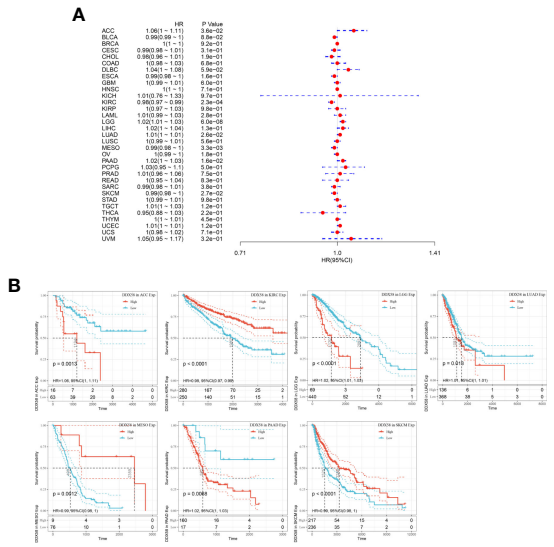


**FIGURE 8** | Relationship between DDX58 expression levels and overall survival of patients **(A)** Forest plot of the risk ratio of DDX58 in human pan-cancer; **(B)** Kaplan-Meier OS curves of DDX58 in the seven most significantly associated tumors.

Our pan-cancer analysis indicates significantly higher DDX58 expression in most cancer types than in normal tissues. Our survival analysis suggests that high DDX58 expression is significantly associated with poor OS, DSS, and PFI in three cancer types: LGG, LUAD, and PAAD. This suggests that DDX58 may act as an oncogene with varying prognostic significance in different cancer types and that DDX58-targeting therapy may provide prognostic benefits to cancer patients. Immune cell reprogramming may be a mechanism of HF-induced tumor growth (18). Our correlation analyses suggest that DDX58 expression is significantly associated with immune cell infiltration in various cancer types, with Macrophage_M1 being the most prominent, indicating that the monocyte/macrophage system may be a key mediator of DDX58 and immune cell infiltration in the tumor microenvironment. Evidence suggests that monocytes/macrophages are involved in tumor growth, metastasis, and tumor vascularization by regulating the tumor microenvironment and that they play a key role in accelerating breast cancer growth after myocardial infarction (18, 53). It can thus be concluded that the centralized regulation of the innate immune system in ischemic HF exerts multiple tumorigenic effects that may lead to myocardial carcinogenesis.

Our correlation analyses also suggest that DDX58 expression significantly correlates with the expression of 47 immune checkpoint genes in most cancer types. Although immunotherapy targeting various immune checkpoints has shown significant efficacy in some cancer types, it is effective only against some "hot" tumors with heavy T cell infiltration. Moreover, the high cardiotoxicity of immune checkpoint inhibitors limits their use. For these reasons, a novel therapeutic approach to enhancing tumor immunogenicity has recently been proposed, namely the activation of RIG-I-mediated innate immunity in the tumor microenvironment (21, 54). It has

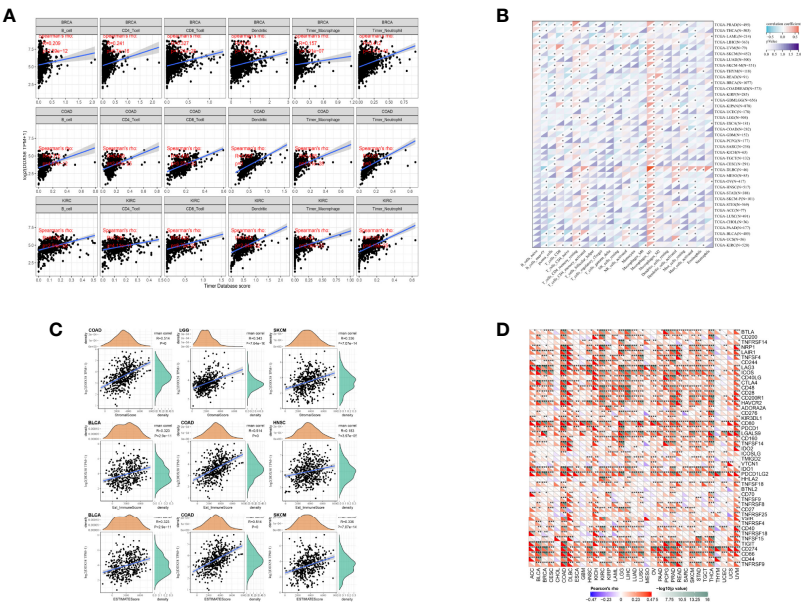


**FIGURE 9** | Correlation between DDX58 expression levels and cancer immune infiltration **(A)** Relationship between DDX58 expression levels and immune infiltration in the three most significantly associated tumors; **(B)** Heat map of the relationship between 22 immune cells and DDX58 in 33 tumors using CIBERSORT analysis; **(C)** Relationship between DDX58 expression levels and immune score, stromal score and ESTIMATE score in the three most significantly associated tumor types, respectively. **(D)** Correlation of DDX58 expression levels with immune checkpoint gene expression. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

also been reported that the RIG-I-mediated innate immune response may serve as a complementary mechanism for enhancing the anticancer efficacy of immune checkpoint therapies (55). All these pieces of evidence suggest that DDX58 is closely related to tumor immunotherapy.

In conclusion, this study shows that DDX58 is an important immune-related gene significantly associated with immune cell infiltration and tumor prognosis. Considering the role of DDX58 in ischemic HF and cancer, it can be hypothesized that the innate immune response is a potential mechanism of ischemic HF–induced tumor growth, in which DDX58-mediated monocyte/macrophage infiltration plays a crucial role. Therefore, immunotherapy targeting DDX58 may reduce the incidence of cancer in ischemic HF patients. Although the comprehensive analysis of DDX58 in this study is based on different databases and algorithms, there are still some limitations. First of all, DDX58 has not been reported in Ischemic HF, and this study only verifies the expression relationship between DDX58 and Ischemic HF, and the potential mechanism of DDX58 in Ischemic HF should be explored and verified in subsequent studies. Secondly, the role of DDX58 in cancer in this study is based on microarray data and bioinformatics analysis, and further *in vivo* or *in vitro* models should be established to elucidate the potential mechanism of DDX58 in Ischemic HF-induced tumor growth, and whether the innate immune response is a key link. Finally, prospective studies should be conducted in patients with Associated HF with cancer to evaluate the efficacy of DDX58 in reducing the cancer rate in patients with Ischemic HF.

In 2012, the European Society of Medical Oncology first published authoritative guidelines for the field of cardio-oncology (56), but in fact it was not until 2016 that the European Society of Cardiology and the American Society of Clinical Oncology issued two guidelines that the field of cardio-oncology had a real programmatic guideline (57, 58). To date, several countries around the world have actively set up independent diagnosis and treatment units of cardio-oncology, but cardio-oncology as a new and interdisciplinary discipline still faces serious challenges. Due to the lack of scientific and effective multidisciplinary diagnosis and treatment mechanism and related research data on cardio-oncology, clinicians face certain risks in the diagnosis and treatment process. The proposal of reverse cardio-oncology offers a whole new direction in cardiac oncology, and the active research on the potential pathophysiological mechanisms between heart failure and cancer and related epidemiological investigations will enrich the knowledge system of cardio-oncology and provide a theoretical basis and effective targets for the treatment of cardiac oncology diseases.

## 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 animal study was reviewed and approved by The Animal Care and Welfare Committee of Guangxi Medical University.

## AUTHOR CONTRIBUTIONS

All authors participated in the analysis process and wrote manuscripts of relevant sections. All authors have read and approved the publication of this study before submission.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.911309/full#supplementary-material

**Supplementary Figure 1 |** Relationship between DDX58 expression levels and disease-specific survival (DSS) **(A)** Forest plot of the risk ratio of DDX58 in human pan-cancer; **(B)** Kaplan-Meier DSS curves of DDX58 in the seven most significantly associated tumors.

**Supplementary Figure 2 |** Relationship between DDX58 expression levels and progression-free interval (PFI). **(A)** Forest plot of the risk ratio of DDX58 in human pan-cancer; **(B)** Kaplan-Meier PFI curves of DDX58 in the seven most significantly associated tumors.

## REFERENCES


1. Zhang Y, Bauersachs J, Langer HF. Immune Mechanisms in Heart Failure. *Eur J Heart Failure* (2017) 19(11):1379–89. doi: 10.1002/ejhf.942
2. Spencer LJ, Degu A, Kalkidan Hassen A, Solomon MA, Cristiana A, Nooshin A, et al. Global, Regional, and National Incidence, Prevalence, and Years Lived With Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990-2017: A Systematic Analysis for the Global Burden of Disease Study 2017. *Lancet* (2018) 392:1789–858. doi: 10.1016/S0140-6736(18)32279-7
3. Christiansen MN, Køber L, Weeke P, Vasan RS, Jeppesen JL, Smith JG, et al. Age-Specific Trends in Incidence, Mortality, and Comorbidities of Heart Failure in Denmark, 1995 to 2012. *Circulation* (2017) 135(13):1214–23. doi: 10.1161/CIRCULATIONAHA.116.025941
4. Dokainish H, Teo K, Zhu J, Roy A, AlHabib KF, ElSayed A, et al. Heart Failure in Africa, Asia, the Middle East and South America: The INTER-CHF Study. *Int J Cardiol* (2016) 204:133–41. doi: 10.1016/j.ijcard.2015.11.183
5. Bocchi EA. Heart Failure in South America. *Curr Cardiol Rev* (2013) 9(2):147–56. doi: 10.2174/1573403x11309020007
6. Sulaiman K, Panduranga P, Al-Zakwani I, Alsheikh-Ali AA, AlHabib KF, Al-Suwaidi J, et al. Clinical Characteristics, Management, and Outcomes of Acute Heart Failure Patients: Observations From the Gulf Acute Heart Failure Registry (Gulf CARE). *Eur J Heart Fail* (2015) 17(4):374–84. doi: 10.1002/ejhf.245
7. Meijers WC, Maglione M, Bakker SJL, Oberhuber R, Kieneker LM, de Jong S, et al. Heart Failure Stimulates Tumor Growth by Circulating Factors. *Circulation* (2018) 138(7):678–91. doi: 10.1161/CIRCULATIONAHA.117.030816

8. de Wit S, de Boer RA. From Studying Heart Disease and Cancer Simultaneously to Reverse Cardio-Oncology. *Circulation* (2021) 144(2):93–5. doi: 10.1161/CIRCULATIONAHA.120.053315
9. Murphy KT. The Pathogenesis and Treatment of Cardiac Atrophy in Cancer Cachexia. *Am J Physiol Heart Circ Physiol* (2016) 310(4):H466–77. doi: 10.1152/ajpheart.00720.2015
10. Karlstaedt A, Zhang X, Vitrac H, Harmancey R, Vasquez H, Wang JH, et al. Oncometabolite D-2-Hydroxyglutarate Impairs α-Ketoglutarate Dehydrogenase and Contractile Function in Rodent Heart. *Proc Natl Acad Sci U S A* (2016) 113(37):10436–41. doi: 10.1073/pnas.1601650113
11. Bertero E, Ameri P, Maack C. Bidirectional Relationship Between Cancer and Heart Failure: Old and New Issues in Cardio-Oncology. *Card Fail Rev* (2019) 5(2):106–11. doi: 10.15420/cfr.2019.1.2
12. Hasin T, Gerber Y, Weston SA, Jiang R, Killian JM, Manemann SM, et al. Heart Failure After Myocardial Infarction Is Associated With Increased Risk of Cancer. *J Am Coll Cardiol* (2016) 68(3):265–71. doi: 10.1016/j.jacc.2016.04.053
13. Banke A, Schou M, Videbaek L, Møller JE, Torp-Pedersen C, Gustafsson F, et al. Incidence of Cancer in Patients With Chronic Heart Failure: A Long-Term Follow-Up Study. *Eur J Heart Fail* (2016) 18(3):260–6. doi: 10.1002/ejhf.472
14. Hasin T, Gerber Y, McNallan SM, Weston SA, Kushwaha SS, Nelson TJ, et al. Patients With Heart Failure Have an Increased Risk of Incident Cancer. *J Am Coll Cardiol* (2013) 62(10):881–6. doi: 10.1016/j.jacc.2013.04.088
15. Ameri P, Canepa M, Anker MS, Belenkov Y, Bergler-Klein J, Cohen-Solal A, et al. Cancer Diagnosis in Patients With Heart Failure: Epidemiology, Clinical Implications and Gaps in Knowledge. *Eur J Heart Fail* (2018) 20(5):879–87. doi: 10.1002/ejhf.1165
16. Zhang Y, Wen W, Liu H. The Role of Immune Cells in Cardiac Remodeling After Myocardial Infarction. *J Cardiovasc Pharmacol* (2020) 76(4):407–13. doi: 10.1097/FJC.0000000000000876
17. Gonzalez H, Hagerling C, Werb Z. Roles of the Immune System in Cancer: From Tumor Initiation to Metastatic Progression. *Genes Dev* (2018) 32(19-20):1267–84. doi: 10.1101/gad.314617.118
18. Koelwyn GJ, Newman AAC, Afonso MS, van Solingen C, Corr EM, Brown EJ, et al. Myocardial Infarction Accelerates Breast Cancer *via* Innate Immune Reprogramming. *Nat Med* (2020) 26(9):1452–58. doi: 10.1038/s41591-020-0964-7
19. Zhu H, Ivanovic M, Nguyen A, Nguyen PK, Wu SM. Immune Checkpoint Inhibitor Cardiotoxicity: Breaking Barriers in the Cardiovascular Immune Landscape. *J Mol Cell Cardiol* (2021) 160:121–7. doi: 10.1016/j.yjmcc.2021.07.006
20. Salem JE, Manouchehri A, Moey M, Lebrun-Vignes B, Bastarache L, Pariente A, et al. Cardiovascular Toxicities Associated With Immune Checkpoint Inhibitors: An Observational, Retrospective, Pharmacovigilance Study. *Lancet Oncol* (2018) 19(12):1579–89. doi: 10.1016/S1470-2045(18)30608-9
21. Elion DL, Cook RS. Harnessing RIG-I and Intrinsic Immunity in the Tumor Microenvironment for Therapeutic Cancer Treatment. *Oncotarget* (2018) 9(48):29007–17. doi: 10.18632/oncotarget.25626
22. Rihawi K, Ricci AD, Rizzo A, Brocchi S, Marasco G, Pastore LV, et al. Tumor-Associated Macrophages and Inflammatory Microenvironment in Gastric Cancer: Novel Translational Implications. *Int J Mol Sci* (2021) 22(8):3805. doi: 10.3390/ijms22083805
23. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia Enables Predictive Modelling of Anticancer Drug Sensitivity. *Nature* (2012) 483(7391):603–7. doi: 10.1038/nature11003
24. GTEx Consortium. The Genotype-Tissue Expression (GTEx) Project. *Nat Genet* (2013) 45(6):580–5. doi: 10.1038/ng.2653
25. Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, et al. InnateDB: Systems Biology of Innate Immunity and Beyond–Recent Updates and Continuing Curation. *Nucleic Acids Res* (2013) 41(Database issue):D1228–33. doi: 10.1093/nar/gks1147
26. Langfelder P, Horvath S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559
27. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6
28. 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
29. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res* (2017) 77(21):e108–10. doi: 10.1158/0008-5472.CAN-17-0307
30. Liu Z, Lu T, Li J, Wang L, Xu K, Dang Q, et al. Clinical Significance and Inflammatory Landscape of Anovel Recurrence-Associated Immune Signature in Stage II/III Colorectal Cancer. *Front Immunol* (2021) 12:702594. doi: 10.3389/fimmu.2021.702594
31. 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
32. Cao Y, Tang W, Tang W. Immune Cell Infiltration Characteristics and Related Core Genes in Lupus Nephritis: Results From Bioinformatic Analysis. *BMC Immunol* (2019) 20(1):37. doi: 10.1186/s12865-019-0316-x
33. Martini E, Kunderfranco P, Peano C, Carullo P, Cremonesi M, Schorn T, et al. Single-Cell Sequencing of Mouse Heart Immune Infiltrate in Pressure Overload-Driven Heart Failure Reveals Extent of Immune Activation. *Circulation* (2019) 140(25):2089–107. doi: 10.1161/CIRCULATIONAHA.119.041694
34. Elgendy IY, Mahtta D, Pepine CJ. Medical Therapy for Heart Failure Caused by Ischemic Heart Disease. *Circ Res* (2019) 124(11):1520–35. doi: 10.1161/CIRCRESAHA.118.313568
35. Avraham S, Abu-Sharki S, Shofti R, Haas T, Korin B, Kalfon R, et al. Early Cardiac Remodeling Promotes Tumor Growth and Metastasis. *Circulation* (2020) 142(7):670–83. doi: 10.1161/CIRCULATIONAHA.120.046471
36. Wang Y, Liu X, Shi H, Yu Y, Yu Y, Li M, et al. NLRP3 Inflammasome, an Immune-Inflammatory Target in Pathogenesis and Treatment of Cardiovascular Diseases. *Clin Transl Med* (2020) 10(1):91–106. doi: 10.1002/ctm2.13
37. Swirski FK, Nahrendorf M. Cardioimmunology: The Immune System in Cardiac Homeostasis and Disease. *Nat Rev Immunol* (2018) 18(12):733–44. doi: 10.1038/s41577-018-0065-8
38. Gröschel C, Sasse A, Monecke S, Röhrborn C, Elsner L, Didié M, et al. $CD8^+$-T Cells With Specificity for a Model Antigen in Cardiomyocytes Can Become Activated After Transverse Aortic Constriction But Do Not Accelerate Progression to Heart Failure. *Front Immunol* (2018) 9:2665. doi: 10.3389/fimmu.2018.02665
39. Fukunaga T, Soejima H, Irie A, Sugamura K, Oe Y, Tanaka T, et al. Relation Between $CD4^+$T-Cell Activation and Severity of Chronic Heart Failure Secondary to Ischemic or Idiopathic Dilated Cardiomyopathy. *Am J Cardiol* (2007) 100(3):483–8. doi: 10.1016/j.amjcard.2007.03.052
40. Yazdanyar A, Aziz MM, Enright PL, Edmundowicz D, Boudreau R, Sutton-Tyrell K, et al. Association Between 6-Minute Walk Test and All-Cause Mortality, Coronary Heart Disease-Specific Mortality, and Incident Coronary Heart Disease. *J Aging Health* (2014) 26(4):583–99. doi: 10.1177/0898264314525665
41. Sager HB, Kessler T, Schunkert H. Monocytes and Macrophages in Cardiac Injury and Repair. *J Thorac Dis* (2017) 9(Suppl 1):S30–5. doi: 10.21037/jtd.2016.11.17
42. Wrigley BJ, Lip GY, Shantsila E. The Role of Monocytes and Inflammation in the Pathophysiology of Heart Failure. *Eur J Heart Fail* (2011) 13(11):1161–71. doi: 10.1093/eurjhf/hfr122
43. Wynn TA, Barron L. Macrophages: Master Regulators of Inflammation and Fibrosis. *Semin Liver Dis* (2010) 30(3):245–57. doi: 10.1055/s-0030-1255354
44. Duncan SE, Gao S, Sarhene M, Coffie JW, Linhua D, Bao X, et al. Macrophage Activities in Myocardial Infarction and Heart Failure. *Cardiol Res Pract* (2020) 2020:4375127. doi: 10.1155/2020/4375127
45. Wei X, Zhang A, Yang W, Fang Y. Depletion of β3-Adrenergic Receptor Relieves Pressure Overload-Induced Cardiac Hypertrophy and Heart Failure *via* Enhancing Innate Immune Response. *BioMed Pharmacother* (2021) 143:112194. doi: 10.1016/j.biopha.2021.112194
46. Morelli M, Galluzzo M, Madonna S, Scarponi C, Scaglione GL, Galluccio T, et al. HLA-Cw6 and Other HLA-C Alleles, as Well as MICB-DT, DDX58, and TYK2 Genetic Variants Associate With Optimal Response to Anti-IL-17A

Treatment in Patients With Psoriasis. *Expert Opin Biol Ther* (2021) 21(2):259–70. doi: 10.1080/14712598.2021.1862082
47. Xian H, Yang S, Jin S, Zhang Y, Cui J. LRRC59 Modulates Type I Interferon Signaling by Restraining the SQSTM1/p62-Mediated Autophagic Degradation of Pattern Recognition Receptor DDX58/RIG-I. *Autophagy* (2020) 16(3):408–18. doi: 10.1080/15548627.2019.1615303
48. Choi Y, Bowman JW, Jung JU. Autophagy During Viral Infection - a Double-Edged Sword. *Nat Rev Microbiol* (2018) 16(6):341–54. doi: 10.1038/s41579-018-0003-6
49. Hansen M, Rubinsztein DC, Walker DW. Autophagy as a Promoter of Longevity: Insights From Model Organisms. *Nat Rev Mol Cell Biol* (2018) 19(9):579–93. doi: 10.1038/s41580-018-0033-y
50. Jin S, Tian S, Luo M, Xie W, Liu T, Duan T, et al. Tetherin Suppresses Type I Interferon Signaling by Targeting MAVS for NDP52-Mediated Selective Autophagic Degradation in Human Cells. *Mol Cell* (2017) 68(2):308–22.e4. doi: 10.1016/j.molcel.2017.09.005
51. Ma L, Shi W, Ma X, Zou M, Chen W, Li W, et al. Comprehensive Analysis of Differential Immunocyte Infiltration and the Potential ceRNA Networks During Epicardial Adipose Tissue Development in Congenital Heart Disease. *J Transl Med* (2020) 18(1):111. doi: 10.1186/s12967-020-02279-y
52. Chattopadhyay S, Sen GC. RIG-I-Like Receptor-Induced IRF3 Mediated Pathway of Apoptosis (RIPA): A New Antiviral Pathway. *Protein Cell* (2017) 8(3):165–8. doi: 10.1007/s13238-016-0334-x
53. Engblom C, Pfirschke C, Pittet MJ. The Role of Myeloid Cells in Cancer Therapies. *Nat Rev Cancer* (2016) 16(7):447–62. doi: 10.1038/nrc.2016.54
54. Rameshbabu S, Labadie BW, Argulian A, Patnaik A. Targeting Innate Immunity in Cancer Therapy. *Vaccines (Basel)* (2021) 9(2):138. doi: 10.3390/vaccines9020138
55. Heidegger S, Wintges A, Stritzke F, Bek S, Steiger K, Koenig PA, et al. RIG-I Activation is Critical for Responsiveness to Checkpoint Blockade. *Sci Immunol* (2019) 4(39):eaau8943. doi: 10.1126/sciimmunol.aau8943
56. Curigliano G, Cardinale D, Suter T, Plataniotis G, de Azambuja E, Sandri MT, et al. Cardiovascular Toxicity Induced by Chemotherapy, Targeted Agents and Radiotherapy: ESMO Clinical Practice Guidelines. *Ann Oncol* (2012) 23 Suppl 7:vii155–66. doi: 10.1093/annonc/mds293
57. Zamorano JL, Lancellotti P, Rodriguez Muñoz D, Aboyans V, Asteggiano R, Galderisi M, et al. ESC Position Paper on Cancer Treatments and Cardiovascular Toxicity Developed Under the Auspices of the ESC Committee for Practice Guidelines: The Task Force for Cancer Treatments and Cardiovascular Toxicity of the European Society of Cardiology (ESC). *Eur Heart J* (2016) 37(36):2768–801. doi: 10.1093/eurheartj/ehw211
58. Armenian SH, Lacchetti C, Barac A, Carver J, Constine LS, Denduluri N, et al. Prevention and Monitoring of Cardiac Dysfunction in Survivors of Adult Cancers: American Society of Clinical Oncology Clinical Practice Guideline. *J Clin Oncol* (2017) 35(8):893–911. doi: 10.1200/JCO.2016.70.5400

**Conflict of Interest:** YZ is employed by Biotrans Technology Co., LTD.

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 01 July 2022
doi: 10.3389/fonc.2022.897495

# Colorectal Cancer Chemotherapy Drug Bevacizumab May Induce Muscle Atrophy Through CDKN1A and TIMP4

*Qun Xu, Jinyou Li, Yue Wu, Wenjing Zhou and Zherong Xu**

*Department of Geriatrics, The First Affiliated Hospital, Zhejiang University, School of Medicine, Hangzhou, China*

The muscle in the organism has the function of regulating metabolism. Long-term muscle inactivity or the occurrence of chronic inflammatory diseases are easy to induce muscle atrophy. Bevacizumab is an antiangiogenic drug that prevents the formation of neovascularization by inhibiting the activation of VEGF signaling pathway. It is used in the first-line treatment of many cancers in clinic. Studies have shown that the use of bevacizumab in the treatment of tumors can cause muscle mass loss and may induce muscle atrophy. Based on bioinformatics analysis, this study sought the relationship and influence mechanism between bevacizumab and muscle atrophy. The differences of gene and sample expression between bevacizumab treated group and control group were studied by RNA sequencing. WGCNA is used to find gene modules related to bevacizumab administration and explore biological functions through metascape. Differential analysis was used to analyze the difference of gene expression between the administration group and the control group in different muscle tissues. The key genes timp4 and CDKN1A were obtained through Venn diagram, and then GSEA was used to explore their biological functions in RNA sequencing data and geo chip data. This study studied the role of bevacizumab in muscle through the above methods, preliminarily determined that timp4 and CDKN1A may be related to muscle atrophy, and further explored their functional mechanism in bevacizumab myotoxicity.

**Keywords: bevacizumab, sarcopenia, TIMP4, CDKN1A, muscles**

OPEN ACCESS

***Edited by:***
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

***Reviewed by:***
*Hailun Zhou,*
*Guangxi Medical University, China*
*Xiao Feng Liao,*
*Hefei University of Technology, China*

***Correspondence:**
*Zherong Xu*
*xzr1978@zju.edu.cn*

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

***Received:*** *16 March 2022*
***Accepted:*** *04 May 2022*
***Published:*** *01 July 2022*

***Citation:***
*Xu Q, Li J, Wu Y, Zhou W and Xu Z (2022) Colorectal Cancer Chemotherapy Drug Bevacizumab May Induce Muscle Atrophy Through CDKN1A and TIMP4. Front. Oncol. 12:897495. doi: 10.3389/fonc.2022.897495*

## INTRODUCTION

Muscle is a major component of lean body mass and plays a vital role in maintaining health. It has been shown that there is a direct or indirect relationship between muscle and strength, energy, mobility, skeletal support, balance, wound healing, immune function, digestive function, and skin health (1).Muscle mass begins to decline from the age of 40 at a rate of 6% per decade and accelerates to a rate of 25–40% per decade above age 70 years. The muscle loss related to age is sarcopenia (2).Sarcopenia in the older adult has been associated with functional impairments, disability, increased risk of falls and fractures, reduced health-related quality of life, and increased risk of death (3). Sarcopenia is considered "primary" when there is no obvious other specific cause,

while sarcopenia is considered "secondary" when other causes other than aging are obvious. Sarcopenia can be secondary to systemic diseases, especially those that can trigger inflammation processes, such as malignant tumors or organ failure (4).Sarcopenia in the oncology setting is an area of growing research interest which is associated with many adverse consequences. Its prevalence is high in adult cancer patients, ranging from 11% to 74% in all adults, and even higher in elderly cancer patients (5).

In many clinical studies, skeletal muscle mass (SMM) changes in cancer patients have been extensively investigated through the use of diagnosis, staging, or follow-up computed tomography (CT) scans. Multiple studies have shown that, among cancer patients, the proportion of patients with sarcopenia increased significantly and the skeletal muscle volume and density was lost after treatment with cytoskeletal disruptors (taxanes), nucleotide analogs (gemcitabine) or kinase inhibitors (bevacizumab), or neoadjuvant chemotherapy (6–10).In particular,colorectal cancer patients prescribed bevacizumab appear to lose weight and muscle over a few months even in the absence of cancer progression (11).In colorectal cancer patients with metastatic disease under bevacizumab-based chemotherapy treatment, Adeline et al. assessed changes in muscle mass over an interval of 70 days and indicated 47% patients showed loss of SMM (12). In metastatic colorectal cancer, significant muscle loss occurred in patients with bevacizumab + capecitabine + oxaliplatin-based combination chemotherapy(CAPOX-B) (13, 14). Collectively, the literature supports a direct association between chemotherapy, especially bevacizumab and muscle atrophy.

Bevacizumab (Avastin®), a VEGF-A-targeting monoclonal antibody, is the first FDA-approved anti-angiogenic drug in clinical practice. Bevacizumab prevents the interaction of VEGF-A with VEGFR by binding to VEGF-A, inhibiting the activation of VEGF signaling pathway, thereby preventing the formation of new blood vessels (15).Bevacizumab inhibits the growth of human tumor cell lines because of this mechanism. In addition to the above application in patients with colorectal cancer, bevacizumab is also used in the treatment of many other cancers.In the treatment of malignant pleural mesothelioma, bevacizumab combined with pemetrexed and cisplatin can significantly improve the overall survival of patients (16). When combined with paclitaxel, it can improve the curative effect of pleural effusion in non-small cell lung cancer (17).In oral squamous cell carcinoma, bevacizumab was found to downregulate biomarker expression and promote cancer cell apoptosis (18). Given its widespread use in cancer treatment, the cause of bevacizumab-induced sarcopenia needs to be explored urgently.

Mechanisms of action of chemotherapies to induce sarcopenia are various.Chemotherapeutic drugs could induce the expression of pro-inflammatory cytokines, active the myostatin pathway, cause mitochondrial damage and induce oxidative stress.They increase protein degradation by activating ubiquitin-proteasome pathway and autophagy-lysosome pathway, and inhibit protein synthesis by silencing the IGF-1/PI3K/Akt/mTOR anabolic pathway (2). As a kinase inhibitor, bevacizumab inhibits the vascular endothelial growth factor receptor and disrupts many of skeletal muscle pathways, such as PI3K and AKT (11, 19). In this study, we will use bioinformatics approaches to find potential genes associated with muscle atrophy by mouse models with bevacizumab and obtain sequencing information,so as to provide useful help for the clinical use of bevacizumab.

# METHODS AND MATERIALS

## Data Source

12 mice were tested to obtain a set of sequencing data. Six of the mice were treated with bevacizumab, and the other six were used as the control group. The gastrocnemius muscle and soleus muscle were collected from 6 cases each, and the muscle tissues were collected for RNA-seq sequencing after the same treatment. Finally, MutilQC (https://multiqc.info/) was used to normalize and summarize the data in different samples. Two datasets, GSE38417 and GSE6011, were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), containing 22 and 37 samples, respectively. Merge GSE38417 and GSE6011 to get a new dataset new_merge.

## Principal Components Analysis

In order to verify the rationality of the data of the 12 mice, PCA was used for cluster analysis to observe the repeatability of the samples. R-packet pheatmap is used to observe the distribution of each sample and gene.

## Weighted Gene Co-Expression Network Analysis

The co-expression of genes in the expression profile was analyzed by R-package WGCNA. The optimal soft threshold is obtained according to the scale-free network(β).According to the similar expression, the genes are divided into different modules, and each color contains different genes. Then analyze the relationship between these gene modules and phenotype, and select the gene module with the most significant correlation with phenotype for subsequent analysis.

## Differential Expression Analysis

R-packet limma packet was used to analyze the differential expression of genes. 12 mouse samples were divided into two independent gene sets GM and SM according to muscle type, and the differences were analyzed respectively to find the genes differentially expressed in the administration group and the control group.| Log2fc | < 1, $P < 0.05$ was defined as the screening threshold of differential genes.

## Gene Expression

T test was used to analyze the expression of genes between different samples, and R package ggplot2 was used to show the expression of genes. Correlations between genes were analyzed by spearman and plotted by the R package ggstatsplot. In addition, Venn map was used to analyze intersecting genes in different datasets.

## Function Cluster Analysis

Metascape was used to analyze the functions and pathways of key module genes to further understand the relationship between genes and diseases. Single Gene Set Enrichment Analysis (GSEA) was used to analyze the Hallmark pathway and Kyoto Encyclopedia of Genes (KEGG) pathway for the involvement of hub genes in muscle atrophy. The screening conditions for important pathways were |NES|<1, NOM P-val<0.05, FDR q-val<0.25.

# RESULTS

## Verification of Sequencing Data

The PCA results showed that the samples in the control group and the administration group could be clearly separated, and the gastrocnemius and soleus muscles could also be clearly separated (**Figure 1A**). Subsequent studies found that both gene correlations and expression distributions were differentiated by sample type (**Figures 1B, C**). Data were normalized for facilitating subsequent studies(**Figure 1D**).

## WGCNA

The optimal soft threshold was determined to be 5 based on standard analysis of scale-free networks (**Figure 2A**). 16 color modules were obtained by WGCNA analysis, and each module contained different genes (**Figure 2B**). By analyzing the correlation between gene modules and traits, it was found that lightcyan1 (cor=0.70, P=0.01), floralwhite (cor=0.83, P=7.7e-4), darkgrey (cor=0.84, P=6.1e-4) had the strongest correlation with bevacizumab administration group (**Figure 2C**). These three gene modules were selected as key modules, containing a total of 280 genes.

## Metascape

As shown in **Figure 3A**,there was the protein-protein interaction network of 280 genes (**Figure 3A**). These genes were subsequently subjected to functional clustering analysis, and most of the genes were enriched in GO functional terms such as negative regulation of protein modification processes, vascular morphogenesis, and negative regulation of cell differentiation (**Figure 3B**). Among the enriched functional terms, 4 functional terms were related to proteolytic pathways. The functional terms related to the proteolytic pathway and the number of genes included in each term are shown in **Table 1**, and there are 39 genes in total.

## Differential Expression Analysis

The differential expression analysis between the control group and the administration group in the gastrocnemius muscle showed that 363 genes were down-regulated and 217 genes were up-regulated (**Figure 4A**). At the same time, the difference analysis between the control group and the administration group in the

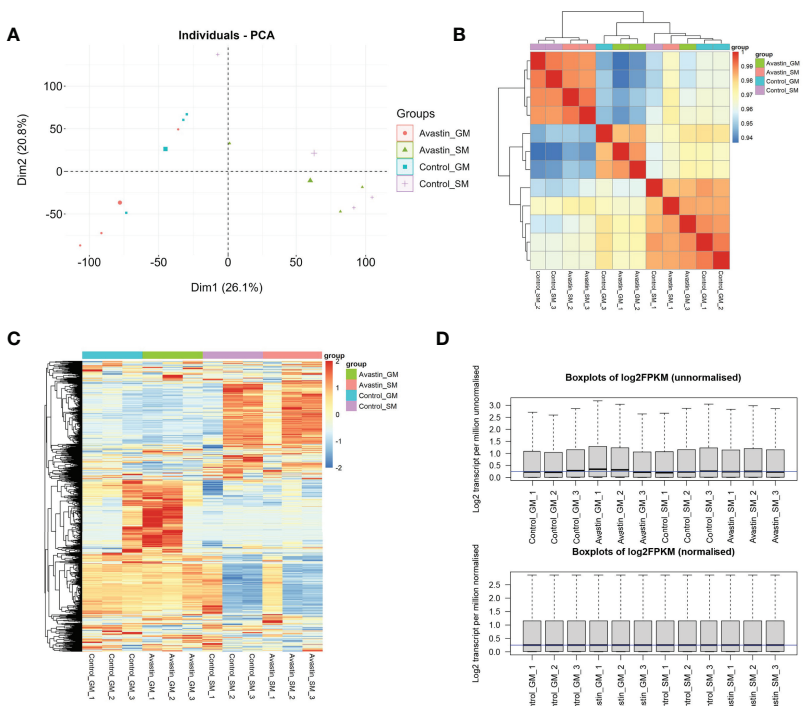



**FIGURE 1** | Validation of sequencing data. **(A)** PCA analysis of data distribution in 12 samples; **(B)** correlation analysis of genes in each sample; **(C)** heat map of gene distribution in samples; **(D)** normalization results of samples.

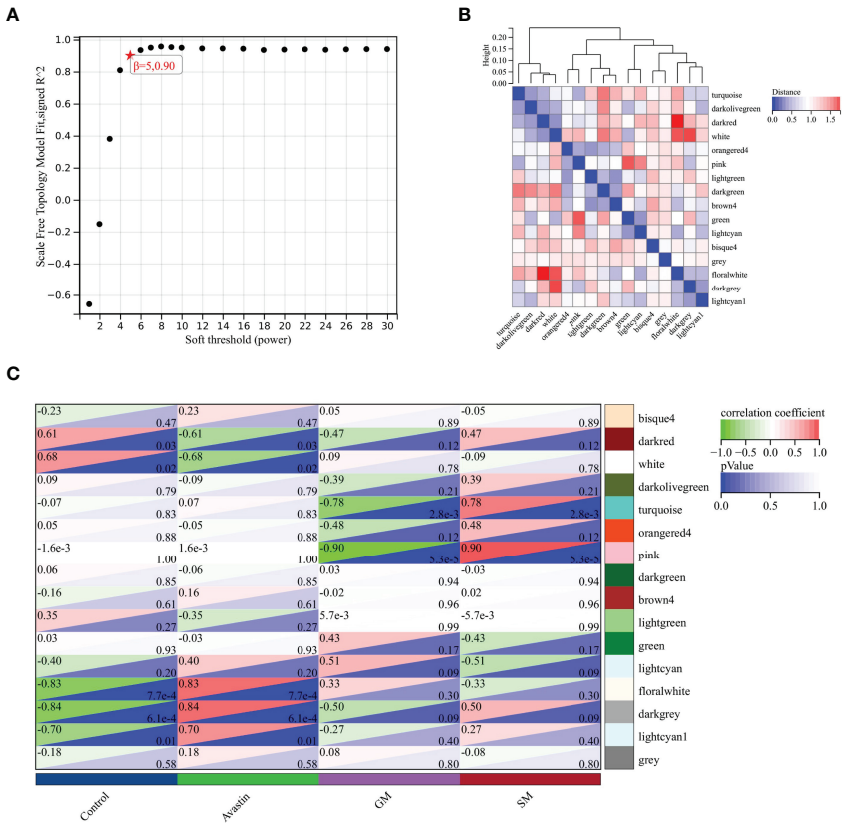



**FIGURE 2** | Selection of key gene modules. **(A)** Determination of soft threshold; **(B)** Heat map of module feature vector clustering; **(C)** Heat map of correlation between gene modules and traits.

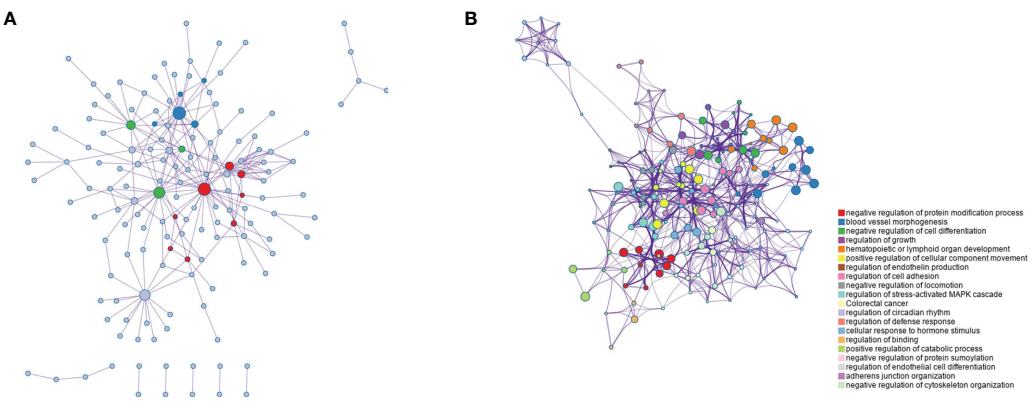



**FIGURE 3** | Function clustering analysis of genes in key modules. **(A)** Co-expression network diagram of 280 key genes; **(B)** Rich term network of genes, showing the top 20 functional terms.

soleus muscle showed that 295 genes were down-regulated and 236 genes were up-regulated (**Figure 4C**). The expressions of up-regulated and down-regulated genes are shown in **Figures 4B, D**, and there are significant differences.

## Gene Expression

The intersection of key gene modules, GM_DEG, SM_DEG and genes related to the proteolytic pathway in WGCNA, two genes Timp4 and Cdkn1a were obtained (**Figure 5A**). Vegfa is the

**TABLE 1** | Functional terms and number of genes related to MAPK pathway in functional clustering results.

| Category | Description | Counts | Hits |
|---|---|---|---|
| GO (BP) | negative regulation of protein modification process | 25 | Ager\|Casp3\|Ctnnb1\|Cdkn1a\|Dmtn\|Fkbp8\|Hhex\|Hmg20b\|Irf1\|Per2\|Inpp5k\|Sin3a\|Xrcc1\|Gadd45g\|Sh3bp5\|Ceacam1\|Tinf2\|Ctdsp2\|Niban1\|Dusp10\|Errfi1\|Rasd2\|Adarb1\|Pias3\|Kdm4a |
| GO (BP) | monoubiquitinated protein deubiquitination | 3 | Taf10\|Sf3b5\|Asxl1 |
| GO (BP) | positive regulation of protein catabolic process | 8 | App\|Plk2\|Tnfsf12\|Apc2\|Tiparp\|Rnf144a\|Lpcat1\|Trib2 |
| GO (BP) | regulation of protein catabolic process | 11 | App\|Plk2\|Tnfsf12\|Apc2\|Gabarapl2\|Tiparp\|Rnf144a\|Timp4\|Lpcat1\|Trib2\|Ccar2 |

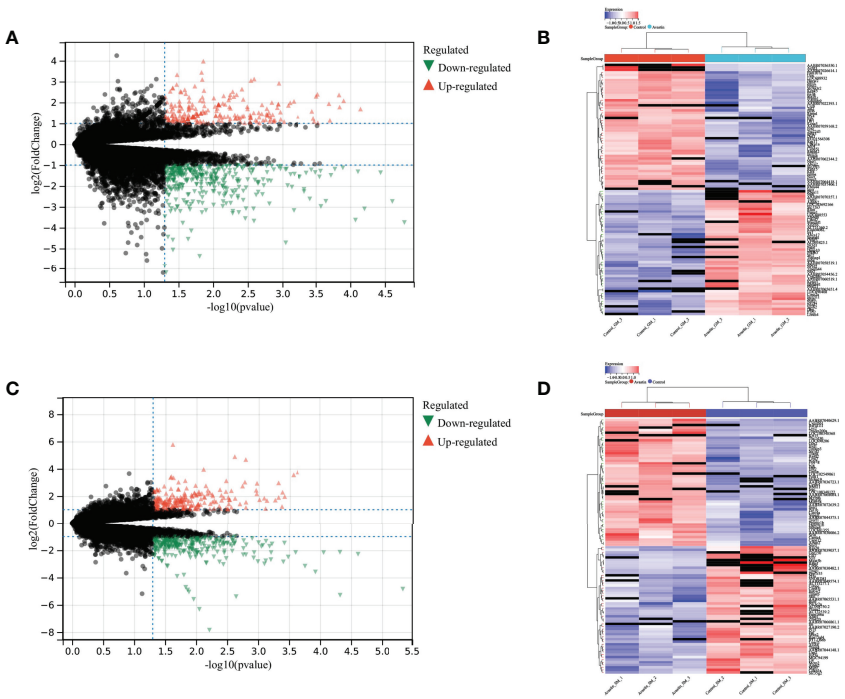



**FIGURE 4** | Differential expression analysis results. Differential expression analysis of genes in gastrocnemius muscle, with **(A)** volcano plot showing the differential expression results of genes, **(B)** heat map showing the distribution of up- and down-regulated genes; Differential expression analysis of genes in soleus muscle, with **(C)** volcano plot showing the differential expression results of genes, and **(D)** the heat map showing the distribution of up- and down-regulated genes.

primary target of bevacizumab. Correlation analysis with Vegfa found that both Timp4 and Cdkn1a were positively correlated with Vegfa (**Figures 5B, C**). The expression comparison showed that Timp4 and Cdkn1a were significantly overexpressed in the control group, but significantly underexpressed in the administration group (**Figures 5D, E**). At the same time, by comparing the expression of Timp4 and Cdkn1a in GSE38417, the results showed the opposite results. Timp4 was underexpressed in the muscle atrophy group, whereas Cdkn1a was overexpressed in the muscle atrophy samples (**Figures 5F–I**).

## Functional Pathways

According to the expression of Timp4 and Cdkn1a, the samples were divided into two groups: high expression and low expression, and GSEA was used to analyze the biological signal pathways of gene enrichment. The results showed that Cdkn1a was significantly enriched in KEGG pathways related to ribosomes, antigen processing and presentation, glutathione metabolism,and spliceosome (**Figure 6A**),as well as Hallmark pathways such as reactive oxidation pathways, WNT_β_catenin_signaling,apoptosis, MYC_target_V1,and oxidative phosphorylation(**Figure 6B**). Timp4 was significantly enriched in Hallmark pathways such as oxidative phosphorylation, MYC_target_V1, apoptosis, reactive oxidativepathways,and angiogenesis (**Figure 6C**), as well as KEGG pathways such as ribosomes, oxidative phosphorylation, Parkinson's disease, Huntington's disease, Alzheimer's disease and Hemer's disease (**Figure 6D**). In the mergeddataset new_merge, Cdkn1a was significantly enriched in Hallmark pathways such as epithelial-mesenchymal transition, TNFA_signaling through NFKB, apoptosis, inflammation and allograft rejection-related pathways

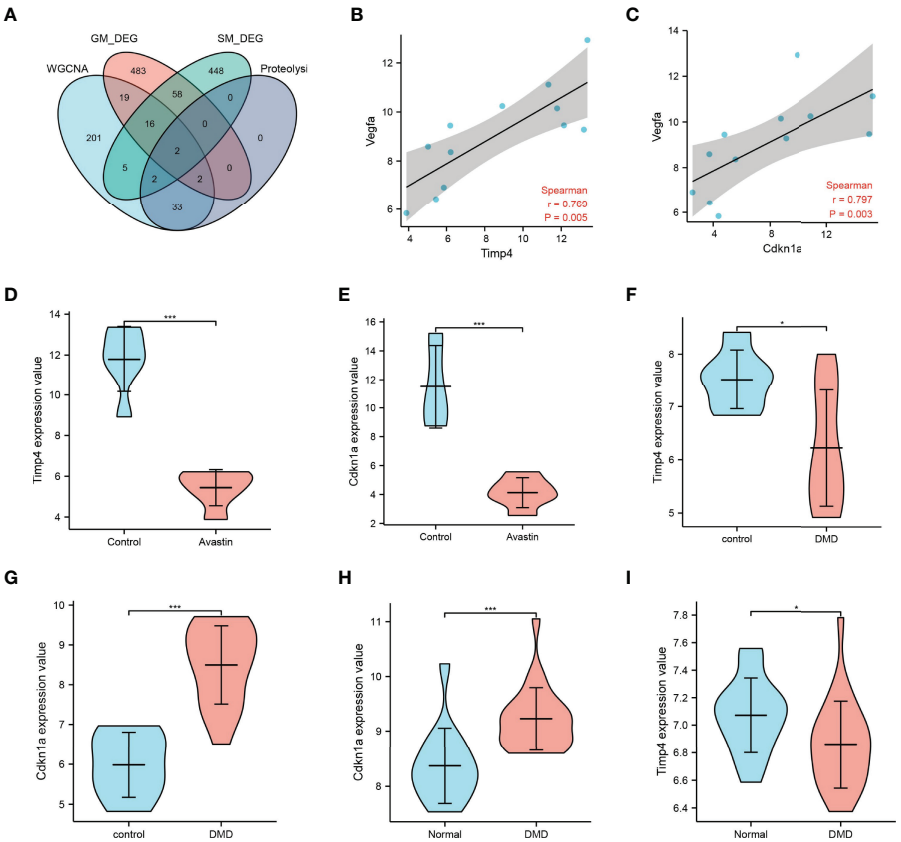



**FIGURE 5** | Map3k6 expression. **(A)** Venn map showing the intersection of different datasets; correlation of **(B)** Timp4 and **(C)** Cdkn1a with Vegfa; **(D)** Timp4 and **(E)** Cdkn1a expression in drug-treated and control groups; **(F)** Timp4 and **(G)** Cdkn1a expression in different groups in GSE38417; **(H)** Timp4 and **(I)** Cdkn1a expression in different groups in GSE6011. *p < 0.05,***p < 0.001.

(**Figure 6E**), as well as KEGG pathways such as the complement and condensation cascades, sphingolipid metabolism, P53_signaling pathway and lysosome(**Figure 6F**). Timp4 was significantly enriched in Hallmark pathways such as inflammation and inflammatory factor signaling pathways, androgen response and KRAS signaling (**Figure 6G**), and KEGG pathways such as PPAR_signaling pathway, cytokine receptor interaction, pyruvate metabolism, and citrate cycle (**Figure 6H**).

## DISCUSSION

Muscle atrophy is an important manifestation of sarcopenia. Previous studies have shown that pathway changes related to protein ubiquitination and energy production are common features during muscle atrophy (20). In addition, recent studies have also shown that mitochondrial dysfunction (21), inflammatory pathways (22), and drug toxicity are also closely related to the occurrence of muscle atrophy. Bevacizumab was originally used as a first-line treatment for metastatic colorectal cancer and is being continuously investigated for the treatment of non-small cell lung cancer, glioblastoma, renal cell carcinoma, ovarian cancer, cervical cancer and others (15). The use of bevacizumab could increases the risk of several toxicities, including common adverse effects such as bleeding, wound healing complications, gastrointestinal perforation, arterial thromboembolism, hypertension, and proteinuria (23). At the same time,multiple studies have also shown that patients with metastatic colorectal cancer experience loss of skeletal muscle mass during treatment with bevacizumab (24, 25). In addition, lower extremity muscle weakness has also been reported in patients with recurrent glioblastoma (26). In this study, the gastrocnemius muscle (n=6) and soleus muscle (n=6) of mice were used as the research objects, and the blank control group and the bevacizumab administration group were set, and a complete gene expression profile was obtained after normalization. After WGCNA and differential analysis, three sets of key genes were obtained, and the genes related to the proteolytic pathway were intersected. Proteolysis is one of the main mechanisms of muscle atrophy, and two overlapping genes, Timp4 and CDKN1A, were obtained through the venn diagram.

The cyclin-dependent kinase inhibitor 1 (CDKN1A) gene encodes the p21 protein, also known as p21Cip1/WAF1. p21 is considered an important regulator involved in multiple cellular

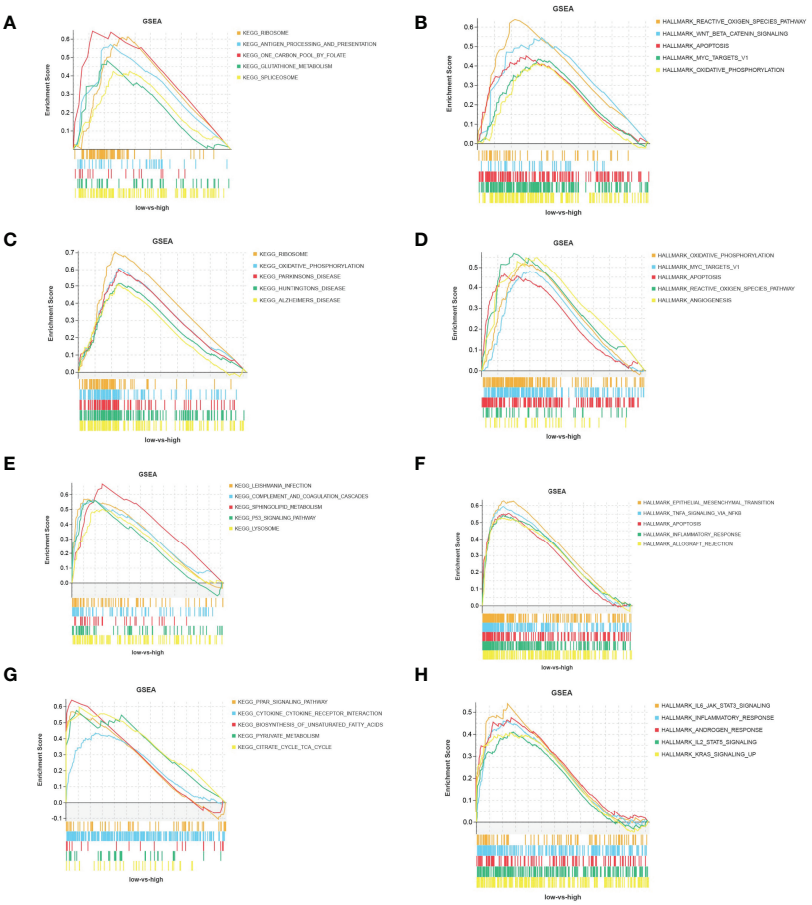



**FIGURE 6** | Functional pathways of Timp4 and Cdkn1a. **(A, C)** KEGG pathway and **(B, D)** Hallmark pathway of Timp4 and Cdkn1a in the bevacizumab-administered group; **(E, G)** KEGG pathway and **(F, H)** Hallmark pathway of Timp4 and Cdkn1a in the combined dataset of GSE6011 and GSE38417.

functions, including G1/S cell cycle progression, cell growth, DNA damage, and cell stemness (27). The expression of CDKN1A is related to the progression of various cancers, and it can be initiated by oncogenes, tumor suppressors, inflammatory cytokines, etc., and can inhibit cell proliferation by binding to CDKS, thus inhibiting the growth function of cancer cells (28, 29). In addition to its role in the occurrence, progression and treatment of cancer, CDKN1A has also been proved to be related to muscle atrophy. Wenjing Ma et al. found that CDKN1A was upregulated in denervated skeletal muscle (30). The increase of CDKN1A expression level can also be found in the muscle atrophy model induced by starvation (31).Similarly, the expression of CDKN1A increased in the model of muscle atrophy induced by hindlimb suspension, but decreased after reloading (32). While in other types of skeletal muscle atrophy, such as Duchenne muscular dystrophy (57), amyotrophic lateral sclerosis, aging and critical illness, CDKN1A mRNA is one of the most highly induced skeletal muscle mRNAs (33). As a well-known cell cycle inhibitor, CDKN1A has also been confirmed to be related to cell proliferation and proliferation regulation of muscle cells (34). It has been reported in the literature that Cdkn1a is an important protein related to muscle development, and it is involved in preventing myoblast proliferation, thereby inhibiting myogenesis or regeneration (35). It can affect the repair of skeletal muscle by regulating the expression of CDKN1A (36).However, Daniel K et al. suggested that p21 may promote muscle atrophy through a cell cycle-independent mechanism in skeletal muscle fibers. They speculate that these mechanisms may lead to cellular changes known to promote muscle wasting, including decreased anabolic signaling, increased procatabolic catabolism, decreased protein synthesis, and impaired mitochondrial function (33).In addition, the high induction of p21 during muscle atrophy leads to the decrease of spermine oxidase. Spermine oxidase inhibition mediated by p21 is considered to be a key step in the pathogenesis of skeletal muscle atrophy (37).

CDKN1A (p21) is a downstream target gene of TP53 (p53) (27).The p53/p21CIP pathway is a key pathway that normally responds to persistent DNA damage, and by which cellular senescence is achieved (38).Cell aging inhibits tumorigenesis in the body by preventing the proliferation of potential cancer cells. On the other hand, aging cells will destroy the integrity of local tissues and lead to some pathological changes, such as sarcopenia (39, 40). Cellular senescence has been identified as a mechanism for the

development of myopathies associated with muscular dystrophy mouse models (41). Cellular senescence may adversely affect sarcopenia through muscle stem cell dysfunction and a senescence-associated secretory phenotype (38). In addition, other factors that contribute to muscle atrophy are also linked to the p53/p21 pathway.1,25-Dihydroxyvitamin D deficiency induces sarcopenia through the induction of skeletal muscle cell senescence *via* the p53-p21 axis (42). The deficiency of peroxiredoxin 6, an antioxidant enzyme involved in maintaining intracellular redox homeostasis, also leads to the increase of p53-p21 pathway and thereby induces muscle atrophy (43). Doxorubicin (DOX), a chemotherapeutic drug, is an effective cell inhibitor. It also causes myocardial and skeletal muscle atrophy by activating p53-p21 signaling pathway (44). Myostatin also induces the high expression of p21. The excessive induction of p21 will lead to the irreversible senescence in quiescent satellite cell and impaire muscle regeneration (58).

The above studies have shown that high expression of CDKN1A (p21) has adverse effects on muscle. In the present study, CDKN1A showed low expression in the normal muscle group and high expression in the muscle atrophy group; however, the opposite situation occurred when the control group and the bevacizumab-administered group were compared. There are also a few studies that support the benefits of high expression of CDKN1A. The decreased apoptotic susceptibility of myoblasts to ROS is regulated at least in part by enhanced p21 promoter activity and nuclear p21 localization in myotubes (45). Synthetic beta-adrenergic agonists (BA) have broad biomedical and agricultural applications in increasing lean body mass, however BA treatment mediate increased mRNA expression of CDKN1A (46). A study by Diane et al. using clenbuterol to induce muscle hypertrophy also showed that BA treatment resulted in a significant upregulation of Cdkn1a mRNA abundance in skeletal muscle.CDKN1A inhibits cyclin-dependent kinase 2 activity, leading to irreversible cell cycle exit and terminal differentiation of myocytes (47).Satellite cells act as a source of new myonuclei during muscle repair and growth, whereas decreased p21Cip1 in aging skeletal muscle delays the withdrawal of satellite cells from the cell cycle and make them fail to differentiate, resulting in sarcopenia and impaired skeletal muscle regeneration (48). Thus the significant upregulation of Cdkn1a mRNA expression may indicate an increased potential for terminal differentiation and recruitment of myogenic precursor cells to support muscle hypertrophy.

Tissue inhibitor of metalloproteinase 4 (TIMP4) belongs to the family of extracellular matrix metalloproteinase inhibitors and regulates extracellular matrix (EMC) turnover (49). Studies have shown that TIMP4 is involved in cell survival, cell proliferation, inflammation, and epithelial-mesenchymal transition (EMT) signaling networks (50), and knockdown of its expression also indirectly promotes cell invasion and migration (51). The upregulation of TIMP4 gene prevents the metastasis of human cervical cancer cells by inhibiting PI3K/Akt/snail signaling pathway and blocking epithelial-mesenchymal transition (EMT) (52). Extracellular TIMP4 is involved in promoting the activation of the PI3K/AKT/mTOR pathway and promoting tumor metastasis, and is a prognostic and predictive marker for triple-negative breast cancer (53). In addition to this, a new study showed that TIMP4 affects lipid metabolism and smooth muscle cell proliferation (54). However, few studies have been conducted between TIMP4 and skeletal muscle atrophy. Proteomic analysis by Huemer et al. revealed that TIMP4 is a novel marker for the combination of low muscle mass and high fat mass (55). TIMP4 is highly expressed in adipose tissue, and in a study using TIMP4-deficient mice exposed to a high-fat diet, it promoted high-fat-induced obesity, fatty liver, and dyslipidemia. Furthermore, TIMP4-deficient mice are protected from skeletal muscle triglyceride accumulation in the quadriceps (49).The biochemical and structural remodeling of ECM is very important for the normal development of skeletal muscle. Through the study of zebrafish tendon junction (MTJ), Emma et al. found thatTIMP4 plays a regulatory role in this process (56). In this study, compared with normal tissues, TIMP4 showed low expression in both the administration group and muscle atrophy tissues. This suggests that the low expression of TIMP4 may be related to muscle atrophy, and bevacizumab may reduce the expression of timp4.Given that its association with muscle atrophy has not been definitively reported, more research is needed to determine.

## CONCLUSION

In conclusion, this study confirmed that bevacizumab may have muscle toxicity through bioinformatics analysis. Timp4 and Cdkn1a were identified as key genes, which are of great significance in mouse muscle atrophy. In addition, the biological effects of timp4 and Cdkn1a in the muscle toxicity of bevacizumab were briefly summarized through the analysis of function and pathway, so as to provide useful thinking for the follow-up use of bevacizumab. However, the specific mechanism of bevacizumab on muscle atrophy still needs to be further verified by *in vitro* experiments.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

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

## FUNDING

This research is supported by Grant No.2020YFC2005601 from the National Key R&D Program of China and Grant No.81771497 from the National Natural Science Foundation of China

## REFERENCES


1. Colloca G, Di Capua B, Bellieni A, Cesari M, Marzetti E, Valentini V, et al. Muscoloskeletal Aging, Sarcopenia and Cancer. *J Geriatric Oncol* (2019) 10 (3):504–9. doi: 10.1016/j.jgo.2018.11.007
2. Bozzetti F. Chemotherapy-Induced Sarcopenia. *Curr Treat Options Oncol* (2020) 21(1):7. doi: 10.1007/s11864-019-0691-9
3. Williams GR, Dunne RF, Giri S, Shachar SS, Caan BJ. Sarcopenia in the Older Adult With Cancer. *J Clin Oncol* (2021) 39(19):2068–78. doi: 10.1200/JCO.21.00102
4. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: Revised European Consensus on Definition and Diagnosis. *Age Ageing* (2019) 48(1):16–31. doi: 10.1093/ageing/afy169
5. Williams GR, Rier HN, McDonald A, Shachar SS. Sarcopenia & Aging in Cancer. *J Geriatric Oncol* (2019) 10(3):374–7. doi: 10.1016/j.jgo.2018.10.009
6. Awad S, Tan BH, Cui H, Bhalla A, Fearon KCH, Parsons SL, et al. Marked Changes in Body Composition Following Neoadjuvant Chemotherapy for Oesophagogastric Cancer. *Clin Nutr* (2012) 31(1):74–7. doi: 10.1016/j.clnu.2011.08.008
7. Goncalves MD, Taylor S, Halpenny DF, Schwitzer E, Gandelman S, Jackson J, et al. Imaging Skeletal Muscle Volume, Density, and FDG Uptake Before and After Induction Therapy for non-Small Cell Lung Cancer. *Clin Radiol* (2018) 73(5):505.e1–8. doi: 10.1016/j.crad.2017.12.004
8. Shachar SS, Deal AM, Weinberg M, Nyrop KA, Williams GR, Nishijima TF, et al. Skeletal Muscle Measures as Predictors of Toxicity, Hospitalization, and Survival in Patients With Metastatic Breast Cancer Receiving Taxane-Based Chemotherapy. *Clin Cancer Res* (2017) 23(3):658–65. doi: 10.1158/1078-0432.CCR-16-0940
9. Kim I, Choi MH, Lee IS, Hong TH, Lee MA. Clinical Significance of Skeletal Muscle Density and Sarcopenia in Patients With Pancreatic Cancer Undergoing First-Line Chemotherapy: A Retrospective Observational Study. *BMC Cancer* (2021) 21(1):77. doi: 10.1186/s12885-020-07753-w
10. Choi Y, Oh D, Kim T, Lee K, Han S, Im S, et al. Skeletal Muscle Depletion Predicts the Prognosis of Patients With Advanced Pancreatic Cancer Undergoing Palliative Chemotherapy, Independent of Body Mass Index. *PloS One* (2015) 10(10):e0139749. doi: 10.1371/journal.pone.0139749
11. Poterucha T, Burnette B, Jatoi A. A Decline in Weight and Attrition of Muscle in Colorectal Cancer Patients Receiving Chemotherapy With Bevacizumab. *Med Oncol* (2012) 29(2):1005–9. doi: 10.1007/s12032-011-9894-z
12. Dolly A, Lecomte T, Bouché O, Borg C, Terrebonne E, Douillard J, et al. Concurrent Losses of Skeletal Muscle Mass, Adipose Tissue and Bone Mineral Density During Bevacizumab / Cytotoxic Chemotherapy Treatment for Metastatic Colorectal Cancer. *Clin Nutr* (2020) 39(11):3319–30. doi: 10.1016/j.clnu.2020.02.017
13. Blauwhoff-Buskermolen S, Versteeg KS, de van der Schueren MA, den Braver NR, Berkhof J, Langius JA, et al. Loss of Muscle Mass During Chemotherapy Is Predictive for Poor Survival of Patients With Metastatic Colorectal Cancer. *J Clin Oncol* (2016) 34(12):1339–44. doi: 10.1200/JCO.2015.63.6043
14. Derksen JWG, Kurk SA, Oskam MJ, Peeters PHM, Punt CJA, Koopman M, et al. Factors Contributing to Cancer-Related Muscle Wasting During First-Line Systemic Treatment for Metastatic Colorectal Cancer. *JNCI Cancer Spectr* (2019) 3(2):pkz014. doi: 10.1093/jncics/pkz016
15. Garcia J, Hurwitz HI, Sandler AB, Miles D, Coleman RL, Deurloo R, et al. Bevacizumab (Avastin®) in Cancer Treatment: A Review of 15 Years of Clinical Experience and Future Outlook. *Cancer Treat Rev* (2020) 86:102017. doi: 10.1016/j.ctrv.2020.102017
16. Zalcman G, Mazieres J, Margery J, Greillier L, Audigier-Valette C, Moro-Sibilot D, et al. Bevacizumab for Newly Diagnosed Pleural Mesothelioma in the Mesothelioma Avastin Cisplatin Pemetrexed Study (MAPS): A Randomised, Controlled, Open-Label, Phase 3 Trial. *Lancet* (2016) 387 (10026):1405–14. doi: 10.1016/S0140-6736(15)01238-6
17. Qi N, Li F, Li X, Kang H, Zhao H, Du N. Combination Use of Paclitaxel and Avastin Enhances Treatment Effect for the NSCLC Patients With Malignant Pleural Effusion. *Medicine* (2016) 95(47):e5392. doi: 10.1097/MD.0000000000005392
18. Emami N, Mohamadnia A, Mirzaei M, Bayat M, Mohammadi F, Bahrami N. miR-155, miR-191, and miR-494 as Diagnostic Biomarkers for Oral Squamous Cell Carcinoma and the Effects of Avastin on These Biomarkers. *J Korean Assoc Oral Maxillofacial Surgeons* (2020) 46(5):341–7. doi: 10.5125/jkaoms.2020.46.5.341
19. Tsavachidou-Fenner D, Tannir N, Tamboli P, Liu W, Petillo D, Teh B, et al. Gene and Protein Expression Markers of Response to Combined Antiangiogenic and Epidermal Growth Factor Targeted Therapy in Renal Cell Carcinoma. *Ann Oncol* (2010) 21(8):1599–606. doi: 10.1093/annonc/mdp600
20. Huang H, Xue J, Zheng J, Tian H, Fang Y, Wang W, et al. Bioinformatic Analysis of the Gene Expression Profile in Muscle Atrophy After Spinal Cord Injury. *Sci Rep* (2021) 11(1):21903. doi: 10.1038/s41598-021-01302-6
21. Hyatt H, Deminice R, Yoshihara T, Powers SK. Mitochondrial Dysfunction Induces Muscle Atrophy During Prolonged Inactivity: A Review of the Causes and Effects. *Arch Biochem Biophysics* (2019) 662:49–60. doi: 10.1016/j.abb.2018.11.005
22. Perry BD, Caldow MK, Brennan-Speranza TC, Sbaraglia M, Jerums G, Garnham A, et al. Muscle Atrophy in Patients With Type 2 Diabetes Mellitus: Roles of Inflammatory Pathways, Physical Activity and Exercise. *Exerc Immunol Rev* (2016) 22:94–109.
23. Gressett SM, Shah SR. Intricacies of Bevacizumab-Induced Toxicities and Their Management. *Ann Pharmacother* (2009) 43(3):490–501. doi: 10.1345/aph.1L426
24. Kurk SA, Peeters PHM, Dorresteijn B, Jong PA, Jourdan M, Creemers GJM, et al. Loss of Skeletal Muscle Index and Survival in Patients With Metastatic Colorectal Cancer: Secondary Analysis of the Phase 3 CAIRO3 Trial. *Cancer Med* (2019) 9(3):1033–43. doi: 10.1002/cam4.2787
25. Kurk SA, Peeters PHM, Dorresteijn B, de Jong PA, Jourdan M, Kuijf HJ, et al. Impact of Different Palliative Systemic Treatments on Skeletal Muscle Mass in Metastatic Colorectal Cancer Patients. *J Cachexia Sarcopenia Muscle* (2018) 9 (5):909–19. doi: 10.1002/jcsm.12337
26. Fallah J, Chaudhary R, Rogers L, Auston WW, Brewer C, Peereboom D, et al. Actr-26. Safety and Efficacy of Bevacizumab Plus Ttfields in Patients With Recurrent Glioblastoma (Gbm): Data From a Phase Ii Clinical Trial. *Neuro-Oncology* (2019) 21(Supplement_6). doi: 10.1093/neuonc/noz175.069
27. Xiao B, Zhao Y, Jia X, Wu J, Wang Y, Huang F. Multifaceted P21 in Carcinogenesis, Stemness of Tumor and Tumor Therapy. *World J Stem Cells* (2020) 12(6):481–7. doi: 10.4252/wjsc.v12.i6.481
28. Li S, Wang C, Yu X, Wu H, Hu J, Wang S, et al. miR-3619-5p Inhibits Prostate Cancer Cell Growth by Activating CDKN1A Expression. *Oncol Rep* (2017) 37 (1):241–8. doi: 10.3892/or.2016.5250
29. Kreis NN, Louwen F, Yuan J. The Multifaceted P21 (Cip1/Waf1/CDKN1A) in Cell Differentiation, Migration and Cancer Therapy. *Cancers (Basel)* (2019) 11 (9):1220. doi: 10.3390/cancers11091220
30. Ma W, Cai Y, Shen Y, Chen X, Zhang L, Ji Y, et al. HDAC4 Knockdown Alleviates Denervation-Induced Muscle Atrophy by Inhibiting Myogenin-Dependent Atrogene Activation. *Front Cell Neurosci* (2021) 15. doi: 10.3389/fncel.2021.663384
31. Ebert SM, Monteys AM, Fox DK, Bongers KS, Shields BE, Malmberg SE, et al. The Transcription Factor ATF4 Promotes Skeletal Myofiber Atrophy During Fasting. *Mol Endocrinol* (2010) 24(4):790–9. doi: 10.1210/me.2009-0345
32. Wang J, Wang X, Feng W. Reloading Promotes Recovery of Disuse Muscle Loss by Inhibiting Tgfβ Pathway Activation in Rats After Hind Limb Suspension. *Am J Phys Med Rehabil* (2017) 96(6):430–7. doi: 10.1097/PHM.0000000000000617
33. Fox DK, Ebert SM, Bongers KS, Dyle MC, Bullard SA, Dierdorff JM, et al. P53 and ATF4 Mediate Distinct and Additive Pathways to Skeletal Muscle Atrophy During Limb Immobilization. *Am J Physiol-Endocrinol Metab* (2014) 307(3):E245–61. doi: 10.1152/ajpendo.00010.2014
34. Wang J, Song C, Cao X, Li H, Cai H, Ma Y, et al. MiR-208b Regulates Cell Cycle and Promotes Skeletal Muscle Cell Proliferation by Targeting CDKN1A. *J Cell Physiol* (2019) 234(4):3720–9. doi: 10.1002/jcp.27146
35. Davidovic L, Durand N, Khalfallah O, Tabet R, Barbry P, Mari B, et al. A Novel Role for the RNA-Binding Protein FXR1P in Myoblasts Cell-Cycle Progression by Modulating P21/Cdkn1a/Cip1/Waf1 mRNA Stability. *PloS Genet* (2013) 9(3):e1003367. doi: 10.1371/journal.pgen.1003367
36. Chunyan F. Role of miR-128, Sp1 and CDKN1A mRNA in Repairing Muscle Injury in a Rat Skeletal Muscle Injury Model. *Indian J Pharm Sci* (2020) 82:15412. doi: 10.36468/pharmaceutical-sciences.spl.124

37. Bongers KS, Fox DK, Kunkel SD, Stebounova LV, Murry DJ, Pufall MA, et al. Spermine Oxidase Maintains Basal Skeletal Muscle Gene Expression and Fiber Size and is Strongly Repressed by Conditions That Cause Skeletal Muscle Atrophy. *Am J Physiol-Endocrinol Metab* (2015) 308(2):E144–58. doi: 10.1152/ajpendo.00472.2014
38. He Y, Xie W, Li H, Jin H, Zhang Y, Li Y. Cellular Senescence in Sarcopenia: Possible Mechanisms and Therapeutic Potential. *Front Cell Dev Biol* (2022) 9. doi: 10.3389/fcell.2021.793088
39. Ma JF, Hall DT, Gallouzi I. The Impact of mRNA Turnover and Translation on Age-Related Muscle Loss. *Ageing Res Rev* (2012) 11(4):432–41. doi: 10.1016/j.arr.2012.05.004
40. Muñoz-Espín D, Serrano M. Cellular Senescence: From Physiology to Pathology. *Nat Rev Mol Cell Biol* (2014) 15(7):482–96. doi: 10.1038/nrm3823
41. Sosa P, Alcalde-Estevez E, Plaza P, Troyano N, Alonso C, Martínez-Arias L, et al. Hyperphosphatemia Promotes Senescence of Myoblasts by Impairing Autophagy Through Ilk Overexpression, A Possible Mechanism Involved in Sarcopenia. *Aging Dis* (2018) 9(5):769. doi: 10.14336/AD.2017.1214
42. Yu S, Ren B, Chen H, Goltzman D, Yan J, Miao D. 1,25-Dihydroxyvitamin D Deficiency Induces Sarcopenia by Inducing Skeletal Muscle Cell Senescence. *Am J Transl Res* (2021) 13(11):12638–49.
43. Pacifici F, Della-Morte D, Piermarini F, Arriga R, Scioli MG, Capuani B, et al. Prdx6 Plays a Main Role in the Crosstalk Between Aging and Metabolic Sarcopenia. *Antioxidants* (2020) 9(4):329. doi: 10.3390/antiox9040329
44. Hulmi JJ, Nissinen TA, Räsänen M, Degerman J, Lautaoja JH, Hemanthakumar KA, et al. Prevention of Chemotherapy-Induced Cachexia by ACVR2B Ligand Blocking has Different Effects on Heart and Skeletal Muscle. *J Cachexia Sarcopenia Muscle* (2018) 9(2):417–32. doi: 10.1002/jcsm.12265
45. Haramizu S, Asano S, Butler DC, Stanton DA, Hajira A, Mohamed JS, et al. Dietary Resveratrol Confers Apoptotic Resistance to Oxidative Stress in Myoblasts. *J Nutr Biochem* (2017) 50:103–15. doi: 10.1016/j.jnutbio.2017.08.008
46. Brown D, Ryan K, Daniel Z, Mareko M, Talbot R, Moreton J, et al. The Beta-Adrenergic Agonist, Ractopamine, Increases Skeletal Muscle Expression of Asparagine Synthetase as Part of an Integrated Stress Response Gene Program. *Sci Rep* (2018) 8(1). doi: 10.1038/s41598-018-34315-9
47. Spurlock DM, McDaneld TG, McIntyre LM. Changes in Skeletal Muscle Gene Expression Following Clenbuterol Administration. *BMC Genomics* (2006) 7(1):320–0. doi: 10.1186/1471-2164-7-320
48. Zwetsloot KA, Childs TE, Gilpin LT, Booth FW. Non-Passaged Muscle Precursor Cells From 32-Month Old Rat Skeletal Muscle Have Delayed Proliferation and Differentiation. *Cell Proliferation* (2013) 46(1):45–57. doi: 10.1111/cpr.12007
49. Sakamuri SSVP, Watts R, Takawale A, Wang X, Hernandez-Anzaldo S, Bahitham W, et al. Absence of Tissue Inhibitor of Metalloproteinase-4 (TIMP4) Ameliorates High Fat Diet-Induced Obesity in Mice Due to Defective Lipid Absorption. *Sci Rep* (2017) 7(1):6210. doi: 10.1038/s41598-017-05951-4
50. Lizarraga F, Espinosa M, Ceballos-Cancino G, Vazquez-Santillan K, Bahena-Ocampo I, Schwarz-Cruz Y Celis A, et al. Tissue Inhibitor of Metalloproteinases-4 (TIMP-4) Regulates Stemness in Cervical Cancer Cells. *Mol Carcinog* (2016) 55(12):1952–61. doi: 10.1002/mc.22442
51. Cheng D, Jiang S, Chen J, Li J, Ao L, Zhang Y. Upregulated Long Noncoding RNA Linc00261 in Pre-Eclampsia and its Effect on Trophoblast Invasion and Migration *via* Regulating miR-558/TIMP4 Signaling Pathway. *J Cell Biochem* (2019) 120(8):13243–53. doi: 10.1002/jcb.28598
52. Nayim P, Mbaveng AT, Sanjukta M, Rikesh J, Kuete V, Sudhir K. CD24 Gene Inhibition and TIMP-4 Gene Upregulation by Imperata Cylindrica's Root Extract Prevents Metastasis of CaSki Cells *via* Inhibiting PI3K/Akt/snail Signaling Pathway and Blocking EMT. *J Ethnopharmacol* (2021) 275:114111. doi: 10.1016/j.jep.2021.114111
53. Wallon UM, Sabol JL, Zemba-Palko V, DuHadaway JS, Sutanto-Ward E, Ali ZA, et al. Abstract P6-08-54: TIMP-4 is a Prognostic and Predictive Marker in Triple-Negative Breast Cancers. *Cancer Res* (2015) 75(9_Supplement):P6–08-54-P6-08-54. doi: 10.1158/1538-7445.SABCS14-P6-08-54
54. Hu M, Jana S, Kilic T, Wang F, Shen M, Winkelaar G, et al. Loss of TIMP4 (Tissue Inhibitor of Metalloproteinase 4) Promotes Atherosclerotic Plaque Deposition in the Abdominal Aorta Despite Suppressed Plasma Cholesterol Levels. *Arteriosclerosis Thrombosis Vasc Biol* (2021) 41(6):1874–89. doi: 10.1161/ATVBAHA.120.315522
55. Huemer MT, Bauer A, Petrera A, Scholz M, Hauck SM, Drey M, et al. Proteomic Profiling of Low Muscle and High Fat Mass: A Machine Learning Approach in the KORA S4/FF4 Study. *J Cachexia Sarcopenia Muscle* (2021) 12(4):1011–23. doi: 10.1002/jcsm.12733
56. Matchett EF, Wang S, Crawford BD. Paralogues of Mmp11 and Timp4 Interact During the Development of the Myotendinous Junction in the Zebrafish Embryo. *J Dev Biol* (2019) 7(4):22. doi: 10.3390/jdb7040022
57. Welle S, Brooks AI, Delehanty JM, Needler N, Bhatt K, Shah B, et al. Skeletal Muscle Gene Expression Profiles in 20–29 Year Old and 65–71 Year Old Women. *Exp Gerontol* (2004) 39(3):369–77. doi: 10.1016/j.exger.2003.11.011
58. Veeranki S, Lominadze D, Tyagi SC. Hyperhomocysteinemia Inhibits Satellite Cell Regenerative Capacity Through P38 Alpha/Beta MAPK Signaling. *Am J Physiol-Heart Circulatory Physiol* (2015) 309(2):H325–34. doi: 10.1152/ajpheart.00099.2015


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 04 July 2022
doi: 10.3389/fonc.2022.901705

# Effects of Tumor-Derived DNA on CXCL12-CXCR4 and CCL21-CCR7 Axes of Hepatocellular Carcinoma Cells and the Regulation of Sinomenine Hydrochloride

*Conghuan Shen*[†], *Jianhua Li*, *Ruidong Li*, *Zhenyu Ma*, *Yifeng Tao*, *Quanbao Zhang*[†] *and Zhengxin Wang*[*]

*General Surgery Department, Affiliated Huashan Hospital of Fudan University, Shanghai, China*

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Changyong E,*
*Jilin University, China*
*Ting Li,*
*Central South University, China*

***Correspondence:**
*Zhengxin Wang*
*wangzhengxin@huashan.org.cn*

[†]*These authors share first authorship*

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

**Received:** *22 March 2022*
**Accepted:** *11 May 2022*
**Published:** *04 July 2022*

**Citation:**
*Shen C, Li J, Li R, Ma Z, Tao Y, Zhang Q and Wang Z (2022) Effects of Tumor-Derived DNA on CXCL12-CXCR4 and CCL21-CCR7 Axes of Hepatocellular Carcinoma Cells and the Regulation of Sinomenine Hydrochloride. Front. Oncol. 12:901705. doi: 10.3389/fonc.2022.901705*

Currently, chemokines and their receptors, CXCL12-CXCR4 and CCL21-CCR7 axes, are deemed vital factors in the modulation of angiogenesis and are crucial for the growth and development of liver cancer. Tumor-derived DNA can be recognized by immune cells to induce an autoimmune response. In this study, we demonstrated the mechanism of tumor-derived DNA on the CXCL12-CXCR4 and CCL21-CCR7 axes of hepatocellular carcinoma (HCC) cells and the regulatory effect of sinomenine hydrochloride. Tumor-derived DNA was separated from HCCLM cell lines. Tumor-derived DNA was transfected into SK-Hep1 cells by Lipofectamine 2000. We found that sinomenine hydrochloride reduced the expression of CXCR4, CXCR12, CCR7, and CCL21 in HCC cells, suppressed the growth and invasion of HCC cells, and increased apoptosis. In contrast to the controls, the protein expressions of CXCR4, CXCL12, CCR7, CCL21, P-ERK1/2, MMP-9, and MMP-2 in SK-Hep1 cells were significantly increased after transfection of tumor-derived DNA, while the increase was reversed by sinobine hydrochloride. Acid sinomenine interferes with tumor-derived DNA and affects ERK/MMP signaling *via* the CXCL12/CXCR4 axis in HCC cells. CXCR4 siRNA and CCR7 siRNA attenuated tumor-derived DNA activation of ERK1/2/MMP2/9 signaling pathways in HCC cells. CXCR4-oe and CCR7-OE enhance the stimulation of erK1/2/MMP2/9 signaling pathway by tumor-derived DNA in HCC cells. Tumor-derived DNA reduced apoptosis and increased invasion of SK-Hep1 cells by CXCL12-CXCR4 axis and CCL21-CCR7 axis, and sinobine hydrochloride reversed this regulation. These results strongly suggest that tumor-derived DNA can increase the growth and invasion of oncocytes *via* the upregulation of the expression of CXCL12-CXCR4 and CCL21-CCR7 axis and through ERK1/2/MMP2/9 signaling pathway in HCC cells, and sinobine hydrochloride can inhibit this signaling pathway, thus inhibiting HCC cells. These results provide new potential therapeutic targets for blocking the progression of HCC induced by CXCL12-CXCR4 axis and CCL21-CCR7.

**Keywords: hepatic carcinoma, CXCL12-CXCR4, CCL21-CCR7, sinomenine hydrochloride, tumor-derived DNA**

## INTRODUCTION

Hepatic carcinoma is one of the most commonly seen malignancies of the digestive system in the world. In 2020, there were 905,677 novel cases of liver carcinoma across the globe, accounting for 4.7% of the novel tumor cases in the whole year, making it the second leading cause of tumor death in the world, with 830,180 deaths (8.3%) in the whole year (1, 2). Hepatocellular carcinoma (HCC), the most commonly seen type of primary hepatic carcinoma, is one of the most commonly seen malignancies in clinical practice, and its related mortality ranks the third among malignant tumors (1). Most patients with HCC have no obvious symptoms at the early stage but are already in the middle and advanced stages when they are treated, and the overall prognosis is still poor (3). It was found that the great rate of cancer relapse and metastases was the main factor leading to unsatisfactory prognoses of HCC sufferers. Hence, new enlightenment pertaining to the mechanisms of HCC is imperative for the determination of new prognosis molecule biomarkers and potentially valid treatment targets to ameliorate the survival of sufferers (4, 5).

When tumor cells are necrotic, normally apoptotic, and killed by chemotherapy and radiotherapy, a large amount of DNA can be repositioned from the nucleus and released into the circulatory system. The released DNA can lead to autoimmune reactions, and the circulating DNA can stimulate immunocytes, like macrophages and natural killer cells, and then induce autoimmune reactions (5, 6). Circulating tumor DNA (ctDNA) is a specific tumor-related DNA fragment that is secreted and released into peripheral blood during tumor cell renewal, necrosis, and apoptosis. ctDNA can reflect tumor cytogenetics and epigenetic information (7). In the whole-cycle management of HCC, ctDNA monitoring can dynamically provide the information of tumor genetic variation, which has become a hot research direction in recent years (8).

Sinomenine is the alkaloid monomer extracted from Chinese traditional medicine. Sinomenine hydrochloride (SIN) is a water-soluble hydrochloride, which has anti-inflammatory and anti-immune pharmacological effects. Recent studies have also found that sinomenine can inhibit the proliferation, block the cell cycle and induce the apoptosis of oncocytes such as hepatic carcinoma, gastric cancer, mammary carcinoma, and osteosarcoma. Li et al. displayed that SIN can reduce breast cancer metastasis by inducing breast cancer cell death and inhibiting inflammatory associated epithelial mesenchymal transition (9). Deng et al. discovered that SIN promotes programmed cell death in kidney cell cancer by increasing autophagic activities *via* the phosphatidylinositol 3-kinase/AKT/rapamycin path (10). In addition, research has revealed the integrated effects of SIN and chemotherapeutic drugs in cancer treatment. Liu et al. found that SIN enhanced doxorubicin sensitivity of colon cancer cells (CACO-2) by downregulating the expression of multiple drug-tolerant protein 1 and cyclooxygenase-2 (11). In addition, the integrated effect of SIN and 5-fluorouracil on esophageal cancer is superior to that of 5-fluorouracil alone and does not aggravate the adverse events of chemotherapies. However, the potential mechanism of SIN in inhibiting HCC cells remains to be fully elucidated.

This study will investigate the effects of CXCL12-CXCR4 and CCL21-CCR7 axis signaling pathways on cancer-derived DNA activation and proliferation induction of HCC cells. Experiments were conducted to elucidate that SIN inhibits tumor-derived DNA tumor-induced HCC cells and induces cell proliferation through CXCL12-CXCR4 and CCL21-CCR7 axis signaling pathways, aiming to provide possible drug selection and relevant preclinical evidence for clinical HCC treatment.

## METHODS

### Reagents and Antibodies

SIN was bought from Zheng Qing Pharmaceutic Group (Hunan, PRC). Human HCC cell line SK-Hep1 (Shanghai Aolu Biotechnology Co., Ltd.), TRIzol reagent, SuperScript III first-strand synthetic system (Invitrogen, America), ABI 7500 fast real-time PCR system (Applied Biosystems, USA), diquinone bicinchoninic acid (BCA) tool (Thermo Fisher Scientific, USA), Fluorchem FC3 system (Proteinsimple, USA), polyvinylidene fluoride (PVDF) membrane, chemiluminescent reagent (Millipore, USA), transpore Chamber (BD Biosciences, USA), diaminophenylindole (DAPI), 3-MA, 4-PBA, LPS, and MDC were bought from Sigma-Aldrich (USA), Cell Counting Kit 8 (CCK-8; Japan Dojindo Company), multi-hole scanning spectrophotometer (US Thermo Scientific Company). Single-person ultra-clean bench (Beijing Liuyi Instrument Factory), CB15 C02 cell incubator (Beijing Liuyi Instrument Factory), Victor3 1420 Multi-lable Counter microplate reading device (BD, USA), HD-3000 gel imaging device (Shang Tian Precision Apparatus Company, PRC), Millipore flow cytometer (Beijing Weixin Yiao Technology Development Co., Ltd.). Phospho-ERK1/2, matrix metalloproteinase (MMP)-2, MMP-9, extracellular signal-regulated kinase 1/2 (ERK1/2), CXCL12, CXCR4, CCL21, CCR7, Caspase 3, and GAPDH antibodies were bought from Abcam (UK). Every second antibody was provided by Santa Cruz Biotechnology, Inc. (USA).

### Cell Culture

Human HCC cell line SK-Hep1 was cultivated in dulbecco's modified eagle medium (DMEM) or Roswell Park Memorial Institute (RPMI) 1640 intermediary with 10% FBS. The medium involved 100 mg/ml penicillin and kyowamycin sulfate. And in 5% $CO_2$, 90% humidity, 37°C incubator. Harvested media and cell lysates are preserved under -20°C for subsequent assays. Mouse liver cancer cell lines (Hepa1-6) were cultivated in DMEM Basic medium, 10% FBS, 100 U/ml penicillin-kyowamycin, 37°C, 5% $CO_2$ cell box. Cellular cultivation intermediary was refreshed every 1–2 days, and cell passage was performed when 90% of cells were fused.

### Tumor-Derived DNA Transfection

Serum-free medium was mixed with Lipofectamine 2000 5 μl every 50 μl and let stand for 5 min. Serum-free medium was mixed with extracted tumor cell (Lewis cell) DNA every 50 μl and let stand for 5 min. The two were fully mixed and let stand for 20 min.

The transfected cells were cleaned two times in serum-free DMEM/F12, and the mixed solution of liposome and DNA was supplemented to the cells. Cells were cultivated under 37°C and 5% $CO_2$ for 4–6 h. The transfection intermediary was removed, and 10% FBS-DMEM/F12 intermediary was supplemented into the cells, and the cells were cultured for 48 h. Real-time PCR was leveraged to identify the expression of 18S mRNA to determine the transfection effect, and then Western blot (WB) or other experiments were performed.

## Transfection of siRNA

SiRNA-CXCR4 plasmid, siRNA-CCR7 plasmid, and negative control plasmid (siRNA-NC) were all provided by Genechem Shanghai, Primer sequences: GATGAGGTCACGGACGATT (5'-GATCCGATGAGGTCACGGACGATTCTATGGACAAA TCGTCCGTGACCTCATCTTTTTTGTCGACA-3', 3'-GCTACTCCAGTGCCTGCTAAGATACCTGTTTAGCAGGC ACTGGAGTAGAAAAAACAGCTGTTCGA-5'). The transfection was performed according to the Lipofectamine™ 2000 (Invitrogen) kit instruction. Trypsinase was digested before transfection, and the cell density was modified to $6 \times 10^6$/ml in RPMI 1640 intermediary involving 10% FBS. Cells were inoculated in 24-well culture dishes nightlong and transfected when the cell culture reached about 75% after fusion. At 48 h after transient transfection, the digested cells were inoculated on 6-well dishes and cultivated with RPMI 1640 intermediary for subsequent assay.

## Transfection of Plasmid DNA (CXCR4/ CCR7 High-Expression Plasmid Screening Validation)

The cell culture medium in the 6-well plate was changed to 1.5 ml DMEM(-), and 12 1.5-ml EP tubes were placed on the test tube shelf. The cells were equally divided into two groups: Lipofectamine®2000 Transfection Reagent group and plasmid DNA group. In this study, 245, 245, 245, 240, 240, and 240 μl DMEM(-) were added to each tube of the Lipofectamine®2000 Transfection Reagent. The required DMEM(-) volume was calculated according to the measured plasmid concentration in the plasmid DNA set. Add 5μL, 5μL, 5μL, 8μL, 8μL, 8μL Lipofectamine®2000 Transfection Reagent to the Lipofectamine®2000 Transfection Reagent group, and mix well. 2, 4, and 6 μg pcmv3-CXCR4-ha plasmids were added to the plasmid DNA groups and mixed and stood at room temperature for 5 min. Subsequently, the plasmid DNA group and the Lipofectamine®2000 Transfection Reagent group were mixed in pairs and stood at room temperature for 20 min to form the composite. The mixture was evenly dropped into the corresponding culture well, and the compound was evenly distributed in the medium by gently shaking the 6-well plate, and the label was made. At 4–6 h later, 200 μl FBS was added to each well. Cells were observed the next day and collected 48 h later.

## Total RNA Extraction and Fluorescence Quantitative PCR

After trypsin digested cells in each experimental group, cells were lysed with TRIzol solution, total RNA was abstracted as per the specification of RNA tool, and RNA was reversely transcribed to cDNA as per the specification of TaqMan reverse transcription kit. According to the instructions of Takara fluorescence quantitative kit, RT-qPCR was performed on Light Cycler480II fluorescence quantitative PCR instrument. GAPDH was utilized as an internal reference, and $2^{-\Delta\Delta Ct}$ was employed to compute the relative expression levels (12). All experiments were repeated in triplicate. Primer sequences are as follows:

Mmp-9, 5'-ATGGAGCTGGAATTGGATGC-3' (forward), 5' -cTAGCCATATCTGTCCT-3 ' (reverse); Mmp-2, 5'-tacaccta taccaagaACTTCCG-3' (forward), 5'-tgTCCGCCAGatGaACC G-3' (reverse); CCR7, 5'-AAGGGTCAGGAGGAAGagGA3' (forward), 5'-GGCTGGTCGTGTTGACCtAT-3' (reverse); GAPDH, 5'-TCACTGCCaagaaga-3' (forward), 5'-TaccagG AaatGAGcTTGA-3' (reverse). Real-time fluorescence quantitative PCR was completed using the QuantiTect SYBR Green RT-PCR tool (QIAGEN, Germany) *via* ABI Prism 7000 sequence detecting device (Applied Biological Systems, Germany). The comparative level of MMP-2/-9 mRNA was normalized to GAPDH mRNA. The assay was carried out in triplicate.

## Detection of Cell Proliferation Activity

The proliferative activity of cells was identified *via* CCK-8 approach. The transfected cells were digested with trypsinase, and the cellular density was modified by medium. Here, $10^4$ cells in each well were seeded to 96-well plates at 0, 24, 48, and 72 h, 10 μl CCK-8 liquor was supplemented, mixed and incubated in an incubator for 4 h. Optical density (OD) was identified at 490 nm *via* an automatic porous spectral photometer (Bio-Rad Laboratories, USA). Cellular proliferation curves were plotted using culture time as abscissa and OD value as ordinate.

## MTT Cell Proliferation Assay

Single-cell suspension was prepared from P3 Bone Mesenchymal Stem Cells (BMSCs), the cellular density was modified ($5 \times 10^4$ cells/ml), and cells were seeded to a 96-well dish ($5 \times 10^3$ cells/ well). Following cell adherence, the cultivation intermediary was cultured for 24 h, and amaranth was supplemented into the cultivation intermediary at a content gradient of 5, 10, and 20 μg/ ml for 48 h. 3-(4,5)-dimethylthiahiazo (-z-y1)-3,5-di-phenytetrazoliumromide (MTT) 20 μl (5 mg/ml) was added, culturing in 37°C incubator was continued for 4 h, the supernatant was absorbed, DMSO was added to each well, 150 μl was vibrated for 10 min, and the absorbance value (OD490 nm) was measured with a microplate reader.

## Transwell Experiment

After serum-free culture for 12 h before inoculation, P3 cells were digested and the cellular concentration was modified to $2 \times 10^5$ cells/ml in FBS medium (0.1%). In this study, 200 μl cellular suspension was supplemented into the upper chamber. In addition, 750 μl FBS intermediary (20%) with or without amaranth (5, 10, 20 μg/ml) was supplemented into the lower chamber, and 3 multiple wells were set for every group. After 10-h culture in the incubator, the chamber was taken out, the cultivation intermediary in the abandoned well was cleaned

two times in PBS, subjected to fixation in 4% PFA for 20 min, cleaned two times in PBS, dyed with Giemsa dye for 20 min, and cleaned two times in PBS. Under an inverted microscope (×400), 5 fields were randomly selected for observation of counting cells.

## Detection of Apoptosis

Annexin V and propidium iodide (PI) apoptotic identification tool was leveraged to identify programmed cell death. Cells in each transfection group were subjected to digestion *via* trypsin, and the cellular concentration was modified to $1 \times 10^6$/ml by resuspension combined with buffer. Annexin V 5 μl and PI 5 μl were supplemented into suspension cells and mixed. The cells should be protected from light for 5–15 min under ambient temperature. After the cellular suspension was mixed, it was identified *via* flow cell technique immediately. The excitation wavelength Ex = 488 nm and emission wavelength Em = 530 nm and the cell suspension could be lightly elastic again to avoid cell clumps before the machine.

## Western Blotting Experiment

The cells were cultured in a 10-cm Petri dish, and after the degree of cell fusion reached over 90%, cell precipitation was collected, and RIPA protein lysate containing Phenylmethanesulfonyl fluoride (PMSF) was added to fully lysate the cells. After the content of every specimen was identified *via* BCA protein quantitative kit, the samples were stored in a refrigerator at -80°C for use. The WB procedure was as follows: 50 μg total protein was electrophoresized in 12% sodium dodecyl sulfate (SDS)-polyacrylamide gel (110 V). The proteins in the gel were moved onto the PVDF film (current: 300 mA, 2 h). Here, 5% powdered skimmed milk was sealed under ambient temperature for 3 h. Tween-20 phosphate buffer (PBST) was washed 3× 5 min, and primary antibody was incubated at 4°C overnight (β-actin, CXCL12, CXCR4, CCL21, CCR7, lysed Caspase 3, caspase 3, The concentrations of ERK and P-ErK were 1:1,000), and the PBST membrane was washed for 3× 5 min. Horseradish peroxidase (HRP)-labeled second antibody (1:10,000) was cultivated under 37°C for 45 min. PBST washed the membrane 4× 5 min. Electrochemiluminescence (ECL) developer was supplemented to the membrane, and the membrane was exposed and photographed in a fully automatic chemiluminescence system.

## Immunofluorescence

After digestion and centrifugation, the cells were seeded to a 6-well dish with cover glass at a rate of $5 \times 10^5$/well, and 2 ml/well of DMEM/F12 medium with 10%FBS was added and placed in an incubator overnight. DMEM/F12 was cleaned twice, SIN was added with a concentration of 0.25 mM, and cultured for 48 h. Lipopolysaccharide (LPS) 2 ml with the concentration of 1 μg/ml was added and cultured for 4 h. The intermediary was discarded and cleaned in PBS 3 times. Here, 2 ml of 4% paraformaldehyde was added into the well and fixed overnight at 4°C. Then, 4% paraformaldehyde was discarded and cleaned with PBS 3 times. Then, 0.1% Triton-100 was added, 2 ml/well, 10 min later, it was washed with PBS 3 times. The primary antibody was added at 4° C overnight, negative control PBS 30 μl, and washed with PBS 3 times. After adding second antibody for 1 h, it was washed with PBS 3 times. DAPI 30 μl/tablet (diluted at 1:2,000) was added, washed with PBS 3 times under ambient temperature for 5 min. Immunofluorescence blocker was used to block the cells, and the cells were analyzed and imaged under a fluorescent microscope.

## Wound Healing Test

The cells were digested with trypsin and inoculated evenly into a six-well tissue culture plate 1 day before the scratch and grew to almost complete confluence within 24 h. A sterile 10-μl tip was used to generate an artificial homogeneous wound on a single layer. After scraping, the cells were cleaned in serum-free intermediary. Pictures of cell migration to the wound were collected at 0 and 24 h, and every assay was completed ≥3 times.

## Statistics

SPSS 11.0 statistic program package was utilized for statistical assays, and the data were presented as average ± SD (X ± S). The mean of multiple samples was studied *via* one-way ANOVA. Least—Significant Difference (LSD) method was utilized to contrast the experiment group with the controls. Snk-q test was employed to contrast the mean number of several samples. $P < 0.05$ had significance on statistics. GraphPad software was used to draw relevant figures (Insightful Science company).

# RESULTS

## CXCL12-CXCR4 Axis Participated in the Effect of Sinomenine Hydrochloride on Hepatocellular Carcinoma Cells

Firstly, we detected the effect of SIN on HCC cell viability. SK-Hep1 cells were exposed to SIN at different levels for 48 h, and cellular activity was evaluated *via* MTT assay. **Figure 1A** shows the roles of SIN in the viability of human liver cancer SK-Hep1 cells. The outcomes revealed that the cell viability of SK-Hep1 was significantly reduced by 0.25-mM sinoline hydrochloride for 48 h. CXCL12/CXCR4 axis is widely believed to be vital for many tumors and metastases, including liver carcinoma (13). To determine how SIN affects the proliferative and invasive activities of HCC cells, we subsequently investigated whether SIN regulates the expression of CXCL12/CXCR4 in HCC. As shown in immunoblotting, the expressions of CXCL12 and CXCR4 were remarkably downregulated in SK-Hep1 cells 48 h after 0.25-mM SIN treatment, suggesting that the CXCL12-CXCR4 axis was affected by sinine in HCC cells (**Figure 1B**). Subsequently, we significantly reduced CXCR4 expression by siCXCR4 in SK-Hep1cells, and then SK-Hep1 cells were treated with diverse levels of sinocine hydrochloride for 48 h. Cellular activity was identified *via* CCK-8 analysis. The outcomes revealed that the cell viability of SK-Hep1 cells with CXCR4 knockdown was significantly reduced by comparison (**Figure 1C**). On the contrary, the activity of SK-Hep1 cells overexpressing CXCR4 was remarkably increased (**Figure 1D**). In addition, SK-Hep1 cells transfected with CXCR4 presented a remarkable increase in cellular growth in the presence of CXCL12 (100 ng/ml). The results showed that the combination of CXCL12 and CXCR4 could induce a significant increase in cell

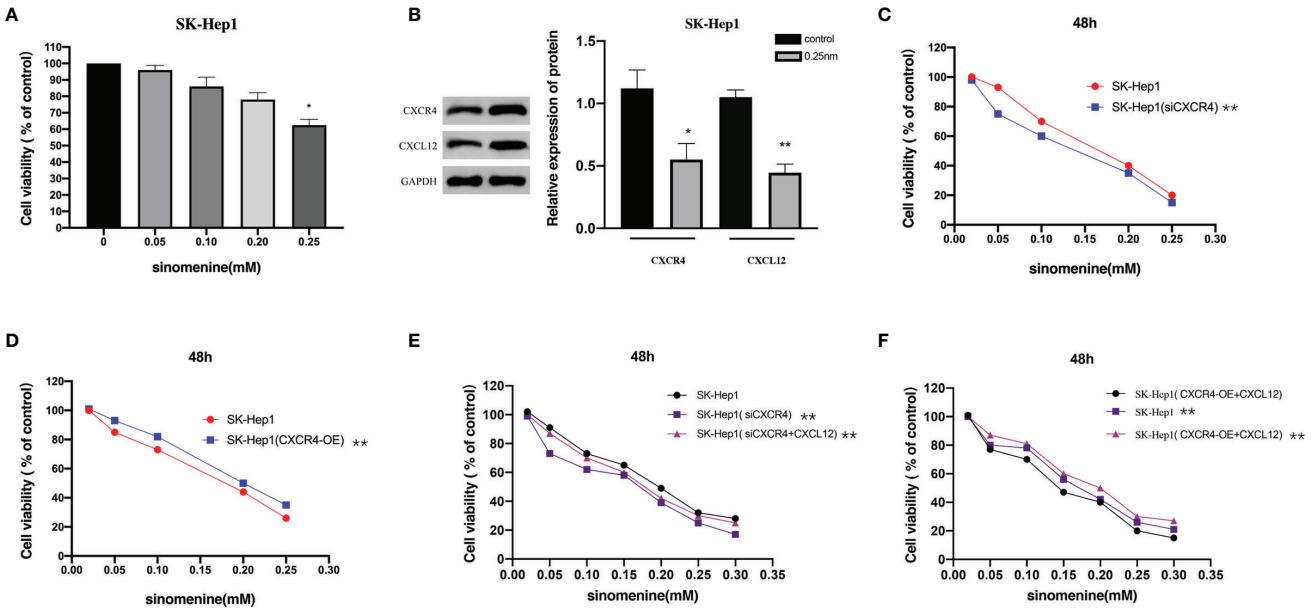



**FIGURE 1** | The CXCL12-CXCR4 axis is involved in the role of sinomenine hydrochloride in the proliferative activity of hepatocellular carcinoma (HCC) cells. **(A)** 3-(4,5)-dimethylthiahiazo (-z-y1)-3,5-di- phenytetrazoliumromide (MTT) assay to assess the viability of SK-Hep1 cells exposed to diverse levels of sinomenine hydrochloride. **(B)** Western blot (WB) to detect the expression of CXCL12 and CXCR4 in SK-Hep1 cells. **(C)** Decrease of CXCR4 expression by siCXCR4 in SK-Hep1 cells, SK-Hep1 cellular activity was identified *via* CCK-8 analysis. **(D)** SK-Hep1 cellular activity was identified in SK-Hep1 cells by overexpressing CXCR4, Cell Counting Kit-8 (CCK-8) assay. **(E, F)** CCK-8 assay measures cell viability. *P < 0.05, **P<0.01.

growth, while SIN repressed the cellular proliferation triggered by combining CXCL12 and CXCR4 to a certain extent (**Figures 1E, F**). These results indicate that SIN suppresses HCC cell growth by regulating the CXCL12-CXCR4 axis.

## CXCL12-CXCR4 Axis Was Involved in the Effect of Sinomenine Hydrochloride on Hepatocellular Carcinoma Cell Apoptosis

Next, we determined whether SIN affected apoptosis of HCC cells through the CXCL12-CXCR4 axis. The expression of apoptosis-related protein Caspase 3 was identified. As unveiled by immunoblotting, the knockout of CXCR4 and sinocine hydrochloride exposure both elevated caspase 3 cleavage in SK-Hep1 cells (**Figure 2A**). Moreover, the highest expression of lysed Caspase 3 was discovered in the sinoline hydrochloride in combination with CXCR4 knockout group in SK-Hep1 cells. By comparison, CXCR4 overexpression decreased sinomenine-triggered cleavage caspase 3 stimulation (**Figure 2B**). The role of CXCR4 modulation in sinomenine-triggered apoptosis in SK-Hep1 cells was studied *via* flow cell technique. SK-Hep1 cells were cultured for 48 hours in the presence of 0.25 mM sinomenine HCl, CXCR4 silenced, or 0.25 mM sinomenine HCl combined with CXCR4 silenced. SK-Hep1 cells were cultured for 48 h under the condition of 0.25-mM SIN and CXCR4 overexpression or 0.25-mM SIN and CXCR4 overexpression. The percentage of apoptosis of SK-Hep1 cells was 8.13 ± 0.45 in SIN group and 10.75 ± 0.29% in SIN in combination with siCXCR4 group, indicating that the knockout of CXCR4 elevated sinomenine-triggered programmed cell death of SK-Hep1 cells (**Figure 2C**). By comparison, the proportion of programmed cell death of SK-Hep1 cells in the SIN group was 30.52 ± 0.74, while the percentage of SK-Hep1 cells in the SIN combined with CXCR4 overexpression group was 26.19% ± 0.28%, indicating that CXCR4 overexpression reduced sinomenine-induced apoptosis of SK-Hep1 cells (**Figure 2D**). Those outcomes suggest that CXCR4 knockout enhances the susceptibility of SK-Hep1 cells to sinoline hydrochloride *via* causing programmed cell death. By comparison, CXCR4 overexpression decreased the susceptibility of HCC cells to sinoline hydrochloride *via* causing programmed cell death.

## CXCL12-CXCR4 Axis Participated in the Effect of Sinomenine Hydrochloride on the Aggression of Hepatocellular Carcinoma Cells

To investigate the effects of CXCR4 on sinoline-induced invasion of SK-Hep1 cells, Transwell invasion assay with matrix gel coating was used. In the SIN (0.25 mM) treatment group, the aggression of SK-Hep1 cells was significantly reduced, while CXCR4 knockdown further intensified the invasion of SK-Hep1 cells induced by SIN (**Figure 3A**). In contrast, CXCR4 overexpression reversed the decreased aggression of SK-Hep1 cells induced by sinoline hydrochloride (**Figure 3B**). Therefore, SIN repressed the activity and invasive ability of HCC cells and increased the apoptosis of HCC cells. In conclusion, CXCL12-

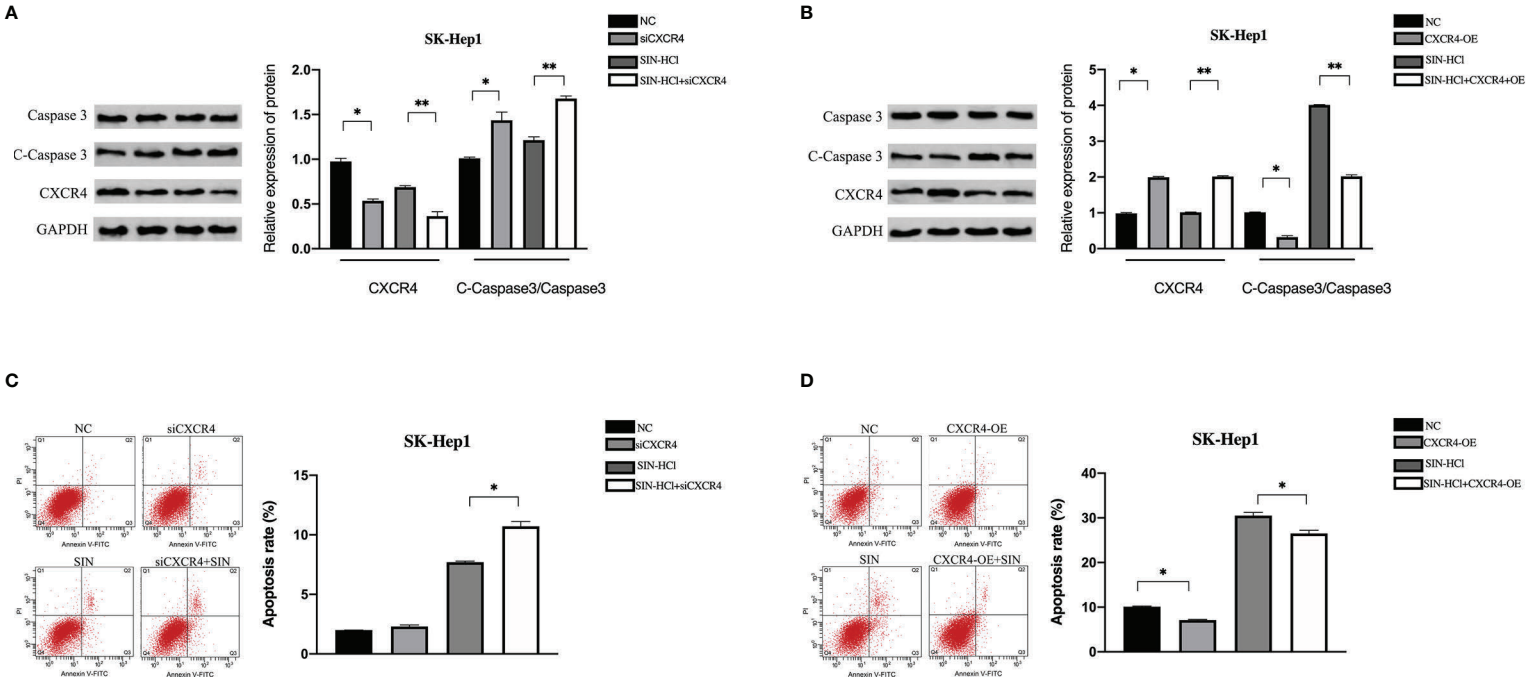



**FIGURE 2** | CXCL12-CXCR4 axis was involved in the effect of sinomenine hydrochloride on HCC cell apoptosis. CXCR4 knockdown and overexpression were detected in SK-Hep1 cells. **(A, B)** WB was utilized to identify the expression of apoptotic protein in SK-Hep1 cells. **(C, D)** Programmed cell death was identified *via* flow cell technique. *P < 0.05, **P < 0.01.

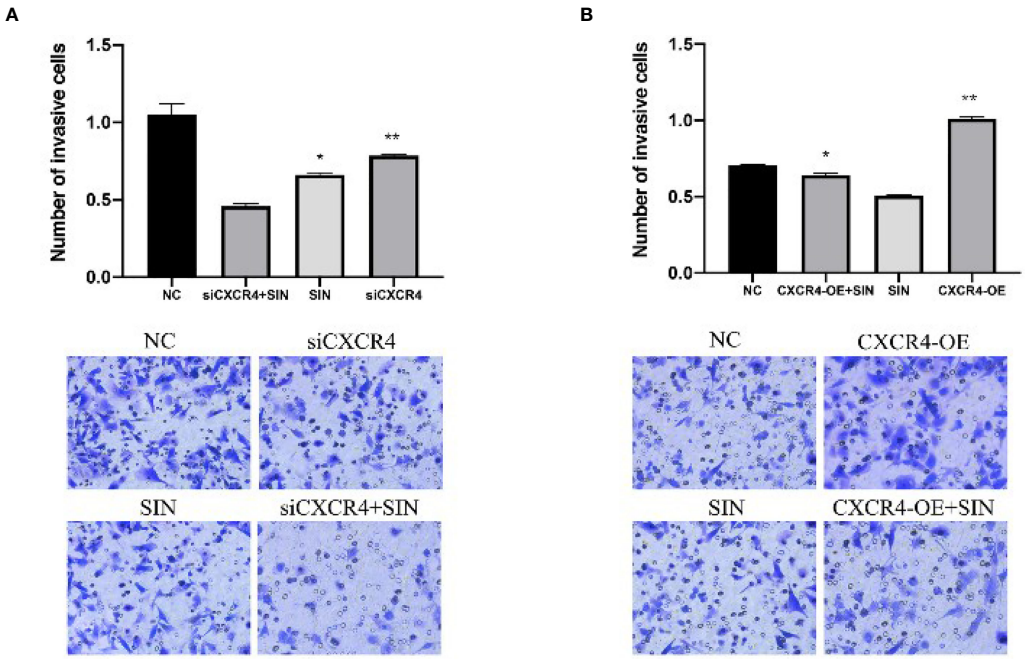



**FIGURE 3** | The CXCL12-CXCR4 axis participates in the effect of sinomenine hydrochloride on HCC cell aggression. CXCR4 knockdown and overexpression in SK-Hep1 cells. **(A, B)** Transwell assay for cell invasion (×200). *P < 0.05, **P < 0.01.

CXCR4 axis is involved in the role of SIN in HCC cellular viability, invasion, and programmed cell death.

## The CCL21-CCR7 Axis Participated in the Effect of Sinomenine Hydrochloride on Hepatocellular Carcinoma Cells

CC chemokine ligand 21/chemokine receptor 7 (CCL21/CCR7) facilitates growth and metastases of a variety of cancers (14). This prompted us to explore if SIN also modulates the expression of CCL21/CCR7 in HCC. WB confirmed that CXCR7 participated in the modulation of HCC cells by SIN. As presented by **Figure 4A**, the expressions of CCR7 and CCL21 were remarkably regulated downward in SK-Hep1 cells after SIN (0.25 mM) treatment. To characterize the effects of CCR7 on HCC, our team downregulated or overexpressed CCR7 in SK-Hep1 cells. CCK-8 showed that SK-Hep1 cells with CCR7 knockdown grew more slowly, while CXCR7-overexpressed cells proliferated more rapidly than control cells (**Figures 4B, C**). In addition, SK-Hep1 cells transfected with CCR7 displayed a remarkable increase in cellular growth in the presence of CCL21 (80 ng/ml). The results showed that the binding of CCL21 and CCR7 could induce a significant increase in cell growth, while SIN repressed the cellular proliferation triggered *via* the binding of CCL21 and CCR7 to a certain extent (**Figures 4D, E**). Those outcomes reveal that SIN also regulates the CCL21-CCR7 axis in HCC cells to suppress the growth of oncocytes. Afterward, whether SIN also influences HCC cellular apoptosis through the CCL21-CCR7 axis was determined. WB analysis showed that both CCR7 silencing and sinobine hydrochloride treatment increased caspase 3 cleavage in SK-Hep1 cells (**Figure 4F**). In addition, the highest expression of lysed Caspase 3 was discovered in the sinoline hydrochloride in combination with CCR7 knockout group in SK-Hep1 cells. By comparison, CCR7 overexpression reduces sinomenine hydrochloride-induced activation of cleaved caspase 3 (**Figure 4G**). The effect of CCR7 regulation on sinoline-triggered programmed cell death in SK-Hep1 cells was studied by flow cytometry (**Figure 4H**). SK-Hep1 cells were cultivated for 48 h under the condition of 0.25 mM SIN, CCR7 knockdown, or 0.25 mM SIN combined with CCR7 knockdown. SK-Hep1 cells were overexpressed with SIN and CCR7 at 0.25 mM or SIN combined with CCR7 at 0.25 mM for 48 h. The percentage of apoptosis of SK-Hep1 cells was 7.52% ± 0.42% in SIN group and 8.02% ± 0.37% in SIN in combination with siCXCR4 group, indicating that CCR7 knockout elevated SIN-triggered programmed cell death of SK-Hep1 cells. Apoptosis assay showed that CCR7 overexpression reduced sinoline-triggered programmed cell death of SK-Hep1 cells. By comparison, apoptosis rate was remarkably elevated when CCR7 was silenced in SK-Hep1 cells. Transwell analysis showed that CCR7 knockdown attenuated the invasiveness of SK-Hep1 cells (**Figure 4I**). On the contrary, CCR7 overexpression increased SIN-induced decreased aggression of SK-Hep1 cells. Our results confirm that the CCL21-CCR7 axis is vital for the regulation of sinoline hydrochloride-affected HCC cell proliferation, invasiveness, and apoptosis.

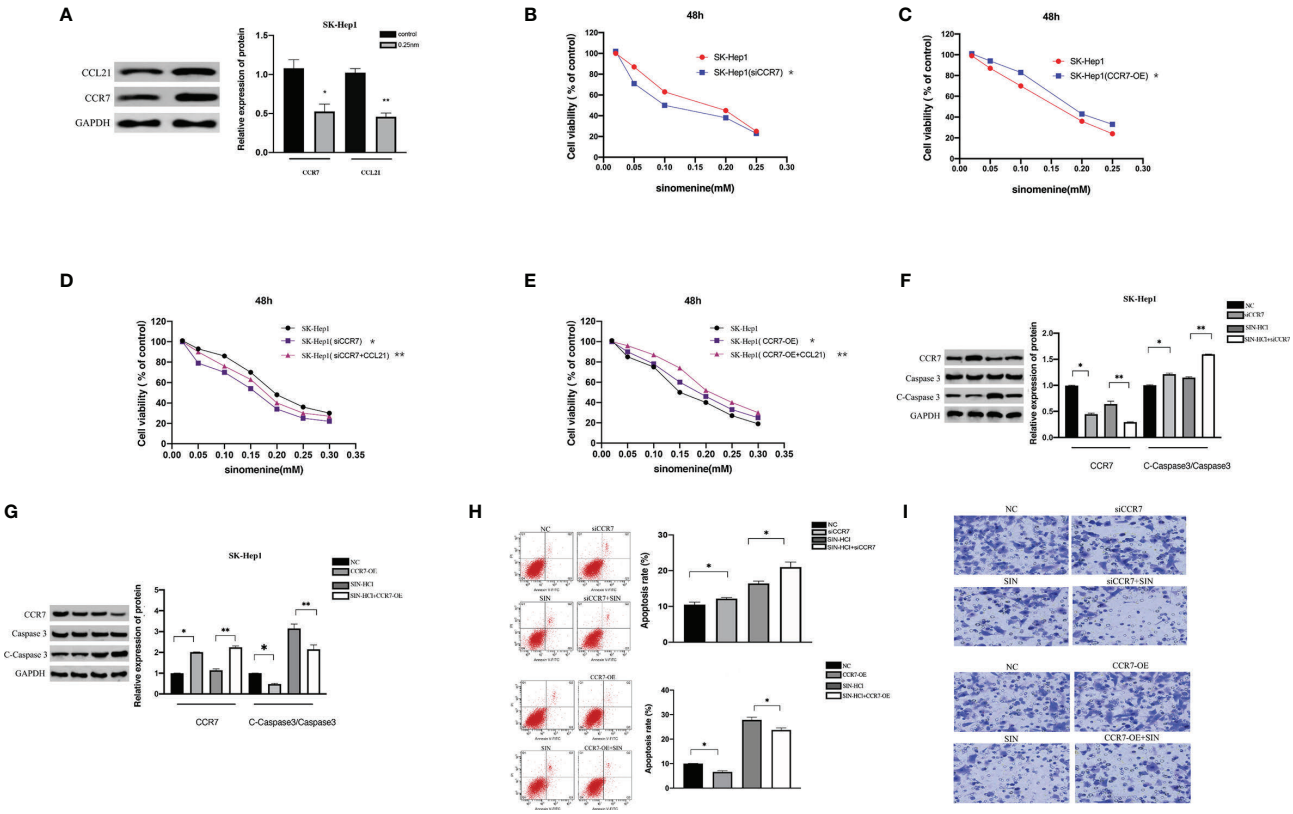



**FIGURE 4** | The CCL21-CCR7 axis participates in the effect of sinomenine hydrochloride on HCC cells. **(A)** WB assay of the expression of CCR7 and CCL21 in SK-Hep1 cells. **(B, C)** CCK-8 experiments examine the effect of CCR7 knockout and CXCR7-overexpressing cells on cell proliferation. **(D, E)** CCK-8 experiments on the effect of CCR7-transfected SK-Hep1 cells on cell proliferation in the presence of CCL21 (80 ng/ml). **(F, G)** WB assay of apoptosis-related protein expression in SK-Hep1 cells. **(H)** Apoptosis was detected by flow cytometry. **(I)** Transwell assay for cell invasion (×200). *P < 0.05, **P < 0.01.

## Tumor-Derived DNA Can Activate CXCL12-CXCR4 and CCL21-CCR7 Axes of Hepatocellular Carcinoma Cells, While Sinoline Hydrochloride Can Inhibit CXCL12-CXCR4 and CCL21-CCR7 Axes of Hepatocellular Carcinoma Cells

Cancer-originated DNA was separated from Lewis cell line. Tumor-derived DNA was transfected into SK-Hep1 cells by Lipofectamine 2000. As revealed by WB analyses, the expression of CXCL12, CXCR4, CCL21, and CCR7 proteins in SK-Hep1 cells was significantly increased after transfection with tumor-derived DNA in contrast to the blank control group (**Figure 5A**). The outcomes showed that tumor-derived DNA activated the CXCL12-CXCR4 and CCL21-CCR7 axes in HCC, while sinoline hydrochloride inhibited the CXCl12-CXCR4 and CCL21-CCR7 axes in SK-Hep1 cells. Next, we investigated whether tumor-derived DNA and SIN participated in the modulation of epithelial–mesenchymal transition (EMT) in HCC cells. Cell morphological results showed that tumor-derived DNA increased the mesenchymal morphology of SK-Hep1 cells, and SIN reduced this morphological change (**Figure 5B**). WB was further investigated to determine whether tumor-derived DNA and SIN affected the protein expression levels of E-cadherin, N-cadherin, and vimentin. As shown in **Figure 5C**, tumor-derived DNA significantly downregulated the expression of E-cadherin and upregulated the expression of N-cadherin and vimentin at the protein level. However, sinobine hydrochloride could reverse the effect of tumor-derived DNA on EMT-related markers in SK-Hep1 cells. Representative immunofluorescence was shown in **Figure 5D**, which showed results similar to WB. Transwell analysis showed that tumor-derived DNA remarkably elevated the invasiveness of SK-Hep1 cells, while sinobine hydrochloride reversed the upregulation of tumor-derived DNA on the invasiveness of SK-Hep1 cells (**Figure 5E**). Scratch test results showed that tumor-derived DNA increased HCC cell migration, while sinoline hydrochloride inhibited HCC cell migration *in vitro* (**Figure 5F**). In conclusion, our results indicate that tumor-derived DNA regulates the CXCL12-CXCR4 axis and CCL21-CCR7 axis, which directly affects the metastasis and aggression of HCC cells, while sinoline hydrochloride can reverse the effects of tumor-derived DNA on HCC cells.

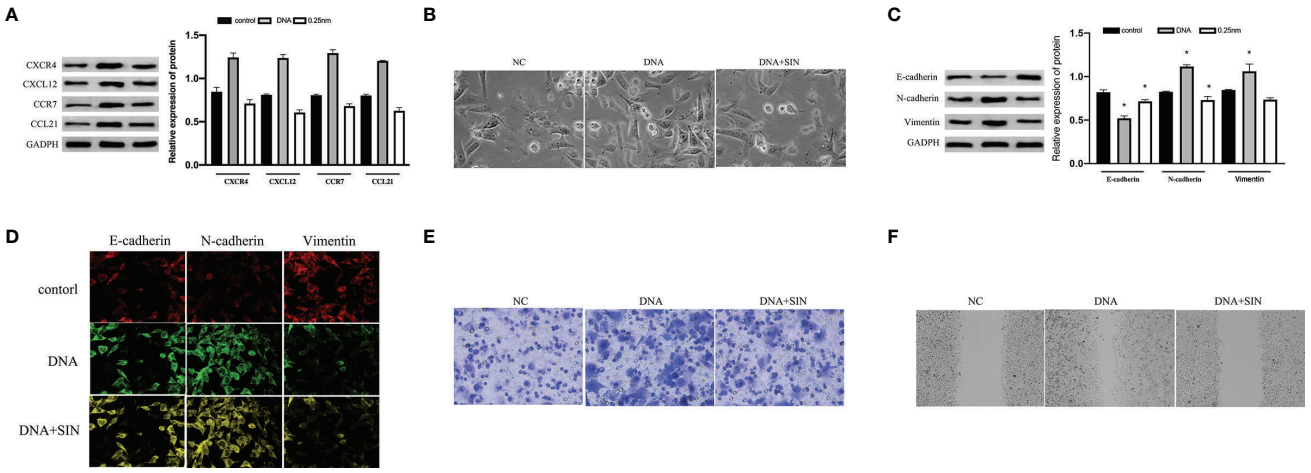



**FIGURE 5** | Tumor-derived DNA can activate the CXCL12-CXCR4 and CCL21-CCR7 axes of HCC cells, while sinomenine hydrochloride has the effect of inhibiting the CXCL12-CXCR4 and CCL21-CCR7 axes of HCC cells. **(A)** The protein expression of CXCL12, CXCR4, CCL21, and CCR7 was examined by Western blot (WB) analysis after transfection of tumor-derived DNA. **(B)** Cytomorphological analysis (×200). **(C)** WB assay to examine the expression of E-cadherin, N-cadherin, and vimentin. **(D)** Immunofluorescence for E-cadherin, N-cadherin, and vimentin (×200). **(E)** Transwell analysis of HCC cell invasion (×200). **(F)** Scratch assay to examine HCC cell migration (×200). *P < 0.05.

## Sinomenine Hydrochloride Interferes With Tumor-Derived DNA and Affects ERK/MMP Signaling Pathway Through CXCL12/CXCR4 Axis in Hepatocellular Carcinoma Cells

The effects of mitogen-activated protein kinase (MAPK)/ERK path on tumor has been widely confirmed, which not only participates in oncocyte proliferative activity and angiogenesis but also is vital for cancer aggression and migration (15). Research has revealed that the CXCL12/CXCR4 axis is implicated in a variety of biofunctions of cancers *via* stimulating ERK1/2 (16). For the sake of exploring the effects and causal links of tumor-derived DNA in HCC cell metastasis. We examined the effect of tumor-derived DNA and sinomenine hydrochloride on the phosphorylation of ERK 1/2 in HCC cells and further studied its regulation of MMP expression in HCC cells. The results showed that tumor-derived DNA-treated SK-Hep1 cells exhibited significantly higher P-ErK1/2 and MMP2/9 activity, while sinobine hydrochloride reduced this increase (**Figure 6A**). In addition, tumor-derived DNA remarkably regulated protein expression levels of P-ERK1/2 and MMP2/9 upward, whereas sinobine hydrochloride reduced this upregulation (**Figure 6B**, $P < 0.05$). Those outcomes revealed that tumor-derived DNA activated ERK/MMP signaling pathway in HCC cells, and sinoline hydrochloride significantly inhibited the ERK/MMP signaling pathway activated by tumor-derived DNA. We further examined whether tumor-derived DNA influenced the ERK/MMP pathway in HCC cells *via* the CXCL12-CXCR4 chemokine axis. As shown in **Figure 6C**, the expressions of P-ErK1/2, MMP2, and MMP9 in SK-Hep1 cells were remarkably reduced after siCXCR4 transfection in contrast to the group without siCXCR4 transfection (DNA group and siCXCR4 +DNA group, DNA+SIN group, and siCXCR4+DNA+SIN group were compared respectively). Compared with the untreated sinomenine group, the protein expressions of P-ErK1/2, MMP2, and MMP9 were further inhibited in the sinomenine group (DNA group and DNA+sinomenine group, siCXCR4+DNA group, and siCXCR4+DNA +sinomenine group). We further transfected SK-Hep1 cell lines with CXCR4 plasmid to determine whether tumor-derived DNA and the effect of sinine hydrochloride on HCC cells were related to the CXCL12-CXCR4 chemokine axis. After transfection of tumor-derived DNA, the expression of CXCL12-CXCR4 axis, P-ERK1/2, and MMP2/9 proteins was further significantly increased under CXCR4 overexpression. SIN inhibited this increase to some extent (**Figure 6D**).

## Sinomenine Hydrochloride Interferes With Tumor-Derived DNA and Affects the ERK/MMP Signaling Pathway Through the CCL21-CCR7 Axis in Hepatocellular Carcinoma Cells

The CXCL12/CXCR4 axis has been reported to be involved in a variety of biofunctions of cancers *via* stimulating ERK1/2. Recent reports suggest that CCR7 facilitates cellular proliferative activity *via* the ERK path. WB analyses were completed to explore whether tumor-derived DNA and sinobine hydrochloride modulate ERK/MMP signaling pathway by regulating the CCL21-CCR7 axis in HCC cells.

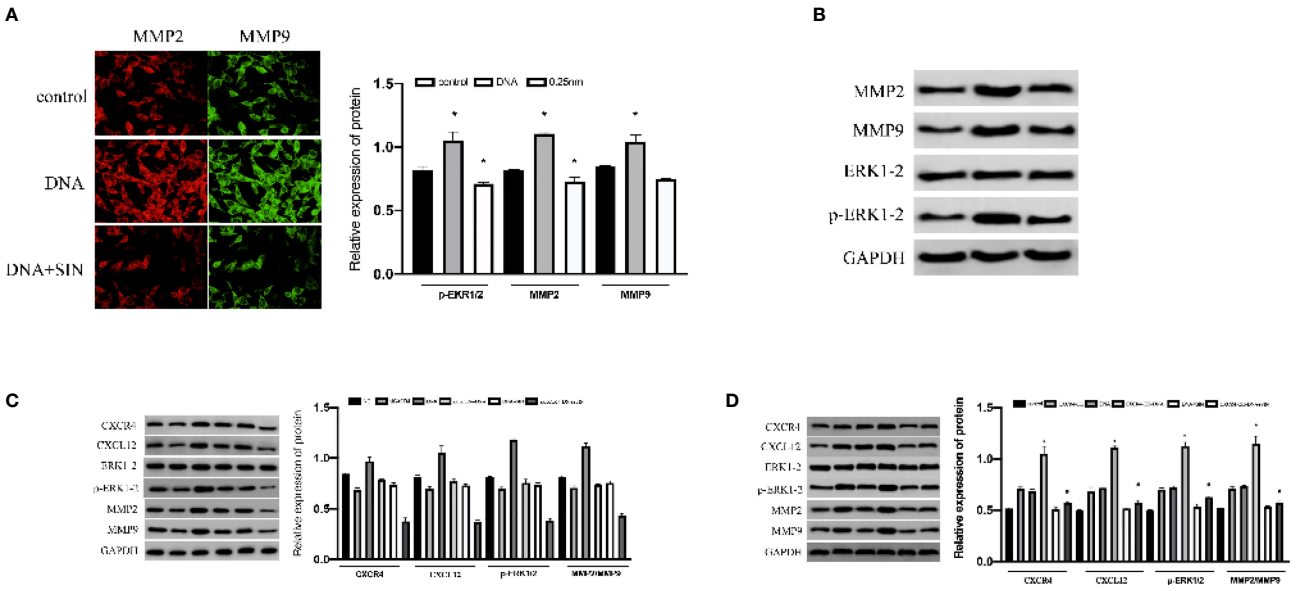



**FIGURE 6** | Interference of tumor-derived DNA with sinomenine hydrochloride affects the ERK/MMP signaling pathway through the CXCL12/CXCR4 axis in HCC cells. **(A)** Immunofluorescence examination of the expression of p-ERK1/2 and MMP2/9 in cells (×200). **(B)** WB assay to examine the expression of p-ERK1/2 and MMP2/9 in cells. **(C)** The expression of p-ERK1/2, MMP2, and MMP9 proteins in SK-Hep1 cells after siCXCR4 transfection. **(D)** The expression of p-ERK1/2, MMP2, and MMP9 proteins in SK-Hep1 cells after CXCR4-OE transfection. *P < 0.05.

To determine whether the CCL21-CCR7 axis participates in the modulation of tumor-derived DNA and sinobine hydrochloride on ERK1/2 and MMP2/9 signaling pathways, CCR7 was knocked down by siRNA in HCC cells to detect the expression changes of ERK1/2 and MMP2/9. The loss of CCR7 expression leads to the decline of P-ERK1/2 and MMP2/9 (**Figure 7A**), suggesting that CCR7 regulates p-ERK1/2 and MMP2/9 signal paths. The expression of P-ERK1/2, MMP2, and MMP9 protein was further inhibited by sinoline hydrochloride treatment. We also verified the above results from the opposite direction by overexpressing CCR7 in HCC cells (**Figure 7B**). The expression of P-ERK1/2 and MMP2/9 was upregulated after transfection with CCR7 plasmid (**Figure 7B**). Those outcomes reveal that sinobine hydrochloride reverses the role of tumor-derived DNA in HCC cells *via* stimulating the ERK1/2/MMP2/9 signal path regulated by the CCL21-CCR7 axis.

## Tumor-Derived DNA Promoted the Invasion of Hepatocellular Carcinoma Cells Through the CXCL12-CXCR4 Axis and the CCL21-CCR7 Axis, and Sinoline Hydrochloride Had a Suppressive Effect on the Invasion of Hepatocellular Carcinoma Cells

To evaluate the role of tumor-derived DNA in tumor metastasis, Transwell matrix gel invasion was assayed using SK-Hep1 cells with CXCR4 overexpression or silencing. The results showed that tumor-derived DNA promoted SK-Hep1 cell invasion, while CXCR4 knockdown reduced SK-Hep1 cell invasion induced by tumor-derived DNA (**Figure 8A**). Similarly, CCR7 knockdown also reduced SK-Hep1 cell invasion by tumor-derived DNA (**Figure 8B**). SIN significantly suppressed the aggression of SK-Hep1 cells, just as it blocked the CXCL12/CXCR4 axis or CCL21-CCR7 axis. However, the addition of CXCR4 overexpression and CXCL12 further enhanced the increased migration capacity of SK-Hep1 cells induced by tumor-derived DNA (**Figure 8C**). Consistent with the above results, the overexpression of CCR7 and the addition of CCL21 significantly enhanced the invasion ability of SK-Hep1 cells (**Figure 8D**). The addition of SIN repressed CXCR4 overexpression or CCR7 overexpression induced enhanced cell migration. These results further confirmed the previous experimental results that tumor-derived DNA enhanced the invasive ability of SK-Hep1 cells through the CXCL12-CXCR4 axis and the CCL21-CCR7 axis.

## Tumor-Derived DNA Reduced Apoptosis of Hepatocellular Carcinoma Cells by CXCL12-CXCR4 Axis and CCL21-CCR7 Axis, and Sinoline Hydrochloride Reversed the Decrease of Apoptosis

This study subsequently evaluated whether the CXCL12-CXCR4 axis and CCL21-CCR7 axis were involved in the regulation of tumor-derived DNA on apoptosis in HCC cells.

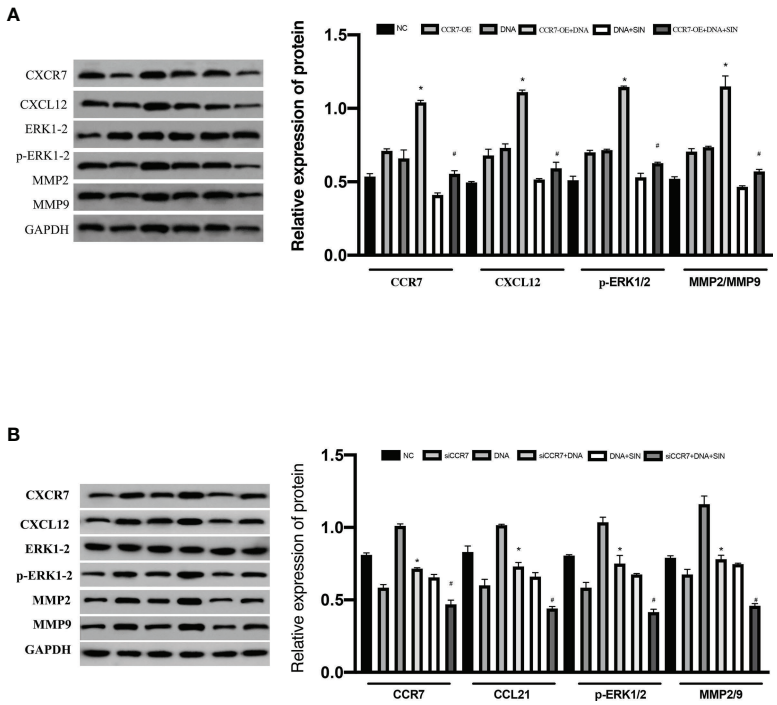



**FIGURE 7** | Interfering with tumor-derived DNA by sinomenine hydrochloride affects the ERK/MMP signaling pathway through the CCL21-CCR7 axis in HCC cells. **(A)** The expression of p-ERK1/2, MMP2, and MMP9 proteins in SK-Hep1 cells after siCCR7 transfection. **(B)** The expression of p-ERK1/2, MMP2, and MMP9 proteins in SK-Hep1 cells after CCR7-OE transfection. *P < 0.05(compared with NC), #P<0.05(compared with DNA+SIN group).

Flow cytometry analysis showed that siCXCR4 or siCCR7 increased programmed cell death of SK-Hep1 cells in the presence of tumor-derived DNA (**Figures 9A, B**). Sinoline hydrochloride combined with siCXCR4 or siCCR7 could further increase the programmed cell death of SK-Hep1 cells. In order to reveal in detail that sinobine hydrochloride reversed tumor-derived DNA-induced apoptosis reduction of HCC cells, we examined the protein Caspase 3 levels, and the results were consistent with flow cytometry results, suggesting that siCXCR4 or siCCR7 reversed tumor-derived DNA-induced apoptosis reduction of HCC cells. The addition of SIN further increased HCC cell apoptosis (**Figures 9C, D**). As shown in **Figures 9E** and **F**, apoptosis of SK-Hep1cells was significantly reduced after treatment with tumor-derived DNA combined with CXCR4 or CCR7 overexpression but increased after treatment with sinocine hydrochloride. WB analyses were completed to identify the level of Caspase 3, a protein associated with apoptosis, and the results were consistent with flow cytometry results, indicating that CXCR4 or CCR7 overexpression reduced tumor-derived DNA-induced apoptosis of HCC cells, while sinocine hydrochloride increased Caspase 3 levels (**Figures 9G, H**). In conclusion, these results further support the previous experimental results that tumor-derived DNA reduces the apoptosis of HCC cells through the CXCL12-CXCR4 and CCL21-CCR7 axes, and the addition of sinobine hydrochloride can reverse this reduction.

## DISCUSSION

HCC is a challenging and lethal malignancy with increasing incidence (17). Exogenous or improperly positioned endogenous DNA can stimulate immune responses, and the molecular mechanisms of DNA recognition have received increasing attention. Tumor-derived DNA is fragmented DNA of tumor-derived DNA found in human blood and other biological fluids (18, 19). It carries the complete genome information of tumor-specific genetic changes, which can be detected by detecting the genetic changes associated with the tumor. In recent years, increasing evidence has shown that tumor DNA plays a key role in the pathogenesis of various tumors (20). The mechanism by which DNA enters the cytoplasm of Dendritic Cells (DCs) under physiological conditions *in vivo* is one of the unresolved questions in our current work. Here, our team intends to study and focus on the mechanism by which tumor-derived DNA acts as a carcinogen of liver cancer to promote cancer progression *in vivo* (21, 22). Chemokine signaling has long been involved in cancer progression and metastases through autocrine or paracrine mechanisms. Importantly, in previous studies, the chemokine receptor CXCR4 was found to be overexpressed in HCC samples and HCC cells and was found to enhance HCC growth and invasiveness (17, 23). CXCR4 overexpression may be attracted to specific organs by

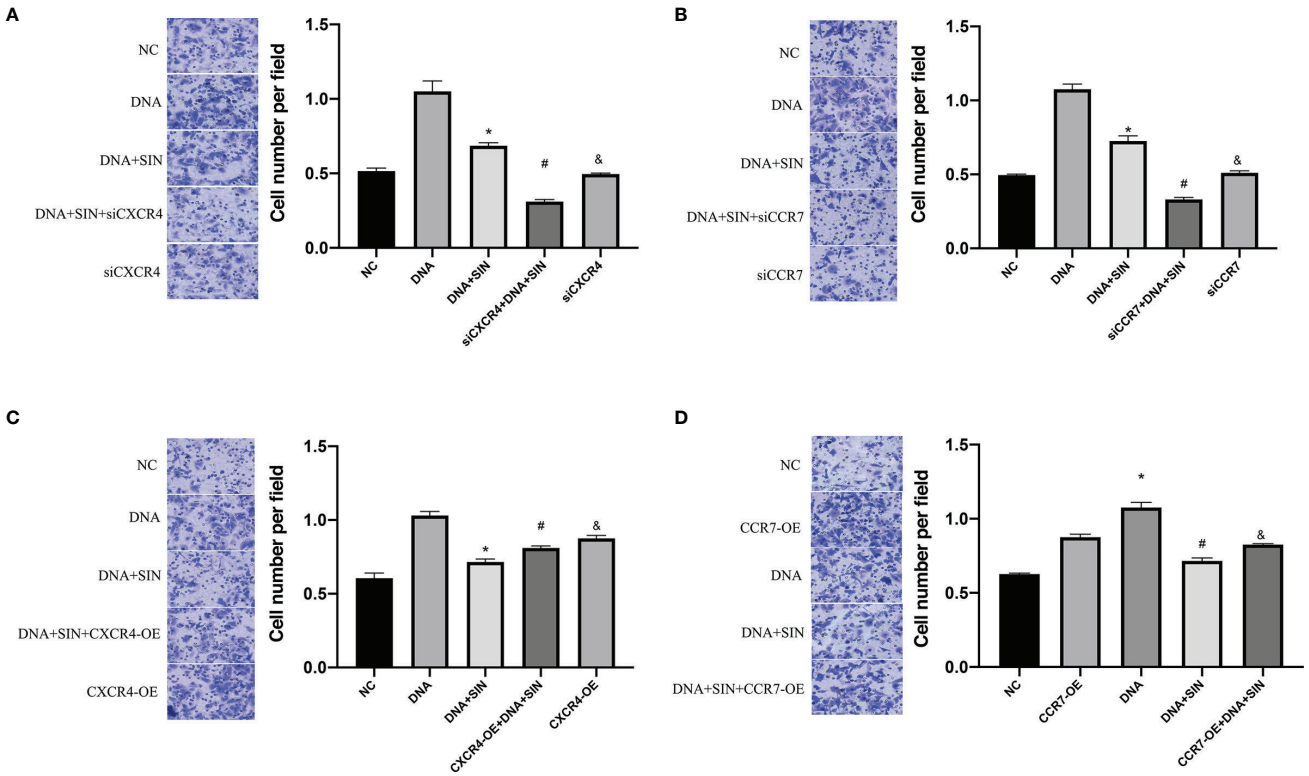



**FIGURE 8** | Tumor-derived DNA promotes cell invasion of HCC cells through the CXCL12-CXCR4 axis and CCL21-CCR7 axis, and sinomenine hydrochloride has an anti-invasive effect on HCC cells. **(A, B)** Transwell Matrigel invasion assay of CXCR4 or CCR7 knockdown on cell invasion (×200). **(C, D)** Transwell Matrigel invasion assay on cell invasive ability of CXCR4 or CCR7 overexpression (×200). *P <0.05(compared with NC), #P<0.05(compared with DNA+SIN group), &P<0.05 (compared with siCXCR4+DNA+SIN, or compared with CXCR4-OE+DNA+SIN).

its ligand CXCL12, which together lead to organ-specific metastasis. The CXCL12-CXCR4 axis promotes tumor growth and angiogenesis and enhances invasion and migration. Our preliminary results indicate that CXCR4 and CXCL12 are highly expressed in HCC cells. Although the CXCL12-CXCR4 axis is critical for HCC, it is involved in different cellular pathways under tissue heterogeneity and different cellular contexts. To date, the biological function of tumor-derived DNA on the CXCL12-CXCR4 axis in HCC has not been reported. Here, we report a novel finding that tumor-derived DNA upregulates the CXCL12-CXCR4 axis in HCC, activates the ERK/MMP signaling pathway, and promotes the proliferation and metastasis of HCC cells.

SIN is an alkaloid monomer extracted from the dried stems of Sinomenine or Sinomenine, is the hydrochloride form of sinomenine. It has a wide range of pharmacological activities and has significant anti-inflammatory, anti-immunity, and antitumor effects. Recently, with the advancement of SIN research, certain new pharmacological effects and mechanisms of sinomenine have been discovered, especially its antitumor effects have been paid more and more attention by scholars at home and abroad. Here we show that SIN can inhibit the upregulation of CXCL12-CXCR4 axis induced by tumor-derived DNA in HCC cells. This research offers enlightenment pertaining to the modulation of CXCL12-CXCR4 axis and HCC cell proliferation and metastasis by SIN. First, the effects of several doses of SIN on HCC cell survival/proliferative activity were evaluated. The outcomes unveiled that SIN exerted suppressive effects on the survival/proliferation of SK-Hep1 cells, and 0.25-mm SIN had the maximum inhibitory effect. Apoptosis is also a major determinant of cell survival/proliferation. Lysed Caspase 3 is considered to be a key executor of apoptosis and can induce apoptosis.

Herein, our team discovered that sinobine hydrochloride elevated the protein level of proapoptotic activator caspase 3, and increased apoptosis was also observed by drain cytology, suggesting that the effect of proapoptotic activator sinobine hydrochloride was related to the regulation of apoptosis-related regulatory factors. CXCL12-CXCR4 axis is overexpressed in HCC cells, and downregulation of CXCL12-CXCR4 axis is beneficial to inhibit cancer growth. The outcomes herein unraveled that SIN reduced the protein levels of CXCL12, CXCR4, CCL21, and CCR7 in HCC cells. Our team discovered that cancer-originated DNA can activate the CXCL12-CXCR4 and CCL21-CCR7 axes in HCC cells, while sinoline treatment

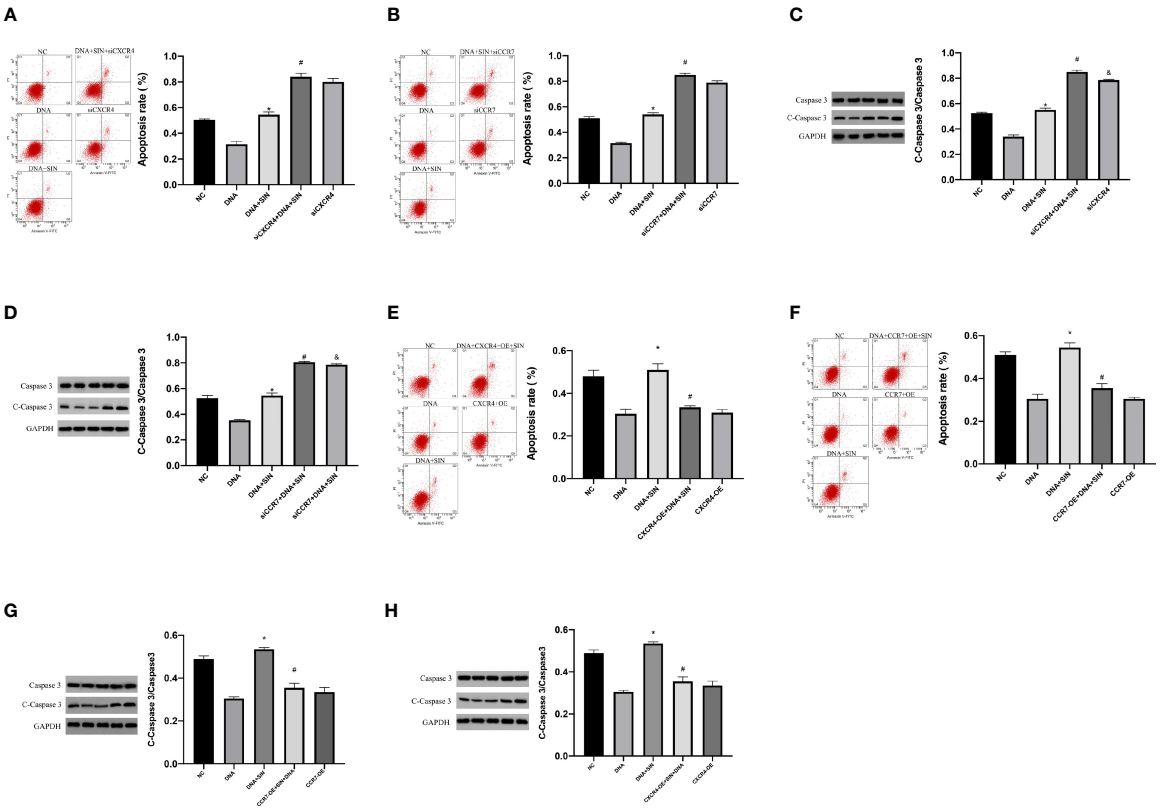



**FIGURE 9** | Tumor-derived DNA reduces apoptosis in HCC cells *via* CXCL12-CXCR4 axis and CCL21-CCR7 axis, and sinomenine hydrochloride reverses the reduction in apoptosis. **(A, B)** Flow cytometry to identify the role of siCXCR4 or siCCR7 in the programmed cell death of SK-Hep1 cells in the presence of tumor-derived DNA. **(C, D)** WB assay to identify the expression of apoptotic protein in cells. **(E, F)** The effect of CXCR4 or CCR7 overexpression on the programmed cell death of SK-Hep1 cells in the presence of tumor-derived DNA was detected by flow cytometry. **(G, H)** WB assay to identify the expression of apoptotic protein in cells. *P <0.05(compared with NC), #P<0.05 (compared with DNA+SIN group), &P<0.05 (compared with siCXCR4+DNA+SIN, or compared with CXCR4-OE +DNA7+SIN).

reduced the expressions of CXCL12, CXCR4, CCL21, and CCR7 in HCC cells. The MAPK pathway is a prominent mechanism for coordinating transcription factors that are activated by ERK kinase to modulate downstream MMP-2 and MMP-9 genetic expression and afterward elevate cancer onset and development. The aggression and migration of cancer are complicated processes with many factors and steps. MMP-mediated decomposition of exocellular matrixes and basement membranes is a pivotal step in cancer aggression and migration. It has been reported that oral sinobine hydrochloride significantly blocks the stimulation of ERK1/2 in the spinal cord (24). In addition, SIN can effectively reduce the metastases of mankind glioblastoma *via* inhibiting the MMP2/9 signaling pathway (25). Our team discovered that cancer-originated DNA can promote the stimulation of ERK1/2 and MMP2/9 in HCC cells, while cotreatment with sinobine hydrochloride reduces the expression of ERK1/2 and MMP2/9. Brand et al. (26) showed that chemokine CXCL12 can induce colon cancer cells to secrete MMP9 through its specific receptor CXCR4. CXCR4 cellular mutations activate Akt and ERK signaling and promote varying degrees of drug resistance to ibrutinib, bendamustine, fludarabine, bortezomib, and edralicib (27). CXCL12 can also enhance MMP9 secretion of non-Hodgkin's lymphoma (NHL) cells and promote extracellular matrix degradation, thus participating in cell invasion (28).

To further study the mechanism of CXCL12-CXCR4 induced by tumor-derived DNA promoting the aggression of HCC cells *in vitro*. Herein, the effects of CXCR4 knockdown and overexpression on the expression of ERK1/2 and MMP2/9 were also observed. It was found that CXCR4 knockdown downregulated the level of ERK1/2 in SK-Hep1 cells, enhanced the expression and secretion of MMP2/9, and degraded exocellular matrixes and basement membranes, thus promoting the invasion and migration of HCC. In addition, the proapoptotic protein Caspase 3 and apoptosis rate were also significantly downregulated in HCC cells after CXCR4 knockdown and sinoline hydrochloride treatment, suggesting that CXCR4 knockdown also promoted HCC cell apoptosis. CXCR4 knockdown and sinobine hydrochloride cotreatment also further decreased the expression of ERK1/2 and MMP2/9, further repressed the proliferative and migratory activities of SK-Hep1 cells, and further facilitated the programmed cell

death of HCC cells. On the contrary, CXCR4 overexpression increased the tumor-derived DNA expression of ERK1/2 and MMP2/9, increased the proliferative and migratory activities of SK-Hep1 cells, and reduced apoptosis. Therefore, regulation of ERK1/2/MMP2/9 signaling pathway may be one of the mechanisms of CXCL12-CXCR4 axis or CCL21-CCR7C axis inhibition and sinoline hydrochloride synergistic anti-HCC. On the contrary, CXCR4 overexpression led to further decreased expression of ERK1/2 and MMP2/9, which also proved that CXCL12-CXCR4 axis or CCL21-CCR7C axis inhibition and sinoline hydrochloride had a synergistic anti-HCC effect.

## CONCLUSION

These results suggest that inhibition of tumor-derived DNA-induced activation of CXCL12-CXCR4 or CCL21-CCR7C axis has a synergistic anti-HCC effect with sinoline hydrochloride.

## 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

We contributed equally for this work. All authors contributed to the article and approved the submitted version.

## FUNDING

Clinical Research Plan of SHDC (No.SHDC2020CR2021B).

## REFERENCES


1. Anwanwan D, Singh SK, Singh S, Saikam V, Singh R. Challenges in Liver Cancer and Possible Treatment Approaches. *Biochim Biophys Acta (BBA) - Rev Cancer* (2019) 1873(1):188314. doi: 10.1016/j.bbcan.2019.188314
2. Zhou J, Zhou F, Wang W, Zhang X, Ji Y, Zhang P, et al. Epidemiological Features of NAFLD From 1999 to 2018 in China. *Hepatology* (2020) 71 (5):1851–64. doi: 10.1002/hep.31150
3. Huschka H, Mihm S. Interferon-Lambda (IFNL) Germline Variations and Their Significance for HCC and PDAC Progression: An Analysis of The Cancer Genome Atlas (TCGA) Data. *BMC Cancer* (2020) 20(11):1131. doi: 10.1186/s12885-020-07589-4
4. Ren FH, Yang H, He R, Lu J, Lin X, Liang H, et al. Analysis of Microarrays of miR-34a and Its Identification of Prospective Target Gene Signature in Hepatocellular Carcinoma. *BMC Cancer* (2018) 18(1):12. doi: 10.1186/s12885-017-3941-x
5. Maruyama H, Yamaguchi T, Nagamatsu H, Shiina S. AI-Based Radiological Imaging for HCC: Current Status and Future of Ultrasound. *Diagnostics (Basel)* (2021) 11(2):292. doi: 10.3390/diagnostics11020292
6. Woo SR, Fuertes MB, Corrales L, Spranger S, Furdyna MJ, Leung MYK, et al. STING-Dependent Cytosolic DNA Sensing Mediates Innate Immune Recognition of Immunogenic Tumors. *Immunity* (2014) 41(5):830–42. doi: 10.1016/j.immuni.2014.10.017
7. Cai C, Guo Z, Cao Y, Zhang W, Chen Y. A Dual Biomarker Detection Platform for Quantitating Circulating Tumor DNA (ctDNA). *Nanotheranostics* (2018) 2(1):12–20. doi: 10.7150/ntno.22419
8. Dhayat SA, Yang Z. Impact of Circulating Tumor DNA in Hepatocellular and Pancreatic Carcinomas. *J Cancer Res Clin Oncol* (2020) 146(3):1625–45. doi: 10.1007/s00432-020-03219-5
9. Li X, Li P, Liu C, Ren Y, Tang X, Wang K, et al. Sinomenine Hydrochloride Inhibits Breast Cancer Metastasis by Attenuating Inflammation-Related Epithelial-Mesenchymal Transition and Cancer Stemness. *Oncotarget.* (2017) 8(8):13560–74. doi: 10.18632/oncotarget.14593
10. Deng F, Ma Y, Liang L, Zhang P, Feng J. The Pro-Apoptosis Effect of Sinomenine in Renal Carcinoma via Inducing Autophagy Through Inactivating PI3K/AKT/mTOR Pathway. *Biomed Pharmacother.* (2018) 97:1269–74. doi: 10.1016/j.biopha.2017.11.064
11. Liu ZQ, Jiang ZH, Liu L, Hu M. Mechanisms Responsible for Poor Oral Bioavailability of Paeoniflorin: Role of Intestinal Disposition and Interactions With Sinomenine. *Pharm Res.* (2006) 23(12):2768–80. doi: 10.1007/s11095-006-9100-8
12. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* (2001) 25(4):402–8. doi: 10.1006/meth.2001.1262
13. Wu X, Zhang H, Sui Z, Wang Y, Yu Z. The Biological Role of the CXCL12/CXCR4 Axis in Esophageal Squamous Cell Carcinoma. *Cancer Biol Med* (2021) 18(2):401–10. doi: 10.20892/j.issn.2095-3941.2020.0140
14. Rizeq B, Malki MI. Cancers The Role of CCL21/CCR7 Chemokine Axis in Breast Cancer Progression. *Cancers* (2020) 12(4):1036. doi: 10.3390/cancers12041036
15. Wei X, Zhang Y, Li C, Ai K, Li K, Li H, et al. The Evolutionarily Conserved MAPK/Erk Signaling Promotes Ancestral T-Cell Immunity in Fish *via* C-Myc-Mediated Glycolysis. *J Biol Chem* (2020) 295(10):3000–16. doi: 10.1074/jbc.RA119.012231
16. Rodríguez-Nieves JA, Patalano SC, Almanza D, Gharaee-Kermani M, Macoska JA. CXCL12/CXCR4 Axis Activation Mediates Prostate Myofibroblast Phenoconversion Through Non-Canonical EGFR/MEK/ERK Signaling. *PloS One* (2016) 11(7):e0159490. doi: 10.1371/journal.pone.0159490
17. Zakharia K, Luther CA, Alsabbak H, Roberts LR. Hepatocellular Carcinoma: Epidemiology, Pathogenesis and Surveillance - Implications for Sub-Saharan Africa. *S Afr Med J* (2018) 108(8b):35–40. doi: 10.7196/SAMJ.2018.v108i8b.13499
18. Soffer A, Eisdorfer SA, Ifrach M, Ilic S, Afek A, Schussheim H, et al. Inferring Primase-DNA Specific Recognition Using a Data Driven Approach. *Nucleic Acids Res* (2021) 49(20):11447–58. doi: 10.1093/nar/gkab956
19. Wang W, Xu J, Chong J, Wang D. Structural Basis of DNA Lesion Recognition for Eukaryotic Transcription-Coupled Nucleotide Excision Repair. *DNA Repair (Amst)* (2018) 71:43–55. doi: 10.1016/j.dnarep.2018.08.006
20. Szymanski JJ, Sundby RT, Jones PA, Srihari D, Earland N, Harris PK, et al. Cell-Free DNA Ultra-Low-Pass Whole Genome Sequencing to Distinguish Malignant Peripheral Nerve Sheath Tumor (MPNST) From its Benign Precursor Lesion: A Cross-Sectional Study. *PloS Med* (2021) 18(8): e1003734. doi: 10.1371/journal.pmed.1003734
21. Zhang L, Wang S, Yang Z, Hoshika S, Xie S, Li J, et al. An Aptamer-Nanotrain Assembled From Six-Letter DNA Delivers Doxorubicin Selectively to Liver Cancer Cells. *Angew Chem Int Ed Engl* (2020) 59(2):663–8. doi: 10.1002/anie.201909691
22. Kong Y, Zhang X, Zhao Y, Xue Y, Zhang Y. Uptake of DNA by Cancer Cells Without a Transfection Reagent. *Biol Res* (2017) 50(1):2. doi: 10.1186/s40659-017-0107-x
23. Wang C, Wang M, Cheng P, Huang H, Dong W, Zhang W, et al. Hepatitis B Virus X Protein Promotes the Stem-Like Properties of OV6(+) Cancer Cells in Hepatocellular Carcinoma. *Cell Death Dis* (2017) 8(1):e2560. doi: 10.1038/cddis.2016.493
24. Komatsu T, Katsuyama S, Takano F, Okamura T, Sakurada C, Tsuzuki M, et al. Possible Involvement of the μ Opioid Receptor in the

Antinociception Induced by Sinomenine on Formalin-Induced Nociceptive Behavior in Mice. *Neurosci Lett* (2019) 699:103–8. doi: 10.1016/j.neulet.2019.01.035
25. Jiang Y, Jiao Y, Liu Y, Zhang M, Wang Z, Li Y, et al. Sinomenine Hydrochloride Inhibits the Metastasis of Human Glioblastoma Cells by Suppressing the Expression of Matrix Metalloproteinase-2/-9 and Reversing the Endogenous and Exogenous Epithelial-Mesenchymal Transition. *Int J Mol Sci* (2018) 19(3):844. doi: 10.3390/ijms19030844
26. Brand S, Dambacher J, Beigel F, Olszak T, Diebold J, Otte J, et al. CXCR4 and CXCL12 are Inversely Expressed in Colorectal Cancer Cells and Modulate Cancer Cell Migration, Invasion and MMP-9 Activation. *Exp Cell Res* (2005) 310(1): 117–30. doi: 10.1016/j.yexcr.2005.07.006
27. Cao Y, Hunter ZR, Liu X, Xu L, Yang G, Chen J, et al. The WHIM-Like CXCR4(S338X) Somatic Mutation Activates AKT and ERK, and Promotes Resistance to Ibrutinib and Other Agents Used in the Treatment of Waldenstrom's Macroglobulinemia. *Leukemia* (2015) 29(1):169–76. doi: 10.1038/leu.2014.187
28. Jalili A, Shirvaikar N, Marquez-Curtis L, Qiu Y, Korol C, Lee H, et al. Fifth Complement Cascade Protein (C5) Cleavage Fragments Disrupt the SDF-1/CXCR4 Axis: Further Evidence That Innate Immunity Orchestrates the Mobilization of Hematopoietic Stem/Progenitor Cells. *Exp Hematol* (2010) 38(4):321–32. doi: 10.1016/j.exphem.2010.02.002

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 07 July 2022
doi: 10.3389/fonc.2022.941283

# Comprehensive Analysis of TRP Channel–Related Genes in Patients With Triple-Negative Breast Cancer for Guiding Prognostic Prediction

*Haojie Zhang*[1,2], *Xiangsheng Zhang*[2], *Xiaohong Wang*[2*], *Hongguang Sun*[2], *Changran Hou*[2], *Yue Yu*[2], *Song Wang*[2], *Fangxu Yin*[2] *and Zhenlin Yang*[2*]

[1] The Second Medical College, Binzhou Medical University, Yantai, China, [2] Department of Thyroid and Breast Surgery, Binzhou Medical University Hospital, Binzhou, China

OPEN ACCESS

**Edited by:**
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

**Reviewed by:**
Yi Li,
Nanjing Medical University, China
Jianyong Zhang,
Zunyi Medical University, China
Jian Wang,
Kunming Medical University, China

***Correspondence:**
Xiaohong Wang
wangxiaohong0607@126.com
Zhenlin Yang
ikb0607@163.com

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

**Received:** 11 May 2022
**Accepted:** 25 May 2022
**Published:** 07 July 2022

**Citation:**
Zhang H, Zhang X, Wang X, Sun H, Hou C, Yu Y, Wang S, Yin F and Yang Z (2022) Comprehensive Analysis of TRP Channel–Related Genes in Patients With Triple-Negative Breast Cancer for Guiding Prognostic Prediction. Front. Oncol. 12:941283. doi: 10.3389/fonc.2022.941283

**Background:** Triple-negative breast cancer (TNBC) is a special subtype of breast cancer. Transient Receptor Potential (TRP) channel superfamily has emerged as a novel and interesting target in a variety of tumors. However, the association of TRP channel–related genes with TNBC is still unclear.

**Methods:** The The Cancer Genome Atlas (TCGA)-TNBC and GSE58812 datasets were downloaded from the public database. The differentially expressed TRP channel–related genes (DETGs) were screened by limma package, and mutations of the above genes were analyzed. Subsequently, new molecular subtypes in TNBC-based DETGs were explored by consensus clustering analysis. In addition, Lasso–Cox regression analysis was used to divide it into two robust risk subtypes: high-risk group and low-risk group. The accuracy and distinguishing ability of above models were verified by a variety of methods, including Kaplan–Meier survival analysis, ROC analysis, calibration curve, and PCA analysis. Meanwhile, CIBERSORT algorithm was used to excavate status of immune-infiltrating cells in TNBC tissues. Last, we explored the therapeutic effect of drugs and underlying mechanisms of risk subgroups by pRRophetic package and GSEA algorithm, respectively.

**Results:** A total of 19 DETGs were identified in 115 TNBC and 113 normal samples from TCGA database. In addition, missense mutation and SNP were the most common variant classification. According to Lasso–Cox regression analysis, the risky formula performed best when nine genes were used: TRPM5, TRPV2, HTR2B, HRH1, P2RY2, MAP2K6, NTRK1, ADCY6, and PRKACB. Subsequently, Kaplan–Meier survival analysis, ROC analysis, calibration curve, and Principal Components Analysis (PCA) analysis showed an excellent accuracy for predicting OS using risky formula in each cohort ($P < 0.05$). Specifically, high-risk group had a shorter OS compared with low-risk group. In addition, T-cell CD4 memory activated and macrophages M1 were enriched in normal tissues, whereas Tregs were increased in tumor tissues. Note that the low-risk group was better therapeutic effect to docetaxel, doxorubicin, cisplatin, paclitaxel, and gemcitabine than the high-risk group ($P < 0.05$). Last, *in vitro* assays, Quantitative Real-time PCR (qRT-PCR)

indicated that TRPM5 was significantly highly expressed in MDA-MB-231 and MDA-MB-468 cells compared with that in MCF-10A cells ($P < 0.01$).

**Conclusion:** We identified a risky formula based on expression of TRP channel–related genes that can predict prognosis, therapeutic effect, and status of tumor microenvironment for patients with TNBC.

**Keywords: TNBC, prognosis, TRP channels, immune, bioinformatic algorithms**

## INTRODUCTION

In 2020, the number of new cases of breast cancer reached 2.26 million (1). At the same time, breast cancer is the most common malignant tumor in women. Triple-negative breast cancer (TNBC) is defined as a type of breast cancer, in which Estrogen Receptor (ER), Progestogen Receptorv (PR), and Human Epidermal GrowthFactor Receptor 2 (HER-2) are all negatively expressed. In addition, TNBC is extremely aggressive and has a high rate of early recurrence when compared with other breast cancer subtypes (2). Despite the fact that chemotherapy cure many individuals with early-stage TNBC, the majority of these patients eventually develop metastases (3). As a result, novel biomarkers or risk stratification system are urgently needed to guide the prognosis of patients with TNBC.

A recent review has highlighted that Transient Receptor Potential (TRP) channel superfamily has emerged as a novel and interesting target for therapeutic intervention in the context of breast cancer (4). The TRP channel superfamily is defined by a six-transmembrane-segment structure that acts as a multi-mode sensor for a variety of stimuli (5). The superfamily is classified into seven families based on their sequence: TRPA, TRPC, TRPM, TRPML, TRPN, TRPP, and TRPV (6). In addition, a new member, TRPS, was identified by Himmel and colleagues (7). Expression and function of TRP channels are linked to cellular processes that drive cancer progression, such as cell proliferation, migration, invasion, and apoptosis, as well as drug sensitivity (8). In fact, TRP channels act as an ion channel, and its changes are significant in most tumors (9). It is important to note that not only are ion channels dysregulated in cancer but also their regulators, effectors, and other interacting genes expression are significantly altered (10).

Given above the evidence, we explored the prognostic value of TRP channel interactors (TRP channel–related genes) in TNBC. In addition, the related genes that may be involved in the progression of TNBC were comprehensively analyzed. In addition, we established a risk stratification system to analyze the survival probability of each TNBC patient, so as to avoid personalized follow-up and treatment for patients.

## MATERIALS AND METHODS

### Datasets

The expression profile data and clinical information of patients with TNBC were obtained using TCGA (115 tumor and 113 normal samples) and Gene Expression Omnibus (GEO) database (GSE58812, 107 tumor samples). It is worth noting that GSE58812 dataset only contains survival data but no detailed clinical information. Moreover, mutation data about TNBC, such as copy number variation (CNV), single-nucleotide variation (SNV), and methylation data were also downloaded from TCGA database. In collection of TRP channel–related genes, gene set (Reactome_TRP_channels) from the molecular signatures database and another gene set (inflammatory mediator regulation of TRP channels) from the KEGG database. Last, 120 TRP channel–related genes were used for bioinformatics analysis.

### Mutation and Differentially Expressed Analysis

The mutation frequency and oncoplot waterfall plot of differentially expressed TRP channel–related genes (DETGs) in patients with TNBC were generated by the maftools package. The difference in TRP channel–related genes expression in tumor and normal tissues was identified using the limma package ($P < 0.05$; $|logFC| > 1$). We then constructed a protein–protein interaction (PPI) network for DETGs using the search tool: Genemania (11).

### Development and Validation of TRP Score for Estimating Survival Risk

We used both cohorts to validate the prediction performance, whereas one of the training set (TCGA-TNBC) was used to construct the prognostic model. First, we identified significant prognostic genes by univariate Cox regression analysis ($P < 0.1$). Subsequently, we used the glmnet package to perform LASSO regression and Cox multivariate regression for screening genes participating in TRP score formula. The formula for calculating TRP risk score is as follows: (gene 1 expression × coefficient) + (gene 2 expression × coefficient) + … + (gene n expression × coefficient). In addition, all cases were divided into two groups (the low-risk group or the high-risk group) according to the median of the TRP risk scores. The OS was compared between the two subgroups *via* Kaplan–Meier analysis. PCA analysis was performed in the stats package. The timeROC package was used to perform a ROC curve analysis. Last, we applied independent prognostic factors determined by multivariate Cox regression to construct a prognostic nomogram using rms package.

### Analysis of Immune Status and Drug Sensitivity

CIBERSORT algorithm (12) to calculate the proportion of different immune cell types based on the expression level of

immune cell-related genes. The output results of 22 infiltrating immune cells were integrated to generate an matrix of immune cell fractions for analysis. In addition, we used pRRophetic package for prediction of clinical chemotherapeutic response from tumor gene expression levels (13).

## Cell Culture

The MDA-MB-231 cell lines were purchased from Nanjing KeyGen, and the MDA-MB-468 cell lines and MCF-10A cell lines were purchased from Wuhan Procell. MDA-MB-231 cells and MDA-MB-468 cells were cultured in L-15 medium supplemented with fetal bovine serum (FBS) and antibiotics. MCF-10A cells were cultured in special culture medium (Procell CM-0525). All cells were tested negative for mycoplasma and maintained in 5% $CO_2$ at 37°C.

## Quantitative Real-Time PCR Analysis

Isolation of total RNA from the three cells (MCF-10A, MDA-MB-231, and MDA-MB-468) was performed according to the instructions for the TRIzol reagent, and the purity and concentration of RNA were determined by MV3000 Micro-Spectrophotometer (260/280 ratio). Equal amounts of RNA were then reversely transcribed to make complementary DNA. Then, we selected and designed primers. The qRT-PCR of TRPM5 was carried out using the primer set (forward: 5′-TTG CTGCCCTAGTGAACCAG-3′, reverse: 5′-GCACGATGTCC TCCCAAGAG-3′). In addition, we used Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH) as an internal control gene; reactions were carried out using the SYBR Green Premix Pro Taq HS qPCR Kit (Accurate Biology, AG11701). Last, a Bio-Rad CFX96 Real-Time PCR Detection System was used for the determination of the target gene expression levels (TRPM5). Each sample was repeated three times. GAPDH was used as an internal reference gene. The relative abundance of each gene mRNA was calculated by the $2^{-\Delta\Delta Ct}$ method.

## Statistical Analyses

The statistical analyses were conducted in the R software (version 4.0.1). Specific statistical methods have been mentioned in the bioinformatics methods above. ***, **, *, and ns refers to $P < 0.001$, $< 0.01$, $< 0.05$, and not significant, respectively.

# RESULTS

## Expression Landscape of TRP Channel–Related Genes in TNBC

We obtained gene set (Reactome_TRP_channels) from the molecular signatures database and another gene set (inflammatory mediator regulation of TRP channels) from the KEGG database. Last, we collected 120 TRP channel–related genes for follow-up analysis. Using limma package to analyze RNA-seq data of TCGA-TNBC cohort, a total of 19 DETGs were identified in 115 TNBC and 113 normal samples (**Figure 1A**). The genes highlighted using red in the heatmap are members of the TRP superfamily, including TRPM5, TRPA1, TRPM2, TRPM8, TRPV3, TRPM3, TRPM1, and TRPC6. Moreover, nine genes were upregulated and 10 downregulated in tumors compared with normal samples (**Figure 1B**). Specifically, the nine upregulated genes were TRPM5, TRPA1, TRPM2, TRPM8, MCOLN2, TRPV3, KNG1, PLA2G4D, and CALML5. The 10 downregulated genes were TRPM3, TRPM1, TRPC6, HTR2A, ITPR1, MAPK10, PIK3R1, ADCY4, ADCY5, and IGF1 (**Figure 1C**). In addition, to further verify the prognostic ability of TRP channel–related genes ($P < 0.05$), univariate Cox regression analysis was used to screen. The results showed that only MAP2K6 was protective factor, whereas TRPM5, HRH1, P2RY2, PLA2G4D, MAPK11, and ADCY6 were risk factors (**Figure 1D**). Notably, in the TRP superfamily, only TRPM5 was differentially expressed in different samples and had prognostic value in TNBC. Therefore, we subsequently carried out *in vitro* assays on TRPM5 in TNBC cell lines.

## Different TRP Molecular Subtypes in Patients With TNBC

To explore the new subtypes in TNBC-based TRP–related molecular, we performed a consensus clustering analysis. In TCGA-TNBC cohort, the intragroup correlations were the highest, and the intergroup correlations were low when $k = 2$, indicating that patients could be well divided into two clusters based on 19 DETGs (**Figure 1E**). Similar, GEO-TNBC cohort ($n = 107$) were divided into three clusters when $k = 3$ (**Figure 1G**). Meanwhile, survival analysis showed significant differences in among different molecular subgroups ($P < 0.05$, **Figures 1F, H**). In addition, there were significant differences in stage and N staging among different subtypes in TCGA-TNBC cohort (**Figure 1I**). Taken together, the significance of TRP channel–related genes on the survival and tumor progression of patients with TNBC was demonstrated from another perspective.

## Mutation Analysis of TRP Channel–Related Genes

To further study the significance of TRP channel–related genes in TNBC, we downloaded mutation data from TCGA database and conducted in-depth analysis on above 19 DETGs. First, we performed a PPI network analysis of 19 DETGs and found that members of the TRP superfamily, including the genes mentioned above, were closely related to each other. Moreover, missense mutation was the most common variant classification. SNP were the most common variant type, and C>T ranked as the top SNV class. It is worth noting that the mutation frequency of PIK3R1 in the DETGs is the highest (**Figure 2A**). We also revealed the status of CNVs of 19 DETGs in patients with TNBC. The results showed that heterozygous amplification and heterozygous deletion were present in the vast majority of genes (**Figure 2B**). Last, we investigated the methylation of these genes with normal samples as controls. Interestingly, except for PLA2G4D, KNG1, CALML5, and IGF1, the other genes had more DNA methylation than the normal samples (**Figure 2C**). In addition, we used GSVA algorithm to estimate the role of each gene in cancer-related pathways, including apoptosis, cell cycle, DNA damage response, Epithelial-mesenchymal Transition

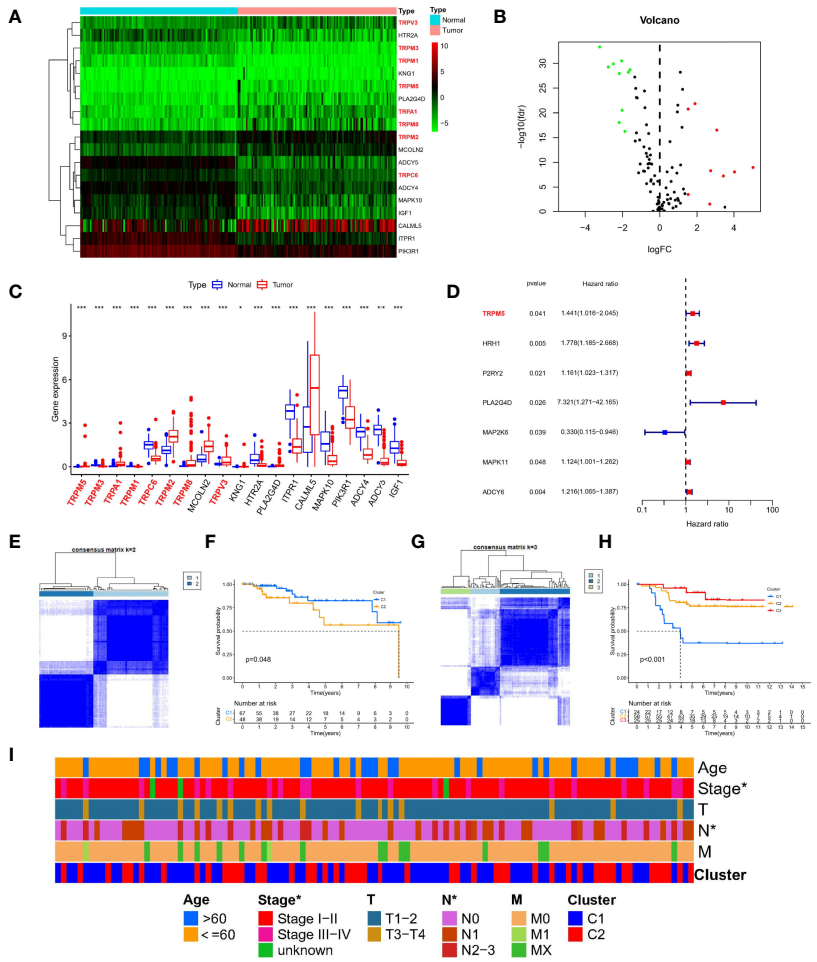



**FIGURE 1** | Expression landscape and molecular subtypes based on TRP channel–related genes. Heatmap **(A)**, volcano plot **(B)**, and boxplot **(C)** of expression of DETGs in tumor and normal samples. **(D)** A forest plot for the result of univariate Cox regression analysis in 120 TRP channel–related genes. **(E)** Consensus clustering analysis in the TCGA cohort based on different expression genes. **(F)** Survival analysis of two clusters in TCGA-TNBC. **(G)** Consensus clustering analysis in GEO cohort based on different expression genes. **(H)** Survival analysis of three clusters in TCGA-TNBC. **(I)** A composite heatmap containing clinical information in the TCGA cohort. *P < 0.05, and ***P < 0.001.

(EMT), hormone AR, hormone ER, Phosphatidylinositol-3-kinase/AKT (PI3K/AKT), RAS/Mitogen Activated Protein Kinase (MAPK), Receptor Tyrosine Kinase (RTK), Tuberous Sclerosis Complex (TSC)/mechanistic Target of Rapamycin (mTOR). The results showed that TRP-related genes may activate most of above pathways, whereas role of inhibition only included a little pathways, such as cell cycle (**Figure S1**).

## Lasso–Cox Regression Analysis Based on TRP Channel–Related Genes

To calculate TRP score for estimating survival risk in patients with TNBC, 120 TRP channel–related genes were selected by univariate Cox regression analysis (P < 0.1), and the genes were further screened by LASSO regression analysis. The prognostic model performed best when nine genes were used (**Figures 2D, E**). Last, multivariate Cox regression analysis was used to calculate the regression coefficients of nine genes (**Figure 2F**). The formula of the model is as follows: TRP risk score = (−0.498711399 × expression level of TRPM6) + (-−0.15894075 × expression level of HTR2C) + (−0.217314244 × expression level of PLA2G4A) + (−0.690911672 × expression level of ASIC4) + (−0.955369151 × expression level of P2RY2) + (−0.949344105 × expression level of MAPK14) + (−0.699670637 × expression level of PLCG2) + (0.673798617 × expression level of SRC).

## Prognostic Efficacy of Risk Score Based on TRP Channel–Related Genes

First, 107 patients from GEO cohort were utilized as the external validation cohort. According to the median score of TCGA-TNBC cohort, 115 patients in the TCGA cohort were divided into the low-risk group (n = 58) and the high-risk group (n = 57), whereas GEO cohort were divided into 80 patients as high risk

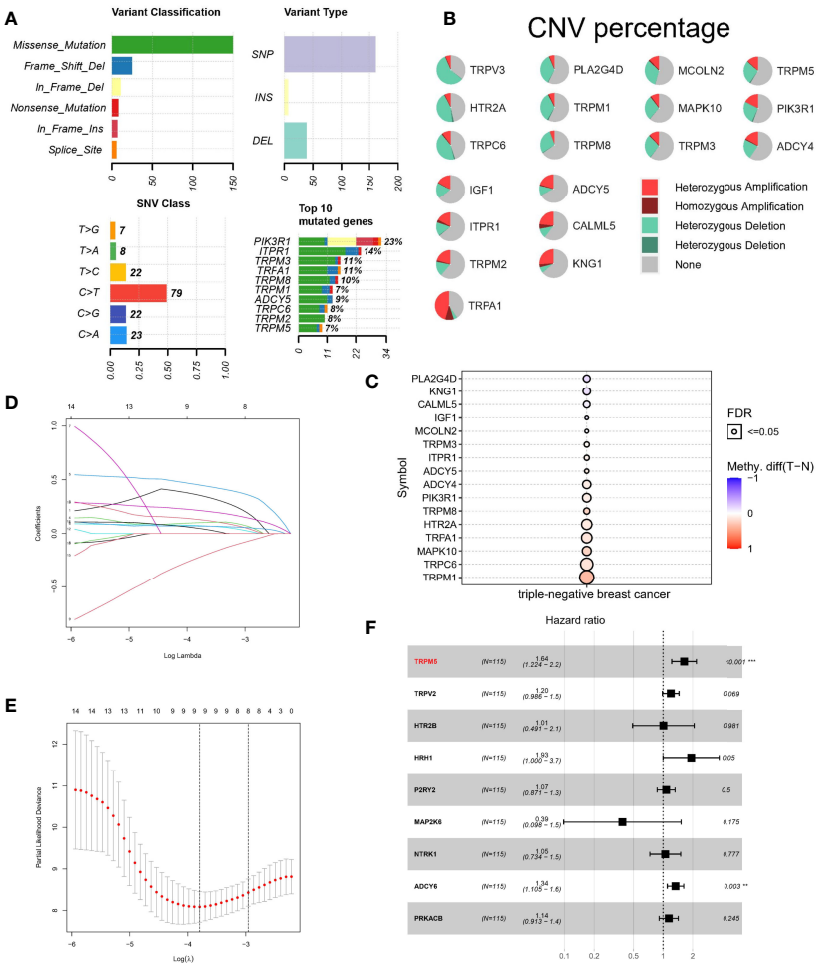



**FIGURE 2** | Mutation landscape and Lasso–Cox regression analysis of TRP channel–related genes. **(A, B)** Landscape of mutations about 19 DETRGs. **(C)** Methylation difference in different tissues. **(D, E)** LASSO regression. **(F)** A forest plot for the result of multivariate Cox regression analysis. **P < 0.01 and ***P < 0.001.

and 27 patients as low risk, respectively. The PCA analysis showed satisfactory separation between the two risk subgroups (**Figures 3A, D**). ROC analysis of the TCGA-TNBC cohort showed that risk score had good predictive effect (AUC of 3 years = 0.841 and Area Under Curve (AUC) of 5 years = 0.867), as shown in **Figure 3B**. Similarly, it also showed excellent predictive ability in GEO-TNBC cohort (AUC of 3 years = 0.605 and AUC of 5 years = 0.685), as shown in **Figure 3E**. In addition, Kaplan–Meier survival analysis also indicated a significant difference in the OS between the low- and high-risk groups (P < 0.05; **Figures 3C, F**). Specifically, the high-risk group has a shorter overall survival time compared with the low-risk group.

## Independent Prognostic Value of the Risk Score

On the basis of RNA-seq combined with clinical information, we performed Cox regression analysis again to evaluate whether the TRP risk score could be used as an independent predictor of patients with TNBC. Univariate Cox regression analysis showed that risk score and other factors, including T staing, N staging, and clinical stage, were significantly correlated with OS (P < 0.05, **Figure 3G**) Multivariate Cox regression analysis showed that only risk score was independent risk factor associated with OS (P < 0.05, **Figure 3H**). In addition, considering the importance of clinical stage in clinical practice, we combined clinical stage and TRP risk score to construct the nomogram (**Figure 3I**), and the calibration curve showed good indicator performance (**Figure 3J**). Meanwhile, there were significant differences in TRP risk score among different clinicopathological factors, including age (P = 0.02, **Figure S2A**), clinical stage (P = 0.046, **Figure S2B**), and T staging (P = 0.03, **Figure S2C**).

## Comprehensive Immune Analysis

TNBC is the most immunogenic subtype of breast cancer, and there is strong evidence that tumor-infiltrating immune cells in

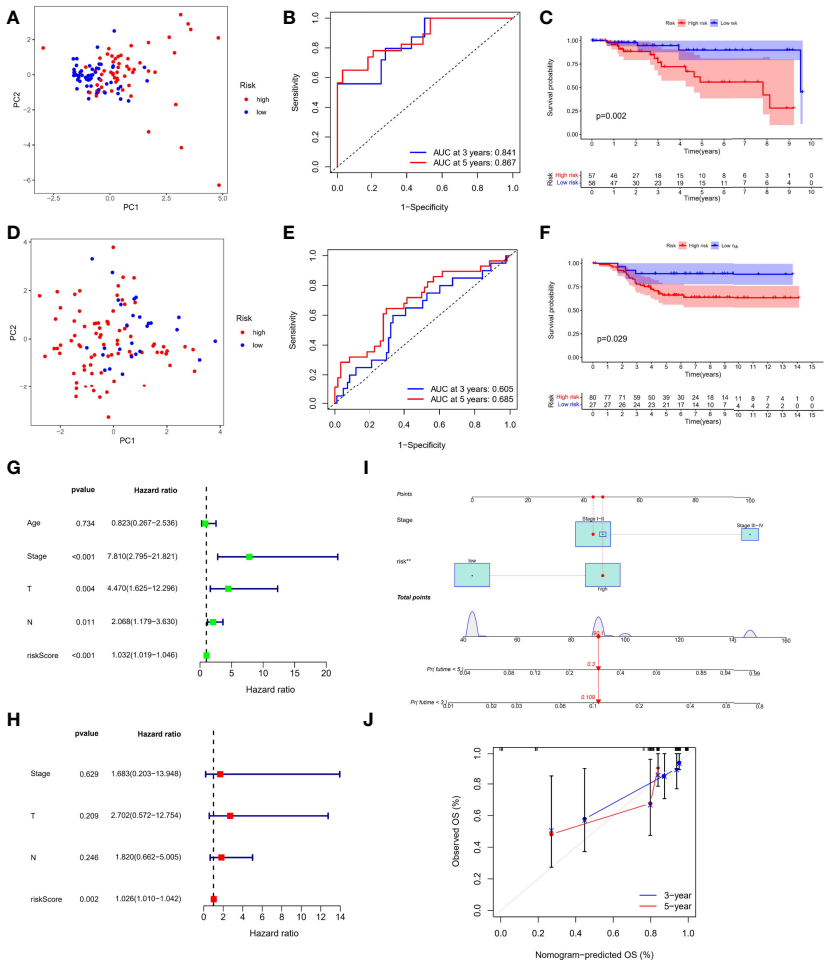



**FIGURE 3** | Prognostic efficacy of risk score based on TRP formula. PCA plot based on the risk score in the TCGA **(A)** and GEO cohort **(D)** Receiver Operating Characteristic (ROC) curves demonstrated the predictive efficiency of the risk score in the TCGA **(B)** and GEO cohort **(E)**. Kaplan–Meier curves in the high- and low-risk groups in the TCGA **(C)** and GEO cohort **(F)**. Univariate **(G)** and multivariate analysis **(H)** for the TCGA cohort. **(I)** Nomogram based on the TRP risk score and clinicopathological parameters. **(J)** Calibration curves of nomogram.

TNBC have prognostic value and are associated with improved clinical outcomes (14). Therefore, we further comprehensively evaluated the guiding role of new subtypes based on TRP risk score in immunotherapy and immune-related mechanism. We explored the distribution of various immune cells, and the results showed macrophages accounted for a larger proportion of patients with TNBC in each cohort (**Figure 4A**). In the difference analysis using Wilcoxon test, boxplots showed that the three common types of immune cells in different tissues were significantly different ($P < 0.05$, **Figure 4B**). Specifically, T-cell CD4 memory activated and macrophages M1 were enriched in normal tissues, whereas Tregs were increased in tumor tissues. Meanwhile, Pearson analysis was used to explore the association of each gene participating in risk score formula with 22 types of immune cells. ADCY6 was negatively correlated with T-cell CD4 memory activated and macrophages M1 ($P < 0.05$). HRH1 was strongest positively correlated with macrophages M2 but strongest negatively correlated with T-cell follicular helper ($P < 0.05$). HTR2B was positively correlated with macrophages M2 but strongest negatively correlated with T-cell follicular helper ($P < 0.05$); other detailed correlation information was shown in **Figure 4C**. In addition, we explored the relationship between risk score and mRNA expression of HLA, immune checkpoints. In HLA analysis, only HLA-DQB1, HLA-E, HLA-DPB1, and HLA-DBM have difference in the different risk group, and high-risk group had a status of HLA overexpression (**Figure 4D**). Meanwhile, the boxplot showed only the immune checkpoints with differences (TNFRSF14, NRP1, LAIR1, CD40LG, CD28, CD200R1, CD160, TNFSF14, TMFRSF25, and TNFRSF4), and they were all upregulated in the high-risk group compared with the low-risk group, as shown in **Figure 4E**. Finally, HRH1 regulated most Transcription Factor (TF) and dominated the transcriptional regulatory network (**Figure S3**).

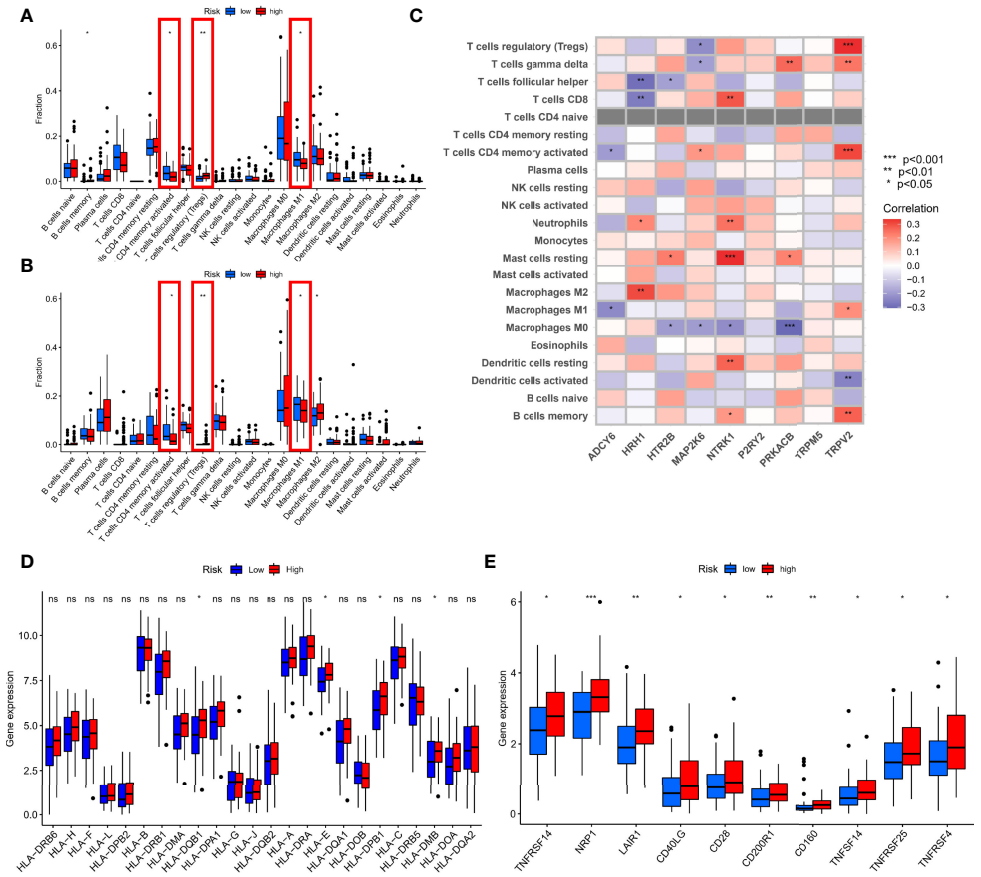



**FIGURE 4** | Comprehensive immune analysis. **(A)** Difference analysis of 22 types of immune cells in the TCGA cohort. **(B)** Difference analysis of 22 types of immune cells in GEO cohort. **(C)** Spearman analysis of 22 types of immune cells and each gene. **(D)** Difference analysis of mRNA expression of HLA. **(E)** Difference analysis of mRNA expression of checkpoints. *P < 0.05, **P < 0.01, ***P < 0.001 and ns, no statistical significance.

## Gene Set Enrichment Analysis in Different Risk Groups

To further explore the underlying mechanisms of different risks caused by disrupted TRP-related genes expression, we performed GSEA analysis to identify 672 Gene Ontology (GO) terms and 30 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways associated with different risk groups. As shown in **Figures 5A, B**, genes in the high-risk group were enriched in KEGG pathways, such as the ECM receptor interaction and focal adhesion, and GO terms, such as epithelial cell differentiation and metal ion homeostasis. In the meantime, as shown in **Figures 5C, D**, genes in the low-risk group were enriched in oxidative phosphorylation, Parkinson's disease, non-coding RNA (ncRNA) processing, ribonucleoprotein complex biogenesis, and ribosome biogenesis.

## Prediction of Therapeutic Effect

In view of the important role of anti–PD-L1 agents and chemotherapy drugs in TNBC in the current clinical guidelines (15), we evaluated the therapeutic effect of drugs by pRRophetic algorithm. Although there was no difference in PD-L1 expression between different risk groups (**Figure 6A**), correlation analysis suggested a weak positive correlation: PDL1 expression increased with increased risk score (r = 0.19, P = 0.045, **Figure 6B**). Moreover, the results of IC50 were also interesting: Low-risk group was better therapeutic effect to docetaxel (**Figure 6C**), doxorubicin (**Figure 6D**), cisplatin (**Figure 6E**), paclitaxel (**Figure 6F**), and gemcitabine (**Figure 6G**) than high-risk group, suggesting the guiding role of new subtypes based on TRP risk score in chemotherapy.

## *In Vitro* Assays for a Hub Member of TRP Superfamily

Notably, in the TRP superfamily, only TRPM5 was differentially expressed in different samples and had prognostic value in TNBC (**Figure 1**), and it also participates in the construction of risk models (**Figure 2C**). Therefore, we made a bold hypothesis: TRPM5 has an underlying meaning in the occurrence and development of TNBC. Therefore, we conducted a preliminary experimental study on the expression of TRPM5. To further confirm the effects of TRPM5 on these cells, qRT-PCR and Western blot were performed. The results showed that the TRPM5 mRNA level of MDA-MB-231 and MDA-

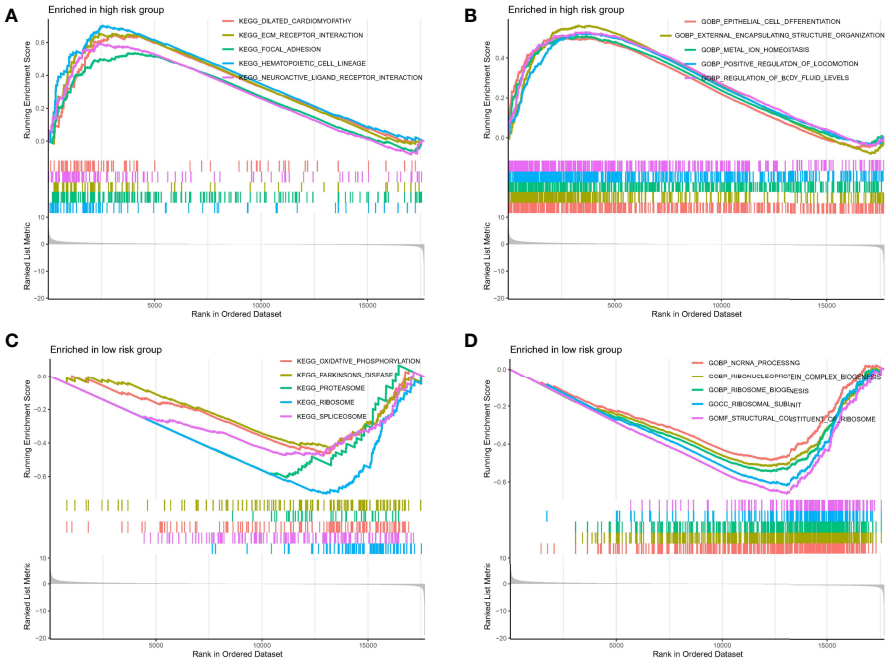



**FIGURE 5** | GSEA analysis. **(A, B)** GSEA analysis in high-risk group. **(C, D)** GSEA analysis in high-risk group.

MB-468 cells was higher than that of McF-10a cells ($P < 0.01$), (**Figures 6H-J**).

## DISCUSSION

Although there have been advancements in treating advanced TNBC, such as immunotherapy, there are still many difficulties for researchers and clinicians to conquer (16). After all, this is an advanced stage of the disease and improving the OS of such patients remains difficult. Therefore, the current goal is to research and develop biomarkers as soon as possible to predict the prognosis of patients with TNBC correctly (17). Ca2+ signaling is involved in many processes that affect the biological progression of cancer (18). Ion channels, especially TRP channel superfamily, are interesting situations for protein or gene expression because they regulate intracellular Ca2+ levels (19). Widespread dysregulation of TRP channel–related genes has been found in several types of cancers (20, 21), including breast cancer (22). Therefore, several TRP channel–related genes have been proposed as biomarkers of tumor progression and response to therapy in a variety of tumors. However, the association of TRP channel–related genes with TNBC is still unclear.

Although researchers have created many various approaches to define and quantify the immunological status of TNBC in contemporary studies (23, 24), attention to the prognostic role of widely expressed mechanosensitive calcium channels remains modest. In the study, we constructed a risky formula based on TRP-related genes. Patients were divided into groups according to high- and low-risk group, which had a value of indicting overall survival. In addition, most importantly, we also discussed the evaluation value in immune microenvironment and clinical application value of TRP risk score. In addition, we conducted *in vitro* experiments in TNBC cell lines on a hub member of TRP superfamily: TRPM5. Notably, TRPM5 also falls into members of the TRPM ("Melastatin") family (25). TRPM5 mediates depolarization of the plasma membrane (26). Selective expression of TRPM5 was originally found in taste buds, suggesting that it may play a role in taste transmission (27). Meanwhile, TRPM5 is an intrinsic signal component of mammalian chemosensory organs (28). Research on TRPM5 in tumors is extremely scarce; although we conducted a simple experimental validation for TRPM5 in TNBC cell lines, further research is still needed in the future.

There are still some limitations of our study that are worth noting. First, the bioinformatics results, for starters, have been validated using TCGA and GEO samples. We were unable to conduct a second external validation, because we lacked the sufficient funding to sequence all patients with TNBC in our hospital. Second, we only used CIBERSORT algorithm to corroborate our findings for the association between status of immune microenvironment and risk score, and we will need to conduct more experiments in the future to confirm our conclusion except for only aiming at TRPM5. In conclusion, this study developed and validated a TRP risky formula that can guide clinical decision in TNBC.

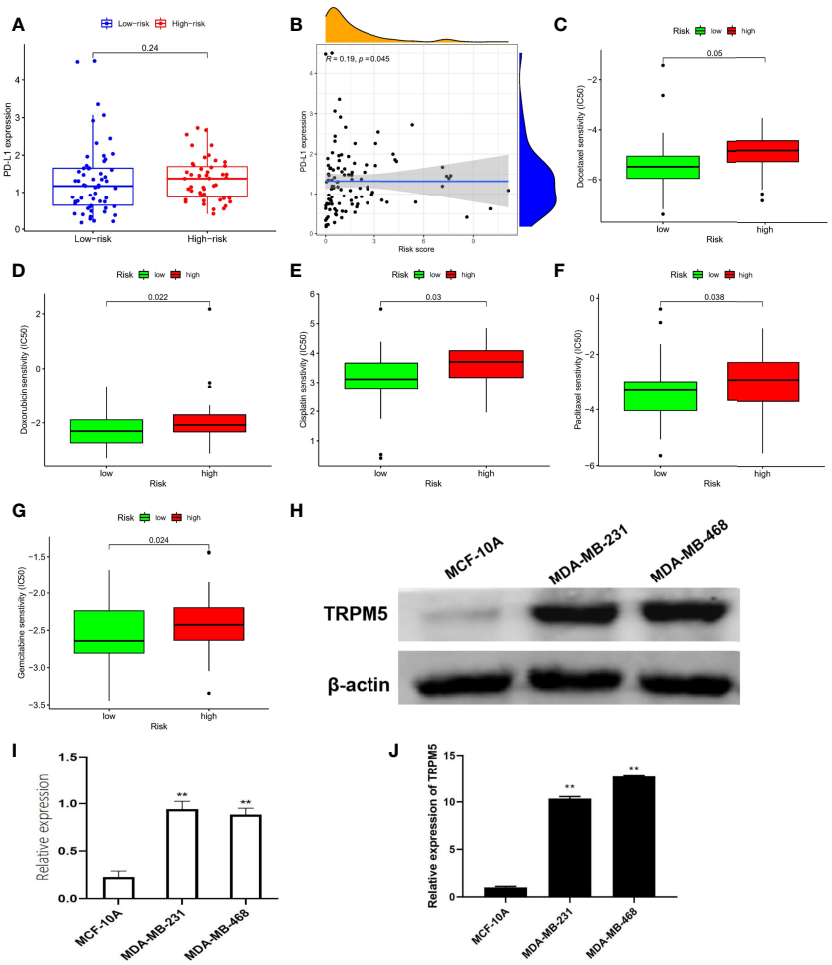



**FIGURE 6** | Prediction of therapeutic effect and expression validation. **(A)** Difference in PD-L1 expression. **(B)** Correlation analysis in PD-L1 with risk score. Difference in IC50 analysis about docetaxel **(C)**, doxorubicin **(D)**, cisplatin **(E)**, paclitaxel **(F)**, and gemcitabine **(G)**. **(H, I)** The protein levels of TRPM5 in MCF-10A, MDA-MB-231, and MDA-MB-468 were detected by Western blot. **(J)** The mRNA levels of TRPM5 in MCF-10A, MDA-MB-231, and MDA-MB-468 were detected by qRT-PCR.**P < 0.01.

## CONCLUSION

We identified a risky formula based on expression of TRP channel–related genes that can predict prognosis, therapeutic effect, and status of tumor microenvironment for patients with TNBC.

## 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

HZ and XZ designed and conceptualized the study. XW and ZY supervised the study. HS, CH, and YY contributed toward data collection and analysis. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by Projects of Binzhou technology development program [No. 2015ZC0301], Scientific Research Staring Foundation of Binzhou Medical University [No. BY2014KYQD36, No. BY2014KJ36 and No. BY2017KJ01], Science and Technology Program of Universities in Shandong Province [No. J15LL51].

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.941283/full#supplementary-material

**Supplementary Figure 1 |** GSVA algorithm to estimate the role of 19 DETGs in cancer-related pathways.

**Supplementary Figure 2 |** A network of TFs-mRNAs.

**Supplementary Figure 3 |** Differential analysis between clinical features and risk groups. Age **(A)**, stage **(B)**, and T staging **(C)**.

## 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. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-Negative Breast Cancer: Challenges and Opportunities of a Heterogeneous Disease. *Nat Rev Clin Oncol* (2016) 13(11):674–90. doi: 10.1038/nrclinonc.2016.66
3. Andre F, Zielinski C. Optimal Strategies for the Treatment of Metastatic Triple-Negative Breast Cancer With Currently Approved Agents. *Ann Oncol* (2012) 23:vi46–51. doi: 10.1093/annonc/mds195
4. Saldías MP, Maureira D, Orellana-Serradell O, Silva I, Lavanderos B, Cruz P, et al. TRP Channels Interactome as a Novel Therapeutic Target in Breast Cancer. *Front Oncol* (2021) 11:621614. doi: 10.3389/fonc.2021.621614
5. Ramsey IS, Delling M, Clapham DE. An Introduction to TRP Channels. *Annu Rev Physiol* (2006) 68:619–47. doi: 10.1146/annurev.physiol.68.040204.100431
6. Venkatachalam K, Montell C. TRP Channels. *Annu Rev Biochem* (2007) 76:387–417. doi: 10.1146/annurev.biochem.75.103004.142819
7. Himmel NJ, Gray TR, Cox DN. Phylogenetics Identifies Two Eumetazoan TRPM Clades and an Eighth TRP Family, TRP Soromelastatin (TRPS). *Mol Biol Evol* (2020) 37(7):2034–44. doi: 10.1093/molbev/msaa065
8. Santoni G, Maggi F, Morelli MB, Santoni M, Marinelli O. Transient Receptor Potential Cation Channels in Cancer Therapy. *Med Sci* (2019) 7(12):108. doi: 10.3390/medsci7120108
9. Litan A, Langhans SA. Cancer as a Channelopathy: Ion Channels and Pumps in Tumor Development and Progression. *Front Cell Neurosci* (2015) 9:86. doi: 10.3389/fncel.2015.00086
10. Prevarskaya N, Skryma R, Shuba Y. Ion Channels in Cancer: Are Cancer Hallmarks Oncochannelopathies? *Physiol Rev* (2018) 98(2):559–621. doi: 10.1152/physrev.00044.2016
11. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA Prediction Server: Biological Network Integration for Gene Prioritization and Predicting Gene Function. *Nucleic Acids Res* (2010) 38 (suppl_2):W214–20. doi: 10.1093/nar/gkq537
12. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells With CIBERSORT. *Methods Mol Biol (Clifton NJ)* (2018) 1711:243–59. doi: 10.1007/978-1-4939-7493-1_12
13. 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
14. Kwa MJ, Adams S. Checkpoint Inhibitors in Triple-Negative Breast Cancer (TNBC): Where to Go From Here. *Cancer* (2018) 124(10):2086–103. doi: 10.1002/cncr.31272
15. Emens LA, Adams S, Cimino-Mathews A, Disis ML, Gatti-Mays ME, Ho AY, et al. Society for Immunotherapy of Cancer (SITC) Clinical Practice Guideline on Immunotherapy for the Treatment of Breast Cancer. *J Immunother Cancer* (2021) 9(8):e002597. doi: 10.1136/jitc-2021-002597
16. Makki J. Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. *Clin Med Insights: Pathol* (2015) 8:23–31. doi: 10.4137/CPath.S31563
17. Wahba HA, El-Hadaad HA. Current Approaches in Treatment of Triple-Negative Breast Cancer. *Cancer Biol Med* (2015) 12(2):106–16. doi: 10.7497/j.issn.2095-3941.2015.0030
18. Berridge MJ, Bootman MD, Roderick HL. Calcium Signalling: Dynamics, Homeostasis and Remodelling. *Nat Rev Mol Cell Biol* (2003) 4(7):517–29. doi: 10.1038/nrm1155
19. Vrenken KS, Jalink K, van Leeuwen FN, Middelbeek J. Beyond Ion-Conduction: Channel-Dependent and-Independent Roles of TRP Channels During Development and Tissue Homeostasis. *Biochim Biophys Acta (BBA)-Molecular Cell Res* (2016) 1863(6):1436–46. doi: 10.1016/j.bbamcr.2015.11.008
20. Bödding M. TRP Proteins and Cancer. *Cell Signalling* (2007) 19(3):617–24. doi: 10.1016/j.cellsig.2006.08.012
21. Yang D, Kim J. Emerging Role of Transient Receptor Potential (TRP) Channels in Cancer Progression. *BMB Rep* (2020) 53(3):125–32. doi: 10.5483/BMBRep.2020.53.3.016
22. Kärki T, Rajakylä EK, Acheva A, Tojkander S. TRPV6 Calcium Channel Directs Homeostasis of the Mammary Epithelial Sheets and Controls Epithelial Mesenchymal Transition. *Sci Rep* (2020) 10(1):14683. doi: 10.1038/s41598-020-71645-z
23. Yu F, Quan F, Xu J, Zhang Y, Xie Y, Zhang J, et al. Breast Cancer Prognosis Signature: Linking Risk Stratification to Disease Subtypes. *Briefings Bioinf* (2019) 20(6):2130–40. doi: 10.1093/bib/bby073
24. 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
25. Hantute-Ghesquier A, Haustrate A, Prevarskaya N. TRPM Family Channels in Cancer. *Pharmaceuticals* (2018) 11(2):58. doi: 10.3390/ph11020058
26. Hofmann T, Chubanov V, Gudermann T, Montell C. TRPM5 is a Voltage-Modulated and Ca2+-Activated Monovalent Selective Cation Channel. *Curr Biol* (2003) 13(13):1153–8. doi: 10.1016/S0960-9822(03)00431-7
27. Pérez CA, Huang L, Rong M, Kozak JA, Preuss AK, Zhang H, et al. A Transient Receptor Potential Channel Expressed in Taste Receptor Cells. *Nat Neurosci* (2002) 5(11):1169–76. doi: 10.1038/nn952
28. Kaske S, Krasteva G, König P, Kummer W, Hofmann T, Gudermann T, et al. TRPM5, a Taste-Signaling Transient Receptor Potential Ion-Channel, is a Ubiquitous Signaling Component in Chemosensory Cells. *BMC Neurosci* (2007) 8:49. doi: 10.1186/1471-2202-8-49


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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

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

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 14 July 2022
doi: 10.3389/fonc.2022.928605

# Genomic Profiling Reveals Novel Predictive Biomarkers for Chemo-Radiotherapy Efficacy and Thoracic Toxicity in Non-Small-Cell Lung Cancer

*Kewen He[1†], Shaotong Zhang[2†], Jiaohui Pang[3], Jiani C. Yin[3], Dianbin Mu[4], Jun Wang[5], Hong Ge[6], Jie Ma[7], Zhe Yang[8], Xiaoli Zheng[6], Lihua Dong[9], Junli Zhang[3], Pengyu Chang[9], Li Li[1], Shanshan Tang[1], Hua Bao[3], Xue Wu[3], Xiaonan Wang[3], Yang Shao[3,10*], Jinming Yu[1*] and Shuanghu Yuan[1*]*

[1] Department of Radiation Oncology, Shandong University Cancer Center, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, [2] Department of Ultrasound, Jinan Central Hospital Affiliated to Shandong First Medical University, Jinan, China, [3] Geneseeq Research Institute, Nanjing Geneseeq Technology Inc., Nanjing, China, [4] Department of Pathology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, China, [5] Department of Radiation Oncology, Fourth Hospital of Hebei Medical University, Shijiazhuang, China, [6] Department of Radiation Oncology, The Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, China, [7] Department of Pathology, The Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, China, [8] Department of Radiation Oncology, Shandong Provincial Hospital, Jinan, China, [9] Department of Radiation Oncology & Therapy, Jilin Provincial Key Laboratory of Radiation Oncology & Therapy, The First Hospital of Jilin University, Jilin, China, [10] School of Public Health, Nanjing Medical University, Nanjing, China

Chemo-radiotherapy (CRT) remains the main treatment modality for non-small-cell lung cancer (NSCLC). However, its clinical efficacy is largely limited by individual variations in radio-sensitivity and radiotherapy-associated toxicity. There is an urgent need to identify genetic determinants that can explain patients' likelihood to develop recurrence and radiotherapy-associated toxicity following CRT. In this study, we performed comprehensive genomic profiling, using a 474-cancer- and radiotherapy-related gene panel, on pretreatment biopsy samples from patients with unresectable stage III NSCLCs who underwent definitive CRT. Patients' baseline clinical characteristics and genomic features, including tumor genetic, genomic and molecular pathway alterations, as well as single nucleotide polymorphisms (SNPs), were correlated with progression-free survival (PFS), overall survival (OS), and radiotherapy-associated pneumonitis and/or esophagitis development after CRT. A total of 122 patients were enrolled between 2014 and 2019, with 84 (69%) squamous cell carcinomas and 38 (31%) adenocarcinomas. Genetic analysis confirmed the association between the KEAP1-NRF2 pathway gene alterations and unfavorable survival outcome, and revealed alterations in *FGFR* family genes, *MET*, *PTEN*, and *NOTCH2* as potential novel and independent risk factors of poor post-CRT survival. Combined analysis of such alterations led to improved stratification of the risk populations. In addition, patients with *EGFR* activating mutations or any oncogenic driver

OPEN ACCESS

**Edited by:**
*Ye Wang,*
*The Second Affiliated Hospital of Medical College of Qingdao University, China*

**Reviewed by:**
*Rafael Rosell,*
*Catalan Institute of Oncology, Spain*
*Le-Tian Huang,*
*Shengjing Hospital of China Medical University, China*

***Correspondence:**
*Shuanghu Yuan*
*yuanshuanghu@sina.com*
*Jinming Yu*
*sdyujinming@163.com*
*Yang Shao*
*yang.shao@geneseeq.com*

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

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

**Received:** *26 April 2022*
**Accepted:** *23 May 2022*
**Published:** *14 July 2022*

**Citation:**
*He K, Zhang S, Pang J, Yin JC, Mu D, Wang J, Ge H, Ma J, Yang Z, Zheng X, Dong L, Zhang J, Chang P, Li L, Tang S, Bao H, Wu X, Wang X, Shao Y, Yu J and Yuan S (2022) Genomic Profiling Reveals Novel Predictive Biomarkers for Chemo-Radiotherapy Efficacy and Thoracic Toxicity in Non-Small-Cell Lung Cancer. Front. Oncol. 12:928605. doi: 10.3389/fonc.2022.928605*

mutations exhibited improved OS. On the other hand, we also identified genetic markers in relation to radiotherapy-associated thoracic toxicity. SNPs in the DNA repair-associated *XRCC5* (rs3835) and *XRCC1* (rs25487) were associated with an increased risk of high-grade esophagitis and pneumonitis respectively. *MTHFR* (rs1801133) and *NQO1* (rs1800566) were additional risk alleles related to higher susceptibility to pneumonitis and esophagitis overall. Moreover, through their roles in genome integrity and replicative fidelity, somatic alterations in *ZNF217* and *POLD1* might also serve as risk predictors of high-grade pneumonitis and esophagitis. Taken together, leveraging targeted next-generating sequencing, we identified a set of novel clinically applicable biomarkers that might enable prediction of survival outcomes and risk of radiotherapy-associated thoracic toxicities. Our findings highlight the value of pre-treatment genetic testing to better inform CRT outcomes and clinical actions in stage III unresectable NSCLCs.

**Keywords: non-small cell lung cancer, radiotherapy, radiation sensitivity, biomarker, genetic variation, prognosis, chemo-radiotherapy, radiotherapy-associated toxicity**

## INTRODUCTION

Lung cancer is the leading cause of cancer-related deaths worldwide and in China, among which approximately 85% of patients have non-small cell lung cancers (NSCLCs) (1, 2). NSCLC is sub-categorized based on histological features into mainly adenocarcinoma (ADC) and squamous cell carcinoma (SCC) (1). Over the past two decades, many therapeutic advances have been made given our deepened understanding of lung cancer etiology. The identification of actionable molecular targets has revolutionized the management of NSCLC, with targeted therapies demonstrating remarkable clinical benefits in patients carrying specific driver mutations (1). Nevertheless, the majority of lung cancer patients still require radiotherapy for cure or palliative care. In particular, for NSCLC patients with unresectable locally advanced tumors, especially SCC, the combination of chemotherapy and thoracic radiation, given either concurrently or sequentially, remains the standard of care (3).

Radiotherapy, together with the radio-sensitizing effect of chemotherapy, results in enhanced anti-tumor efficacy, although at the expense of significant normal tissue toxicity. Radiotherapy-induced lung injury (known as radiation pneumonitis in an early phase and pulmonary fibrosis in the late phase), as well as esophagitis are common adverse events following thoracic radiation (4, 5). There are considerable variations between patients in their likelihood to develop severe adverse events following a given dose of radiation, which consequently, limits the maximum dose that can be administered to the majority (6). Similarly, there are substantial differences in individual response to chemo-radiotherapy (CRT) and the risk of resistance development. It has been long recognized that genetic variations between individuals or tumors are major contributors to the differences in radio-sensitivity and risks of developing radiotherapy-associated toxicity, and thus, must be taken into consideration for personalized radiotherapy dose-prescription.

Our understanding of differential response to radiotherapy begins with the discovery of several genetic syndromes caused by mutations in the DNA repair pathways, which can lead to life-threatening radiotherapy toxicity (7, 8). Subsequently, a number of other molecular processes, such as scavenging of reactive oxygen species (ROS), apoptosis, proliferation and inflammatory response, have been implicated in the development of radiation-induced toxicity (9–12). While multiple approaches, including candidate gene approach and genome-wide association studies, were undertaken to identify genetic variants that might explain the differences in individual response to radiotherapy, no robust biomarkers with convincing clinical applicability have been identified (13). In addition, a lung tissue-specific, comprehensive analysis for personalized radiotherapy is still lacking. Here, taking advantage of next-generation sequencing (NGS) technology, we performed comprehensive genomic profiling on 474 cancer- and radio-sensitivity-related genes of the tumor biopsies from 122 unresectable stage III NSCLC patients prior to radiation therapy, and identified a set of promising biomarkers for predicting radiation survival and toxicity, which may prove beneficial for guiding clinical treatment decision-making.

## MATERIAL AND METHODS

### Patient Enrollment

The patients with NSCLC in the study were treated with CRT at the multiple centers between October 2014 and March 2019. Eligible patients for this study were determined based on the following criteria: histological diagnosis of unresectable stage IIIA-C NSCLC based on the tumor, node and metastasis (TNM) staging system without severe pleural or pericardial effusion, age older than 18 years, adequate lung, bone marrow, renal, hepatic, and cardiac function, and no history of systemic treatment or radiotherapy for thoracic cancers. The study was approved by the Ethical Review Board of the Oncology Center of Shandong

Provincial Hospital, and all patients provided written informed consent.

## Treatment and Assessments

All patients in this study received standard definitive CRT (dCRT). A median of five cycles of cisplatin- or paclitaxel-based chemotherapy were given concurrently or sequentially with radiotherapy. The choice of chemotherapy regimen was left to the investigator's discretion. Three-dimensional conformal radiation therapy (3D-CRT) or intensity-modulated radiation therapy (IMRT) was administered at a total dose of 50-70 Gy.

The follow-up of all patients was conducted 1 month after radiotherapy, and then every 3 months during the first year. After that, the patients were followed up every 3-6 months. Radiotherapy-associated thoracic toxicities were graded according to the toxicity criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC) (14). For toxicity criteria of pneumonitis, grade 1 includes mild symptoms of dry cough or dyspnea on exertion; grade 2 includes persistent cough requiring narcotic or antitussive agents, or dyspnea with minimal effort but not at rest; grade 3 includes severe cough unresponsive to narcotic antitussive agent or dyspnea at rest, clinical or radiological evidence of acute pneumonitis, or requirement of intermittent oxygen or steroids; and grade 4 includes severe respiratory insufficiency or continuous oxygen or assisted ventilation. For esophagitis, grade 1 includes mild dysphagia or odynophagia, requirement of topical anesthetic, non-narcotic analgesics, or soft diet; grade 2 includes moderate dysphagia or odynophagia, requirement of narcotic analgesics, puree or liquid diet; grade 3 includes severe dysphagia or odynophagia with dehydration or weight loss >15% from pretreatment baseline, requirement of nasogastric feeding tube, intravenous fluids, or hyperalimentation; grade 4 includes complete obstruction, ulceration, perforation, or fistula. Treatment responses were assessed using CT imaging at each follow-up and compared to the images at baseline or from the last follow-up and were evaluated according to the Response Evaluation Criteria in Solid Tumors (RECIST), version 1.1. Progression-free survival (PFS) was defined as the time from the beginning of treatment to disease progression. Patients who had not progressed were censored at the date of their last scan. Overall survival (OS) was defined from the beginning of treatment to the time of death from any cause or the last follow-up.

## DNA Extraction and Library Preparation

All tumor samples were formalin-fixed paraffin-embedded (FFPE, 10 μm) and were obtained from original biopsies prior to any treatment. At least 10% tumor content of all samples, as determined by pathologists, was required. NGS was performed in a CLIA-certified and CAP-accredited laboratory (Nanjing Geneseeq Technology Inc., Nanjing, China). Genomic DNA was extracted from de-paraffinized FFPE sections using QIAamp DNA FFPE Tissue Kit (Qiagen) according to the manufacturer's instructions. Quantity and quality of DNA were assessed using Qubit 3.0 fluorometer and Nanodrop 2000 (ThermoFisher), respectively. DNA was fragmented into 350 bp using the Covaris M220 sonication system and purified with Agencourt AMPure XP beads (Beckman Coulter).

DNA (50 ng) libraries were prepared with KAPA hyper library preparation kit (KAPA Biosystems). Libraries with different indices were pooled for targeted enrichment with IDT xGen Lockdown Reagents and a customized enrichment panel (IDT) covering 474 cancer-related genes with whole-exon coverage, including those that have been implicated in radiotherapy response and/or radiotherapy-associated adverse effects (Radiotron®, Nanjing Geneseeq Technology Inc., Nanjing; **Supplementary Table 1**). Libraries were captured with Dynabeads M-270 (Life Technologies) and xGen Lockdown hybridization and wash kit (IDT). The captured library was further on-beads PCR amplified with Illumina p5 (5' AAT GAT ACG GCG ACC GA 3') and p7 (5' CAA GCA GAA GAC GGC ATA CGA GAT 3') primers in KAPA Hifi HotStart ReadyMix (KAPA Biosystems) and purified with Agencourt AMPure XP beads. Sequencing libraries were sized on the Agilent Bioanalyzer 2100 (Agilent Technologies) and their concentrations analyzed by qPCR with KAPA Library Quantification kit (KAPA Biosystems). The final libraries were sequenced on an Illumina Hiseq 4000 platform to a mean coverage depth of ~350.

## Sequencing Data Analysis

NGS read preprocessing, including quality control of FASTQ files and removing leading/trailing low quality (quality reading below 15) or N bases, was conducted with Trimmomatic (15). Qualified pair-end reads were aligned to the reference human genome hg19 with Burrows-Wheeler Aligner (v0.7.12) (16). PCR deduplication was performed using Picard and local realignment around indels and base quality score recalibration was performed using GATK3. Samples with mean dedup depth of less than 30X were also removed. Cross-sample contamination was quantified by using ContEst (Broad Institute). Single nucleotide polymorphisms (SNPs) were identified if present in >1% population frequency in the 1000g, genomAD, or ExAC databases. Somatic single nucleotide variants (SNVs) and indels were identified using VarScan2 (17) with the following parameters: i) minimum read depth=20; ii) minimum variant supporting reads=5, mapped to both strands; iii) minimum base quality=15; iv) strand bias no greater than 10%. Somatic variants were further filtered through an internally collected list of recurrent sequencing errors and if present in >1% population frequency in the 1000g, genomAD or ExAC database. Copy number variations (CNVs) were detected using CNVkit (18). CNV gain and loss were identified if depth ratio were above 1.6 or below 0.6, respectively. Final list of mutations was annotated using vcf2maf (call VEP for annotation). Panel tumor mutational burden (TMB) was counted by summing all base substitutions and indels in the coding region of targeted genes, including synonymous alterations to reduce sampling noise and excluding known driver mutations as they are over-represented in the Panel, as previously described (19).

## Statistical Analysis

For comparisons of proportion between groups, Fisher's exact tests were performed. For non-normally distributed data, such as TMB, differences between two groups were evaluated with the

non-parametric Mann-Whitney/Wilcoxon rank-sum test. For survival analyses, Kaplan-Meier curves were estimated using the log-rank test, and hazard ratios (HRs) for PFS and OS were calculated by Cox proportional hazards model. Multivariable survival analysis was performed using the Cox regression model. A two-sided P value of less than 0.05 was considered to be statistically significant unless otherwise indicated. All statistical analyses were done in R (v.3.5.2).

# RESULTS

## Patient Overview

We retrospectively performed analyses on 122 patients with unresectable stage III NSCLC, who underwent dCRT. Patients' baseline characteristics were summarized in **Table 1**. Median age of the study cohort was 62 years. Histological subtypes included 84 SCC (68.9%) and 38 ADC (31.1%). Consistent with a higher proportion of SCC patients, there were more male patients (87.7%) and former smokers (74.6%) in the study cohort. 51.6% (63/122) patients received concurrent dCRT and the remaining (48.4%) received sequential dCRT. At data cutoff, the median follow-up time was 30.1 months. Median PFS and OS of the study cohort were 11.4 and 34.6 months, respectively (**Supplementary Figure 1**). A total of 51 (41.8%) patients developed grade 2 or higher toxicity, with 39 patients developed grade 2 or higher pneumonitis and 16 cases of esophagitis.

Genomic profiling on baseline tissues using a targeted NGS panel covering 474 cancer- and radiotherapy-associated genes (**Supplementary Table 1**) revealed the mutational landscape of the cohort (**Figure 1A**). The most frequently altered genes were *TP53* (ADC, 78.9%; SCC, 94.0%), *MCL1* (ADC, 52.6%; SCC, 62.7%), *MYC* (ADC, 23.7%; SCC, 38.5%), *NOS2* (ADC, 26.3%; SCC, 28.9%) and *EGFR* (ADC, 31.6%; SCC 24.1%). Notably, we observed a high frequency of *TP53* and a low frequency of *EGFR* mutations in our ADC patients, which was consistent with an enrichment of former smokers in our study. The median TMB of the cohort is 13.4 mutations/Mb, with no significant difference observed between ADC (12.4 mutations/Mb) and SCC (13.9 mutations/Mb; *P*=0.68).

## Predictive Markers for Survival Outcome Following CRT

First, we examined potential associations between clinical characteristics and survival outcome following dCRT. No significant difference in survival outcomes was observed comparing patients with different histological subtypes (PFS, *P*=0.27; OS, *P*=0.76). Patients treated with concurrent and sequential dCRT also had similar PFS outcome (*P*=0.56, **Supplementary Figure 2A**), albeit a small trend towards increased OS in patients treated with concurrent dCRT (HR [95% CI] =0.63 [0.37-1.1], *P*=0.10, **Supplementary Figure 2B**). Patients treated with 3D-conformal RT had lower risk of progression compared with those treated with intensity-modulated RT (HR [95% CI] =0.53 [0.31-0.90], *P*=0.01, **Supplementary Figure 2C**), which did not translate into an OS difference (*P*=0.92, **Supplementary Figure 2D**). Patients with smoking histories had worse outcome compared with never smokers (PFS, HR [95% CI] =1.95 [1.15-3.32], *P*=0.01; OS, HR [95% CI] =1.58 [0.79-3.16], *P*=0.19, **Supplementary Figures 2E, F**). A higher overall dose was associated with a trend towards improved PFS and prolonged OS (PFS, HR [95% CI] =0.73 [0.45-1.18], *P*=0.2; OS, HR [95% CI] =0.56 [0.31-1.0], *P*=0.05, **Supplementary Figures 2G, H**).

**TABLE 1** | Clinical characteristics of the patients.

| Characteristics | No. of patients (%) |
|---|---|
| **Sex** | |
| Male | 107 (87.7%) |
| Female | 15 (12.3%) |
| **Age** | |
| <65 | 73 (59.8%) |
| ≥65 | 49 (40.2%) |
| **Median age (range)** | 62 (33–84) |
| **Histological subtypes** | |
| Adenocarcinoma | 38 (31.1%) |
| Squamous cell carcinoma | 84 (68.9%) |
| **Smoking history** | |
| Former smokers | 91 (74.6%) |
| Never smokers | 31 (25.4%) |
| **Treatment regimen** | |
| SCRT | 59 (48.4%) |
| CCRT | 63 (51.6%) |
| **Radiation type** | |
| 3DCRT | 30 (24.6%) |
| IMRT | 92 (75.4%) |
| Radiation dose | |
| <60 Gy | 28 (23.0%) |
| =60 Gy | 66 (54.0%) |
| >60 Gy | 28 (23.0%) |
| **RT-related pneumonitis** | |
| Grade 0-1 | 83 (68.0%) |
| Grade 2 | 18 (14.8%) |
| Grade 3 | 19 (15.6%) |
| Grade 4 | 2 (1.6%) |
| **RT-related esophagitis** | |
| Grade 0-1 | 106 (86.9%) |
| Grade 2 | 8 (6.55%) |
| Grade 3 | 8 (6.55%) |
| Grade 4 | 0 (0%) |

*SCRT, sequential chemoradiotherapy; CCRT, concurrent chemoradiotherapy*

Next, we explored the associations between individual genomic alterations (non-synonymous alterations that occur in at least 5% of the cohort) and dCRT survival outcome. We identified 19 patients with mutations in Kelch Like ECH Associated Protein 1 (*KEAP1*, **Figures 1A, B**). KEAP1 is an E3 ubiquitin ligase that functions as a sensor for oxidative stress and negatively regulates NRF2, a transcription factor upstream of cytoprotective and antioxidant genes, in the absence of stress (20). Of the 19 *KEAP1*-mutant patients, three carried nonsense mutations and the rest carried missense mutations that were all except three predicted to be deleterious or potentially damaging to protein function by SIFT or PolyPhen. Specifically, we detected three mutations in the BTB domain and four in the intervening BACK (BTB and C-terminal Kelch) domain, both of which mediate its interaction with Cullin 3 (Cul3) for protein

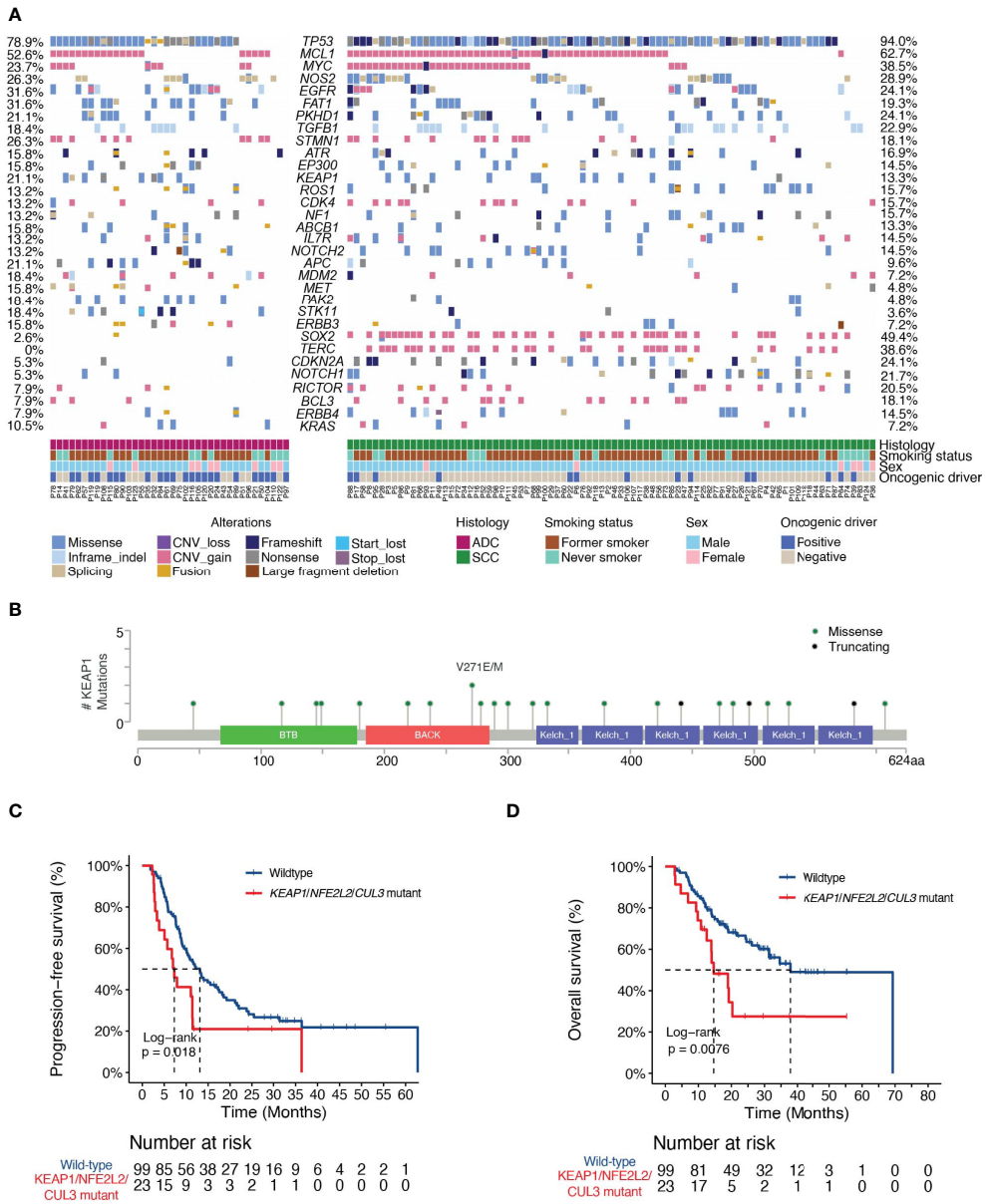



**FIGURE 1** | Landscape of genetic variations and the associations of dCRT survival outcomes with the KEAP1-NRF2 pathway. **(A)** The distribution of various genetic variations in each patient was shown. Clinical characteristics of each patient were shown at the bottom. ADC, adenocarcinoma; SCC, squamous cell carcinoma. **(B)** Lollipop plot showing the distribution of *KEAP1* mutations in the study cohort. **(C, D)** Kaplan-Meier estimates of **(C)** PFS and **(D)** OS in the full analysis set comparing patients with and without KEAP1-NRF2 pathway gene mutations. HR denotes hazard ratio; CI denotes confidence interval. Tick marks indicate censored data.

ubiquitination, as well as ten mutations located in the Kelch repeat domains that mediates interaction with NRF2 (**Figure 1B**). *KEAP1* mutations have been reported to correlate with an increased rate of local recurrence in NSCLC patients treated with radiotherapy (21). Indeed, our data independently showed that patients with *KEAP1* mutations had shorter median PFS (6.7 months vs. 12.2 months; HR [95% CI] = 2.17 [1.24-3.81], $P$ = 0.006, **Supplementary Figure 3A**) and OS (18.8 vs. 37.8 months; HR [95% CI] =2.37 [1.23-4.55], $P$=0.008, **Supplementary Figure 3B**) compared with those with the wild-type gene. Considering patients with deleterious mutations in genes in the KEAP1-NRF2 pathway (*KEAP1*, *NFE2L2* or *CUL3*) showed a consistent increase in the risk of disease progression (HR [95% CI] =1.86 [1.1-3.15], $P$=0.02, **Figure 1C**), as well as decreased OS (HR=2.27 [1.22-4.23], $P$=0.008, **Figure 1D**).

Univariate analysis revealed additional associations of survival outcome following dCRT with variations in several key genes that play important roles in lung cancer initiation and progression. The *MET* oncogene, which encodes a receptor

tyrosine kinase, has become an important target for the treatment of NSCLC. We identified 11 patients with *MET* alterations, of which one had *MET* amplification, four carried exon 14 skipping mutations, three had *MET* fusions, and one patient with both amplification and an exon 14 skipping mutation (**Figure 1A**). The presence of *MET* alterations had a negative impact on disease progression, with shorter PFS than those with the wild-type gene (HR [95% CI] =2.33 [1.2-4.52], *P*=0.01, **Figure 2A**). No significant difference in OS was found comparing patients with *MET* alterations and those with the wild-type gene (HR [95% CI] =1.29 [0.55-3.03], *P*=0.56, **Figure 2B**). *PTEN* is an important tumor suppressor gene in lung cancers. Patients with deleterious mutations in *PTEN*, including nonsense, frameshift and splicing alterations exhibited a higher progression risk than those without (PFS, HR [95% CI] =2.19 [1.12-4.27], *P*=0.02, **Figure 2C**). Similarly, no OS difference was found comparing patients with and without *PTEN* alterations (HR [95% CI] =1.57 [0.67-3.68], *P*=0.29, **Figure 2D**), suggesting that both *MET* and *PTEN* alterations might serve as predictive markers of dCRT response. Preclinical studies have suggested that the NOTCH signaling pathway might also promote radiation resistance (22). Mutations in *NOTCH2*

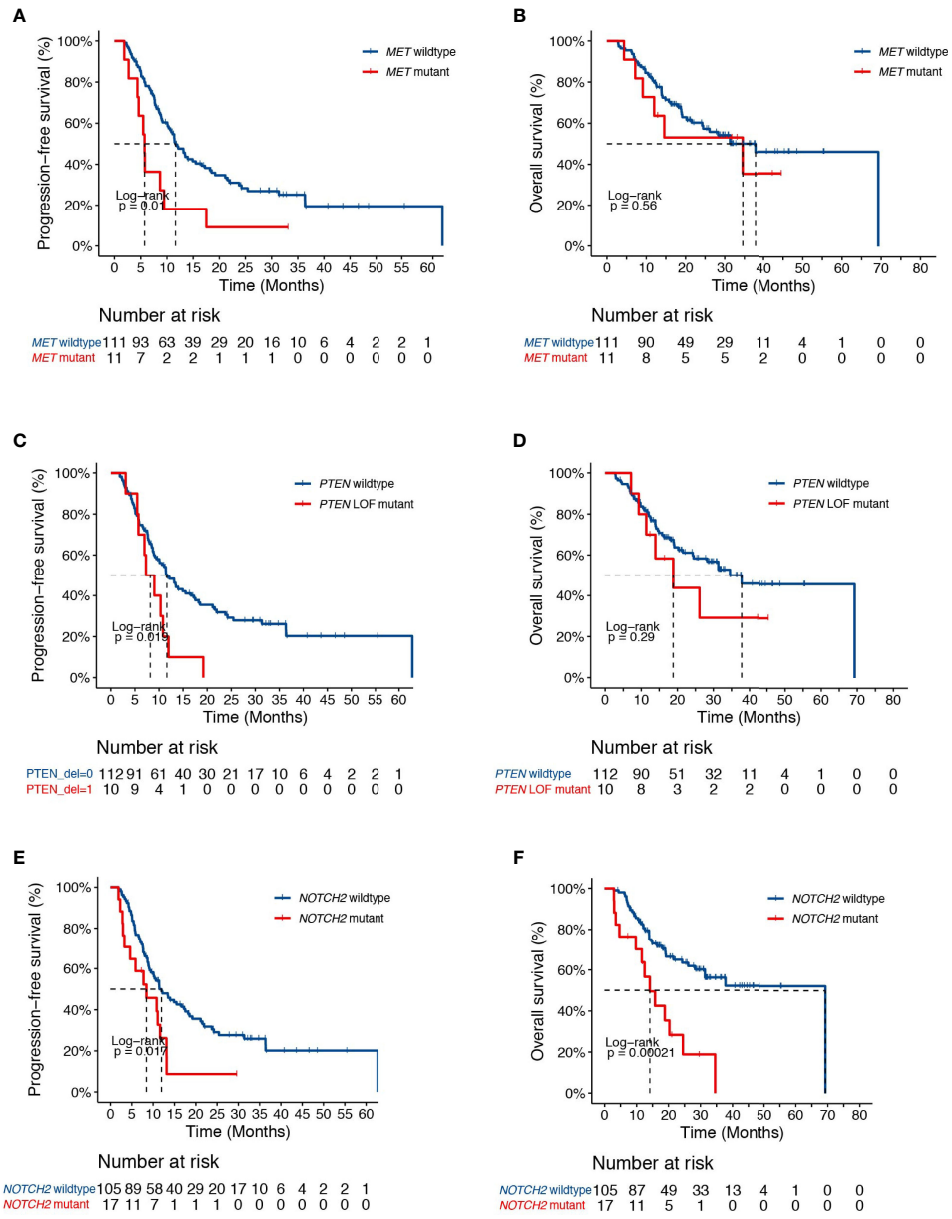



**FIGURE 2** | Clinical associates of dCRT survival outcomes in NSCLC. **(A, B)** Kaplan-Meier estimates of **(A)** PFS and **(B)** OS in the full analysis set comparing patients with and without *MET* alterations. **(C, D)** Kaplan-Meier estimates of **(C)** PFS and **(D)** OS in the full analysis set comparing patients with and without *PTEN* deleterious mutations. **(E, F)** Kaplan-Meier estimates of **(E)** PFS or **(F)** OS in the full analysis set comparing patients with and without *NOTCH2* alterations.

were found to correlate with unfavorable survival outcome (PFS, HR [95% CI] = 2.0 [1.12-3.57], *P*=0.02; OS, HR [95% CI] = 3.12 [1.65-5.89], *P*=0.0002, **Figures 2E, F**).

Genetic analysis also revealed alterations in *FGFR1* in association with higher risk of progression in our study cohort (HR [95% CI] = 2.44 [1.16-5.14], *P*=0.015, **Figure 3A**). No difference in survival was observed between *FGFR1* wildtype and mutant patients (HR [95% CI] = 1.41 [0.56-3.56], *P*=0.46, **Figure 3B**). Given that FGFR signaling is often dysregulated in NSCLC and have been implicated in radiation resistance in preclinical studies (23, 24), we sought to further test the association between FGFR family genes and patient survival. Indeed, we found that genetic alterations in the FGFR family receptors, including *FGFR1-4*, were associated with earlier progression (PFS, HR [95% CI] = 1.72 [1.06-2.79], *P*=0.03; OS, HR [95% CI] = 2.04 [1.14-3.65], *P*=0.01, **Figures 3C, D**).

Based on these findings, we also assessed if other oncogenic mutations might influence patient outcome. Overall, the presence of oncogenic driver mutations in key lung cancer targets, such as activating mutations in *EGFR*, *ERBB2*, *KRAS*, *MET*, as well as *ALK*, *RET*, and *ROS1* fusions, had no impact on patient progression (HR [95% CI] =0.87 [0.55-1.38], *P*=0.56, **Supplementary Figure 4A**). Interestingly, carriers of oncogenic mutations had significantly improved OS compared with those without (HR [95% CI] =0.52 [0.27-0.99], *P*=0.04, **Supplementary Figure 4B**). Similarly, patients with activating *EGFR* mutations alone had no significant impact on PFS but demonstrated increased OS compared with *EGFR* wild-type patients (PFS, HR [95% CI] =0.77 [0.46-1.26], *P*=0.3; OS, HR [95% CI] =0.42 [0.20-0.89], *P*=0.02, **Supplementary Figures 4C, D**), which is likely explained by the potential use of subsequent targeted therapies in these patients. We also examined the effect of TMB on patient recurrence, and no significant association was identified.

By adjusting for differences in clinical characteristics, including types of RT, smoking histories and overall dose, multivariate Cox analysis showed that alterations in *MET* and deleterious mutations in *PTEN*, as well as the FGFR pathway gene alterations remained independent predictive factors for reduced PFS following CRT (**Figure 4A**). On the other hand, the associations of OS with alterations in *NOTCH2* and those in genes in the FGFR and the KEAP1-NRF2 pathways remained independent (**Figure 4B**). Subgroup analysis considering all patients who had received the recommended doses of 60-66Gy was performed and we found that deleterious mutations in *PTEN* and *FGFR1* mutations remained independently associated with poorer PFS (**Supplementary Table 2**). In addition, mutations in *NOTCH2* remained independent predictors of unfavorable OS outcome (**Supplementary Table 2**). Building on these associations and the largely exclusive nature of *KEAP1*, *MET*, *PTEN*, *NOTCH2* and *FGFR* family gene alterations, we next

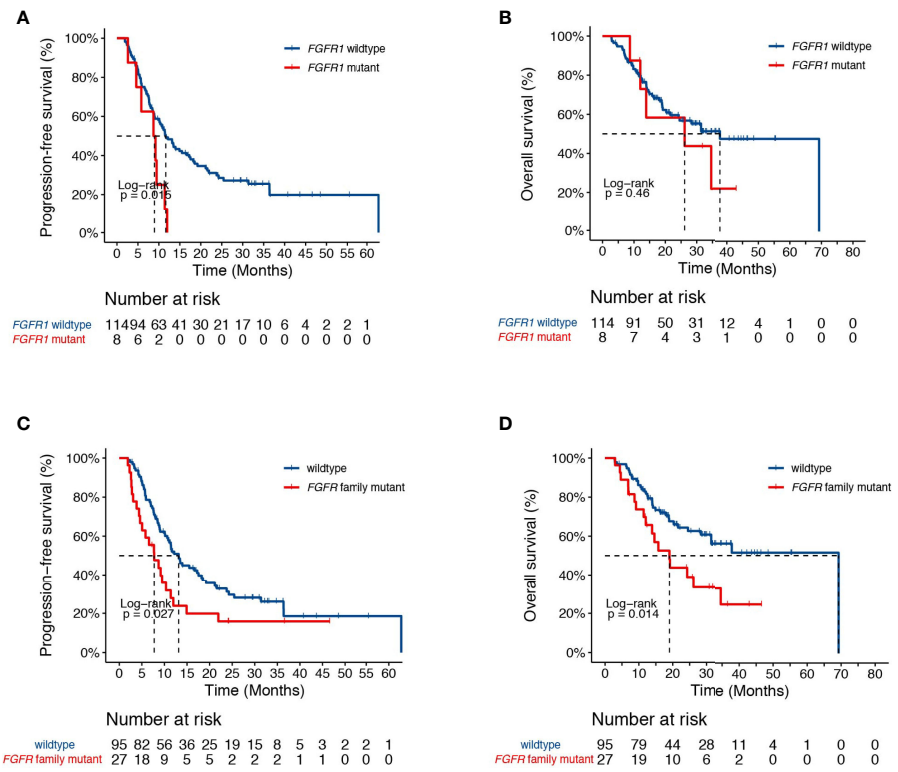



**FIGURE 3** | Associations of dCRT survival outcomes with the FGFR pathway. **(A, B)** Kaplan-Meier estimates of **(A)** PFS and **(B)** OS in the full analysis set comparing patients with and without *FGFR1* alterations. **(C, D)** Kaplan-Meier estimates of **(C)** PFS and **(D)** OS in the full analysis set comparing patients with and without alterations in the *FGFR* family genes.

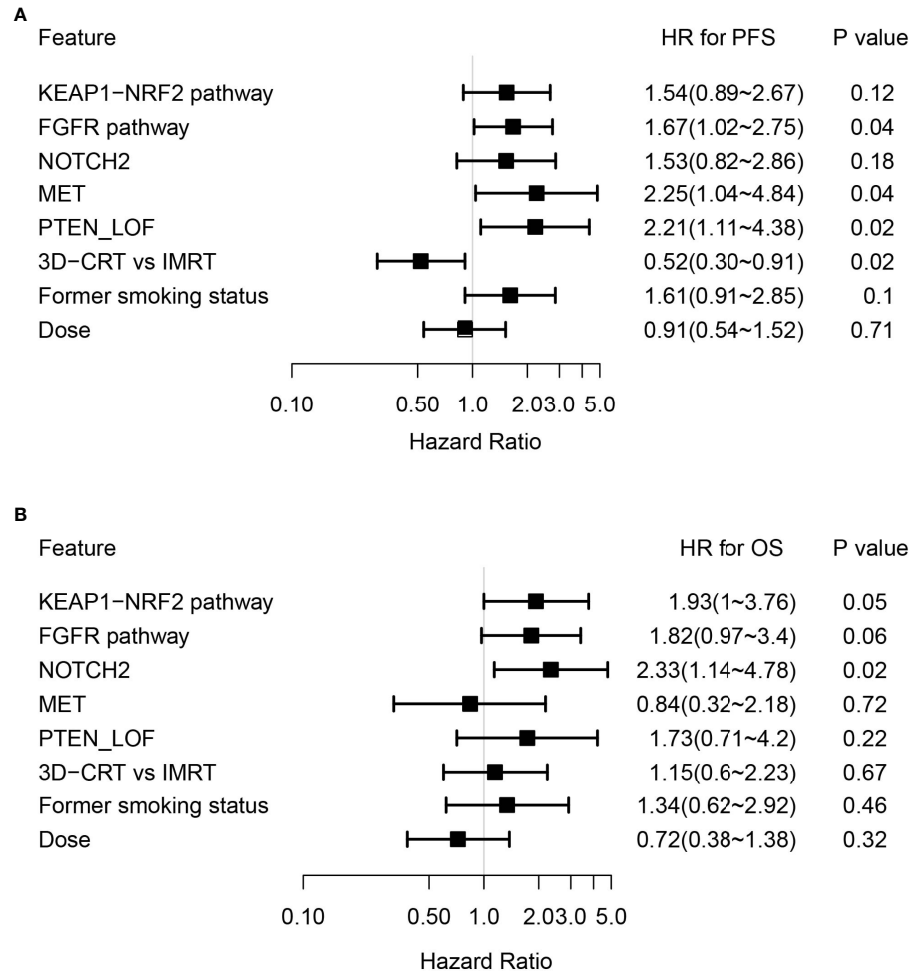



**FIGURE 4** | Multivariate Cox analysis of genetic features associated with survival outcomes. **(A, B)** Forest plots showing key genetic and clinical features in association with **(A)** PFS and **(B)** OS following dCRT treatment by multivariate analysis.

sought to improve the stratification of NSCLC patients with differential survival outcome. Combined analysis of the mutant subgroup showed markedly improved risk stratification of our patients with significant PFS (HR=2.09, 95% CI=1.36-3.2, *P*=0.0006) and OS (HR= 2.73, 95% CI=1.54-4.85, *P*=0.0003) differences (**Figures 5A, B**). In addition, multivariate analysis incorporating all relevant clinical characteristics, including histology, modes of CRT, smoking status and overall dose, showed that the mutant subgroup remained a strong independent predictor of survival outcomes (**Figures 6A, B**).

## SNPs and Somatic Mutations Predictive of Radiation Toxicity

To identify potential risk factors that could explain individual variations in their likelihood to develop radiotherapy-associated toxicity, we first examined the potential effects of various clinical characteristics and treatment regimens. None of the clinical features, including age, smoking status, dose, sequential or concurrent combinations of chemo- and radiotherapy, or delivery methods, had a significant influence on the development of radiation-induced thoracic toxicity, including grade 2 or higher pneumonitis and esophagitis. In recognition that radiotherapy-associated toxicity are manifested as damages to the normal tissue surrounding the site of lesion, we analyzed the association between SNPs and the incidence of radiation-induced thoracic toxicity. Consistent with existing studies demonstrating the associations of radiotherapy-associated toxicity with SNPs in genes in the DNA damage repair, oxidative reduction and metabolic pathways, we identified SNPs in X-ray repair cross-complementing 1 (*XRCC1*, rs25487, c.1196A>G; OR=2.31 [95%CI, 1.0-5.56]; *P*=0.05) and *XRCC5* (rs3835, c.2110-2408G>A; OR=3.59 [95% CI, 0.93-12.96]; *P*=0.03), which conferred increased risks of radiation-induced pneumonitis and esophagitis, respectively (**Figures 7A, B**). Further analysis revealed a stronger association between the *XRCC5* allele with severe (grade 3 or higher) esophagitis (OR=5.71 [95% CI, 1.30-25.0]; *P*=0.03). In addition to these two SNPs, *MTHFR* (rs1801133, c.665C>T) and NAD(P)H

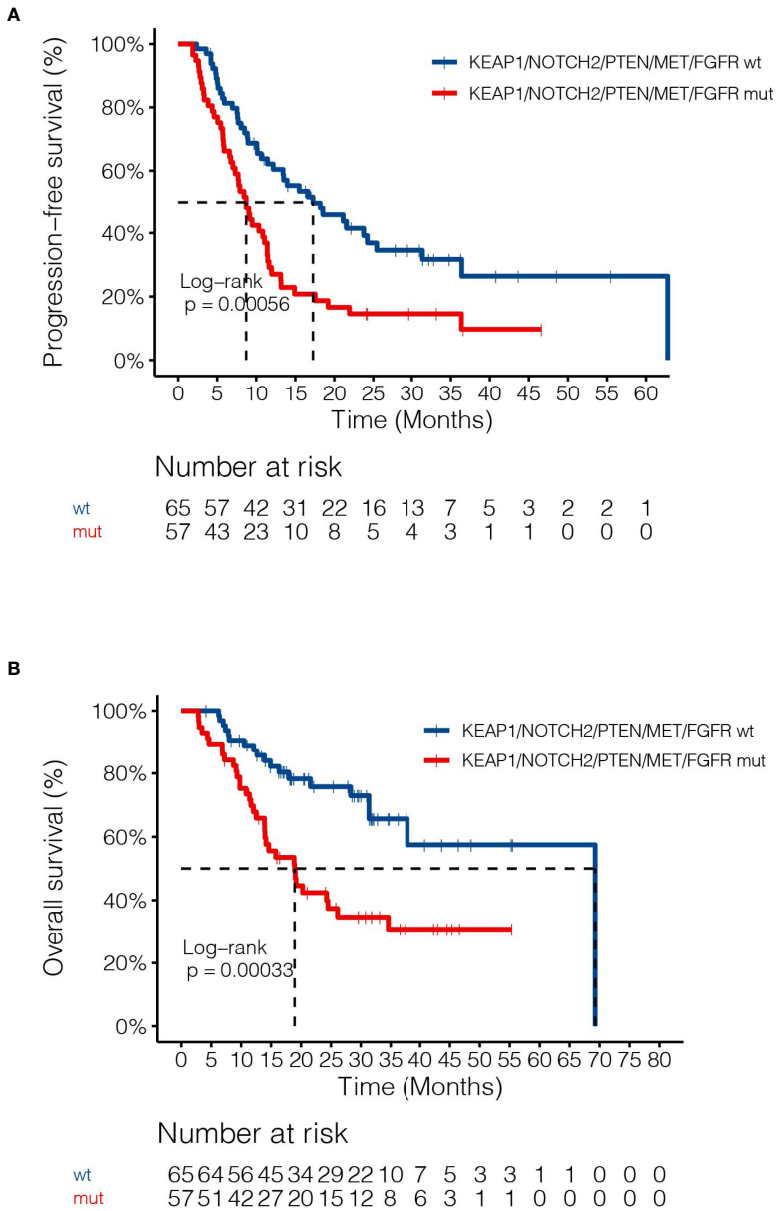



**FIGURE 5** | Associations of dCRT survival outcomes with the mutant subgroup. **(A, B)** Kaplan-Meier estimates of **(A)** PFS and **(B)** OS in the full analysis set comparing patients harboring any of the *MET*, *NOTCH2* and *PTEN* loss of function, as well as FGFR and KEAP1-NRF2 pathway gene alterations and those without.

Quinone Dehydrogenase 1 (*NQO1*, rs1800566, c.559C>T) were associated with trends towards higher incidence of radiotherapy-associated toxicity overall (**Figure 7C**).

We also observed enrichments of several somatic aberrations in patients who developed grade 2 or higher pneumonitis and/or esophagitis (**Figure 7D**). The zinc-finger protein 217 (*ZNF217*) gene is frequently amplified in human cancers (25, 26). It encodes a Kruppel-like transcription factor that mediates complex molecular processes through the regulation of gene expression. A total of 11 patients carried *ZNF217* amplifications, ten of which developed grade 2 or higher pneumonitis and esophagitis (90.9% vs. 36.9%, OR=16.7 [95% CI, 2.24-748.4], *P*=0.0007, **Figure 7D**). Of these, four developed radiation-induced esophagitis (36.4% vs. 10.8%, OR=4.6 [95% CI, 0.86-21.7], *P*=0.04) and eight developed pneumonitis (72.7% vs. 27.9%, OR=6.8 [95% CI, 1.5-42.1], *P*=0.005). Similarly, *ZNF217* amplification was associated with severe pneumonitis and esophagitis (grade 3 or higher; OR=7.5 [95% CI, 2.0-28.0], *P*=0.003), as well as severe radiation-induced pneumonitis (OR=4.9 [95% CI, 1.3-18.2], *P*=0.02). In addition, we found that mutations in *POLD1*, encoding the DNA polymerase delta 1 that is a key protein for ensuring the replicative fidelity of DNA,

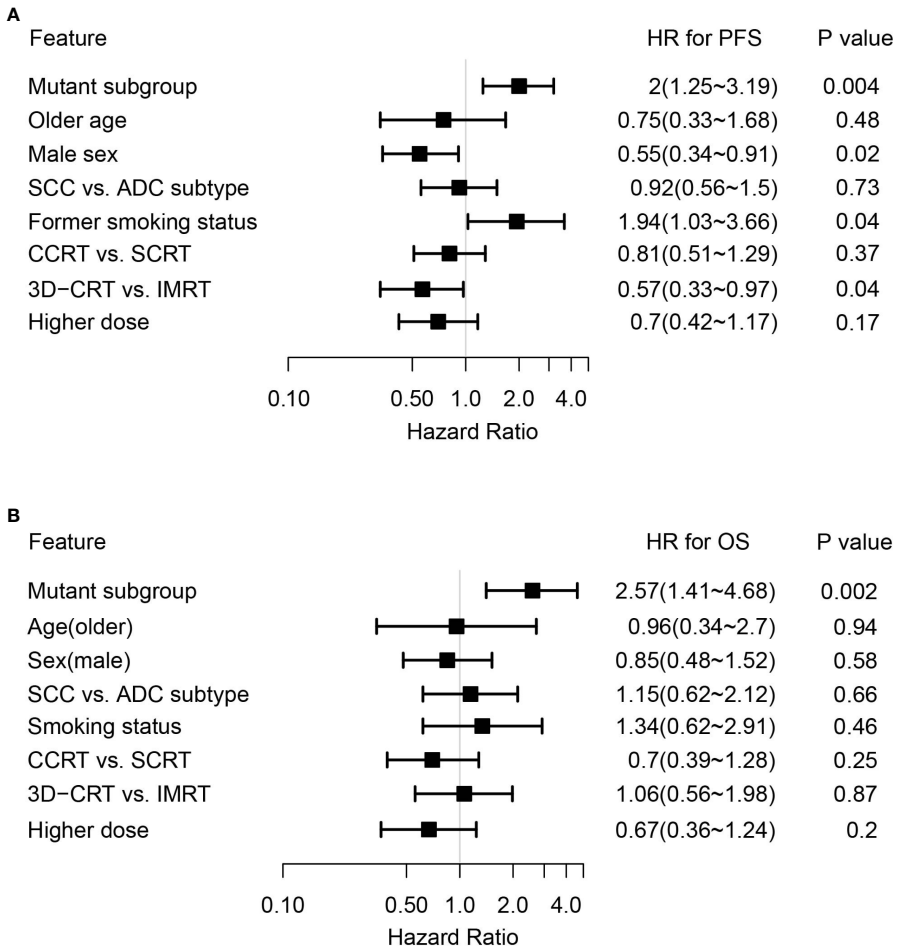



**FIGURE 6** | Multivariate Cox analysis of clinical features associated with survival outcomes in the genetically altered subgroup. **(A, B)** Forest plots showing the mutant subgroup and all relevant clinical features in association with **(A)** PFS and **(B)** OS following dCRT treatment by multivariate analysis. Mutant subgroup: patients with any of the *MET*, *NOTCH2* and *PTEN* loss of function, as well as FGFR and KEAP1-NRF2 pathway gene alterations; ADC: adenocarcinoma; SCC: squamous cell carcinoma.

were also associated with an increase in overall toxicity risk (pneumonitis and/or esophagitis, 77.8% vs. 38.9%, OR=5.4 [95% CI, 0.97-55.7], *P*=0.03, **Figure 7D**). Six (out of nine) *POLD1* mutations were predicted to have functional consequences (SIFT score ≤ 0.01). Of the seven patients who developed toxicity, six had pneumonitis and one had esophagitis. It is worth noting that somatic alterations in both genes showed stronger associations with radiation pneumonitis, as compared to esophagitis, likely reflecting interactions of these tumor cells with the local microenvironment.

## DISCUSSION

As we enter the era of personalized medicine, there is an area of unmet needs for identifying genetic determinants that explain individual differences in dCRT response. The development of disease recurrence and severe radiation-induced toxicity following dCRT could negatively impact patients' survival outcome and quality of life. In this study, by comprehensive profiling of 122 unresectable stage IIIA-C NSCLC patients who underwent dCRT, we identified a number of highly relevant and novel genetic and pathway-level features that might serve as potential biomarkers for predicting response to CRT.

The KEAP1-NRF2 pathway is often altered in NSCLC. In line with previous studies that have demonstrated the potential prognostic or predictive value of the dysregulation in the KEAP1-NRF2 pathway in NSCLC patients following chemotherapy and/or radiotherapy (20, 21, 27, 28), our findings independently confirmed its association with poor outcome in NSCLC patients following dCRT. Multiple others signaling pathways, including the MAPK, PI3K/AKT, FGFR and NOTCH pathways, have been implicated in radio-resistance, through their regulation of cellular proliferation, differentiation, apoptosis, invasion and maintenance of cancer stem cells (22, 23, 29–32). However, most findings were based on preclinical studies using *in vitro* or animal models. Here, we provided the first

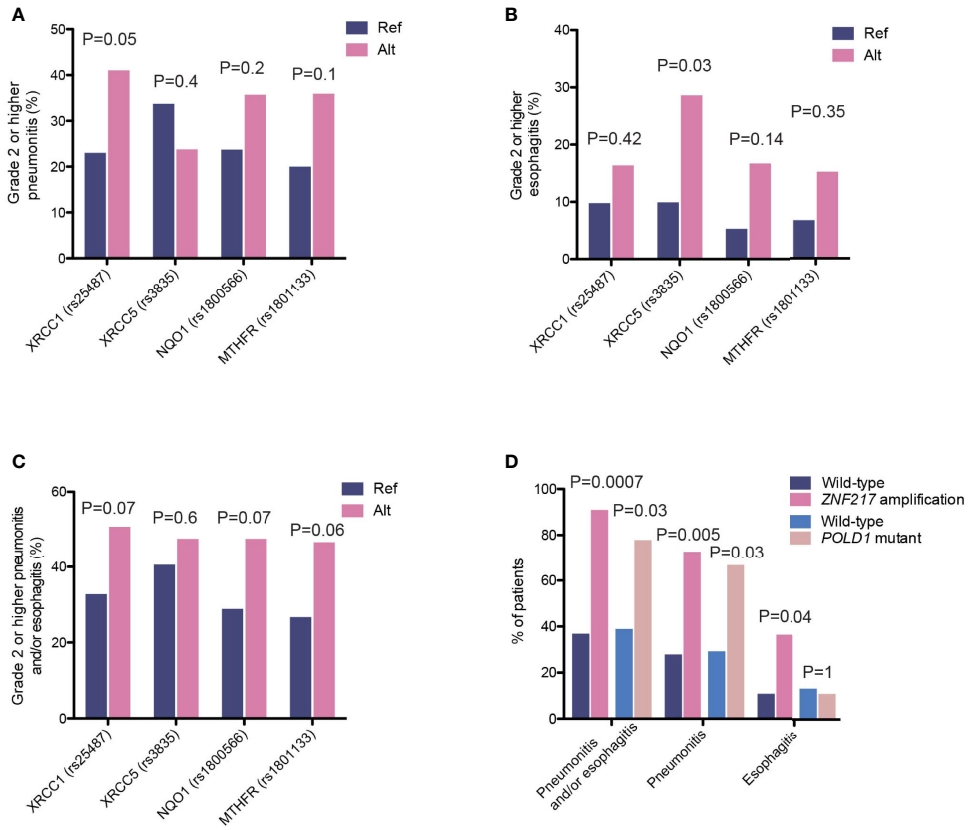



**FIGURE 7** | Genetic variants associated with incidence of high-grade radiation toxicity. **(A–C)** The proportions of patients carrying the indicated polymorphisms that developed high-grade **(A)** radiation-induced pneumonitis, **(B)** radiation-induced esophagitis, and **(C)** overall pneumonitis and esophagitis. **(D)** The proportions of patients carrying *ZNF217* amplification or *POLD1* mutations that developed high-grade radiation-induced toxicity events as indicated.

clinical evidence showing the associations of dCRT recurrence with several highly functionally relevant lung cancer genes, including *MET*, *PTEN*, *FGFR1-4*, and *NOTCH2*. Of these, *MET*, *PTEN* and *FGFR*s remained independent predictors of PFS by multivariate analysis. In particular, MET is commonly activated in NSCLC, which can be a result of gene amplification and exon 14 skipping mutations. In our study, the majority of *MET*-altered cases could lead to activation of the protein, which included *MET* exon 14 skipping mutations, gene rearrangement and amplification. While the presence of *MET* alterations was associated with worse outcome when treated with dCRT, no difference in OS was detected. Despite insufficient clinical information on the subsequent lines of treatment, these *MET*-altered cases might have derived long-term benefit from MET-targeted therapies. Similarly, while the presence of other potentially targetable driver mutations, such as *EGFR* activating mutations, had no impact on disease progression, it was associated with improved OS, which is also likely explained by the later use of targeted therapies in these patients. Thus, our data suggest that *MET* alteration is likely a negative predictive, rather than prognostic, factor of dCRT recurrence.

The FGFR family receptors activate multiple signaling pathways, including the RAS/MAPK, PI3K, and STAT pathways, which play important roles in cancer initiation and development (33). Emerging evidence suggest that FGFR may be implicated in variable response to radiotherapy. Pre-clinical studies in NSCLC cell lines, xenograft models and genetically engineered mouse models have shown that FGFR inhibition can enhance radiation response, which may be through the upregulation of cellular apoptosis and autophagy (23) and/or polarization of tumor-associated macrophages towards the M1 phenotype (24). In our study, we provide the first clinical evidence for the role of FGFR in mediating dCRT response. Patients with FGFR family gene alterations demonstrated reduced PFS outcome. Furthermore, the negative association of FGFRs with PFS also translated into poor OS outcome.

In addition to disease recurrence, a subset of patients would also develop severe, and often long-term, radiation-induced toxicities, which can negatively impact their quality of life. Understanding the genetic basis underlying individual differences in the development of dCRT-associated adverse events would allow for the risk stratification of patients and consequently personalized dCRT regimens, which would maximize tumor control while minimizing damage to the local tissue. The involvement of SNPs in the various damage and stress-response genes in mediating radiation-induced toxicities

has been extensively studied in multiple cancer types (9–12, 34–37). Early association studies have employed a candidate gene approach, which has led to the identification of several key genes that may serve as potential predictors of radiation-induced toxicities, including *ATM*, base excision repair genes (*XRCC1-5*), mismatch repair genes (*MSH2*, *MLH1*), and oxidative damage-detoxification genes (*GSTM1*, *GSTT1*) (38–45). As we shift from candidate gene approach to genome-wide association studies, such as the multi-centered RAPPER (Radiogenomics: Assessment of Polymorphisms for Predicting the Effects of Radiotherapy) study (46), more genetic polymorphisms associated with radio-toxicity have been identified. However, these studies are often underpowered and difficult to replicate due to the small effect size of individual SNPs on radiotoxicity, and rarely lead to clinically useful biomarkers. A gene-expression-based radio-sensitivity model has been developed (47–50) and clinically validated in multiple cancer types (51–55). These data provide the prescription framework for genomic-based radiotherapy and emphasize the importance of multi-gene testing as response to CRT is dependent on the combined influence of genetic variations at multiple loci. However, it remains challenging to adopt RNA-sequencing routinely into clinical practice, particularly on formalin-fixed paraffin-embedded samples.

Taking advantage of the NGS technology, our study independently verified the predictive potential of the DNA damage repair and oxidative stress pathway gene variants for radiotherapy-associated toxicity. Specifically, we identified polymorphisms in *XRCC1*/*XRCC5* (x-ray repair cross-complementing 1/5), encoding two key genes responsible for base excision repair, that were associated with differential risks of high-grade toxicity. The *XRCC1* rs25487 allele has been associated with severe oral mucositis in oropharyngeal carcinoma patients treated with radiotherapy (45). The *XRCC5* rs3835 allele has been implicated in the development of severe radiation pneumonitis in NSCLC patients (56). In our study, no significant association between this particular allele with radiation pneumonitis was observed. Instead, we showed that *XRCC5* might increase the risk of developing severe radiation esophagitis. In addition, we identified two risk alleles in *MTHFR* and *NQO1* that were associated with radiotoxicity. *MTHFR* encodes a methylenetetrahydrofolate reductase, which participates in folate metabolism and the regulation of DNA methylation and repair (57, 58). On the other hand, NQO1 is involved in the regulation of reactive oxygen species and continued oxidative stress can also induce DNA damage and chronic inflammation (59, 60). Combinatorial testing for these genetic variations might be useful for identifying patients who are susceptible to radiation toxicity. However, large-scale studies are needed to fully assess the predictive potential of these particular polymorphisms or variations in DNA damage repair and oxidative stress pathways.

Solid tumors often exhibit complex interactions with their surrounding tissues *via* stromal components, the vasculature and immune cells, among others (61). However, it has never been reported that somatic mutations could influence a patient's likelihood of developing radiation toxicity. Here, we report the associations of *ZNF217* amplification and *POLD1* mutations with increased likelihoods of developing radiation toxicity. In particular, *ZNF217* amplification was associated increased risks of developing severe (grade 3 or higher) pneumonitis and esophagitis. *ZNF217* is commonly amplified in human cancers (25, 26). While the presence of *ZNF217* amplification itself may indicate loss of genome integrity, there may also be functional consequences given that *ZNF217* encodes a transcription factor that mediates a diverse array of cellular processes through the regulation of various target gene expressions (62). Importantly, ZNF217 may have a role in DNA damage repair as it has been shown to repress the levels of BRCA1 (63). Likewise, the role of POLD1 in controlling replicative fidelity has been firmly established (64, 65). Thus, we speculate that *ZNF217* amplification and/or *POLD1* mutations in the tumor may affect overall genome stability and lead to the generation of tumor-specific neoantigens and consequently extensive lymphocyte infiltration (66). Conceivably, this could also exacerbate inflammation in the surrounding normal tissue following radiation. Interestingly, similar to our work in small-cell lung cancer (67), somatic alterations are more likely to affect radiotherapy-associated toxicity at the site of the lesion, as they were more commonly associated with pneumonitis than with esophagitis. Due to the retrospective nature of this current study, our cohort consisted of relatively high proportions of SCC and former smokers, as well as widely varied radiation doses which is associated with radiotherapy toxicity (68). Thus, the predictive value of these novel variations' merits further investigation. Nevertheless, our data indicate that extra caution should be exercised when giving radiotherapy to NSCLC patients carrying such mutations.

Our observations from clinical data of genetic associations with CRT survival outcome and toxicity provide a set of candidate predictive biomarkers present in normal and also tumor tissues. The mechanisms by which some of these genetic variants act to promote development of adverse response or cancer recurrence remain to be elucidated, although it is likely through their combined influence on important oncogenic signaling pathways, as well as DNA damage repair, oxidative and inflammatory response pathways. Due to the lack of sufficient data from published work or public databases such as TCGA, future work should involve validation of these potential biomarkers in a larger set of cohorts and generation of multifactorial prediction models of the expected treatment outcome. Taken together, our results demonstrate the clinical utility of NGS panels in identifying predictive biomarkers for response to CRT and suggest that testing for these susceptibility loci would prove beneficial in improving personalized CRT in NSCLC patients.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: Genome Sequence Archive for Human (GSA- Human); HRA002626.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Review Board of the Oncology Center of Shandong Provincial Hospital. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

Conceptualization: JY, SY, YS, XW, and XNW. Investigation: JY, SY, YS. Resources: DM, HG, JW, JM, ZY, XZ, LD, PC. Data curation: KH, ST, LL. Data analysis and interpretation: JP, JCY, JZ, DM and JM. Manuscript writing: All authors. Manuscript review & editing: All authors. Funding acquisition: JY. Supervision: JY, SY, YS. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by Natural Science Foundation of China (grant No. NSFC81872475 and NSFC82073345 to SY; NSFC81874254 to PC); and Research Unit of Radiation Oncology, Chinese Academy of Medical Sciences (2019RU071), the foundation of National Natural Science Foundation of China (NSFC81627901, NSFC81972863 and NSFC82030082), the foundation of Natural Science Foundation of Shandong (ZR201911040452) and the Academic Promotion Program of Shandong First Medical University (2019ZL002) to JY.

## ACKNOWLEDGMENT

The authors thank all investigators and research staff at the following hospitals and research sites: Fujian Zhangzhou Hospital, Nanjing Gulou Hospital, Cancer Hospital of The University of Chinese Academy of Science (Zhejiang Cancer Hospital), Wuhan University Renmin Hospital, Yantai Yuhuangding Hospital, Liuzhou Gongren Hospital, Huazhong University of Science and Technology for providing cases and samples.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.928605/full#supplementary-material

## REFERENCES

1. Herbst RS, Morgensztern D, Boshoff C. The Biology and Management of non-Small Cell Lung Cancer. *Nature* (2018) 553(7689):446–54. doi: 10.1038/nature25183
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer Statistics in China, 2015. *CA Cancer J Clin* (2016) 66(2):115–32. doi: 10.3322/caac.21338
3. Dillman RO, Seagren SL, Propert KJ, Guerra J, Eaton WL, Perry MC, et al. A Randomized Trial of Induction Chemotherapy Plus High-Dose Radiation Vesus Radiation Alone in Stage III non-Small-Cell Lung Cancer. *N Engl J Med* (1990) 323:940–5. doi: 10.1056/NEJM199010043231403
4. Arroyo-Hernandez M, Maldonado F, Lozano-Ruiz F, Munoz-Montano W, Nunez-Baez M, Arrieta O. Radiation-Induced Lung Injury: Current Evidence. *BMC Pulm Med* (2021) 21(1):9. doi: 10.1186/s12890-020-01376-4
5. Daly ME, Singh N, Ismaila N, Antonoff MB, Arenberg DA, Bradley J, et al. Management of Stage III Non-Small-Cell Lung Cancer: ASCO Guideline. *J Clin Oncol* (2022) 40(12):1356–84. doi: 10.1200/JCO.21.02528
6. Barnett GC, West CML, Dunning AM, Elliott RM, Coles CE, Pharoah PDP, et al. Normal Tissue Reactions to Radiotherapy: Towards Tailoring Treatment Dose by Genotype. *Nat Rev Cancer* (2009) 9(2):134–42. doi: 10.1038/nrc2587
7. Taylor AM, Harnden DG, Arlett CF, Harcourt SA, Lehmann AR, Stevens S, et al. Ataxia Telangiectasia: A Human Mutation With Abnormal Radiation Sensitivity. *Nature* (1975) 258:427–9. doi: 10.1038/258427a0
8. Kleinerman RA. Radiation-Sensitive Genetically Susceptible Pediatric Sub-Populations. *Pediatr Radiol* (2009) 39(Suppl 1):S27–31. doi: 10.1007/s00247-008-1015-6
9. Dayal R, Singh A, Pandey A, Mishra KP. Reactive Oxygen Species as Mediator of Tumor Radiosensitivity. *J Cancer Res Ther* (2014) 10(4):811–8. doi: 10.4103/0973-1482.146073
10. Multhoff G, Radons J. Radiation, Inflammation, and Immune Responses in Cancer. *Front Oncol* (2012) 2:58. doi: 10.3389/fonc.2012.00058
11. Tang Y, Liu B, Li J, Wu H, Yang J, Zhou X, et al. Genetic Variants in PI3K/AKT Pathway are Associated With Severe Radiation Pneumonitis in Lung Cancer Patients Treated With Radiation Therapy. *Cancer Med* (2016) 5(1):24–32. doi: 10.1002/cam4.564
12. Cordes N, Hildebrandt MAT, Komaki R, Liao Z, Gu J, Chang JY, et al. Genetic Variants in Inflammation-Related Genes Are Associated With Radiation-Induced Toxicity Following Treatment for Non-Small Cell Lung Cancer. *PloS One* (2010) 5(8):e12402. doi: 10.1371/journal.pone.0012402
13. He K, Zhang S, Shao LL, Yin JC, Wu X, Shao YW, et al. Developing More Sensitive Genomic Approaches to Detect Radioresponse in Precision Radiation Oncology: From Tissue DNA Analysis to Circulating Tumor DNA. *Cancer Lett* (2020) 472:108–18. doi: 10.1016/j.canlet.2019.12.004
14. Cox JD, Stetz J, Pajak TF. Toxicity Criteria of the Radiation Therapy Oncology Group (RTOG) and the European Organization for Research and Treatment of Cancer (EORTC). *Int J Radiat Oncol Biol Phys* (1995) 31(5):1341–6. doi: 10.1016/0360-3016(95)00060-C
15. Bolger AM, Lohse M, Usadel B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* (2014) 30(15):2114–20. doi: 10.1093/bioinformatics/btu170
16. Li H, Durbin R. Fast and Accurate Short Read Alignment With Burrows-Wheeler Transform. *Bioinformatics* (2009) 25(14):1754–60. doi: 10.1093/bioinformatics/btp324
17. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: Somatic Mutation and Copy Number Alteration Discovery in Cancer by Exome Sequencing. *Genome Res* (2012) 22(3):568–76. doi: 10.1101/gr.129684.111
18. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization From Targeted DNA Sequencing. *PloS Comput Biol* (2016) 12(4):e1004873. doi: 10.1371/journal.pcbi.1004873
19. Fang W, Ma Y, Yin JC, Hong S, Zhou H, Wang A, et al. Comprehensive Genomic Profiling Identifies Novel Genetic Predictors of Response to Anti-PD-(L)1 Therapies in Non-Small Cell Lung Cancer. *Clin Cancer Res* (2019) 25(16):5015–26. doi: 10.1158/1078-0432.CCR-19-0585
20. Kansanen E, Kuosmanen SM, Leinonen H, Levonen A-L. The Keap1-Nrf2 Pathway: Mechanisms of Activation and Dysregulation in Cancer. *Redox Biol* (2013) 1(1):45–9. doi: 10.1016/j.redox.2012.10.001

21. Jeong Y, Hoang NT, Lovejoy A, Stehr H, Newman AM, Gentles AJ, et al. Role of KEAP1/NRF2 and TP53 Mutations in Lung Squamous Cell Carcinoma Development and Radiation Resistance. *Cancer Discov* (2017) 7(1):86–101. doi: 10.1158/2159-8290.CD-16-0127
22. Yahyanejad S, Theys J, Vooijs M. Targeting Notch to Overcome Radiation Resistance. *Oncotarget* (2015) 7(7):7610–28. doi: 10.18632/oncotarget.6714
23. SenthilKumar G, Fisher MM, Skiba JH, Miller MC, Brennan SR, Kaushik S, et al. FGFR Inhibition Enhances Sensitivity to Radiation in Non–Small Cell Lung Cancer. *Mol Cancer Ther* (2020) 19(6):1255–65. doi: 10.1158/1535-7163.MCT-19-0931
24. Im JH, Buzzelli JN, Jones K, Franchini F, Gordon-Weeks A, Markelc B, et al. FGF2 Alters Macrophage Polarization, Tumour Immunity and Growth and can be Targeted During Radiotherapy. *Nat Commun* (2020) 11(1):4064. doi: 10.1038/s41467-020-17914-x
25. Cohen PA, Donini CF, Nguyen NT, Lincet H, Vendrell JA. The Dark Side of ZNF217, a Key Regulator of Tumorigenesis With Powerful Biomarker Value. *Oncotarget* (2015) 6(39):41566–81. doi: 10.18632/oncotarget.5893
26. Lee D-F, Walsh MJ, Aguiló F. ZNF217/ZFP217 Meets Chromatin and RNA. *Trends Biochem Sci* (2016) 41(12):986–8. doi: 10.1016/j.tibs.2016.07.013
27. Cameron BD, Sekhar KR, Ofori M, Freeman ML. The Role of Nrf2 in the Response to Normal Tissue Radiation Injury. *Radiat Res* (2018) 190(2):99. doi: 10.1667/RR15059.1
28. Goeman F, De Nicola F, Scalera S, Sperati F, Gallo E, Ciuffreda L, et al. Mutations in the KEAP1-NFE2L2 Pathway Define a Molecular Subset of Rapidly Progressing Lung Adenocarcinoma. *J Thorac Oncol* (2019) 14 (11):1924–34. doi: 10.1016/j.jtho.2019.07.003
29. Hassan KA, Wang L, Korkaya H, Chen G, Maillard I, Beer DG, et al. Notch Pathway Activity Identifies Cells With Cancer Stem Cell–like Properties and Correlates With Worse Survival in Lung Adenocarcinoma. *Clin Cancer Res* (2013) 19(8):1972–80. doi: 10.1158/1078-0432.CCR-12-0370
30. Hasslacher S, Schneele L, Stroh S, Langhans J, Zeiler K, Kattner P, et al. Inhibition of PI3K Signalling Increases the Efficiency of Radiotherapy in Glioblastoma Cells. *Int J Oncol* (2018) 53(5):1881–96. doi: 10.3892/ijo.2018.4528
31. Sklar MD. The Ras Oncogenes Increase the Intrinsic Resistance of NIH 3T3 Cells to Ionizing Radiation. *Science* (1988) 239(4840):645–7. doi: 10.1126/science.3277276
32. Dent P, Yacoub A, Fisher PB, Hagan MP, Grant S. MAPK Pathways in Radiation Responses. *Oncogene* (2003) 22(37):5885–96. doi: 10.1038/sj.onc.1206701
33. Turner N, Grose R. Fibroblast Growth Factor Signalling: From Development to Cancer. *Nat Rev Cancer* (2010) 10(2):116–29. doi: 10.1038/nrc2780
34. Barnett GC, Coles CE, Elliott RM, Baynes C, Luccarini C, Conroy D, et al. Independent Validation of Genes and Polymorphisms Reported to be Associated With Radiation Toxicity: A Prospective Analysis Study. *Lancet Oncol* (2012) 13(1):65–77. doi: 10.1016/S1470-2045(11)70302-3
35. Alsner J, Andreassen CN, Overgaard J. Genetic Markers for Prediction of Normal Tissue Toxicity After Radiotherapy. *Semin Radiat Oncol* (2008) 18 (2):126–35. doi: 10.1016/j.semradonc.2007.10.004
36. Georgakilas AG, Pavlopoulou A, Louka M, Nikitaki Z, Vorgias CE, Bagos PG, et al. Emerging Molecular Networks Common in Ionizing Radiation, Immune and Inflammatory Responses by Employing Bioinformatics Approaches. *Cancer Lett* (2015) 368(2):164–72. doi: 10.1016/j.canlet.2015.03.021
37. Rosenstein B. Identification of SNPs Associated With Susceptibility for Development of Adverse Reactions to Radiotherapy. *Pharmacogenomics* (2011) 12(2):267–75. doi: 10.2217/pgs.10.186
38. Zhang L, Yang M, Bi N, Fang M, Sun T, Ji W, et al. ATM Polymorphisms Are Associated With Risk of Radiation-Induced Pneumonitis. *Int J Radiat Oncol Biol Physics* (2010) 77(5):1360–8. doi: 10.1016/j.ijrobp.2009.07.1675
39. Xiong H, Liao Z, Liu Z, Xu T, Wang Q, Liu H, et al. ATM Polymorphisms Predict Severe Radiation Pneumonitis in Patients With Non-Small Cell Lung Cancer Treated With Definitive Radiation Therapy. *Int J Radiat Oncol Biol Physics* (2013) 85(4):1066–73. doi: 10.1016/j.ijrobp.2012.09.024
40. Seibold P, Behrens S, Schmezer P, Helmbold I, Barnett G, Coles C, et al. XRCC1 Polymorphism Associated With Late Toxicity After Radiation Therapy in Breast Cancer Patients. *Int J Radiat Oncol Biol Phys* (2015) 92 (5):1084–92. doi: 10.1016/j.ijrobp.2015.04.011
41. Moullan N, Cox DG, Angele S, Romestaing P, Gerard JP, Hall J. Polymorphisms in the DNA Repair Gene XRCC1, Breast Cancer Risk, and Response to Radiotherapy. *Cancer Epidemiol Biomark Prev* (2003) 12:1168–74.
42. Yin M, Liao Z, Liu Z, Wang L-E, Gomez D, Komaki R, et al. Functional Polymorphisms of Base Excision Repair Genes XRCC1 and APEX1 Predict Risk of Radiation Pneumonitis in Patients With Non–Small Cell Lung Cancer Treated With Definitive Radiation Therapy. *Int J Radiat Oncol Biol Physics* (2011) 81(3):e67–73. doi: 10.1016/j.ijrobp.2010.11.079
43. Mangoni M, Bisanzi S, Carozzi F, Sani C, Biti G, Livi L, et al. Association Between Genetic Polymorphisms in the XRCC1, XRCC3, XPD, GSTM1, GSTT1, MSH2, MLH1, MSH3, and MGMT Genes and Radiosensitivity in Breast Cancer Patients. *Int J Radiat Oncol Biol Physics* (2011) 81(1):52–8. doi: 10.1016/j.ijrobp.2010.04.023
44. Venkatesh GH, Manjunath VB, Mumbrekar KD, Negi H, Fernandes DJ, Sharan K, et al. Polymorphisms in Radio-Responsive Genes and Its Association With Acute Toxicity Among Head and Neck Cancer Patients. *PloS One* (2014) 9(3):e89079. doi: 10.1371/journal.pone.0089079
45. Gupta A, Mathew D, Bhat SA, Ghoshal S, Pal A. Genetic Variants of DNA Repair Genes as Predictors of Radiation-Induced Subcutaneous Fibrosis in Oropharyngeal Carcinoma. *Front Oncol* (2021) 11:652049. doi: 10.3389/fonc.2021.652049
46. Burnet NG, Barnett GC, Elliott RM, Dearnaley DP, Pharoah PD, Dunning AM, et al. RAPPER: The Radiogenomics of Radiation Toxicity. *Clin Oncol (R Coll Radiol)* (2013) 25(7):431–4. doi: 10.1016/j.clon.2013.04.001
47. Eschrich S, Zhang H, Zhao H, Boulware D, Lee JH, Bloom G, et al. Systems Biology Modeling of the Radiation Sensitivity Network: A Biomarker Discovery Platform. *Int J Radiat Oncol Biol Phys* (2009) 75(2):497–505. doi: 10.1016/j.ijrobp.2009.05.056
48. Eschrich SA, Pramana J, Zhang H, Zhao H, Boulware D, Lee JH, et al. A Gene Expression Model of Intrinsic Tumor Radiosensitivity: Prediction of Response and Prognosis After Chemoradiation. *Int J Radiat Oncol Biol Phys* (2009) 75 (2):489–96. doi: 10.1016/j.ijrobp.2009.06.014
49. Torres-Roca JF, Eschrich S, Zhao H, Bloom G, Sung J, McCarthy S, et al. Prediction of Radiation Sensitivity Using a Gene Expression Classifier. *Cancer Res* (2005) 65(16):7169–76. doi: 10.1158/0008-5472.CAN-05-0656
50. Scott JG, Berglund A, Schell MJ, Mihaylov I, Fulp WJ, Yue B, et al. A Genome-Based Model for Adjusting Radiotherapy Dose (GARD): A Retrospective, Cohort-Based Study. *Lancet Oncol* (2017) 18(2):202–11. doi: 10.1016/S1470-2045(16)30648-9
51. Eschrich SA, Fulp WJ, Pawitan Y, Foekens JA, Smid M, Martens JW, et al. Validation of a Radiosensitivity Molecular Signature in Breast Cancer. *Clin Cancer Res* (2012) 18(18):5134–43. doi: 10.1158/1078-0432.CCR-12-0891
52. Ahmed KA, Chinnaiyan P, Fulp WJ, Eschrich S, Torres-Roca JF, Caudell JJ. The Radiosensitivity Index Predicts for Overall Survival in Glioblastoma. *Oncotarget* (2015) 6:34413–22. doi: 10.18632/oncotarget.5437
53. Strom T, Hoffe SE, Fulp W, Frakes J, Coppola D, Springett GM, et al. Radiosensitivity Index Predicts for Survival With Adjuvant Radiation in Resectable Pancreatic Cancer. *Radiother Oncol* (2015) 117(1):159–64. doi: 10.1016/j.radonc.2015.07.018
54. Torres-Roca JF, Fulp WJ, Caudell JJ, Servant N, Bollet MA, van de Vijver M, et al. Integration of a Radiosensitivity Molecular Signature Into the Assessment of Local Recurrence Risk in Breast Cancer. *Int J Radiat Oncol Biol Phys* (2015) 93(3):631–8. doi: 10.1016/j.ijrobp.2015.06.021
55. Ahmed KA, Fulp WJ, Berglund AE, Hoffe SE, Dilling TJ, Eschrich SA, et al. Differences Between Colon Cancer Primaries and Metastases Using a Molecular Assay for Tumor Radiation Sensitivity Suggest Implications for Potential Oligometastatic SBRT Patient Selection. *Int J Radiat Oncol Biol Phys* (2015) 92(4):837–42. doi: 10.1016/j.ijrobp.2015.01.036
56. Yin M, Liao Z, Liu Z, Wang L-E, O'Reilly M, Gomez D, et al. Genetic Variants of the Nonhomologous End Joining Genelig4and Severe Radiation Pneumonitis in Nonsmall Cell Lung Cancer Patients Treated With Definitive Radiotherapy. *Cancer* (2012) 118(2):528–35. doi: 10.1002/cncr.26214
57. Zhang Y, Li Z, Zhang J, Li H, Qiao Y, Huang C, et al. Genetic Variants in MTHFR Gene Predict ≥ 2 Radiation Pneumonitis in Esophageal Squamous Cell Carcinoma Patients Treated With Thoracic Radiotherapy. *PloS One* (2017) 12(1):e0169147. doi: 10.1371/journal.pone.0169147

58. Mak RH, Alexander BM, Asomaning K, Heist RS, Liu CY, Su L, et al. A Single-Nucleotide Polymorphism in the Methylene Tetrahydrofolate Reductase (MTHFR) Gene is Associated With Risk of Radiation Pneumonitis in Lung Cancer Patients Treated With Thoracic Radiation Therapy. *Cancer* (2012) 118 (14):3654–65. doi: 10.1002/cncr.26667
59. Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. Oxidative Stress, Inflammation, and Cancer: How are They Linked? *Free Radic Biol Med* (2010) 49(11):1603–16. doi: 10.1016/j.freeradbiomed.2010.09.006
60. Dinkova-Kostova AT, Talalay P. NAD(P)H:quinone Acceptor Oxidoreductase 1 (NQO1), a Multifunctional Antioxidant Enzyme and Exceptionally Versatile Cytoprotector. *Arch Biochem Biophys* (2010) 501 (1):116–23. doi: 10.1016/j.abb.2010.03.019
61. Egeblad MN, Nakasone ES, Werb Z. Tumors as Organs: Complex Tissues That Interface With the Entire Organism. *Dev Cell Rev* (2010) 18(6):884–901. doi: 10.1016/j.devcel.2010.05.012
62. Thillainadesan G, Isovic M, Loney E, Andrews J, Tini M, Torchia J. Genome Analysis Identifies the P15ink4b Tumor Suppressor as a Direct Target of the ZNF217/CoREST Complex. *Mol Cell Biol* (2008) 28(19):6066–77. doi: 10.1128/MCB.00246-08
63. Banck MS, Li S, Nishio H, Wang C, Beutler AS, Walsh MJ. The ZNF217 Oncogene is a Candidate Organizer of Repressive Histone Modifiers. *Epigenetics* (2009) 4(2):100–6. doi: 10.4161/epi.4.2.7953
64. Wang F, Zhao Q, Wang YN, Jin Y, He MM, Liu ZX, et al. Evaluation of POLE and POLD1 Mutations as Biomarkers for Immunotherapy Outcomes Across Multiple Cancer Types. *JAMA Oncol* (2019) 5(10):1504–6. doi: 10.1001/jamaoncol.2019.2963
65. Church DN, Briggs SE, Palles C, Domingo E, Kearsey SJ, Grimes JM, et al. DNA Polymerase Epsilon and Delta Exonuclease Domain Mutations in Endometrial Cancer. *Hum Mol Genet* (2013) 22(14):2820–8. doi: 10.1093/hmg/ddt131
66. Nolan E, Savas P, Policheni AN, Darcy PK, Vaillant F, Mintoff CP, et al. Combined Immune Checkpoint Blockade as a Therapeutic Strategy for BRCA1-Mutated Breast Cancer. *Sci Trans Med* (2017) 9(393):eaal4922. doi: 10.1126/scitranslmed.aal4922
67. Li L, Tang S, Yin JC, Dong L, Yang Z, Liu Y, et al. Comprehensive Next-Generation Sequencing Reveals Novel Predictive Biomarkers of Recurrence and Thoracic Toxicity Risks After Chemoradiation Therapy in Limited Stage Small Cell Lung Cancer. *Int J Radiat Oncol Biol Phys* (2022) 112(5):1165–76. doi: 10.1016/j.ijrobp.2021.12.009
68. He J, Dong G, Deng Y, He J, Xiu Z, Feng F. Comparison of Application Value of Different Radiation Dose Evaluation Methods in Evaluating Radiation Dose of Adult Thoracic and Abdominal CT Scan. *Front Surg* (2022) 9. doi: 10.3389/fsurg.2022.860968

**Conflict of Interest:** JP, JY, JZ, HB, XW, XNW and YS are staff of Nanjing Geneseeq Technology Inc.

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

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 He, Zhang, Pang, Yin, Mu, Wang, Ge, Ma, Yang, Zheng, Dong, Zhang, Chang, Li, Tang, Bao, Wu, Wang, Shao, Yu and Yuan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 25 July 2022
DOI 10.3389/fonc.2022.903484

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Qianqian Song,
Wake Forest School of Medicine, United States
Huimei Wang,
Fudan University, China

*CORRESPONDENCE
Jinhui Zhu
2512016@zju.edu.cn

†These authors share first authorship

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

RECEIVED 24 March 2022
ACCEPTED 27 June 2022
PUBLISHED 25 July 2022

CITATION
Fu R, Shao Q, Yang B, Chen Y, Ye Q, Chen X and Zhu J (2022) MiR-520a-5p/PPP5C regulation pattern is identified as the key to gemcitabine resistance in pancreatic cancer. *Front. Oncol.* 12:903484. doi: 10.3389/fonc.2022.903484

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

# MiR-520a-5p/PPP5C regulation pattern is identified as the key to gemcitabine resistance in pancreatic cancer

Ruibiao Fu†, Qian Shao, Bin Yang, Yan Chen, Qinghuang Ye, Xi Chen† and Jinhui Zhu*

Department of Hepato-Pancreato-Biliary (HPB) Surgery, Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

**Objective:** To explore the effects of the expression level of miR-520-5p/PPP5C in pancreatic cancer cells and exosomes on cell viability, angiogenesis, autophagy, which involved in the mechanism of gemcitabine resistance in pancreatic cancer.

**Methods:** APSC-1 cell line was treated with gemcitabine, after which its exosomes were extracted
for NTA assay. Subsequently, the drug resistance of APSC-1 cells was assayed using CCK8, as well as the activity of HUVEC cells treated with exosomes from each group of APSC-1 cells after drug resistance treatment as well as overexpression treatment. Five groups of HUVEC cells treated with exosomes were subjected to *in vitro* tubule formation assay. levels of PPP5C in each group of ASPC-1 cells and their exosomes, levels of overexpressed PPP5C, and related exosomal proteins were examined by WB. mRNA expression levels of PPP5C and levels of miR-520a were examined by qPCR The relationship between miR-520a-5p and PPP5C was investigated. After that, the autophagy of PPP5C was detected. Finally, it was analyzed by TCGA database for survival prognosis analysis.

**Results:** APSC-1 cells had an IC50 value of 227.1 μM for gemcitabine, elevated PPP5C expression, drug resistance, and enhanced HUVEC cell activity; exosomes CD9, CD63, and CD81 were significantly expressed in all groups; meanwhile, enhanced PPP5C expression not only promoted *in vitro* tubule formation but also increased autophagy levels; meanwhile, its relationship with miR-520-5p and There was a targeted inhibitory relationship between its level and miR-520-5p and PPP5C, and its elevated level also led to a decrease in the survival level of patients over 3-5 years.

**Conclusion:** PPP5C has a prognostic role in pancreatic cancer by promoting the value-added and invasion of pancreatic cancer cells, and a targeted inhibitory relationship between miR-520-5p and PPP5C was found.

KEYWORDS

**Pancreatic cancer, PPP5C, miR-520-5p, gemcitabine, autophagy**

# 1 Introduction

Pancreatic cancer is one of the most lethal human tumors with a poor prognosis and a five-year survival rate of less than 10% (1). The incidence and mortality of this disease is predicted to increase dramatically worldwide by 2040 (2). The extreme lack of treatment options, the frequent occurrence of chemotherapy resistance, and its metastatic nature make the prognosis of pancreatic cancer clinical treatment very poor (3–6). And since the occurrence of PPP5C is closely related to the abnormal activation of tumor driver genes, the exploration of pancreatic cancer-related genes has been of increasing interest to researchers in recent years (7).

Through genome-wide association studies, increasing evidence has identified significant associations between pancreatic cancer susceptibility, variants in relevant genes (at least 23 genome-wide significant susceptibility loci), and pancreatic cancer risk (8, 9). However, because of the small sample size of genome-wide association studies related to pancreatic cancer, compared with, for example, breast and colorectal cancer, there are fewer associated genes and fewer variant sites in pancreatic cancer (8). One of these, a member of the phosphoprotein phosphatase (PPP) family of Ser/Thr phosphatases, Ser/Thr protein phosphatase 5 (PPP5C), is encoded by the PPP5C gene (10). Substrates of PPP5C include the glucocorticoid receptor (GR), the tumor suppressor p53, Hsp90, and the co-chaperone Cdc37 (11). Because of this property, PPP5C has been linked to asthma, cardiac contractility and heart failure, diabetes, lipid metabolism, and obesity (12–14). On the cancer side, the role of PPP5C in proliferation and cell survival and its unique structure make it a potentially attractive therapeutic target, and elevated PPP5C expression has been found to increase proliferation in most cells in breast and renal cancer clinical studies (15). The role of PPP5C in pancreatic cancer is similar to that in other cancers: it leads to the progression of pancreatic cancer by enabling tumors to develop resistance to gemcitabine (16). Therefore, targeting the inhibition of PPP5C is particularly important in the treatment of pancreatic cancer, while miRNAs regulate the properties of mRNA expression by binding to its 3'-UTR (17, 18). A type of miRNA, miR-520a-5p in NSCLC, can reverse cancer progression in a near step through enrichment, and a study showed that miR-520a-5p is a tumor suppressor in NSCLC and plays an important role in angiogenesis (19). However, the role of this type of miRNA in pancreatic cancer, as well as its relationship with PPP5C, has been less well-studied.

This study will investigate the expression levels of PPP5C in cells as well as exosomes and the effect of its expression levels on gemcitabine resistance, cellular activity, angiogenesis, and cellular autophagy levels in pancreatic cancer, analyze its targeting relationship with miR-520-5p, and obtain the effect of PPP5C on survival over 3-5 years by bioinformatics analysis.

# 2 Materials and methods

## 2.1 Methods

### 2.1.1 Construction of gemcitabine-resistant APSC-1 cell line

APSC-1 cells were treated with different doses of gemcitabine: 0 μM, 0.05 μM, 0.5 μM, 5 μM, 50 μM and 500 μM for 48 h. After these treatments, the cell lines were tested for drug resistance.

### 2.1.2 Exosome extraction

Supernatant exosomes were extracted from APSC-1 cells and the particles obtained from supernatants of untreated APSC-1 cells and gemcitabine-resistant APSC-1 cells were analyzed using nanoparticle tracking analysis.

### 2.1.3 Cell viability assay

Treated cells were added to CCK-8 and incubated at 37°C for 4 h afterwards. The absorbance value OD450 of each well was measured using an enzyme marker.

### 2.1.4 *In vitro* tubule formation assay

HUVEC cells were treated with each group of APSC-1 cell exosomes and divided into five groups for culture: HUVEC, HUVEC+ APSC-1-Exo, HUVEC+Ge-APSC-1-Exo, HUVEC+ APSC-1-R-Exo, and HUVEC+Ge-APSC-1-R-Exo. The above groups were assayed using CCK-8 cell activity. *In vitro* tubule formation assay was performed on five groups of HUVEC cells treated with exosomes.

### 2.1.5 Overexpression of PPP5C construct

HUVEC and APSC-1 were overexpressed and divided into: HUVEC+control plasmid, HUVEC+PPP5C overexpression plasmid; APSC-1+control plasmid, APSC-1+PPP5C overexpression plasmid.

### 2.1.6 WB

To detect the protein expression level of PPP5C using WB, the protein was first extracted from the cells or exosomes, followed by electrophoresis operation. After electrophoresis, a formal immunoblotting operation is performed to obtain the relevant film. The relevant primer sequences are shown in Table 1.

### 2.1.7 QPCR

The differentially expressed miRNAs, hsa-miR-520a-5p, hsa-miR-7-1-3p, and hsa-miR-874-3p, were screened by whole-transcriptome sequencing. qPCR was then used to detect the results of these miRNAs in APSC-1 and its

TABLE 1 Primer sequence list.

| Gene | Primer | Sequence (5'-3') |
|---|---|---|
| PPP5C | Forward | 5'-AAGTTCTACAGCCAGGCCAT-3' |
| | Reverse | 5'-ATCCTTGTCATGGGGCTTCA-3' |
| GAPDH | Forward | 5'-TCAAGAAGGTGGTGAAGCAGG-3 |
| | Reverse | 5'-TCAAAGGTGGAGGAGTGGGT-3' |

exosomes, and those with consistent expression changes in cells and exosomes, and those with sequencing results, were selected as indicators for subsequent studies. The RNA expression level of PPP5C, which also needs to be detected by qPCR, was selected as an indicator for subsequent studies. Relevant primer sequences (Table 2).

### 2.1.8 Firefly luciferase

After performing cell culture and transfection using the kit, firefly and sea kidney luciferase activities were analyzed using a dual luciferase reporter assay (Promega) to verify the correspondence between hsa-miR-520a-5p and its target gene PPP5C (Luciferase Reporter assays).

### 2.1.9 Autophagic flow assay

After overexpression treatment of HUVEC and APSC-1, the expression levels of LC3 II/I, p62 in them were detected using WB. Also for the treated cells, 2 μL of m RFP- GFP-LC3 adenovirus transfection was added to each group for 24 to 36 h. The corresponding stimuli were added to intervene for 12 h, respectively, and sent to the confocal microscope room for observation and photographic preservation to detect the cellular autophagy level.

### 2.1.10 Bioinformatics analysis

The effect of PPP5C on the survival prognosis of pancreatic cancer was analyzed using the TCGA database.

## 3 Results

### 3.1 PPP5C overexpression increases pancreatic cancer cell activity

After CCK-8 assay, the IC50 value of gemcitabine in ASPC-1 cells was found to be 227.1 μM (Figure 1A). After exosome treatment, HUVEC cell viability was increased in the ASPC-1-Exo-treated group compared with the control HUVEC cells; cell viability was further upregulated in the Ge-ASPC-1-Exo-treated group and ASPC-1-R-Exo-treated group compared with the ASPC-1-Exo-treated group; Ge-ASPC-1-R-Exo-treated group was further upregulated compared with the ASPC-1-Exo-treated group; and Ge-ASPC-1-Exo-R-Exo-treated group was further upregulated compared with the ASPC-1-Exo-treated group. The cell viability was further increased in the Ge-ASPC-1-R-Exo treated group compared with the ASPC-1-R-Exo treated group (Figure 1B). After overexpression treatment, cell activity was increased after PPP5C overexpression treatment compared to the control group (Figure 1C).

### 3.2 Elevated expression of PPP5C promotes tubule formation *in vitro*

After examination, compared with control HUVEC cells, *in vitro* tubule formation was increased in HUVEC+ASPC-1-Exo group; compared with HUVEC+ASPC-1-Exo, *in vitro* tubule formation was further increased in HUVEC+Ge-

TABLE 2 Primer sequences used for gene detection.

| Gene | Primer | Sequence (5'-3') |
|---|---|---|
| U6 | Forward | CGCTTCGGCAGCACATATAC |
| | Reverse | AAATATGGAACGCTTCACGA |
| hsa-miR-520a-5p | loop primer | GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGAAAGTA |
| | Forward | TGCGCCTCCAGAGGGAAGTAC |
| hsa-miR-7-1-3p | loop primer | GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTATGGCAG |
| | Forward | TGCGCCAACAAATCACAGTCTG |
| hsa-miR-874-3p | loop primer | GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACTCGGTCCC |
| | Forward | TGCGCCTGCCCTGGCCCGAGGG |

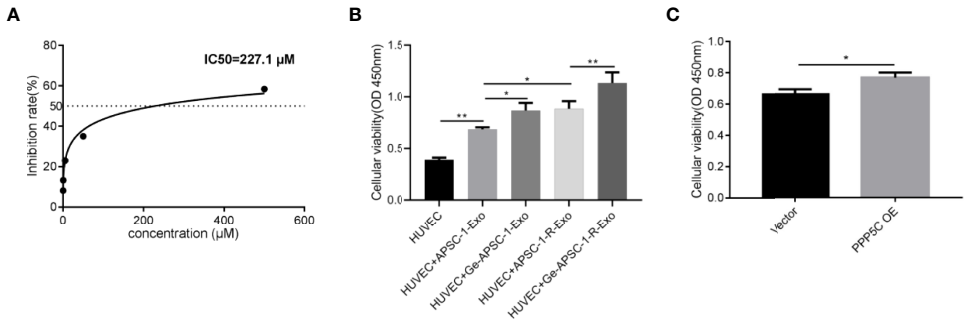



FIGURE 1
Drug resistance and cellular activity: **(A)** IC50 values of gemcitabine in ASPC-1 cells; **(B)** CCK8 detection of cell proliferation; **(C)** CCK8 detection of cell proliferation after overexpression treatment.

ASPC-1-Exo group and HUVEC+ASPC-1-R-Exo treated group; HUVEC+Ge- ASPC-1-R-Exo group further increased *in vitro* tubule formation compared with HUVEC+ASPC-1-R-Exo treatment (Figures 2A, B). And after overexpression treatment, *in vitro* tubule formation was increased after PPP5C overexpression treatment compared to the control group (Figures 2C, D).

## 3.3 Effect of cellular autophagy level

From WB experiments, it was found that LC3 II/I expression was up-regulated and p62 expression was down-regulated in the overexpression PPP5C-treated group compared with the control HUVEC cells, and LC3 II/I expression was up-regulated and p62 expression was down-regulated in the overexpression PPP5C-

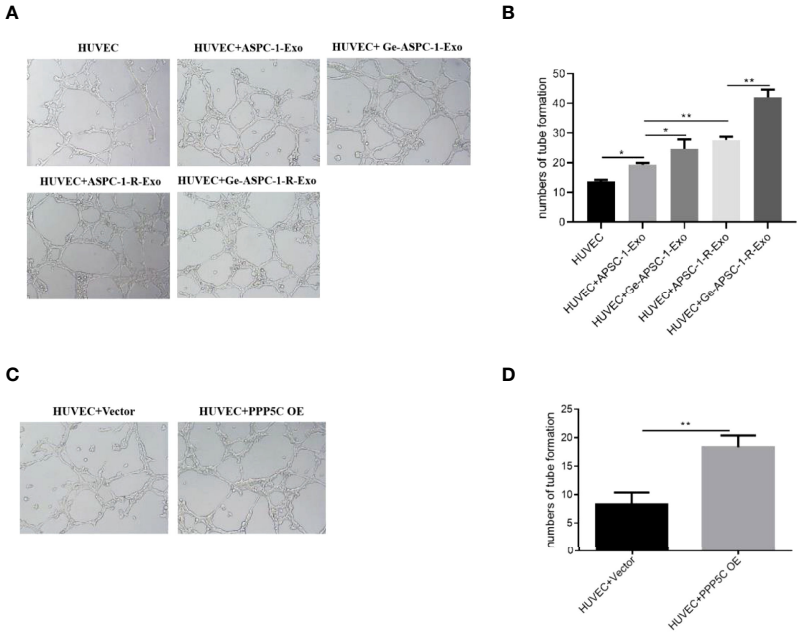



FIGURE 2
Tubule formation capacity assay **(A)** *in vitro* tubule formation capacity assay for each group (experimental plots); **(B)** *in vitro* tubule formation capacity assay for each group (bar graph); **(C)** *in vitro* tubule formation capacity assay for each group after overexpression treatment (experimental plots); **(D)** *in vitro* tubule formation capacity assay for each group after overexpression treatment (bar graph). Note: *p < 0.05, **p < 0.01.

treated group compared with the control APSC-1 cells (Figure 3A). From the immunofluorescence autophagy level, it was found that the autophagic puncta were increased in the overexpression PPP5C-treated group compared with the control HUVEC cells, and the autophagic puncta were increased in the overexpression PPP5C-treated group compared with the control APSC-1 cells (Figure 3B).

## 3.4 Gemcitabine treatment increased the expression of PPP5C

After WB as well as qPCR assays, it was found that only PP5C (i.e., PPP5C) expression was increased in the gemcitabine-treated group and in the resistant cell line compared with the control ASPC-1 cells, and PP5C was further upregulated in the resistant cell line treated with gemcitabine, while mRNA expression was also upregulated (Figures 4A, B). Exosomes CD9, CD63 and CD81 were significantly expressed in all groups (Figure 4C). Among the exosomes, the results showed that the particle size of the highest concentration in the supernatant of the ASPC-1 group was 76 nm, the average particle size of the cellular exosomes was 124.2 nm, and the concentration was 1.67×109 particles/mL; the particle size of the highest concentration in the supernatant of the ASPC-1+ Gemcitabine group was 64 nm, the average particle size of the cellular exosomes was 118.1 nm, and the concentration was 1.67×109 particles/mL. The highest particle size in the supernatant of ASPC-1-R cells was 50 nm, the average particle size of exosomes was 91.1 nm, and the concentration was 1.42×109 particles/mL; the highest particle size in the supernatant of ASPC-1-R+ Gemcitabine group was 69 nm, and the concentration of exosomes was 1.42×109 particles/mL. The highest particle size was 69 nm, and the average particle size of cellular exosomes was 93.6 nm and the concentration was 1.53×109 particles/mL (Figure 4D). After exosome treatment, PPP5C expression increased in the ASPC-1-Exo-treated group compared with the control ASPC-1 cells; PPP5C expression further increased in the Ge-ASPC-1-Exo-treated and ASPC-1-R-Exo-treated groups compared with the ASPC-1-Exo-treated group; PPP5C expression increased in the Ge-ASPC-1-R-Exo-treated group compared with the ASPC-1-Exo-treated group; PPP5C expression increased in the ASPC-1-Exo-treated group compared with the ASPC-1-Exo-treated group. PPP5C was further increased in the ASPC-1-R-Exo treated group compared to the ASPC-1-R-Exo treated group (Figure 2E). After overexpression treatment, PPP5C expression was significantly increased in HUVEC cells transfected with PPP5C overexpression vector compared with the control group (Figure 2F).

## 3.5 miR-520a targeted PPP5C inhibition

Compared with control ASPC-1 cells and their exosomes, hsa-miR-520a-5p expression was decreased both in gemcitabine-administered treated and resistant cells and in treated exosomes (Figures 5A, B). Dual luciferase assays showed that hsa-miR-520a-5p was targeted to PPP5C (Figure 5C). In contrast, among HUVEC cells, hsa-miR-520a-5p expression was significantly upregulated after hsa-miR-520a-5p mimics treatment compared to control HUVEC, while PPP5C expression was significantly decreased, while compared to the exosome-treated group of resistant cells, exosome treatment after drug administration caused hsa-miR-520a-5p expression decreased. (Figures 5D, E).

## 3.6 Elevated PPP5C expression levels make tumor patients have worse prognosis

After analysis of the TCGA database, it was found that the higher the expression level of PPP5C among tumors, the worse

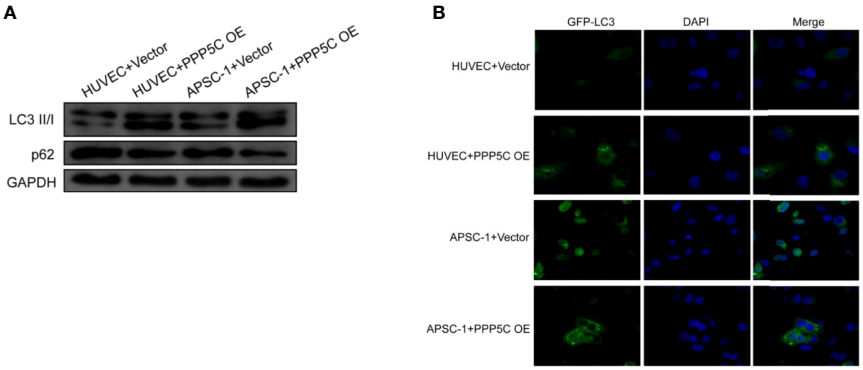



FIGURE 3
Autophagy levels: **(A)** WB autophagy level assay in each group of cells; **(B)** immunofluorescence autophagy level assay in each group of cells. .

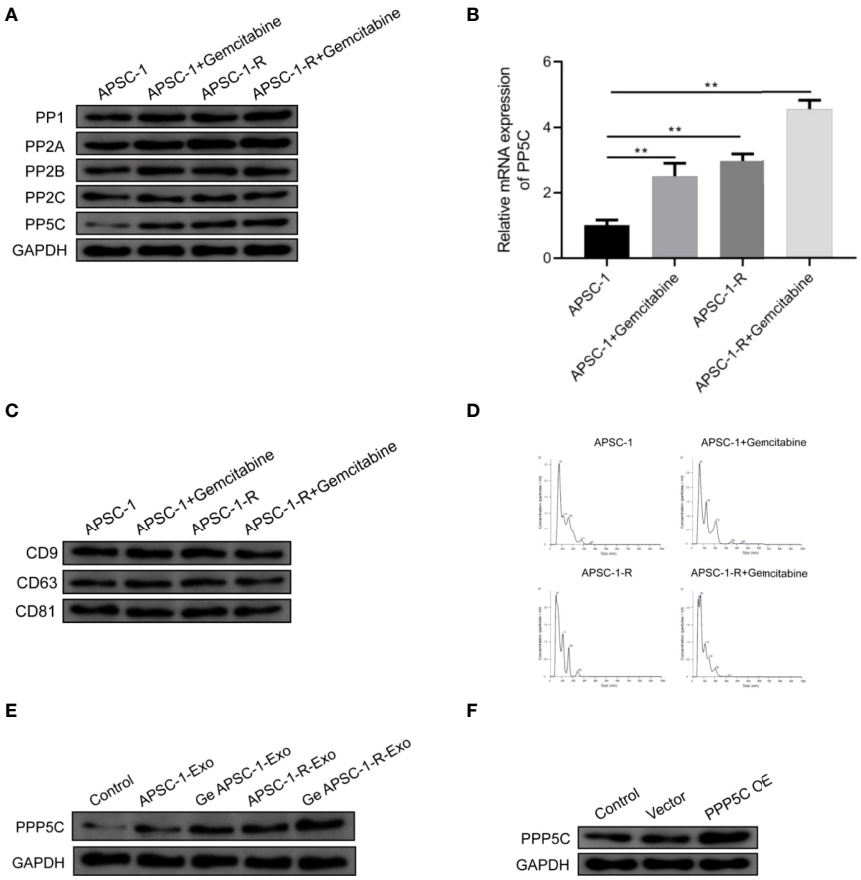



FIGURE 4
Expression of PPP5C as well as exosomal proteins: **(A)** serine/threonine protein phosphatase expression assay; **(B)** PP5C mRNA content in each group of ASPC-1 cells; **(C)** exosome markers in each group; **(D)** exosome diameter distribution in each group of cells; **(E)** PPP5C expression in each group of ASPC-1 exosomes; (F) PPP5C in each group after overexpression.

the prognosis of tumor patients (Figures 6A, B), and the ROC curve analysis revealed that PPP5C could be used more accurately to predict the survival rate of patients within 3-5 years (Figure 6C).

## 4 Discussion

In recent years, research on the molecular target therapy of tumors has gradually come to the forefront. In contrast, PPP5C mentioned in this study is a cancer-promoting factor in many cancers (20, 21). It has been less studied in pancreatic cancer and is only known to make tumors resistant to gemcitabine (16). In this section, we would analyze our results through the lens of relevant literature in order to find the immune-related genes of colon cancer, establish a prognostic risk score model, and verify its accuracy.

Members of the Ser/Thr protein phosphatase gene family, such as PPP1c, PPP2c, and PPP3c, play an important role in various cancers (22). If PPP5C is dephosphorylated at the P site, its highly conserved catalytic core and bimetallic system (M1/M2) will change to substrate-binding and hydrolysis sites. This will lead to the overexpression of PPP5C, which is one of the reasons for cell proliferation and the progression of various cancers (23). As an essential regulator of hormone and stress-related signal transduction (24), inhibition of its activity or expression will cause cell cycle arrest, hinder mitosis, and eventually lead to apoptosis (25). In this study, through preliminary analysis of the TCGA database, we found no significant difference in the expression of PPP5C between pancreatic cancer and normal tissues. However, the prognosis analysis showed that the higher the PPP5C level, the lower the five-year survival rate and the worse the patient's prognosis. So, how to study the role played by PPP5C in pancreatic cancer? In a study on tumor brain metastasis, it was found that biomarker discovery was facilitated by raw letter analysis of available case information combined with clinical trials of tumor tissue (26). Another study on the expression characteristics and clinical significance of ubiquitin-specific proteases (USPs) in

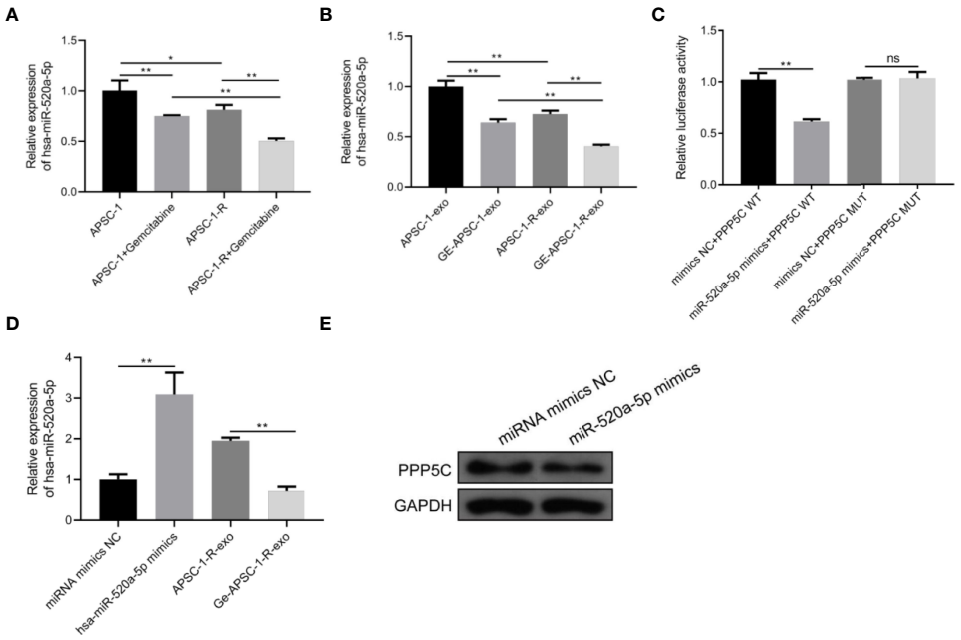



FIGURE 5
Mechanism of PPP5C expression imbalance: **(A)** hsa-miR-520a-5p expression in each group of ASPC-1 cells; **(B)** hsa-miR-520a-5p expression in each group of ASPC-1 exosomes; **(C)** luciferase activity assay; **(D)** hsa-miR-520a-5p overexpression versus hsa-miR-520a-5p in the exosomes of drug-resistant cells. miR-520a-5p expression assay; **(E)** expression level of PPP5C after hsa-miR-520a-5p overexpression. Note: * indicates $p < 0.05$, *$p < 0.05$, **$p < 0.01$.

hepatocellular carcinoma (HCC) identified USPs as one of the eight signature genes in hepatocellular carcinoma and a potential molecular target for HCC development and progression by using raw signal analysis combined with WB and immunohistochemistry (27). With reference to the methodology used in this study, we performed a related experiment on PPP5C. In the experiments related to AsPC-1 cells, it was found that the protein and mRNA of PP5c in the gemcitabine treatment group and drug-resistant cell lines were upregulated, while in the control AsPC-1 cells and in the exosomes, the expression level of PPP5C was the same. After overexpression, with the increase of PPP5C expression level in the cell experiment, the cell activity of PPP5C increased, and the increment speed accelerated. In the tumor microenvironment, exosomes, as well as other extracellular vesicles and cytokines, act as intercellular communication agents in various biological activities that contribute to drug resistance in cancer cells and are potential targets and candidate biomarkers of drug efficacy for reversing chemoresistance in HCC patients in many cancers (28). According to Zhu etal. (16), the inhibition of PPP5C increases the expression of related apoptotic markers, indicating the apoptosis of corresponding pancreatic cancer cells. This may enhance the sensitivity of PC cells to gemcitabine. This is similar to this study.

We also conducted an experiment on the effect of PPP5C on the angiogenesis of cancer cells and autophagy. The results show that the increased expression of PPP5C promotes angiogenesis and autophagy. After the tumor develops to an advanced stage, autophagy, as a dynamic degradation and circulatory process, contributes to the survival and growth of established tumors and can effectively promote metastasis and cancer invasiveness (29). Combined with our results on autophagy, we can conclude that PPP5C can promote the metastasis of pancreatic cancer cells. Furthermore, our angiogenesis experiments suggest that PPP5C can promote further deterioration of the tumor. This is sufficient to confirm the prognostic analysis results that the higher the level of PPP5C, the lower the survival rate within five years and the worse the prognosis of patients. In this connection, our study effectively proves that the role of PPP5C in pancreatic cancer is similar to that in other cancers.

At the same time, compared with previous studies on PPP5C and pancreatic cancer, this study examined the expression of miR-520a-5p in pancreatic cancer and its relationship with PPP5C. Although miR-520a-5p exists as a pathogenic factor in psoriasis and cardiomyocyte injury, studies on various human cancers, such as colorectal cancer and hepatocellular carcinoma, have found that miR-520a-5p inhibits tumor development (30). In this context, let us look at the results of the miRNA experiment in this study. Although the databases we used do not support the relationship between miR-520a-5p and PPP5C, there is a lack of relevant literature. From the results of the double luciferase experiment in this

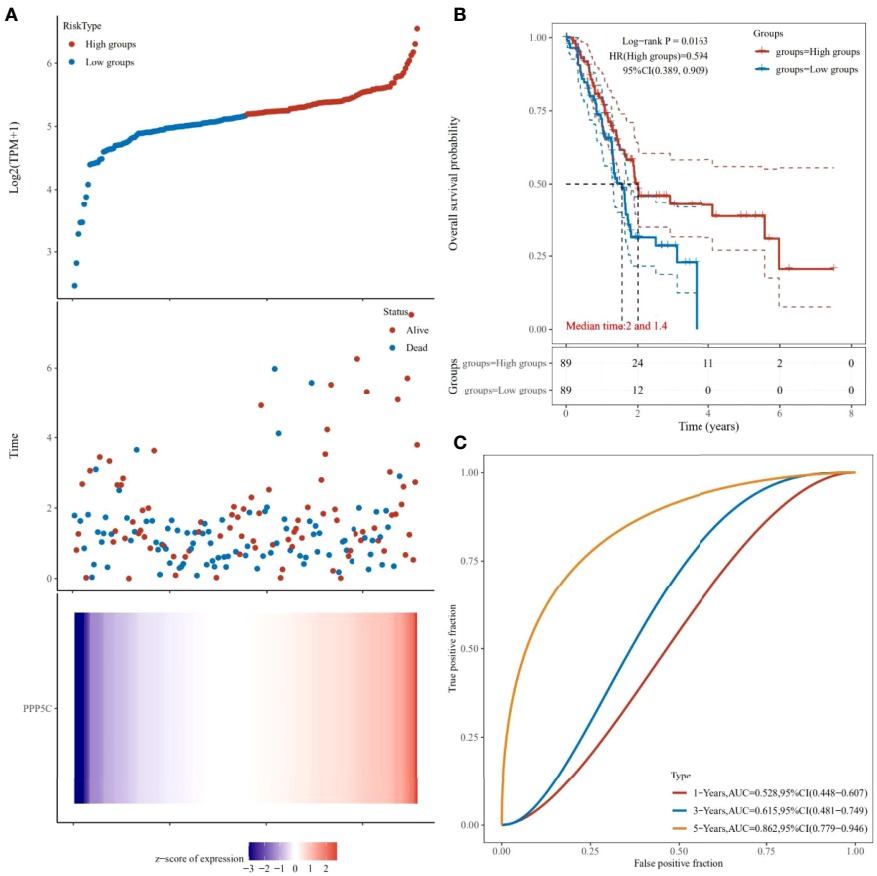



FIGURE 6
Prognostic analysis of PPP5C in pancreatic cancer patients from TCGA database: **(A)** Distribution of high and low risk samples. **(B)** Survival curves. **(C)** ROC curves.

study, miR-520a had a targeted inhibitory effect on PPP5C after overexpression. At the same time, we also found that its expression decreased in drug-resistant strains. From these results, we can conclude that miR-520a-5p can inhibit the growth of pancreatic cancer by targeting PPP5C inhibition.

In conclusion, this study explored the role of PPP5C in pancreatic cancer by combining bioinformatics analysis and experiments and found the target relationship between miR-520-5p and PPP5C. This provides a new molecular mechanism for pancreatic cancer research and offers new directions for targeted gene 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 author/s.

## Author contributions

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

## Funding

This work was supported by Zhejiang Province Bureau of Health under grant 2020366835 and Funds of Science Technology Department of Zhejiang Province under grant 2020C03074.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Zhao DD, Zhao X, Li WT. Identification of differentially expressed metastatic genes and their signatures to predict the overall survival of uveal melanoma patients by bioinformatics analysis. *Int J Ophthalmol* (2020) 13(7):1046. doi: 10.18240/ijo.2020.07.05

2. Shields CL, Kaliki S, Shah SU, Luo W, Furuta M, Shields JA. Iris melanoma: features and prognosis in 317 children and adults. *J Am Assoc Pediatr Ophthalmol Strabismus* (2012) 16(1):10–6. doi: 10.1016/j.jaapos.2011.10.012

3. Henderson E, Margo CE. Iris melanoma. *Arch Pathol Lab Med* (2008) 132 (2):268–72. doi: 10.5858/2008-132-268-IM

4. Nickla DL, Wallman J. The multifunctional choroid. *Prog Retin Eye Res* (2010) 29(2):144–68. doi: 10.1016/j.preteyeres.2009.12.002

5. Gajdzis M, Kaczmarek R, Gajdzis P. Novel prognostic immunohistochemical markers in uveal melanoma-literature review. *Cancers* (2021) 13(16):4031. doi: 10.3390/cancers13164031

6. Singh AD, Turell ME, Topham AK. Uveal melanoma: trends in incidence, treatment, and survival. *Ophthalmology* (2011) 118(9):1881–5. doi: 10.1016/j.ophtha.2011.01.040

7. Shain AH, Bagger MM, Yu R, Chang D, Liu S, Vemula S, et al. The genetic evolution of metastatic uveal melanoma. *Nat Genet* (2019) 51(7):1123–30. doi: 10.1038/s41588-019-0440-9

8. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol Res* (2019) 7 (5):737–50. doi: 10.1158/2326-6066.CIR-18-0436

9. Liu F, Yang Z, Zheng L, Shao W, Cui X, Wang Y, et al. A tumor progression related 7-gene signature indicates prognosis and tumor immune characteristics of gastric cancer. *Front Oncol* (2021) 11:2090. doi: 10.3389/fonc.2021.690129

10. Li J, Zhou D, Qiu W, Shi Y, Yang JJ, Chen S, et al. Application of weighted gene co-expression network analysis for data from paired design. *Sci Rep* (2018) 8 (1):1–8. doi: 10.1038/s41598-017-18705-z

11. Ma C, Cui Z, Wang Y, Zhang L, Wen J, Guo H, et al. Bioinformatics analysis reveals TSPAN1 as a candidate biomarker of progression and prognosis in pancreatic cancer. *Bosnian J Basic Med Sci* (2021) 21(1):47. doi: 10.17305/bjbms.2020.5096

12. Shields CL, Furuta M, Thangappan A, Nagori S, Mashayekhi A, Lally DR, et al. Metastasis of uveal melanoma millimeter-by-millimeter in 8033 consecutive eyes. *Arch Ophthalmol* (2009) 127(8):989–98. doi: 10.1001/archophthalmol.2009.208

13. Xie J, Wu Z, Xu X, Liang G, Xu J. Screening and identification of key genes and pathways in metastatic uveal melanoma based on gene expression using bioinformatic analysis. *Medicine (Baltimore)* (2017) 99(43):e22974. doi: 10.1097/MD.0000000000022974

14. Wu MY, Lai TT, Liao WT, Li CJ. Clinicopathological and prognostic significance and molecular mechanisms governing uveal melanoma. *Ther Adv Med Oncol* (2020) 12:1758835920917566. doi: 10.1177/1758835920917566

15. Thijs JL, Herath A, de Bruin-Weller MS, Hijnen D. Multiplex platform technology and bioinformatics are essential for development of biomarkers in atopic dermatitis. *J Allergy Clin Immunol* (2017) 139(3):1065. doi: 10.1016/j.jaci.2016.08.023

16. Zhu J, Ji Y, Yu Y, Jin Y, Zhang X, Zhou J, et al. Knockdown of serine/threonine protein phosphatase 5 enhances gemcitabine sensitivity by promoting apoptosis in pancreatic cancer cells *in vitro*. *Oncol Lett* (2018) 15(6):8761–9. doi: 10.3892/ol.2018.8363

17. Wu S, Chen H, Zuo L, Jiang H, Yan H. Suppression of long noncoding RNA MALAT1 inhibits the development of uveal melanoma *via* microRNA-608-mediated inhibition of HOXC4. *Am J Physiol Cell Physiol* (2020) 318(5):C903–12. doi: 10.1152/ajpcell.00262.2019

18. Zhang L, He X, Li F, Pan H, Huang X, Wen X, et al. The miR-181 family promotes cell cycle by targeting CTDSPL, a phosphatase-like tumor suppressor in uveal melanoma. *J Exp Clin Cancer Res* (2018) 37(1):1–13. doi: 10.1186/s13046-018-0679-5

19. Xu X, Tao R, Sun L, Ji X. Exosome-transferred hsa_circ_0014235 promotes DDP chemoresistance and deteriorates the development of non-small cell lung cancer by mediating the miR-520a-5p/CDK4 pathway. *Cancer Cell Int* (2020) 20 (1):1–15. doi: 10.1186/s12935-020-01642-9

20. Zhi X, Zhang H, He C, Wei Y, Bian L, Li G. Serine/threonine protein phosphatase-5 accelerates cell growth and migration in human glioma. *Cell Mol Neurobiol* (2015) 35(5):669–77. doi: 10.1007/s10571-015-0162-1

21. Zheng X, Zhang L, Jin B, Zhang F, Zhang D, Cui L. Knockdown of protein phosphatase 5 inhibits ovarian cancer growth *in vitro*. *Oncol Lett* (2016) 11(1):168–72. doi: 10.3892/ol.2015.3828

22. Von Kriegsheim A, Pitt A, Grindlay GJ, Kolch W, Dhillon AS. Regulation of the raf–MEK–ERK pathway by protein phosphatase 5. *Nat Cell Biol* (2006) 8 (9):1011–6. doi: 10.1038/ncb1465

23. Salter EA, Wierzbicki A, Honkanen RE. Quantum-based modeling of dephosphorylation in the catalytic site of serine/threonine protein phosphatase-5 (PPP5C). *Catalysts* (2020) 10(6):674. doi: 10.3390/catal10060674

24. Xie J, Han M, Zhang M, Deng H, Wu W. PP5 (PPP5C) is a phosphatase of Dvl2. *Sci Rep* (2018) 8(1):1–15. doi: 10.1038/s41598-018-21124-3

25. D'Arcy BM, Prakash A, Honkanen RE. Targeting phosphatases in cancer: suppression of many versus the ablation of one. *Oncotarget* (2019) 10(61):6543.

26. Su J, Song Q, Qasem S, O'Neill S, Lee J, Furdui CM, et al. Multi-omics analysis of brain metastasis outcomes following craniotomy. *Front Oncol* (2021) 10:615472. doi: 10.3389/fonc.2020.615472

27. Ni W, Bian S, Zhu M, Song Q, Zhang J, Xiao M, et al. Identification and validation of ubiquitin-specific proteases as a novel prognostic signature for hepatocellular carcinoma. *Front Oncol* (2021) 11. doi: 10.3389/fonc.2021.629327

28. Zhang J, Song Q, Wu M, Zheng W. The emerging roles of exosomes in the chemoresistance of hepatocellular carcinoma. *Curr Med Chem* (2021) 28(1):93–109.

29. Li X, He S, Ma B. Autophagy and autophagy-related proteins in cancer. *Mol Cancer* (2020) 19(1):1–16. doi: 10.1186/s12943-020-1138-4

30. Jin M, Shi C, Hua Q, Li T, Yang C, Wu Y, et al. High circ-SEC31A expression predicts unfavorable prognoses in non-small cell lung cancer by regulating the miR-520a-5p/GOT-2 axis. *Aging (albany NY)* (2020) 12 (11):10381. doi: 10.18632/aging.103264

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 26 July 2022
DOI 10.3389/fonc.2022.965088

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Zilei Chen,
Osnabrück University, Germany
Qihong Zhao,
Anhui Medical University, China

*CORRESPONDENCE
Xu Lu
jackkrissheng123@163.com

†These authors share first authorship

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

RECEIVED 09 June 2022
ACCEPTED 04 July 2022
PUBLISHED 26 July 2022

CITATION
Lu X, Chen X, Wang X, Qing J, Li J and Pan Y (2022) Construction of lncRNA and mRNA co-expression network associated with nasopharyngeal carcinoma progression.
*Front. Oncol.* 12:965088.
doi: 10.3389/fonc.2022.965088

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

# Construction of lncRNA and mRNA co-expression network associated with nasopharyngeal carcinoma progression

Xu Lu[1*†], Xing Chen[1†], Xinke Wang[2], Jing Qing[1], Ji Li[1] and Yunyun Pan[1]

[1]Ningbo First Hospital, Ningbo, China, [2]Ninghai County Third Hospital, Ningbo, China

Nasopharyngeal carcinoma is a type of head and neck cancer with a high incidence in men. In the past decades, the survival rate of NPC has remained around 70%, but it often leads to treatment failure due to its distant metastasis or recurrence. The lncRNA-mRNA regulatory network has not been fully elucidated. We downloaded the NPC-related gene expression datasets GSE53819 and GSE12452 from the Gene Expression Omnibus database; GSE53819 included 18 NPC tissues and 18 normal tissues, and GSE12452 included 31 NPC tissues and 10 normal tissues. Weighted gene co-expression network analysis was performed on mRNA and lncRNA to screen out modules that were highly correlated with tumor progression. The two datasets were subjected to differential analysis after removing batch effects, and then Venn diagrams were used to screen for overlapping genes in the module genes and differential genes. The lncRNA-mRNA co-expression network was then constructed, and key mRNAs were identified by MCODE analysis and expression analysis. GSEA analysis and qRT-PCR were performed on key mRNAs. Through a series of analyses, we speculated that BTK, CD72, PTPN6, and VAV1 may be independent predictors of the prognosis of NPC patients.Taken together, our study provides potential candidate biomarkers for NPC diagnosis, prognosis, or precise treatment.

KEYWORDS

**lncRNA, nasopharyngeal carcinoma, SMAD5-AS1, co-expression network, WGCNA**

## Introduction

Nasopharyngeal carcinoma is a type of head and neck cancer (1). NPC is highly metastatic, often metastasizing to local and distant lymph nodes, bone, lung, and liver (2, 3). NPC is relatively rare in North America and Europe but is commonly found in many places (4). NPC has significant gender differences, with men more likely to develop NPC

(5). In addition to genetic susceptibility and environmental factors, NPC is caused by genetic and epigenetic alterations, as well as dense lymphatic infiltration of the primary tumor (6). With advances in radiation therapy techniques in recent decades, NPC has maintained a survival rate of about 70% with radiation therapy-based combination therapy, but tumor recurrence or distant metastases occur within a few years after treatment, leading to treatment failure (7). It is necessary to study the underlying molecular mechanisms of NPC occurrence to provide better biomarkers for the diagnosis and treatment of NPC. Recent researched have identified a variety of lncRNAs that are closely associated with various processes of tumorigenesis and progression (8, 9).

LncRNAs are long-stranded noncoding RNAs that play key roles in various cellular and physiological processes and are aberrantly expressed in various cancers (10). MEG3 overexpression promotes BCa cell apoptosis and inhibits cell proliferation (11). lncRNA ANRIL and lncRNA n375709 have also been shown to play a role in NPC by Zou et al. and Ren et al. (12, 13). However, our understanding of cancer-related lncRNAs is still very limited. One of the basic molecular mechanisms of lncRNAs is as competing endogenous RNAs, which are mutually regulated with microRNAs (miRNAs) and jointly participate in the expression of target gene mRNAs, playing this important role in tumor development. Therefore, cancer-critical functional lncRNAs can be identified by lncRNA-induced transcriptional disruption of target gene mRNAs (14, 15). The lncRNA-mRNA regulatory network related to NPC progression has been rarely reported due to the lack of common analysis of lncRNA and mRNA expression levels in NPC.

In this study, based on the NPC-related gene expression datasets GSE53819 and GSE12452, we performed weighted gene co-expression network analysis and differential analysis on the samples from them. The lncRNA-mRNA co-expression network related to NPC progression was constructed to explain the roles of NPC-associated mRNAs and lncRNAs. This finding gives potential biomarkers for NPC diagnosis, prognosis, or precise treatment.

# Methodology

## Data sources

We used the keyword "nasopharyngeal carcinoma" to find relevant datasets in the GEO database, and then manually reviewed and selected cohorts containing lncRNA and mRNA expression and clinical survival information, including 18 NPC tissues and 18 normal tissues, and the GSE12452 dataset (platform GPL570, Human mRNA) from the University of Wisconsin-Madison, including 31 NPC tissues and 10 normal tissues. Both datasets contain lncRNA and mRNA genes.

## WGCNA

Pearson correlation coefficients between genes were obtained from the differentially expressed mRNAs and lncRNAs used in WGCNA using the WGCNA database of R software; subsequently, an appropriate soft threshold β was chosen to ensure that the network was not scalable. A gene network is built to transform the adjacency matrix into a topological overlap matrix to generate a hierarchical clustering tree of genes. Highly correlated co-expressed gene modules were identified using the DynamicTreeCut method with thresholds set to cut height = 0.25 and minSize = 150. Pearson correlation test can analyze the relationship between module eigen genes and clinical features.

## Differential expression analysis

The GSE53819 dataset and GSE12452 dataset were subjected to batch effect removal using the R package, and the final matrix with batch effects removed was obtained. The differential expression of gene in this matrix was then conducted using the limma package of R software. "adj. $P<0.05$ and $FC=0.07$" were defined as the thresholds for lncRNA and mRNA differential expression screening, and heat maps and volcano maps were plotted using the pheatmap and ggplot2 packages, respectively.

## Venn analysis

The modules and differential genes screened by WGCNA were screened for overlapping genes using online calculations and plotting custom Venn diagrams.

## Protein-protein interaction network construction

Protein-protein interaction network construction was performed using the Metascape tool, and the MCODE algorithm was used to cluster the PPI network, identify potential protein complexes, and screen for hub genes in the PPI network. The mRNAs in the largest sub-networks were screened as key mRNAs. lncRNA and hub mRNA co-expression networks were constructed in Cytoscape software.

## GSEA

To find the effect of gene expression on NPC, we divided the samples from the combined expression profile into high and low expression parts and used GSEA analysis to obtain the key gene-related pathways. P-value < 0.05, FDR < 0.25, NES > 1 or less than -1 were used as screening conditions.

## Cell culture

The human nasopharyngeal carcinoma cell line SUNE-1 and human nasopharyngeal carcinoma epithelial cells NP69 were purchased from Cell Bank of Type Culture Collection supplemented with 10% fetal bovine serum, 100 U/mL penicillin, and 100 μ g/mL streptomycin. NP69 cells were cultured in a keratinocyte/serum-free medium. All cell lines were cultured in a humidified incubator at 37°C, 5% $CO_2$.

## qRT-PCR

Total RNA was extracted from NP69 and SUNE-1 cells using TRIzol reagent, and lncRNA and mRNA were reverse transcribed to cDNA using FastQuant RT kit including gDNase and SuperReal Premix Plus-SYBR Green. Green to reverse transcribe lncRNA and mRNA to cDNA. real-time qPCR assays were performed using the miScript SYBR Green PCR Kit. The relative expression of the genes was calculated by the 2 -ΔΔCT method. The experiment was repeated three times and the mean was taken. gAPDH was used as an endogenous control for mRNA normalization. The primers are displayed in Table 1.

TABLE 1 qRT-PCR primers.

| Genes | Primer sequences (5'- 3') |
|---|---|
| BTK | F:CCAATGGCTGCCTCCTGAACTAC<br>R:TCGGTGAAGGAACTGCTTTGACTC |
| CD72 | F:CATCTCCAGCAGGTTAGGACA<br>R:CGGGCACTTGAACATTCTC |
| PTPN6 | F:TGCAGGGACGTGACAGTAAC<br>R:TGACACGAGTGTTCTCCTGC |
| VAV1 | F:TCAGTGCGTGAACGAGGTCAAG<br>R:CCATAGTGAGCCAGAGACTGGT |
| GAPDH | F:GTCAACGGATTTGGTCTGTATT<br>R: AGTCTTCTGGGTGGCAGTGAT |

# Results

## Gene co-expression modules

To explore the co-expression patterns of mRNA and lncRNA in NPC, we performed WGCNA analysis on the GSE12452 dataset and GSE53819 dataset, respectively. To ensure scale-free networks, we chose soft thresholds of β = 6, and β = 5, respectively (Figures 1A, D), used WGCNA packages as soft threshold power to generate hierarchical clustering trees (Figures 1B, E), and then we built co-expression networks of associations between clinical features and these modules (Figures 1C, F). The blue module and the turquoise module of GSE12452 were significantly associated with tumor progression, and these two modules were defined as SUR1 modules. the brown module and the turquoise module of GSE12452 were significantly associated with tumor progression, and these two modules were defined as SUR2 modules.

## Differential expression analysis

We performed batch effect removal on the GSE12452 dataset and the GSE53819 dataset and performed differential expression analysis on the matrix with batch effect removed, and obtained a total of 1646 differential genes (Figures 2A, B). The SUR1 and SUR2 modular genes and differential genes obtained from WGCNA analysis were screened using a Venn diagram to identify 410 overlapping genes (Figure 2C).

## lncRNA-mRNA co-expression network

There were 8 lncRNAs in the genes screened by the Wayne diagram, and the remaining 402 mRNAs were their target genes. The co-expression patterns of lncRNA-mRNAs in SUR1 and SUR2 modules were analyzed to construct co-expression networks, and 2203 co-expression relationships were obtained (Figure 3A). Among these 8 lncRNAs SMAD5-AS1 was significantly associated with the progression of NPC, and therefore SMAD5-AS1 was considered a key lncRNA. SMAD5-AS1 and 106 co-expressed mRNAs were got by degree screening (Figure 3B). The 106 mRNAs were analyzed using the Metascape tool to obtain PPI networks for further visualization of gene information and network construction (Figure 3C). The MCODE algorithm in Metascape clustered the PPI networks and screened the pivotal genes in the PPI networks (Figure 3D). Six mRNAs in the largest sub-network were regarded as key mRNAs. The co-expression network of SMAD5-AS1 and six hub mRNAs is displayed in (Figure 3E).

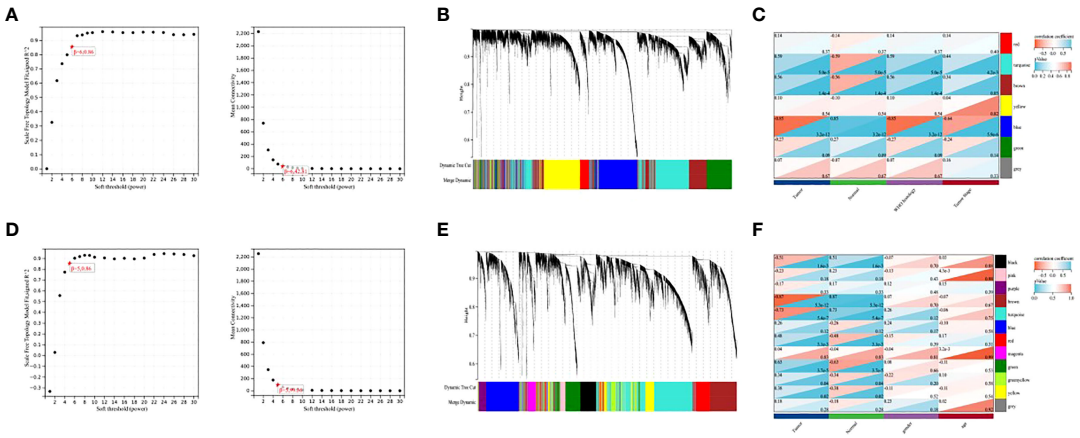



FIGURE 1
Identification of gene co-expression modules. **(A, D)** Network topologies with different soft threshold power; **(B, E)** Gene tree graph obtained based on overlapping clustering of the same topologies; **(C, F)** Correlation of each module with clinical information.

## Expression analysis

To verify whether the screened hub genes were significantly related to NPC, the expression of genes in NPC tissues and normal tissues were analyzed, and the expression of BTK, CD72, PTPN6, and VAV1 was highly significant in cancer and normal tissues, and the expression in cancer tissues was lower than that in normal tissues (Figure 4).

## GSEA

To detect the effect of genes on NPC, we classified the samples into two parts of high and low expression according to the expression of BTK, CD72, PTPN6, and VAV1, and analyzed the samples by GSEA. The top two most abundant signaling pathways or biological processes were listed according to the scores.GSEA confirmed that BTK and CD72 were mainly enriched in the IGA-producing intestinal immune network, B-cell receptor signaling pathway; PTPN6 and VAV1 were mainly associated with primary immunodeficiency, the IGA-producing intestinal immune network (Figure 5).

## qRT-PCR

We examined the expression of BTK, CD72, PTPN6, and VAV1 in NPC cell lines by qRT-PCR, and the expression levels of BTK, CD72, PTPN6, and VAV11 were lower in human nasopharyngeal carcinoma SUNE-1 cells compared with human nasopharyngeal epithelial cells NP69 ($P<0.05$) (Figure 6).

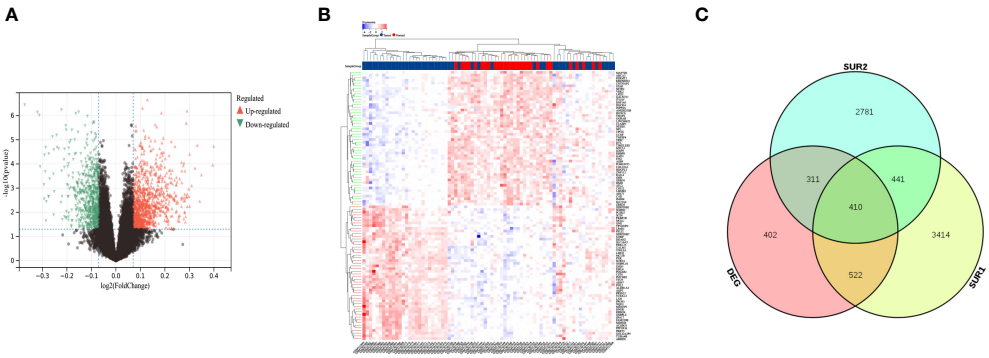



FIGURE 2
Identification of differentially expressed mRNAs and lncRNAs. **(A)** Volcano map showing differentially expressed lncRNA and mRNA; **(B)** Heat map showing differentially expressed lncRNA and mRNA; **(C)** Venn diagram screening SUR1, SUR2 module genes and overlapping genes of differential genes.

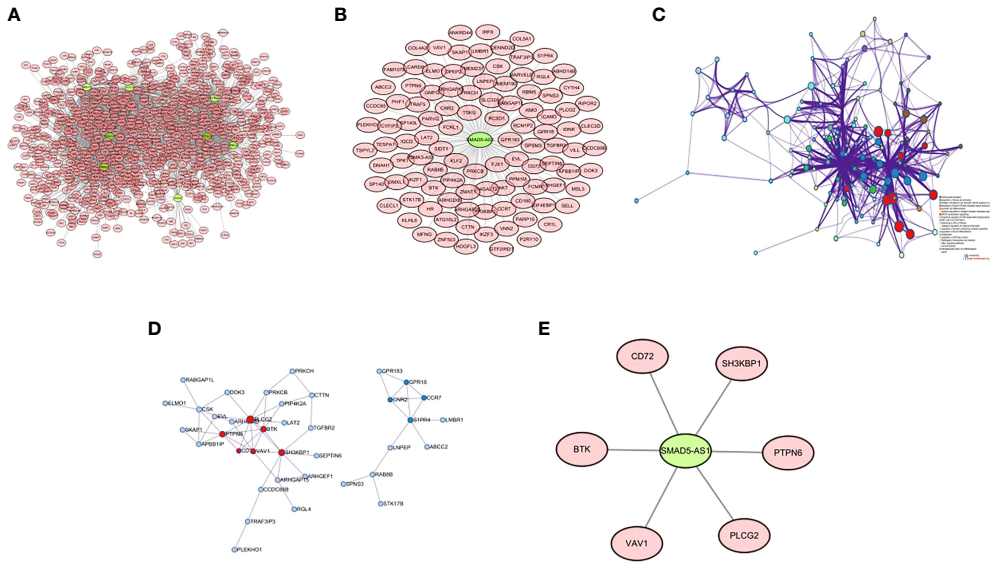



FIGURE 3
Construction of lncRNA-mRNA co-expression network and PPI network. **(A)** lncRNA-mRNA co-expression network of 8 lncRNAs in SUR1 and SUR2 modules; **(B)** key lncRNA SMAD5-AS1 and mRNA co-expression network; **(C)** PPI network of co-expressed mRNAs in the co-expression network; **(D)** MCODE plug-in screening the highest-scoring sub-network; **(E)** SMAD5-AS1 and hub mRNA co-expression networks.

## Discussion

NPC is a common cancer of the head and neck (16). Currently, NPC is mostly squamous cell carcinoma, and the treatment of choice for NPC is radiotherapy (17, 18). Despite the continuous advances in radiotherapy, its distant metastasis remains the main reason for treatment failure, and the discovery of more efficient treatment methods is an essential future research direction. Thus, it is clear that understanding the etiology and mechanisms of NPC progression is important for the prevention and treatment of NPC.

In recent years, the direction of treatment for nasopharyngeal carcinoma has trended toward targeted therapy, less toxic and more effective forms of chemotherapy, and new technologies (19).

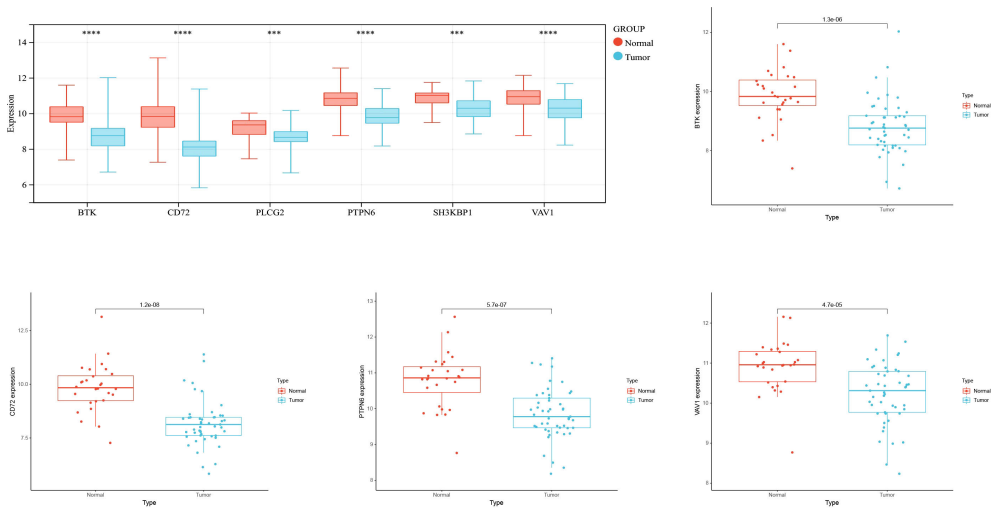



FIGURE 4
Expression analysis. ***indicates P < 0.001, ****indicates P < 0.0001.

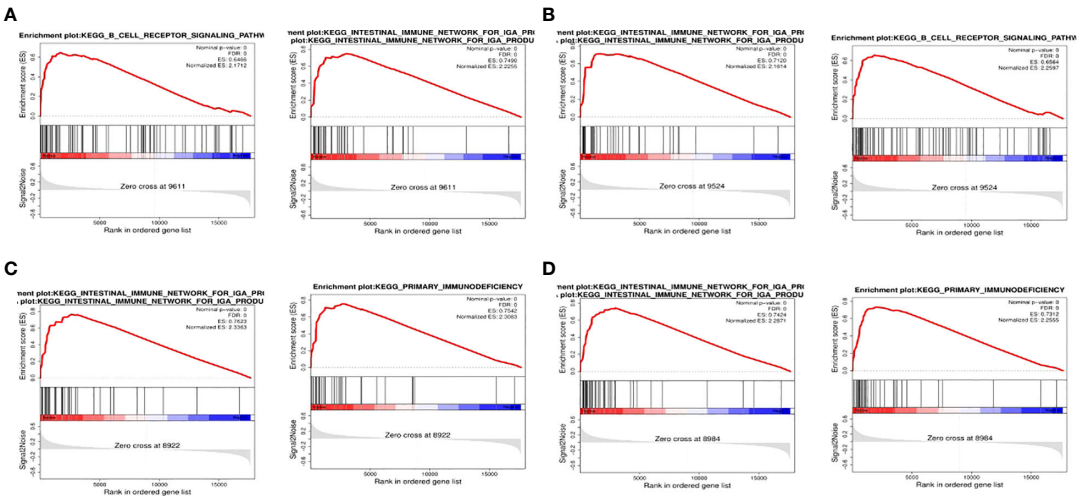



FIGURE 5
GSEA. **(A)** GSEA results of the top two correlations between BTK and signaling pathway; **(B)** GSEA results of the top two correlations between CD72 and signaling pathway; **(C)** GSEA results of the top two correlations between PTPN6 and signaling pathway; **(D)** GSEA results of the top two correlations between VAV1 and signaling pathway.

lncRNAs, as regulators of biological functions, play key roles in various cellular and physiological processes (20) and are good choices in gene therapy. However, the lack of a comprehensive database providing resources for experimental validation of lncRNA function has become the most significant challenge in lncRNA-based therapeutic modalities.

In ceRNA theory, some lncRNAs can be co-expressed with the corresponding coding genes (15). We analyzed the lncRNA-mRNA co-expression network in NPC. We used the GSE12452 dataset and GSE53819 to build co-expression networks of clinical features and inter-module associations and selected the modules most relevant to NPC occurrence. Differential analysis was performed on both datasets after removing batch effects, and then overlapping genes between module genes and differential genes were screened out. Among the eight lncRNAs screened, SMAD5-AS1 was significantly associated with the biological progression of NPC cells (21), so it was considered a key lncRNA to construct a lncRNA-mRNA co-expression network related to NPC progression. Six hub mRNAs (BTK, CD72, PTPN6, VAV1, PLCG2, SH3KBP1) were then identified by the MCODE algorithm. The expression of genes in NPC tissues and normal tissues was observed, and it was found that the expression of BTK, CD72, PTPN6, and VAV1 was different in cancer and normal tissues, and the expression was lower in both cancer and normal tissues.

Bruton's tyrosine kinase is a component of the B-cell receptor (BCR) signaling body (22). Activated BCR signaling contributes to the development of B-cell malignancies (23) and plays a role in the pathobiology of other hematologic malignancies such as chronic lymphocytic leukemia (CLL) (24). BTK, through the first-in-class inhibitor ibrutinib whose efficacy is clinically validated as a target for B-cell malignancies (25, 26). Our GSEA results also suggest that BTK is closely associated with the B-cell receptor signaling pathway. It was

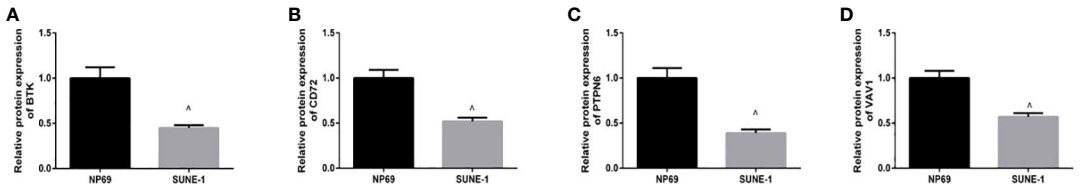



FIGURE 6
qRT-PCR. **(A)** Expression level of BTK; **(B)** expression level of CD72; **(C)** expression level of PTPN6; **(D)** expression level of VAV1. ^ indicates $P<0.05$ compared with the NP69 group.

shown that BTK may be involved in the radioresistance process of NPC cells (27). This suggests to us that BTK may be a new key biomarker for NPC. CD72 is a co-receptor of BCR and an important regulator in the pathogenesis of several immune diseases; it plays a role in various B cell biological processes, including proliferation, apoptosis, and differentiation. In patients with Systemic Lupus Erythematosus, low expression of CD72 on B cells is negatively correlated with patient disease activity (SLEDAI) (28). CD72 is lowly expressed in multiple sclerosis and its ligand CD100 is increased in T cells (29). CD72 is lowly expressed in NPC tissues, suggesting that our low CD72 expression may be related to poor prognosis in NPC patients. Protein tyrosine phosphatase non-receptor type 6 is a key regulatory protein in cell signaling that regulates cell death and inflammation (30); It has different regulatory mechanisms and effects on cell cycle and cell proliferation in different tumors (31) and upregulated in colon cancer (32). We demonstrated low expression of PTPN6 in NPC by dataset analysis and qRT-PCR analysis, suggesting that our low expression of PTPN6 may be associated with poor prognosis in NPC patients. VAV1 is frequently mutated and overexpressed in hematopoietic malignancies and various cancers (33) and is accompanied by B-cell lymphoma. It was found that there is potential crosstalk between epithelial cells of VAV1 (which secrete CSF-1) and lymphocytes expressing CSF-1R, which leads to B-cell lymphoma (34). Through this mechanism, VAV1 promotes tumor propagation. No correlation between VAV1 and NPC progression has been found in existing studies, and our findings suggest that VAV1 is lowly expressed in NPC and that VAV1 is mainly associated with pathways of primary immunodeficiency, suggesting that VAV1 may be a new prognostic independent predictor for NPC patients.

Since lncRNAs can post-transcriptionally regulate target mRNAs, it is important to predict the interaction between the two and explore potential mechanisms of the pathological process through regulatory networks. Zhang J et al. characterized the regulatory networks of lncRNAs and mRNAs in cutaneous melanoma (SKCM) by analyzing expression profile data and identified potential therapeutic targets (35) A study by Ying-Juan Zheng found that SMAD5-AS1 knockdown inhibited EMT, cell proliferation, migration, and invasion in NPC by upregulating miR-106a-5p and downregulating SMAD5 (21). In our study, a novel NPC-regulated lncRNA SMAD5-AS1 mRNA axis was identified, which provides a reference for exploring the mechanism of NPC progression.

Taken together, our comprehensive analysis provides new way into the lncRNA-mRNA co-expression network in NPC progression. It suggests that lncRNA plays an important role in NPC and that lncRNA-SMAD5-AS1 may serve as a biomarker for predicting NPC prognosis. In addition, it may mediate the biological functions of NPC cells through co-expression networks involving BTK, CD72, PTPN6, and VAV1, which in turn affect the NPC process. The results obtained in this study may provide the necessary theory for future researches on the role of lncRNAs in NPC.

## 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

We contributed equally for this work.

## Funding

This work was supported by Ningbo First Hospital (2021KY981).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Gong D, Zhu H, Zeng L, Hu R, Hu J, Ding J. Overexpression of HOXA10 promotes the growth and metastasis of nasopharyngeal carcinoma. *Exp Biol Med Maywood* (2021) 24623:2454–62. doi: 10.1177/15353702211030854

2. Dubrulle F, Souillard R, Hermans R. Extension patterns of nasopharyngeal carcinoma. *Eur Radiol* (2007) 1710:2622–30. doi: 10.1007/s00330-007-0616-z

3. Dumrongpisutikul N, Luangcharuthorn K. Imaging characteristics of nasopharyngeal carcinoma for predicting distant metastasis. *Clin Radiol* (2019) 7410:818.e9–818.e15. doi: 10.1016/j.crad.2019.06.031

4. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* (2015) 652:87–108. doi: 10.3322/caac.21262

5. Tang LL, Chen WQ, Xue WQ, He YQ, Zheng RS, Zeng YX, et al. Global trends in incidence and mortality of nasopharyngeal carcinoma. *Cancer Lett* (2016) 3741:22–30. doi: 10.1016/j.canlet.2016.01.040

6. Yin H, Qu J, Peng Q, Gan R. Molecular mechanisms of EBV-driven cell cycle progression and oncogenesis. *Med Microbiol Immunol* (2019) 2085:573–83. doi: 10.1007/s00430-018-0570-1

7. Li LN, Xiao T, Yi HM, Zheng Z, Qu JQ, Huang W, et al. MiR-125b increases nasopharyngeal carcinoma radioresistance by targeting A20/NF-κB signaling pathway. *Mol Cancer Ther* (2017) 1610:2094–106. doi: 10.1158/1535-7163.MCT-17-0385

8. Li JK, Chen C, Liu JY, Shi JZ, Liu SP, Liu B, et al. Long noncoding RNA MRCCAT1 promotes metastasis of clear cell renal cell carcinoma *via* inhibiting NPR3 and activating p38-MAPK signaling. *Mol Cancer* (2017) 161:111. doi: 10.1186/s12943-017-0681-0

9. Mondal T, Juvvuna PK, Kirkeby A, Mitra S, Kosalai ST, Traxler L, et al. Sense-antisense lncRNA pair encoded by locus 6p22.3 determines neuroblastoma susceptibility *via* the USP36-CHD7-SOX9 regulatory axis. *Cancer Cell* (2018) 333:417–434.e7. doi: 10.1016/j.ccell.2018.01.020

10. Bhan A, Mandal SS. LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. *Biochim Biophys Acta* (2015) 18561:151–64. doi: 10.1016/j.bbcan.2015.07.001

11. Zhang X, Gejman R, Mahta A, Zhong Y, Rice KA, Zhou Y, et al. Maternally expressed gene 3, an imprinted noncoding RNA gene, is associated with meningioma pathogenesis and progression. *Cancer Res* (2010) 706:2350–8. doi: 10.1158/0008-5472.CAN-09-3885

12. Zou ZW, Ma C, Medoro L, Chen L, Wang B, Gupta R, et al. LncRNA ANRIL is up-regulated in nasopharyngeal carcinoma and promotes the cancer progression *via* increasing proliferation, reprograming cell glucose metabolism and inducing side-population stem-like cancer cells. *Oncotarget* (2016) 738:61741–54. doi: 10.18632/oncotarget.11437

13. Ren S, Li G, Liu C, Cai T, Su Z, Wei M, et al. Next generation deep sequencing identified a novel lncRNA n375709 associated with paclitaxel resistance in nasopharyngeal carcinoma. *Oncol Rep* (2016) 364:1861–7. doi: 10.3892/or.2016.4981

14. Liu Y, Zhang R, Qiu F, Li K, Zhou Y, Shang D, et al. Construction of a lncRNA-PCG bipartite network and identification of cancer-related lncRNAs: a case study in prostate cancer. *Mol Biosyst* (2015) 112:384–93. doi: 10.1039/c4mb00439f

15. Wang P, Ning S, Zhang Y, Li R, Ye J, Zhao Z, et al. Identification of lncRNA-associated competing triplets reveals global patterns and prognostic markers for cancer. *Nucleic Acids Res* (2015) 437:3478–89. doi: 10.1093/nar/gkv233

16. Zhou J, Zhang B, Zhang X, Wang C, Xu Y. Identification of a 3-miRNA signature associated with the prediction of prognosis in nasopharyngeal carcinoma. *Front Oncol* (2022) 11:823603. doi: 10.3389/fonc.2021.823603

17. Guo R, Mao YP, Tang LL, Chen L, Sun Y, Ma J. The evolution of nasopharyngeal carcinoma staging. *Br J Radiol* (2019) 921102:20190244. doi: 10.1259/bjr.20190244

18. Jiang W, Liu N, Chen XZ, Sun Y, Li B, Ren XY, et al. Genome-wide identification of a methylation gene panel as a prognostic biomarker in nasopharyngeal carcinoma. *Mol Cancer Ther* (2015) 1412:2864–73. doi: 10.1158/1535-7163.MCT-15-0260

19. Liu T, Li J, Wu X, Zhang S, Lu Z, Li G, et al. Transferrin-targeting redox hyperbranched polyamido amine.-functionalized graphene oxide for sensitized chemotherapy combined with gene therapy to nasopharyngeal carcinoma. *Drug Deliv* (2019) 261:744–55. doi: 10.1080/10717544.2019.1642421

20. Fan Y, Shen B, Tan M, Mu X, Qin Y, Zhang F, et al. Long non-coding RNA UCA1 increases chemoresistance of bladder cancer cells by regulating wnt signaling. *FEBS J* (2014) 2817:1750–8. doi: 10.1111/febs.12737

21. Zheng YJ, Zhao JY, Liang TS, Wang P, Wang J, Yang DK, et al. Long noncoding RNA SMAD5-AS1 acts as a microRNA-106a-5p sponge to promote epithelial mesenchymal transition in nasopharyngeal carcinoma. *FASEB J* (2019) 3311:12915–28. doi: 10.1096/fj.201900803R

22. Niemann CU, Wiestner A. B-cell receptor signaling as a driver of lymphoma development and evolution. *Semin Cancer Biol* (2013) 236:410–21. doi: 10.1016/j.semcancer.2013.09.001

23. Honigberg LA, Smith AM, Sirisawad M, Verner E, Loury D, Chang B, et al. The bruton tyrosine kinase inhibitor PCI-32765 blocks b-cell activation and is efficacious in models of autoimmune disease and b-cell malignancy. *Proc Natl Acad Sci U S A* (2010) 10729:13075–80. doi: 10.1073/pnas.1004594107

24. Herman SE, Gordon AL, Hertlein E, Ramanunni A, Zhang X, Jaglowski S, et al. Bruton tyrosine kinase represents a promising therapeutic target for treatment of chronic lymphocytic leukemia and is effectively targeted by PCI-32765. *Blood* (2011) 11723:6287–96. doi: 10.1182/blood-2011-01-328484

25. Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, Blum KA, et al. Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N Engl J Med* (2013) 3691:32–42. doi: 10.1056/NEJMoa1215637

26. Wang ML, Rule S, Martin P, Goy A, Auer R, Kahl BS, et al. Targeting BTK with ibrutinib in relapsed or refractory mantle-cell lymphoma. *N Engl J Med* (2013) 3696:507–16. doi: 10.1056/NEJMoa1306220

27. Shen Y, Ma Y, Xie J, Lin L, Shi Y, Li X, et al. A regulatory role for CD72 expression on b cells and increased soluble CD72 in primary sjogren's syndrome. *BMC Immunol* (2020) 211:21. doi: 10.1186/s12865-020-00351-2

28. Vadasz Z, Haj T, Balbir A, Peri R, Rosner I, Slobodin G, et al. A regulatory role for CD72 expression on b cells in systemic lupus erythematosus. *Semin Arthritis Rheumatol* (2014) 436:767–71. doi: 10.1016/j.semarthrit.2013.11.010

29. Kuklina EM, Baidina TV, Danchenko IY, Nekrasova IV. Semaforin Sema4D in the immune system in multiple sclerosis. *Bull Exp Biol Med* (2014) 1572:234–7. doi: 10.1007/s10517-014-2533-x

30. Kiratikanon S, Chattipakorn SC, Chattipakorn N, Kumfu S. The regulatory effects of PTPN6 on inflammatory process: Reports from mice to men. *Arch Biochem Biophys* (2022) 721:109189. doi: 10.1016/j.abb.2022.109189

31. Wen LZ, Ding K, Wang ZR, Ding CH, Lei SJ, Liu JP, et al. SHP-1 acts as a tumor suppressor in hepatocarcinogenesis and HCC progression. *Cancer Res* (2018) 7816:4680–91. doi: 10.1158/0008-5472.CAN-17-3896

32. Liu G, Zhang Y, Huang Y, Yuan X, Cao Z, Zhao Z. PTPN6-EGFR protein complex: A novel target for colon cancer metastasis. *J Oncol* (2022) 2022:7391069. doi: 10.1155/2022/7391069

33. Prieto-Sánchez RM, Hernández JA, García JL, Gutiérrez NC, San Miguel J, Bustelo XR, et al. Overexpression of the VAV proto-oncogene product is associated with b-cell chronic lymphocytic leukaemia displaying loss on 13q. *Br J Haematol* (2006) 1336:642–5. doi: 10.1111/j.1365-2141.2006.06094.x

34. Shalom B, Farago M, Salaymeh Y, Sebban S, Pikarsky E, Katzav S. Vav1 promotes b-cell lymphoma development. *Cells* (2022) 116:949. doi: 10.3390/cells11060949

35. Zhang J, Liu H, Zhang W, Li Y, Fan Z, Jiang H, et al. Identification of lncRNA-mRNA regulatory module to explore the pathogenesis and prognosis of melanoma. *Front Cell Dev Biol* (2020) 8:615671. doi: 10.3389/fcell.2020.615671

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 02 August 2022
DOI 10.3389/fonc.2022.972744

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Jia Shi,
Wuxi Maternity and Child Health Care Hospital, China
JinKu Zhang,
Hebei University, China
Shaocong Mo,
Fudan University, China

*CORRESPONDENCE
Min-Bin Chen
cmb1981@163.com
Yi-Ping Du
duyp1989@126.com

†These authors have contributed equally to this work

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

RECEIVED 19 June 2022
ACCEPTED 05 July 2022
PUBLISHED 02 August 2022

CITATION
Zhu X-R, Zhu J-Q, Chen Y-F, Liu Y-Y, Lu J-J, Sun J, Peng S-Q, Chen M-B and Du Y-P (2022) Bioinformatics analysis and experimental verification of the prognostic and biological significance mediated by fatty acid metabolism related genes for hepatocellular carcinoma. *Front. Oncol.* 12:972744. doi: 10.3389/fonc.2022.972744

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

# Bioinformatics analysis and experimental verification of the prognostic and biological significance mediated by fatty acid metabolism related genes for hepatocellular carcinoma

Xiao-Ren Zhu[1,2†], Jia-Qi Zhu[3,4†], Yu-Fei Chen[5†], Yuan-Yuan Liu[6], Jing-Jing Lu[1], Jun Sun[2], Shi-Qing Peng[1,2], Min-Bin Chen[1*] and Yi-Ping Du[1*]

[1]Department of Radiotherapy and Oncology, Affiliated Kunshan Hospital of Jiangsu University, Kunshan, China, [2]Department of Medical Oncology, Affiliated Kunshan Hospital of Jiangsu University, Medical School of Jiangsu University, Kunshan, China, [3]Department of Thoracic Surgery, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong, China, [4]Nantong Key Laboratory of Translational Medicine in Cardiothoracic Diseases and Research Institution of Translational Medicine in Cardiothoracic Diseases, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong, China, [5]Department of Hepatology, Infectious Diseases Hospital Affiliated with Soochow University, Suzhou, China, [6]Clinical Research and Lab Center, Affiliated Kunshan Hospital of Jiangsu University, Kunshan, China

**Background:** Liver cancer is among the leading causes of death related to cancer around the world. The most frequent type of human liver cancer is hepatocellular carcinoma (HCC). Fatty acid (FA) metabolism is an emerging hallmark that plays a promoting role in numerous malignancies. This study aimed to discover a FA metabolism-related risk signature and formulate a better model for HCC patients' prognosis prediction.

**Methods:** We collected mRNA expression data and clinical parameters of patients with HCC using the TCGA databases, and the differential FA metabolism-related genes were explored. To create a risk prognostic model, we carried out the consensus clustering as well as univariate and multivariate Cox regression analyses. 16 genes were used to establish a prognostic model, which was then validated in the ICGC dataset. The accuracy of the model was performed using receiver operating characteristic (ROC) analyses, decision curve analysis (DCA) and nomogram. The immune cell infiltration level of risk genes was evaluated with single-sample GSEA (ssGSEA) algorithm. To reflect the response to immunotherapy, immunophenoscore (IPS) was obtained from TCGA-LIHC. Then, the expression of the candidate risk genes ($p < 0.05$) was validated by qRT-PCR, Western blotting and single-cell transcriptomics. Cellular function assays were performed to revealed the biological function of HAVCR1.

**Results:** According to the TCGA-LIHC cohort analysis, the majority of the FA metabolism-related genes were expressed differentially in the HCC and normal tissues. The prognosis of patients with high-risk scores was observed to be worse. Multivariate COX regression analysis confirmed that the model can be employed as an independent prognosis factor for HCC patients. Furthermore, ssGSEA analysis revealed a link between the model and the levels of immune cell infiltration. Our model scoring mechanism also provides a high predictive value in HCC patients receiving anti-PDL1 immunotherapy. One of the FA metabolism-related genes, HAVCR1, displays a significant differential expression between normal and HCC cell lines. Hepatocellular carcinoma cells (Huh7, and HepG2) proliferation, motility, and invasion were all remarkably inhibited by HAVCR1 siRNA.

**Conclusion:** Our study identified a novel FA metabolism-related prognostic model, revealing a better potential treatment and prevention strategy for HCC.

KEYWORDS

**fatty acid metabolism, hepatocellular carcinoma, prognosis model, tumor microenvironment, immunotherapy, HAVCR1**

## Introduction

Hepatocellular carcinoma (HCC) is the most widely known malignancy, resulting in significant human mortalities (1). The 5-year overall survival (OS) of patients with HCC has decreased by 20% globally and by 12% in Asian countries (2). Patients with advanced metastatic and/or recurrent HCCs have failed to gain benefit over the long term from standard HCC treatments such as surgical resection and liver transplantation (3). Therefore, it is crucial to explore novel molecularly-targeted therapies and new prognostic factors for HCC patients (4–6). Very, recently, multiple personalized molecular subtypes of HCC have been reported. Fu et al. constructed a novel predictive model for the prognosis of patients with HCC based on pyroptosis-related genes by categorizing HCC patients into two subgroups from the TCGA dataset (7). However, the accuracy of prognosis for HCC patients is still poor. Hence, more efficient prognostic factors must be explored.

**Abbreviations:** HCC, hepatocellular carcinoma; FA, fatty acid; TCGA, The Cancer Genome Atlas; ICGC, The International Cancer Genome Consortium; DEGs, differentially expressed genes; PPI, protein-protein interaction network; NES, normalized enrichment score; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; OS, overall survival; ROC, receiver operating characteristic; DCA, decision curve analysis; GSEA, gene-set enrichment analysis; GSVA, gene-set variation analysis; TME, Tumor Microenvironment; TIME, Tumor Immune Microenvironment; PD1, programmed cell death protein 1; PD-L1, Programmed death-ligand 1; HR, Hazard ratio.

Metabolic abnormalities are a typical characteristic of cancer (8). Cancer cells have unique metabolic characteristics that distinguish them from normal cells. When carcinogenic signals are blocked, cancer cells may be able to survive in the adverse microenvironments by metabolic reprogramming (9). Increasing evidence supports the critical involvement of metabolic reprogramming in tumor onset and progression (10–12). FA metabolism disorder has become a typical cancer cell characteristic (13). Many cellular biological processes require FAs, including membrane formation, signaling molecule release, and energy storage. FAs are essential for cancer formation and progression, according to several studies (14, 15). Wang et al. demonstrated that abnormal activation of various oncogenic signaling cascades promotes HCC development by regulating the lipid-metabolizing enzyme expression and/or activity, as well as FA metabolism reprogramming (16). Although FA metabolism has been associated with HCC oncogenesis, its correlation with the progression and clinical prognosis of HCC is yet unknown. Thus, the identification of novel and valuable characteristic molecular models linked to FA metabolism may shed light on the anti-HCC strategy.

The objective of this research was to develop a novel prognostic model according to DEGs linked to FA metabolism and to explore its relationship with clinicopathological parameters and OS in HCC. In addition, the correlation between the tumor immune microenvironment (TIME) and model genes was explored. The potential of our model in guidance of anti-PD-1 immunotherapy was also investigated.

Furthermore, the hub gene HAVCR1 was selected for further functional validation of HCC cells *in vitro* based on its expression level. Collectively, our study provides new insight into the relationship between FA metabolism and HCC.

# Materials and methods

## Data acquisition

The RNA-seq data (FPKM format), including 374 HCC and 50 normal tissues, along with clinical data, were provided by the TCGA database. The LIRI-JP cohort data set, containing transcriptomics data from 231 HCC patients, was retrieved from the ICGC database. cBioPortal shows genetic alterations in 16-risk-gene (17). The HPA database is composed of numerous sections that integrate several omics technologies for researchers (18–20). IHC images showed the difference in HAVCR1 protein expression between HCC tissues and normal tissues.

## Analysis of differential expression of genes related to FA metabolism

Table 1 lists 30 FA metabolism-related genes that were retrieved for the study. Differentially expressed genes (DEGs) in HCC were identified by using the "Limma" package of R ($p < 0.05$) according to the screening criteria.

We utilized the STRING database to create a protein-protein interaction network (PPI) regarding DEGs, and Cytoscape helped us visualize the interactive network data. The mutual regulatory relationship between DEGs was demonstrated by R (version 4.1.2).

## Consensus clustering and functional enrichment analysis

We employed the R package "Consensus ClusterPlus", for consensus clustering, and the mRNA expression data of 30 genes, which were highly correlated with FA metabolism, were classified into several molecular subtypes *via* the K-means clustering (21). 1,000 iterations were carried out to make sure the classification was accurate. DEGs were screened for further analysis based on the samples from the prior cluster analysis ($|log2FC| > 1$, adjusted $p < 0.05$). R package "GOplot" and "ggplot2" were used to perform GO and KEGG analyses between the two groups (22).

## Construction and validation of FA metabolism-related prognostic model

Initially, we used the univariate cox analysis for DEGs to identify and screen out the genes that were associated with prognosis in HCC patients ($p < 1*10-6$). These genes were then subjected to multivariate Cox analysis to identify 16 genes linked to prognosis, and the prognostic risk assessment construction formula is as stated below:

Risk score = coefficient1 * expression of gene1 +… + coefficientN * expression of geneN

As per the median risk scores, patients were subsequently classified into two groups. The OS of HCC patients in both groups was assessed through KM analysis using the "Survival" and "SurvMiner" R packages. The "Rtsne" package was used to perform PCA analysis on 16 prognostic genes to reduce the dimension of complex data. For identification and comparison of potential prognostic factors, the univariate Cox analysis was used, while a multivariate Cox analysis was performed to test whether the risk score was an independent prognostic factor. The accuracy evaluation of the prognostic model was done by ROC curves using the R packages "SurvivalROC" and "timeROC". Using the "rms" and "survival" packages, a predictive nomogram was developed according to the risk score and clinicopathological parameters (23). Using the "ggDCA" package, the DCA model intuitively described the relationship between risk score and other parameters.

TABLE 1 30 FA metabolism-related genes.

| |
|---|
| SREBF1 |
| ELOVL5 |
| SPTLC3 |
| ACSL1 |
| FABP5 |
| FASN |
| AMACR |
| SPOP |
| DGAT1 |
| ABHD5 |
| PDHA1 |
| ELOVL3 |
| ACOX1 |
| OXSM |
| ACAA2 |
| ADH5 |
| ACACB |
| CPT2 |
| SLC17A2 |
| CPT1C |
| ACADL |
| HADH |
| ELOVL6 |
| ADH6 |
| SIRT1 |
| ACADM |
| ACSL6 |
| FABP4 |
| SCD |
| CD36 |

## GSEA and GSVA

GSEA was employed in our study to investigate potential signaling pathways between the two groups to show a possible

molecular mechanism underlying the prognostic difference. The Type and Replacement were set to "phenotype" and "1000", respectively. "c2.cp.kegg.v6.2.-symbols" was downloaded to perform GSVA, which converts gene alterations into signaling pathway changes. To investigate potential changes in biological function and then annotate different risk genes, the GSVA algorithm and the "clusterProfiler" R package were employed.

## Evaluation of tumor immune microenvironment

Initially, expression data (ESTIMATE) was utilized to quantify the ratio of stromal cells to immune cells in malignancies, which was then used to estimate the TIME in HCC samples indirectly (24). The normalized enrichment score (NES) was then used to calculate the levels of immune function pathway enrichment. The scores of TIME cells were evaluated by the ssGSEA algorithm. The CIBERSORT algorithm was utilized to evaluate the relative proportion of 22 immune cells in the two groups with the aid of R 4.1.2. Finally, with the help of the "reshape2" and "ggpubr"R package the immune score, stromal score, and ESTIMATE score were obtained.

## Prediction of immunotherapeutic sensitivity

The tumor immune exclusion score can be used to indicate how well HCC patients respond to immunotherapy. The immunophenoscore (IPS), which assesses the tumor immunogenicity determinants based on machine learning, is a biomarker for the response to immunotherapy. To reflect the response of different groups to immunotherapy, IPS were obtained from TCGA-LIHC.

## Single-cell analysis

We used Tumor Immune Single-cell Hub (TISCH) pipline (25) to characterize LIHC tumor microenvironment at single-cell resolution (http://tisch.comp-genomics.org/). A total of 1,944,551 cells from 76 datasets across 28 cancer types and 101,195 cells from 3 PBMC datasets are retained in TISCH database. In GSE125449 database, we used TISCH pipeline to annotate the cell types (cell-type annotations provided by the original studies, marker-based annotation method employed in MAESTRO using the DEGs between clusters, InferCNV method). Finally, we annotated eight different cell clusters, including fibroblasts, endothelial cells, exhausted CD8 T cells (CD8Tex), Plasma cells, B cells, malignant cells, Monocytes or Macrophages (Mono/Macro), hepatic progenitor. Gene expression was compared between different cells. In addition, due to the HAVCR1 was not annotated in GSE125449 dataset, we used CellMarker database to search the cell location (http://yikedaxue.slwshop.cn/search.php?quickSearchInfo=HAVCR1#framekuang).

## Cell culture

The Institute of Neuroscience, Soochow University, provided the L-O2 cell line and human hepatocellular carcinoma cell lines (Huh7, HepG2). In a humidified atmosphere (37°C with 5% CO2), L-O2 cells were cultured in RPMI1640 with 20% FBS, whereas Huh7 and HepG2 were cultured in DMEM with 10% FBS.

## Cell transfection

Cells were transfected using CP Reagent (Ribo-Bio, Guangzhou, China). HAVCR1 expression was knocked down using two different types of siRNAs. The following are the HAVCR1 siRNA sequences: si-HAVCR1#1: GACGGCCAATACCACTAAA, si-HAVCR1#2: CGACTGTTCTGACGACAAT. As a negative control group (si-con), nonspecific siRNA was used. After 48 hours of transfection, the cells were collected. The efficiency measurement was carried out by qRT-PCR and Western blot.

## Cellular function assays

Cell suspension (1000 cells/well) was inoculated in a DMEM medium containing 10% FBS in 96-well plates. We incubated the culture plates for 24, 48 and 72 hours at 37°C and 5% CO2. The CCK8 assay was used to assess cell viability. Optical densities of CCK8 were measured using a microplate reader at 450 nm. EdU and phagokinetic track motility assays has been described in detail in our previous article (26). Transwell chambers (24-well, 12μm pore size, BIOFIL, China) were used to detect the migration of HuH7 and HepG2 after HAVCR1 silencing. The lower chamber was added with 0.6 mL DMEM with 20% FBS, whereas the upper chamber was added with around $7 \times 10^4$ cells resuspended in Basic DMEM media and incubated overnight at 37°C with 5% $CO_2$. Cells were fixed with 4% paraformaldehyde and stained with 2.5% crystal violet 24 hours later. Three microscopic views were selected randomly and counted by ImageJ. Transwell chambers (24-well,12 μm pore size, BIOFIL, China) were also used in a cell invasion assay. The upper chamber was precoated with 250 μg/mL Matrigel (BD Bioscience) and was left uncoated for migration. Subsequent steps are similar to the "Transwell Assay".

## qRT-PCR assay

HCC cells were seeded into 6-well plates at a density of $1\times10^5$ cells in each well. To extract total RNA, lysis buffer was

added to the culture. The QuantiTect Reverse Transcription Kit was then used to reverse transcribe the RNA into cDNA. qRT-PCR was performed through an SYBR Green PCR kit (Ribo-Bio, Guangzhou). The $2^{-\Delta\Delta Ct}$ method was used for the quantification of targeted mRNA. As an internal control, β-Actin mRNA was tested. Table 2 lists the gene primer sequences in detail.

## Western blotting

In this procedure, 10% SDS–polyacrylamide gel electrophoresis (SDS-PAGE) was used to separate aliquots of 20 mg of protein from each treatment, which were then transferred to the polyvinylidene difluoride (PVDF) membrane (Millipore, Bedford, MA). After a 2-hour blocking procedure with 10% instant nonfat dry milk (BD, USA), membranes were incubated with specific antibodies overnight at 4°C followed by 100 minutes at 20°C with HRP-conjugated secondary antibodies. The next Western blotting protocols were reported previously (26). Data quantification was performed by ImageJ. The primary antibodies include anti-TTK (1:1000, BOSTER), anti-KIF2C (1:1000, BOSTER), anti-MMP1 (1:1000, BOSTER), anti-HAVCR1 (1:1000, BOSTER), anti-β-actin (1:2000, Abcam).

## Statistical analysis

All bioinformatics analyses were performed using the R platform (v.4.1.2). The data were presented as mean ± SD. One-way analysis of variance (ANOVA) was used, followed by Student's t-test. Statistical significance was defined as a P-value of less than 0.05 ($P < 0.05$).

TABLE 2 Premier sequences for qRT-PCR analysis.

| Premier | Sequences (5′–3′) |
|---|---|
| ANLN-F | TGCCAGGCGAGAGAATCTTC |
| ANLN-R | CGCTTAGCATGAGTCATAGACCT |
| UCK2-F | GCCCTTCCTTATAGGCGTCAG |
| UCK2-R | CTTCTGGCGATAGTCCTACTTC |
| LPCAT1-F | CGCCTCACTCGTCCTACTTC |
| LPCAT1-R | TTCCCCAGATCGGGATGTCTC |
| TTK-F | GTGGAGCAGTACCACTAGAAATG |
| TTK-R | CCCAAGTGAACCGGAAAATGA |
| KIF2C-F | CTCAGTTCGGAGGAAATCATGTC |
| KIF2C-R | TGCTCTTCGATAGGATCAGTCA |
| HAVCR1-F | TGGCAGATTCTGTAGGCTGGTT |
| HAVCR1-R | AGAGAACATGAGCCTCTATTCCA |
| MMP1-F | AAAATTACACGCCAGATTTGCC |
| MMP1-R | GGTGTGACATTACTCCAGAGTTG |
| CBX2-F | GCCCAGCACTGGACAGAAC |
| CBX2-R | CACTGTGACGGTGATGAGGTT |
| ACTB-F | TCAAGATCATTGCTCCTCCTGAG |
| ACTB-R | ACATCTGCTGGAAGGTGGACA |

# Results

## Identification of FA metabolism-related DEGs between normal and HCC tissues

A flow chart of the overall research is shown in Figure S1. 30 genes were chosen based on their roles in previous studies (27–33). The majority of FA metabolism-related genes were identified as DEGs ($p < 0.05$) using Heatmap analysis (Figure 1A). The protein-protein interaction (PPI) analysis was carried out on these DEGs using the Homo sapiens data set (with a confidence of 0.9) to better comprehend their interactions. The PPI network retained 23 hub DEGs that had complicated regulatory correlations (Figure 1B). Moreover, the correlation of these DEGs was analyzed and presented in Figure 1C (cutoff >0.4). We preliminarily concluded that the majority of these FA metabolism-related DEGs affect the tumorigenesis and tumor progression through mutual positive regulation.

## Clustering, construction of HCC classification according to genes associated with FA metabolism

Based on the expression matrix of 30 FA metabolism-related DEGs, 2 clusters were identified using unsupervised clustering methods (k=2, Figures 2A–C). There were significant differences in OS time among them (Figure 2D). Thus, identifying prognostic genes related to FA metabolism was crucial. The DEGs between the two subtypes were then screened and obtained for subsequent analysis ($|\log2FC| > 1$, p-value < 0.05). DEGs expression profiles and clinicopathologic parameters were shown on the heatmap. As expected, the expression of most DEGs, as well as the number of patients with stage III-IV, were significantly greater in cluster1 (Figure 2E).

## Functional analyses and the tumor immune microenvironment between FA metabolism-related subtypes

In the TCGA-LIHC cohort, DEGs between the two subtypes were analyzed to study the biological role and pathways of FA metabolism-related genes in more detail. As shown, GO analysis was divided into several parts: biological process (BP), cellular component (CC) and molecular function (MF) (Figure 3A).

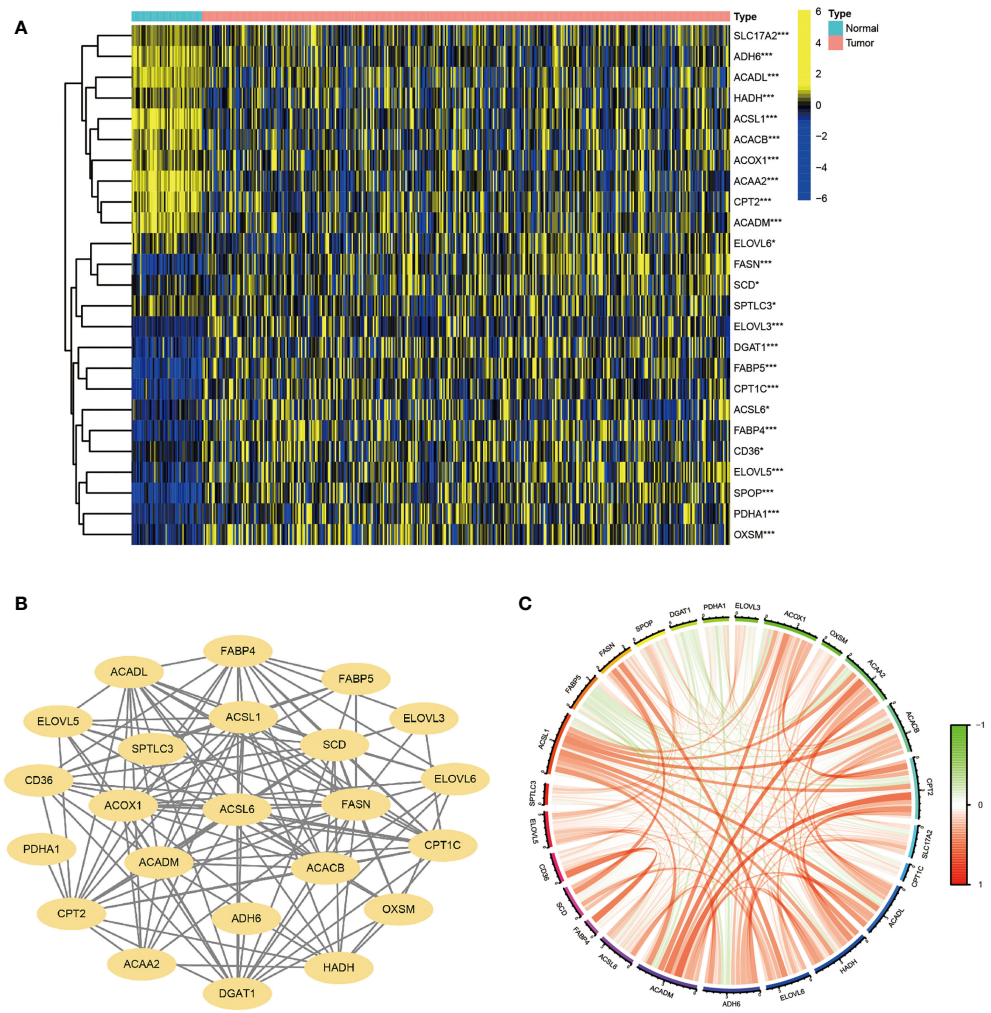



FIGURE 1
Identification of FA Metabolism-Related DEGs between Normal and HCC Tissues. The landscape of FA metabolism-related genes of HCC patients from TCGA database. **(A)** Heatmap showed DEG expression in two different tissues. **(B)** PPI network of the interactions. **(C)** The correlation network of these genes (*p < 0.05, **p < 0.01, ***p < 0.001).

Notably, DEGs were also abundantly enriched in numerous immune responses, including neutrophil activation involved in immune response, neutrophil degranulation, T cell proliferation, mononuclear cell proliferation, regulation of leukocyte proliferation, lymphocyte apoptotic process, B cell apoptotic process, T-helper 1 cell differentiation (Table 3). Further, KEGG enrichment analyses indicated that the role of the Ribosome, Retinol metabolism, PPAR signaling pathway, Fatty acid metabolism, Cell cycle were enriched in both cohorts (Figure 3B). The tumor immune microenvironment (TIME) has profound implications for tumor diagnosis, patient survival outcomes, and sensitivity to clinical treatment (34). By analyzing the relationship between DEGs and TIME in the two subtypes, the potential immune mechanism of FA metabolism affecting the tumorigenesis of HCC was revealed. Results showed that both the subtypes were significantly associated with the immune scores. Extraordinarily, the patients with higher infiltration levels of immunosuppressive cells such as activated B cell, activated CD4 T cell, Mast cell and MDSC were more prone to cluster1 (Figure 3C). These findings confirmed that the expression of FA metabolism-related DEGs is associated with the prognosis and the TIME in HCC patients.

## Establishment of a prognostic risk model in TCGA-LIHC cohort

The prognostic value of risk characteristics of FA metabolism was explored considering the complex regulation. To identify prognostic genes in the TCGA-LIHC, researchers used a univariate Cox regression analysis, which

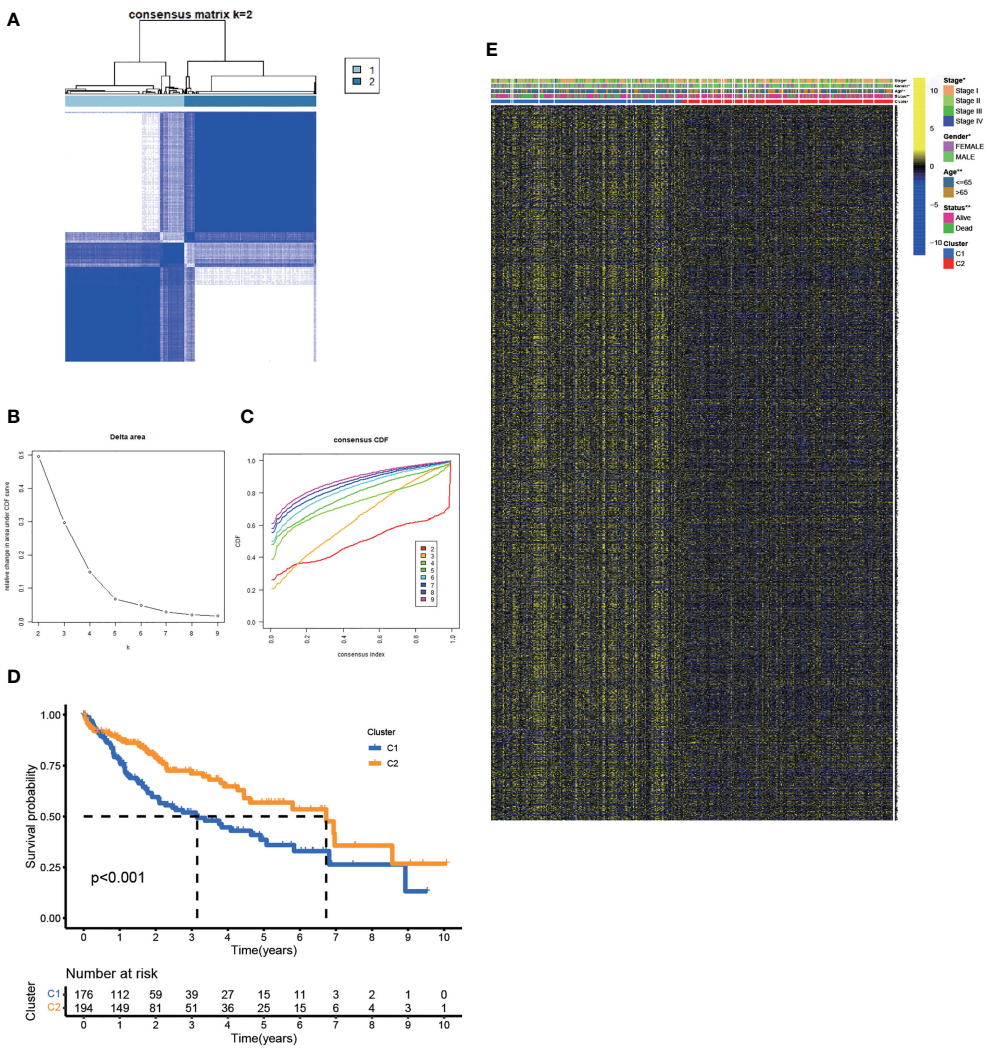
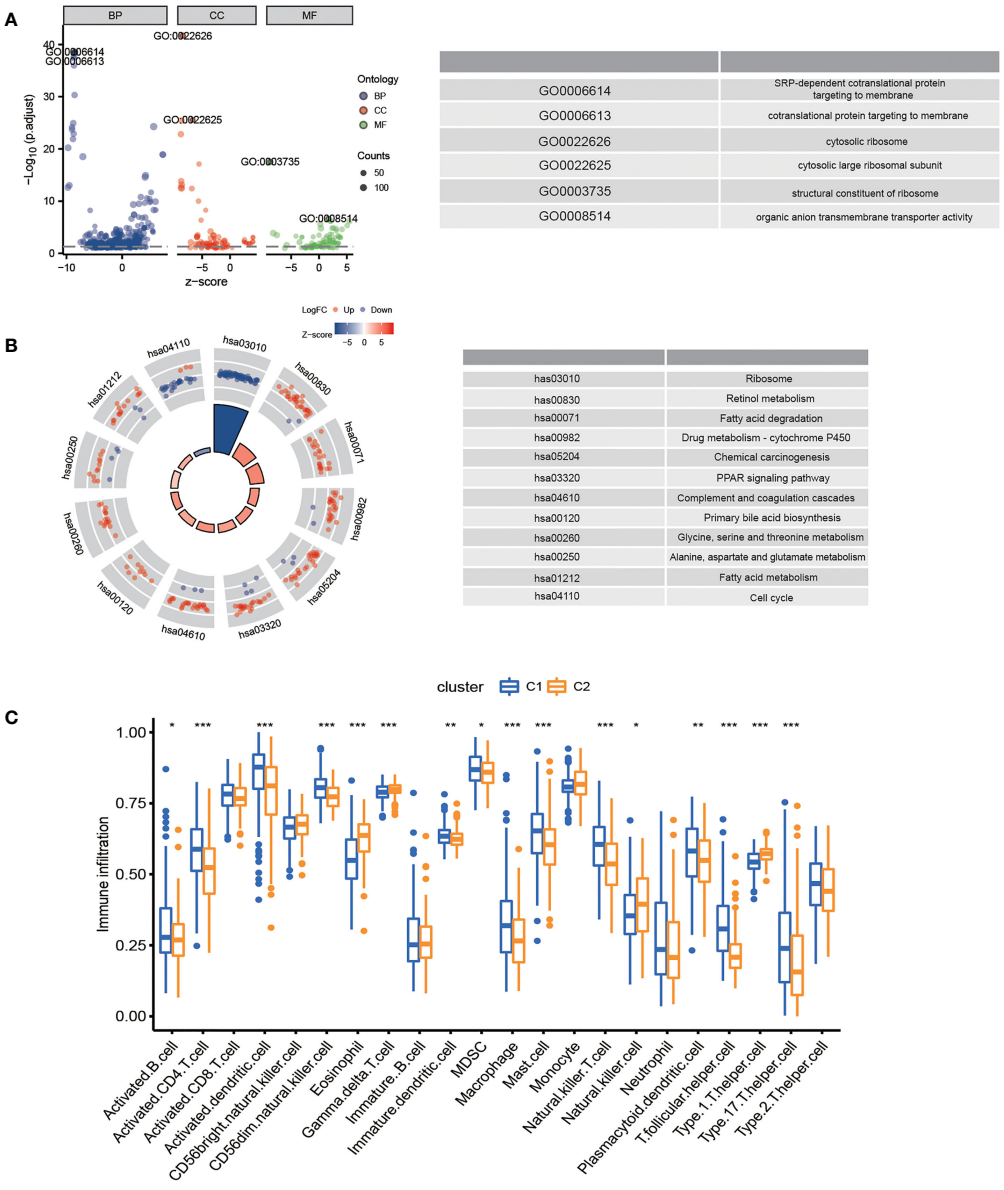



FIGURE 2
Clustering, Construction of HCC Classification According to Genes associated with FA Metabolism. **(A)** Patients were classified into two subtypes according to the consensus clustering matrix. **(B, C)** Consensus clustering model. **(D)** The KM analysis of the OS in the two subtypes. **(E)** Heatmap showed the correlation between the subtypes and clinicopathologic parameters (*p < 0.05, **p < 0.01, ***p < 0.001).

revealed that all genes were high-risk genes for HCC prognosis (Figure 4A). Multivariate Cox regression was used on the TCGA-LIHC cohort to further narrow down the potential gene range for developing a prognostic model. Genes including ANLN, UCK2, LPCAT1, TTK, CCNB1, KIF2C, HAVCR1, MMP1, PIGU, CENPA, CENPO, CDCA8, CBX2, KIAA1841, KIF18A, and CEP55, with their coefficients were subsequently maintained (Figure 4B). After exploring the prognosis of 16-gene, we used cBioPortal to analyze its mutation in HCC. As displayed in Figure 4C, all of these genes had great genetic variations, of which amplification was the most common variation characteristic. In addition, there was a significant positive correlation between these 16 genes (Figure 4D). As per the median risk score, we separated the patients into high- and low-risk groups, these two groups could be well-separated according to the PCA and t-SNE analysis (Figures 4E, F). According to the KM analysis, the association between high-risk score patients and poor prognosis was significant (Figure 4G). Furthermore, the number of mortalities showed an increasing trend with increasing risk scores (Figure 4H). The receiver operating characteristic (ROC) curve, also known as the sensitivity curve, was constructed to evaluate the model's accuracy and feasibility in predicting patients' survival (35), suggesting that the model exhibited a great predictive capability (AUC=0.811). Besides, the ROC curve also indicated the effectiveness of the FA metabolism-related signature in predicting the 1-, 3-, and 5-year survival rates in patients with HCC (Figure 4I).




| | |
|---|---|
| GO0006614 | SRP-dependent cotranslational protein targeting to membrane |
| GO0006613 | cotranslational protein targeting to membrane |
| GO0022626 | cytosolic ribosome |
| GO0022625 | cytosolic large ribosomal subunit |
| GO0003735 | structural constituent of ribosome |
| GO0008514 | organic anion transmembrane transporter activity |

| | |
|---|---|
| has03010 | Ribosome |
| has00830 | Retinol metabolism |
| hsa00071 | Fatty acid degradation |
| hsa00982 | Drug metabolism - cytochrome P450 |
| hsa05204 | Chemical carcinogenesis |
| hsa03320 | PPAR signaling pathway |
| hsa04610 | Complement and coagulation cascades |
| hsa00120 | Primary bile acid biosynthesis |
| hsa00260 | Glycine, serine and threonine metabolism |
| hsa00250 | Alanine, aspartate and glutamate metabolism |
| hsa01212 | Fatty acid metabolism |
| hsa04110 | Cell cycle |

FIGURE 3
Functional Analyses and the Tumor Immune Microenvironment Between FA metabolism-related Subtypes. Potential biological pathways and tumor microenvironment affected by DEGs. GO **(A)** and KEGG **(B)** enrichment of DEGs. **(C)** Comparison of the ssGSEA scores between two subtypes (*p < 0.05, **p < 0.01, ***p < 0.001).

## Association between risk genes and clinicopathologic parameters

Next, the association between the model and clinical parameters in HCC patients was explored. We discovered that these risk genes were substantially expressed in the high-risk group, as shown by the heatmap that shows risk gene expression profiles and clinicopathologic parameters. Moreover, there were significant differences between tumor stage and grade (Figure S2A). Notably, patients having higher risk scores may be in higher stages, whether AJCC stage or T stage (Figure S2B). Then, we divided the patients into several subgroups based on different clinical parameters such as gender (female *vs* male), age (> 65 *vs* ≤ 65), AJCC stage (I-II *vs* III-IV), and T stage (T1 *vs* T2-4). The KM analysis revealed that the high-risk patients had a lowered survival rate in all conditions (Figures S2C–F). Overall, the model constructed is highly correlated with clinicopathologic parameters and can guide the prognosis of HCC patients.

TABLE 3 Immunologic signature associated biological processes enriched in DEGs groups.

| Gene set name | Size | Zscore | P.adjust | Qvalue |
|---|---|---|---|---|
| GO- neutrophil activation involved in immune response | 82 | -5.08 | 0.008 | 0.007 |
| GO- neutrophil degranulation | 81 | -5.222 | 0.01 | 0.009 |
| GO- T cell proliferation | 37 | -4.768 | 0.012 | 0.01 |
| GO- mononuclear cell proliferation | 50 | -5.94 | 0.014 | 0.012 |
| GO- regulation of leukocyte proliferation | 41 | -5.154 | 0.025 | 0.021 |
| GO- lymphocyte apoptotic process | 17 | -3.638 | 0.028 | 0.023 |
| GO- B cell apoptotic process | 8 | -2.828 | 0.048 | 0.041 |
| GO- T-helper 1 cell differentiation | 7 | -1.89 | 0.046 | 0.039 |

## Independent prognostic value of the model

The efficiency of the model was tested in the TCGA-LIHC cohort. Based on univariate COX analysis, a high-risk score was shown to be correlated with poor prognosis significantly ($p < 0.001$, HR = 1.194, 95% CI: 1.139 - 1.251). The other variable associated with a worse prognosis was a stage (Figure 5A). According to multivariate Cox analysis, a higher risk score was revealed to be independently associated with poorer survival, supporting its potential for being an independent prognostic factor for HCC ($p < 0.001$, HR = 1.217, 95% CI: 1.166 - 1.217) (Figure 5B). Notably, DCA, a novel method that is used to assess clinical predictive models, diagnostic tests, and molecular markers (36), showed that our risk model achieves greater net benefit than any one single independent clinical parameter (Figure 5C). Additionally, the nomogram (C-index > 0.7) based on the clinical parameters and risk scores could effectively predict the probability of the1-, 2-, and 3- years OS (Figure 5D). Calibration curve results verified high agreement between nomogram predictions and actual observations (Figure 5E).

## Validation of the risk signature in ICGC-LIRI cohort

To further verify the model's predictive accuracy, 231 HCC cases were extracted from the ICGC database to establish a test cohort. In the LIRI-JP cohort, 16 risk genes were all up-regulated, as shown in Figure S3A. Then, using these risk genes to separate the LIRI-JP cohort into two groups, we observed that patients in the high-risk group had higher mortality and shorter survival periods (Figure S3B). Based on the KM analysis, the OS of patients belonging to the low-risk group was higher (Figure S3C) 1-, 3-, and 5-year AUC values were 0.619, 0.595, and 0.950, respectively (Figure S3D), and the PCA plot validated that the high- and low-risk groups could be separated (Figure S3E). The association of high- and low-risk scores and OS in HCC patients was further validated in various clinical parameter subgroups. Based on the KM analysis, patients with high-risk scores had a lower survival rate when they were female, over 65, in stages I-II, and had primary malignancy (Figure S3F). All these suggest the reliability of the model.

## The potential molecular mechanism of the model

GSEA was applied to analyze the transcript message of HCC patients. Interestingly, the activity of metabolic pathways such as cytochrome P450 drug metabolism, FA metabolism, and retinol metabolism, were shown to be enriched in the low-risk group, according to KEGG enrichment analysis (Figure S4A). Cell cycle, DNA replication, ECM receptor interaction, neuroactive ligand-receptor interaction, and oocyte meiosis were enriched in the high-risk group (Figure S4B). Then, we performed GSVA enrichment to further explore potential signaling pathways. As shown in Figure S4C, these model genes were significantly enriched in most signaling oncogenic pathways and were positively correlated, including WNT, VEGF, Notch, and mTOR signaling pathways. In some pathways, such as PPAR and ADIPOCYTOKINE signaling pathways, these genes show consistent negative correlations. In addition, immune-related pathways, such as T cell receptor signaling pathway, B cell receptor signaling pathway were also enriched. These results provided a novel strategy for our subsequent research to find potential therapeutic targets.

## TME infiltration and immunotherapy

Based on our findings above, we suggested this prognostic model is closely correlated with immune infiltration. Using ssGSEA, we systematically evaluated 13 types of immune function pathways to further assess the immune status-related association between the two groups. The risk score was highly associated with Type II IFN Response, Type I IFN Response, MHC class I, and Cytolytic activity,

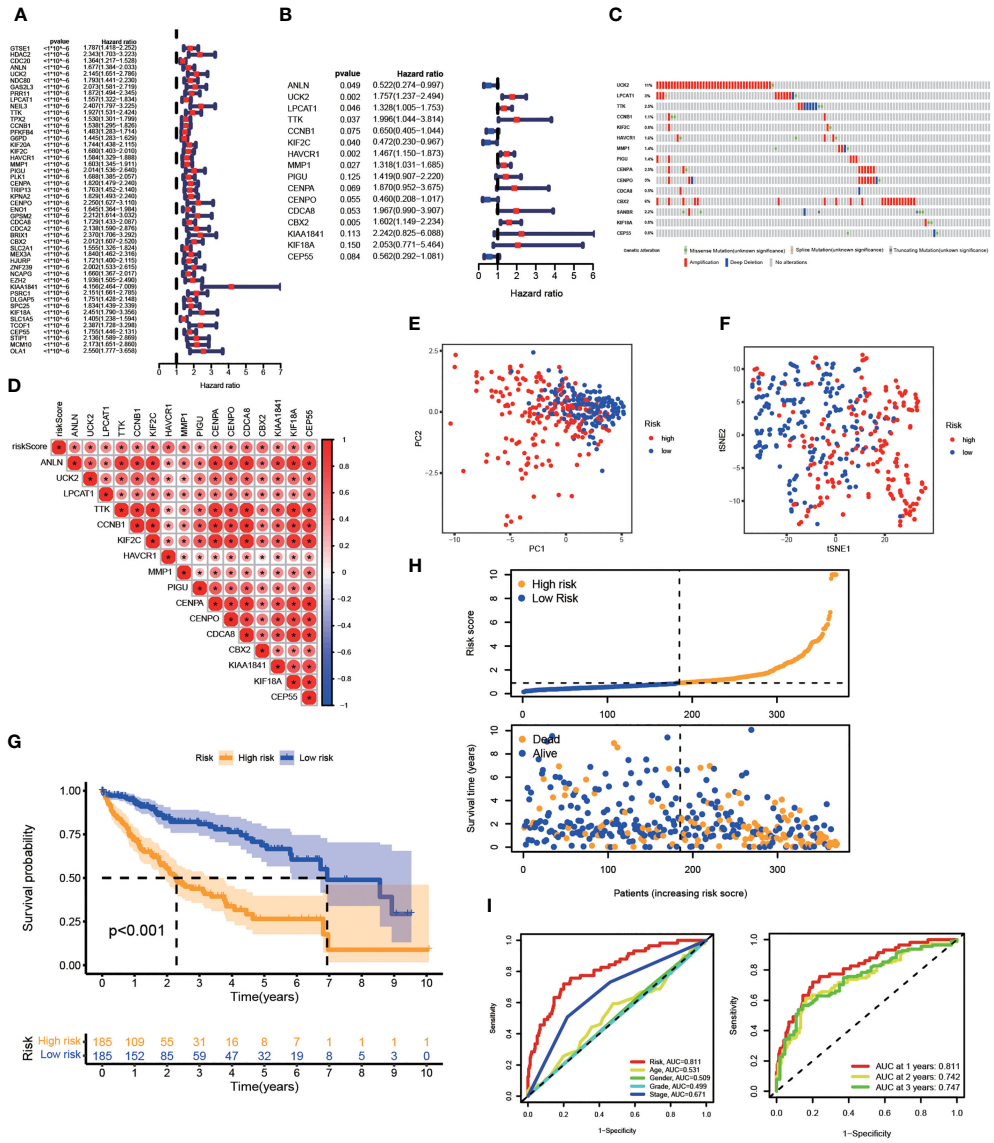



FIGURE 4
Establishment of A Prognostic Risk Model in TCGA-LIHC Cohort. Construction of a FA metabolism-related model in TCGA. **(A)** Univariate Cox regression analysis to find FA metabolism-related prognostic genes. **(B)** Multivariate Cox regression analysis to identify genes linked to the prognosis of FA metabolism. **(C)** The distribution of 16-gene genomic alterations in the TCGA-LIHC dataset. **(D)** Spearman correlation analysis of 16 genes. PCA analysis **(E)** and t-SNE analysis **(F)** based on risk scores. **(G)** The KM analysis of the OS based on risk scores. **(H)** Survival status distribution of these two groups. **(I)** ROC curve showing the accuracy of risk scores on the clinical parameters and year survival rate (*p < 0.05, **p < 0.01, ***p < 0.001).

according to Heatmap analysis results (Figure 6A). The distribution of immune cells calculated by XCELL, TIMER, QUANTISEQ, MCPCOUNTER, EPIC, and CIBERSORT was also explored. HCC patients belonging to the high-risk group had higher proportions of immune cells including B Cell, T Cell, Macrophage, and so on (Figure 6B). The correlation between these 16 risk genes and immune cells was also performed by the CIBERSORT algorithm and presented in Figure 6C. The stromal score, immune score, and estimate score were then generated using the ESTIMATE algorithm. Furthermore, low-risk group patients had a higher stromal score ($p < 0.05$), immune score ($p < 0.001$) and estimate score ($p < 0.001$) (Figure 6D). Immunotherapy is a new type of therapy for a variety of cancers, including HCC. Regarding the response to immunotherapy in these two groups of HCC patients, we found a higher Exclusion score in the high-risk group, indicating a worse effect on receiving immunotherapy in the high-risk group (Figure 6E). Anti-PD-1/PD-L1 therapies have emerged as an effective treatment option, especially in HCC (37).

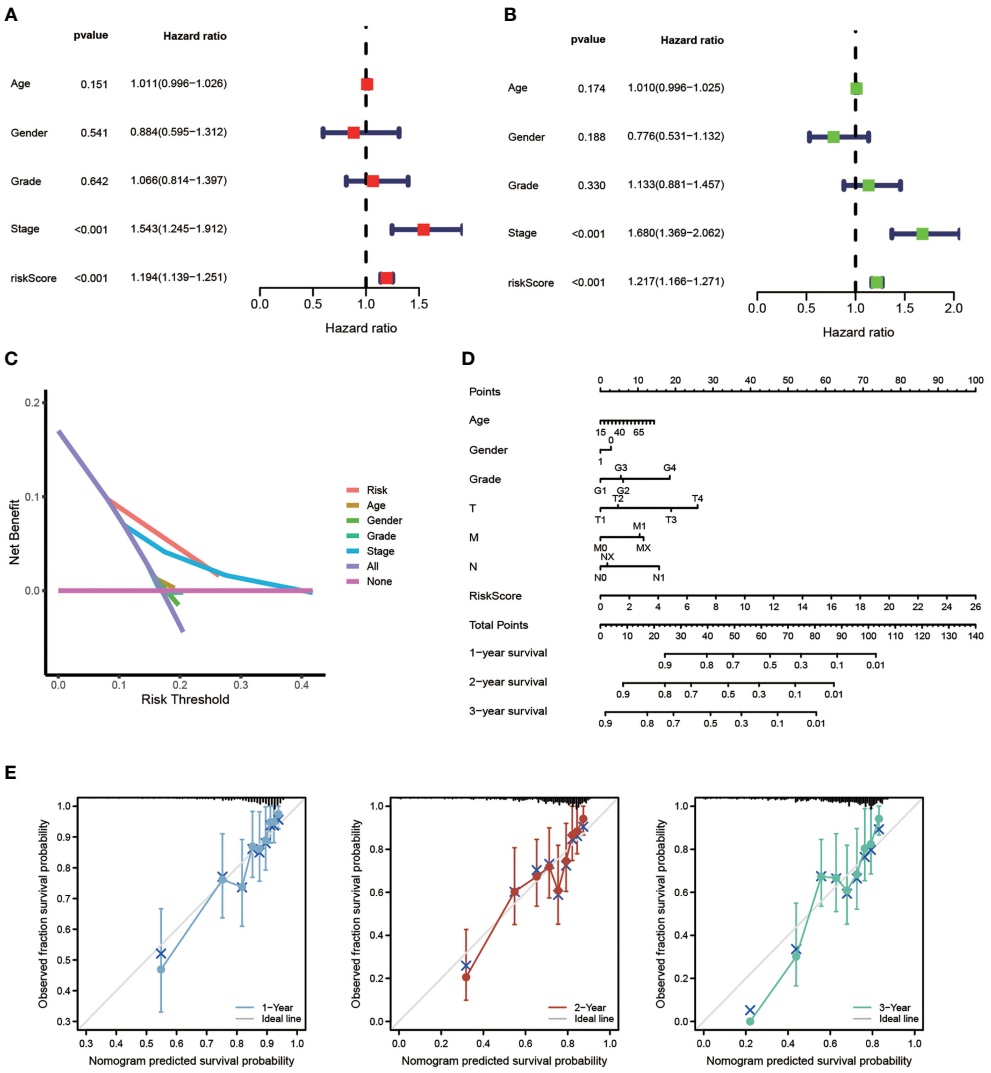



FIGURE 5
Independent Prognostic Value of The Model. Univariate Cox **(A)** and Multivariate Cox **(B)** analysis to assess the independence of the model. **(C)** DCA was performed to present the net benefit of risk score compared to clinical parameters. **(D)** Nomogram survival prediction of HCC patients with risk score. **(E)** Calibration plot of the nomogram (*p < 0.05, **p < 0.01, ***p < 0.001).

Significant differences in immunotherapy scoring mechanisms were revealed in all four immunotherapy regimens; a lower risk score suggests a greater anti- PD-1/L1 therapeutic efficacy (Figure 6F). FA metabolism-related prognostic model was associated with anti-PD-1/L1 immunotherapy, as expected, and can potentially predict immunotherapy response.

## Differential expression of independent prognostic genes and validation

Subsequently, eight candidate genes including ANLN, UCK2, LPCAT1, TTK, KIF2C, HAVCR1, MMP1, and CBX2 were selected based on multivariate Cox regression analyses ($p < 0.05$). Notably, the correlation between these genes was significantly positive (Figure 7A). We further explored the expression of these prognostic genes in 50 pairs of samples from the TCGA-LIGC cohort, and the results revealed that significant elevation of the mRNA expression levels of these genes in HCC tissue (Figure 7B). In addition, the ROC curve displays a favorable predictive value of these independent genes over 1, 3, and 5 years (Figure 7C). To verify mRNA expression in HCC, we performed qRT-PCR in human liver epithelial (LO2) and two HCC cell lines (Huh7, HepG2) (Figure 7D). To test TTK, KIF2C, HAVCR1, and MMP1 protein expressions, Western blotting assays were then performed and the results confirmed the up-regulation of all these

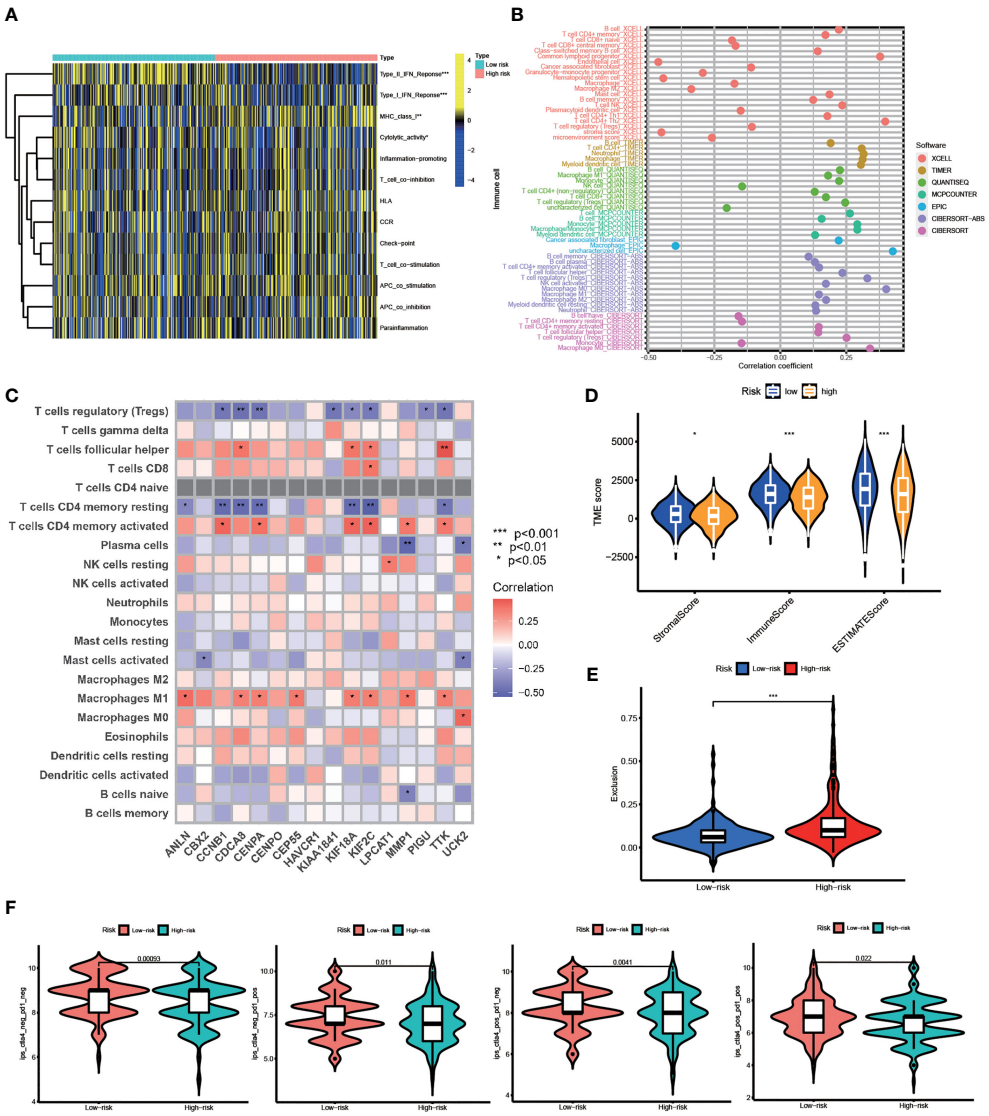



FIGURE 6
TME Infiltration and Immunotherapy. **(A)** Heatmap of the distribution of 13 types of immune function pathways between two groups. **(B)** Immune cell infiltration analysis based on different algorithms. **(C)** Correlation between immune-related cells and 16 genes. **(D)** TME score, including Stromal score, Immune score and Estimates score. **(E)** Exclusion score calculated by TIDE algorithm. **(F)** Immunotherapy score between two groups (po, positive; neg, negative) (*p < 0.05, **p < 0.01, ***p < 0.001).

proteins in HCC cell lines (Figure 7E). Extraordinarily, we found dramatically increased HAVCR1 mRNA and protein expression in HCC cell lines. The protein levels of HAVCR1 were validated by the HPA database (Figure 7F).

## Single cell analysis

In GSE125449 database, we used TISCH pipeline to annotate the cell types (cell-type annotations provided by the original studies, marker-based annotation method employed in MAESTRO using the DEGs between clusters, InferCNV method). Finally, we annotated eight different cell clusters (Figure S5A), including fibroblasts, endothelial cells, exhausted CD8 T cells (CD8Tex), plasma cells, B cells, malignant cells, monocytes or macrophages (mono/macro), hepatic progenitor. Subsequently, we explored the expression of our risk genes in different cell types. Unfortunately, the HAVCR1 was not annotated in this dataset. Hence, we used CellMarker database to search the cellular location, and we found HAVCR1 was

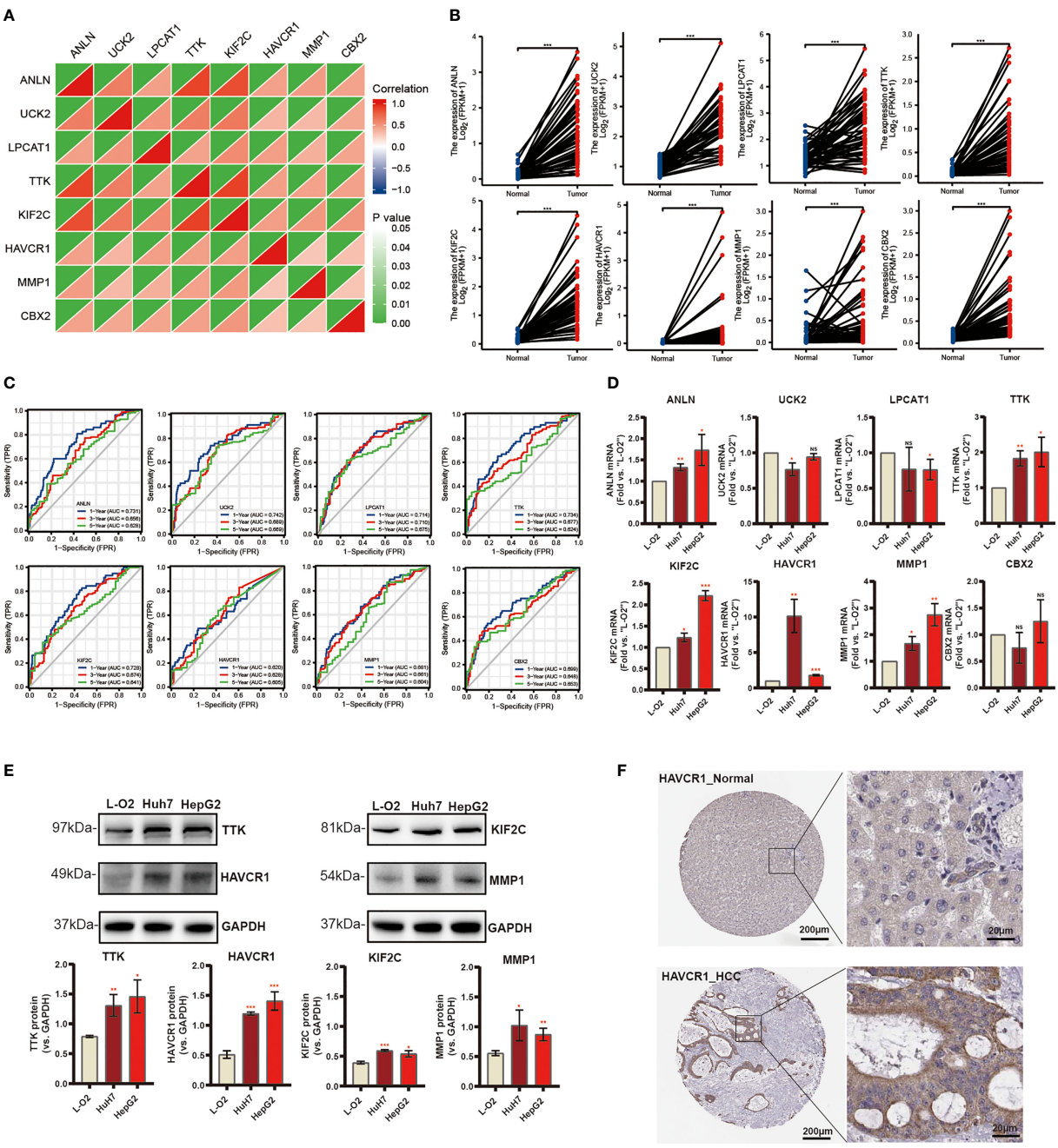



FIGURE 7
Differential Expression of Independent Prognostic Genes and Validation. **(A)** Spearman correlation analysis of eight candidate genes in the TCGA-LIHC cohort. **(B)** Eight risk genes expression in paired tissues from TCGA database. **(C)** The ROC curve of independent risk genes in TCGA-LIHC cohort. **(D)** The mRNA levels of quantified using qRT-PCR analysis in human liver cell line and two HCC cell lines. **(E)** The protein expression of TTK, KIF2C, HAVCR1, and MMP1. **(F)** Immunohistochemistry of the HAVCR1 from the HPA database (*p < 0.05, **p < 0.01, ***p < 0.001).

mainly expressed in normal liver cells. Moreover, the TISCH results showed that most of the genes were not significantly expressed in B cells, and only LPCAT1 and CBX2 were significantly expressed in B cells (Figure S5B). Among them, the ANLN, UCK2, KIF2A were more evenly distributed in cell types other than B cells. TTK was significantly expressed in CD8Tex cells, hepatic progenitor, malignant, and mono/macro cells. Finally, MMP1 was only significantly expressed in

TABLE 4 Association of HAVCR1 expression with clinicopathological parameters in TCGA-LIHC.

| Characteristic | Low expression of HAVCR1 | High expression of HAVCR1 | p |
|---|---|---|---|
| n | 187 | 187 | |
| Age, n (%) | | | 0.643 |
| <=60 | 86 (23.1%) | 91 (24.4%) | |
| >60 | 101 (27.1%) | 95 (25.5%) | |
| Gender, n (%) | | | 0.185 |
| Female | 54 (14.4%) | 67 (17.9%) | |
| Male | 133 (35.6%) | 120 (32.1%) | |
| Pathologic stage, n (%) | | | 0.436 |
| Stage I | 89 (25.4%) | 84 (24%) | |
| Stage II | 44 (12.6%) | 43 (12.3%) | |
| Stage III | 38 (10.9%) | 47 (13.4%) | |
| Stage IV | 1 (0.3%) | 4 (1.1%) | |
| AFP(ng/ml), n (%) | | | < 0.001*** |
| <=400 | 122 (43.6%) | 93 (33.2%) | |
| >400 | 19 (6.8%) | 46 (16.4%) | |
| Histologic grade, n (%) | | | 0.003** |
| G1 | 38 (10.3%) | 17 (4.6%) | |
| G2 | 90 (24.4%) | 88 (23.8%) | |
| G3 | 52 (14.1%) | 72 (19.5%) | |
| G4 | 3 (0.8%) | 9 (2.4%) | |
| Age, median (IQR) | 61 (52.5, 69) | 61 (51, 68) | 0.290 |

endothelial cells, fibroblasts, hepatic progenitor and malignant cells. Taken together, our data showed risk genes were not only expressed in malignant cells, and different genes had expression heterogeneity.

## HAVCR1 silencing inhibits HCC cell proliferation, motility, and invasion

Considering the model was strongly associated with the HCC, the independent prognostic genes may have a greater impact on the biological function of HCC cells. We selected HAVCR1 with the largest expression difference to further verify our hypothesis. The relationship between HAVCR1 expression and prognosis of HCC was validated in ICGG database (Figure S6). Moreover, HAVCR1 expression was positively correlated with histologic grade (Table 4). HAVCR1 siRNA was transduced to Huh7 cells and HepG2 cells. Robust decrease of HAVCR1 mRNA and protein levels by HAVCR1 siRNA was confirmed by qRT-PCR and Western Blotting assay (Figures 8A, B). CCK-8 assay results showed a significant reduction in viability by HAVCR1 siRNA in HuH7 and HepG2 cells (Figure 8C). Huh7 and HepG2 cell proliferation was largely inhibited by decreased EdU-positive nuclei ratio after siRNA-mediated knockdown of HAVCR1 (Figure 8D). In addition, the phagokinetic track motility assay results confirmed that cell motility was significantly inhibited by siRNA (Figure 8E). Moreover, using the "Transwell" assay, it was shown that HCC cell migration was attenuated (Figure 8F). A significant decrease in the HCC cell invasion was also shown by "Matrigel Transwell" assays (Figure 8G). These findings implied that HAVCR1 siRNA can inhibit the biological function in HCC, but the detailed mechanism needs to be further illuminated.

# Discussion

Recently, the HCC incidence has been rising (38). Despite the advances in cancer prevention, early screening, and current treatment options, the prognosis for HCC is extremely poor (39). The current diagnostic options are not sensitive and accurate enough (40). Thus, it is extremely necessary to explore novel and efficient markers of diagnosis and prognosis for improving the OS of HCC patients (41). Increasing studies have shown that metabolic dysregulation is one of the main characteristics of malignant tumor cells, leading to growth, angiogenesis, proliferation, and invasion (42–44). FA metabolism, as an important part of energy metabolism, is involved in multiple biological processes for promoting tumorigenesis and progression (13). FA metabolism has been shown to play a key role in the onset and progression of HCC (39). Most research focused only on a single regulatory factor of FA metabolism in HCC (10, 29, 45), however, integrated models

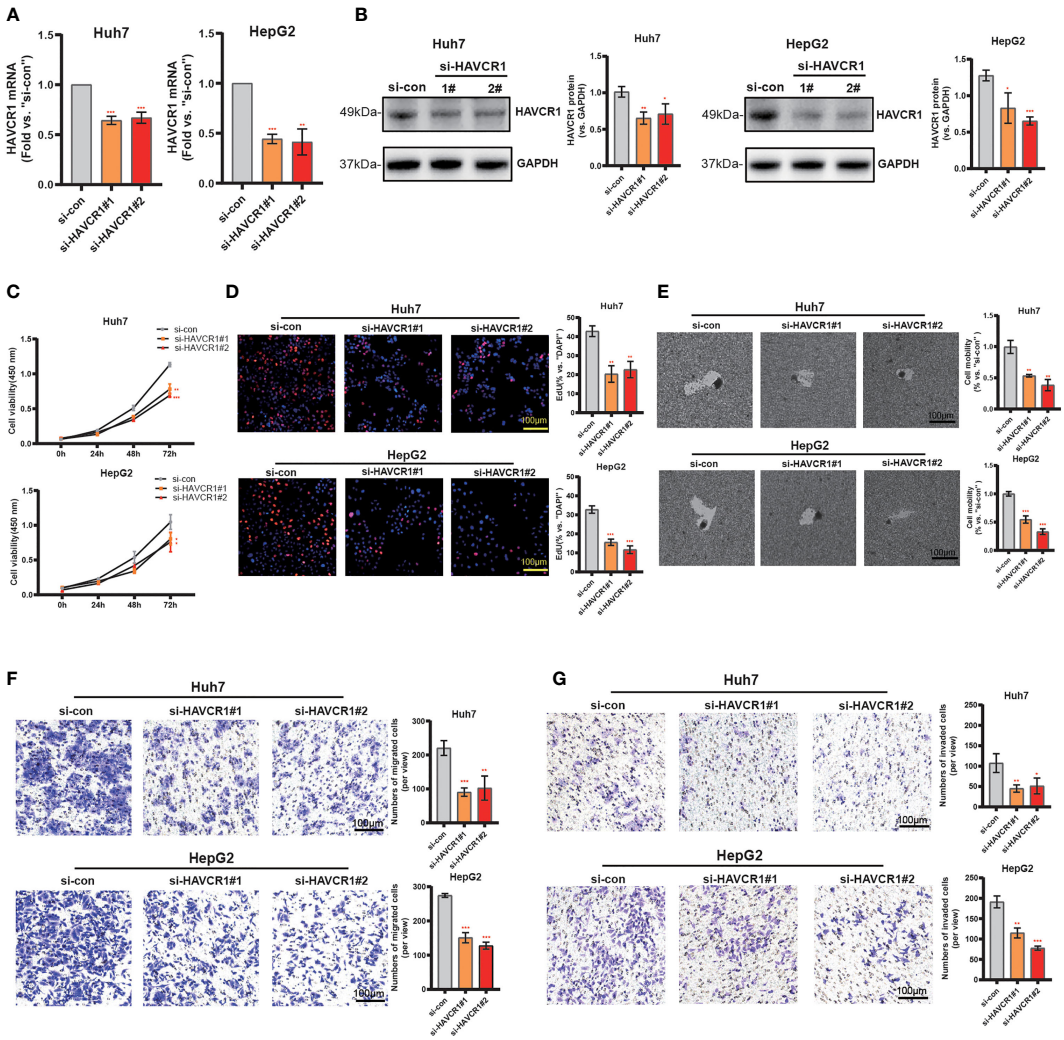



FIGURE 8
HAVCR1 Silencing Inhibits HCC Cell Proliferation, Migration, and Invasion. Established human HCC cell lines (HuH7 and HepG2) (**A–F**), bearing the HAVCR1 siRNA ("si-HAVCR1#1" and "si-HAVCR1#2"). qRT-PCR **(A)** and Western blotting **(B)** were employed for assessing the HAVCR1 mRNA and protein expression after si-HAVCR transfection. CCK-8 **(C)** and EdU assay **(D)** were used to test the proliferation of HuH7 and HepG2 cells. Phagokinetic track motility assay **(E)** was used measure the motility of cells. Cell migration and invasion were measured by "Transwell" **(F)** and "Matrigel Transwell" assays **(G)** (*p < 0.05, **p < 0.01, ***p < 0.001).

of multiple important genes involved in FA metabolism are needed. He et al. had explored FA metabolism-related risk genes in HCC by constructing a prognostic model but failed to further explore the role of these genes in the onset and progression of HCC (46). Identification of key molecular markers associated with FA metabolism and clarification of their roles in the progression of HCC is necessary.

Our present study first systematically investigated 30 genes highly associated with FA metabolism in patients with HCC and revealed that 14 genes among them were significantly upregulated, while 10 genes were downregulated. Most of these genes were positively correlated. Then, we identified two clusters based on these FA metabolism-related genes by performing consensus clustering, these two clusters showed significant survival differences. DEGs were then compared between the two clusters, with the results revealing enrichment of DEGs in immune processes. Currently, the investigation of FA metabolism-related genes in the TIME in HCC is insufficient. Our study showed that patients with few immunosuppressive cells favored cluster 2 compared to those with more immunosuppressive cells. High infiltration of immunosuppressive cells suggested the tumor microenvironment was inhibited, leading to the poor prognosis of HCC. As the consensus clustering was based on 30 FA metabolism-related genes expression, we inferred that FA metabolism was closely related to the prognosis and TIME of HCC patients.

DEGs were analyzed using univariate and multivariate Cox regression analyses to develop a 16-gene risk model to further investigate the prognostic significance of FA metabolism in HCC. The model possessed great predictive accuracy and could guide the prognosis among the patients with different clinical parameters. These results were validated in the ICGC external validation dataset. GSEA enrichment was performed to further explore potential molecular mechanism of the model, several pathways (e.g., Cell cycle, DNA replication, ECM receptor interaction, and so on) involving the tumorigenesis were enriched in the high-risk group. It is widely recognized that abnormal cell cycle and DNA replication was considered a biomarker of HCC (47, 48). GSVA enrichment showed the model genes were significantly enriched in most signaling oncogenic pathways and were positively correlated, suggesting these risk genes play the oncogenic role in HCC. All these findings provide ideas for our future mechanism research.

Robust evidence has shown an intimate relationship between FA metabolism and tumor immunity (49, 50). The immune infiltration status of 22 immune cells was analyzed by the ssGSEA algorithm, revealing that several immune cells, including B Cell, T Cell, and macrophage were associated with the risk score significantly. Subsequently, the risk score was highly associated with Type I and II IFN Response. The activation of IFN-I on liver cells controls glucose homeostasis and lipid metabolism which supports cell proliferation and tumorigenesis (51). Interestingly, the correlation between IRF-1, IRF-2, and PD-L1 was significantly positive. Overexpression of IRF-2 could down-regulate PD-L1 promoter activity and protein levels which was induced by IFN-γ (52). Anti-PD-1/PD-L1 therapy has improved outcomes in a range of advanced malignancies, including HCC, since its discovery. It is worth noting that although immunotherapy has many advantages in anti-cancer treatment, its efficacy shows strong individual variability (53). Our findings revealed a significant correlation between risk score and immunotherapy efficacy. A low-risk score indicates a better therapeutic effect of anti-PD-1/PD-L1 therapies. The prognostic model could effectively predict the suitability of HCC patients for anti-PD1/PDL1 immunotherapy, further supporting that FA metabolism is indispensable in shaping individual TIME characterizations.

Next, we chose eight risk genes including ANLN, UCK2, LPCAT1, TTK, KIF2C, HAVCR1, MMP1, and CBX2 based on multivariate Cox regression analyses above ($p < 0.05$). These eight genes RNA-seq paired sample data obtained from TCGA, showed higher mRNA levels in the HCC tissues compared to the normal tissues, and the ROC curve of these independent genes indicated a favorable predictive value over 1, 3, and 5 years. Later, qRT-PCR and Western blotting assays showed that mRNA and protein expression of TTK, KIF2C, HAVCR1, and MMP1 mRNA between tumors and normal tissues are significantly different. Threonine and tyrosine protein kinase (TTK), which is also known as monopolar spindle 1 (Mps1), acts as an oncogenic gene in a variety of cancers (54, 55). Previous studies indicate that KIF2C promotes the growth, invasion, and metastasis of HCC by mediating the Ras/MAPK signaling pathway (56). MMP1, which is an interstitial collagenase, has been implicated in the proliferation and metastasis in a variety of malignancies (57–59). HAVCR1 also known as T-cell immunoglobulin mucin domains (TIM)-1, is overexpressed in renal cell carcinoma (60), human colorectal cancer (61), and gastric adenocarcinomas (62), promoting the occurrence and progression of tumors. Mori ever reported HAVCR1 could mediate FA uptake to promote progress of kidney disease (63). Here, we focused on the HAVCR1, which shows the most significant difference in expression between normal liver cells and HCC cell lines. Functional experiments found that siRNA-induced HAVCR1 silencing robustly inhibited HCC cell growth, proliferation, migration, and invasion, suggesting that HAVCR1 could play an oncogenic role in HCC, however, the detailed underlying molecular mechanisms need to be further elucidated.

There are some unique superiorities in our research. Specifically, our study systematically evaluated the expression and prognostic value of FA metabolism-related genes in HCC. A better prognostic model consisting of 16 genes was established and further validated in the ICGC dataset. We found that HCC patients with high-risk scores had significantly poor prognoses and were highly correlated with clinicopathologic parameters. Notably, our model predicts patient survival with higher accuracy than previous models. The model genes were significantly enriched in most signaling oncogenic pathways and were significantly associated with tumor immunity and can predict the efficacy of immunotherapy. In addition, we studied the effect of HAVCR1 on the biological functions of HCC cells, finding that HAVCR1 silencing inhibits HCC cell growth, proliferation, and motility.

Exploring the prognostic value of FA metabolism-related genes also lays a foundation for our future mechanism research. Nevertheless, there are some limitations of our study that must be considered. Firstly, all analyses were performed using TCGA and ICGC databases, and more clinical patient data is needed to confirm its accuracy. Secondly, further experiments are required to investigate the correlation between our prognostic model and the tumor microenvironment. Finally, we only have preliminarily explored the effects of HAVCR1 on HCC cell functions by siRNA silencing, more genetic modifications need to be performed to further confirm the role of HAVCR1 on HCC cells, and the underlying molecular mechanisms need to be further elucidated.

## Conclusion

We constructed a FA metabolism-related prognostic model of genes that possessed predictive accuracy based on the data of the

LIHC cohort available in the TCGA database. The prognostic model was also significantly associated with tumor immunity and can predict the efficacy of anti-PD-1 immunotherapy. HAVCR1, a gene highly related to FA metabolism, has been proven to promote the growth, proliferation, migration, and invasion of HCC cells. The FA metabolism-related signatures could provide further possibilities to predict the progression and prognosis. Our study provides a novel idea for future research on personalized treatment strategies for HCC patients.

## Data availability statement

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

## Ethics statement

This study was approved by Ethics Committee of Affiliated Kunshan Hospital of Jiangsu University.

## Author contributions

X-RZ and J-QZ performed formal analysis and drafted the manuscript. Y-FC is responsible for project management. J-JL and JS participated in software and data analysis. Y-YL and S-QP conducted the experiment. Article was written, edited, and reviewed by X-RZ and Y-YL. M-BC and Y-PD guided and revised the manuscript. All authors read and approved the final submission. X-RZ, J-QZ and Y-FC contributed equally to this article.

## Funding

This project was supported by grants from the National Natural Science Foundation of China (82072712) and Suzhou Science and Technology plan project (KJXW2019064).

## Acknowledgments

We thank all our authors listed in 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

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.2022.972744/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Overall design of the study.

SUPPLEMENTARY FIGURE 2
Association Between Risk Genes and Clinicopathologic Parameters. **(A)** Heatmap indicating the links between clinicopathological parameters and different risk groups. PCA analysis. **(B)** Risk scores are classified by cluster, gender, stage and T stage. KM analysis of the OS based on gender **(C)**, age **(D)**, Stage **(E)**, T stage **(F)** (*p < 0.05, **p < 0.01, ***p < 0.001).

SUPPLEMENTARY FIGURE 3
Validation of The Risk Signature in ICGC-LIRI Cohort. **(A)** The volcano plot shows 16 risk genes in the ICGC-LIRI cohort. **(B)** Survival status distribution in ICGC-LIRI cohort. **(C)** KM analysis of the OS based on risk scores in the ICGC-LIRI cohort. The ROC curve **(D)** and PCA analysis **(E)** in the ICGC-LIRI cohort. (F) KM analyses based on different clinical parameters (*p < 0.05, **p < 0.01, ***p < 0.001).

SUPPLEMENTARY FIGURE 4
The Potential Molecular Mechanism of the Model. GSEA in low-risk groups **(A)** and high-risk groups **(B)**. **(C)** GVSA of 16 risk genes (*p < 0.05, **p < 0.01, ***p < 0.001).

SUPPLEMENTARY FIGURE 5
Single cell analysis of different genes expression.

SUPPLEMENTARY FIGURE 6
KM analysis comparing the high and low expression of HAVCR1 in HCC in the ICGC-LIRI cohort.

# References


1. McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of hepatocellular carcinoma. *Hepatology* (2021) 73 Suppl 1:4–13. doi: 10.1002/hep.31288

2. Craig AJ, von Felden J, Garcia-Lezana T, Sarcognato S, Villanueva A. Tumour evolution in hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol* (2020) 17 (3):139–52. doi: 10.1038/s41575-019-0229-4

3. Wang XK, Liao XW, Zhou X, Han CY, Chen ZJ, Yang CK, et al. Oncogene UBE2I enhances cellular invasion, migration and proliferation abilities *via* autophagy-related pathway resulting in poor prognosis in hepatocellular carcinoma. *Am J Cancer Res* (2020) 10(12):4178–97.

4. Ding J, He X, Luo W, Zhou W, Chen R, Cao G, et al. Development and validation of a pyroptosis-related signature for predicting prognosis in hepatocellular carcinoma. *Front Genet* (2022) 13:801419. doi: 10.3389/fgene.2022.801419

5. Li J, Tao H, Wang W, Li J, Zhang E. The detection and verification of two heterogeneous subgroups and a risk model based on ferroptosis-related genes in hepatocellular carcinoma. *J Oncol* (2022) 2022:1182383. doi: 10.1155/2022/1182383

6. Yin L, Zhou L, Gao S, Feng Y, Zhu H, Xiang J, et al. Classification of hepatocellular carcinoma based on N6-Methylandenosine-Related lncRNAs profiling. *Front Mol Biosci* (2022) 9:807418. doi: 10.3389/fmolb.2022.807418

7. Fu XW, Song CQ. Identification and validation of pyroptosis-related gene signature to predict prognosis and reveal immune infiltration in hepatocellular carcinoma. *Front Cell Dev Biol* (2021) 9:748039. doi: 10.3389/fcell.2021.748039

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

9. Ding C, Shan Z, Li M, Chen H, Li X, Jin Z. Characterization of the fatty acid metabolism in colorectal cancer to guide clinical therapy. *Mol Ther Oncolytics* (2021) 20:532–44. doi: 10.1016/j.omto.2021.02.010

10. Wu D, Yang Y, Hou Y, Zhao Z, Liang N, Yuan P, et al. Increased mitochondrial fission drives the reprogramming of fatty acid metabolism in hepatocellular carcinoma cells through suppression of sirtuin 1. *Cancer Commun (Lond)* (2022) 42(1):37–55. doi: 10.1002/cac2.12247

11. Sheng Y, Chen Y, Zeng Z, Wu W, Wang J, Ma Y, et al. Identification of pyruvate carboxylase as the cellular target of natural bibenzyls with potent anticancer activity against hepatocellular carcinoma *via* metabolic reprogramming. *J Med Chem* (2022) 65(1):460–84. doi: 10.1021/acs.jmedchem.1c01605

12. Martin SD, McGee SL. Metabolic reprogramming in type 2 diabetes and the development of breast cancer. *J Endocrinol* (2018) 237(2):R35–46. doi: 10.1530/JOE-18-0037

13. Munir R, Lisec J, Swinnen JV, Zaidi N. Too complex to fail? targeting fatty acid metabolism for cancer therapy. *Prog Lipid Res* (2022) 85:101143. doi: 10.1016/j.plipres.2021.101143

14. Dong SR, Ju XL, Yang WZ. STAT5A reprograms fatty acid metabolism and promotes tumorigenesis of gastric cancer cells. *Eur Rev Med Pharmacol Sci* (2019) 23(19):8360–70. doi: 10.26355/eurrev_201910_19147

15. Yoon H, Lee S. Fatty acid metabolism in ovarian cancer: Therapeutic implications. *Int J Mol Sci* (2022) 23(4):2170. doi: 10.3390/ijms23042170

16. Wang M, Han J, Xing H, Zhang H, Li Z, Liang L, et al. Dysregulated fatty acid metabolism in hepatocellular carcinoma. *Hepat Oncol* (2016) 3(4):241–51. doi: 10.2217/hep-2016-0012

17. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* (2013) 6(269):pl1. doi: 10.1126/scisignal.2004088

18. Digre A, Lindskog C. The human protein atlas-spatial localization of the human proteome in health and disease. *Protein Sci* (2021) 30(1):218–33. doi: 10.1002/pro.3987

19. Ponten F, Jirstrom K, Uhlen M. The human protein atlas–a tool for pathology. *J Pathol* (2008) 216(4):387–93. doi: 10.1002/path.2440

20. Yang X, Zhao Y, Shao Q, Jiang G. Cytochrome b561 serves as a potential prognostic biomarker and target for breast cancer. *Int J Gen Med* (2021) 14:10447–64. doi: 10.2147/IJGM.S338878

21. Wilkerson MD, Hayes DN. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26 (12):1572–3. doi: 10.1093/bioinformatics/btq170

22. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an r package for comparing biological themes among gene clusters. *OMICS* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118

23. Wang S, Yang L, Ci B, Maclean M, Gerber DE, Xiao G, et al. Development and validation of a nomogram prognostic model for SCLC patients. *J Thorac Oncol* (2018) 13(9):1338–48. doi: 10.1016/j.jtho.2018.05.037

24. Wu J, Li L, Zhang H, Zhao Y, Zhang H, Wu S, et al. A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma. *Oncogene* (2021) 40 (26):4413–24. doi: 10.1038/s41388-021-01853-y

25. Sun D, Wang J, Han Y, Dong X, Ge J, Zheng R, et al. TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. *Nucleic Acids Res* (2021) 49(D1): D1420–30. doi: 10.1093/nar/gkaa1020

26. Zhu XR, Peng SQ, Wang L, Chen XY, Feng CX, Liu YY, et al. Identification of phosphoenolpyruvate carboxykinase 1 as a potential therapeutic target for pancreatic cancer. *Cell Death Dis* (2021) 12(10):918. doi: 10.1038/s41419-021-04201-w

27. Wang ZJ, Jiang Z, Chen XN. SLC17A2 expression correlates with prognosis and immune infiltrates in hepatocellular carcinoma. *Comb Chem High Throughput Screen* (2022). doi: 10.21203/rs.3.rs-573999/v1

28. Aljomah G, Baker SS, Liu W, Kozielski R, Oluwole J, Lupu B, et al. Induction of CYP2E1 in non-alcoholic fatty liver diseases. *Exp Mol Pathol* (2015) 99(3):677–81. doi: 10.1016/j.yexmp.2015.11.008

29. Zhang T, Zhang Y, Liu J, Ma Y, Ye Q, Yan X, et al. MicroRNA-377-3p inhibits hepatocellular carcinoma growth and metastasis through negative regulation of CPT1C-mediated fatty acid oxidation. *Cancer Metab* (2022) 10 (1):2. doi: 10.1186/s40170-021-00276-3

30. Wang H, Zhou Y, Xu H, Wang X, Zhang Y, Shang R, et al. Therapeutic efficacy of FASN inhibition in preclinical models of HCC. *Hepatology* (2022). doi: 10.1002/hep.32359

31. Attal N, Sullivan MT, Girardi CA, Thompson KJ, McKillop IH. Fatty acid binding protein-4 promotes alcohol-dependent hepatosteatosis and hepatocellular carcinoma progression. *Transl Oncol* (2021) 14(1):100975. doi: 10.1016/j.tranon.2020.100975

32. Luo X, Zheng E, Wei L, Zeng H, Qin H, Zhang X, et al. The fatty acid receptor CD36 promotes HCC progression through activating Src/PI3K/AKT axis-dependent aerobic glycolysis. *Cell Death Dis* (2021) 12(4):328. doi: 10.1038/s41419-021-03596-w

33. Ma APY, Yeung CLS, Tey SK, Mao X, Wong SWK, Ng TH, et al. Suppression of ACADM-mediated fatty acid oxidation promotes hepatocellular carcinoma *via* aberrant CAV1/SREBP1 signaling. *Cancer Res* (2021) 81(13):3679–92. doi: 10.1158/0008-5472.CAN-20-3944

34. Liu K, Cui JJ, Zhan Y, Ouyang QY, Lu QS, Yang DH, et al. Reprogramming the tumor microenvironment by genome editing for precision cancer therapy. *Mol Cancer* (2022) 21(1):98. doi: 10.1186/s12943-022-01561-5

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

36. Vickers AJ, Woo S. Decision curve analysis in the evaluation of radiology research. *Eur Radiol* (2022) 1–3. doi: 10.1007/s00330-022-08685-8

37. Wu X, Gu Z, Chen Y, Chen B, Chen W, Weng L, et al. Application of PD-1 blockade in cancer immunotherapy. *Comput Struct Biotechnol J* (2019) 17:661–74. doi: 10.1016/j.csbj.2019.03.006

38. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin* (2022) 72(1):7–33. doi: 10.3322/caac.21708

39. 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* (2022) 12(2):558–80. doi: 10.1016/j.apsb.2021.09.019

40. Huang Q, Chen Y, Lin K, Sun C, Zheng S, Chen J, et al. Redefining hepatocellular carcinoma staging systems based on the bile duct invasion status: A multicenter study. *Front Oncol* (2021) 11:673285. doi: 10.3389/fonc.2021.673285

41. Reig M, Forner A, Rimola J, Ferrer-Fabrega J, Burrel M, Garcia-Criado A, et al. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J Hepatol* (2022) 76(3):681–93. doi: 10.1016/j.jhep.2021.11.018

42. Sondergaard JN, Sommerauer C, Atanasoai I, Hinte LC, Geng K, Guiducci G, et al. CCT3-LINC00326 axis regulates hepatocarcinogenic lipid metabolism. *Gut* (2022). doi: 10.1136/gutjnl-2021-325109

43. Siddiqui S, Glauben R. Fatty acid metabolism in myeloid-derived suppressor cells and tumor-associated macrophages: Key factor in cancer immune evasion. *Cancers (Basel)* (2022) 14(1):250. doi: 10.3390/cancers14010250

44. Qi W, Zhu F, Wang M, Teng Z, Xu R, Xi Y, et al. The antitumoral effect of Paris saponin II on head and neck squamous cell carcinomas mediated *via* the nitric oxide metabolic pathway. *Front Cell Dev Biol* (2021) 9:803981. doi: 10.3389/fcell.2021.803981

45. Li J, Huang Q, Long X, Zhang J, Huang X, Aa J, et al. CD147 reprograms fatty acid metabolism in hepatocellular carcinoma cells through Akt/mTOR/SREBP1c and P38/PPARalpha pathways. *J Hepatol* (2015) 63(6):1378–89. doi: 10.1016/j.jhep.2015.07.039

46. He D, Cai L, Huang W, Weng Q, Lin X, You M, et al. Prognostic value of fatty acid metabolism-related genes in patients with hepatocellular carcinoma. *Aging (Albany NY)* (2021) 13(13):17847–63. doi: 10.18632/aging.203288

47. Cui XH, Peng QJ, Li RZ, Lyu XJ, Zhu CF, Qin XH. Cell division cycle associated 8: A novel diagnostic and prognostic biomarker for hepatocellular carcinoma. *J Cell Mol Med* (2021) 25(24):11097–112. doi: 10.1111/jcmm.17032

48. Guo Y, Wang J, Benedict B, Yang C, van Gemert F, Ma X, et al. Targeting CDC7 potentiates ATR-CHK1 signaling inhibition through induction of DNA replication stress in liver cancer. *Genome Med* (2021) 13(1):166. doi: 10.1186/s13073-021-00981-0

49. Lochner M, Berod L, Sparwasser T. Fatty acid metabolism in the regulation of T cell function. *Trends Immunol* (2015) 36(2):81–91. doi: 10.1016/j.it.2014.12.005

50. Zhou X, Zhu X, Zeng H. Fatty acid metabolism in adaptive immunity. *FEBS J* (2021). doi: 10.1111/febs.16296

51. Wu MS, Kuo YP, Lo YC, Tsai DJ, Lai CY, Chuang TH, et al. Type I interferon signaling accelerates liver regeneration by metabolic modulation in noninfectious conditions. *Am J Pathol* (2021) 191(6):1036–48. doi: 10.1016/j.ajpath.2021.03.006

52. Yan Y, Zheng L, Du Q, Yan B, Geller DA. Interferon regulatory factor 1 (IRF-1) and IRF-2 regulate PD-L1 expression in hepatocellular carcinoma (HCC) cells. *Cancer Immunol Immunother* (2020) 69(9):1891–903. doi: 10.1007/s00262-020-02586-9

53. Pare L, Pascual T, Segui E, Teixido C, Gonzalez-Cao M, Galvan P, et al. Association between PD1 mRNA and response to anti-PD1 monotherapy across multiple cancer types. *Ann Oncol* (2018) 29(10):2121–8. doi: 10.1093/annonc/mdy335

54. Suyal G, Pandey P, Saraya A, Sharma R. Tumour suppressor role of microRNA-335-5p in esophageal squamous cell carcinoma by targeting TTK (Mps1). *Exp Mol Pathol* (2022) 124:104738. doi: 10.1016/j.yexmp.2021.104738

55. Xia P, Liang J, Jin D, Jin Z. Reversine inhibits proliferation, invasion and migration and induces cell apoptosis in gastric cancer cells by downregulating TTK. *Exp Ther Med* (2021) 22(3):929. doi: 10.3892/etm.2021.10361

56. Mo S, Fang D, Zhao S, Thai Hoa PT, Zhou C, Liang T, et al. Down regulated oncogene KIF2C inhibits growth, invasion, and metastasis of hepatocellular carcinoma through the Ras/MAPK signaling pathway and epithelial-to-mesenchymal transition. *Ann Transl Med* (2022) 10(3):151. doi: 10.21037/atm-21-6240

57. Ng L, Wong SK, Huang Z, Lam CS, Chow AK, Foo DC, et al. CD26 induces colorectal cancer angiogenesis and metastasis through CAV1/MMP1 signaling. *Int J Mol Sci* (2022) 23(3):1181. doi: 10.3390/ijms23031181

58. Gabasa M, Radisky ES, Ikemori R, Bertolini G, Arshakyan M, Hockla A, et al. MMP1 drives tumor progression in large cell carcinoma of the lung through fibroblast senescence. *Cancer Lett* (2021) 507:1–12. doi: 10.1016/j.canlet.2021.01.028

59. 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

60. Cuadros T, Trilla E, Sarro E, Vila MR, Vilardell J, de Torres I, et al. HAVCR/KIM-1 activates the IL-6/STAT-3 pathway in clear cell renal cell carcinoma and determines tumor progression and patient outcome. *Cancer Res* (2014) 74(5):1416–28. doi: 10.1158/0008-5472.CAN-13-1671

61. Wang Y, Martin TA, Jiang WG. HAVcR-1 expression in human colorectal cancer and its effects on colorectal cancer cells *in vitro*. *Anticancer Res* (2013) 33(1):207–14.

62. Xue J, Li Y, Yi J, Jiang H. HAVCR1 affects the MEK/ERK pathway in gastric adenocarcinomas and influences tumor progression and patient outcome. *Gastroenterol Res Pract* (2019) 2019:6746970. doi: 10.1155/2019/6746970

63. Mori Y, Ajay AK, Chang JH, Mou S, Zhao H, Kishi S, et al. KIM-1 mediates fatty acid uptake by renal tubular cells to promote progressive diabetic kidney disease. *Cell Metab* (2021) 33(5):1042–1061 e1047. doi: 10.1016/j.cmet.2021.04.004

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 09 August 2022
DOI 10.3389/fonc.2022.952521

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Limin Xu,
Huzhou Central Hospital, China
Shengbao Wang,
Lanzhou University Second Hospital, China

*CORRESPONDENCE
Zhiqing Lin
zhiqinglinly@outlook.com

†These authors share first authorship

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

RECEIVED 25 May 2022
ACCEPTED 30 June 2022
PUBLISHED 09 August 2022

CITATION
Zhou S, Wang H, Huang Y, Wu Y and Lin Z (2022) The global change of gene expression pattern caused by PTEN mutation affects the prognosis of glioblastoma.
*Front. Oncol.* 12:952521.
doi: 10.3389/fonc.2022.952521

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

# The global change of gene expression pattern caused by PTEN mutation affects the prognosis of glioblastoma

Shengjun Zhou†, Haifeng Wang†, Yi Huang, Yiwen Wu and Zhiqing Lin*

Department of Neurosurgery, Ningbo City First Hospital, Ningbo, China

Glioblastoma (GBM), an aggressive primary tumor, is common in humans, accounting for 12–15% of all intracranial tumors, and has median survival of fewer than 15 months. Since a growing body of evidence suggests that conventional drugs are ineffective against GBM, our goal is to find emerging therapies that play a role in its treatment. This research constructs a risk model to predict the prognosis of GBM patients. A set of genes associated with GBM was taken from a GBM gene data bank, and clinical information on patients with GBM was retrieved from the Cancer Genome Atlas (TCGA) data bank. One-way Cox and Kaplan–Meier analyses were performed to identify genes in relation to prognosis. Groups were classified into high and low expression level of PTEN expression. Prognosis-related genes were further identified, and multi-factor Cox regression analysis was used to build risk score equations for the prognostic model to construct a survival prognostic model. The area under the ROC curve suggested that the pattern had high accuracy. When combined with nomogram analysis, GJB2 was considered an independent predictor of GBM prognosis. This study provides a potential prognostic predictive biological marker for GBM patients and confirms that GJB2 is a key gene for GBM progression.

KEYWORDS

**prognosis, LASSO, pan-cancer analysis, GBM, glioblastoma**

## Introduction

Glioblastoma (GBM) is an aggressive primary tumor that is the most aggressive and common type of brain tumor in humans (1), with a median survival of fewer than 15 months (2). It consists of highly malignant cells with metastatic and angiogenic properties that lead to resistance to agents such as temozolomide, although none of these agents significantly improves overall survival (3). The conventional treatment is a combination of chemotherapy and surgical resection followed by radiotherapy and adjuvant chemotherapy. This regimen

has been effective in improving overall survival but largely fails to prevent recurrence because surgical treatment fails to completely eradicate GBM cells, which remain surrounded by scattered GBM-infiltrating cells (4). Therefore, the identification of reliable prognostic markers has become crucial for GBM treatment.

Phosphatase and tensin homolog (PTEN), a tumor suppressor gene, is closely involved in cell translation, proliferation, and tumorigenesis. PTEN gene mutations are frequently found in the genetic landscape of high-grade gliomas, and is hallmarks of glioma malignancy,they influence cell proliferation, proangiogenetic pathways, and antitumoral immune response (5).This study evaluated the prognosis and expression of GBM patients by analyzing the mutation spectrum of PTEN-related microenvironment.

We collected a set of genes related to GBM. Candidate genes were obtained by differentially expressed genes. One-way Cox and Kaplan–Meier analyses were performed to identify the genes. Combined with LASSO regression analysis to establish prognostic features and further identify prognosis-related genes, a survival prognostic model was constructed, and the area under the ROC curve suggested that the model had high accuracy. Combined with nomogram analysis, the prognostic significance of genetic features in GBM was assessed.

# Materials and method

## Data acquisition

Expression matrices of all tumors, including the clinically relevant pathological features of tumor tissue samples, were obtained from the Cancer Genome Atlas (TCGA) database and evaluated based on the data completeness of clinical samples and degree of matching with sequenced samples. Duplicate and censored samples and cases without clinical findings were excluded.

## Variance analysis

The Limma package of R software can screen for differentially expressed genes between GBM and para cancer with |logFC| = 0.3785 (adj. *P*-value <0.05) and plotted volcanoes.To further confirm the potential functions of potential targets,the data were analyzed by functional enrichment. ClusterProfiler program package in R software was used to analyze the GO Function of potential mRNA and KEGG pathway gene ontology (GO) containing Molecular functions (MF), Biological process (BP) and Cellular Component (CC).

## Kaplan–Meier survival analysis

The effect of genetic characteristics on prognosis was verified by One-way Cox. Kaplan–Meier survival curves were plotted to compare the survival of patients in the high and low expression level groups.

## Establishment and analysis of risk prediction models

LASSO regression analysis was applied to narrow the range of prognosis-related genes and ensure the stability of the results. The median risk score of each sample was calculated using the risk score formula as the threshold, and patients were classified into high expression level and low expression level groups according to the expression level of PTEN. Survival curves were plotted. In addition, the working curves (ROC) of the subjects were plotted, and the area under the curve (AUC) was calculated to assess the predictive validity of the model.

## Nomogram analysis

Independent prognostic factors for GBM were determined by univariate and multifactorial analyses. Column plots and calibration curve plots predicted the predictive power of survival at 1, 2, and 3 years.

## Pan-cancer analysis

Data from normal and TCGA tumor tissues in the GTEx database were combined to analyze the differences in their gene expression. The amount of tumor mutations in each tumor sample was counted separately, and the relationship between gene expression and tumor mutational burden was analyzed.

# Results

## Mutation landscape

PTEN gene mutation data, clinical data, and transcriptome data from the TCGA data bank and somatic mutations in GBM patients were downloaded. Mutation analysis revealed that PTEN had a high mutation rate of 30.03% in GBM, which was the highest among TCGA tumors (Figure 1A). PTEN mutation frequency ranked first (34.1%) among all mutated genes in GBM (Figure 1B).

## Differentially expressed gene screening

In | logFC | = 0.58 under the condition of the filter, identified the genes from the TCGA - GBM data too little number, change | logFC | = 0.3785, identified 45 differentially expressed genes

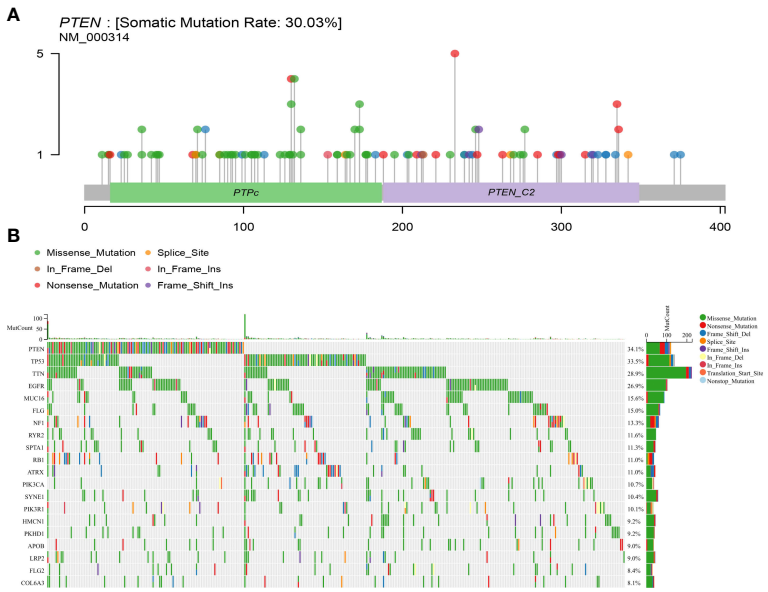



FIGURE 1
Somatic mutation rate and mutation landscape map of PTEN in GBM. **(A)** Lollipop plot of PTEN mutation distribution in GBM; **(B)** oncoplot showing the somatic landscape of the GBM cohort.

(adj. P-value <0.05). There were 5 down-regulated genes and 40 up-regulated genes (Figures 2A, B). The results of differentially up-regulated genes and down-regulated genes, KEGG pathway enrichment, and GO term enrichment are shown in Figure 2C.

## Prognostic analysis

To understand the effect of each gene on prognosis, one-way Cox regression was used to validate the patients into high and low expression level groups by scoring each gene's median expression level as the median value. It was further observed that only KIRREL2, TENT5B, DIRAS3, SDC1, GJB2, DDIT4L, HOXA10, and H2AW were prognostically significant in GBM (Figure 3A). The relationship between the prognosis of GBM and the expression levels of the above genes was further analyzed by plotting Kaplan–Meier survival curves. It was found that the higher the expression of DDIT4L, GJB2, KIRREL2, DIRAS3, HOXA10, SDC1, and TENT5B, the worse the prognosis, and the lower the expression of H2AW, the worse the prognosis (Figure 3B).

## Expression analysis

The expression of DDIT4L, DIRAS3, GJB2, H2AW, HOXA10, KIRREL2, SDC1, and TENT5B in the PTEN mutant group, PTEN wild group, and normal group were observed, and the box plot identified that DDIT4L, DIRAS3, GJB2, HOXA10, SDC1, and TENT5B were up-regulated in PTEN mutant group and PTEN wild group. H2AW and KIRREL2 were down-regulated in PTEN mutant group and PTEN wild group, compared with the normal group (Figure 4).

## Construction of the LASSO prognostic model

LASSO regression analysis was performed to further narrow the range of prognosis-related genes and ensure the stability of the results (Figures 5A, B). The risk score was calculated for each sample using the following risk score formula:

$$\text{lambda.min} = 0.0982$$

$$\text{Riskscore} = (0.0267) \times \text{SDC1} + (0.0704) \times \text{GJB2} + (0.0625) \times \text{DDIT4L} + (-0.0792) \times \text{H2AW}$$

The median expression level score was used as a threshold for dividing the patients into high and low expression level groups and plotting the survival curves, which showed that the survival prognosis of the high expression level group was significantly worse than that of the low expression level group (Figure 5C). The ROC curves of 1, 3, and 5-year survival of GBM patients predicted by this risk score model were plotted (Figure 5D), and their AUC areas were 0.694,0.651, and 0.763, respectively.

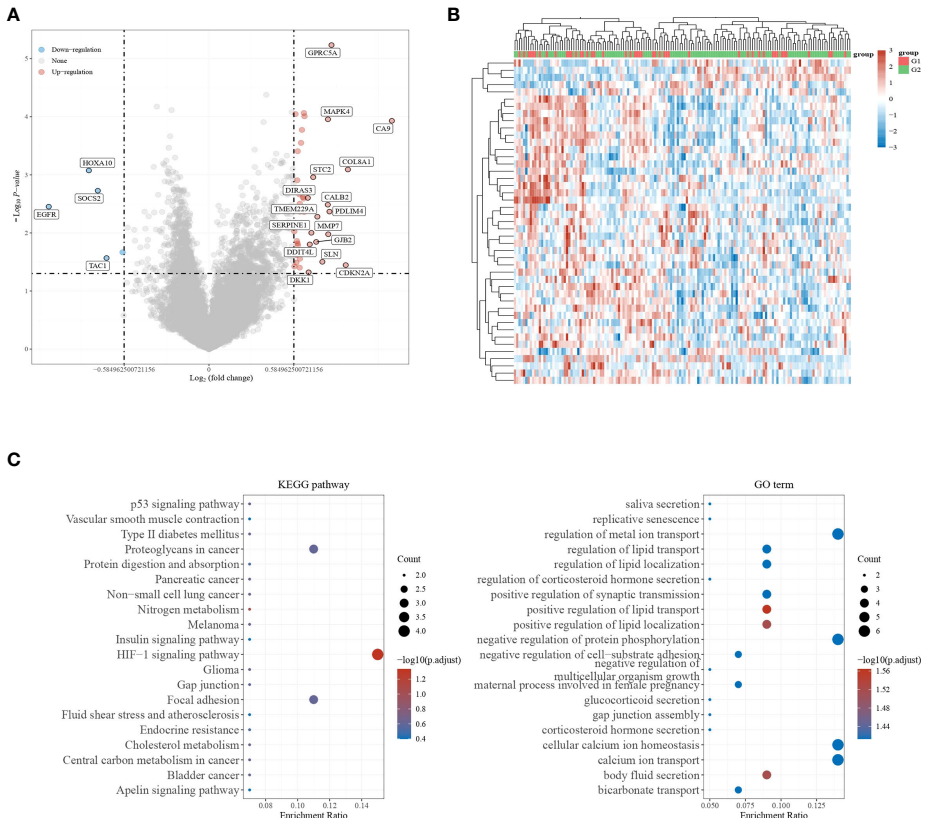



FIGURE 2
Screening for differentially expressed genes. **(A, B)** Volcano plot showing differentially expressed genes in GBM; **(C)** KEGG pathway enrichment results and GO term enrichment results.

## Nomogram analysis

The GJB2 gene was shown to be a prognostic factor for GBM by univariate and multifactorial analyses (Figures 6A, B). Column and calibration curve plots indicated the predictive power of survival at 1, 2, and 3 years. The C-index was 0.583 and the *P*-value was 0.02.

## Expression of GJB2 in GBM

The expression of GJB2 in tumors was analyzed by integrating the data from TCGA, and the results showed that GJB2 was highly expressed in BLCA, BRCA, CESC, COAD, DLBC, ESCA, GBM, HNSC, KIRC, KIRP, LGG, LIHC, LUAD, LUSC, OV, PAAD, PRAD, STAD, TGCT, THCA, UCEC, and UCS, and there was low expression in CHOL, KICH, and READ (Figure 7A).

One-way Cox regression showed prognostic significance of GJB2 in KICH and PCPG (Figure 7B).

The relationship between GJB2 gene expression and immune checkpoint gene expression was analyzed, and CD276 was highly correlated with multiple cancers (Figure 7C).

Immune mutation load correlation indicated that GJB2 was significantly associated with immune mutation load in CESC, COAD, KIRC, KIRP, LAML, LUAD, PAAD, SKCM, STAD, and UCEC (Figure 7D).

GJB2 was correlated with microsatellite instability in COAD, DLBC, ESCA, HNSC, KIRC, LIHC, and PRAD (Figure 7E).

# Discussion

As contemporary living standards and the material conditions of people continue to improve, health issues are receiving more and more attention. However, the incidence of cancer is increasing year by year, threatening human health. Among the cancers, GBM, which is classified as a grade IV diffuse glioma, is the primary tumor in adults and has a terrible prognosis. It has a high rate of recurrence and can spread rapidly to other parts of the brain (6), causing thousands of deaths worldwide each year. GBM consists of highly malignant cells that have metastatic and angiogenic properties that lead to resistance to agents such as temozolomide (3). The average survival of GBM patients remains below 20 months using current

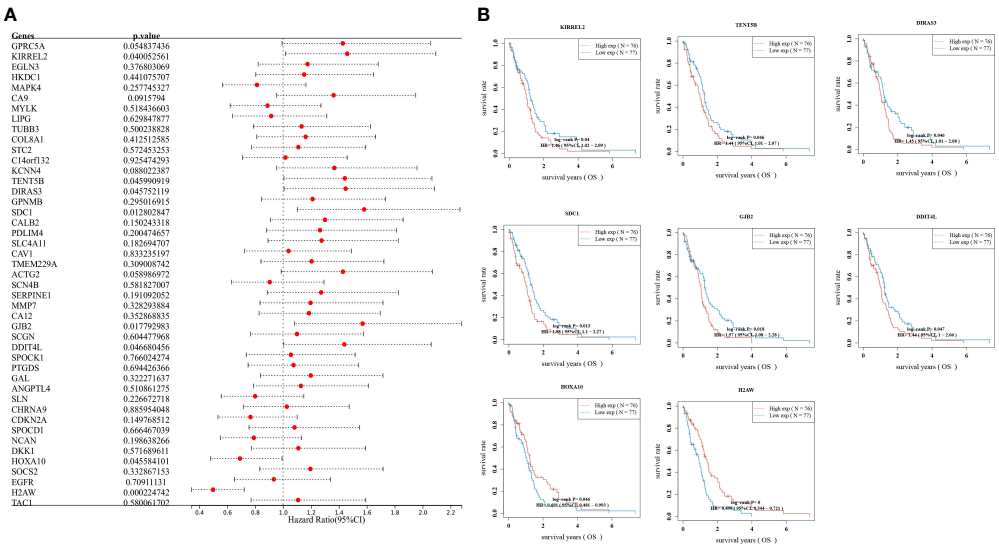



FIGURE 3
Prognostic analysis, **(A)** Validation of gene effects on prognosis using one-way Cox regression, presented using forest plots; **(B)** Kaplan–Meier survival analysis of the relationship between high and low expression levels of DDIT4L, DIRAS3, GJB2, H2AW, HOXA10, KIRREL2, SDC1, and TENT5B genes and GBM prognosis.

therapies (7), and conventional systemic chemotherapeutic agents used to treat GBM are ineffective. Several studies have found that the extremely high recurrence rate of GBM is associated with the expression of strongly proliferative genes in cells (8). Since this process usually involves multiple genes (9), we aimed to explore the prognostic genes that may play a role in the treatment of GBM.

PTEN, one of the most frequently mutated genes in human cancers (10), is a tumor suppressor with growth and survival regulatory functions (11). It controls many processes, including survival, proliferation, differentiation, energy metabolism, and deregulation of cell structure and mobility (12, 13). Deletion or mutation of PTEN leads to increased cell proliferation and decreased cell death and tumor development (14). Somatic loss-of-function mutations in PTEN can cause or drive the malignant phenotype of various human cancers (13), which is consistent with our findings. GJB2 is considered an oncogene and is related to tumor growth, EMT, and lymph node metastasis in a variety of cancers (15–18). Mutations in the GJB2 gene are a major cause of autosomal recessive hereditary nonsyndromic hearing loss (ARNSHL) in many populations (19). GJB2 expression is elevated in many tumor cell lines, tumor tissues (20), and breast cancer (21). It has been shown that GJB2 is an independent prognostic

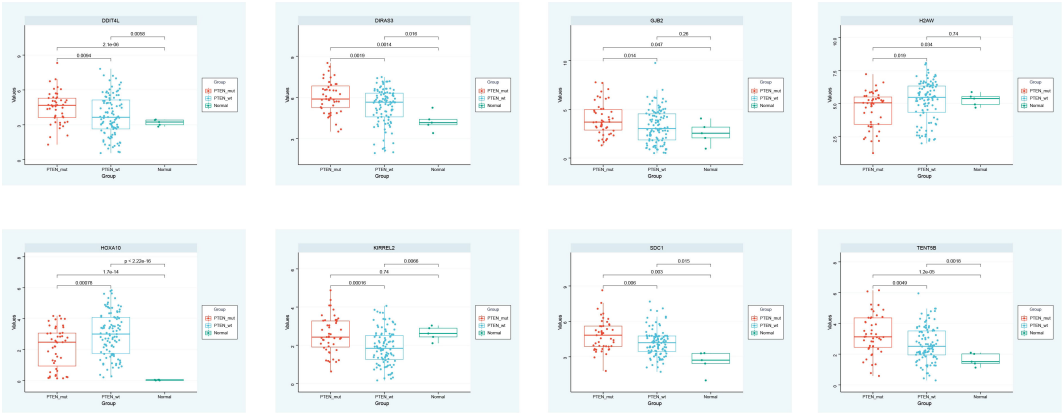



FIGURE 4
Expression analysis.

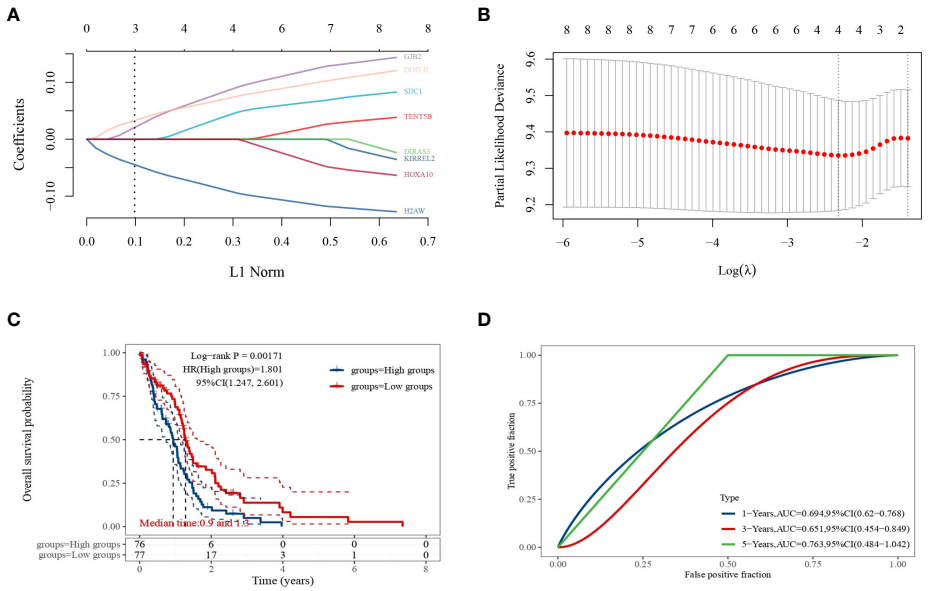



FIGURE 5
LASSO regression analysis. **(A)** Coefficients of selected features expressed as λ parameters; **(B)** partial likelihood deviation plotted with log(l) using LASSO regression analysis model; **(C)** Kaplan–Meier survival curves for patients in the high and low expression level groups; **(D)** time-dependent ROC curve plot.

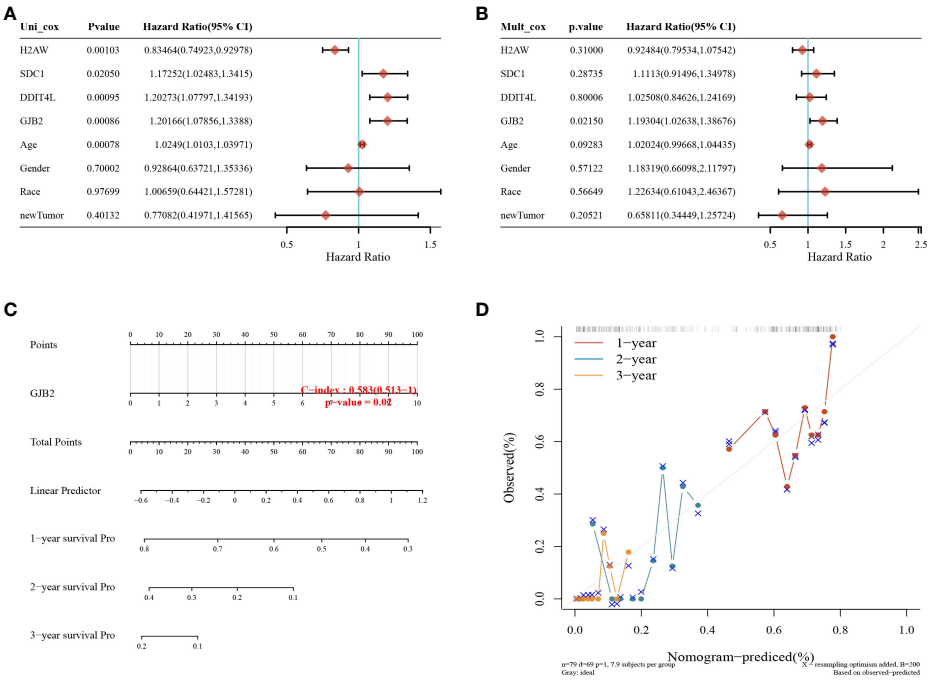



FIGURE 6
Nomogram analysis. **(A)** Univariate and **(B)** multifactorial Cox analyses showing that the GJB2 gene is an independent prognostic factor for GBM; **(C)** columnar plots predicting 1, 2, and 3-year overall survival in GBM patients; **(D)** calibration curves for the overall survival columnar plot model.

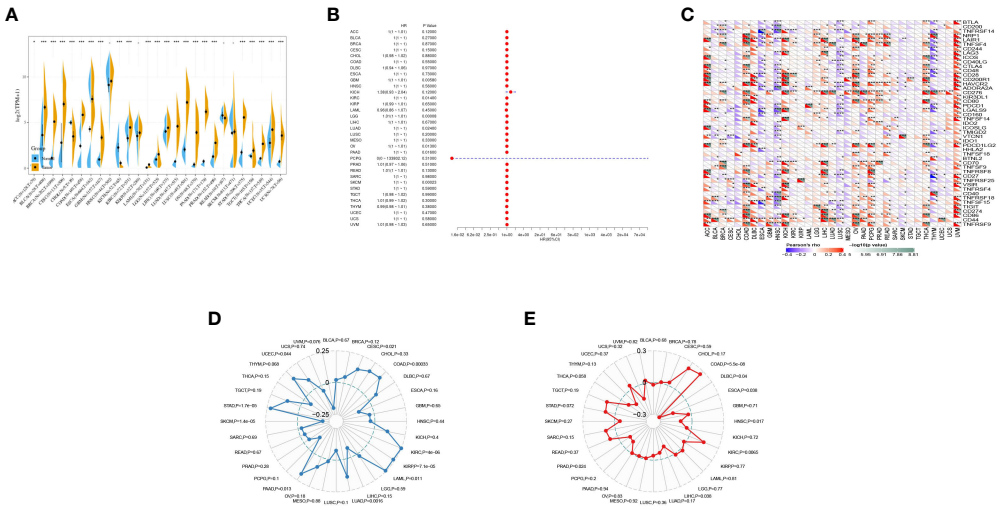



FIGURE 7
Pan-cancer analysis of GJB2 in multiple tumors. **(A)** Expression of GJB2 in multiple tumors; **(B)** prognostic analysis of GJB2 in multiple tumors; **(C)** correlation of GJB2 with immune checkpoints in multiple tumors; **(D)** correlation of GJB2 with immune mutation load in multiple tumors; **(E)** correlation of GJB2 with microsatellites in multiple tumors. *P<0.05, **P<0.01,***P<0.001.

biomarker for LUAD, and patients with GJB2 overexpression have shorter overall survival. GJB2 may also be a potential prognostic factor for KIRC, as shown by pan-cancer analysis (22).

We identified genes related to the prognosis of GBM, namely GJB2. We first screened differentially expressed genes by differential analysis and performed one-way Cox regression and Kaplan–Meier analyses to identify genes related to prognosis. The groups were classified into high and low expression level, based on PTEN expression. Combined with LASSO regression analysis to establish prognostic characteristics and further identify prognosis-related genes, multi-factor Cox regression analysis was used to establish risk score equations for the prognostic model to construct the survival prognostic model, and the area under the ROC curve proved that the model had high accuracy. Combined with nomogram analysis, GJB2 was considered an independent predictor of GBM prognosis.

Many studies on the genetic correlation of GBM prognosis have been conducted. For example, Yanxin Li et al. found that the prognosis of GBM patients was poorer when the HOXD10 gene was highly expressed, and HOXD10 may play different roles at different stages of GBM development (23). To the best of our knowledge, the GJB2 gene that we screened is a new GBM biomarker, and no previous reports of this gene associated with the development and progression of GBM have appeared.

In summary, this study provides a predictive biological marker for GBM patients and confirms GJB2 as a key gene for GBM progression. These conclusions may provide a direction for prognosis prediction and treatment of GBM patients.

## Data availability statement

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

## Funding

This study was supported by the grants from the Ningbo medical and health brand discipline (PPXK2018-04) and Ningbo Science and Technology Innovation 2025 Major Project (2020Z094, 2022Z125), Key Laboratory of Precision Medicine for Atherosclerotic Diseases of Zhejiang Province (2022E10026).

## Author contributions

We contributed equally for this work.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Yang Y, Tu Y, Lu J, Chen Q, Zhu Z, Peng W, et al. PT109, a novel multi-kinase inhibitor, suppresses glioblastoma multiforme through cell reprogramming: Involvement of PTBP1/PKM1/2 pathway. *Eur J Pharmacol* (2022) 23:174837. doi: 10.1016/j.ejphar.2022.174837

2. Shergalis A, Bankhead A 3rd, Luesakul U, Muangsin N, Neamati N. Current challenges and opportunities in treating glioblastoma. *Pharmacol Rev* (2018) 703:412–45. doi: 10.1124/pr.117.014944

3. Carlsson SK, Brothers SP, Wahlestedt C. Emerging cellular therapies for glioblastoma multiforme. *Cureus* (2018) 10(3):e2305. doi: 10.7759/cureus.2305

4. Vasilev A, Sofi R, Rahman R, Smith SJ, Teschemacher AG, Kasparov S, et al. Using light for therapy of glioblastoma multiforme (GBM). *Brain Sci* (2020) 10 (3):75. doi: 10.3390/brainsci10020075

5. Giotta Lucifero A, Luzzi S. Immune landscape in PTEN-related glioma microenvironment: A bioinformatic analysis. *Brain Sci* (2022) 12(4):501. doi: 10.3390/brainsci12040501

6. Batash R, Asna N, Schaffer P, Francis N, Schaffer M, et al. Glioblastoma multiforme, diagnosis and treatment; recent literature review. *Curr Medicin Chem* (2017) 24(27):3002–9. doi: 10.2174/0929867324666170516123206

7. Wang S, Yao F, Lu X, Li Q, Su Z, Lee JH, et al. Temozolomide promotes immune escape of GBM cells *via* upregulating PD-L1. *Am J Cancer Res* (2019) 9 (6):1161–71.

8. Lara-Velazquez M, Zarco N, Carrano A, Phillipps J, Norton ES, Schiapparelli P, et al. Alpha 1-antichymotrypsin contributes to stem cell characteristics and enhances tumorigenicity of glioblastoma. *Neuro-oncology* (2021) 23(4):599–610. doi: 10.1093/neuonc/noaa264

9. Malik V, Garg S, Afzal S, Dhanjal JK, Yun CO, Kaul SC, et al. Bioinformatics and molecular insights to anti-metastasis activity of triethylene glycol derivatives. *Int J Mol Sci* (2020) 21(15):5463. doi: 10.3390/ijms21155463

10. Zheng L, Liang H, Zhang Q, Shen Z, Sun Y, Zhao X, et al. circPTEN1, a circular RNA generated from PTEN, suppresses cancer progression through inhibition of TGF-β/Smad signaling. *Mol Cancer* (2022) 21(1):41. doi: 10.1186/s12943-022-01495-y

11. Chen C-Y, Chen J, He L, Stiles BL. PTEN: Tumor suppressor and metabolic regulator. *Front Endocrinol* (2018) 9:338. doi: 10.3389/fendo.2018.00338

12. Chow JT, Salmena L. Recent advances in PTEN signalling axes in cancer. *Facult Rev* (2020) 9:31. doi: 10.12703/r/9-31

13. Chen L, Deyin G. The functions of tumor suppressor PTEN in innate and adaptive immunity. *Cell Mol Immunol* (2017) 14(7):581–9. doi: 10.1038/cmi.2017.30

14. Song MS, Salmena L, Pandolfi PP. The functions and regulation of the PTEN tumour suppressor. *Nat Rev Mol Cell Biol* (2012) 13(5):283–96. doi: 10.1038/nrm3330

15. Ezumi K, Yamamoto H, Murata K, Higashiyama M, Damdinsuren B, Nakamura Y, et al. Aberrant expression of connexin 26 is associated with lung metastasis of colorectal cancer. *Clin Cancer Res* (2008) 14(3):677–84. doi: 10.1158/1078-0432.CCR-07-1184

16. Inose T, Kato H, Kimura H, Faried A, Tanaka N, Sakai M, et al. Correlation between connexin 26 expression and poor prognosis of esophageal squamous cell carcinoma. *Ann Surg Oncol* (2009) 16(6):1704–10. doi: 10.1245/s10434-009-0443-3

17. Naoi Y, Miyoshi Y, Taguchi T, Kim SJ, Arai T, Maruyama N, et al. Connexin26 expression is associated with aggressive phenotype in human papillary and follicular thyroid cancers. *Cancer Lett* (2008) 262(2):248–56. doi: 10.1016/j.canlet.2007.12.008

18. Wu J-I, Wang LH. Emerging roles of gap junction proteins connexins in cancer metastasis, chemoresistance and clinical application. *J Biomed Sci* (2019) 26 (1):8. doi: 10.1186/s12929-019-0497-x

19. Koohiyan M, Koohian F, Azadegan-Dehkordi F. GJB2-related hearing loss in central Iran: Review of the spectrum and frequency of gene mutations. *Ann Hum Genet* (2020) 84(2):107–13. doi: 10.1111/ahg.12354

20. Zhu T, Gao YF, Chen YX, Wang ZB, Yin JY, Mao XY, et al. Genome-scale analysis identifies GJB2 and ERO1LB as prognosis markers in patients with pancreatic cancer. *Oncotarget* (2017) 8(13):21281–9. doi: 10.18632/oncotarget.15068

21. Shettar A, Damineni S, Mukherjee G, Kondaiah P. Gap junction β- 2 expression is negatively associated with the estrogen receptor status in breast cancer tissues and is a regulator of breast tumorigenesis. *Oncol Rep* (2018) 40 (6):3645–53. doi: 10.3892/or.2018.6764

22. Lu Ao, Shi Y, Liu Y, Lin J, Zhang H, Guo Y, et al. Integrative analyses identified ion channel genes GJB2 and SCNN1B as prognostic biomarkers and therapeutic targets for lung adenocarcinoma. *Lung Cancer (Amsterd Netherl)* (2021) 158:29–39. doi: 10.1016/j.lungcan.2021.06.001

23. Li Y, Ma K, Xie Q, Zhang X, Zhang X, Chen K, et al. Identification of HOXD10 as a marker of poor prognosis in glioblastoma multiforme. *OncoTarget Ther* (2021) 14:5183–95. doi: 10.2147/OTT.S336225

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 August 2022
DOI 10.3389/fonc.2022.918230

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Xu Jia,
Affiliated Hospital of Guizhou Medical University, China
Yuxuan Song,
Peking University People's Hospital, China
Svetlana Saakyan,
Helmholtz Moscow Research Institute of Eye Diseases (NMITS GB), Russia

*CORRESPONDENCE
Sha Wang
wangsha_1982@csu.edu.cn

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

RECEIVED 12 April 2022
ACCEPTED 24 June 2022
PUBLISHED 16 August 2022

CITATION
Wang W and Wang S (2022) The prognostic value of immune-related genes *AZGP1*, *SLCO5A1*, and *CTF1* in Uveal melanoma.
*Front. Oncol.* 12:918230.
doi: 10.3389/fonc.2022.918230

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

# The prognostic value of immune-related genes *AZGP1*, *SLCO5A1*, and *CTF1* in Uveal melanoma

Wanpeng Wang[1,2,3] and Sha Wang[1,2,3]*

[1]Eye Center of Xiangya Hospital, Central South University, Changsha, China, [2]Hunan Key Laboratory of Ophthalmology, Changsha, China, [3]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Changsha, China

**Objective:** Uveal melanoma (UM) is an aggressive malignancy with a poor prognosis and no available effective treatment. Therefore, exploring a potential prognostic marker for UM could provide new possibilities for early detection, recurrence, and treatment.

**Methods:** In this study, we used "ConsensusClusterPlus" to classify patients with UM into subgroups, screened for significant differences in immune prognostic factors between subgroups, selected three genes using LASSO (Least absolute shrinkage and selection operator) regression to construct a risk model, and performed tumor immune cell infiltration analysis on the risk model. infiltration analysis, and then verified the heterogeneous role of the 3 core genes in other cancers by pan-cancer analysis and validate its expression by RT-qPCR in normal and tumor cells.

**Results:** We consistently categorized 80 UM patients into two subgroups after the immunogenetic set, where the UM1 subgroup had a better prognosis than the UM2 subgroup, and used 3 immune-related genes AZGP1, SLCO5A1, and CTF1 to derive risk scores as independent prognostic markers and predictors of UM clinicopathological features. We found significant differences in overall survival (OS) between low- and high-risk groups, and prognostic models were negatively correlated with B cell and myeloid dendritic cell and positively correlated with CD8+ T cell AZGP1 and CTF1 were significantly upregulated in UM cells compared with normal UM cells.

**Conclusion:** Immunogens are significantly associated with the prognosis of UM, and further classification based on genetic characteristics may help to develop immunotherapeutic strategies and provide new approaches to develop customized treatment strategies for patients.

KEYWORDS

**bioinformatics, uveal melanoma, immunity, prognosis, UM**

## Introduction

Uveal melanoma (UM) is a malignancy of pigment cell derivation occurring in the eye with a poor prognosis and susceptibility to metastasis, as well as a mortality rate of up to 50% (1, 2). Tumor metastasis can occur at any time (3). Once metastases are detected, the median survival time of UM patients is about 12 months (4, 5). Therefore, the search for a stable clinical indicator and molecular biomarker that can predict patient prognosis has become a hot topic in the UM treatment domain (6).

The tumor microenvironment (TME) has been found to play a key role in cancer progression and treatment response (7, 8). Prognostic or predictive biomarkers associated with the TME may hold great promise in identifying molecular targets and guiding patient management. The current therapeutic options for the treatment of metastatic UM include liver directed therapies and systemic targeted immunotherapy. Notably, immunotherapy may be a potential option for alternative or adjuvant therapy, even in prophylactic settings. High-risk UMs that metastasize usually contain macrophages and lymphocytes. These lymphocytes are usually regulatory T cells that can suppress the immune response, but UM may be particularly sensitive to T cell-based immunotherapy. Another treatment option for patients with metastatic UM is targeted therapy with T cells that target tumor-associated antigens. Some studies have demonstrated the potential role of PRAME (preferentially expressed antigen in melanoma)-targeted immunotherapy in patients with metastatic UM (9). Although substantial progress has been attained in some important genes and pathways involved in the diagnosis and treatment of UM, the prognosis of UM is still poor. Therefore, there remains an urgent need to develop an immune-related prognostic marker for UM.

Unsupervised class discovery is a data mining technique used to detect unknown groups of possible items based on intrinsic features and no external information. Consensus clustering methods provide quantitative and visual stability evidence for estimating the number of unsupervised classes in a dataset (10). In recent years, consensus clustering has been increasingly used in cancer research to classify cancers into different subgroups by consensus clustering and to explore the differences in clinical features and heterogeneous expression of genes between subgroups. Twenty-two common epithelial–mesenchymal transition (EMT)-related genes that were differentially expressed in gliomas were divided into two subgroups by consensus clustering, and then seven EMT-related genes were used to derive risk scores as independent prognostic markers and predictors of glioma clinicopathological features (11). In lung adenocarcinoma (LUAD), molecular subtypes of LUAD were identified based on tumor invasion-related genes and a 5-gene signature prognostic stratification system (12). Therefore, different UM subgroups can be developed by consensus clustering to further explore immunotherapeutic strategies.

This study aimed to classify patients with UM into subgroups based on immune gene sets using a hierarchical clustering approach to screen prognostic factors by taking the intersections of differentially expressed genes (DEGs) and immune-related genes between subgroups for prognostic analysis and then constructing prognostic risk models by LASSO to obtain the three core immune prognostic genes AZGP1, SLCO5A1, and CTF1. These were subsequently subjected to pan-cancer analysis to verify their heterogeneous roles in other cancers of this kind and to come up with new targets for the diagnosis and treatment of UM.

## Materials and Methods

### Data sources and statistical methods

Tumor RNAseq (RNA-sequencing) data were obtained from The Cancer Genome Atlas (TCGA) database (IlluminaHiSeq: log2-normalized_count+ 1). A list of immune-related genes was rooted in the Immunology Database and Analysis Portal (ImmPort) database. Data on immune infiltrating cell scores in six from UM were downloaded from the TIMER database. Subsequent statistical methods were implemented using the R software without special instructions.

### Hierarchical clustering

UM patients were divided into different subgroups according to the expression of the immune gene set and the relationship between different subgroups, and immunity was observed. The maximum number of clusters was 6. The optimal number of clusters was inferred by choosing the appropriate K value, and the clustering heat map was analyzed using the R package pheatmap. The immune infiltration was reliably estimated using the R packages ggplot2 and pheatmap, and the distribution of immune checkpoint-associated genes was observed.

**Abbreviations:** BLCA, Bladder Urothelial Carcinoma; BRCA, Breast invasive carcinoma; CHOL, Cholangiocarcinoma; COAD, Colon adenocarcinoma; ESCA, Esophageal carcinoma; GBM, Glioblastoma multiforme; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney Chromophobe KIRC Kidney renal clear cell carcinoma; LGG, Brain Lower Grade Glioma; LIHC, Liver hepatocellular carcinoma; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; PAAD, Pancreatic adenocarcinoma; PRAD, Prostate adenocarcinoma; READ, Rectum adenocarcinoma; STAD, Stomach adenocarcinoma; THCA, Thyroid carcinoma; UCEC, Uterine Corpus Endometrial Carcinoma.

## Functional enrichment analysis

Differential analysis of immune-related genes was constructed using the R package limma, with $p < 0.05$ and |fold change|>2 as the screening conditions represented by a volcano plot. Meanwhile, the intersection of the screened DEGs with immune-related genes was taken using Venn diagrams. Gene Ontology (GO) functional enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the intersected genes were performed using the R package cluster profile. For gene set functional enrichment analysis, the h.all.v7.4.symbols.gmt subset was downloaded from the Molecular Signatures Database (http://www.gsea-msigdb.org/gsea/downloads.jsp) and used as a background to map genes to the The enrichment analysis was performed using the R package clusterProfiler (version 3.14.3) to obtain the results of gene set enrichment. A minimum gene set of 5 and a maximum gene set of 5,000 were set, and a $p$ value <0.05 and an FDR <0.25 were considered statistically significant.

## Survival analysis

The RNAseq data (level 3) and corresponding clinical information were collected from patients with UM in TCGA database. Log-rank was used to test the difference in survival between the two parts mentioned above for Kaplan–Meier (KM) survival analysis. For the KM curves, $p$-values and hazard ratios (HRs) with 95% confidence intervals (CIs) were derived *via* log-rank test and univariate Cox regression. $p < 0.05$ was considered significant.

## Prognostic signature model

The relationship between prognostic immune-related gene expression and overall survival (OS) was first assessed. A prognostic risk prediction model for UM was developed. UM patients were divided into high- and low-risk groups using the median risk score as a cutoff. KM curves were plotted to compare the OS between the high- and low-risk groups. The reactive oxygen species (ROC) survival analysis was performed, and the decision curve analysis was performed using the "rmda" package. The relationship between the risk score model and tumor immune infiltrating cells was also investigated using and found to be statistically significant at $p < 0.05$.

## Pan-cancer analysis

The RNAseq data (level 3) of different tumor tissues and tumor paracancer tissues were obtained from TCGA database. Prognostic analysis was performed using univariate Cox regression analysis, and the forest plot was used to show the $p$-value, HR, and 95% CI through the "forestplot" R package. The rank-sum test was used to detect the difference between the two data parts, and $p < 0.05$ was significant.

## Cell culture

Human UM cell line OM431 and normal UM cells changed to: Human retinal pigment epithelial cell line (hTERT RPE-1 U1L) were purchased from ATCC (American Type Culture Collection) and cultured in DMEM (Dulbecco's Modified Eagle Medium) containing 100 ml/L fetal bovine serum, 100 U/ml penicillin, and 100 μg/ml streptomycin at 37°C with 5% $CO_2$ by volume.

## RT-qPCR

Total RNA was extracted from groups of cells using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's guidelines. cDNA was synthesized from RNA using the PrimeScript RT kit (Takara, Dalian, China). cDNA was amplified and quantified using SYBR Green mix (Takara) in an Applied Biosystems 7500 instrument. The primer sequences used in this study are shown in Table 1. $-2^{\Delta\Delta Ct}$ method was used to determine the relative gene expression.

# Results

## Consensus clustering of immune-related uveal melanoma

To investigate the role of immune-related genes in UM, we used the ConsensusClusterPlus software package to cluster patients according to their immune-related gene expression profiles. The consensus cumulative distribution function (CDF) plots showed a good proportion of disambiguation clusters

TABLE 1 Primer sequences.

| Name of primer | Sequences |
|---|---|
| AZGP1-F | 5'-TACAACGACAGTAACGGGTCT-3' |
| AZGP1-R | 5'-TATTTCCAGAATGCTCCGCTG-3' |
| SLCO5A1-F | 5'-TGCCTCTACGTGGTCCTCAC-3' |
| SLCO5A1-R | 5'-TTACGCTGCTCAGGTACCCA-3' |
| CTF1-F | 5'-CCCTCCTCGTCTGCATGGTA-3' |
| CTF1-R | 5'-GAGGCCAAAGGGAACTGAGG-3' |
| GAPDH-F | 5'-CAAGCAACTGTCCCTGAG-3' |
| GAPDH-R | 5'-TAGACAGAAGGTGGCACA-3' |

(Figures 1A–C) when the number of clusters was two, dividing the patients into two independent groups (Figures 1D, E). We evaluated these two clusters, and the results showed that they remain separable (Figure 1F). We compared the differences in prognostic and clinical characteristics of patients belonging to these two UM subtypes. We found that the patients in part 2 had a significantly worse prognosis compared with that of the patients in part 1. A comparison of the clinical baseline information between the two groups revealed significant differences in survival status between the two parts (Table 2). Thus, based on the immune gene set, the sample could be divided into two subtypes through consensus clustering. Notably, significant differences in clinical prognosis survival were noted between the two subtypes (Figure 1G, $p = 0.00418$).

## Analysis of subgroup gene expression differences in immune-related uveal melanoma

To explore the hidden mechanisms driving the differences in clinical immune characteristics and biological functions between the two subgroups, we analyzed the differences in their mRNA expression profiles in TCGA database. The DEGs were analyzed for immune-related C1 and C2 UM subtypes according to |Log2| FC||>2, adj. $p < 0.05$ as the screening threshold. Four upregulated and 115 downregulated genes were obtained (Figure 2A). The heat map shows the upregulated and downregulated DEGs (Figure 2B). GO (using BP only), KEGG, and Hallmark were selected as the databases for the functional gene enrichment analysis of these DEGs by Metascape (Figures 2C–E). The parameters for enrichment analysis are as follows: the number of Min Overlap genes was 3, the $p$-value cutoff was 0.01, and the Min Enrichment was 1.5.

## Immune-related prognostic model for uveal melanoma

We obtained 58 identical genes (Figure 3A) by taking the intersection of the subgroup DEGs, prognosis-related genes, and immune-related genes. Then, prognostic characteristics were established (Figures 3B, C) for the 58 genes based on LASSO Cox analysis, which were determined by Riskscore = (-0.0548)*AZGP1+(0.0207)*SLCO5A1+(- 0.4091)*. CTF1 divided the patients with UM into high-risk and low-risk parts (Figure 4A). Their survival status is shown in Figure 4B. There was a significant difference in survival between the high- and low-risk groups ($p = 0.00368$, Figure 3D). Prognostic survival prediction was performed for

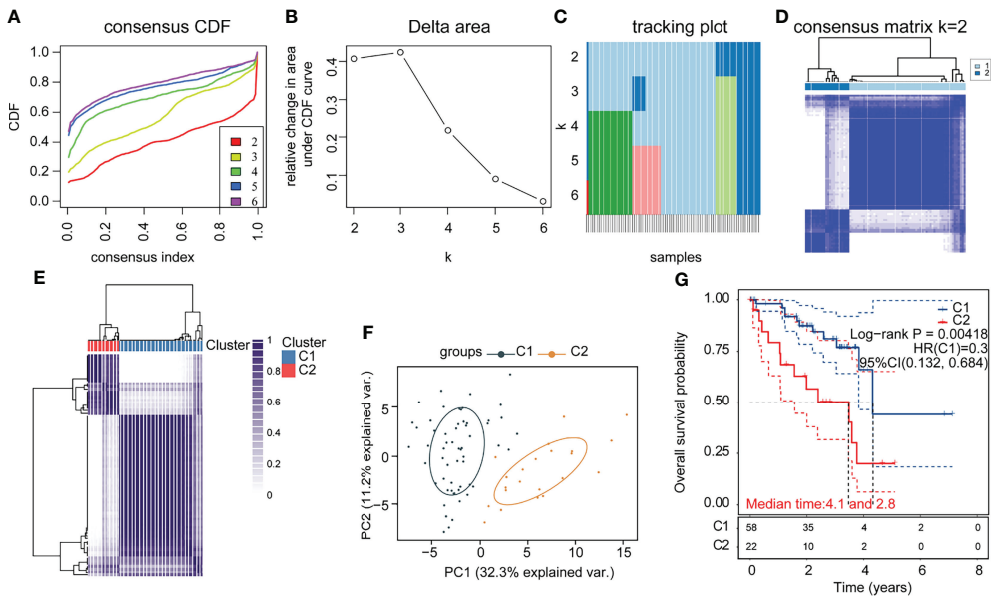



FIGURE 1
Consensus clustering of immune-related uveal melanoma (UM) **(A)** CDF plot; **(B)** CDF Delta area plot; **(C)** clustered sample distribution plot (the color indicates the sample distribution cohort, the vertical coordinate indicates the number of clusters, and the horizontal coordinate indicates the samples); **(D)** consensus matrix heat map with two sample clusters defined (consensus range of 0–1, with 0 representing white, meaning that the samples are not clustered, and 1 representing blue, meaning that the samples are always clustered); **(E)** sample correlation heat map (range of 0–1 and the larger the value, the higher the correlation); **(F)** PCA (Principal Component Analysis), clustering plot between samples; **(G)** KM (Kaplan-Meier) survival curve of the two subgroups (statistically significant at $p < 0.05$).

TABLE 2 Comparison of clinical data between subgroups.

| Character | C1 | C2 | *P*-value |
|---|---|---|---|
| Status | | | 0.004 |
| Alive | 47 | 10 | |
| Dead | 11 | 12 | |
| Mean (SD) | 50.8(14.2) | 63.8 (13.4) | |
| Age | | | 0.397 |
| Median [MIN, MAX] | 63.5[22, 86] | 60[37, 86] | |
| Gender | | | 0.344 |
| FEMALE | 23 | 12 | |
| MALE | 35 | 10 | |
| pT-stage | | | 0.97 |
| T2a | 3 | | |
| T3 | 6 | 2 | |
| T3a | 18 | 6 | |
| T3b | 3 | 1 | |
| T4 | 1 | 1 | |
| T4a | 17 | 6 | |
| T4b | 7 | 2 | |
| T4c | 1 | 1 | |
| T4d | 1 | 1 | |
| T4e | 1 | | |
| T2 | | 1 | |
| T2b | | 1 | |
| pN-stage | | | 0.646 |
| N0 | 56 | 20 | |
| NX | 2 | 2 | |
| pM-stage | | | 0.637 |
| M0 | 54 | 19 | |
| M1b | 2 | | |
| MX | 2 | 2 | |
| M1 | | 1 | |
| pTNM-stage | | | 0.85 |
| IIA | 9 | 3 | |
| IIB | 20 | 7 | |
| IIIA | 19 | 7 | |
| IIIB | 8 | 2 | |
| IV | 2 | 2 | |
| IIIC | | 1 | |
| New tumor event type | | | 0.903 |
| Metastasis | 9 | 7 | |
| Primary | 2 | 3 | |
| Recurrence | 1 | | |
| History of neoadjuvant treatment therapy type | | | |
| No neoadjuvant | 58 | 22 | |
| Chemotherapy | 2 | 2 | |

1, 3, and 5 years, and this prognostic model showed good sensitivity and specificity (Figure 3E). Based on the foregoing, AZGP1 and CTF1 were good prognostic genes, and SLCO5A1 was the main risk gene.

Subsequently, we analyzed the correlation between this prognostic signature model and tumor immune cell infiltration (ICI). The results are shown in Figure 4B, indicating a negative correlation with B cells and myeloid dendritic cells ($p < 0.05$) and a positive correlation with CD8+ T cells ($p < 0.05$).

## Three core gene prognosis, clinical characteristics, and model accuracy analysis

Based on the prognostic signature model, CTF1, SLCO5A1, and AZGP1 are the core genes of 59 intersecting genes. To verify the clinical significance of these three core genes for UM, we used OS as an index to determine the correlation between gene expression and prognosis. The prognostic relevance was clarified by KM survival curves.

We used OS as an index to judge the prognostic relevance of gene expression, the KM survival curve to clarify its prognostic relevance, and ROC (subject operating curve) as a tool to judge the accuracy of the model. The results showed that high expressions of AZGP1 and CTF1 had better OS than low expressions in UM (Figures 5A, C), and a high expression of SLCO5A1 had worse OS than a low expression (Figure 5B). The ROC curve confirmed the accuracy of the model (Figures 5D–F). We then further validated the correlation between the high and low expressions of these three genes on clinical characteristics, and the results showed that the high and low expressions of AZGP1, SLCO5A1, and CTF1 were strongly correlated with survival status (Figures 5G–I). These results suggest that the expressions of the three core genes AZGP1, SLCO5A1, and CTF1 affect the clinical prognosis of UM.

## Analysis of the correlation between the three core genes and the pathway

To further explore the mechanistic roles of the three core genes in UM, we performed a pathway correlation analysis on the three core genes CTF1, SLCO5A1, and AZGP1. The results showed that the top 20 correlated pathways were mainly focused on immune-related pathway signals, among which CTF1 and AZGP1 were negatively correlated (Supplementary Tables 1, 3) whereas SLCO5A1 was positively correlated (Supplementary Table 2).

## Expression of three core genes in uveal melanoma cells

We then examined the mRNA levels of AZGP1, SLCO5A1, and CTF1 in UM, and the results are shown in Figure 6. AZGP1

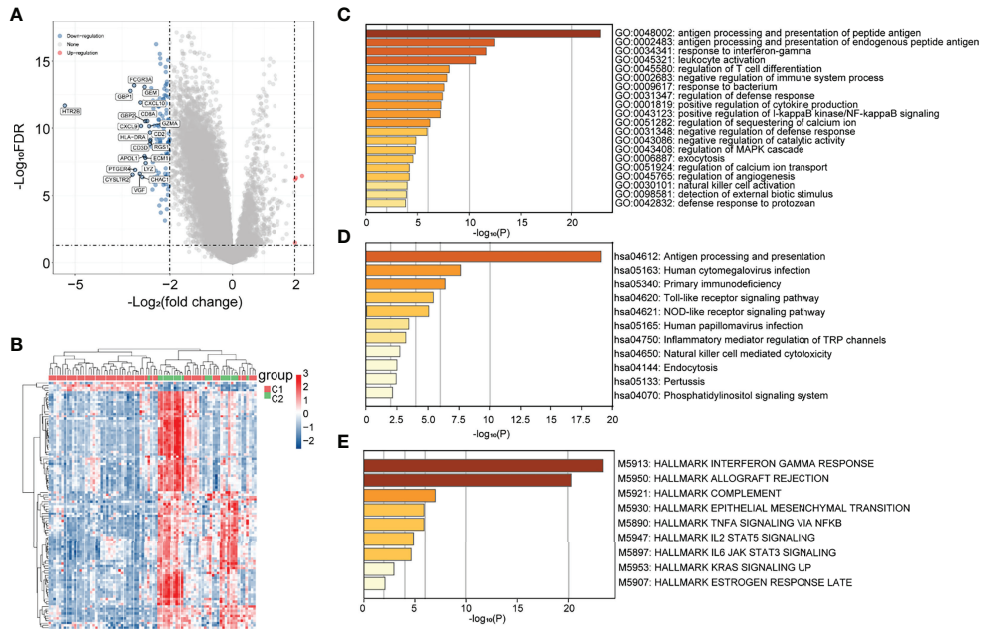



FIGURE 2
Differential analysis of subgroup gene expression in immune-related uveal melanoma (UM). **(A)** Volcano map; **(B)** heat map; **(C–E)** GO (BP only), KEGG, and Hallmark pathway enrichment maps.

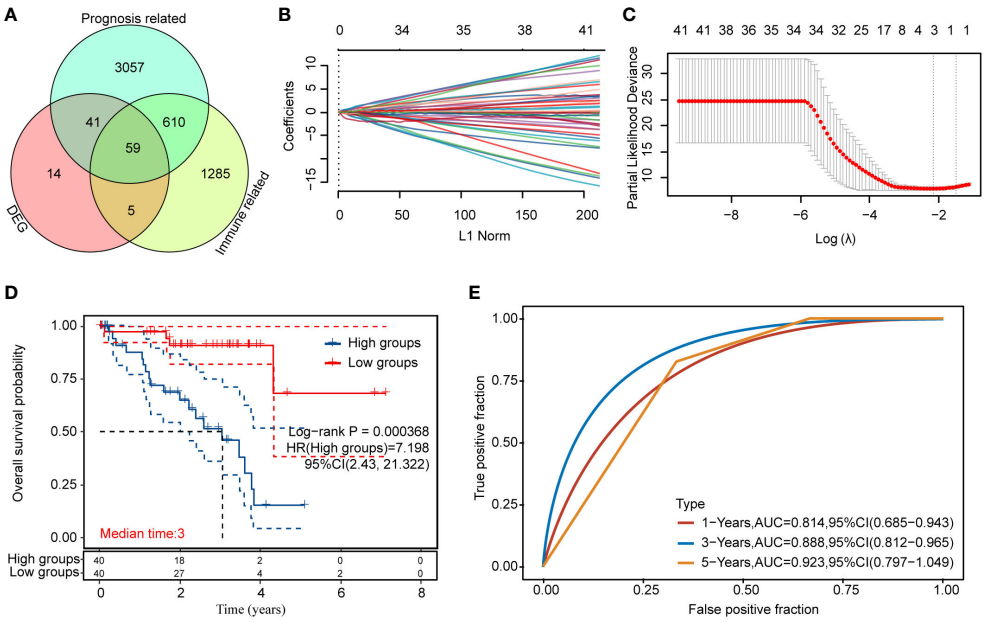



FIGURE 3
Immune-related prognosis model construction for DEGs **(A)** Intersection of the DEGs, prognosis-related genes, and immune-related genes; **(B, C)** distribution of LASSO coefficients of immune prognosis common genes of DEGs to obtain the adjustment parameter λ.min = 0.000368; **(D)** KM survival curves for high- and low-risk groups; **(E)** ROC curves at 1, 3, and 5 years (statistically significant at $p < 0.05$).

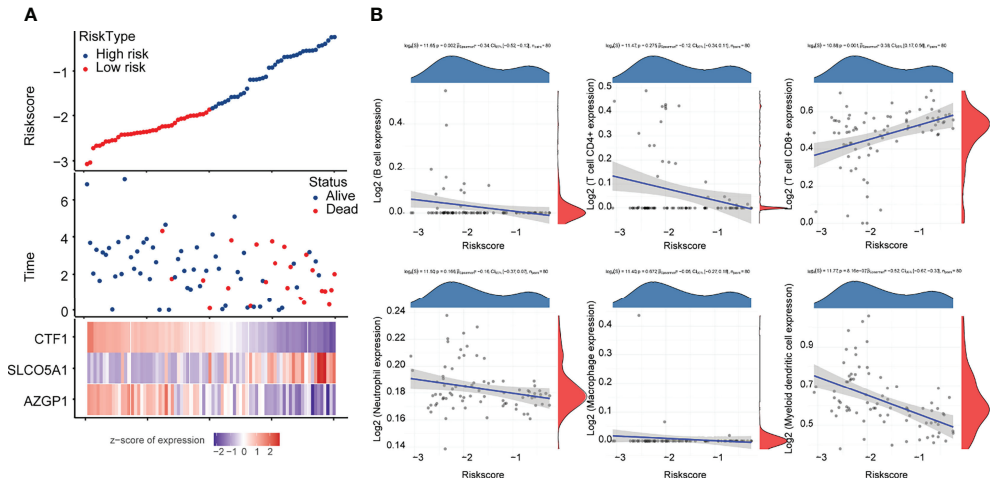



FIGURE 4
Correlation analysis of the prognostic signature model with tumor immune cell infiltration **(A)** Survival status and duration of patients with UM; **(B)** letters A–G represent B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, myeloid dendritic cells, and uncharacterized cells, respectively. The horizontal coordinates in the figure represent the model score distribution, whereas the vertical coordinates are the immune score distribution. The density curves on the right side represent the immune score distribution trends. The upper side density curve is the trend of distribution of one gene or model score. The uppermost value represents the correlation *p*-value, correlation coefficient, and correlation calculation method, with $p < 0.05$ considered as statistically significant.

and CTF1 were significantly upregulated in UM cells compared with hTERT RPE-1 cell (Figure 6A, C).

## Pan-cancer analysis of the three core genes

To evaluate the heterogeneous roles of the three core genes in different cancers, we performed a pan-cancer analysis. First, we compared the gene expression profile data of the three genes in different cancer tissues and normal tissues adjacent to the cancer. Compared with the paracancerous tissues, CTF1 was significantly decreased in BLCA(Bladder Urothelial Carcinoma), BRCA(Breast invasive carcinoma), ESCA(Esophageal carcinoma), KICH (Kidney Chromophobe), LIHC(Liver hepatocellular carcinoma), LUAD(Lung adenocarcinoma), LUSC(Lung squamous cell carcinoma), PRAD(Prostate adenocarcinoma), STAD(Stomach adenocarcinoma), and UCEC(Pancreatic adenocarcinoma); significantly increased in CHOL(Cholangiocarcinoma), GBM (Glioblastoma multiforme), KIRP(Kidney renal papillary cell carcinoma), and THCA(Thyroid carcinoma); and had no significant effect on ESCA(2Esophageal carcinoma), HNSC (Head and Neck squamous cell carcinoma), LGG(Brain Lower Grade Glioma), and PAAD(Pancreatic adenocarcinoma) (Figure 7A). Meanwhile, SLCO5A1 was significantly lower in KICH(Kidney Chromophobe); significantly higher in BRCA (Breast invasive carcinoma), CHOL(Cholangiocarcinoma), COAD(Colon adenocarcinoma), ESCA(Esophageal carcinoma), KIRC(Kidney renal clear cell carcinoma), KIRP(Kidney renal papillary cell carcinoma), LGG(Brain Lower Grade Glioma), LIHC(Liver hepatocellular carcinoma), LUAD(Lung adenocarcinoma), LUSC(Lung squamous cell carcinoma), READ(Rectum adenocarcinoma), STAD(Stomach adenocarcinoma), and UCEC(Uterine Corpus Endometrial Carcinoma); and significantly higher in BLCA(Bladder Urothelial Carcinoma), GBM(Glioblastoma multiforme), HNSC (Head and Neck squamous cell carcinoma), PAAD(Pancreatic adenocarcinoma), PRAD(Prostate adenocarcinoma), and THCA (Thyroid carcinoma) (Figure 7B). Finally, AZGP1 was significantly decreased in BLCA(Bladder Urothelial Carcinoma), CHOL(Cholangiocarcinoma), ESCA(Esophageal carcinoma), HNSC(Head and Neck squamous cell carcinoma), KIRC(Kidney renal clear cell carcinoma), KIRP(Kidney renal papillary cell carcinoma), LIHC(Liver hepatocellular carcinoma), LUAD(Lung adenocarcinoma), LUSC(Lung squamous cell carcinoma), THCA, and UCEC(Pancreatic adenocarcinoma); significantly increased in COAD(Colon adenocarcinoma) and READ(Rectum adenocarcinoma); and had no significant effect on BRCA(Breast invasive carcinoma), GBM (Glioblastoma multiforme), KICH (Kidney Chromophobe), LGG(Brain Lower Grade Glioma), PAAD(Pancreatic adenocarcinoma), and STAD(Stomach adenocarcinoma) Figure 7C. Subsequently, we evaluated the prognostic effects of the three core genes on these cancers and found that CTF1 had a significant effect Htabon the prognosis of ACC(Adrenocortical carcinoma), COAD(Colon adenocarcinoma), ESCA(Esophageal carcinoma), GBM (Glioblastoma multiforme), LGG(Brain Lower Grade Glioma), LUAD(Lung adenocarcinoma), MESO(Mesothelioma), UCEC

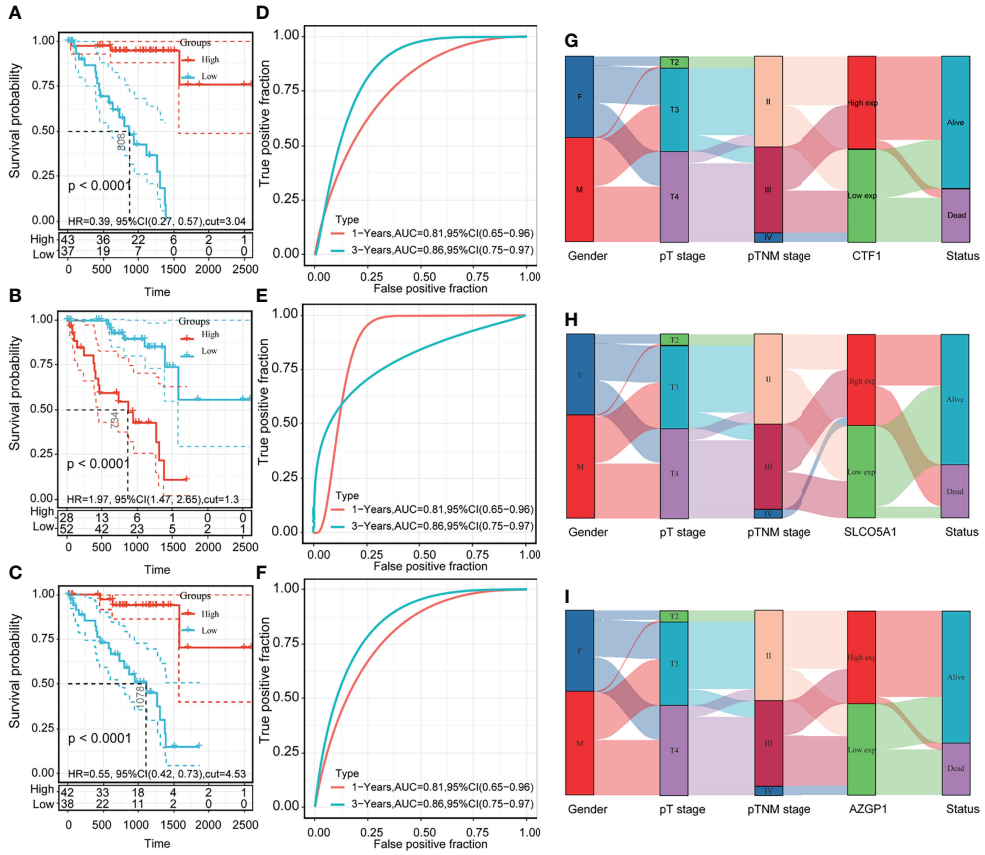



FIGURE 5
Prognosis of core genes, clinical characteristics, and model accuracy analysis. **(A–C)** KM survival curves of high and low expressions of AZGP1, SLCO5A1, and CTF1; **(D–F)** 1-year and 3-year ROC curves of high and low expressions of AZGP1, SLCO5A1, and CTF1; **(G–I)** Sankey plots of high and low expressions of AZGP1, SLCO5A1, and CTF1 with clinical characteristics; statistically significant at $p < 0.05$.

(Pancreatic adenocarcinoma), and UVM(Uveal Melanoma); SLCO5A1 had a significant effect on the prognosis of KIRC (Kidney renal clear cell carcinoma), LGG, PAAD(Pancreatic adenocarcinoma), and UVM(Uveal Melanoma); and AZGP1 had a significant effect on the prognosis of KIRC(Kidney renal clear cell carcinoma), KIRP(Kidney renal papillary cell carcinoma), UCEC(Pancreatic adenocarcinoma), and UVM (Uveal Melanoma) (Figures 7D–F, all $p < 0.05$).

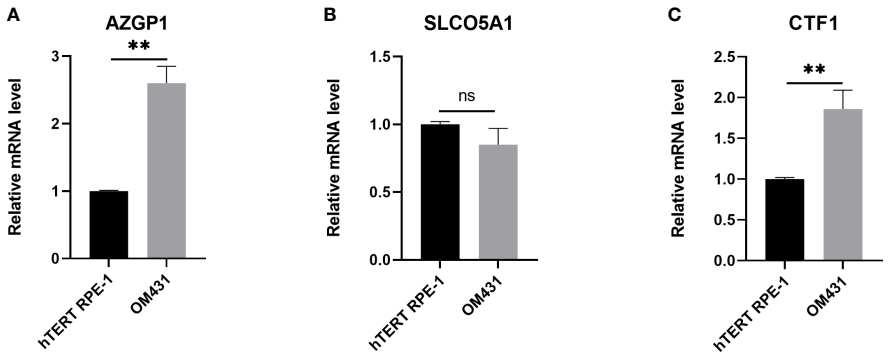



FIGURE 6
Verification of the expression of the three core genes in cells **(A–C)** The expression of AZGP1, SLCO5A1, and CTF1 in Human retinal pigment epithelial cell line(hTERT RPE-1) and uveal melanoma cells (OM431) was detected by RT-qPCR, cellular experiments were performed in three independent replicates, data in the figures are all measures in the form of mean ± standard deviation, data were analyzed by t-test,** indicates $p < 0.01$.

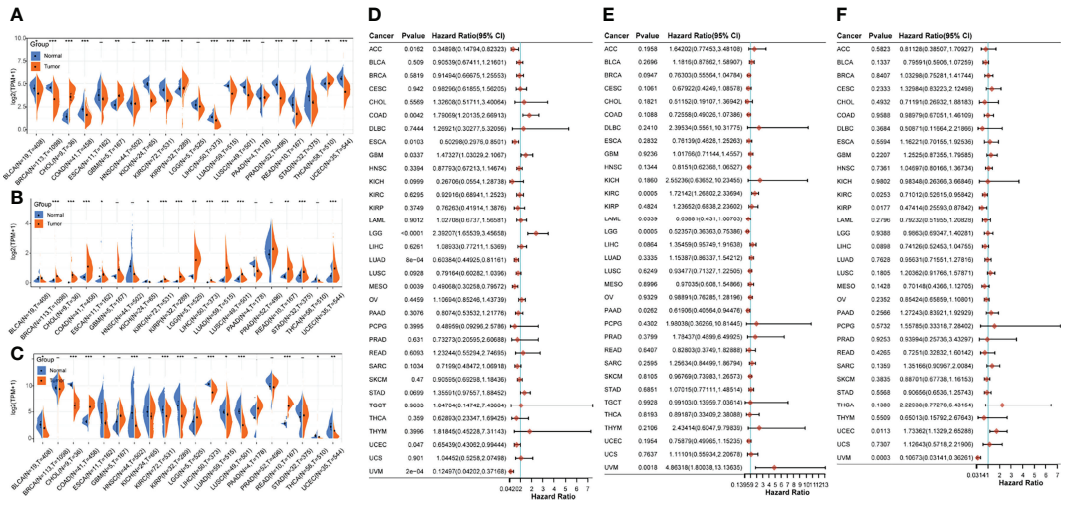



FIGURE 7
Pan-cancer analysis of core genes. **(A–C)** Expressions of AZGP1, SLCO5A1, and CTF1 in different cancers and paracancerous tissues; **(D–F)**: prognostic survival effects of AZGP1, SLCO5A1, and CTF1 in different cancers; * indicates $p < 0.05$, ** indicates $p < 0.01$, and *** indicates $p < 0.001$, with $p < 0.05$ being significant.

## Discussion

In this study, we obtained two subgroups of UM through hierarchical clustering for which we performed differential gene analysis, as well as GO, KEGG, and Hallmark enrichment analyses focused on the immune response signaling pathways. Subsequent prognostic survival analysis of these differential genes identified 53 immune prognostic factors. Three core genes, namely, CTF1, SLCO5A1, and AZGP1, were obtained for the LASSO construct risk model. The pathway correlation analysis demonstrated that these three factors were associated with the immune response. Pan-cancer analysis further demonstrated the heterogeneous role of these three core genes.

Identification of novel features in UM based on immune-related genes provides new directions for assessing immune efficacy. Further evaluation of these classifications based on genetic features may help to develop immunotherapeutic strategies and improve the sensitivity to different subtypes of UM. In this study, we analyzed TCGA dataset and classified patients into two different immune subtypes, with significant differences in prognostic and clinical characteristics between the two groups. Due to the heterogeneity in gene expression and prognosis between the two immune subtypes, we hypothesized that integrated tumor ICI analysis and immune gene expression pattern assessment would be a new approach to develop patient-tailored and customized treatment strategies.

Understanding the TME structures and applying the insights gained to drug design have become a hot topic in modern cancer research (13). CD8+ T cells have been found to have a favorable prognostic role in many cancers (14). The immunosuppressive factors present in the TME can suppress T cell-mediated responses and rewire their activity to benefit tumors in which T-cell infiltration is related to poor prognosis (15). Studies have shown that metalloproteinases are effective targets for cancer therapy (16). Aside from acting as simple effectors of angiogenesis and metastasis, they also play a role in the regulation of immune responses. It has recently been shown that UM high-risk populations are related to a significantly enhanced metalloproteinase profile, which may play a dominant role in driving UM metastasis. CTF1 is a mitogenic cytokine of the interleukin 6 family. Meanwhile, CTF1 contains the transmembrane and cytoplasmic structural domains of E-calmodulin and is produced upon cleavage. It can act as a cleavage product of E-calmodulin, in addition to CTF itself, which can also act as a downstream signaling molecule. Studies have shown that the MMP28–E-calmodulin–CTF association mechanism promotes colorectal cancer progression and metastasis (17).

One of the factors associated with increased immune infiltration is chromosome 3 monosomy, which is considered a negative prognostic factor, while the early acquisition of chromosome 8 is associated with macrophage infiltration (18). SLCO5A1 is located on chromosome 8, but it remains unknown whether it can influence chromosomal variants leading to the development and progression of UM. A recent study has shown that SLCO5A1 is a negative prognostic factor in UM. Moreover, in another study, AZGP1 was found to be a good prognostic factor, which coincides with our results, while SLCO5A1 and AZGP1 are closely associated with the abundance of neutrophils and CD8+ T cells (19).

The present study has some limitations. The current results need to be validated in immunotherapy clinical trials with larger UM cohorts, which will confirm the utility of classification in clinical evaluation and decision-making. In addition, the present study was based on transcriptome expression profiles of UM tissues from TCGA database and did not have a controlled study of normal uveal melanocytes, which may not allow for an accurate prediction. Therefore, it is important to better understand the circulating biomarkers released into the bloodstream from tumor cells and tumor-associated immune cells. Further *in vivo* and *in vitro* experiments should investigate the potential functional and mechanistic differences between subtypes. Finally, although we did a pan-cancer analysis of 3 core genes, the results of this study and ICI scores may be applicable to other cancers based on the heterogeneous functional role of genes in different cancers, which need further investigation.

In our study, we analyzed both the expression and prognosis of these three core genes in other cancers and found that CTF1, SLCO5A1, and AZGP1 presented different functional roles in other cancers. A recent study reported that the systematic screening of differentially expressed circRNAs (DEcircRNAs), miRNAs (DEmiRNAs), and mRNAs (DEGs) associated with LUAD identified a ceRNA prognostic regulatory network consisting of 1 circRNA, 2 miRNAs, and 7 mRNAs, in which CTF1 as a good prognostic gene is regarded as a drug target (20). In addition, it was shown that CTF1/N-Cad (CTF1) is a product of extracellular metalloproteinase (MMP), which cleaves near the interface between the extracellular and transmembrane regions of N-calcine mucin, and that the extracellular and intracellular cleavages of N-calcine mucin may be involved in elevated MMP-9 expression and enhanced invasion of human nasopharyngeal carcinoma cells (21). In some studies, the SNP(single nucleotide polymorphism) in SLCO5A1 was correlated with the clinical staging of prostate cancer (22). Similarly, in some studies, SLCO5A1 expression was found to decrease during differentiation from monocytes to macrophages but increase during differentiation from monocytes to mature dendritic cells (23). This differential expression profile may be related to the fact that SLCO5A1 plays an important role in the immune function involved in chylomicronoma. In our data, AZGP1 was considered to have a significant impact on KIRC, KIRP, UCEC, and UVM prognosis, but it has been shown that AZGP1 could be a feasible candidate biomarker for colorectal cancer based on an analysis of the GSE21962, GSE24551, and GSE29638 datasets in the GEO(Gene Expression Omnibus) database (24). Patients with oral squamous cell carcinoma (OSCC) have low salivary levels of AZGP1 (25)and low mRNA levels of AZGP1 in OSCC tumor tissues (26). The role of AZGP1 in inhibiting cell invasion and migration (27, 28) suggests its correlation with poor disease response. Similarly, the low mRNA/protein expression of AZGP1 is associated with disease progression and poor survival in pancreatic cancer (29, 30). The above data underscore the heterogeneous role that the three core genes, CTF1, SLCO5A1, and AZGP1, present in other cancers.

In summary, in our study, based on TCGA database, we obtained three core immune-related prognostic factors for UM, namely, CTF1, SLCO5A1, and AZGP1, through bioinformatics methods. Notably, the immune infiltration analysis proved that the risk model that we constructed was related to B cells, myeloid dendritic cells, and CD8+ T cells and that the three core genes mentioned above were mainly the focus of immune-related pathway signaling. These results could drive new therapies for UM treatment.

## Data availability statement

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

## Author contributions

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

## Funding

The project was sponsored by grants from Hunan Provincial Natural Science Foundation of China (2020JJ4912).

## 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.2022.918230/full#supplementary-material.

# References


1. Amaro A, Gangemi R, Piaggio F, Angelini G, Barisione G, Ferrini S, et al. The biology of uveal melanoma. *Cancer Metastasis Rev* (2017) 36(1):109–40. doi: 10.1007/s10555-017-9663-3

2. Kaliki S, Shields CL. Uveal melanoma: Relatively rare but deadly cancer. *Eye (Lond)* (2017) 31(2):241–57. doi: 10.1038/eye.2016.275

3. Stalhammar G. Forty-year prognosis after plaque brachytherapy of uveal melanoma. *Sci Rep* (2020) 10(1):11297. doi: 10.1038/s41598-020-68232-7

4. Rantala ES, Hernberg M, Kivela TT. Overall survival after treatment for metastatic uveal melanoma: A systematic review and meta-analysis. *Melanoma Res* (2019) 29(6):561–8. doi: 10.1097/CMR.0000000000000575

5. Rodriguez-Vidal C, Fernandez-Diaz D, Fernandez-Marta B, Lago-Baameiro N, Pardo M, Silva P, et al. Treatment of metastatic uveal melanoma: Systematic review. *Cancers (Basel)* (2020) 12(9):2557. doi: 10.3390/cancers12092557

6. Jager MJ, Dogrusoz M, Woodman SE. Uveal melanoma: Identifying immunological and chemotherapeutic targets to treat metastases. *Asia Pac J Ophthalmol (Phila)* (2017) 6(2):179–85. doi: 10.22608/APO.201782

7. Ravindran S, Rasool S, Maccalli C. The cross talk between cancer stem Cells/ Cancer initiating cells and tumor microenvironment: The missing piece of the puzzle for the efficient targeting of these cells with immunotherapy. *Cancer Microenviron* (2019) 12(2-3):133–48. doi: 10.1007/s12307-019-00233-1

8. Cully M. Tumour microenvironment: Fibroblast subtype provides niche for cancer stem cells. *Nat Rev Cancer* (2018) 18(3):136. doi: 10.1038/nrc.2018.18

9. Gezgin G, Luk SJ, Cao J, Dogrusöz M, van der Steen DM, Hagedoorn RS, et al. PRAME as a potential target for immunotherapy in metastatic uveal melanoma. *JAMA Ophthalmol* (2017) 135(6):541–9. doi: 10.1001/jamaophthalmol.2017.0729

10. Wilkerson MD, Hayes DN. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26 (12):1572–3. doi: 10.1093/bioinformatics/btq170

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

12. Yu P, Tong L, Song Y, Qu H, Chen Y. Systematic profiling of invasion-related gene signature predicts prognostic features of lung adenocarcinoma. *J Cell Mol Med* (2021) 25(13):6388–402. doi: 10.1111/jcmm.16619

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

14. Talmadge JE. Immune cell infiltration of primary and metastatic lesions: Mechanisms and clinical impact. *Semin Cancer Biol* (2011) 21(2):131–8. doi: 10.1016/j.semcancer.2010.12.002

15. Kaliki S, Shields CL, Shields JA. Uveal melanoma: Estimating prognosis. *Indian J Ophthalmol* (2015) 63(2):93–102. doi: 10.4103/0301-4738.154367

16. Webb AH, Gao BT, Goldsmith ZK, Irvine AS, Saleh N, Lee RP, et al. Inhibition of MMP-2 and MMP-9 decreases cellular migration, and angiogenesis in in vitro models of retinoblastoma. *BMC Cancer* (2017) 17(1):434. doi: 10.1186/ s12885-017-3418-y

17. Drury J, Rychahou PG, Kelson CO, Geisen ME, Wu Y, He D, et al. Upregulation of CD36, a fatty acid translocase, promotes colorectal cancer metastasis by increasing MMP28 and decreasing e-cadherin expression. *Cancers (Basel)* (2022) 14(1):252. doi: 10.3390/cancers14010252

18. Gezgin G, Dogrusöz M, van Essen TH, Kroes WGM, Luyten GPM, van der Velden PA, et al. Genetic evolution of uveal melanoma guides the development of an inflammatory microenvironment. *Cancer Immunol Immunother* (2017) 66 (7):903–12. doi: 10.1007/s00262-017-1991-1

19. Luo H, Ma C. Identification of prognostic genes in uveal melanoma microenvironment. *PloS One* (2020) 15(11):e0242263. doi: 10.1371/ journal.pone.0242263

20. Gao L, Zhang L. Construction and comprehensive analysis of a ceRNA network to reveal potential prognostic biomarkers for lung adenocarcinoma. *BMC Cancer* (2021) 21(1):849. doi: 10.1186/s12885-021-08462-8

21. Hsu CC, Huang SF, Wang JS, Chu WK, Nien JE, Chen WS, et al. Interplay of n-cadherin and matrix metalloproteinase 9 enhances human nasopharyngeal carcinoma cell invasion. *BMC Cancer* (2016) 16(1):800. doi: 10.1186/s12885-016-2846-4

22. Perlman D, Chikarmane H, Halvorson HO. Improved resolution of DNA fragments in polysaccharide-supplemented agarose gels. *Anal Biochem* (1987) 163 (1):247–54. doi: 10.1016/0003-2697(87)90120-5

23. Sebastian K, Detro-Dassen S, Rinis N, Fahrenkamp D, Müller-Newen G, Merk HF, et al. Characterization of SLCO5A1/OATP5A1, a solute carrier transport protein with non-classical function. *PloS One* (2013) 8(12):e83257. doi: 10.1371/ journal.pone.0083257

24. Ding D, Han S, Zhang H, He Y, Li Y. Predictive biomarkers of colorectal cancer. *Comput Biol Chem* (2019) 83:107106. doi: 10.1016/j.compbiolchem.2019.107106

25. Jain A, Kotimoole CN, Ghoshal S, Bakshi J, Chatterjee A, Prasad TSK, et al. Identification of potential salivary biomarker panels for oral squamous cell carcinoma. *Sci Rep* (2021) 11(1):3365. doi: 10.1038/s41598-021-82635-0

26. Suhr ML, Dysvik B, Bruland O, Warnakulasuriya S, Amaratunga AN, Jonassen I, et al. Gene expression profile of oral squamous cell carcinomas from Sri Lankan betel quid users. *Oncol Rep* (2007) 18(5):1061–75.

27. Feng M, Feng J, Chen W, Wang W, Wu X, Zhang J, et al. Lipocalin2 suppresses metastasis of colorectal cancer by attenuating NF-kappaB-Dependent activation of snail and epithelial mesenchymal transition. *Mol Cancer* (2016) 15 (1):77. doi: 10.1186/s12943-016-0564-9

28. Kong B, Michalski CW, Hong X, Valkovskaya N, Rieder S, Abiatari I, et al. AZGP1 is a tumor suppressor in pancreatic cancer inducing mesenchymal-to-Epithelial transdifferentiation by inhibiting TGF-Beta-Mediated ERK signaling. *Oncogene* (2010) 29(37):5146–58. doi: 10.1038/onc.2010.258

29. Burdelski C, Kleinhans S, Kluth M, Hube-Magg C, Minner S, Koop C, et al. Reduced AZGP1 expression is an independent predictor of early PSA recurrence and associated with ERG-fusion positive and PTEN deleted prostate cancers. *Int J Cancer* (2016) 138(5):1199–206. doi: 10.1002/ijc.29860

30. Zhang AY, Grogan JS, Mahon KL, Rasiah K, Sved P, Eisinger DR, et al. A prospective multicentre phase III validation study of AZGP1 as a biomarker in localized prostate cancer. *Ann Oncol* (2017) 28(8):1903–9. doi: 10.1093/annonc/mdx247

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 01 September 2022
DOI 10.3389/fonc.2022.986990

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Pinggui Lei,
Affiliated Hospital of Guizhou Medical University, China
Zheng Wang,
Southeast University, China

*CORRESPONDENCE
Shengbao Wang
wsb887600@163.com

†These authors share first authorship

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

RECEIVED 05 July 2022
ACCEPTED 01 August 2022
PUBLISHED 01 September 2022

CITATION
Lv W, Shi L, Pan J and Wang S (2022) Comprehensive prognostic and immunological analysis of CCT2 in pan-cancer.
*Front. Oncol.* 12:986990.
doi: 10.3389/fonc.2022.986990

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

# Comprehensive prognostic and immunological analysis of CCT2 in pan-cancer

Wenming Lv[1†], Lin Shi[2†], Jiebing Pan[3] and Shengbao Wang[4*]

[1]Department of Neurology, Second Hospital of Lanzhou University, Lanzhou, China, [2]Department of Hematology, Peking University International Hospital, Beijing, China, [3]Department of General Surgery, Lanzhou University Second Hospital, Lanzhou, China, [4]Emergency Center of the Second Hospital of Lanzhou University, Lanzhou, China

CCT2 acts as a molecular chaperone protein that assists in the proper folding of proteins, thus ensuring a dynamic balance of cellular homeostasis. Despite increasing evidence supporting the important role of CCT2 in the tumorigenesis of certain cancers, few articles that provide a systematic pan-cancer analysis of CCT2 have been published. Hence, to evaluate the expression status and prognostic significance of CCT2 in pan-cancers, an analysis of the relationship between CCT2 and different tumor immune cell infiltrations was conducted using datasets from the Cancer Genome Atlas, Cancer Cell Lineage Encyclopedia, and so on. In most cancers, CCT2 expression was high and was associated with poor prognosis. Moreover, CCT2 gene expression was negatively correlated with infiltration of most immune cells in 10 cancer types, and CCT2 expression was related to tumor mutation burden and microsatellite instability. The role that CCT2 plays in tumorigenesis and tumor immunity suggests that it can serve as a prognostic marker in many cancers.

KEYWORDS

**CCT2, pan-cancer, prognosis, immune infiltration, tumor microenvironment**

## Introduction

Chaperone proteins are a family of proteins that wrap around a substrate and help it to fold in an ATP-dependent manner and are classified into group I (heat shock protein 60 or GroEL) and group II (TCP-1 chaperone protein complex) (1). CCT is an important molecule for the synthesis of TCP-1 ring-containing chaperone protein complex (TRiC) with two ring structures (2). One-tenth of the proteins in cells are folded by TRiC, including actin and microtubulin, and it regulates the expression of tumor-associated proteins and cell cycle, which are aberrantly expressed in many tumors and are potential targets for treatment (3). CCT plays a role in the binding and hydrolysis of ATP as well as the recognition and folding of substrates as part of TRiC (4, 5). There are eight different

subunits (CCT1–CCT8), each of which has different substrate recognition and ATP hydrolysis properties (6).

The CCT family of genes is closely associated with the development of tumors according to several studies (7, 8). Hepatocellular carcinoma, for example, exhibits high levels of TCP1/CCT2-CCT8 expression and low levels of CCT6B, which leads to the abnormal regulation of Myc target genes and hypoxia-inducible factor target genes as well as cell cycle abnormalities (3). A study by Carr et al. (9) examined the effects of the CCT family genes on the development of hepatocellular carcinoma by examining the protein levels of CCT subunits in hepatocellular carcinoma, prostate cancer, and lung cancer, where higher levels of CCT2 were detected than in normal tissue. In gallbladder cancer, positive CCT2 expression was negatively correlated with low postoperative patient survival and positively correlated with high mortality (10). In our previous study, we found that CCT2 expression showed an increasing trend in normal, ulcerative colitis, and colon cancer tissues by analyzing data from The Cancer Genome Atlas (TCGA) database and Gene Expression Omnibus (GEO) database and predicted the possibility that CCT2 may be closely associated with the development of colon inflammation or cancer and that CCT2 is a prognostic factor for colon cancer (11). Therefore, a combination of previous studies and our previous study started from CCT2 and cited the clinical data of various human cancers, genomic variants, mRNA expression, miRNA expression, methylation, and other data included in TCGA database as well as protein expression in different human tissues and organs included in the Human Protein Atlas (HPA) database in order to explore the function of the role CCT2 plays in various cancers. The aim is to investigate the function of CCT2 in various cancers.

In summary, CCT2 expression has been shown to be related to poor prognosis in numerous types of tumors. In distinction, most reports to this point are restricted to investigations on the role of CCT2 in specific styles of cancer. Pan-cancer research on the association between CCT2 and varied cancers has not been reported. Therefore, we tend to analyze the CCT2 expression levels and their association with the prognosis of various styles of malignancies supported by multiple databases of TCGA, HPA, and so on.

# Methods

## Data sources

Gene data in tumors and normal tissues were obtained from The Cancer Genome Atlas database and the Genotype–Tissue Expression project database, and the data for tumor cell lines were from the Cancer Cell Line Encyclopedia database. TCGA (The Cancer Genome Atlas, www.cancer.gov) database was used for the expression differential analysis and validation of CCT2 expression correlation with clinical features, HPA (www.proteinatlas.org) database was used to detect CCT2 protein expression analysis, xCell database was used for the association of CCT2 expression with immune cell infiltration, and GEO database was used to download and analyze the transcriptome data of colon cancer GSE143985 and thyroid cancer GSE33630.

## CCT2 expression analysis

The tumor data were obtained from the TCCGA database, the paraneoplastic data were obtained from the TCGA dataset, and the expression differences of various tumors were analyzed by R software. The expression levels of CCT2 in different tissues and different tumor tissues were analyzed using the Kruskal–Wallis test.

## Prognostic analysis in pan-cancer

One-way Cox multivariate analysis was utilized to calculate the correlation between CCT2 expression and patient survival in 33 tumors, and Kaplan–Meier survival analysis was employed to compare the link between high and low CCT2 expression levels and tumor prognosis.

## Correlation analysis of immune microenvironment

The correlation between CCT2 expression and immune cell scores was analyzed by downloading the scores of six immune infiltrating cells from the xCell for 33 styles of cancer. The immune score and the stromal score of every tumor sample were analyzed using the R software package ESTIMATE to determine the correlation between CCT2 expression and immune score in 33 tumors. The correlation was considered significant and positive once $p$ <0.05 and $R$ >0.20.

## Immunohistochemical staining

Immunohistochemistry (IHC) pictures of CCT2 protein expression in normal and tumor tissues were downloaded from HPA to conduct a differential analysis of CCT2 expression at the protein level.

## Analysis of the relationship between CCT2 and clinical phenotype

Patient survival and clinical phenotype data were uploaded from the TCGA database. Two clinical phenotypes—tumor

staging and grading—were selected, and their relationship with CCT2 expression was analyzed using the R package "limma" and "ggpubr", the result of which was considered significant at $p < 0.05$.

## Correlation of CCT2 expression with tumor mutation burden and tumor microsatellite instability

The tumor mutation burden (TMB) scores were calculated using Perl scripts and corrected by dividing by the full length of exons. The MSI scores were determined for all samples, supported with corporal mutation data obtained from TCGA and an analysis of the link between CCT2 expression and TMB and MSI using Spearman's rank coefficient of correlation. Therefore, the resulting square measurements, shown as heat maps, using the R package "reshape2" and "RColorBrewer" were generated.

## Significance of CCT2 expression

Gene set enrichment analysis (GSEA) was performed to study the function of CCT2 in tumors. The Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets were downloaded from the official GSEA website, and functional analysis was performed using the R package "limma".

## RT-qPCR

Total RNA was isolated from cultured cells using Trizol reagent (Invitrogen) according to the instructions. One microgram of RNA was reverse-transcribed to cDNA using PrimerScript RT Master Mix (Takara, Dalian, China). RT-qPCR experiments were then performed according to the instructions using SYBR Premix Ex Taq (Takara, Dalian, China) on ABI 7500 (ABI, America) for qRT-PCR experiments, with GAPDH as an internal reference, and the relative expression of genes was calculated by the $2^{-\triangle\triangle Ct}$ method, respectively. The PCR primers are shown in Table 1.

TABLE 1 Primer sequence for RT-qPCR.

| Gene | | Primer sequence |
|---|---|---|
| CCT2 | Forward | 5'-GCACTACCTCTGTTACCGTTTT-3' |
| | Reverse | 5'-CTTCTCTCCAACCCGCTATGA-3' |
| GAPDH | Forward | 5'-CGGAGTCAACGGATTTGGTCGTAT-3' |
| | Reverse | 5'-AGCCTTCTCCATGGTGGTGAAGAC-3' |

# Results

## CCT2 pan-cancer expression

Compared with normal sample tissues, CCT2 was expressed in BLCA, BRCA, CHOL, CESC, COAD, ESCA, GBM, HNSC, KIRP, LIHC, LUAD, LUSC, PRAD, READ, STAD, and UCEC, which was upregulated in LAML and downregulated in KICH and THCA (Figure 1A). In addition, we also investigated CCT2 protein expression in the HPA cohort, which was expressed at varying degrees in all but glioma, pancreatic, renal, and prostate cancers (Figure 1B), a representative immunohistochemical map of which is shown in Figure 1C.

## The prognostic value of CCT2

The Cox regression analysis figured out that CCT2 expression was related to OS in 11 cancers: ACC, BRCA, HNSC, KICH, LICH, LUAD, MESO, OV, SARC, SKCM, and THYM (Figure 2A). The Kaplan–Meier survival curves showed that the CCT2 expression levels were significantly associated with prognosis in ACC, BRCA, KICH, LICH, LUAD, MESO, OV, SARC, and THYM, where a high CCT2 expression was associated with poor prognosis in cancers other than OV (Figure 2B).

## Pan-cancer analysis of CCT2 expression and clinicopathological correlation

To explore the correlation between CCT2 expression and the clinicopathological features of cancer, we evaluated CCT2 expression in patients with different grades and stages of cancer. The results revealed that CCT2 expression in BRCA, PRAD, and THYM was significantly different from tumor grade (Figure 3A). CCT2 expression in BRCA and THYM was significantly different from tumor stage (Figure 3B). There was a significant difference in CCT2 expression between T1 and T 2 stages in BRAC and an increase in CCT2 expression with stage in PRAD (Figure 3C). No significant correlation between tumor stage classification and CCT2 expression was found in patients with other cancers.

## Relationship between CCT2 expression level and tumor immune cell infiltration level

We analyzed the relationship between CCT2 expression levels and the infiltration levels of 26 immune-related cells. The results showed that the level of immune cell infiltration

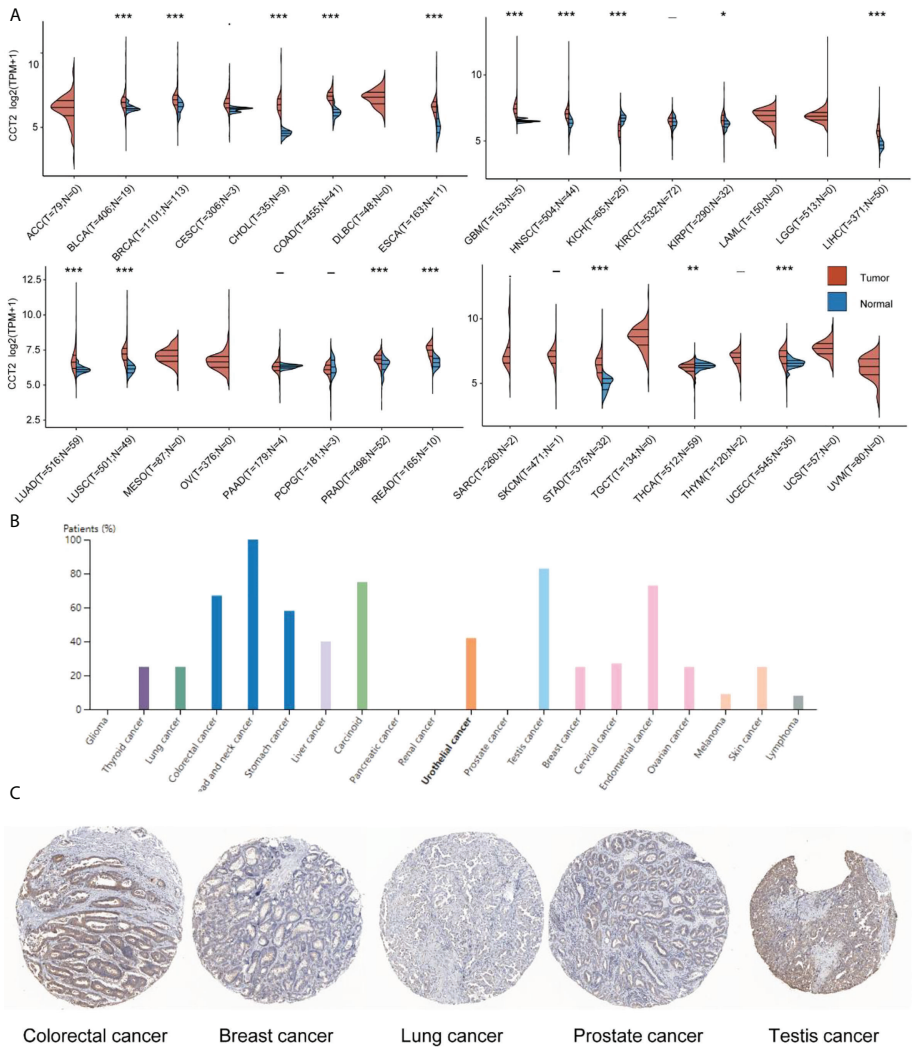



FIGURE 1
CCT2 expression in human pan-cancer. **(A)** Differential expression of CCT2 mRNA in tumor and normal tissues in the TCGA cohort. **(B)** Twenty cancer types in which CCT2 protein pachytene was present. **(C)** Representative immunohistochemical staining for CCT2 in Human Protein Atlas. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

was significantly correlated with CCT2 expression in most cancer types, and CCT2 expression was negatively correlated with the level of most immune cell infiltration processes, but the CCT2 expression levels were negatively correlated with the level of T cell CD4+ Th2 and common lymphoid progenito cell infiltration (Figure 4A). In addition, we also analyzed the relationship between CCT2 expression and immune-related genes in 33 tumors. The results are shown in Figure 4B, where most of the immune genes were positively correlated with the CCT2 expression profile in 33 tumors. The immune checkpoint results showed that the CCT2 expression levels showed a positive correlation with the immune checkpoints in most tumors, except for THYM and GBM (Figure 4C, $p < 0.05$).

## Relationship between CCT2 gene expression and immune neoantigens, TMB, and microsatellite instability

We analyzed the correlation between the CCT2 expression levels and immune neoantigens, TMB and MSI and found that all three were fundamentally associated with sensitivity to immune checkpoint inhibitors. Figure 5 shows the correlation between CCT2 expression and tumor immune neoantigens. It was shown that immune neoantigens were significantly positively correlated with BRCA, KIRC, KIRP, STAD, HNSC, PRAD, and LGG and significantly negatively correlated with THCA. Considering that TMB and MSI play an important role

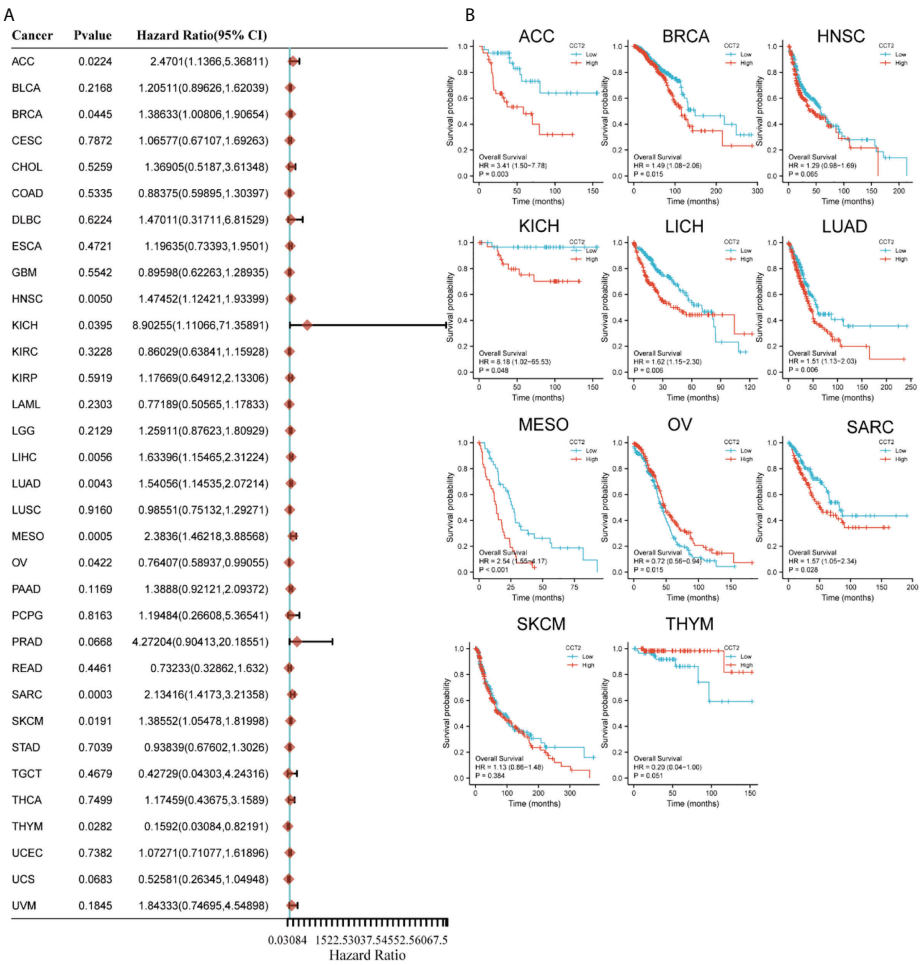



FIGURE 2
Relationship between CCT2 expression and overall survival time. **(A)** Forest plot of overall survival (OS) association of 33 tumors. **(B)** Kaplan–Meier analysis of CCT2 expression and OS.

in the process of tumor immunotherapy, we also examined the relationship of pan-cancer CCT2 expression with TMB and MSI, and the results are shown in Figure 6. CCT2 expression showed a positive correlation with TMB in most tumors and a negative correlation with BLCA, OC, ESCA, CESC, THCA, UVM, and THYM (Figure 6A). Similarly, CCT2 expression showed a positive correlation in MSI in most tumors and a negative correlation in BLCA, OC, ESCA, CESC, THCA, UVM, and THYM and in negative correlation in OV, SKCM, TGCT, ACC, THCA, BLCA, LUSC, HNSC, GBM, LAML, SARC, PRAD, LUAD, LGG, CHOL, PCPG, and DLBC (Figure 6B).

## GSEA analysis of the high and low expression of CCT2 gene in tumors

To study the effect of gene expression levels on tumor, we classified the samples into high and low parts and analyzed the enrichment of the KEGG and HALLMARK pathways between the two groups of high- and low-expression part by GSEA; the results are shown in Figure 7. The most significant top three pathways are visualized as follows: CCT2 gene mainly loads PYRIMIDINE METABOLISM, PURINE METABOLISM, HEDGEHOG SIGNALING, and other pathways.

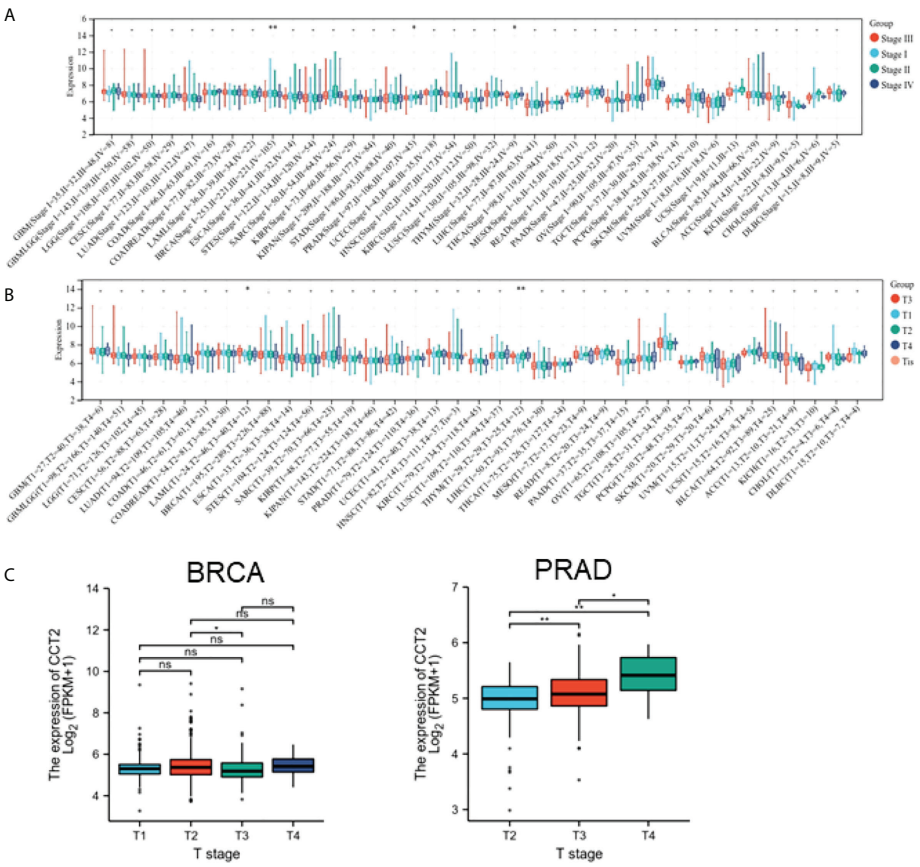



FIGURE 3
Association between CCT2 expression and clinical characteristics. **(A)** Association between CCT2 expression and tumor stage grading. **(B)** Association between CCT2 expression and BRCA and PRAD tumor stages. **(C)** Expression of BRCA and PRAD in different tumor stages. *$p < 0.05$, **$p < 0.01$. NS, no significant

## External dataset validates the heterogeneous role of CCT2 in cancer

Meanwhile, in our published study, we found that Liu et al. (15) analyzed the expression and prognosis of CCT2 in various cancers and focused on the role played by CCT2 in breast cancer. In our study, we systematically analyzed the prognosis of CCT2 expression in various cancers and elaborated the correlation between CCT2 expression and tumor immunity. We also performed validation in the GEO dataset, which showed that CCT2 was highly expressed in colorectal cancer and less in thyroid cancer compared with normal tissue (Figures 8A, B). We also verified the correlation between CCT2 expression in colon cancer and immune genes (Figure 8C). Subsequently, we also examined the expression levels of CCT2 in colon cancer cells (HCT116), thyroid cancer cells (TPC-1), and normal cell lines (NCM460, Nthy–cri3-1). The results showed that CCT2 was highly expressed in colon cancer cells and less expressed in thyroid cancer cells (Figures 8D, E).

# Discussion

Human TCP1-containing chaperonin containing TCP1 subunit 2 (CCT2) is an isoform of heat shock protein 60 in eukaryotic cells that is involved in the metabolism of many cells and is highly expressed in many malignant tissues, where tumorigenesis, development, and prognosis are closely related (7, 12, 13). In our study, CCT2 was highly expressed in 27 tumors, and this was confirmed by IHC results. In other studies, the expression levels of TRiC subunits TCP1, CCT2, CCT 3, CCT 4, CCT 5, CCT 6A, CCT 7, and CCT 8 were significantly upregulated in HCC (14). PAAD, READ, STAD, and UCEC are significantly upregulated in CCT2 expression than in adjacent normal tissues (9, 15).

Our analysis of Kaplan–Meier survival of tumors by TCGA data displayed that a high CCT2 expression was related to poor prognosis in ACC, GBMLGG, LIHC, LUAD, MESO, SARC, and THYM. Similarly, one study reported that, in breast cancer, a high expression of CCT2 predicted its poor prognosis (15). In gallbladder cancer, it was shown by multifactorial analysis

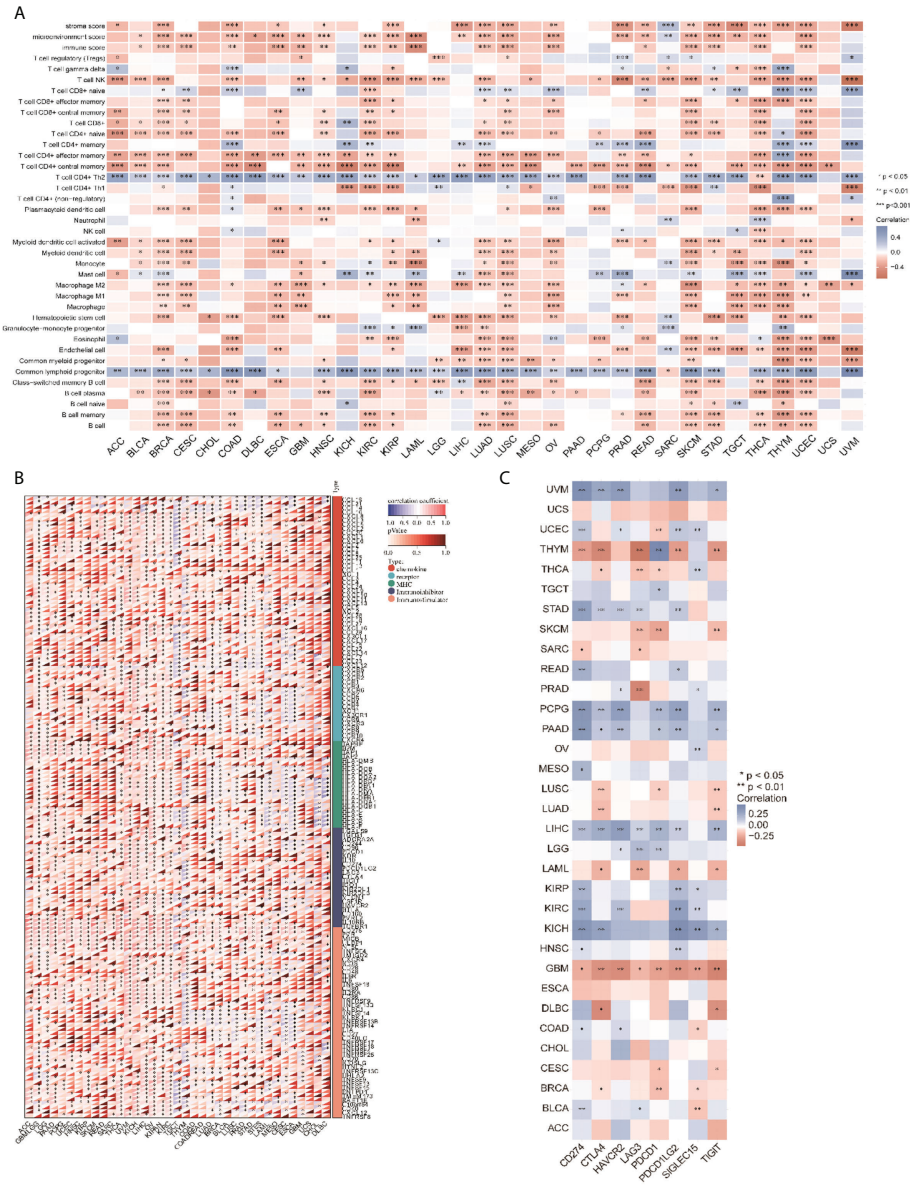



FIGURE 4
CCT2 expression and cancer immune correlation. **(A)** CCT2 expression levels in different tumors correlated with the infiltration levels of 26 immune-related cells. **(B)** CCT2 expression levels in different tumors correlated with immune genes. **(C)** CCT2 expression levels in different tumors correlated with immune checkpoints.

that a positive CCT2 expression was negatively correlated with low postoperative patient survival and positively correlated with high mortality (10). A study by Showalter et al. (16) overexpressed CCT2 by 1.3–1.8-fold in breast cancer cells by a lentiviral vector and found that cells overexpressing CCT2 were more aggressive and had a higher proliferation index, while CCT2 depletion in a homozygous mouse model of triple-negative breast cancer prevented tumor growth. Similarly, in a study by Guest et al. (17), they found that TCP1 and CCT2 are repeatedly altered in breast cancer and that TCP1 and CCT2 are required for the growth/survival of breast cancer cells *in vitro* and are determinants of overall survival in breast cancer patients. It has also been shown that TSPAN31 is highly expressed in gastric cancer (GC) tissues and that a high expression of TSPAN31 leads to a poor prognosis in GC patients. TSPAN31 regulates GC cell proliferation, migration, and apoptosis, and this regulatory mechanism is achieved by the co-expression of TSPAN31 (12). These data suggest that the

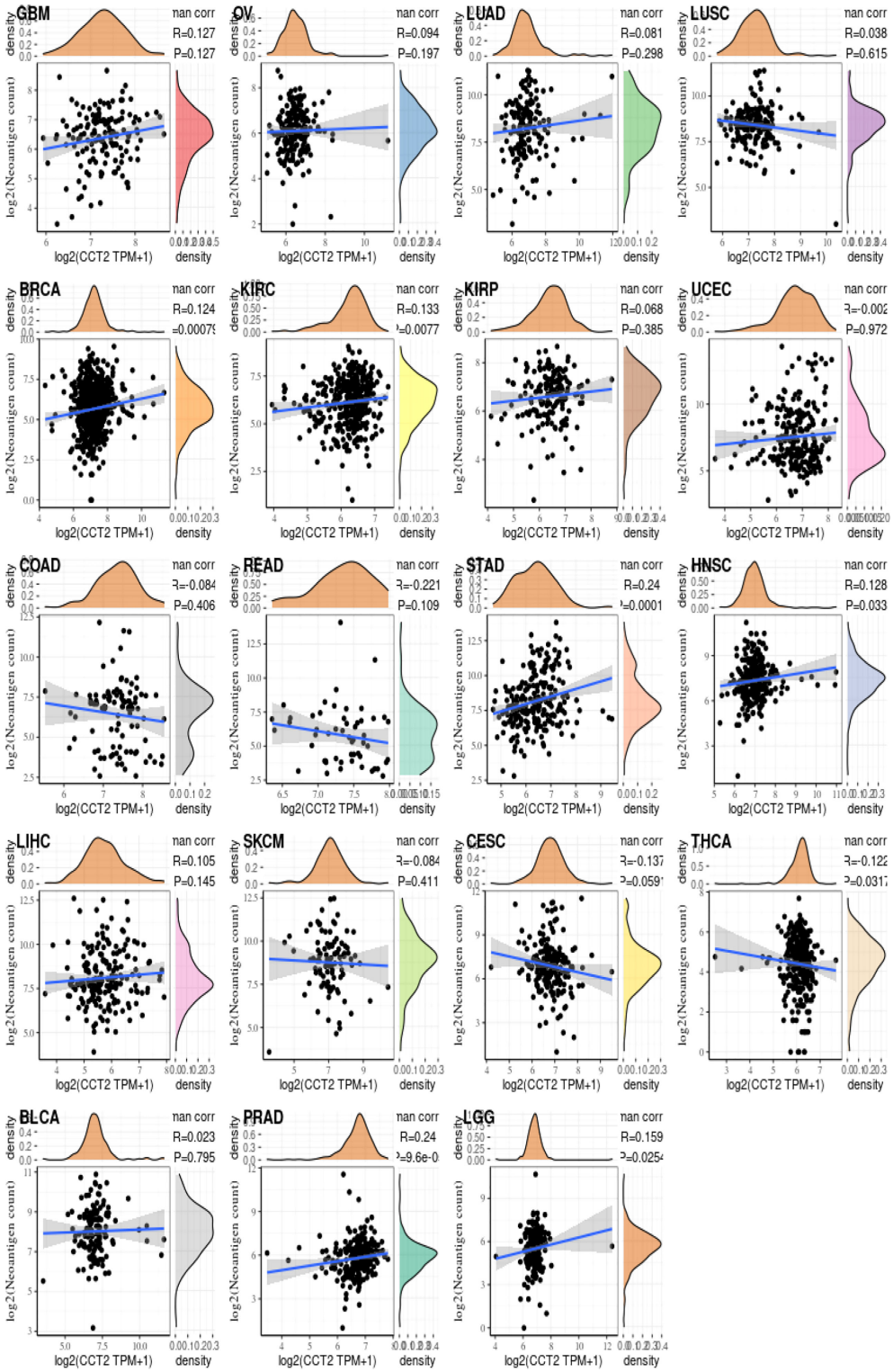



FIGURE 5
CCT2 expression correlates with immune neoantigens.

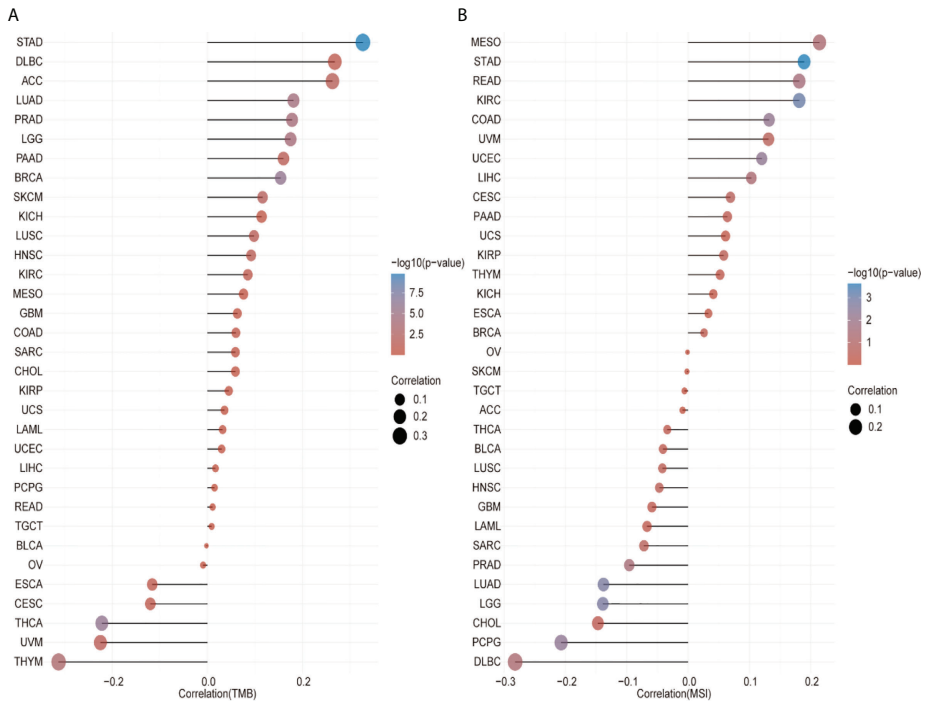



FIGURE 6
CCT2 expression correlates with tumor mutation burden (TMB) and microsatellite instability (MSI). **(A)** Lollipop plot of correlation between CCT2 expression and TMB level in pan-cancer analysis. **(B)** Lollipop plot of correlation between CCT2 expression and MSI level in pan-cancer analysis.

CCT2 subunit is an important component of chaperone protein activity and is required for some tumorigenesis.

A growing body of proof support TMB as a possible biomarker of response to immune checkpoint inhibitors in most cancers (18, 19). These studies counsel that a better burden of nonsynonymous mutations in tumors promotes inflated neoantigen formation, creating tumors that are a lot immunogenic and therefore up to the clinical response to therapy (20, 21). Within the present study, we tend to evaluate the association between CCT2 and TMB and established that CCT2 was not related to most tumor TMBs, except STAD, SKCM, PRAD, LUAD, LGG, HNSC, COAD, CESC, BRCA, and THYM. Tumor immune infiltrating cells play a very important role within the immune regulation of tumor tissues (22, 23). An increasing range of studies has found that tumor immune infiltrating cells square measure closely related to immune checkpoint suppression and prognostic efficaciousness (24–27). To explore the connection between CCT2 expression and multiple infiltrating lymphocytes, we have a tendency to analyze the relative fraction of infiltrating immune cell sorts in 33 cancer sorts exploitation CIBERSORT. We found these associations again to be tumor type dependent. In addition, our enrichment analysis suggests that CCT2 can potentially influence cancer etiology or pathogenesis through HEDGEHOG SIGNALING. It has been shown that, during hypoxia in colorectal cancer, hypoxia activates the hedgehog pathway, CCT2 helps protein folding by binding to the oncogenic protein gli1, and a high expression of CCT2 and glii-1 enhances tumor invasion and migration *in vivo* and *in vitro* (28). The above-mentioned data suggest that CCT2 expression levels are closely associated with the immune infiltration of tumor cells and affect patient prognosis.

A fundamental challenge in the diagnosis and treatment of cancer is currently to detect changes in gene expression during tumorigenesis and progression and their relationship with prognosis. The heterogeneity of different molecular signaling pathways in tumor progression and postoperative recurrence in low- and high-grade tumors is a hallmark of bladder cancer, which is also useful in assessing tumor prognosis and patient survival outcomes. The comprehensive molecular characterization of multiple cancer types and corresponding patient clinical data collected through TCGA, the HAP database collecting protein expression levels in each tissue, and the XCELL database analyzes the correlation of gene expression with immune cell infiltration, thus providing the possibility of genomic studies of cancer.

In this study, we tend to analyze the expression of CCT2 in normal and neoplastic tissues. We also evaluated the prognostic worth of CCT2 in pan-cancer as supported by the TCGA dataset. Subsequently, we analyzed the relationship between CCT2 expression levels and immune cell infiltration among

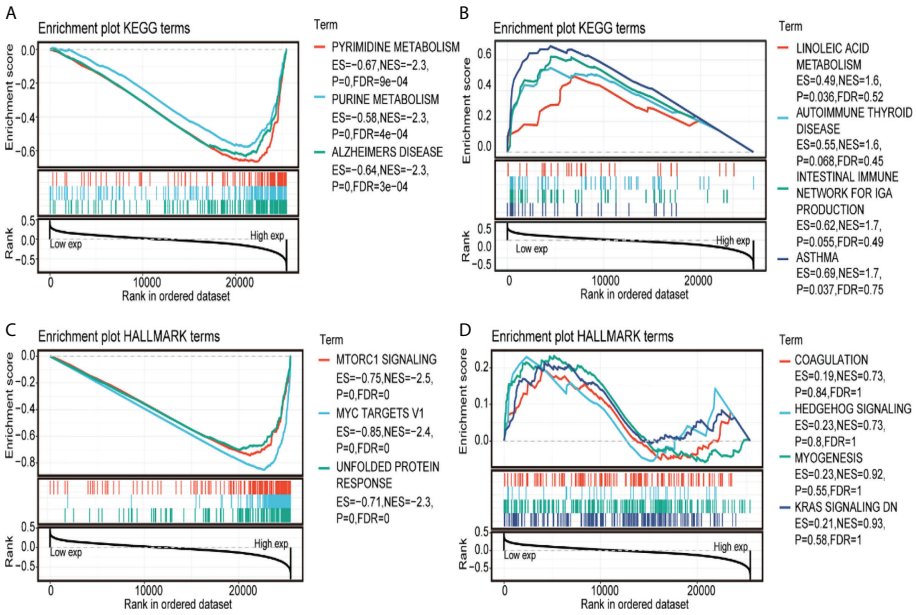



FIGURE 7
Gene set enrichment analysis (GSEA). **(A, B)** Kyoto Encyclopedia of Genes and Genomes annotation of CCT2 with GSEA. **(C, D)** HALLMARK annotation of CCT2 with GSEA.

others. The results recommend that CCT2 can be considered an associate independent prognostic factor for multiple tumors, in which the CCT2 expression levels are completely different in several tumors and predict different prognostic outcomes; however, this requires investigating the particular role of CCT2 in every cancer in several tumors. CCT2 expression is associated with TMB, MSI, and immune cell infiltration. Its effect on tumor immunity also varies by tumor type. These

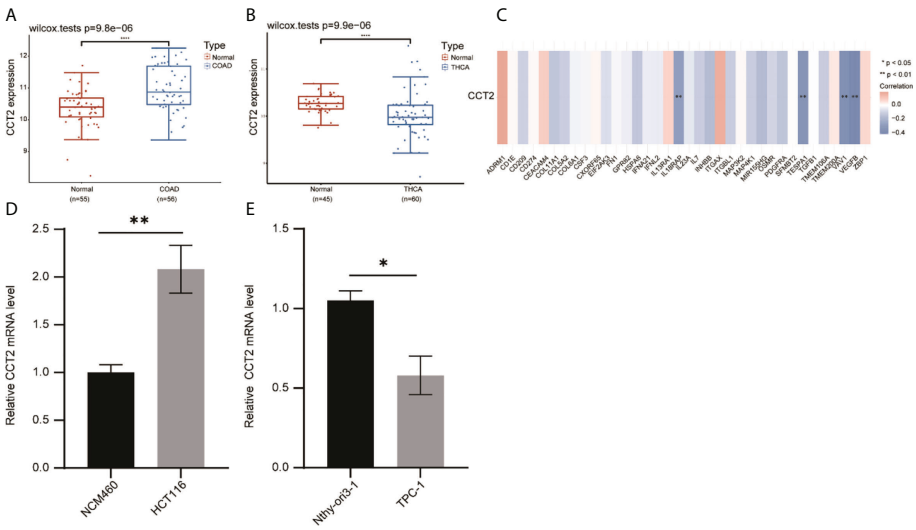



FIGURE 8
Validation of the correlation between CCT2 expression in tumors and immune genes. **(A)** Expression of CCT2 in colon cancer. **(B)** Expression of CCT2 in thyroid cancer. **(C)** Correlation between CCT2 expression and some immune genes in colon cancer. **(D, E)** RT-qPCR detection of CCT2 expression in colon cancer cells and thyroid cancer cells. *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$.

findings may lead to the realization of a more precise and personalized immunotherapy in the future.

## Data availability statement

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

## Author contributions

WL: conception and design,writing the article,data collection. LS: analysis and interpretation. JP: data collection, analysis and interpretation, writing the article. SW: writing the article, analysis and interpretation. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Natural Science Foundation of Gansu Province (20JR10RA724), the Second Hospital of Lanzhou University Cuiying Science and Technology Innovation Program Project (CY2020-MS06), and the Science and Technology Plan Project of Lanzhou City (2020-ZD-91).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Hou JY, Wu HY, He RQ, Lin P, Dang YW, Chen G. Clinical and prognostic value of chaperonin containing T-complex 1 subunit 3 in hepatocellular carcinoma: A study based on microarray and RNA-sequencing with 4272 cases. *Pathol Res Pract* (2019) 215(1):177–94. doi: 10.1016/j.prp.2018.11.006

2. Jin M, Liu C, Han W, Cong Y. TRiC/CCT chaperonin: Structure and function. *Subcell Biochem* (2019) 93:625–54. doi: 10.1007/978-3-030-28151-9_19

3. Yao L, Zou X, Liu L. The TCP1 ring complex is associated with malignancy and poor prognosis in hepatocellular carcinoma. *Int J Clin Exp Pathol* (2019) 12 (9):3329–43.

4. Gestaut D, Limatola A, Joachimiak L, Frydman J. The ATP-powered gymnastics of TRiC/CCT: An asymmetric protein folding machine with a symmetric origin story. *Curr Opin Struct Biol* (2019) 55:50–8. doi: 10.1016/j.sbi.2019.03.002

5. Willison KR. The substrate specificity of eukaryotic cytosolic chaperonin CCT. *Philos Trans R Soc Lond B Biol Sci* (2018) 373(1749):20170192. doi: 10.1098/rstb.2017.0192

6. Sergeeva OA, Chen B, Haase-Pettingell C, Ludtke SJ, Chiu W, King JA. Human CCT4 and CCT5 chaperonin subunits expressed in escherichia coli form biologically active homo-oligomers. *J Biol Chem* (2013) 288(24):17734–44. doi: 10.1074/jbc.M112.443929

7. Macario AJL, Conway de Macario E. Chaperonins in cancer: Expression, function, and migration in extracellular vesicles. *Semin Cancer Biol* (2021) S1044–579X(21):00159–0. doi: 10.1016/j.semcancer.2021.05.029

8. Klimczak M, Biecek P, Zylicz A, Zylicz M. Heat shock proteins create a signature to predict the clinical outcome in breast cancer. *Sci Rep* (2019) 9(1):7507. doi: 10.1038/s41598-019-43556-1

9. Carr AC, Khaled AS, Bassiouni R, Flores O, Nierenberg D, Bhatti H, et al. Targeting chaperonin containing TCP1 (CCT) as a molecular therapeutic for small cell lung cancer. *Oncotarget* (2017) 8(66):110273–88. doi: 10.18632/oncotarget.22681

10. Zou Q, Yang ZL, Yuan Y, Li JH, Liang LF, Zeng GX, et al. Clinicopathological features and CCT2 and PDIA2 expression in gallbladder squamous/adenosquamous carcinoma and gallbladder adenocarcinoma. *World J Surg Oncol* (2013) 11:143. doi: 10.1186/1477-7819-11-143

11. Wang S, Zhen L, Li X, Fu X, Li P, Zhang D. Search for key genes and functional pathways of ulcerative colitis to colon cancer based on bioinformatics. *Front Oncol* (2022) 12:857148. doi: 10.3389/fonc.2022.857148

12. Ma X, Qiu S, Tang X, Song Q, Wang P, Wang J, et al. TSPAN31 regulates the proliferation, migration, and apoptosis of gastric cancer cells through the METTL1/CCT2 pathway. *Transl Oncol* (2022) 20:101423. doi: 10.1016/j.tranon.2022.101423

13. Chang YX, Lin YF, Chen CL, Huang MS, Hsiao M, Liang PH. Chaperonin-containing TCP-1 promotes cancer chemoresistance and metastasis through the AKT-GSK3beta-beta-Catenin and XIAP-survivin pathways. *Cancers (Basel)* (2020) 12(12):3865. doi: 10.3390/cancers12123865

14. Liu J, Huang L, Zhu Y, He Y, Zhang W, Lei T, et al. Exploring the expression and prognostic value of the TCP1 ring complex in hepatocellular carcinoma and overexpressing its subunit 5 promotes HCC tumorigenesis. *Front Oncol* (2021) 11:739660. doi: 10.3389/fonc.2021.739660

15. Liu Q, Qi Y, Kong X, Wang X, Zhang W, Zhai J, et al. Molecular and clinical characterization of CCT2 expression and prognosis *via* Large-scale transcriptome profile of breast cancer. *Front Oncol* (2021) 11:614497. doi: 10.3389/fonc.2021.614497

16. Showalter AE, Martini AC, Nierenberg D, Hosang K, Fahmi NA, Gopalan P, et al. Investigating chaperonin-containing TCP-1 subunit 2 as an essential component of the chaperonin complex for tumorigenesis. *Sci Rep* (2020) 10 (1):798. doi: 10.1038/s41598-020-57602-w

17. Guest ST, Kratche ZR, Bollig-Fischer A, Haddad R, Ethier SP. Two members of the TRiC chaperonin complex, CCT2 and TCP1 are essential for survival of breast cancer cells and are linked to driving oncogenes. *Exp Cell Res* (2015) 332 (2):223–35. doi: 10.1016/j.yexcr.2015.02.005

18. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. *Ann Oncol* (2019) 30(1):44–56. doi: 10.1093/annonc/mdy495

19. Sha D, Jin Z, Budczies J, Kluck K, Stenzinger A, Sinicrope FA. Tumor mutational burden as a predictive biomarker in solid tumors. *Cancer Discovery* (2020) 10(12):1808–25. doi: 10.1158/2159-8290.CD-20-0522

20. Liu L, Bai X, Wang J, Tang XR, Wu DH, Du SS, et al. Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. *Clin Cancer Res* (2019) 25(24):7413–23. doi: 10.1158/1078-0432.CCR-19-0558

21. 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

22. Mao X, Xu J, Wang W, Liang C, Hua J, Liu J, et al. Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: New

findings and future perspectives. *Mol Cancer* (2021) 20(1):131. doi: 10.1186/s12943-021-01428-1

23. Garaud S, Buisseret L, Solinas C, Gu-Trantien C, de Wind A, Van den Eynden G, et al. Tumor infiltrating b-cells signal functional humoral immune responses in breast cancer. *JCI Insight* (2019) 5(18):e129641. doi: 10.1172/jci.insight.129641

24. Zhang Y, Zhang Z. The history and advances in cancer immunotherapy: Understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell Mol Immunol* (2020) 17(8):807–21. doi: 10.1038/s41423-020-0488-6

25. Liu X, Bao X, Hu M, Chang H, Jiao M, Cheng J, et al. Inhibition of PCSK9 potentiates immune checkpoint therapy for cancer. *Nature* (2020) 588(7839):693–8. doi: 10.1038/s41586-020-2911-7

26. Kim W, Chu TH, Nienhüser H, Jiang Z, Del Portillo A, Remotti HE, et al. PD-1 signaling promotes tumor-infiltrating myeloid-derived suppressor cells and gastric tumorigenesis in mice. *Gastroenterology* (2021) 160(3):781–96. doi: 10.1053/j.gastro.2020.10.036

27. Srivastava S, Furlan SN, Jaeger-Ruckstuhl CA, Sarvothama M, Berger C, Smythe KS, et al. Immunogenic chemotherapy enhances recruitment of CAR-T cells to lung tumors and improves antitumor efficacy when combined with checkpoint blockade. *Cancer Cell* (2021) 39(2):193–208.e10. doi: 10.1016/j.ccell.2020.11.005

28. Park SH, Jeong S, Kim BR, Jeong YA, Kim JL, Na YJ, et al. Activating CCT2 triggers gli-1 activation during hypoxic condition in colorectal cancer. *Oncogene* (2020) 39(1):136–50. doi: 10.1038/s41388-019-0972-6

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fonc.2022.965427

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Claudia Peitzsch,
Center for Regenerative Therapies Dresden, Technical University Dresden, Germany
Junjie Gu,
Nanjing Medical University, China

*CORRESPONDENCE
Yi Yu
172428315@qq.com
Jinquan Liu
609149209@qq.com

†These authors have contributed equally to this work

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

RECEIVED 09 June 2022
ACCEPTED 15 August 2022
PUBLISHED 02 September 2022

CITATION
He L, Liang M, Guo W, Liu J and Yu Y (2022) HOXA1 is a radioresistance marker in multiple cancer types. *Front. Oncol.* 12:965427. doi: 10.3389/fonc.2022.965427

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

# HOXA1 is a radioresistance marker in multiple cancer types

Lu He[1†], Min Liang[2,3†], Weisheng Guo[1], Jinquan Liu[1*] and Yi Yu[1*]

[1]Department of Radiotherapy, Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, China, [2]Department of Oncology, The Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou Medical University, Guangzhou, China, [3]Key Laboratory of Biological Targeting Diagnosis, Therapy and Rehabilitation of Guangdong Higher Education Institutes, The Fifth Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

Radiotherapy is an important therapeutic method for patients with cancer. However, radioresistance can cause treatment failure. Thus, there is an urgent need to investigate mechanisms of radioresistance and identity markers that could be used to predict radioresistance and prognosis of post-radiotherapy cancer patients. In the present study, we propose HOXA1 as a candidate biomarker of intrinsic radioresistance in multiple cancer types. By analyzing data from The Cancer Genome Atlas (TCGA), we found that HOXA1 was aberrantly upregulated in multiple cancers, and that elevated HOXA1 was significantly associated with poor prognosis of post-radiotherapy head and neck squamous cell carcinoma (HNSCC) and low-grade glioma (LGG) patients. Correlation analysis showed that HOXA1 expression was positively correlated with expression of EGFR, CDK6, and CAV1, which have been reported to enhance radioresistance. In addition, gene set enrichment analysis (GSEA) showed that the oxidative phosphorylation gene set was negatively enriched in HOXA1 high-expression samples in both HNSCC and LGG. Moreover, immunohistochemical assays indicated that high HOXA1 expression was significantly correlated with a high recurrence rate of nasopharyngeal carcinoma (NPC) after radiotherapy. Further *in vitro* experiments demonstrated that HOXA1 knockdown markedly attenuated the DNA repair capacity of NPC cells and sensibilized NPC cells to irradiation. Taken together, the results of this study demonstrate that HOXA1 has potential to be a predictive marker for radioresistance and post-radiotherapy prognosis that could help to guide individualized treatment in multiple cancer types.

KEYWORDS

**radioresistance, HOXA1, biomarker, transcription factor, radiotherapy**

# Introduction

About half of cancer patients receive radiotherapy at some point after their diagnosis of cancer (1). For patients with nasopharyngeal carcinoma (NPC), radiotherapy is the preferred radical treatment method because of the anatomical location of the cancer (2). In addition, radiotherapy is an important adjuvant therapy method used to prevent in situ recurrence of many types of cancer, including head and neck squamous cell carcinoma (HNSCC) and glioma (3, 4).

Recent advances in radiotherapy have led to significant benefits for many cancer patients, enabling improved tumor control with reduced toxic effects (5). However, the response to radiotherapy varies owing to inter-individual differences in the radiosensitivity of tumor cells. For example, residual tumor cells might persist in some patients even after irradiation with a total dose of 80 Gy, whereas other patients might achieve complete remission after only 40 Gy irradiation (6). In addition, radioresistance, which induces local treatment failure and results in residual or recurrent tumors, remains an important problem in radiotherapy. For NPC, 7.4–14% of patients who receive radiotherapy develop local recurrence or persistent disease (7, 8). Therefore, it is important to understand the mechanisms of tumor radioresistance to enable development of resistance-reversal strategies and identification of radioresistance biomarkers that could be used to guide individualized radiotherapy.

The homeobox genes encode a highly conserved subgroup of homeobox transcription factors, which regulate multiple processes including development, differentiation, apoptosis, motility, angiogenesis, and even carcinogenesis (9). Homeobox A1 (HOXA1) is a member of the A cluster of homeobox transcription factors. HOXA1 expression is extremely low during normal growth and differentiation but is detectable in a variety of human cancer lesions (10). Studies have reported that HOXA1 functions as an oncogene. In breast cancer, HOXA1 mediates oncogenic transformation via activation of the p44/42 MAP kinase pathway and is related to endocrine therapy resistance (11, 12). Moreover, elevated HOXA1 expression promotes cell proliferation in gastric cancer and drives tumor growth and metastasis in melanoma (13, 14). In addition, aberrant upregulation of HOXA1 is correlated with poor prognosis of patients with hepatocellular carcinoma, prostate cancer, and HNSCC (15–18). Although these studies suggest that upregulation of HOXA1 promotes cancer progression, the role of HOXA1 in determining the biological properties of malignant tumors is not fully understood, nor has the correlation between HOXA1 expression and cancer radioresistance been reported.

In this study, we identified HOXA1 as a new intrinsic radioresistance marker of multiple cancer types. Reanalysis of data from The Cancer Genome Atlas (TCGA) and the results of *in vitro* experiments suggest that high expression of HOXA1 contributes to radioresistance in NPC, HNSCC, and LGG.

# Materials and methods

## Cell culture and RNA interference

CNE1 (CVCL_6888) and HNE1 (CVCL_0308) human NPC cell lines were obtained from the State Key Laboratory of Oncology in South China, Sun Yat-sen University Cancer Centre. The cells were confirmed to be negative for mycoplasma with a PCR Mycoplasma Detection Kit (Beyotime, Shanghai, China) and cultured in RPMI 1640 supplemented with 10% fetal bovine serum (Thermo Fisher Scientific, Waltham, MA, USA) at 37°C under 5% CO2.

For RNA interference, cells were transfected with a chemically synthesized scrambled short interfering RNA (siRNA) or HOXA1 siRNA using Lipofectamine 3000 (Thermo Fisher Scientific) in accordance with the manufacturer's instructions. The HOXA1 siRNA was synthesized by Ribobio Co. Ltd., and the sequence (sense: 5′ GUUCCUUUCAGAUGACCUU 3′) was as previously reported. The scrambled control siRNA was purchased from Ribobio Co. Ltd. The knockdown efficiency was validated by quantitative PCR (qPCR) assays and confirmed to be more than 75%.

## Cell irradiation and clonogenic survival assay

Cell irradiation and clonogenic survival assays were performed as previous reported (19). The indicated cells were seeded into 3.5cm culture dishes at different densities (100, 100, 200, 1000, 10000 and 100000 cells for 0, 0.5, 1, 2, 4, 6 and 8 Gy irradiation, respectively). After the cells became adherent, the cells were irradiated at defined doses using a Rad Source (Rad Source Tech, Suwanee, GA, USA) R2000 X-ray irradiator (1.1 Gy/min., 160 kV, 25 mA, 0.3 mm copper filters). After 14 days incubation, the cells were fixed and stained with crystal violet stain. Then the Colonies were scored. The plating efficiency (PE) was calculated by dividing the number of counted colonies by the number of cells plated. The surviving fractions (SF) were then calculated by dividing the PE by the PE of the non-irradiated control. The radiation dose-clonogenic survival curves were fit to a linear-quadratic model as previously described. The curves were compared using the extra sum-of-squares F test in GraphPad Prism 8.0 (GraphPad, La Jolla, CA, USA).

## Patient specimens

Paraffin-embedded NPC specimens used for immunohistochemistry assays were obtained from 70 NPC patients without distant metastases who had undergone

conventional radiotherapy at the Affiliated Cancer Hospital of Guangzhou Medical University from 2013 to 2015 (Supplementary Table 1). All patients were monitored for more than 5 years, during which time 20 of them experienced local recurrence. This study was approved by the Ethics Committee of Affiliated Cancer Hospital of Guangzhou Medical University (ZN2022-22).

## Immunohistochemistry assays

Formalin-fixed, paraffin-embedded NPC tissue specimens were analyzed by immunohistochemistry as previously described (20). Briefly, a primary antibody against HOXA1 (ab230513; rabbit polyclonal; working dilution 1:100; Abcam, Cambridge, UK) was used for immunohistochemistry assays, and a non-biotin horseradish peroxidase detection system (DAKO, Glostrup, Denmark) was used to detect the expression level of HOXA1 protein. Both the extent and intensity of immunostaining were taken into consideration when analyzing the data. The intensity of staining was scored from 0 to 3, and the extent of staining was scored from 0% to 100%. The final quantitation of each stain was obtained by multiplying the two scores. HOXA1 expression was classified as high if the score was higher than 1.5, whereas scores of 1.5 or less indicated low expression.

## Cell proliferation assays

Three thousand of the indicated cells were seeded per well in 96-well plates. after 24h, 48 and 72h of incubation, cell proliferation was assessed using CCK8 reagents (Dojindo, Kumamoto, Japan) according to the manufacturer's instructions. Three replicates were used for each experiment.

## Immunofluorescence assays

Cells were trypsinized and seeded on glass chamber slides. After 24 h, the cells were irradiated at a dose of 2 Gy. Immunofluorescence analysis of γH2AX was performed 0.5 and 24 h after irradiation. The immunofluorescence analysis was performed as previously described (19). Briefly, the cells were fixed with cold methanol for 10 min, followed by blocking in blocking buffer for 30 min. Next, the cells were incubated with a rabbit monoclonal antibody against phospho-H2A.X (1:200; Cell Signaling Technology; 20E3, #9718) and anti-rabbit Alexa 555-conjugated secondary antibody (Thermo Fisher Scientific). Then, nuclei were counterstained with DAPI solution (Thermo Fisher Scientific), and the slides were mounted with antifade mounting solution (Thermo Fisher Scientific). Images were taken using a Zeiss LSM800 confocal imaging system. Cells with more than ten γH2AX foci were defined as γH2AX-positive cells. Five random fields were examined to estimate the number of γH2AX-positive cells per field for each slide.

## Real-time qPCR

Total mRNA of cells was isolated using TRIzol reagent (Thermo Fisher Scientific). One microgram of RNA was used for cDNA synthesis with a Transcriptor First Stand cDNA Synthesis Kit (Thermo Fisher Scientific). Real-time qPCR assays were performed according to the manufacturer's instructions using SYBR Green mix and LightCycler480 II system (Roche, Basel, Switzerland). The results were normalized to GAPDH expression using the 2-ΔΔCT method. Samples were run in triplicate. The following primers were used: HOXA1 forward, TCCTGGAATACCCCATACTTAGC; HOXA1 reverse, GCACGACTGGAAAGTTGTAATCC. GAPDH forward, GGAGCGAGATCCCTCCAAAAT; GAPDH reverse, GGCTGTTGTCATACTTCTCATGG. EGFR forward, AGGCACGAGTAACAAGCTCAC; EGFR reverse, ATGAGGACATAACCAGCCACC. CDK6 forward, CCAGATGGCTCTAACCTCAGT; CDK6 reverse, AACTTCCACGAAAAAGAGGCTT. CAV1 forward, GCGACCCTAAACACCTCAAC; CAV1 reverse, ATGCCGTCAAAACTGTGTGTC.

## Public data analysis

TCGA and Genotype-Tissue Expression data were acquired from UCSC Xena (21). Expression and survival data for HNSCC and low-grade glioma (LGG) cases were obtained from the TCGA Research Network (http://cancergenome.nih.gov/). In the HNSCC cohort, 295 patients had undergone postoperative radiotherapy, 149 patients did not undergo radiotherapy, and relevant information was lacking for 56 patients. In the LGG cohort, 302 patients had undergone postoperative radiotherapy, 178 patients did not undergo radiotherapy, and relevant information was lacking for 30 patients. We only included patients for whom adequate information was available in this study. Expression data were analyzed using R (version 3.6.3, http://www.r-project.org/) and R package DESeq2 (22); and gene ontology (GO) analysis and gene set enrichment analysis (GSEA) were performed using R and R package ClusterProfiler (23). The protein–protein interaction (PPI) network of the overlapping different expression genes was constructed based on the STRING online database (24).

## Statistics

At least three replicates of each experiment were conducted. Student's t-tests or chi-squared tests were used to compare

differences, as appropriate. Survival curves were constructed using the Kaplan–Meier method and analyzed by log-rank test or Cox regression model. The correlation coefficients between HOXA1 and EGFR, HOXA1 and CDK6, and HOXA1 and CAV1 expression levels were calculated using Pearson's correlation test. The two-tailed chi-squared test was used to analyze the association of HOXA1 expression with clinical parameters. Dose–survival curves were compared using the extra sum-of-squares F-test in GraphPad Prism 8.0 (GraphPad Software). P<0.05 was considered to indicate a significant difference.

# Results

## Pattern and prognostic significance of HOXA1 expression in human cancers

Expression analysis based on TCGA data indicated that HOXA1 was significantly upregulated in most cancer types (Figure 1A), including CESC (cervical squamous cell carcinoma and endocervical adenocarcinoma), LGG (low-grade glioma), GBM (Glioblastoma multiforme), and HNSCC. The results also indicated that HOXA1 was downregulated in BRCA (breast-invasive carcinoma), KICH (kidney chromophobe), OV (ovarian serous cystadenocarcinoma), PRAD (prostate adenocarcinoma), THCA (thyroid carcinoma), and UCEC (uterine corpus endometrial carcinoma) compared with adjacent non-tumor tissues (Figure 1A).

Moreover, we analyzed the prognostic value of HOXA1 expression in different human cancers using TCGA data (Figure 1B). We found that high HOXA1 expression was a risk factor for poor prognosis in multiple cancers, including CESC, LGG, and UVM (uveal melanoma).

Taken together, these results suggest that high HOXA1 expression could serve as a biomarker for poor prognosis in several cancers, and that the prognostic significance of HOXA1 expression depended on the cancer type.

## High HOXA1 expression predicted radioresistance in HNSCC and LGG

We also found that high HOXA1 expression was associated with radioresistance in HNSCC and LGG. Based on TCGA data, we found that high HOXA1 expression was correlated with poor

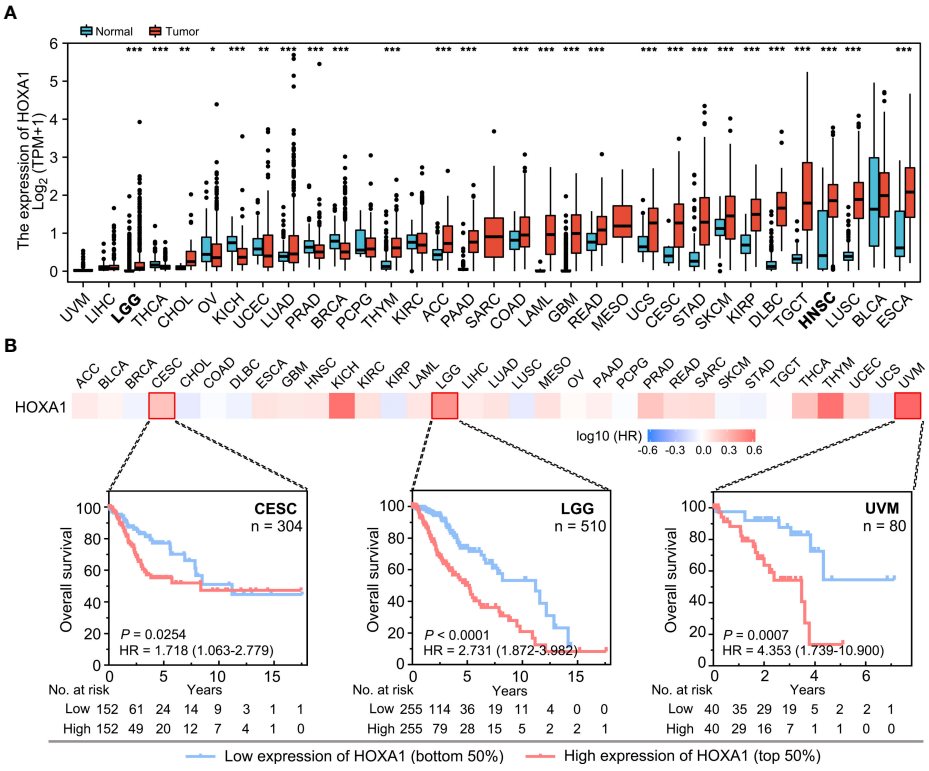



FIGURE 1
HOXA1 expression and the prognostic significance of HOXA1 in different cancer types. **(A)** HOXA1 expression in different cancer types. Data are from TCGA. **(B)** Heatmap of hazard ratios for overall survival according to HOXA1 expression in different cancers. Kaplan–Meier survival curves for CESC, LGG, and UVM are shown. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

prognosis and a high recurrence rate in HNSCC and LGG patients who had undergone postoperative radiotherapy (Figures 2A, B). However, HOXA1 expression did not have prognostic significance in HNSCC and LGG patients that did not undergo radiotherapy (Figures 2A, B). These results demonstrate that HOXA1 promotes radioresistance and increases local recurrence rates in patients with certain tumor types.

## Transcriptomes of HNSCC and LGG tissues with varying HOXA1 expression

HOXA1 has been identified as a transcription factor that regulates the expression of various genes (25). To further investigate the downstream genes of HOXA1 and the molecular mechanism by which HOXA1 affects radiosensitivity, we identified the differentially expressed genes (DEGs) between samples with high (highest quartile) and low (lowest quartile) HOXA1 expression in HNSCC and LGG based on TCGA data, respectively. We found 424 significantly upregulated (log2 fold change>1; P<0.05) and 3186 significantly downregulated (log2 fold change<−1; P<0.05) genes in the HNSCC cohort (Figure 3A). On the other hand, we found 4121 significantly upregulated (log2 fold change>1; P<0.05) and 2584 significantly downregulated (log2 fold change<−1; P<0.05) genes in the LGG cohort (Figure 3B). Among these DEGs, we identified 71 overlapping upregulated DEGs and 304 downregulated DEGs (Figure 3C). Based on these data, we found that HOXA1 expression was significantly correlated with some well-reported radioresistance-associated genes including EGFR, CDK6, and CAV1 (Figures 3D, E) (26–28). This suggests that the overlapping DEGs could be key downstream genes of HOXA1 that affect the radiosensitivity of cancer cells.

## HOXA1-associated signaling pathway

To investigate the HOXA1-associated signaling pathway, we performed GO enrichment analysis on the overlapping DEGs (P<0.05) of the HNSCC and LGG cohorts. Several terms associated with DNA damage and repair, including signal transduction in response to DNA damage, DNA integrity checkpoint, and DNA damage checkpoint, were enriched in the overlapping upregulated genes (Figure 4A), whereas terms including co-translational protein targeting to membrane, protein targeting to ER, and protein localization to

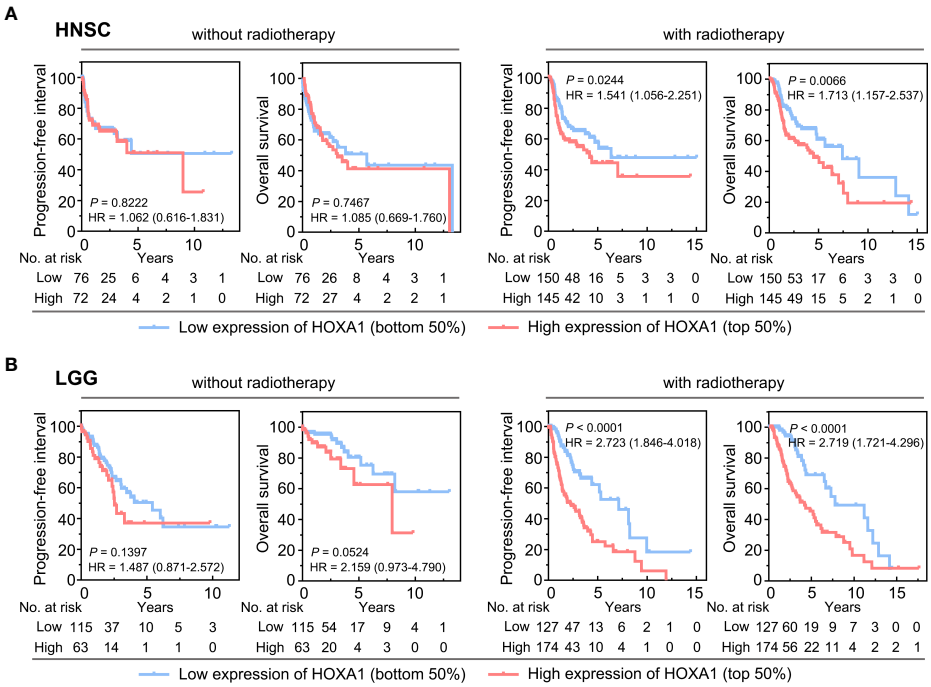



FIGURE 2
HOXA1 expression had prognostic significance in HNSCC and LGG patients receiving postoperative radiotherapy. **(A)** Overall survival (OS) and progression-free interval (PFI) curves of HNSCC patients with and without postoperative radiotherapy. **(B)** OS and PFI curves of LGG patients with and without postoperative radiotherapy.

10.3389/fonc.2022.965427

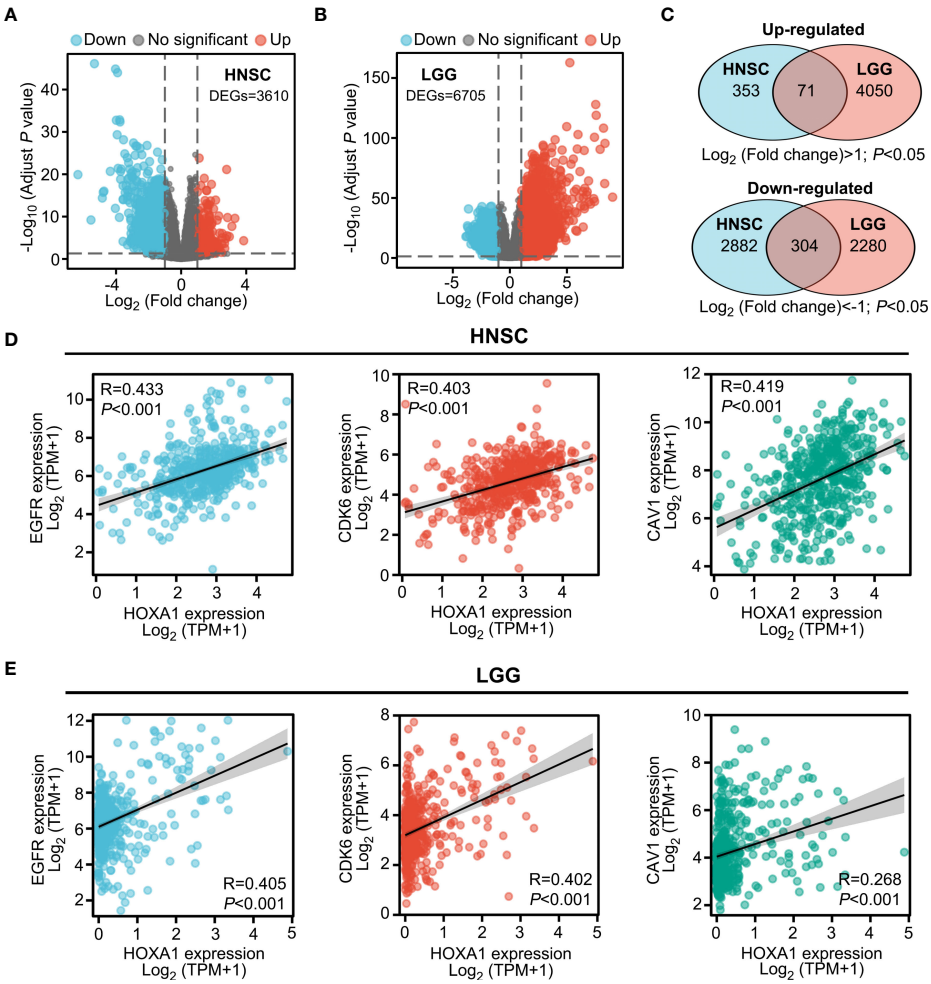



FIGURE 3
Transcriptomes of HNSCC and LGG tissues with varying HOXA1 expression. **(A, B)** Volcano plot showing the DEGs ($P < 0.05$, $\log_2$ fold change< −1 or $\log_2$ fold change >1) between high HOXA1 expression (top 25%) and low HOXA1 expression (bottom 25%) tumor tissue samples from HNSCC or LGG patients. **(C)** Venn plot depicting the overlapping genes with similar trends of expression changes in HNSCC and LGG samples with varying HOXA1 expression. **(D)** HOXA1 expression was significantly correlated with EGFR, CDK6, and CAV1 expression in HNSCC tissues. **(E)** HOXA1 expression was significantly correlated with EGFR, CDK6, and CAV1 expression in LGG tissues.

endoplasmic reticulum were enriched in the overlapping downregulated genes (Figure 4B).

We also performed GSEA using MSigDB hallmark gene sets (Figure 4C). The oxidative phosphorylation gene set was negatively enriched in HOXA1 high-expression samples in both HNSCC and LGG cohorts (Figure 4D, E). Previous studies have reported that enhanced oxidative phosphorylation could enhance the radiosensitivity of glioma cells (29). Hence, HOXA1 could induce radioresistance of cancer cells via inhibition of oxidative phosphorylation. In addition, we constructed a PPI network containing the most significant overlapping DEGs (log2 fold change>1 or log2 fold change<−1; P<0.05) to identify the most significant clusters of the overlapping DEGs of HNSCC and LGG (Supplementary Figure 1).

## Knockdown of HOXA1 enhanced the radiosensitivity of NPC cells

Radiotherapy is the principal treatment modality for NPC. To determine whether HOXA1 affected the radiosensitivity of NPC, we performed immunohistochemistry assays to examine HOXA1 expression in 20 NPC samples from patients who experienced local recurrence within 5 years of radiotherapy and 50 samples from patients who achieved long-term recurrence-free survival (Figure 5A). Our results indicated that high HOXA1 expression was significantly correlated with poor overall survival and high recurrence risk of NPC after radiotherapy (Figure 5B, C). This confirmed that HOXA1 enhanced radioresistance of cancer cells. Moreover, we

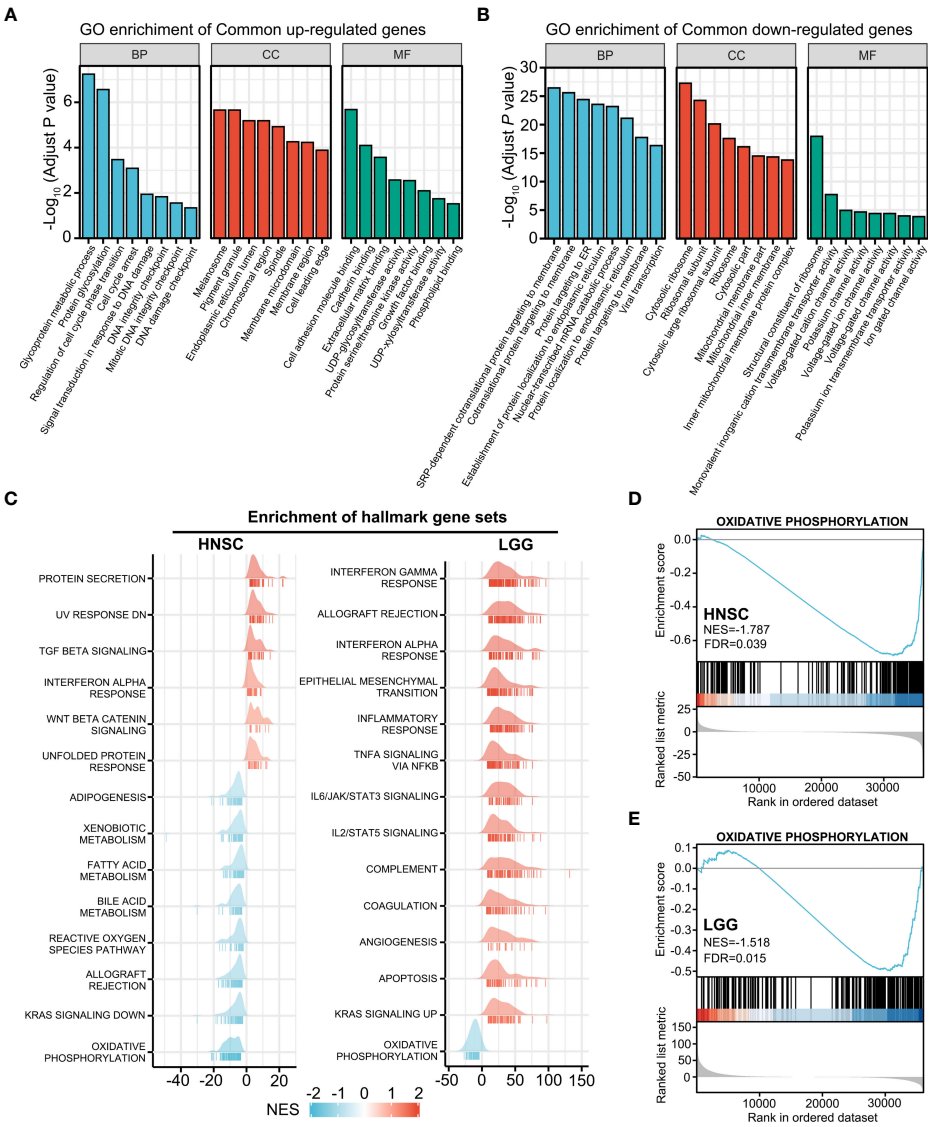



FIGURE 4
HOXA1-associated signaling pathway. **(A)** GO enrichment analysis of overlapping upregulated genes ($P < 0.05$) of the HNSCC and LGG cohorts. **(B)** GO enrichment analysis of the overlapping downregulated genes ($P < 0.05$) of the HNSCC and LGG cohorts. **(C)** Ridgeline plots depicting the significantly enriched signaling pathways involving the HOXA1-associated DEGs in HNSCC and LGG as revealed by GSEA performed using 50 hallmark gene sets. **(D, E)** GSEA plots showing that the oxidative phosphorylation gene set was negatively enriched in HNSCC and LGG samples with high HOXA1 expression.

knocked down HOXA1 in CNE1 and HNE1 NPC cells using a siRNA targeted against HOXA1. HOXA1 knockdown significantly reduced the expression levels of EGFR, CDK6, and CAV1 (Figure 5D, F). We also assessed the proliferation capacity of the cells and found that HOXA1 knockdown reduced the proliferation of CNE1 and HNE1 cells (Figures 5E, G). Next, we detected the radiosensitivity of these cells via colony formation assays. The dose–survival curves suggested that HOXA1 knockdown significantly sensitized these NPC cells to irradiation (Figures 5H, I).

## HOXA1 knockdown attenuated the DNA repair capacity of NPC cells

DNA double-strand breaks (DSBs) are the main lesions induced by irradiation; therefore, capacity to repair DSBs is closely related to radiosensitivity (19). In addition, our results indicated that several genes associated with DNA repair were enriched in the samples with high HOXA1 expression (Figure 4A). Thus, we detected the DNA damage response induced by 2 Gy irradiation in control and HOXA1-knockdown

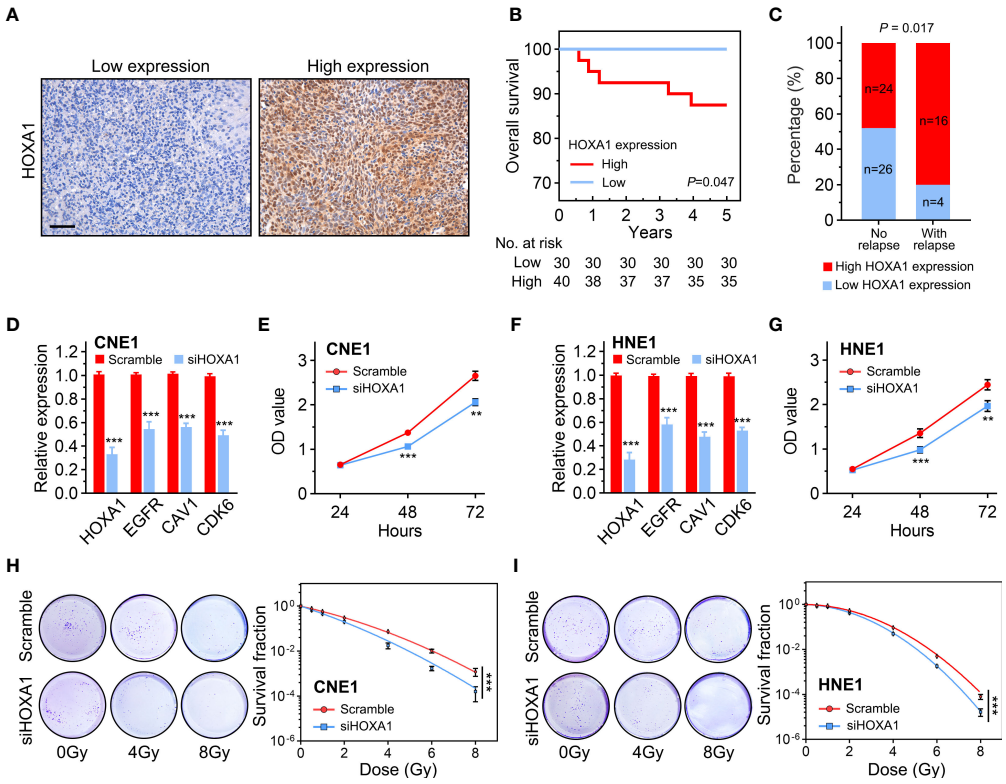



FIGURE 5
HOXA1 enhanced the radioresistance of NPC cells. **(A)** Representative immunohistochemistry images showing low and high HOXA1 expression in NPC tissues. Scale bars, 50 μm. **(B)** Overall survival curves of NPC patients. **(C)** Cumulative bar chart showing that high HOXA1 expression was correlated with high risk of local relapse in NPC patients. **(D)** knockdown of HOXA1 expression downregulated EGFR, CAV1 and CDK6 expression in CNE1 cells. **(E)** knockdown of HOXA1 expression reduced the proliferation of CNE1 cells. **(F)** knockdown of HOXA1 expression downregulated EGFR, CAV1 and CDK6 expression in HNE1 cells. **(G)** knockdown of HOXA1 expression reduced the proliferation of HNE1 cells. **(H, I)** Dose-survival curves showing that knockdown of HOXA1 expression enhanced the radiosensitivity of CNE1 and HNE1 NPC cells. $^{**}P < 0.01$; $^{***}P < 0.001$.

NPC cells. We conducted immunofluorescence assays 0.5 and 24 h after irradiation to examine the phosphorylation of H2A.X at Ser139 (γH2AX, a biomarker of DSBs). The results showed that HOXA1 knockdown did not affect formation of γH2AX foci induced by irradiation. However, HOXA1 knockdown significantly delayed absorption of γH2AX foci (Figure 6). These results suggest that HOXA1 could affect radiosensitivity by regulating DNA repair.

## Discussion

In this study, we found that HOXA1 has potential to serve as a radioresistance-predictive biomarker in multiple cancer types. We found that high HOXA1 expression was associated with poor prognosis and high recurrence risk in HNSCC and LGG patients who had undergone postoperative radiotherapy, and in NPC patients who had undergone radical radiotherapy, whereas HOXA1 had no predictive value with respect to prognosis in patients who underwent surgery alone. Many studies have reported that elevated HOXA1 expression is correlated with poor prognosis in multiple cancer types (30, 31). However, little was known about whether high HOXA1 expression promotes the radioresistance of cancer cells. Our study provides novel insights that could inform further research on the role of HOXA1 in radioresistance.

HOXA1 is a well-reported oncogene that is upregulated in human malignancies including HNSCC, breast cancer, and non-small-cell lung cancer (17, 32, 33). Nevertheless, the mechanism underlying the upregulation of HOXA1 in cancer cells was not completely clear. There is no evidence that mutation leads to the abnormal HOXA1 expression observed in tumor tissues. However, epigenetic modifications of the HOXA1 promoter could be a mechanism that induces abnormal expression of HOXA1. In GBM, long noncoding RNA HOTAIRM1 was shown to mediate histone methylation modifications in the

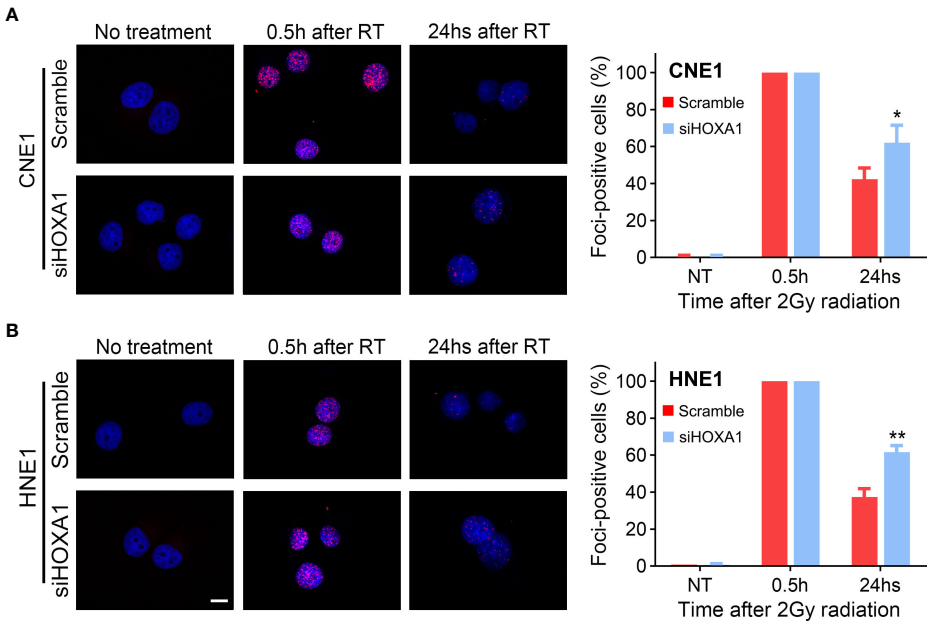



FIGURE 6
HOXA1 affected DNA repair capacity of NPC cells. enhanced the radioresistance of NPC cells. Immunofluorescence staining for γH2AX indicated that HOXA1 knockdown delayed absorption of γH2AX foci induced by 2Gy irradiation in CNE1 **(A)** and HNE1 cells **(B)**. **(B)** Cumulative bar chart showing that high HOXA1 expression was correlated with high risk of local relapse in NPC patients. Representative immunofluorescence images and quantification of the percentage of γH2AX foci-positive cells were shown. A positive cell was defined by the presence of more than 10 γH2AX foci. Scale bars, 10 μm. *$P < 0.05$, **$P < 0.001$.

promoter region that increased HOXA1 expression (30). In HNSCC, studies have reported that HOXA1 expression levels were negatively correlated with the DNA methylation level of the HOXA1 promoter (17).

We also performed *in vitro* experiments to confirm the radiosensitivity regulation function of HOXA1 in NPC cells. Our results confirmed that knockdown of HOXA1 expression significantly sensitized NPC cells to irradiation. Few studies have reported that HOXA1 regulates the radiosensitivity of cancer cells. However, it has been reported that high HOXA1 expression is correlated with cisplatin resistance of lung adenocarcinoma (34). In addition, many studies have reported that HOXA1 promoted proliferation and invasion of different cancer types (13, 30, 35). In our study, we also observed that HOXA1 knockdown reduced the proliferation capacity of NPC cells, consistent with previous reports.

Mechanistically, HOXA1 could regulate radiosensitivity via effects on the DNA repair capacity of cancer cells. Our results indicated that several terms associated with DNA repair were significantly enriched in HOXA1 high expression samples. *In vitro* experiments also demonstrated that knockdown of HOXA1 expression hindered the repair of DNA DSBs induced by 2 Gy irradiation in NPC cells. Previous studies reported that expression of some metabolism-related genes was correlated with radiosensitivity in HNSCC (36, 37). In this study, we also found that the oxidative phosphorylation gene set was significantly negatively enriched in HOXA1 high-expression samples in both HNSCC and LGG. It had been reported that enhancing oxidative phosphorylation by reversing the Warburg effect sensitized glioma cells to irradiation (29). Thus, HOXA1 may also induce radioresistance via effects on oxidative phosphorylation.

In addition, we found that HOXA1 regulated the expression of some radioresistance-associated genes including EGFR, CDK6, and CAV1. Our results suggest that HOXA1 expression is positively correlated with expression of EGFR, CDK6, and CAV1 in both HNSCC and LGG. Knockdown of HOXA1 also decreased the expression of these three genes in NPC cells. Moreover, we found some putative HOXA1 binding sites in the promoter regions of these three genes using the JASPAR database (Supplementary Table 2). Studies have reported that these three genes were associated with radioresistance. EGFR induced radioresistance of glioma cells via triggering the PI3K–Akt and MEK–ERK pathways and attenuated cetuximab-mediated radiosensitization of squamous cell carcinoma cells via the JIP-4/JNK2 signaling pathway (26, 38). CDK6 expression was correlated with radioresistance in HNSCC and NPC. Inhibition of CDK6 or targeting CDK6-associated signaling pathways enhanced the radiosensitivity of

HNSCC or NPC cells (27, 39). CAV1 increased oxidative stress protection and DNA repair, and its expression was correlated with radioresistance in rhabdomyosarcoma, pancreatic cancer, and lung cancer (28, 40, 41).

Taken as a whole, our data suggest that HOXA1 has potential as a predictive marker for intrinsic radioresistance in HNSCC, LGG, and NPC. Mechanistically, HOXA1 affected radiosensitivity via regulation of the DNA repair capacity of cancer cells. Knockdown of HOXA1 attenuated the DNA repair capacity and enhanced the radiosensitivity of NPC cells. Our results suggest that HOXA1 could be used to predict radioresistance and guide individualized treatment in multiple cancer types. However, more detailed basic and clinical studies are needed to confirm the regulatory function and clinical applications of HOXA1.

## Data availability statement

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

## Author contributions

YY and JL designed the study and experiments. LH wrote the manuscript. LH, ML and WG carried out the experiments. LH and ML performed the data analysis. All authors contributed to the article and approved the submitted version.

## Funding

This study was supported by Guangzhou Municipal Science and Technology Project (202201010965), National Natural Science Foundation of China (81602067), and Guangzhou Key Medical Discipline Construction Project.

## 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.2022.965427/full#supplementary-material

## References

1. Barton MB, Jacob S, Shafiq J, Wong K, Thompson SR, Hanna TP, et al. Estimating the demand for radiotherapy from the evidence: A review of changes from 2003 to 2012. *Radiother Oncol* (2014) 112(1):140–4. doi: 10.1016/j.radonc.2014.03.024

2. Wong KCW, Hui EP, Lo KW, Lam WKJ, Johnson D, Li L, et al. Nasopharyngeal carcinoma: An evolving paradigm. *Nat Rev Clin Oncol* (2021) 18(11):679–95. doi: 10.1038/s41571-021-00524-x

3. Ionna F, Bossi P, Guida A, Alberti A, Muto P, Salzano G, et al. Recurrent/Metastatic squamous cell carcinoma of the head and neck: A big and intriguing challenge which may be resolved by integrated treatments combining locoregional and systemic therapies. *Cancers* (2021) 13(10):2371. doi: 10.3390/cancers13102371

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

5. Chandra RA, Keane FK, Voncken FEM, Thomas CRJr. Contemporary radiotherapy: Present and future. *Lancet* (2021) 398(10295):171–84. doi: 10.1016/s0140-6736(21)00233-6

6. Li ZQ, Xia YF, Liu Q, Yi W, Liu XF, Han F, et al. Radiotherapy-related typing in 842 patients in canton with nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Phys* (2006) 66(4):1011–6. doi: 10.1016/j.ijrobp.2006.06.028

7. Lee AWM, Ng WT, Chan JYW, Corry J, Mäkitie A, Mendenhall WM, et al. Management of locally recurrent nasopharyngeal carcinoma. *Cancer Treat Rev* (2019) 79:101890. doi: 10.1016/j.ctrv.2019.101890

8. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. *Lancet* (2019) 394(10192):64–80. doi: 10.1016/s0140-6736(19)30956-0

9. Shah N, Sukumar S. The hox genes and their roles in oncogenesis. *Nat Rev Cancer* (2010) 10(5):361–71. doi: 10.1038/nrc2826

10. Cho HS, Toyokawa G, Daigo Y, Hayami S, Masuda K, Ikawa N, et al. The JmjC domain-containing histone demethylase KDM3A is a positive regulator of the G1/S transition in cancer cells *via* transcriptional regulation of the HOXA1 gene. *Int J Cancer* (2012) 131(3):E179–89. doi: 10.1002/ijc.26501

11. Mohankumar KM, Xu XQ, Zhu T, Kannan N, Miller LD, Liu ET, et al. HOXA1-stimulated oncogenicity is mediated by selective upregulation of components of the p44/42 MAP kinase pathway in human mammary carcinoma cells. *Oncogene* (2007) 26(27):3998–4008. doi: 10.1038/sj.onc.1210180

12. Belpaire M, Ewbank B, Taminiau A, Bridoux L, Deneyer N, Marchese D, et al. HOXA1 is an antagonist of ERα in breast cancer. *Front Oncol* (2021) 11:609521. doi: 10.3389/fonc.2021.609521

13. Yuan C, Zhu X, Han Y, Song C, Liu C, Lu S, et al. Elevated HOXA1 expression correlates with accelerated tumor cell proliferation and poor prognosis in gastric cancer partly *via* cyclin D1. *J Exp Clin Cancer Res* (2016) 35:15. doi: 10.1186/s13046-016-0294-2

14. Wardwell-Ozgo J, Dogruluk T, Gifford A, Zhang Y, Heffernan TP, van Doorn R, et al. HOXA1 drives melanoma tumor growth and metastasis and elicits an invasion gene expression signature that prognosticates clinical outcome. *Oncogene* (2014) 33(8):1017–26. doi: 10.1038/onc.2013.30

15. Zhang Y, Pan Q, Shao Z. Tumor-suppressive role of microRNA-202-3p in hepatocellular carcinoma through the KDM3A/HOXA1/MEIS3 pathway. *Front Cell Dev Biol* (2020) 8:556004. doi: 10.3389/fcell.2020.556004

16. Wang H, Liu G, Shen D, Ye H, Huang J, Jiao L, et al. HOXA1 enhances the cell proliferation, invasion and metastasis of prostate cancer cells. *Oncol Rep* (2015) 34(3):1203–10. doi: 10.3892/or.2015.4085

17. Li H, Wang X, Zhang M, Wang M, Zhang J, Ma S. Identification of HOXA1 as a novel biomarker in prognosis of head and neck squamous cell carcinoma. *Front Mol Biosci* (2020) 7:602068. doi: 10.3389/fmolb.2020.602068

18. Bitu CC, Destro MF, Carrera M, da Silva SD, Graner E, Kowalski LP, et al. HOXA1 is overexpressed in oral squamous cell carcinomas and its expression is correlated with poor prognosis. *BMC Cancer* (2012) 12:146. doi: 10.1186/1471-2407-12-146

19. Tu Z, Xu B, Qu C, Tao Y, Chen C, Hua W, et al. BRCC3 acts as a prognostic marker in nasopharyngeal carcinoma patients treated with radiotherapy and mediates radiation resistance *in vitro*. *Radiat Oncol* (2015) 10:123. doi: 10.1186/s13014-015-0427-3

20. He L, Zhou X, Qu C, Hu L, Tang Y, Zhang Q, et al. Musashi2 predicts poor prognosis and invasion in hepatocellular carcinoma by driving epithelial-mesenchymal transition. *J Cell Mol Med* (2014) 18(1):49–58. doi: 10.1111/jcmm.12158

21. Goldman M, Craft B, Swatloski T, Ellrott K, Cline M, Diekhans M, et al. The UCSC cancer genomics browser: Update 2013. *Nucleic Acids Res* (2013) 41 (Database issue):D949–54. doi: 10.1093/nar/gks1008

22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* (2014) 15(12):550. doi: 10.1186/s13059-014-0550-8

23. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an r package for comparing biological themes among gene clusters. *Omics* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118

24. Wang X, Yue J, Ren X, Wang Y, Tan M, Li B, et al. Modularity analysis based on predicted protein-protein interactions provides new insights into pathogenicity and cellular process of escherichia coli O157:H7. *Theor Biol Med Model* (2011) 8:47. doi: 10.1186/1742-4682-8-47

25. Chen W, Liu J, Ge F, Chen Z, Qu M, Nan K, et al. Long noncoding RNA HOTAIRM1 promotes immunosuppression in sepsis by inducing T cell exhaustion. *J Immunol* (2022) 208(3):618–32. doi: 10.4049/jimmunol.2100709

26. Fedrigo CA, Grivicich I, Schunemann DP, Chemale IM, dos Santos D, Jacovas T, et al. Radioresistance of human glioma spheroids and expression of HSP70, p53 and EGFr. *Radiat Oncol* (2011) 6:156. doi: 10.1186/1748-717x-6-156

27. Göttgens EL, Bussink J, Leszczynska KB, Peters H, Span PN, Hammond EM. Inhibition of CDK4/CDK6 enhances radiosensitivity of HPV negative head and neck squamous cell carcinomas. *Int J Radiat Oncol Biol Phys* (2019) 105(3):548–58. doi: 10.1016/j.ijrobp.2019.06.2531

28. Codenotti S, Marampon F, Triggiani L, Bonù ML, Magrini SM, Ceccaroli P, et al. Caveolin-1 promotes radioresistance in rhabdomyosarcoma through increased oxidative stress protection and DNA repair. *Cancer Lett* (2021) 505:1–12. doi: 10.1016/j.canlet.2021.02.005

29. Cook KM, Shen H, McKelvey KJ, Gee HE, Hau E. Targeting glucose metabolism of cancer cells with dichloroacetate to radiosensitize high-grade gliomas. *Int J Mol Sci* (2021) 22(14):7265. doi: 10.3390/ijms22147265

30. Li Q, Dong C, Cui J, Wang Y, Hong X. Over-expressed lncRNA HOTAIRM1 promotes tumor growth and invasion through up-regulating HOXA1 and sequestering G9a/EZH2/Dnmts away from the HOXA1 gene in glioblastoma multiforme. *J Exp Clin Cancer Res* (2018) 37(1):265. doi: 10.1186/s13046-018-0941-x

31. Zhao F, Tian H, Liu X, Guan Y, Zhu Y, Ren P, et al. Homeobox A1 facilitates immune escape and alleviates oxidative stress in lung adenocarcinoma. *Oxid Med Cell Longev* (2022) 2022:4102666. doi: 10.1155/2022/4102666

32. Abe M, Hamada J, Takahashi O, Takahashi Y, Tada M, Miyamoto M, et al. Disordered expression of HOX genes in human non-small cell lung cancer. *Oncol Rep* (2006) 15(4):797–802. doi: 10.3892/or.15.4.797

33. Belpaire M, Taminiau A, Geerts D, Rezsohazy R. HOXA1, a breast cancer oncogene. *Biochim Biophys Acta Rev Cancer* (2022) 1877(4):188747. doi: 10.1016/j.bbcan.2022.188747

34. Sui Q, Chen Z, Hu Z, Huang Y, Liang J, Bi G, et al. Cisplatin resistance-related multi-omics differences and the establishment of machine learning models. *J Trans Med* (2022) 20(1):171. doi: 10.1186/s12967-022-03372-0

35. He W, Huang Y, Jiang CC, Zhu Y, Wang L, Zhang W, et al. miR-100 inhibits cell growth and proliferation by targeting HOXA1 in nasopharyngeal carcinoma. *Onco Targets Ther* (2020) 13:593–602. doi: 10.2147/ott.S228783

36. Cruz-Gregorio A, Martínez-Ramírez I, Pedraza-Chaverri J, Lizano M. Reprogramming of energy metabolism in response to radiotherapy in head and neck squamous cell carcinoma. *Cancers* (2019) 11(2):182. doi: 10.3390/cancers11020182

37. Li Y, Weng Y, Pan Y, Huang Z, Chen X, Hong W, et al. A novel prognostic signature based on metabolism-related genes to predict survival and guide personalized treatment for head and neck squamous carcinoma. *Front Oncol* (2021) 11:685026. doi: 10.3389/fonc.2021.685026

38. Eke I, Schneider L, Förster C, Zips D, Kunz-Schughart LA, Cordes N. EGFR/JIP-4/JNK2 signaling attenuates cetuximab-mediated radiosensitization of squamous cell carcinoma cells. *Cancer Res* (2013) 73(1):297–306. doi: 10.1158/0008-5472.Can-12-2021

39. Huang D, Zhu X, Wang Y, Yu H, Pu Y. Long non-coding RNA FAM133B-2 represses the radio-resistance of nasopharyngeal cancer cells by targeting miR-34a-5p/CDK6 axis. *Aging (Albany NY)* (2020) 12(17):16936–50. doi: 10.18632/aging.103600

40. Leiser D, Samanta S, Eley J, Strauss J, Creed M, Kingsbury T, et al. Role of caveolin-1 as a biomarker for radiation resistance and tumor aggression in lung cancer. *PloS One* (2021) 16(11):e0258951. doi: 10.1371/journal.pone.0258951

41. Hehlgans S, Eke I, Storch K, Haase M, Baretton GB, Cordes N. Caveolin-1 mediated radioresistance of 3D grown pancreatic cancer cells. *Radiother Oncol* (2009) 92(3):362–70. doi: 10.1016/j.radonc.2009.07.004

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 12 September 2022
DOI 10.3389/fimmu.2022.934078

Check for updates

OPEN ACCESS

EDITED BY
Xinmin Li,
University of California, Los Angeles, United States

REVIEWED BY
Ganesan Ramamoorthi,
Moffitt Cancer Center, United States
Enrique J. Arenas,
Vall d'Hebron Institute of Oncology (VHIO), Spain

*CORRESPONDENCE
Wei Zhao
zhaowei23@mail.sysu.edu.cn
Changye Zou
zouchy@mail.sysu.edu.cn

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

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

RECEIVED 03 May 2022
ACCEPTED 19 August 2022
PUBLISHED 12 September 2022

CITATION
Yang M, Wang F, Lu G, Cheng M, Zhao W and Zou C (2022) Single-cell transcriptome analysis reveals T-cell exhaustion in denosumab-treated giant cell tumor of bone.
*Front. Immunol.* 13:934078.
doi: 10.3389/fimmu.2022.934078

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

# Single-cell transcriptome analysis reveals T-cell exhaustion in denosumab-treated giant cell tumor of bone

Meiling Yang[1†], Fen Wang[2†], Guohao Lu[3], Mingzhe Cheng[4], Wei Zhao[1,3*] and Changye Zou[4*]

[1]Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [2]Pathologica Department, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China, [3]Key Laboratory of Stem Cells and Tissue Engineering (Sun Yat-Sen University), Ministry of Education, Guangzhou, China, [4]Musculoskeletal Oncology Department, the First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China

Denosumab (DMAB), a human monoclonal antibody against the receptor activator of the nuclear factor-kappa B ligand, is used for the treatment for unresectable giant cell tumor of bone (GCTB). However, little is known about the molecular and functional characteristics of GCTB-infiltrating lymphocytes after DMAB treatment. Here, we performed single-cell RNA sequencing and immunostaining assays to delineate the immune landscape of GCTB in the presence and absence of DMAB. We found that exhausted $CD8^+$ T cells were preferentially enriched in DMAB-treated GCTB. A distinct M2-skewed type of tumor-associated macrophages (TAMs) comprises the majority of GCTB TAMs. We identified cytokines, including interleukin-10, and inhibitory receptors of M2 TAMs as important mediators of $CD8^+$ T cell exhaustion. We further revealed that DMAB treatment notably increased the expression levels of periostin (POSTN) in GCTB cells. Furthermore, POSTN expression was transcriptionally regulated by c-FOS signaling and correlated with GCTB recurrence in patients after DMAB treatment. Collectively, our findings reveal that $CD8^+$ T-cells undergo unappreciated exhaustion during DMAB therapy and that GCTB cell-derived POSTN educates TAMs and establishes a microenvironmental niche that facilitates GCTB recurrence.

KEYWORDS

**Denosumab, giant cell tumor of bone, T-cell exhaustion, periostin, single-cell RNA-seq, RANKL**

# Introduction

Giant cell tumor of bone (GCTB) is a locally aggressive bone tumor. It is mainly composed of two cell types: active osteoclast-like giant cells expressing the receptor activator of nuclear factor-kappa B (RANK), and neoplastic mononuclear stromal cells expressing the RANK ligand (RANKL) (1–3). The standard treatment for GCTB is the surgical resection of the tumor, including en bloc resection and extensive curettage. Curettage combined with local adjuvants should be the first choice for preserving a functional joint, although it shows a high recurrence rate of GCTB (4). Denosumab (DMAB), a human monoclonal antibody against RANKL, has been approved by the US Food and Drug Administration and the European Medicines Agency, and specifically inhibits RANKL-mediated formation and activation of osteoclast-like giant cells (4–6). Numerous clinical trials have shown that DMAB is correlated with a beneficial tumor response, but there are many controversies regarding its safety (5, 7–11). Moreover, the exact molecular basis and factors affecting the efficacy of DMAB remain poorly understood.

In addition to promoting osteoclast formation and activation, the RANKL–RANK signaling pathway plays important roles in lymph node development (12), lymphocyte differentiation, T-cell activation, dendritic cell survival, and immune tolerance induction (13, 14). RANKL-deficient mice do not exhibit lymph node metastasis (15). During initial T-cell receptor activation, T cells may provide RANKL directly to dendritic cells (DCs) to promote long-term interactions. In the absence of RANKL–RANK engagement, DCs may be at an increased risk of apoptosis, leading to reduced T cell activation (16–19). Thus, inhibition of RANKL could increase the immune escape as a result of T cell inactivation. Moreover, RANK expression in tumor-associated myeloid immune cells, such as DCs, TAMs, and myeloid-derived suppressor cells, is observed in diverse mouse tumor models and human tumors. The blocking of the RANKL–RANK interaction could either promote or suppress the antitumor immunity, depending on the specific phase and activated pathways (20). The potential effects of DMAB on the functions of myeloid immune cells present in the tumor microenvironment (TME) of GCTB remain unclear.

Single-cell RNA sequencing (scRNA-seq) has recently been used to characterize subsets of immune cells in tumors and their corresponding transcriptome changes upon treatment (21–24). Previous scRNA-seq data related to GCTB revealed the heterogeneity of osteoclasts and immune cells in GCTB (25). In the present study, we used scRNA-seq and immunostaining assays to detect the dynamic changes in the immune cells in GCTB and to unveil $CD8^+$ T-cell exhaustion associated with DMAB therapy.

# Materials and methods

## Human GCTB samples

This study was approved by the Medical Ethics Committee of the First Affiliated Hospital of Sun Yat-Sen University. All patients provided informed consent preoperatively. For scRNA-seq, two patients with or without DMAB treatment were pathologically diagnosed with GCTB at the First Affiliated Hospital of Sun Yat-Sen University. Formalin-fixed paraffin embedded (FFPE) archival tissue blocks used for Immunohistochemistry (IHC) and immunofluorescence (IF) staining assays were collected from GCTB tumors with or without DMAB treatment.

## Single-cell isolation and scRNA-seq

Fresh tumor tissues were surgically removed from GCTB patients and minced into 2-4 mm pieces. The pieces were transferred to the tube with digestive enzyme from Tumor Dissociation Kit (Cat# 130-095-929, Miltenyi Biotec) and incubated at 37 °C for 30 min on a shaker. After digestion, 2% FBS was added to neutralize enzyme lysate, and the tissues were filtered through a 70 μm filter. Subsequently, the samples were centrifuged at 350 × g for 5 min and the supernatants were discarded. To remove red blood cells, the cell pellets were suspended in red blood cell lysis buffer (Beyotime) for 30 s. The solution was then centrifuged at 350 × g for 5 min and resuspended in Dulbecco's phosphate-buffered saline (DPBS; Thermo Fisher). The samples were stained with trypan blue (Solarbio) and the cellular viability was evaluated. Finally, single cells were encapsulated into emulsion droplets using the Chromium Controller (10× Genomics). The scRNA-seq libraries were constructed following the manufacturer's instructions (10× Genomics) and then used for sequencing.

## Pre-processing of scRNA-seq data

Raw data were processed using Cell Ranger (v3.0.2) to align reads, generate feature-barcode matrices, and perform gene expression analysis. We used the mkfastq pipelines to make fastq files and used the cell count pipelines for alignment (with reference genome Hg19), filtering, barcode counting, and UMI counting.

The raw output data were processed with the Seurat package (version 4.0.4; http://satijalab.org/seurat/) in R software. We filtered out the cells with less than 200 genes and the percent of mitochondrial genes over 25% of total expressed genes. Genes detected in fewer than 10 cells were also excluded. Feature counts

for each cell were divided by the total counts for that cell and multiplied by the scale factor (10,000), and then natural-log transformed (default Seurat approach). To adjust for technical variation and batch effects between samples, we used the standard anchor-based workflow for dataset integration in Seurat (26). The merged dataset included 13,857 cells and 20,438 detected genes across the two samples. Integration-transformed expression values were used only for dimension reduction and clustering. The original lognormalized expression values were used for all differential expression and gene set level analyses.

## Unsupervised clustering and identification of cell types

We used the Seurat package to perform unsupervised clustering. We performed principal component analysis (PCA) on the integration-transformed expression matrix using highly variable genes identified by "FindVariableFeatures" function. Following the results of PCA, the appropriate principal components (PCs) were selected for clustering with the specific resolution parameters. The identified clusters were visualized on the 2D map produced with the t-SNE method. The cell groups were annotated based on the well-known cellular markers from the literature (27, 28).

For the clustering of all cells, the top 15 PCs were selected with a resolution parameter equal to 0.8. For the clustering of myeloid cells, the top 15 PCs were selected with a resolution parameter equal to 0.6. For the clustering of TILs, we used the top 15 PCs with a resolution parameter equal to 2, but a group of cells expressed both macrophage and T cell marker genes. We eliminated this group of cells because it was most likely doublets and then reclustering TILs with the top 15 PCs and a resolution parameter equal to 1.2.

## Identification of tumor cells

Tumor cells were identified using cluster-level marker genes expression (BGLAP, RUNX2, TNFSF11, IBSP) and inferred CNV profiles. We performed InferCNV (inferCNV of the Trinity CTAT Project. https://github.com/broadinstitute/inferCNV) within each sample, using TILs as reference group and tumor cells as observation group. We identified large-scale chromosomal copy number variants, either gains or losses, in tumor cells, in addition to the expression of RUNX2. The raw single-cell gene expression data was extracted from the Seurat object according to the software recommendation. The inferCNV analysis was performed with default parameters including a value of 0.1 for "cutoff".

## IHC and IF staining assays

IHC and IF staining of FFPE GCTB specimens were performed according to the previous published paper (29). Briefly, all sections were deparaffinized, rehydrated and antigen repaired. For IHC staining of RANKL, RUNX2 and POSTN, endogenous peroxidase was blocked using 3% hydrogen peroxide for 30 min. Samples were blocked with goat serum at room temperature (RT) for 1 h and incubated separately with RANKL antibody (Cell Signaling Technology, 1:200), RUNX2 antibody (ab76956, Abcam, 1:200), RANK antibody (Cell Signaling Technology) and POSTN antibody (ab14041, Abcam, 1:800) overnight at 4°C, followed by horseradish peroxidase (HRP)-linked secondary antibodies and DAB staining (#8059, Cell Signaling Technology). Counterstaining was done with hematoxylin.

The rate of RANKL-, RANK-, or RUNX2-positive cells was counted, based on all cells of the whole tissue slide. The scoring system of POSTN staining was blind observed at 5 fields of the sections under a 20-fold microscope, and scored the immunohistochemical staining intensity and positive areas. Staining intensity score: 0 points for no staining, 1 point for light yellow, 2 points for tawny and 3 points for brown; Positive area percentage score: 0 points for no positive area, 1 point for < 30%, 2 points for 30%~60%, and 3 points for more than 60%. The product of the score values was used as the score of immunostaining score, with a total of 7 score level (0, 1, 2, 3, 4, 6, and 9). The patients were divided into high expression group and low expression group according to the median patient score.

For IF staining of CD8 and LAG3, samples were blocked with 5% donkey serum at RT for 1 h and incubated with CD8 antibody (ab199016, Abcam, 1:100) and LAG3 antibody (ab180187, Abcam, 1:5000) overnight at 4°C. After washing, samples were incubated for 1 h at RT with fluorescently labeled secondary antibodies including donkey anti-mouse Alexa Fluor 488 (ab150105, Abcam, 1:500) and donkey anti-rabbit Alexa Fluor 555 (ab150074, Abcam, 1:500). Nuclei were counterstained with Hoechst33342 (H3570, Thermo Fisher Scientific, 1:500). To remove unwanted fluorescence in tissue sections due to aldehyde fixation, red-blood cells, and structural elements, we used Vector TrueVIEW Autofluorescence Quenching Kit (SP-8400, VECTOR). To ensure representativeness, the whole tissue slide was observed, and 5 fields with high infiltration of $CD8^+$ T cells were selected for photographing with laser confocal scanning microscope (LSM780) and the rate of $CD8^+LAG3^+$ T cells in $CD8^+$ T cells was calculated.

## Cell-cell interaction analysis

To analyze cell-cell interactions between different cell types, we used CellChat (version 1.1.2; https://github.com/sqjin/CellChat/) (30) to identify significant ligand-receptor pairs within no DMAB and DMAB samples. For both no DMAB and DMAB samples, the cell type specific ligand-receptor interactions were identified based on the specific expression of a receptor by one cell type and a ligand by another cell type. We used the "netVisual_bubble" function to show the cell communication mediated by exhaustion-related ligand-receptor pairs, and the "netVisual_chord_gene" function to visualize cell-cell communication for the enhanced signaling pathways to macrophages in GCTB after DMAB treatment.

## Definition of cell scores and signature

To evaluate the potential functions of a cell group of interest from No DMAB and DMAB samples, we calculated the scores of functional feature sets for the cell group, using the "AddModuleScore" function in Seurat at single cell level. The average expression levels of the functional feature sets were subtracted by the aggregated expression of control feature sets. All analysed genes were binned based on averaged expression, and the control features were randomly selected from each bin.

The functional gene sets including M1/M2 polarization and anti-inflammatory for macrophages, co-stimulatory and exhausted for $CD8^+$ T cells. The involved gene sets were listed in the supplementary material (Supplementary Table S1) (27).

## Trajectory analysis of single cells

The single-cell pseudotime trajectories were generated with the Monocle package (v2.20.0) in R (31–33). The gene-cell matrix in the raw counts derived from the Seurat RNA assays were used as the inputs. The "newCellDataSet" function was applied to create an object for subsequent analysis. We Filtered out low-quality genes by using "detectGenes" function with the parameters "min_expr = 0.1". We reduced data dimensionality by using the "reduceDimension" function with the parameter reduction_method = "DDRTree". The cells were ordered and visualized with the "plot_cell_trajectory" function. Genes of interest that changed along with the pseudotime were visualized with the "plot_pseudotime_heatmap" and the genes were clustered into subgroups according to the gene expression patterns. To identify the genes that separate cells into branches, the branch expression analysis modeling (BEAM) analysis were performed. Genes of interest, resulting from the BEAM analysis, were visualized with the "plot_genes_branched_heatmap" function and "plot_genes_branched_pseudotime" function. In the single-cell pseudotime trajectories analysis of $CD8^+$T cells, we excluded the proliferative subgroup CD8 Ki67, because its high proliferative status would affect the trajectory analysis.

## DEGs identification

DEGs of tumor cells between No DMAB and DMAB samples were identified using "FindMarkers" functions in Seurat with the Wilcoxon Rank-Sum test (logfc.threshold = 0.25).

## ChIP-qPCR assay

GCTB tissues from three patients were cross-linked with 1% (v/v) formaldehyde at room temperature for 10 min and then processed using the Magna ChIP G-Chromatin Immunoprecipitation Kit (Sigma-Aldrich, 17-611) according to manufacturer's instructions. The subsequent qPCR was conducted using the 2 × ChamQ Universal SYBR qPCR Master Mix (Vazyme, China). The primer sequences for POSTN were as follows: forward, TGAGACTTAAACATGCAGTGAGT; Reverse, ACATTGAGCTACTTTTCCTTTTCAT.

## Statistics

Comparisons of gene expression or functional signature between two groups of cells were performed using unpaired two-tailed student's t test. Statistical analyses and presentation were performed using R. Comparisons of the fractions of $RANKL^+$ or $RUNX2^+$ cells in paired No DMAB and DMAB were performed using wilcoxon matched-pairs signed rank tests. Comparisons of the infiltration ratio of $CD8^+LAG3^+$ T cells in unpaired No DMAB and DMAB were performed using unpaired two-tailed t test. Statistical analyses and presentation were performed using Graphpad prism8. Other statistical tests used in figures were shown in figure legends.

# Results

## scRNA-seq profiling of the TME in GCTB with or without DMAB treatment

To reveal the TME in DMAB-treated GCTB, we collected surgical tumor specimens from untreated (No DMAB) and DMAB-treated (DMAB) GCTB patients for scRNA-seq. After the initial quality control assessment, single-cell transcriptomes were obtained from 13,857 cells. In addition, there were 8 paired GCTB samples in immunohistochemistry (IHC) assays for validation of RANKL and RUNX family transcription factor 2 (RUNX2) expression, 6 unpaired GCTB samples (3 treated vs. 3

untreated) in IF assays for validation of CD8 and lymphocyte activating 3 (LAG3) expression, and 32 GCTB samples (9 no DMAB vs. 23 DMAB; 14 no recurrence vs. 9 recurrence within 23 DMAB) in IHC assays for validation of POSTN expression (Figure 1A).

Seurat was used for cell classification and marker gene identification. Seven main clusters were identified and visualized using the t-distributed stochastic neighbor embedding (t-SNE) method (Figure 1B). They were as follows (1): myeloid cells highly expressing CD74, CD14, and lysozyme (2); TILs, including T and NK cells, specifically expressing CD3D, CD3E, and NKG7 (3); osteoclast-like giant cells with high expression of ACP5, CTSK, and MMP9 (4); tumor cells highly expressing BGLAP, RUNX2, RANKL, and IBSP (5); fibroblasts specifically expressing FN1, LUM, and DCN (6); endotheliocytes specifically expressing VWF, RAMP2, and CDH5; and (7) pericytes highly expressing RGS5, ACTA2, and MCAM (Figure 1C; Supplementary Figure S1A). To ensure the correct definition of tumor cells, we applied the inferCNV algorithm to analyze the copy number variations (CNVs) of

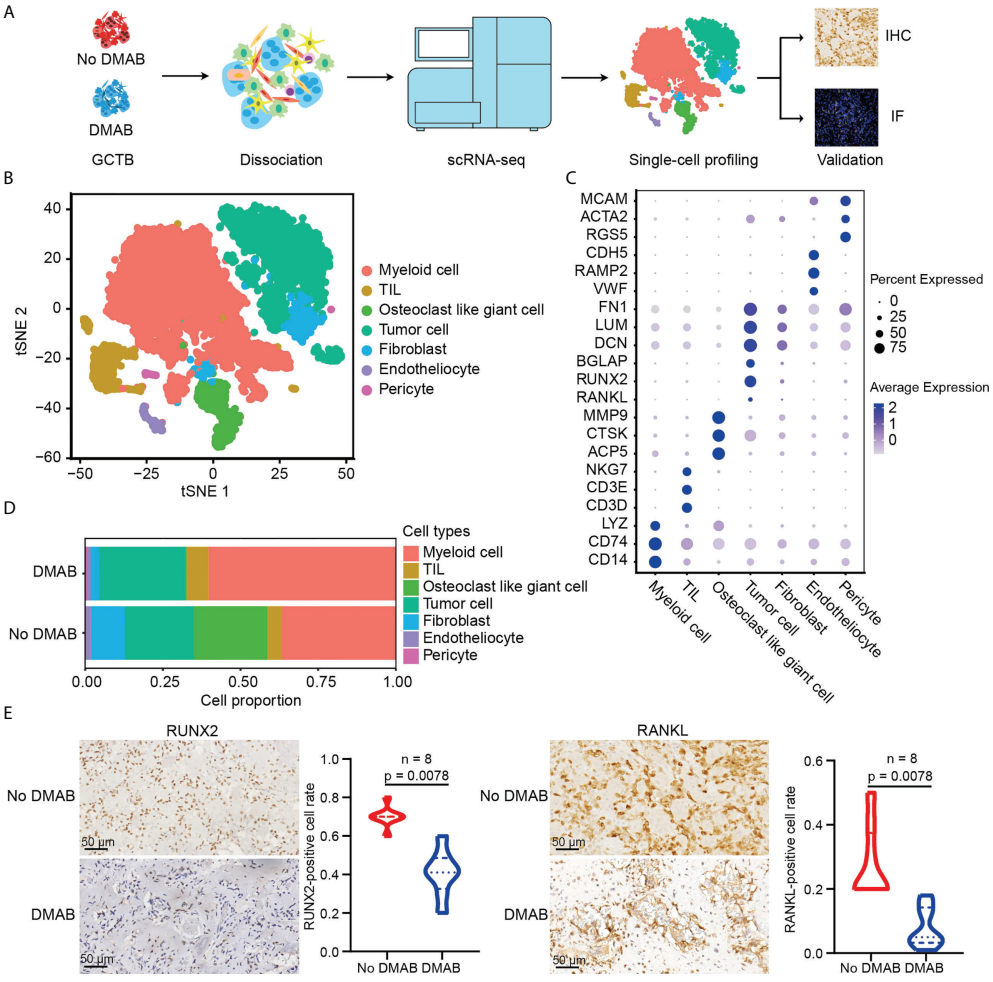



FIGURE 1
Single-cell RNA sequencing (scRNA-seq) profiling of the untreated and DMAB-treated giant cell tumor of bone (GCTB) tumor microenvironments (TMEs). **(A)** Schematic representation of the experimental design. Single-cell suspensions were collected from GCTB tumors of two patients followed by scRNA-seq on 10× Genomics platform. A total of 13,857 qualified single cells were recovered. **(B)** T-distributed stochastic neighbor embedding (t-SNE) plot showing the annotation and color codes for cell types in the GCTB ecosystem. **(C)** Dot plot showing the expression levels of marker genes in the indicated cell types. The size of the dot represents the proportion of cells expressing the particular marker within the group and the spectrum of color indicates the mean expression levels of the markers. Blue color dots indicate the highly expressed genes, while gray color dots indicate the low expressed genes. **(D)** Histogram showing the proportion of all cell types in untreated and denosumab (DMAB)-treated samples. **(E)** Representative images of immunohistochemistry (IHC) staining in formalin-fixed paraffin-embedded (FFPE) tissues, indicating RUNX family transcription factor 2 (RUNX2) $^{+}$ cells and RANKL$^{+}$ cells in paired No DMAB and DMAB samples (n = 8). Scale bar, 50 μm. Violin plot presenting the fractions of RUNX2$^{+}$ cells and RANKL$^{+}$ cells in paired No DMAB and DMAB samples based on IHC staining results. Statistical analyses are paired Wilcoxon tests.

tumor cells using TILs as control cells (data not shown), and confirmed that the defined tumor cells had obvious CNVs (Supplementary Figure S1B).

We noticed that almost all types of cell populations were present in both DMAB-treated and untreated samples, except for osteoclast-like giant cells, which were significantly missing in the DMAB sample (Figure 1D). In addition, we found that the expression of osteoblast-related genes RUNX2 and RANKL was reduced in tumor cells after DMAB treatment (Figure 1E; Supplementary Figure S1C). The expression of RANK, receptor of RANKL, was not significantly changed after DMAB treatment (Supplementary Figures S1D, E).

## DMAB treatment is associated with the exhaustion of CD8$^+$ LAG3$^+$ T cells in GCTB

The re-clustering of TILs revealed 11 populations, including five subtypes of CD8$^+$ T cells (CD8 KLRC1, CD8 GZMB, CD8 GZMH, CD8 Ki67, and CD8 LAG3), four clusters of CD4$^+$ T cells (CD4 CCR7, CD4 IL7R, CD4 CD40LG, and Treg FOXP3), one NK subtype, and one unknown cluster with high COL1A2 expression (CD45 COL1A2) (Figure 2A). CD8$^+$ T cells were identified by the high expression levels of CD3D and CD8A. They also highly expressed the genes associated with cytotoxicity (GZMA, GZMK, and NKG7). CD8 Ki67 cells displayed high expression of proliferative genes (MKI67, PCNA, and TOP2A) and moderate expression of exhaustion-related markers (LAG3, TIGHT, and PDCD1), suggesting that these cells represent an early exhausted state. CD8 LAG3 cells showed the highest expression levels of T cell exhaustion markers (LAG3, TIGHT, PDCD1, HAVCR2, and CTLA4), suggesting that these cells were exhausted (Figures 2B; Supplementary Figure S2A). In addition, we found that cytotoxic CD4$^+$ T cells had increased infiltration and enhanced toxicity following DMAB treatment (Supplementary Figures S2B, C).

To further investigate the effect of DMAB on the functional characteristics of CD8$^+$ T cells, we calculated the costimulatory and exhaustion scores of CD8$^+$ T cells by analyzing the expression of related genes. Interestingly, we found that costimulatory and exhausted scores and related genes increased significantly in DMAB-treated samples compared to untreated GCTB (Figure 2C; Supplementary Figures S2D, E). Additionally, the number of exhausted CD8$^+$ cells that expressed high exhaustion scores (Figure 2D) was higher in DMAB-treated samples than in untreated samples (Figure 2E). LAG3 is a recently recognized immune checkpoint, and its high expression correlates with T-cell exhaustion. We further verified the increased abundance of exhausted CD8$^+$ T cells (CD8$^+$LAG3$^+$) in DMAB versus No DMAB using IF staining (Figure 2F). These results suggest that DMAB treatment correlates with CD8$^+$ T cell exhaustion in GCTB.

## Dynamic cell transitions of CD8$^+$ T cells in GCTB with or without DMAB treatment

Next, we explored the dynamic cell transitions and immune states in CD8$^+$ T cells by inferring state trajectories using Monocle. This analysis showed that CD8 KLRC1 cells were at the beginning of the trajectory, whereas CD8 LAG3 cells were in a terminal state (Figure 3A, upper part). We identified three sets of differentially expressed genes along the CD8$^+$ T cell trajectory. The first set, consisting of naive T cell markers (CCR6 and LEF1), decreased along the trajectory, while the second set, consisting of effector genes (IFNG) and cytotoxic genes (GZMB, GZMK, and GZMA), increased from the middle to the end of the trajectory. The last set, consisting of exhausted markers (LAG3, CTLA4, TIGHT, and PDCD1), increased towards the end of the trajectory (Supplementary Figures S3A, B). CD8$^+$ T cells in the No DMAB sample only occupied branch one (cell fate 1), whereas CD8$^+$ T cells in the DMAB sample covered the entire trajectory, including the two branches (cell fate 1 and cell fate 2) (Figure 3A, lower part). Based on BEAM, we found that tissue-resident memory T cell-related genes (CD69, ITGA1, and FOSB) were upregulated along cell fate 1, suggesting that this branch shifted towards memory T cells. Activation genes (HLA-DMA, HLA-DPA1, HLA-DQA1, and HLA-DRA) and exhausted genes (CTLA4, TOX, and ENTPD1) were both upregulated along cell fate 2 (Figure 3B; Supplementary Figure S3C), suggesting that CD8$^+$ T cells in this branch were first activated and then subjected to exhaustion.

To determine the mechanism that induces CD8$^+$ T cell exhaustion, we used CellChat for cell communication analysis. The analysis showed that DMAB treatment enhanced multiple exhaustion-related pathways, including PVR–TIGIT, PDCD1LG2-PDCD1, LGALS9-HAVCR2, IL10-IL10R and CD86-CTLA4, in various cell populations (Figure 3C). Intriguingly, TAMs targeted the exhausted subpopulation with the highest communication probability in the DMAB sample, suggesting that TAMs may contribute to the exhaustion of CD8$^+$ T cells after DMAB treatment.

## Periostin enhances the M2-like phenotype of TAMs in DMAB treated GCTB

Next, we performed unsupervised clustering of myeloid cells in GCTB. Ten clusters emerged within the myeloid lineage, including five clusters for TAMs (Macro1–Macro5), three for DCs (DC CD1C, DC LAMP3, and DC Cycling), one for monocytes, and one for neutrophils (Figure 4A). TAMs were identified by the high expression levels of CD68 and CD163. We could not clearly distinguish M1 and M2 TAMs using known

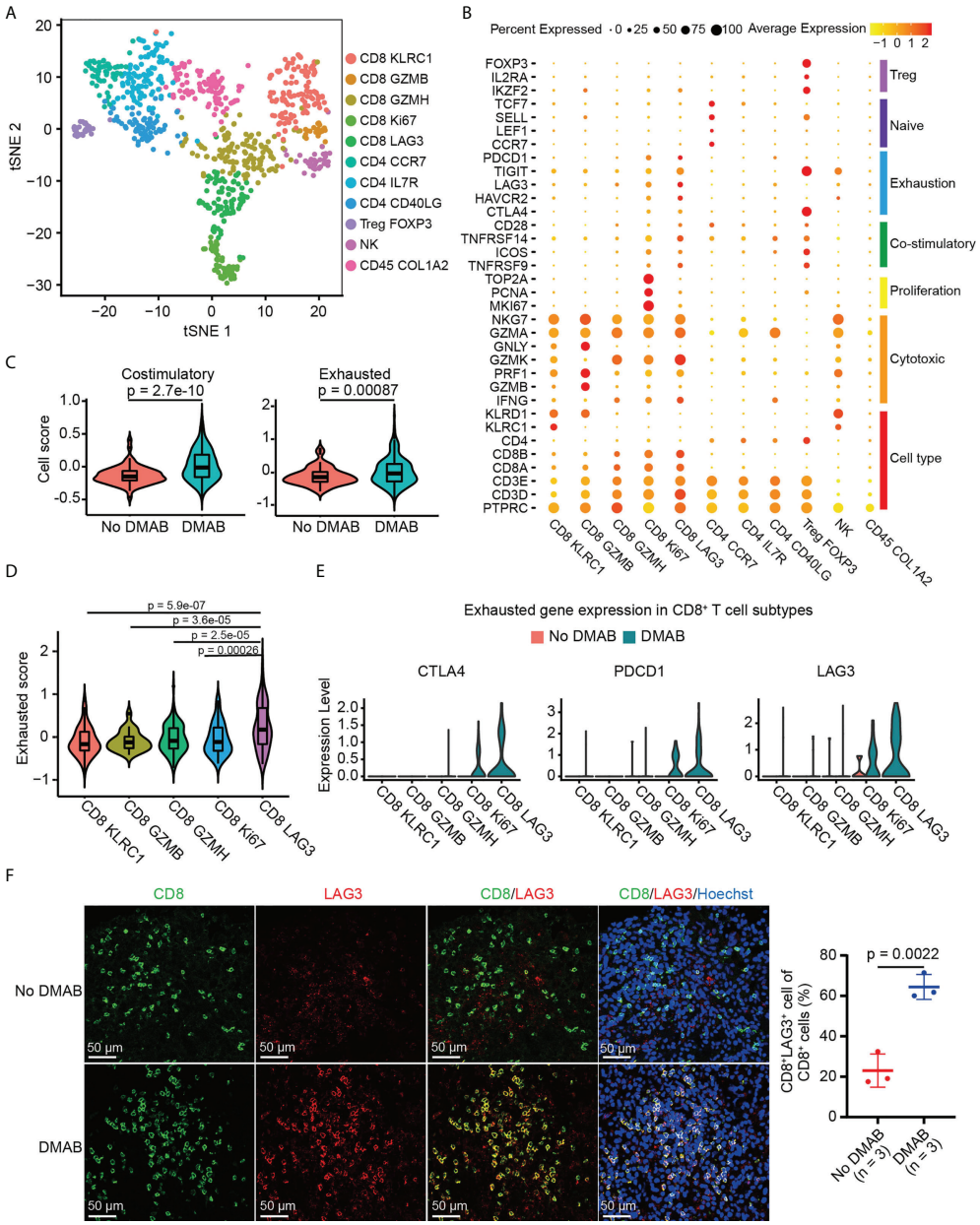



FIGURE 2
DMAB treatment promotes the exhaustion of CD8$^+$ lymphocyte activating 3 (LAG3) $^+$ T cells in GCTB. **(A)** t-SNE plot showing the subtypes of tumor-infiltrating lymphocytes (TILs) derived from untreated and DMAB-treated patients with GCTB. **(B)** Dot plot showing the expression levels of selected gene sets in each subtype of TILs, including Treg, naive, exhaustion, costimulatory, proliferation, and cytotoxic cell types. **(C)** Violin plot showing the costimulatory and exhausted scores of CD8$^+$ T cells from the No DMAB (red) and DMAB (blue) samples. The p values are calculated by Student's *t* test. **(D)** Violin plot indicating the exhausted scores in CD8$^+$ T cell subtypes. The p values are calculated by Student's *t* test. **(E)** Violin plot showing the expression levels of selected exhausted genes in CD8$^+$ T cell subtypes. Red, No DMAB; blue, DMAB. **(F)** Immunofluorescence **(IF)** staining of CD8 and LAG3 antibodies, showing the infiltration of CD8$^+$LAG3$^+$ T cells in unpaired patients with GCTB with (n = 3) or without (n = 3) DMAB. Scale bars = 50 μm. Based on the IF staining results, the infiltration ratio of CD8$^+$LAG3$^+$ T cells was statistically analyzed using an unpaired *t* test.

marker genes, CD86 (M1) and CD163 (M2), as they were both expressed in these cells (Supplementary Figure S4A). However, by calculating M1 and M2 scores for each cell using related gene sets (Supplementary Table S1), we observed that clusters Macro1-Macro4 exhibited an M2-like phenotype, whereas cluster Macro5 exhibited an M1-like phenotype (Supplementary Figure S4B). DCs were identified by the expression of CD1C and CD1A. DC cycling showed high

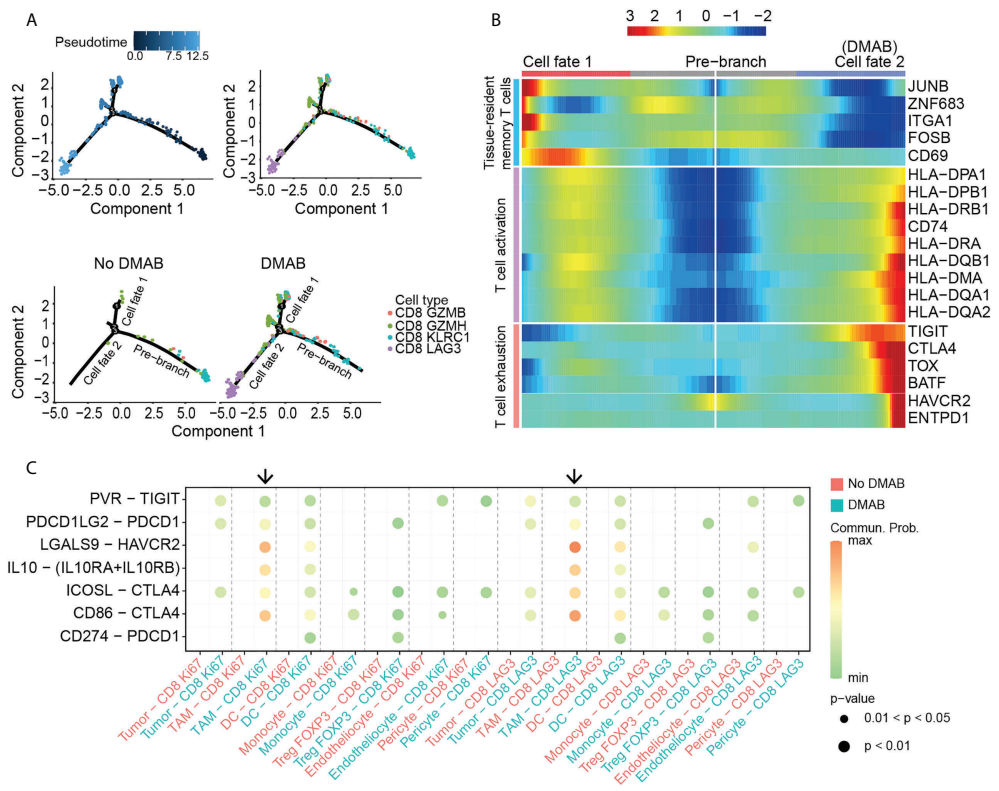



FIGURE 3
TAM induces the exhaustion of CD8$^+$ T cells after DMAB treatment. **(A)** Pseudotime-ordered analysis of CD8$^+$ T cells from No DMAB and DMAB samples. CD8$^+$ T cell subtypes are labeled with colors. Each dot indicates a single cell. **(B)** Pseudotime heatmap showing the dynamic changes in gene expression along cell fate 1 and cell fate 2 according to the BEAM analysis. The selected genes are associated with tissue-resident memory T cells, T cell activation, and T cell exhaustion. **(C)** Bubble chart showing exhaustion-related signaling targeting exhausted subtypes (CD8 Ki67 and CD8 LAG3) based on selected ligand and receptor pairs, as calculated by CellChat. Dot color reflects communication probabilities and dot size represents computed p-values. Empty space means that the communication probability is zero. Black arrows indicate the communication between TAMs and exhausted subtypes in the DMAB sample (blue). The p values were computed using a one-sided permutation test.

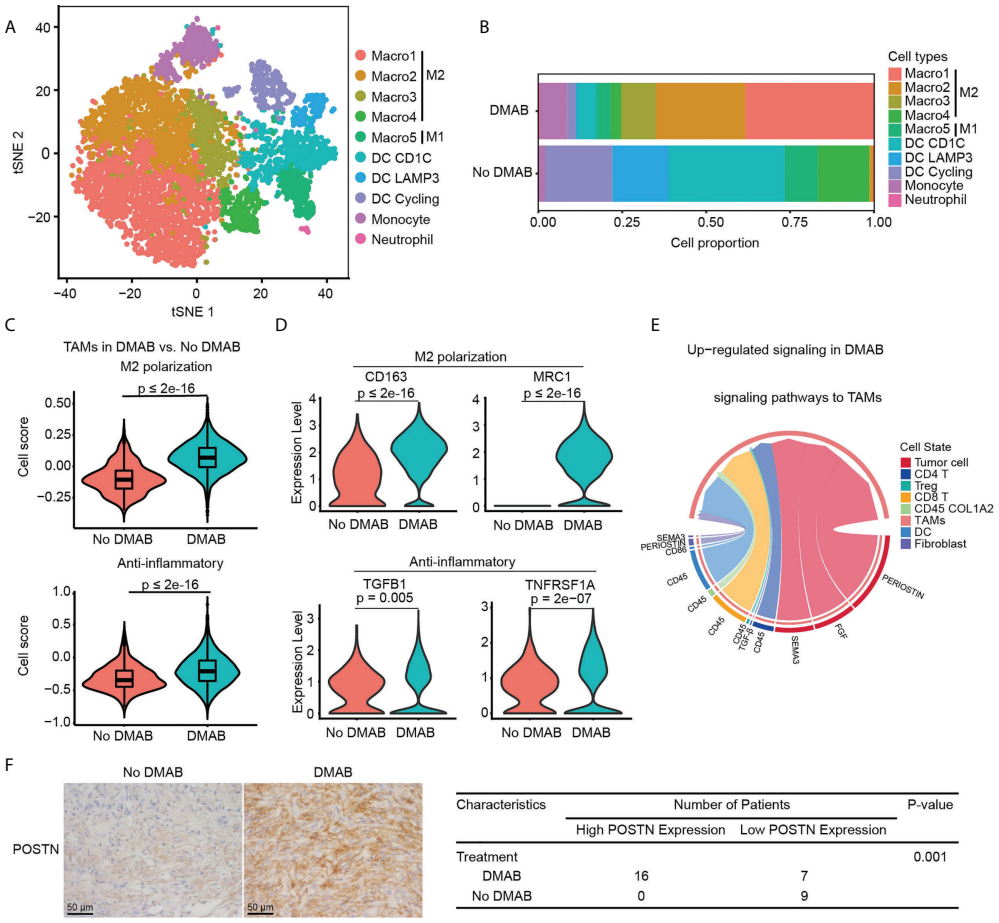
expression of proliferation-related genes (MKI67 and TOP2A) (Supplementary Figure S4A).

We found that the TAM subpopulations showed treatment bias. The proportion of TAMs was significantly higher in DMAB than in No DMAB samples, especially in M2-like TAMs (Figure 4B). To further investigate the effect of DMAB on the functional characteristics of TAMs, we calculated the M2 polarization and anti-inflammatory scores of TAMs by analyzing the expression of related genes. We found that the M2 polarization and anti-inflammatory scores increased significantly after DMAB treatment (Figures 4C). Consistently, the expression of M2 related genes (e.g., CD163 and MRC1) and anti-inflammatory related genes (e.g., TGFB and TNFRSF1A) was higher in DMAB-treated GCTB than in untreated GCTB (Figure 4D).

To determine the underlying mechanism by which GCTB facilitates the M2-like phenotype in TAMs, we used CellChat to analyze cell communication between TAMs and GCTB cells. We found that the POSTN pathway (POSTN-ITGAV/ITGB5) was enhanced from GCTB tumor cells to TAMs after DMAB treatment (Figure 4E; Supplementary Figure S4C). In addition, POSTN expression levels were upregulated in DMAB-treated GCTB compared to untreated control (Figure 4F). These analyses indicated that POSTN secreted by GCTB cells likely promoted the M2-like phenotype in TAMs.

## POSTN expression is associated with the relapse of GCTB after DMAB treatment

We further investigated the gene expression patterns of GCTB cells in the DMAB-treated and untreated samples (Figure 5A). The analysis revealed that activator protein 1 transcription factors (c-Fos, c-Jun, and FOSB) were significantly upregulated in DMAB-treated GCTB cells (Figure 5B). Furthermore, chromatin immunoprecipitation




| Characteristics | Number of Patients | | P-value |
|---|---|---|---|
| | High POSTN Expression | Low POSTN Expression | |
| Treatment | | | 0.001 |
| DMAB | 16 | 7 | |
| No DMAB | 0 | 9 | |

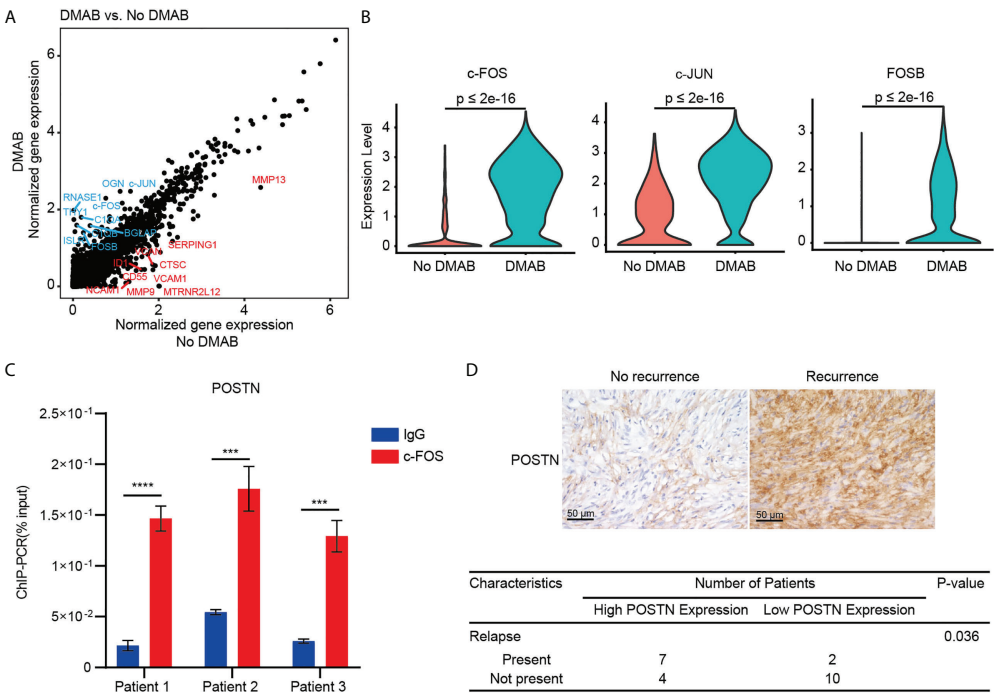
FIGURE 4
DMAB treatment promotes the M2 subtype of TAMs *via* periostin (POSTN). **(A)** t-SNE plot showing the subtypes of myeloid-derived cells derived from No DMAB and DMAB-treated GCTB patients. **(B)** Histogram showing the proportion of myeloid subgroups in No DMAB and DMAB-treated samples. **(C)** Violin plot showing the M2 polarization and anti-inflammatory scores of TAMs from the No DMAB (red) and DMAB-treated (blue) samples. The p values were calculated using a Student's *t* test. **(D)** Violin plot showing the expression levels of selected M2 polarization and anti-inflammatory genes in TAMs from the No DMAB (red) and DMAB-treated (blue) samples. The p values were calculated using a Student's *t* test. **(E)** Chord diagram showing the significantly upregulated signaling pathways in TAMs in GCTB after DMAB treatment. **(F)** Representative IHC images of POSTN in GCTB without DMAB (n =9) versus GCTB with DMAB (n = 23). Scale bar, 50 μm. Statistical table of 32 GCTB samples by POSTN staining intensity. Significance was determined by a fisher exact test.

(ChIP) assay results demonstrated that c-FOS binds to the POSTN promoter in GCTB (Figure 5C).

POSTN is a secreted extracellular matrix protein, which is usually associated with poor prognosis in cancers. Indeed, we found that high POSTN expression was associated with poor prognosis in kidney papillary cell carcinoma, liver cancer, and lung cancer (Supplementary Figures S5A–C). Furthermore, POSTN is highly expressed in bone metastasis of prostate cancer and breast cancer (Figure S5D, S5E, and S5F). To determine the correlation between POSTN expression and relapse of GCTB in patients receiving DMAB, we stained DMAB-treated GCTB clinical specimens with POSTN with or without recurrence. POSTN expression levels were upregulated in DMAB-treated patients with relapse. Tumor relapse was positively correlated with POSTN staining (Figure 5D). These results demonstrate that POSTN expression levels have important prognostic significance for GCTB patients treated with DMAB.

## Discussion

Little is known about the mechanisms mediating the adverse effects of DMAB therapy in patients with GCTB. Here, we depicted the cellular landscape and transcriptional profiles of GCTB with or without DMAB therapy, encompassing immune cells and tumor cells. We revealed an unappreciated T-cell dysfunction in DMAB-treated GCTB, whereby T cells undergo a transition from activated




| Characteristics | Number of Patients | | P-value |
|---|---|---|---|
| | High POSTN Expression | Low POSTN Expression | |
| Relapse | | | 0.036 |
| Present | 7 | 2 | |
| Not present | 4 | 10 | |

FIGURE 5
C-FOS-mediated POSTN expression is associated with the relapse of GCTB after DMAB treatment. **(A)** Scatter plot showing the DEGs of tumor cells from DMAB-treated sample versus No DMAB sample. The top 10 DEGs were labeled in blue (upregulated in DMAB-treated samples) or in red (upregulated in No DMAB samples). **(B)** Violin plot showing the expression levels of activator protein 1 (AP-1) transcription factor (c-FOS, c-JUN, and FOSB) in tumor cells from the No DMAB (red) and DMAB-treated (blue) samples. The p values were calculated using a Student's t test. **(C)** ChIP-qPCR analysis of c-FOS enrichment at the POSTN promoter region in 3 GCTB tissues. Significance was determined by a Student's t test. **(D)** Representative IHC images of POSTN in GCTB without recurrence (n = 14) versus GCTB with recurrence (n = 9) after DMAB treatment. Scale bar, 50 μm. Statistical table of 23 patients with GCTB by POSTN staining intensity. Significance was determined by a fisher exact test. *$P$ < 0.05; **$P$ < 0.01; ***$P$ < 0.001 and ****$P$ < 0.0001. All the results were obtained from three independent experiments. Values are presented as mean ± SD.

to exhausted T cells. Moreover, we identified the upregulation of POSTN as a possible mechanism of T cell exhaustion and proposed its relevance to recurrence in patients with GCTB treated with DMAB (Figure 6). This study provides a comprehensive cellular interaction atlas of DMAB-treated GCTB and a framework for improving DMAB therapy in the future.

DMAB may have dual functions in immunomodulation. Previous preclinical data showed that DMAB treatment increased tumor-infiltrating lymphocytes (TILs) in breast cancer, indicating that DMAB might improve the response to immunotherapy in patients with breast cancer. Based on this observation, new clinical trials for DMAB combined with immune checkpoint inhibitors should be conducted for breast cancer (34). In the present study, we used scRNA-seq combined with immunostaining to demonstrate that DMAB induces T-cell exhaustion in GCTB. In chronic infections and cancer, T cell exhaustion is caused by long-term exposure to persistent antigen (35).

To investigate the causes of T-cell exhaustion in GCTB with DMAB therapy, we analyzed the transcriptome of immune cells from GCTB and found that $CD8^+$ T cell exhaustion integrates stimuli from altered tumor microenvironments. $CD8^+$ T cells treated with DMAB showed increased costimulatory markers, indicating that DMAB may promote the continuous exposure of $CD8^+$ T cells to the persisting antigen, especially in patients who receive DMAB at a higher frequency and longer time. Additional signals from inhibitory receptors (PD1, LAG3, and CTLA4) may also contribute to $CD8^+$ T cell exhaustion in GCTB. Our understanding of the mechanisms by which inhibitory receptors control T-cell exhaustion in GCTB remains unclear. Our scRNA-seq data revealed that these inhibitory receptors might cause transient intracellular attenuation of positive signals and the induction of inhibitory gene expression to facilitate T-cell exhaustion. Because T cell exhaustion is reversible, agents that target inhibitory receptors may improve DMAB therapy.

The efficacy of DMAB in GCTB may be correlated with the M1-M2 TAM polarization status. RANK-RANKL induces macrophage differentiation into osteoclasts. RANKL also triggers M1 polarization of macrophages during bone formation (36). However, the role of RANK-RANKL in TAM remains unclear. A previous study showed that the loss of RANK

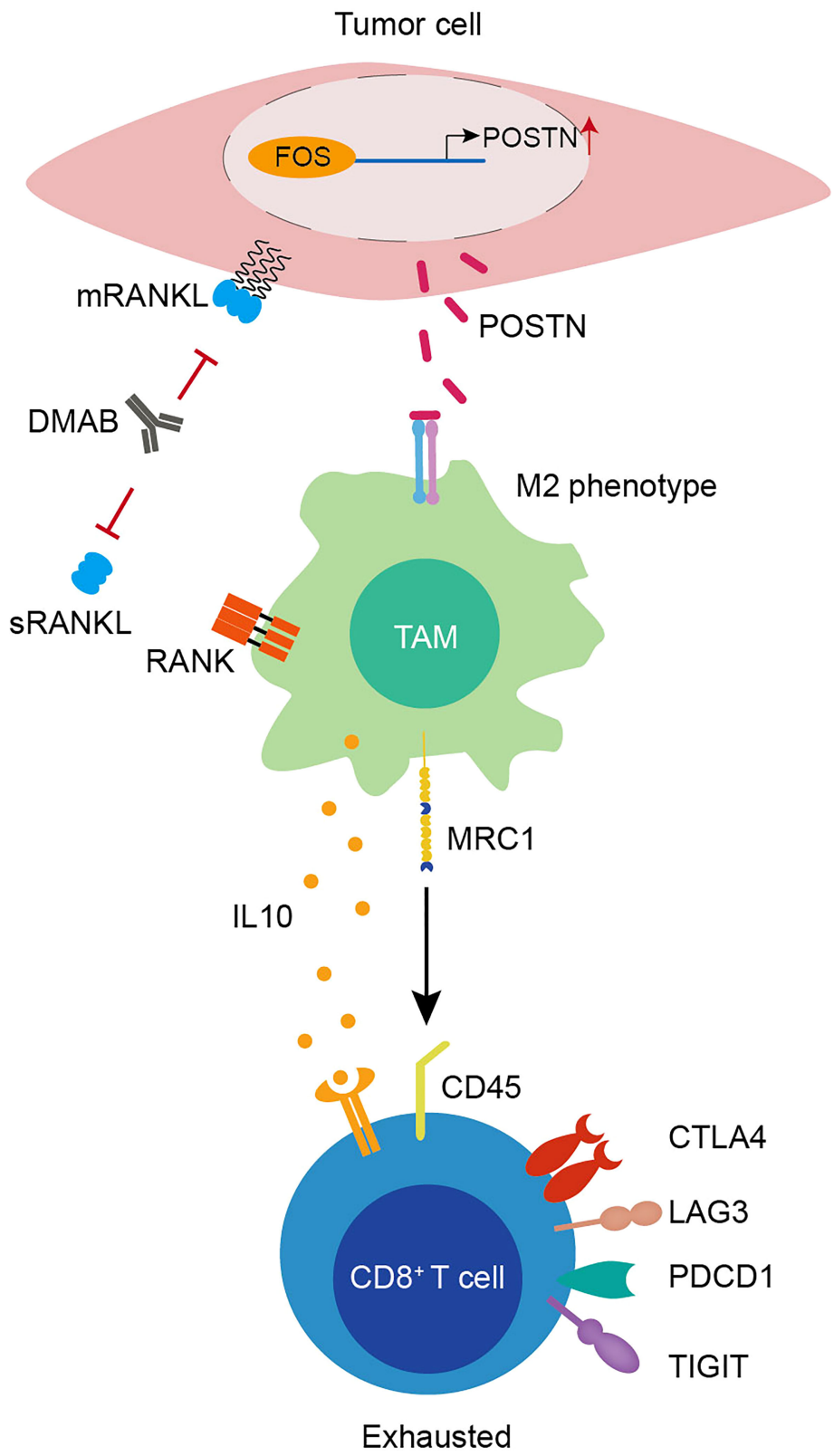



FIGURE 6
Graphical illustration of T-cell exhaustion in DMAB-treated GCTB based on bioinformatic prediction.

signaling in mouse breast cancer cells reduces TAM and tumor-associated neutrophil infiltration (34). In our study, we showed that M1 and M2 type TAMs coexist in GCTB, whereas the majority of TAMs exhibit an M2-like phenotype. This is consistent with previous reports (37–40). Intriguingly, inhibition of RANK-RANKL signaling by DMAB accompanied by M1 TAM peaking early and switching towards M2 TAM in the TME of GCTB may facilitate the escape of T-cell immune surveillance by GCTB cells. Furthermore, M2-like TAMs were composed of four distinct subsets with different transcriptome profiles. However, the four subsets have overall similarity to M2 TAMs, such as secretion of IL-10, which is an important extrinsic cytokine involved in T cell exhaustion. Although the M1 to M2 phenotypic transition of TAM can be promoted by GCTB-secreted POSTN, the precise origin and function of M2-like TAM in GCTB still require further investigation.

It has been well demonstrated that POSTN expression is deregulated in malignant transformation. High POSTN expression levels are usually associated with aggressive tumor behavior and poor prognosis in cancer (41–44). For example, using immunohistochemical analyses, Hu et al. showed that POSTN expression was higher in osteosarcoma than in osteochondroma. Osteosarcoma patients with high levels of POSTN had a worse prognosis than those with low POSTN expression (45). Our findings revealed that POSTN expression was correlated with the recurrence of GCTB in patients receiving DMAB therapy, suggesting that POSTN levels could be a useful prognostic biomarker in GCTB. Recent studies have also shown that POSTN plays an important role in cancer treatment resistance. Liu et al. showed that POSTN confers gemcitabine resistance in pancreatic cancer cells (46). Hu et al. demonstrated the effect of POSTN on cisplatin resistance in NSCLC cells (47). Sung et al. also found a correlation between cisplatin resistance and POSTN in patients with ovarian cancer (48). Recombinant POSTN promotes resistance to carboplatin and paclitaxel in ovarian cancer cells (49). Taken together, these findings, including our work, suggest that POSTN targeting could be a new therapeutic approach to overcome DMAB therapy failure in GCTB.

## Data availability statement

The data presented in the study are deposited in the NCBIGene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/), accession number GSE212341.

## Ethics statement

The studies involving human participants were reviewed and approved by IEC for Clinical Research and Animal Trials of the First Affiliated Hospital of Sun Yat-sen University. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

MY performed the scRNA-seq analyses, IHC staining assays, and wrote the manuscript. FW performed the IHC staining assays. GL performed ChIP-qPCR assays. MC interpreted the data and revised the manuscript. WZ and CZ designed the experiments, interpreted the data, wrote the manuscript, and provided supervision. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the National Natural Science Foundation of China (81972507, 81972651, and 82172698), and CSCO-GCT research Project (Y-2020GCT-003).

## Acknowledgments

We thank Dr. Meng Zhao for valuable suggestions.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.934078/full#supplementary-material

# References


1. Huang L, Xu J, Wood DJ, Zheng MH. Gene expression of osteoprotegerin ligand, osteoprotegerin, and receptor activator of NF-kappaB in giant cell tumor of bone: possible involvement in tumor cell-induced osteoclast-like cell formation. *Am J Pathol* (2000) 156(3):761–7. doi: 10.1016/s0002-9440(10)64942-5

2. Roux S, Amazit L, Meduri G, Guiochon-Mantel A, Milgrom E, Mariette X. RANK (receptor activator of nuclear factor kappa b) and RANK ligand are expressed in giant cell tumors of bone. *Am J Clin Pathol* (2002) 117(2):210–6. doi: 10.1309/bpet-f2pe-p2bd-j3p3

3. Atkins GJ, Kostakis P, Vincent C, Farrugia AN, Houchins JP, Findlay DM, et al. RANK expression as a cell surface marker of human osteoclast precursors in peripheral blood, bone marrow, and giant cell tumors of bone. *J Bone Mineral Res* (2006) 21(9):1339–49. doi: 10.1359/jbmr.060604

4. van der Heijden L, Dijkstra PDS, Blay J-Y, Gelderblom H. Giant cell tumour of bone in the denosumab era. *Eur J Cancer* (2017) 77:75–83. doi: 10.1016/j.ejca.2017.02.021

5. Chawla S, Henshaw R, Seeger L, Choy E, Blay J-Y, Ferrari S, et al. Safety and efficacy of denosumab for adults and skeletally mature adolescents with giant cell tumour of bone: interim analysis of an open-label, parallel-group, phase 2 study. *Lancet Oncol* (2013) 14(9):901–8. doi: 10.1016/S1470-2045(13)70277-8

6. Noh BJ, Park YK. Giant cell tumor of bone: updated molecular pathogenesis and tumor biology. *Hum Pathol* (2018) 81:1–8. doi: 10.1016/j.humpath.2018.06.017

7. Thomas D, Henshaw R, Skubitz K, Chawla S, Staddon A, Blay J-Y, et al. Denosumab in patients with giant-cell tumour of bone: an open-label, phase 2 study. *Lancet Oncol* (2010) 11(3):275–80. doi: 10.1016/S1470-2045(10)70010-3

8. Chawla S, Blay J-Y, Rutkowski P, Le Cesne A, Reichardt P, Gelderblom H, et al. Denosumab in patients with giant-cell tumour of bone: a multicentre, open-label, phase 2 study. *Lancet Oncol* (2019) 20(12):1719–29. doi: 10.1016/S1470-2045(19)30663-1

9. Agarwal MG, Gundavda MK, Gupta R, Reddy R. Does denosumab change the giant cell tumor treatment strategy? lessons learned from early experience. *Clin Orthop Relat Res* (2018) 476(9):1773–82. doi: 10.1007/s11999.0000000000000243

10. Aponte-Tinao LA, Piuzzi NS, Roitman P, Farfalli GL. A high-grade sarcoma arising in a patient with recurrent benign giant cell tumor of the proximal tibia while receiving treatment with denosumab. *Clin Orthop Relat Res* (2015) 473 (9):3050–5. doi: 10.1007/s11999-015-4249-2

11. Errani C, Tsukamoto S, Leone G, Righi A, Akahane M, Tanaka Y, et al. Denosumab may increase the risk of local recurrence in patients with giant-cell tumor of bone treated with curettage. *J Bone Joint Surg Am* (2018) 100(6):496–504. doi: 10.2106/JBJS.17.00057

12. Dougall WC, Glaccum M, Charrier K, Rohrbach K, Brasel K, De Smedt T, et al. RANK is essential for osteoclast and lymph node development. *Genes Dev* (1999) 13(18):2412–24. doi: 10.1101/gad.13.18.2412

13. Cheng M, Fong L. Effects of RANKL-targeted therapy in immunity and cancer. *Front Oncol* (2014) 3:329. doi: 10.3389/fonc.2013.00329

14. Leibbrandt A, Penninger JM. RANK/RANKL: regulators of immune responses and bone physiology. *Ann New York Acad Sci* (2008) 1143(1):123–50. doi: 10.1196/annals.1443.016

15. Kong Y-Y, Yoshida H, Sarosi I, Tan H-L, Timms E, Capparelli C, et al. OPGL is a key regulator of osteoclastogenesis, lymphocyte development and lymph-node organogenesis. *Nature* (1999) 397(6717):315–23. doi: 10.1038/16852

16. Josien R, Li HL, Ingulli E, Sarma S, Wong BR, Vologodskaia M, et al. TRANCE, a tumor necrosis factor family member, enhances the longevity and adjuvant properties of dendritic cells in vivo. *J Exp Med* (2000) 191(3):495–502. doi: 10.1084/jem.191.3.495

17. Bachmann MF, Wong BR, Josien R, Steinman RM, Oxenius A, Choi Y. TRANCE, a tumor necrosis factor family member critical for CD40 ligand-independent T helper cell activation. *J Exp Med* (1999) 189(7):1025–31. doi: 10.1084/jem.189.7.1025

18. Wiethe C, Dittmar K, Doan T, Lindenmaier W, Tindle R. Enhanced effector and memory CTL responses generated by incorporation of receptor activator of NF-kappa b (RANK)/RANK ligand costimulatory molecules into dendritic cell immunogens expressing a human tumor-specific antigen. *J Immunol* (2003) 171 (8):4121–30. doi: 10.4049/jimmunol.171.8.4121

19. Anderson DM, Maraskovsky E, Billingsley WL, Dougall WC, Tometsko ME, Roux ER, et al. A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature* (1997) 390(6656):175–9. doi: 10.1038/36593

20. Ahern E, Smyth MJ, Dougall WC, Teng MWL. Roles of the RANKL-RANK axis in antitumour immunity - implications for therapy. *Nat Rev Clin Oncol* (2018) 15(11):676–93. doi: 10.1038/s41571-018-0095-y

21. Cheng S, Li Z, Gao R, Xing B, Gao Y, Yang Y, et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* (2021) 184(3):792–809.e23. doi: 10.1016/j.cell.2021.01.010

22. Zheng L, Qin S, Si W, Wang A, Xing B, Gao R, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science* (2021) 374(6574):abe6474. doi: 10.1126/science.abe6474

23. Bi K, He MX, Bakouny Z, Kanodia A, Napolitano S, Wu J, et al. Tumor and immune reprogramming during immunotherapy in advanced renal cell carcinoma. *Cancer Cell* (2021) 39(5):649–61 e5. doi: 10.1016/j.ccell.2021.02.015

24. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, et al. Therapy-induced evolution of human lung cancer revealed by single-cell RNA sequencing. *Cell* (2020) 182(5):1232–51 e22. doi: 10.1016/j.cell.2020.07.017

25. Feng W, He M, Jiang X, Liu H, Xie T, Qin Z, et al. Single-cell RNA sequencing reveals the migration of osteoclasts in giant cell tumor of bone. *Front Oncol* (2021) 11:715552. doi: 10.3389/fonc.2021.715552

26. Hao Y, Hao S, Andersen-Nissen E, Mauck WM3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell* (2021) 184(13):3573–87 e29. doi: 10.1016/j.cell.2021.04.048

27. 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):404–21 e16. doi: 10.1016/j.cell.2020.11.041

28. Zhou Y, Yang D, Yang Q, Lv X, Huang W, Zhou Z, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. *Nat Commun* (2020) 11(1):6322. doi: 10.1038/s41467-020-20059-6

29. Lu B, Zou C, Yang M, He Y, He J, Zhang C, et al. Pharmacological inhibition of core regulatory circuitry liquid-liquid phase separation suppresses metastasis and chemoresistance in osteosarcoma. *Advanced Sci (Weinheim Baden-Wurttemberg Germany)* (2021) 8(20):e2101895. doi: 10.1002/advs.202101895

30. Jin SQ, Guerrero-Juarez CF, Zhang LH, 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

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

32. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* (2014) 32(4):381–6. doi: 10.1038/nbt.2859

33. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-cell mRNA quantification and differential analysis with census. *Nat Methods* (2017) 14(3):309–15. doi: 10.1038/nmeth.4150

34. Gómez-Aleza C, Nguyen B, Yoldi G, Ciscar M, Barranco A, Hernández-Jiménez E, et al. Inhibition of RANK signaling in breast cancer induces an anti-tumor immune response orchestrated by CD8+ T cells. *Nat Commun* (2020) 11 (1):6335. doi: 10.1038/s41467-020-20138-8

35. Wherry EJ, Kurachi M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol* (2015) 15(8):486–99. doi: 10.1038/nri3862

36. Huang R, Wang X, Zhou Y, Xiao Y. RANKL-induced M1 macrophages are involved in bone formation. *Bone Res* (2017) 5(1):17019. doi: 10.1038/boneres.2017.19

37. Jin S, Li R, Chen MY, Yu C, Tang LQ, Liu YM, et al. Single-cell transcriptomic analysis defines the interplay between tumor cells, viral infection, and the microenvironment in nasopharyngeal carcinoma. *Cell Res* (2020) 30 (11):950–65. doi: 10.1038/s41422-020-00402-8

38. Zhang Q, He Y, Luo N, Patel SJ, Han Y, Gao R, et al. Landscape and dynamics of single immune cells in hepatocellular carcinoma. *Cell* (2019) 179 (4):829–45.e20. doi: 10.1016/j.cell.2019.10.003

39. Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* (2018) 174(5):1293–308.e36. doi: 10.1016/j.cell.2018.05.060

40. Müller S, Kohanbash G, Liu SJ, Alvarado B, Carrera D, Bhaduri A, et al. Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. *Genome Biol* (2017) 18(1):1–14. doi: 10.1186/s13059-017-1362-4

41. González-González L, Alonso J. Periostin: A matricellular protein with multiple functions in cancer development and progression. *Front Oncol* (2018) 8:225. doi: 10.3389/fonc.2018.00225

42. Ratajczak-Wielgomas K, Grzegrzolka J, Piotrowska A, Matkowski R, Wojnar A, Rys J, et al. Expression of periostin in breast cancer cells. *Int J Oncol* (2017) 51(4):1300–10. doi: 10.3892/ijo.2017.4109

43. Lincoln V, Chao L, Woodley DT, Murrell D, Kim M, O'Toole EA, et al. Over-expression of stromal periostin correlates with poor prognosis of cutaneous squamous cell carcinomas. *Exp Dermatol* (2021) 30(5):698–704. doi: 10.1111/exd.14281

44. Zhang W, Wu FH, Wang J, Wei HZ, Wang TW, Zhang YC. Periostin: a predictable molecule to prognosis and chemotherapy responses of gastrointestinal and hepato-biliary-pancreatic malignant tumors? *Neoplasma* (2022) 69 (3):491–503. doi: 10.4149/neo_2021_211025N1521

45. Hu F, Shang X-F, Wang W, Jiang W, Fang C, Tan D, et al. High-level expression of periostin is significantly correlated with tumour angiogenesis and poor prognosis in osteosarcoma. *Int J Exp Pathol* (2016) 97(1):86–92. doi: 10.1111/iep.12171

46. Liu Y, Li F, Gao F, Xing L, Qin P, Liang X, et al. Periostin promotes the chemotherapy resistance to gemcitabine in pancreatic cancer. *Tumor Biol* (2016) 37 (11):15283–91. doi: 10.1007/s13277-016-5321-6

47. Hu W, Jin P, Liu W. Periostin contributes to cisplatin resistance in human non-small cell lung cancer A549 cells *via* activation of Stat3 and akt and upregulation of survivin. *Cell Physiol Biochem* (2016) 38(3):1199–208. doi: 10.1159/000443068

48. Sung PL, Jan YH, Lin SC, Huang CC, Lin H, Wen KC, et al. Periostin in tumor microenvironment is associated with poor prognosis and platinum resistance in epithelial ovarian carcinoma. *Oncotarget* (2016) 7(4):4036–47. doi: 10.18632/oncotarget.6700

49. Ryner L, Guan Y, Firestein R, Xiao Y, Choi Y, Rabe C, et al. Upregulation of periostin and reactive stroma is associated with primary chemoresistance and predicts clinical outcomes in epithelial ovarian cancer. *Clin Cancer Res* (2015) 21 (13):2941–51. doi: 10.1158/1078-0432.ccr-14-3111

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 15 September 2022
DOI 10.3389/fonc.2022.928022

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Shuying Shen,
Sir Run Run Shaw Hospital, China
Xiaowen Zhao,
Qingdao Central Hospital, China

*CORRESPONDENCE
Xinhua Zhu
drzhuxh@njglyy.com

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

RECEIVED 25 April 2022
ACCEPTED 15 July 2022
PUBLISHED 15 September 2022

CITATION
Pan J, Hu Y, Yuan C, Wu Y and Zhu X (2022) lncRNA NEAT1 promotes the proliferation and metastasis of hepatocellular carcinoma by regulating the FOXP3/PKM2 axis. *Front. Oncol.* 12:928022. doi: 10.3389/fonc.2022.928022

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

# lncRNA NEAT1 promotes the proliferation and metastasis of hepatocellular carcinoma by regulating the FOXP3/PKM2 axis

Junping Pan[1], Yingzhe Hu[2], Chenlu Yuan[2], Yafu Wu[1] and Xinhua Zhu[1*]

[1]Department of Hepatobiliary Surgery, Affiliated Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China, [2]Department of Hepatobiliary Surgery, Nanjing Drum Tower Hospital Clinical College of Traditional Chinese and Western Medicine, Nanjing University of Chinese Medicine, Nanjing, China

**Objective:** Hepatocellular carcinoma (HCC) is a malignant tumor. The occurrence of HCC is involved in the alteration of a variety of oncogenes or tumor suppressor genes, but the specific molecular mechanism remains unknown. This research proved the effects of long non-coding RNA NEAT1 (lncRNA NEAT1) on the viability, proliferation, migration, and invasion of hepatocellular carcinoma cells and explored the mechanism behind these effects.

**Methods:** NEAT1 in 97H and Huh7 cell lines was overexpressed or knocked down, respectively. The expression of FOXP3 and its target gene PKM2 was hinged on qRT-PCR and Western blot, respectively. RNA pulldown and RNA immunoprecipitation experiments were carried out to detect the interaction between NEAT1 and proteins. Finally, the effect of NEAT1 on the tumor volume of HCC was verified by animal experiments.

**Results:** A series of experiments have shown that NEAT1 knockdown can inhibit the viability, proliferation, migration, and invasion of HCC cells; NEAT1 can bind FOXP3 to promote PKM2 transcription; PKM2 knockdown can inhibit the viability, proliferation, migration, and invasion of HCC cells; and PKM2 knockdown reversed the function of NEAT1.

**Conclusion:** lncRNA NEAT1 can promote the malignant behavior of HCC cells, while silencing of NEAT1 can inhibit that behavior of HCC cells. Mechanically, NEAT1 promotes the transcriptional activation of PKM2 by binding FOXP3, and PKM2 knockout reverses the function of NEAT1.

KEYWORDS

**hepatocellular carcinoma, NEAT1, FOXP3, PKM2, long non-coding RNA**

# Introduction

Hepatocellular carcinoma (HCC) is the common reason of cancer death (1–3). At present, the treatment of HCC has made great progress, but the treatment in preventing metastasis and recurrence is still very limited (4). The pathogenesis of HCC is very complex, involving multiple genes and signaling pathways. Exploring the mechanisms of HCC progression is necessary to identify new therapeutic approaches (5).

Long non-coding RNA (lncRNA) is a structure with a 5'-cap and 3'-polyadenylated tail (6). lncRNA could bind with transcription factors, activators, inhibitors, or other target genes to regulate transcription and expression. lncRNAs are involved in a variety of biological processes (7, 8). lncRNA NEAT1 belongs to the lncRNA (9, 10). At present, lncRNA NEAT1 expression changes have been found in a variety of human tumors, and it is related to tumor metastasis (11), occurrence, and development (12). However, studies on lncRNA NEAT1 in HCC have been limited.

Earlier studies showed that FOXP3 was expressed in T cells (13). Later, it was also found to be expressed in kidney, breast, melanoma, and other cancer cells and tissues (14). FOXP3 regulates the expression of downstream genes by binding with the promoter region or other transcription factors and is involved in the biological functions of tumor cells (15). The study of FOXP3-related functions may display a target for the targeted therapy of HCC.

Pyruvate kinase has four isoenzymes, namely, PKM1, PKM2, PKL, and PKR. During tumor formation, other isoenzymes disappear and PKM2 is highly expressed, laying a new biological foundation for cancer treatment (16, 17). PKM2 is a pivotal protein associated with microenvironmental sugar metabolism, cell growth signals, and oxidative stress in cancer cells. In terms of aerobic glycolysis of cancer cells, PKM2 acts as a biofunctional sensor and regulator of glycolysis in the form of high-activity tetramer and low-activity dimer to determine whether it is metabolized into lactic acid or synthesized into biological macromolecules.

To find the lncRNA NEAT1 mechanism *in vitro* and *in vivo*, studies have shown that NEAT1, by binding FOXP3, regulates the transcriptional activation of PKM2. This study extends our understanding of the role lncRNAs play in HCC development and may provide novel targets for the clinical therapy of HCC.

# Materials and methods

## Materials

The primary antibodies used in the experiment were purchased from BD, USA, and the secondary antibodies were purchased from Invitrogen, USA. PcDNA3.1 carriers are provided by Shanghai Gemal Biotechnology Co., Ltd., Shanghai, China. Lipofectamine™ 2000 transfection reagent was purchased from BD, USA. The RT-PCR instrument was purchased from BD, USA, and the HBS-1096B microplate was purchased from Nanjing Detie Experimental Equipment Co., Ltd. The Western blot electrophoresis equipment was purchased from Bio-Rad, USA.

## Experimental subject

In this study, HCC tissues were obtained from patients undergoing surgery at the Affiliated Drum Tower Hospital ($n$ = 56), all of whom had received preoperative radiotherapy. These tissues were quick-frozen in liquid nitrogen and stored at −80°C until needed. Informed consent was been obtained from all patients. The scheme was approved by our research ethics committee.

## Cell culture, transfection, and grouping

Hepatocellular carcinoma cell lines including 97H and Huh7 were from the medical school and cultured in RPMI 1640 medium at 37°C in a 5% $CO_2$ incubator. The siRNAs targeting NEAT1, FOXP3, and PKM2 as well as the negative control were purchased from GenePharma. The overexpressing vector of NEAT1 and PKM2 and the vector control were from GenePharma. The knockdown vector of Psilencer/PKM2 was synthesized by Hanbio. The plasmids were transfected into HCC cells using Lipofectamine 3000. The transfection concentration was 300 nmol/well. The cells were divided into the following groups: pcDNA3.1 (empty carrier), pcDNA3.1/NEAT1 (transfected NEAT1 plasmid), Si-nc (transfected with NEAT1 negative plasmid), Si-NEAT1 1# (transfected NEAT1 low-expression plasmid 1#), Si-NEAT1 2# (transfected NEAT1 low-expression plasmid 2#), vector (empty carrier), NEAT1 (transfected NEAT1 overexpression plasmid), Si-NEAT1 (transfected with NEAT1 low-expression plasmid), Si-FOXP3 1# (transfected FOXP3 low-expression plasmid 1#), Si-FOXP3 2# (transfected FOXP3 low-expression plasmid 2#), NEAT1 + Si-nc (NEAT1 group continued to transfect FOXP3-negative plasmid), NEAT1 + Si-FOXP3 (NEAT1 group continued to transfect FOXP3 low-expression plasmid), Ad-vector (adenovirus transfected with empty vectors), Ad-NEAT1 (adenovirus transfected with NEAT1), pcDNA3.1/PKM2 (transfected PKM2 plasmid), Si-PKM2 (transfected with PKM2 low-expression plasmid), and NEAT1 + Si-PKM2 (NEAT1 group continued to transfect PKM2 low-expression plasmid).

## MTT

HCC cells were inoculated into 96-well plates at a density of $1 \times 10^4$ cells/well and incubated overnight at 37°C in a 5% $CO_2$ incubator. Cells were treated with 0, 4, 20, 100, 500, and 2,500 μg/ml of the chemotherapeutic drug TMZ (Kaigi Company, Nanjing, China) in a medium containing 10% FBS for 24 h. Meanwhile, the $IC_{50}$ of TMZ was detected. Then, 10 μl of MTT (0.5 mg/ml) was added to each well and incubated for 4 h. After removing the supernatant, 200 μl of dimethyl sulfoxide (DMSO) was added to stop the reaction and incubated at 37°C for 15 min. The absorbance at 490 nm was measured with a Bio-Rad microplate meter (Bio-Rad, Hercules, CA, USA).

## The cell colony formation experiment

After different treatments, the cells were digested and cultured in six-well plates for 2 weeks. Afterward, the cells were fixed with 10% formaldehyde and stained with 1% crystal violet. After washing, the stained cell colonies were observed and imaged (Leica, Germany). The number of cell clones was counted.

## Luciferase report

The PKM2 promoter was amplified from the cDNA of the HEK293T cell and subcloned into the luciferase reporter vector. HEK293T cells were co-transfected with 100 ng of luciferase vectors and other plasmids (50 nM). Cells were collected for 48 h.

## Biotinylated pulldown assay

The biotinylated pulldown assay was based on previous instructions. In brief, the biotin-coupled RNA complexes were captured and transfected with 50 μM of the biotinylated NEAT1 probe or the control probe. Streptavidin beads (100 μl) were added to each reaction tube, and the biotin-coupled RNA complexes were assembled at room temperature. Western blot was used to evaluate the expression of proteins bound to NEAT1.

## Transwell assay

For the migration assay, transfected $3 \times 10^4$ HCC cells were inoculated on the upper surface of the Transwell compartment. RPMI 1640 medium containing 12% FBS was added to the lower cell chamber. The cells were incubated in an incubator at 37°C for 24 h. The cells under the septum were then fixed and stained with 0.2% crystal violet solution for 10 min. Ten random fields of view (×200) were examined and the number of cells under the septum was recorded. Regarding cell invasion, a Transwell chamber with a precoated 100-μl Matrigel was used.

## RNA extraction and qRT-PCR

RNA was extracted and separated using the All-in-One miRNA extraction kit and the All-in-One miRNA qRT-PCR assay kit following the instructions of the manufacturer. After RNA was extracted, the RNA concentration was determined by a NanoDrop 1000 spectrophotometer. Real-time PCR reaction was performed using an SYBR Green Mix kit and a FAST7500 Real-time PCR system. The gene sequences are shown in Table 1. The $2^{-\Delta\Delta Ct}$ quantitative method was used to calculate the relative expression of the genes.

## Western blot assay

Cells were collected, lysed, and denatured. The Bradford method was performed to evaluate the amount of proteins. Forty micrograms of protein was separated using sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE) electrophoresis. The protein was transferred to a polyvinylidene fluoride membrane using the electrotransfer method after being blocked with skimmed milk for 2 h. Subsequently, the blots were visualized with primary antibody incubated overnight at 4°C, followed by goat anti-rabbit secondary antibody incubation for 2 h. The blots were visualized with the ECL Fluorescence Detection Kit. Photographs were taken and the ImageJ software was used to calculate the gray value of each band. The primary antibodies used were as follows: FOXP3 (ab20034, 1:1,000, Abcam), PKM2 (ab137852, 1:1,000, Abcam), and GAPDH (ab8245, 1:1,000, Abcam).

## RNA immunoprecipitation assay

The RNA immunoprecipitation (RIP) assay was conducted by adopting the Magna RIP™ RNA-Binding Protein immunoprecipitation kit (Millipore, Billerica, MA, USA). The

TABLE 1 qRT-PCR primer.

| Gene | Primer sequence (5'–3') |
|---|---|
| *NEAT1* | F: GTTCCGTGCTTCCTCTTCTG<br>R: GTGTCCTCCGACTTTACCAG |
| *FOXP3* | F: CTCCAATCCCTGCCCTTGACC<br>R: ACATCATCGCCCGGTTTCCA |
| *PKM2* | F: GCTGCCATCTACCACTTGC<br>R: CCAGACTTGGTGAGGACGATT |

collected cells were then lysed with lysate buffer, and the magnetic beads targeting argonaute-2 (Ago2, AB13537, Abcam) antibodies or immunoglobulin G (IgG, AB172730, Abcam) antibodies were incubated together with the RIP buffer. Then, the lysate was added and incubated at 4°C overnight. These magnetic beads were incubated with protease K, followed by total RNA isolation for subsequent qRT-PCR assay.

## Chromatin immunoprecipitation

For chromatin immunoprecipitation (ChIP) measurement, the SimpleChIP Enzymatic Chromatin IP Kit (9003; Cell Signaling Technology) was used. The cells were crosslinked in 1% formaldehyde for 10 min. Immunoprecipitation of the DNA–protein complexes was performed using the PKM2 antibody (AB137852, 1:1,000, Abcam) or rabbit IgG (Cell Signaling Technology, 2729). The bound DNA fragments were further analyzed by qRT-PCR and 2% agarose gel electrophoresis.

## Immunofluorescence

When cells grew on the flap and fused to 95%–100%, they were removed from the incubator and washed three times with 1× PBS for 10 min. The cells were then fixed with 4% formaldehyde for 20–30 min and washed three times with 1× PBS for 10 min. The cells were permeated by 0.2% Triton X-100 for 2–5 min and washed with 1× PBS three times for 10 min. BSA (5%) was used for sealing for 30 min. The sample was combined with the primary antibody and put in a wet box overnight at 4°C. After incubation of the sample with the secondary antibody in the dark for 30 min, it was washed with 1× PBS three times for 10 min. Glycerin (95%) was used to seal the film and the image was viewed under a fluorescence microscope.

## Immunohistochemical staining

Paraffin sections were dewaxed with water. The sections were incubated in 3% $H_2O_2$ for 5–10 min. The sections were washed with distilled water and soaked in PBS for 5 min. Normal goat serum (5%–10%; PBS dilution) was used for sealing. The sections were incubated for 10 min. The primary antibody was added and incubated at 37°C for 1–2 h or 4°C overnight. PBS was used for flushing three times for 5 min. An appropriate amount of biotin-labeled secondary antibody working fluid was added and incubated at 37°C for 10–30 min. An appropriate amount of horseradish enzyme or alkaline phosphatase-labeled *Streptomyces ovalbumin* working fluid was added and incubated at 37°C for 10–30 min. DAB or NBT/BCIP was used for color rendering for 3–15 min. The sections were thoroughly rinsed with tap water.

## Xenograft model in nude mice

The cells were washed with PBS and centrifuged. The cells were washed twice in a serum-free medium, and then the cells were resuspended in the serum-free medium for cell counting so that the 200-μl suspension contained $1 \times 10^7$ cells.

BALB/C nude mice were purchased and raised in SPF conditions. The animals were observed every day with unlimited access to water and food. After adapting to the new environment in the animal house at 20°C–25°C for a week, each 4–6-week-old mouse was subcutaneously inoculated with HCC cells expressing stable NEAT1 and HCC cells with NEAT1 knockout in the right axillary abdominal wall. Tumor growth was observed daily, and tumor volume was measured with a vernier caliper after tumor growth. Tumor volume = $(a \times b^2)/2$ ($a$ represents the long diameter of the tumor, and $b$ represents the short diameter of the tumor). The relative volume (RTV) of the tumor was calculated once every 2–3 days of observation. The average volume of transplanted tumors in each group was used to plot the growth curve of the transplanted tumors. Nude mice were photographed 30 days after inoculation and the tumor was removed after death. The weight of the transplanted tumors was weighed, and the tumors were surgically removed and photographed.

## Statistical analysis

The measured data were expressed as mean ± standard deviation ($x \pm s$). The Student's $t$-test was used for the comparison between two groups. Comparisons between three components were first made using ANOVA, and $P < 0.05$ was considered significant.

# Results

## NEAT1 knockdown can inhibit the viability, proliferation, migration, and invasion of hepatocellular carcinoma cells

To study the biological function of NEAT1, we first overexpressed NEAT1. Overexpression or the interference efficiency of the plasmids was detected by qPCR. Overexpressing or silencing plasmids effectively increased or reduced the expression level of NEAT1 (Figure 1A). Based on preliminary data, the effect of NEAT1 on the viability and proliferation ability of HCC cells was determined by MTT and the colony formation assay. Overexpression of NEAT1 accelerated the viability and proliferation of the 97H and Huh7 cell lines, while knockdown of NEAT1 inhibited the viability and proliferation of the HCC cell lines (Figures 1B, C). The Transwell assay was further carried out to evaluate the effect of NEAT1 on the migration ability of HCC cells. Overexpression of NEAT1 enhanced the migration formation

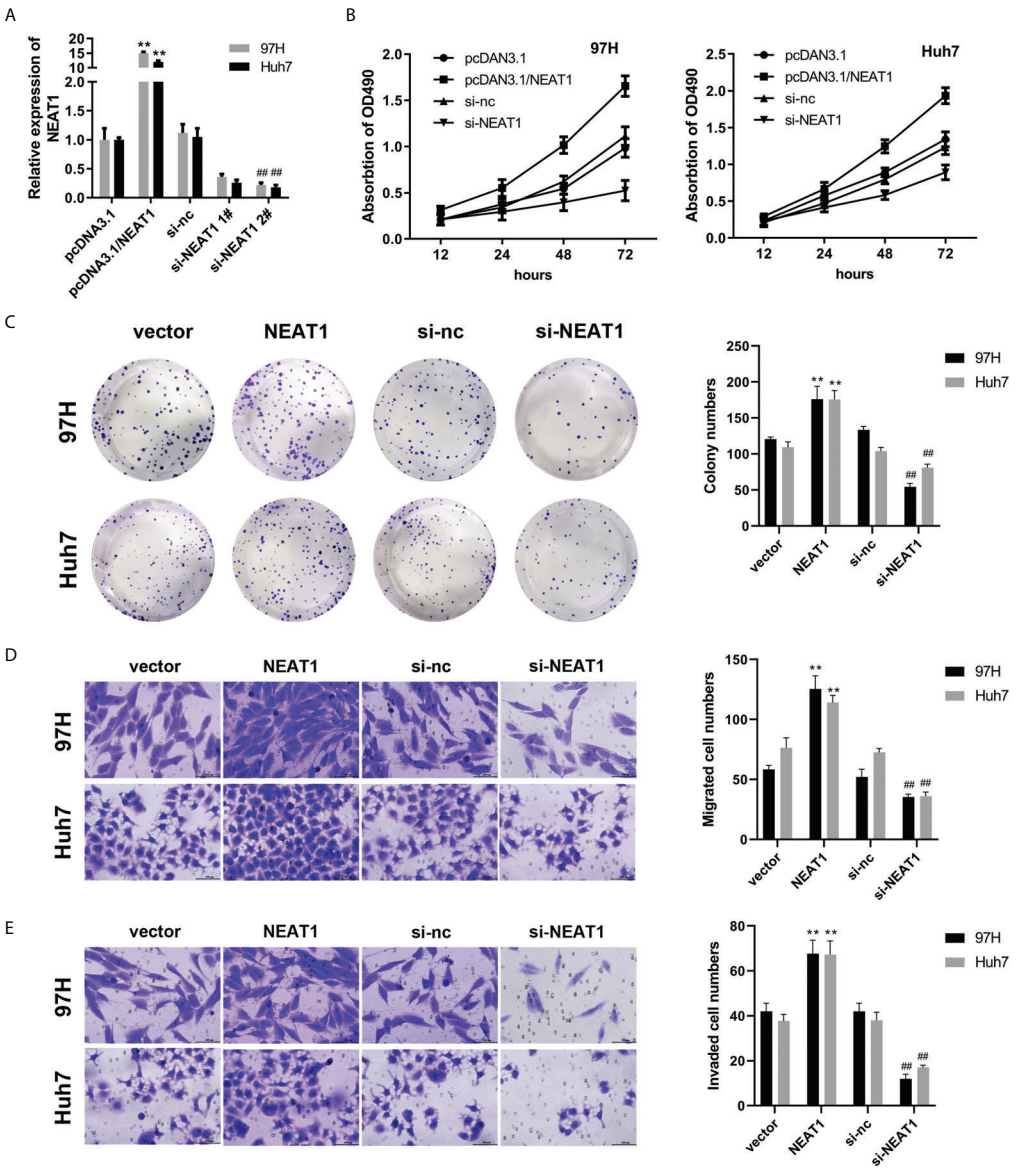



FIGURE 1
Knockdown of NEAT1 can inhibit the viability, proliferation, migration, and invasion of hepatocellular carcinoma cells. **(A)** qPCR was used to detect the overexpression or interference efficiency of NEAT1. **(B, C)** The MTT and colony formation assay were used to detect cell viability and proliferation. **(D, E)** A Transwell assay was performed to evaluate the migration and invasion of HCC cells. ** indicates $P$ <0.05 compared with pcDNA3.1 or vector; ## indicates $P$ <0.05 compared with Si-nc.

ability of the HCC cell lines. At the same time, knockdown of NEAT1 inhibited the migration of the HCC cell lines (Figures 1D, E).

## NEAT1 binds FOXP3 to promote PKM2 transcription

To study the downstream binding proteins of NEAT1, we used WB and mass spectrometry to analyze the binding proteins of NEAT1. The experimental results showed that NEAT1 could bind to FOXP3 (Figure 2A). The RNA pulldown experiment further confirmed the binding between them (Figure 2B). The RIP experimental results were the same as the RNA pulldown results, and the binding of NEAT1 and FOXP3 was verified (Figures 2C, D). Bioinformatics predicted the binding site between NEAT1 and FOXP3. The following RIP experiment showed that NEAT1 binds with FOXP3 at the 900–1,200 fragment site (Figure 2E). Bioinformatics analysis of the

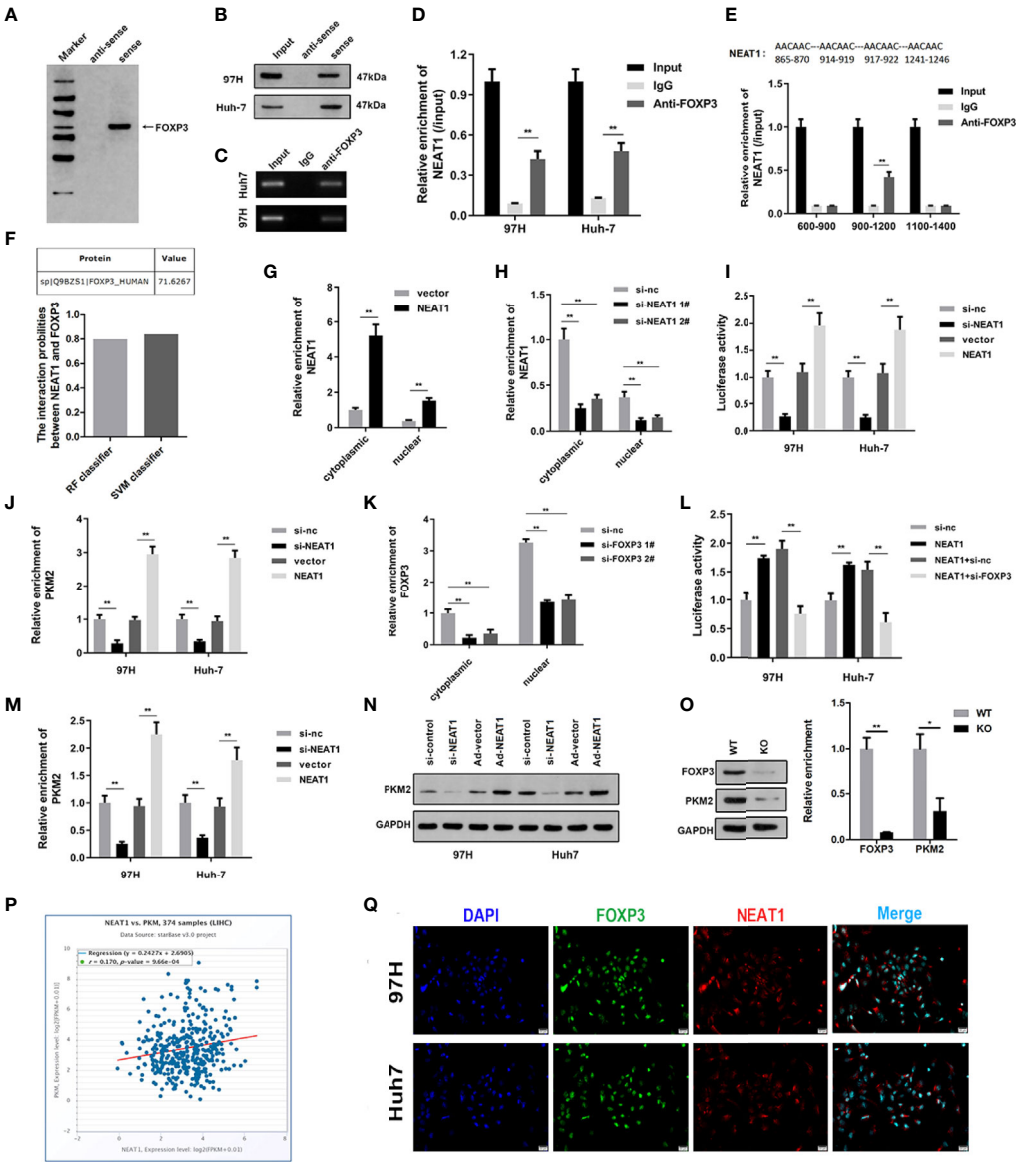



FIGURE 2
NEAT1 binds FOXP3 to promote PKM2 transcription. **(A)** RNA pulldown, silver staining, and mass spectrometry analysis were carried out to investigate the interaction between NEAT1 and FOXP3. **(B)** RNA pulldown was performed to verify the binding between NEAT1 and FOXP3. **(C, D)** RIP experiments confirmed the binding between NEAT1 and FOXP3. **(E)** Bioinformatics predicted the binding sites of NEAT1 and FOXP3, and RIP again found that NEAT1 was the binding site in the 900–1,200 fragment of FOXP3. **(F)** Bioinformatics catRAPID predicted the combination score of NEAT1 and FOXP3, proving that the possibility of the combination was high. **(G, H)** After overexpression or knockdown, qPCR was used to assess the expression of NEAT1 in the cytoplasm and nucleus. **(I)** A luciferase reporter experiment was performed to investigate whether NEAT1 overexpression or knockdown could change the transcription activity of PKM2. **(J)** A ChiP assay was performed to detect the interaction between FOXP3 and PKM2 under NEAT1 knockdown or overexpression. **(K)** qPCR was used to assess the expression of FOXP3 in the plasma and nucleus. **(L)** A luciferase reporter experiment was performed to investigate the transcription activity of PKM2 after the transfection of pcDNA3.1/NEAT1, si-FOXP3, or their combination. **(M, N)** qPCR and Western blot were used to detect the effect of NEAT1 overexpression or knockdown on PKM2 expression. **(O)** After the knockout of FOXP3 by Crispr/cas9, PKM2 expression was also weakened, confirming the regulatory relationship between FOXP3 and PKM2. **(P)** Bioinformatics analysis confirmed the positive correlation between NEAT1 and PKM2 expression. **(Q)** An immunofluorescence confocal co-localization experiment indicated the co-localization of NEAT1 and FOXP3. * indicates $P < 0.05$; ** indicates $P < 0.01$.

combined score using the catRAPID software showed a high probability, with the predicted value reaching 71.6267 (Figure 2F). After overexpression or knockdown, the level of NEAT1 in the cytoplasm and nucleus was detected to prove the efficiency of the plasmids. The experimental results showed that overexpression or knockdown plasmids could effectively promote or inhibit the expression of NEAT1 (Figures 2G, H). Overexpression of NEAT1 can promote the transcription activity of PKM2, while NEAT1 knockdown can inhibit it (Figure 2I). The ChiP experiments proved that overexpression of NEAT1 could promote the interaction between PKM2 and FOXP3, while NEAT1 knockdown inhibited it (Figure 2J). FOXP3 was silenced by the transfection of siRNAs. The qPCR results indicated that siRNA for FOXP3 significantly reduced the expression level of FOXP3 present in both the cytoplasm and nucleus (Figure 2K). Knockdown of FOXP3 reversed the effect of NEAT1 on the transcriptional activity of FOXP3 (Figure 2L). Overexpression of NEAT1 promoted the mRNA and protein expression of PKM2, whereas knockdown of NEAT1 suppressed its expression (Figures 2M, N). Moreover, after the knockout of FOXP3 by Crispr/cas9, PKM2 expression was also weakened, confirming the regulatory relationship between FOXP3 and PKM2 (Figure 2O). There was a positive correlation between NEAT1 and PKM2 expression (Figure 2P). indicating the co-localization of NEAT1 and FOXP3 (Figure 2Q).

## PKM2 is highly expressed in HCC, and knockdown of PKM2 can inhibit the viability, proliferation, migration, and invasion of HCC cells

Previous studies have shown that PKM2 is highly expressed during tumor formation. To investigate the effect of PKM2 on HCC, further studies were carried out. We first analyzed the expression of PKM2. The expression of PKM2 from the GEO database of the GSE6764 and GSE14520 datasets was analyzed, and PKM2 was highly expressed in HCC (Figures 3A, B). We performed qPCR, Western blot, and immunohistochemical staining on PKM2 in the tissues of HCC patients. PKM2 was highly expressed in HCC (Figures 3C, D). Prognostic analysis of the genes in the GSE14520 dataset showed that patients with high PKM2 expression had a shorter survival period (Figure 3E). qPCR was used to detect PKM2 expression after transfection of overexpressing or silencing plasmids (Figure 3F). Furthermore, MTT and colony formation experiments were used to detect the effect of PKM2 on HCC. Overexpression of PKM2 promoted the viability and proliferation of HCC, while PKM2 knockdown inhibited the viability and proliferation of HCC (Figure 3G, H). To further evaluate the function of PKM2, we used the Transwell experiment to assess the effect of PKM2 on the migration ability of HCC cells. Overexpression of PKM2 could enhance the migration formation ability of the HCC cell lines (Figures 3I, J).

## PKM2 knockdown reversed the function of NEAT1

To find the relationship between NEAT1 and PKM2, we applied the rescue experiment to study the regulatory relationship between the two genes. qPCR was used to detect PKM2 overexpression or interference efficiency. The results showed that the overexpressing plasmid or knockdown plasmid could effectively promote or inhibit the expression of PKM2, respectively (Figure 4A). To determine the viability and proliferation of HCC cells, overexpression of NEAT1 promoted the viability and proliferation of HCC cells, while knockdown of PKM2 reversed the viability and proliferation ability of NEAT1 on HCC cells (Figures 4B, C). Overexpression of NEAT1 promoted the migratory and invasive ability of HCC cells, whereas PKM2 knockdown reversed the migratory and invasive ability of NEAT1 on HCC cells (Figures 4D, E).

## *In-vivo* experiments were conducted to investigate the regulatory effects of NEAT1 and PKM2 on the growth of hepatocellular carcinoma cells

To prove the effects of NEAT1 and PKM2 on HCC, we built a nude mouse xenograft model. Overexpression of NEAT1 promoted tumor growth, while PKM2 knockdown reversed the oncogenic effect of NEAT1 (Figures 5A, B). The results of tumor weight were consistent with the results of tumor volume. Overexpression of NEAT1 increased the weight of tumors, while PKM2 knockdown reversed this effect of NEAT1 (Figure 5C). The expression of Ki67 and PKM2 in tumors was detected by IHC. NEAT1 promoted the expression of Ki67 and PKM2. Knockdown of PKM2 attenuated their expression compared to the NEAT1 group (Figure 5D). Furthermore, the liver metastasis of the HCC tumor was detected. As shown in Figures 5E, F, NEAT1 promoted the liver metastasis of HCC, while PKM2 knockdown reversed this effect of NEAT1.

# Discussion

Primary HCC is a malignant tumor all over the world (18, 19). At present, the postoperative metastasis and recurrence of HCC make the overall treatment effect still unsatisfactory. lncRNA plays a crucial role in normal human and disease development (20–22). The expression of lncRNA is tissue-specific and is usually dysregulated in many types of tumors. It has been confirmed that some lncRNAs are related to tumor recurrence and poor prognosis. lncRNA can regulate tumor suppressor or oncoprogenitor genes through epigenetic modification, shear, RNA degradation, and post-translational

10.3389/fonc.2022.928022

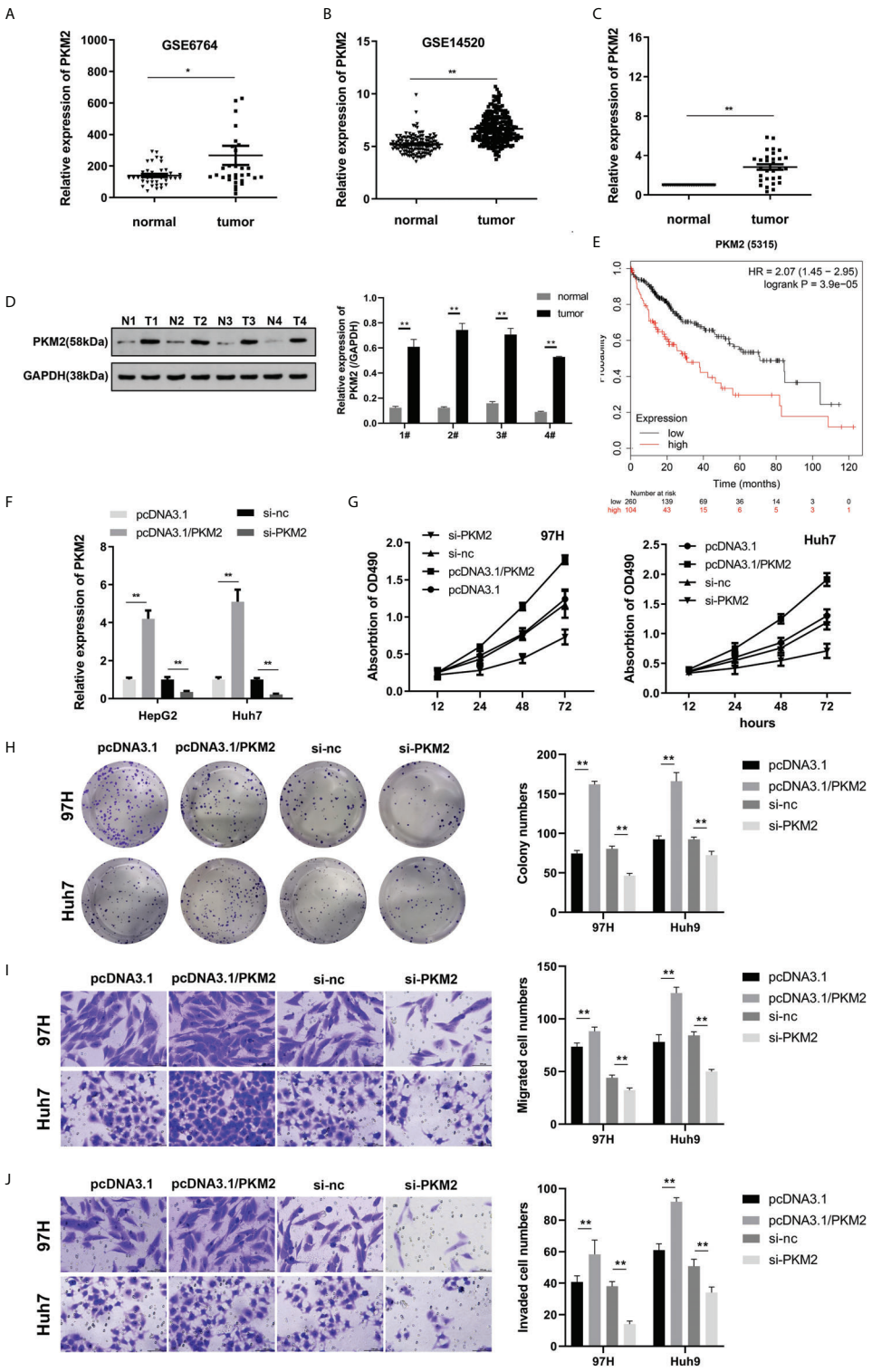



FIGURE 3
PKM2 is highly expressed in HCC, and knockdown of PKM2 can inhibit the viability, proliferation, migration, and invasion of HCC cells. **(A, B)** The expression of PKM2 based on the GSE6764 and GSE14520 datasets from the GEO database was analyzed. **(C)** The expression of PKM2 in the tissues was detected using qPCR. **(D)** Western blot was used to detect PKM2 expression in tumor tissues. **(E)** Survival analysis confirmed that PKM2 was associated with the prognosis of HCC. **(F)** qPCR was performed to detect PKM2 after overexpression or knockdown. **(G)** Cell viability was assessed using the MTT assay. **(H)** Clone formation detection of proliferation was carried out. **(I, J)** A Transwell assay was carried out to evaluate cell migration and invasion. * indicates $P$ <0.05; ** indicates $P$ <0.01.

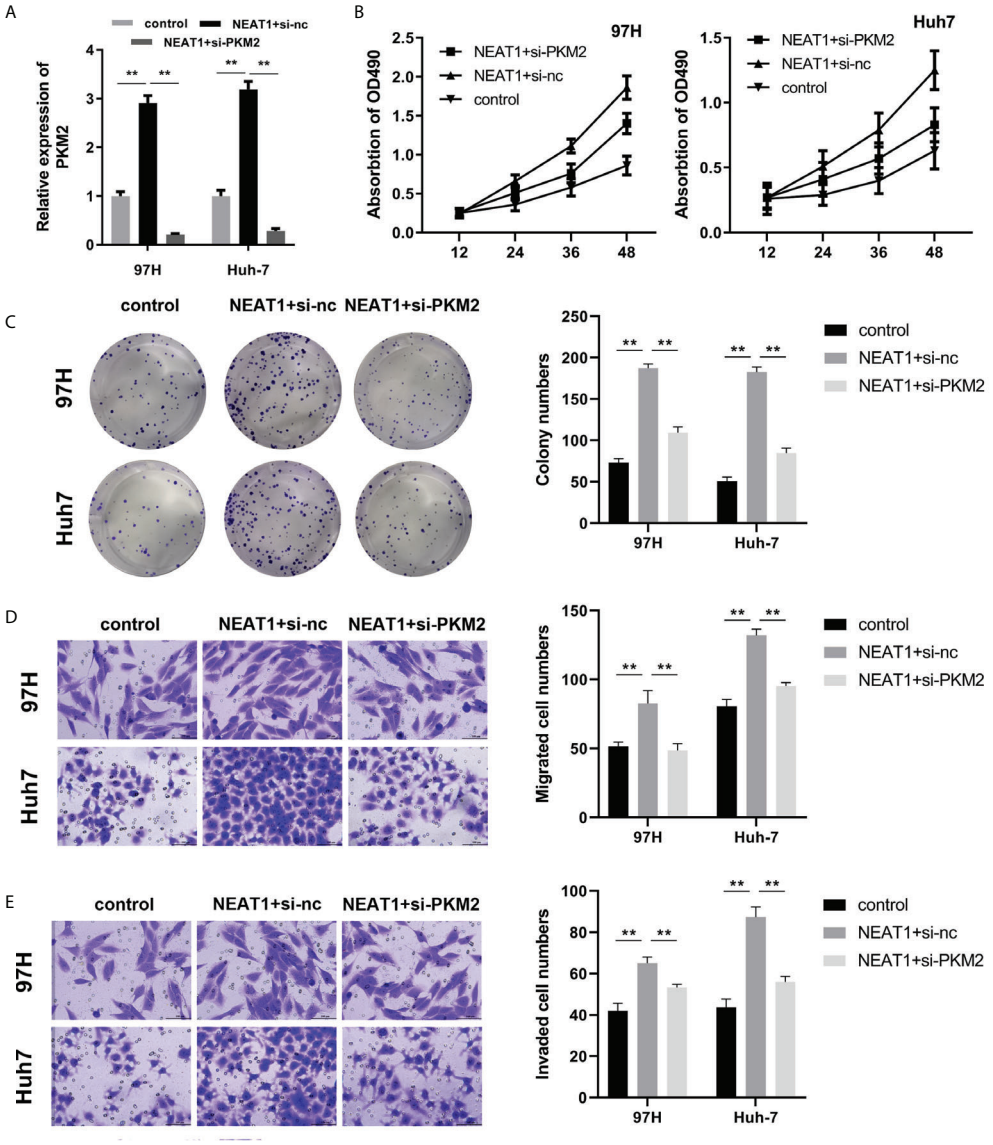



FIGURE 4
PKM2 silencing reversed the function of NEAT1. **(A)** qPCR was used to detect PKM2 expression. **(B)** The MTT assay for viability detection was performed. **(C)** A clone formation assay was used to detect proliferation. **(D, E)** A Transwell assay was carried out to evaluate cell migration and invasion. ** indicates $P < 0.01$.

modification, thus participating in tumor biological processes such as tumor generation and metastasis.

lncRNA NEAT1 was initially found in chromosome II of multiple secretory adenoma type I of familial tumor syndrome. lncRNA NEAT1 was related to TNM staging, lymph node metastasis, distant metastasis, poor prognosis, and other clinicopathological features. After silencing this gene, the viability, proliferation, and invasion of tumor cells were significantly inhibited. Chai et al. reported that the high expression of lncRNA NEAT1 in ovarian cancer tissues was not only related to clinical staging and lymph node metastasis but also promoted the malignant biological behavior of tumor cells after the high expression, while the opposite result was found after NEAT1 was knocked out. In addition, Ke et al. found that lncRNA NEAT1 regulated the survival of breast cancer cells.

As in HCC, lncRNA NEAT1 has been proved to accelerate the invasion and migration of HCC cells (23–25). However, the mechanism underlying the effect of NEAT1 remains very limited. In this study, lncRNA NEAT1 promoted the viability, proliferation, and metastasis of HCC *in vitro* and *in vivo* as previously reported. Mechanically, NEAT1 can bind with FOXP3 directly in HCC cells.

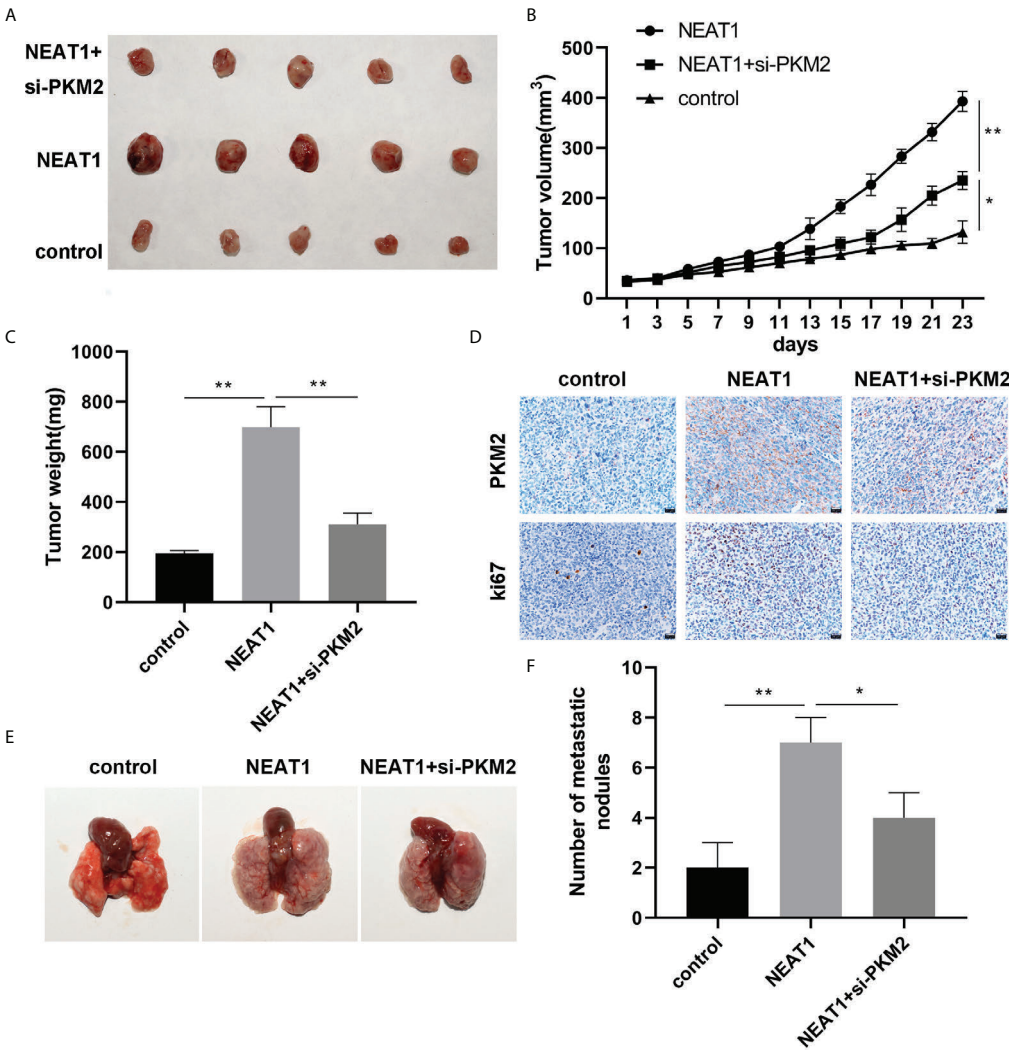



FIGURE 5
*In-vivo* experiments were conducted to investigate the effects of NEAT1 and PKM2 on the growth of HCC. **(A)** A picture of the tumor is shown. **(B)** The curve of tumor growth. **(C)** The weight of the tumors was assessed. **(D)** Histochemistry was used to detect the expression of Ki67 and PKM2 in the tumor tissues, proving the conclusion of the cell part. **(E, F)** Lung metastasis was analyzed. * indicates $P$ <0.05; ** indicates $P$ <0.01.

FOXP3 is the most specific biomarker of Treg and is involved in the maintenance and immunosuppression of Treg cells (26). The relationship between tumor cells and FOXP3 has received increasing attention. FOXP3 is expressed in many tumor cell lines and tumor cells in tumor tissues. Related research findings found that downregulation of FOXP3 inhibited tumor cell invasion by reducing MMP-9 and MMP-2. With FOXP3 knockout, tumor cells secreted less IL-10 and TGF-β1, and T-cell survival was significantly upregulated, suggesting that FOXP3 plays an important role in malignant phenotypes, particularly during invasion and immune escape (27). In HCC, the main research on FOXP3 is still focused on the immune system. For instance, $CD4^+CD25^+$ regulatory T cells in the liver of patients with hepatocellular carcinoma were notably increased. In addition, $CD4^+FOXP3^+$ regulatory T cells activated the CD39/ENTPD1 pathway and promoted hepatic metastatic tumor growth in mice. We performed RNA pulldown and RIP experiments which confirmed that NEAT1 could bind with the FOXP3 protein directly. As a transcription factor, FOXP3 modulated the expression of numerous genes transcriptionally. We screened and found that FOXP3 may regulate the transcriptional activation of PKM2 by binding with NEAT1.

In terms of growth signals of cancer cells, there is a dynamic balance between the tetramer and dimer of PKM2 (28, 29), which is regulated by the signals of the tumor suppressor and oncogenic proteins (P53, cMYC), in phosphorylation and acetylation. Anastasiou et al. reported that increased PKM2 expression could reduce ROS production (30). The more

malignant the tumor cells were, the higher the glycolysis level was, and the stronger the expression and activity of hypoxia-inducible factor (HIF)-1 were, which promoted the malignant tumor progression (31).

lncRNA NEAT1 was highly expressed in HCC cell lines. Silencing the expression of lncRNA NEAT1 can inhibit the viability, proliferation, migration, and invasion of HCC cells, and its mechanism is related to the regulation of the FOXP3/PKM2 signaling pathway. The basic research on the biological function of lncRNA NEAT1 will provide a basis for elucidating the pathogenesis of HCC and developing new therapeutic approaches.

## 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

All experiments were performed in adherence with the ARRIVE guidelines and the National Institutes of Health (NIH Publication, 8th Edition, 2011) guidelines for the use of laboratory animals. The animals care and experimental protocols were examined and verifed by Laboratory Animal Ethics Committee of Nanjing Drum Tower Hospital affiliated to Nanjing University School.

## Author contributions

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

## Funding

This work was supported by the Jiangsu Province Government Foundation (No. 2016-WSW-067).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Znaor A, Chimed T, Laversanne M, Tudev U, Sanjaajamts E, Sandagdorj T, et al. The public health challenge of HCC in Mongolia. *Lancet Gastroenterol Hepatol* (2018) 3:660–2. doi: 10.1016/S2468-1253(18)30243-7

2. Lokody I. International HCC association annual conference. *Lancet Oncol* (2015) 16:1289. doi: 10.1016/S1470-2045(15)00324-1

3. Meng J, Ai X, Lei Y, Zhong W, Qian B, Qiao K, et al. USP5 promotes epithelial-mesenchymal transition by stabilizing SLUG in hepatocellular carcinoma. *Theranostics* (2019) 9:573–87. doi: 10.7150/thno.27654

4. Xiao T, Zhong W, Zhao J, Qian B, Liu H, Chen S, et al. Polyphyllin I suppresses the formation of vasculogenic mimicry *via* Twist1/VE-cadherin pathway. *Cell Death Dis* (2018) 9:906. doi: 10.1038/s41419-018-0902-5

5. Kota J, Chivukula RR, O'Donnell KA, Wentzel EA, Montgomery CL, Hwang HW, et al. Therapeutic microRNA delivery suppresses tumorigenesis in a murine HCC model. *Cell* (2009) 137:1005–17. doi: 10.1016/j.cell.2009.04.021

6. Hua JT, Ahmed M, Guo H, Zhang Y, Chen S, Soares F, et al. Risk SNP-mediated promoter-enhancer switching drives prostate cancer through lncRNA PCAT19. *Cell* (2018) 174:564–75.e18. doi: 10.1016/j.cell.2018.06.014

7. Zhang N, Shuai K, Cheng J, Yang W, Kan Z. LncRNA linc01116 prometes glioma cell migration and invasion by modulation of radixin targeted by miR-31. *Int J Clin Exp Pathol* (2019) 12:1078–86.

8. Zhang M, Jiang X, Jiang S, Guo Z, Zhou Q, He J. LncRNA FOXD2-AS1 regulates miR-25-3p/Sema4c axis to promote the invasion and migration of colorectal cancer cells. *Cancer Manag Res* (2019) 11:10633–9. doi: 10.2147/CMAR.S228628

9. Zhang P, Cao L, Zhou R, Yang X, Wu M. The lncRNA Neat1 promotes activation of inflammasomes in macrophages. *Nat Commun* (2019) 10:1495. doi: 10.1038/s41467-019-09482-6

10. Adriaens C, Standaert L, Barra J, Latil M, Verfaillie A, Kalev P, et al. p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. *Nat Med* (2016) 22:861–8. doi: 10.1038/nm.4135

11. Zheng X, Zhang Y, Liu Y, Fang L, Li L, Sun J, et al. HIF-2alpha activated lncRNA NEAT1 promotes hepatocellular carcinoma cell invasion and metastasis by affecting the epithelial-mesenchymal transition. *J Cell Biochem* (2018) 119:3247–56. doi: 10.1002/jcb.26481

12. Wang L, Yang D, Tian R, Zhang H. NEAT1 promotes retinoblastoma progression via modulating miR-124. *J Cell Biochem* (2019)120(9):15585–93. doi: 10.1002/jcb.28825

13. Sakaguchi S, Yamaguchi T, Nomura T, Ono M. Regulatory T cells and immune tolerance. *Cell* (2008) 133:775–87. doi: 10.1016/j.cell.2008.05.009

14. Bennett CL, Christie J, Ramsdell F, Brunkow ME, Ferguson PJ, Whitesell L, et al. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nat Genet* (2001) 27:20–1. doi: 10.1038/83713

15. McInnes N, Sadlon TJ, Brown CY, Pederson S, Beyer M, Schultze JL, et al. FOXP3 and FOXP3-regulated microRNAs suppress SATB1 in breast cancer cells. *Oncogene* (2012) 31:1045–54. doi: 10.1038/onc.2011.293

16. Marsolier J, Perichon M, Weitzman JB, Medjkane S. Secreted parasite Pin1 isomerase stabilizes host PKM2 to reprogram host cell metabolism. *Commun Biol* (2019) 2:152. doi: 10.1038/s42003-019-0386-6

17. Wong TL, Ng KY, Tan KV, Chan LH, Zhou L, Che N, et al. CRAF methylation by PRMT6 regulates aerobic glycolysis driven hepatocarcinogenesis *via* ERK-dependent PKM2 nuclear relocalization and activation. *Hepatology* (2019) 71(4):1279–96. doi: 10.1002/hep.30923

18. Liu J, Fan L, Yu H, Zhang J, He Y, Feng D, et al. Endoplasmic reticulum stress causes HCC cells to release exosomal miR-23a-3p and up-regulate programmed death ligand 1 expression in macrophages. *Hepatology* (2019) 70:241–58. doi: 10.1002/hep.30607

19. Singh V, Yeoh BS, Chassaing B, Xiao X, Saha P, Aguilera Olvera R, et al. Dysregulated microbial fermentation of soluble fiber induces cholestatic HCC. *Cell* (2018) 175:679–94.e22. doi: 10.1016/j.cell.2018.09.004

20. Zou G, Zhong W, Wu F, Wang X, Liu L. Inhibition of lncRNA Neat1 by catalpol *via* suppressing transcriptional activity of NF-kappaB attenuates cardiomyocyte apoptosis. *Cell Cycle* (2019) 18:3432–41. doi: 10.1080/15384101.2019.1673619

21. Zhou ZW, Zheng LJ, Ren X, Li AP, Zhou WS. LncRNA NEAT1 facilitates survival and angiogenesis in oxygen-glucose deprivation (OGD)-induced brain microvascular endothelial cells (BMECs) *via* targeting miR-377 and upregulating SIRT1, VEGFA, and BCL-XL. *Brain Res* (2019) 1707:90–8. doi: 10.1016/j.brainres.2018.10.031

22. Zhou W, Chen X, Hu Q, Chen X, Chen Y, Huang L. Galectin-3 activates TLR4/NF-kappaB signaling to promote lung adenocarcinoma cell proliferation through activating lncRNA-NEAT1 expression. *BMC Cancer* (2018) 18:580. doi: 10.1186/s12885-018-4461-z

23. Quan D, Chen K, Zhang J, Guan Y, Yang D, Wu H, et al. Identification of lncRNA NEAT1/miR-21/RRM2 axis as a novel biomarker in breast cancer. *J Cell Physiol* (2020) 235:3372–81. doi: 10.1002/jcp.29225

24. Ling ZA, Xiong DD, Meng RM, Cen JM, Zhao N, Chen G, et al. LncRNA NEAT1 promotes deterioration of hepatocellular carcinoma based on *In vitro* experiments, data mining, and RT-qPCR analysis. *Cell Physiol Biochem* (2018) 48:540–55. doi: 10.1159/000491811

25. Jiang X, Li D, Shen W, Shen X, Liu Y. LncRNA NEAT1 promotes hypoxia-induced renal tubular epithelial apoptosis through downregulating miR-27a-3p. *J Cell Biochem* (2019) 120:16273–82. doi: 10.1002/jcb.28909

26. Wiedmann MW, Mossner J. Molecular targeted therapy of biliary tract cancer–results of the first clinical studies. *Curr Drug Targets* (2010) 11:834–50. doi: 10.2174/138945010791320818

27. Ma GF, Chen SY, Sun ZR, Miao Q, Liu YM, Zeng XQ, et al. FOXP3 inhibits proliferation and induces apoptosis of gastric cancer cells by activating the apoptotic signaling pathway. *Biochem Biophys Res Commun* (2013) 430:804–9. doi: 10.1016/j.bbrc.2012.11.065

28. Shi X, You L, Luo RY. Glycolytic reprogramming in cancer cells: PKM2 dimer predominance induced by pulsatile PFK-1 activity. *Phys Biol* (2019) 16:066007. doi: 10.1088/1478-3975/ab3f5a

29. Chen S, Youhong T, Tan Y, He Y, Ban Y, Cai J, et al. EGFR-PKM2 signaling promotes the metastatic potential of nasopharyngeal carcinoma through induction of FOSL1 and ANTXR2. *Carcinogenesis* (2019) 41(6):723–33. doi: 10.1093/carcin/bgz180

30. Anastasiou D, Poulogiannis G, Asara JM, Boxer MB, Jiang JK, Shen M, et al. Inhibition of pyruvate kinase M2 by reactive oxygen species contributes to cellular antioxidant responses. *Science* (2011) 334:1278–83. doi: 10.1126/science.1211485

31. Sonveaux P, Vegran F, Schroeder T, Wergin MC, Verrax J, Rabbani ZN, et al. Targeting lactate-fueled respiration selectively kills hypoxic tumor cells in mice. *J Clin Invest* (2008) 118:3930–42. doi: 10.1172/JCI36843

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 21 September 2022
DOI 10.3389/fonc.2022.965136

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Kai Xia,
Tongji University, China
Ping Wang,
Zhejiang University, China

*CORRESPONDENCE
Meidi Yan
2510700535@qq.com

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

RECEIVED 09 June 2022
ACCEPTED 08 August 2022
PUBLISHED 21 September 2022

CITATION
Dai J, Chen Q, Li Q, Chen M, Sun H and Yan M (2022) DIRAS3, GPR171 and RAC2 were identified as the key molecular patterns associated with brain metastasis of breast cancer. *Front. Oncol.* 12:965136. doi: 10.3389/fonc.2022.965136

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

# DIRAS3, GPR171 and RAC2 were identified as the key molecular patterns associated with brain metastasis of breast cancer

Ji Dai, Qi Chen, Guoqing Li, Mengze Chen, Haohang Sun and Meidi Yan*

General Surgery I (Thyroid, Breast, Vascular, Hernia Surgery), People's Hospital of Zhenhai, Ningbo, China

**Objective:** Brain metastasis is a primary cause of morbidity and mortality in breast cancer patients. Therefore, elucidation and understanding of the underlying mechanisms are essential for the development of new therapeutic strategies.

**Methods:** Differential gene analysis was performed for those with and without distant metastasis in The Cancer Genome Atlas (TCGA) database and those with and without recurrence in the brain in the dataset GSE12276. The differentially expressed genes procured from the two databases were intersected to obtain the intersecting genes associated with brain metastasis. Thereafter, the intersecting genes were subjected to LASSO model construction to screen for prognostic genes. The expression of the obtained genes in metastatic breast cancer was observed, and survival analysis was performed. Finally, GSEA analysis of the obtained genes was performed, and the relationship between them and immune cells was explored.

**Results:** A total of 335 differential genes for the occurrence of distant metastases were obtained based on the TCGA database. A total of 1070 differential genes for recurrence to the brain were obtained based on the dataset GSE12276. The Venn diagram showed 24 intersecting genes associated with brain metastasis. The LASSO prognostic model contained a total of five genes (GBP2, GPR171, DIRAS3, RAC2, and CACNA1D). Expression difference analysis showed that GBP2, GPR171, DIRAS3, and RAC2 were significantly down-regulated in expression in metastatic breast cancer compared with primary breast cancer tumors. Only GPR171, DIRAS3, and RAC2 were strongly correlated with the overall survival of breast cancer patients. Their correlation analysis with immune cells showed that the correlation coefficient between the expression levels of DIRAS3 and immune cells was low, and the expression levels of GPR171 and RAC2 were more closely correlated with B cells and macrophages.

**Conclusions:** The expression of DIRAS3, GPR171 and RAC2, genes associated with brain metastasis, was reduced in metastatic breast cancer, and GPR171

was found to promote brain metastasis of breast cancer cells by inducing B cells and thereby.

KEYWORDS

breast cancer, brain metastasis, bioinformatic analysis, immune cells, DIRAS3

# Introduction

Breast cancer is a life-threatening disease and a major cause of death in women. Among all malignant diseases, breast cancer is considered to be one of the leading causes of death in postmenopausal women (1). Data show that 8 million people died from malignant diseases in 2008, and this number is expected to reach 11 million by 2030 (2). Depending on the cellular origin involved, breast cancer can be divided into two main categories: carcinoma and sarcoma. Carcinoma accounts for the majority of breast cancers, while sarcomas, such as lobular tumors and angiosarcomas, are rare. Carcinoma is a breast cancer caused by the epithelial component of the breast, which consists of cells arranged in lobules and terminal ducts (3). Over the past two decades, research related to breast cancer has led to astonishing advancements in our understanding of breast cancer, resulting in further skilled treatments. However, women are still diagnosed at advanced stages due to negligence in self-examination of the breast and clinical examination (4).

Advanced breast cancer is usually metastatic, and metastases can be found in axillary lymph nodes or distant sites, such as the lung, liver, bone, and brain. Distant metastasis is a complex multistep process in which tumor cells detach from the primary tumor, infiltrate the body circulation, survive in circulation, evade immune attack, adhere to capillaries, and exude before they colonize distant organs (5). Even after the primary tumor is removed, tiny tumor cells or microscopic metastases may remain in the body, allowing cancer to recur and spread in more than 30% of breast cancer patients (6). The bone is the most affected metastatic site, while the brain is the least affected metastatic site (7). Nevertheless, brain metastases from breast cancer require attention. Although there are several available treatments for brain metastases, such as chemotherapy, radiotherapy, and targeted therapy, the survival rate of breast cancer patients with brain metastases remains low. Therefore, elucidating and understanding the underlying mechanisms are essential for the development of new therapeutic strategies.

# Methods

## Data collection

Raw counts of all breast cancer RNA sequencing data and corresponding clinical information were obtained from TCGA database (https://portal.gdc.com). Samples were grouped according to clinical information with or without distant metastasis, of which 22 samples developed distant metastasis (M1) and 907 samples did not (M0).

The dataset GSE12276 microarray raw data were downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). These contained gene expression profiles and 204 breast cancer samples with 188 samples without recurrence in the brain and 16 samples with recurrence in the brain.

## Differential gene analysis

Differentially expressed genes with or without distant metastasis and with or without recurrence in the brain were identified in both databases using the R package limma method, with $|\log2\ FC|<0.3785$, $p<0.05$ as the differential gene screening threshold. The Venn diagram shows the intersection of genes that metastasize and recur to the brain.

## Prognostic modeling

Feature selection was performed using the R package glmnet for the minimum absolute shrinkage and the LASSO regression algorithm using 10-fold cross-validation. For Kaplan–Meier curves, log-rank tests and univariate Cox proportional hazards regression were derived and included p-values and hazard ratios (HRs) with 95% confidence intervals (CIs). TimeROC analysis was used to compare the predictive accuracy and risk scores of the genes. All of the above analyses were performed using the R package. $P< 0.05$ was considered statistically significant.

## Kaplan–Meier survival curve analysis

Survival analysis was performed using the Kaplan–Meier plotter (kmplot.com/analysis). To assess the prognostic values of specific genes, patient samples were divided into two clusters (high versus low expression) based on the median expression of the genes. Log-rank p-values and HR with 95% CIs were determined. $P< 0.05$ was considered statistically significant.

## GSEA analysis

Samples were divided into high and low expression groups based on the median gene expression values, and then enrichment analysis of the KEGG pathway was performed using GSEA. FDR< 0.25, NOM p-value< 0.05 and |NES|> 1 were considered significant enrichment. The top three ranked signaling pathways were selected for this study.

# Results

## Differential gene analysis

We first performed differential gene analysis for breast cancer patients with or without distant metastasis based on TCGA database, with |log2 FC|<0.3785, p< 0.05 as the screening threshold and obtained a total of 335 differential genes, including 163 up-regulated and 172 down-regulated genes. In addition, we performed differential gene analysis for breast cancer with or without recurrence in the brain in the GEO dataset GSE12276 using |log2 FC|<0.3785, p<0.05 as the screening threshold and obtained a total of 1070 differential genes, including 403 up-regulated and 667 down-regulated genes. The Venn diagram shows 24 intersecting genes associated with brain metastasis (see Figure 1).

## Prognostic model construction

We then constructed the prognostic model based on LASSO for the 24 genes obtained from the intersection. A RiskScore formula with five genes was obtained as follows:

$$\begin{aligned}\text{RiskScore} = {} & (-0.0164) * \text{GBP2} + (-0.0589) * \text{GPR171} \\ & + (-0.0187) * \text{DIRAS3} \\ & + (-0.085) * \text{RAC2} + (-0.0379) * \text{CACNA1D}\end{aligned}$$

The analysis showed that the prognosis of low expression of the gene in this model was significantly better than that of high expression. The AUC areas of this model were 0.655, 0.663, and 0.617 at one, three, and five years, respectively (see Figure 2).

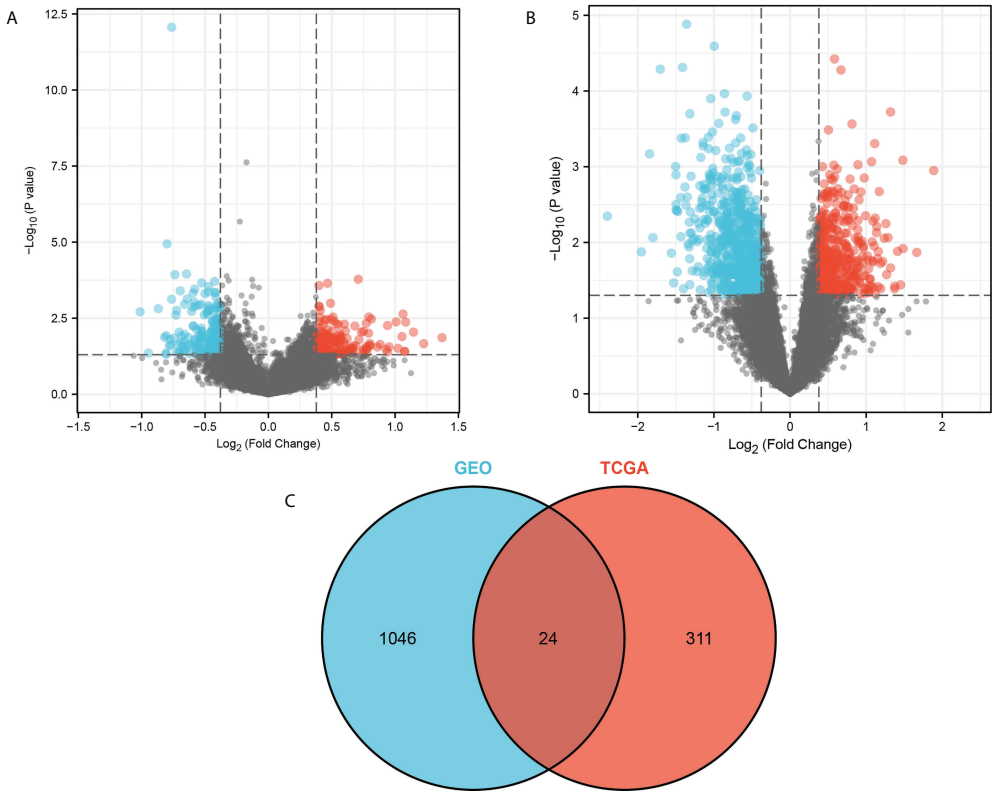



FIGURE 1
Differential gene analysis. **(A)** Differential gene analysis based on TCGA database is shown with a volcano plot. **(B)** Differential gene analysis based on the GEO dataset GSE12276 is shown with a volcano plot. **(C)** Venn diagram showing intersecting genes. The blue color in the volcano plot indicates down-regulated genes, while the red color indicates up-regulated genes.

## Expression difference analysis

We performed expression analyses of five genes (GBP2, GPR171, DIRAS3, RAC2, CACNA1D) obtained from the LASSO model to observe their expressions in primary breast cancer and metastatic breast cancer. The analysis showed that only GBP2, GPR171, DIRAS3, and RAC2 were significantly down-regulated in expression in metastatic breast cancer compared with primary breast cancer tumors (see Figure 3).

## KM survival analysis

We also performed a prognostic analysis of genes that have significant differences in expression in metastatic breast cancer. The analysis showed that GBP2 was not significantly different from the overall survival of breast cancer patients, while GPR171, DIRAS3, and RAC2 were strongly associated with the overall survival of breast cancer patients (see Figure 4).

## GSEA analysis

To further explore the biological pathways most relevant to the pathogenesis of breast cancer brain metastasis, we performed GSEA analysis on GPR171, DIRAS3, and RAC2. FDR< 0.25, NOM p-value< 0.05, and |NES|> 1 were considered to be significantly enriched. Analysis revealed that the gene sets associated with ECM-receptor interactions, complement and cohesion cascades, and adherent spot signaling were differentially enriched in the DIRAS3 high-expression phenotype. Moreover, genomes

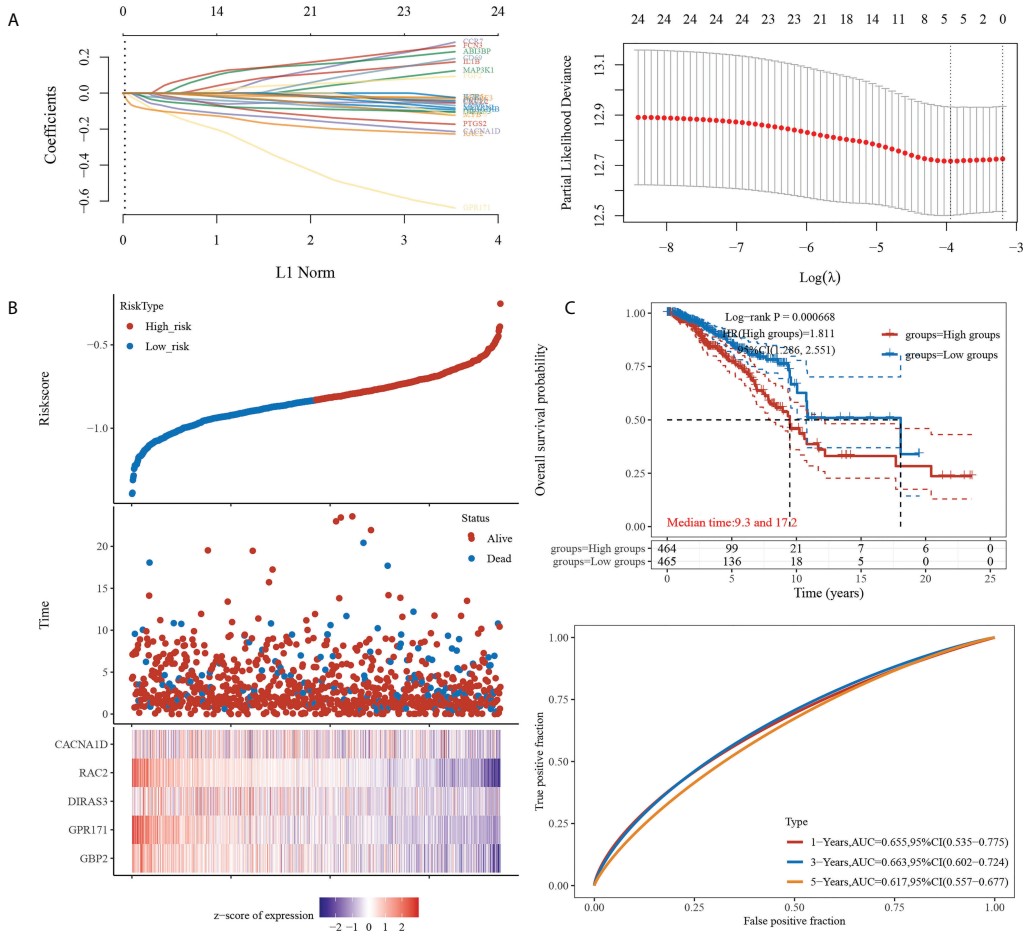



FIGURE 2
Prognostic model construction. **(A)** (Left) Lambda parameters are shown, with the horizontal axis representing the value of the independent variable lambda and the vertical axis representing the coefficient of the independent variable. (Right) Partial likelihood deviation versus log(λ). **(B)** (Upper) Scatterplot of RiskScore from low to high. (Middle) Scatterplot distribution of survival time and survival status corresponding to different sample RiskScores. (Lower) Heat map of gene expression in this model. **(C)** (Upper) Distribution of KM survival curves of this model in TCGA dataset. (Lower) ROC curves with AUC for one, three, and five years of this model.

associated with allograft rejection, autoimmune thyroid disease, and graft-versus-host disease signaling were differentially enriched in the GPR171 high-expression phenotype. Genomes associated with autoimmune thyroid disease, allograft rejection, and cell adhesion molecule signaling were differentially enriched in the RAC2 high-expression phenotype (see Figure 5).

## Gene and immune cell correlation analysis

Finally, we explored the correlation between gene expression levels and immune cells. The results of the correlation heat map showed that DIRAS3 expression levels were weakly correlated with the levels of multiple immune cells, while the expression levels of GPR171 and RAC2 were only moderately correlated with the levels of B cells and macrophages. Among them, GPR171 had the highest correlation coefficient with B cells (R=0.62) (see Figure 6).

# Discussion

Metastasis is a multistep process that requires uncontrolled tumor growth, penetration of the basement membrane, and new angiogenesis. In the bloodstream, circulating tumor cells extravasate into distant organs by penetrating the endothelium or, in the case of brain metastases, the blood–brain barrier (BBB), to form small colonies in the target organ, survive the apoptotic trail, and finally regrow at secondary sites (8). Brain metastases are a primary cause of morbidity and mortality in breast cancer patients. In autopsy studies, 15% to 35% of breast cancer patients are found to have brain metastases, and not all of these have clinical manifestations prior to death (9, 10). Brain metastases are critical to patient survival, with a median survival of approximately 15 months after diagnosis of brain metastases in breast cancer patients (11). Therefore, unraveling the molecular mechanisms of metastasis formation is vital for the formulation of potential therapeutic interventions.

GPR171 is a GPCR receptor that is closely related to the P2Y receptor, a group of GPCR receptors known to be important for immune response (12). It has been shown that GPR171 is a pro-

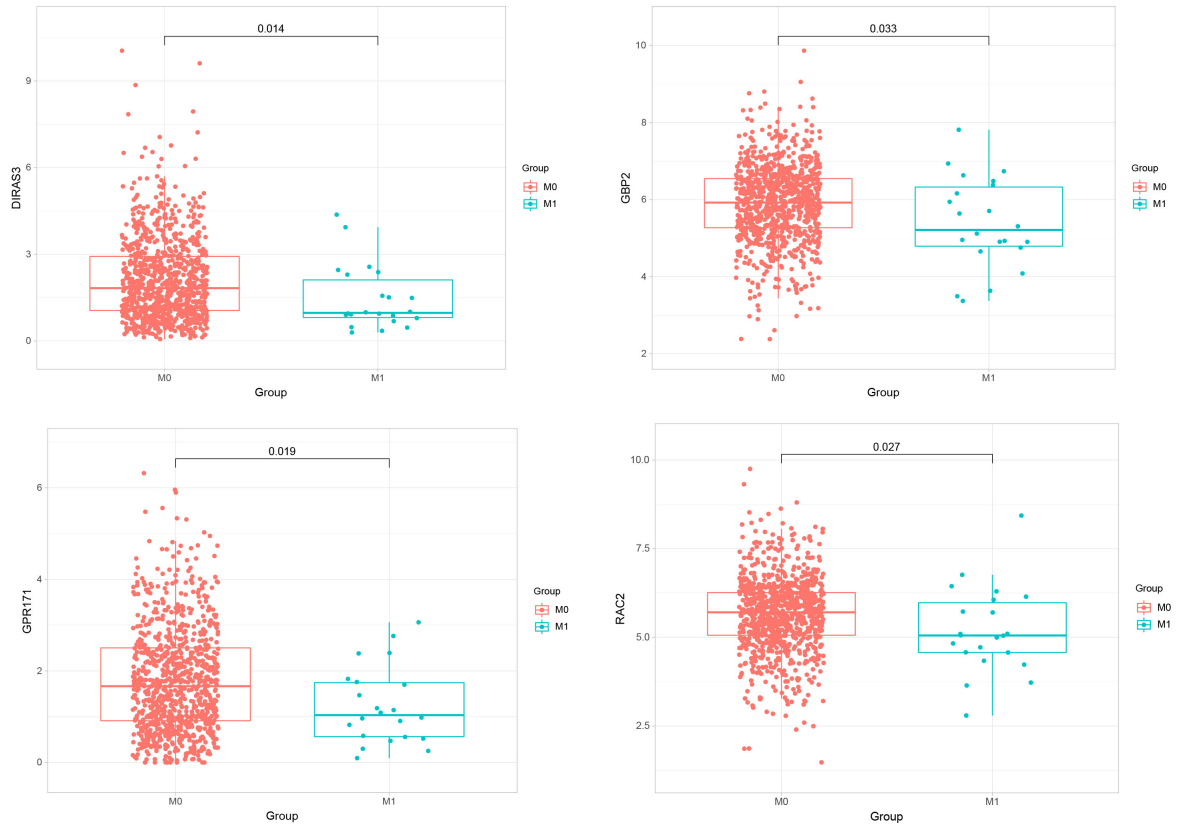



FIGURE 3
Expressions of GBP2, GPR171, DIRAS3, and RAC2 in primary breast cancer and metastatic breast cancer.

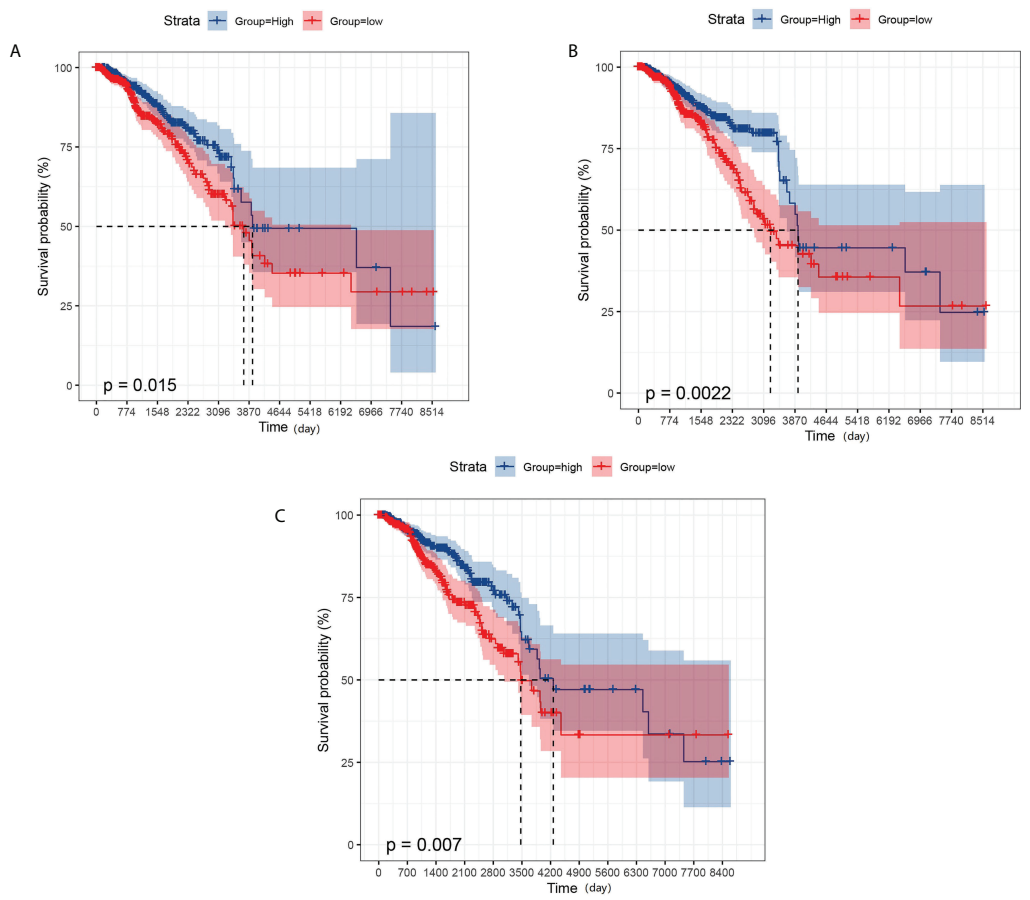



FIGURE 4
KM survival analysis. **(A)** Effect of the expression level of GPR171 on overall survival of breast cancer patients. **(B)** Effect of the expression level of RAC2 on overall survival of breast cancer patients. **(C)** Effect of the expression level of DIRAS3 on overall survival of breast cancer patients.

oncogene that induces proliferation, invasion, and migration of tumor cells (13). DIRAS3 is an oncogene involved in tumor development and autophagy, and low expression of DIRAS3 is associated with high malignancy in ovarian, breast, and prostate cancers (14). Furthermore, in ovarian and breast cancers, DIRAS3 inhibits cell migration, induces autophagy, and increases sensitivity to chemotherapy (15, 16). RAC2 is a GTPase with a molecular weight of 21 kDa; it contains the catalytic subunit of NADPH oxidase (17). In addition, RAC2 has been shown to be associated with epithelial cell polarization, a process that is closely related to changes in intercellular junctions, cytoskeletal distribution, and organelle repositioning (18). It has been reported that RAC2 may play a key role in the regulation of the actin cytoskeleton during breast cancer metastasis and that its downregulation is associated with the invasive and metastatic capacities of human cancers (19, 20). All of the foregoing could indicate that DIRAS3, GPR171, and RAC2 play important roles in tumorigenesis and tumor progression. Notably, in the present study, our results also showed that DIRAS3, GPR171, and RAC2 were less expressed in tumors that developed metastatic disease compared with primary breast cancer tumors. Moreover, the expression levels of DIRAS3, GPR171, and RAC2 were strongly correlated with the overall survival of breast cancer patients. Unfortunately, the expression levels of DIRAS3, GPR171, and RAC2 were not found to be significantly associated with the overall survival of patients with metastatic breast cancer, probably because the sample size was small and, therefore, did not show significance. In addition, GSEA analysis showed that DIRAS3, GPR171, and RAC2 are involved in signaling pathways such as ECM–receptor interaction, complement and cohesion cascades, adhesion plaques, allograft rejection, autoimmune thyroid disease, graft-versus-host disease, and cell adhesion molecules.

Tumor cells require considerable potential to multiply and transform into large tumors. The immune system usually tries to identify cancer cells and deoxyribonucleic acid-damaged cells and destroys them (21). Breast cancer may be the result of a malfunction of this useful immune defense and surveillance.

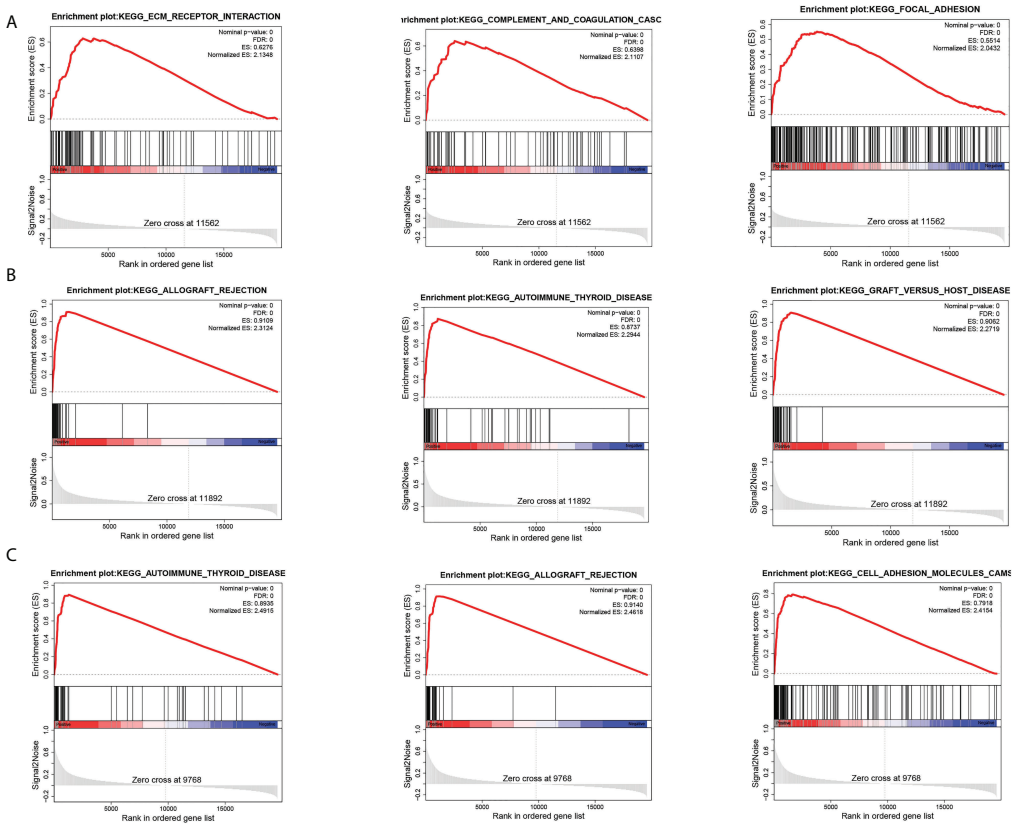



FIGURE 5
GSEA analysis. **(A)** Top three pathways enriched in DIRAS3 high-expression phenotype. **(B)** Top three pathways enriched in GPR171 high-expression phenotype. **(C)** Top three pathways enriched in RAC2 high-expression phenotype.

Therefore, in the present study, we also analyzed the correlation between the expression levels of DIRAS3, GPR171, and RAC2 and immune cells. The results showed that the correlation between the expression levels of DIRAS3 and immune cells did not show the desired results. However, the expression levels of GPR171 and RAC2 correlated more closely with B cells and macrophages. Among them, GPR171 had the highest correlation coefficient with B cells. It has been indicated that GPR171 expression is inducible in T cells and suppresses T cell-mediated immune responses through GPR171 signaling. Furthermore, the disruption of GPR171 signaling promoted T cell-mediated antitumor immunity (22). It is known that disruption of the Blood Brain Barrier by CNS tumors and changes in the composition of the extracellular matrix can allow leakage of Blood Brain Barrier at the tumor site (23). An intact brain contains almost no lymphocytes, but some studies

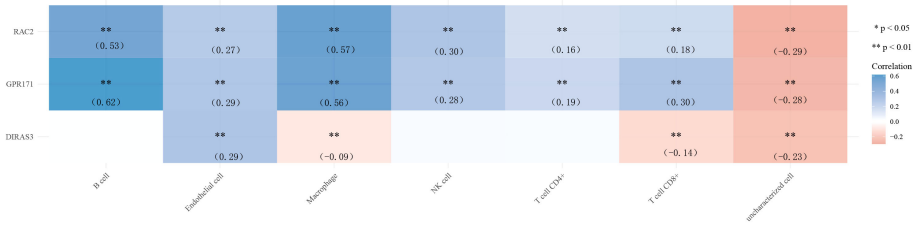



FIGURE 6
Heat map of correlation between the expression levels of DIRAS3, GPR171, and RAC2 and immune cells. * p< 0.05, ** p< 0.01.

have shown that T and B cells have been observed in the setting of brain metastases (24). Therefore, we speculate that GPR171 could promote brain metastasis of breast cancer cells by inducing B cells, but the exact mechanism remains to be investigated.

In conclusion, the expressions of DIRAS3, GPR171, and RAC2 genes associated with brain metastasis were reduced in metastatic breast cancer and were strongly associated with overall breast cancer survival. Our study also showed that the expression level of GPR171 was significantly correlated with B cells, suggesting that GPR171 could promote brain metastasis of breast cancer cells by inducing B cells. Our findings lay the foundation for understanding the molecular basis of breast cancer metastasis to distant sites, which will have the potential to contribute to future therapeutic research and provide new directions for exploring new drugs or therapies for breast cancer.

## Data availability statement

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

## Author contributions

JD: Conceptualization, Methodology, Writing - Original Draft. QC: Methodology, Formal analysis, Writing - Review & Editing. GL: Writing - Review & Editing, Visualization, Investigation. MC: Data Curation, Methodology, Writing - Original Draft. HS: Writing - Review & Editing, Formal analysis, Resources. MY: Writing - Original Draft, Funding acquisition, Resources.

## Funding

This work was supported by Ningbo Natural Science Foundation Project (2019A610302) and Zhejiang Province Medical and Health Clinical Research Application Project (2022KY1173).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Wu M, Li Q, Wang H. Identification of novel biomarkers associated with the prognosis and potential pathogenesis of breast cancer *via* integrated bioinformatics analysis. *Technol Cancer Res Treat* (2021) 20:1533033821992081. doi: 10.1177/1533033821992081

2. Wang Y, Li Y, Liu B, Song A. Identifying breast cancer subtypes associated modules and biomarkers by integrated bioinformatics analysis. *Biosci Rep* (2021) 41(1):BSR20203200. doi: 10.1042/BSR20203200

3. Wei L, Wang Y, Zhou D, Li X, Wang Z, Yao G, et al. Bioinformatics analysis on enrichment analysis of potential hub genes in breast cancer. *Transl Cancer Res* (2021) 10(5):2399. doi: 10.21037/tcr-21-749

4. Caswell-Jin JL, Plevritis SK, Tian L, Cadham CJ, Xu C, Stout NK, et al. Change in survival in metastatic breast cancer with treatment advances: Meta-analysis and systematic review. *JNCI Cancer Spectr* (2018) 2:pky062. doi: 10.1093/jncics/pky062

5. Chen W, Hoffmann AD, Liu H, Liu X. Organotropism: new insights into molecular mechanisms of breast cancer metastasis. *NPJ Precis Oncol*(2018) 2:1–12. doi: 10.1038/s41698-018-0047-0

6. Jin L, Han B, Siegel E, Cui Y, Giuliano A, Cui X. Breast cancer lung metastasis: Molecular biology and therapeutic implications. *Cancer Biol Ther* (2018) 19:858–68. doi: 10.1080/15384047.2018.1456599

7. Riggio AI, Varley KE, Welm AL. The lingering mysteries of metastatic recurrence in breast cancer. *Br J Cancer* (2021) 124:13–26. doi: 10.1038/s41416-020-01161-4

8. Conrad C, Götte M, Schlomann U, Roessler M, Pagenstecher A, Anderson P, et al. ADAM8 expression in breast cancer derived brain metastases: Functional implications on MMP-9 expression and transendothelial migration in breast cancer cells. *Int J Cancer* (2018) 142:779–91. doi: 10.1002/ijc.31090

9. Martin AM, Cagney DN, Catalano PJ, Warren LE, Bellon JR, Punglia RS, et al. Brain metastases in newly diagnosed breast cancer: a population-based study. *JAMA Oncol* (2017) 3:1069–77. doi: 10.1001/jamaoncol.2017.0001

10. Shah N, Mohammad AS, Saralkar P, Sprowls SA, Vickers SD, John D, et al. Investigational chemotherapy and novel pharmacokinetic mechanisms for the treatment of breast cancer brain metastases. *Pharmacol Res* (2018) 132:47–68. doi: 10.1016/j.phrs.2018.03.021

11. Ogiya R, Niikura N, Kumaki N, Yasojima H, Iwasa T, Kanbayashi C, et al. Comparison of immune microenvironments between primary tumors and brain metastases in patients with breast cancer. *Oncotarget* (2017) 8:103671. doi: 10.18632/oncotarget.22110

12. Rayah A, Kanellopoulos JM, Di Virgilio FJM. P2 receptors and immunity. *Microbes Infect* (2012) 14:1254–62. doi: 10.1016/j.micinf.2012.07.006

13. Dho SH, Lee K-P, Jeong D, Kim C-J, Chung K-S, Kim JY, et al. GPR171 expression enhances proliferation and metastasis of lung cancer cells. *Oncotarget* (2016) 7:7856. doi: 10.18632/oncotarget.6856

14. Qiu J, Li X, He Y, Sun D, Li W, Xin Y. Distinct subgroup of the ras family member 3 (DIRAS3) expression impairs metastasis and induces autophagy of gastric cancer cells in mice. *J Cancer Res Clin Oncol* (2018) 144:1869–86. doi: 10.1007/s00432-018-2708-3

15. Zou C-F, Jia L, Jin H, Yao M, Zhao N, Huan J, et al. Re-expression of ARHI (DIRAS3) induces autophagy in breast cancer cells and enhances the inhibitory effect of paclitaxel. *BMC Cancer* (2011) 11:1–11. doi: 10.1186/1471-2407-11-22

16. Washington M, Suh G, Orozco A, Sutton M, Yang H, Wang Y, et al. ARHI (DIRAS3)-mediated autophagy-associated cell death enhances chemosensitivity to cisplatin in ovarian cancer cell lines and xenografts. *Cell Death Dis* (2015) 6:e1836–6. doi: 10.1038/cddis.2015.208

17. Liu Y, Cheng G, Song Z, Xu T, Ruan H, Cao Q, et al. RAC2 acts as a prognostic biomarker and promotes the progression of clear cell renal cell carcinoma. *Int J Oncol* (2019) 55:645–56. doi: 10.3892/ijo.2019.4849

18. Tan S, Pei W, Huang H, Zhou G, Hu W. Additive effects of simulated microgravity and ionizing radiation in cell death, induction of ROS and expression of RAC2 in human bronchial epithelial cells. *NPJ Microgravity* (2020) 6:1–6. doi: 10.1038/s41526-020-00123-7

19. Li H, Yang L, Fu H, Yan J, Wang Y, Guo H, et al. Association between Gαi2 and ELMO1/Dock180 connects chemokine signalling with rac activation and metastasis. *Nat Commun* (2013) 4:1–12. doi: 10.1038/ncomms2680

20. Gildea JJ, Seraj MJ, Oxford G, Harding MA, Hampton GM, Moskaluk CA, et al. RhoGDI2 is an invasion and metastasis suppressor gene in human cancer. *Cancer Res* (2002) 62:6418–23.

21. Akram M, Iqbal M, Daniyal M, Khan AU. Awareness and current knowledge of breast cancer. *Biol Res* (2017) 50:1–23. doi: 10.1186/s40659-017-0140-9

22. Fujiwara Y, Torphy RJ, Sun Y, Miller EN, Ho F, Borcherding N, et al. The GPR171 pathway suppresses T cell activation and limits antitumor immunity. *Nat Commun* (2021) 12:1–14. doi: 10.1038/s41467-021-26135-9

23. Galea I, Bechmann I, Perry VH. What is immune privilege (not)? *Trends Immunol* (2007) 28:12–8. doi: 10.1016/j.it.2006.11.004

24. Berghoff AS, Lassmann H, Preusser M, Höftberger RJC. Characterization of the inflammatory response to solid cancer metastases in the human brain. *Clin Exp Metastasis* (2013) 30:69–81. doi: 10.1007/s10585-012-9510-4

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 29 September 2022
DOI 10.3389/fonc.2022.917897

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Liang Liu,
Fudan University, China
Chenyan Chang,
Clemson University, United States

*CORRESPONDENCE
Yuxiang Zhao
yuxiang_zhao@biotranstech.com
Yifu Gong
gongyifu@nbu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 11 April 2022
ACCEPTED 03 June 2022
PUBLISHED 29 September 2022

CITATION
Liu H, Zhang J, Wei C, Liu Z, Zhou W, Yang P, Gong Y and Zhao Y (2022) Prognostic signature construction of energy metabolism-related genes in pancreatic cancer.
*Front. Oncol.* 12:917897.
doi: 10.3389/fonc.2022.917897

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

# Prognostic signature construction of energy metabolism-related genes in pancreatic cancer

Hao Liu[1,2,3,4†], Jianhua Zhang[4], Chaoguang Wei[4], Zhao Liu[3], Wei Zhou[2], Pan Yang[3], Yifu Gong[1*] and Yuxiang Zhao[3,4*†]

[1]Key Laboratory of Marine Biotechnology of Zhejiang Province, School of Marine Sciences, Ningbo University, Ningbo, China, [2]College of Pharmaceutical Science, Zhejiang Chinese Medical University, Hangzhou, China, [3]United New Drug Research and Development Center, Hunan Biotrans Technology Co., LTD., Changsha, China, [4]Institute of Bioengineering, Biotrans Technology Co., LTD., Shanghai, China

Pancreatic cancer is the 7th leading cause of cancer death worldwide, and its incidence and mortality rate have been on the rise in recent years in Western developed countries. The specificity of the disease and the lack of appropriate treatments have resulted in a 5-year overall survival rate of only 9%. In this study, we conducted a study based on the TCGA database and GEO database and analyzed using the energy metabolism gene set to establish a prognostic model with the least absolute shrinkage and selection operator to identify 7-genes prognostic signature, and the gene expression was verified by Real-time PCR. The model was validated using a risk score calculation, and the OS rates of the 7 genes were analyzed using one-way Cox regression. The prognostic relationship between vesicle-associated membrane protein 2 (VAMP2) and pancreatic cancer patients was analyzed by OS and progression-free survival, and the prognosis was found to be significantly worse in the high-expression group. A Nomogram showed that VAMP2 was an independent prognostic factor in pancreatic cancer. Gene set enrichment analysis showed that VAMP2 upregulation was enriched in pathways associated with immune response and that VAMP2 downregulation was enriched in metabolism-related pathways. The association of VAMP2 with immune cell infiltration was analyzed for the enrichment results, and VAMP2 was found to be positively associated with all 6 immune cells. The results of this study suggest that VAMP2 is an independent prognostic factor associated with energy metabolism in pancreatic cancer and may be involved in the immune response.

KEYWORDS

**VAMP2, pancreatic cancer, energy metabolism, prognosis signature, consensus clustering**

# Introduction

Pancreatic cancer ranks as the 14th most common cancer worldwide and is the 7th leading cause of cancer death (1). The incidence of pancreatic cancer is mainly concentrated in developed regions and is slightly higher in men than in women (2). According to Globolcan estimates, there will be more than 495,000 new cases and 466,000 deaths in 2020 (3), with a 5-year survival rate of only 9%. The low survival rate is mainly due to the fact that patients often present with symptoms at an advanced stage and to the lack of appropriate diagnostic tools and treatment measures (4). Patients with pancreatic cancer are usually classified as resectable, borderline resectable, locally advanced, or metastatic according to the degree of disease progression, with surgical resection being the only curative option (5). However, the vast majority of patients are diagnosed with inoperable advanced or metastatic disease, which to some extent reduces the prognostic survival time of pancreatic cancer patients (6). The development of pancreatic cancer is commonly associated with diabetes, as well as obesity, chronic pancreatitis, alcohol abuse, and genetic susceptibility (1).

The growth of cancer cells requires a large supply of energy, and to meet the demands of rapid growth, cancer cells reprogram their energy metabolism (7). This phenomenon is called "energy metabolic reprogramming" and is associated with the malignant biological behavior of pancreatic cancer (8). There are various ways to reprogram energy metabolism. Warburg suggested that cancer cells ferment glucose through glycolysis to obtain energy for growth (9). It is also possible to provide biomolecules for cell replication through the pentose phosphate and serine pathways, as well as using glutamine and lipids to promote their own proliferation (10). In the metastatic process of pancreatic cancer cells, metabolic reprogramming provides energy through aerobic glycolysis and oxidative phosphorylation, among other mechanisms (11). Cancers are heterogeneous diseases with complex and diverse metabolic patterns and the ability to improvise. Metabolic alterations contribute to the regulation of apoptosis and angiogenesis and confer a resistance phenotype (12). This resistance is reflected in pancreatic cancer in terms of drug resistance, which leads to poor treatment response (13). Metabolic reprogramming in pancreatic cancer is also associated with chemotherapy, radiotherapy, and immunotherapy, which can lead to poor prognosis (14).

The aim of this study was to identify prognostic features associated with energy metabolism in pancreatic cancer and to provide suggestions on the direction of pancreatic cancer treatment. We established 7-genes prognostic signature using the least absolute shrinkage and selection operator (LASSO) model and identified vesicle-associated membrane protein 2 (VAMP2) as a new energy metabolism-related prognostic biomarker for pancreatic cancer.

# Data and methods

## Data sources

Clinical information and the gene expression profiles of pancreatic cancer were obtained from the TCGA database (https://portal.gdc.cancer.gov/), containing 178 tumor samples and 4 normal samples. Since there were too few paraneoplastic samples in TCGA, 328 paraneoplastic samples were obtained from the GTEx database for subsequent analysis. Two energy metabolism-related gene sets containing 156 genes were downloaded from MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/). GSE57495, GSE11838, GSE15932 and GSE62165 were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), containing clinical information and genes expression profiles. Immunoscore data were obtained from the TIMER (http://timer.cistrome.org/).

## Consensus clustering

Consensus clustering was performed based on the TCGA-PAAD dataset and energy metabolism-related genes to compare the clinical information of different subgroups. Consistency analysis was performed using the R package ConsensusClusterPlus with a threshold of 2 clusters and 100 repetitions of 80% of the samples drawn. The clustering heat map was drawn using the R package pheatmap.

## Selection of prognosis-related energy metabolism genes

Genes associated with overall survival (OS) in pancreatic cancer patients were analyzed using a univariate COX model, and a forest plot was drawn using the R package forestplot to represent the top 20 most significant genes. The genes with significant prognosis were intersected with energy metabolism genes and visualized by a Venn plot.

## Construction of energy metabolism-related prognostic gene signature

Energy metabolism-related prognostic gene signature was constructed based on associations of gene expression levels with energy metabolism-related genes. The LASSO regression algorithm was used for gene signature selection, 10-fold cross-validation was applied, and a risk score model was constructed. Grouping was performed based on the best risk score cut-off value, and the expression and survival differences between the high- and low-risk groups were analyzed. The relationship

between risk scores and clinical traits was analyzed by the univariate and multivariable Cox model. The TCGA dataset was used as the training set and GSE57495 was used as the validation set; P<0.05 was considered statistically significant.

## Gene expression validation and prognostic correlation analysis

The univariate and multivariable Cox model were used to analyze the relationship between clinical traits and OS in pancreatic cancer patients. R package ggplot2 was used to perform the box plot of VAMP2 expression under different clinical traits groups. The OS and progression-free survival (PFS) of prognostic traits were analyzed using the R package Survival, and the receiver operating characteristic (ROC) curve was plotted by timeROC. Human epidermoid tumor pancreatic ductal tumor cell line PANC-1, human pancreatic adenocarcinoma cell line BxPC-3 and human normal pancreatic ductal cell line hTERT-HPNE were used to verify the gene expression of the 7-genes prognostic signature by RT-PCR. The primer sequences were listed in Tab S1.

## Gene set enrichment analysis

GSEA was used to analyze functional differences between groups when considering gene expression (15). The KEGG database was used as the functional gene set for GSEA, with the cut-off threshold |NES|>1, NOM p-val<0.05, and FDR q-val<0.25.

## Immunocorrelation analysis

The relationship between the expression of VAMP2 and immune cells infiltration level was analyzed using Spearman, and immune cell correlation plots were performed with the R package ggstatsplot.

# Results

## Consensus cluster analysis of energy metabolism-related genes in pancreatic cancer

The pancreatic cancer patients in TCGA were grouped based on the expression of energy metabolism-related genes and were divided into 2 subgroups (Figures 1A–C). The expression of energy metabolism-related genes in the 2 subgroups is shown in Figure 1D. A comparison of clinical information between the 2 subgroups revealed significant differences in performance in age, T stage, and tumor stage (Table 1).

## Construction of energy metabolism-related gene signature

A univariate Cox analysis showed a total of 3138 genes associated with prognostic OS in pancreatic cancer (Figure 2A). These genes were intersected with energy metabolism-related genes to obtain 27 signature genes (Figure 2B). The 27 characteristic genes were used to construct the LASSO model, and 7 genes were obtained when the minimum characteristic coefficient (λ)=0.0918 (Figures 3A, B). The data risk scores in the set were calculated using the following risk score formula:

$$\begin{aligned}\text{RiskScore} = {} & (-0.0133) * \text{ACACB} + (0.0839) * \text{GNA15} \\ & + (-0.3206) * \text{GNB3} + (-0.0897) * \text{GNG7} \\ & + (0.0601) * \text{IQGAP1} + (-0.0145) * \text{STXBP1} \\ & + (-0.036) * \text{VAMP2}\end{aligned}$$

The best cut-off value for the training set was −0.63. KM curves showed significant prognostic differences between the

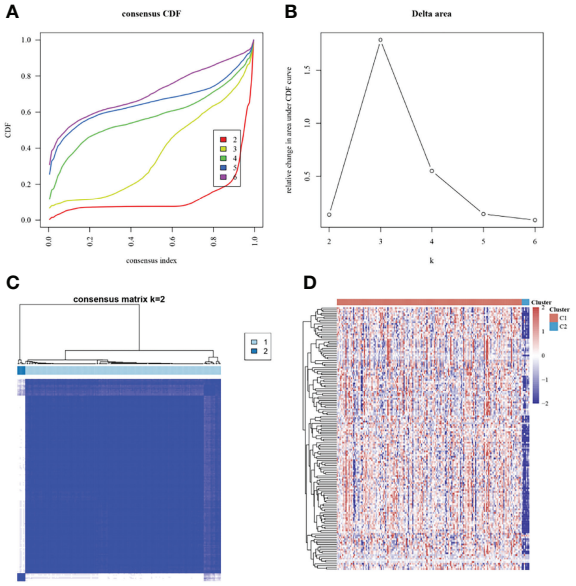



FIGURE 1
Consensus clustering analysis based on energy metabolism-related genes. **(A)**: Cumulative distribution function (CDF) curve; **(B)**: CDF delta area curve; **(C)**: Consensus matrix when K=2; **(D)**: Heat map of energy metabolism-related genes in different subgroups, with red indicating high expression, and blue indicating low expression.

TABLE 1 Comparison of clinical information.

| Types | Charar | C1 | C2 | P value |
|---|---|---|---|---|
| Status | Alive | 82 | 4 | |
| | Dead | 89 | 3 | 0.927 |
| Age | Mean (SD) | 64.5 (10.9) | 71.3 (5.8) | |
| | Median [MIN, MAX] | 65 [35,88] | 72 [63,79] | 0.02 |
| Gender | FEMALE | 77 | 3 | |
| | MALE | 94 | 4 | 1 |
| Race | ASIAN | 10 | 1 | |
| | BLACK | 6 | | |
| | WHITE | 151 | 6 | 0.948 |
| pT_stage | T1 | 7 | | |
| | T2 | 23 | 1 | |
| | T3 | 137 | 5 | |
| | T4 | 2 | 1 | |
| | TX | 1 | | 0.037 |
| pN_stage | N0 | 50 | | |
| | N1 | 112 | 7 | |
| | N1b | 4 | | |
| | NX | 4 | | |
| pM_stage | M0 | 74 | 5 | |
| | M1 | 5 | | |
| | MX | 92 | 2 | 0.313 |
| pTNM_stage | I | 1 | | |
| | IA | 5 | | |
| | IB | 15 | | |
| | IIA | 28 | | |
| | IIB | 112 | 6 | |
| | III | 2 | 1 | |
| | IV | 5 | | 0.414 |
| Grade | G1 | 29 | 2 | |
| | G2 | 94 | 1 | |
| | G3 | 46 | 2 | |
| | G4 | 1 | 1 | |
| | GX | 1 | 1 | 0 |
| new_tumor_event_type | Metastasis | 54 | 3 | |
| | Metastasis:Recurrence | 2 | | |
| | Primary | 2 | | |
| | Recurrence | 19 | 1 | 1 |
| Smoking | Non-smoking | 63 | 2 | |
| | Smoking | 75 | 4 | 0.861 |

high- and low-risk groups in the training set (Figure 3C), and ROC analysis indicated AUCs of 0.93, 0.69, and 0.79 at 1, 3, and 5 years, respectively (Figure 3D). The model was validated using GSE57495 and showed significantly worse prognostic survival in the high-risk group when grouped using an optimal cut-off value of −1.92 (Figure 3E) and ROC showing AUCs of 0.65 and 0.69 at 1 and 3 years, respectively (Figure 3F).

## Validation of the 7 genes energy metabolism-related prognostic signature model

To validate the robustness of the LASSO model, P-values, HR values, and 95% CI of each clinical trait and risk score were analyzed by univariate and multifactorial Cox regression in the

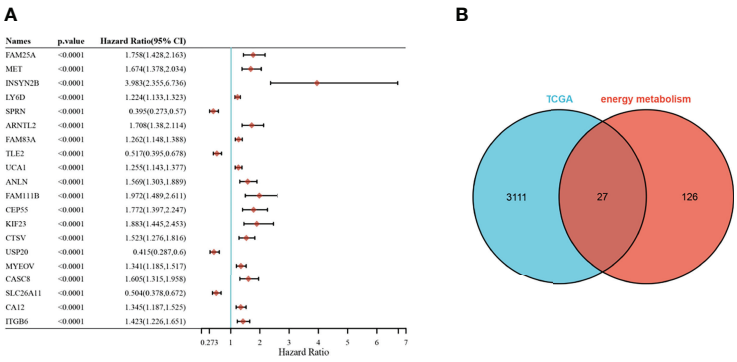



FIGURE 2
Selection of prognosis-related energy metabolism genes. **(A)**: Forest plot of the top 20 significant prognostic genes; **(B)**: Venn plot.

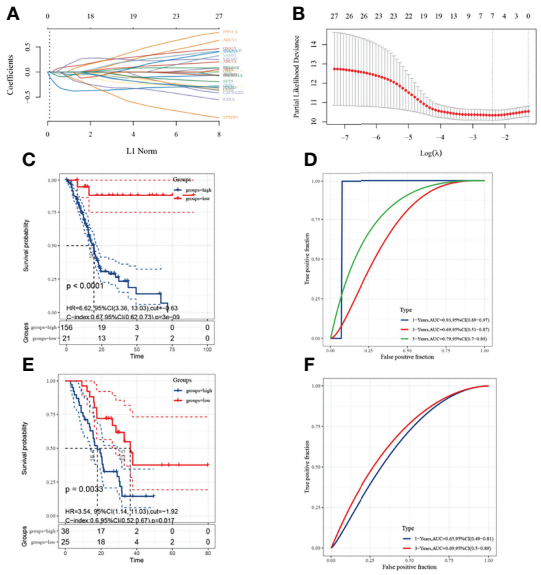



FIGURE 3
Construction of prognostic 7-gene signature with LASSO model. **(A)**: The coefficients of selected features are shown by lambda parameter. The horizontal axis represents the value of the independent variable lambda, and the vertical axis indicates the coefficient of the independent variable; **(B)**: Partial likelihood bias and log; **(C, D)**: KM survival curve and ROC curve based on the 7-gene signature in the training set; **(E, F)**: KM survival curve and ROC curve based on the 7-gene signature in the validation set. Due to the microarray samples, only the area under the curve (AUC) for 1 and 3 years is available in the ROC curves.

TCGA dataset, and the risk score was found to be an independent prognostic predictor for pancreatic cancer patients (Figures 4A, B). The high and low expression groups were classified according to the median risk score, and the differences in the high and low expression of risk score among different clinical traits were observed. The results showed that the high risk score group had high expression, which was associated with the degree of tumor differentiation (Figures 4C–J).

## Expression of 7 energy metabolism-related gene traits

A one-way Cox regression analysis of the 7 energy metabolism-related gene features revealed that all were

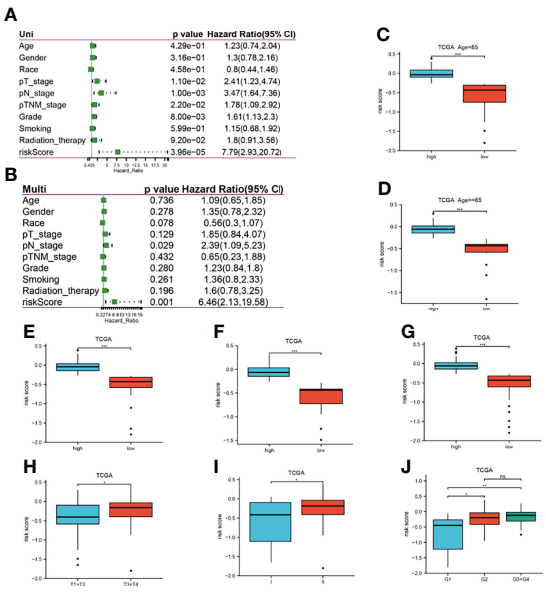



FIGURE 4
Correlation between risk scores and clinical traits in TCGA. **(A, B)**: P-value, HR, 95% CI for univariate and multifactorial Cox analysis of clinical traits and risk scores; **(C–J)**: Differences in expression of high and low risk scores in different clinical traits, respectively, age<65, age≥65, male, female, Caucasian, T stage, pTNM stage, and tumor grade. *P<0.05, **P<0.01, ***P<0.001, ns, no significant.

associated with prognosis in pancreatic cancer patients (Figure 5A). Subsequent analysis showed that all 7 trait genes were significantly differentially expressed in cancer versus paracancer (Figures 5B, C). However, the expression of the 7 energy-related signature genes differed across clinical traits, with VAMP2 expression being more prominent and significant for almost every clinical trait (Figures 5D–H). Expression analysis was performed for VAMP2, and differential expression was found in cancer versus paracancer and for each clinical trait (Figure 6). The expression levels of 7 genes in the cell lines were shown in Figure S1, which was consistent with the results in TCGA database.

## VAMP2 can be used as an independent predictor of pancreatic cancer

The relationships between VAMP2 and both prognostic OS and PFS of pancreatic cancer patients were analyzed by grouping with median expression values. OS and PFS survival were significantly higher in the high expression group than in the low expression group (Figures 7A, C), and had better predictive ability (Figures 7B, D). Figures 7E, F shows that the higher the expression of VAMP2, the higher the tumor grade and the worse the prognosis for survival. Combined with the analysis of VAMP2 expression and other clinical traits, VAMP2 was found to be an independent prognostic factor for pancreatic cancer patients (Figures 7G, H). A nomogram dependent on OS-independent prognostic parameters in pancreatic cancer patients was also constructed (Figures 7I, J).

## Involvement of VAMP2 in immune response

Based on the median expression value of VAMP2 for grouping, the functional pathways of the high- and low-expression groups were enriched using GSEA. The enrichment is shown in Figure 8. All pathways were ranked according to P-values. Table 2 shows the top 10 significantly enriched pathways. The results show that when VAMP2 was highly expressed, it was mainly enriched in pathways related to immune response, and when VAMP2 was lowly expressed, it was mainly enriched in biometabolic pathways.

## VAMP2 is involved in immune infiltration in pancreatic cancer

Since VAMP2, when highly expressed, was significantly enriched in immune-related pathways in pancreatic cancer, we performed an immune correlation analysis of VAMP2. As Figure 9A shows, VAMP2 in pancreatic cancer was positively

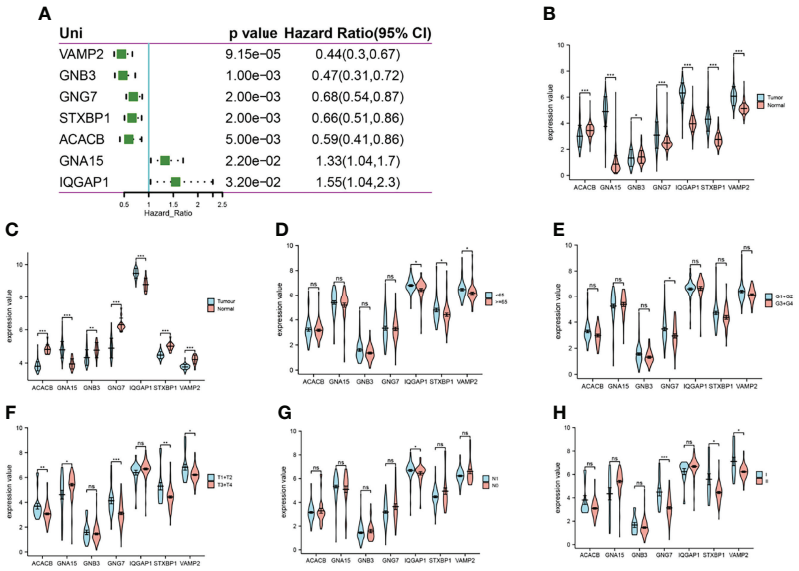



FIGURE 5
Expression of 7 genes. **(A)**: one-way Cox regression analysis of the prognosis of 7 energy metabolism-related genes in pancreatic cancer; **(B)**: expression differences between cancer and paracancer in TCGA; **(C)**: expression differences between cancer and paracancer in GSE62165; **(D–H)**: expression differences of genes in different age stages, grade stage, T stage, N stage, and pTNM stage in TCGA, respectively. *p<0.05, **p<0.01, ***p<0.001 compared with the control group. ns, no significant.

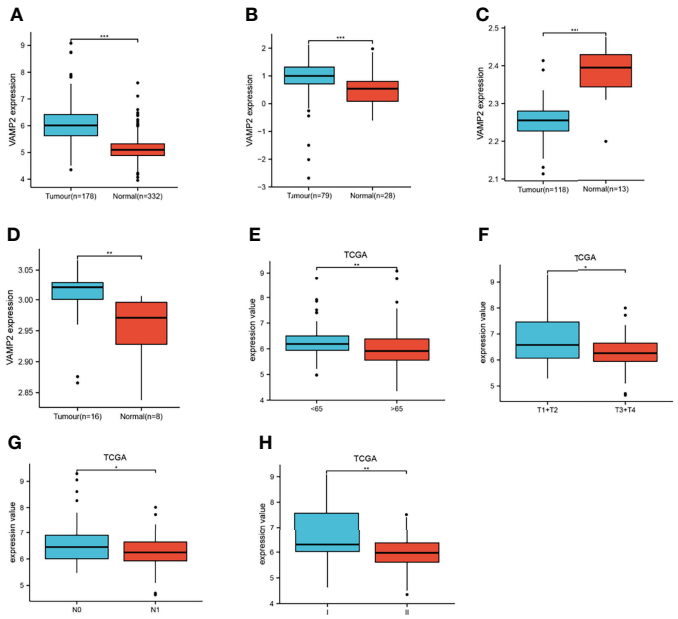



FIGURE 6
Expression of VAMP2. **(A)**: TCGA dataset as the training set, **(B–D)**: GSE11838, GSE62165, and GSE15932 as the validation set for analyzing the expression of VAMP2 in cancer versus paracancer and observing the expression differences between age, T stage, N stage, and pTNM stage **(E–H)**. *P<0.05, **P<0.01, ***P<0.001.

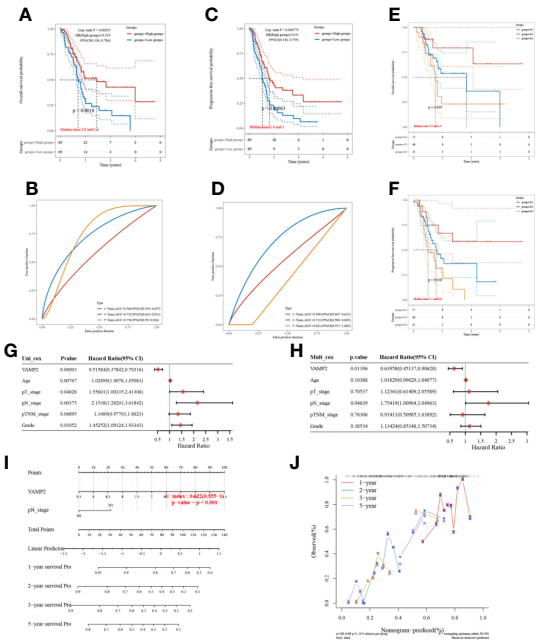



FIGURE 7
Prognosis of VAMP2. **(A, B)**: OS; **(C, D)**: PFSKM survival curves and ROC prediction model; **(E, F)**: KM curves for different tumor grades with OS and PFS, respectively; **(G, H)**: single-factor Cox regression analysis and multi-factor Cox regression analysis based on OS; **(I)**: column line graph; **(J)**: calibration curve.

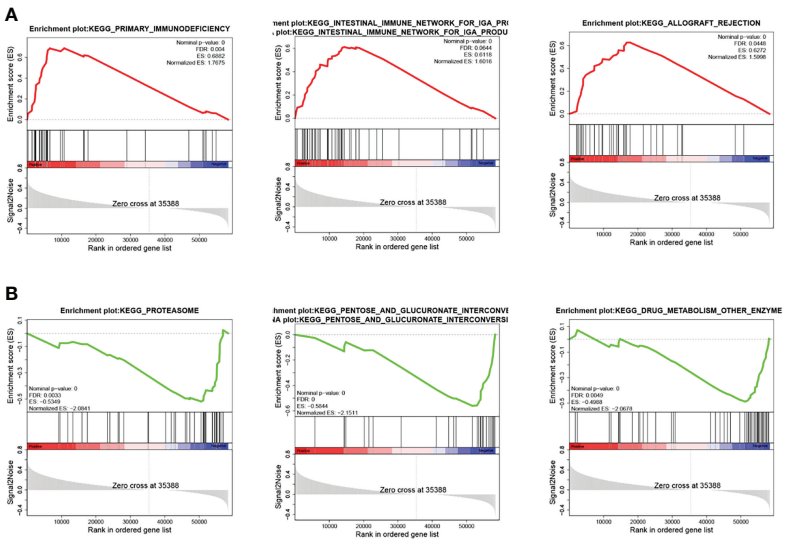



FIGURE 8
Pathway enrichment of VAMP2. **(A, B)**: KEGG pathway enrichment in high- and low-expression groups, respectively; hallmark pathway enrichment in high- and low-expression groups, respectively.

TABLE 2 KEGG pathway enrichment of top10.

| | NAME | ES | NES | NOM p-val | FDR q-val |
|---|---|---|---|---|---|
| High_exp | primary immunodeficiency | 0.688 | 1.767 | 0 | 0.004 |
| | intestinal immune network for IGA production | 0.612 | 1.602 | 0 | 0.064 |
| | allograft rejection | 0.627 | 1.600 | 0 | 0.045 |
| | chemokine signaling pathway | 0.552 | 1.569 | 0 | 0.043 |
| | neuroactive ligand receptor interaction | 0.540 | 1.553 | 0 | 0.046 |
| | T cell receptor signaling pathway | 0.552 | 1.540 | 0 | 0.048 |
| | cell adhesion molecules cams | 0.518 | 1.459 | 0 | 0.085 |
| | cytokine cytokine receptor interaction | 0.493 | 1.417 | 0 | 0.104 |
| | calcium signaling pathway | 0.499 | 1.413 | 0 | 0.103 |
| | mapk signaling pathway | 0.455 | 1.307 | 0 | 0.232 |
| Low_exp | pentose and glucuronate interconversions | -0.584 | -2.151 | 0 | 0.000 |
| | proteasome | -0.535 | -2.084 | 0 | 0.003 |
| | drug metabolism other enzymes | -0.499 | -2.068 | 0 | 0.005 |
| | ascorbate and aldarate metabolism | -0.544 | -2.021 | 0 | 0.004 |
| | O glycan biosynthesis | -0.557 | -1.989 | 0 | 0.005 |
| | metabolism of xenobiotics by cytochrome p450 | -0.440 | -1.945 | 0 | 0.006 |
| | retinol metabolism | -0.419 | -1.863 | 0 | 0.011 |
| | linoleic acid metabolism | -0.498 | -1.826 | 0 | 0.012 |
| | starch and sucrose metabolism | -0.419 | -1.782 | 0 | 0.015 |
| | pentose phosphate pathway | -0.484 | -1.766 | 0 | 0.015 |

correlated with all 6 types of immune infiltrating cells in TIMER. The strongest correlation was with macrophages. The samples were separated into high and low VAMP2 expression groups, and 5 immune cell scores were found to correlate with VAMP2 expression (Figure 9B). The heat map demonstrates the expression trends of different immune cell scores in different samples (Figure 9C), and Figure 9D shows the highest percentage abundance of myeloid dendritic cells.

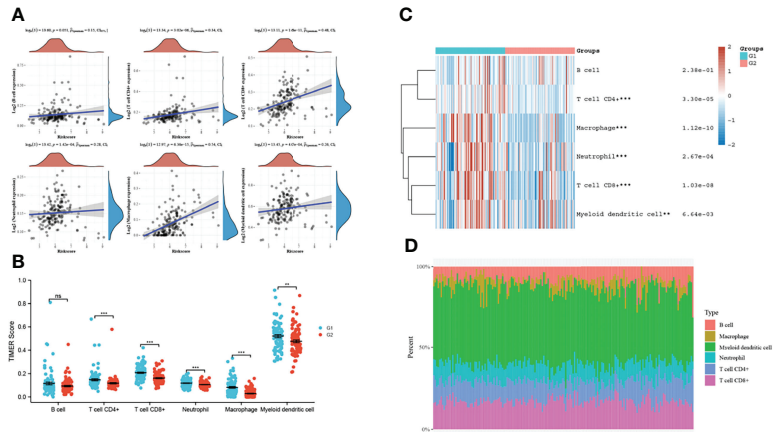



FIGURE 9
VAMP2 and immune correlation. **(A)**: relationship between VAMP2 and TIMER immune infiltrating cells; **(B)**: the relationship between VAMP2 and each immune cell score; **(C)**: heat map of immune cell scores; **(D)**: percentage abundance of tumor-infiltrating immune cells in each sample. **p<0.01, ***p<0.001 compared with the two groups. ns, no significant.

## Discussion

The current modality that effectively prolongs the prognostic survival of patients with pancreatic cancer is surgical resection plus adjuvant chemotherapy (16). However, the vast majority of patients are not suitable for surgical resection and have a high recurrence rate, resulting in a median survival rate of only 2–2.5 years (17). As an aggressive cancer, pancreatic cancer has a high metabolism, which means that an adequate energy supply is required to meet the growth of cancer cells (18). Energy metabolic reprogramming, which maintains the energy balance during cancer cell growth, proliferation, and migration, is an emerging hallmark of cancer (19). Pancreatic cancer relies mainly on glutamine to maintain cell proliferation and survival. It also uses the glycolytic pathway to metabolize glucose, thereby producing ATP (20). It has been reported that energy metabolism can lead to the expression or silencing of specific oncogenes, resulting in abnormal cell proliferation, cycle arrest, and cellular senescence (21). LASSO prognostic modeling of energy metabolism-related genes revealed 7 prognostic features most associated with OS in pancreatic cancer patients. The risk scores calculated using the 7 prognostic features were determined as independent prognostic factors for pancreatic cancer patients and were validated using the GSE57495 data. These 7 prognostic features are GNB3, which has been shown to affect OS in pancreatic cancer (22); GNG7, which can be used as a therapeutic target in pancreatic cancer (23); IQGAP1, the overexpression of which promotes pancreatic cancer progression (24), and 4 other genes (ACACB GNA15, STXBP1, VAMP2) that have not yet been reported in pancreatic cancer.

VAMP2 is a protein-encoding gene that belongs to the family of synaptic vesicle proteins (VAMPs) responsible for intracellular transport and extracellular secretion of vesicles (25). VAMP2 is an abundant synaptic vesicle protein that is closely associated with cancer cell adhesion, survival, and migration. Reduced expression of VAMP2 can lead to useless protein degradation and abnormal patterns of unwanted protein degradation (26). Recently, it has been shown that VAMP2 is significantly expressed in bladder cancer and increases in a stage-dependent manner according to tumor stage (27). VAMP2 acts as a downstream target and plays a pro-tumorigenic role in liver cancer (25). It also affects ovarian cancer prognosis and tumor progression (28). In addition, VAMP2 can act as a fusion gene and play an oncogenic role in non-small cell lung adenocarcinoma (29). In the present study, VAMP2 was screened as an energy metabolism-related feature, and GSEA showed that down-regulated VAMP2 was mainly enriched in glucose metabolism-related pathways, such as the pentose and glucuronide interconversion pathway and the pentose phosphate pathway (PPP). The PPP is a major regulator of cellular redox homeostasis and biosynthesis and is an important component of glucose metabolism (30). It also supports the glycolytic process in cancer cells, helping to meet their anabolic demands and counteract oxidative stress (31). PPP flux plays a role in promoting cancer cell survival and proliferation and is associated with the progression of hepatocellular carcinoma, lung cancer, and breast cancer (32). This suggests that VAMP2 downregulation may also affect pancreatic cancer progression by regulating the PPP.

GSEA showed that upregulated VAMP2 was mainly enriched in immune response-related pathways, such as chemokine signaling pathways and cytokine–cytokine receptor interactions. Based on this result, we analyzed VAMP2 along with immune cells and immune scores and found that VAMP2

was positively correlated with all 6 immune cells in TIMER and showed differential expression in multiple immune cell scores. The tumor microenvironment (TME) has multiple components, among which immune cells and cytokines are important and inextricably linked to tumor progression (33). Cytokines are molecular messengers of innate and adaptive immunity that allow immune cells to communicate in a paracrine or autocrine manner (34). Cytokines inhibit tumor cell growth by suppressing proliferation, promoting apoptosis, or stimulating the toxic activity of immune cells against tumor cells (35). Pro-inflammatory cytokines can promote cancer immunotherapy, acting at each stage of the cancer immune cycle (36). It has been shown that VAMP2 is the main type of VAMP that is functionally involved in antibody secretion (37). VAMP2 is a key protein in the SNARE complex that mediates the release of neurotransmitters from synaptic vesicles by neurons (38). Complexes of VAMP2 can lead to inflammatory pain in the dorsal horn of the spinal cord, which implies that VAMP2 may have an inflammatory role (39). The GSEA results suggest that it may be possible to mediate the immune response of immune cells by up-regulating the expression of VAMP2 to produce a therapeutic effect on pancreatic cancer.

In this study, we used bioinformatics to identify prognostic genes associated with energy metabolism in pancreatic cancer. We then constructed prognostic models to identify signature genes and validated them using an external validation set. The results of an expression analysis of different clinical traits led us to focus on VAMP2, and a prognostic analysis confirmed that VAMP2 is an independent prognostic factor in pancreatic cancer. GSEA demonstrated the VAMP2-enriched KEGG pathway and preliminary analysis of the association between VAMP2 and immune response. In conclusion, we believe that this study will provide new knowledge for the precise treatment of pancreatic cancer and provide a new strategy for predicting the survival of pancreatic cancer patients based on the expression of energy metabolism-related genes.

## 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

YZ and YG designed this study. JF, JZ, CW and ZL performed the analysis. HL and YZ wrote the manuscript. WZ provided constructive suggestions for the structure of the manuscript and the visualization of the analysis. All authors contributed to the article and approved the submitted version.

## Funding

This study was supported by Opening Project of Zhejiang Provincial Preponderant and Characteristic Subject of Key University (Traditional Chinese Pharmacology), Zhejiang Chinese Medical University (No. ZYAOXYB2019002) and the K.C.Wong Magna Fund in Ningbo University.

## Conflict of interest

HL, YZ, JZ, CW, ZL and PY are employed by Biotrans Technology Co., LTD.

The remaining 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.2022.917897/full#supplementary-material

## References

1. McGuigan A, Kelly P, Turkington RC, Jones C, Coleman HG, McCain RS. Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. *World J Gastroenterol* (2018) 24(43):4846. doi: 10.3748/wjg.v24.i43.4846

2. Khalaf N, El-Serag HB, Abrams HR, Thrift AP. Burden of pancreatic cancer: From epidemiology to practice. *Clin Gastroenterol Hepatol* (2021) 19(5):876–84. doi: 10.1016/j.cgh.2020.02.054

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

4. Rawla P, Sunkara T, Gaduputi V. Epidemiology of pancreatic cancer: Global trends, etiology and risk factors. *World J Oncol* (2019) 10(1):10. doi: 10.14740/wjon1166

5. Mizrahi JD, Surana R, Valle JW, Shroff RT. Pancreatic cancer. *Lancet* (2020) 395(10242):2008–20. doi: 10.1016/S0140-6736(20)30974-0

6. Tempero MA. NCCN guidelines updates: Pancreatic cancer. *J Natl Compr Cancer Netw.* (2019) 17(5.5):603–5. doi: 10.6004/jnccn.2019.5007

7. Zhu X, Chen H-H, Gao C-Y, Zhang X-X, Jiang J-X, Zhang Y, et al. Energy metabolism in cancer stem cells. *World J Stem Cells* (2020) 12(6):448. doi: 10.4252/wjsc.v12.i6.448

8. Xu R, Yang J, Ren B, Wang H, Yang G, Chen Y, et al. Reprogramming of amino acid metabolism in pancreatic cancer: Recent advances and therapeutic strategies. *Front Oncol* (2020) 10. doi: 10.3389/fonc.2020.572722

9. House SW, Warburg O, Burk D, Schade AL. On respiratory impairment in cancer cells. *Science* (1956) 124(3215):267–72.

10. Lin-Chong S, Ping G. Regulation of cancer metabolic reprogramming on immune microenvironment. *Prog Biochem biophys.* (2017) 44(8):688–96.

11. Yang J, Ren B, Yang G, Wang H, Chen G, You L, et al. The enhancement of glycolysis regulates pancreatic cancer metastasis. *Cell Mol Life Sci* (2020) 77 (2):305–21. doi: 10.1007/s00018-019-03278-z

12. Grasso C, Jansen G, Giovannetti E. Drug resistance in pancreatic cancer: Impact of altered energy metabolism. *Crit Rev oncol./hematol.* (2017) 114:139–52. doi: 10.1016/j.critrevonc.2017.03.026

13. Wang C, Liu B, Xu X, Zhuang B, Li H, Yin J, et al. Toward targeted therapy in chemotherapy-resistant pancreatic cancer with a smart triptolide nanomedicine. *Oncotarget* (2016) 7(7):8360. doi: 10.18632/oncotarget.7073

14. Qin C, Yang G, Yang J, Ren B, Wang H, Chen G, et al. Metabolism of pancreatic cancer: Paving the way to better anticancer strategies. *Mol Cancer* (2020) 19(1):1–19. doi: 10.1186/s12943-020-01169-7

15. Zhu N, Hou J, Ma G, Guo S, Zhao C, Chen B. Co-Expression network analysis identifies a gene signature as a predictive biomarker for energy metabolism in osteosarcoma. *Cancer Cell Int* (2020) 20(1):1–16. doi: 10.1186/s12935-020-01352-2

16. Strobel O, Neoptolemos J, Jäger D, Büchler MW. Optimizing the outcomes of pancreatic cancer surgery. *Nat Rev Clin Oncol* (2019) 16(1):11–26. doi: 10.1038/s41571-018-0112-1

17. Riquelme E, Zhang Y, Zhang L, Montiel M, Zoltan M, Dong W, et al. Tumor microbiome diversity and composition influence pancreatic cancer outcomes. *Cell* (2019) 178(4):795–806.e712. doi: 10.1016/j.cell.2019.07.008

18. Biancur DE, Kimmelman AC. The plasticity of pancreatic cancer metabolism in tumor progression and therapeutic resistance. *Biochim Biophys Acta (BBA)-Rev. Cancer* (2018) 1870(1):67–75. doi: 10.1016/j.bbcan.2018.04.011

19. Fumarola C, Petronini PG, Alfieri R. Impairing energy metabolism in solid tumors through agents targeting oncogenic signaling pathways. *Biochem Pharmacol* (2018) 151:114–25. doi: 10.1016/j.bcp.2018.03.006

20. Roy M, Finley SD. Computational model predicts the effects of targeting cellular metabolism in pancreatic cancer. *Front Physiol* (2017) 8:217. doi: 10.3389/fphys.2017.00217

21. Valle-Mendiola A, Soto-Cruz I. Energy metabolism in cancer: The roles of STAT3 and STAT5 in the regulation of metabolism-related genes. *Cancers* (2020) 12(1):124. doi: 10.3390/cancers12010124

22. Ye Y, Chen Z, Shen Y, Qin Y, Wang H. Development and validation of a four-lipid metabolism gene signature for diagnosis of pancreatic cancer. *FEBS Open Bio.* (2021) 11(11):3153–70. doi: 10.1002/2211-5463.13074

23. Karavadi B, Nisha H. Homology modeling and docking studies to identify the targets in pancreatic cancer. *Res J Pharm Technol* (2017) 10(7):2032–40.

24. Zhang B, Cheng X, Zhan S, Jin X, Liu T. MIB1 upregulates IQGAP1 and promotes pancreatic cancer progression by inducing ST7 degradation. *Mol Oncol* (2021) 15(11):3062–75. doi: 10.1002/1878-0261.12955/v2/response1

25. Wang G, Fang X, Han M, Wang X, Huang Q. MicroRNA-493-5p promotes apoptosis and suppresses proliferation and invasion in liver cancer cells by targeting VAMP2. *Int J Mol Med* (2018) 41(3):1740–8. doi: 10.3892/ijmm.2018.3358

26. Truckenbrodt S, Viplav A, Jähne S, Vogts A, Denker A, Wildhagen H, et al. Newly produced synaptic vesicle proteins are preferentially used in synaptic transmission. *EMBO J* (2018) 37(15):e98044. doi: 10.15252/embj.201798044

27. Raja SA, Abbas S, Shah STA, Tariq A, Bibi N, Yousuf A, et al. Increased expression levels of syntaxin 1a and synaptobrevin 2/Vesicle-associated membrane protein-2 are associated with the progression of bladder cancer. *Genet Mol Biol* (2019) 42:40–7. doi: 10.1590/1678-4685-gmb-2017-0339

28. Sun D, Liu J, Zhou L. Upregulation of circular RNA Circ–FAM53B predicts adverse prognosis and accelerates the progression of ovarian cancer *via* the Mir–646/VAMP2 and Mir–647/MDM2 signaling pathways. *Oncol Rep* (2019) 42 (6):2728–37. doi: 10.3892/or.2019.7366

29. Jung Y, Yong S, Kim P, Lee H-Y, Jung Y, Keum J, et al. VAMP2–NRG1 fusion gene is a novel oncogenic driver of non–Small-Cell lung adenocarcinoma. *J Thorac Oncol* (2015) 10(7):1107–11. doi: 10.1097/JTO.0000000000000544

30. Ge T, Yang J, Zhou S, Wang Y, Li Y, Tong X. The role of the pentose phosphate pathway in diabetes and cancer. *Front Endocrinol* (2020) 11:365. doi: 10.3389/fendo.2020.00365

31. Patra KC, Hay N. The pentose phosphate pathway and cancer. *Trends Biochem Sci* (2014) 39(8):347–54. doi: 10.1016/j.tibs.2014.06.005

32. Jin L, Zhou Y. Crucial role of the pentose phosphate pathway in malignant tumors. *Oncol Lett* (2019) 17(5):4213–21. doi: 10.3892/ol.2019.10112

33. Xia A, Zhang Y, Xu J, Yin T, Lu X-J. T Cell dysfunction in cancer immunity and immunotherapy. *Front Immunol* (2019) 10, 1719. doi: 10.3389/fimmu.2019.01719

34. Conlon KC, Miljkovic MD, Waldmann TA. Cytokines in the treatment of cancer. *J Interferon Cytokine Res* (2019) 39(1):6–21. doi: 10.1089/jir.2018.0019

35. Berraondo P, Sanmamed MF, Ochoa MC, Etxeberria I, Aznar MA, Pérez-Gracia JL, et al. Cytokines in clinical cancer immunotherapy. *Br J Cancer* (2019) 120(1):6–15. doi: 10.1038/s41416-018-0328-y

36. Chen DS, Mellman I. Elements of cancer immunity and the cancer–immune set point. *Nature* (2017) 541(7637):321–30. doi: 10.1038/nature21349

37. Gómez-Jaramillo L, Romero-García R, Jiménez-Gómez G, Riegle L, Ramos-Amaya AB, Brieva JA, et al. VAMP2 is implicated in the secretion of antibodies by human plasma cells and can be replaced by other synaptobrevins. *Cell Mol Immunol* (2018) 15(4):353–66. doi: 10.1038/cmi.2016.46

38. Agliardi C, Meloni M, Guerini FR, Zanzottera M, Bolognesi E, Baglio F, et al. Oligomeric α-syn and SNARE complex proteins in peripheral extracellular vesicles of neural origin are biomarkers for parkinson's disease. *Neurobiol Dis* (2021) 148:105185. doi: 10.1016/j.nbd.2020.105185

39. Duan X-L, Guo Z, He Y-T, Li Y-X, Liu Y-N, Bai H-H, et al. SNAP25/syntaxin4/VAMP2/Munc18-1 complexes in spinal dorsal horn contributed to inflammatory pain. *Neuroscience* (2020) 429:203–12. doi: 10.1016/j.neuroscience.2020.01.003

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 22 November 2022
DOI 10.3389/fonc.2022.1017745

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Lixia Yuan,
Zhejiang Pharmaceutical College, China
Jihong Yang,
Zhejiang University, China

*CORRESPONDENCE
Tao Fu
Futao916@163.com
Lei Guo
lincguo@100biotech.com

†These authors have contributed equally to this work

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

RECEIVED 12 August 2022
ACCEPTED 19 October 2022
PUBLISHED 22 November 2022

CITATION
Li C, Chen Z, Gao J, Tang T, Zhou L, Zhang G, Zhang D, Shen C, Guo L and Fu T (2022) MIR4435-2HG in exosomes promotes gastric carcinogenesis by inducing M2 polarization in macrophages. *Front. Oncol.* 12:1017745. doi: 10.3389/fonc.2022.1017745

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

# MIR4435-2HG in exosomes promotes gastric carcinogenesis by inducing M2 polarization in macrophages

Chaofeng Li[1†], Zhengju Chen[2,3†], Jinli Gao[4†], Tao Tang[1], Lei Zhou[1], Guochao Zhang[1], Dongdong Zhang[5], Chao Shen[5], Lei Guo[3*] and Tao Fu[1,6*]

[1]Gastrointestinal Surgery Department, China-Japan Friendship Hospital, Beijing, China, [2]Nanchang Institute of Technology, College of Medicine, China. Pooling Medical Research Institutes, Hangzhou, China, [3]Pooling Medical Research Institutes, Hangzhou, Beijing, China, [4]Department of Pathology, East Hospital, Tongji University, Shanghai, China, [5]Gastrointestinal Surgery, Peking University International Hospital, Beijing, China, [6]Department of Gastrointestinal Surgery, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education), Peking University Cancer Hospital and Institute, Beijing, China

Gastric cancer (GC) is a cancer with a high mortality rate. lncRNAs play a role in regulating GC tumorigenesis. In this paper, we analyzed differentially expressed lncRNAs between GC and adjacent normal tissues using multiple bioinformatics tools to identify new potential targets in GC. Cell viability and migration ability were detected using the Cell Counting Kit-8 (CCK-8) and transwell assays, MIR4435-2HG was negatively correlated with the survival rate of GC patients, and by inhibiting the activity of MIR4435-2HG, the viability and migration ability of GC cells could be reduced. In addition, RT- qPCR and western blot to detect gene and protein level expression, transmission electron microscopy and nanoparticle tracking analysis (NTA) to study the efficiency of exosome isolation, and flow cytometry to observe cell differentiation were employed, delivery of MIR4435-2HG shRNA *via* MKN45 cell-derived exosomes significantly reversed the MKN45 exosome-induced M2 polarization in macrophages. Furthermore, the low expression of MIR4435-2HG in MKN45 cell-derived exosomes inhibited the Jagged1/Notch and JAK1/STAT3 pathways in macrophages; MIR4435-2HG downregulated exosomes were found to significantly inhibit GC tumor growth *in vivo* by establishing a mouse model. In short, MKN45 cell-derived exosomes deliver lncRNA MIR4435-2HG, which promotes gastric carcinogenesis by inducing macrophage M2 polarization.

KEYWORDS

GC, MKN45, MIR4435-2HG, Jagged1/Notch, JAK1/STAT3

## Introduction

Gastric cancer is one of the most common high-mortality cancers. Accounting for 8% of all cancer types in terms of both cancer incidence and mortality (1, 2). The main cause of its poor prognosis is distant metastasis, which accounts for approximately 90% of cancer-related deaths (3). Surgical resection with or without chemotherapy is usually used clinically for GC, but treatment outcomes have not met expectations so far (4). Therefore, it is an important and urgent matter to find the pathogenesis of GC-related molecules and promote targeted therapy.

Long-stranded noncoding RNAs are noncoding RNAs greater than 200 bp in length with limited protein-coding potential (5), and several existing findings have shown that lncRNAs play a role in epigenetic, transcriptional, and posttranscriptional levels, as well as in cancer development (6, 7), and lncRNAs play a key role in gastric carcinogenesis (8, 9).

Exosomes are a type of vesicle secreted from the mammalian intracellular to the extracellular, consisting of membranes of multivesicular bodies, which can participate in cellular communication by transferring proteins and nucleic acids (10–13) and are an efficient drug carrier in clinical applications (14). Exosomes can play a role in tumorigenesis through communication with tumor cells and macrophages and exhibit immunosuppressive effects to some extent (15, 16). In contrast, macrophages (especially M2 macrophages) can suppress the immune response in tumor development and thus can promote drug resistance in cancer therapy (17, 18), which is widely used in clinical practice. Small interfering RNA (siRNA) is a therapeutic agent that can treat a variety of diseases, but safe, efficient, and targeted delivery of siRNA is still one of the main challenges (19). This paper focuses on exosome-targeted delivery of siRNA.

In this study, we investigated the differentially expressed lncRNAs between GC and adjacent normal tissues to find GC-associated lncRNAs and investigated the exosomal transport of siRNA, which could represent a breakthrough for GC clinical treatment and research. We found that MIR4435-2HG in exosomes is a key factor in regulating the promotion of macrophage M2 polarization and mapped the macrophage M2 polarization pathway based on this theory (Figure 1).

## Materials and methods

### GC lncRNA differential expression analysis

The GC lncRNA dataset was got from the Cancer Genome Atlas database, and differential analysis was performed using the

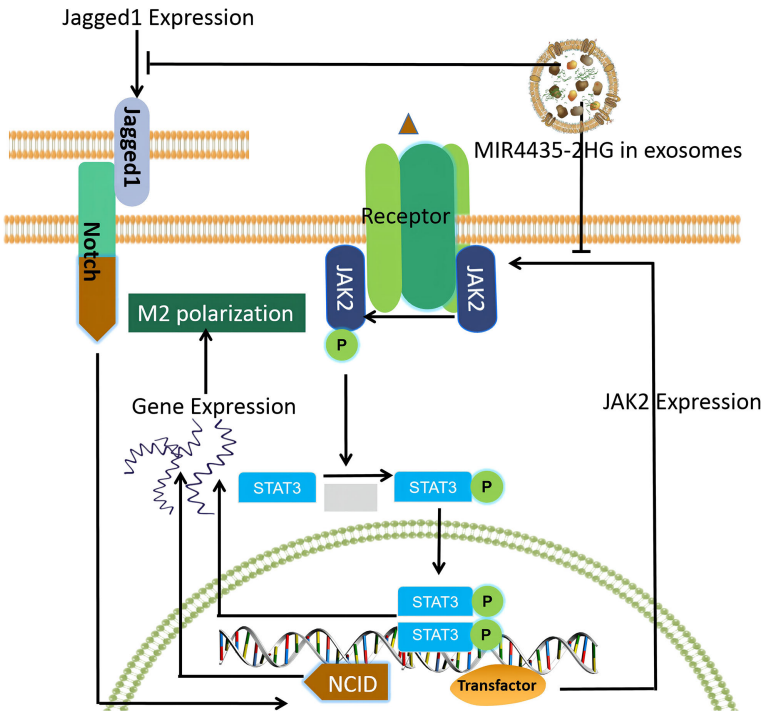



FIGURE 1
Macrophage M2 polarization pathway mapped based on our findings.

R software limma package. The results are presented in volcano plots with a threshold of |log2FC| > 1, P < 0.05. The most significantly differentially expressed gene was selected as the key lncRNA. The R software ggplot2 package was employed to analyze the differential expression of this lncRNA in tumor nd normal samples. The results are illustrated by box plots. P < 0.05 was statistically significant.

## Cell culture

Human normal gastric epithelial cell lines (GES-1), GC cell lines (MKN-45AGS and SNU-5), macrophages (THP-1), and 293T cells were placed in DMEM solution (Thermo Fisher Scientific, Waltham, MA, USA) with 10% FBS and 1% penicillin and streptomycin and incubated at a constant temperature of 37°C.

## Cell transfection

MIR4435-2HG shRNA1 and shRNA2, nontargeting sequences from Hanbio Biotechnology Co., Ltd, and OE from Shanghai GenePharma Co., Ltd were packaged into lentiviral vectors. The lentiviral vector DNA was transfected into 293T cells and incubated at 37°C after transfection. The supernatant was filtered into pellets, and the lentiviral pellets were infected with GC cells. qPCR was used to verify the transfection efficiency.

## Extraction and identification of exosomes

Cells were first cultured in a complementary medium, and when the cells reach 80% fusion with the culture medium, the medium was replaced with an FBS-free medium for 2 days, the supernatant was collected by centrifugation and filtration (Millipore, USA), and the supernatant was collected by ultracentrifugation (Beckman Coulter) to obtain exosomes. The subgroups were named KN45-Exo-NC, MKN45-Exo-MIR4435-2HG OE or MKN45-Exo-MIR4435-2HGshRNA2, respectively. Exosomes were identified using transmission electron microscopy and immunoblotting.

## Macrophage isolation and culture

Macrophages were treated with 100 ng/ml of PMA (Peprotech) and then with PBS, MKN45-Exo-NC, MKN45-Exo-MIR4435-2HG OE, or MKN45-Exo-MIR4435-2HG shRNA2 for 24 h. Macrophages were isolated using a flow cytometer (New Jersey, USA) for the macrophage surface markers CD206 (eBioscience) and CD68 (eBioscience).

## Reverse transcription-quantitative PCR (RT-qPCR)

The total RNA of samples was extracted from cells using TRIzol reagents and reverse transcribed into cDNA using the PrimeScript RT kit, and qPCR was performed using the SYBR premix Ex Taq II kit to detect gene expression. The expression of MIR4435-2HG, arginase-1, and iNOS genes was measured using β-actin as an internal reference (Table 1). The relative expression of the genes was calculated using the 2-ΔΔCT method.

## Cell Counting Kit-8 (CCK-8) assay

GC cells (5 × 103 cells/well) were plated in 96-well plates, and NC, MIR4435-2HG shRNA2, and MIR4435-2HG OE were treated at 37°C with 5% CO2 for 0, 24, 48, and 72 h. Ten microliters of CCK-8 reagent was added and incubated at 37°C for 2 h. The absorbance of each well was measured at 450 nm using an enzyme marker (Thermo Fisher Scientific).

## Western blotting

Total proteins were extracted using RIPA lysate (Beyotime) and quantified using the bicinchoninic acid protein kit (Thermo Fisher Scientific). Proteins were separated by 10% SDS-PAGE, transferred to PVDF membranes, and blocked with 5% skim milk at room temperature for 1 h. The following primary antibodies were used with β-actin as an internal reference: CD63 (Abcam; ab134045), TSG101 (Abcam; ab125011), Notch1 (Abcam; ab52627), Notch2 (Abcam;ab118824), vimentin (CST; #5741), N-cadherin (CST; #13116), E-cadherin (CST; #14472), STAT3 (Abcam; ab68153), p-STAT3 (Abcam;

TABLE 1 Gene primer sequences.

| Symbol | Forward | Reverse |
|---|---|---|
| MIR4435-2HG | 5'-TGGGCAATAGGCGATACGAT-3' | 5'-TGGGCAATAGGCGATACGAT-3' |
| Arginase-1 | 5'-CACCTGGGCAAGGATTCA-3' | 5'-CTCAACTGGTGTCGTGGAGTC-3' |
| iNOS | 5'-CTCAACTGGTGTCGTGGAGTC-3' | 5'-ACAGCACACCGTAGTCTCGG-3' |
| β-actin | 5'-GTCCACCGCAAATGCTTCTA-3' | 5'-TGCTGTCACCTTCACCGTTC-3' |

ab267373), JAK1 (CST; #29261), p-JAK1 (CST; #3331), Jagged1 (CST; #2620), Hes1 (CST; #11988), Hes5 (Abcam; ab194111), and β-actin (Abcam; ab8226). After initial incubation, membranes were incubated with HRP-coupled secondary antibodies for 1 h at room temperature. Protein bands were visualized using an ECL kit. Densitometric analysis was performed using IPP6.0 (Image-ProPlus6.0).

## Immunofluorescence

Cells were inoculated into 24-well plates overnight, prefixed with 4% paraformaldehyde for 10 min, and then fixed in methanol for 10 min. Cells were stained with Phalloidin-iFluor 488 reagent (ab176753) and the PKH26 Red Cell Membrane Staining Kit (D0030, Solarbio), respectively. Phalloidin-iFluor 488 stains the cytoskeleton and PKH26 stains the cell membrane, enabling localization of the exosome.

## Cell migration assay

GC cells were inoculated in the upper chamber of the medium containing 1% FBS and the density was adjusted to approximately $1.0\times10^6$ cells per chamber. RPMI 1640 medium containing 10% FBS was added to the lower chamber. After incubation at 37°C for 24 h, the transfer chamber was rinsed twice (5 min each time) with PBS. The cells were fixed with 5% glutaraldehyde at 4°C, stained with 0.1% crystal violet for 30 min, and the transwell chambers were washed twice with PBS, followed by observation under a microscope. The number of migrating cells was considered to be a reflection of migratory capacity.

## *In vivo* study

Thirty BALB/c nude mice (6-8 weeks old) purchased from VitalLiver (Beijing, China) were housed in a dedicated SPF facility for culturing. MKN45 cells were co-cultured with macrophages; macrophages-exo, macrophagesexo-NC, macrophagesexo-MIR4435-2HG OE, or macrophagesexo-MIR4435-2HG shRNA2 was transplanted into BALB/c nude mice, and tumor size was measured regularly.

## Statistical analysis

Three independent replicate experiments were performed for each group of samples, and the results are presented as mean ± standard deviation (SD). Statistical methods included the t-test (with 2 groups) and one-way ANOVA (with 3 or more groups), and Tukey's test was used. $P < 0.05$ was statistically significant.

# Results

## GC lncRNA differential expression results

Through differential analysis of GC and normal samples next to cancer, 847 genes were obtained, including 636 upregulated genes and 211 downregulated genes (Figure 2A). The expression results of MIR4435-2HG in tumors compared with normal samples showed that MIR4435-2HG was highly expressed in GC tissues (Figure 2B).

## Overexpression of MIR4435-2HG significantly promoted the proliferation of gastric cancer cells

The expression of MIR4435-2HG in various cells was detected using the RT-qPCR method. The results showed that MIR4435-2HG expression was significantly higher in MKN-45 and AGS cells than in GES-1 cells (Figure 3A). In addition, MIR4435-2HG OE significantly upregulated MIR4435-2HG expression in MKN-45 and AGS cells (Figure 3B) but was suppressed in the presence of MIR4435-2HG shRNA (Figure 3C), and GC cells were more sensitive to MIR4435-2HG shRNA2. MIR4435-2HG shRNA2 was chosen as the subject for the follow-up study. High MIR4435-2HG expression could promote the proliferation of MKN-45 and AGS cells (Figures 3D, E). In summary, overexpression of MIR4435-2HG significantly promoted the proliferation of GC cells.

## Overexpression of MIR4435-2HG increases the migratory ability of GC cells

Cell migration was examined using the transwell assay, and the results showed that overexpression of MIR4435-2HG significantly increased the migration of GC cells, whereas downregulation of MIR4435-2HG inhibited the migration of GC cells (Figures 4A, B). In addition, the expressions of N-cadherin and vimentin in GC cells were significantly upregulated by MIR4435-2HG overexpression but downregulated in the presence of MIR4435-2HG shRNA. In contrast, silencing of MIR4435-2HG significantly upregulated the E-cadherin protein level, whereas upregulation of MIR4435-2HG had an inhibitory effect on E-cadherin (Figure 4C). In conclusion,

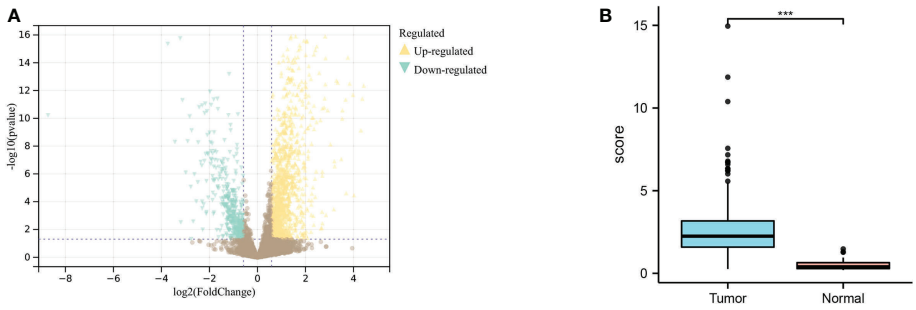



FIGURE 2
Differential analysis of TCGA dataset for gastric cancer. **(A)** Differential expression results of the TCGA dataset; yellow indicates upregulated genes, blue indicates downregulated genes. **(B)**. Analysis of MIR4435-2HG expression in gastric cancer tissues and normal tissues adjacent to cancer, blue indicates tumor, and red indicates normal tissue. *** indicates P<0.001.

overexpression of MIR4435-2HG increased the migration of GC cells by promoting the EMT process.

## Results of exosome extraction and identification

Round particles with diameters of 30 to 150 nm were observed by TEM, and nanoparticle tracking analysis (NTA) yielded a size distribution close to that of TEM (Figure 5A). The expression of exosomal proteins (CD63 and TSG101) was significantly higher in GC cell exosomes compared with GC cells (Figure 5B). The expression of MIR4435-2HG in MKN45-Exo was significantly higher than that of GES-1-Exo (Figure 5C), whereas MIR4435-2HG overexpression or knockdown in MKN45 cell exosomes had a limited effect on exosomal protein levels (Figure 5D). MIR4435-2HG OE significantly upregulated MIR4435-2HG expression in MKN45 cells and MKN45 cell exosomes, and MIR4435-2HG shRNA2 significantly suppressed its expression (Figure 5E). Tumor-derived exosomes labeled with fluorescent PKH26 were internalized by unstained macrophages when co-cultured with macrophages (Figure 5F). Exosomes from MKN45 cells delivered MIR4435-2HG to PMA-treated THP-1 cells (Figure 5G). Exosomes were successfully isolated from GC cells.

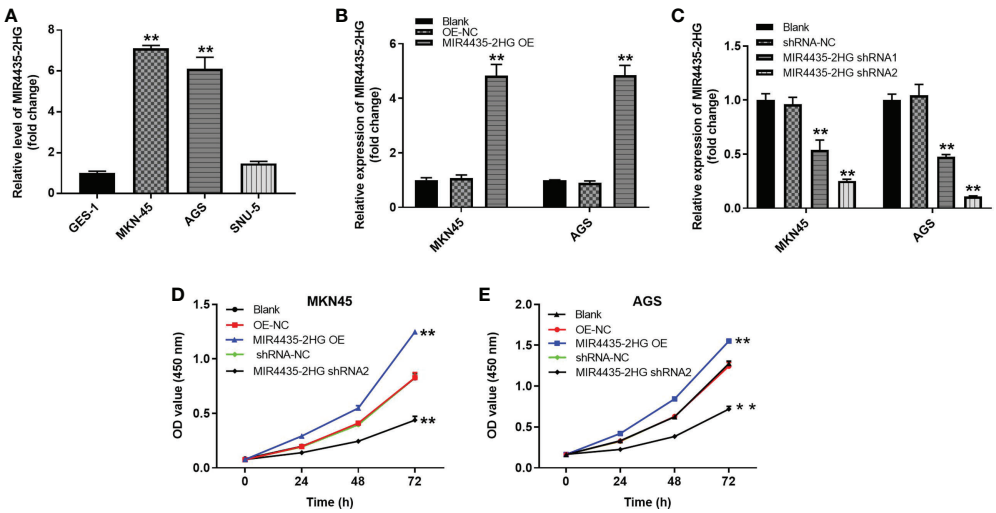



FIGURE 3
MIR4435-2HG expression results. **(A)** Expression of MIR4435-2HG in GES-1, MKN45, AGS and SNU-5 cells by RT-Qpcr. **(B)** MKN45 or AGS cells were transfected with MIR4435-2HG OE. Then, the expression of MIR4435-2HG in GC cells was detected by RT-qPCR. **(C)** MKN45 or AGS cells were transfected with MIR4435-2HG shRNA1 or shRNA2. Then, the expression of MIR4435-2HG in GC cells was detected by RT-qPCR. **(D, E)** The viability of gastric cancer cells was tested by CCK8 assay. **P < 0.01 compared with control (OE-NC; shRNA-NC).

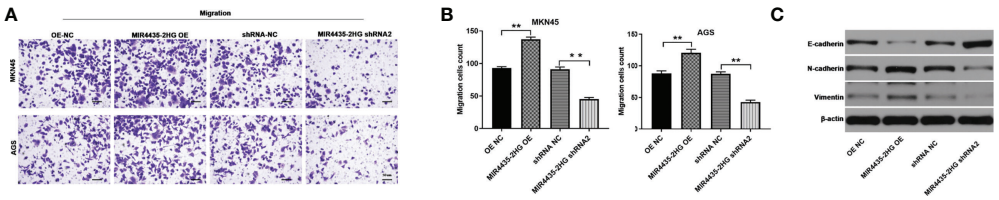



FIGURE 4
Overexpression of MIR4435-2HG increased the migration of gastric cancer cells. **(A, B)** The migration of gastric cancer cells was measured by the transwell assay. **(C)** The protein levels of E-cadherin, N-cadherin, and vimentin in GC cells were detected by western blot. **P < 0.01 compared with the control group.

## Overexpression of exosomal MIR4435-2HG promotes M2 polarization in macrophages

To find the immunomodulatory effect of MKN45 cell-derived exosomes on macrophages, THP-1 cells were treated with 50 μg/ml of exosomes excreted from MKN45 cells. The distribution of CD86 (M1 phenotype marker) in THP-1 cells was significantly reduce, and THP-1 cells and macrophages exhibited the CD86low/CD206 high phenotype when cells were incubated with exosomes containing high levels of MIR4435-2HG. Additionally, exosomes from THP-1 cells significantly upregulated the distribution rate of CD206 (M2 marker) in THP-1 cells, which was partially reversed by exosomes downregulated with MIR4435-2HG (Figures 6A, B). Exosomes of MKN45 origin significantly upregulated arginase-1 levels, whereas exosomes downregulated with MIR4435-2HG reversed the effect of exosomes from MKN45 cells, and the exosomes had a limited effect on iNOS expression (Figure 6C). The expression of M2 markers (IL-10 and TGF-β) was significantly increased in macrophages incubated with MKN45 cell-derived exosomes compared to the PBS group, which was further exacerbated by MIR4435-2HG upregulated exosomes, and conversely, MIR4435-2HG downregulated exosomes partially reversed the effect of MKN45 cell-derived exosomes (Figure 6D). In summary, overexpression of exosomal MIR4435-2HG promotes M2 polarization in macrophages.

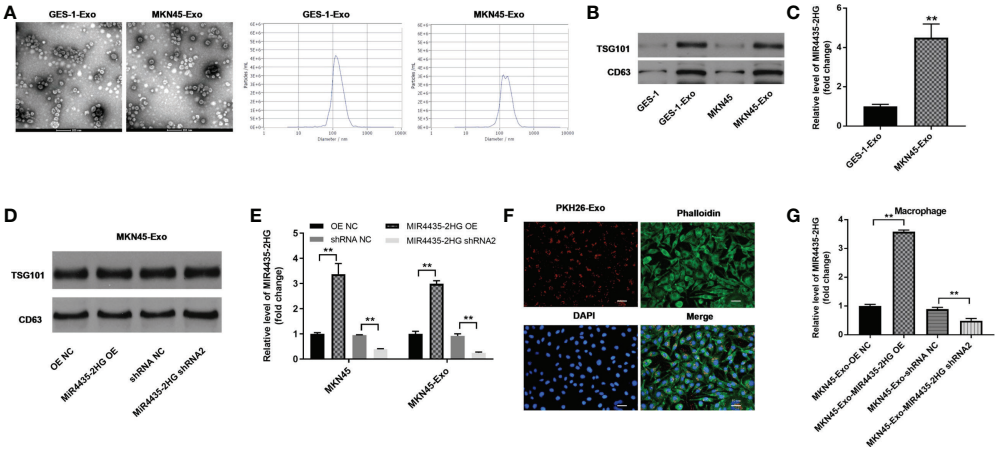



FIGURE 5
Results of exosome extraction and identification. **(A)** The separation efficiency of exosomes was examined by TEM. The particle sizes of exosomes were measured by NTA. **(B)** The expressions of TSG101 and CD63 in GES-1, MKN45 cells or exosomes derived from GES-1 or MKN45 cells were detected by WB. **(C)** The levels of MIR4435-2HG in exosomes derived from GES-1 or MKN45 cells were measured by RT-qPCR. **(D)** The expression of TSG101 in exosomes derived from MKN45 cells was detected by western blot. **(E)** The expression of MIR4435-2HG in MKN45 or exosomes derived from MKN45 cells was detected by RT-qPCR. **(F)** THP-1 cells were treated with 100 ng/ml of PMA and co-cultured with MKN45 cell-derived exosomes, and the location of exosomes was observed by immunofluorescence staining. **(G)** Macrophages were co-cultured with MKN45-Exo-OE NC, MKN45-Exo-MIR4435-2HG OE, MKN45-Exo-shRNA NC, or MKN45-Exo-MIR4435-2HG shRNA2. Then, the expression of MIR4435-2HG in macrophages was detected by RT-qPCR. **P < 0.01 compared with the control group.

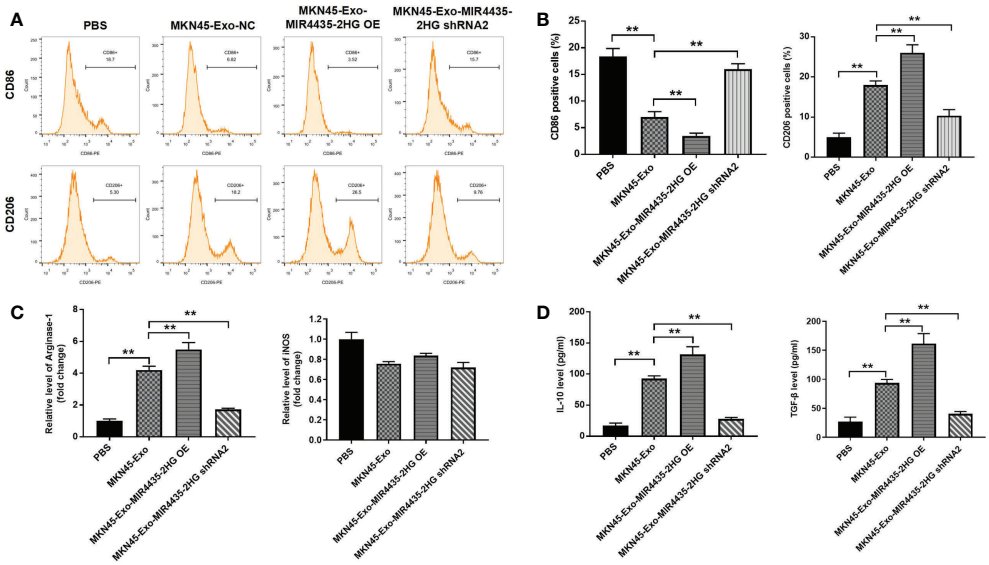



FIGURE 6
Exosome MIR4435-2HG promotes M2 polarization in macrophages. **(A, B)** Macrophages were co-cultured with MKN45-Exo-NC, MKN45-Exo-MIR4435-2HG OE, or MKN45-Exo-MIR4435-2HG shRNA2 for 24 h, and then the distribution rate of CD86 or CD206 in macrophages was detected by flow cytometry. **(C)** The levels of Arginase-1 and iNOS in macrophages were tested by RT-qPCR. **(D)** The levels of TGF-β and IL-10 in supernatants of macrophages were detected by ELISA. **P < 0.01.

## Exosome-induced M2 macrophages promote the migration of GC cells by facilitating the EMT process

Tumor cell-derived exosome-induced M2 macrophages significantly increased the migration of GC cells. Meanwhile, MIR4435-2HG upregulated exosomes aggravated this phenomenon, whereas MIR4435-2HG downregulated exosomes reversed it (Figure 7A). MKN45 cells exhibited a spindle-shaped morphology when co-cultured with M2 macrophages, which was reversed by MIR4435-2HG downregulated exosomes (Figure 7B). N-cadherin and vimentin expression was significantly increased in MKN45 cells after co-culturing with exosome-induced M2 macrophages, whereas MIR4435-2HG downregulated exosomes caused significant downregulation of N-cadherin and vimentin. The opposite data were obtained on E-cadherin expression (Figure 7C). In conclusion, exosome-induced M2 macrophages promote the migration of GC cells by increasing the EMT process.

## MIR4435-2HG in exosomes promotes M2 polarization of macrophages by regulating Jagged1/Notch and JAK1/STAT3 axes

Western blot results showed that exosomes from MKN45 cells significantly upregulated the protein levels of Notch1, Notch2, Jagged1, Hes1, and Hes5, whereas MIR4435-2HG downregulated exosomes partially inhibited the effect of exosomes on these proteins (Figure 8A). When MIR4435-2HG was downregulated in exosomes, its induction of the JAK1/STAT3 signaling pathway in macrophages was inhibited (Figure 8B). Taken together, exosomes from MKN45 cells promote M2 polarization in macrophages by regulating the Jagged1/Notch and JAK1/STAT3 axes.

## Upregulation of MIR4435-2HG in exosomes significantly promotes tumor growth in GC

Results from xenograft mouse models showed that exosomes from MKN45 cells significantly increased tumor size in nude mice, which was partially reversed by exosomes downregulated by MIR4435-2HG (Figures 9A, B). Exosomes from MKN45 cells greatly increased tumor weight in nude mice, which was partially reversed by exosomes downregulated by the MIR4435-2HG phenomenon (Figure 9C). MKN45 cell-derived exosomes significantly upregulated N-cadherin and vimentin expression in mouse tissues, whereas MIR4435-2HG downregulated exosomes reversed this phenomenon. Additionally, MKN45 cell-derived exosomes inhibited E-cadherin expression in nude mice, whereas MIR4435- 2HG OE upregulated exosomes further enhanced the role of exosomes (Figure 9D). Taken together, MIR4435-2HG upregulation in exosomes induced M2 polarization in

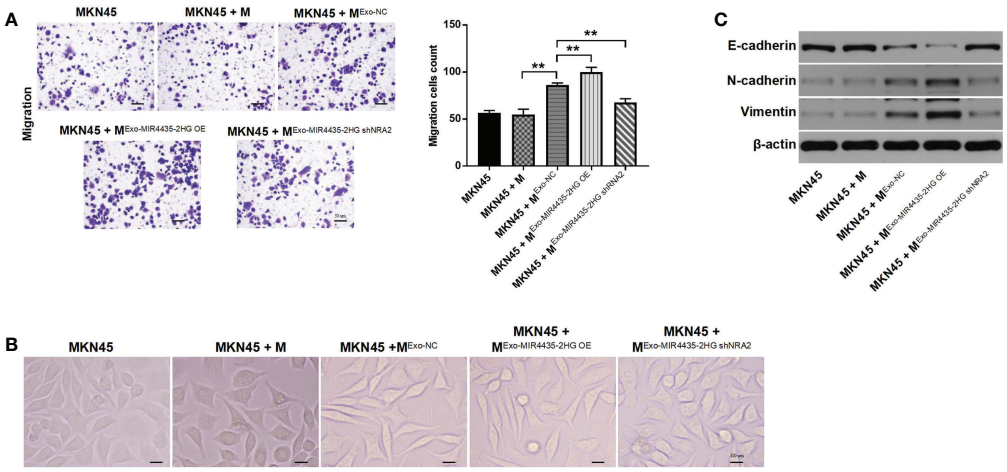



FIGURE 7
Exosome-induced M2 macrophages promote the migration of gastric cancer cells by inhibiting the EMT process. **(A)** MKN45 cells were co-cultured with macrophages, ExoNC-treated macrophages, Exo-MIR4435-2HGOE-treated macrophages, and Exo-MIR4435-2HGshRNA2-treated macrophages for 24 h, and the migration of MKN45 cells was detected by the transwell method. **(B)** The morphology of MKN45 cells was observed under a microscope. **(C)** The E-cadherin, N-cadherin, and vimentin protein levels in macrophages were detected by western blot. **P < 0.01.

macrophages, which significantly promoted the growth of GC tumors.

## Discussion

RNA interference is a phenomenon that can reverse the silencing of any gene, and clinical delivery materials are generally selected to transport siRNA to sites of action in target tissue cells (20). Exosomes, also known as signalosomes, can be expressed on exosomal membranes or packaged through ligands and adhesion molecules, thereby performing signaling (21). In this study, we focused on the role and potential value of MIR4435-2HG in GC, starting from the exosomes of GC cells.

The results showed that MIR4435-2HG upregulation could promote tumorigenesis in GC, and a previous study found that

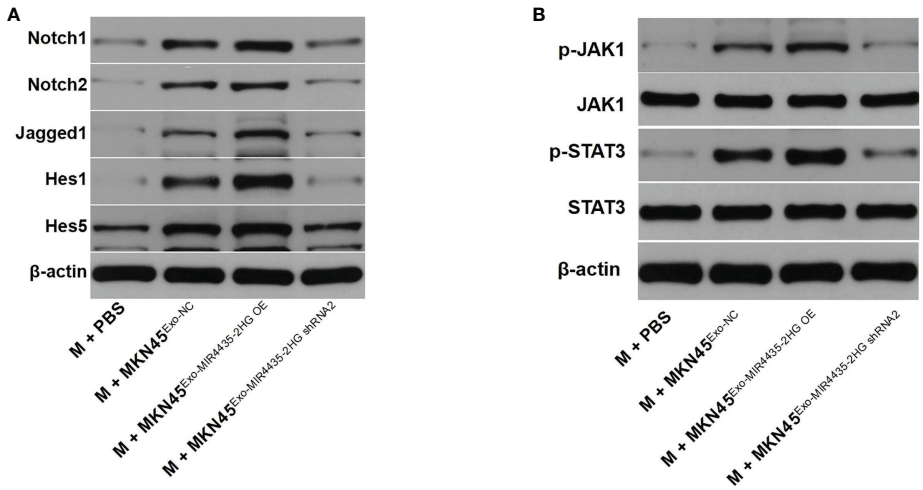



FIGURE 8
Exosomes from MKN45 cells promote M2 polarization in macrophages by regulating the Jagged1/Notch and JAK1/STAT3 axes. **(A)** The Jagged1, Notch1, Notch2, Hes1, and Hes5 protein levels in macrophages were detected by WB. **(B)** WB detection of p-JAK1, JAK1, p-STAT3, and STAT3 protein levels in macrophages.

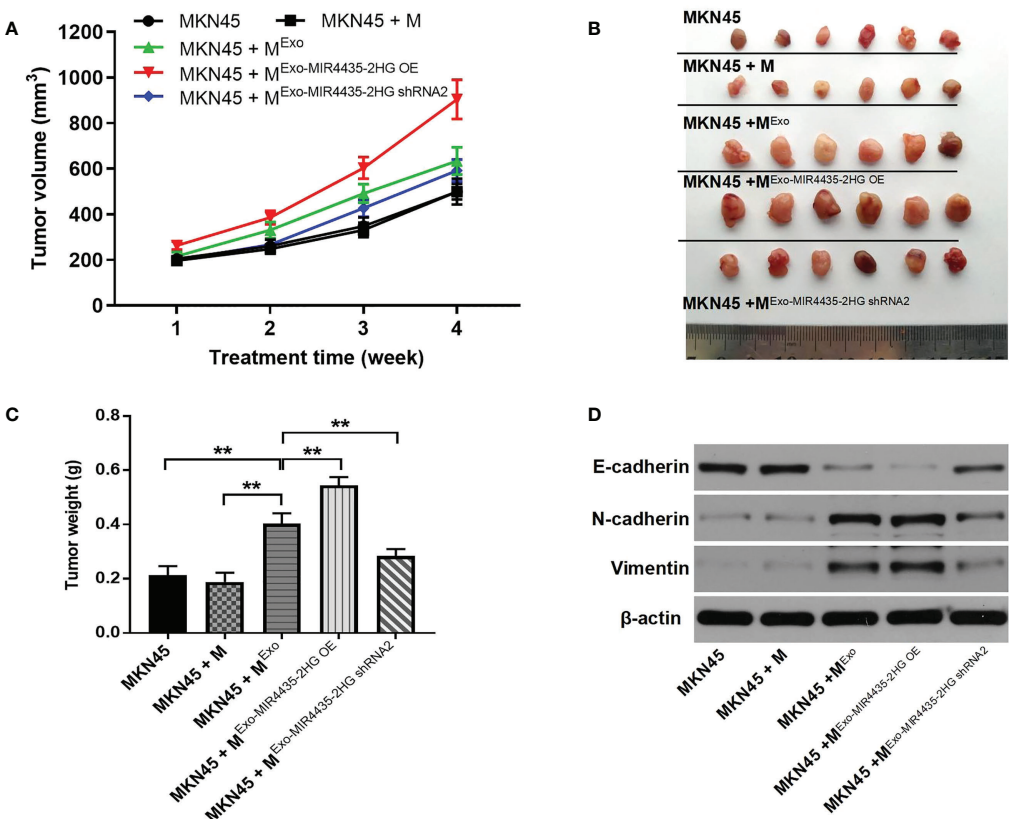



FIGURE 9
M2 m acrophages induced by exosomes significantly promotes the tumor growth of gastric cancer. **(A)** Tumor volumes of nude mice were tested. **(B)** Tumor tissues of nude mice were pictured. **(C)** Tumor weights of nude mice were measured. **(D)** The expression of E-cadherin, Ncadherin, and vimentin in tissues of nude mice were detected by WB. **P < 0.01 compared with MKN45.

MIR4435-2HG could promote tumorigenesis in GC cells (22), which was consistent with our findings. In addition, we found that exosomal delivery of MIR4435-2HG from GC cells promotes the polarization of M2 macrophages, which in turn promotes GC development, and we speculate that exosomal MIR4435-2HG might act as an oncogene in GC and can be considered a marker of cancer tumorigenesis. Some studies have found that MIR4435-2HG is associated with various cancers, such as cervical cancer, GC, and renal clear cell carcinoma (23), which is consistent with our speculation. We also found that exosomes might increase migration of GC cells *in vitro* and *in vivo* when they induce selective activation of macrophages toward the M2 phenotype, and we hypothesized that macrophage M2 polarization might promote gastric carcinogenesis. One study found that macrophage M2 polarization can induce colorectal carcinogenesis by secreting CXCL13 (24). In addition, exosomes can be considered important mediators of the tumor microenvironment, and they can act as important messengers to regulate the cross-talk between different cells.

Jagged1 (JAG1) is an important Notch ligand, and JAG1/Notch signaling controls oncogenic processes in multiple cell types and is related to poor clinical prognosis (25). We found that exosomes from MKN45 cells can promote M2 polarization in macrophages by regulating the Jagged1/Notch or JAK1/STAT3 axis.

The JAK1/STAT3 axis is associated with phosphate metabolism, phosphorylation, and nuclear accumulation of STAT3 when JAK1 expression is upregulated (26), and the JAK1/STAT3 signaling pathway has been found to promote M2 polarization in macrophages (27), which is similar to our findings. The JAK1/STAT3 signaling pathway might be an M2 macrophage polarization-mediating factor.

In conclusion, we found that MKN45 cell-derived exosomes could induce M2 macrophage polarization and promote GC tumorigenesis; therefore, MIR4435-2HG could be used as a biomarker for GC and a new target for GC 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 author.

## Author contributions

CL: Conceptualization, Resources, Original Draft. ZC: Conceptualization, Resources, Review and Editing. JG: Conceptualization, Resources, Review and Editing. TT: Methodology, Original Draft, Review and Editing. LZ: Methodology, Original Draft, Review and Editing GZ: Methodology, Resources, Review and Editing. DZ: Validation, Resources, Supervision. CS: Validation, Original Draft, Supervision. LG: Validation, Original Draft, Project administration. TF: Formal analysis, Original Draft, Funding acquisition.

## Funding

National Natural Science Foundation of China (82002477).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Garcia-Perez J, Lope V, Fernandez de Larrea-Baz N, Molina AJ, Tardon A, Alguacil J, et al. Risk of gastric cancer in the environs of industrial facilities in the MCC-Spain study. *Environ Pollut* (2021) 278:116854. doi: 10.1016/j.envpol.2021.116854

2. Guan D, Li C, Li Y, Li Y, Wang G, Gao F, et al. The DpdtbA induced EMT inhibition in gastric cancer cell lines was through ferritinophagy-mediated activation of p53 and PHD2/hif-1alpha pathway. *J Inorg Biochem* (2021) 218:111413. doi: 10.1016/j.jinorgbio.2021.111413

3. Matsui R, Inaki N, Tsuji T. The impact of the preoperative hand grip strength on the long-term outcomes after gastrectomy for advanced gastric cancer. *Surg Today* (2021) 51(7):1179–87. doi: 10.1007/s00595-021-02256-y

4. Wang X, Chen Y, Gao Y, Zhang H, Guan Z, Dong Z, et al. Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat Commun* (2021) 12(1):1637. doi: 10.1038/s41467-021-21674-7

5. Zhong X, Wen X, Chen L, Gu N, Yu X, Sui K. LncRNA KCNQ1OT1 promotes progression of gastric cancer via the miR-145-5p/ARF6 axis. *J Gene Med* (2021) 23(5):e3330. doi: 10.1002/jgm.3330

6. Wu Y, Deng J, Lai S, You Y, Wu J. A risk score model with five long non-codingRNAs for predicting prognosis in gastric cancer: an integrated analysis combining TCGA and GEO datasets. *PeerJ.* (2021) 9:e10556. doi: 10.7717/peerj.10556

7. Chen W, Zhang K, Yang Y, Guo Z, Wang X, Teng B, et al. MEF2A-mediated lncRNA HCP5 inhibits gastric cancer progression *via* MiR-106b-5p/p21 axis. *Int J Biol Sci* (2021) 17(2):623–34. doi: 10.7150/ijbs.55020

8. Liang LC, Liu LQ, Liu L, Liu DL, He YR, Wan X, et al. Long non-coding RNA BX357664 inhibits gastric cancer progression by sponging miR-183a-3p to regulate the PTEN expression and PI3K/AKT pathway. *Food Chem Toxicol* (2021) 150:112069. doi: 10.1016/j.fct.2021.112069

9. Qu F, Zhu B, Hu YL, Mao QS, Feng Y. LncRNA HOXA-AS3 promotes gastric cancer progression by regulating miR-29a-3p/LTbetaR and activating NF-kappaB signaling. *Cancer Cell Int* (2021) 21(1):118. doi: 10.1186/s12935-021-01827-w

10. Zhang J, Li S, Li L, Li M, Guo C, Yao J, et al. Exosome and exosomal microRNA: trafficking, sorting, and function. *Genomics Proteomics Bioinf* (2015) 13(1):17–24. doi: 10.1016/j.gpb.2015.02.001

11. Zhang H, Wang L, Li C, Yu Y, Yi Y, Wang J, et al. Exosome-induced regulation in inflammatory bowel disease. *Front Immunol* (2019) 10:1464. doi: 10.3389/fimmu.2019.01464

12. Yang B, Chen Y, Shi J. Exosome biochemistry and advanced nanotechnology for next-generation theranostic platforms. *Adv Mater* (2019) 31 (2):e1802896. doi: 10.1002/adma.201802896

13. Zhang L, Yu D. Exosomes in cancer development, metastasis, and immunity. *Biochim Biophys Acta Rev Cancer* (2019) 1871(2):455–68. doi: 10.1016/j.bbcan.2019.04.004

14. Batrakova EV, Kim MS. Using exosomes, naturally-equipped nanocarriers, for drug delivery. *J Control Release* (2015) 219:396–405. doi: 10.1016/j.jconrel.2015.07.030

15. Mashouri L, Yousefi H, Aref AR, Ahadi AM, Molaei F, Alahari SK. Exosomes: composition, biogenesis, and mechanisms in cancer metastasis and drug resistance. *Mol Cancer* (2019) 18(1):75. doi: 10.1186/s12943-019-0991-5

16. Milane L, Singh A, Mattheolabakis G, Suresh M, Amiji MM. Exosome mediated communication within the tumor microenvironment. *J Control Release* (2015) 219:278–94. doi: 10.1016/j.jconrel.2015.06.029

17. Su MJ, Aldawsari H, Amiji M. Pancreatic cancer cell exosome-mediated macrophage reprogramming and the role of MicroRNAs 155 and 125b2 transfection using nanoparticle delivery systems. *Sci Rep* (2016) 6:30110. doi: 10.1038/srep30110

18. Lan J, Sun L, Xu F, Liu L, Hu F, Song D, et al. M2 macrophage-derived exosomes promote cell migration and invasion in colon cancer. *Cancer Res* (2019) 79(1):146–58. doi: 10.1158/0008-5472.CAN-18-0014

19. Ma S, Liu M, Xu Z, Li Y, Guo H, Ge Y, et al. A double feedback loop mediated by microRNA-23a/27a/24-2 regulates M1 versus M2 macrophage polarization and thus regulates cancer progression. *Oncotarget.* (2016) 7 (12):13502–19. doi: 10.18632/oncotarget.6284

20. Kanasty R, Dorkin JR, Vegas A, Anderson D. Delivery materials for siRNA therapeutics. *Nat Mater* (2013) 12(11):967–77. doi: 10.1038/nmat3765

21. Shao J, Zaro J, Shen Y. Advances in exosome-based drug delivery and tumor targeting: From tissue distribution to intracellular fate. *Int J Nanomed* (2020) 15:9355–71. doi: 10.2147/IJN.S281890

22. Wang H, Wu M, Lu Y, He K, Cai X, Yu X, et al. LncRNA MIR4435-2HG targets desmoplakin and promotes growth and metastasis of gastric cancer by activating wnt/β-catenin signaling. *Aging (Albany NY).* (2019) 11(17):6657–73. doi: 10.18632/aging.102164

23. Zhang M, Yu X, Zhang Q, Sun Z, He Y, Guo W. MIR4435-2HG: A newly proposed lncRNA in human cancer. *Biomed Pharmacother* (2022) Volume 150:112971. doi: 10.1016/j.biopha.2022.112971

24. Zhao S, Mi Y, Guan B, Zheng B, Wei P, Gu Y, et al. Tumor-derived exosomal miR-934 induces macrophage M2 polarization to promote liver metastasis of colorectal cancer. *J Hematol Oncol* (2020) 13(1):156. doi: 10.1186/s13045-020-00991-2. Erratum in: J Hematol Oncol. 2021 Feb 23;14(1):33.

25. Xiu MX, Liu YM, Kuang BH. The oncogenic role of Jagged1/Notch signaling in cancer. *BioMed Pharmacother* (2020) 129:110416. doi: 10.1016/j.biopha.2020.110416

26. Zeng R, Lu X, Lin J, Ron Z, Fang J, Liu Z, et al. FOXM1 activates JAK1/STAT3 pathway in human osteoarthritis cartilage cell inflammatory reaction. *Exp Biol Med (Maywood)* (2021) 246(6):644–53. doi: 10.1177/1535370220974933

27. Ma S, Liu M, Xu Z, Li Y, Guo H, Ge Y, et al. A double feedback loop mediated by microRNA-23a/27a/24-2 regulates M1 versus M2 macrophage polarization and thus regulates cancer progression. *Oncotarget.* (2016) 7 (12):13502–19. doi: 10.18632/oncotarget.6284

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 29 November 2022
DOI 10.3389/fonc.2022.1004511

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Yuqing Hang,
Salk Institute for Biological Studies, United States
Jin Gong,
The First Affiliated Hospital of Yangtze University, China

*CORRESPONDENCE
Jianzhong Peng
jianzhongpeng@outlook.com
Peng Xiao
tulipxp@zju.edu.cn

†These authors share first authorship

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

RECEIVED 27 July 2022
ACCEPTED 26 October 2022
PUBLISHED 29 November 2022

CITATION
Peng J, Wang T, Yue C, Luo X and Xiao P (2022) PGAM5: A necroptosis gene associated with poor tumor prognosis that promotes cutaneous melanoma progression. *Front. Oncol.* 12:1004511. doi: 10.3389/fonc.2022.1004511

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

# PGAM5: A necroptosis gene associated with poor tumor prognosis that promotes cutaneous melanoma progression

Jianzhong Peng[1*†], Tao Wang[1], Chao Yue[1], Xianyan Luo[1] and Peng Xiao[2,3*]

[1]Department of Dermatologic Surgery, Hangzhou Third People's Hospital, Hangzhou, Zhejiang, China, [2]Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, China, [3]Institute of Immunology, Zhejiang University School of Medicine, Hangzhou, China

Cutaneous melanoma is the deadliest type of skin cancer, and its highly aggressive and metastatic nature leads to an extremely poor prognosis. Necrotizing apoptosis, a specific form of programmed cell death, has been extensively studied in recent years. In this study, we analyzed the relationship between necroptosis-related functional genes and cutaneous melanoma in order to identify the biomarkers associated with the prognosis and progression of cutaneous melanoma. Cutaneous melanoma samples were classified into three subgroups on the basis of a necroptosis gene set. These subgroups were subjected to a prognostic survival analysis, and the greatest differences were observed between subgroups C1 and C3. Between these subgroups, 28 necrotizing apoptosis-related genes were significantly differently expressed. Among these, 16 necrotizing apoptosis-related genes were associated with cutaneous melanoma prognosis. Downscaling analysis and prognostic modeling using the least absolute shrinkage and selection operator analysis yielded nine pivotal genes and revealed phosphoglycerate translocase 5 (PGAM5) as the key gene. Then, qRT-PCR was used to verify the expression level of PGAM5. The results showed that PGAM5 was highly expressed in cutaneous melanoma tissues. In this study, a bioinformatics approach was used to identify PGAM5, a biomarker whose high expression is associated with the poor prognosis of cutaneous melanoma.

KEYWORDS

**SKCM, necrotizing apoptosis, PGAM5, prognosis, biomarker**

# Introduction

Melanocytes produce melanin, a pigment that help protect the skin against UV damage, and then they transfer this pigment to keratin-forming cells (1). However, when the skin is exposed to UV light or other stimuli for a long time, melanocytes may be transformed into melanoma cells due to the abnormal changes that occur at the molecular and biochemical levels (2). Cutaneous melanoma (SKCM) is the most common form of melanoma, and its incidence has steadily increased worldwide in recent years (3). Although the incidence of cutaneous melanoma is not as high as that of other cancers, the aggressive nature of cutaneous melanoma and its high mortality rate render it the most lethal type of skin cancer (4). The Global Cancer Statistics 2020 has revealed that there are over 320,000 new cutaneous melanoma cases and nearly 60,000 deaths (5). Differences in ethnic skin phenotype and in the amount of sunlight exposure, along with gender and age specificity, lead to variations in the prevalence of cutaneous melanoma across countries (6). Studies have shown that overexposure to UV radiation is the main cause of cutaneous melanoma development; in addition, the number of melanocytic nevi, genetic susceptibility, and family history are important factors contributing to the development of cutaneous melanoma (7, 8).

"Early detection, early treatment" has been emphasized in oncology treatment because most tumors can be potentially cured in early stages but show poor prognosis in late stages (9). Most newly diagnosed cutaneous melanoma cases are in their early stages; usually, these cases can be treated with surgical excision, and they are cured almost completely (10). However, the risk of recurrence remains one of the major concerns for patients with melanoma, and survival rates drop dramatically when the disease metastasizes (11). Cutaneous melanoma is a malignant neoplasm that accounts for more than 80% of all mortality in skin cancer patients (12). The highly aggressive and metastatic nature of cutaneous melanoma results in an extremely poor prognosis, with a five-year overall survival (OS) rate of not exceeding 15% for patients with advanced disease (13, 14). Therefore, the search for biomarkers associated with cutaneous melanoma progression and prognosis can help in further understanding the mechanisms of cutaneous melanoma development and progression, and it can provide useful targets for clinical treatment.

Necroptosis is a new form of programmed cell death activated by necrosomes, and it involves serine/threonine-protein kinase 1 (RIPK1) and RIPK3, as well as pseudokinase mixed-spectrum kinase structural domain-like protein (MLKL) (15). The release of cell contents into the microenvironment following necroptosis can cause a severe inflammatory response; various diseases, including certain cancers, have been found to be associated with the *in vivo* effects of necroptosis (16). Necroptosis has been speculated to play a significant regulatory role in melanoma (17). In this study, we aimed to clarify the role and function of necroptosis-related genes in the precision treatment of cutaneous melanoma through bioinformatics analysis. First, tumor samples were classified into subtypes through consensus clustering, and the core genes were identified through differential expression analysis, OS analysis, and least absolute shrinkage and selection operator (LASSO) analysis to resolve the tumor heterogeneity in necroptosis-related cutaneous melanoma. Then, *in vitro* experiments were conducted to validate the expression of the key necrotizing apoptosis-associated genes.

# Materials and methods

## Data and material sources

Clinical information corresponding to the RNA sequencing dataset for cutaneous melanoma was obtained from the The Cancer Genome Atlas (TCGA) database that includes 470 patient samples. Data on paraneoplastic skin melanoma samples were obtained from the The Genotype-Tissue Expression (GTEx) database, which includes 1809 cases. The gene sets associated with necroptosis were obtained from the Molecular Signature Database (MSigDB) (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp), from which 33 genes associated with humans were screened; then, a necroptosis-related functional gene set was constructed.

## Consensus clustering

Subtype grouping of cutaneous melanoma samples was performed using the constructed necroptosis gene set, with a maximum number of clusters of 6, by using the R package ConsensusClusterPlus; the clustering heat map was analyzed by the R package pheatmap. The appropriate number of clusters by which to divide the samples was selected, and then the subgroups were subjected to a survival analysis using the R packages survival and survminer. The survival status of each subgroup was demonstrated by a KM survival curve, in which different groups were subjected to a log-rank test.

## Preliminary gene screening

The two subgroups with large differences in survival status based on the KM survival curve were compared in terms of the differential expression of necrotizing apoptotic genes. Box line plots were drawn using the R package ggplot2, and the Wilcoxon rank-sum test was employed to test the significance of gene expression between the two subgroups. Also, observation of factors with prognostic impact in the cutaneous melanoma samples was achieved using the R package SURVIVAL. One-

way cox was used to analyze the genes with significant prognostic features, and forest plots were drawn using the R package forestplot to show the P-value, risk factor (HR), and 95% confidence interval (95% CI) for the top 20 significant genes. Intersecting genes were identified using Venn diagrams to select the prognostic correlates that were significantly expressed in cutaneous melanoma.

## LASSO

A dimensionality reduction analysis was performed, and prognostic models were constructed using LASSO. A LASSO-Cox regression analysis was performed using the R package glmnet with a 10-fold cross-validation. Log-rank was used in the KM survival analysis, KM curves were used to compare the survival differences between the two subgroups, and timeROC was performed to compare the predictive accuracy and risk scores of the signature genes. A one-way Cox analysis was performed for the signature genes, and forest plots were drawn using the R package forestplot to show the P-value, HR, and 95% CI for each gene; the significant genes were selected as key genes. The expression of the pivotal genes was verified in the TCGA database using the R package ggplot2 to observe their differential expression in cutaneous melanoma tissues and paracancerous tissues.

## Sample source and ethical review

The samples used in this study were obtained from Hangzhou Third People's Hospital, and three pairs of carcinoma and adjacent samples were collected from three patients with cutaneous melanoma. This study was reviewed and approved by the Medical Ethics Review Committee of Hangzhou Third People's Hospital (Approval No.: 2022KA032). The review process follows international ethical guidelines and relevant domestic laws and regulations.

## Quantitative reverse transcription-polymerase chain reaction

Three pairs of fresh cutaneous melanoma samples were collected for the external validation of the key genes. Total RNA was extracted from fresh cutaneous melanoma tissues using a Trizol kit (Beyotime), and cDNA was synthesized using a high-capacity cDNA reverse transcription kit (Thermo Fisher Scientific). The following primer sequences were used, with GAPDH as the endogenous control: phosphoglycerate translocase 5 (PGAM5) (forward primer: TCGTCCATTCGTCTATGACGC; reverse primer: GGCTTCCAATGAGACACGG); GAPDH (forward primer: CTGGGCTACACTGAGCACC; reverse primer: AAGTGGTCGTTGAGGGCAATG). The result is calculated by $2^{-\Delta\Delta ct}$.

## Statistical analysis

Data were analyzed using the SPSS software (V 17.0) and the prism software (V 6.01). A t-test was used to compare the differences between the two groups, and $P<0.05$ indicated statistical significance.

# Results

## Consensus clustering

Subgroup typing of cutaneous melanoma based on the set of necroptosis-related functional genes was performed through consensus clustering using similarity features, and the relative change in the area under the CDF curve was most significant when $k$=3 (Figures 1A, B). The samples were divided into three subgroups, each containing a different number of samples: C1 (86 cases), C2 (232 cases), and C3 (152 cases) (Figure 1C). To understand the differences between the three subgroups in terms of survival, we compared the prognosis based on the KM survival curves (Figure 1D). A significant difference in prognosis was observed, and the most significant difference was found between C1 and C3.

## Necrotizing apoptosis-related genes with prognostic significance

Of the 33 necrotizing apoptosis-related genes, 28 genes were significantly differentially expressed between the C1 and C3 subgroups (Figures 2A, B). Subsequent gene-wide survival analysis of cutaneous melanoma showed that a total of 3066 genes were associated with prognosis (Figure 2C) (see Supplementary Table 1 for details). When the differentially expressed necroapoptosis-related genes and the genes with prognostic significance for cutaneous melanoma were graphically represented using Venn diagrams, 16 overlapping genes were obtained (Figure 2D), and these genes were associated with both prognostic and necroapoptotic functions.

## PGAM5 as the pivotal gene

LASSO-based dimensionality reduction and prognostic model construction were employed to further analyze the pivotal genes associated with necroptosis in cutaneous melanoma (Figures 3A, B). The minimum value for the independent variable λ was 0.0285, and the risk score

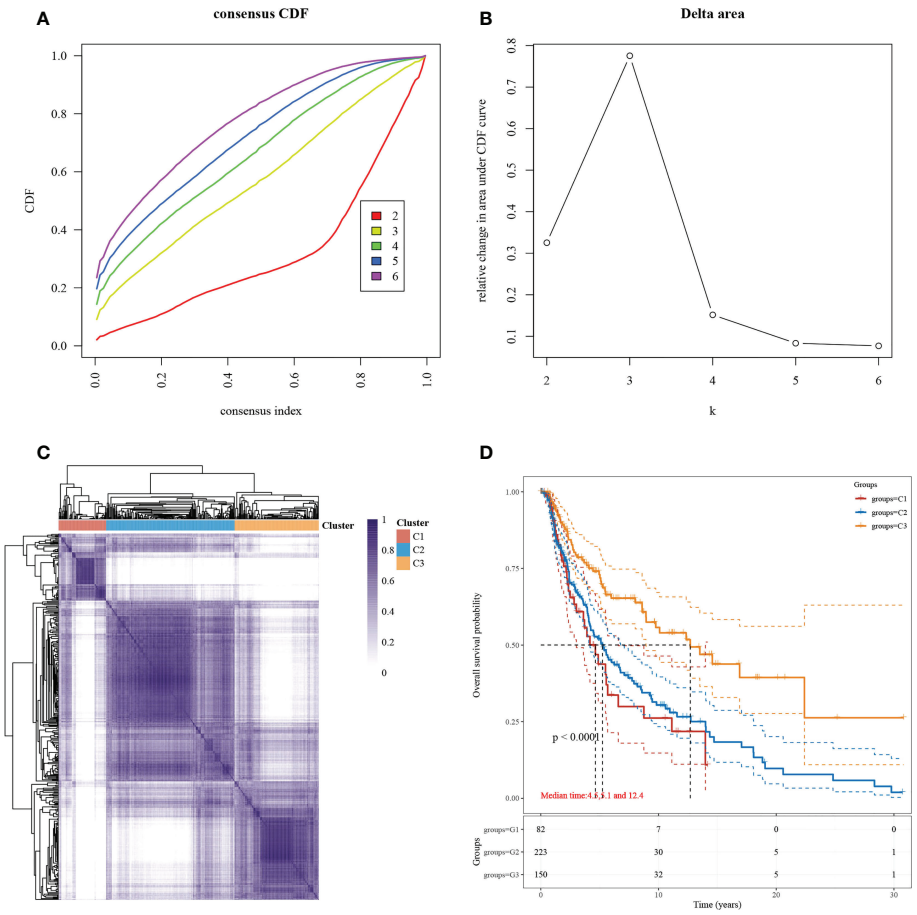



FIGURE 1
Subgroup typing based on necroptosis-associated gene sets. **(A)** CDF curve. Delta area curve for the consistent clustering and the area under the curve when samples were classified into different subtypes. **(B)** CDF delta area curve. The horizontal axis indicates a category number *k*, and the vertical axis indicates the relative change in area under the CDF curve. **(C)** A heat map showing consistent clustering results. Different colors indicate different subgroups. **(D)** KM survival curve distribution, indicating the different subgroups between prognosis comparisons between different subgroups.

calculation formula was as follows: Riskscore=(-0.1351)*ZBP1 +(-0.1491)*MLKL+(-0.2293)*NLRP6+(0.1086)*BOK+(-0.0261) *CASP8+(-0.004)*FAS+(0.1684)*PPIF+(-0.0248)*TLR3 +(0.0672)*PGAM5.

The samples were divided into high-risk and low-risk groups according to risk score values (Figure 3C), and their survival times and survival status are presented in scatter plots (Figure 3D). The risk score value of each sample is calculated according to the risk scoring formula, and the samples are divided into high risk group and low risk group according to the median value. The KM survival curves demonstrated that survival was higher and the median survival time was longer in the low-risk score group than in the high-risk score group (Figure 3E). The ROC curves demonstrated the strong prognostic predictive power of the risk score model, with AUC values of 0.695, 0.68, and 0.723 at 1, 3, and 5 years, respectively (Figure 3F). A one-way Cox regression analysis of the genetic characteristics selected for the model was performed, and PGAM5 was selected for the follow-up study based on the results (Figure 3G).

## Association between high PGAM5 expression and poor prognosis in cutaneous melanoma

To understand the expression and prognostic relationship between PGAM5 and cutaneous melanoma, we conducted an analysis using the TGCA database and GTEX database. The results showed that PGAM5 was differentially expressed in cutaneous melanoma tissues and paraneoplastic tissues and was highly expressed in tumor tissues (Figure 4A). PGAM5 expression in cutaneous melanoma was assessed by qRT-PCR, and the results showed that PGAM5 was overexpressed in

10.3389/fonc.2022.1004511

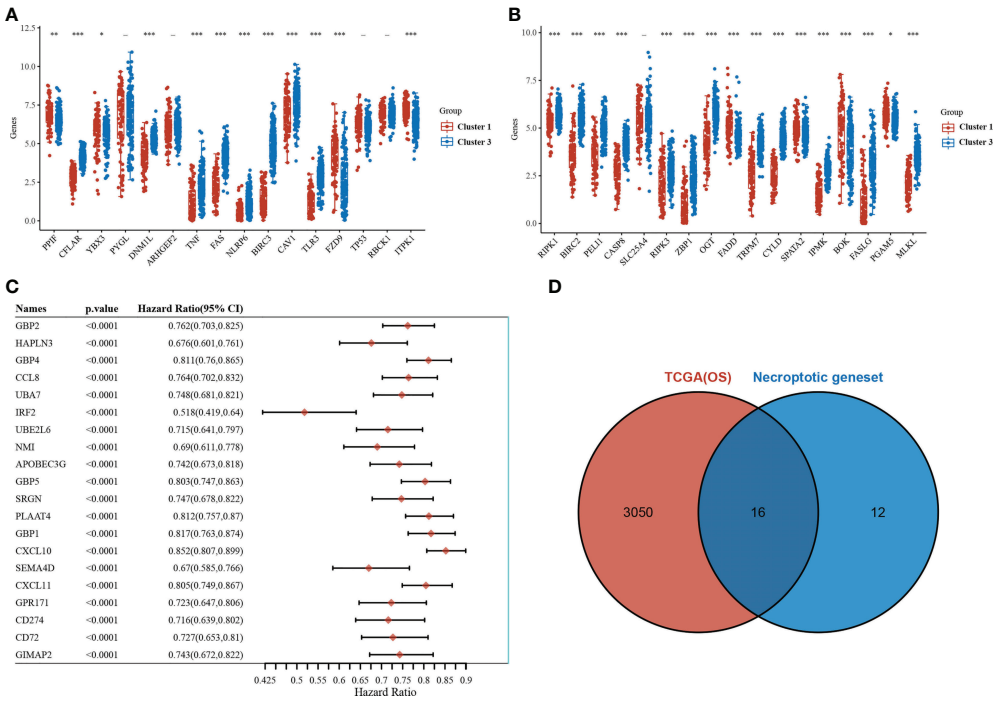



FIGURE 2
Necroptosis-related genes associated with prognosis. **(A, B)** Differential expression of necroptosis-related genes between the C1 and C3 subgroups. **(C)** Results of the whole genome survival analysis of cutaneous melanoma showing the top 20 genes. **(D)** Venn diagram. *p<0.05, **p<0.01, ***p<0.001.

cutaneous melanoma tissues compared with its expression in normal tissues (Figure 4B). The KM curves showed that high PGAM5 expression was associated with poor prognosis (Figure 4C); also, it showed the differences of the pTNM stages in terms of survival, wherein the survival rate decreased with higher disease stage (Figure 4D).

## Discussion

Unlike apoptosis and cell necrosis, necroptosis is a regulated form of necrotic cell death (18). When cells are unable to enter the apoptotic pathway, the necroptosis pathway can be activated by the stimulation of the proinflammatory cytokine tumor necrosis factor-alpha (TNF-α) (19). The loss of function of necroptosis-related genes in cutaneous melanoma mediates tumor immune escape (20), and certain compounds that effectively inhibit the proliferation of cutaneous melanoma cells in *in vitro* experiments act through the very pathway of necroptosis (21). RIPK1, a regulator of necroptosis, is known to be activated by its interaction with RIPK3, promoting the necroptotic process (22). Several genes have been identified to be involved in the necroapoptotic process; for example, RIPK3 deletion has been shown to predict the necroapoptosis resistance process in malignant melanoma (23). In this study, we conducted a series of bioinformatics analyses based on 33 human-associated necroapoptosis-related genes, aiming to identify the biomarkers related to the prognosis of cutaneous melanoma and to explore their biological functions.

Melanoma is the most aggressive type of skin cancer, and its progression and survival depend on mechanisms that favor survival, such as manipulation of cell death pathways and immune evasion by cancer cell lines (24). Necroptosis can sensitize tumor cells to anticancer drugs, and induction of this process is beneficial for killing tumor cells and is a promising tool for cancer therapy (25). Based on the expression level of the necroptosis-related gene set and the prognostic significance between this gene set and cutaneous melanoma, PGAM5 was selected as a key gene in this study. PGAM5 is a protein phosphatase that is localized in the mitochondria through its N terminus (26, 27). PGAM5 is involved in cellular activities related to the control of signal transduction pathways (28), and it can participate in apoptotic and necroptotic pathways by inducing mitochondrial autophagy after mitochondrial damage

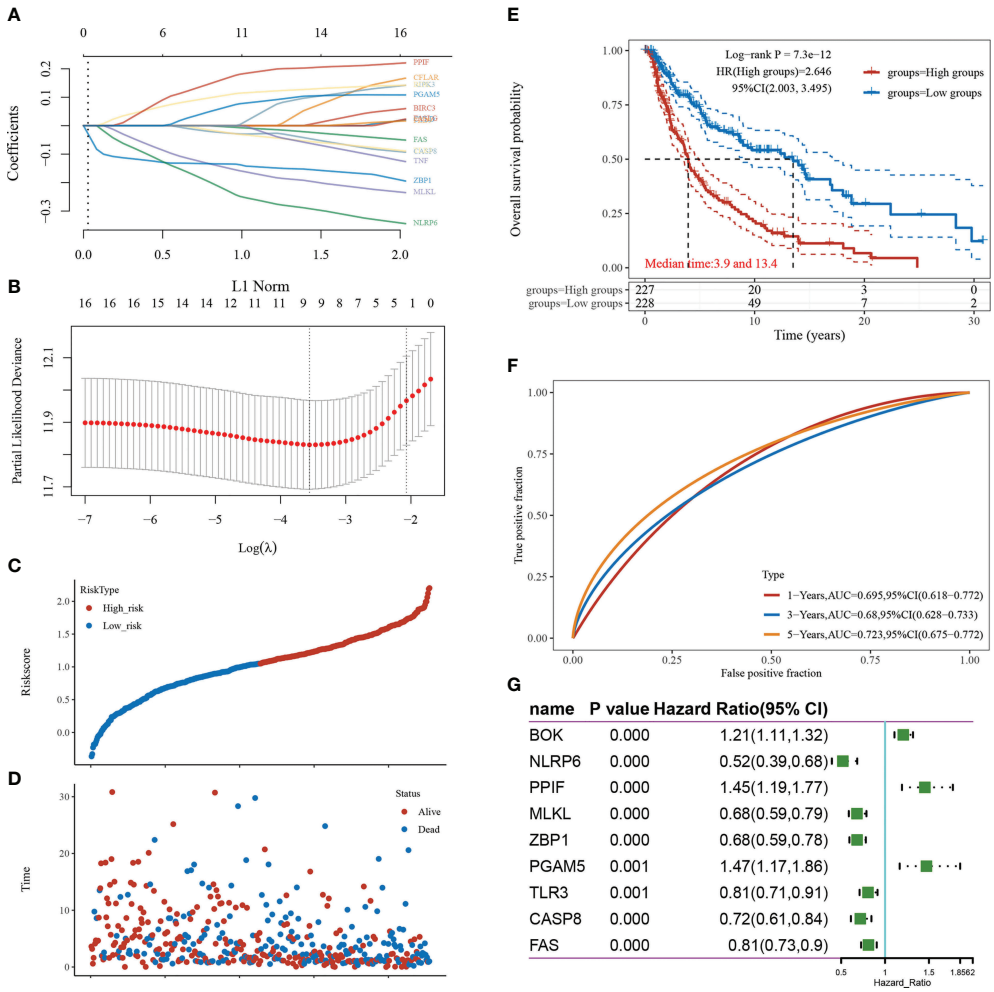



FIGURE 3
Screening of pivotal genes. **(A)** The coefficients for the selected features are presented as the λ parameter. **(B)** Partial likelihood deviation plotted against log(λ) using the LASSO-Cox regression model. **(C)** Scatter plot of risk scores from low to high. **(D)** Scatter plot distribution of survival time and survival status corresponding to the risk scores of different samples. **(E)** KM survival curve plot based on the high and low grouping of risk scores, excluding some of the missing values containing samples with missing values. **(F)** ROC curve. **(G)** Forest plot showing the P-value, HR, and 95% CI of each gene.

(29). The lack of translocase activity leads to the involvement of PGAM5 in the regulation of mitochondrial dynamics and programmed cell death, usually through protein–protein interactions and through a specific Ser/Thr/His protein phosphatase activity (30). Evidence has shown that PGAM5 expression is associated with necroptosis (31, 32). Research on the mechanism of necroptosis has increased in the recent years, and PGAM5 has been found to be involved in some cancers. For example, both the mRNA and protein expression of PGAM5 were much higher in lung cancer tissues than in normal tissues (33). Moreover, PGAM5 depletion inhibits hepatocellular carcinoma cell growth and promotes apoptosis (34).

In cancer, necroptosis is a double-edged sword. For some types of cancer, necroptosis occurs when apoptosis fails, thereby preventing further tumor progression; however, necroptosis can also trigger an inflammatory response that promotes cancer metastasis and immunosuppression (35, 36). Necroptosis is usually regulated by three pro-necrotic molecules, namely, RIPK1, RIPK3, and MLKL, and the use of pro-necrotic molecules for gene expression, for the analysis of prognostic impact, and for the evaluation of responsiveness to anticancer therapy has become a research hotspot in recent years (37). No direct evidence has ever established a link between PGAM5 and cutaneous melanoma. In this study, we found for the first time that PGAM5 was

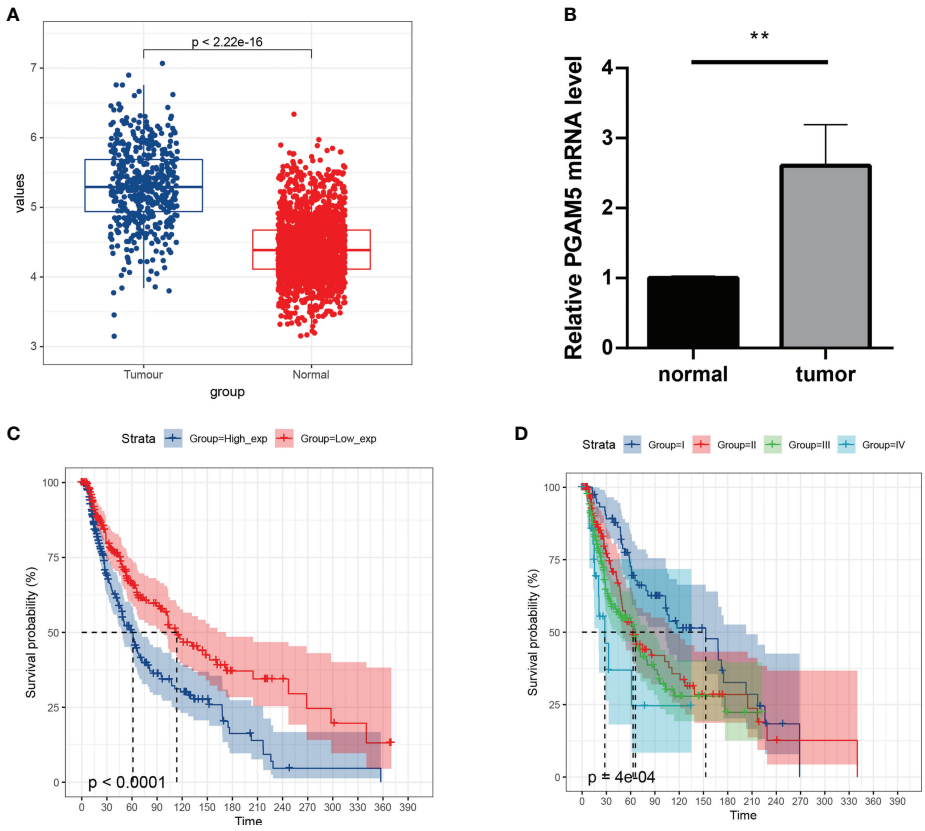



FIGURE 4
Expression and prognosis of pivotal genes. **(A)** Distribution of PGAM5 expression in cancer and paraneoplastic tissues. **(B)** Expression level of PGAM5 as determined by qRT-PCR. **(C)** KM survival curve showing the prognosis of high and low PGAM5 expression. **(D)** Prognosis of PGAM5 expression in relation to pTNM staging. **P<0.01.

overexpressed in cutaneous melanoma tissues compared with its expression levels in non-tumor tissues (P<0.001) and that PGAM5 leads to a poor prognosis. In other types of tumors or diseases, PGAM5 is used as a mediator of signaling pathways, such as the RIP1/RIP3/PGAM5 pathway in breast cancer (38), the PGAM5-CypD pathway in prolactinoma (39), and the Ripk3/Pgam5 signaling pathway in septic cardiomyopathy, to induce necroptosis (40). The results of this study showed that PGAM5 was overexpressed in cutaneous melanoma and that it played an active role in tumor progression. This implies that overexpression of PGAM5, a necroptosis-associated gene, promotes cutaneous melanoma progression.

In conclusion, this study used bioinformatics analysis and *in vitro* experiments to identify the prognosis-related biomarkers for cutaneous melanoma. The results showed that PGAM5 was highly expressed in cutaneous melanoma and that it led to a poor prognosis.

## Data availability statement

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

## Author contributions

We contributed equally for this work. All authors contributed to the article and approved the submitted version.

## Funding

Hangzhou medical key discipline construction project (No [2021]21-3).

## 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.2022.1004511/full#supplementary-material

## References

1. Kawakami A, Fisher DE. The master role of microphthalmia-associated transcription factor in melanocyte and melanoma biology. *Lab Invest* (2017) 97 (6):649–56. doi: 10.1038/labinvest.2017.9

2. Matthews NH, Li W-Q, Qureshi AA, Weinstock MA, Cho E. *Epidemiology of melanoma*. Exon Publications (2017) Epidemiology of melanoma. Exon Publications p. 3–22.

3. Curado M-P, Edwards B, Shin HR, Storm H, Ferlay J, Heanue M, et al. *Cancer incidence in five continents, volume IX*. IARC Press, International Agency for Research on Cancer (2007). Cancer incidence in five continents, Volume IX. IARC Press, International Agency for Research on Cancer

4. Walker GJ, Hayward NK. Pathways to melanoma development: lessons from the mouse. *J Invest Dermatol* (2002) 119(4):783–92. doi: 10.1046/j.1523-1747.2002.00217.x

5. 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

6. Leonardi GC, Candido S, Falzone L, Spandidos DA, Libra M. Cutaneous melanoma and the immunotherapy revolution. *Int J Oncol* (2020) 57(3):609–18. doi: 10.3892/ijo.2020.5088

7. Falzone L, Marconi A, Loreto C, Franco S, Spandidos DA, Libra M. Occupational exposure to carcinogens: Benzene, pesticides and fibers. *Mol Med Rep* (2016) 14(5):4467–74. doi: 10.3892/mmr.2016.5791

8. Hawkes JE, Truong A, Meyer LJ. Genetic predisposition to melanoma. In: *Semin Oncol*. (2016) 43(5):591-597. doi: 10.1053/j.seminoncol.2016.08.003

9. Miller R, Walker S, Shui I, Brandtmüller A, Cadwell K, Scherrer E. Epidemiology and survival outcomes in stages II and III cutaneous melanoma: A systematic review. *Melanoma Manage* (2020) 7(1):MMT39. doi: 10.2217/mmt-2019-0022

10. Ross MI, Gershenwald JE. Evidence-based treatment of early-stage melanoma. *J Surg Oncol* (2011) 104(4):341–53. doi: 10.1002/jso.21962

11. Greenhaw BN, Covington KR, Kurley SJ, Yeniay Y, Cao NA, Plasseraud KM, et al. Molecular risk prediction in cutaneous melanoma: a meta-analysis of the 31-gene expression profile prognostic test in 1,479 patients. *J Am Acad Dermatol* (2020) 83(3):745–53. doi: 10.1016/j.jaad.2020.03.053

12. Guo W, Zhu L, Zhu R, Chen Q, Wang Q, Chen J-Q. A four-DNA methylation biomarker is a superior predictor of survival of patients with cutaneous melanoma. *Elife* (2019) 8:e44310. doi: 10.7554/eLife.44310

13. Goel N, Ward WH, Jian QY, Farma JM. *Short-term and long-term management of melanoma*. Exon Publications (2017) p. 161–74.

14. Rausch MP, Hastings KT. *Immune checkpoint inhibitors in the treatment of melanoma: from basic science to clinical application*. Exon Publications (2017) p. 121–42.

15. Schwabe RF, Luedde T. Apoptosis and necroptosis in the liver: a matter of life and death. *Nat Rev Gastroenterol Hepatol* (2018) 15(12):738–52. doi: 10.1038/s41575-018-0065-y

16. Heckmann BL, Tummers B, Green DR. Crashing the computer: apoptosis vs. necroptosis in neuroinflammation. *Cell Death Differ* (2019) 26(1):41–52. doi: 10.1038/s41418-018-0195-3

17. Ashrafizadeh M, Mohammadinejad R, Tavakol S, Ahmadi Z, Roomiani S, Katebi M. Autophagy, anoikis, ferroptosis, necroptosis, and endoplasmic reticulum stress: Potential applications in melanoma therapy. *J Cell Physiol* (2019) 234 (11):19471–9. doi: 10.1002/jcp.28740

18. Chen Q, Kang J, Fu C. The independence of and associations among apoptosis, autophagy, and necrosis. *Signal Transduct Targeted Ther* (2018) 3 (1):1–11. doi: 10.1038/s41392-018-0018-5

19. Zhu K, Liang W, Ma Z, Xu D, Cao S, Lu X, et al. Necroptosis promotes cell-autonomous activation of proinflammatory cytokine gene expression. *Cell Death Dis* (2018) 9(5):1–16. doi: 10.1038/s41419-018-0524-y

20. Workenhe ST, Nguyen A, Bakhshinyan D, Wei J, Hare DN, MacNeill KL, et al. *De novo* necroptosis creates an inflammatory environment mediating tumor susceptibility to immune checkpoint inhibitors. *Commun Biol* (2020) 3(1):1–11. doi: 10.1038/s42003-020-01362-w

21. Liu N, Li Y, Chen G, Ge K. Evodiamine induces reactive oxygen species –dependent apoptosis and necroptosis in human melanoma A–375 cells. *Oncol Lett* (2020) 20(4):1–. doi: 10.3892/ol.2020.11983

22. Geng J, Ito Y, Shi L, Amin P, Chu J, Ouchida AT, et al. Regulation of RIPK1 activation by TAK1-mediated phosphorylation dictates apoptosis and necroptosis. *Nat Commun* (2017) 8(1):1–12. doi: 10.1038/s41467-017-00406-w

23. Geserick P, Wang J, Schilling R, Horn S, Harris P, Bertin J, et al. Absence of RIPK3 predicts necroptosis resistance in malignant melanoma. *Cell Death Dis* (2015) 6(9):e1884–e. doi: 10.1038/cddis.2015.240

24. Broussard L, Howland A, Ryu S, Song K, Norris D, Armstrong CA, et al. Melanoma cell death mechanisms. *Chonnam Med J* (2018) 54(3):135–42. doi: 10.4068/cmj.2018.54.3.135

25. Cho YS, Park HL. Exploitation of necroptosis for treatment of caspase –compromised cancers. *Oncol Lett* (2017) 14(2):1207–14. doi: 10.3892/ol.2017.6285

26. Yu B, Ma J, Li J, Wang D, Wang Z, Wang S. Mitochondrial phosphatase PGAM5 modulates cellular senescence by regulating mitochondrial dynamics. *Nat Commun* (2020) 11(1):1–17. doi: 10.1038/s41467-020-16312-7

27. Sugawara S, Kanamaru Y, Sekine S, Maekawa L, Takahashi A, Yamamoto T, et al. The mitochondrial protein PGAM5 suppresses energy consumption in brown adipocytes by repressing expression of uncoupling protein 1: PGAM5 negatively regulates brown adipocyte function. *J Biol Chem* (2020) 295(17):5588–601. doi: 10.1074/jbc.RA119.011508

28. Zhu Y, Gu L, Lin X, Liu C, Lu B, Cui K, et al. Dynamic regulation of ME1 phosphorylation and acetylation affects lipid metabolism and colorectal tumorigenesis. *Mol Cell* (2020) 77(1):138–49.e5. doi: 10.1016/j.molcel.2019.10.015

29. Bernkopf DB, Jalal K, Brückner M, Knaup KX, Gentzel M, Schambony A, et al. Pgam5 released from damaged mitochondria induces mitochondrial biogenesis *via* wnt signaling. *J Cell Biol* (2018) 217(4):1383–94. doi: 10.1083/jcb.201708191

30. Cheng M, Lin N, Dong D, Ma J, Su J, Sun L. PGAM5: a crucial role in mitochondrial dynamics and programmed cell death. *Eur J Cell Biol* (2020) 100 (1):151144. doi: 10.1016/j.ejcb.2020.151144

31. Wang Y-S, Yu P, Wang Y, Zhang J, Hang W, Yin Z-X, et al. AMP-activated protein kinase protects against necroptosis *via* regulation of Keap1-PGAM5 complex. *Int J Cardiol* (2018) 259:153–62. doi: 10.1016/j.ijcard.2018.01.036

32. Dong X-H, Liu H, Zhang M-Z, Zhao P-X, Liu S, Hao Y, et al. Postconditioning with inhaled hydrogen attenuates skin ischemia/reperfusion injury through the RIP-MLKL-PGAM5/Drp1 necrotic pathway. *Am J Trans Res* (2019) 11(1):499.

33. Kwong FNK, Nicholson A, Pavlidis S, Adcock I, Chung K. PGAM5 expression and macrophage signatures in non-small cell lung cancer associated with chronic obstructive pulmonary disease (COPD). *BMC Cancer* (2018) 18(1):1–9. doi: 10.1186/s12885-018-5140-9

34. Cheng J, Qian D, Ding X, Song T, Cai M, Xie D, et al. High PGAM5 expression induces chemoresistance by enhancing bcl-xL-mediated anti-apoptotic signaling and predicts poor prognosis in hepatocellular carcinoma patients. *Cell Death Dis* (2018) 9(10):1–16. doi: 10.1038/s41419-018-1017-8

35. Gong Y, Fan Z, Luo G, Yang C, Huang Q, Fan K, et al. The role of necroptosis in cancer biology and therapy. *Mol Cancer* (2019) 18(1):1–17. doi: 10.1186/s12943-019-1029-8

36. Najafov A, Chen H, Yuan J. Necroptosis and cancer. *Trends Cancer* (2017) 3 (4):294–301. doi: 10.1016/j.trecan.2017.03.002

37. Sprooten J, De Wijngaert P, Vanmeerbeek I, Martin S, Vangheluwe P, Schlenner S, et al. Necroptosis in immuno-oncology and cancer immunotherapy. *Cells* (2020) 9(8):1823. doi: 10.3390/cells9081823

38. Li Y, Gong P, Kong C, Tian X. Bufalin engages in RIP1-dependent and ROS-dependent programmed necroptosis in breast cancer cells by targeting the RIP1/RIP3/PGAM5 pathway. *Anti-cancer Drugs* (2019) 30(7):706–13. doi: 10.1097/CAD.0000000000000770

39. Zhang S-L, Tang H-B, Hu J-T, Zang Z-L, Ding X, Li S, et al. PGAM5-CypD pathway is involved in bromocriptine-induced RIP3/MLKL-dependent necroptosis of prolactinoma cells. *Biomed Pharmacother* (2019) 111:638–48. doi: 10.1016/j.biopha.2018.12.128

40. Fu G, Wang B, He B, Feng M, Yu Y. LPS induces cardiomyocyte necroptosis through the Ripk3/Pgam5 signaling pathway. *J Receptors Signal Transduct* (2021) 41(1):32–7. doi: 10.1080/10799893.2020.1783682

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 02 December 2022
DOI 10.3389/fonc.2022.1014997

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Jun Wang,
Fourth Hospital of Hebei Medical University, China
Masayuki Chida,
Dokkyo Medical University, Japan

*CORRESPONDENCE
Jiajun Du
dujiajun@sdu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 09 August 2022
ACCEPTED 14 November 2022
PUBLISHED 02 December 2022

CITATION
Wang Y, Wang G, Zheng H, Liu J, Ma G, Huang G, Song Q and Du J (2022) Distinct gene mutation profiles among multiple and single primary lung adenocarcinoma.
*Front. Oncol.* 12:1014997.
doi: 10.3389/fonc.2022.1014997

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

# Distinct gene mutation profiles among multiple and single primary lung adenocarcinoma

Yadong Wang[1†], Guanghui Wang[1,2†], Haotian Zheng[1], Jichang Liu[1], Guoyuan Ma[2], Gemu Huang[3], Qingtao Song[3] and Jiajun Du[1,2*]

[1]Institute of Oncology, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, China, [2]Department of Thoracic Surgery, Shandong Provincial Hospital, Shandong University, Jinan, Shandong, China, [3]Research and Development Department, Amoy Diagnostics Co., LTD., Xiamen, Fujian, China

With the development of technologies, multiple primary lung cancer (MPLC) has been detected more frequently. Although large-scale genomics studies have made significant progress, the aberrant gene mutation in MPLC is largely unclear. In this study, 141 and 44 lesions from single and multiple primary lung adenocarcinoma (SP- and MP-LUAD) were analyzed. DNA and RNA were extracted from formalin-fixed, paraffin-embedded tumor tissue and sequenced by using the next-generation sequencing-based YuanSu450TM gene panel. We systematically analyzed the clinical features and gene mutations of these lesions, and found that there were six genes differently mutated in MP-LUAD and SP-LUAD lesions, including RBM10, CDK4, ATRX, NTRK1, PREX2, SS18. Data from the cBioPortal database indicated that mutation of these genes was related to some clinical characteristics, such as TMB, tumor type, et al. Besides, heterogeneity analysis suggested that different lesions could be tracked back to monophyletic relationships. We compared the mutation landscape of MP-LUAD and SP-LUAD and identified six differentially mutated genes (RBM10, CDK4, ATRX, NTRK1, PREX2, SS18), and certain SNV loci in TP53 and EGFR which might play key roles in lineage decomposition in multifocal samples. These findings may provide insight into personalized prognosis prediction and new therapies for MP-LUAD patients.

KEYWORDS

**lung adenocarcinoma, multiple primary lung cancer, multifocal lung cancer, mutation, next generation sequencing**

## Introduction

Lung cancer is still the first cause of oncological death. Lung cancer can be divided into two broad categories according to histology: small-cell lung cancer and non-small-cell lung cancer (NSCLC) which accounts for approximately 85%. subdivided into adenocarcinoma (LUAD; 60%), squamous cell carcinoma (LUSC; 30–35%), large cell carcinoma, and other rare tumors, including adenosquamous carcinoma and in recent years, the ratio of LUAD among NSCLC is rising (1).

With the development of technologies, especially high-resolution computed tomography (HRCT), and the conduction of early lung cancer screening, multiple primary lung cancer (MPLC) has been detected more frequently (2–4). In particularly, adenocarcinoma was shown to be the most common pathological type in MPLC (5).

MPLC is a unique type of lung cancer, defined by presence of at least two independent primary tumors. MPLC was divided into synchronous MPLC (sMPLC) and metachronous MPLC (mMPLC) according to the diagnosis interval (6–8). According to current research, surgical resection was still recommended as the first choice for certain MPLC (9–12). Nowadays, with improved research methods such as machine learning, some studies would combine CT findings and gene sequencing technologies to diagnose and distinguish multiple primary lung cancers from pulmonary metastasis (8, 13–16). Understanding the molecular determinants of MPLC is one of the critical challenges in oncology.

Cancer is a genetic disease. Rapid advancing in next-generation sequencing technology and The Cancer Genome Atlas (TCGA) have profiled and analyzed the molecular aberrations at the DNA, RNA levels (17, 18). Genome instability and mutation is one of the hallmarks of cancer (19, 20). The accumulation of somatic mutations in the DNA affected neoplastic transformation, including driver mutations, mutations that directly affect tumor growth, such as TP53, epidermal growth factor receptor (EGFR) or RAS, and passenger mutations, which do not directly impact the growth of the cancer cell (21, 22).

In lung adenocarcinoma, several large-scale genomics studies have analyzed the genomic mutational landscape and some important targets, including EGFR, ALK, RAS, TP53 were identified to be associated with overall survival or treatments (17, 23–26). Besides, the most common therapeutic targets are EGFR and BRAF mutations and ALK and ROS1 rearrangements (17). Despite this progression, the aberrant gene mutation in MPLC compared with single primary lung cancer (SPLC) is largely unclear.

In this study, we included 141 single primary lung adenocarcinoma (SP-LUAD) patients and 44 multiple primary lung adenocarcinoma (MP-LUAD) patients to analysis the mutational landscape and there was an apparently difference between these two groups. RBM10, CDK4, ATRK, NTRK1, PREX2 and SS18 were identified as the significantly differential gene. The relationship between these genes and clinical characters was conducted using cBioPortal database. The study of associated gene mutations in MPLC will provide new insight into the mechanism, potential therapeutic targets, promote the prognosis and survival in clinic.

## Material and methods

### Patient cohort description

Primary lung cancers were collected between January 2018 and December 2020 in our institution. After sample collection, surgical specimens and biopsy tissues were snap-frozen in liquid nitrogen within 30 minutes of resection. Genomic DNA was extracted from all included samples. 141 SP-LUAD and 44 MP-LUAD patients were included in this study. Among these MP-LUAD specimens, 73 lesions were analyzed, including 9 lesions from 3 three-primary cases, 46 lesions from 23 dual-primary cases and 18 lesions from the other 18 MP-LUAD cases.

MP-LUAD were defined according to Warren and Gate's criteria: 1) each tumor had to show definite features of malignancy; 2) each cancer had to be anatomically separate and distinct; 3) the possibility that one cancer was a recurrence or metastatic lesion of the first cancer had to be ruled out; and 4) the subsequent primary malignancies had to be present in either the same or different organs (27).

### Tumor processing and DNA extraction

Before DNA extraction, 4 μm formalin-fixed paraffin-embedded (FFPE) specimens were taken. Histopathological examination confirmed that the area of each specimen was greater than 1 cm2 and the tumor cell density was greater than 20%.

**Abbreviations:** NSCLC, non-small-cell lung cancer; LUAD, adenocarcinoma; LUSC, squamous cell carcinoma; HRCT, high-resolution computed tomography; MPLC, multiple primary lung cancer; SPLC, single primary lung cancer; TCGA, The Cancer Genome Atlas; EGFR, epidermal growth factor receptor; MP-LUAD, multiple primary lung adenocarcinoma; SP-LUAD, single primary lung adenocarcinoma; FFPE, formalin-fixed paraffin-embedded; CAP, College of American Pathologists; CLIA, Clinical Laboratory Improvement Amendments; SNV, single nucleotide variation; INDELS, short and long insertion/deletion; CNV, copy number variation; IGV, Integrative Genomics Viewer; TMB, Tumor mutation burden; GSEA, Gene Set Enrichment Analysis; TFD, time to the first discovery; AIS, adenocarcinoma in situ; MIA, microinvasive adenocarcinoma; RBP, RNA-binding protein; MAPK, mitogen-activated protein kinase; PI3K, phosphoinositide 3-kinase; RBM10, RNA-binding motif protein 10; CDK4, Cyclin-dependent kinase 4; ATRX, Alpha thalassemia/mental retardation syndrome X-linked chromatin remodeler; NTRK1, Neurotrophic receptor tyrosine kinase 1; PREX2, Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 2; PTEN, phosphatase and tensin homolog.

According to the manufacturer's instructions, 0.5-2 μg of cancer tissue DNA was extracted from 4 μm FFPE tumor samples using a DNA extraction kit (QIAamp DNA FFPE tissue kit).

## Library construction

Libraries were constructed using Roche's KAPA Hyper Prep Kit according to the manufacturer's instructions. This custom hybrid capture panel includes more than 23,660 individually synthesized 5'-DNA biotin-labeled 120 bp oligonucleotides to target the approximately 2.6 Mb human genome, which contains 7029 extra coding nucleotides of 468 cancer-related genes exons and selected introns of 39 genes that are frequently rearranged in solid tumors.

Hybridization capture was performed according to the protocol of "IDT Company xGen LOCKDOWN probe and reagent hybridization capture DNA library", and sequenced on Illumina Nextseq500 with an average coverage of 1000 times. According to this method, paired paired-end sequencing (2 × 150 bp) was obtained by OrigiMed (OrigiMed, College of American Pathologists (CAP) and Clinical Laboratory Improvement Amendments (CLIA) accredited laboratory, Shanghai, China) Comprehensive genomic profiling, including single nucleotide variation (SNV), short and long insertion/deletion (INDELS), copy number variation (CNV), gene rearrangements and gene fusions.

## Next-generation sequencing

DNAs of both FFPE tumor tissues and matched blood were obtained by using QIAamp DNA FFPE Tissue Kit and QIAamp DNA Blood Midi Kit (Qiagen, Hilden, Germany), respectively, and sequenced by using the next-generation sequencing-based YuanSu450TM gene panel of OrigiMed (Shanghai, China), from where the laboratory was certified by CAP and CLIA. The genes were captured and sequenced with a mean depth of 800× by using Illumina Nova (Illumina, Inc., CA).

## Mutation analysis

Genomic alterations were identified as following: SNVs were identified by MuTect (v1.7). Insertion-deletions (Indels) were identified by using PINDEL (V0.2.5). The functional impact of genomic alterations was annotated by SnpEff3.0. CNV regions were identified by Control FREEC (v9.7) with the following parameters: window = 50 000 and step = 10 000. Gene fusions were detected through an in house developed pipeline. Gene rearrangements were assessed by Integrative Genomics Viewer (IGV). Known somatic mutations in the Catalog of Somatic Mutations in Cancer (COSMIC) and known germline polymorphisms in the U.S. National Center for Biotechnology Information's Single Nucleotide Polymorphism Database (dbSNP) were not counted. Tumor mutation burden (TMB) was calculated by counting the coding somatic mutations, including SNVs and Indels, per megabase of the sequence examined in each patient. The waterfall chart was drawn using the R report "ComplexHeatmap (version=2.2.0)", and the difference in gene mutation frequency between single foci and multiple foci was compared using the fisher test.

## Phylogenetic analysis

The phylogenetic analysis was conducted using LICHeE as reported before (28). Briefly, the branch evolution of the tumor within each patient was inferred by comparing the list of mutations in each tumor region. Regions containing all mutations observed in another region are indicated as their ancestors. If no such region exists, putative precursors are inferred from a set of changes common to multiple regions. Regions that did not change were considered to be parallel branches, although alternative dendrograms could be formed by assuming that these regions are ancestors of regions with mutations.

## Data source

The mutated gene data was recruited from cBioPortal database (http://www.cbioportal.org), originally developed at Memorial Sloan Kettering Cancer Center, an open-access resource for the interactive exploration of multidimensional cancer genomics data sets (29). The database can provide visualization (the associations between genes and clinical characteristics), analysis and downloads of large-scale cancer genomics data sets (the different expressed genes data). The somatic mutation data was acquired from the Cancer Genome Atlas (TCGA) database (https://cancergenome.nih.gov/).

## Gene set enrichment analysis

Gene Set Enrichment Analysis (GSEA) was used to analyze the correlation between gene differential expression with cellular pathway (KEGG pathways) and cellular functions (GO-molecular function). "clusterProfiler", "org.Hs.eg.db", "enrichplot", "ggplot2" R packages were applied to perform GSEA and "maftools" for co-occurrence of gene mutations. Adjusted P value < 0.05 was considered as statistically significant.

## Statistical analysis

All tests were performed with the R environment version 4.0.2 (Vienna, Austria), SPSS 25.0 (Chicago, US) or GraphPad Prism 7.0 (San Diego, CA). Comparisons of clinical

characteristics between paired primary tumors and metastases were based on Student' s t test, Chi-square test and Fisher's exact test. The nonparametric Wilcoxon rank-sum test was applied for comparison of mutation counts and branch lengths. If not noted otherwise, the tests applied were two-sided. As per the convention, $p<0.05$ was considered statistically significant.

# Results

## Patient characteristics

There were 141 SP-LUAD patients and 44 MP-LUAD patients in the study. Table 1 listed all of the patients' information. There were 86 males and 99 females in this study, with an average age of 57.7 and 59.3 years for the SP- and MP-LUAD cohorts, respectively. Weight Loss was a statistically significant factor across all these factors, indicating that weight loss may be more common in MP-LUAD. Furthermore, each lesion in these cohorts had a different histologic class. In the MP-LUAD cohort, the ratios of lepidic, adenocarcinoma *in situ* (AIS), and microinvasive adenocarcinoma (MIA) were higher, while acinar, solid, papillary, and micropapillary were lower.

In Table 2, the right upper lobe lesion accounted for the majority of the cases. There appeared to be no link between smoking or drinking habits and these two types of patients (Table 3), and the amount of smoking packs (daily amount of cigarettes*years of smoking/20) was barely correlated with the largest lesion diameter of MP-LUAD (R square: 0.125) (Figure 1).

## Mutation profile of the SP-LUAD and MP-LUAD cohort

We aimed to analyzed the mutational landscape in the SP-LUAD and MP-LUAD cohort and identified the top 30 most frequently mutant genes in SP-LUAD cohort, including EGFR, TP53, LRP1B, FRS2, RBM10, MDM2, PIK3CA, RB1, KRAS, SPTA1, ERBB2, GNAS, NFKBIA, NKX2-1, FAT3, ATM, FOS, HDAC9, RET, SDHA, ALK, CTNNB1, EPHA3, GLI2, KMT2D, LRP2, RICTOR, SMAD4, TERT, FAM135B (Figure 2).

The detail of each altered gene was shown in Figure 2A. Apparently, the ratio of each mutated gene in MP-LUAD cohort was not consistent with those in SP-LUAD cohort (Figures 2B, C). Then we compared the frequency between these two cohorts, and six genes (RBM10, CDK4, ATRX, NTRK1, PREX2, SS18) were shown to have significantly different mutation frequencies ($p=0.0089$, 0.0469, 0.0191, 0.0456, 0.0456, 0.0375, respectively) (Figures 2D, E). The frequency of two cohorts was shown in Figures 2F, G. The MP-LUAD cohort had a higher mutation frequency.

The gene variant type was shown in Figure 3A, B. In the MP-LUAD cohort, the Splicing Site variant was the most common while in frame deletion in the SP- LUAD cohort. Figures 3C, D shows the co-occurrence of mutant genes. The co-occurrence signature in MP-LUAD differed from that in SP-LUAD. Only three genetic changes were found to be prevalent (Figure 3E). The co-mutated genes of the SP- and MP-LUAD cohorts in EGFR-, KRAS-, and TP53-mutant patients were almost completely different, as shown in Supplementary Table 1.

## The associations between the six different mutated genes and patients' clinical characteristics

To further investigate the characteristics of these six genes in the public database, we analyzed the data from the cBioPortal database and TCGA database, including the genetic alteration ratio(Figure 4A), gene location on chromosomes(Figure 4B), the CNV alteration frequency(Figure 4C) and the clinical correlations (Figure 4D).The investigation of CNV alteration frequency showed most were focused on the amplification in copy number, among which the NTRK1 and CDK4 had a higher frequency(Figure 4C). The parameters (Histology, Sex, Tumor type, TMB, Mutation Count, Age) were included (Figure 4D and Supplementary Figure 1) and the results suggested that TMB ($p<0.001$) and mutation count ($p<0.001$) in six-gene altered group was much higher. The single gene analysis of ATRX ($p<0.001$, $p<0.001$), PREX2($p<0.001$, $p<0.001$) and SS18 ($p<0.001$, $p<0.001$) got the similar results. SS18 was unique among the genes studied since it was the only one in which the mutation status was linked to Tumor Type ($p=0.0137$). There were differences in sex distribution between the mutated and non-mutated groups, but the differences between the two groups were not statistically significant ($p=0.397$, 0.628, 0.0866, 0.945, 0.602, 0.865). Apart from that, the mutations of PREX2 ($p=0.0156$) and SS18 ($p=0.0199$) were linked to Fraction Genome Alteration and Ragnum Hypoxia Score, respectively (Supplementary Figures 1C, D). However, we found that these genes were unrelated to the patients' overall survival (Supplementary Figure 2).

## Gene expression differences associated with six gene mutation.

To investigate the Gene expression differences associated with six gene mutation, data from cBioPortal database was recruited, and the volcano plot of differentially expressed genes was shown in Figure 5A. Besides, based on these differentially expressed genes, GSEA was performed using R tool. The most enriched KEGG pathways and GO-molecular function terms were summarized in Figures 5B–D and the analysis of single gene was shown in Supplementary Figures 3, 4.

10.3389/fonc.2022.1014997

TABLE 1 Clinical characteristics of all patients.

| Characteristics | SP-LUAD | | MP-LUAD | | p |
|---|---|---|---|---|---|
| | n. | Ratio | n. | Ratio | |
| **Gender** | | | | | |
| Male | 68 | 0.482 | 18 | 0.409 | 0.396 |
| Female | 73 | 0.518 | 26 | 0.591 | |
| **Age** | 57.7 ± 10.336 | | 59.32 ± 6.653 | | 0.3308 |
| **TFD** | | | | | |
| <6 | 104 | 0.738 | 30 | 0.682 | 0.470 |
| ≥6 | 37 | 0.262 | 14 | 0.318 | |
| **Symptom** | | | | | |
| Pos. | 50 | 0.355 | 13 | 0.295 | 0.470 |
| Neg. | 91 | 0.645 | 31 | 0.705 | |
| **Chest Discomfort** | | | | | |
| Pos. | 24 | 0.170 | 11 | 0.250 | 0.238 |
| Neg. | 117 | 0.830 | 33 | 0.750 | |
| **Weight Loss** | | | | | |
| Pos. | 6 | 0.043 | 6 | 0.136 | 0.027 |
| Neg. | 135 | 0.957 | 38 | 0.864 | |
| **History of Other Tumor** | | | | | |
| Pos. | 9 | 0.064 | 2 | 0.045 | 1 |
| Neg. | 132 | 0.936 | 42 | 0.955 | |
| **Chronic Disease** | | | | | |
| Pos. | 68 | 0.482 | 24 | 0.545 | 0.464 |
| Neg. | 73 | 0.518 | 20 | 0.455 | |
| **History of Surgery** | | | | | |
| Pos. | 52 | 0.369 | 13 | 0.295 | 0.374 |
| Neg. | 89 | 0.631 | 31 | 0.705 | |
| **History of Trauma** | | | | | |
| Pos. | 7 | 0.050 | 3 | 0.068 | 0.704 |
| Neg. | 134 | 0.950 | 41 | 0.932 | |
| **Allergic History** | | | | | |
| Pos. | 11 | 0.078 | 6 | 0.136 | 0.242 |
| Neg. | 130 | 0.922 | 38 | 0.864 | |
| **Marriage** | | | | | |
| Married | 139 | 0.986 | 41 | 0.932 | 0.088 |
| Widowed | 2 | 0.014 | 3 | 0.068 | |
| **Family History of Cancer** | | | | | |
| Pos. | 17 | 0.121 | 8 | 0.182 | 0.300 |
| Neg. | 124 | 0.879 | 36 | 0.818 | |
| **Family History of Lung Cancer** | | | | | |
| Pos. | 11 | 0.078 | 5 | 0.114 | 0.463 |
| Neg. | 130 | 0.922 | 39 | 0.886 | |

Pos., Positive; Neg., Negative; TFD, Time to the First Discovery.

10.3389/fonc.2022.1014997

TABLE 2 Pathological correlation analysis.

| Characteristics | SP-LUAD | | MP-LUAD | | p |
|---|---|---|---|---|---|
| | n. | Ratio | n. | Ratio | |
| **Length** | 2.45035 ± 1.39988 | | 2.17954 ± 1.2159 | | 0.252 |
| **Subtype** | | | | | |
| Acinar | 80 | 0.567 | 33 | 0.452 | <0.0001 |
| Solid | 14 | 0.099 | 1 | 0.014 | |
| Lepidic | 12 | 0.085 | 13 | 0.178 | |
| Papillary | 8 | 0.057 | 3 | 0.041 | |
| Micropapillary | 6 | 0.043 | 2 | 0.027 | |
| Invasive Mucinous | 4 | 0.028 | 0 | 0.000 | |
| Poorly Differentiated | 1 | 0.007 | 2 | 0.027 | |
| AIS | 1 | 0.007 | 4 | 0.055 | |
| IA | 2 | 0.014 | 9 | 0.123 | |
| MIA | 5 | 0.035 | 6 | 0.082 | |
| LUAD | 8 | 0.057 | 0 | 0.000 | |
| **Site** | | | | | |
| Right Upper Lobe | 51 | 0.362 | 31 | 0.425 | 0.085 |
| Right Middle Lobe | 8 | 0.057 | 10 | 0.137 | |
| Right Lower Lobe | 22 | 0.156 | 10 | 0.137 | |
| Left Upper Lobe | 27 | 0.191 | 14 | 0.192 | |
| Left Lower Lobe | 33 | 0.234 | 8 | 0.110 | |

## Heterogeneity of MP-LUAD and phylogenetic reconstruction of MP-LUAD

To investigate the heterogeneity of different lesions of MP-LUAD, we performed heterogeneity analysis. The percentage of mutated genes in each lesion was summarized in Figures 6, 7.

As we could see that distinct lesions from one patient were virtually entirely different. Only one case, who carried three lesions, shared one SNV, EGFR L858R, as shown in Figure 6C, while the others were completely distinct (Figures 6A, B). The phylogenetic trees of all two-lesion-cases listed in Figure 7 indicated that there were still some relationships among two lesions in most cases. The number in the circle denoted the number of SNVs that had been adjusted, while the number near the arrow signified the SNVs' contributions. Monophyletic connections might be traced back to several lesions. We evaluated the contribution value and

TABLE 3 Clinical correlation analysis of smoking and drinking.

| Characteristics | SP-LUAD | | MP-LUAD | | p |
|---|---|---|---|---|---|
| | n. | Ratio | n. | Ratio | |
| **Smoke** | | | | | |
| Ever | 13 | 0.092 | 3 | 0.068 | 0.668 |
| Smoker | 29 | 0.206 | 12 | 0.273 | |
| Never | 99 | 0.702 | 29 | 0.659 | |
| **Drink** | | | | | |
| Ever | 2 | 0.014 | 1 | 0.023 | 0.816 |
| Regular | 23 | 0.163 | 5 | 0.114 | |
| Sometimes | 26 | 0.184 | 8 | 0.182 | |
| Never | 90 | 0.638 | 30 | 0.682 | |

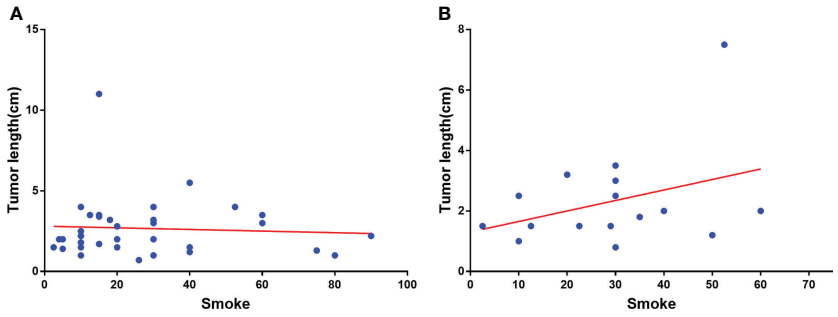



FIGURE 1
Correlation analysis of smoking amount with the largest lesion diameter of SP-**(A)** and MP-LUAD **(B)**.

amino acid changes associated with these genes and discovered that five TP53 and EGFR mutations were the most significant (Supplementary Table 2).

## Discussion

MPLC has become increasingly common as medical technology has advanced, but despite recent advances in large-scale genomics investigations, little is known about its gene signature, which remains a major problem of therapy failure and poor long-term survival.

In the present study, we conducted a mutation analysis and compared the difference between MP-LUAD and SP-LUAD. Obviously, the frequency of each mutated gene in MPL-LUAD cohort was not consistent with these in SP-LUAD cohort The related clinical characteristics was shown in Table 1. It seemed that while there were no significantly difference of age, chest discomfort and family history between these two groups, MP-LUAD patients were related to a higher ratio of weight loss (Table 1). One of the clinical characteristics called Time to the First Discovery (TFD) attracted us, which reflected the length of medical history. Although there was no statistical difference of TFD, probably limited by sample size, MP-LUAD patients showed little symptoms (cough, expectoration, and hemoptysis) which might reduce their willingness to further therapy and showed a longer TFD (Table 1), and this result might suggest that we had to pay more attention to TFD.

The amount of smoking was barely correlated with the diameter in MP-LUAD patients (Figure 1B). The TMB in the MP-LUAD cohort (3.381 ± 0.3586) was greater than the SP-LUAD cohort (2.434 ± 0.2615), according to our findings (Figure 2G).

Importantly, we found there were 6 genes with significantly different mutation frequency among these patients, including RBM10, CDK4, ATRX, NTRK1, PREX2, SS18. It was obviously that the mutation frequency of RBM10 in MP-LUAD cases was the highest and followed by CDK4, ATRX, NTRK1, PREX2, SS18.

RBM10 (RNA-binding motif protein 10), a member of the RNA-binding protein (RBP) family, is located at p11.3 on the X chromosome and also an alternative RNA splicing factor that participates in the regulation of gene expression (30). Recently, RBM10 was reported to function as an oncogene in LUAD by activating EGFR, mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase (PI3K)-AKT pathways and inhibition of apoptotic pathways (31), consist with the result in this study shown in Figure 5C and Supplementary Figure 3. Apart from that, Zhao, Jiawei et al. found that RBM10 mutations contributed to lung adenocarcinoma pathogenesis by deregulating splicing (32). Vinayanuwattikun, Chanida et al. revealed that the number of RBM10 mutations was higher in invasive lung adenocarcinoma (33), which suggested that RBM10 contributed to the LUAD progression and might also explain the higher mutation frequency of RBM10 in MP-LUAD.

CDK4 (Cyclin-dependent kinase 4) is a well-recognized cyclin-dependent kinase that specifically regulate cellular transition from the G1 phase to S phase of cell cycle together with CDK6 (34–36). As Figure 4A shows, the most frequent alteration of gene CDK4 is Amplification. CDK4 Amplification was seen in several tumors, such as head and neck mucosal melanoma (37), urinary bladder cancer (38), liposarcomas (39), melanoma (40) and lung cancer (41–43). Dysregulation of the cyclin D–CDK4/6–INK4–Rb pathway results in increased proliferation and due to the importance of CDK4/6 activity in cancer cells, CDK4/6 inhibitors seem as promising treatment (44). However, CDK4 amplification may reduce sensitivity to CDK4/6 inhibition in some cases (45). As Supplementary Figure 3 showed, mutation of gene CDK4 was related to "regulation of response to drug", and it was reported that in lung cancer, amplification of CDK4 was significant in *de novo* EGFR TKI resistance (43).

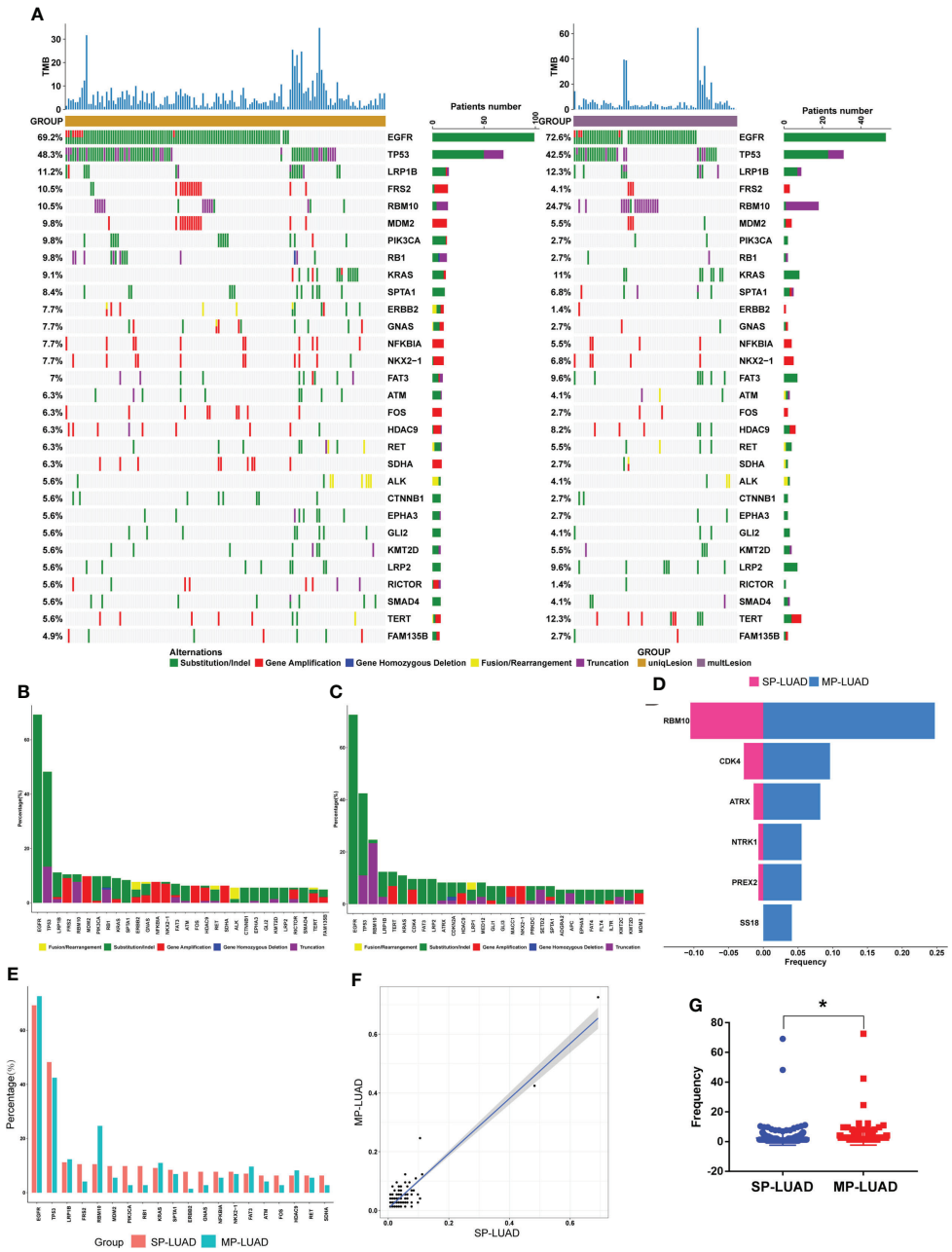



FIGURE 2
Comparation of SP- and MP-LUAD lesions. **(A)**The waterfall plot of tumor somatic mutation established based on the SP-LUAD (left) and MP-LUAD (right) cohort. **(B, C, E)** The difference of mutation frequency of common mutated genes between SP- and MP-LUAD lesions. **(D)** Six genes statistically different mutated among these two cohorts, including RBM10, CDK4, ATRX, NTRK1, PREX2, SS18. **(F, G)** Mutation Frequency of SP- and MP-LUAD lesions. * p<0.05.

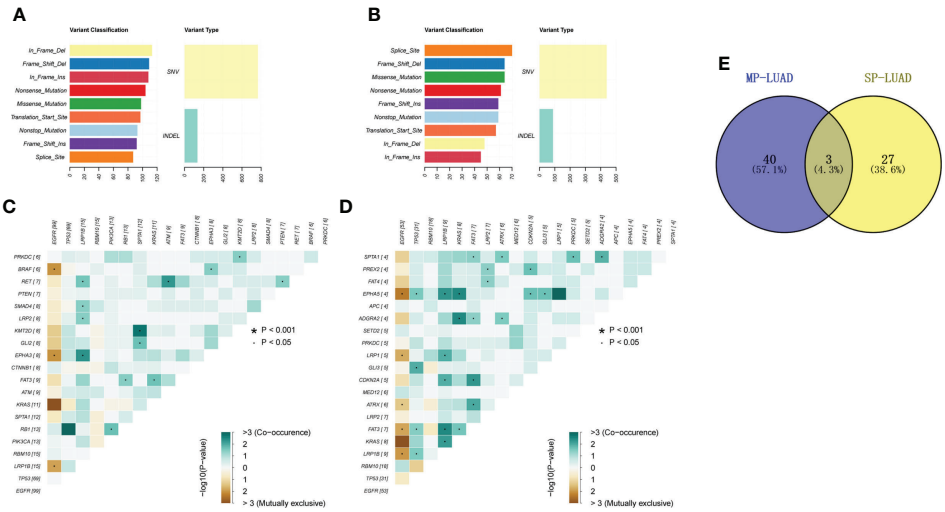



FIGURE 3
**(A, B)** Mutation types of SP-**(A)** and MP-LUAD **(B)** lesions. **(C, D)** Co-mutation analysis of SP-**(C)** and MP-LUAD **(D)** lesions. **(E)** The common co-mutated genes among SP- and MP-LUAD lesions. * p<0.05.

Gene ATRX (Alpha thalassemia/mental retardation syndrome X-linked chromatin remodeler) was first discovered in the X-linked mental retardation syndrome (ATRX syndrome) patients (46). And nowadays, its role in cancer is emerging. Araujo-Castro Marta et al. reported that ATRX mutation was linked to a shorter disease-specific survival (47). ATRX protein is one of the SWI/SNF2 (SWItch/Sucrose Non-Fermentable) family of chromatin remodeling proteins, and maintains genomic stability through its deposition of the replication-independent histone variant H3.3 at telomeres and pericentromeric heterochromatin (48). As shown in Figure 4A, Nonsense Mutation of ATRX was the most frequent, and ATRX loss has been shown to promote ALT, DNA damage and replicative stress (49–51). The GSEA analysis indicated that mutation of gene ATRX was related to "humoral immune response" (Supplementary Figure 3), and now it has been reported in many tumors including pleomorphic Sarcomas (52), glioma (53, 54), gastric cancer patients (55), while the relationship between mutation of ATRX with cytochrome P450 was not reported.

NTRK1 (Neurotrophic receptor tyrosine kinase 1) was originally identified as a fusion oncogene, trkA (tropomyosin receptor kinase) (56). The NTRK1 gene belongs to nerve growth factor receptor genes family, which mainly expressed in neuronal system (57). Several fusion partner genes of NTRK1 were reported in the past few years in thyroid cancer, glioblastoma and lung cancer (58). Many drugs have been developed for the treatment of NTRK1-rearanged cancers. However, it is worth noting that fusions of the NTRK1 genes with CD74 and MPRIP genes were identified in only 3% of some American patients and none of others (59). In this study, we identified two NTRK1 gene alterations including Gene Amplification (1.37% in MP-LUAD) and Substitution (0.7% and 4.11% in SP- and MP-LUAD, respectively) (Figures 2B, C) and it showed a highest gene alteration ratio and CNV frequency in the public database (Figures 4A, C). However, in the clinical correlation analysis, NTRK1 alteration was shown to be statistically unrelated to any clinical factor (Figure 4D). Mutations of gene NTRK1 were also reported in lung cancer (60, 61), while the probable mechanism how these mutations excluding Rearrangement mutation contributed to lung cancer progression was still unknown.

PREX2 (Phosphatidylinositol 3,4,5-trisphosphate-dependent Rac exchanger 2) is considered to be an oncogene for the PREX2 protein' inhibition of phosphatase and tensin homolog (PTEN) and thus activation of PI3K signaling pathway (62–64). The somatic mutation of PREX2 has been reported in several cancers including hepatocellular cancer (65), breast cancer (66), Melanoma (67, 68) and lung cancer (69). In the study of hepatocellular cancer, most mutant forms of PREX2, had an extended half-life compared with wild-type PREX2, and mutated PREX2 also promoted migration and activated the AKT pathway (65). In lung cancer, PREX2 played an important role in mediating the activation of PI3K/Akt signaling pathway (64), which might provide the evidence for the higher mutation frequency in MP-LUAD cases.

SS18 is the only gene that alternated in MP-LUAD (2 and 1 cases for Rearrangement and Amplification, respectively) rather than SP-LUAD. Fusion of SS18 was frequently detected in synovial sarcoma (70–73). In the GSEA analysis, SS18 was shown to be associated with nuclear division and cell adhesion molecules (Supplementary Figure 3) indicating that SS18 might

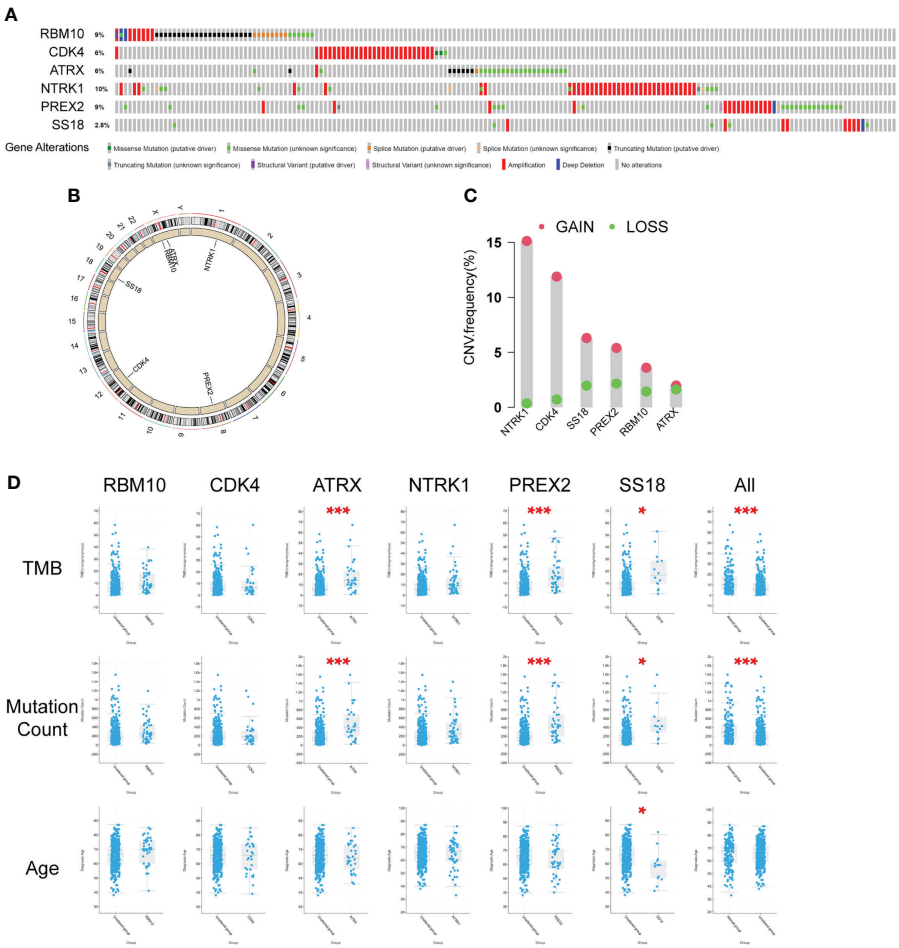



FIGURE 4
Characteristics of the six differently mutated genes. **(A)** Gene alterations obtained from cBioPortal database. **(B)** Gene location on chromosomes. **(C)** Gene CNV analysis using TCGA data. **(D)** Clinical correlation analysis of six genes using cBioPortal database. * p<0.05, *** p<0.001.

promote proliferation and metastasis of cancer cells. Knowledge about the actual oncogenic signals effected by SS18-fusion protein in lung cancer is still limited. In the study of synovial sarcoma, the SS18-SSX fusion protein would induce aberrant YAP/TAZ signals (71) and associated with SWI/SNF and Polycomb chromatin complexes to dysregulate gene expression (74–76), which might provide ideas for the aberrant alternation of SS18 in MP-LUAD.

Co-occurrence of genetic abnormalities was found to impact the response of lung cancer to several anticancer therapy (77). There were 40 co-occurring genomic changes in MP-LUAD patients and only three in common, as shown in Figure 3E and Supplementary Table 1. In the MP-LUAD cohort, the Splicing Site variation was the most common. The co-mutated genes in EGFR-, KRAS-, and TP53-mutant LUAD were virtually entirely different, as indicated in Supplementary Table 1. KRAS and LRP1B were all mutational exclusivity in EGFR-mutated LUAD patients, whereas genes like ATRX, EPHA5, and LRP1 were enriched in MP-LUAD patients. LRP1B was shown to be co-mutated in TP53- and KRAS-mutated MP-LUAD patients, which is comparable to a previous finding (78). Besides, LRP1B was reported to be associated with outcomes to immune checkpoint inhibitors, especially co-mutation of FAT3 and LRP1B in LUAD (79, 80).

Finally, we performed an analysis of the multiple lesions from MP-LUAD patients. According the results in Figures 6, 7, through comparing the contribution value and amino acid variations, we discovered that TP53 and EGFR were the most significantly different mutated genes that might contribute to the oncogenesis of MP-LUAD, and that certain SNV loci in TP53 (F143S, H193R) and EGFR (L747 P753 delinsS, S768 D770dup,

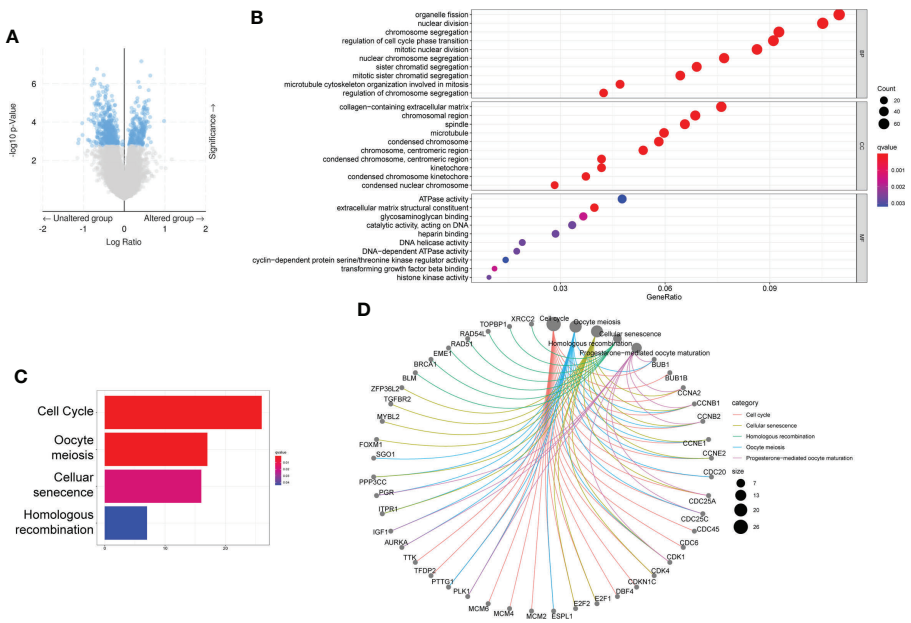



FIGURE 5
Differential expressed gene based on the mutation status of the six genes. **(A)** Volcano figure of differently expressed genes from cBioPortal database. **(B–D)** GO **(B)** and KEGG **(C, D)** analysis based on the differently expressed genes.

T790M) might play key roles in lineage decomposition in multifocal samples (Supplementary Table 2).

The present study has several limitations. First, the sample size is too small, which may lead us to neglect some factors; Second, not all the samples were fresh-frozen and we could not guarantee the purity of these samples even though with a strict pathologic evaluation; Third, some MP-LUAD patients who did not undergo surgical treatment were not included, which might lead to selection and information biases.

## Conclusion

In conclusion, we conducted a mutation analysis of SP- and MP-LUAD patients and identified genomic alterations and evolutionary trajectories underlying MP-LUAD. These findings will provide new insights into the oncogenesis of MP-LUAD and useful information for development novel approach to target MP-LUAD. Nevertheless, further research is required to elucidate the functions of these genes and their pathways.

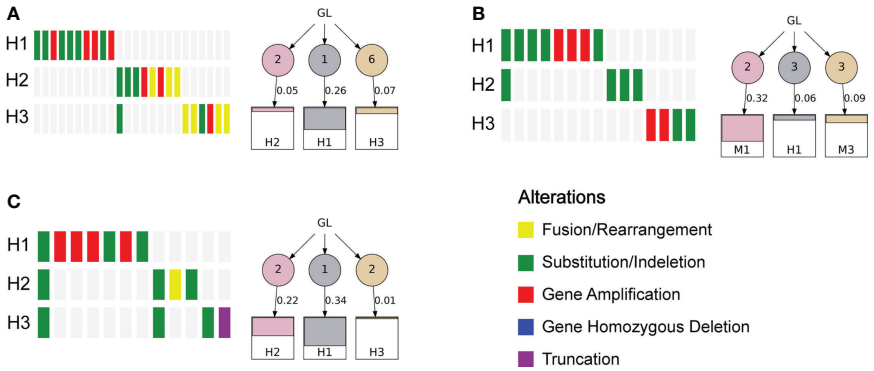



FIGURE 6
Gene alterations and polygenetic tree of the triple-primary LUAD lesions.

10.3389/fonc.2022.1014997

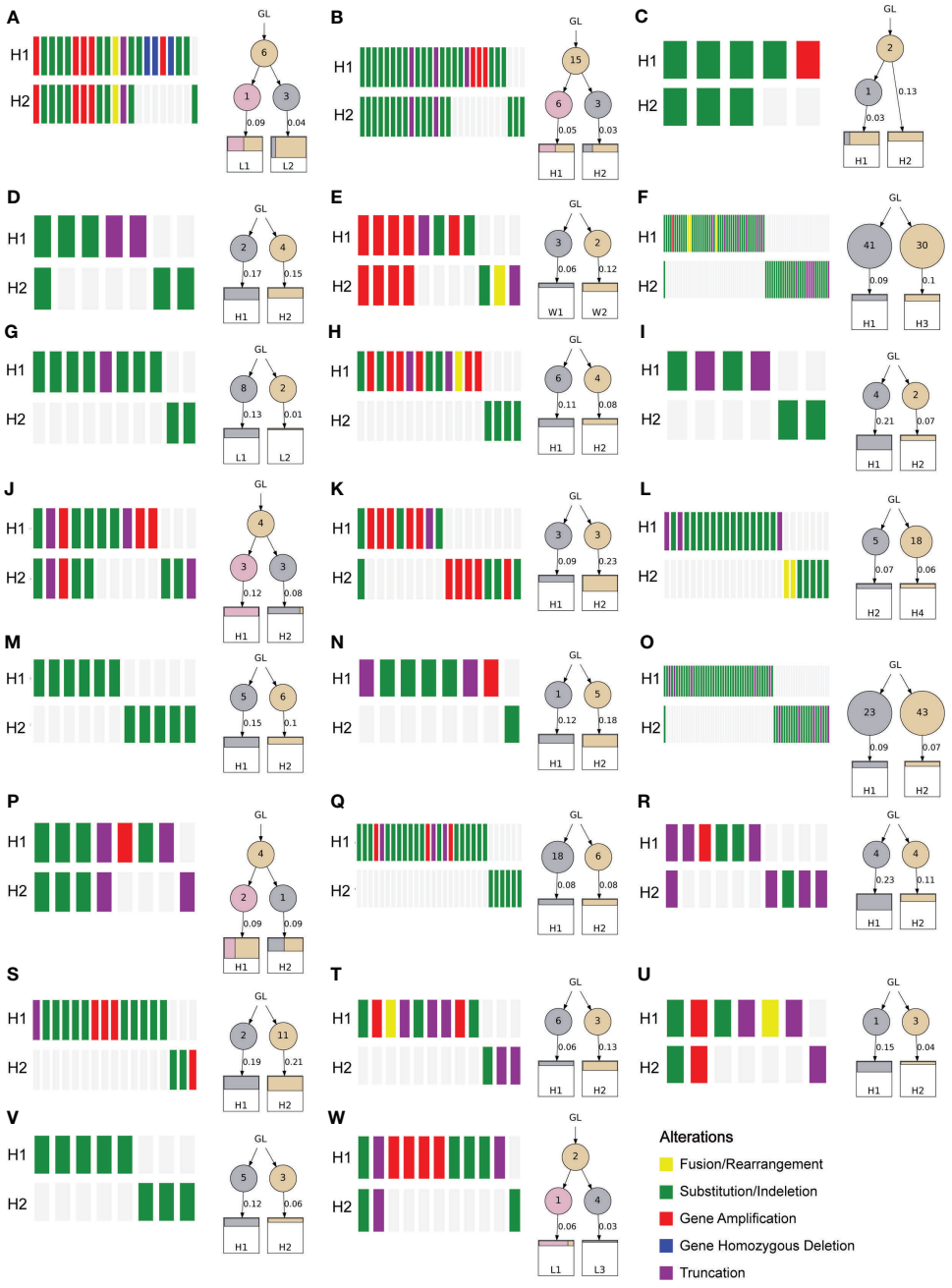



FIGURE 7
Gene alterations and polygenetic tree of the dual-primary LUAD lesions.

## Data availability statement

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

## Ethics statement

The studies involving human participants were reviewed and approved by The ethics committees of Shandong Provincial Hospital(No. 2022-159). The patients/participants provided their written informed consent to participate in this study.

## Author contributions

JD contributed to the designing and supervising the study, and correspondence. Co-authors YW and GW analyzed the data and completed the manuscript. HZ, JL, GM, GH, QS helped to collect specimens, search for references and offered guidelines of statistical methods. All authors contributed to the article and approved the submitted version.

## Funding

This work is supported by Shandong Provincial Natural Science Foundation (Grant No. ZR2021MH192 and ZR2020QH215).

## Acknowledgments

Thanks for everyone who helped to complete this study.

## Conflict of interest

Authors GH and QS are employed by Amoy Diagnostics Co., LTD.

The remaining 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.2022.1014997/full#supplementary-material

## References

1. Huang C, Qu X, Du J. Proportion of lung adenocarcinoma in female never-smokers has increased dramatically over the past 28 years. *J Thorac Dis* (2019) 11 (7):2685–8. doi: 10.21037/jtd.2019.07.08

2. Li M, Wan Y, Zhang L, Zhou L-N, Shi Z, Zhang R, et al. Synchronous multiple lung cancers presenting as multifocal pure ground glass nodules: are whole-body positron emission tomography/computed tomography and brain enhanced magnetic resonance imaging necessary? *Transl Lung Cancer Res* (2019) 8(5):649–57. doi: 10.21037/tlcr.2019.09.10

3. Tie H, Luo J, Shi R, Li Z, Chen D, Wu Q. Characteristics and prognosis of synchronous multiple primary lung cancer after surgical treatment: A systematic review and meta-analysis of current evidence. *Cancer Med* (2021) 10(2):507–20. doi: 10.1002/cam4.3614

4. Zhou LN, Wu N, Zhao SJ, Yang L, Wang YL, Wang RB, et al. [HRCT features differentiate synchronous multiple primary lung adenocarcinomas from intrapulmonary metastases]. *Zhonghua Zhong Liu Za Zhi* (2020) 42(6):449–55. doi: 10.3760/cma.j.cn112152-20200227-00126

5. Chen K, Chen W, Cai J, Yang F, Lou F, Wang X, et al. Favorable prognosis and high discrepancy of genetic features in surgical patients with multiple primary lung cancers. *J Thorac Cardiovasc Surg* (2018) 155(1):371–379. doi: 10.1016/j.jtcvs.2017.08.141

6. Pfannschmidt J, Dienemann H. Surgical treatment of oligometastatic non-small cell lung cancer. *Lung Cancer* (2010) 69(3):251–8. doi: 10.1016/j.lungcan.2010.05.003

7. Martini N, Melamed MR. Multiple primary lung cancers. *J Thorac Cardiovasc Surg* (1975) 70(4):606–12. doi: 10.1016/S0022-5223(19)40289-4

8. Wang J, Zhang T, Ding H, Dong G, Xu L, Jiang F. [Research progress in distinguishing methods of simultaneous multiple primary lung cancer and intrapulmonary metastasis]. *Zhongguo Fei Ai Za Zhi* (2021) 24(5):365–71. doi: 10.3779/j.issn.1009-3419.2021.101.12

9. Zuin A, Andriolo LG, Marulli G, Schiavon M, Nicotra S, Calabrese F, et al. Is lobectomy really more effective than sublobar resection in the surgical treatment of second primary lung cancer? *Eur J Cardiothorac Surg* (2013) 44(2):e120–5. doi: 10.1093/ejcts/ezt219

10. Xiao F, Liu D, Guo Y, Shi B, Song Z, Tian Y, et al. Survival rate and prognostic factors of surgically resected clinically synchronous multiple primary non-small cell lung cancer and further differentiation from intrapulmonary metastasis. *J Thorac Dis* (2017) 9(4):990–1001. doi: 10.21037/jtd.2017.03.59

11. Wang Y, Wang L, Zhang D. [Diagnosis and surgical treatment of double primary lung cancer]. *Zhonghua Zhong Liu Za Zhi* (2001) 23(5):428–30.

12. Nie Y, Wang X, Yang F, Zhou Z, Wang J, Chen K. Surgical prognosis of synchronous multiple primary lung cancer: Systematic review and meta-analysis. *Clin Lung Cancer* (2021) 22(4):341–350. doi: 10.1016/j.cllc.2020.10.022

13. Mansuet-Lupo A, Barritault M, Alifano M, Janet-Vendroux A, Zarmaev M, Biton J, et al. Proposal for a combined histomolecular algorithm to distinguish multiple primary adenocarcinomas from intrapulmonary metastasis in patients with multiple lung tumors. *J Thorac Oncol* (2019) 14(5):844–56. doi: 10.1016/j.jtho.2019.01.017

14. Ma J-W, Fu Y-L, Miao L, Li M. Imaging of synchronous multiple primary lung adenocarcinoma with concomitant and mutations: a case report and review of the literature. *Transl Cancer Res* (2021) 10(5):2509–15. doi: 10.21037/tcr-20-3258

15. Goodwin D, Rathi V, Conron M, Wright GM. Genomic and clinical significance of multiple primary lung cancers as determined by next-generation sequencing. *J Thorac Oncol* (2021) 16(7):1166–75. doi: 10.1016/j.jtho.2021.03.018

16. Belardinilli F, Pernazza A, Mahdavian Y, Cerbelli B, Bassi M, Gradilone A, et al. A multidisciplinary approach for the differential diagnosis between multiple primary lung adenocarcinomas and intrapulmonary metastases. *Pathol Res Pract* (2021) 220:153387. doi: 10.1016/j.prp.2021.153387

17. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* (2014) 511(7511):543–50. doi: 10.1038/nature13385

18. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* (2012) 489(7417):519–25. doi: 10.1038/nature11404

19. Hanahan D. Hallmarks of cancer: New dimensions. *Cancer Discovery* (2022) 12(1):31–46. doi: 10.1158/2159-8290.CD-21-1059

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

21. Kinzler KW, Vogelstein B. Landscaping the cancer terrain. *Science.* (1998) 280(5366):1036–7. doi: 10.1126/science.280.5366.1036

22. Wood LD, Parsons DW, Jones S, Lin J, Sjöblom T, Leary RJ, et al. The genomic landscapes of human breast and colorectal cancers. *Science.* (2007) 318 (5853):1108–13. doi: 10.1126/science.1145720

23. Chen J, Yang H, Teo ASM, Amer LB, Sherbaf FG, Tan CQ, et al. Genomic landscape of lung adenocarcinoma in East asians. *Nat Genet* (2020) 52(2):177–86. doi: 10.1038/s41588-019-0569-6

24. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell.* (2012) 150(6):1107–20. doi: 10.1016/j.cell.2012.08.029

25. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* (2016) 48(6):607–16. doi: 10.1038/ng.3564

26. Kadara H, Choi M, Zhang J, Parra ER, Rodriguez-Canales J, Gaffney SG, et al. Whole-exome sequencing and immune profiling of early-stage lung adenocarcinoma with fully annotated clinical follow-up. *Ann Oncol* (2017) 28 (1):75–82. doi: 10.1093/annonc/mdw436

27. Li F, Zhong W-Z, Niu F-Y, Zhao N, Yang J-J, Yan H-H, et al. Multiple primary malignancies involving lung cancer. *BMC Cancer* (2015) 15:696. doi: 10.1186/s12885-015-1733-8

28. Popic V, Salari R, Hajirasouliha I, Kashef-Haghighi D, West RB, Batzoglou S. Fast and scalable inference of multi-sample cancer lineages. *Genome Biol* (2015) 16:91. doi: 10.1186/s13059-015-0647-8

29. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* (2012) 2(5):401–4. doi: 10.1158/2159-8290.CD-12-0095

30. Sun Y, Bao Y, Han W, Song F, Shen X, Zhao J, et al. Autoregulation of RBM10 and cross-regulation of RBM10/RBM5 *via* alternative splicing-coupled nonsense-mediated decay. *Nucleic Acids Res* (2017) 45(14):8524–40. doi: 10.1093/nar/gkx508

31. Sun X, Jia M, Sun W, Feng L, Gu C, Wu T. Functional role of RBM10 in lung adenocarcinoma proliferation. *Int J Oncol* (2019) 54(2):467–78. doi: 10.3892/ijo.2018.4643

32. Zhao J, Sun Y, Huang Y, Song F, Huang Z, Bao Y, et al. Functional analysis reveals that RBM10 mutations contribute to lung adenocarcinoma pathogenesis by deregulating splicing. *Sci Rep* (2017) 7:40488. doi: 10.1038/srep40488

33. Vinayanuwattikun C, Le Calvez-Kelm F, Abedi-Ardekani B, Zaridze D, Mukeria A, Voegele C, et al. Elucidating genomic characteristics of lung cancer progression from *In situ* to invasive adenocarcinoma. *Sci Rep* (2016) 6:31628. doi: 10.1038/srep31628

34. Fassl A, Geng Y, Sicinski P. CDK4 and CDK6 kinases: From basic science to cancer therapy. *Science* (2022) 375(6577):eabc1495. doi: 10.1126/science.abc1495

35. Fry DW, Harvey PJ, Keller PR, Elliott WL, Meade M, Trachet E, et al. Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. *Mol Cancer Ther* (2004) 3 (11):1427–38. doi: 10.1158/1535-7163.1427.3.11

36. Goel S, DeCristo MJ, McAllister SS, Zhao JJ. CDK4/6 inhibition in cancer: Beyond cell cycle arrest. *Trends Cell Biol* (2018) 28(11):911–25. doi: 10.1016/j.tcb.2018.07.002

37. Lyu J, Miao Y, Yu F, Chang C, Guo W, Zhu H. CDK4 and TERT amplification in head and neck mucosal melanoma. *J Oral Pathol Med* (2021) 50 (10):971–8. doi: 10.1111/jop.13180

38. Simon R, Struckmann K, Schraml P, Wagner U, Forster T, Moch H, et al. Amplification pattern of 12q13-q15 genes (MDM2, CDK4, GLI) in urinary bladder cancer. *Oncogene.* (2002) 21(16):2476–83. doi: 10.1038/sj.onc.1205304

39. Italiano A, Bianchini L, Gjernes E, Keslair F, Ranchere-Vince D, Dumollard J-M, et al. Clinical and biological significance of CDK4 amplification in well-differentiated and dedifferentiated liposarcomas. *Clin Cancer Res* (2009) 15 (18):5696–703. doi: 10.1158/1078-0432.CCR-08-3185

40. Muthusamy V, Hobbs C, Nogueira C, Cordon-Cardo C, McKee PH, Chin L, et al. Amplification of CDK4 and MDM2 in malignant melanoma. *Genes Chromosomes Cancer* (2006) 45(5):447–54. doi: 10.1002/gcc.20310

41. Schrock AB, Frampton GM, Suh J, Chalmers ZR, Rosenzweig M, Erlich RL, et al. Characterization of 298 patients with lung cancer harboring MET exon 14 skipping alterations. *J Thorac Oncol* (2016) 11(9):1493–502. doi: 10.1016/j.jtho.2016.06.004

42. Yu HA, Suzawa K, Jordan E, Zehir A, Ni A, Kim R, et al. Concurrent alterations in EGFR-mutant lung cancers associated with resistance to EGFR kinase inhibitors and characterization of MTOR as a mediator of resistance. *Clin Cancer Res* (2018) 24(13):3108–18. doi: 10.1158/1078-0432.CCR-17-2961

43. Sitthideatphaiboon P, Teerapakpinyo C, Korphaisarn K, Leelayuwatanakul N, Pornpatrananrak N, Poungvarin N, et al. Co-Occurrence CDK4/6 amplification serves as biomarkers of *de novo* EGFR TKI resistance in sensitizing EGFR mutation non-small cell lung cancer. *Sci Rep* (2022) 12(1):2167. doi: 10.1038/s41598-022-06239-y

44. Hamilton E, Infante JR. Targeting CDK4/6 in patients with cancer. *Cancer Treat Rev* (2016) 45:129–38. doi: 10.1016/j.ctrv.2016.03.002

45. Olanich ME, Sun W, Hewitt SM, Abdullaev Z, Pack SD, Barr FG. CDK4 amplification reduces sensitivity to CDK4/6 inhibition in fusion-positive rhabdomyosarcoma. *Clin Cancer Res* (2015) 21(21):4947–59. doi: 10.1158/1078-0432.CCR-14-2955

46. Gibbons RJ, Picketts DJ, Villard L, Higgs DR. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell.* (1995) 80(6):837–45. doi: 10.1016/0092-8674(95)90287-2

47. Araujo-Castro M, Pascual-Corrales E, Molina-Cerrillo J, Moreno Mata N, Alonso-Gordoa T. Bronchial carcinoids: From molecular background to treatment approach. *Cancers (Basel)* (2022) 14(3):520. doi: 10.3390/cancers14030520

48. Pekmezci M, Rice T, Molinaro AM, Walsh KM, Decker PA, Hansen H, et al. Adult infiltrating gliomas with WHO 2016 integrated diagnosis: additional prognostic roles of ATRX and TERT. *Acta Neuropathol* (2017) 133(6):1001–16. doi: 10.1007/s00401-017-1690-1

49. Ramamoorthy M, Smith S. Loss of ATRX suppresses resolution of telomere cohesion to control recombination in ALT cancer cells. *Cancer Cell* (2015) 28 (3):357–69. doi: 10.1016/j.ccell.2015.08.003

50. Cai J, Chen J, Zhang W, Yang P, Zhang C, Li M, et al. Loss of ATRX, associated with DNA methylation pattern of chromosome end, impacted biological behaviors of astrocytic tumors. *Oncotarget.* (2015) 6(20):18105–15. doi: 10.18632/oncotarget.3906

51. Haase S, Garcia-Fabiani MB, Carney S, Altshuler D, Núñez FJ, Méndez FM, et al. Mutant ATRX: Uncovering a new therapeutic target for glioma. *Expert Opin Ther Targets* (2018) 22(7):599–613. doi: 10.1080/14728222.2018.1487953

52. Darmusey L, Pérot G, Thébault N, Le Guellec S, Desplat N, Gaston L, et al. ATRX alteration contributes to tumor growth and immune escape in pleomorphic sarcomas. *Cancers (Basel)* (2021) 13(9):2151. doi: 10.3390/cancers13092151

53. Hu C, Wang K, Damon C, Fu Y, Ma T, Kratz L, et al. ATRX loss promotes immunosuppressive mechanisms in IDH1 mutant glioma. *Neuro Oncol* (2021) 24 (6):888–900. doi: 10.1093/neuonc/noab292

54. Babikir H, Wang L, Shamardani K, Catalan F, Sudhir S, Aghi MK, et al. ATRX regulates glial identity and the tumor microenvironment in IDH-mutant glioma. *Genome Biol* (2021) 22(1):311. doi: 10.1186/s13059-021-02535-4

55. Ge Y, Wei F, Du G, Fei G, Li W, Li X, et al. The association of sex-biased ATRX mutation in female gastric cancer patients with enhanced immunotherapy-related anticancer immunity. *BMC Cancer* (2021) 21(1):240. doi: 10.1186/s12885-021-07978-3

56. Martin-Zanca D, Hughes SH, Barbacid M. A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences. *Nature.* (1986) 319(6056):743–8. doi: 10.1038/319743a0

57. Snider WD. Functions of the neurotrophins during nervous system development: what the knockouts are teaching us. *Cell.* (1994) 77(5):627–38. doi: 10.1016/0092-8674(94)90048-5

58. Amatu A, Sartore-Bianchi A, Siena S. Gene fusions as novel targets of cancer therapy across multiple tumour types. *ESMO Open* (2016) 1(2):e000023. doi: 10.1136/esmoopen-2015-000023

59. Vaishnavi A, Capelletti M, Le AT, Kako S, Butaney M, Ercan D, et al. Oncogenic and drug-sensitive NTRK1 rearrangements in lung cancer. *Nat Med* (2013) 19(11):1469–72. doi: 10.1038/nm.3352

60. Pu X, Zhu L, Li F, Zheng J, Wu H, Fu Y, et al. Target molecular treatment markers in intrahepatic cholangiocarcinoma based on Chinese population. *Pathol Res Pract* (2020) 216(9):153116. doi: 10.1016/j.prp.2020.153116

61. Li H, Yan S, Liu Y, Ma L, Liu X, Liu Y, et al. Analysis of NTRK mutation and clinicopathologic factors in lung cancer patients in northeast China. *Int J Biol Markers* (2020) 35(3):36–40. doi: 10.1177/1724600820949883

62. Rosenfeldt H, Vázquez-Prado J, Gutkind JS. P-REX2, a novel PI-3-kinase sensitive rac exchange factor. *FEBS Lett* (2004) 572(1-3):167–71. doi: 10.1016/j.febslet.2004.06.097

63. Pandiella A, Montero JC. Molecular pathways: P-Rex in cancer. *Clin Cancer Res* (2013) 19(17):4564–9. doi: 10.1158/1078-0432.CCR-12-1662

64. Yeung YT, Fan S, Lu B, Yin S, Yang S, Nie W, et al. CELF2 suppresses non-small cell lung carcinoma growth by inhibiting the PREX2-PTEN interaction. *Carcinogenesis.* (2020) 41(3):377–89. doi: 10.1093/carcin/bgz113

65. Yang M-H, Yen C-H, Chen Y-F, Fang C-C, Li C-H, Lee K-J, et al. Somatic mutations of PREX2 gene in patients with hepatocellular carcinoma. *Sci Rep* (2019) 9(1):2552. doi: 10.1038/s41598-018-36810-5

66. Srijakotre N, Man J, Ooms LM, Lucato CM, Ellisdon AM, Mitchell CA. P-Rex1 and p-Rex2 RacGEFs and cancer. *Biochem Soc Trans* (2017) 45(4):963–77. doi: 10.1042/BST20160269

67. Berger MF, Hodis E, Heffernan TP, Deribe YL, Lawrence MS, Protopopov A, et al. Melanoma genome sequencing reveals frequent PREX2 mutations. *Nature.* (2012) 485(7399):502–6. doi: 10.1038/nature11071

68. Horrigan SK, Courville P, Sampey D, Zhou F, Cai S. Replication study: Melanoma genome sequencing reveals frequent mutations. *Elife.* (2017) 6:e21634. doi: 10.7554/eLife.21634

69. Wang H, Shan Q, Guo J, Han X, Zhao C, Li H, et al. PDL1 high expression without TP53, KEAP1 and EPHA5 mutations could better predict survival for patients with NSCLC receiving atezolizumab. *Lung Cancer* (2021) 151:76–83. doi: 10.1016/j.lungcan.2020.11.006

70. Jerby-Arnon L, Neftel C, Shore ME, Weisman HR, Mathewson ND, McBride MJ, et al. Opposing immune and genetic mechanisms shape oncogenic programs in synovial sarcoma. *Nat Med* (2021) 27(2):289–300. doi: 10.1038/s41591-020-01212-6

71. Isfort I, Cyra M, Elges S, Kailayangiri S, Altvater B, Rossig C, et al. SS18-SSX-Dependent YAP/TAZ signaling in synovial sarcoma. *Clin Cancer Res* (2019) 25 (12):3718–31. doi: 10.1158/1078-0432.CCR-17-3553

72. Stacchiotti S, Van Tine BA. Synovial sarcoma: Current concepts and future perspectives. *J Clin Oncol* (2018) 36(2):180–7. doi: 10.1200/JCO.2017.75.1941

73. Li J, Mulvihill TS, Li L, Barrott JJ, Nelson ML, Wagner L, et al. A role for SMARCB1 in synovial sarcomagenesis reveals that SS18-SSX induces canonical BAF destruction. *Cancer Discovery* (2021) 11(10):2620–37. doi: 10.1158/2159-8290.CD-20-1219

74. Kadoch C, Crabtree GR. Reversible disruption of mSWI/SNF (BAF) complexes by the SS18-SSX oncogenic fusion in synovial sarcoma. *Cell.* (2013) 153(1):71–85. doi: 10.1016/j.cell.2013.02.036

75. Banito A, Li X, Laporte AN, Roe J-S, Sanchez-Vega F, Huang C-H, et al. The SS18-SSX oncoprotein hijacks KDM2B-PRC1.1 to drive synovial sarcoma. *Cancer Cell* (2018) 34(2):346–8. doi: 10.1016/j.ccell.2018.07.006

76. McBride MJ, Pulice JL, Beird HC, Ingram DR, D'Avino AR, Shern JF, et al. The SS18-SSX fusion oncoprotein hijacks BAF complex targeting and function to drive synovial sarcoma. *Cancer Cell* (2018) 33(6):1128–1141. doi: 10.1016/j.ccell.2018.05.002

77. Skoulidis F, Heymach JV. Co-Occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nat Rev Cancer* (2019) 19(9):495–509. doi: 10.1038/s41568-019-0179-8

78. Hu C, Zhao L, Liu W, Fan S, Liu J, Liu Y, et al. Genomic profiles and their associations with TMB, PD-L1 expression, and immune cell infiltration landscapes in synchronous multiple primary lung cancers. *J Immunother Cancer* (2021) 9(12):: e003773. doi: 10.1136/jitc-2021-003773

79. Brown LC, Tucker MD, Sedhom R, Schwartz EB, Zhu J, Kao C, et al. LRP1B mutations are associated with favorable outcomes to immune checkpoint inhibitors across multiple cancer types. *J Immunother Cancer* (2021) 9(3):e001792. doi: 10.1136/jitc-2020-001792

80. Zhu M, Zhang L, Cui H, Zhao Q, Wang H, Zhai B, et al. Co-Mutation of FAT3 and LRP1B in lung adenocarcinoma defines a unique subset correlated with the efficacy of immunotherapy. *Front Immunol* (2021) 12:800951. doi: 10.3389/fimmu.2021.800951

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 14 December 2022
DOI 10.3389/fonc.2022.1059591

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Hao Shen,
Huazhong University of Science and Technology, China
Meifang Xiao,
Hainan Maternal and Child Health Hospital, China

*CORRESPONDENCE
Xue Sun
hbsysxue@163.com

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

RECEIVED 01 October 2022
ACCEPTED 24 November 2022
PUBLISHED 14 December 2022

CITATION
Tian Y, Xie T and Sun X (2022) Analysis of the regulatory mechanisms of prognostic immune factors in thyroid cancer.
*Front. Oncol.* 12:1059591.
doi: 10.3389/fonc.2022.1059591

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

# Analysis of the regulatory mechanisms of prognostic immune factors in thyroid cancer

Yin Tian[1], Tao Xie[2] and Xue Sun[3]*

[1]Department of Pediatric Surgery, Jingzhou Central Hospital, Jingzhou Hospital Affiliated to Yangtze University, Jingzhou, Hubei, China, [2]Department of Anesthesiology, Jingzhou Central Hospital, Jingzhou Hospital Affiliated to Yangtze University, Jingzhou, Hubei, China, [3]Department of Clinical Nutrition, Sinopharm Dongfeng General Hospital, Hubei University of Medicine, Shiyan, Hubei, China

**Objective:** To explore the regulatory mechanism of immune prognostic factors in thyroid cancer.

**Methods:** Based on the TCGA database and GEO database, this study used bioinformatics methods to study the potential regulatory mechanism of thyroid cancer prognosis, analyzed the differentially expressed genes and differential miRNAs between thyroid cancer and normal paracancerous tissues by R software, and constructed lasso risk factors. The immune prognostic factors of thyroid cancer were obtained from the model, and the miRDB website was used to predict the possibility of differential miRNA target binding of the immune prognostic factors and correlation analysis was performed, and finally verified by cell experiments.

**Results:** There were 1413 differentially expressed genes between thyroid cancer and normal paracancerous tissues, among which 21 immune-related genes were prognostic factors with significant differences in expression; lasso risk model obtained AKAP12, APOC1, TIMP3, ADAMTS9, ANK2, HTRA3, SYNDIG1 , ADAMTS5 and DACT1 were nine prognostic factors. A total of 58 differential miRNAs were found in thyroid cancer tissues and non-cancerous tissues. The possibility of differential miRNA targeting and binding of immune prognostic factors on the miRDB website and cell experiments was analyzed.

**Conclusions:** The potential miRNA regulatory mechanism of immune prognostic factors in thyroid cancer has been explored.

KEYWORDS

**thyroid cancer, immune gene, prognosis, miRNA, targeted combination**

# Introduction

Thyroid cancer is a common endocrine malignancy (1, 2). Although THCA has a low mortality rate and a relatively good prognosis, the existing therapeutic modalities for locally advanced or recurrent metastatic THCA are still ineffective in improving the prognosis of patients. Therefore, new therapeutic modalities targeting the molecular mechanisms of thyroid carcinogenesis and progression are still being explored.

Recent studies have revealed that normal cells must possess the intrinsic characteristics of tumor cells to develop into cancer cells; most importantly, tumor cells can evade antitumor immune responses and form tumor foci (3). Antitumor immunity is an essential mechanism for the host's inhibition of tumorigenesis. However, tumor cells can evade the body's immune attack, survive, and proliferate to form tumor foci (4, 5). Tumor cells can play a crucial role in disease development through different immune mechanisms. Immune-related genes have been shown to play a role in cancer development (6–8) and, in some studies, can also can predict the prognosis of cancer patients and as therapeutic targets (9–12). In addition to being an effective prognostic biomarker by constructing the immune prognostic features of cancer using bioinformatics, it is also a promising and novel therapeutic target. Wang et al. (13) found that an immune-related prognostic signature consisting of *SLC10A2, FGF2, CCL28, NDRG1, ESM1, UCN, UTS2*, and *TRDC* could effectively predict 3- and 5-year overall survival rates. Heterogeneous expression of immune-related genes can be used as a prognostic factor, but the regulatory mechanism for the heterogeneous expression of immune factors is unknown.

Micro ribonucleic acids (miRNAs) are conserved endogenous short-stranded noncoding ribonucleic acid (RNA) molecules consisting of 21-23 nucleotides. miRNAs have been shown to target multiple messenger RNAs (mRNAs) and are involved in almost every biological process (14, 15). In recent years, studies targeting the role of miRNAs in cancer based on miRNA expression profiles have increased annually. Many miRNAs are down- or up-regulated in human cancers and act as oncogenic or intra-cancer suppressors (16). In ovarian cancer, many miRNAs regulate the epithelial-mesenchymal transition program (17). At the level of molecular mechanisms, miR-125a and miR-125b can regulate immune cell development and function, thus acting as tumor suppressors or promoters (18).

In this study, a total of nine prognostic immune factors, *AKAP12, APOC1, TIMP3, ADAMTS9, ANK2, HTRA3, SYNDIG1, ADAMTS5*, and *DACT1*, were obtained by screening differentially-expressed genes and differential miRNAs using LASSO analysis and analyzing the differential miRNA. The target-binding relationship between differential miRNAs and prognostic immune factors was also analyzed. We validated them using correlation analysis to find the regulatory mechanism of immune prognostic factors in THCA.

# Materials and methods

## Data sources

The data sources for this study were the TCGA and GEO databases. Those databases included RNA sequencing (RNAseq) data and clinicopathological information from 510 THCA samples; mRNA data from 58 normal tissues adjacent to the cancer were obtained. The tissue from 29 THCA patients and five non-THCA patients was obtained from the GEO database GSE103996 dataset miRNA expression profile. Immune-related genes was got from the Immunology Database and Analysis Portal database (https://www.immport.org).

## Differential expression analysis

Differential expression genes(DEGs) and differentially-expressed miRNAs of THCA tissues and paraneoplastic tissues were analyzed differentially using the R package limma package, where differentially-expressed genes were screened at $p < 0.05$, $|\log2FC|>1$, and differentially-expressed miRNAs were screened at $p < 0.05$, $|\log2FC|>1$, and represented in volcano and heat maps using Venn diagrams to take the intersection of the screened DEGs with immune-related genes.

## Prognostic analysis

The transcriptional expression profile data and information of thyroid cancer patients obtained from the TCGA database were plotted using Kaplan–Meier curves with R software. P-values and hazard ratios with 95% confidence intervals (CI) were derived using a log-rank test and univariate Cox regression; the data were presented as forest plots.

## The prognostic signature model

The relationship between prognostic immune-related gene expression and overall survival (OS) was assessed using LASSO COX analysis. the prognostic risk prediction model for THCA was based on the LASSO risk score calculation formula. patients with THCA were divided into high-risk and low-risk groups. KM curves were plotted to compare OS in the high-risk and low-risk fractions. receiver operating characteristic survival analysis was performed using the R package SURVIVAL; decision curve analysis was performed using the risk model decision analysis (rmda) package. The relationship between risk score models and tumor immune infiltrating cells was also observed.

## Single gene expression analysis

RNAseq data in level 3 HTSeq-FPKM format or miRNAseq data in level 3 BCGSC miRNA Profiling from the TCGA THCA (thyroid cancer) project. The miRNAseq data in fragments per kilobase per million (FPKM) or reads per million mapped reads (RPM) format was log2-transformed; data significance was $p < 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, and $p < 0.001$.

## Target gene miRNA prediction and correlation analysis

Target gene miRNA prediction was performed using the microRNA target prediction database (miRDB) website (http://mirdb.org/). Transcriptional profiles and miRNA expression profile data were obtained from the TCGA database. Correlation plots between the two were implemented using the R ggstatsplot package. Spearman's correlation analysis can perform correlations between quantitative variables, which were statistically significant with a p-value less than 0.05.

## Cell culture and RT-qPCR

RT-qPCR assays for human normal thyroid cell line Nthy-ori 3-1 and human THCA cell line ACT-1 were purchased from the American Type Culture Collection. All cell lines were cultured in RPMI-1640 containing 10% fetal bovine serum (FBS, Gibco), 1% penicillin/streptomycin, and maintained in a 37°C, 5% $CO_2$ incubator. Total RNA was extracted from each group of cells using TRIzol reagent (Invitrogen, Carlsbad, CA) following the manufacturer's guidelines. RNA was synthesized into cDNA using Primescript RT Reagent Kit (Takara, Dalian, China). cDNA amplification and quantification were performed on an Applied Biosystems 7500 instrument using SYBR Green mix (Takara). The relative expression of mRNA was calculated using the $2^{-\Delta\Delta Ct}$ method, and the experiment was repeated three times independently for each sample. miRNA was used as an internal reference for U6, and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an internal reference for genes. The PCR primers are shown in Table 1.

TABLE 1 Sequence of primers.

| Primer name | Sequences |
|---|---|
| ADAMTS5-F | GAACATCGACCAACTCTACTCCG |
| ADAMTS5-R | CAATGCCCACCGAACCATCT |
| ADAMTS9-F | TGGGTTTTCCAGTTTTCAG |
| ADAMTS9-R | GTTGATGCTAAAACGACCC |
| ANK2-F | ACCTGCGATACAGCTTGGAG |
| ANK2-R | AGAGTGTGAGACCTGTCGGA |
| AKAP12-F | CTGCCTTGGGAGTTTGCC |
| AKAP12-R | GGGTTACGCCTTCCCCAAG |
| APOC1-F | ACCCACTT AGAGTTGTGAGCCC |
| APOC1-R | CAGACCACCTTAGTCCCTTTCC |
| DACT1-F | AGATATCCCCTTGGCACCCT |
| DACT1-R | TTCAGTGAGAGTCCACCACA |
| HTRA3-F | TGACCAGTCCGCGGTACAAG |
| HTRA3-R | TTGGAGCTGGAGACCACGTG |
| TIMP3-F | CTCGAGCAAGGAGGAACTTGGGTG |
| TIMP3-R | GCGGCCGCAATACAGAAGTGTCT |
| SYNDIG1-F | CCTTGTCCCGGAGCCCA |
| SYNDIG1-R | ACAGACGTGGAGCACTGAAG |
| miR-15a-5p-F | CGCGTAGCAGCACATAATGG |
| miR-15a-5p-R | AGTGCAGGGTCCGAGGTATT |
| miR-30c-5p-F | GCGCGTGTAAACATCCTACACT |
| miR-30c-5p-R | AGTGCAGGGTCCGAGGTATT |
| miR-34b-5p-F | GGGTAGGCAGTGTCATTAGC |
| miR-34b-5p-R | AACAACCAACACAACCCAAC |
| GAPDH-F | GGAGCGAGATCCCTCCAAAAT |
| GAPDH-R | GGCTGTTGTCATACTTCTCATGG |
| U6-F | GCTTCGGCAGCACATATACTAAAAT |
| U6-R | CGCTTCACGAATTTGCGTGTCAT |

## Statistical analysis

Cellular tests were performed using the IBM Statistical Package for the Social Sciences (SPSS) 24.0 and GraphPad Prism8.0.1 software to process the data. Normally distributed measures were expressed as ($x \pm s$), and an independent samples t-test was used to compare the two parts. A one-way analysis of variance can comparison between multiple parts, and Tukey's multiple comparisons test was used for *post hoc* tests. P was a two-sided test, and differences were statistically significant at $p < 0.05$.

# Results

## Immunoprognostic analysis of differentially significant genes in thyroid cancer

We obtained transcriptomic data from 510 thyroid cancer patients and 58 paraneoplastic tissues from the TCGA database for differential gene analysis; a total of 676 up-regulated genes and 737 downregulated genes were obtained by $|\log2FC|>1$ with $p < 0.05$. Figure 1A shows the volcano plot, and Figure 1B shows the heat map. We then obtained 264 identical genes by taking the intersection of the DEGs and immune gene sets (Figure 1C). We performed a prognostic analysis of these 264 genes based on individual gene expression using both one-way Cox and

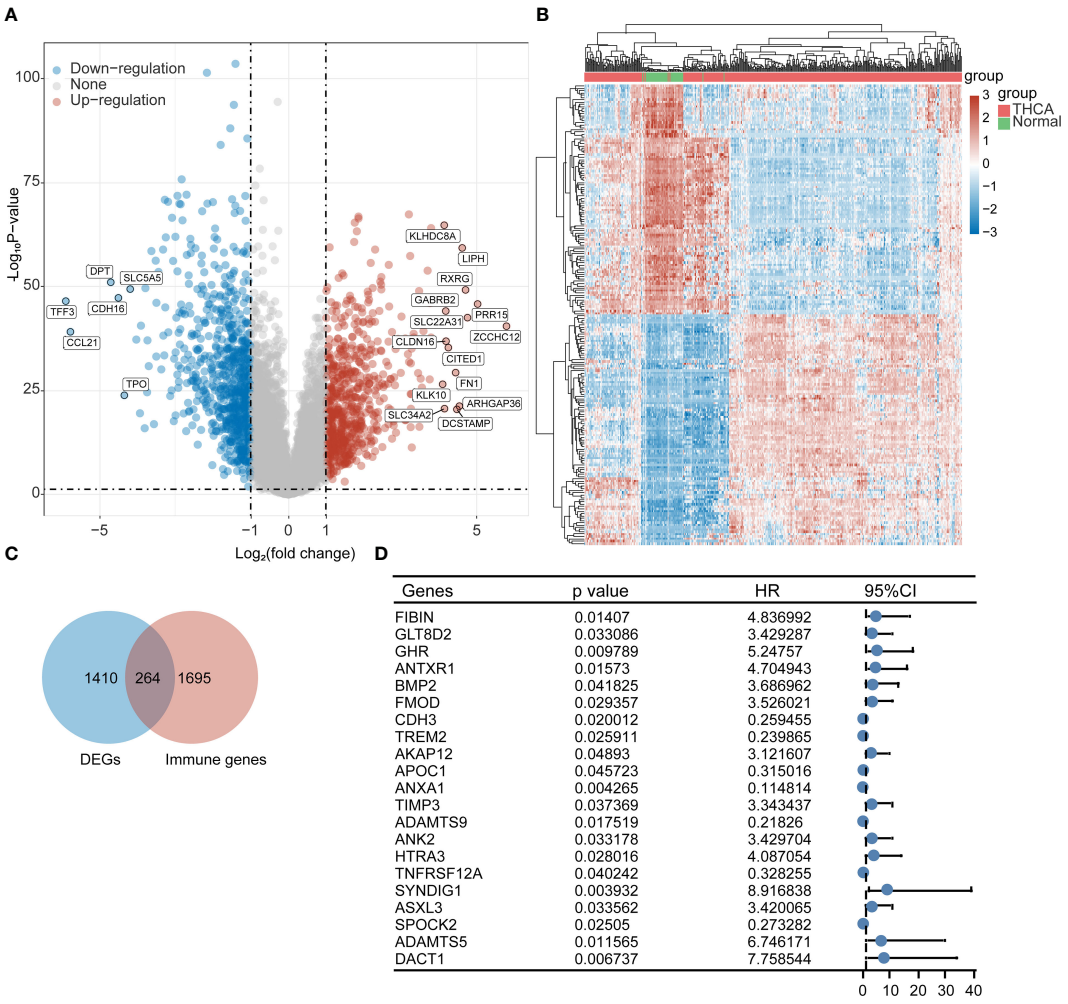



FIGURE 1
Screening of immune-related prognostic genes in DEGs. **(A)** Volcano plot, **(B)** Heat map, **(C)** Wayne plot, and **(D)** 21 differential genes with prognostic value.

log-rank tests. We obtained a total of *FIBIN, LT8D2, GHR, ANTXR1, BMP2, FMOD, CDH3, TREM2, AKAP12, APOC1, ANXA1, TIMP3, ADAMTS9 ANK2, HTRA3, TNFRSF12A, SYNDIG1, ASXL3, SPOCK2, ADAMTS5*, and *DACT1* 21 prognostic genes with significantly different immune-related expression (Figure 1D).

## Immunogenic prognostic model construction

To further explore the clinical value of significant differences in immune-related expression, prognostic features were established for 21 immune-related genes based on LASSO Cox analysis (Figures 2A, B). Using risk score Riskscore = (0.0819) **FIBIN*+(0.0058)**AKAP12*+(-0.0814)**APOC1*+(0.1218)**TIMP3* +(-0.1791)**ADAMTS9*+(0.0335)**ANK2*+(0.3115)**HTRA3* +(0.2709)**SYNDIG1*+(0.0237)**ADAMTS5*+(0.1317)**DACT1*, THCA patients were divided into high and low-risk groups (Figure 2C); Survival status is shown in Figure 6D. The figure shows the difference in survival between the high and low-risk parts (p = 0.00613). Prognostic survival was predicted for 1, 3, and 5 years; the model showed good sensitivity (Figure 6E). In summary, nine prognostic factors were obtained for *AKAP12, APOC1, TIMP3, ADAMTS9, ANK2, HTRA3, SYNDIG1, ADAMTS5*, and *DACT1*.

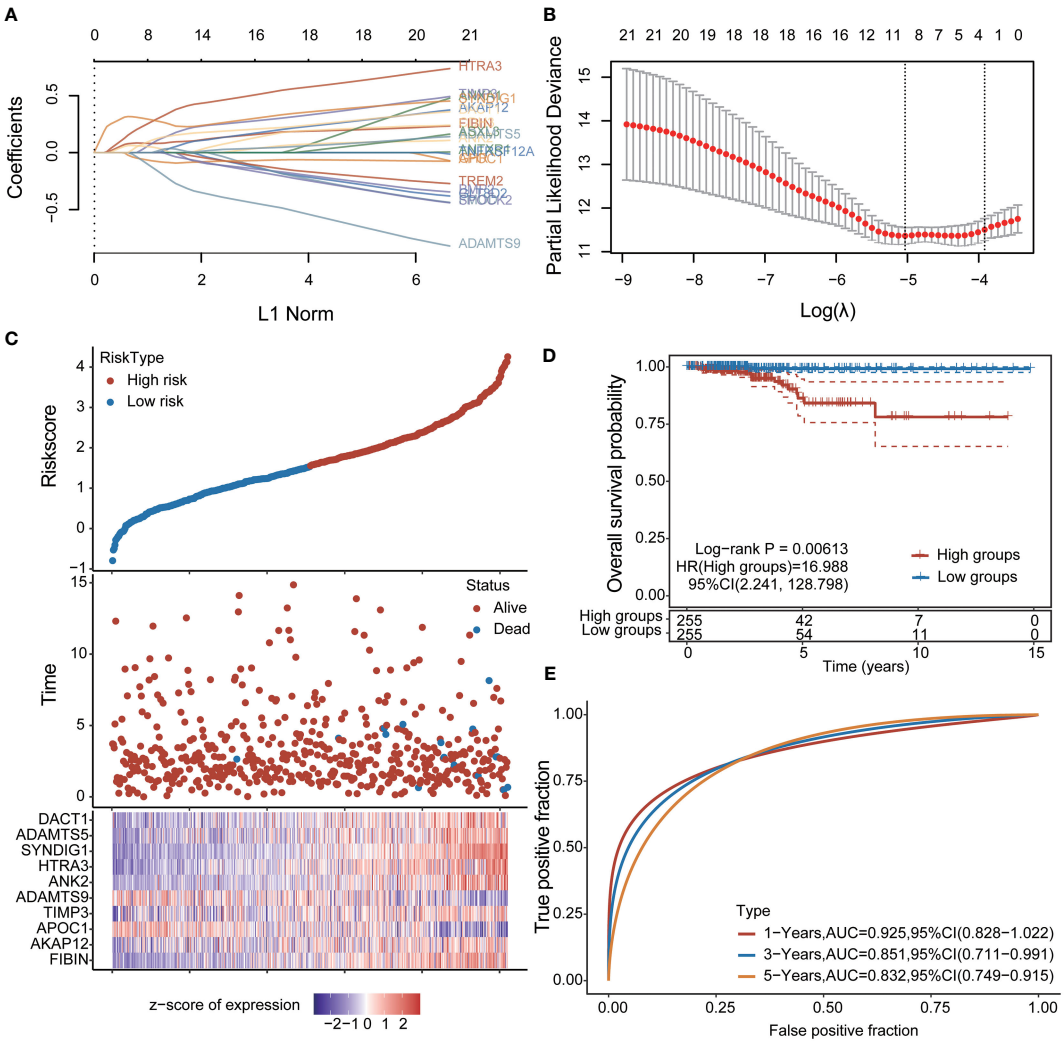



FIGURE 2
Construction of prognostic features of immune-related genes in thyroid cancer **(A, B)** distribution of LASSO coefficients of 21 pivot genes, the adjustment parameter λ.min= 0.0065 is obtained, and the vertical black dotted line in B defines the optimal λ value; **(C)** Thyroid cancer Risk score (Riskscore) distribution of patients, survival status and duration of thyroid cancer patients **(D)** KM survival curves of high and low-risk groups; **(E)**: ROC curves of 1 year, 3 years and 5 years.

## Significant differences exist in miRNAs regulating the heterogeneous expression of prognostic factors

To understand the heterogeneous expression mechanism of immune prognostic factors in thyroid cancer and adjacent tissues, we introduced the miRNA transcriptome information of thyroid cancer and non-cancer tissues into the GSE103996 dataset of the GEO database, and obtained 58 differential miRNAs through analysis. Among them, 21 were up-regulated and 37 were down-regulated. The heat map is shown in Figure 3B. We then screened for miRNAs that bind to postimmune factors, and their targeting relationships are shown in the Sankey diagram in Figure 3C.

## Targeted binding miRNA expression to prognostic factors and prognosis

We verified the expression of miRNAs that bind to prognostic immune factors in thyroid cancer tissues and paracancerous tissues in the THCAGA database as follows: miR-21-5p, miR-181a-5p, miR-181b-5p, miR-34a-5p, miR-15-5p, miR-182-5p, and miR-181-2-3p were significantly up-

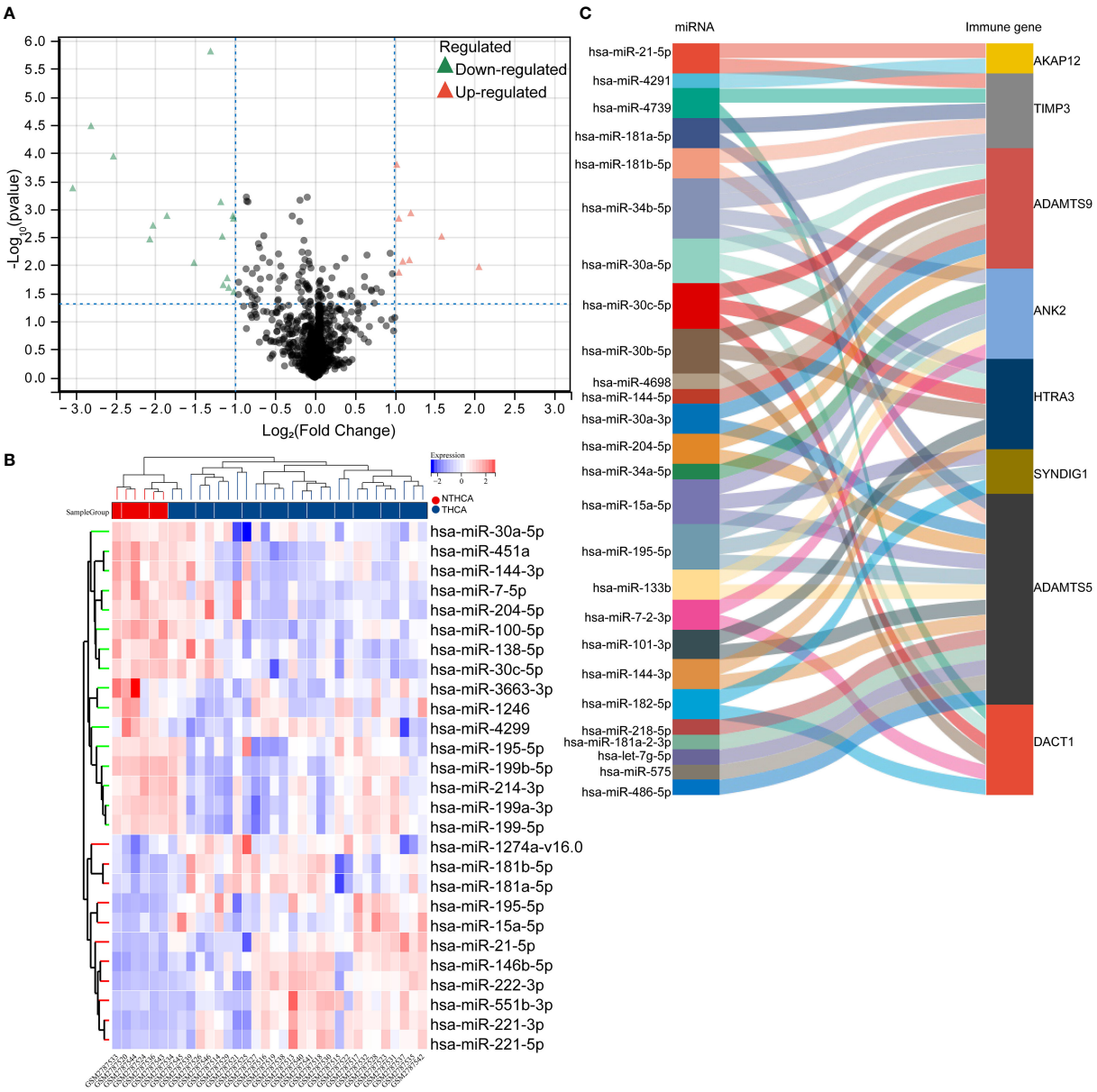



FIGURE 3
Targeted regulation of differentially significant miRNAs that regulate heterogeneous expression of prognostic factors. **(A, B)** volcano plot of differentially significantly expressed miRNAs in the GSE103996 dataset, heat map; **(C)** Sankey plot of miRNAs that target predicted binding to prognostic factors.

regulated tissues. miR-34b-5p, miR-30c-5p, miR-34b-5p, miR-144-5p, miR-30a-3p, miR-204-5p, miR-195-5p, miR-133b, miR-7-2-3p, miR-101-3p, miR-144-3p, miR-218-5p, let-7g-5p, and miR-486-5p were significantly downregulated in tumor tissues. miR-4793, miR-4698, and miR-575 were not significantly changed (Figures 4A–Y).

Subsequently, we verified the prognostic impact of the expression profiles of these miRNAs on THCA. The results showed that among these significantly different miRNAs, only miR-181-2-3p significantly impacted the prognosis of THCA (Figures 5A–V).

## Correlation between miRNA and prognostic immune factors

To further validate the mechanisms of miRNA regulation involved in the expression of prognostic factors, we analyzed the

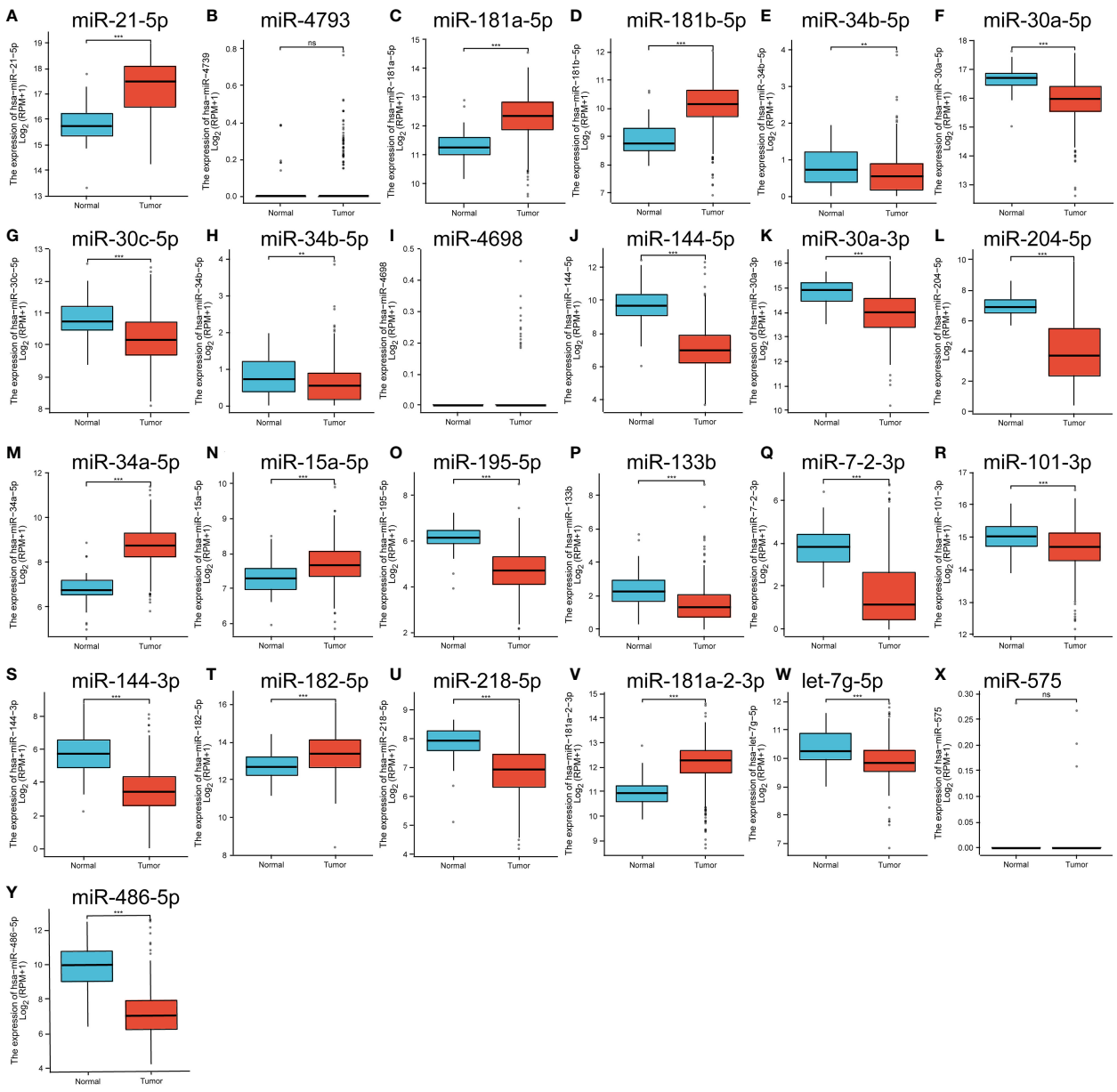



FIGURE 4
**(A–Y)** is the expression of significantly different miRNAs in the GEO database that have a targeted binding relationship with the prognostic factors in the TCGA database. ns means no statistical significance, ** means p<0.01, *** means p<0.00.1.

targeted regulatory relationships between miRNAs and prognostic factors. The results are shown in Figure 6. The figure shows the following: *ADMTS5* was significantly and negatively correlated with miR-15a-5p, miR-101-3p, miR-181a-5p, and miR-181b-5p. *ADMTS9* was significantly and negatively correlated with miR-30a-5p, miR-30a-3p, miR-30c-5p, and miR-144-5p. *AKAP12* was significantly and negatively correlated with miR-21-5p. *ANK2* was significantly and negatively correlated with miR-15a-5p and miR-34a-5p. *DACT1* was significantly and negatively associated with miR-182-5p. *SYNDIG1* was significantly and negatively related to miR-15a-5p and miR-182-5p. *TIMP3* was significantly and negatively related to miR-21-5p, miR-181a-5p, and miR-181b-5p.

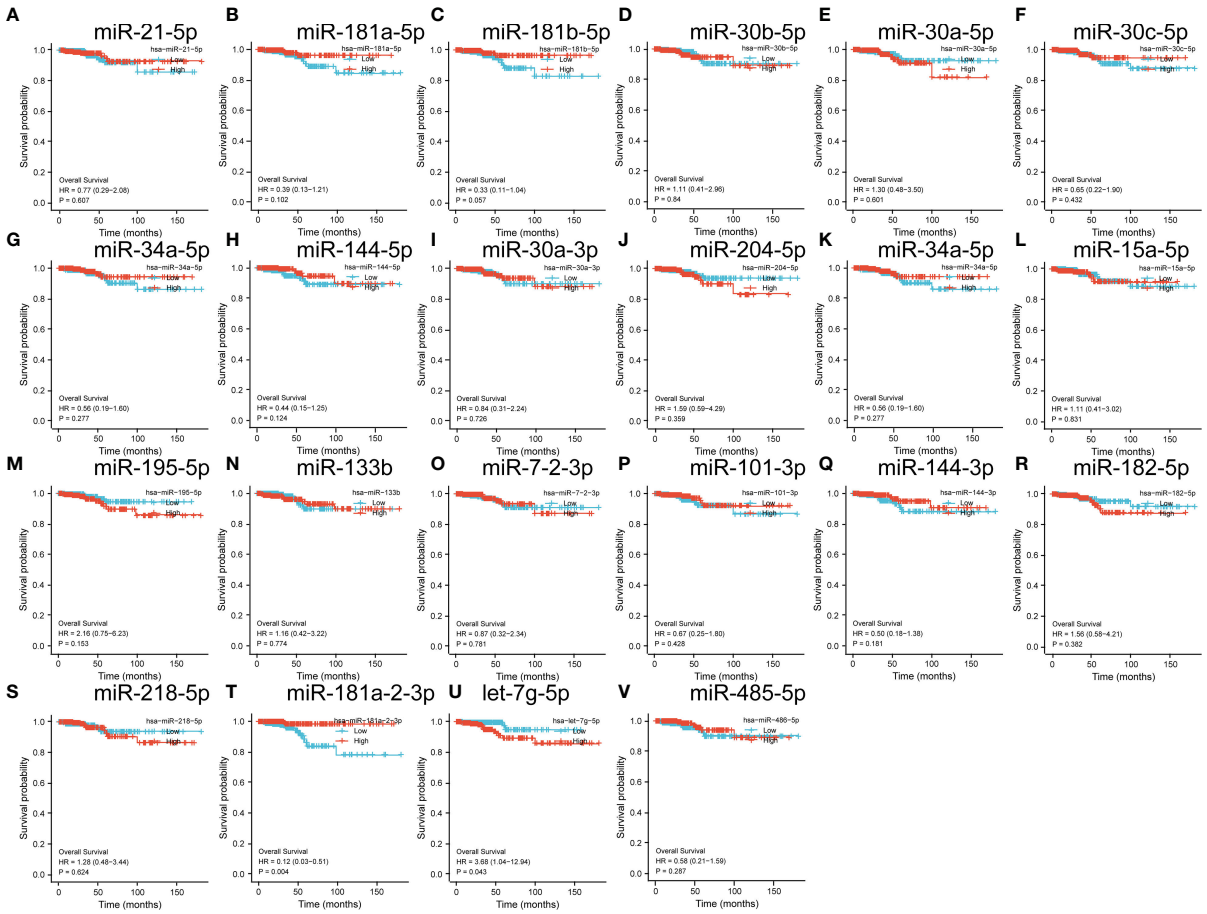



FIGURE 5
**(A–V)** is a significantly different miRNA in the GEO database that has a target binding relationship with a prognostic factor and the impact on the prognosis of thyroid cancer.

## Correlation of experimental cellular miRNAs with target genes

First, we verified the expression of 9 immune-related prognostic factors, AKAP12, APOC1, TIMP3, ADAMTS9, ANK2, HTRA3, SYNDIG1, ADAMTS5 and DACT1, in 11 thyroid cancer cell lines in the CCLE database. Eleven thyroid cancer cell lines expressed to varying degrees (Figures 7A–I). We verified the expression of immunoprognostic factors in the cells, as shown in Figure 8A. Compared with the normal human thyroid cell line Nthy-ori 3-1, the human THCA cell line ACT-1 showed significant upregulation of *ADAMTS9, APOC1*, and miR-15a-5p. *AKAP12, TIMP3, ANK2, HTRA3, SYNDIG1, ADAMTS5, DACT1*, miR-30c-5p, and miR-34b-5p showed significant down-regulation (Figure 7A). We then compared the correlation of miR-15a-5p, miR-30c-5p, and miR-34b-5p with their target genes and found that miR-15a-5p showed a significant negative correlation with *ADAMTS5*, *ANK2*, and *SYNDIG1*. miR-30c-5p and miR-34b-5p showed a significant negative correlation with *ADAMTS9* (Figure 8B).

## Discussion

Tumor immunity is associated with proto-oncogene activation and tumor immune escape; therefore, understanding tumor-associated immune genes is crucial for tumor immunity research. In recent years, THCA immune studies have focused primarily on developing immune prognostic models (19–21), and less research has been involved in the regulatory mechanisms of prognostic immune factors. This study mainly analyzed the relationship between immune prognostic factors and miRNAs in THCA. We provided a theoretical basis for further understanding of the immune gene regulation

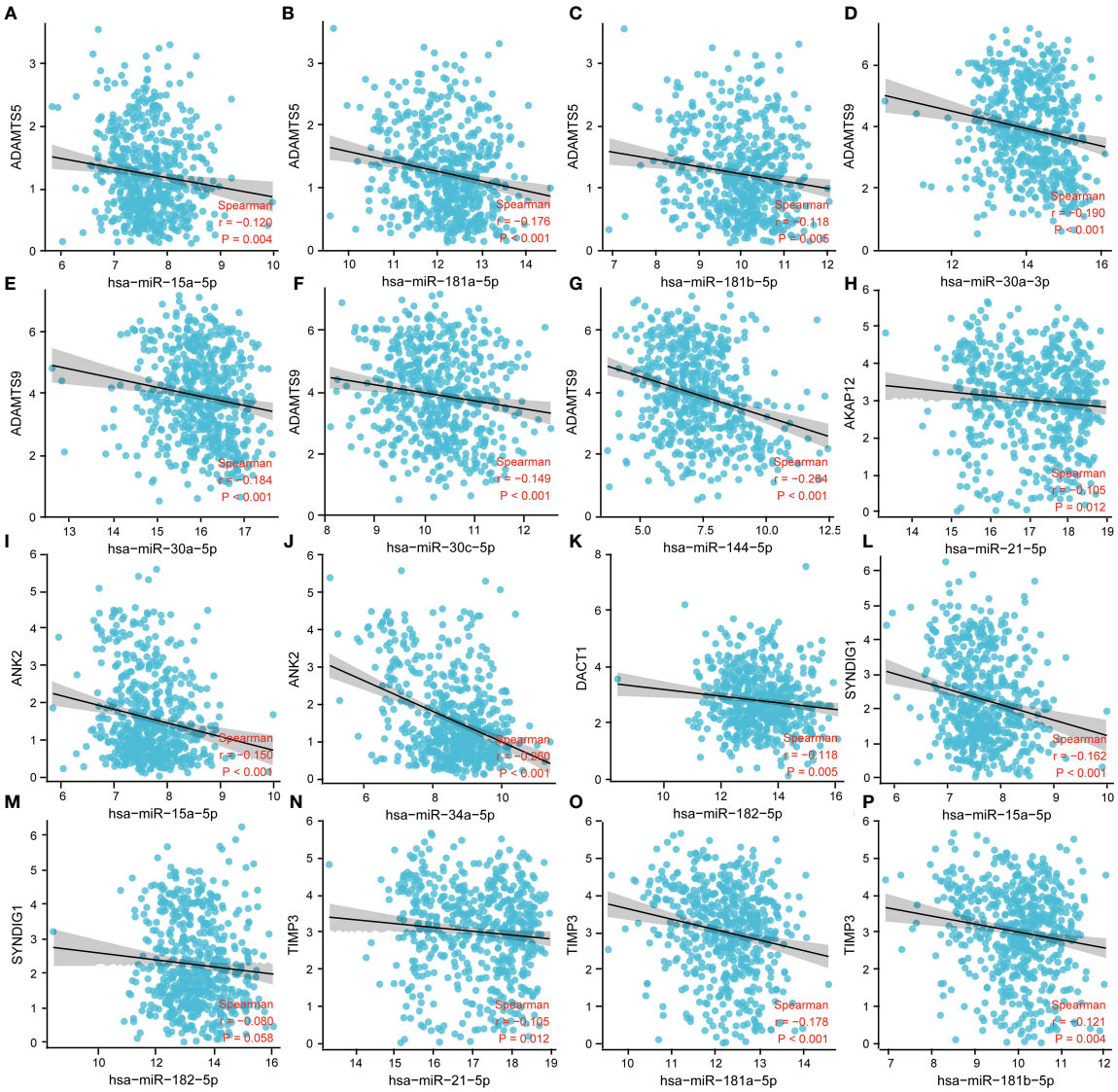



FIGURE 6
Correlation analysis between miRNA and prognostic immune factors. **(A–C)** *ADMTS5* with miR-15a-5p, miR-101-3p, miR-181a-5p, miR-181b-5p; **(D–G)** *ADMTS9* with miR-30a-5p, miR-30a-3p, miR-30c-5p, miR-144-5p; **(H)** AKAP12 with miR-21-5p. **(I, J)** *ANK2* with miR-15a-5p, miR-34a-5p; **(K)** *DACT1* with miR-182-5p; **(L, M)** *SYNDIG1* with miR-15a-5p, miR-182-5p; and **(N–P)** *TIMP3* with miR-21-5p, miR-181a-5p, and miR-181b-5p. $P < 0.05$ was statistically significant.

mechanism of THCA and a reference for assessing the prognosis of THCA by analyzing the correlation between the two.

This study analyzed the relationship between immune genes and prognosis using THCAGA and immunoinformatics databases. Twenty-one genes with significant differences in the immune prognosis of THCA were screened, and nine prognostic immune factors were identified by constructing a LASSO risk model (*AKAP12, APOC1, TIMP3, ADAMTS9, ANK2, HTRA3, SYNDIG1, ADAMTS5*, and *DACT1*). This risk model predicted OS after five years with high accuracy (AUC = 0.832).

The nine immune genes in the model have biological characteristics and are related to the clinical prognosis of tumor patients. AKAP12 is an A-kinase scaffold protein with a characteristic binding structural domain of the protein kinase A regulatory subunit. AKAP12 deletion is associated with increased cancer susceptibility (22). *AKAP1* expression is associated with *APOC1* overexpression, promotes tumor progression, and has a poorer prognosis for patient survival (23). *TIMP3* is a metalloproteinase that belongs to the TIMP family. It has a high affinity for proteoglycans in the ECM and its broadest range of substrates, including all MMPs, ADAMs (a disintegrin and metalloproteinase), and ADAMTSs (ADAMs with thrombospondin motifs) (24).

*TIMP-3* has been shown to have anti-metastatic effects by inhibiting matrix metalloproteinases and members of the ADAM

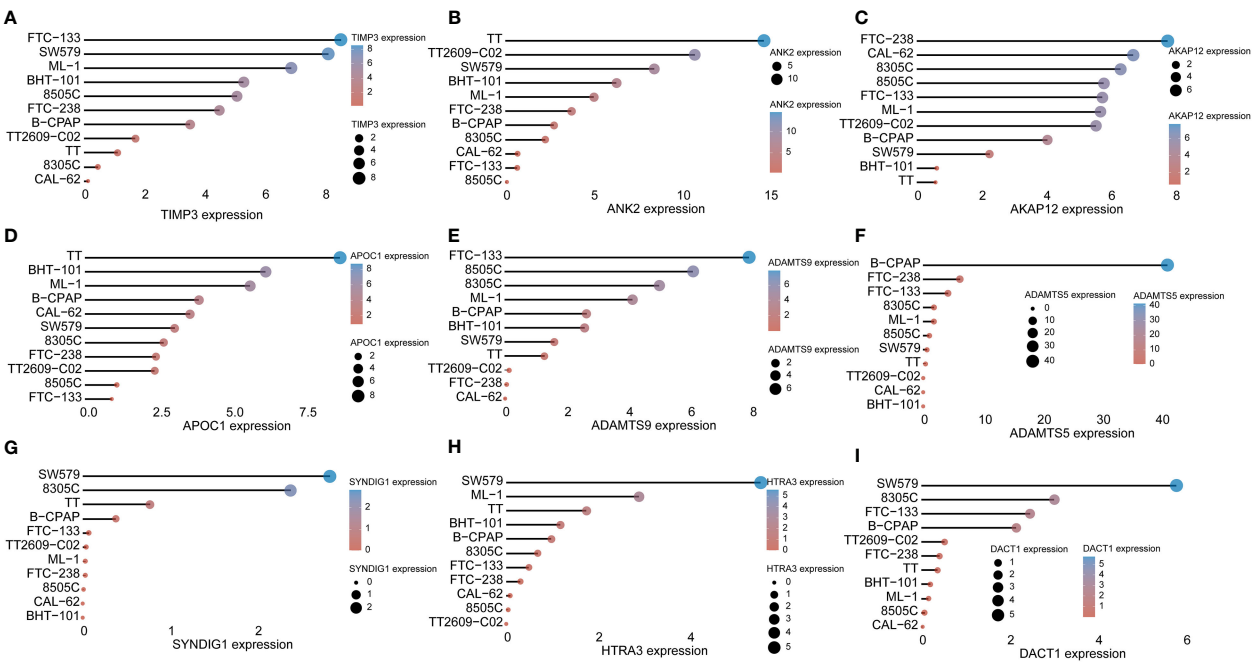



FIGURE 7
Immune prognostic factors expressed in CCLE data Verification **(A-I)** AKAP12, APOC1, TIMP3, ADAMTS9, ANK2, HTRA3, SYNDIG1, ADAMTS5, and DACT1 Nine immune-related prognostic factors in TT, BHT-101, ML-1, B-CPAP, the expression of 11 thyroid cancer cell lines including CAL-62, SW579, 8305C, FTC-238, TT2609-CO2, 8505C, and FTC-133.

family and down-regulating angiogenesis (25). *ADAMTS5* and *ADAMTS9* belong to the ADAMTS (disintegrin-like and metalloproteinase with platelet abasicin motif) proteins. They are extracellular zinc metalloproteinases that play an essential role in extracellular matrix assembly and degradation, connective tissue structure, angiogenesis generation, and cell migration (26). *ADAMTSL5* plays a role in maintaining the function of critical oncogenic signaling pathways, suggesting that it may act as a significant regulator of tumorigenicity and drug resistance (27). adamts9 regulates cancer cell growth and metastasis (28) and is associated with patient survival (29).

*ANK2* has a high mutation frequency in some cancers (30) and silencing of *ANK2* expression reduces the growth and invasion of cancer cell type. *HTRA3* has been implicated as a tumor suppressor in cancer progression in several cancer types, *HTRA3* expression is negatively correlated with adaptive immune cell abundance (T helper cell 17 cells) and positively correlated with innate immune cells (natural killer cells, macrophages, etc.); abundance is positively correlated (31). *SYNDIG1* is a prognostic immune factor in diffuse large B-cell lymphoma and breast cancer (32, 33). *DACT1* belongs to the DACT (Disheveled-associated antagonist of β-catenin) family and is a methylation biomarker for DACT1 in esophageal squamous cell carcinoma (34).

The above data suggest that the nine immune genes in our constructed risk model can exist as prognostic factors in some tumors, reflecting the accuracy of our constructed risk model from the side.

In many studies, the regulatory mechanisms of prognostic factors remain unclear. Based on this starting point, this study analyzed the correlation between immune prognostic factors and miRNAs through TCGA and GEO databases to find the miRNA mechanism of immune prognostic factors in thyroid cancer species.

It has been shown that miRNAs can regulate the target genes involved in the prognosis of tumor patients. For example, in nasopharyngeal carcinoma, the negative correlation between miRNA-19a-3p and *PDCD5* expression levels, miRNA-19a-3p targeting to suppress PDCD5 expression, miRNA-19a-3p levels correlated with N classification and clinical stage of nasopharyngeal carcinoma patients, and PDCD5 levels correlated with T classification, pathological grade, and clinical stage. Survival analysis showed that the expression of high levels of miRNA-19a-3p or low levels of PDCD5 had a poorer prognosis in patients with nasopharyngeal carcinoma (35).

Our study identified 16 groups of miRNAs that showed negative correlations with prognostic immune factors, with specific target-binding relationships confirmed in other cancers. For example, *TIMP3* was downregulated in cervical cancer and was related to poor prognosis in cervical cancer patients, and miR-21-5p target binding to *TIMP3* modulated the development of cervical cancer (36). In gastric cancer, the miR-

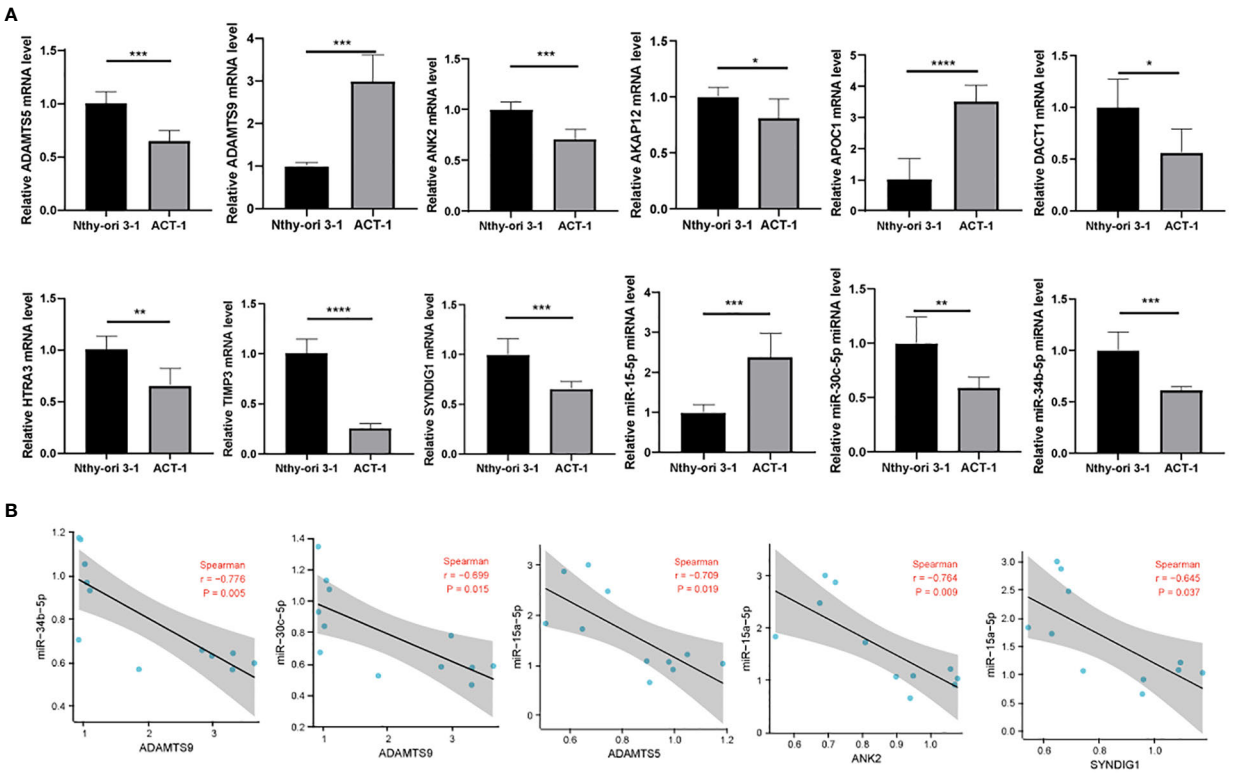



FIGURE 8
Correlation of experimental cellular miRNAs with target genes. **(A)** RT-qPCR detection of *ADAMTS5, ADAMTS9, ANK2, AKAP12, APOC1, DACT1, HTRA3, TIMP3, SYNDIG1*, miR-15a-5p, miR-30c-5p, and miR-34b-5p expression levels; **(B)** Correlation analysis of miR-15a-5p, miR-30c-5p, and miR-34b-5p and their target genes. * means p <0.05; ** means p < 0.01; *** means p < 0.001; **** means p < 0.0001.

21-5p binding relationship with *TIMP3* is related to drug resistance in gastric cancer (37). In prostate cancer, the miR-181b-5p-*TIMP3* axis regulates prostate cancer proliferation, migration, and invasive ability (38).

This study also found that by constructing a miRNA regulatory network with certain selected candidate factors (39–42), we could provide a new approach to cancer diagnosis and treatment. The miRNA regulatory network of prognostic immune factors was screened. We also validated at the cellular level that miR-15a-5p showed a significant negative correlation with *ADAMTS5, ANK2*, and *SYNDIG1*. miR-30c-5p and miR-34b-5p showed a significant negative correlation with *ADAMTS9*, but no functional gene. However, no further validation of the functional role of the genes was performed, which will be the direction of our later research.

In summary, nine immunoprognostic factors (AKAP12, APOC1, TIMP3, ADAMTS9, ANK2, HTRA3, SYNDIG1, ADAMTS5, and DACT1) were screened as biomolecular markers for predicting THCA. A prognostic assessment model and risk score system were constructed to predict THCA 5-year OS rates with high and low-risk groups using risk scores. The 5-year OS rate of THCA was predicted with high accuracy. By constructing a prognostic immune factor miRNA regulatory network, we can provide a theoretical basis for the prognostic immune regulatory mechanism of THCA.

## Data availability statement

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

## Author contributions

YT writing the article, writing the revision of article, collect data. TX writing the article, designing the methods, analysis and interpretation. XS writing the article, writing the revision of article, designing the methodology. All authors contributed to the article and approved the submitted version.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

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

2. Franchini F, Palatucci G, Colao A, Ungaro P, Macchia PE, Nettore IC, et al. Obesity and thyroid cancer risk: An update. *Int J Environ Res Public Health* (2022) 19(3):1116. doi: 10.3390/ijerph19031116

3. Pulendran B, Davis MM. The science and medicine of human immunology. *Science* (2020) 369(6511). doi: 10.1126/science.aay4014

4. St Paul M, Ohashi PS. The roles of CD8(+) T cell subsets in antitumor immunity. *Trends Cell Biol* (2020) 30(9):695–704. doi: 10.1016/j.tcb.2020.06.003

5. Ikeda H, Togashi Y. Aging, cancer, and antitumor immunity. *Int J Clin Oncol* (2022) 27(2):316–22. doi: 10.1007/s10147-021-01913-z

6. Oliva M, Spreafico A, Taberna M, Alemany L, Coburn B, Mesia R, et al. Immune biomarkers of response to immune-checkpoint inhibitors in head and neck squamous cell carcinoma. *Ann Oncol* (2019) 30(1):57–67. doi: 10.1093/annonc/mdy507

7. Akbulut H. Immune gene therapy of cancer. *Turk J Med Sci* (2020) 50(SI-2):1679–90. doi: 10.3906/sag-2005-327

8. 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

9. Sun S, Guo W, Wang Z, Wang X, Zhang G, Zhang H, et al. Development and validation of an immune-related prognostic signature in lung adenocarcinoma. *Cancer Med* (2020) 9(16):5960–75. doi: 10.1002/cam4.3240

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

11. Kong Y, Feng ZC, Zhang YL, Liu XF, Ma Y, Zhao ZM, et al. Identification of immune-related genes contributing to the development of glioblastoma using weighted gene Co-expression network analysis. *Front Immunol* (2020) 11:1281. doi: 10.3389/fimmu.2020.01281

12. Xu H, Zhang A, Han X, Li Y, Zhang Z, Song L, et al. ITGB2 as a prognostic indicator and a predictive marker for immunotherapy in gliomas. *Cancer Immunol Immunother* (2022) 71(3):645–60. doi: 10.1007/s00262-021-03022-2

13. Wang J, Yu S, Chen G, Kang M, Jin X, Huang Y, et al. A novel prognostic signature of immune-related genes for patients with colorectal cancer. *J Cell Mol Med* (2020) 24(15):8491–504. doi: 10.1111/jcmm.15443

14. Annese T, Tamma R, De Giorgis M, Ribatti D. microRNAs biogenesis, functions and role in tumor angiogenesis. *Front Oncol* (2020) 10:581007. doi: 10.3389/fonc.2020.581007

15. Chauhan N, Dhasmana A, Jaggi M, Chauhan SC, Yallapu MM. miR-205: A potential biomedicine for cancer therapy. *Cells* (2020) 9(9):1957. doi: 10.3390/cells9091957

16. Ali Syeda Z, Langden SSS, Munkhzul C, Lee M, Song SJ. Regulatory mechanism of MicroRNA expression in cancer. *Int J Mol Sci* (2020) 21(5):1723. doi: 10.3390/ijms21051723

17. Ghafouri-Fard S, Shoorei H, Taheri M. miRNA profile in ovarian cancer. *Exp Mol Pathol* (2020) 113:104381. doi: 10.1016/j.yexmp.2020.104381

18. Wang JK, Wang Z, Li G. MicroRNA-125 in immunity and cancer. *Cancer Lett* (2019) 454:134–45. doi: 10.1016/j.canlet.2019.04.015

19. Gan X, Guo M, Chen Z, Li Y, Shen F, Feng J, et al. Development and validation of a three-immune-related gene signature prognostic risk model in papillary thyroid carcinoma. *J Endocrinol Invest* (2021) 44(10):2153–63. doi: 10.1007/s40618-021-01514-7

20. Lin R, Fogarty CE, Ma B, Li H, Ni G, Liu X, et al. Identification of ferroptosis genes in immune infiltration and prognosis in thyroid papillary carcinoma using network analysis. *BMC Genomics* (2021) 22(1):576. doi: 10.1186/s12864-021-07895-6

21. Qin R, Li C, Wang X, Zhong Z, Sun C. Identification and validation of an immune-related prognostic signature and key gene in papillary thyroid carcinoma. *Cancer Cell Int* (2021) 21(1):378. doi: 10.1186/s12935-021-02066-9

22. Hayashi M, Nomoto S, Kanda M, Okamura Y, Nishikawa Y, Yamada S, et al. Identification of the a kinase anchor protein 12 (AKAP12) gene as a candidate tumor suppressor of hepatocellular carcinoma. *J Surg Oncol* (2012) 105(4):381–6. doi: 10.1002/jso.22135

23. Li YL, Wu LW, Zeng LH, Zhang ZY, Wang W, Zhang C, et al. ApoC1 promotes the metastasis of clear cell renal cell carcinoma *via* activation of STAT3. *Oncogene* (2020) 39(39):6203–17. doi: 10.1038/s41388-020-01428-3

24. Fan D, Kassiri Z. Biology of tissue inhibitor of metalloproteinase 3 (TIMP3), and its therapeutic implications in cardiovascular pathology. *Front Physiol* (2020) 11:661. doi: 10.3389/fphys.2020.00661

25. Rai GP, Baird SK. Tissue inhibitor of matrix metalloproteinase-3 has both anti-metastatic and anti-tumourigenic properties. *Clin Exp Metastasis* (2020) 37 (1):69–76. doi: 10.1007/s10585-019-10017-y

26. Mead TJ, Apte SS. ADAMTS proteins in human disorders. *Matrix Biol* (2018) 71-72:225–39. doi: 10.1016/j.matbio.2018.06.002

27. Arechederra M, Bazai SK, Abdouni A, Sequera C, Mead TJ, Richelme S, et al. ADAMTSL5 is an epigenetically activated gene underlying tumorigenesis and drug resistance in hepatocellular carcinoma. *J Hepatol* (2021) 74(4):893–906. doi: 10.1016/j.jhep.2020.11.008

28. Chen L, Tang J, Feng Y, Li S, Xiang Q, He X, et al. ADAMTS9 is silenced by epigenetic disruption in colorectal cancer and inhibits cell growth and metastasis by regulating Akt/p53 signaling. *Cell Physiol Biochem* (2017) 44(4):1370–80. doi: 10.1159/000485534

29. Du W, Wang S, Zhou Q, Li X, Chu J, Chang Z, et al. ADAMTS9 is a functional tumor suppressor through inhibiting AKT/mTOR pathway and associated with poor survival in gastric cancer. *Oncogene* (2013) 32(28):3319–28. doi: 10.1038/onc.2012.359

30. Jeong S, Park YJ, Yun W, Lee ST, Choi JR, Suh C, et al. Genetic heterogeneity and prognostic impact of recurrent ANK2 and TP53 mutations in mantle cell lymphoma: A multi-centre cohort study. *Sci Rep* (2020) 10(1):13359. doi: 10.1038/s41598-020-70310-9

31. Ji C, Sun LS, Xing F, Niu N, Gao HL, Dai JW, et al. HTRA3 is a prognostic biomarker and associated with immune infiltrates in gastric cancer. *Front Oncol* (2020) 10:603480. doi: 10.3389/fonc.2020.603480

32. Mu S, Shi D, Ai L, Fan F, Peng F, Sun C, et al. International prognostic index-based immune prognostic model for diffuse Large b-cell lymphoma. *Front Immunol* (2021) 12:732006. doi: 10.3389/fimmu.2021.732006

33. Cui Q, Tang J, Zhang D, Kong D, Liao X, Ren J, et al. A prognostic eight-gene expression signature for patients with breast cancer receiving adjuvant chemotherapy. *J Cell Biochem* (2019). doi: 10.1002/jcb.29550

34. Guo YL, Shan BE, Guo W, Dong ZM, Zhou Z, Shen SP, et al. Aberrant methylation of DACT1 and DACT2 are associated with tumor progression and poor prognosis in esophageal squamous cell carcinoma. *J BioMed Sci* (2017) 24 (1):6. doi: 10.1186/s12929-016-0308-6

35. Zhong JH, Zhong JJ, Shi YN, Hu XM, Xu MJ, Guo Z, et al. Prognostic potentials of miRNA-19a-3p and PDCD5 in nasopharynx carcinoma. *Eur Rev Med Pharmacol Sci* (2020) 24(21):11114–9. doi: 10.26355/eurrev_202011_23598

36. Gao Y, Zou T, Liang W, Zhang Z, Qie M. Long non-coding RNA HAND2-AS1 delays cervical cancer progression via its regulation on the microRNA-21-5p/TIMP3/VEGFA axis. *Cancer Gene Ther* (2021) 28(6):619–33. doi: 10.1038/s41417-020-00243-y

37. Chen J, Zhou C, Li J, Xiang X, Zhang L, Deng J, et al. miR215p confers doxorubicin resistance in gastric cancer cells by targeting PTEN and TIMP3. *Int J Mol Med* (2018) 41(4):1855–66. doi: 10.3892/ijmm.2018.3405

38. Xie X, Sun FK, Huang X, Wang CH, Dai J, Zhao JP, et al. A circular RNA, circSMARCA5, inhibits prostate cancer proliferative, migrative, and invasive capabilities via the miR-181b-5p/miR-17-3p-TIMP3 axis. *Aging (Albany NY)* (2021) 13(15):19908–19. doi: 10.18632/aging.203408

39. Ran X, Luo J, Zuo C, Huang Y, Sui Y, Cen J, et al. Developing metabolic gene signatures to predict intrahepatic cholangiocarcinoma prognosis and mining a miRNA regulatory network. *J Clin Lab Anal* (2022) 36(1):e24107. doi: 10.1002/jcla.24107

40. Gao L, Zhou L, Huang X. Identification of novel kinase-transcription factor-mRNA-miRNA regulatory network in nasopharyngeal carcinoma by bioinformatics analysis. *Int J Gen Med* (2021) 14:7453–69. doi: 10.2147/IJGM.S327657

41. Yang T, Miao X, Bai Z, Tu J, Shen S, Niu H, et al. A novel mRNA-miRNA regulatory Sub-network associated with prognosis of metastatic clear cell renal cell carcinoma. *Front Oncol* (2020) 10:593601. doi: 10.3389/fonc.2020.593601

42. Zhu F, Liu Z, Zhou Q, Fan J, Zhou D, Xing L, et al. Identification of mRNA prognostic markers for TGCT by integration of Co-expression and CeRNA network. *Front Endocrinol (Lausanne)* (2021) 12:743155. doi: 10.3389/fendo.2021.743155

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 14 December 2022
DOI 10.3389/fonc.2022.1056623

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Li Qiu,
Tianjin Medical University Cancer Institute and Hospital, China
Li Youkong,
Jingzhou Central Hospital, China

*CORRESPONDENCE
Sanyang Tan
tansanyanglove@163.com

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

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

RECEIVED 29 September 2022
ACCEPTED 23 November 2022
PUBLISHED 14 December 2022

CITATION
He Y, Lin Z and Tan S (2022) Identification of prognosis-related gene features in low-grade glioma based on ssGSEA.
*Front. Oncol.* 12:1056623.
doi: 10.3389/fonc.2022.1056623

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

# Identification of prognosis-related gene features in low-grade glioma based on ssGSEA

Yuanzhi He[1†], Zhangping Lin[2†] and Sanyang Tan[3*]

[1]Department of Neurosurgery, Wuhan Children's Hospital (Wuhan Maternal and Child Healthcare Hospital), Tongji Medical College, Huazhong University of Science & Technology, Wuhan, Hubei, China, [2]Clinical Laboratory, Hainan Women and Children's Medical Center, Haikou, Hainan, China, [3]Clinical Laboratory, Haikou Hospital of, The Maternal and Child Health, Haikou, Hainan, China

Low-grade gliomas (LGG) are commonly seen in clinical practice, and the prognosis is often poor. Therefore, the determination of immune-related risk scores and immune-related targets for predicting prognoses in patients with LGG is crucial. A single-sample gene set enrichment analysis (ssGSEA) was performed on 22 immune gene sets to calculate immune-based prognostic scores. The prognostic value of the 22 immune cells for predicting overall survival (OS) was assessed using the least absolute shrinkage and selection operator (LASSO) and univariate and multivariate Cox analyses. Subsequently, we constructed a validated effector T-cell risk score (TCRS) to identify the immune subtypes and inflammatory immune features of LGG patients. We divided an LGG patient into a high-risk–score group and a low-risk–score group based on the optimal cutoff value. Kaplan–Meier survival curve showed that patients in the low-risk–score group had higher OS. We then identified the differentially expressed genes (DEGs) between the high-risk–score group and low-risk-score group and obtained 799 upregulated genes and 348 downregulated genes. The analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) show that DEGs were mainly concentrated in immune-related processes. In order to further explore the immune-related genes related to prognosis, we constructed a protein–protein interaction (PPI) network using Cytoscape and then identified the 50 most crucial genes. Subsequently, nine DEGs were found to be significantly associated with OS based on univariate and multivariate Cox analyses. It was further confirmed that CD2, SPN, IL18, PTPRC, GZMA, and TLR7 were independent prognostic factors for LGG through batch survival analysis and a nomogram prediction model. In addition, we used an RT-qPCR assay to validate the bioinformatics results. The results showed that CD2, SPN, IL18, PTPRC, GZMA, and TLR7 were highly expressed in LGG. Our study can provide a reference value for the prediction of prognosis in LGG patients and may help in the clinical development of effective therapeutic agents.

KEYWORDS

**low-grade glioma, immune infiltration, T-cell risk score, crucial gene, prognosis**

# Introduction

A glioma is a tumor arising from the carcinogenesis of glial cells in the brain and spinal cord and is characterized by a high incidence, a low cure rate, aggressive growth, a high malignancy, and a significantly higher incidence in men than in women. Gliomas are commonly found in clinical practice, accounting for approximately 81% of intracranial malignancies (1). Most types of gliomas have a poor prognosis due to their malignant biological behavior (2). Low-grade gliomas (LGG) are gliomas of low malignancy, and even though the prognosis is better than that of glioblastoma, many LGG progress to high-grade gliomas due to the great heterogeneity between different LGG, and thus the prognosis of LGG patients is poor (3, 4). Currently, clinical treatment for LGG includes surgical resection, radiotherapy, chemotherapy, electric field therapy, and supportive therapy, all of which help to prolong survival time and improve quality of life (5). However, reliable biomarkers that can predict poor prognoses in patients with LGG are uncommon in the diagnosis and treatment of LGG. Therefore, a search for effective prognostic predictors and therapeutic targets for LGG is necessary.

The development of LGG is closely related to immunity, and LGG cells can secrete a large number of cytokines that promote the entry of various immune cells into the tumor, which in turn creates a tumor microenvironment (6). The tumor microenvironment is involved in LGG growth, recurrence, invasion, and response to therapy, and immune cells are closely associated with poor prognoses in LGG patients (7). Currently, immuno-oncology is of great clinical interest due to its specific benefits in the treatment of a variety of cancers. Immune-related genes and immune-infiltrating cells play an integral role in the tumor microenvironment, helping to determine prognosis and providing an impetus for immunotherapy (8). Therefore, it is crucial to find immune-related risk scores and immune-related targets for predicting prognosis.

Currently, there is no report on the bioinformatics analysis of immune-related targets for predicting LGG prognosis based on single-sample gene set enrichment analysis (ssGSEA) and constructing T-cell risk score (TCRS), as described in the previous articles. In the present study, immune cell abundance in LGG samples was explored using ssGSEA to construct an effector TCRS. The prognostic value of immune cells for overall survival (OS) time prediction was assessed to elucidate functional differences between high- and low-risk groups distinguished by TCRS. Subsequently, a protein–protein interaction (PPI) network, univariate and multivariate Cox analyses, batch survival analysis, and nomogram prediction model can identify the crucial genes related to LGG prognosis and resolve the mode of action of crucial genes in LGG, providing a reference for improving the prognosis of LGG patients.

# Data

## Data sources

RNA-seq data from the LGG patient sample ($n$ = 512) and corresponding clinical information can be obtained from The Cancer Genome Atlas (TCGA) for inclusion in subsequent analyses.

## Prognostic value of immune cells for predicting overall survival

The ssGSEA algorithm is based on a set of 22 immune genes, including genes associated with different immune cell types, functions, pathways, and checkpoints. The ssGSEA results were analyzed using the R package "GSVA" to identify different levels of infiltration of immune cell types, immune-related functions, and immune-related pathways in LGG expression profiles. The image clustering heat map was drawn using the R package "Pheatmap."

The prognostic value of 22 immune cells for predicting OS was assessed using least absolute shrinkage and selection operator (LASSO) and univariate and multivariate Cox analyses. The TCRS was determined based on the most prognostically significant immune cells, and the optimal cutoff value of the TCRS was calculated using the R package "ggrisk." The LGG samples were divided into a high-risk–score group and a low-risk–score group based on the optimal cutoff value. Kaplan–Meier and ROC analyses were used to compare the survival of patients in both groups. Based on the ESTIMATE algorithm, the R package "ESTIMATE" was used to calculate the stroma score, immune score, and ESTIMATE score of LGG samples in different risk score groups. The box plot was drawn using the R package "ggpubr." The relationship between different risk scores and changes in the expression of each immune gene was analyzed.

## Differential analysis

Differential analysis was performed on the high-risk–score group and a low-risk–score group using the R package "limma" according to the screening threshold of |Log2FC|>1, adj.$p < 0.05$. The R package "clusterProfiler" was used to select datasets from the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO).

## Protein–protein interaction networks

PPI networks were identified using the STRING database, and the biomolecular interaction networks of related genes were visualized using Cytoscape. A modular analysis of the network was performed using the MCODE plugin to screen the crucial

genes in the PPI network. The prognostic value of the differentially expressed genes (DEGs) was analyzed using univariate and multivariate Cox analyses to obtain the genes related to the prognosis of the LGG patient sample, and $p < 0.05$ was considered to be significant.

## Identification of low-grade glioma prognosis-associated genes

A univariate Cox regression analysis was performed on the LGG patient sample DEGs to screen for genes associated with prognosis using the log-rank test. The prognostic differences between the high gene expression group and low gene expression group were analyzed using the R package "Kaplan–Meier," and $p < 0.05$ was considered to be significant.

## Nomogram construction

Independent prognostic factors were determined based on univariate and multivariate Cox regression models. Column plots were constructed using the R packages "rms" and "survival." A calibration curve was used to assess the accuracy of the line graphs, and $p < 0.05$ was considered to be significant.

## Cell culture

Human glioma cell lines HS683 and the normal human glial cell line HEB were selected and provided by Xiangya Medical College of Changsha Central South University (China). The cells were cultured in DMEM supplemented with 1% streptomycin/penicillin and 10% FBS under saturated humidity, 37°C, and 5% $CO_2$.

## Real-time quantitative polymerase chain reaction

RNA was extracted using a TRIzol kit (Invitrogen, USA), and cDNA was synthesized using a reverse transcription kit. Real-time quantitative polymerase chain reaction (RT-qPCR) was performed using a SYBR Premix Ex Taq kit (TaKaRa, Japan). The primer sequences used are listed in Table 1. Gene expression levels were quantified using the $2^{-\Delta\Delta CT}$ method. The experiment was repeated three times.

# Results

## Prognostic value of immune cells for predicting overall survival

The ssGSEA method was applied to the transcriptome of LGG samples to find the distribution of 22 immune cell types (Figure 1A). Based on the LASSO model, the TCRS can be calculated as follows when lambda.min = 0.0257: RiskScore = 6.74934179805055 * CD4_naive + 6.83207575533917 * Tr1 − 1.22576340775653 * Th1 + 1.31766352136828 * NKT + 2.38814433064296 * B_cell + 3.26414059327394 * Monocyte + 3.2630292767646 * CD4_T (Figure 1B). Furthermore, using univariate and multivariate Cox analyses, the optimal cutoff value of 2.649 was obtained, and the results showed that Tr1, CD4_naive, and B_cell were related to the OS of patients with LGG (Figure 1C). The LGG patient sample was divided into a high-risk–score group and a low-risk–score group based on the optimal cutoff value (Figure 1D). Kaplan–Meier survival curves figured out that patients in the low-risk–score group had greater OS than those in the high-risk–score group (Figure 1E, $p < 0.05$). The ROC curves showed that TCRS had a better predictive value for 1-, 3-, and 5-year OS (Figure 1F). To verify the validity of the above model, the ESTIMATE algorithm was used to calculate the stromal score, immune score, and ESTIMATE score of LGG samples. The results showed that the stromal score, immune score, and ESTIMATE score were higher in the high-risk–score group than in the low-risk–score group (Figure 1G, $p < 0.05$). There was a significant difference in the expression of immune genes between the high-risk–score group and the low-risk–score group (Figure 1H, $p < 0.05$).

## Analysis of differences

Differential analysis was performed on the high-risk–score group and the low-risk–score group, and 799 upregulated genes

TABLE 1 RT-qPCR primer sequence.

| Gene | Forward 5′→3′ | Reverse 5′→3′ |
|---|---|---|
| CD2 | TCAAGAGAGGGTCTCAAAACCA | CCATTCATTACCTCACAGGTCAG |
| SPN | GCTGGTGGTAAGCCCAGAC | GGCTCGCTAGTAGAGACCAAA |
| IL18 | TCTTCATTGACCAAGGAAATCGG | TCCGGGGTGCATTATCTCTAC |
| PTPRC | ACCACAAGTTTACTAACGCAAGT | TTTGAGGGGGATTCCAGGTAAT |
| GZMA | TCTCTCTCAGTTGTCGTTTCTCT | GCAGTCAACACCCAGTCTTTTG |
| TLR7 | TCCTTGGGGCTAGATGGTTTC | TCCACGATCACATGGTTCTTTG |
| GAPDH | CTGGGCTACACTGAGCACC | AAGTGGTCGTTGAGGGCAATG |

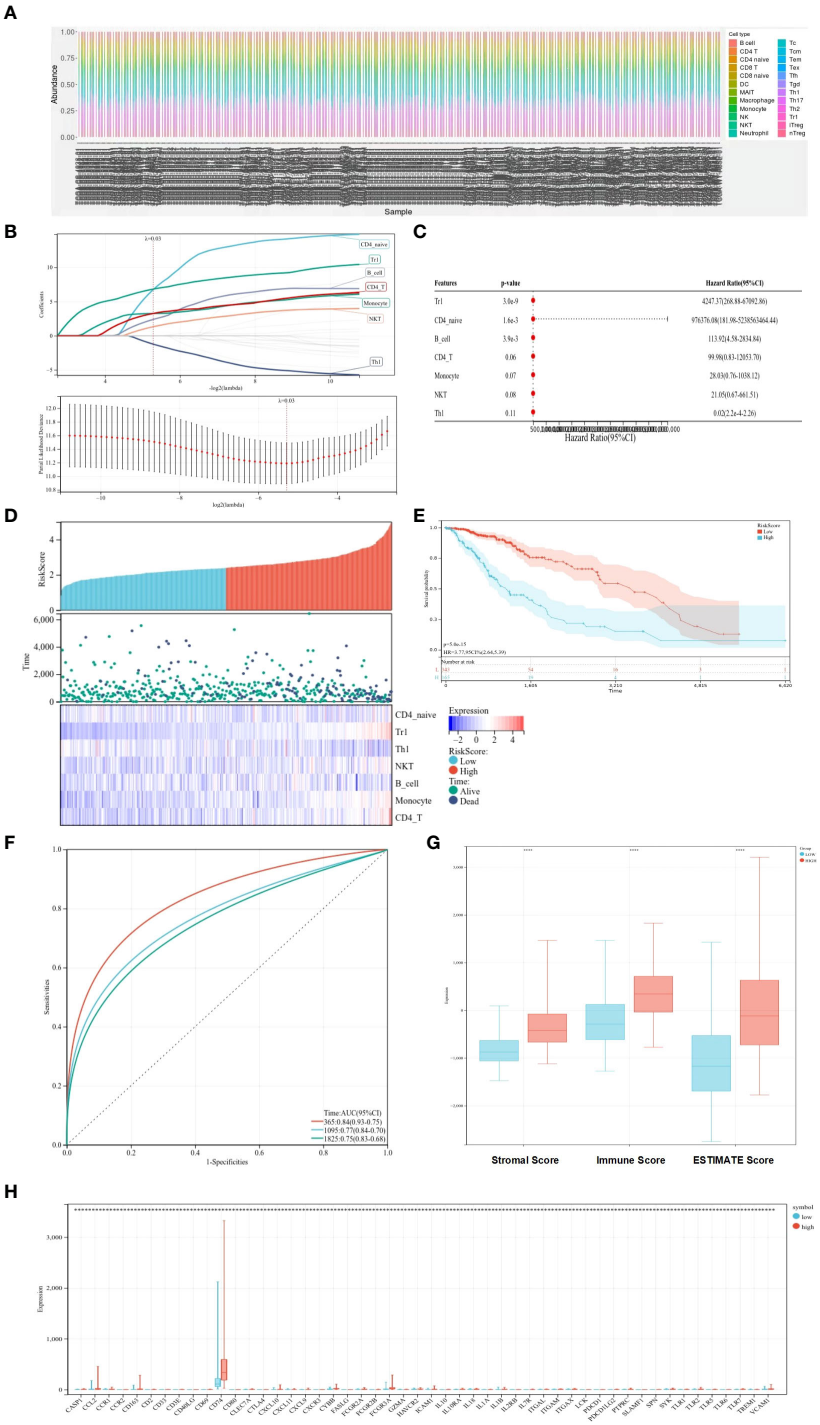



FIGURE 1
Prognostic value of immune cells for predicting OS. **(A)** Distribution of immune cell types. **(B)** LASSO coefficient profiles. **(C)** Univariate and multifactorial Cox analyses used to identify OS-related immune cells. **(D)** Heat map of survival status and expression in different risk score groups. **(E)** Kaplan–Meier survival curves showing survival in the different risk score groups. **(F)** ROC curve showing the predictive value of TCRS for 1-, 3-, and 5-year OS. **(G)** Box plots showing differences in the stromal score, immune score, and ESTIMATE score between different risk score groups. **(H)** Expression of immune genes in different risk-score groups compared with low group. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

and 348 downregulated genes were identified (Figures 2A, B). KEGG analysis showed that DEGs were enriched mainly in *Staphylococcus aureus* infection, complement and coagulation cascades, tuberculosis, and other pathways (Figure 2C). GO analysis showed that DEGs were enriched mainly in the immune system process, immune response, and other terms (Figure 2D).

## Establishment of protein–protein interaction network

The most important module in the PPI network consisted of 50 nodes and 868 edges (Figures 3A, B). The prognostic value of the 50 DEGs was analyzed using univariate and multivariate Cox analyses, and a total of nine prognostic-related genes were obtained, including CD2, SPN, IL18, CLEC7A, PTPRC, TLR2, GZMA, CD163, and TLR7 (Tables 2, 3). The expression of all nine genes was different in the different risk-score groups (Figure 3C, $p < 0.05$).

## Batch survival analysis to identify low-grade glioma prognosis-related genes

Kaplan–Meier survival curves showed that OS was significantly worse in the group with a high expression of CD2, SPN, IL18, CLEC7A, PTPRC, TLR2, GZMA, CD163, and TLR7 (Figure 4, $p < 0.05$).

## Construction of nomogram prognostic model

A total of nine prognosis-related genes, age, sex, and race were used as variables, and univariate and multivariate Cox

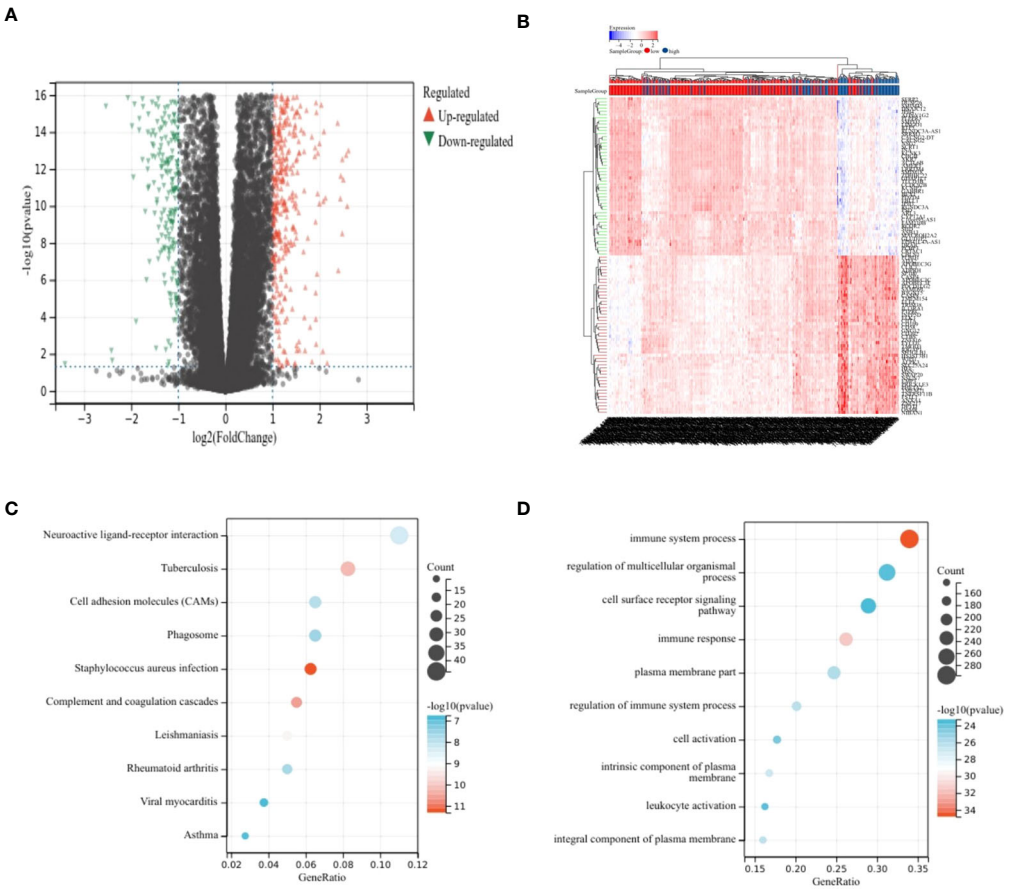



FIGURE 2
Analysis of differences. **(A, B)** Volcano plot showing DEGs in different risk score groups. **(C)** Bubble plot from KEGG analysis showing major enrichment of DEGs. **(D)** Bubble plot from GO analysis showing major enrichment of DEGs.

10.3389/fonc.2022.1056623

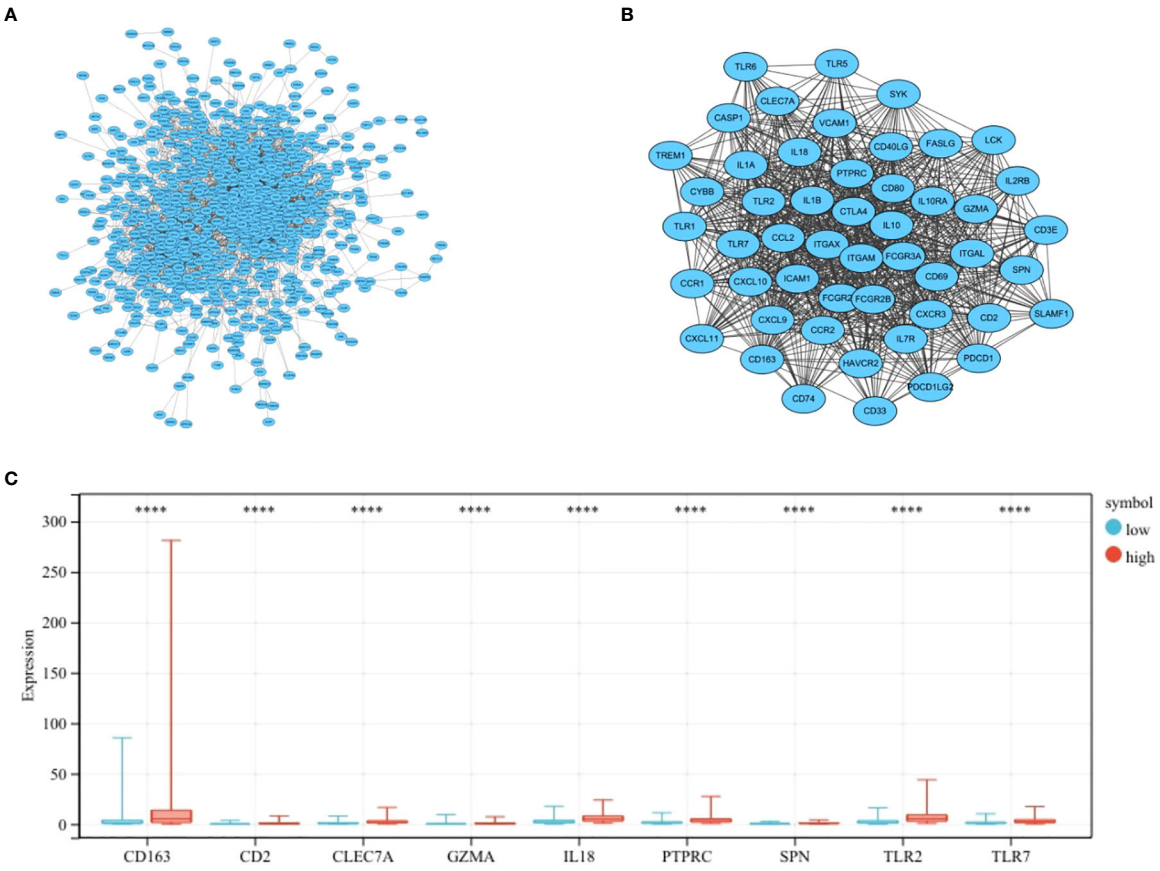



FIGURE 3
Establishment of the PPI network. **(A)** PPI network showing protein interactions between nodes. **(B)** Most important modules in the PPI network. **(C)** Expression of nine prognosis-related genes in the different risk-score groups compared to low group. ****$p < 0.0001$.

TABLE 2 Univariate Cox regression analysis.

| Tag | HR | Lower 95% | Upper 95% | Likelihood | Log-rank | Wald |
|---|---|---|---|---|---|---|
| PDCD1LG2 | 1.520116292 | 1.369792153 | 1.686937347 | 7.34E−10 | 8.80E−17 | 3.20E−15 |
| CD2 | 1.452943277 | 1.311213594 | 1.609992586 | 2.66E−08 | 1.18E−15 | 9.75E−13 |
| CD40LG | 75.90563174 | 23.15108433 | 248.8723573 | 2.02E−08 | 3.88E−15 | 8.93E−13 |
| CASP1 | 1.213978947 | 1.154007576 | 1.277066905 | 4.51E−10 | 7.95E−15 | 6.31E−14 |
| LCK | 3.6385036 | 2.5422305 | 5.207516961 | 2.25E−08 | 9.35E−15 | 1.66E−12 |
| SPN | 2.622728033 | 2.023027009 | 3.400202915 | 9.09E−11 | 5.14E−14 | 3.36E−13 |
| CD3E | 1.662195837 | 1.43668872 | 1.923099249 | 9.79E−08 | 1.11E−13 | 8.44E−12 |
| CXCR3 | 7.589077041 | 4.178056523 | 13.78489975 | 6.80E−08 | 3.83E−13 | 2.83E−11 |
| CXCL11 | 1.181252395 | 1.120596448 | 1.245191542 | 2.70E−06 | 1.01E−12 | 5.89E−10 |
| FCGR2B | 1.976307546 | 1.61004357 | 2.425891813 | 1.48E−06 | 2.41E−12 | 7.31E−11 |
| CCR2 | 4.555315166 | 2.836910836 | 7.314609962 | 3.53E−07 | 2.11E−11 | 3.49E−10 |
| SLAMF1 | 84.49490994 | 19.95938776 | 357.6958317 | 1.67E−06 | 7.08E−11 | 1.68E−09 |
| FASLG | 132.1604317 | 24.87050552 | 702.29291 | 5.23E−06 | 9.26E−10 | 9.99E−09 |
| IL7R | 1.275505655 | 1.157951312 | 1.404994027 | 0.000466979 | 2.21E−09 | 8.11E−07 |
| IL18 | 1.142776013 | 1.092357372 | 1.195521767 | 1.01E−07 | 2.84E−09 | 6.75E−09 |

*(Continued)*

TABLE 2 Continued

| Tag | HR | Lower 95% | Upper 95% | Likelihood | Log-rank | Wald |
|---|---|---|---|---|---|---|
| FCGR2A | 1.084849239 | 1.055299647 | 1.115226254 | 8.86E−07 | 4.07E−09 | 7.47E−09 |
| CLEC7A | 1.216203839 | 1.13925352 | 1.29835173 | 2.67E−06 | 4.27E−09 | 4.37E−09 |
| CD80 | 4.563367614 | 2.473199935 | 8.419992129 | 0.000541267 | 7.15E−09 | 1.19E−06 |
| IL10 | 1.972319731 | 1.548931374 | 2.511438005 | 4.07E−05 | 1.49E−08 | 3.61E−08 |
| ITGAX | 1.116224672 | 1.073805236 | 1.160319839 | 7.99E−07 | 1.49E−08 | 2.66E−08 |
| PDCD1 | 2.892951433 | 1.970626433 | 4.24695815 | 1.00E−05 | 2.01E−08 | 5.86E−08 |
| CXCL10 | 1.027817602 | 1.016827153 | 1.038926841 | 9.43E−05 | 3.30E−08 | 5.67E−07 |
| CD69 | 1.39826594 | 1.230533292 | 1.588862041 | 1.36E−05 | 1.79E−07 | 2.72E−07 |
| PTPRC | 1.111688395 | 1.066629168 | 1.158651128 | 2.33E−05 | 3.24E−07 | 5.29E−07 |
| ITGAL | 1.210034564 | 1.121282744 | 1.305811272 | 2.64E−05 | 5.14E−07 | 9.33E−07 |
| CD74 | 1.000982762 | 1.000574112 | 1.001391579 | 4.37E−05 | 1.55E−06 | 2.42E−06 |
| HAVCR2 | 1.060795489 | 1.035001734 | 1.087232063 | 1.48E−05 | 1.93E−06 | 2.61E−06 |
| TLR2 | 1.061560111 | 1.034981622 | 1.088821139 | 0.000124234 | 3.77E−06 | 3.88E−06 |
| CXCL9 | 1.085836678 | 1.042788787 | 1.130661652 | 0.00175701 | 1.48E−05 | 6.61E−05 |
| TLR1 | 1.22612022 | 1.11399963 | 1.349525397 | 0.0001822 | 2.26E−05 | 3.09E−05 |
| TLR6 | 1.706751561 | 1.32659079 | 2.195854905 | 0.000152853 | 2.84E−05 | 3.21E−05 |
| GZMA | 1.207293281 | 1.097999303 | 1.327466295 | 0.001108677 | 4.71E−05 | 9.98E−05 |
| IL2RB | 1.601590896 | 1.2462979 | 2.058170359 | 0.001969111 | 0.000133686 | 0.000232795 |
| CTLA4 | 2.287753151 | 1.470110919 | 3.560149383 | 0.004282951 | 0.0001396 | 0.000244672 |
| CCR1 | 1.057674123 | 1.027356743 | 1.08888617 | 0.001114016 | 0.000179795 | 0.000157572 |
| CD163 | 1.008029664 | 1.003566842 | 1.012512332 | 0.004829108 | 0.000181982 | 0.00041132 |
| VCAM1 | 1.018221134 | 1.007751055 | 1.028799993 | 0.002950449 | 0.000549683 | 0.000616844 |
| ITGAM | 1.124513823 | 1.051511646 | 1.202584244 | 0.001181054 | 0.000597246 | 0.000611085 |
| IL10RA | 1.076085559 | 1.031286061 | 1.122831167 | 0.003457336 | 0.000630622 | 0.000725157 |
| CD33 | 1.585029901 | 1.213991796 | 2.069470154 | 0.001298347 | 0.000677602 | 0.000711565 |
| IL1A | 1.439369434 | 1.15739573 | 1.790039752 | 0.005193204 | 0.000983137 | 0.001060541 |
| CCL2 | 1.008871682 | 1.003453323 | 1.014319299 | 0.005649978 | 0.00101128 | 0.001306026 |
| FCGR3A | 1.006650517 | 1.002547706 | 1.010770119 | 0.005185133 | 0.001310516 | 0.001467292 |
| CYBB | 1.01873458 | 1.006276509 | 1.031346887 | 0.006484694 | 0.003259481 | 0.003110249 |
| ICAM1 | 1.025125663 | 1.006887071 | 1.043694626 | 0.019545403 | 0.005367287 | 0.006742245 |
| TLR7 | 1.099084271 | 1.024661061 | 1.178912989 | 0.013378943 | 0.008035934 | 0.008267125 |
| SYK | 1.060559802 | 1.015264986 | 1.10787539 | 0.011451597 | 0.008064287 | 0.008284398 |
| TLR5 | 1.185162561 | 1.033717809 | 1.358794715 | 0.024620574 | 0.015399611 | 0.014876759 |
| TREM1 | 1.142466409 | 1.003308082 | 1.300925927 | 0.087955159 | 0.04046097 | 0.044452378 |

analyses showed that CD2, SPN, IL18, PTPRC, GZMA, and TLR7 were independent prognostic factors for LGG (Figures 5A, B, $p < 0.05$). The nomogram prediction model had good predictive power for 1-, 3-, and 5-year prognoses in LGG patients (Figures 5C, D).

## Expression validation of crucial genes

CD2, SPN, IL18, PTPRC, GZMA, and TLR7 genes were validated using an RT-qPCR assay. The results showed that the expression of CD2, SPN, IL18, PTPRC, GZMA, and TLR7 was upregulated in the tumor group compared with the normal group (Figure 6, $p < 0.05$).

# Discussion

LGG have become a focus of brain tumor research. Surgical resection is currently the main treatment modality for LGG, but because LGG frequently presents with infiltrative growth and is mostly found in functional areas of the brain, extended resection by surgery is greatly limited, and the LGG patients still do not have an ideal prognosis (9, 10). As a result, there is a need to construct effective prognostic prediction models for LGG. Immune infiltration plays a role in the development of LGG (11). Because LGG survive in a complex tumor microenvironment, it brings a serious challenge for clinical assessment and treatment of LGG.

Wu et al. constructed an immune risk score signature (IRSS) using the LASSO model, and the IRSS included six relevant

TABLE 3 Multivariate Cox regression analysis.

| Tag | Exp(coef) | *p*-value | Lower 95% | Upper 95% | Coef |
|---|---|---|---|---|---|
| CD2 | 4.411465527 | 0.036529593 | 1.0974606 | 17.73278066 | 1.484206953 |
| SPN | 5.049279253 | 0.001727158 | 1.833888852 | 13.90227164 | 1.619245511 |
| IL18 | 1.253022008 | 0.005611449 | 1.068163609 | 1.469872349 | 0.22555824 |
| CLEC7A | 0.672503094 | 0.042080937 | 0.458728497 | 0.985899969 | −0.396748567 |
| PTPRC | 1.476656053 | 0.003460154 | 1.137091104 | 1.917623919 | 0.389780108 |
| TLR2 | 0.839361265 | 0.021333758 | 0.723100033 | 0.974315172 | −0.175114075 |
| GZMA | 0.147221226 | 0.000464471 | 0.050361944 | 0.430366421 | −1.915818884 |
| CD163 | 1.029691184 | 0.010471213 | 1.006880376 | 1.05301877 | 0.029258936 |
| TLR7 | 0.656053832 | 0.039488752 | 0.439226813 | 0.979918843 | −0.421512432 |

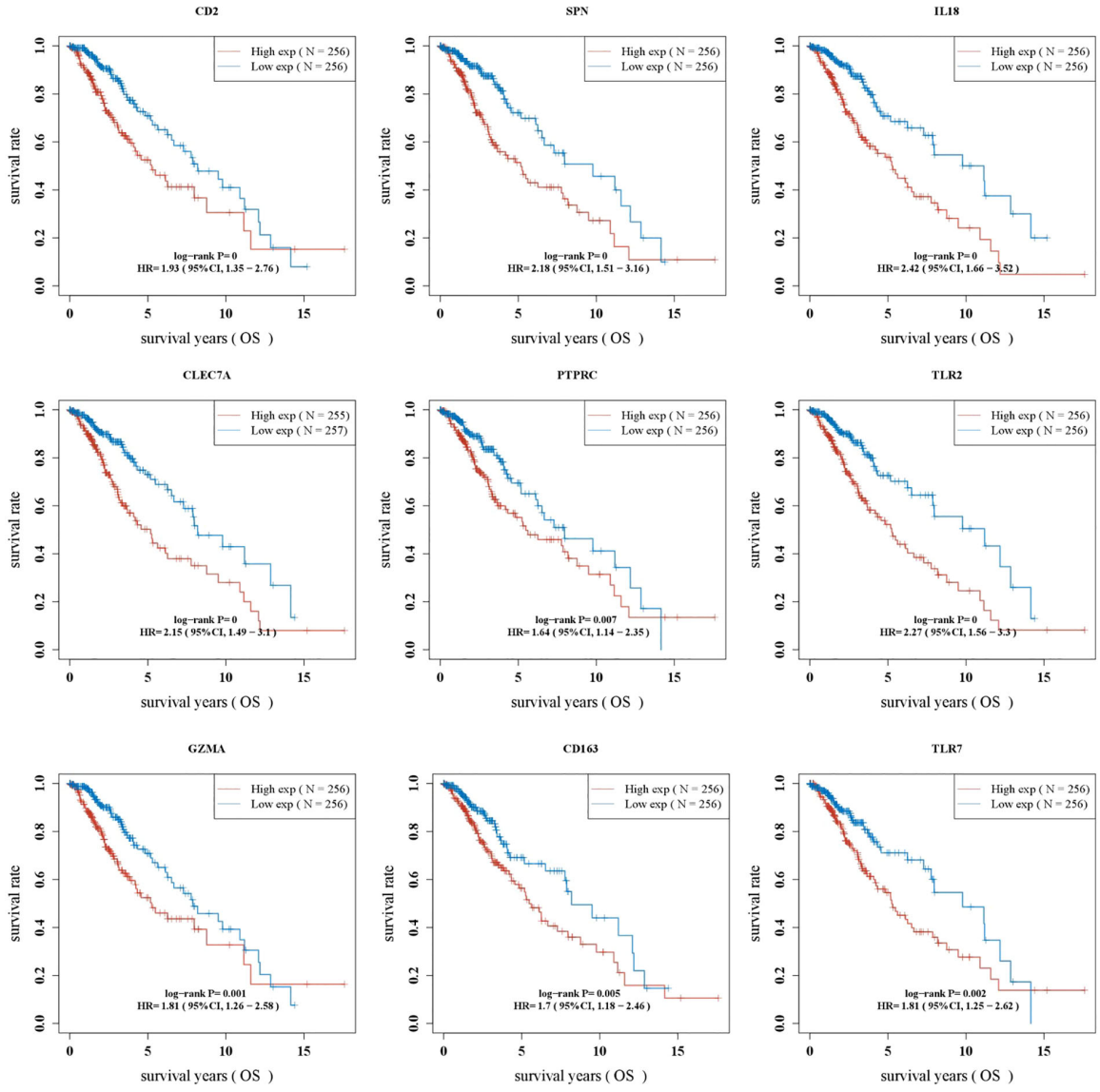



FIGURE 4
Batch survival analysis identifying LGG prognosis-related genes.

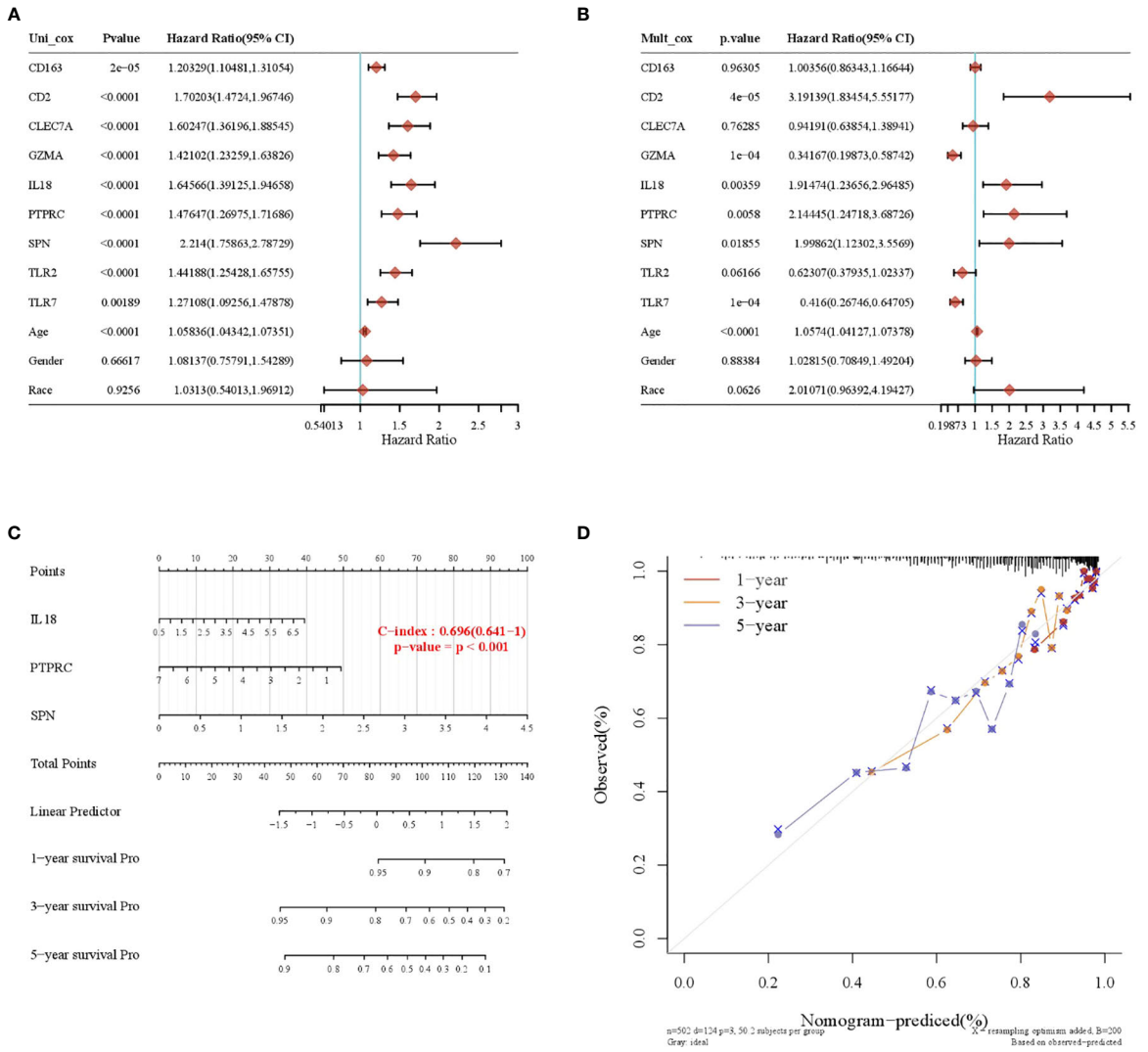



FIGURE 5
Construction of a nomogram prognostic model. **(A)** Univariate Cox regression model. **(B)** Multivariate Cox regression model. **(C)** Column-line plot for predicting survival time. **(D)** Calibration curve for predicting a 1-, 3-, and 5-year prognosis.

immune genes that were good predictors of prognosis in LGG patients. Moreover, the immune infiltration results showed that the genetic profile correlated with innate immune cytopenia (12). Zhang et al. found that using LASSO and multivariate Cox regression analyses, they were able to obtain six immune genes that comprise a risk model and may be involved in the process of neoantigen presence and triggering immune responses (13). The present study differs from previous LGG-related literature in that we obtained crucial genes that may be associated with LGG prognosis based on ssGSEA and by constructing TCRS. In this study, ssGSEA was performed on 22 immune gene sets to calculate immune-based prognostic scores. The prognostic value of the 22 immune cells for predicting OS was assessed using LASSO and univariate and multivariate Cox analyses. Subsequently, we constructed a validated TCRS to identify immune subtypes and inflammatory immune features in LGG patients. Trl, CD4_naive, and B_cells were found to be related to OS in LGG patients by LASSO and by univariate and multifactorial Cox analyses. We divided the LGG patient sample into a high-risk–score and low-risk–score group according to the optimal cutoff value. Kaplan–Meier survival curves displayed that patients in the low-risk–score group had higher OS. The ROC curve showed that TCRS was able to identify the immune subtype of LGG and had a better predictive value for 1-, 3-, and 5-year OS. We then identified DEGs in the high-risk–score and low-risk–score groups and obtained 799 upregulated genes and 348 downregulated genes. KEGG and GO analyses showed that DEGs were enriched mainly in immune-related

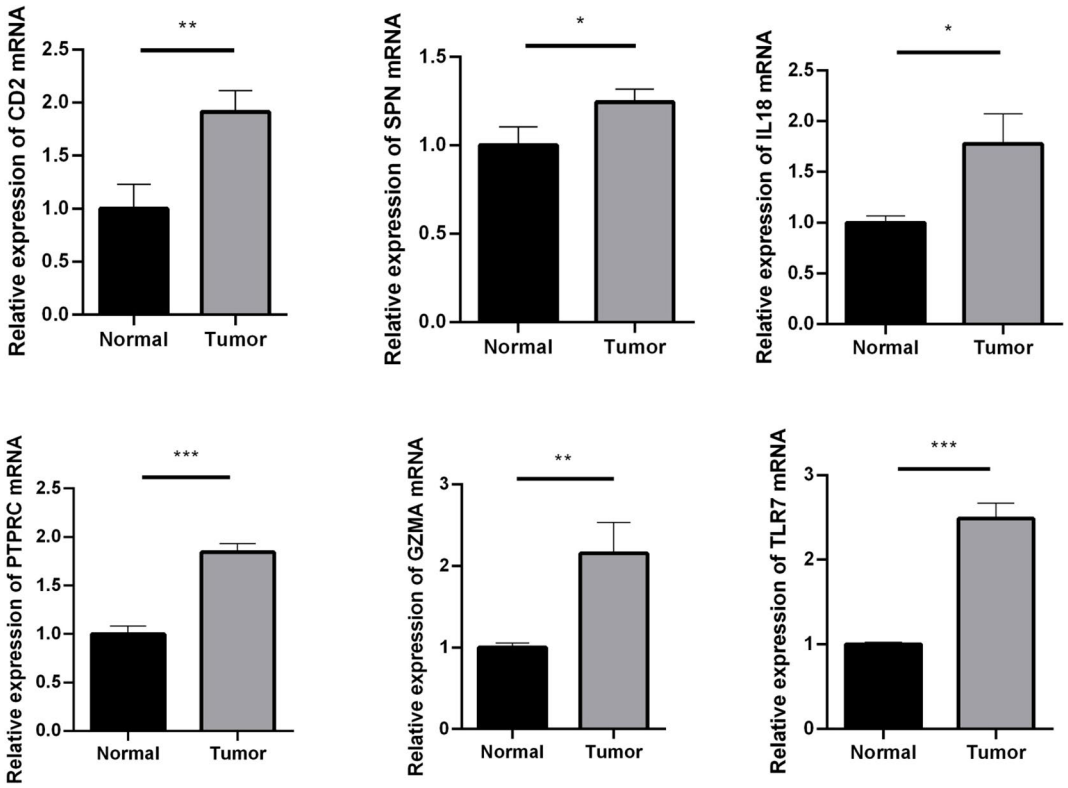



FIGURE 6
RT-qPCR assay used to verify the expression of crucial genes. Compared with the normal group. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

processes. We constructed a PPI network using Cytoscape and then identified the top 50 crucial genes. Subsequently, nine DEGs were found to be significantly related to OS based on univariate and multivariate Cox analyses. OS was significantly worse in the high-expression group of CD2, SPN, IL18, CLEC7A, PTPRC, TLR2, GZMA, CD163, and TLR7 compared to the low-expression group. This indicates that these nine crucial genes may be related to the process of immune cells affecting OS. Finally, we constructed a prognostic nomogram model that revealed CD2, SPN, IL18, PTPRC, GZMA, and TLR7 to be independent prognostic factors for LGG. Columnar plots and ROC curves were used to verify that the model was reasonably accurate in predicting the prognosis of LGG patients at 1, 3, and 5 years. In addition, this study used an RT-qPCR assay to verify the bioinformatics results, revealing that CD2, SPN, IL18, PTPRC, GZMA, and TLR7 were highly expressed in LGG.

CD2 is expressed on the surface of all peripheral blood T cells, more than 95% of human thymocytes, most NK cells, and some malignant B cells and may indirectly reflect the immune function of the body's cells (14). Chen's team found that CD2 was upregulated in breast cancer samples and that CD2 immunomodulation contributed to the mitigation of disease progression and could be used as an immunomodulatory agent in clinical treatment (15). SPN, alias CD43, encodes a glycoprotein that is expressed on the membrane surface of normal and tumorigenic T cells. SPN can regulate intercellular adhesion, intracellular signaling, cell proliferation, and apoptosis (16). Gao's team found that miR-129-5p was beneficial in delaying the malignant progression of clear cell renal carcinoma by targeting the downregulation of SPN (17). IL18 is a proinflammatory cytokine with important functions, such as induction of angiogenesis and regulation of immune function, and is involved in the progress of many inflammatory diseases, immune disorders, and tumors (18). Park's team speculated that IL18 contributes to the poor prognosis of triple-negative breast cancer patients by inducing immunosuppression of PD-1 expression on NK cells (19). PTPRC, alias CD45, is an antigen of leukocytes that is common on their surface. PTPRC acts as a key molecule of signal transduction on cell membranes and positively regulates T-cell antigen receptor signaling (20, 21). PTPRC can affect the processes of cell growth, differentiation, and mitosis, and it has been suggested that PTPRC may have a role in regulating the MAPK/ERK signaling pathway, with implications for cervical carcinogenesis and patient prognosis (22). GZMA is a serine protease that is mainly secreted by NK cells and cytotoxic T lymphocytes and delivered to bacterial or virally infected target

cells (23). GZMA mediates apoptosis and cell scorching, induces the release of inflammatory factors, is involved in the body's defense against pathogenic bacterial infections, and is associated with the development of certain autoimmune diseases (24). Santiago's team showed that GZMA can be involved in tumor development and is a potential prognostic target for various cancers; this may be due to the ability of extracellular GZMA to promote the production of NF-κB-dependent IL6 in macrophages (25). TLR7 is an important pattern recognition receptor in natural immunity, playing a vital role in the body's resistance to pathogenic infections and acting as a key line of defense for the immune system (26). Studies have shown that TLR7 can be used as a reliable marker of poor prognosis, which reveals the high expression of TLR7 in patients with non-small cell lung cancer, which is related to the inflammatory process of TLR7 signaling (27).

In summary, CD2, SPN, IL18, PTPRC, GZMA, and TLR7 were identified as independent prognostic factors for LGG, and these genes may be potential indicators of the regulation of the immune microenvironment. This may contribute to the clinical development of more effective therapeutic agents. To ensure the accuracy of the results, a larger sample is needed, and other datasets should be used for validation. In addition, the mechanisms of action of CD2, SPN, IL18, PTPRC, GZMA, and TLR7 in LGG require further exploration.

## Data availability statement

Publicly available datasets were analyzed in this study. These data can be found in the TCGA database (https://tcga-data.nci.nih.gov/tcga/).

## Author contributions

YH: Conceptualization, Methodology, Formal analysis, Writing - Original Draft. ZL: Methodology, Formal analysis, Writing - Review & Editing. ST: Conceptualization, Validation, Data Curation, Writing - Original Draft. All authors contributed to the article and approved the submitted version.

## Funding

This project is supported by Hainan Province Clinical Medical Center (QWYH202175).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Ostrom QT, Bauchet L, Davis FG, Deltour I, Fisher JL, Langer CE, et al. The epidemiology of glioma in adults: a "state of the science" review. *Neuro-oncology* (2014) 16(7):896–913. doi: 10.1093/neuonc/nou087

2. Molinaro AM, Taylor JW, Wiencke JK, Wrensch MR. Genetic and molecular epidemiology of adult diffuse glioma. *Nat Rev Neurol* (2019) 15(7):405–17. doi: 10.1038/s41582-019-0220-2

3. Tom MC, Cahill DP, Buckner JC, Dietrich J, Parsons MW, Yu JS. Management for different glioma subtypes: are all low-grade gliomas created equal? *Am Soc Clin Oncol Educ Book* (2019) 39:133–45. doi: 10.1200/EDBK_238353

4. Brown TJ, Bota DA, van Den Bent MJ, Brown PD, Maher E, Aregawi D, et al. Management of low-grade glioma: a systematic review and meta-analysis. *Neuro-oncol practice.* (2019) 6(4):249–58. doi: 10.1093/nop/npy034

5. Morshed RA, Young JS, Hervey-Jumper SL, Berger MS. The management of low-grade gliomas in adults. *J Neurosurg Sci* (2019) 63(4):450–7. doi: 10.23736/S0390-5616.19.04701-5

6. Gieryng A, Pszczolkowska D, Walentynowicz KA, Rajan WD, Kaminska B. Immune microenvironment of gliomas. *Lab Invest* (2017) 97(5):498–518. doi: 10.1038/labinvest.2017.19

7. Miyauchi JT, Tsirka SE. Advances in immunotherapeutic research for glioma therapy. *J neurology.* (2018) 265(4):741–56. doi: 10.1007/s00415-017-8695-5

8. Lim M, Xia Y, Bettegowda C, Weller M. Current state of immunotherapy for glioblastoma. *Nat Rev Clin Oncol* (2018) 15(7):422–42. doi: 10.1038/s41571-018-0003-5

9. Delgado-López P, Corrales-García E, Martino J, Lastra-Aras E, Dueñas-Polo M. Diffuse low-grade glioma: a review on the new molecular classification, natural history and current management strategies. *Clin Trans Oncol* (2017) 19(8):931–44. doi: 10.1007/s12094-017-1631-4

10. Jooma R, Waqas M, Khan I. Diffuse low-grade glioma–changing concepts in diagnosis and management: a review. *Asian J neurosurg* (2019) 14(02):356–63. doi: 10.4103/ajns.AJNS_24_18

11. Haddad AF, Young JS, Oh JY, Okada H, Aghi MK. The immunology of low-grade gliomas. *Neurosurg Focus.* (2022) 52(2):E2. doi: 10.3171/2021.11.FOCUS21587

12. Wu Y, Peng Z, Gu S, Wang H, Xiang W. A risk score signature consisting of six immune genes predicts overall survival in patients with lower-grade gliomas. *Comput Math Methods Med* (2022) 2022:2558548. doi: 10.1155/2022/2558548

13. Zhang M, Wang X, Chen X, Zhang Q, Hong J. Novel immune-related gene signature for risk stratification and prognosis of survival in lower-grade glioma. *Front Genet* (2020) 11:363. doi: 10.3389/fgene.2020.00363

14. Demetriou P, Abu-Shah E, Valvo S, McCuaig S, Mayya V, Kvalvaag A, et al. A dynamic CD2-rich compartment at the outer edge of the immunological synapse boosts and integrates signals. *Nat Immunol* (2020) 21(10):1232–43. doi: 10.1038/s41590-020-0770-x

15. Chen Y, Meng Z, Zhang L, Liu F. CD2 is a novel immune-related prognostic biomarker of invasive breast carcinoma that modulates the tumor microenvironment. *Front Immunol* (2021) 12:664845. doi: 10.3389/fimmu.2021.664845

16. Tuccillo FM, de Laurentiis A, Palmieri C, Fiume G, Bonelli P, Borrelli A, et al. Aberrant glycosylation as biomarker for cancer: Focus on CD43. *BioMed Res Int* (2014) 2014:742831. doi: 10.1155/2014/742831

17. Gao B, Wang L, Zhang N, Han M, Zhang Y, Liu H, et al. miR-129-5p inhibits clear cell renal cell carcinoma cell proliferation, migration and invasion by targeting SPN. *Cancer Cell Int* (2021) 21(1):1–9. doi: 10.1186/s12935-021-01820-3

18. Haghshenas MR, Saffarian A, Khademolhosseini A, Dehghanian AR, Ghaderi A, Jahromi AS. Simultaneous increase in serum levels of IL-37 and IL-18 binding protein in low-grade and high-grade brain tumors. *Asian Pacific J Cancer Prev* (2022) 23(8):2851–6. doi: 10.31557/APJCP.2022.23.8.2851

19. Park IH, Yang HN, Lee KJ, Kim T-S, Lee ES, Jung S-Y, et al. Tumor-derived IL-18 induces PD-1 expression on immunosuppressive NK cells in triple-negative breast cancer. *Oncotarget* (2017) 8(20):32722. doi: 10.18632/oncotarget.16281

20. Junghans V, Santos AM, Lui Y, Davis SJ, Jönsson P. Dimensions and interactions of large T-cell surface proteins. *Front Immunol* (2018) 9:2215. doi: 10.3389/fimmu.2018.02215

21. Courtney AH, Shvets AA, Lu W, Griffante G, Mollenauer M, Horkova V, et al. CD45 functions as a signaling gatekeeper in T cells. *Sci Signaling* (2019) 12 (604):eaaw8151. doi: 10.1126/scisignal.aaw8151

22. Chen S, Gao C, Wu Y, Huang Z. Identification of prognostic miRNA signature and lymph node metastasis-related key genes in cervical cancer. *Front Pharmacol* (2020) 11:544. doi: 10.3389/fphar.2020.00544

23. Shimizu K, Yamasaki S, Sakurai M, Yumoto N, Ikeda M, Mishima-Tsumagari C, et al. Granzyme a stimulates pDCs to promote adaptive immunity *via* induction of type I IFN. *Front Immunol* (2019) 10:1450. doi: 10.3389/fimmu.2019.01450

24. Velaga S, Ukena SN, Dringenberg U, Alter C, Pardo J, Kershaw O, et al. Granzyme a is required for regulatory T-cell mediated prevention of gastrointestinal graft-versus-host disease. *PloS One* (2015) 10(4):e0124927. doi: 10.1371/journal.pone.0124927

25. Santiago L, Castro M, Sanz-Pamplona R, Garzón M, Ramirez-Labrada A, Tapia E, et al. Extracellular granzyme a promotes colorectal cancer development by enhancing gut inflammation. *Cell Rep* (2020) 32(1):107847. doi: 10.1016/j.celrep.2020.107847

26. Schön M, Schön M. TLR7 and TLR8 as targets in cancer therapy. *Oncogene* (2008) 27(2):190–9. doi: 10.1038/sj.onc.1210913

27. Chatterjee S, Crozet L, Damotte D, Iribarren K, Schramm C, Alifano M, et al. TLR7 promotes tumor progression, chemotherapy resistance, and poor clinical outcomes in non–small cell lung CancerRole of TLR7 in physiopathology of lung cancer. *Cancer Res* (2014) 74(18):5008–18. doi: 10.1158/0008-5472.CAN-13-2698

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 January 2023
DOI 10.3389/fonc.2022.1057383

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Taotao Liang,
Henan Provincial Cancer Hospital, China
Sajad Najafi,
Shahid Beheshti University of Medical Sciences, Iran

*CORRESPONDENCE
Yun Jiang
1462500228@qq.com

†These authors share first authorship

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

RECEIVED 29 September 2022
ACCEPTED 19 October 2022
PUBLISHED 16 January 2023

CITATION
Tian Y, Chen L'e and Jiang Y (2023) LASSO-based screening for potential prognostic biomarkers associated with glioblastoma.
*Front. Oncol.* 12:1057383.
doi: 10.3389/fonc.2022.1057383

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

# LASSO-based screening for potential prognostic biomarkers associated with glioblastoma

Yin Tian[1†], Li'e Chen[2†] and Yun Jiang[3*]

[1]Department of Pediatric Surgery, Jingzhou Central Hospital, Jingzhou Hospital Affiliated to Yangtze University, Jingzhou, Hubei Province, China, [2]Department of Pathology, Sanya Central Hospital (Hainan Third People's Hospital), Sanya, Hainan Province, China, [3]Department of Ultrasound Diagnosis, Hubei Provincial Hospital of Integrated Chinese & Western Medicine, Wuhan, Hubei Province, China

**Background:** Glioblastoma is the most common malignancy of the neuroepithelium, yet existing research on this tumor is limited. LASSO is an algorithm of selected feature coefficients by which genes associated with glioblastoma prognosis can be obtained.

**Methods:** Glioblastoma-related data were selected from the Cancer Genome Atlas (TCGA) database, and information was obtained for 158 samples, including 153 cancer samples and five samples of paracancerous tissue. In addition, 2,642 normal samples were selected from the Genotype-Tissue Expression (GTEx) database. Whole-gene bulk survival analysis and differential expression analysis were performed on glioblastoma genes, and their intersections were taken. Finally, we determined which genes are associated with glioma prognosis. The STRING database was used to analyze the interaction network between genes, and the MCODE plugin under Cytoscape was used to identify the highest-scoring clusters. LASSO prognostic analysis was performed to identify the key genes. Gene expression validation allowed us to obtain genes with significant expression differences in glioblastoma cancer samples and paracancer samples, and glioblastoma independent prognostic factors could be derived by univariate and multivariate Cox analyses. GO functional enrichment analysis was performed, and the expression of the screened genes was detected using qRT-PCR.

**Results:** Whole-gene bulk survival analysis of glioblastoma genes yielded 607 genes associated with glioblastoma prognosis, differential expression analysis yielded 8,801 genes, and the intersection of prognostic genes with differentially expressed genes (DEG) yielded 323 intersecting genes. PPI analysis of the intersecting genes revealed that the genes were significantly enriched in functions such as the formation of a pool of free 40S subunits and placenta development, and the highest-scoring clusters were obtained using the MCODE plug-in. Eight genes associated with glioblastoma prognosis were identified based on LASSO analysis: RPS10, RPS11, RPS19, RSL24D1, RPL39L, EIF3E, NUDT5, and RPF1. All eight genes were found to be highly expressed in

the tumor by gene expression verification, and univariate and multivariate Cox analyses were performed on these eight genes to identify RPL39L and NUDT5 as two independent prognostic factors associated with glioblastoma. Both RPL39L and NUDT5 were highly expressed in glioblastoma cells.

**Conclusion:** Two independent prognostic factors in glioblastoma, RPL39L and NUDT5, were identified.

KEYWORDS

GBM, LASSO, independent prognostic factors, RPL39L, NUDT5

## Introduction

Glioblastoma is a malignant primary brain tumor disease and is among the most common malignant tumors of the central nervous system (CNS) (1). It accounts for 30% of all brain and CNS tumors and 80% of all malignant tumors (2). Despite the current multimodal treatment modalities, patients with glioblastoma have a poor prognosis, with a median survival time of only 14.6 months (3). The median age at diagnosis is 64 years (4), and the five-year mortality rate is higher than 90% (5). It has been shown that the incidence of glioblastoma is higher in men than in women (4, 6), the incidence is higher in developed countries than in developing countries (4), and the incidence of glioblastoma is higher in Asians, Latinos, and whites (7). Glioblastomas are aggressive tumors with a median survival time of only three months if left untreated (8). Surgery, radiation, and chemotherapy are available to improve the survival rates of glioblastoma patients.

The clinical treatment of glioblastoma may be facilitated by identifying genes and independent factors associated with glioblastoma prognosis. In this study, glioblastoma related data were selected from the Cancer Genome Atlas (TCGA) database, which contains 153 tumor samples and 2,647 paraneoplastic and GTEx undiseased tissues. Whole-gene bulk survival analysis of glioblastoma genes was performed to identify genes associated with glioblastoma prognosis. Differential expression analysis of the genes was conducted after whole-gene bulk survival analysis was performed with genes from normal samples. The interaction network between genes was analyzed using Metascape, and the highest- scoring clusters were identified using the MCODE plug-in. GO functional enrichment analysis was performed, and LASSO prognostic analysis was performed to identify key genes. Gene expression validation was performed to identify genes with significant expression differences in glioblastoma cancer samples compared with paracancerous samples, from which independent prognostic factors associated with glioblastoma prognosis were identified. Finally, qRT-PCR was used to verify the expression of the genes.

Due to the heterogeneity and complex pathogenesis of glioblastoma, the disease is still incurable (9). In this study, we explore the prognosis-related genes and independent prognostic factors of glioblastoma based on the LASSO bioinformatics analysis method.

## Methods and materials

### Sample source

This study is an exploratory study based on TCGA (https://portal.gdc.cancer.gov/) dataset, which contains 153 tumor samples, 5 paracancerous tissues, and 2647 normal samples of undiseased tissues retrieved from Genotype-Tissue Expression (GTEx, https://commonfund.nih.gov/GTEx) undiseased tissues.The 153 tumor samples were analyzed along with five samples of paracancerous tissues after obtaining gene expression matrices and clinical information data to identify genes associated with glioblastoma prognosis.

### Glioblastoma whole-gene bulk survival analysis

RNA-seq data of glioblastoma and corresponding clinical sample information were obtained from the TCGA, and the whole-gene bulk survival of the dataset was analyzed using the R package ggplot.2 Images were plotted using the R package forestplot. The data obtained from the all-gene bulk survival analysis could be used for subsequent analyses. A value of $p<0.05$ was considered statistically significant.

## Glioblastoma differential expression gene analysis

Clinical information was used to classify the samples into disease and control groups, and the Limma software package for R computing was used to analyze the differential expression of the study mRNAs. The results of the differential expression analysis for each data set are presented by volcano plots, and a Venn diagram shows the overlapping portion of differentially expressed genes (DEG) in the two groups; the overlapping genes can be used for subsequent analysis. The screening threshold for DEG is P<0.05, |log2FC|>1.

## Glioblastoma protein–protein interaction network analysis

To further investigate the genes associated with glioblastoma functional pathways, we performed PPI analysis on overlapping genes. Overlapping genes were analyzed using STRING (https://cn.string-db.org/) to obtain a protein-protein interaction relationship network. The relationship network was imported into Cytoscape for visualization, and the densely connected network components were identified using the MCODE plug-in to obtain seven gene modules, and the highest-scoring gene module was selected for subsequent analysis. Since the relationship network map drawn by STRING was more complex, the protein-protein interaction network map was redrawn using Metascape (https://metascape.org/).

## Glioblastoma prognostic risk model construction

LASSO (least absolute shrinkage and selection operator) is a risk-scoring model based on prognostic factors (10). The analysis used the R software survival package to conduct a multivariate Cox regression analysis, followed by an iterative analysis using the STEP function, thus selecting the optimal model as the final model. The prognostic risk model was constructed using the highest-scoring gene module obtained by the PPI algorithm to obtain a risk scoring formula. For Kaplan–Meier curves, p-values and hazard ratios (HR) with 95% confidence intervals (CI) were derived by the log-rank test and univariate cox regression. A value of p<0.05 was considered statistically significant.

## Comparison of glioblastoma gene expression

A comparative expression analysis of genes in the LASSO formula was implemented using the R software ggplot2 package to compare the distribution of the same gene in tumor tissue and normal tissue. A value of p<0.05 was considered statistically significant.

## Finding independent prognostic factors for glioblastoma

To further investigate the independent prognostic factors of glioblastoma, univariate Cox and multivariate cox regression analyses of genes were performed using the R software forestplot package; the results are presented using forest plots. A factor was considered to be an independent prognostic factor when the P value for both univariate cox and multivariate cox was less than 0.05.

## Cell culture

Human microglia HMC3 cells (Procell, CL-0620) were used as the experimental group (Experimental), and human U251 glioblastoma cells (Procell, CL-0237) comprised the control group (Control). HMC3 cells were cultured in high-glucose Dulbecco's modified Eagle medium (DMEM; Hyclone, Shanghai, CHN) and supplemented with 1% penicillin–streptomycin (Solarbio Life Sciences, Beijing, CHN) and 15% fetal bovine serum (FBS; GIBCO, Grand Island, USA). U251 cells were cultured in penicillin (100 μg/ml) and streptomycin (100 μg/ml) in high-sugar DMEM and 10% FBS. All cell lines were grown at 37°C in 5% CO2.

## Detection of RPL39L and NUDT5 gene expression by qRT-PCR

Total RNA was extracted from both groups of cells using TRIzol reagent (Invitrogen, USA). The extracted mRNA was reverse transcribed into cDNA using SuperReal PreMix Plus (SYBR Green) (FP205-02, Tiangen, China), and gene expression was detected using qRT-PCR. The relative expression of the genes was calculated using the $2^{-\Delta\Delta CT}$ method. The experiment was repeated three times to determine the average. The expression levels of RPL39L and NUDT5 were detected using GAPDH as an internal reference. The primer sequences used are shown in Table 1.

# Results

## Glioblastoma whole-gene bulk survival analysis and differentially expressed gene analysis

A total of 607 genes significantly associated with prognosis were obtained from a whole-gene bulk survival analysis of 153 glioblastoma tumor samples. To further analyze the distribution

TABLE 1 Primer sequences of genes.

| Gene | Forward | Reverse |
|---|---|---|
| RPL39L | 5'-CAAAATCGTCCCATCCCC-3' | 5'-TTCTTCTCCAATGCCTCCTT-3' |
| NUDT5 | 5'-GTTCTCCAGCGGTCTGTATG-3' | 5'-CTTCGGCCTTGCGTTTTCG-3' |
| GAPDH | 5'-CCTCTCCAGAACATCATCC-3' | 5'-GTGTCGCTGTTGAAGTCAG-3' |

of genes in glioblastoma compared with their distribution in paraneoplastic and normal tissues, a differential analysis was performed. The results showed that there were 7,919 upregulated genes and 882 downregulated genes (Figure 1A). The differential expression heat map demonstrated the expression trends of the largest 50 upregulated and 50 downregulated genes in different tissues (Figure 1B). Three hundred and twenty-three genes that overlapped two groups were identified using a Venn diagram (Figure 1C).

## Results of PPI analysis

Functional enrichment analysis of 323 DEGs using Metascape revealed that the genes were significantly enriched in functions such as the formation of a pool of free 40S subunits and placenta development (Figure 2A). Protein–protein interaction expression was constructed for 323 genes to identify a gene-to-gene interaction linkage, as illustrated in Figure 2B. Nineteen gene modules were obtained *via* the MCODE algorithm, and the highest-scoring gene module was selected, which contained 17 nodes and 124 edges (Figure 2C).

## LASSO analysis to identify genes associated with prognosis

Based on the above study, 17 genes associated with glioblastoma were identified, and a prognostic model was constructed using multivariate cox regression analysis. The

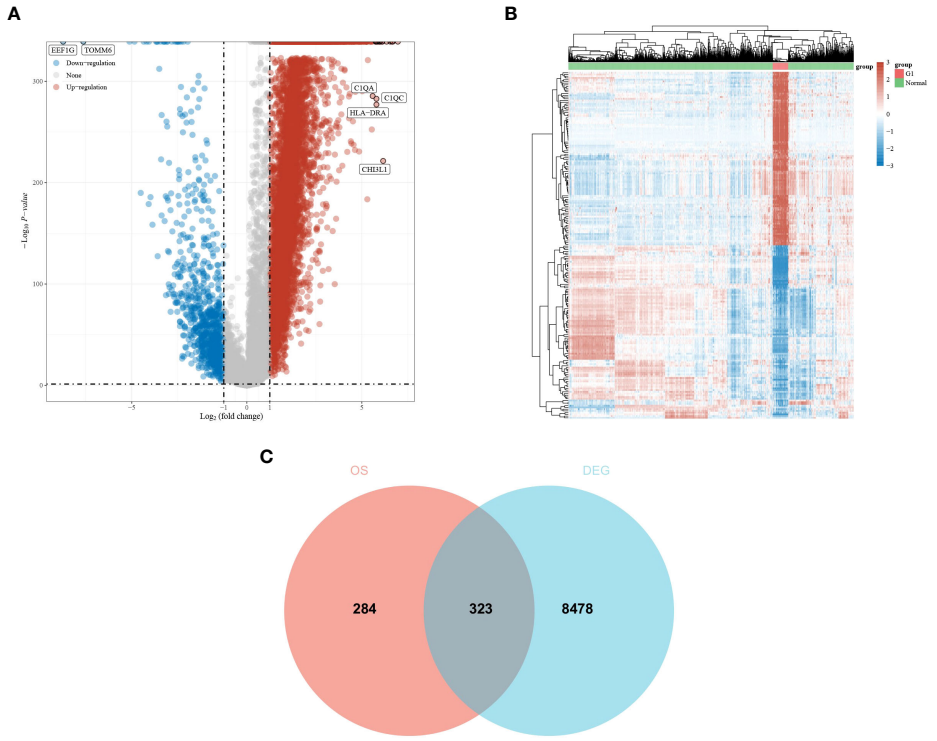



FIGURE 1
Differential expression analysis of glioma dataset to obtain DEGs. **(A, B)** Differential expression analysis volcano and heat map. The volcano plot in red indicates up-regulated genes, and that in blue indicates down-regulated genes. **(C)** An intersection of prognostic genes obtained from whole-gene batch survival with DEGs was taken. Red indicates prognostic genes, and blue indicates DEGs.

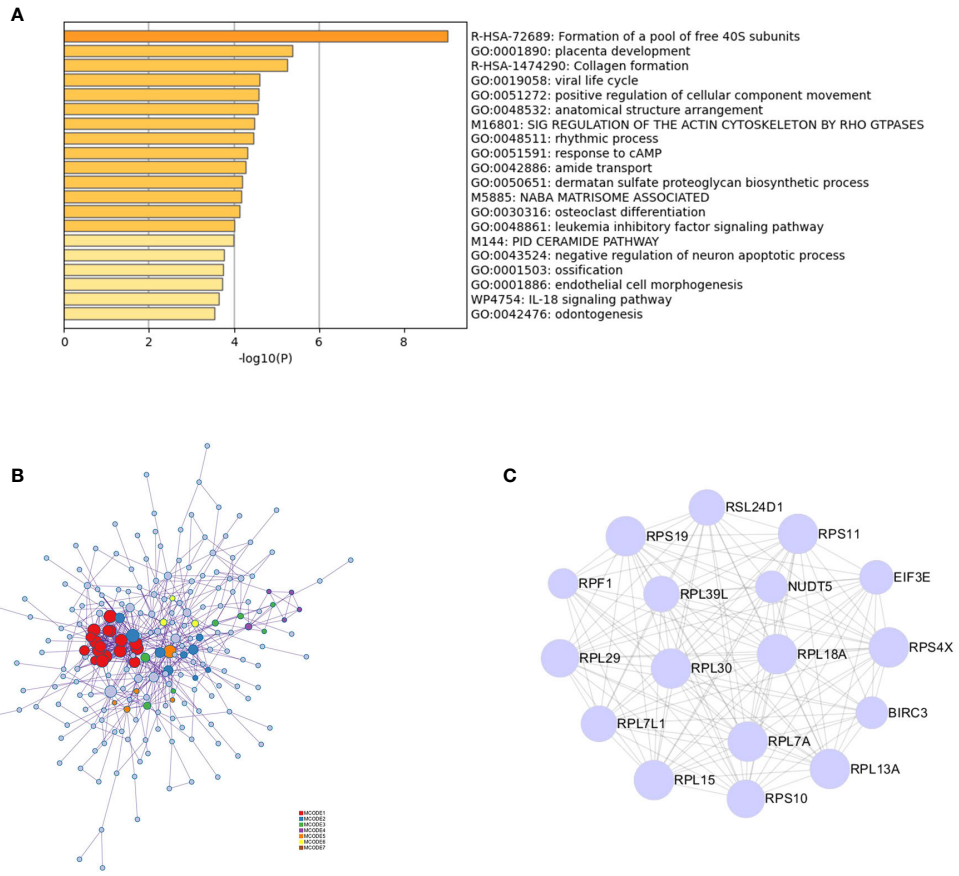



FIGURE 2
PPI analysis of intersecting genes. **(A)** Functional and pathway enrichment analysis of 323 genes using Metascape. **(B)** Analysis of the interactions between 323 genes using Metascape. Modules based on the MCODE algorithm are shown in different colors. **(C)** Analysis based on the MCODE algorithm in Cytoscape to obtain the highest scoring gene modules

coefficients of the selected features were shown by the λ parameter; partial likelihood deviations were plotted against log(λ) using the LASSO Cox regression model (Figures 3A, B). The risk score formula is as follows: Riskscore = (-0.2159) *RPS10 + (-0.0121)*RPS19 + (0.2469)*RPL39L + (-0.0443) *RPS11 + (-0.0357)*RSL24D1 + (-0.3645)*RPF1 + (-0.6918) *NUDT5 + (-0.0858)*EIF3E. Based on the calculation of the risk score formula, the sample was divided into high-risk and low-risk groups; the distribution of the sample is illustrated in Figure 3C. Figure 3D demonstrates that the low-risk group is effective for prognosis. Figure 3E demonstrates that the model is accurate in predicting the three- and five-year survival of glioblastomas.

## Independent prognostic factors for glioblastoma

Gene expression analysis of eight genes obtained from LASSO that were associated with glioblastoma prognosis was performed, and all eight genes were found to be highly expressed in the tumors (Figures 4A–H). These genes are presented as box line plots. An independent prognostic analysis of these eight genes was performed to obtain the results of univariate Figure 5A and multivariate Cox Figure 5B analyses, which are presented as forest plots. As can be seen in the figure, two genes, RPL39L and NUDT5, were significant in both univariate and multivariate Cox regression; therefore, these two genes can be considered independent prognostic factors of glioblastoma.

## Results of qRT-PCR

The expression of RPL39L and NUDT5 in glioblastoma cells and normal cells was detected using qRT-PCR. The results showed that both genes were highly expressed in glioblastoma cells (Figures 6A, B).

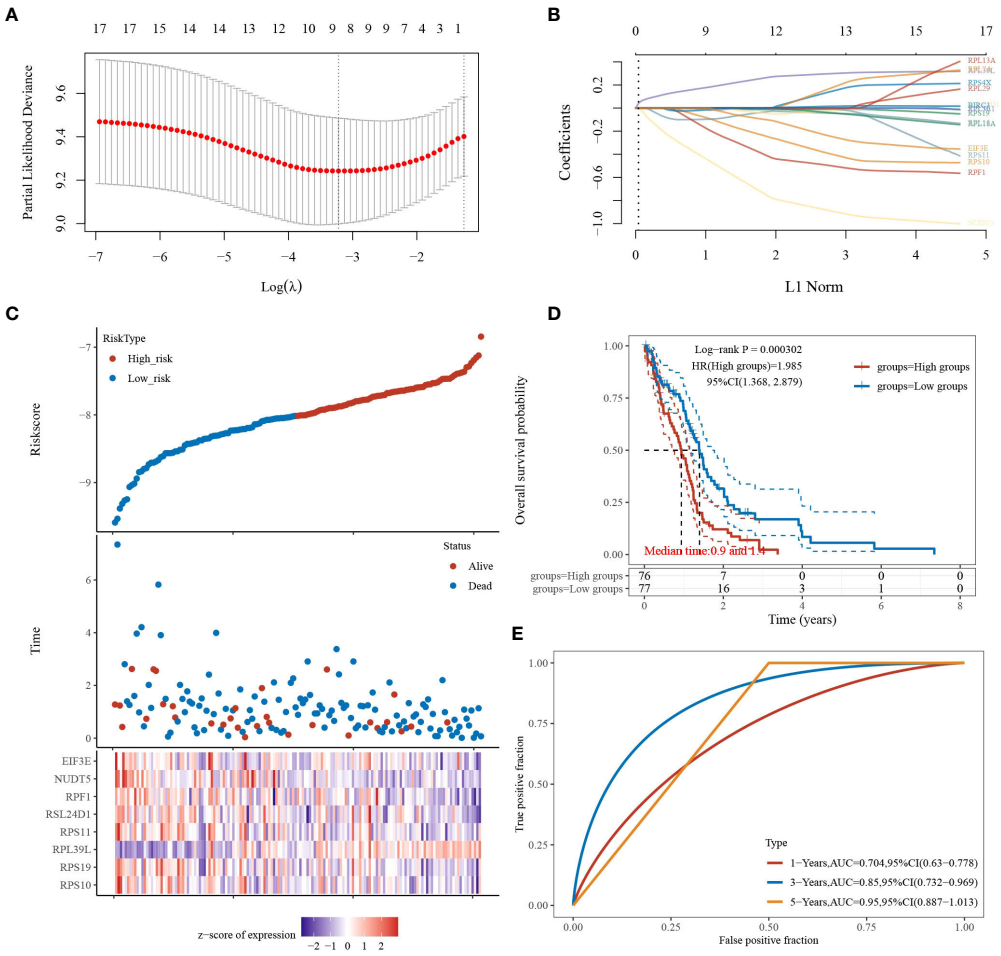



FIGURE 3
LASSO analysis of the 17 genes derived from PPI analysis. **(A)** The coefficients of selected features are shown by the lambda parameter, the horizontal coordinate represents the value of the independent variable lambda, and the vertical coordinate represents the coefficient of the independent variable. **(B)** The partial likelihood deviation was plotted against log(λ) using the LASSO Cox regression model. **(C)** Risk score, survival time, and survival status in the selected dataset, presented using scatter plots and heat maps. **(D)** KM survival curves with median risk values as groupings. **(E)** ROC curves for the risk model.

## Discussion

In recent years, the number of cancer patients has increased dramatically, and finding a breakthrough treatment for cancer has become urgent (11). Malignant glioblastoma is the most common type of primary brain tumor in adults and is associated with a disproportionate amount of cancer-related morbidity and mortality (12), making it particularly important to find ways to treat glioblastoma tumors. The rising trend in the incidence of glioblastoma has been accompanied by an increase in concern (2). At present, the diagnosis of glioblastoma still depends mainly on pathological features and medical imaging, such as CT, MRI, DSA, PET, and SPECT, which need to be verified by a surgeon (6). In addition, numerous studies have been conducted on glioblastoma biomarkers, such as the specificity of circRNAs in glioblastoma, which is expected to be a new biomarker for the development of glioblastoma (13) and to lead to a future cure for glioblastoma tumors. In this study, information from clinical sample data was used to study some of the genes that may be correlated with glioblastoma prognosis. Genes significantly associated with glioblastoma prognosis can be identified through a whole-gene batch survival analysis of clinical samples, and then key genes can be further identified by PPI and LASSO. Then, an expression validation analysis of these genes can be performed to identify potential prognostic biomarkers associated with glioblastoma. Finally, independent prognostic factors for glioblastoma can be obtained via an independent prognostic analysis of these genes.

In this study, we used the clinical data of glioblastoma patients available from the TCGA and identified genes

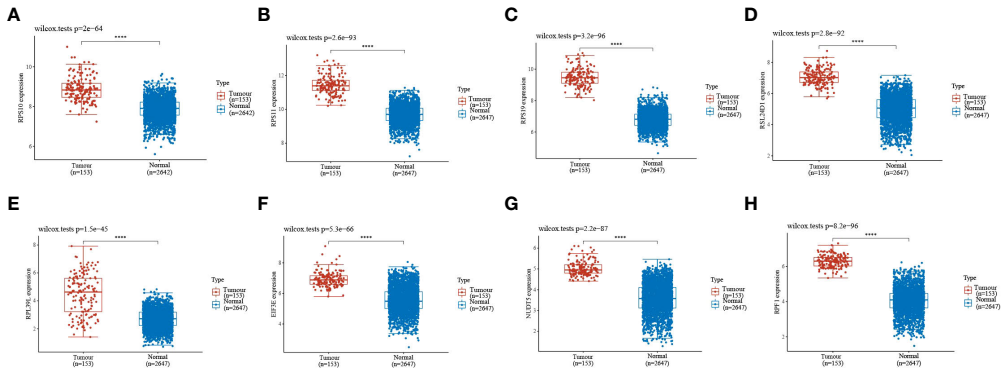



FIGURE 4
Expression validation of the eight genes in the LASSO formula. **(A–H)** The expression validation of the eight genes in the LASSO formula showed that all eight genes were highly expressed in gliomas. Red indicates tumor samples, and blue indicates normal samples. Prognostic gene expression validation. ****indicates P<0.0001.

associated with glioblastoma prognosis by performing a whole-gene bulk survival analysis. We also investigated the functional enrichment pathways of these genes and found that they were significantly enriched in functions such as building free 40S subunits as well as in placental development. A network of gene–gene interactions was also constructed, and the highest scoring motif modules were further analyzed with the help of algorithms. Eight genes—RPS10, RPS11, RPS19, RSL24D1, RPL39L, EIF3E, NUDT5, and RPF—were found to have an expression in the prognosis of glioblastoma. Finally, univariate cox and multivariate cox analyses were performed, which identified RPL39L and NUDT5 as independent prognostic factors for glioblastoma. The results were verified by qRT-PCR experiments.

Ribosomal proteins are synthesized in the cytoplasm by RNA polymerase II and then imported into the nucleus, where they are assembled into small and large ribosomal subunits (14, 15). The small ribosomal subunit contains an 18S rRNA and approximately 32 ribosomal proteins (RPS proteins), and the large ribosomal subunit 60S consists of one of 5S, 5.8S and 28S rRNAs and approximately 47 ribosomal proteins (RPL proteins) (16). Long-term studies have shown that ribosomal proteins not only constitute ribosomes as structural proteins but also play important roles in the cell cycle, proliferation, apoptosis/death, tumorigenesis, DNA repair, and other responses (17).

From the analysis conducted in this study, it is clear that most of the genes associated with glioblastoma prognosis are ribosomal protein genes, represented by RPS10 and RPS11. The RPS10 gene encodes the RPS10 protein, which is part of the small subunit of the mitochondrial ribosome (18) and is involved in ribosome biogenesis, as well as participating in cellular transformation mechanisms (16). RPS10 encodes the 165 amino-acid-long RPS10 protein, which is a component of the 40S ribosomal subunit (19) and can cross-link to the eukaryotic initiation factor 3 (eIF3) of translation. It has been shown that the RPS10 protein is part of the structural domain

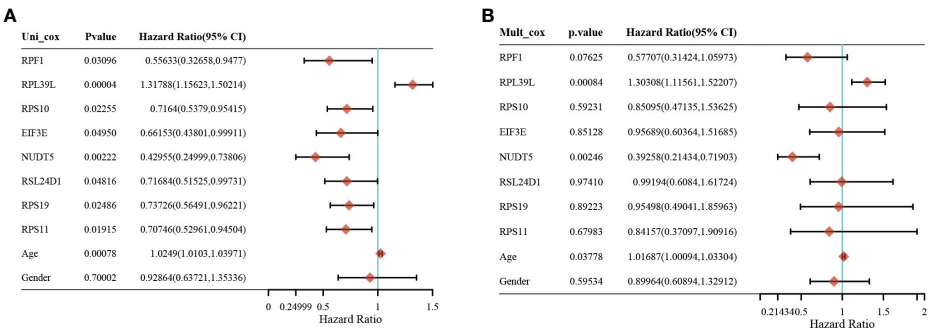



FIGURE 5
Screening of independent prognostic factors for glioma based on univariate and multivariate cox analyses. **(A)** Univariate cox analysis of prognostic genes. **(B)** Multivariate cox analysis of prognostic genes.

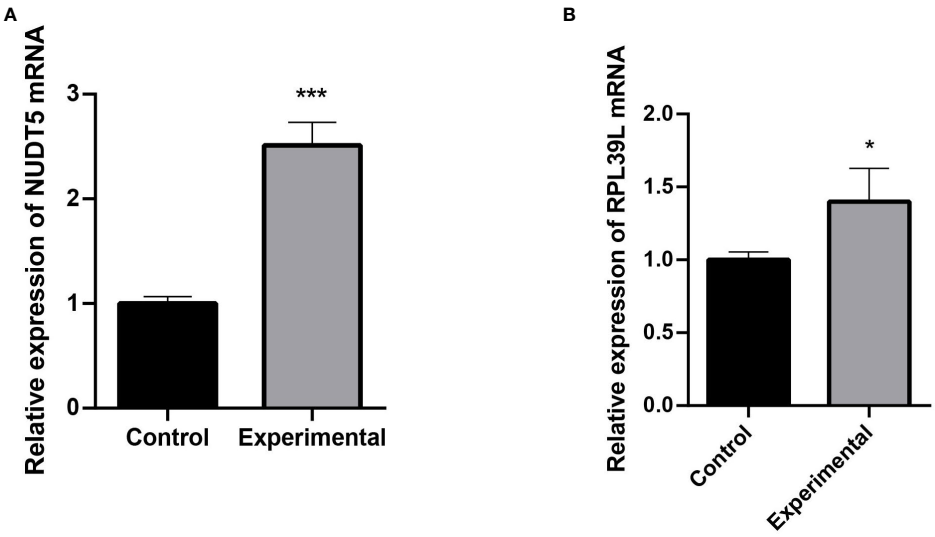



FIGURE 6
Detection of RPL39L and NUDT5 expression using qRT-PCR. **(A)** Relative expression of NUDT5 mRNA. **(B)** Relative expression of RPL39L mRNA. *** indicates P<0.001 compared with the control group; * indicates P<0.05 compared with the control group.

involved in the binding of the initiation factor to the 40S subunit at the onset of translation (20). RPS11 encodes the RPS11 protein, which is overexpressed in various malignancies and is associated with tumor recurrence (17). Elevated levels of RPS11 have been found to be associated with a poor prognosis in patients with glioblastoma (21). RPL19 is upregulated in many prostate cancers, and its downregulation leads to a milder malignant phenotype *in vivo*, suggesting a functional role in promoting tumorigenesis (22). It has been shown that RPL19 is associated with glioblastoma prognosis (23). nudt5 encodes NUDT5 hydrolase, which is associated with breast cancer prognosis (24). It can inhibit the propagation of HeLa cells and T47D cells (25). No studies have been conducted on NUDT5 related to glioblastoma.

GO analysis of important pathways has identified the composition of the free 40s subunit pool and placental development, among others, and studies have shown that 40S subunits can enter the free and membrane-bound polyribosomes from the cytoplasmic pool of newly made free natural subunits (26), thus affecting protein synthesis and cell development. There is a link between the placenta and cancer cells at the molecular level (27), and the placenta acts as a transport function between the mother and the fetus, during development. Genes influence tumorigenesis by affecting the 40s subunit and placental development, among others.

In conclusion, this study obtained two independent prognostic markers associated with glioblastoma prognosis, which could play a central role in the prognosis of glioblastoma, were obtained by bioinformatics using glioblastoma tumor samples and normal samples.

## 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

YT: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing. LC: Conceptualization, Formal analysis, Writing - Original Draft, Writing - Review & Editing. All authors contributed to the article and approved the submitted version.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

# References


1. Liu M, Xu Z, Du Z, Wu B, Jin T, Xu K, et al. The identification of key genes and pathways in glioma by bioinformatics analysis. *J Immunol Res* (2017) 2017:1278081. doi: 10.1155/2017/1278081

2. Geng RX, Li N, Xu Y, Liu JH, Yuan FE, Sun Q, et al. Identification of core biomarkers associated with outcome in glioma: Evidence from bioinformatics analysis. *Dis Mark* (2018) 2018:3215958. doi: 10.1155/2018/3215958

3. Li X, Kong S, Cao Y. miR-1254 inhibits progression of glioma *in vivo* and *in vitro* by targeting CSF-1. *J Cell Mol Med* (2020) 24(5):3128–38. doi: 10.1111/jcmm.14981

4. Thakkar JP, Dolecek TA, Horbinski C, Ostrom QT, Lightner DD, Barnholtz-Sloan JS, et al. Epidemiologic and molecular prognostic review of glioblastoma. *Cancer Epidemiol Biomarkers Prev* (2014) 23(10):1985–96. doi: 10.1158/1055-9965.EPI-14-0275

5. Carlsson SK, Brothers SP, Wahlestedt C. Emerging treatment strategies for glioblastoma multiforme. *EMBO Mol Med* (2014) 6(11):1359–70. doi: 10.15252/emmm.201302627

6. Ohgaki H, Kleihues P. Epidemiology and etiology of gliomas. *Acta Neuropathol.* (2005) 109(1):93–108. doi: 10.1007/s00401-005-0991-y

7. Iacob G, Dinca EB. Current data and strategy in glioblastoma multiforme. *J Med Life* (2009) 2(4):386–93.

8. Malmström A, Grønberg BH, Marosi C, Stupp R, Frappaz D, Schultz H, et al. Nordic Clinical brain tumour study group (NCBTSG). temozolomide versus standard 6-week radiotherapy versus hypofractionated radiotherapy in patients older than 60 years with glioblastoma: the Nordic randomised, phase 3 trial. *Lancet Oncol* (2012) 13(9):916–26. doi: 10.1016/S1470-2045(12)70265-6

9. Hanif F, Muzaffar K, Perveen K, Malhi SM, Simjee S. Glioblastoma multiforme: A review of its epidemiology and pathogenesis through clinical presentation and treatment. *Asian Pac J Cancer Prev* (2017) 18(1):3–9. doi: 10.22034/APJCP.2017.18.1.3

10. Zhang C, Gou X, He W, Yang H, Yin H. A glycolysis-based 4-mRNA signature correlates with the prognosis and cell cycle process in patients with bladder cancer. *Cancer Cell Int* (2020) 20:177. doi: 10.1186/s12935-020-01255-2

11. Wen PY, Kesari S. Malignant gliomas in adults. *N Engl J Med* (2008) 359(5):492–507. doi: 10.1056/NEJMra0708126

12. Guo X, Piao H. Research progress of circRNAs in glioblastoma. *Front Cell Dev Biol* (2021) 9:791892. doi: 10.3389/fcell.2021.791892

13. Mischel PS, Cloughesy TF. Targeted molecular therapy of GBM. *Brain Pathol* (2003) 13(1):52–61. doi: 10.1111/j.1750-3639.2003.tb00006.x

14. Rodnina MV, Wintermeyer W. Recent mechanistic insights into eukaryotic ribosomes. *Curr Opin Cell Biol* (2009) 21(3):435–43. doi: 10.1016/j.ceb.2009.01.023

15. Tschochner H, Hurt E. Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends Cell Biol* (2003) 13(5):255–63. doi: 10.1016/s0962-8924(03)00054-0

16. Abbas W, Dichamp I, Herbein G. The HIV-1 nef protein interacts with two components of the 40S small ribosomal subunit, the RPS10 protein and the 18S rRNA. *Virol J* (2012) 9:103. doi: 10.1186/1743-422X-9-103

17. Zhou C, Sun J, Zheng Z, Weng J, Atyah M, Zhou Q, et al. High RPS11 level in hepatocellular carcinoma associates with poor prognosis after curative resection. *Ann Transl Med* (2020) 8(7):466. doi: 10.21037/atm.2020.03.92

18. Majewski P, Wołoszyńska M, Jańska H. Developmentally early and late onset of Rps10 silencing in arabidopsis thaliana: genetic and environmental regulation. *J Exp Bot* (2009) 60(4):1163–78. doi: 10.1093/jxb/ern362

19. Doherty L, Sheen MR, Vlachos A, Choesmel V, O'Donohue MF, Clinton C, et al. Ribosomal protein genes RPS10 and RPS26 are commonly mutated in diamond-blackfan anemia. *Am J Hum Genet* (2010) 86(2):222–8. doi: 10.1016/j.ajhg.2009.12.015

20. Westermann P, Nygård O. The spatial arrangement of the complex between eukaryotic initiation factor eIF-3 and 40 s ribosomal subunit. cross-linking between factor and ribosomal proteins. *Biochim Biophys Acta* (1983) 741(1):103–8. doi: 10.1016/0167-4781(83)90015-5

21. Yong WH, Shabihkhani M, Telesca D, Yang S, Tso JL, Menjivar JC, et al. Ribosomal proteins RPS11 and RPS20, two stress-response markers of glioblastoma stem cells, are novel predictors of poor prognosis in glioblastoma patients. *PloS One* (2015) 10(10):e0141334. doi: 10.1371/journal.pone.0141334

22. Bee A, Brewer D, Beesley C, Dodson A, Forootan S, Dickinson T, et al. siRNA knockdown of ribosomal protein gene RPL19 abrogates the aggressive phenotype of human prostate cancer. *PloS One* (2011) 6(7):e22672. doi: 10.1371/journal.pone.0022672

23. Wang Z, Gao L, Guo X, Lian W, Deng K, Xing B. Development and validation of a novel DNA methylation-driven gene based molecular classification and predictive model for overall survival and immunotherapy response in patients with glioblastoma: A multiomic analysis. *Front Cell Dev Biol* (2020) 8:576996. doi: 10.3389/fcell.2020.576996

24. Tong XY, Quan Y, Zhang HY. NUDT5 as a novel drug target and prognostic biomarker for ER-positive breast cancer. *Drug Discovery Today* (2021) 26(3):620–5. doi: 10.1016/j.drudis.2020.11.031

25. Li J, Yang CC, Tian XY, Li YX, Cui J, Chen Z, et al. MutT-related proteins are novel progression and prognostic markers for colorectal cancer. *Oncotarget* (2017) 8(62):105714–26. doi: 10.18632/oncotarget.22393

26. Mechler B, Vassalli P. Membrane-bound ribosomes of myeloma cells. II. kinetic studies on the entry of newly made ribosomal subunits into the free and the membrane-bound ribosomal particles. *J Cell Biol* (1975) 67(1):16–24. doi: 10.1083/jcb.67.1.16

27. Costanzo V, Bardelli A, Siena S, Abrignani S. Exploring the links between cancer and placenta development. *Open Biol* (2018) 8(6):180081. doi: 10.1098/rsob.180081

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 13 February 2023
DOI 10.3389/fonc.2022.1014902

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Xianfeng Jiang,
Sir Run Run Shaw Hospital, China
Jiawei Xu,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Jian Hu
dr_hujian@zju.edu.cn
Weiyu Shen
shenweiyu@nbu.edu.cn

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

RECEIVED 09 August 2022
ACCEPTED 28 December 2022
PUBLISHED 13 February 2023

CITATION
He J, Liang G, Tian H, Wang Y, Yu L, Lv W, Hu J and Shen W (2023) Wnt signaling pathway-related gene PRICKLE1 is a prognostic biomarker for esophageal squamous cell carcinoma.
*Front. Oncol.* 12:1014902.
doi: 10.3389/fonc.2022.1014902

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

# Wnt signaling pathway-related gene PRICKLE1 is a prognostic biomarker for esophageal squamous cell carcinoma

Jinxian He[1,2], Gaofeng Liang[2], Hui Tian[2], Yiqing Wang[1], Li Yu[1], Wang Lv[1], Jian Hu[1]* and Weiyu Shen[2]*

[1]Department of Thoracic Surgery, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China, [2]Department of Thoracic Surgery, Ningbo Medical Center Lihuili Hospital, Ningbo University, Ningbo, Zhejiang, China

Esophageal squamous cell carcinoma (ESCC) has become a major health risk to human health. Although significant clinical progress has been made in the treatment of ESCC, the prognosis of patients still needs to be improved. Therefore, it is important to screen effective molecular indicators for the prognosis of ESCC. In this study, the intersection of up-regulated genes, down-regulated genes, and Wnt signaling pathway-related genes in ESCC was taken, and 47 overlapping genes were found. PRICKLE1 was determined to be an independent prognostic factor in ESCC based on univariate and multifactorial COX risk regression models. Kaplan-Meier survival curves showed that patients in the PRICKLE1 high expression group had significantly better overall survival. In addition, we performed various experiments to examine the effects of PRICKLE1 overexpression on proliferation, migration, and apoptosis of ESCC cells. The experimental results showed that the PRICKLE1-OE group had reduced cell viability, significantly lower migration ability and significantly higher apoptosis rate compared to the NC group.Therefore, we hypothesized that high PRICKLE1 expression could be used to predict the survival rate of ESCC patients, which could be used as an independent prognostic indicator for ESCC patients and provide potential applications for ESCC clinical treatment.

KEYWORDS

**esophageal squamous cell carcinoma, PRICKLE1, Wnt signaling pathway, overall survival, proliferation, cell migration, apoptosis**

## 1 Introduction

The incidence of esophageal cancer has shown a significantly increasing trend. The main manifestations of esophageal cancer are damage to the esophageal wall and narrowing of the esophageal lumen, and patients often have symptoms such as difficulty in swallowing, dryness, and tightness in the throat (1). Esophageal cancer is very easy to invade and infiltrate and grow, and the symptoms are insidious, so the first diagnosis is often a patient in the

middle or late stage, so the annual survival rate is not high (2). In 2020, the global cancer incidence of esophageal cancer is reported to be 600,000, ranking eighth, and the number of global deaths from esophageal cancer is reported to be 540,000, ranking sixth, making esophageal cancer a major health hazard to human beings (3). The causes of esophageal cancer are still unclear, and it is believed that the occurrence of the disease is influenced by a combination of factors, such as genetics, lifestyle habits, dietary structure, alcoholism and smoking, and environmental characteristics, which are all related to the occurrence of esophageal cancer (4, 5). The histological typing of esophageal cancer is mainly ESCC andesophageal adenocarcinoma, among which, the pathological type of esophageal cancer in China is mainly ESCC, which is more common than esophageal adenocarcinoma (6). Although significant clinical progress has been made in the treatment of ESCC in recent years, the prognosis of patients still needs to be improved. Studies have shown that the occurrence and progression of ESCC involve multi-gene and multi-pathway alterations (7). Therefore, it is important to screen effective molecular indicators for population screening and prognosis determination of ESCC, which can help prolong the survival time of patients.

The Wnt signaling pathway is one of the five major cell signaling pathways and consists of a series of proteins encoded by oncogenes and oncogenes, which are regulated by several key proteins that are interconnected and mutually regulated (8). The Wnt signaling pathway is evolutionarily conserved and is involved in regulating cell growth and development, differentiation, proliferation, migration, and other growth processes (9). Abnormalities in the Wnt signaling pathway activation may lead to malignant transformation of cells and tumor development, invasion, and metastasis, and are closely related to poor patient prognosis (10). In recent years, this pathway has been a hot topic of research in cancers such as gastric, hepatocellular, colorectal, breast, and pancreatic cancers, and certain results have been achieved (11, 12). This study aimed to identify Wnt signaling pathway genes associated with ESCC prognosis, construct a Nomogram model, analyze single-gene prognosis, and experimentally validate them to improve patients' quality of life.

# 2 Methodology

## 2.1 Data sources

RNA-seq data of ESCC and its clinical information were downloaded from The Cancer Genome Atlas database, containing ESCC samples (N=82) and paracancerous tissue samples (N=11), clinical information included data on the age, race, clinical stage, and overall survival (OS) of the sample, with OS defined as the time from any cause or last follow-up diagnosis to death.Wnt signaling pathway-related genes (N=167) were downloaded from the MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp).

## 2.2 Differential expression analysis

Using P<0.05 and |Log2FC|>1 as the screening threshold for differentially expressed genes (DEGs), differential analysis of ESCC cancer and paraneoplastic tissues was performed using the R software "limma" package. Venn diagrams were used to intersect the DEGs and Wnt signaling pathway-related genes to obtain the differentially expressed Wnt signaling pathway-related genes in ESCC.

## 2.3 Identification of prognosis-related genes

One-way COX regression analysis was performed on DEGs related to the Wnt signaling pathway, genes related to prognosis were screened using the Log-rank test, and forestplot was plotted using the R package "forestplot".

## 2.4 Nomogram construction

Independent influences on patient prognosis were determined based on univariate and multifactorial COX risk regression models. Based on the independent influences on prognosis, a Nomogram was constructed using the R package "rms" and a line graph was drawn.

## 2.5 Single-gene prognostic analysis

The overall survival of patients in the single gene high and low expression groups was analyzed by the R package Kaplan-Meier and Log-rank test. ROC curves were used to assess the predictive performance of prognostic genes for 1, 2, and 3-year survival of patients.

## 2.6 Cell culture and processing

The human esophageal cancer KYSE-450 cell line and KYSE-410 cell line and human normal esophageal squamous epithelial cell Het-1A was inoculated in RPMI 1640 culture medium (containing 10% fetal bovine serum, penicillin 100 U/mL, and streptomycin 100 μg/mL) and cultured at 37°C, 5% $CO_2$, and saturated humidity, and the culture medium was changed once in 2 d. When the cell growth density reached about 90%, the cells were passaged at a ratio of 1:2. Cells in the logarithmic growth phase were taken for experiments.

Cells were transfected with lentiviral overexpression vector or empty vector using Lipofectamine 2000 reagent. The cells were transfected with the PRICKLE1-OE group and control group respectively, and the transfected cells were incubated at 37°C for 24 h. Refer to the kit instructions for the specific transfection operation.

## 2.7 Real-time quantitative polymerase chain reaction

Total RNA was extracted using the TRIzol kit (Invitrogen, USA). cDNA was synthesized using the reverse transcription kit. qPCR was performed using the SYBR Premix Ex Taq kit (TaKaRa, Japan). GAPDH was used as an internal reference, and the primer sequences were as in Table 1. follow the kit instructions. The experiment was repeated 3 times.

TABLE 1 qPCR primer sequence.

| Gene | Forward 5′-3′ | Reverse 5′-3′ |
|---|---|---|
| PRICKLE1 | TTTGCTTGCTTACCAGAGGAAA | ACTGGCAATACCGTACCTCAT |
| GAPDH | CTGGGCTACACTGAGCACC | AAGTGGTCGTTGAGGGCAATG |

## 2.8 Immunoblotting assay

Cells were lysed using cell lysis buffer. total protein concentration was determined by BCA method. Proteins were separated using SDS-PAGE electrophoresis and transferred to PVDF membranes. After sealing with 5% skim milk, the membranes were incubated with primary antibodies overnight. The primary antibody consisted of Anti-β Catenin (ab32572, abcam). Membranes were washed 3 times with TBST and incubated with horseradish peroxidase-coupled secondary antibodies at room temperature. Target bands were detected using ECL. GAPDH (ab8245, abcam) was used as an internal control. Repeat the experiment 3 times.

## 2.9 CCK-8

Esophageal cancer cells overexpressing PRICKLE1 were digested and inoculated into 96-well plates, and the cell density was adjusted to $1*10^4$ cells per well using culture medium, with 3 replicate wells per group. Add 10 μl CCK-8 solution to each well, incubate in the incubator for 2 h, and detect the absorbance at 450 nm using an enzyme marker. The percentage of cell viability was obtained.

## 2.10 Cell migration

A marker pen was used to draw horizontal lines evenly across the back of the 6-well plate, approximately every 0.5~1 cm, across the wells, with at least 5 lines across each well. About $5\times10^5$ cells were added to the wells, and the next day the marker was used with the tip of the gun compared to a straightedge and scratched as far down as possible to the horizontal line behind. Wash the cells with PBS 3 times, remove the scratched down cells, add serum-free medium; put into 37°C, 5% $CO_2$ incubator and incubate; take samples according to 0h and 24h, observe the migration of cells at specific locations with inverted microscope and take pictures.

## 2.11 Apoptosis

Collect the cells, resuspend $1\times10^5$ cells in 200μL Binding Buffer, and add 4μL 0.5 mg/mL PI and 2μL Annexin V-FITC solution. Incubate at room temperature for 15 min at avoidance of light, and perform fluorescence detection by flow cytometry.

## 2.12 Statistical processing

GraphPad Prism 6 was used for data processing. The measurement data were expressed as mean ± standard deviation, and the data were compared between groups using a t-tset. all experiments were repeated three times. $p<0.05$ was considered to be a significant difference.

# 3 Results

## 3.1 Differential expression analysis

Differential analysis of ESCC samples and paracancer tissue samples yielded 3451 up-regulated genes and 761 down-regulated genes (Figures 1A, B). The intersection of up-regulated genes, down-regulated genes, and Wnt signaling pathway-related genes in ESCC was taken, and 47 overlapping genes were found (Figure 1C).

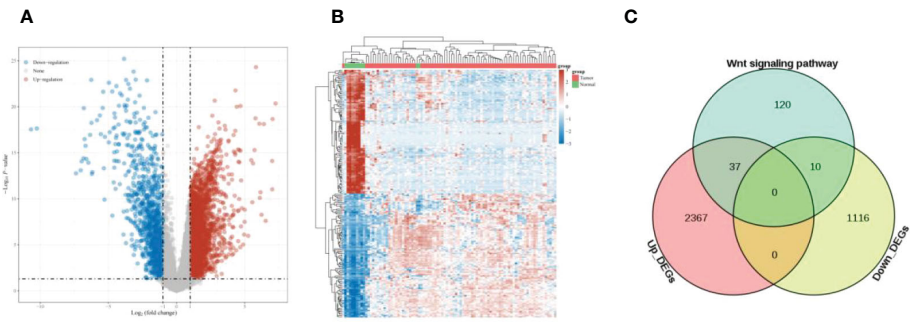



FIGURE 1
Differential expression analysis **(A)** Volcano plot showing DEGs of ESCC; **(B)** Heat map showing the distribution of DEGs of ESCC; **(C)** Venn diagram, overlapping genes were obtained.

## 3.2 Identification of prognosis-related genes

The results of one-way COX regression analysis showed that prickle sense protein 1 (PRICKLE1) was a prognosis-related Wnt signaling pathway gene (Figure 2, P<0.05).

## 3.3 Nomogram construction

One-way COX risk regression analysis showed that PRICKLE1 was a prognosis-associated Wnt signaling pathway gene (Figure 3A, P<0.05). Multi-factor COX risk regression analysis showed that PRICKLE1 may be an independent prognostic factor (Figure 3B, P<0.05). Nomogram prediction model showed that PRICKLE1 was an independent prognostic factor in ESCC, and PRICKLE1 had the best predictive ability for a 1-year prognosis in ESCC patients (Figures 3C, D).

## 3.4 Single-gene prognostic analysis

The samples were classified into the PRICKLE1 high expression group and PRICKLE1 low expression group according to the amount of individual gene expression (Figure 4A). Kaplan-Meier survival curves showed that the overall survival of patients in the PRICKLE1 high expression group was significantly better (Figure 4B, P<0.05). ROC showed that PRICKLE1 predicted 1-year prognosis of ESCC patients with AUC value was the largest (Figure 4C).

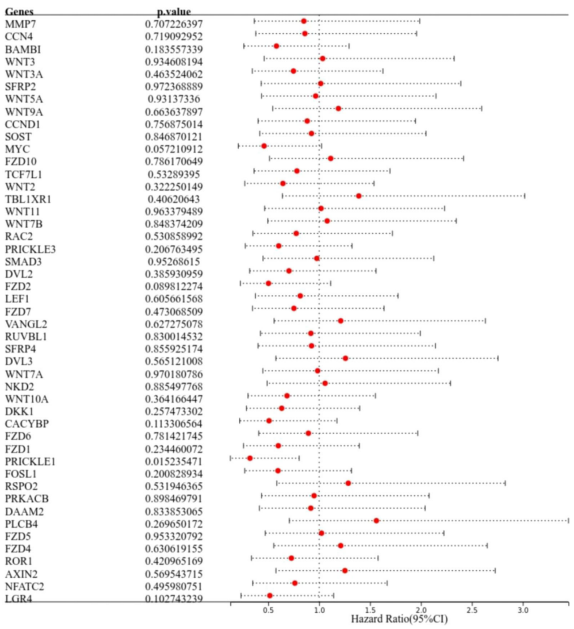



FIGURE 2
Identification of prognosis-related genes. The forest plot shows PRICKLE1 as a prognosis-related Wnt signaling pathway gene.

## 3.5 PRICKLE1 overexpression inhibits cell migration and promotes apoptosis

Through experimental validation, we found that the expression of PRICKLE1 was reduced in esophageal cancer; PRICKLE1 was able to attenuate Wnt/β-catenin signaling; the cell viability was reduced, the migration ability was significantly lower and the apoptosis rate was significantly higher in the PRICKLE1-OE group compared with the NC group (Figures 5A–F, P<0.05).

# 4 Discussion

The early symptoms of ESCC patients are not obvious and easily ignored by patients and physicians, and patients are often in the middle to late stages when they are clinically diagnosed with ESCC, which is characterized by progressive dysphagia as the main clinical manifestation and a low survival rate (13). Although the techniques for treating ESCC have improved significantly in recent years, the five-year survival rate of the disease is still less than 20% (14). Therefore, there is a clinical need to actively investigate in depth the prognosis-related biomarkers of ESCC to provide important clues and directions for targeted therapy. In recent years, molecular biology has become a hot spot for clinical research, especially the Wnt signaling pathway in tumors (15).

The Wnt signaling pathway is a key pathway in cell development and regulation of growth, which is not only important in early embryonic development, organ formation, and tissue regeneration but also closely related to the cellular carcinogenesis process (16). In a normally growing mature organism, the Wnt signaling pathway is turned off. A mutation in one of the signaling members of the Wnt signaling pathway leads to overactivation of the pathway, resulting in abnormal proliferation and differentiation of many cells, which in turn leads to tumor formation and promotes tumor development (17). It has been shown that abnormalities in the Wnt signaling pathway exist in a variety of human tumors, such as breast, gastric, hepatocellular, colorectal, and prostate cancers (18). The Wnt signaling pathway is a major mechanism in tumor biology, and the activation of this pathway involves various processes such as Wnt2, β-catenin, GSK3β, and Axin. Nakajima M et al. suggested that Axin-2 acts as a negative regulator in the Wnt signaling pathway and is involved in physiological processes such as embryonic development, cell differentiation, and organ formation. Among 81 ESCC patients, five ESCC patients had Axin gene polymorphism as well as reduced Axin protein expression, and low expression of Axin was associated with the clinicopathological characteristics of ESCC patients and was beneficial in predicting the prognosis of patients (19). chu CY's team concluded that miRNAs related to the Wnt signaling pathway can regulate the development of ESCC cells and affect the prognosis of patients, and Wnt signaling pathway-related genes could be used as potential therapeutic targets for ESCC (20). Deng F's team found that Wnt2 and β-catenin were highly expressed and GSK3β was lowly expressed in ESCC samples; and the positive rate of Wnt2 was positively correlated with the malignancy of ESCC, the expression of GSK3β was significantly correlated with the

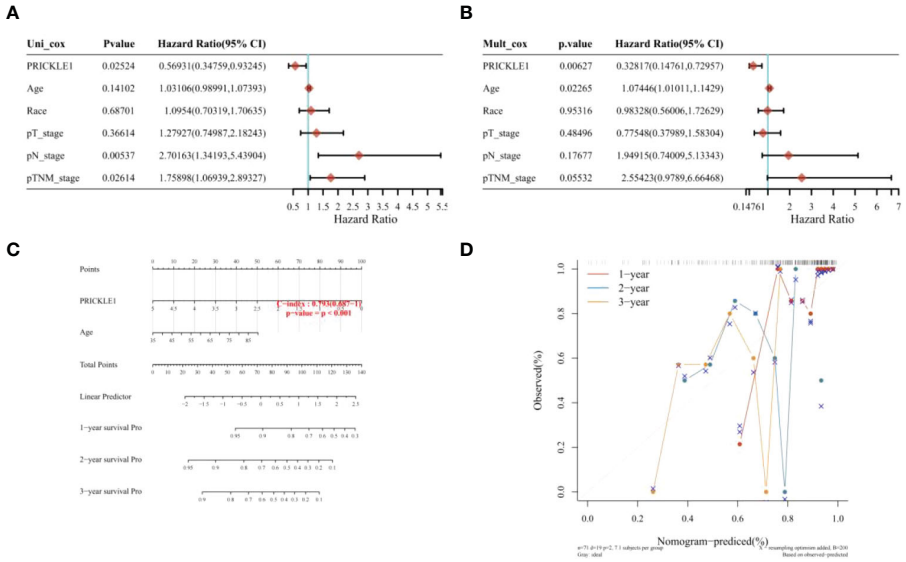



FIGURE 3
Nomogram construction **(A)** single-factor COX risk regression model; **(B)** multi-factor COX risk regression model; **(C)** column line graph predicting patient survival; **(D)** calibration curve showing the predictive performance of PRICKLE1 on patient survival at 1, 2 and 3 years.

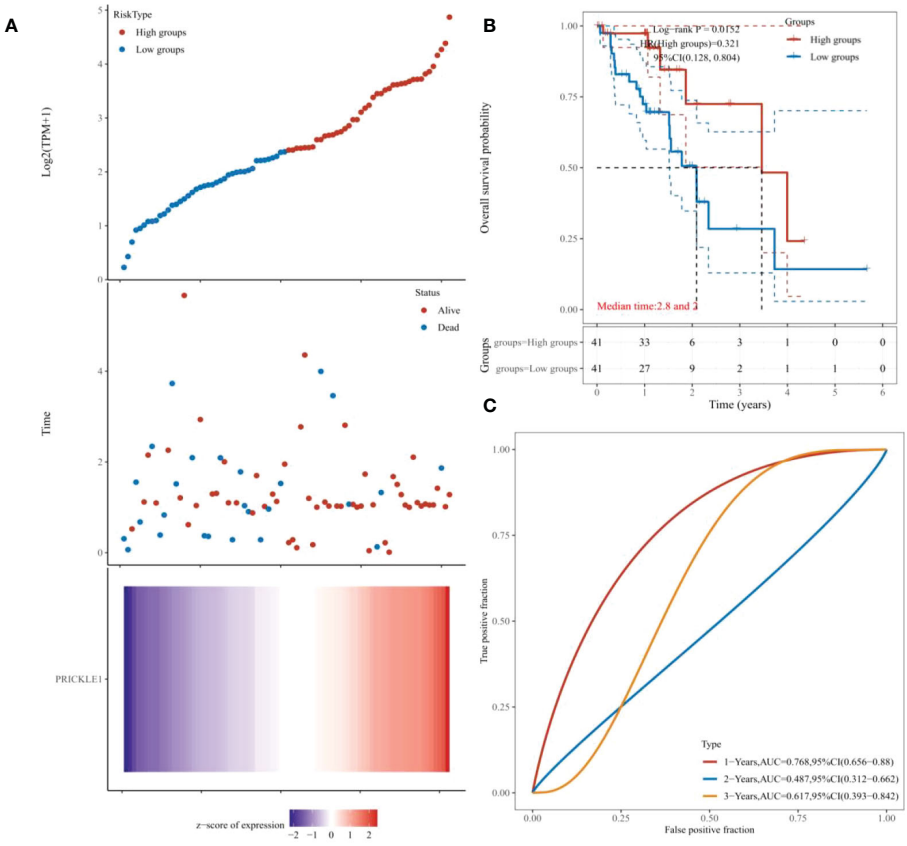



FIGURE 4
Single-gene prognostic analysis **(A)** Grouping according to individual gene expression; **(B)** Kaplan-Meier survival curves showing patient survival; **(C)** ROC showing the predictive performance of PRICKLE1 on patient survival at 1, 2, and 3 years.

10.3389/fonc.2022.1014902

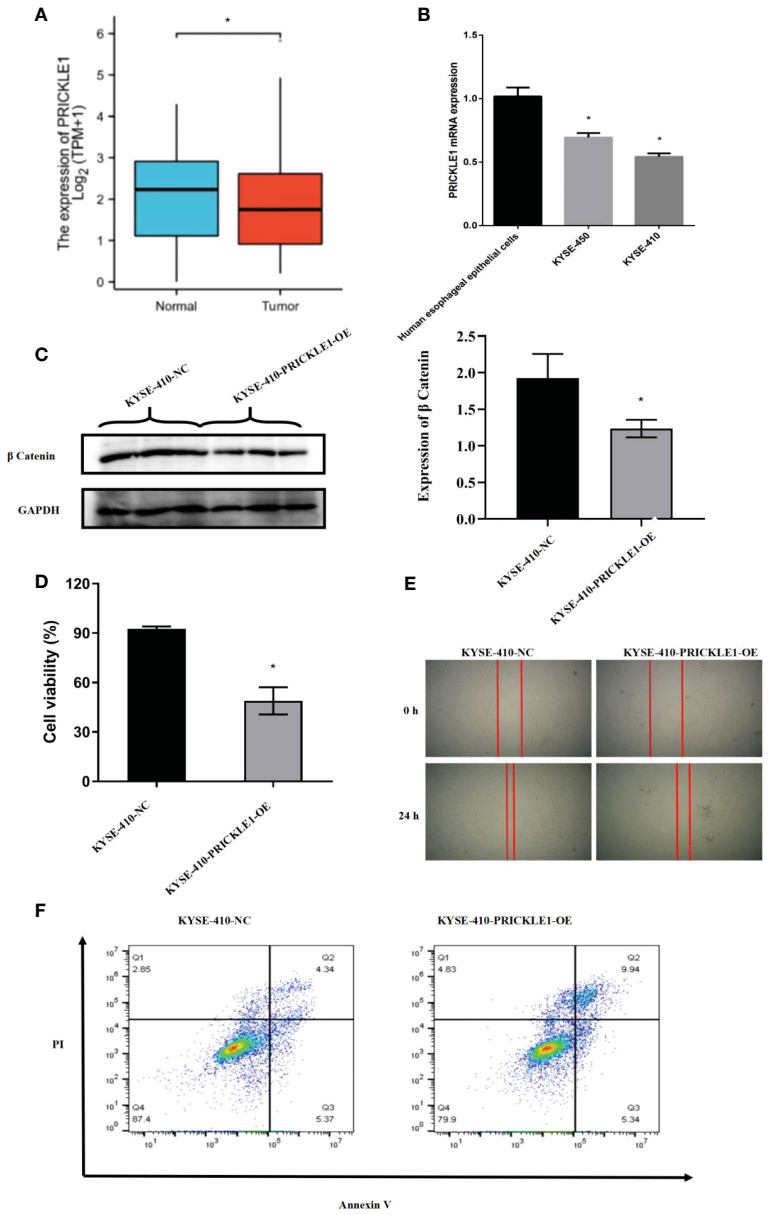



FIGURE 5
Experimental validation **(A)** Expression of PRICKLE1 in normal human esophageal epithelial cells and esophageal cancer cells; **(B)** qPCR showing mRNA expression in normal human esophageal epithelial cells and esophageal cancer cells; **(C)** WB showing the activation level of Wnt/b-catenin signaling pathway in NC group and PRICKLE1-OE group; **(D)** CCK8 showing cell viability in NC group and PRICKLE1-OE group; **(E)** Cell scratch assay showing cell migration in the NC group and PRICKLE1-OE groups; **(F)** Flow cytometry showing apoptosis in the NC group and PRICKLE1-OE groups, with the sum of cells in Q2 and Q3 in the NC group being 9.71% and the sum of cells in Q2 and Q3 in the PRICKLE1-OE group being 15.28%. Compared with the Human esophageal epithelial cells group, *P<0.05; Compared with the NC group, *P<0.05.

tumor location of ESCC, and β-catenin was closely associated with the clinicopathological process in GSK3β-negative ESCC samples (21).

PRICKLE1, a nuclear receptor, is a member of the non-classical Wnt/planar cell polarity (PCP) pathway, where Wnt/PCP signaling controls tissue polarity and cell motility and mediates collective migration events.PRICKLE1 is also a negative regulator in the Wnt signaling pathway, which regulates cells mainly through ubiquitination/deubiquitination and serine/threonine kinase phosphorylation (22).In various cell lines, PRICKLE1 regulates directed cell migration, while at the molecular level, PRICKLE1 regulates the subcellular localization of related proteins as a way to coordinate directed cell migration (23). In different tumor cells, PRICKLE1 has different regulatory roles. On the one hand, it has been found that PRICKLE1 is highly expressed in breast cancer and its upregulation correlates with increased phosphorylation of AKT and its downstream components. PRICKLE1 enhances the migration ability and proliferation of cancer cells by acting on cytoskeleton-related interacting proteins and can be used as a marker of poor

prognosis in breast cancer patients (24). Zhou R's group reported that gastric cancer patients with PRICKLE1 expression were elevated, which may be because PRICKLE1 affects cytoskeletal reorganization by activating mTOR signaling, thus enhancing migration and invasion of gastric cancer cells (25). On the other hand, PRICKLE1 can attenuate Wnt/β-catenin signaling through the degradation inactivation of β-catenin and may act as a tumor suppressor in some cancers (26). It has been suggested that PRICKLE1 is a disheveled (DVL)-associated protein and exerts tumor suppressive effects by antagonizing DVL recruitment; PRICKLE1 expression is reduced in colorectal cancer tissues compared with normal tissues, and PRICKLE1 deficiency may allow further tumor progression (27). In addition, in colorectal cancer patients, PRICKLE1 can directly interact with DVL to mediate ubiquitin-proteasome pathway degradation. Also, PRICKLE1 can negatively regulate Wnt/β-catenin activity, which in turn affects the proliferative activity of colorectal cancer cells (28).

In this study, based on univariate and multifactorial COX risk regression analysis, we observed that PRICKLE1 was an independent prognostic factor in ESCC and that PRICKLE1 had the best predictive power for 1-year prognosis in ESCC patients. In addition, survival analysis showed that patients in the PRICKLE1 high expression group had significantly better overall survival and better prognostic significance in ESCC patients. Through experimental validation, we found that the expression of PRICKLE1 was reduced in esophageal cancer; PRICKLE1 was able to attenuate Wnt/β-catenin signaling; and the cell viability was reduced, the migration ability was significantly reduced and the apoptosis rate was significantly higher in the PRICKLE1-OE group compared with the NC group.This indicates that PRICKLE1 is closely related to the development of ESCC cells. as a class of non-classical Wnt signaling pathway core proteins, PRICKLE1 can effectively inhibit the proliferation and migration of ESCC and increase apoptosis by attenuating the effect of Wnt/β-catenin signaling.

In summary, high PRICKLE1 expression can be used to predict the survival of ESCC patients and can be used as a prognostic indicator for ESCC patients, providing a potential application for ESCC clinical treatment. The present study identified the relationship between PRICKLE1 and the prognosis of ESCC patients, but further elucidation of more detailed regulatory mechanisms is still needed in future studies.In addition, this study has not determined whether Wnt signaling pathway inhibitors are more effective in ESCC patients with high PRICKLE1 expression, and we await further studies in the future.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

## Author contributions

JH, GL, and HT: writing the article, conception and design, data collection; YW: writing the article, data collection, critical revision of the article; LY: conception and design, analysis and interpretation, data collection; WL: analysis and interpretation, data collection, writing the article; JH: obtaining funding, critical revision of the article, writing the article; WS: critical revision of the article, data collection, writing the article. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Ningbo Clinical Research Center for thoracic & breast neoplasms (2021L002), NINGBO Medical & Health Leading Academic Discipline Project (2022-F02), the major science and technology projects of Zhejiang, China (2020C03058), and Zhejiang Lung Cancer Center (JBZX-202007).

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Li S, Chen H, Man J, Zhang T, Yin X, He Q, et al. Changing trends in the disease burden of esophageal cancer in China from 1990 to 2017 and its predicted level in 25 years. *Cancer Med* (2021) 10(5):1889–99. doi: 10.1002/cam4.3775

2. Kakeji Y, Oshikiri T, Takiguchi G, Kanaji S, Matsuda T, Nakamura T, et al. Multimodality approaches to control esophageal cancer: development of chemoradiotherapy, chemotherapy, and immunotherapy. *Esophagus.* (2021) 18(1):25–32. doi: 10.1007/s10388-020-00782-1

3. Huang J, Koulaouzidis A, Marlicz W, Lok V, Chu C, Ngai CH, et al. Global burden, risk factors, and trends of esophageal cancer: an analysis of cancer registries from 48 countries. *Cancers.* (2021) 13(1):141. doi: 10.3390/cancers13010141

4. Li J, Xu J, Zheng Y, Gao Y, He S, Li H, et al. Esophageal cancer: Epidemiology, risk factors and screening. *Chin J Cancer Res* (2021) 33(5):535. doi: 10.21147/j.issn.1000-9604.2021.05.01

5. Li B, Liu Y, Peng J, Sun C, Rang W. Trends of esophageal cancer incidence and mortality and its influencing factors in China. *Risk Manage Healthcare Policy* (2021) 14:4809. doi: 10.2147/RMHP.S312790

6. He Z, Ke Y. Precision screening for esophageal squamous cell carcinoma in China. *Chin J Cancer Res* (2020) 32(6):673. doi: 10.21147/j.issn.1000-9604.2020.06.01

7. Cui Y, Chen H, Xi R, Cui H, Zhao Y, Xu E, et al. Whole-genome sequencing of 508 patients identifies key molecular features associated with poor prognosis in esophageal squamous cell carcinoma. *Cell Res* (2020) 30(10):902–13. doi: 10.1038/s41422-020-0333-6

8. Gou Y, Jin D, He S, Han S, Bai Q. RNF168 is highly expressed in esophageal squamous cell carcinoma and contributes to the malignant behaviors in association with the wnt/β-catenin signaling pathway. *Aging (Albany NY)* (2021) 13(4):5403. doi: 10.18632/aging.202471

9. Zhao YL, Li JB, Li YJ, Li SJ, Zhou SH, Xia H. Capn4 promotes esophageal squamous cell carcinoma metastasis by regulating ZEB1 through the wnt/β-catenin signaling pathway. *Thorac Cancer* (2019) 10(1):24–32. doi: 10.1111/1759-7714.12893

10. Taciak B, Pruszynska I, Kiraga L, Bialasek M, Król M. Wnt signaling pathway in development and cancer. *J Physiol Pharmacol* (2018) 69(2):185–96. doi: 10.26402/jpp.2018.2.07

11. Ghosh N, Hossain U, Mandal A, Sil PC. The wnt signaling pathway: a potential therapeutic target against cancer. *Ann New York Acad Sci* (2019) 1443(1):54–74. doi: 10.1111/nyas.14027

12. Jackstadt R, Hodder MC, Sansom OJ. WNT and β-catenin in cancer: Genes and therapy. *Annu Rev Cancer Biol* (2020) 4:177–96. doi: 10.1146/annurev-cancerbio-030419-033628

13. Codipilly DC, Qin Y, Dawsey SM, Kisiel J, Topazian M, Ahlquist D, et al. Screening for esophageal squamous cell carcinoma: recent advances. *Gastrointestinal endoscopy* (2018) 88(3):413–26. doi: 10.1016/j.gie.2018.04.2352

14. Hamai Y, Emi M, Ibuki Y, Murakami Y, Nishibuchi I, Nagata Y, et al. Early recurrence and cancer death after trimodal therapy for esophageal squamous cell carcinoma. *Anticancer Res* (2019) 39(3):1433–40. doi: 10.21873/anticanres.13259

15. Abbaszadegan MR, Riahi A, Forghanifard MM, Moghbeli M. WNT and NOTCH signaling pathways as activators for epidermal growth factor receptor in esophageal squamous cell carcinoma. *Cell Mol Biol letters* (2018) 23(1):1–9. doi: 10.1186/s11658-018-0109-x

16. Wu Y-F, Wang C-Y, Tang W-C, Lee Y-C, Ta HDK, Lin L-C, et al. Expression profile and prognostic value of wnt signaling pathway molecules in colorectal cancer. *Biomedicines.* (2021) 9(10):1331. doi: 10.3390/biomedicines9101331

17. Zhang Y, Wang X. Targeting the wnt/β-catenin signaling pathway in cancer. *J Hematol Oncol* (2020) 13(1):1–16. doi: 10.1186/s13045-020-00990-3

18. Jung Y-S, Park J-I. Wnt signaling in cancer: therapeutic targeting of wnt signaling beyond β-catenin and the destruction complex. *Exp Mol Med* (2020) 52(2):183–91. doi: 10.1038/s12276-020-0380-6

19. Nakajima M, Fukuchi M, Miyazaki T, Masuda N, Kato H, Kuwano H. Reduced expression of axin correlates with tumour progression of oesophageal squamous cell carcinoma. *Br J cancer* (2003) 88(11):1734–9. doi: 10.1038/sj.bjc.6600941

20. Chu C-Y, Wang R, Liu X-L. Roles of wnt/β-catenin signaling pathway related microRNAs in esophageal cancer. *World J Clin Cases* (2022) 10(9):2678. doi: 10.12998/wjcc.v10.i9.2678

21. Deng F, Zhou K, Cui W, Liu D, Ma Y. Clinicopathological significance of wnt/β-catenin signaling pathway in esophageal squamous cell carcinoma. *Int J Clin Exp pathol* (2015) 8(3):3045.

22. Daulat AM, Finetti P, Revinski D, Silveira Wagner M, Camoin L, Audebert S, et al. ECT2 associated to PRICKLE1 are poor-prognosis markers in triple-negative breast cancer. *Br J cancer* (2019) 120(9):931–40. doi: 10.1038/s41416-019-0448-z

23. Jiang D, He Y, Mo Q, Liu E, Li X, Huang L, et al. PRICKLE1, a Wnt/PCP signaling component, is overexpressed and associated with inferior prognosis in acute myeloid leukemia. *J Trans Med* (2021) 19(1):1–17. doi: 10.1186/s12967-021-02873-8

24. Daulat AM, Bertucci F, Audebert S, Sergé A, Finetti P, Josselin E, et al. PRICKLE1 contributes to cancer cell dissemination through its interaction with mTORC2. *Dev Cell* (2016) 37(4):311–25. doi: 10.1016/j.devcel.2016.04.011

25. Zhou R, Tang Z, Li H, Wang X, Sun Y. PRICKLE1 promotes gastric cancer metastasis by activating mTOR signaling. *Am J Trans Res* (2021) 13(5):4266.

26. Pehlivan M, Cerci B, Sercan HO. The intracellular negative regulator genes of the wnt signaling in imatinib treatment. (2021). doi: 10.5455/annalsmedres.2021.01.111

27. Galamb O, Kalmár A, Péterfia B, Csabai I, Bodor A, Ribli D, et al. Aberrant DNA methylation of WNT pathway genes in the development and progression of CIMP-negative colorectal cancer. *Epigenetics* (2016) 11(8):588–602. doi: 10.1080/15592294.2016.1190894

28. Groulx J-F, Giroux V, Beauséjour M, Boudjadi S, Basora N, Carrier JC, et al. Integrin α6A splice variant regulates proliferation and the wnt/β-catenin pathway in human colorectal cancer cells. *Carcinogenesis* (2014) 35(6):1217–27. doi: 10.1093/carcin/bgu006

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 06 March 2023
DOI 10.3389/fonc.2022.1059547

Check for updates

OPEN ACCESS

EDITED BY
Ye Wang,
The Second Affiliated Hospital of Medical College of Qingdao University, China

REVIEWED BY
Fengxuan Tian,
First Affiliated Hospital of Xinjiang Medical University, China
Yuqing Hang,
Salk Institute for Biological Studies, United States

*CORRESPONDENCE
Shengbao Wang
wsb887600@163.com

†These authors share first authorship

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

RECEIVED 01 October 2022
ACCEPTED 23 December 2022
PUBLISHED 06 March 2023

CITATION
Feng H, Wang L, Liu J and Wang S (2023) The bioinformatic approach identifies *PARM1* as a new potential prognostic factor in osteosarcoma. *Front. Oncol.* 12:1059547. doi: 10.3389/fonc.2022.1059547

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

# The bioinformatic approach identifies *PARM1* as a new potential prognostic factor in osteosarcoma

Haijun Feng[1†], Liping Wang[1†], Jie Liu[2] and Shengbao Wang[3*]

[1]Department of Orthopedics, Second Hospital of Lanzhou University, Lanzhou, Gansu, China, [2]Department of Neurosurgery, Liaocheng Second People's Hospital, Liaocheng, Shandong, China, [3]Second Hospital of Lanzhou University, Lanzhou, Gansu, China

**Objective:** To explore the key factors affecting the prognosis of osteosarcoma patients.

**Methods:** Based on the GEO dataset and differential expression analysis of normal and osteosarcoma tissues, the gene modules related to the prognosis of osteosarcoma patients were screened by WGCNA, and intersecting genes were taken with differential genes, and the risk prognosis model of osteosarcoma patients was constructed by LASSO regression analysis of intersecting genes, and the prognosis-related factors of osteosarcoma patients were obtained by survival analysis, followed by target for validation, and finally, the expression of prognostic factors and their effects on osteosarcoma cell migration were verified by cellular assays and lentiviral transfection experiments.

**Results:** The prognosis-related gene module of osteosarcoma patients were intersected with differential genes to obtain a total of 9 common genes. PARM1 was found to be a prognostic factor in osteosarcoma patients by LASSO regression analysis, followed by cellular assays to verify that PARM1 was lowly expressed in osteosarcoma cells and that overexpression of PARM1 in osteosarcoma cells inhibited cell migration. Pan-cancer analysis showed that PARM1 was lowly expressed in most cancers and that low expression of PARM1 predicted poor prognosis for patients.

**Conclusion:** The data from this study suggest that PARM1 is closely associated with the prognosis of osteosarcoma patients, and PARM1 may serve as a novel potential prognostic target for osteosarcoma, providing a heartfelt direction for the prevention and treatment of osteosarcoma.

KEYWORDS

**osteosarcoma, prognosis, PARM1, bioinformatics, GEO**

# 1 Introduction

Osteosarcoma is common primary malignant bone tumors of mesenchymal (1, 2). Osteosarcoma occurs in the blood-rich epiphyses of long bones of the extremities, such as the distal femur and proximal tibia. The clinical symptoms of osteosarcoma are usually local swelling and pain, occasionally accompanied by joint mobility disorders, and it has a high tendency to metastasize, and is highly susceptible to early metastases, especially pulmonary metastases, and the disease progresses rapidly (3, 4). The treatment of osteosarcoma has evolved from early surgical amputation alone to comprehensive treatment with neoadjuvant chemotherapy combined with limb-preserving surgery (5, 6). The 5-year survival rate of patients with limited osteosarcoma has improved from 20% to 60%-70% with this combined treatment modality, but further improvement in treatment outcomes is bottlenecked at this stage, while the survival rate of patients with metastatic and recurrent osteosarcoma remains low (7–9). Therefore, it is still urgent to deeply investigate the complex cellular mechanisms and molecular signaling mechanisms of pathogenesis and progression in osteosarcoma.

With the development of technology in recent years, high-throughput sequencing technology has been widely used and is an important tool for diseases such as cancer. The GEO database contains a large amount of oncogene expression data (10–12). For example, Guo et al. (13) identified pivotal genes and pathways in a rat model of renal ischemia-reperfusion injury based on bioinformatics analysis of GEO microarray datasets and integration of gene expression profiles. Ma et al. (14) screened for differential mRNA expression between brain tissue and blood by mining Alzheimer's patient data in the GEO dataset and found that ITGB1 and RAB7A could be used as biomarkers for Alzheimer's disease. In addition, Li et al. (15) identified *MMP9, CD74, SPP1, CXCL12, TYROBP, FCER1G* et al. as central genes for osteosarcoma initiation through the GEO database.

To screen for differentially expressed genes in osteosarcoma cells compared to normal cells, find the biological role of differential genes in cancer cells, verify the correlation between differential genes and survival prognosis of osteosarcoma patients, and provide a theoretical basis for identifying targets for clinical osteosarcoma treatment.

# 2 Methods

## 2.1 Data source

All the data in this study come from the target database (https://ocg.cancer.gov/programs/target), and the GEO database (https://www.ncbi.nlm.nih.gov/geo/), the target database contains the mRNA expression profile and clinical data of 98 osteosarcoma patients were obtained from the GEO database GSE16088 and GSE16091 data sets. The mRNA expression profile data and clinical data of 6 normal controls and 48 osteosarcoma patients were obtained. Normalize and debatch the obtained GEO data, standardize the samples, and view the clustering of the sample groups through the PCA graph and the UMAP graph.

## 2.2 Expression difference analysis

The DESeq2 package of R software was used for expression profile difference analysis as well as gene volcano and heat map plotting. Corrected P values and |log2FC| were calculated, with $P<0.05$ and $|log2FC|\geq1.0$ for differentially expressed genes. The Gene Ontology (GO) analysis was visualized using the ClusterProfiler package in R software, and a false discovery rate <0.01 was statistically significant.

## 2.3 WGCNA

The GSE16091 dataset included expression profiles, and survival events of 34 osteosarcoma tissues, which were suitable for the construction of weighted gene co-expression networks. GSE160306 and GSE160308 gene expression data matrices were constructed using the R package " WGCNA", and the top 25% of genes with the largest differences in the samples were selected as the input dataset for the subsequent WGCNA. A sample hierarchical clustering approach was used to detect and remove abnormal samples before selecting the appropriate soft threshold function. The neighbor-joining matrix and topological overlap matrix were constructed to calculate the corresponding dissimilarities and complete the identification of gene trees and modules. The minimum module size was 20. then the highly similar modules were merged by clustering and fusing module feature genes. The degree of difference is less than 0.25. Average linkage hierarchical clustering was established, similar genes were divided into a module, and osteosarcoma grading was used as phenotype data to screen the optimal module.

## 2.4 Survival analysis

The impact of prognostic survival of osteosarcoma was assessed by overall survival (OS). All survival analyses were performed using COX multiple survival regression model analysis as the primary method and the Kaplan-Meier method as an adjunct. The endpoint of overall survival was defined as the time from the time of randomization to death from any cause,

and all methods and R packages were performed using R software version v4.0.3. $p < 0.05$ was statistically significant.

## 2.5 Prognostic feature models

The collection between gene expression and OS was assessed using LASSO COX analysis. A prognostic risk prediction model for osteosarcoma was built. Patients with osteosarcoma were divided into high and low-risk groups based on the median risk score. KM curves were plotted to compare OS between the high-risk and low-risk parts, and ROC survival analysis was performed using the R package SURVIVAL, and the "rmda" package was used for decision curve analysis.

## 2.6 Cell culture and transfection

The derived human osteoblast cell line hfob 1.19 and osteosarcoma cell line saos-2 were cultured in a 1:1 mixture of Ham's F12 medium and Dulbecco's modified Eagle's medium containing 2.5 mM L -gu Aminoamide (ATCC, LGS standard, PL), supplemented with 100 U/mL penicillin, 100 U/mL streptomycin (Sigma-Aldrich, Warsaw, Poland), 0.3 mg/mL G418 (Sigma-Aldrich, Warsaw, Poland) and 10% fetal bovine serum (v/v, FBS, Gibco, Thermo Fisher Scientific, Warsaw, Poland). Cells were grown at 34°C, 5% $CO_2$. The pc DNA3.1 transfection plasmid was purchased from Genechem Co., Ltd. (Shanghai, China). Osaos-2 cells were seeded in 6-well plates ($2\times10^5$/well), and after incubation for 24 hours, they were transfected with LiPofectamine 2000 (corresponding plasmid 100 nM), and the operation was performed according to the instructions. Subsequent experiments were performed 48 hours after transfection.

## 2.7 Quantitative reverse transcription PCR (RT-qPCR)

Total RNA was extracted from each group of cells using TRIzol reagent (Invitrogen). cDNA was synthesized from RNA using a Primescript RT kit (Takara, Japan). The cDNA was amplified and quantified using SYBR Green mix (Finnish Finnzymes) in an Applied Biosystems 7500 instrument. The primer sequences used are shown in Table 1. $-2^{\Delta\Delta Ct}$ method could determine the gene expression.

## 2.8 Western Blot (WB)

Cells were collected, lysed in lysate (phosphatase inhibitor, protease inhibitor and PMSF), and protein quantification was performed by BCA method (Thermo Fisher Scientific, China). 10-20 μg protein was loaded on 8%-12% 30% acrylamide-Bis gel, then transferred through 0.22 μm pore PVDF membrane (Merck Millipore, USA), and blocked with 5% nonfat milk powder for 1 h., respectively, add primary antibodies (Rabbit polyclonal to IFIT1, 1:1000, ab236256; Rabbit monoclonal [EPR10009] to PARM-1, 1:1000, ab168369, abcam) and incubate overnight at 4°C. The next day, secondary antibodies (Goat Anti-Rabbit actin, 1:2000, ab8226, abcam) were added and incubated at room temperature for 1 h, and then luminescent solution (Thermo Fisher Scientific, China) was added for exposure and color development. Image J software was used for analysis, and the relative protein content was expressed as the gray value of the corresponding protein band/the gray value of the actin protein band, which was repeated three times.

## 2.9 Transwell

Digest the cells and adjust to 1×106 cells/mL in culture medium without serum. Matrigel (BD Biosciences, San Jose, CA, USA) was added to the upper chamber and incubated at 37°C for 2 h. Add 100 μL of cell suspension to the upper chamber and 600 μL of cell culture medium with FBS to the bottom chamber. Subsequently, the cells were incubated at 37°Cand 5% $CO_2$ in an ambient environment for 24 hours. Detach the upper chamber, remove the cells, and fix with 4% paraformaldehyde. Cells were washed with phosphate-buffered saline (PBS) and stained with crystal violet. Stained cells were observed with a microscope (Olympus, Tokyo, Japan).

## 2.10 Statistical analysis of data

SPSS 24.0 and GraphPad Prism 8.0.1 software can process the data. Normally distributed measures are expressed as ($x \pm s$), and comparisons between two parts were made by independent samples t-test, and P was a two-sided test, and differences were statistically significant at $P < 0.05$.

# 3 Results

## 3.1 Differential gene screening

We downloaded the dataset of GSE16088 from the GEO database, which included 14 cases of osteosarcoma tissues and 6 cases of normal tissues, first, we analyzed the datasets, which showed good normalization and differences between samples (Figures 1A–C). We then analyzed the differentially expressed genes in normal and osteosarcoma tissues, and there were 5005 differentially expressed genes (2719 significantly up-regulated

TABLE 1 Primer sequences.

| Name | sequence |
|---|---|
| *IFIT1*-F | TCAGAAGTCTAGTCACTTGGGG |
| *IFIT1*-R | ACACCTTCGCCCTTTCATTTC |
| *PARM1*-F | TCTTAACTGCAGGATGGAGGG |
| *PARM1*-R | TAGTCCAGATGGTGGTCGGT |
| GAPDH-F | TTGATTTTGGAGGGATCTCGCTC |
| GAPDH-R | GAGTCAACGGATTTGGTCGTATTG |

genes, 2286 down-regulated genes), Figures 1D, E show the differential gene volcano and heat map.

## 3.2 WGCNA identification of genes associated with osteosarcoma survival orientation

To identify genes associated with osteosarcoma survival orientation, we constructed co-expression networks by WGCNA in GSE16091, and in this study, we chose β=6 as the threshold to construct scaling-free networks (Figures 2A, B), and a total of 25 gene expression modules were identified after using merged dynamic tree cuts (Figure 2C), by constructing a random gene network map (Figure 2D). By calculating the correlation between module feature genes and clinical features, the skyblue3 module was most strongly correlated with the survival of osteosarcoma patients (Figure 2E), and the neighbor-joining matrix was clustered with the clinical features of the samples in Figure 2F, and the scatter plot of skyblue3 module genes in Figure 2G. We then obtained a total of 9 genes (*FI27 ISG15, IFI6, IFI44L, PARM1, MX1, IFIH1, IFIT1, RSAD2*) by taking the intersection of skyblue3 module genes and DEGs Figure 2H.

## 3.3 Prognostic risk construction based on intersecting genes

To explore the prognostic impact of intersecting genes of modular genes and DEGs on osteosarcoma, we did prognostic models based on LASSO analysis for these nine genes, obtained the value of the independent variable lambda and the coefficient of the independent variable (Figure 3A), and plotted the partial likelihood deviation versus log(λ) (Figure 3B). Patients with osteosarcoma were subsequently divided into high and low-risk groups (Figure 3C) by the hazard score Riskscore=(-0.8635)* *IFIT1*+(-0.481)* *PARM1*, and their survival status was presented in Figure 3C. Figure 3D shows the difference in survival between the high and low-risk parts (p=4.3e-3). Prognostic survival was predicted for 1, 3, and 5 years, and this showed good sensitivity and specificity (Figure 3E).

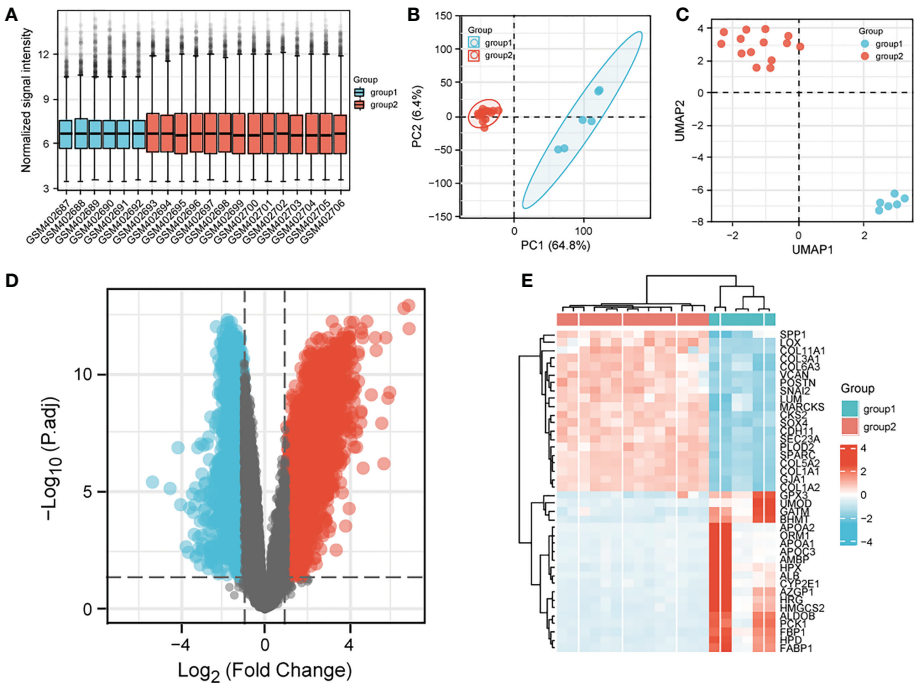



FIGURE 1
Analysis of differential genes in the GSE16088 dataset. **(A)** sample normalized box plot, vertical coordinates indicate sample normalized intensity; **(B)** sample PCA plot; **(C)** UMAP clustering plot; **(D)** volcano plot; **(E)** heat map; Top20 differential genes are shown. Normal in the figure indicates normal tissue and the OS indicates osteosarcoma tissue.

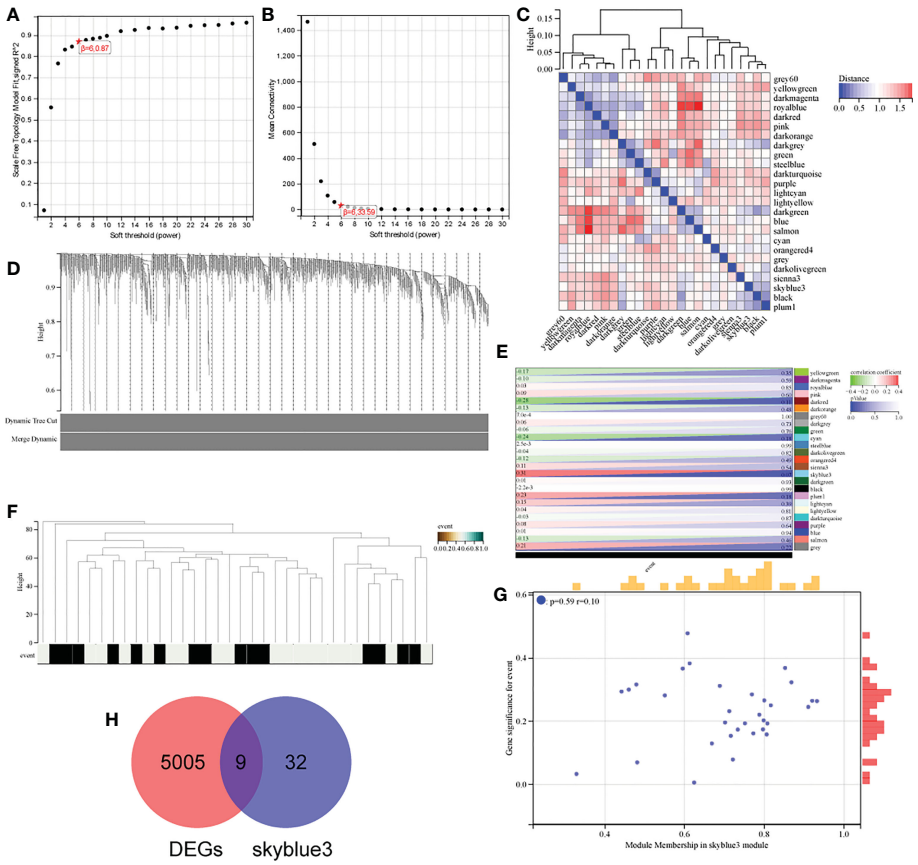



FIGURE 2
Identification of genes associated with osteosarcoma survival orientation. **(A, B)** analysis of the average connectivity of the scale-free fit index and various soft threshold functions. Assessment of scale-free topology at β = 6; **(C)** Gene module distribution mapping; **(D)** Gene vs. sample clustering map; **(E)** Heat map of module feature genes associated with individual osteosarcoma survival; **(F)** Tree and clinical feature heat map; **(G)** Scatter plot of skyblue3 module genes; **(H)** Venn diagram of skyblue3 module genes and DEGs.

## 3.4 Effect of *IFIT1* and *PARM1* on the prognosis of patients with osteosarcoma

In order to further investigate the role of *IFIT1* and *PARM1* in patients with osteosarcoma, we analyzed the expression and prognostic effects of *IFIT1* and *PARM1* in patients with osteosarcoma, using the risk scores of *IFIT1* and *PARM1*. *IFIT1* expression was significantly increased in osteosarcoma tissues compared to normal tissues (Figure 4A), and patients with low *IFIT1* expression had a better prognosis (Figure 4B, p=0.07). *PARM1* was lowly expressed in osteosarcoma tissues (Figure 4C), and low expression of *IFIT1* predicted a poor prognosis for patients (Figure 4D).

## 3.5 Cellular experiments to verify the expression of *IFIT1* and *PARM1*

To further verify the heterogeneity of *IFIT1* and *PARM1* in normal and osteosarcoma tissues, we cultured the osteoblast cell line hfob 1.19 and the osteosarcoma cell line saos-2, and detected the expression of *IFIT1* and *PARM1* by RT-qPCR, and the results showed that compared with osteoblasts, osteosarcoma cells had high expression of *IFIT1* (Figure 5A), *PARM1* was lowly expressed (Figure 5B). Then we detected the protein levels of IFIT1 and *PARM1* in the osteoblast cell line hfob 1.19 and the osteosarcoma cell line saos-2 by WB. The results showed that the expression of *IFIT1* was high and the expression of *PARM1* was low in the osteosarcoma cell line saos-2 (Figure 5C). The above data show that *IFIT1* is highly expressed and *PARM1* is lowly expressed in osteosarcoma tissues. Subsequently, we overexpressed PARM1 in saos-2 cells, the transfection efficiency is shown in Figure 5D, and on this basis, we detected the cell migration ability. The results showed that overexpressing PARM1 in osteosarcoma cells significantly reduced the migration ability of cells (Figure 5E, p<0.01).

## 3.6 Target database verification

In order to further verify the correctness of the constructed risk model, we constructed a prognostic risk model based on the

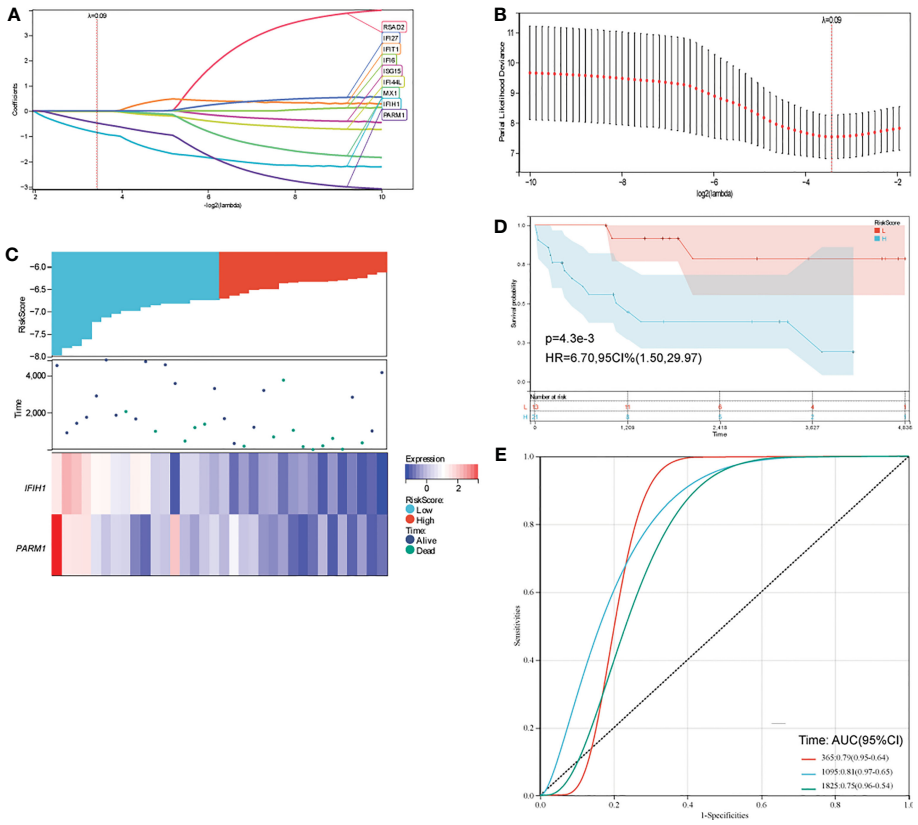



FIGURE 3
**(A, B)** distribution of LASSO coefficients for key genes to obtain the adjustment parameter l.min= 0.09, and the vertical black dashed line in B defines the optimal l value; **(C)** Riskscore distribution, survival status and duration, and gene expression of prognostic characteristics of osteosarcoma patients are shown from top to bottom; **(D)** KM survival curves for high and low-risk groups; **(E)** ROC curves at 1, 3 and 5 years.

target using 9 genes, *FI27 ISG15, IFI6, IFI44L, PARM1, MX1, IFIH1, IFIT1, RSAD2*, and obtained the independent variable lambda values and the coefficients of the independent variable (Figure 6A), and plotted the partial likelihood deviation versus log(λ) (Figure 6B). Subsequently, patients with osteosarcoma were divided into high-risk group and low-risk group by the risk score Riskscore=(-0.0921)**PARM1*+(-0.1418)**IFIH1* (Figure 6C), and their survival status was shown in Figure 3C. Figure 6D shows the difference in survival between high and low risk groups (P=0.00725). Prognostic survival was predicted for 1, 3, and 5 years, and the prognostic model showed good sensitivity and specificity (Figure 6E). Subsequently, we analyzed the effect of high or low expression of PARM1 on the prognosis of patients with osteosarcoma in the target database. Figure 7A shows the distribution of patients with osteosarcoma. The KM curve shows that patients with high PARM1 expression have a better prognosis (Figure 7B), and the ROC curve shows that patients with osteosarcoma have a better prognosis. Good sensitivity (Figure 7C).

## 3.7 *PARM1* pan-cancer expression and prognostic analysis

In the previous data, we confirmed that low expression of *PARM1* predicted poor prognosis in osteosarcoma patients. To further elaborate on the role of *PARM1* in cancer, we performed a pan-cancer expression and prognosis analysis of *PARM1*. As shown in Figure 8A, *PARM1* was lowly expressed in BLCA, BRCA, CESC, COAD, GBM, HNSC, KICH, KIRC, KIRP, LUAD, LUSC, PRAD, PEAD, THCA, and UCEC, and highly expressed in CHOL and LIHC. In addition, the expression level of *PARM1* was closely correlated with the prognosis of COAD, KIRC, LUAD, and THYM, where low expression of *PARM1* in COAD, KIRC, and LUAD predicted poor patient prognosis (Figure 8B).

# 4 Discussion

Abnormal gene expression is a vital factor in the development and metastasis of osteosarcoma, and the genes

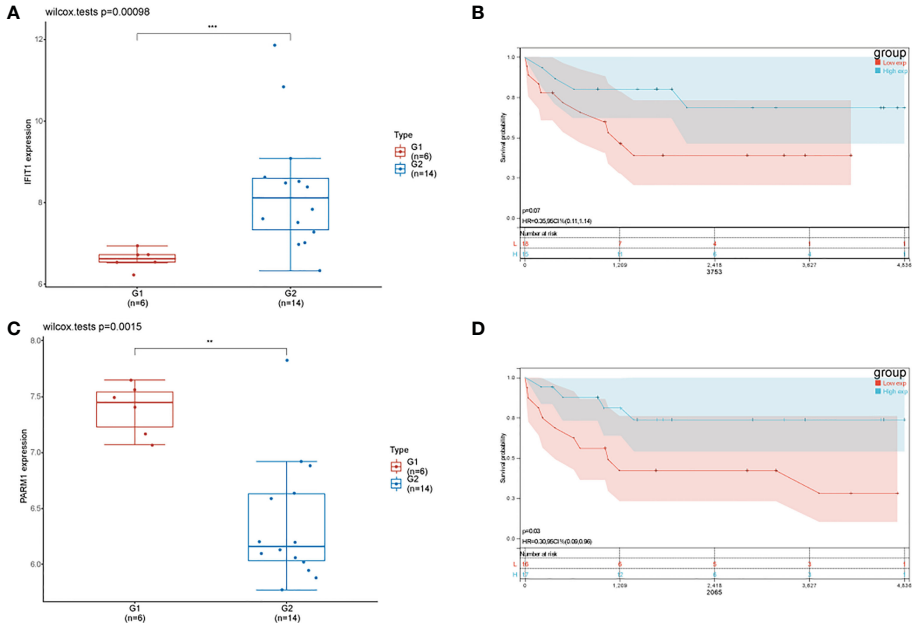



FIGURE 4
Effect of *IFIT1* and *PARM1* on the prognosis of osteosarcoma patients. **(A)** Expression of *IFIT1* in normal and osteosarcoma tissues in GEO dataset; **(B)** KM curve showing the effect of high and low expression of *IFIT1* on the prognosis of osteosarcoma patients; **(C)** Expression of *PARM1* in normal and osteosarcoma tissues in GEO dataset; **(D)** KM curve showing the effect of high and low expression of PARM1 on the prognosis of osteosarcoma patients; ** means p<0.01, *** means p<0.001.

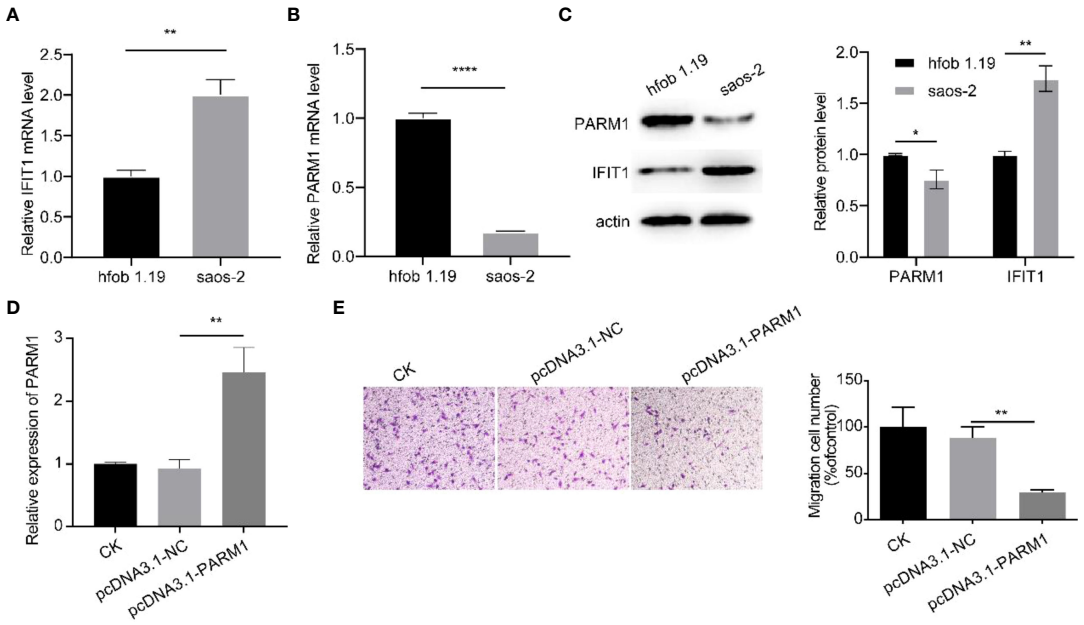



FIGURE 5
Cellular experiments to verify the expression of *IFIT1* and *PARM1*. **(A, B)**: RT-qPCR to detect the mRNA levels of *IFIT1* and *PARM1* in normal and osteosarcoma tissues. **(C)** WB detection of IFIT1 and *PARM1* protein levels in normal and osteosarcoma tissues. **(D)** RT-qPCR detection of mRAN levels of PARM1 in different groups; **(E)**: transwell detection of cell migration ability in different groups; ** means p<0.01, *** means p<0.001, **** means p<0.0001.

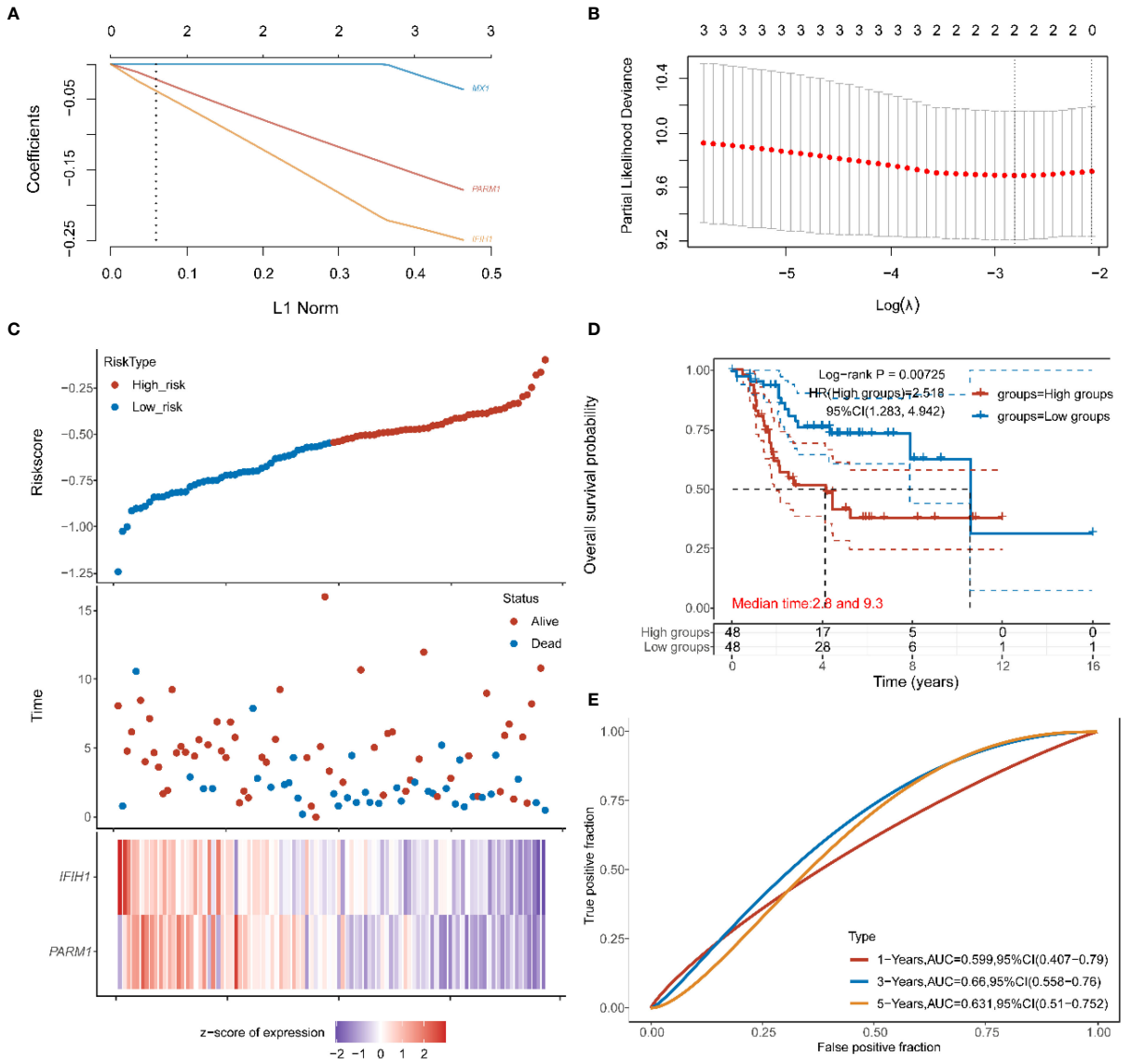



FIGURE 6
**(A, B)** in the target data: the distribution map of the LASSO coefficient of the key gene, and the adjusted parameter l.min= 0.0599, the vertical black dotted line in **(B)** defines the optimal l value; **(C)** Riskscore distribution, survival status and duration, and gene expression of prognostic characteristics of osteosarcoma patients are shown from top to bottom; **(D)** KM survival curves of high and low risk groups; **(E)** ROC curves of 1 year, 3 years and 5 years.

that produce abnormal expression in this complex biological process are necessarily not single, so it is important to analyze and find new molecular markers that affect the progression of osteosarcoma (16, 17). In the past decades of research, the exploration of aberrantly expressed genes in osteosarcoma has never stopped. To find the role of ferroptosis death-related genes in the development of osteosarcoma, Ding et al. (18) classified osteosarcoma by the method of non-negative matrix decomposition clustering and found significant differences in prognosis and immune status between two subgroups, and found that ferroptosis death-related genes could influence the immune status of the tumor microenvironment, thus affecting the development and prognosis of osteosarcoma. In addition, lipid metabolism genes were used to cluster osteosarcoma patients into two subgroups by consensus, and it was found that the expression of lipid metabolism genes correlated with the immune microenvironment and could be used to accurately predict the osteosarcoma prognosis (19).

In this study, we first compared the differential genes between normal tissues and osteosarcoma tissues in the

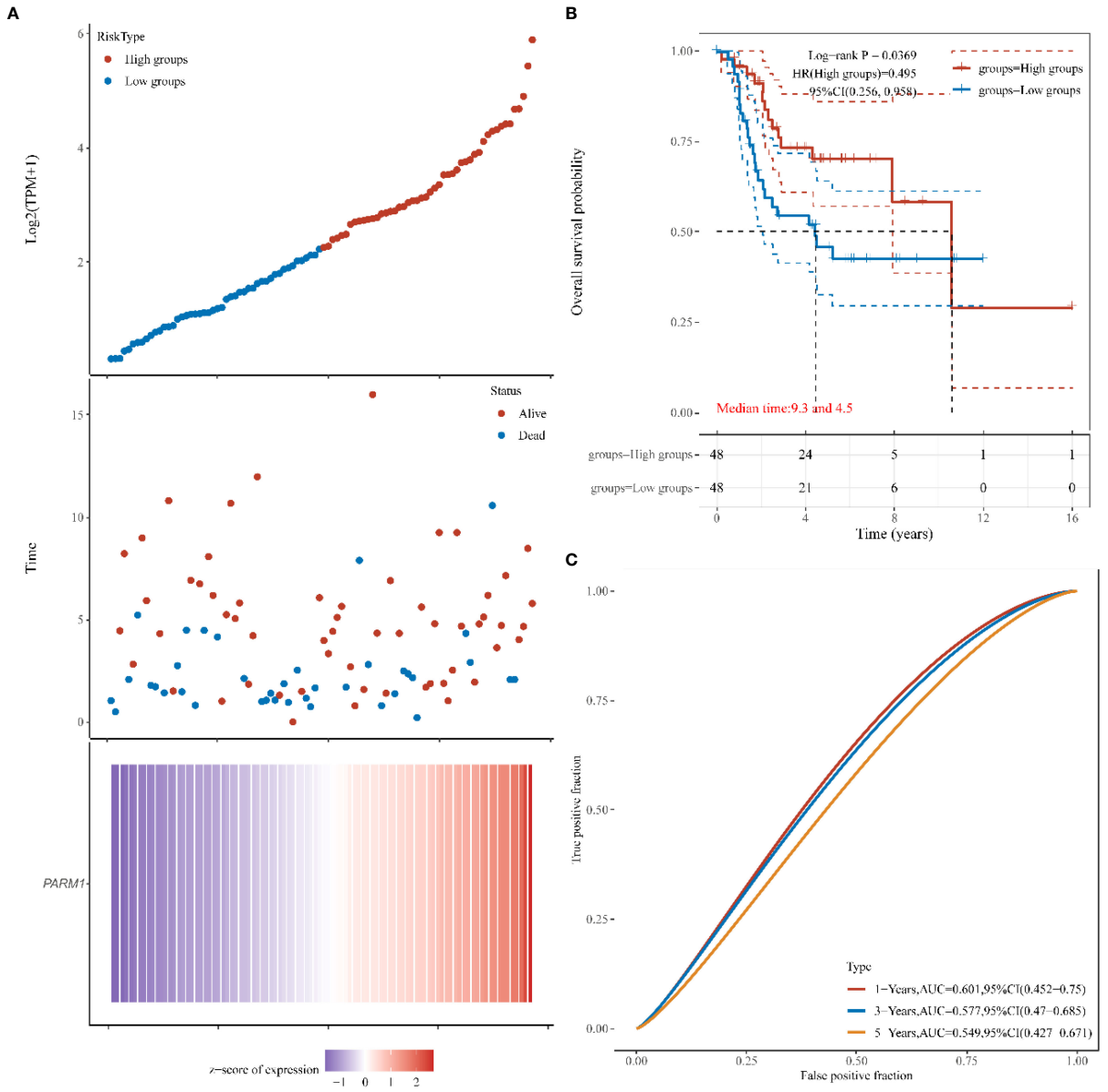



FIGURE 7
The effect of high or low expression of *PARM1* on the prognosis of osteosarcoma patients. **(A)**: *PARM1* expression distribution and survival distribution of osteosarcoma patients; **(B)**: *PARM1* high and low expression survival curve; **(C)**: ROC curves of 1 year, 3 years and 5 years.

GSE16088 set, and found 5005 cases of differential genes. In order to further explore the impact of differential genes on the survival of osteosarcoma. We performed WGCNA analysis on 34 patients in the GSE16091 data set to screen the genes related to the survival of osteosarcoma patients. The results were intersected with the differential genes and found that a total of 9 genes were highly correlated with the survival of osteosarcoma. Then the prognostic model was constructed by LASSO and Survival analysis identified *PARM1* as a favorable factor for osteosarcoma patients. In order to further verify the accuracy of our model construction, we built a prognostic risk model in the target database, and further tested that low expression of *PARM1* indicates poor prognosis of osteosarcoma patients, which further verified the accuracy of our model.

In addition, cellular experiments were performed to verify *PARM1* was lowly expressed in the osteosarcoma cell line saos-2. Studies have shown that *PARM1* is considered a new potential oncogene (20). Prostate androgen-regulated mucin-like protein 1 is an integral secreted protein that accumulates mainly in the Golgi apparatus as well as in early and late endosomes, and *PARM-1* increases ERK1/2, AKT, and STAT3 phosphorylation. In colorectal cancer, *PARM1* may be its potential novel

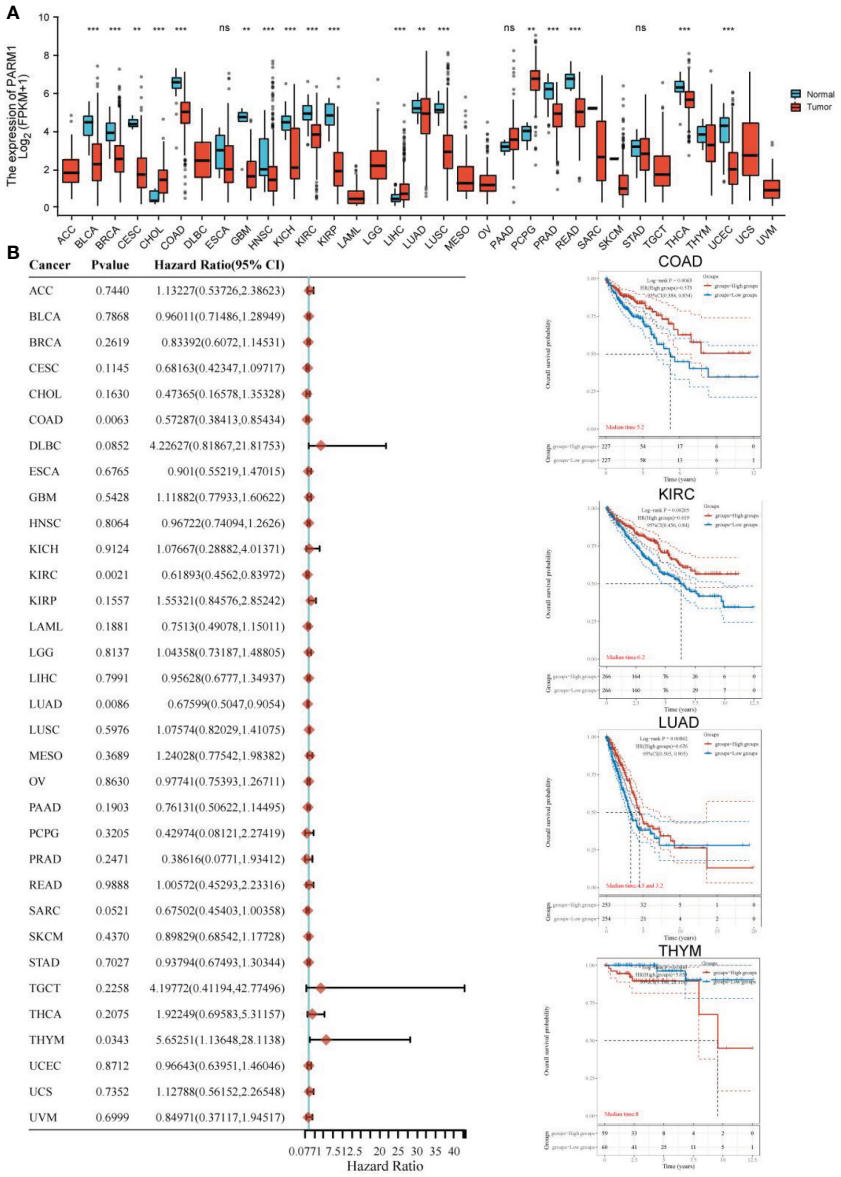



FIGURE 8
*PARM1* pan-cancer expression and prognosis analysis. **(A)** Expression of *PARM1* in cancer; **(B)** Effect of PARM1 expression level on cancer prognosis; ** means p<0.01, *** means p<0.001.

prognostic biomarker (21). In prostate cancer cells, *PARM1* ectopic expression increased cell proliferation (22). In addition, *PARM1* expression is stimulated by pro-inflammatory cytokines and is associated with endoplasmic reticulum stress (23).

We performed a pan-cancer expression and prognostic analysis of *PARM1* and found that *PARM1* showed a trend of low expression in most cancers, and low expression of *PARM1* in COAD, KIRC, and LUAD predicted poor prognosis of patients, which also confirmed *PARM1* as a poor prognostic factor. In conclusion, in our study, we identified *PARM1* as a poor prognostic factor in osteosarcoma by bioinformatics based on the GEO database, and low expression of *PARM1* predicted poor prognosis in osteosarcoma patients. Although we verified the heterogeneous expression of *PARM1* in normal tissues and osteosarcoma cells in cellular experiments, to further explore the mechanism of action of *PARM1* affecting the prognosis of osteosarcoma patients.

## Data availability statement

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

## Author contributions

HF: Writing the article, data collection, conception and design. LW: Conception and design, analysis and interpretation, critical revision of the article. JL: Writing the article, critical revision of the article, methodology. SW: Formal analysis, original draft, critical revision of the article. All authors contributed to the article and approved the submitted version.

## Conflict of interest

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

## Publisher's note

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.2022.1059547/full#supplementary-material

## References

1. Simpson E, Brown HL. Understanding osteosarcomas. *JAAPA* (2018) 31 (8):15–9. doi: 10.1097/01.JAA.0000541477.24116.8d

2. Kager L, Tamamyan G, Bielack S. Novel insights and therapeutic interventions for pediatric osteosarcoma. *Future Oncol (London England)* (2017) 13(4):357–68. doi: 10.2217/fon-2016-0261

3. Nørregaard KS, Jürgensen HJ, Gårdsvoll H, Engelholm LH, Behrendt N, Søe K, et al. Osteosarcoma and metastasis associated bone degradation-a tale of osteoclast and malignant cell cooperativity. *Int J Mol Sci* (2021) 22(13):6865. doi: 10.3390/ijms22136865

4. Yang C, Tian Y, Zhao F, Chen Z, Su P, Li Y, et al. Bone microenvironment and osteosarcoma metastasis. *Int J Mol Sci* (2020) 21(19):6985. doi: 10.3390/ijms21196985

5. Casali PG, Bielack S, Abecassis N, Aro HT, Bauer S, Biagini R, et al. Bone sarcomas: ESMO-PaedCan-EURACAN clinical practice guidelines for diagnosis, treatment and follow-up. annals of oncology : official journal of the European society for medical oncology. *Ann Oncol* (2018) 29(Suppl 4):iv79–95. doi: 10.1093/annonc/mdy310

6. Jing S, Ding F, Yuan Y, An J, He Q. Efficacy of neoadjuvant chemotherapy plus limb-sparing surgery for osteosarcoma and its impact on long-term quality of life. *Evid Based Complement Alternat Med* (2022) 2022:1693824. doi: 10.1155/2022/1693824

7. Belayneh R, Fourman MS, Bhogal S, Weiss KR. Update on osteosarcoma. *Curr Oncol Rep* (2021) 23(6):71. doi: 10.1007/s11912-021-01053-7

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

9. Mailankody S, Kumar VS, Khan SA, Banavali SD, Bajpai J. Resource-appropriate selection of osteosarcoma treatment protocols in low- and middle-income countries. *Pediatr Blood Cancer* (2022) 69(3):e29540. doi: 10.1002/pbc.29540

10. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res* (2013) 41(Database issue):D991–5. doi: 10.1093/nar/gks1193

11. Toro-Domínguez D, Martorell-Marugán J, López-Domínguez R, García-Moreno A, González-Rumayor V, Alarcón-Riquelme ME, et al. ImaGEO: integrative gene expression meta-analysis from GEO database. *Bioinformatics* (2019) 35(5):880–2. doi: 10.1093/bioinformatics/bty721

12. Wang Z, Lachmann A, Ma'ayan A. Mining data and metadata from the gene expression omnibus. *Biophys Rev* (2019) 11(1):103–10. doi: 10.1007/s12551-018-0490-8

13. Guo A, Wang W, Shi H, Wang J, Liu T. Identification of hub genes and pathways in a rat model of renal ischemia-reperfusion injury using bioinformatics analysis of the gene expression omnibus (GEO) dataset and integration of gene expression profiles. *Med Sci Monit* (2019) 25:8403–11. doi: 10.12659/MSM.920364

14. Ma G, Liu M, Du K, Zhong X, Gong S, Jiao L, et al. Differential expression of mRNAs in the brain tissues of patients with alzheimer's disease based on GEO expression profile and its clinical significance. *BioMed Res Int* (2019) 2019:8179145. doi: 10.1155/2019/8179145

15. Li W, Ding Z, Wang D, Li C, Pan Y, Zhao Y, et al. Ten-gene signature reveals the significance of clinical prognosis and immuno-correlation of osteosarcoma and study on novel skeleton inhibitors regarding MMP9. *Cancer Cell Int* (2021) 21 (1):377. doi: 10.1186/s12935-021-02041-4

16. Alfranca A, Martinez-Cruzado L, Tornin J, Abarrategi A, Amaral T, de Alava E, et al. Bone microenvironment signals in osteosarcoma development. *Cell Mol Life Sci* (2015) 72(16):3097–113. doi: 10.1007/s00018-015-1918-y

17. Shi D, Wu F, Mu S, Hu B, Zhong B, Gao F, et al. LncRNA AFAP1-AS1 promotes tumorigenesis and epithelial-mesenchymal transition of osteosarcoma through RhoC/ROCK1/p38MAPK/Twist1 signaling pathway. *J Exp Clin Cancer Res* (2019) 38(1):375. doi: 10.1186/s13046-019-1363-0

18. Lei T, Qian H, Lei P, Hu Y. Ferroptosis-related gene signature associates with immunity and predicts prognosis accurately in patients with osteosarcoma. *Cancer Sci* (2021) 112(11):4785–98. doi: 10.1111/cas.15131

19. Qian H, Lei T, Hu Y, Lei P. Expression of lipid-metabolism genes is correlated with immune microenvironment and predicts prognosis in osteosarcoma. *Front Cell Dev Biol* (2021) 9:673827. doi: 10.3389/fcell.2021.673827

20. Charfi C, Levros LC Jr, Edouard E, Rassart E. Characterization and identification of PARM-1 as a new potential oncogene. *Mol Cancer* (2013) 12:84. doi: 10.1186/1476-4598-12-84

21. Liu YR, Hu Y, Zeng Y, Li ZX, Zhang HB, Deng JL, et al. Neurexophilin and PC-esterase domain family member 4 (NXPE4) and prostate androgen-regulated mucin-like protein 1 (*PARM1*) as prognostic biomarkers for colorectal cancer. *J Cell Biochem* (2019) 120(10):18041–52. doi: 10.1002/jcb.29107

22. Fladeby C, Gupta SN, Barois N, Lorenzo PI, Simpson JC, Saatcioglu F, et al. Human PARM-1 is a novel mucin-like, androgen-regulated gene exhibiting proliferative effects in prostate cancer cells. *Int J Cancer* (2008) 122(6):1229–35. doi: 10.1002/ijc.23185

23. Isodono K, Takahashi T, Imoto H, Nakanishi N, Ogata T, Asada S, et al. PARM-1 is an endoplasmic reticulum molecule involved in endoplasmic reticulum stress-induced apoptosis in rat cardiac myocytes. *PloS One* (2010) 5(3):e9746. doi: 10.1371/journal.pone.0009746

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 17 March 2023
DOI 10.3389/fonc.2023.1068053

Check for updates

OPEN ACCESS

EDITED BY
Anton A. Buzdin,
European Organisation for Research and Treatment of Cancer, Belgium

REVIEWED BY
Yunfeng Feng,
Qinghai University Medical College, China
Ahmet Acar,
Middle East Technical University, Türkiye

*CORRESPONDENCE
Simon Walker-Samuel
simon.walkersamuel@ucl.ac.uk

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

RECEIVED 12 October 2022
ACCEPTED 09 February 2023
PUBLISHED 17 March 2023

CITATION
Gourmet LE and Walker-Samuel S (2023) The role of physics in multiomics and cancer evolution.
*Front. Oncol.* 13:1068053.
doi: 10.3389/fonc.2023.1068053

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

# The role of physics in multiomics and cancer evolution

Lucie E. Gourmet[1,2] and Simon Walker-Samuel[1,2]*

[1]Centre for Advanced Biomedical Imaging, Division of Medicine, University College London, London, United Kingdom, [2]Centre for Computational Medicine, Division of Medicine, University College London, London, United Kingdom

Complex interactions between the physical environment and phenotype of a tumour, and genomics, transcriptomics, proteomics and epigenomics, are increasingly known to have a significant influence on cancer development, progression and evolution. For example, mechanical stress can alter both genome maintenance and histone modifications, which consequently affect transcription and the epigenome. Increased stiffness has been linked to genetic heterogeneity and is responsible for heterochromatin accumulations. Stiffness thereby leads to deregulation in gene expression, disrupts the proteome and can impact angiogenesis. Several studies have shown how the physics of cancer can influence diverse cancer hallmarks such as resistance to cell death, angiogenesis and evasion from immune destruction. In this review, we will explain the role that physics of cancer plays in cancer evolution and explore how multiomics are being used to elucidate the mechanisms underpinning them.

KEYWORDS

**cancer, physics, evolution, multiomics, angiogenesis**

## Introduction

Cancer initially emerges from the accumulation of key mutations in somatic cells (1). The phenotypic changes that emerge from these mutations can provide a selective advantage over other cells that, through competition, can lead to clonal expansion (1). These processes are increasingly being studied and understood through an evolutionary lens (1). It has been established that tumours typically harbour a limited set of driver mutations that are positively selected through evolutionary pressures, alongside multiple passenger mutations that occur randomly (2). Likewise, deleterious (or disadvantageous) mutations are removed through negative selection (2). Cancer evolution has been shown to be dominated by positive selection and driver mutations have become an active area of cancer research (3). Meanwhile, negative selection affects cell survival and immune response (4). It is now becoming clear that a complex interplay exists between genetics and the physical properties of the tumour microenvironment, both of which can exert an influence in cancer evolution (5). This interplay can be seen as a two-way street: genetics and the physical environment are mechanistically coupled. This is reflected in the revised

hallmarks of cancer in which physical mechanisms have been acknowledged such as phenotypic plasticity (6).

Biomechanics plays a key role in the development of cancer, such as *via* interactions between the extracellular matrix (ECM), cell membranes, cytoskeleton and blood vessels. This broad set of microstructural components are referred to as the physical microenvironment, and plays a fundamental role in the propagation of mechanical stress and fluid dynamics, both of which can have a direct impact on tumour growth and progression (7, 8). Here, we aim to summarise current understandings of the interaction between these physical properties and tumour 'omics' (genomics, transcriptomics, proteomics and epigenomics), alongside how they are incorporated into the emerging field of cancer evolution. We discuss the impact of physical constraints on cell fate and the spatial distribution of tumour cells. We finally focus on angiogenesis, and describe how this hallmark of cancer provides a good example of how the physical microenvironment and cancer biology intertwine.

# Biomechanical properties of tumours

## Mechanical stress, tumour proliferation and invasion

When external forces are applied to the surface of an object (e.g. a cell membrane), stress is defined as the resulting internal resistance to deformation (9). For a cell, the effect of stress depends on several factors, such as the duration of the force, its magnitude and direction, and the biomechanical properties of the cell and its surroundings. Stress directed towards the outside of a tumour can lead to disruption of the surrounding stromal tissue and an increase in ECM tension (10). Interestingly, in response to ECM stiffness, this tensile stress can result in the contraction of cellular actomyosin which can affect cell motility (10). Tensile stress appears to promote cell division, according to experiments in which collagen incisions lead to the relaxation of ECM tension and resulted in decreased cell invasion (11). These experiments therefore demonstrated that collagen contraction in the ECM induced by cancer cells can contribute to cancer invasion. In contrast, compressive forces can reduce cancer cell proliferation by limiting volume expansion (12). These forces can cause a decrease in cell volume, followed by an increase in the production of $p27^{Kip1}$ which controls the cell cycle. As such, $p27^{Kip1}$ therefore inhibits proliferation but in a reversible manner: after releasing mechanical stress, the number of cells overexpressing p27Kip1 decreases. External compressive forces can also cause blood vessels to be compressed, which can limit the delivery of nutrients to tumour cells, including oxygen, and result in regional hypoxia. Hypoxia can have multiple outcomes, depending on its severity and duration, including limiting progression (13), or, conversely, driving metastasis by contributing to angiogenesis and epithelial-to-mesenchymal transition (13). As indicated by this example, the interaction between mechanical stress and tumour proliferation is highly complex and multifactorial (13).

The propensity of a tumour to change in shape when exposed to mechanical stress is given by its stiffness (or its reciprocal property, elasticity). Elevated tissue stiffness has been reported in multiple cancers, including breast and brain, and was shown to promote cancer cell invasion by affecting pathways involving IDH1, a known cancer driver gene (5). Raised ECM stiffness is thought to be caused by an increase in protein deposition by cancer-associated fibroblasts, which are created from normal fibroblasts in response to carcinogenesis. The exact origin of cancer fibroblasts is unclear, but they are thought to emerge from the expansion of fibroblasts at the periphery of the tumour. ECM stiffness is further promoted through collagen crosslinking by lysyl oxidase and parallel reorientation of collagen fibres (9). Increased tissue stiffening has been shown to affect multiple other factors, such as cellular differentiation and vessel permeability (5). Moreover, experiments using a collagen matrix with graded directional stiffness showed that stiffness guides cell migration and directs cancer cells towards intravasation sites where new blood vessels are formed, which implicates the involvement of mechanical stress in cancer metastasis (14).

## The role of cell morphology and microarchitecture

The shape of a cell is determined by several biomechanical processes, alongside genetic influences (15). Likewise, cell shape and biomechanics can also impact the structure and function of tissue on a range of length scales (Figure 1). Cell shape also influences several pathological changes, as it has been shown to directly regulate proliferation, differentiation, and survival (15). For example, the localisation of the transcription factor NF-κB was found to be sensitive to cell–cell contact and cell area (15). Moreover, high levels of NF-κB in the nucleus were associated with having few cell neighbours and displaying a mesenchymal-type shape.

Cell morphology was also found to be moderated by specific oncogenes (Ras, Akt and Mek) and can distinguish early and late cancers (16). Using machine learning analysis of images of actin cytoskeleton and cell outline, they found morphological signatures of cancerous transformation. Interestingly, most shape measures recorded were useful parameters for cell classification, with an accuracy ranging from 69.5% to 97.5%. Similar shape changes were associated with cancer progression across different cell types (osteoblastoma, breast cancer, retina). Therefore, this study not only shows that cellular morphology reflects the 'internal' state of a cell, but it also illustrates how this interplay can be exploited to better understand cancer evolution. Another study investigated the difference in morphological behaviour between metastatic and healthy cells (17). They developed a potential tool for early cancer detection which confirms the fact that features involving cell geometry and cellular protrusions help distinguish normal from cancer cells.

Mechanics has a further role in cellular organisation, for example in phase separation. Phase separation is the process responsible for cellular compartmentalization, involving the

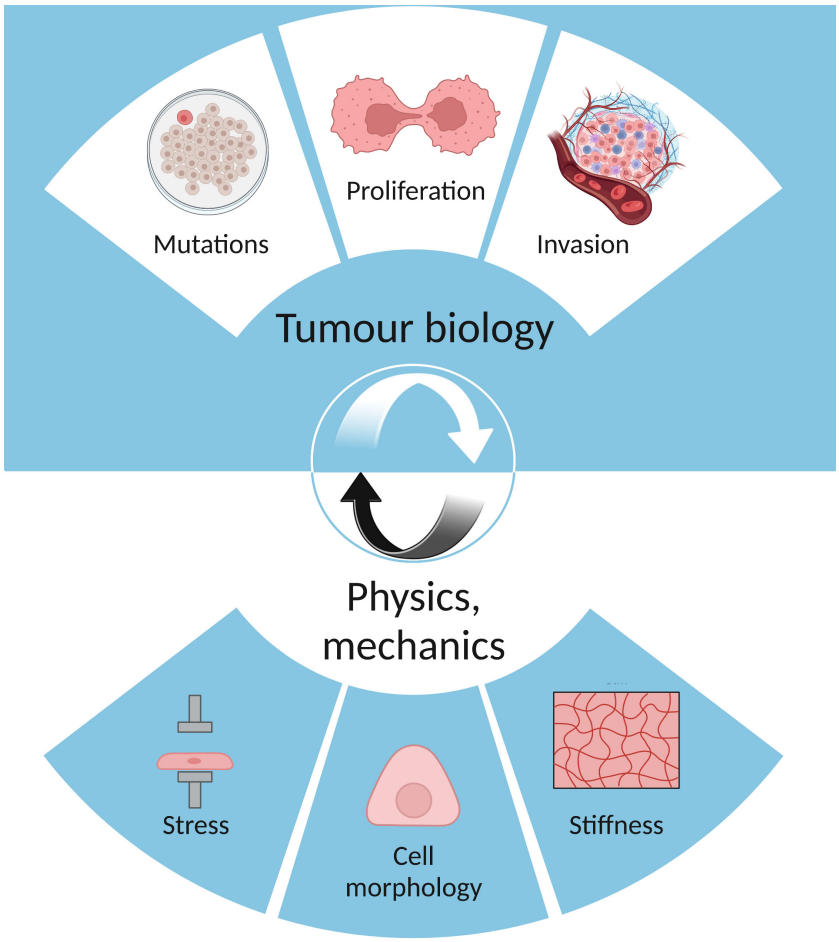



FIGURE 1
The interplay between tumour biology (mutations, proliferation, invasion) and physics (stress, cell morphology, stiffness).

assembly of condensates which are membraneless organelles such as stress granules and polycomb bodies (18). Dysregulated components in cancer can affect the location and the regulation of condensates, thereby leading to altered gene expression (19). Thus, this example demonstrates a physical process affecting the evolution of cancer and of multiomic factors.

## The influence of physical forces on cancer progression

### Nuclear deformation and genome integrity are closely linked

Cells are mechanosensitive which means that they can respond to applied forces *via* cell-cell/cell-ECM adhesions and ion channels that are sensitive to stress (20). However, few studies have investigated whether biomechanical properties directly influence mutational burden. Nuclear pore complexes modulate the import of transcription factors when nuclei are under solid stress and deformed (21). Nava et al. found that the genome is protected from mechanical stress, as nuclear deformation is counteracted *via* a calcium-dependent nuclear softening driven by loss of H3K9me3-marked heterochromatin (22). This chromatin mechanoprotection is significant as it prevents large-scale genetic changes, such as chromosomal breaks or rearrangements, and means that chromosome condensation can be modified to adapt to physical pressures. Less condensed chromosome regions enable gene expression, which positions physical stress as a factor that directly affects transcription. Thus, this protection mechanism is important for both genome integrity and gene expression. Moreover, increased ECM stiffness has been shown to induce DNA damage in mammary epithelial cells through reactive aldehyde species (23). They showed that DNA damage accumulation was due to a decrease in the clearance of these reactive aldehyde species and downregulation of aldehyde dehydrogenases, which is known to counteract oxidative stress. ECM stiffness was also reported to drive genomic heterogeneity in MYCN-amplified neuroblastoma cell lines as distribution of ECM tension is translated into genotypic variations *via* mechanotransduction, the ability of transforming mechanical signals into biological signals (23). This implies that the

mutational status of a tumour in part reflects the state of the tumour microenvironment, just as the microenvironment reflects mutational status. Thus, this further highlights how the interplay between the internal state of the tumour and the physical environment is bi-directional.

## Transcription is responsive to the physical environment

Fluid shear stress disrupts genome integrity and can also lead to changes in gene expression. Tajik et al. demonstrated that the stretching of chromatin *via* an external force result in transcription upregulation (24). The stress angle of the material used to apply pressure on the surface of a cell determines the extent of chromatin stretching and consequently affects the magnitude of DHFR transcription upregulation. Several models have been proposed to explain how mechanotransduction systems directly affect transcription, such as YAP (Yes-associated protein) and TAZ (transcriptional coactivator with PDZ-binding motif) translocating from the cytoplasm to the nucleus. Dupont et al. showed that YAP/TAZ transcriptional activity is regulated by ECM stiffness by growing mammary epithelial cells on cells on ECM of high versus low stiffness and analysing their gene expression (25). They also found that YAP/TAZ depends on cell geometry as the localisation of YAP/TAZ changed from predominantly nuclear in stiff ECM, to predominantly cytoplasmic in cells in soft ECM. Conversely, YAP/TAZ levels affect cell behaviours such as apoptosis (25). These findings concur with other research which found that YAP/TAZ respond to cell crowding and that cells under tensile stresses activated YAP/TAZ thereby leading to cell proliferation (26). Under stiff ECM conditions, the activation of YAP and TAZ transcription is dependent on actin and integrin, molecules that are associated with movement (27). Integrin also has a direct role in cancer gene expression as αvβ3 integrins were found to induce the transcription of the proapoptotic gene PUMA (p53-upregulated modulator of apoptosis) which promotes tumour stemness (28). Overall, these studies describe a direct link between transcriptomics, genomics and the physics of the tumour microenvironment.

## Epigenetic processes are regulated by fluid dynamics and stiffness

Epigenetics refers to heritable or environmental changes in gene expression which are not due to mutations. For example, DNA methylation, histone modifications and RNA mechanisms alter gene expression without modifying the DNA sequence itself. Haemodynamic forces including laminar and oscillatory flow, and cyclic strain have been shown to induce epigenetic modifications in vascular cells (29). For example, shear stress can modulate chromatin remodelling on histone H3 and H4 which appears to control the expression of endothelial nitric-oxide synthase (30). This finding shows that specific histone modifications have an important role in vascular gene expression, thereby affecting blood vessel formation in a pathological context such as cancer. Moreover, the endothelial-specific microRNA miR-92a was shown to promote fluid flow-stimulated angiogenesis but its expression is induced by the mechano-sensitive zinc finger transcription factor klf2a (31). This microRNA facilitates angiogenic sprouting and is dependent on mechanical factors. Epigenomic studies showed that ECM stiffness drives epigenetic changes with experiments using mechanosensitive breast cancer 3D culture model (32). Stowers et al. demonstrated that stiff matrices generated an increase in nuclear wrinkling compared to cells in soft matrices. In stiff matrices, heterochromatin condensed accumulations were found at the nuclear periphery and heterochromatin thickness increased compared to soft matrices, indicating that chromatin state was broadly misregulated. Their experiment showed the direct impact that stiffness has on chromosomal structure and gene expression. Interestingly, these chromatin changes were associated with a tumorigenic phenotype in the mammary epithelium which highlights the importance of the interplay between epigenomics and tumour mechanics (Figure 2).

# The consequences for tumour evolution

## Physical changes affect cell fate

The physical traits of cancer arise through conventional biological processes, but the reverse is also true; mutations, transcription, protein expression and epigenetics are each affected by the physical characteristics of the microenvironment. For example, ECM remodeling is due to the activation of matrix metalloproteinases, which are initially inert due to the interaction between the pro-domain and the catalytic site of these proteins (33). However, physical alterations in cancer can disrupt this interaction and enable migration and metastasis (33).

Epithelial-mesenchymal transition (EMT) is the process in which epithelial cells become mesenchymal cells through the loss of cell polarity and adhesion. This transdifferentiation has been implicated in cancer as mesenchymal cells are more motile, which enables invasion and metastasis. EMT is also affected by the physical traits of tumours, as ECM stiffness promotes EMT *via* a EPHA2/LYN complex and a TWIST1-G3BP2 pathway (34, 35). In addition, tissue geometry has been shown to have an impact on EMT, as changing the shape of the epithelial sheet altered the spatial pattern of EMT and EMT-permissive regions were found to experience the highest mechanical stress (36). Cancer stem cells, which have been linked to therapy resistance and cancer recurrence, are also subject to the mechanical forces of the tumour microenvironment. ECM stiffness increases significantly from the centre outwards, which has implications for the understanding of cancer stem cell development (37). Indeed, the most invasive and metastatic cancer stem cells were found to be mainly concentrated in the invasive outer edge of a tumour (38). The microenvironment of this area enables the clonal expansion of the cancer stem cells present (38). Furthermore, hydrodynamic shear stress was found to

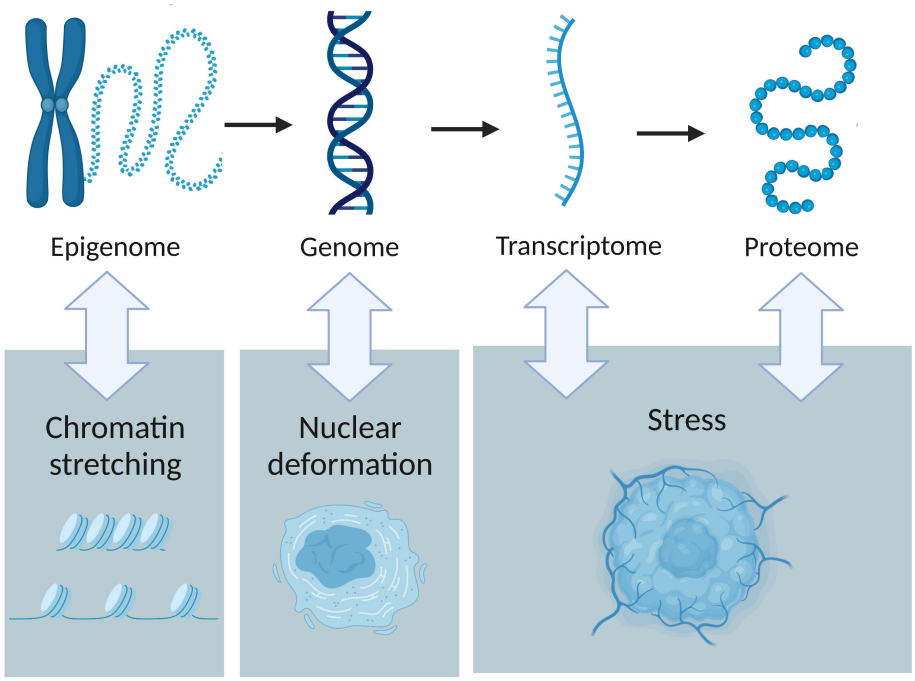



FIGURE 2
The interactions between tumour omics and the physical traits of cancer. Epigenetics is influenced by chromatin stretching, nuclear deformation leads to genome instability and stress affects both transcription and protein production.

promote the conversion of circulating tumour cells to cancer stem-like cells in the blood circulation *via* metabolic reprogramming (glycolysis and amino acid exchange) from the tumour niche (39). Thus, tumour mechanics specific to blood vessels provide an advantage to the circulating cells as they gain new oncogenic properties. This finding supports the fact that the location of a cancer cell matters because its environment can affect its behaviour.

## Spatial context matters for cancer progression

Many studies have used multi-omics approaches to study cancer progression but Srivastava et al. uncovered salient features of both imaging and omics data (40). They compared the performance of transcriptomics, proteomics and tissue images in predicting four domains: tumour stage, oestrogen receptor status, American Joint Committee on Cancer (AJCC) staging and PAM50 subtype. They found that histology images were best for predicting AJCC and tumour stage while transcriptomics data was better at characterising Estrogen Receptor (ER) status and PAM50 subtypes. The ability to compare imaging data with multi-omics data is important because it shows that some methods are more relevant based on the aim of a study. This type of study is interesting, but it requires having both imaging and sequencing data from the same patient, which is rarely the case.

The architecture of normal tissue has been shown to influence the fitness of cancer cells and determine the mode of evolution as a result of competition for space (41). In fact, this study showed that spatial limitations and cell mixing rates are the main factors influencing the acquisition of cancer driving mutations. Thus, the outcome of a selective advantage at a cellular level depends on the environmental competitive context, which highlights the importance of spatial constraints. The different modes of tumour growth (surface or volume) were found to impact the extent and the distribution of clonal diversity over time (42). Interestingly, this study reveals that tumour roundness may be representative of clonal diversity: when the tumour from their cellular automaton model (representing clear cell renal cell carcinoma) grows, it eventually loses its round shape but recovers it mirroring the increase and decrease of clonal diversity during tumour evolution. Noble et al. also showed that spatial structure governs the mode of tumour evolution by analysing four types of tissue structure (non-spatial, gland fission, invasive glandular, and boundary growth) (43). The different tumour architectures influenced the distribution of mutation frequency and the dynamics of clonal diversity. For example, hepatocellular carcinoma (whose architecture is described as boundary growth, where proliferation is confined to the boundary) was found to promote genetic drift and was subject to a mutation burden increasing from the tumour core to its boundary. For invasive glandular tumours (which represent most solid tumours), small increases of cell fitness are sufficient to trigger a clonal expansion. The evolution of invasive glandular tumours can be described as a branching process, where mutations create multiple new clones without necessarily removing old clones. The conclusion from these findings is that alterations of tumour architecture during cancer do impact the mode of tumour evolution.

Because physical and spatial characteristics of tumours affect cancer evolution, they are interesting in the context of multi-omics. To enrich our current understanding of cancer, we can investigate the interplay of different cancer hallmarks, such as angiogenesis and genome instability. In this example, one could wonder whether blood vessels impact the mutational diversity of a tumour. Angiogenesis provides nutrients to cancer cells and enables the circulation of cancer cells which makes it worth investigating.

## The interplay between angiogenesis and tumour evolution

The interaction between blood vessels, tumour biomechanics and genomics is highly complex, which makes it challenging to investigate. However, angiogenesis offers a useful system to highlight the complex relationship between genetics and physical microenvironment.

Blood vessel formation in cancer occurs in response to the need for nutrients and oxygen to sustain homeostasis and growth. Sherwood et al. found that rapidly growing tumours are dependent on angiogenesis whereas dormant tumours are not heavily vascularised (44). They also showed that inhibition of angiogenic signalling pathways prevents vessel formation and leads to tumour dormancy (45). The first step in tumour vascularisation is the angiogenic switch, during which pro-angiogenic signalling such as VEGF becomes dominant over anti-angiogenic factors (46). Blood formation can be initiated in several ways, including sprouting angiogenesis, intussusceptive angiogenesis, vasculogenesis, endothelial progenitor cells, vasculogenic mimicry and transdifferentiation of cancer cells (47). These mechanisms result in very different vessel morphologies, with differing flow and delivery characteristics, and can therefore provide a range of environments for cancers to evolve within. Certain cancer types have corresponding characteristic vascular phenotypes: intussusceptive angiogenesis has been observed in melanoma and adenocarcinoma, while vasculogenesis has been reported in gliomas (47). For example, vascular mimicry has been linked to aggressive tumours and reported following anti angiogenic therapy, making it an alternative neovascularisation process used to survive treatment. It also has been associated with poor prognosis in colorectal cancers and glioblastomas (47). The different types of angiogenesis are important regarding treatment options as it was shown that inhibition of vasculogenesis prevents glioma recurrence, not sprouting angiogenesis (48).

### Blood vessels impact tumour heterogeneity

Computational simulations of the microenvironment of tumours showed that small tumours are spatially heterogeneous when vascularised (49). This finding implies that blood vessels lead to high evolutionary pressure and selection of different clones depending on the microenvironment even in an early cancer stage, when the tumour is still small. Moreover, Sartakhti et al. used an economic model to simulate angiogenesis, where the secretion of pro-angiogenic factors were considered to be public goods that can be exploited by free-riders (i.e. cells that cease to contribute to angiogenesis formation) (50). In this approach, a clone's fitness is determined by collective interactions and a critical mass of cells cooperating is necessary to sustain angiogenesis. Their results revealed that spatial heterogeneity may be a main driver underlying the emergence of angiogenic clones and showed that sprouting of new blood vessels from surrounding vascular tissue is directed towards the centre of the tumour. This is in accordance with studies which already reported this pattern in Lewis lung carcinoma and mammary carcinoma (51, 52). The architecture and morphology of angiogenesis is therefore critical in cancer evolution and has important implications for cancer treatments (Figure 3).

### Timing is crucial for angiogenic morphology

In addition to spatial heterogeneity, blood vessels are also dependent on time-based properties as Bentley and Chakravartula highlighted that endothelial cell rearrangements are regulated by the speed of selection (53). Vessel sprouts are generated in response to hypoxia, with leading tip cells inhibiting stalk cells and being selected to lead sprout growth by a central pattern generating (CPG) mechanism involving Dll4-Notch lateral inhibition in a feedback loop with VEGF – VEGFR signalling (53). Slowing down the CPG mechanism, meaning that cells take longer to decide on their movement, leads to a sparser branching phenotype. Synchronised selection switches vessels from branching to expansion and accelerating selection preserves sprout extension in proliferative tissues. Thus, tissue-derived factors such as VEGF, and active perception *via* filopodia extensions can vary the speed and timing of cell migration. Temporal variability has many implications, including choosing a convenient targeting time for therapeutic normalisation. The environmental conditions are important as well since they influence the eco-evolutionary causes of temporal changes in angiogenesis (54). In this review, angiogenesis is assimilated to an ecoevolutionary process called niche construction and is compared to organisms improving their environment, for example dams constructed by beavers. Temporal fluctuations are thought to either promote a plastic phenotype or select for speciation. In the former case, the cancer cell is viewed as a jack of all trades which can adapt to different microenvironments or treatments easily. On the other hand, cancer cells can specialise through foraging and metabolism. These adaptive strategies are subject to trade-offs which are common in nature, which is why eco-evolutionary dynamics are studied to get a better understanding of cancer.

## Conclusion

Genomics is fundamental to the understanding of cancer evolution, but tumours are also subject to mechanical, compressive, tensile and shear stresses that can play a significant role. For example, raised interstitial fluid pressure can promote

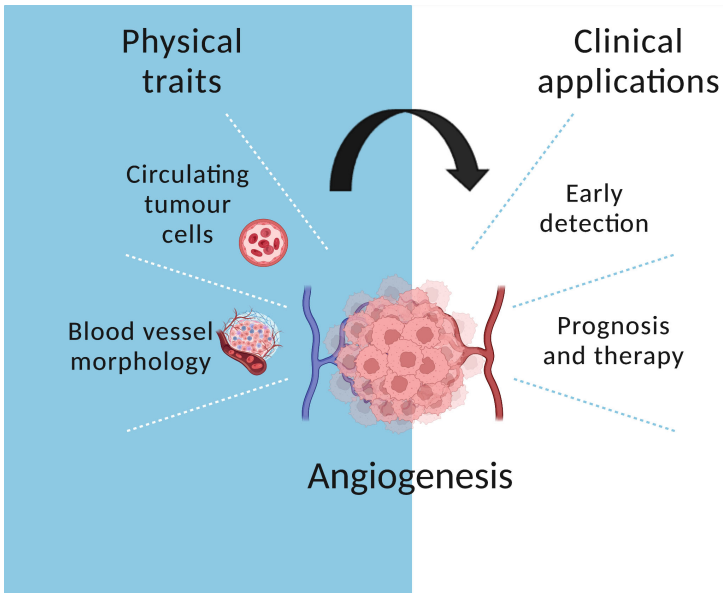



FIGURE 3
The role of angiogenesis in cancer evolution. Circulating tumour cells are released in the bloodstream which enables early detection. In addition, blood vessel morphology informs prognosis and treatment.

cancer cell invasion in the surrounding tissue, and matrix stiffness differs between the centre and the border of the tumour, which helps metastasis by guiding cell migration. Cell geometry can inform on cell function, indicate whether specific mutations are present and predict malignancy. These mechanical characteristics form a complex dynamic system, with interactions between physical features, genetic mutation and expression, epigenetics and protein transcription. Stiffness and shear stress have been shown to lead to DNA damage. The transcription of YAP and TAZ is dependent on tensile stress, and chromatin stretching was shown to upregulate transcription. In addition to the structure of chromatin, epigenetic modifications are essential as shear stress was found to promote angiogenesis by affecting histones.

Computational models and simulations help us predict the impact of tissue architecture, but quantitative imaging analysis is needed to confirm their hypotheses. The physical and omic interactions matter for cancer evolution because they alter cell behaviour, facilitating the development of stem cells and mesenchymal cells. It also helps the circulation of tumour cells by increasing tumour shedding and enhancing angiogenesis. Angiogenesis plays a pivotal role as the mechanism underlying blood vessel formation leads to diverse morphologies and is subject to temporal variations influencing cancer prognosis and therapy. We can view cancer as an ecological system and use strategies found in nature to model cancer evolution. Investigating the impact of angiogenesis on other cancer hallmarks such as mutations and immune evasion will provide a deeper understanding of this disease. Moreover, studying angiogenesis is useful to elucidate the interplay between physics and biology. To improve our current detection methods, we need to continue interdisciplinary research to understand how these diverse aspects of cancer fit together.

## Author contributions

LG conducted the literature review and SW-S provided feedback, reviewed the article. All authors contributed to the article and approved the submitted version.

## Funding

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) (EP/W007096/1); and Cancer Research UK (C44767/A29458 and C23017/A27935).

## Acknowledgments

I would like to thank the EPSRC for helping with the publication fees, the international Alliance for Cancer Early Detection for supporting my PhD and biorender (used to create the figures). Moreover, I would like to acknowledge my family, my friends and CABI lab members for their continuous support.

## Conflict of interest

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Nowell PC. The clonal evolution of tumor cell populations. *Science* (1979) 194 (4260):23–8. doi: 10.1126/science.959840

2. Iranzo J, Martincorena I, Koonin EV. Cancer-mutation network and the number and specificity of driver mutations. *Proc Natl Acad Sci USA*. (2018) 115(26):E6010–9. doi: 10.1073/pnas.1803155115.

3. Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, van Loo P, et al. Universal patterns of selection in cancer and somatic tissues. *Cell* (2017) 171(5):1029–1041.e21. doi: 10.1016/j.cell.2017.09.042

4. Zapata L, Pich O, Serrano L, Kondrashov FA, Ossowski S, Schaefer MH. Negative selection in tumor genome evolution acts on essential cellular functions and the immunopeptidome. *Genome Biol* (2018) 19(1):67. doi: 10.1186/s13059-018-1434-0

5. Nia HT, Munn LL, Jain RK. Physical traits of cancer. *Science* (2020) 30:370(6516). doi: 10.1126/science.aaz0868

6. Hanahan D. Hallmarks of cancer: New dimensions. *Cancer Discovery* (2022) 12 (1):31–46. doi: 10.1158/2159-8290.CD-21-1059

7. Helmlinger G, Netti PA, Lichtenbeld HC, Melder RJ, Jain RK. Solid stress inhibits the growth of multicellular tumor spheroids. *Nat Biotechnol* (1997) 15(8):778–83. doi: 10.1038/nbt0897-778

8. Chauhan VP, Boucher Y, Ferrone CR, Roberge S, Martin JD, Stylianopoulos T, et al. Compression of pancreatic tumor blood vessels by hyaluronan is caused by solid stress and not interstitial fluid pressure. *Cancer Cell* (2014) 26(1):14–5. doi: 10.1016/j.ccr.2014.06.003

9. Northcott JM, Dean IS, Mouw JK, Weaver VM. Feeling stress: The mechanics of cancer progression and aggression. *Front Cell Dev Biol* (2018) 6. doi: 10.3389/fcell.2018.00017

10. Jain RK, Martin JD, Stylianopoulos T. The role of mechanical forces in tumor growth and therapy. *Annu Rev BioMed Eng* (2014) 16(1):321–46. doi: 10.1146/annurev-bioeng-071813-105259

11. Kopanska KS, Alcheikh Y, Staneva R, Vignjevic D, Betz T. Tensile forces originating from cancer spheroids facilitate tumor invasion. *PloS One* (2016) 11(6): e0156442. doi: 10.1371/journal.pone.0156442

12. Delarue M, Montel F, Vignjevic D, Prost J, Joanny JF, Cappello G. Compressive stress inhibits proliferation in tumor spheroids through a volume limitation. *Biophys J* (2014) 107(8):1821–8. doi: 10.1016/j.bpj.2014.08.031

13. Muz B, de la Puente P, Azab F, Azab AK. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia* (2015) 3:83–92. doi: 10.2147/HP.S93413

14. Hadjipanayi E, Mudera V, Brown RA. Guiding cell migration in 3D: A collagen matrix with graded directional stiffness. *Cell Motil Cytoskeleton* (2009) 66(3):121–8. doi: 10.1002/cm.20331

15. Sero JE, Sailem HZ, Ardy RC, Almuttaqi H, Zhang T, Bakal C. Cell shape and the microenvironment regulate nuclear translocation of NF -κB in breast epithelial and tumor cells. *Mol Syst Biol* (2015) 11(3):790. doi: 10.15252/msb.20145644

16. Alizadeh E, Castle J, Quirk A, Taylor CDL, Xu W, Prasad A. Cellular morphological features are predictive markers of cancer cell state. *Comput Biol Med* (2020) 126:104044. doi: 10.1016/j.compbiomed.2020.104044

17. Hasan MR, Hassan N, Khan R, Kim YT, Iqbal SM. Classification of cancer cells using computational analysis of dynamic morphology. *Comput Methods Programs Biomed* (2018) 156:105–12. doi: 10.1016/j.cmpb.2017.12.003

18. Peng Q, Tan S, Xia L, Wu N, Oyang L, Tang Y, et al. Phase separation in cancer: From the impacts and mechanisms to treatment potentials. *Int J Biol Sci* (2022) 18 (13):5103–22. doi: 10.7150/ijbs.75410

19. Boija A, Klein IA, Young RA. Biomolecular condensates and cancer. *Cancer Cell* (2021) 39(2):174–92. doi: 10.1016/j.ccell.2020.12.003

20. Kirby TJ, Lammerding J. Emerging views of the nucleus as a cellular mechanosensor. *Nat Cell Biol* (2018) 20(4):373–81. doi: 10.1038/s41556-018-0038-y

21. Nava MM, Miroshnikova YA, Biggs LC, Whitefield DB, Metge F, Boucas J, et al. Heterochromatin-driven nuclear softening protects the genome against mechanical stress-induced damage. *Cell* (2020) 181(4):800–17. doi: 10.1016/j.cell.2020.03.052

22. Wood A, Sun H, Jones M, Percival H, Broadberry E, Zindy E, et al. Increased microenvironment stiffness leads to altered aldehyde metabolism and DNA damage in mammary epithelial cells through a RhoA-dependent mechanism. *bioRvix* (2020). doi: 10.1101/2020.10.06.327726

23. López-Carrasco A, Martín-Vañó S, Burgos-Panadero R, Monferrer E, Berbegall AP, Fernández-Blanco B, et al. Impact of extracellular matrix stiffness on genomic heterogeneity in MYCN-amplified neuroblastoma cell line. *J Exp Clin Cancer Res* (2020) 39(1):226. doi: 10.1186/s13046-020-01729-1

24. Tajik A, Zhang Y, Wei F, Sun J, Jia Q, Zhou W, et al. Transcription upregulation *via* force-induced direct stretching of chromatin. *Nat Mater* (2016) 15(12):1287–96. doi: 10.1038/nmat4729

25. Dupont S, Morsut L, Aragona M, Enzo E, Giulitti S, Cordenonsi M, et al. Role of YAP/TAZ in mechanotransduction. *Nature* (2011) 474(7350):179–83. doi: 10.1038/nature10137

26. Aragona M, Panciera T, Manfrin A, Giulitti S, Michielin F, Elvassore N, et al. A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. *Cell* (2013) 154(5):1047–59. doi: 10.1016/j.cell.2013.07.042

27. Mohammadi H, Sahai E. Mechanisms and impact of altered tumour mechanics. *Nat Cell Biol* (2018) 20(7):766–74. doi: 10.1038/s41556-018-0131-2

28. Sun Q, Lesperance J, Wettersten H, Luterstein E, DeRose YS, Welm A, et al. Proapoptotic PUMA targets stem-like breast cancer cells to suppress metastasis. *J Clin Invest* (2017) 128(1):531–44. doi: 10.1172/JCI93707

29. Chen L, Wei S, Chiu J. Mechanical regulation of epigenetics in vascular biology and pathobiology. *J Cell Mol Med* (2013) 17(4):437–48. doi: 10.1111/jcmm.12031

30. Fish JE, Matouk CC, Rachlis A, Lin S, Tai SC, D'Abreo C, et al. The expression of endothelial nitric-oxide synthase is controlled by a cell-specific histone code. *J Biol Chem* (2005) 280(26):24824–38. doi: 10.1074/jbc.M502115200

31. Nicoli S, Standley C, Walker P, Hurlstone A, Fogarty KE, Lawson ND. MicroRNA-mediated integration of haemodynamics and vegf signalling during angiogenesis. *Nature* (2010) 464(7292):1196–200. doi: 10.1038/nature08889

32. Stowers RS, Shcherbina A, Israeli J, Gruber JJ, Chang J, Nam S, et al. Matrix stiffness induces a tumorigenic phenotype in mammary epithelium through changes in chromatin accessibility. *Nat BioMed Eng* (2019) 3(12):1009–19. doi: 10.1038/s41551-019-0420-5

33. Popova NV, Jücker M. The functional role of extracellular matrix proteins in cancer. *Cancers (Basel)* (2022) 14(1):238. doi: 10.3390/cancers14010238

34. Fattet L, Jung HY, Matsumoto MW, Aubol BE, Kumar A, Adams JA, et al. Matrix rigidity controls epithelial-mesenchymal plasticity and tumor metastasis *via* a mechanoresponsive EPHA2/LYN complex. *Dev Cell* (2020) 54(3):302–316.e7. doi: 10.1016/j.devcel.2020.05.031

35. Wei SC, Fattet L, Tsai JH, Guo Y, Pai VH, Majeski HE, et al. Matrix stiffness drives epithelial–mesenchymal transition and tumour metastasis through a TWIST1–G3BP2 mechanotransduction pathway. *Nat Cell Biol* (2015) 17(5):678–88. doi: 10.1038/ncb3157

36. Gomez EW, Chen QK, Gjorevski N, Nelson CM. Tissue geometry patterns epithelial-mesenchymal transition via intercellular mechanotransduction. *J Cell Biochem* (2010) 110(1):44–51. doi: 10.1002/jcb.22545

37. Masuzaki R, Tateishi R, Yoshida H, Sato T, Ohki T, Goto T, et al. Assessing liver tumor stiffness by transient elastography. *Hepatol Int* (2007) 1(3):394–7. doi: 10.1007/s12072-007-9012-7

38. Zheng YW, Tsuchida T, Shimao T, Li B, Takebe T, Zhang RR, et al. The CD133 $^{+}$ CD44 $^{+}$ precancerous subpopulation of oval cells is a therapeutic target for hepatocellular carcinoma. *Stem Cells Dev* (2014) 23(18):2237–49. doi: 10.1089/scd.2013.0577

39. Bertero T, Gaggioli C. Mechanical forces rewire metabolism in the tumor niche. *Mol Cell Oncol* (2019) 6(3):1592945. doi: 10.1080/23723556.2019.1592945

40. Srivastava A, Kulkarni C, Mallick P, Huang K, Machiraju R. Building trans-omics evidence: using imaging and "omics" to characterize cancer profiles. *Pac Symp Biocomput* (2018) 23:377–87. doi: 10.1142/9789813235533_0035

41. West J, Schenck RO, Gatenbee C, Robertson-Tessi M, Anderson ARA. Normal tissue architecture determines the evolutionary course of cancer. *Nat Commun* (2021) 12(1):2060. doi: 10.1038/s41467-021-22123-1

42. Fu X, Zhao Y, Lopez JI, Rowan A, Au L, Fendler A, et al. Spatial patterns of tumour growth impact clonal diversification in a computational model and the TRACERx renal study. *Nat Ecol Evol* (2022) 6(1):88–102. doi: 10.1038/s41559-021-01586-x

43. Noble R, Burri D, le Sueur C, Lemant J, Viossat Y, Kather JN, et al. Spatial structure governs the mode of tumour evolution. *Nat Ecol Evol* (2022) 6(2):207–17. doi: 10.1038/s41559-021-01615-9

44. Sherwood LM, Parris EE, Folkman J. Tumor angiogenesis: Therapeutic implications. *New Engl J Med* (1971) 285(21):1182–6. doi: 10.1056/NEJM197111182852108

45. Folkman J, Merler E, Abernathy C, Williams G. ISOLATION OF a TUMOR FACTOR RESPONSIBLE FOR ANGIOGENESIS. *J Exp Med* (1971) 133(2):275–88. doi: 10.1084/jem.133.2.275

46. Hanahan D, Folkman J. Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell* (1996) 86(3):353–64. doi: 10.1016/S0092-8674(00)80108-7

47. Lugano R, Ramachandran M, Dimberg A. Tumor angiogenesis: causes, consequences, challenges and opportunities. *Cell Mol Life Sci* (2020) 77(9):1745–70. doi: 10.1007/s00018-019-03351-7

48. Kioi M, Vogel H, Schultz G, Hoffman RM, Harsh GR, Brown JM. Inhibition of vasculogenesis, but not angiogenesis, prevents the recurrence of glioblastoma after irradiation in mice. *J Clin Invest* (2010) 120(3):694–705. doi: 10.1172/JCI40283

49. Fredrich T, Rieger H, Chignola R, Milotti E. Fine-grained simulations of the microenvironment of vascularized tumours. *Sci Rep* (2019) 9(1):11698. doi: 10.1038/s41598-019-48252-8

50. Salimi Sartakhti J, Manshaei MH, Basanta D, Sadeghi M. Evolutionary emergence of angiogenesis in avascular tumors using a spatial public goods game. *PloS One* (2017) 12(4):e0175063. doi: 10.1371/journal.pone.0175063

51. Morikawa S, Baluk P, Kaidoh T, Haskell A, Jain RK, McDonald DM. Abnormalities in pericytes on blood vessels and endothelial sprouts in tumors. *Am J Pathol* (2002) 160(3):985–1000. doi: 10.1016/S0002-9440(10)64920-6

52. Goel S, Duda DG, Xu L, Munn LL, Boucher Y, Fukumura D, et al. Normalization of the vasculature for treatment of cancer and other diseases. *Physiol Rev* (2011) 91 (3):1071–121. doi: 10.1152/physrev.00038.2010

53. Bentley K, Chakravartula S. The temporal basis of angiogenesis. *Philos Trans R Soc B: Biol Sci* (2017) 372(1720):20150522. doi: 10.1098/rstb.2015.0522

54. Gillies RJ, Brown JS, Anderson ARA, Gatenby RA. Eco-evolutionary causes and consequences of temporal changes in intratumoural blood flow. *Nat Rev Cancer* (2018) 18(9):576–85. doi: 10.1038/s41568-018-0030-7

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 23 March 2023
DOI 10.3389/fonc.2023.916568

Check for updates

OPEN ACCESS

EDITED BY
Xinmin Li,
University of California, Los Angeles,
United States

REVIEWED BY
Jianyu Rao,
University of California, Los Angeles,
United States
Weikuan Gu,
University of Tennessee Health Science
Center (UTHSC), United States

*CORRESPONDENCE
Guanghui Zhong
zgh20040712@126.com
Xin Peng
pengx@nit.zju.edu.cn
Chunlin Zhang
qdzcl2011@163.com

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

RECEIVED 09 April 2022
ACCEPTED 09 February 2023
PUBLISHED 23 March 2023

CITATION
Liu Z, Zhao X, Wang R, Tang X, Zhao Y,
Zhong G, Peng X and Zhang C (2023)
Heterogeneous pattern of gene expression
driven by *TTN* mutation is involved in the
construction of a prognosis model of lung
squamous cell carcinoma.
*Front. Oncol.* 13:916568.
doi: 10.3389/fonc.2023.916568

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

# Heterogeneous pattern of gene expression driven by *TTN* mutation is involved in the construction of a prognosis model of lung squamous cell carcinoma

Zhao Liu[1,2,3], Xiaowen Zhao[4], Ruihong Wang[4], Xiangyue Tang[2,3], Yuxiang Zhao[1,2,3], Guanghui Zhong[1*], Xin Peng[1*] and Chunlin Zhang[4*]

[1]Ningbo Municipal Hospital of Traditional Chinese Medicine (TCM), Affiliated Hospital of Zhejiang Chinese Medical University, Ningbo, China, [2]United New Drug Research and Development Center, Biotrans Technology Co., LTD., Ningbo, China, [3]Institute of Bioengineering, Biotrans Technology Co., LTD., Shanghai, China, [4]Qingdao Central Hospital, The Second Affiliated Hospital of Medical College of Qingdao University, Qingdao, China

**Objective:** To investigate the impact that *TTN* mutation had on the gene heterogeneity expression and prognosis in patients with lung adenocarcinoma.

**Methods:** In this study, the Cancer Genome Atlas (TCGA) dataset was used to analyze the *TTN* mutations in lung adenocarcinoma. Lung adenocarcinoma data was collected from the TCGA database, clinical information of patients was analyzed, and bioinformatics statistical methods were applied for mutation analysis and prognosis survival analysis. The results were verified using the GEO dataset.

**Results:** The incidence of *TTN* mutations in lung adenocarcinoma was found to be 73%, and it was related to the prognosis of lung adenocarcinoma. Ten genes were screened with significant contributions to prognosis. A prognosis model was constructed and verified by LASSO COX analysis in the TCGA and GEO datasets based on these ten beneficial factors. The independent prognostic factor H2BC9 for *TTN* mutation-driven gene heterogeneity expression was screened through multi-factor COX regression analysis.

**Conclusion:** Our data showed that the gene heterogeneity expression, which was driven by *TTN* mutations, prolonged the survival of lung adenocarcinoma patients and provided valuable clues for the prognosis of *TTN* gene mutations in lung adenocarcinoma.

KEYWORDS

**bioinformatics, lung squamous carcinoma, TTN, gene mutation, prognostic survival**

# Introduction

Lung cancer remains the leading type of cancer-related death, and gene mutations have a very important impact on prognosis in NSCLC (1). TTN and TP53 are two of the most commonly mutated genes in NSCLC samples from The Cancer Genome Atlas (TCGA) official website. TP53 plays a crucial biological role in the development of various tumors and has received widespread attention (2–4). Co-mutation analyses of TP53 and cancer driver genes in clinical studies can often be used to provide specific molecular typing for molecularly targeted therapies (5). The coding gene TTN, also known as cardiomyopathy dilated 1G, is the longest-reported gene to date (6). TTN has a high mutation rate in several types of tumor tissue (7). However, TTN is controversial as it is considered a tumor-associated gene. This controversy focuses on the large and complex DNA structure and the false-positive results associated with tumor heterogeneity. In addition, the biological function of TTN mutations in the development of cancer has been questioned, as TTN mutations have been strongly associated with cardiac and skeletal muscle diseases but rarely with oncological diseases (6, 8, 9). To date, only a few studies have specifically explored the relevance of TTN to tumors. The complexity of the TTN gene structure compared to TP53 and the aforementioned controversies regarding TTN have greatly hindered researchers from analyzing this gene in depth. To date, studies related to TTN mutations in tumors are rare.

Using the TCGA database, this paper focuses on the correlation between TTN mutations and the clinicopathological characteristics of patients with LUSC and the impact of these mutations on the survival of these patients. It also focuses on the impact of TTN co-mutations on the prognosis survival of these patients to identify the main prognostic factors that affect TTN mutations in LUSC and thus provide new possible treatment targets for the disease. The main prognostic factors affecting TTN mutations in LUSC are SLC34A2, BARX1, and HOXC10. In this study, we elucidated the prognostic impact of heterogeneous gene expression caused by TTN mutations on LUSC and screened the main prognostic factors of heterogeneous gene expression driven by TTN mutations using the least absolute shrinkage and selection operator (LASSO) Cox analysis. The results could provide new possible targets for the prevention and treatment of LUSC.

# Methods and materials

## Data sources

RNA sequencing data (level 3), expression profile data, and corresponding clinical information were obtained from a TCGA dataset (https://portal.gdc.com) for 488 patients with LUSC. The GSE73403 datasets was from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo).

## Gene mutation analysis

The data of 488 patients with LUSC got from TCGA were analyzed, and somatic mutations in patients with LUSC were visualized using the maftools package in R software.

## Clinical correlation analysis

A univariate analysis of the clinicopathological characteristics was performed using the chi-square test and nonparametric tests. This was followed by screening for statistically significant independent influences associated with *TTN* mutations using binary logistic regression models. P values < 0.05 were considered significant.

## Survival analysis

RNAseq and corresponding clinical information of 393 *TTN* mutations in LUSC were came from TCGA dataset (https://portal.gdc.com). Progression-free survival (PFS), disease-free survival (DFS), and disease-specific survival (DSS) were assessed by overall survival (OS). The impact of *TTN* mutation on the prognostic survival of LUSC was assessed. Survival analyzes were performed using Cox multiple survival regression model analysis as the primary method, and the Kaplan–Meier (KM) method as an adjunct method. The endpoint of OS was defined as the time from randomization to death from any cause. The endpoint of DFS was defined as the time from the start of the initial treatment to the onset of a new tumor-related event or death from any cause. All the above analysis methods were implemented with the R package. P values < 0.05 were considered significant.

## Differential expression analysis

The DESeq2 package of R software was used for expression profiling, differential analysis, gene volcano mapping, and heat mapping. For gene microarray screening, corrected p-values and $|\log1.5FC|$ $p < 0.05$ and $|\log1.5FC| \geq 1.0$ were calculated for differentially expressed genes. Gene ontology (GO) functional enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis enrichment analysis were visualized using the ClusterProfiler package in R software, and a false discovery rate of < 0.01 was considered significant.

## Construction of prognostic signature models

The relationship between prognostic immune-related gene expression and OS was analyzed. A prognostic risk prediction model for LUSC was calculated. Patients with LUSC were divided into high-risk

and low-risk groups using the median risk score as the cut-off. KM curves were plotted to compare these groups. Receiver operating characteristic (ROC) survival analysis was performed using the R package SURVIVAL, and decision curve analysis was performed using the rmda package. The association between risk score models and tumor immune-infiltrating cells was also investigated using Spearman correlation analysis, and statistical significance was set to $P < 0.05$.

# Results

## Mutation landscape in LUSC tissues

To make sure the frequency of mutations in genes in LUSC, we imported data from the TCGA database for analysis. A total of 488 patients with LUSC were imported, containing 478 samples with detected mutations, of which the mapping samples contained a total of 478 (97.75%). The results showed that the *TTN* mutation rate was 73.00% (Figures 1A, B).

## Mutations in the *TTN* gene lead to differences in clinical characteristics of LUSC patients

Subsequently, our univariate analysis of *TTN* gene mutations and clinicopathological characteristics showed that survival status, pN stage, smoking history, and history of neoadjuvant treatment were associated with *TTN* mutations, while gender, age, pathological stage, new tumor event type, and radiation/chemotherapy were not associated with *TTN* mutations in LUSC (Table 1).

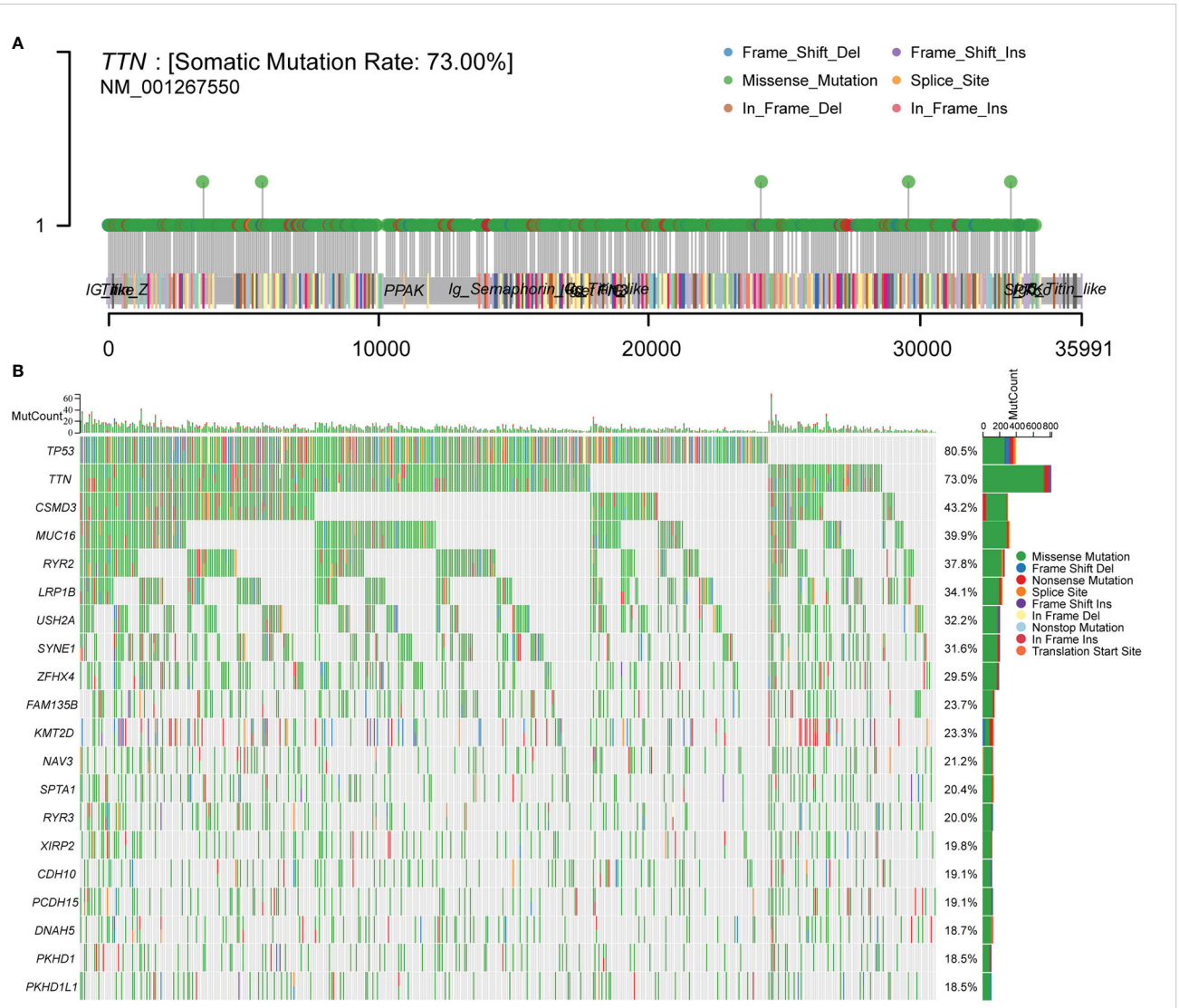



FIGURE 1
Downloading and visualizing somatic mutations in patients with LUSC using the maftools package in R software for TCGA data. **(A)** Lollipop plot of *TTN* gene mutation distribution with 73% mutation rate in somatic cells. **(B)** Somatic cell landscape of LUSC patient cohort with genes sorted by mutation frequency and samples sorted by disease histology; sidebar plot showing $\log_{10}$ transformed *Q*-values estimated by MutSigCV. The waterfall plot shows mutation information for each gene in each sample.

10.3389/fonc.2023.916568

TABLE 1 Correlation analysis between *TTN* mutation and clinicopathological characteristics of LUSC.

| | Character | *TTN* mutant | *TTN* WT | *P*-value |
|---|---|---|---|---|
| Survival status | Alive | 233 | 45 | |
| | Dead | 160 | 51 | 0.037 |
| Age | Mean (SD) | 66.9 (8.5) | 68.1 (9.2) | |
| | Median [MIN, MAX] | 68 [39, 84] | 69 [41, 90] | 0.243 |
| Gender | Female | 96 | 31 | |
| | | 297 | 65 | 0.148 |
| pT stage | T1 | 40 | 8 | |
| | T1a | 19 | 5 | |
| | T1b | 35 | 5 | |
| | T2 | 128 | 39 | |
| | T2a | 72 | 14 | |
| | T2b | 24 | 10 | |
| | T3 | 58 | 10 | |
| | T4 | 17 | 5 | 0.44 |
| pN stage | N0 | 248 | 62 | |
| | N1 | 106 | 22 | |
| | N2 | 32 | 8 | |
| | N3 | 5 | | |
| | NX | 2 | 4 | 0.031 |
| pM stage | M0 | 323 | 78 | |
| | M1 | 5 | | |
| | M1a | 1 | | |
| | M1b | 1 | | 0.709 |
| pTNM stage | I | 2 | 1 | |
| | IA | 72 | 16 | |
| | IB | 113 | 35 | |
| | II | 3 | | |
| | IIA | 54 | 10 | |
| | IIB | 71 | 19 | |
| | III | 2 | 1 | |
| | IIIA | 54 | 8 | |
| | IIIB | 13 | 4 | |
| | IV | 7 | | 0.657 |
| New tumor event type | Metastasis | 25 | 10 | |
| | Metastasis: Primary | 2 | | |
| | Metastasis: Recurrence | 1 | 1 | |
| | Primary | 7 | 3 | |
| | Recurrence | | 5 | 0.619 |

*(Continued)*

TABLE 1 Continued

| | Character | *TTN* mutant | *TTN* WT | *P*-value |
|---|---|---|---|---|
| Smoking | Non-smoking | 10 | 8 | |
| | Smoking | 74 | 85 | 0.016 |
| Radiation therapy | Non-radiation | 111 | 25 | |
| | Radiation | 14 | 1 | 0.435 |
| History of neoadjuvant treatment | Neoadjuvant | 1 | 3 | |
| | No neoadjuvant | 391 | 93 | |
| | Yes, Pharmaceutical Treatment Prior to Resection | 1 | | 0.03 |
| Therapy type | Ancillary: Chemotherapy | 1 | | |
| | Chemotherapy | 109 | 21 | |
| | Chemotherapy: | 2 | 1 | |
| | Chemotherapy: Other. specify in notes:Vaccine | 1 | | |
| | Chemotherapy: Targeted Molecular therapy | 1 | | |
| | Chemotherapy: Vaccine | 1 | | |
| | Other. specify in notes | 1 | | |
| | Vaccine | 1 | | 0.995 |

## *TTN* mutations are associated with longer survival in LUSC

There were 211 deaths among 489 patients with LUSC (Table 1); according to the Cox proportional risk regression model, *TTN* mutation was associated with longer OS (multivariate Cox model P=0.000212, HR=0.547 [95% CI, 0.398–0.753], Figure 2A), PFS (multivariate Cox model P=0.0323, HR=0.649 [95% CI, 0.437–0.964], Figure 2B), and DSS (multivariate Cox model P=0.0181, HR=0.555 [95% CI, 0.341–0.904], Figure 2D), *TTN* mutations was not associated with DFS (Figure 2C). Therefore, it can be concluded that *TTN* mutations are associated with survival in LUSC.

## *TTN* mutations in LUSC drive heterogeneous gene expression

To determine the role of gene expression heterogeneity driven by *TTN* mutations in LUSC in the construction of prognostic models, we screened differentially expressed genes (DEGs) in 393 LUSC *TTN* mutant tissues and 96 LUSC tissues in TCGA. A total of 50 up-regulated genes and 7 down-regulated genes were screened, and the volcano are shown in Figures 3A, B showed the gene expression heat map of each sample, respectively. For DEGs, GO function enrichment and KEGG differential analyzes were performed on DEGs (Figure 3C).

## *TTN* mutation prognosis model was built based on 10 prognostically favorable factors

To further discuss the main genes that prolong the survival of lung adenocarcinoma patients with *TTN* mutations, 57 significant differential genes were analyzed for their correlation with the prognosis of lung adenocarcinoma patients, and 10 significant differential genes with prognostic differences were obtained. Among them, low expression of *C12orf56, DLX5, DSC3, FEZF1, GSTA1, H2BC9, IGSF11* and high expression of *EREG, CEACAM6, SLC34A2* in lung adenocarcinoma were found to predict poor prognosis for patients (Figure 4A). The results of differential expression analysis showed that *C12orf56, DLX5, DSC3, FEZF1, GSTA1, H2BC9, IGSF11* were highly expressed in *TTN* mutation patients, and *EREG, CEACAM6, SLC34A2* were lowly expressed in *TTN* mutation patients (Figure 4B). These data indicate that these 10 significant differential expression prognostic genes are beneficial factors for the prognosis and survival of patients with prolonged *TTN* mutations.

Subsequently, a prognostic feature was established based on the 10 beneficial factors using LASSO Cox analysis (Figures 5A, B). The lung adenocarcinoma patients were divided into high-risk and low-risk groups using the risk score Riskscore = (-4e-04)**DLX5* + (-0.0772) **FEZF1* + (-0.0933)**H2BC9* + (0.0637)**EREG* + (0.0298)**SLC34A2* (Figure 5C), and their survival status was displayed in Figure 5D. The difference in survival between the high-risk and low-risk groups was shown in Figure 5E (P=0.00619). Prognostic survival prediction was performed for 1 year, 3 years, and 5 years, and the prognostic model showed good sensitivity and specificity (Figure 5F).

## Validation of the constructed prognostic model by an external independent dataset

To validate the accuracy of our screened independent prognostic factors, we constructed risk prognosis models in

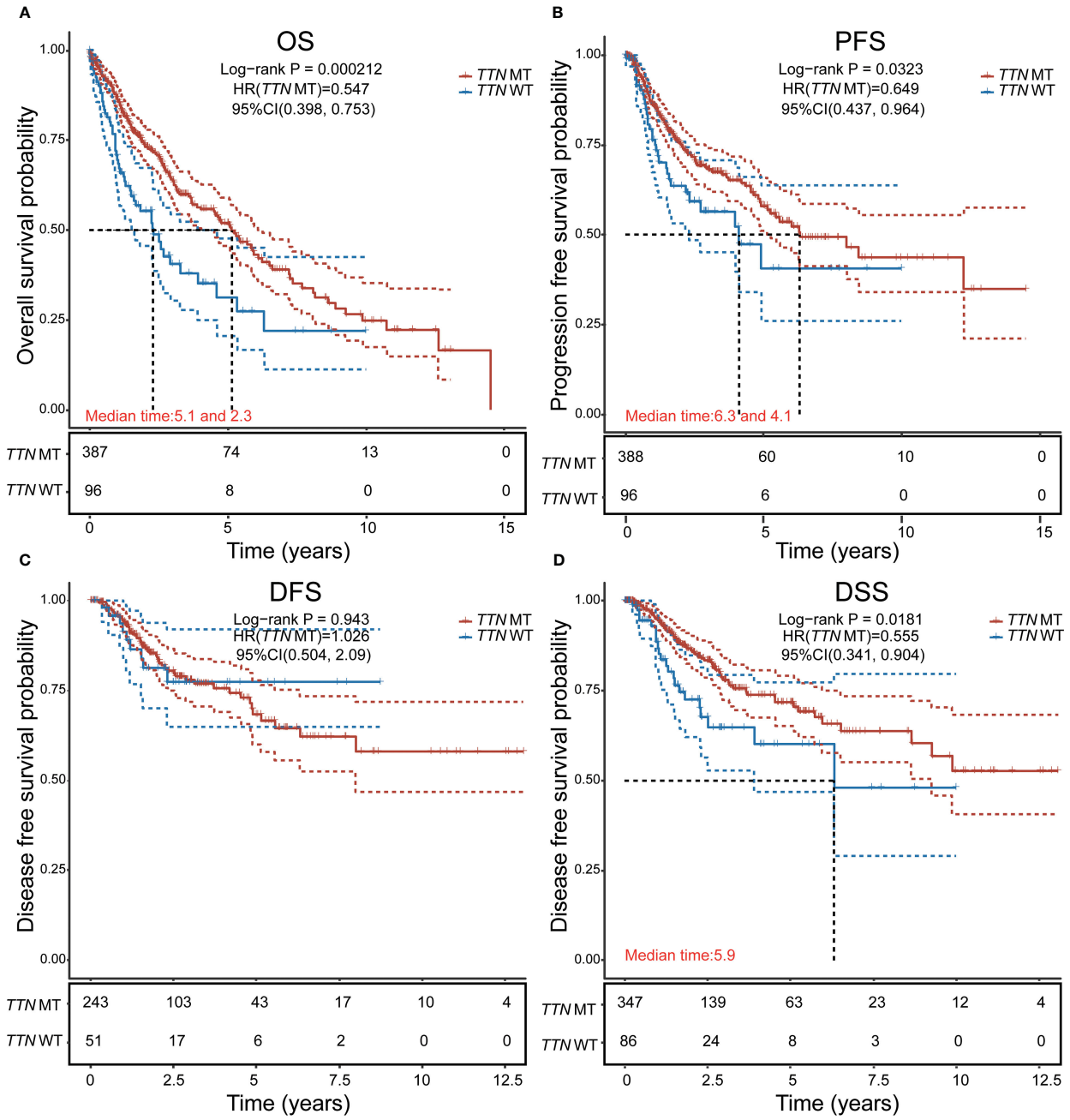



FIGURE 2
Analysis of the effect of *TTN* mutation on prognostic survival in LUSC using the Cox regression model. **(A)** Overall survival (OS). **(B)** Progression-free survival (PFS). **(C)** Disease-free survival (DFS). **(D)** disease-specific survival (DSS). HR represents the risk coefficient of the high-expression group relative to the samples in the low-expression group; 95% CI represents the HR confidence interval; Median time represents the survival rate of the two groups (high-expression group and low-expression group) at 50% (i.e., median survival time) corresponding to the time in years; and the test between the two groups was performed by log-rank and was statistically significant at $P < 0.05$.

GSE73403 datasets and obtained. RiskScore=(-0.61828)**IGSF11* +(-0.13167)**DLX5*+(0.31607)**EREG*+(-0.41127)**SLC34A2* risk assessment models (Figures 6A, B), which divided patients into two groups of high risk and low risk (Figures 6C, D), with significant differences in prognosis (Figure 6E). Prognostic survival was predicted for 1, 3, and 5 years, and this prognostic model showed preferably sensitivity and specificity (Figure 6F).

## H2BC9 was identified as an independent prognostic factor for lung squamous cell carcinoma with *TTN* mutation

To determine the prognostic factors of *TTN* mutations on OS in LUSC, we performed univariate and multivariate Cox regression analysis of prognostic correlates, and the results are shown in Figures

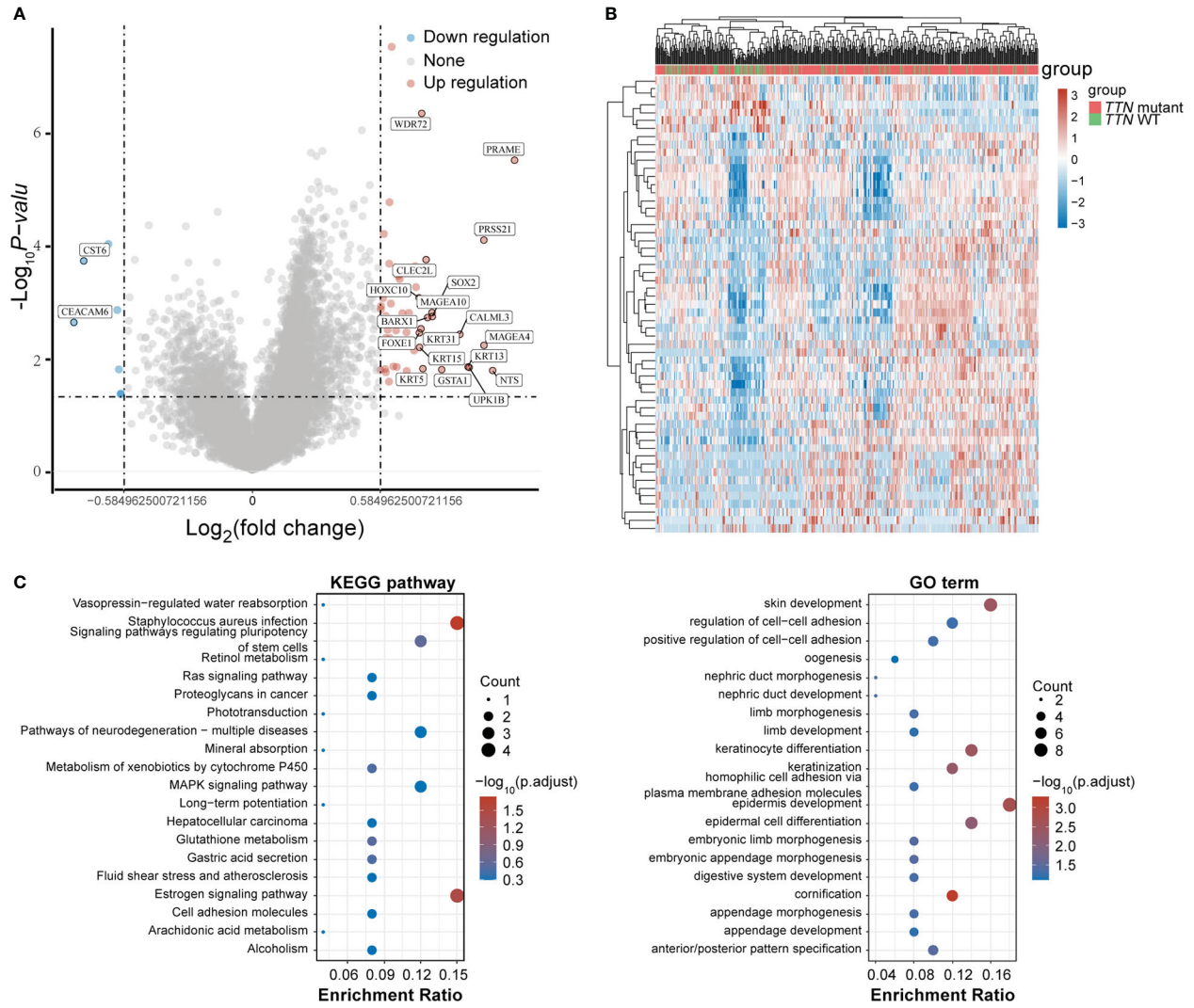



FIGURE 3
DEGs screening. **(A)** Volcano plot. Red dots represent significantly differentially upregulated genes, blue dots represent significantly differentially downregulated genes, and gray dots represent insignificant genes. **(B)** Heat map of the top 50 significantly DEGs. Different colors represent expression trends in different tissues. **(C)** GO functional enrichment and KEGG pathway analysis of upregulated genes.

7A, B. In the univariate Cox regression analysis, *DLX5, FEZF1, H2BC9, EREG, SLC34A2*, pT-stage and pTNM-stage were significantly correlated with the OS of the *TTN* mutations in LUSC patients (Figure 6A). Then, *DLX5, FEZF1, H2BC9, EREG, SLC34A2*, pT-stage and pTNM-stage were analyzed to determine the independent prognostic factors based on a regression analysis of *TTN* mutations in patients with LUSC by multivariateial Cox analysis. H2BC9 was found to be significantly associated with the OS of the *TTN* mutation in the patients (Figure 7B). All independent factors were combined to create nomogram prediction plots predicting 1-, 3-, and 5-year OS (Figures 7C, D).

## Discussion

Lung carcinogenesis and progression are complex biological processes with multi-gene involvement and multiple steps. Few driver mutations have been identified in LUSC that can be used as drug targets, and, unlike with lung adenocarcinoma, the process of development does not depend on classical driver mutations, but more on epigenetic alterations (10–12). Therefore, because of its unique molecular pathology, LUSC cannot benefit significantly from molecularly targeted therapies as lung adenocarcinoma can. The TCGA database maps a large number of human tumor genomic variants through genomic analysis techniques. By mining this database with large samples of data, valuable research subjects can be derived and further targeted studies can be conducted. This database can provide a theoretical basis for clinical treatment. In LUSC, exploring the molecular-level determinants of the immunotherapeutic response is challenging, and many studies are underway to address this issue. The results of the TCGA study of mutations in the genome of 178 LUSC patients revealed potential therapeutic targets mainly associated with PI3K/AKT signaling pathways and signaling pathways mediated by

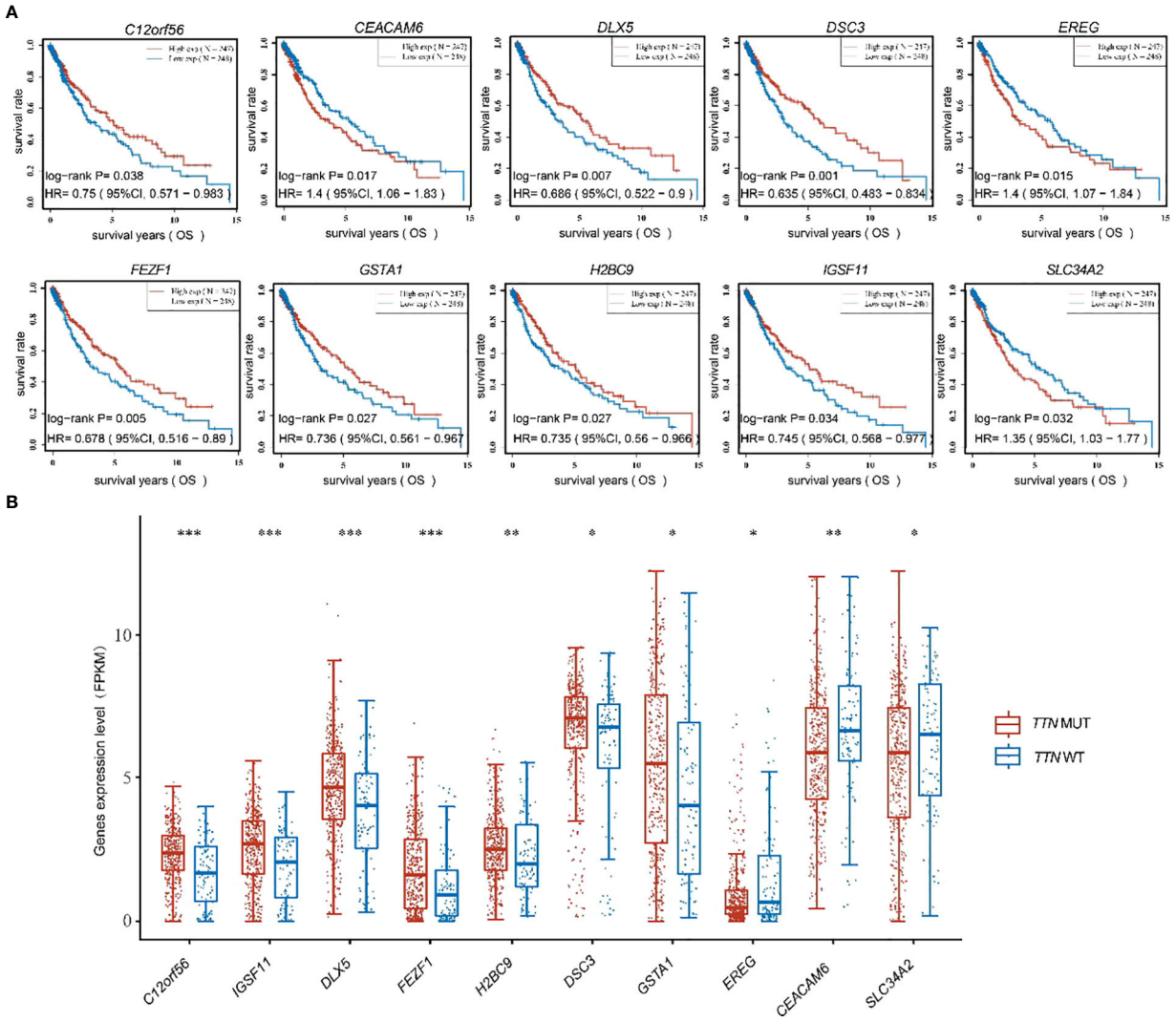



FIGURE 4
Figure 4 was shown to present the expression of ten beneficial factors and their impact on patient prognosis. **(A)** The high or low expression of the ten beneficial factors (*C12orf56, DLX5, DSC3, FEZF1, GSTA1, H2BC9, IGSF11, EREG, CEACAM6, SLC34A2*) was demonstrated to have an impact on the prognosis of lung squamous cell carcinoma patients. **(B)** The expression of the ten beneficial factors (*C12orf56, DLX5, DSC3, FEZF1, GSTA1, H2BC9, IGSF11, EREG, CEACAM6, SLC34A2*) in patients with *TTN* mutations and wild-type *TTN* patients was shown to be presented. t test was used for comparison between groups, with * indicating p<0.05, ** indicating p<0.01, and *** indicating p<0.001.

receptor tyrosine kinases (RTKs), including *PIK3CA* mutations (16% of cases); *FGFR* amplification or mutations (18% of cases); *PTEN* mutations or deletions (15% of cases); *EGFR* amplification and *PDGFRA* amplification or mutation (9% each); and *ERBB2* amplification, *DDR2* mutation, and *BRAF* mutation (4% each) (13).

In recent years, it has been shown that single or double mutations in *TP53* and *TTN*, two genes with high mutation frequency in NSCLC, have different prognostic effects, and *TTN* gene mutations were significantly associated with LUSC (14). Meanwhile, a study analyzed somatic mutations in LUSC samples based on datasets obtained from TCGA and the International Cancer Genome Consortium (ICGC) database and found that *TTN* mutations were associated with tumor mutational burden (TMB) and positively correlated with LUSC prognosis (15). And in a study by Xie et al. (16), *TTN* with mutations in higher TMB was found to be associated with prognostic outcome, and *TTN* mutations were found to enhance anti-tumor immune responses according to the CIBERSORT algorithm. In the above study, the prognostic impact of *TTN* mutation or double mutation with *TP53* on LUSC was systematically described, and the relationship between *TTN* mutated genes and TMB was analyzed based on TCGA and ICGC databases. In our study, we investigated the prognostic impact of heterogeneous expression of genes due to *TTN* mutations on LUSC from the TCGA database and screened for their independent prognostic factors.

In our study, it was found that the prolonged survival of lung squamous cell carcinoma patients was caused by *TTN* mutation. Based on this, the heterogeneous expression of genes driven by *TTN* mutation was analyzed, and 57 differential genes were identified. After performing a prognostic analysis of the differential genes combined with their expression in *TTN* wild-type and *TTN* mutation patients, 10 genes were found to have contributed

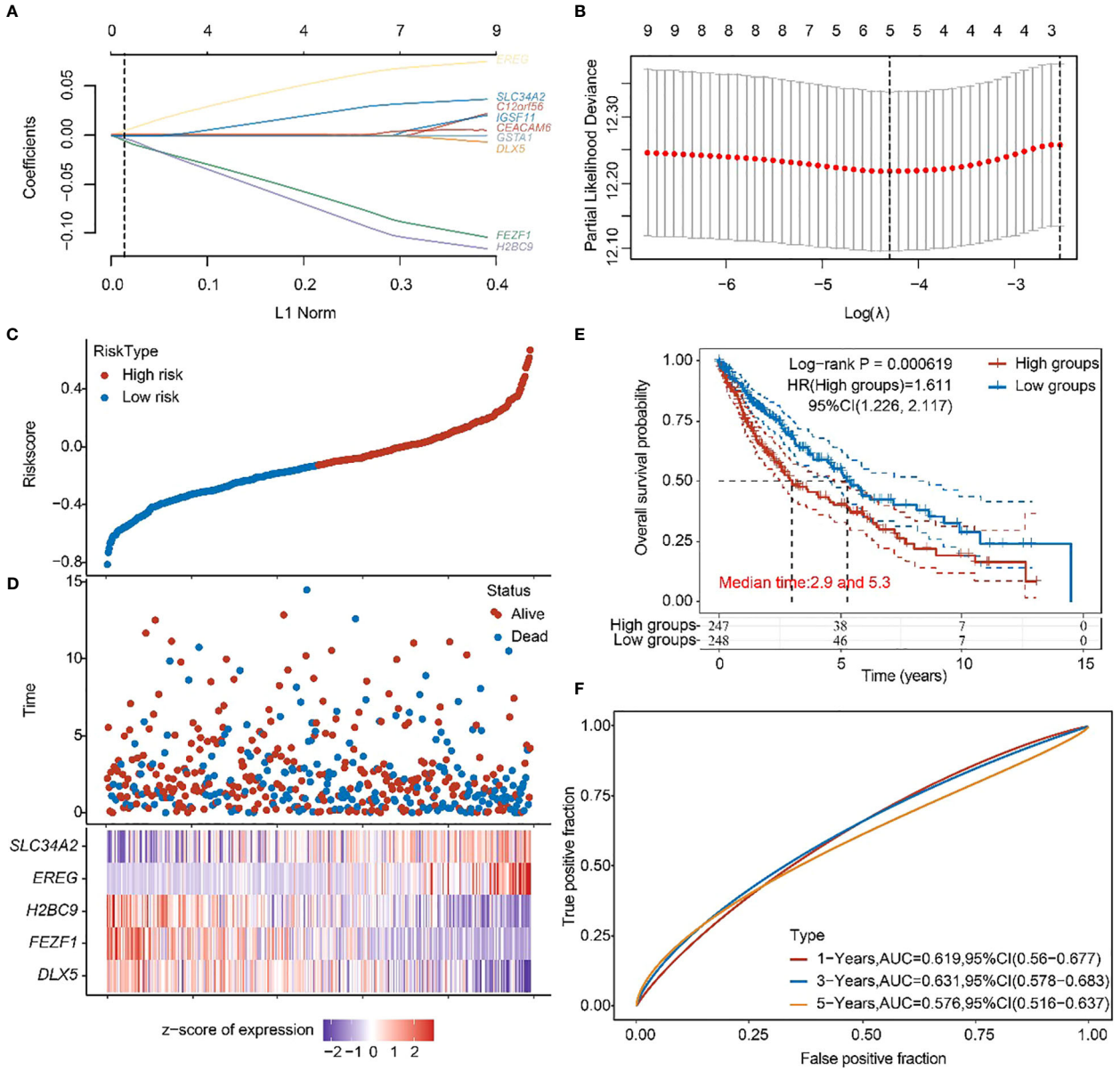



FIGURE 5
Construction of prognostic features of *TTN* mutation DEGs in LUSC. **(A, B)** Distribution of LASSO coefficients of DEGs to obtain the adjustment parameter. λ.min = 0.0329, the vertical black dashed line in **B** defines the optimal λ value. **(C)** Distribution of risk score of *TTN* mutation in LUSC. **(D)** Survival status and duration of *TTN* mutation in LUSC patients. **(E)** KM survival curves for high- and low-risk groups. **(F)** 1-year, 3-year, and 5-year ROC curves.

significantly to the prognostic potential of prolonged survival. A prognostic model was then constructed based on these 10 prognostic beneficial factors, and a prognostic model with *DLX5, FEZF1, H2BC9, EREG, and SLC34A2* as the main factors was obtained.

*SLC34A2* can be highly expressed in the lung (17, 18). In human lungs, *SLC34A2* is expressed only in type II alveolar epithelial cells (AT-II) and is required for the synthesis of AT-II lung surfactant (19). AT-II cells may transform into cancer stem cells in response to exogenous or endogenous factors and eventually induce oncogenic effects and NSCLC carcinogenesis (20–22). *SLC34A2* may have a physiological function in AT-II, and its mutation or abnormal expression will certainly affect the normal function of AT-II, which is associated with lung tumorigenesis. Also, recent studies have reported that *SLC34A2* plays a key role in lung cancer. *SLC34A2* expression was increased during fetal lung development and early embryonic development, but it was decreased in NSCLC tissues compared with surrounding normal lung tissue (23). In addition, studies have reported that *SLC34A2* may inhibit the growth, invasion, and migration of lung cancer cells through PI3K/Akt and Ras/Raf/MEK signaling pathways, by acting as a tumor suppressor (23, 24). In some cancers, *SLC34A2* is considered a potential prognostic marker of tumor disease. *SLC34A2* mutations in breast and thymic tumors reduce the life expectancy of patients, and *SLC34A2* overexpression was found to reduce the life

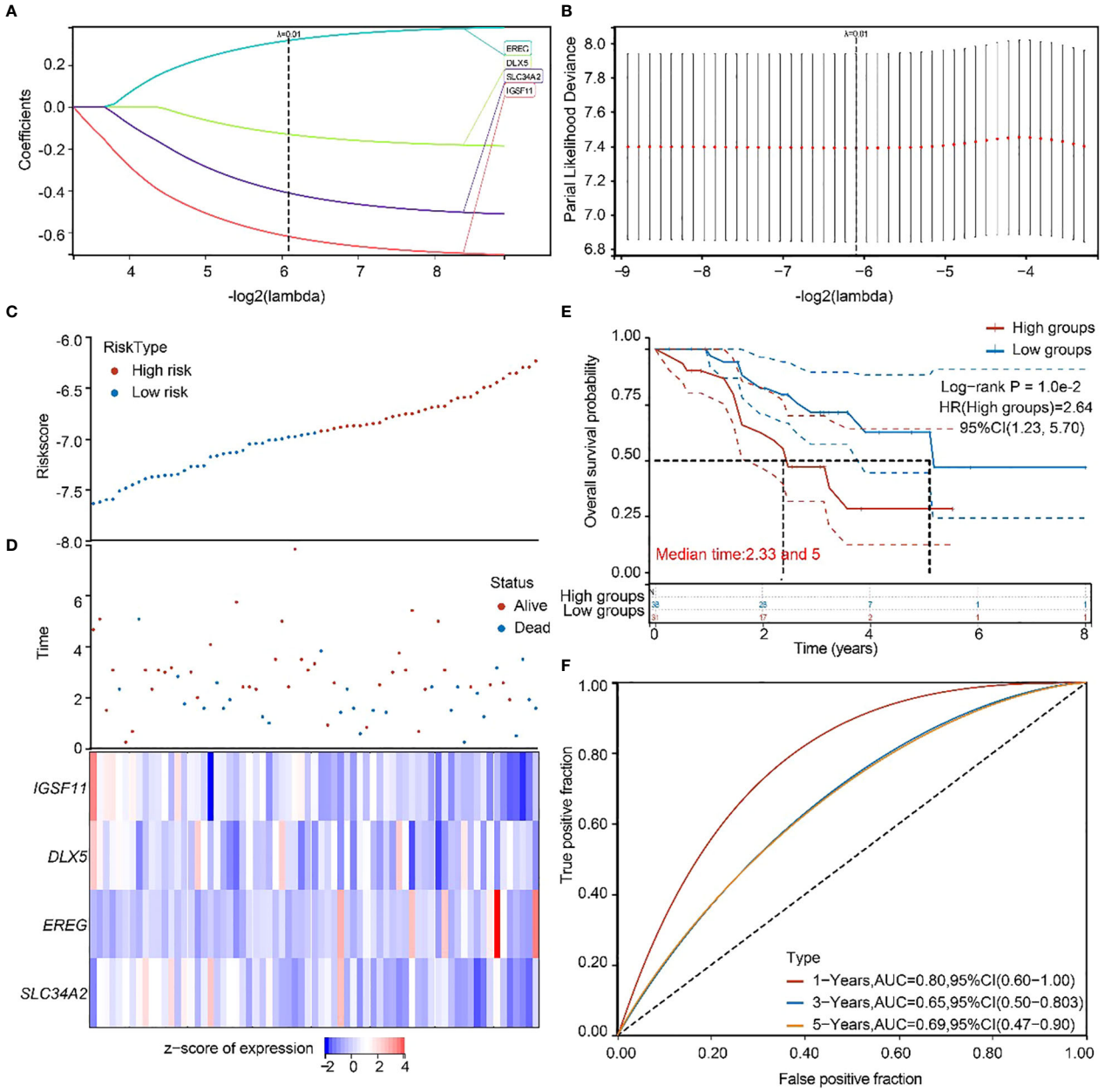



FIGURE 6
GEO independent dataset to validate the risk model. **(A, B)** Distribution of LASSO coefficients in the GSE73403 datasets, obtained by adjusting the parameters λ.min = 0.0146. **(C, D)** Distribution of risk scores of the constructed models for the GSE73403 datasets. **(E)** KM survival curves for the high-and low-risk groups. **(F)** 1-year, 3-year, and 5-year ROC curves.

expectancy of patients with brain tumors, ovarian tumors, and pancreatic tumors (25). High *SLC34A2* expression was present in approximately 75% of lung adenocarcinoma samples and was associated with significantly improved overall patient survival (26). A recent study showed that *SLC34A2* (hazard ratio: 0.86, P = 0.03, coefficient = -0.1551) could be a potential prognostic immune-related gene by validating the characteristics of immune infiltration analysis in patients with NSCLC and its correlation with survival outcomes (27). It has been revealed by other studies that the high expression of *EREG* in lung adenocarcinoma is detrimental to the survival of individuals (28). In gastric cancer, it was found that the EREG mRNA expression was an independent predictor of poor survival, and that the overall survival was significantly reduced in patients with gastric cancer with high expression (29). *DLX* homeobox genes, which belong to the NK family, play a dual role in development and cancer (30). In human ovarian cancer cells, it has been found that DLX5 is essential for regulating AKT signaling, thereby promoting cell proliferation and survival. *FEZF1*, which is a highly conserved transcription factor belonging to the C2H2 zinc finger protein family, has been shown to reduce cell proliferation and migration in human cervical cancer cell lines when it is knockdown and to act as an independent predictor of cervical cancer recurrence (31). *H2BC9*, which belongs to the H2B aggregator histone gene family, is commonly recognized as a molecular biomarker and an independent

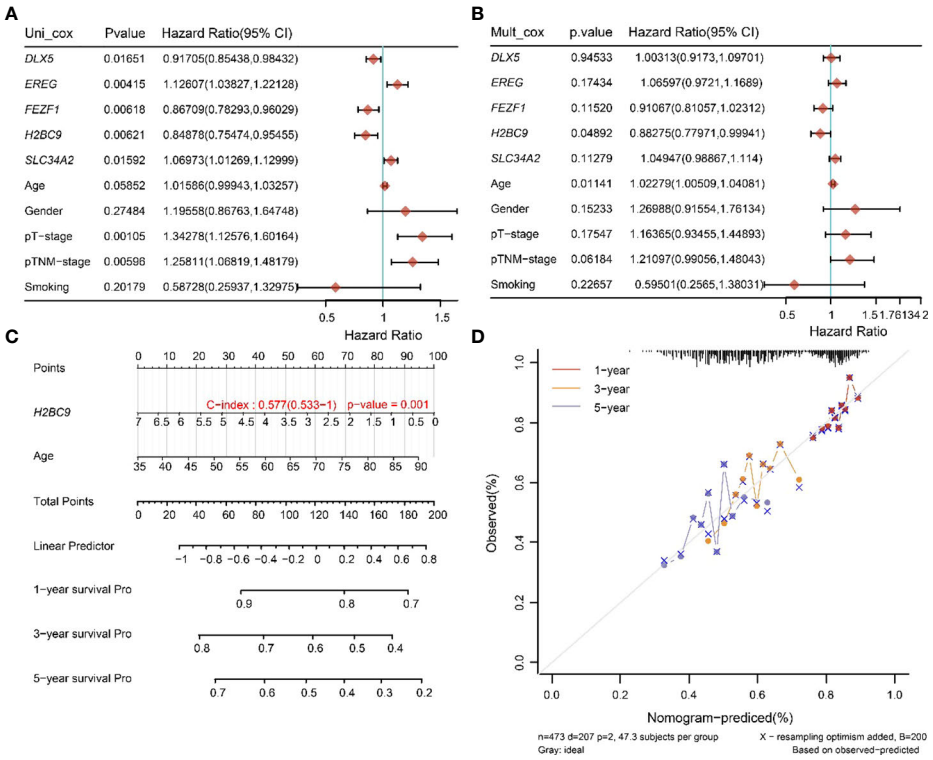



FIGURE 7
Univariate and multivariate Cox screening prognostic factors. **(A, B)** Analysis of clinical factors and risk scores using univariate Cox analysis, and multivariate Cox analysis was used to analyze significant factors. **(C)** Construction of clinical diagnosis column line graphs based on clinical characteristics and risk scores. **(D)** Calibration plots for predicting recurrence at 1, 3, and 5 years. The x-axis indicates the recurrence probability predicted from the column line graphs, and the y-axis indicates the actual recurrence probability.

prognostic factor for *H2BC9* cervical cancer in cervical cancer and glioma (32, 33).

In conclusion, a study was conducted in which the gene mutations in lung squamous cell carcinoma with *TTN* mutations were analyzed. Ten prognostic beneficial factors of the heterogeneous expression of lung squamous cell carcinoma genes driven by *TTN* mutations were screened for and *DLX5, FEZF1, H2BC9, EREG*, and *SLC34A2* were identified as the main factors of the prognostic models. These factors were validated on external data sets. However, it was noted that a public database was used for the analysis and the screened prognostic factors were not experimentally validated. Additionally, data from a large clinical cohort was lacking for testing. It was planned that in a follow-up study, data collected from multiple cases would be analyzed, the correlation of *TTN* mutation-driven gene heterogeneous expression would be validated based on clinical trials, and the mechanism of *TTN* mutation regulation of gene expression would be investigated through cellular assays.

In conclusion, in our study, we identified *SLC34A2*, *BARX1*, and *HOXC10* as the major prognosis-related genes by constructing a prognostic model constructed by the differential expression of *TTN* mutation-driven genes through bioinformatics analysis. These findings could provide prognostic biomarkers for LUSC and possible new targets for its clinical 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.

## Author contributions

YZ, GZ, XP and CZ designed the study. ZL, XZ and XT completed the analysis. ZL and XT wrote the manuscript. RW proofread the manuscript and put forward constructive opinions, and all the authors have read and approved the submission of the manuscript. All authors contributed to the article and approved the submitted version.

## Conflict of interest

ZL, XT and YZ were employed by Biotrans Technology Co., LTD.

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

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

1. Bray F, Ferlay F, Soerjomataram I, Siegel SL, 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.

2. Harrison PT, Vyse S, Huang PH. Rare epidermal growth factor receptor (EGFR) mutations in non-small cell lung cancer. *Semin Cancer Biol* (2020) 61:167–79. doi: 10.1016/j.semcancer.2019.09.015

3. Yang C, Huang X, Li Y, Chen J, Lv Y, Dai S. Prognosis and personalized treatment prediction in TP53-mutant hepatocellular carcinoma: An in silico strategy towards precision oncology. *Brief Bioinform* (2021) 22(3):bbaa164. doi: 10.1093/bib/bbaa164

4. Duffy MJ, Synnott NC, O'Grady S, Crown J. Targeting p53 for the treatment of cancer. *Semin Cancer Biol* (2022) 79:58–67. doi: 10.1016/j.semcancer.2020.07.005

5. VanderLaan PA, Rangachari D, Mockus SM, Spotlow V, Reddi HV, Malcolm J, et al. Mutations in TP53, PIK3CA, PTEN and other genes in EGFR mutated lung cancers: Correlation with clinical outcomes. *Lung Cancer* (2017) 106:17–21. doi: 10.1016/j.lungcan.2017.01.011

6. Chauveau C, Rowell J, Ferreiro A. A rising titan: TTN review and mutation update. *Hum Mutat* (2014) 35(9):1046–59. doi: 10.1002/humu.22611

7. Kim N, Hong Y, Kwon D, Yoon S. Somatic mutaome profile in human cancer tissues. *Genomics Inform* (2013) 11(4):239–44. doi: 10.5808/GI.2013.11.4.239

8. Rees M, Nikoopour R, Fukuzawa A, Kho AL, Fernandez-Garcia MA, Wraige E, et al. Making sense of missense variants in TTN-related congenital myopathies. *Acta Neuropathol* (2021) 141(3):431–53. doi: 10.1007/s00401-020-02257-0

9. Tharp CA, Haywood ME, Sbaizero O, Taylor MRG, Mestroni L. The giant protein titin's role in cardiomyopathy: Genetic, transcriptional, and post-translational modifications of TTN and their contribution to cardiac disease. *Front Physiol* (2019) 10:1436. doi: 10.3389/fphys.2019.01436

10. Wang BY, Huang JY, Chen HC, Lin CH, Lin SH, Hung WH, et al. The comparison between adenocarcinoma and squamous cell carcinoma in lung cancer patients. *J Cancer Res Clin Oncol* (2020) 146(1):43–52. doi: 10.1007/s00432-019-03079-8

11. Ruiz-Cordero R, Devine WP. Targeted therapy and checkpoint immunotherapy in lung cancer. *Surg Pathol Clin* (2020) 13(1):17–33. doi: 10.1016/j.path.2019.11.002

12. Zeng Z, Yang F, Wang Y, Zhao H, Wei F, Zhang P, et al. Significantly different immunological score in lung adenocarcinoma and squamous cell carcinoma and a proposal for a new immune staging system. *Oncoimmunology* (2020) 9(1):1828538. doi: 10.1080/2162402X.2020.1828538

13. Rooney M, Devarakonda S, Govindan R. Genomics of squamous cell lung cancer. *Oncologist* (2013) 18(6):707–16. doi: 10.1634/theoncologist.2013-0063

14. Cheng X, Yin H, Fu J, Chen C, An J, Guan J, et al. Aggregate analysis based on TCGA: TTN missense mutation correlates with favorable prognosis in lung squamous cell carcinoma. *J Cancer Res Clin Oncol* (2019) 145(4):1027–35. doi: 10.1007/s00432-019-02861-y

15. Zou S, Ye J, Hu S, Wei Y, Xu J. Mutations in the TTN gene are a prognostic factor for patients with lung squamous cell carcinomas. *Int J Gen Med* (2022) 15:19–31. doi: 10.2147/IJGM.S343259

16. Xie X, Tang Y, Sheng J, Shu P, Zhu X, Cai X, et al. Titin mutation is associated with tumor mutation burden and promotes antitumor immunity in lung squamous cell carcinoma. *Front Cell Dev Biol* (2021) 9:761758. doi: 10.3389/fcell.2021.761758

17. Wagner CA, Hernando N, Forster IC, Biber J. The SLC34 family of sodium-dependent phosphate transporters. *Pflugers Arch* (2014) 466(1):139–53. doi: 10.1007/s00424-013-1418-6

18. Takeda E, Yamamoto H, Nashiki K, Sato T, Arai H, Taketani Y. Inorganic phosphate homeostasis and the role of dietary phosphorus. *J Cell Mol Med* (2004) 8 (2):191–200. doi: 10.1111/j.1582-4934.2004.tb00274.x

19. Hashimoto M, Wang DY, Kamo T, Zhu Y, Tsujiuchi T, Konishi Y, et al. Isolation and localization of type IIb Na/Pi cotransporter in the developing rat lung. *Am J Pathol* (2000) 157(1):21–7. doi: 10.1016/S0002-9440(10)64512-9

20. Jacob A, Vedaie M, Roberts DA, Thomas DC, Villacorta-Martin C, Alysandratos KD, et al. Derivation of self-renewing lung alveolar epithelial type II cells from human pluripotent stem cells. *Nat Protoc* (2019) 14(12):3303–32. doi: 10.1038/s41596-019-0220-0

21. Desai TJ, Brownfield DG, Krasnow MA. Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* (2014) 507(7491):190–4. doi: 10.1038/nature12930

22. Cho HC, Lai CY, Shao LE, Yu J. Identification of tumorigenic cells in Kras (G12D)-induced lung adenocarcinoma. *Cancer Res* (2011) 71(23):7250–8. doi: 10.1158/0008-5472

23. Kopantzev EP, Monastyrskaya GS, Vinogradova TV, Zinovyeva MV, Kostina MB, Filyukova OB, et al. Differences in gene expression levels between early and later stages of human lung development are opposite to those between normal lung tissue and non-small lung cell carcinoma. *Lung Cancer* (2008) 62(1):23–34. doi: 10.1016/j.lungcan.2008.02.011

24. Wang Y, Yang W, Pu Q, Yang Y, Ye S, Ma Q, et al. The effects and mechanisms of SLC34A2 in tumorigenesis and progression of human non-small cell lung cancer. *J BioMed Sci* (2015) 22(1):52. doi: 10.1186/s12929-015-0158-7

25. Vlasenkova R, Nurgalieva A, Akberova N, Bogdanov M, Kiyamova R. Characterization of SLC34A2 as a potential prognostic marker of oncological diseases. *Biomolecules* (2021) 11(12):1878. doi: 10.3390/biom11121878

26. Zhang Z, Ye S, Zhang M, Wu J, Yan H, Li X, et al. High expression of SLC34A2 is a favorable prognostic marker in lung adenocarcinoma patients. *Tumour Biol* (2017) 39(7):1010428317720212. doi: 10.1177/1010428317720212

27. Mi K, Chen F, Qian Z, Chen J, Lv D, Zhang C, et al. Characterizing heterogeneity of non-small cell lung tumour microenvironment to identify signature prognostic genes. *J Cell Mol Med* (2020) 24(24):14608–18. doi: 10.1111/jcmm.16092

28. Wang W, Xu C, Ren Y, Wang S, Liao C, Fu X, et al. A novel cancer stemness-related signature for predicting prognosis in patients with colon adenocarcinoma. *Stem Cells Int* (2021) 2021:7036059. doi: 10.1155/2021/7036059

29. Suematsu H, Hashimoto I, Hiroshima Y, Watanabe H, Kano K, Takahashi K, et al. Clinical significance of EREG gene expression in gastric cancer tissue after curative surgery. *Anticancer Res* (2022) 42(8):3873–8. doi: 10.21873/anticanres.15880

30. Tan Y, Testa JR. DLX Genes: Roles in development and cancer. *Cancers (Basel)* (2021) 13(12):3005.

31. Lan Y, Xiao X, Luo Y, He Z, Song X. FEZF1 is an independent predictive factor for recurrence and promotes cell proliferation and migration in cervical cancer. *J Cancer* (2018) 9(21):3929–38. doi: 10.7150/jca.26073

32. Li X, Tian R, Gao H, Yang Y, Williams BRG, Gantier MP, et al. Identification of a histone family gene signature for predicting the prognosis of cervical cancer patients. *Sci Rep* (2017) 7(1):16495. doi: 10.1038/s41598-017-16472-5

33. Jia J, Han Z, Wang X, Zheng X, Wang S, Cui Y. H2B gene family: A prognostic biomarker and correlates with immune infiltration in glioma. *Front Oncol* (2022) 12:966817. doi: 10.3389/fonc.2022.966817

# 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