EARLY DETECTION AND DIAGNOSIS OF CANCER

EDITED BY: Jian-Bing Fan, Jin Jen, Neeraj S. Salathia, Youping Deng, Rahul Kumar and Lei Wei

PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-870-9
DOI 10.3389/978-2-88974-870-9

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

EARLY DETECTION AND DIAGNOSIS OF CANCER

Topic Editors:
Jian-Bing Fan, Illumina, United States
Jin Jen, Mayo Clinic, United States
Neeraj S. Salathia, Bristol Myers Squibb, United States
Youping Deng, Rush University Medical Center, United States
Rahul Kumar, Indian Institute of Technology Hyderabad, India
Lei Wei, University at Buffalo, United States

Jian-Bing Fan is a professor at Southern Medical University, China, and founder of AnchorDx. Jin Jen has a joint appointment with the Mayo Clinic and Bristol-Myers Squibb. Neeraj Salathia is employed by Bristol-Myers Squibb. All other Topic Editors declare no competing interests with regard to the Research Topic subject.

Citation: Fan, J.-B., Jen, J., Salathia, N. S., Deng, Y., Kumar, R., Wei, L., eds. (2022). Early Detection and Diagnosis of Cancer. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88974-870-9

Table of Contents

frontiers in Genetics

EDITORIAL
published: 14 March 2022
doi: 10.3389/fgene.2022.875421

Editorial: Early Detection and Diagnosis of Cancer

Rahul Kumar[1]*, *Youping Deng*[2], *Jian-Bing Fan*[3] and *Lei Wei*[4]

[1]Computational Genomics and Transcriptomics Laboratory, Department of Biotechnology, Indian Institute of Technology Hyderabad, Telangana, India, [2]Department of Quantitative Health Sciences, John A. Burns School of Medicine, University of Hawaii at Manoa, Honolulu, HI, United States, [3]AnchorDx Medical Co., Guangzhou, China, [4]Department of Biostatistics and Bioinformatics, Roswell Park Comprehensive Cancer Center, Buffalo, NY, United States

Keywords: cancer treatment, early detection, diagnosis, survival, prognosis

Editorial on the Research Topic

Early Detection and Diagnosis of Cancer

The efficacy of the cancer treatment largely depends on the stage of its detection. Detecting cancer at an early stage can significantly improve cancer treatment. Unfortunately, early-stage cancer detection is limited by various factors and is still a moonshot. However, recent advancements in the next generation sequencing technologies led to the identification of biomarkers with high prognostic value. This research topic compiled some of the recent research work done in this direction.

Zhu et al. identified five genes (i.e., AP2, CDH1, DACT2, HIN1, and RASSF1A), which were found to be frequently methylated in thyroid cancer. They further used this panel of genes to understand the epigenetic heterogeneity in thyroid cancer and found the association between epigenetic heterogeneity and cancer development and progression.

In a bioinformatics-based analysis, Hu et al. found the association between high expression of PLEC1 and poor prognosis of gastric cancer (GC). They also found SNPs (rs3765524 C > T, rs2274223 A > G and rs3781264 T > C) associated with increased risk of GC. Chemotherapy resistance is a major obstacle in GC patients treated with fluorouracil and cisplatin. Liu et al. identified nine autophagy-related genes (ARGs) signatures associated with the prognosis of chemotherapy patients using TCGA data. They also illustrated the role of increased expression of CXCR4 and its association with survival. Zhang et al. identified a prognostic signature of ten differentially expressed lncRNAs to predict the prognosis of GC. These lncRNAs are related to inflammation and can be explored for the therapeutic opportunity.

Xiao et al. studied the prognostic potential of endoplasmic reticulum (ER) stress-related genes, ATF6, EMC6, XBP1 and CHOP, and apoptosis-related gene, APAF1 in pancreatic cancer (PC). They found that ATF6 upregulation and EMC6 and APAF1 downregulation were significantly associated with poor survival of PC patients.

Wu et al. studied the role of centromere protein N (CENPN) in the pathology of gliomas using clinical and gene expression data from TCGA and CGGA databases. They found the elevated expression of CENPN associated with unfavorable outcomes and further validated their findings in independent clinical specimens. GSEA and ssGSEA analysis revealed the association of CENPN with inflammation, immune-related signaling, and infiltration of immune cells.

Zhao et al. analyzed the pan-cancer expression data and found that secreted phosphoprotein (SPP1) is overexpressed in most of the cancer types including cervical cancer. They also reproduced their findings in GEO database. Based on their analysis, they concluded that SPP1 is significantly overexpressed in cervical cancer than normal cervical epithelial tissues and may be explored as a promising prognostic biomarker for cervical cancer.

OPEN ACCESS

Edited and reviewed by:
Stephen J. Bush,
University of Oxford, United Kingdom

***Correspondence:**
Rahul Kumar
rahulk@bt.iith.ac.in

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *14 February 2022*
Accepted: *16 February 2022*
Published: *14 March 2022*

Citation:
Kumar R, Deng Y, Fan J-B and Wei L (2022) Editorial: Early Detection and Diagnosis of Cancer. Front. Genet. 13:875421. doi: 10.3389/fgene.2022.875421

Pu et al. carried out a comprehensive bioinformatics analysis of RNA sequencing data and clinical traits of hepatocellular carcinoma (HCC) patients from TCGA. They found six genes, CEP55, DEPDC1, KIF23, CLSPN, MYBL2, and RACGAP1, closely associated with prognosis and immune infiltration. They also proposed the potential of these genes as therapeutic targets or prognostic biomarkers in HCC.

Zheng et al. studied the glycolysis-related lncRNAs using transcriptome and clinical data of bladder cancer (BCa) from TCGA. They identified 59 differentially expressed glycolysis-related lncRNAs, nine (AC099850.3, AL589843.1, MAFG-DT, AC011503.2, NR2F1-AS1, AC078778.1, ZNF667-AS1, MNX1-AS1, and AC105942.1) of which were found to have prognostic significance. These nine lncRNAs can be utilized as biomarkers to distinguish high-risk and low-risk BCa patients.

These nine research studies published under this research topic identified potential biomarkers of high prognostic and diagnostic significance and advanced the field of biomarker development.

AUTHOR CONTRIBUTIONS

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

Conflict of Interest: Author J-BF was employed by AnchorDx Medical Co.

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 20 May 2021
doi: 10.3389/fgene.2021.591623

Identification of Prognostic Biomarkers and Correlation With Immune Infiltrates in Hepatocellular Carcinoma Based on a Competing Endogenous RNA Network

Zhangya Pu[1], Yuanyuan Zhu[2], Xiaofang Wang[1], Yun Zhong[1], Fang Peng[1,2] and Yiya Zhang[3,4]**

[1] Department of Infectious Diseases, Hunan Key Laboratory of Viral Hepatitis, Xiangya Hospital, Central South University, Changsha, China, [2] NHC Key Laboratory of Cancer Proteomics, Xiangya Hospital, Central South University, Changsha, China, [3] National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China, [4] Department of Dermatology, Xiangya Hospital, Changsha, China

OPEN ACCESS

Edited by:
Jian-Bing Fan,
Illumina, United States

Reviewed by:
Jun Meng,
Dalian University of Technology, China
Shaoqiu Chen,
University of Hawai'i at Mānoa, United States
John A. Burns,
University of Hawai'i at Mānoa, United States

***Correspondence:**
Fang Peng
pengfang@csu.edu.cn
Yiya Zhang
yiya0108@csu.edu.cn

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 05 August 2020
Accepted: 22 April 2021
Published: 20 May 2021

Citation:
Pu Z, Zhu Y, Wang X, Zhong Y, Peng F and Zhang Y (2021) Identification of Prognostic Biomarkers and Correlation With Immune Infiltrates in Hepatocellular Carcinoma Based on a Competing Endogenous RNA Network. Front. Genet. 12:591623. doi: 10.3389/fgene.2021.591623

Background: Hepatocellular carcinoma (HCC) is one of the most common malignant tumors worldwide. Recently, competing endogenous RNAs (ceRNA) have revealed a significant role in the progression of HCC. Herein, we aimed to construct a ceRNA network to identify potential biomarkers and illustrate its correlation with immune infiltration in HCC.

Methods: RNA sequencing data and clinical traits of HCC patients were downloaded from TCGA. The limma R package was used to identify differentially expressed (DE) RNAs. The predicted prognostic model was established using univariate and multivariate Cox regression. A K-M curve, TISIDB and GEPIA website were utilized for survival analysis. Functional annotation was determined using Enrichr and Reactome. Protein-to-protein network analysis was implemented using SRTNG and Cytoscape. Hub gene expression was validated by quantitative polymerase chain reaction, Oncomine and the Hunan Protein Atlas database. Immune infiltration was analyzed by TIMMER, and Drugbank was exploited to identify bioactive compounds.

Results: The predicted model that was established revealed significant efficacy with 3- and 5-years of the area under ROC at 0.804 and 0.744, respectively. Eleven DEmiRNAs were screened out by a K-M survival analysis. Then, we constructed a ceRNA network, including 56 DElncRNAs, 6 DEmiRNAs, and 28 DEmRNAs. The 28 DEmRNAs were enriched in cancer-related pathways, for example, the TNF signaling pathway. Moreover, six hub genes, CEP55, DEPDC1, KIF23, CLSPN, MYBL2, and RACGAP1, were all overexpressed in HCC tissues and independently correlated with survival rate. Furthermore, expression of hub genes was related to immune cell infiltration

in HCC, including B cells, $CD8^+$ T cells, $CD4^+$ T cells, monocytes, macrophages, neutrophils, and dendritic cells.

Conclusion: The findings from this study demonstrate that CEP55, DEPDC1, KIF23, CLSPN, MYBL2, and RACGAP1 are closely associated with prognosis and immune infiltration, representing potential therapeutic targets or prognostic biomarkers in HCC.

Keywords: hepatocellular carcinoma, competing endogenous RNA network, immune infiltration, prognostic prediction model, biomarkers

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most universally malignant tumors in the world, with increasing morbidity and mortality (Shi et al., 2016; Li et al., 2017). Currently, it is the fifth most common cancer and the fourth leading cause of cancer-related death worldwide (Long et al., 2019; Gu et al., 2020). Tumorigenesis of HCC is correlated with several liver primary diseases, such as viral infections, including Hepatitis B virus (HBV), Hepatitis C virus (HCV), and other kinds of hepatotropic viruses, as well as alcoholic liver diseases, dietary aflatoxin exposure diabetes, and other diseases (Fu et al., 2018; Li et al., 2019a). Despite continuous improvement in the methods of diagnosis and treatment, HCC remains a global clinical challenge due to its poor prognosis and low rate of 5-year survival (Yue et al., 2019; Zhang R. et al., 2020). Therefore, individual strategies based on identifying early potential prognostic biomarkers and novel therapeutic targets are urgently needed.

The correlation between protein-coding messenger RNA (mRNA) and non-coding RNA (ncRNA), including long non-coding RNAs (lncRNA) and microRNAs (miRNAs), is complicated and obscure. In 2011, the competing endogenous RNA (ceRNA) hypothesis was elucidated for the first time by Salmena et al. (2011) demonstrating that ncRNAs not only directly take part in the regulation of targeted gene expression but also absorb corresponding miRNAs as natural sponges due to their typically containing more than one miRNA response element (MRE) that competes with mRNA. Recently, increasing evidence indicates that the regulatory network comprised of lncRNA-miRNA-mRNA plays an important role in the physiology and development of various tumors, including HCC, gallbladder cancer, gastric cancer, and others (Long et al., 2019; Yue et al., 2019; Gu et al., 2020). Zhang et al. (2016) indicated that lncRNA-correlated ceRNA networks are involved in diverse biological cancer pathways in glioblastoma. Wang J.J. et al. (2019) identified six lncRNAs, including LINC00536 and MIR7-3HG, that have a significant effect on overall survival in breast cancer. Nevertheless, current studies based on ceRNA networks in multiple databases for HCC are insufficient.

Recently, increasing attention has been paid to immunotherapy research in various cancers, especially in advanced stages, including for mesothelioma, HCC, and others. However, the benefits from immunotherapies are diverse in various tumors and are difficult to evaluate due to a lack of trustworthy immune-related biomarkers (Pan et al., 2019; Chen et al., 2020). Immune-related cells infiltration into the tumor microenvironment (TME) is a key reason leading to immune responses at primary and secondary tumor sites, which is tightly regulated by various mediators, such as chemokines. Several studies have indicated that a variety of different immune cells, including $CD4^+$and $CD8^+$T -cells, dendritic cells, and tumor-associated macrophages (TAMs), have been identified in different cancers, such as prostate cancer, HCC and others (Pan et al., 2019; Zhang Y. et al., 2020). Chen et al. (2015) also demonstrated that $CD4^+$and $CD8^+$T-cells could be recruited into the TME after CXCR4 inhibition in sorafenib-treated HCC in a mouse model. Therefore, there is an impending requirement to identify potential predictors related to immune cell infiltration to enhance the efficacy of individual immunotherapeutic treatment in tumors.

Study design is recapitulated in **Figure 1**. First, differentially expressed RNAs were analyzed in 371 cases of HCC and 50 normal liver tissues from The Cancer Genome Atlas (TCGA). Subsequently, a nomogram predicted model based on 23 miRNAs was established and revealed high performance. Next, we constructed a ceRNA network composed of 56 DELs, 6 DEMs, and 28 DEGs to illustrate preliminary interactions between mRNAs and ncRNAs. DEGs correlating to the ceRNA network were submitted for Gene Ontology (GO) and pathway enrichment analysis to clarify the underlying molecular mechanism in HCC. Finally, six hub genes, including CEP55, DEPDC1, CLSPN, KIF23, MYBL2, and RACGAP1, were identified by protein-to-protein (PPI) analysis and were all overexpressed in HCC samples with significant poor prognosis in HCC patients. Meanwhile they were also closely associated with immune infiltration in HCC. In summary, we believe these genes represent potential prognostic markers, immune-related biomarkers and immunological therapeutic targets for HCC treatment that deserved to be further explored in the future.

Abbreviations: lncRNAs, long non-coding RNAs; ceRNAs, competing endogenous RNAs; HCC, hepatocellular carcinoma; RFS, recurrence-free survival; OS, overall survival; TCGA, the Cancer Genome Atlas; DElncRNAs, differentially expressed lncRNA; DEmiRNAs, differentially expressed miRNA; DEmRNAs, differentially expressed mRNA; MREs, miRNA-response elements; GO, gene ontology; ROC, receiver operating characteristic; BP, biological process; CC, cellular component; MF, molecular function; EMT, Epithelial-mesenchymal transformation; RS, risk score.

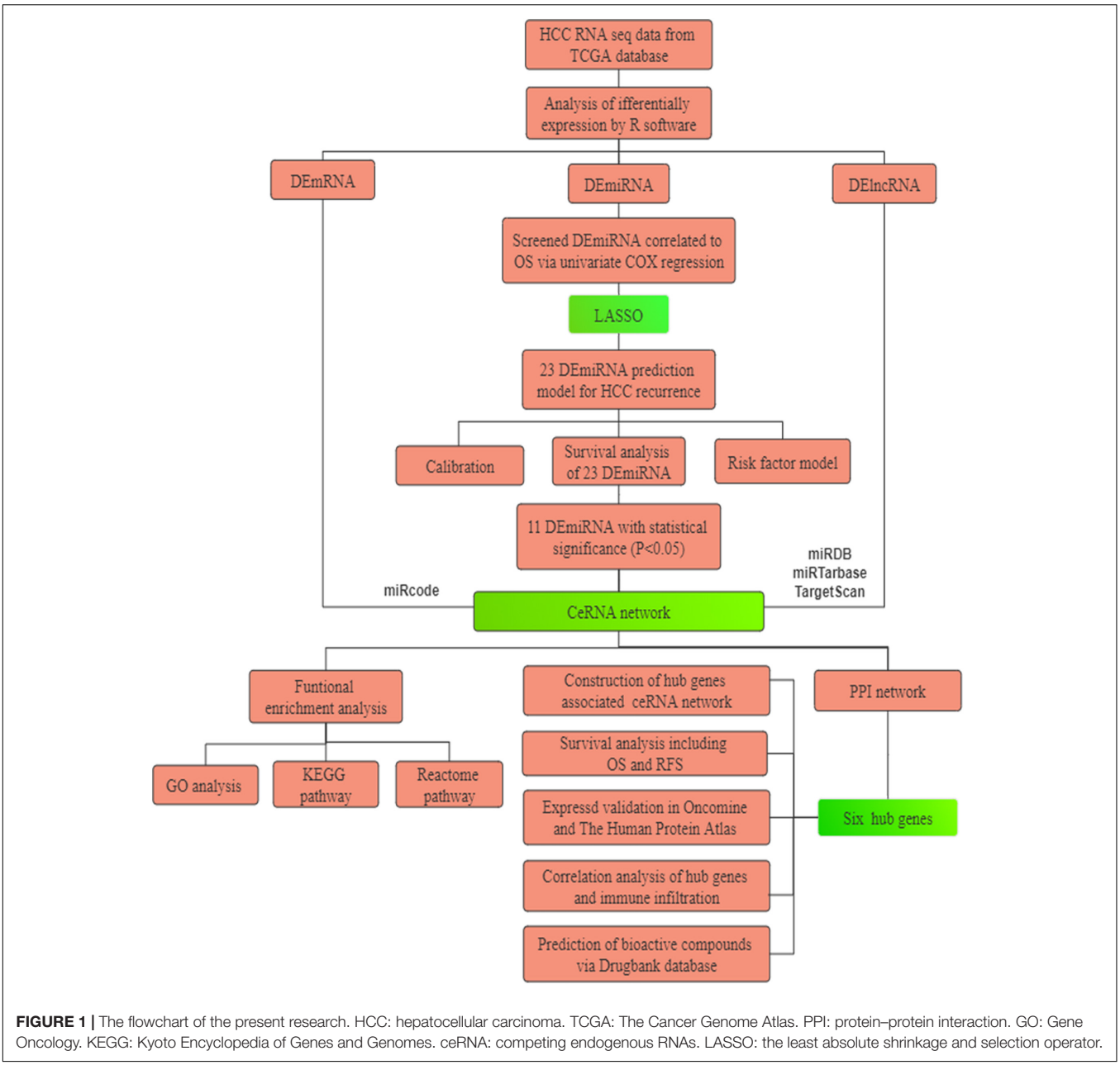


FIGURE 1 | The flowchart of the present research. HCC: hepatocellular carcinoma. TCGA: The Cancer Genome Atlas. PPI: protein–protein interaction. GO: Gene Oncology. KEGG: Kyoto Encyclopedia of Genes and Genomes. ceRNA: competing endogenous RNAs. LASSO: the least absolute shrinkage and selection operator.

MATERIALS AND METHODS

Data Acquisition From TCGA Database

A total of 371 HCC patients were included in this study. RNA sequencing (RNA-Seq), including lncRNA, mRNA (Illumina HiSeq RNA-Seq platform), and miRNA sequence data (Illumina HiSeq miRNA-Seq platform) from 371 HCC samples with survival data and 50 adjacent non-tumorous samples including the corresponding paired 50 HCC samples were downloaded from TCGA database ([1]version 10.1, release time: May 15, 2019). Approval was not required by ethics committee, and the present study conformed to the publication guidelines by TCGA. The RNA sequence data were annotated based on the Ensemble gene ID. Log_2 transformation was performed on all gene expression profiling. Then, Limma package (Version: 3.38.3) in R software (Version:3.5.2) was used to normalize the original data.

Identification of Differentially Expressed (DE) RNAs

The expression profile of RNA sequencing data retrieved from TCGA was analyzed using the limma package of R software[2] with the criterial of $|\log_2$ fold change$| > 2$ and the adjusted

[1]https://portal.gdc.cancer.gov/

[2]https://www.r-project.org/

false discovery rate (FDR) of $P < 0.05$. Screened DE RNAs, including differentially expressed lncRNAs (DELs), differentially expressed miRNAs (DEMs), and differentially expressed mRNAs (DEGs), were used for subsequent analysis. Heat maps and volcano plots for DE RNAs were created using the heatmap package of R software.

Univariate and Multivariate Cox Regression Analysis

Univariate Cox regression was used to screen for potential prognostic miRNAs correlated with overall survival in HCC patients. The least absolute shrinkage and selection operator (LASSO) detects the most influential variables because it analyzes all independent variables simultaneously. According to the principle of a penalty following a regularization path, the coefficients of less influential variables would trend toward zero. The glmnet package was used to perform LASSO algorithm with the criterial of $P < 0.05$. Multivariate Cox regression analysis via survival R package was utilized to establish a prognostic predictive model visualized by nomogram to show the correlation of the expression of DEMs and survival rate of specific HCC patients. The forest plot was created to display the results of multivariate Cox regression using the forestplot package.

Evaluation of miRNA-Based Clinical Predictive Model

To evaluate the predictive performance of the prognostic model based on DEMs, first, a calibration curve of 3- and 5-year survival rates was determined to assess agreement between the predictive model and actual survival time. Moreover, the area under the curve (AUC) was calculated according to the time-dependent receiver operating characteristic analysis (ROC). Additionally, the risk score formula was performed to calculate total risk scores for individual patients based on the coefficient for each DEM. The risk score formula was built according to the following method: total risk score = sum of each coefficient × transcriptional expressed value of DEM. Then, HCC patients were divided into high- and low-risk groups by the median risk scores, regarded as the cutoff value. The difference in survival rate between the two groups was also evaluated. The correlation between expression levels of DEMs and OS in HCC patients was calculated by Kaplan-Meier (K-M) survival analysis using the survival package of R software according to the X tile method with a cutoff P-value < 0.05.

Establishment of the ceRNA Regulatory Network

A co-expressed regulatory network comprised of DELs, DEMs, and DEGs was established to explore the potential functions of these DE RNAs in HCC. The interaction between DEMs and DELs was confirmed using the miRcode database[3], which not only includes putative target sites of miRNAs from the integrated and searchable map but also contains conserved microRNA families annotated by the ENCyclopedia of DNA Elements (ENCODE) (Jeggari et al., 2012). DEM targets were predicted from three databases, including miRDB[4], miRTarBase[5], and TargetScan[6](Agarwal et al., 2015; Wong and Wang, 2015; Chou et al., 2018). Overlapping DEGs were selected for constructing the ceRNA network. Cytoscape[7] software was used to visualize the expression correlation of DE RNAs.

Functional Annotation and PPI Network Analysis

Gene Ontology (GO) analysis of differentially expressed genes, including biological process (BP), molecular function (MF) and cell components (CC), and Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis, were enriched using the Enrich online tool[8]. The Reactome pathway was determined by the Reactome website[9]. The online STRING[10] tool was used to construct the protein-to-protein (PPI) interaction network for DEGs involved in the ceRNA network and was visualized by Cytoscape. Hub genes were defined as the top six genes with the highest degree of connections to others via the CytoHubba plug-in of Cytoscape.

Correlation of Hub Genes and Immune Infiltration Analysis

TIMMER is a comprehensive online database used for systematic analysis of the correlation of immune infiltration and gene markers of interest in 32 cancer types from TCGA[11]. The abundance of tumor-infiltrating immune cells (TIICs) from gene profiles is evaluated based on the statistical method of deconvolution published previously (Aran et al., 2015; Li et al., 2016). We analyzed expressed levels of hub genes in various tumors and the relationship between the expression of hub genes and immune infiltration, including B cells, CD4 + T cells, CD8 + T cells, neutrophils, macrophages, and dendritic cells. Furthermore, the correlation of hub gene expression and gene markers associated with tumor immune infiltration of monocytes, tumor associated macrophages (TAMs), M1 and M2 macrophages was performed via correlation modules. These gene markers were identified in previous studies (Sousa and Maatta, 2016; Danaher et al., 2017; Siemers et al., 2017). The strength of the correlation was evaluated by Spearman's algorithm divided into five levels: very weak (0.00-0.19), weak (0.20-0.39), moderate (0.40-0.59), strong (0.60-0.79), and very strong (0.80-1.0). The cutoff criteria for statistical significance was P-value < 0.05. In addition, the SCNA module was used to identify differences in tumor infiltration of hub genes in HCC with different somatic copy number alterations.

[3]http://www.mircode.org/

[4]http://www.mirdb.org/

[5]http://mirtarbase.mbc.nctu.edu.tw

[6]http://www.targetscan.org

[7]https://cytoscape.org/

[8]http://www.enrichnet.org/

[9]https://reactome.org/

[10]https://string-db.org/

[11]https://cistrome.shinyapps.io/timer/

RESULTS

Identification of Differentially Expressed RNAs

A total of 371 HCC samples and 50 non-tumor tissues were included in this study. Differentially expressed RNAs were analyzed using the limma R package with a screening cutoff threshold of $|\log_2\text{-fold change}| > 2$ and an adjusted P-value < 0.05. 1999 DEGs, 251 DEMs, and 1092 DELs were identified as differentially expressed RNAs. Among them, there were 1794 DEGs, 229 DEMs, and 1034 DELs upregulated. Moreover, differential expression of DEGs, DEMs and DELs is displayed by hierarchical clustering and volcano plots (**Figure 2**).

Establishment of miRNA-Based Prognostic Predictive Model in HCC

A total of 42 DEMs survival-related miRNAs were screened by univariate Cox regression from 251 DEMs identified in the HCC cohort (**Supplementary Tables 1**, **2**). Next, these miRNAs were included in the LASSO analysis to calculate the corresponding coefficients. Twenty-three DEMs significantly correlated with survival were selected out (**Supplementary Figure 1** and **Supplementary Table 3**). A simple-to-use nomogram predictive model was established to describe correlation of the expression of each miRNA and the 3- and 5-year overall survival rate of HCC patients based on multivariate Cox regression (**Figure 3**). Meanwhile, the 3- and 5-year calibration curves were drawn, showing good consistency between predicted survival probability and the actual survival rate (**Figures 4A,B**). The ROC curve also exhibited great reliability for the nomogram prediction model in discriminating tumors from normal tissues with the area under curve (AUC) of 3- and 5-year being 0.804 and 0.744, respectively (**Figure 4C**). Furthermore, the risk score (RS) of each patient in the HCC cohort was calculated, and the patients were subsequently divided into high-risk and low-risk groups according to the mean RS. K-M analysis indicated that patients in the high-risk group exhibited decreased survival compared to the low-risk group, with a log-rank P-value < 0.05 (**Figure 4D**).

Construction of the lncRNA-miRNA-mRNA Regulatory Network

K-M analysis performed for 23 DEMs significantly related with OS revealed that 11 DEMs, including 8 that were upregulated, were independently statistically significant (**Supplementary Table 4**). Results showed that HCC patients with highly expressed DEMs had a shorter survival time than those with lower expression with a P-value < 0.05 (**Supplementary Figure 2**). Recently, increasing evidence has indicated that miRNAs play a significant role in the development and metastasis of tumors. To reveal potential signaling pathways regulated by these DEMs in HCC, the DIANA-miRPath database was exploited and revealed enrichment of cancer-related signaling pathways, such as PI3K-AKT, NF-kappa B, VEGF, and others (**Supplementary Figure 3**).

The lncRNA targeted by the 11 DEMs was screened based on the interactions with 1092 DELs aforementioned. Six DEMs were targeted by 56 DELs according to the miRcode database. Next, targets of the 6 DEMs were predicted using miRTarBas, miRDB, and TargetScan databases. The overlapping 574 mRNAs predicted in all three databases were further intersected with 1999 DEGs identified in the HCC cohort, and only 28 DEGs existed in both groups (**Supplementary Figure 4** and **Supplementary Table 5**). The representative interactions among 56 DELs, 6 DEMs, and 28 DEGs are summarized in **Supplementary Table 6**, and the ceRNA regulatory network, including gene nodes and preliminary interactions, was visualized using Cytoscape software (**Figure 5A**).

Function and Pathway Enrichment Analysis of DEGs Involved in the ceRNA Network

Gene ontology and pathway enrichment analyses were performed to elucidate the functions of 28 DEGs in the ceRNA network correlated with the progression of HCC. Functional annotation of biological process (BP), cellular component (CC), and molecular function (MF), as well as KEGG pathway enrichment, were performed on the Enrichr comprehensive database, in which the top 10 highly enriched items for BP, CC, MF, and KEGG pathway are shown based on a P-value < 0.05 (**Figures 5B-E**). Notably, all top 10 items were closely related to cancer-related pathways, such as TNF signaling, breast cancer, small cell lung cancer and others. Moreover, Reactome pathway analysis was also developed to identify possible metabolic pathways in which the 28 DEGs are involved. A total of 28 pathways were identified, and the top 15 highly enriched pathways are presented in **Figure 5F**.

Construction of PPI Network and Identification of Hub Genes

To further investigate the function of 28 DEGs associated with the ceRNA network at the protein level, we established a protein-to-protein interaction (PPI) network composed of 109 nodes and 218 degrees to visualize detailed interactions (**Figure 6A**). Considering the significance of hub genes in the ceRNA network, the CytoHubba plugin in Cytoscape software was exploited to identify hub genes by evaluating the number of degrees and connections. Finally, six hub genes, CEP55, DEPDC1, MYBL2, RACGAP1, CLSPN, and KIF23, were identified, which were all upregulated in HCC cohort (**Figure 6B**). The filled color of nodes from red to yellow indicates the degree of connectivity of hub genes with others gradually decreases. GO and pathway enrichment, including KEGG and Reactome analyses, were also performed (**Supplementary Figure 5**). Additionally, the sub ceRNA network, including 46 DELs, 3 DEMs (has-mir-30d, hasa-mir-195, has-mir-301a), and 6 hub genes, was built to delineate correlations among the DELs, DEMs, and hub genes (**Figure 6C** and **Supplementary Table 7**).

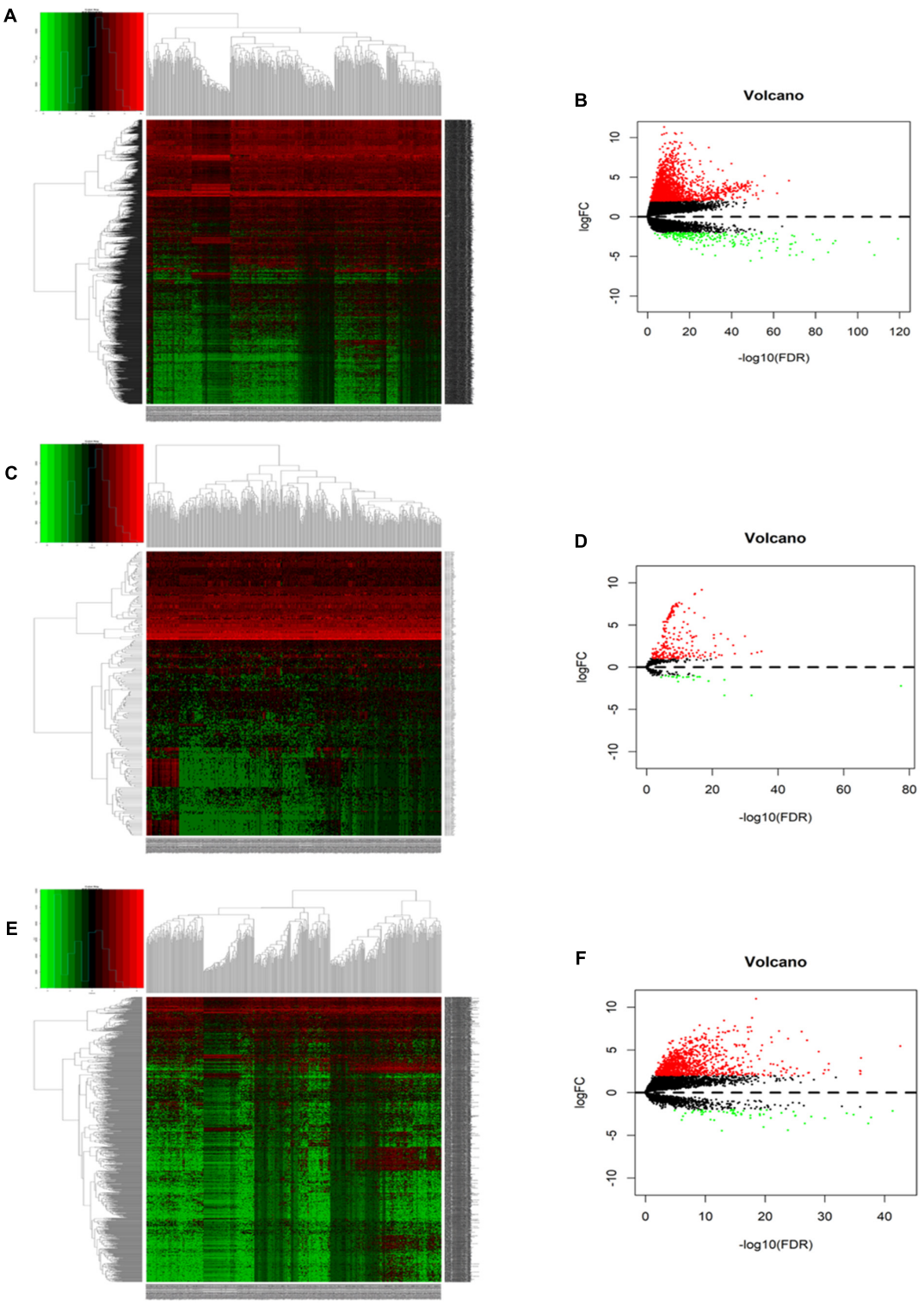


FIGURE 2 | The hierarchical clustering heatmaps and volcano plots for all screened differentially expressed mRNA, miRNA and lncRNA in HCC based on TCGA data. Heatmaps located in the left panels represent differential expressed (DE) mRNAs **(A)**, miRNAs **(C)**, and lncRNAs **(E)**. Volcano plots located in the right panels indicate DEmRNAs **(B)**, DEmiRNAs **(D)**, and DElncRNAs **(F)** with the cutoff criteria of fold change ≥ 2 and P-value < 0.05. Red color: upregulated, green: downregulated, gray: not statistic expressed.

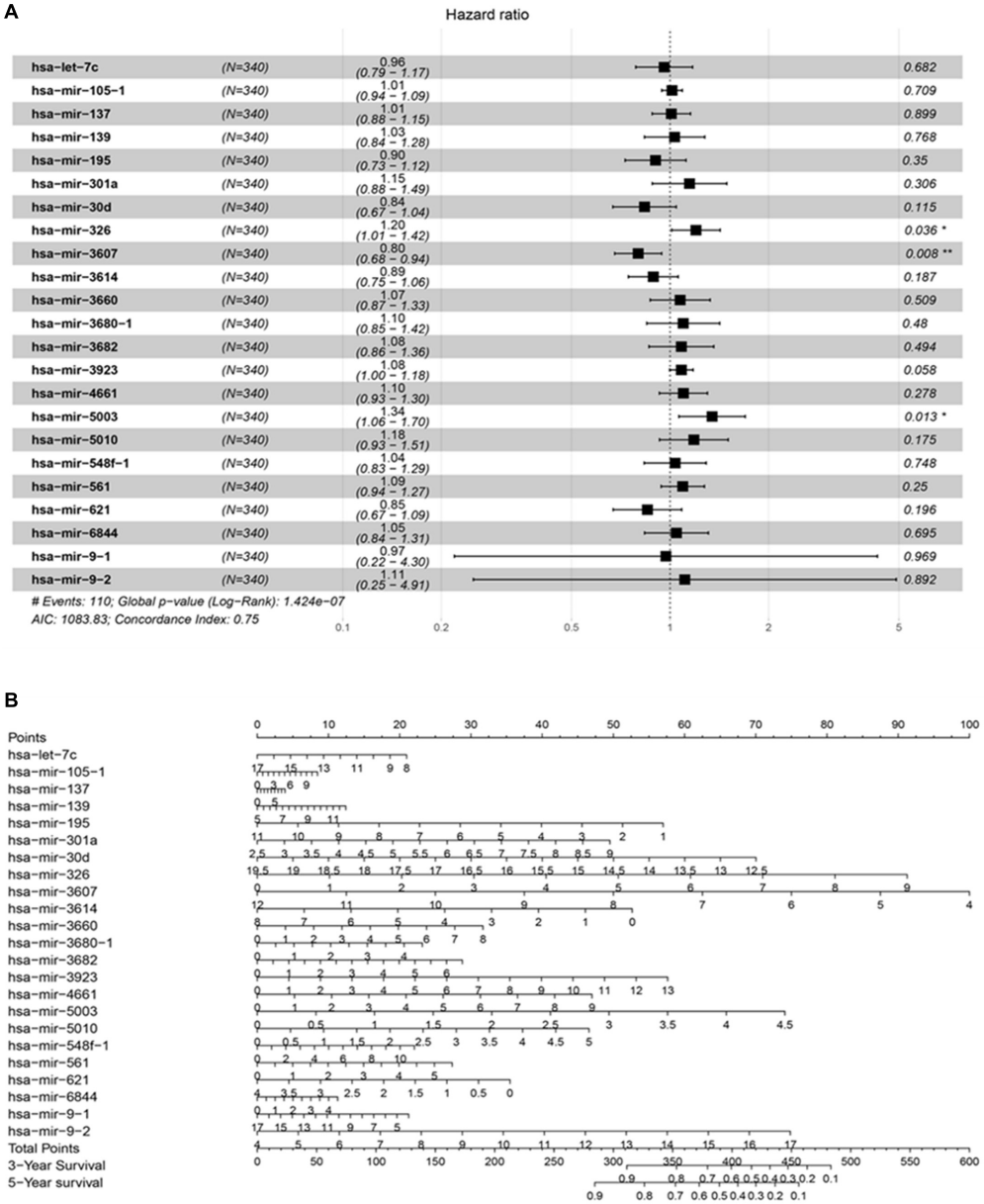


FIGURE 3 | The prognostic predictive model for HCC. **(A)** The analyzed result of multivariate Cox proportional hazards regression involved with recurrence in TCGA HCC cohort. The middle point of the line indicates the hazard ratio (HR), and the whole length on behalf of the 95% CI for each DEmiRNA. **(B)** Nomogram based on differentially expressed miRNAs to predict survival in HCC asymptomatic individuals. The prognostic model aim to estimate the survival rate for individual patient, meanwhile reveal the upregulated or downregulated type for each miRNA. At first, draw a line straight upwards from each miRNA to obtain the points from the points axis. Repeat this step until the total scores were gained for 23 miRNAs. Then, after calculating the overall points according to the total points axis, draw a line straight down to the 3-year and 5-year survival axis from the location of total point axis based on the obtained overall scores to indicate the rate for the specific patient (for e.g., the 3-year survival rate is 60% if a patient get the total points of 400).

Validation of Survival Analysis and Expression of Hub Genes

The correlation between expression of hub genes and OS (**Supplementary Figures 6A-F**) and recurrence-free survival (RFS) (**Supplementary Figures 6G-L**) was assessed using the GEPIA website. HCC patients with high expression levels of hub genes had a lower survival rate with respect to both OS and RFS. Among these, CEP55 had the highest prognostic *P*-value (0.00033 and 0.00063) correlated to overall and recurrence-free survival rates, respectively. Meanwhile, TISIDB tool that integrated five comprehensive public databases of HGNC, NCBI, Ensembl, Uniprot, and GeneCards was also exploited to evaluate the correlation to further confirm (Ru et al., 2019). It showed us consistent result that six hub genes were independently

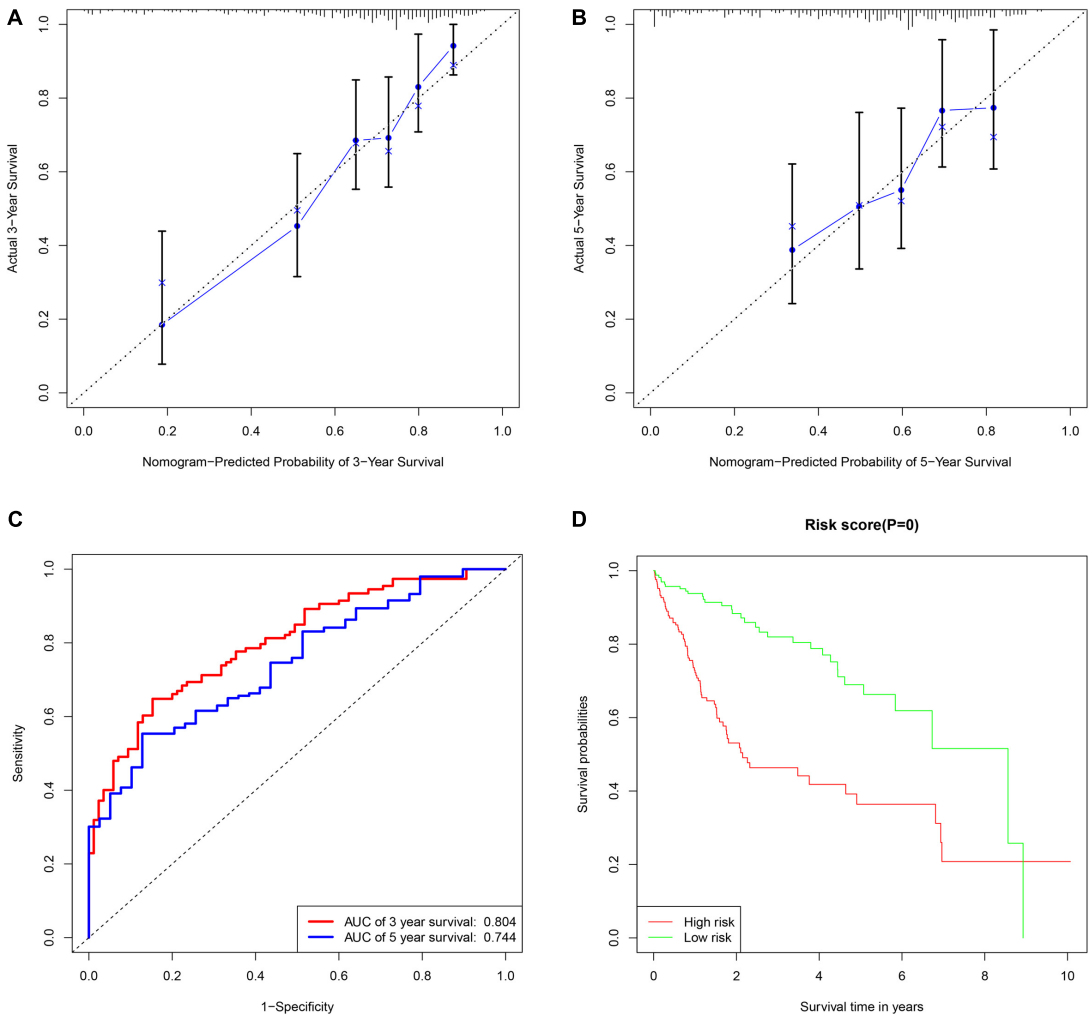


FIGURE 4 | The assessment of miRNAs-based clinical prediction model. The calibration curves according to nomogram model to estimate the survival rate at 3-year **(A)** and 5-year **(B)**, the *X* and *Y*-axis represent predicted and actual survival time respectively. The efficacy of prognostic model of 3- and 5-years survival rate based on time dependent receiver operated characteristic curves **(C)**. The Kaplan-Meier curve of overall survival time between the high- and low- risk groups stratified by the mean of total risk scores **(D)**.

associated with poor prognosis in HCC patients with statistical significance (**Supplementary Figures 6M–R**). mRNA levels of hub genes in various tumor and normal tissues were examined using the TIMMER database. Results indicated that expression of hub genes was higher in various tumors than in corresponding normal tissues including breast cancer, colorectal cancer, gastric cancer and others (**Supplementary Figure 7**). Next, we validated transcriptional expression of hub genes in another HCC cohort from Oncomine database, which demonstrated overexpression of genes in the tumor group compared to non-tumor tissues as well with cutoff *P*-value < 0.01 (**Figures 7A–F**). Moreover, the mRNA expression of hub genes was detected in MHCC-LM3 and Huh7 HCC cell lines, and L02 normal liver cell line by quantitative polymerase chain reaction (qPCR), which showed the same conclusion with Oncomine database (**Figures 7G–L**). Additionally, immunohistochemical data from the Human Protein Atlas was used to verify protein expression of hub genes. Data for CLSPN was lacking in the database, and expression of DEPDC1 in both HCC and normal samples was not detected. However, the staining intensity or the range of positive areas of CEP55, KIF23, MYBL2, and RACGAP1 was higher in tumor samples (medium or high levels) than in non-tumor tissue (**Figures 7M,N**).

Correlation of Hub Gene Expression and Immune Infiltration

Recently, increasing evidence has demonstrated that tumor-infiltrating lymphocytes play a significant role in predicting lymph node status and survival in tumors. Therefore, we investigated whether expression of hub genes was related to immune-infiltrating levels in HCC using the TIMMER database. Results revealed that six hub genes were significantly related to immune infiltration of B cells, $CD4^+$ T cells, $CD8^+$T cells,

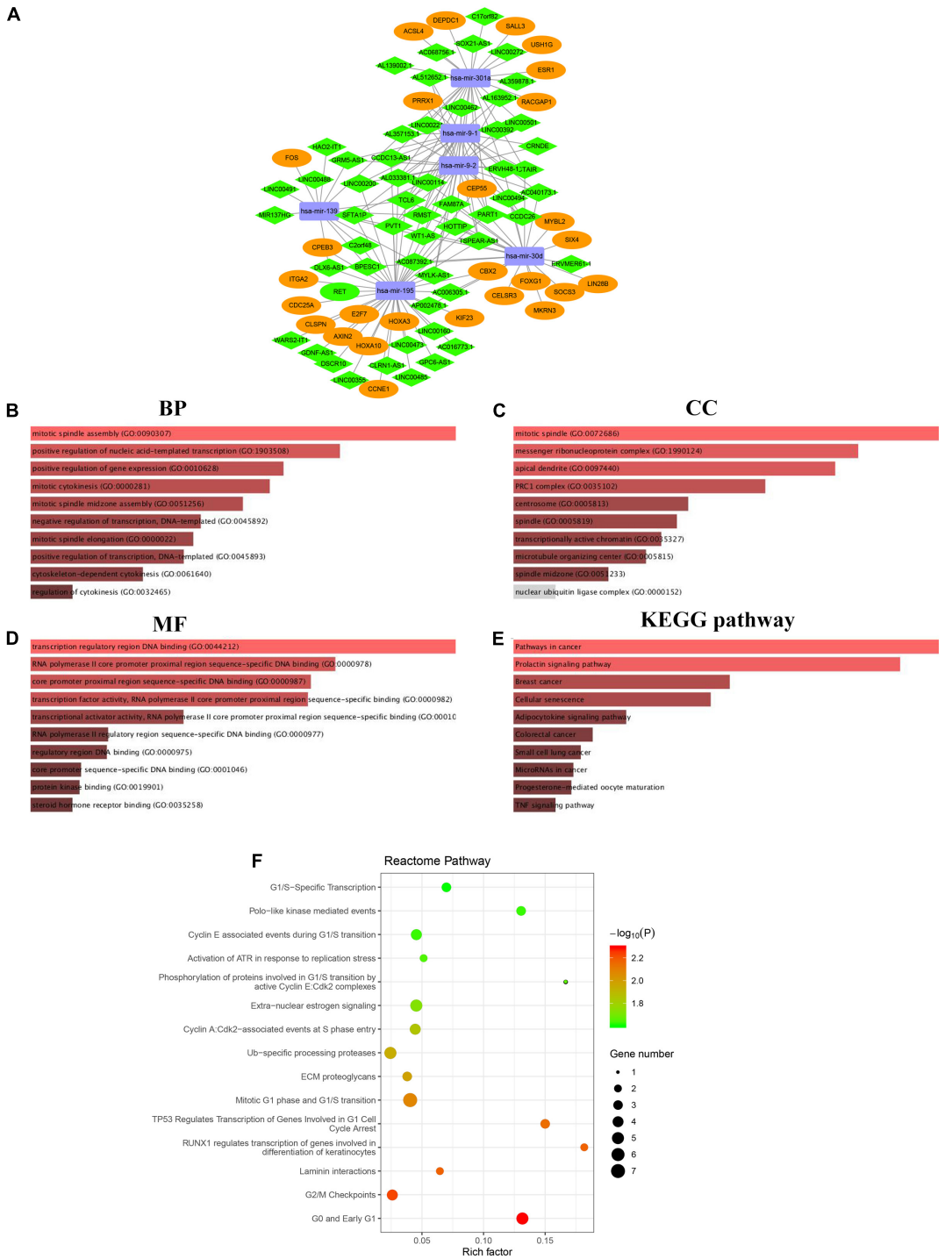


FIGURE 5 | The lncRNA-miRNA-mRNA regulatory network in HCC cohort visualized by Cytoscape software 3.6.1 **(A).** Rectangle represent the 6 DEMs, diamond represent the 56 DELs, the ellipse represent the 28 DEGs. The functional enrichment analysis of DEGs correlated to ceRNA network **(B–F).** Top 10 biological process (BP) terms **(B)**. Top 10 cell components (CC) terms **(C)**. Top 10 molecular functions (MF) terms **(D)**. Top 10 significantly KEGG pathways **(E)**. Top 15 enriched Reactome pathways **(F)**.

macrophages, neutrophils, and dendritic cells in HCC, with statistical *P*-values of <0.0001. However, expression of CLSPN (cor = 0.08, *P* = 1.39e-01), CEP55 (cor = 0.018, *P* = 7.33e-01), and MYBL2 (cor = 0.098, *P* = 6.97e-02) had no significant correlation with tumor purity (**Supplementary Figure 8**). Next, to further investigate the correlation between expression of six hub genes

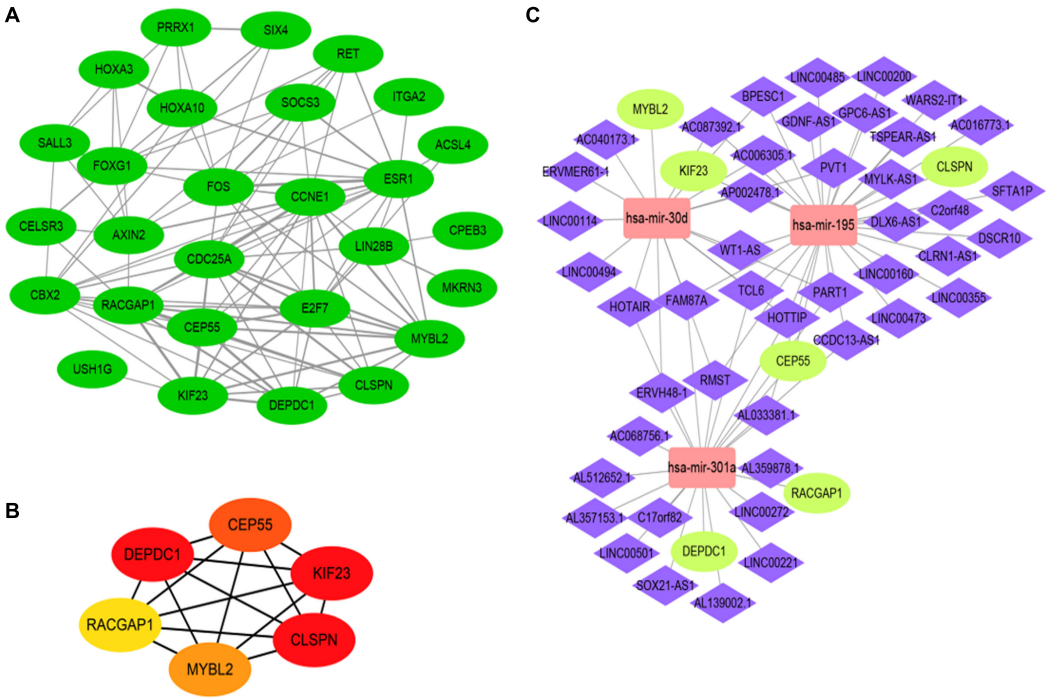


FIGURE 6 | The construction of hub gene associated with ceRNA network based on analysis of protein to protein interaction (PPI) network. The PPI network of DEGs **(A)**. PPI network of 6 hub genes, the color of nodes from red to yellow indicates that the connected degrees between each molecule with others decrease gradually **(B)**. the hub gene ceRNA regulatory network including 3 DEMs (rectangle) and 6 hub DEGs (ellipse) as well as 46 lncRNA (diamond) **(C)**.

and diverse immune infiltrating cells, the relationship between hub genes and a series of immune-related markers in various immune cells of HCC in TIMER database was focused. The analyzed immune markers included $CD8^+$T cell, T cell (general), B cell, monocyte, tumor-associated macrophage (TAM), M1 Macrophage, M2 Macrophage, Neutrophils, Dendritic cell, Treg (**Table 1**). The results revealed that expressed levels of six hub genes were significantly related to most immune markers of various immune cell types. Importantly, we found that the expressed level of most markers of monocytes, TAM, M1and M2 macrophages showed strong correlation with hub genes of CEP55, CLSPN, DEPDC1, KIF23, and RACGAP1 except for MYBL2 which had no distinct relationship with gene markers of PTGS2, CD163, VSIG4, and MS4A4A (**Table 1**). Specially, we displayed that CD86, CSF1R of monocyte, CD68, IL10 of TAM, IRF5, PTSG2 of M1 macrophage, CD163, VSIG4, MS4A4A of M2 macrophage are significantly associated with hub genes in HCC (**Figure 8**). Furthermore, we also analyzed the correlation between hub genes and immune markers mentioned above of monocyte, TAM, M1 and M2 macrophage in HCC from GEPIA database. The correlation results are similar to those in TMER (**Table 2**). These findings suggest that hub genes are correlated with immune infiltration and may regulate macrophage polarization in HCC. Additionally, whether immune infiltrating levels of each immune subset are related to differential copy number of hub genes was also evaluated. However, no significant relationship was observed between most immune cells and hub genes (**Supplementary Figure 9**).

Identification of Bioactive Compounds Targeting Hub Genes

Finally, we predicted potential bioactive compounds targeting hub genes using the Drugbank database, which is a comprehensive, freely accessible, online database including both drugs and drug target information. A total of 15 compounds targeting hub genes were identified, including CEP55 (Irdabisant, CEP-9722, CEP-1347, CEP-37440, Cefapirin), CLSPN (Calfactant, Calusterone), and MYBL2 (Clotrimazole, Propafenone, Letrozole, Sildenafil, Ranitidine, Valproic acid, Esomeprazole, and Pregabalin) (**Table 3**). Except for Calfactant, the 3D chemical structure of the other compounds is presented in **Supplementary Figure 10**. These results could provide new insight into potential novel therapeutic targets for HCC in the future.

DISCUSSION

Hepatocellular carcinoma is one of the most common malignant tumors in the world. In recent years, the prevalence of HCC is gradually increasing, especially in non-traditional high incidence areas, such as the United States and Europe. Most HCC patients are likely to be diagnosed at an advanced stage since HCC is asymptomatic at early stages, and effective biomarkers for early diagnosis and prognostic prediction are lacking (Lei et al., 2019; Li et al., 2019b; Liu et al., 2019). Currently, the main treatments for HCC include radiofrequency ablation, surgical

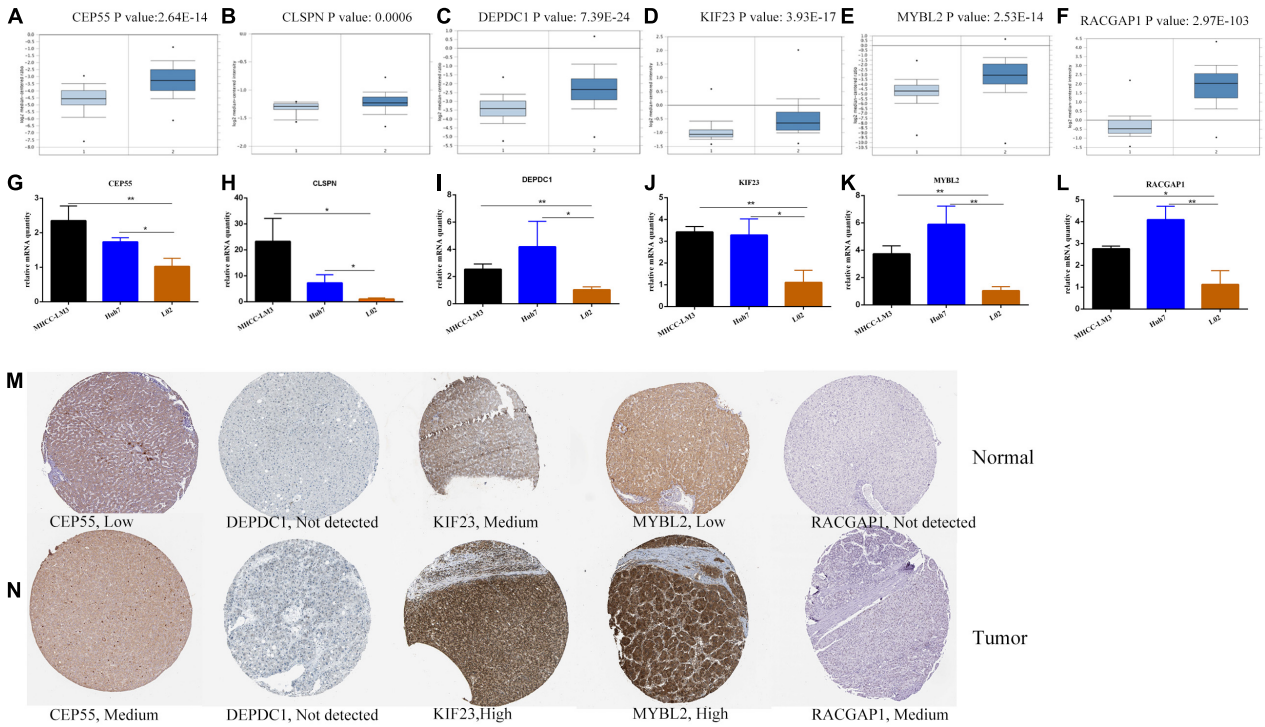


FIGURE 7 | The transcriptional expression of six hub genes in HCC (right panels) and normal liver samples (left panels) based on Oncomine database **(A–F)**. The relative mRNA expression of six hub genes in HCC cell lines and liver cell line **(G–L).**$^{*}P < 0.05$, $^{**}P < 0.01$. The protein expression of hub genes based on Human Protein Atlas database **(M–N).** The protein expressed level of DEPDC1 was not detected in both tumor and normal liver tissues and the expressed data of CLSPN was lacking in database.

resection, immunotherapy, and liver implantation, among others. However, the clinical efficacy of treatments for specific patients is not satisfactory due to poor therapeutic targets, tumor immune escape and complications, leading to a low 5-year survival rate (Yue et al., 2019; Zhang et al., 2019; Zhang Y. et al., 2020). In past decades, studies focused on immunotherapy for various cancers have obtained meaningful breakthroughs, especially in melanoma, non-small cell lung cancer, and others. Different types of immune cells infiltrate into the tumor microenvironment, which is a crucial reason for effective immune responses (Pan et al., 2019; Chen et al., 2020). It has been reported that immune cells, including natural killer cells, $CD4^{+}$ and $CD8^{+}$T-cells, TAMs, and dendritic cells, and others, are detected in cancer tissues, including HCC. Furthermore, these immune infiltrating cells are regulated by various mediators, such as chemokines. However, the mechanism of immune infiltration in the development of tumors is not completely understood (Sousa and Maatta, 2016; Siemers et al., 2017). Therefore, it is urgent and significant to elucidate the molecular mechanism and to identify immune-related signatures in HCC, which will aid in identifying new therapeutic targets and prognostic markers to increase the clinical efficacy and 5-year survival rate of HCC patients.

Currently, the ceRNA hypothesis of crosstalk between ncRNAs and mRNAs has received much attention and is considered a new measure of gene regulation at the posttranscriptional level, which provides new insight into revealing mechanisms of tumorigenesis and identifying potential diagnostic and prognostic biomarkers. A growing number of published studies have demonstrated that many predictive signatures are detected in various tumors based on ceRNA network analysis (Xiong et al., 2018; Wang M. et al., 2019; Xu et al., 2019; Wu et al., 2020; Xiao, 2020). MiRNAs, included in ncRNAs, are identified to be evolutionarily conserved, with an average length of 22-nt, and may bind to the 3′ untranslated region (3′UTR) of the targeted mRNAs according to the principle of complementary base pairing. An increasing body of evidence has demonstrated that dysregulated miRNAs play a crucial role in the initiation, progression, and therapy of various tumors (Long et al., 2019; Lou et al., 2019; Zhang et al., 2019). Baolei et al. revealed that miRNA-124 is a negative regulator of HCC with respect to proliferation and invasion by downregulating lncRNA-UCA1 (Zhao B. et al., 2019). Baltruskeviciene et al. (2017) found that downregulated expression of miRNA-148a and miRNA-625-3p is related to tumor budding in colorectal cancer, and EMT was considered a possible molecular mechanism.

In the present study, a prognostic predictive model was established based on 23 DEMs and exhibited great performance with the area under ROC, which was 0.804 for 3-year and 0.744 for 5-year survival. Moreover, we constructed a ceRNA network, including 56 DELs, 6 DEMs (hsa-mir-9-1,hsa-mir-9-2, hsa-mir-30d, hsa-mir-139, hsa-mir-195, hsa-mir-301a), and 28 DEGs, which identified several potential prognostic

TABLE 1 | The correlation between six hub genes and related genes and markers of immune cells for HCC in TIMER.

Immune cell	Gene marker	CEP55		CLSPN		DEPDC1		KIF23		MYBL2		RACGAP1	
		Cor	P	Cor	P	Cor	P	Cor	P	Cor	P	Cor	P
$CD8^+$ T cell	CD8A	0.297	****	0.221	****	0.185	***	0.227	****	0.211	****	0.152	**
	CD8B	0.267	****	0.159	**	0.174	***	0.187	***	0.235	****	0.115	*
T cell	CD3D	0.372	****	0.209	****	0.24	***	0.273	****	0.365	****	0.152	**
(general)	CD3E	0.315	****	0.194	***	0.164	**	0.219	****	0.233	****	0.109	*
	CD2	0.323	****	0.197	***	0.171	***	0.23	****	0.254	****	0.104	*
B cell	CD19	0.32	****	0.237	****	0.238	****	0.263	****	0.309	****	0.221	****
	CD79A	0.279	****	0.14	**	0.108	*	0.16	**	0.201	****	0.061	0.238
Monocyte	CD86	0.438	****	0.417	****	0.298	****	0.358	****	321	****	0.291	****
	CSF1R	0.294	****	0.295	****	0.155	**	0.213	****	0.169	**	0.177	***
TAM	CD68	0.349	****	0.334	****	0.179	***	0.25	****	0.27	****	0.239	****
	IL10	0.328	****	0.338	****	0.236	****	0.269	****	0.229	****	0.225	****
M1 Macrophage	IRF5	0.451	****	0.43	****	0.346	****	0.472	****	0.372	****	0.403	****
	PTGS2	0.227	****	0.261	****	0.082	0.114	0.188	***	0.062	0.231	0.147	**
M2 Macrophage	CD163	0.154	**	0.286	****	0.118	*	0.122	*	0.046	0.372	0.138	**
	VSIG4	0.198	***	0.269	****	0.129	*	0.142	*	0.086	0.979	0.12	*
	MS4A4A	0.205	****	0.262	****	0.124	*	0.153	**	0.073	0.162	0.125	*
Neutrophils	ITGAM	0.403	****	0.457	****	0.324	****	0.369	****	0.284	****	0.312	****
	CCR7	0.192	***	0.163	**	0.05	0.341	0.11	*	0.078	0.133	0.056	0.285
Dendritic cell	HLA-DPB1	0.278	****	0.243	****	0.15	**	0.209	****	0.187	***	0.16	**
	HLA-DQB1	0.248	****	0.195	***	0.144	**	0.167	**	0.164	**	0.099	*
	HLA-DRA	0.209	****	0.303	****	0.188	***	0.23	****	0.166	**	0.201	****
	CD1C	0.205	****	0.171	***	0.063	0.23	0.156	**	0.116	*	0.115	*
	NRP1	0.319	****	0.417	****	0.185	***	0.324	****	0.118	*	0.442	**
	ITGAX	0.464	****	0.4444	****	0.322	****	0.404	*	0.35	****	0.314	****
Treg	FOXP3	0.201	***	0.338	****	0.236	****	0.222	****	0.094	0.0695	0.19	***
	CCR8	0.481	****	0.541	****	0.404	****	0.459	****	0.33	****	0.403	****
	STAT5B	0.203	****	0.49	****	0.266	*****	0.332	****	0.11	*	0.459	**
	TGFβ	0.418	****	0.307	****	0.195	***	0.334	****	0.308	****	0.271	****

*Annotation: Cor, R-value of Spearman's correlation; *P < 0.05; **P < 0.01; ***P < 0.001; and ****P < 0.0001.*

signatures for HCC. The finally screened 6 DEMs by ceRNA network were independently predicted factors related to poor prognosis in HCC patients. The major pathways that 6 DEMs participated were enriched in cancer-related regulatory signals. These findings here maybe give us a new avenue to early detect HCC patients, and evaluate clinical prognosis before or after receiving treatment via detecting the expressed level of specific DEMs mentioned above. U Lehmann et al. reported that hypermethylation of hsa-mir-9-1 is related to the development of breast cancer. Patients with pre-invasive intraductal lesions were detected by hypermethylated hsa-mir-9-1 (Lehmann et al., 2008). Notably, hsa-mir-195 belongs to the miR-195 family and is located on chromosome 17p13.1, and is correlated with proliferation and angiogenesis in prostate tumors by downregulating expression of the PRR11 gene (Cai et al., 2018). In Sannigrahi et al. (2017) research, they indicated that hsa-mir-139 was downregulated by HPV-16, leading to activation of HPV-16 oncogenic pathways and carcinogenesis of HPV-16 induced cervical and head and neck cancers. Hsa-mir-301a was found to play an oncogenic role in the occurrence and development of laryngeal squamous cell carcinoma (LSCC) by directly targeting the tumor suppressor gene Smad44 and downregulating its expression (Lu et al., 2015). Through analysis of SNP-array data generated from 8 medulloblastoma cell lines, Lu et al. (2009) found that hsa-mir-30d is overexpressed, however, the potentially involved biological processes and molecular mechanisms are still not clarified. Therefore, according to current studies, it has revealed that the 6 DEMs exhibited various biological functions in several cancer types, but the research progress is still limited especially in HCC. Further studies focused on both functions and molecular mechanisms in HCC were necessary. The discoveries in this study would be helpful to provide some clues for proceeding with the process.

In this study, six hub genes (including CEP55, DEPDC1, CLSPN, KIF23, MYBL2, and RACGAP1) were identified after comprehensive bioinformatic analysis in HCC cohort from TCGA, and were all overexpressed in HCC patients, which was validated in mRNA and protein expressed level based on various HCC cohorts from Oncomine and experimental data in HCC cell lines and normal liver cell line by qPCR method. Meanwhile, these hub genes were independently predicted factors associated with poor prognosis in HCC patients in various databases of GEPIA and TISIDB. The consistent results

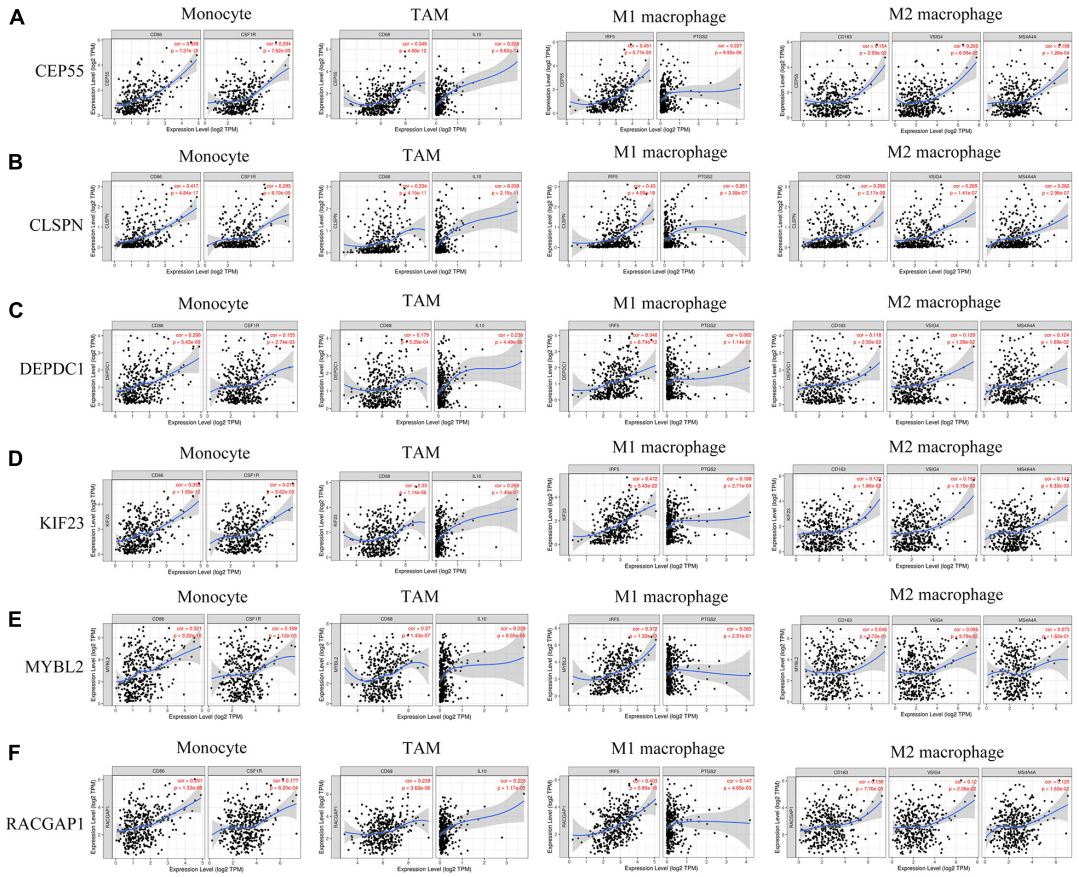


FIGURE 8 | (A–F) Six hub genes expression correlated with macrophage polarization in HCC. Markers include CD86 and CSF1R of monocytes; CD68 and IL10 of TAMs (tumor-associated macrophages); IRF5 and PTGS2 of M1 macrophages; and CD163, VSIG4, and MS4A4A of M2 macrophages.

from databases and experimental approaches, to large extent, enhance the reliability of these findings. It suggested that the six hub genes would be potential biomarkers for evaluating clinical prognosis in HCC patients, and also provide some valuable clues for further study involved in the progression of HCC. Except for RACGAP1 that was the first time found to correlate to the progression in HCC, the other hub genes have also been reported in several tumors associated with biological process or regulatory pathways, although nowadays existing data is limited and more in-depth investigations are required. It was reported that CEP55 could promote proliferation and invasion in the progress of osteosarcoma. In another study, data showed that CEPP55 facilitates the EMT process in renal cell (RCC) cancer and participates in the AKT pathway (Xu et al., 2018; Chen et al., 2019). DEPDC1 has been identified in various cancers, including HCC, breast cancer, and prostate cancer, among others, and is positively involved with multiple tumorous biological processes, including proliferation, invasion etc. (Huang et al., 2017; Guo et al., 2019; Zhao H. et al., 2019). CLSPN was demonstrated to be overexpressed in RCC as assessed by immunohistochemistry in 95 RCC cases. It was also found that CLSPN activates the AKT signaling pathway and is co-expressed with several known tumor-related genes, such as programmed death ligand-1 (Kobayashi et al., 2020). Previous studies have reported that KIF23 is a kinesin-like motor protein and has two splice variants, KIF V1 and KIFV2, which are overexpressed in HCC samples but were not detected in non-tumor tissues. HCC patients with positive identified KIF V1 had a better overall 5-year survival rate than those with no KIF V1. However, there was no significant correlation between expression of KIF V2 and overall survival in patients (Sun et al., 2015; Wei et al., 2019). These discoveries by Xiaotong et al. conflicted with results in this study, indicating that additional investigations should be performed to further elucidate the role of KIF23 in the tumorigenesis of HCC (Sun et al., 2015). MYBL2 (MYB proto-oncogene like 2) is included in the family of MYB transcription factors and was overexpressed in breast cancer. Jianlin et al. revealed that overexpressed MYBL2 in breast cancer promotes growth and metastasis (Chen and Chen, 2018).

Additionally, in our study, we also found that all six hub genes were positively correlated with immune cell infiltration of B cells, $CD4^+$ T cells, $CD8^+$T cells, macrophages, neutrophils, and dendritic cells in HCC. Furthermore, our analysis revealed that the expressed level of hub genes was correlated with

TABLE 2 | The correlation between six hub genes and gene markers of monocyte and macrophages in HCC from GEPIA.

Immune cell type		Gene marker	CEP55		CLSPN		DEPDC1		KIF23		MYBL2		RACGAP1	
			R	P	R	P	R	P	R	P	R	P	R	P
Tumor	Monocyte	CD86	0.53	*	0.44	*	0.3	****	0.52	*	0.25	****	0.38	****
		CSF1R	0.44	*	0.35	****	0.22	****	0.45	*	0.15	**	0.29	****
	TAM	CD68	0.33	****	0.3	****	0.18	***	0.33	****	0.2	***	0.27	****
		IL10	0.57	*	0.29	****	0.18	***	0.34	****	0.12	*	0.35	****
	M1 Macrophage	IRF5	0.34	****	0.35	****	0.27	****	0.34	****	0.36	****	0.38	****
		PTGS2	0.066	0.21	0.12	*	0.1	*	0.089	0.088	-0.013	0.8	0.048	0.36
	M2 Macrophage	CD163	0.46	*	0.26	****	0.16	**	0.43	*	0.13	*	0.24	****
		VSIG4	0.46	*	0.32	****	0.21	****	0.49	*	0.15	**	0.26	****
		MS4A4A	0.41	*	0.26	****	0.16	**	0.39	****	0.092	0.079	0.23	****
Normal	Monocyte	CD86	0.43	**	0.17	0.24	0.13	0.38	0.42	**	0.3	*	0.35	*
		CSF1R	0.41	**	0.26	0.064	0.1	0.48	0.53	****	0.33	*	0.42	**
	TAM	CD68	0.34	*	0.17	0.24	0.026	0.86	0.42	**	0.3	*	0.36	**
		IL10	0.25	0.07	0.1	0.47	0.02	0.89	0.19	0.2	0.15	0.3	0.36	*
	M1 Macrophage	IRF5	0.29	0.1	0.051	0.73	0.043	0.77	0.24	0.093	0.15	0.29	0.23	0.11
		PTGS2	0.02	0.89	-0.033	0.82	-0.098	0.5	0.064	0.66	0.00029	1	0.049	0.73
	M2 Macrophage	CD163	0.36	*	0.13	0.38	0.013	0.93	0.44	**	0.27	0.058	0.28	*
		VSIG4	0.2	0.15	0.097	0.5	0.00088	1	0.38	**	0.2	0.15	0.24	0.099
		MS4A4A	0.37	**	0.14	0.34	0.078	0.59	0.46	***	27	0.055	0.38	**

*Annotation: Tumor, correlation analysis in HCC tissue of TCGA. Normal, correlation analysis in normal liver tissue of TCGA. R, R-value of Pearson's correlation. *P < 0.05; **P < 0.01; ***P < 0.001; and ****P < 0.0001.*

TABLE 3 | The predicted compounds target hub genes via Drugbank database.

Hub genes	Predicted compounds	Accession number	Type	Classification in database	Chemical formula	Numbers of Clinical trial
CEP55	Irdabisant	DB12900	small molecule	Investigational	C18H23N3O2	1
	CEP-9722	DB14882	small molecule	Investigational	C24H26N4O3	3
	CEP-1347	DB05403	small molecule	Investigational	C33H33N3O5S2	1
	CEP-37440	DB13060	small molecule	Investigational	C30H38ClN7O3	1
	Cefapirin	DB01139	small molecule	Approved	C17H17N3O6S2	Not available
CLSPN	Calfactant	DB06415	small molecule	Approved	not found	13
	Calusterone	DB01564	small molecule	Experimental	C21H32O2	Not available
MYBL2	Clotrimazole	DB00257	small molecule	Approved	C22H17ClN2	37
	Propafenone	DB01182	small molecule	Approved	C21H27NO3	22
	Letrozole	DB01006	small molecule	Approved, Investigational	C17H11N5	308
	Sildenafil	DB00203	small molecule	Approved, Investigational	C22H30N6O4S	290
	Ranitidine	DB00863	small molecule	Approved, Withdrawn	C13H22N4O3S	51
	Valproic acid	DB00313	small molecule	Approved, Investigational	C8H16O2	244
	Esomeprazole	DB00736	small molecule	Approved, Investigational	C17H19N3O3S	245
	Pregabalin	DB00230	small molecule	Approved, Investigational	C8H17NO2	306

No compounds were detected in Drugbank database for hub genes of DEPDC1, KIF23, and RAGCAP1.

most immune-related markers of various immune cell types. Interestingly, hub genes of CEP55, CLSPN, DEPDC1, KIF23, and RACGAP1 were significantly related to markers of macrophages, showing its correlation to macrophage polarization. Until now, the correlation of six hub genes and immune infiltration in tumors has not been reported yet. Through an integrated bioinformatics analysis based on a ceRNA network, we identified several new biomarkers related to immune infiltration in HCC,

which may represent potential prognostic and therapeutic targets. However, more in-depth experimental studies are required to further verify the function and elaborate on the underlying mechanisms in HCC. Finally, based on analysis of the Drugbank database, we found a total of 15 bioactive compounds targeting CEP55, CLSPN, and MYBL2, providing some new directions for drug development for HCC treatments in the future. Meanwhile, it suggested these genes could be potential values for predicting the efficacy of clinical treatment.

CONCLUSION

In all, a simple-to-use nomogram predictive model was established based on miRNAs revealed great performance. We also constructed a ceRNA regulatory network to better understand the interactions between mRNAs and ncRNAs in HCC. Moreover, six hub genes were identified through PPI network analysis, all of which are overexpressed in HCC and are associated with survival. Besides, the expression of hub genes was independently correlated to poor prognosis in HCC patients, and was closely associated with immune infiltration in HCC. We believe these genes may be involved in the development of HCC and may represent potential prognostic biomarkers and individual therapeutic targets. However, further in-depth experiments are required to clarify detailed functions and mechanisms.

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

FP and YiZ designed the experiments and revised the manuscript. ZP analyzed the data and wrote the manuscript. XW, YunZ, and YuaZ searched and helped to analyze the data. All authors read and consent the final manuscript.

ACKNOWLEDGMENTS

We appreciate the TCGA team of the National Cancer Institute for allowing us to analyze the data. This manuscript has been released as a pre-print at Research Square.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.591623/full#supplementary-material

REFERENCES

Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4, 299–316. doi: 10.7554/eLife.05005

Aran, D., Sirota, M., and Butte, A. J. (2015). Systematic pan-cancer analysis of tumour purity. *Nat. Commun.* 6:8971. doi: 10.1038/ncomms9971

Baltruskeviciene, E., Schveigert, D., Stankevicius, V., Mickys, U., Zvirblis, T., Bublevic, J., et al. (2017). Down-regulation of miRNA-148a and miRNA-625-3p in colorectal cancer is associated with tumor budding. *BMC Cancer* 17:607. doi: 10.1186/s12885-017-3575-z

Cai, C., He, H., Duan, X., Wu, W., Mai, Z., Zhang, T., et al. (2018). miR-195 inhibits cell proliferation and angiogenesis in human prostate cancer by downregulating PRR11 expression. *Oncol. Rep.* 39, 1658–1670. doi: 10.3892/or.2018.6240

Chen, H., Zhu, D., Zheng, Z., Cai, Y., Chen, Z., and Xie, W. (2019). CEP55 promotes epithelial-mesenchymal transition in renal cell carcinoma through PI3K/AKT/mTOR pathway. *Clin. Transl. Oncol.* 21, 939–949. doi: 10.1007/s12094-018-02012-8

Chen, J., and Chen, X. (2018). MYBL2 is targeted by miR-143-3p and regulates breast cancer cell proliferation and apoptosis. *Oncol. Res.* 26, 913–922. doi: 10.3727/096504017X15135941182107

Chen, J., Wang, Z., Wang, W., Ren, S., Xue, J., Zhong, L., et al. (2020). SYT16 is a prognostic biomarker and correlated with immune infiltrates in glioma: a study based on TCGA data. *Int. Immunopharmacol.* 84:106490. doi: 10.1016/j.intimp.2020.106490

Chen, Y., Ramjiawan, R. R., Reiberger, T., Ng, M. R., Hato, T., Huang, Y., et al. (2015). CXCR4 inhibition in tumor microenvironment facilitates anti-programmed death receptor-1 immunotherapy in sorafenib-treated hepatocellular carcinoma in mice. *Hepatology* 61, 1591–1602. doi: 10.1002/hep.27665

Chou, C. H., Shrestha, S., Yang, C. D., Chang, N. W., Lin, Y. L., Liao, K. W., et al. (2018). miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 46, D296–D302. doi: 10.1093/nar/gkx1067

Danaher, P., Warren, S., Dennis, L., D'amico, L., White, A., Disis, M. L., et al. (2017). Gene expression markers of tumor infiltrating leukocytes. *J. Immunother. Cancer* 5:18. doi: 10.1186/s40425-017-0215-8

Fu, S., Wang, J., Hu, X., Zhou, R. R., Fu, Y., Tang, D., et al. (2018). Crosstalk between hepatitis B virus X and high-mobility group box 1 facilitates autophagy in hepatocytes. *Mol. Oncol.* 12, 322–338. doi: 10.1002/1878-0261.12165

Gu, X., Li, H., Sha, L., and Zhao, W. (2020). Construction and comprehensive analyses of a competing endogenous RNA network in tumor-node-metastasis stage I hepatocellular carcinoma. *Biomed. Res. Int.* 2020:5831064. doi: 10.1155/2020/5831064

Guo, W., Li, H., Liu, H., Ma, X., Yang, S., and Wang, Z. (2019). DEPDC1 drives hepatocellular carcinoma cell proliferation, invasion and angiogenesis by regulating the CCL20/CCR6 signaling pathway. *Oncol. Rep.* 42, 1075–1089. doi: 10.3892/or.2019.7221

Huang, L., Chen, K., Cai, Z. P., Chen, F. C., Shen, H. Y., Zhao, W. H., et al. (2017). DEPDC1 promotes cell proliferation and tumor growth via activation of E2F signaling in prostate cancer. *Biochem. Biophys. Res. Commun.* 490, 707–712. doi: 10.1016/j.bbrc.2017.06.105

Jeggari, A., Marks, D. S., and Larsson, E. (2012). miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics* 28, 2062–2063. doi: 10.1093/bioinformatics/bts344

Kobayashi, G., Sentani, K., Babasaki, T., Sekino, Y., Shigematsu, Y., Hayashi, T., et al. (2020). Claspin overexpression is associated with high-grade histology and poor prognosis in renal cell carcinoma. *Cancer Sci.* 111, 1020–1027. doi: 10.1111/cas.14299

Lehmann, U., Hasemeier, B., Christgen, M., Muller, M., Romermann, D., Langer, F., et al. (2008). Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer. *J. Pathol.* 214, 17–24. doi: 10.1002/path.2251

Lei, B., Zhou, J., Xuan, X., Tian, Z., Zhang, M., Gao, W., et al. (2019). Circular RNA expression profiles of peripheral blood mononuclear cells in hepatocellular carcinoma patients by sequence analysis. *Cancer Med.* 8, 1423–1433. doi: 10.1002/cam4.2010

Li, B., Severson, E., Pignon, J. C., Zhao, H., Li, T., Novak, J., et al. (2016). Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol.* 17:174. doi: 10.1186/s13059-016-1028-7

Li, S., Huang, Y., Huang, Y., Fu, Y., Tang, D., Kang, R., et al. (2017). The long non-coding RNA TP73-AS1 modulates HCC cell proliferation through miR-200a-dependent HMGB1/RAGE regulation. *J. Exp. Clin. Cancer Res.* 36:51. doi: 10.1186/s13046-017-0519-z

Li, Y., Fu, Y., Hu, X., Sun, L., Tang, D., Li, N., et al. (2019a). The HBx-CTTN interaction promotes cell proliferation and migration of hepatocellular carcinoma via CREB1. *Cell Death Dis.* 10:405. doi: 10.1038/s41419-019-1650-x

Li, Y., Ma, B., Yin, Z., Liu, P., Liu, J., Li, J., et al. (2019b). Competing endogenous RNA network and prognostic nomograms for hepatocellular carcinoma patients who underwent R0 resection. *J. Cell. Physiol.* 234, 20342–20353. doi: 10.1002/jcp.28634

Liu, J., Li, W., Zhang, J., Ma, Z., Wu, X., and Tang, L. (2019). Identification of key genes and long non-coding RNA associated ceRNA networks in hepatocellular carcinoma. *PeerJ.* 7:e8021. doi: 10.7717/peerj.8021

Long, J., Bai, Y., Yang, X., Lin, J., Yang, X., Wang, D., et al. (2019). Construction and comprehensive analysis of a ceRNA network to reveal potential prognostic biomarkers for hepatocellular carcinoma. *Cancer Cell Int.* 19:90. doi: 10.1186/s12935-019-0817-y

Lou, W., Liu, J., Ding, B., Chen, D., Xu, L., Ding, J., et al. (2019). Identification of potential miRNA-mRNA regulatory network contributing to pathogenesis of HBV-related HCC. *J. Transl. Med.* 17:7. doi: 10.1186/s12967-018-1761-7

Lu, Y., Gao, W., Zhang, C., Wen, S., Huangfu, H., Kang, J., et al. (2015). Hsa-miR-301a-3p acts as an oncogene in laryngeal squamous cell carcinoma via target regulation of Smad4. *J. Cancer* 6, 1260–1275. doi: 10.7150/jca.12659

Lu, Y., Ryan, S. L., Elliott, D. J., Bignell, G. R., Futreal, P. A., Ellison, D. W., et al. (2009). Amplification and overexpression of Hsa-miR-30b, Hsa-miR-30d and KHDRBS3 at 8q24.22-q24.23 in medulloblastoma. *PLoS One* 4:e6159. doi: 10.1371/journal.pone.0006159

Pan, J. H., Zhou, H., Cooper, L., Huang, J. L., Zhu, S. B., Zhao, X. X., et al. (2019). LAYN Is a prognostic biomarker and correlated with immune infiltrates in gastric and colon cancers. *Front. Immunol.* 10:6. doi: 10.3389/fimmu.2019.00006

Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., et al. (2019). TISIDB: an integrated repository portal for tumor–immune system interactions. *Bioinformatics* 35, 4200–4202. doi: 10.1093/bioinformatics/btz210

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: the rosetta stone of a hidden RNA language? *Cell* 146, 353–358. doi: 10.1016/j.cell.2011.07.014

Sannigrahi, M. K., Sharma, R., Singh, V., Panda, N. K., Rattan, V., and Khullar, M. (2017). Role of host miRNA Hsa-miR-139-3p in HPV-16-induced carcinomas. *Clin. Cancer Res.* 23, 3884–3895. doi: 10.1158/1078-0432.CCR-16-2936

Shi, L., Peng, F., Tao, Y., Fan, X., and Li, N. (2016). Roles of long noncoding RNAs in hepatocellular carcinoma. *Virus Res.* 223, 131–139. doi: 10.1016/j.virusres.2016.06.008

Siemers, N. O., Holloway, J. L., Chang, H., Chasalow, S. D., Ross-Macdonald, P. B., Voliva, C. F., et al. (2017). Genome-wide association analysis identifies genetic correlates of immune infiltrates in solid tumors. *PLoS One* 12:e0179726. doi: 10.1371/journal.pone.0179726

Sousa, S., and Maatta, J. (2016). The role of tumour-associated macrophages in bone metastasis. *J. Bone Oncol.* 5, 135–138. doi: 10.1016/j.jbo.2016.03.004

Sun, X., Jin, Z., Song, X., Wang, J., Li, Y., Qian, X., et al. (2015). Evaluation of KIF23 variant 1 expression and relevance as a novel prognostic factor in patients with hepatocellular carcinoma. *BMC Cancer* 15:961. doi: 10.1186/s12885-015-1987-1

Wang, J. J., Huang, Y. Q., Song, W., Li, Y. F., Wang, H., Wang, W. J., et al. (2019). Comprehensive analysis of the lncRNAassociated competing endogenous RNA network in breast cancer. *Oncol. Rep.* 42, 2572–2582. doi: 10.3892/or.2019.7374

Wang, M., Mao, C., Ouyang, L., Liu, Y., Lai, W., Liu, N., et al. (2019). Long noncoding RNA LINC00336 inhibits ferroptosis in lung cancer by functioning as a competing endogenous RNA. *Cell Death Differ.* 26, 2329–2343. doi: 10.1038/s41418-019-0304-y

Wei, B., Kong, W., Mou, X., and Wang, S. (2019). Comprehensive analysis of tumor immune infiltration associated with endogenous competitive RNA networks in lung adenocarcinoma. *Pathol. Res. Pract.* 215, 159–170. doi: 10.1016/j.prp.2018.10.032

Wong, N., and Wang, X. (2015). miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* 43, D146–D152. doi: 10.1093/nar/gku1104

Wu, Z. H., Cai, F., and Zhong, Y. (2020). Comprehensive analysis of the expression and prognosis for GBPs in head and neck squamous cell carcinoma. *Sci. Rep.* 10:6085. doi: 10.1038/s41598-020-63246-7

Xiao, Y. (2020). Construction of a circRNA-miRNA-mRNA network to explore the pathogenesis and treatment of pancreatic ductal adenocarcinoma. *J. Cell Biochem.* 121, 394–406. doi: 10.1002/jcb.29194

Xiong, D. D., Dang, Y. W., Lin, P., Wen, D. Y., He, R. Q., Luo, D. Z., et al. (2018). A circRNA-miRNA-mRNA network identification for exploring underlying pathogenesis and therapy strategy of hepatocellular carcinoma. *J. Transl. Med.* 16:220. doi: 10.1186/s12967-018-1593-5

Xu, F., Zhao, Y., Qin, G., Huan, Y., Li, L., and Gao, W. (2019). Comprehensive analysis of competing endogenous RNA networks associated with cholangiocarcinoma. *Exp. Ther. Med.* 18, 4103–4112. doi: 10.3892/etm.2019.8052

Xu, L., Xia, C., Sheng, F., Sun, Q., Xiong, J., and Wang, S. (2018). CEP55 promotes the proliferation and invasion of tumour cells via the AKT signalling pathway in osteosarcoma. *Carcinogenesis* 39, 623–631. doi: 10.1093/carcin/bgy017

Yue, C., Ren, Y., Ge, H., Liang, C., Xu, Y., Li, G., et al. (2019). Comprehensive analysis of potential prognostic genes for the construction of a competing endogenous RNA regulatory network in hepatocellular carcinoma. *Onco. Targets Ther.* 12, 561–576. doi: 10.2147/OTT.S188913

Zhang, K., Li, Q., Kang, X., Wang, Y., and Wang, S. (2016). Identification and functional characterization of lncRNAs acting as ceRNA involved in the malignant progression of glioblastoma multiforme. *Oncol. Rep.* 36, 2911–2925. doi: 10.3892/or.2016.5070

Zhang, R., Jiang, Y. Y., Xiao, K., Huang, X. Q., Wang, J., and Chen, S. Y. (2020). Candidate lncRNA-miRNA-mRNA network in predicting hepatocarcinogenesis with cirrhosis: an integrated bioinformatics analysis. *J. Cancer Res. Clin. Oncol.* 146, 87–96. doi: 10.1007/s00432-019-03090-z

Zhang, Y., Zhang, L., Xu, Y., Wu, X., Zhou, Y., and Mo, J. (2020). Immune-related long noncoding RNA signature for predicting survival and immune checkpoint blockade in hepatocellular carcinoma. *J. Cell. Physiol.* 235, 9304–9316. doi: 10.1002/jcp.29730

Zhang, Z., Tang, D., Wang, B., Wang, Z., and Liu, M. (2019). Analysis of miRNA-mRNA regulatory network revealed key genes induced by aflatoxin B1 exposure in primary human hepatocytes. *Mol. Genet. Genomic Med.* 7:e971. doi: 10.1002/mgg3.971

Zhao, B., Lu, Y., Cao, X., Zhu, W., Kong, L., Ji, H., et al. (2019). MiRNA-124 inhibits the proliferation, migration and invasion of cancer cell in hepatocellular carcinoma by downregulating lncRNA-UCA1. *Onco Targets Ther.* 12, 4509–4516. doi: 10.2147/OTT.S205169

Zhao, H., Yu, M., Sui, L., Gong, B., Zhou, B., Chen, J., et al. (2019). High expression of DEPDC1 promotes malignant phenotypes of breast cancer cells and predicts poor prognosis in patients with breast cancer. *Front. Oncol.* 9:262. doi: 10.3389/fonc.2019.00262

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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 19 August 2021
doi: 10.3389/fgene.2021.720421

Identification of a Novel Glycolysis-Related LncRNA Signature for Predicting Overall Survival in Patients With Bladder Cancer

Zhenming Zheng[1,2,3†], Cong Lai[1,2,3†], Wenshuang Li[1,2,3], Caixia Zhang[1,2,3], Kaiqun Ma[1,2,3] and Yousheng Yao[1,2,3]*

[1] Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China, [2] Department of Urology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China, [3] Guangdong Provincial Clinical Research Center for Urological Diseases, Guangzhou, China

OPEN ACCESS

Edited by:
Jian-Bing Fan,
Illumina, United States

Reviewed by:
Benkang Shi,
Shandong University, China
Xiangfu Zhou,
Third Affiliated Hospital of Sun Yat-sen University, China

***Correspondence:**
Yousheng Yao
yaoyoush@mail.sysu.edu.cn

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

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 04 June 2021
Accepted: 30 July 2021
Published: 19 August 2021

Citation:
Zheng Z, Lai C, Li W, Zhang C, Ma K and Yao Y (2021) Identification of a Novel Glycolysis-Related LncRNA Signature for Predicting Overall Survival in Patients With Bladder Cancer. Front. Genet. 12:720421. doi: 10.3389/fgene.2021.720421

Background: Both lncRNAs and glycolysis are considered to be key influencing factors in the progression of bladder cancer (BCa). Studies have shown that glycolysis-related lncRNAs are an important factor affecting the overall survival and prognosis of patients with bladder cancer. In this study, a prognostic model of BCa patients was constructed based on glycolysis-related lncRNAs to provide a point of reference for clinical diagnosis and treatment decisions.

Methods: The transcriptome, clinical data, and glycolysis-related pathway gene sets of BCa patients were obtained from The Cancer Genome Atlas (TCGA) database and the Gene Set Enrichment Analysis (GSEA) official website. Next, differentially expressed glycolysis-related lncRNAs were screened out, glycolysis-related lncRNAs with prognostic significance were identified through LASSO regression analysis, and a risk scoring model was constructed through multivariate Cox regression analysis. Then, based on the median of the risk scores, all BCa patients were divided into either a high-risk or low-risk group. Kaplan-Meier (KM) survival analysis and the receiver operating characteristic (ROC) curve were used to evaluate the predictive power of the model. A nomogram prognostic model was then constructed based on clinical indicators and risk scores. A calibration chart, clinical decision curve, and ROC curve analysis were used to evaluate the predictive performance of the model, and the risk score of the prognostic model was verified using the TCGA data set. Finally, Gene Set Enrichment Analysis (GSEA) was performed on glycolysis-related lncRNAs.

Results: A total of 59 differentially expressed glycolysis-related lncRNAs were obtained from 411 bladder tumor tissues and 19 pericarcinomatous tissues, and 9 of those glycolysis-related lncRNAs (AC099850.3, AL589843.1, MAFG-DT, AC011503.2, NR2F1-AS1, AC078778.1, ZNF667-AS1, MNX1-AS1, and AC105942.1) were found to have prognostic significance. A signature was then constructed for predicting survival in

BCa based on those 9 glycolysis-related lncRNAs. ROC curve analysis and a nomogram verified the accuracy of the signature.

Conclusion: Through this study, a novel prognostic prediction model for BCa was established based on 9 glycolysis-related lncRNAs that could effectively distinguish high-risk and low-risk BCa patients, and also provide a new point of reference for clinicians to make individualized treatment and review plans for patients with different levels of risk.

Keywords: bladder cancer, glycolysis, lncRNAs, prognosis, signature

INTRODUCTION

Bladder cancer (BCa) is one of the most prevalent malignancies of the urinary system, according to statistics from 185 countries worldwide, in 2020, it ranked 9th in incidence and 13th in mortality (Sung et al., 2021). BCa is a heterogeneous disease and can be divided into two clinical types depending on whether it infiltrates the muscle layer: 25% of cases are muscle invasive BCa (MIBC), and 75% are non-muscle invasive BCa (NMIBC) (Cumberbatch et al., 2018). NMIBC has a low mortality rate, while 70% of these tumors will recur and 15% will progress in stage and grade. In contrast, the long-term survival rate of MIBC is unsatisfactory because it is prone to develop distant metastasis (Knowles and Hurst, 2015). Recent studies suggested that the poor prognosis of MIBC may also be related to the different molecular mechanisms of its occurrence and progression; some scholars even believed that MIBC and NMIBC were not the same kind of disease (Gui et al., 2011). With the increasing in-depth study of molecular biological mechanisms and the rapid development of gene detection technology, molecular typing and prognostic evaluation of BCa by genetic testing is becoming a new diagnostic and therapeutic target.

Glycolysis is a metabolic pathway that provides energy to the body. Changes in the glycolytic state may signify the presence of cancer. One study by Otto Warburg found that the glycolysis of tumor cells produced more lactic acid and less ATP, which resulted in tissue hypoxia, even in an environment with a high oxygen content. This phenomenon is known as the "Warburg effect" (Warburg, 1956). Enhanced or abnormally activated glycolysis promoted tumor progression, which has been proven in breast cancer, colorectal cancer, liver cancer, etc. (Chen et al., 2019a; Weng et al., 2020; Zhang et al., 2020). In BCa, an increase of the Warburg effect has been proven to contribute to the aggressiveness of BCa tumor cells and accelerated the proliferation rate of BCa cells (Ritterson Lew et al., 2015). Therefore, changes in glycolytic status may be an emerging marker of malignancy and a potential prognostic target for patients with BCa.

LncRNAs are a class of nucleotides with a transcriptional length of more than 200 nucleotides. Although lncRNAs do not possess protein-coding ability, they can regulate gene expression through multiple layers and pathways, thus affecting the occurrence and progression of tumors, and can be used as markers for tumor diagnosis and prognosis (Evans et al., 2016). There are more than 50,000 lncRNAs in the human genome, and they can play different roles in different tumors, acting as tumor-suppressor or tumor-promoter genes (Derrien et al., 2012). To effectively predict the prognosis of patients with BCa, the sensitivity and specificity of a single lncRNA was not sufficient; it was necessary to combine the high or low expression of multiple lncRNAs to have stronger diagnostic value.

Hyperglycolysis is a common phenomenon in the rapid proliferation of tumor cells and is also a marker of tumorigenesis. Recently, increasing amounts of data have shown that lncRNAs were a crucial regulator of glycolysis, and some studies have attempted to prevent tumor progression by inhibiting the activity of key enzymes in the tumor glycolysis pathway (Yang et al., 2014; Chen et al., 2019a; Liao et al., 2019; Wang et al., 2019a). The functional study of glycolysis-related lncRNAs has found that different lncRNAs could promote glycolysis in various ways, which in turn promoted tumorigenesis, progression, and metastasis. For example, lncRNA-SNHG7 regulated by c-MYC could increase glycolysis by inhibiting the expression of miR-34a-5p, thus promoting the progression of breast cancer (Zhang et al., 2019), while the lncRNA LINRIS stabilized IGF2BP2 and promoted aerobic glycolysis in colorectal cancer (Wang et al., 2019b).

The functional study of glycolysis-related lncRNAs in bladder cancer has found that lncRNA-SLC16A1-AS1 could be used as a target of E2F1 and a co-activator to induce metabolic reprogramming in the progression of bladder cancer (Logotheti et al., 2020), and lncRNA UCA1 up-regulated hexokinase 2 in bladder tumor cells to promote glycolysis through the mTOR–STAT3/microRNA143 pathway (Li et al., 2014). Prognostic models constructed by differentially expressed glycolysis-related lncRNAs have been proven to have good predictive performance in gastrointestinal tumors, breast cancer, gliomas, and other tumors. In bladder cancer, there is currently no known glycolysis-related lncRNA prognostic model.

This study aimed to clarify the relationship between glycolysis-related lncRNAs and the prognosis of BCa. First, a total of 9 glycolysis-related lncRNAs were identified to have prognostic significance. Then, based on these 9 glycolysis-related lncRNAs, a novel prognostic prediction model for BCa was constructed and verified, which proved that the model possessed good predictive power. Finally, gene function enrichment analysis was performed on the differentially expressed glycolysis-related

lncRNAs identified through this study to explore their potential functions in the progression of BCa.

MATERIALS AND METHODS

Sample Sources and Processing

Transcriptome sequencing data and clinical data of BCa patients were downloaded from The Cancer Genome Atlas (TCGA) database[1], and gene sets of glycolysis related functions and glycolysis related pathways were obtained from the Gene Set Enrichment Analysis (GSEA) official website[2]. The "edgeR" package of R software was used to normalize the entire data set, set $|\log2FC| > 1$ and FDR < 0.05 as the threshold to construct the volcano map, and obtain the tumor tissue and pericarcinomatous differentially expressed glycolysis-related genes and lncRNA. The correlation between the lncRNAs and glycolysis-related genes was calculated using Pearson correlation. If the correlation coefficient of lncRNA $|R2| > 0.3$ and $P < 0.05$, it was considered to be related to glycolysis and used for further analysis.

Construction of the Risk Score Model

Univariate Cox regression and Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis were used to screen out 9 lncRNAs with prognostic significance. Genes with $P < 0.05$ were considered to be independent glycolysis-related lncRNAs with prognostic significance. Multivariate Cox regression was used to calculate their respective coefficients (βI). Then, the formula of risk scoring model composed of βI and lncRNA expression level (EXPI) was obtained as follows: Risk score = $\sum_{i=1}^{9}(\beta_i * Exp_i)$. Then the risk score formula was used to calculate the risk score for each BCa patient. Subsequently, the median of the risk score was set as the cutoff value, and all eligible patients with bladder cancer were divided into high-risk and low-risk groups. Next, Kaplan-Meier survival analysis was performed to compare the differences in Overall Survival (OS) between the two groups of BCa patients. Univariate Cox regression and multivariate Cox regression analysis were used to evaluate whether risk score, age, gender, and TNM were independent prognostic factors for BCa. Clinical survival analysis was used to evaluate the predictive power of the risk scoring model for different clinical subgroups. ROC curve analysis was used to evaluate the efficiency of the prognostic model. These results were also verified by the validation data.

Establishment and Evaluation of the Prognostic Model

A nomogram was then constructed to obtain the 3-year and 5-year OS of BCa patients. This nomogram included the results of multivariate Cox regression, risk scores, and clinical information such as gender, age, and TNM stage. The calibration chart generated by the "rms" package of the R software verified the predictive performance of the nomogram, and the ROC curve analysis was used to evaluate the accuracy of the nomogram. Then, a decision curve analysis (DCA) was performed to verify the predictive effect of the prognostic model.

Gene Set Enrichment Analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was performed on the differentially expressed genes of BCa patients between the high-risk and low-risk groups. By analyzing the tumor signaling pathways and tumor progression processes involved in the differentially expressed genes, their possible roles and mechanisms in the occurrence and progression of BCa was further explored.

RESULTS

Differentially Expressed Glycolysis-Related Genes and LncRNAs in Bladder Cancer

The flow chart in **Figure 1** shows the general step-by-step process of this study. First, the gene expression data, lncRNA sequencing data and clinical data of the corresponding BCa patients (411 tumor tissues, 19 pericarcinomatous tissues) were downloaded from the TCGA database, and glycolysis-related gene sets (GLYCOLYTIC_PRCOESS, HALLMARK_GLYCOLYSIS, REACTOME_GLYCOLYSIS, GLYCOLYSIS_GLUCONEOGENESIS, and BIOCARTA_GLYCOLYSIS_PATHWAY) were downloaded from the GSEA database. The "edgeR" package of R software was used to normalize the entire data set, Set $|\log2FC| > 1$ and false discovery rate (FDR) < 0.05 as the threshold to obtain glycolysis-related genes and lncRNAs that were differentially expressed between tumors and pericarcinomatous tissues (**Figure 2**).

Next, Pearson correlation analysis was used to analyze the correlation between the differentially expressed glycolysis-related genes and lncRNAs, and to set correlation coefficients $|R2| > 0.3$ and $P < 0.05$ for lncRNAs that were considered to be related to glycolysis. A total of 59 glycolysis-related lncRNAs were identified as differentially expressed in BCa tumors and pericarcinomatous tissues (**Supplementary Table 1**). Subsequently, univariate Cox regression analysis was performed, through which 10 glycolysis-related lncRNAs related to the prognosis of BCa were identified (**Figure 3A**). Lastly, 9 key glycolysis-related lncRNAs associated with BCa prognosis were further screened by LASSO regression (**Figures 3B,C**).

Construction of a Risk Scoring Model for 9 Glycolysis-Related LncRNAs

The data of bladder cancer patients with a follow-up time of more than 30 days was selected to construct a risk scoring model, where their clinical parameters and pathological stages were showed in **Supplementary Table 2**. Patients were divided into training set and validation set at a ratio of 2:1. The risk score formula (Risk score = $\sum_{i=1}^{9}(\beta_i * Exp_i)$) was used to calculate the risk scores of all BCa patients. The median of the risk score was set as the cutoff value, and all eligible patients with BCa were divided into either high-risk or low-risk groups. Results of Kaplan-Meier survival analysis showed that the Overall Survival

[1] https://cancergenome.nih.gov/

[2] www.gsea-msigdb.org/

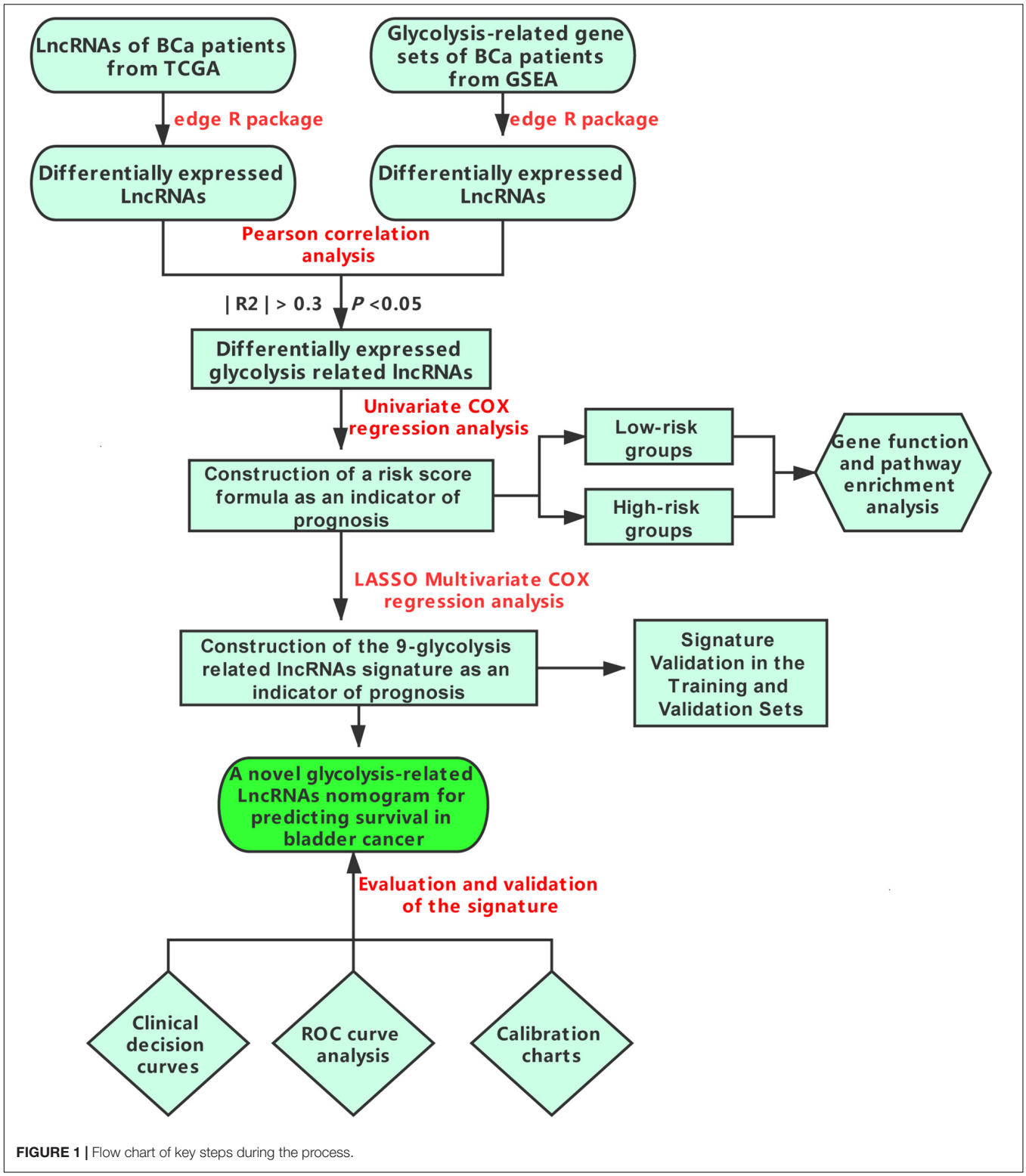


FIGURE 1 | Flow chart of key steps during the process.

(OS) of the high-risk group of BCa patients was significantly lower than that of the low-risk group, both in the training set and the validation set (P = 1.546e-10, **Figures 4A,B**). The glycolysis-related lncRNAs AC011503.2, AC078778.1, and ZNF667-AS1 were expressed at a lesser extent in tumor tissues of patients with BCa and were thus considered to be protective factors for the prognosis of bladder cancer, while glycolysis-related lncRNA AC099850.3, AL589843.1, MAFG -DT, NR2F1-AS1, MNX1-AS1, and AC105942.1 were highly expressed in tumor tissues and were therefore considered to be risk factors for the

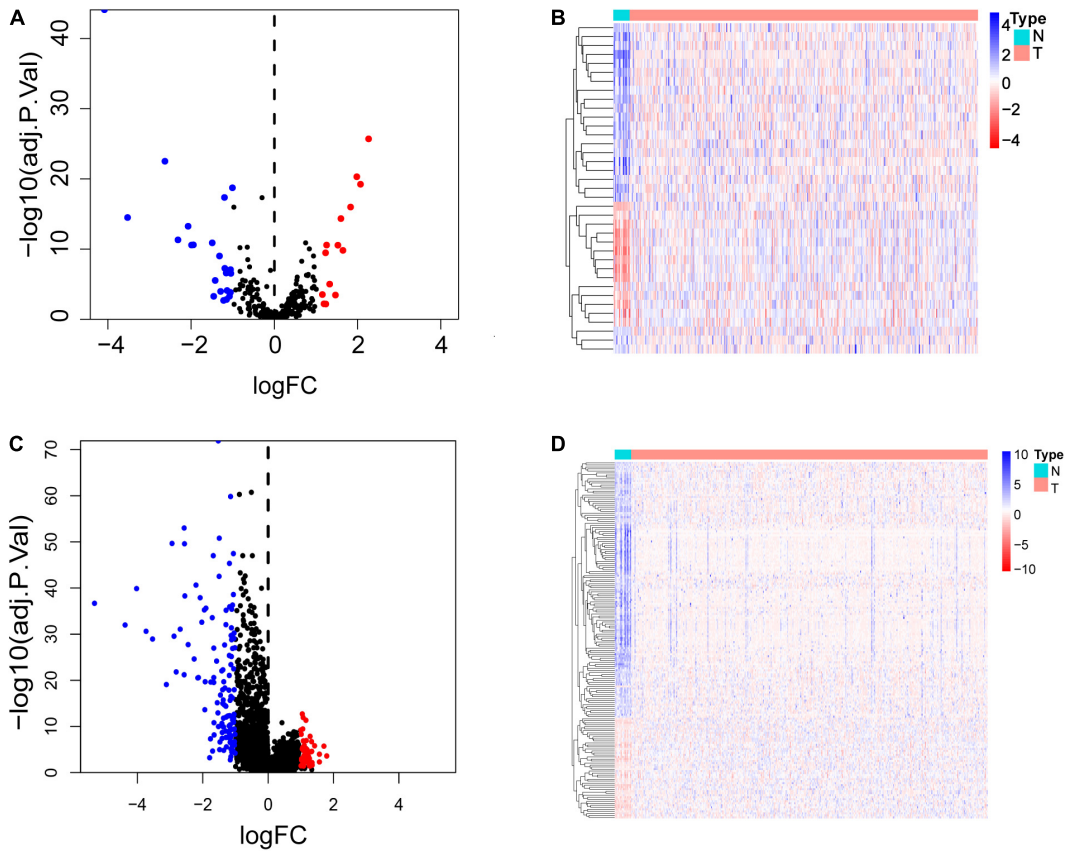


FIGURE 2 | Differentially expressed glycolysis-related genes and lncRNAs in bladder cancer compared to pericarcinomatous tissues. **(A,B)** Volcano plot and heat map of glycolysis-related genes differentially expressed in bladder cancer tissue and pericarcinomatous tissue. **(C,D)** Volcano plot and heat map of lncRNAs differentially expressed in bladder cancer tissues and pericarcinomatous tissue.

prognosis of BCa. ROC curve analysis proved that the model has a good predictive performance (3-year AUC = 0.753, 5-year AUC = 0.797) (**Figure 4C**). We evaluated the survival times of patients in the high- and low-risk groups and found that mortality rates for patients with high-risk scores were higher than those with low-risk scores (**Figures 4E,G**). Heatmap analysis was performed to reveal the expression profiles of the 9 glycolysis-related lncRNAs in high-risk and low-risk groups (**Figure 4I**). These results were then validated using validation set data (**Figures 4B,D,F,H,J**).

Construction and Evaluation of Prognostic Model of 9 Glycolysis-Related LncRNAs

A nomogram was then constructed to obtain the 3-year and 5-year OS of BCa patients (**Figure 5A**). This nomogram included the results of multivariate Cox regression, risk scores, and clinical information such as gender, age, and TNM stage. The calibration chart (**Figures 5D,F**) generated by using the "rms" package of the R software helped verify the performance of the nomogram. Next, the ROC curve (**Figure 5B**) was used to obtain the area under the curve (3-year AUC = 0.781, 5-year AUC = 0.821) to evaluate the accuracy of the nomogram. These results further proved that the prognostic signatures of the 9 glycolysis-related lncRNAs were independent prognostic factors for bladder cancer. Subsequently, the C index was calculated (training set 0.79, validation set 0.724), and Decision Curve Analysis (**Figures 5H,I**) was conducted to verify the predictive effect of the prognostic model. Finally, the validation set data was used to verify these results (**Figures 5C,E,G**). Subgroup analysis results (**Figure 6**) showed that the model had good predictive efficacy in all age groups, in the T1-T2 and T3-T4 stages, with or without lymph node metastasis, and with or without distant metastasis ($P < 0.05$) (**Figures 6A–I**). There was no significant difference observed in the female group ($P = 0.107$) (**Figure 6J**). Possibly because of the small number of cases in the group.

Gene Function and Pathway Enrichment Analysis of 9 Glycolysis Related LncRNAs Signatures

Kyoto Encyclopedia of Genes and Genomes enrichment analysis was performed to explore the possible roles and mechanisms of the differentially expressed glycolysis-related lncRNAs in the occurrence and progression of tumors in the high-risk group and the low-risk group of BCa patients, Analysis results (**Figures 7A–L**) showed that these glycolysis-related

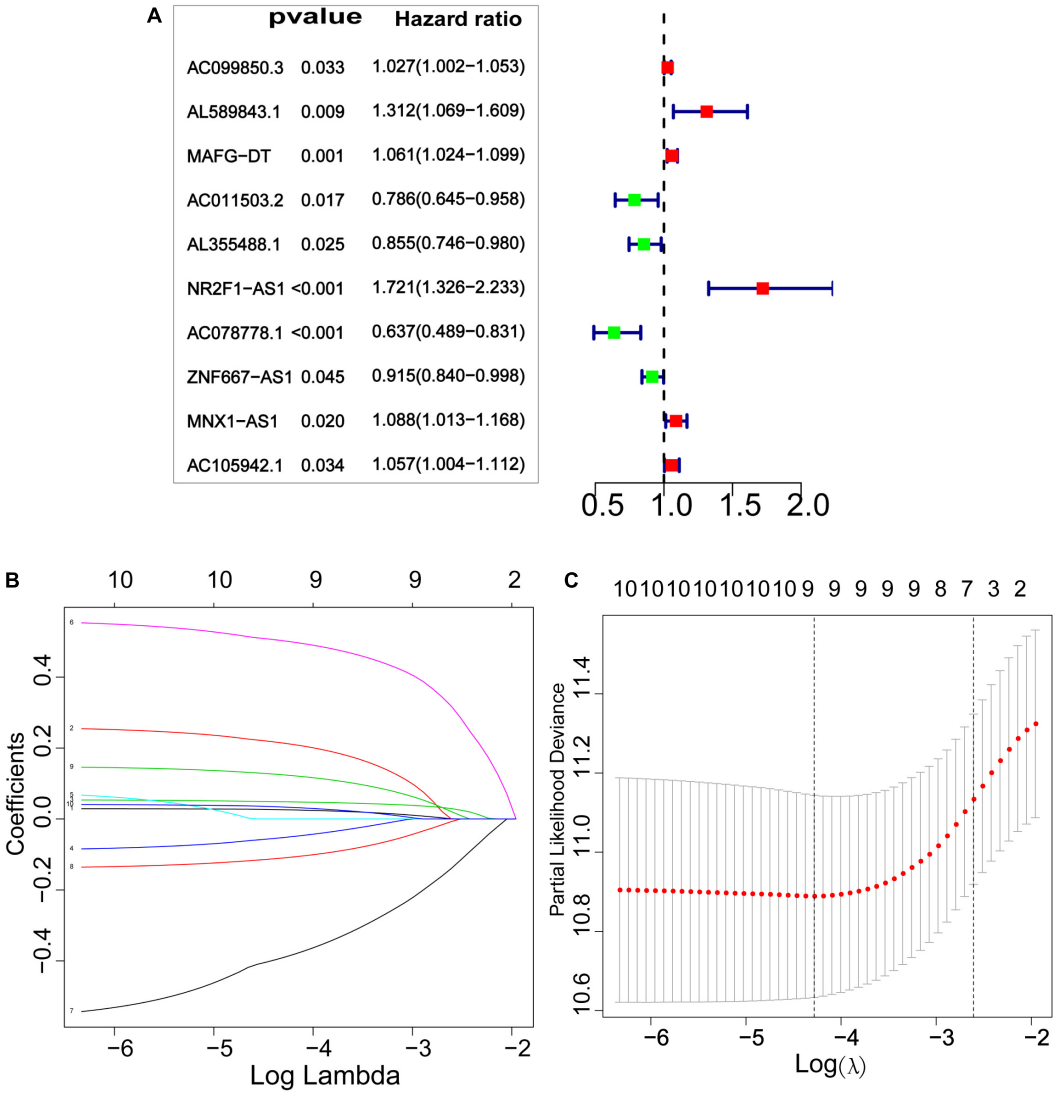


FIGURE 3 | Screening glycolysis-related lncRNAs with prognostic value. **(A)** Risk ratio Forest plot identified 10 glycolysis-related lncRNAs with prognostic significance ($P < 0.05$). **(B,C)** LASSO regression further screened out 9 glycolysis-related lncRNAs that were significantly related to prognosis.

lncRNAs were involved in multiple signaling pathways, such as: B-cell receptor, T-cell receptor, Hedgehog, MAPK, WNT, Calcium, and Chemokine. They were also involved in oxidative phosphorylation, ECM receptor interaction, and the development of basal cell carcinoma. These findings may help researchers determine the direction of further in-depth research in order to continue studying the mechanism of how glycolysis-related lncRNAs affects the progression of BCa.

DISCUSSION

Accurately predicting the prognosis of BCa patients is of great concern for clinicians in order to develop personalized treatment and review plans. For patients with a high risk of recurrence, we must be more cautious in the treatment strategy of bladder preservation, and for low-risk patients, we need to preserve the bladder as much as possible for the patients to help improve their postoperative quality of life. To determine whether a BCa patient was of a high-risk type, current clinical practice was mainly based on TNM staging, postoperative pathological diagnosis, and molecular classification, but there was no quantitative risk score or predictive model. Prognostic models constructed by differentially expressed glycolysis-related lncRNAs have been proven to have good predictive performance in gastrointestinal tumors, breast cancer, gliomas, and other tumors. Ho et al. (2021) constructed a glycolysis-related lncRNA prognosis model that could identify a subgroup of patients with different prognosis and different degrees of immune infiltration. In bladder cancer, there is currently no known glycolysis-related lncRNA prognostic model.

This prediction model was established based on the following theoretical basis and previous research results: First, an increasing amount of evidence shows that changes in glucose metabolism

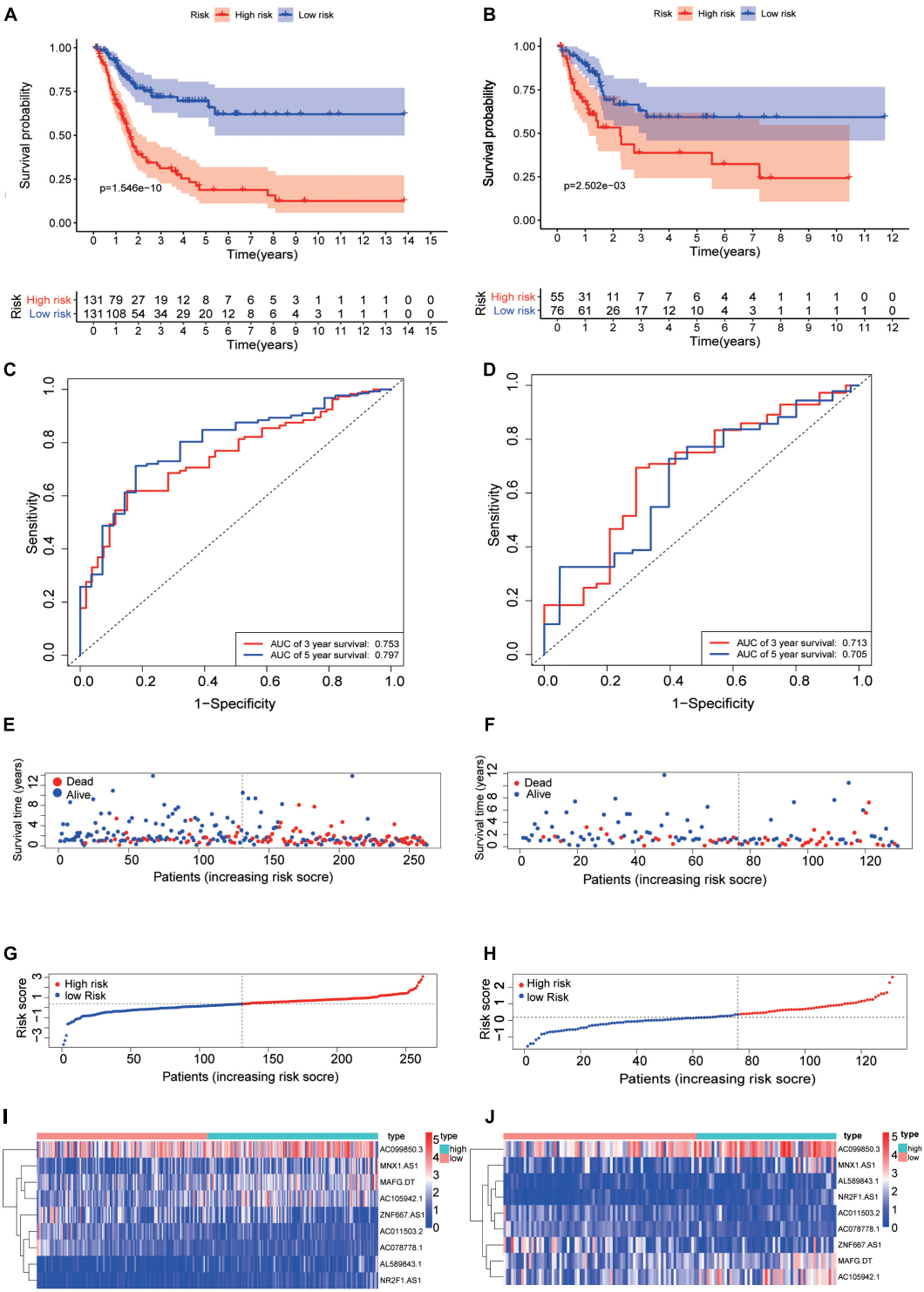


FIGURE 4 | Construction of risk score model. Kaplan-Meier survival analysis showed that the Overall Survival (OS) of patients in the high-risk training set **(A)** and validation set **(B)** was significantly lower than that in the low-risk group. The ROC curve of the training set **(C)** showed that the AUCs of the 3-year and 5-year OS were 0.753 and 0.797, respectively, while the validation set **(D)** were 0.713 and 0.705, respectively. Survival rate and survival status of bladder cancer patients in the training set **(E)** and validation set **(F)**. The distribution of 9 glycolysis-related lncRNAs risk scores for each patient in the training set **(G)** and validation set **(H)**. Heatmap of 9 glycolysis-related lncRNAs in the low-risk group and the high-risk group in the training set **(I)** and validation set **(J)**.

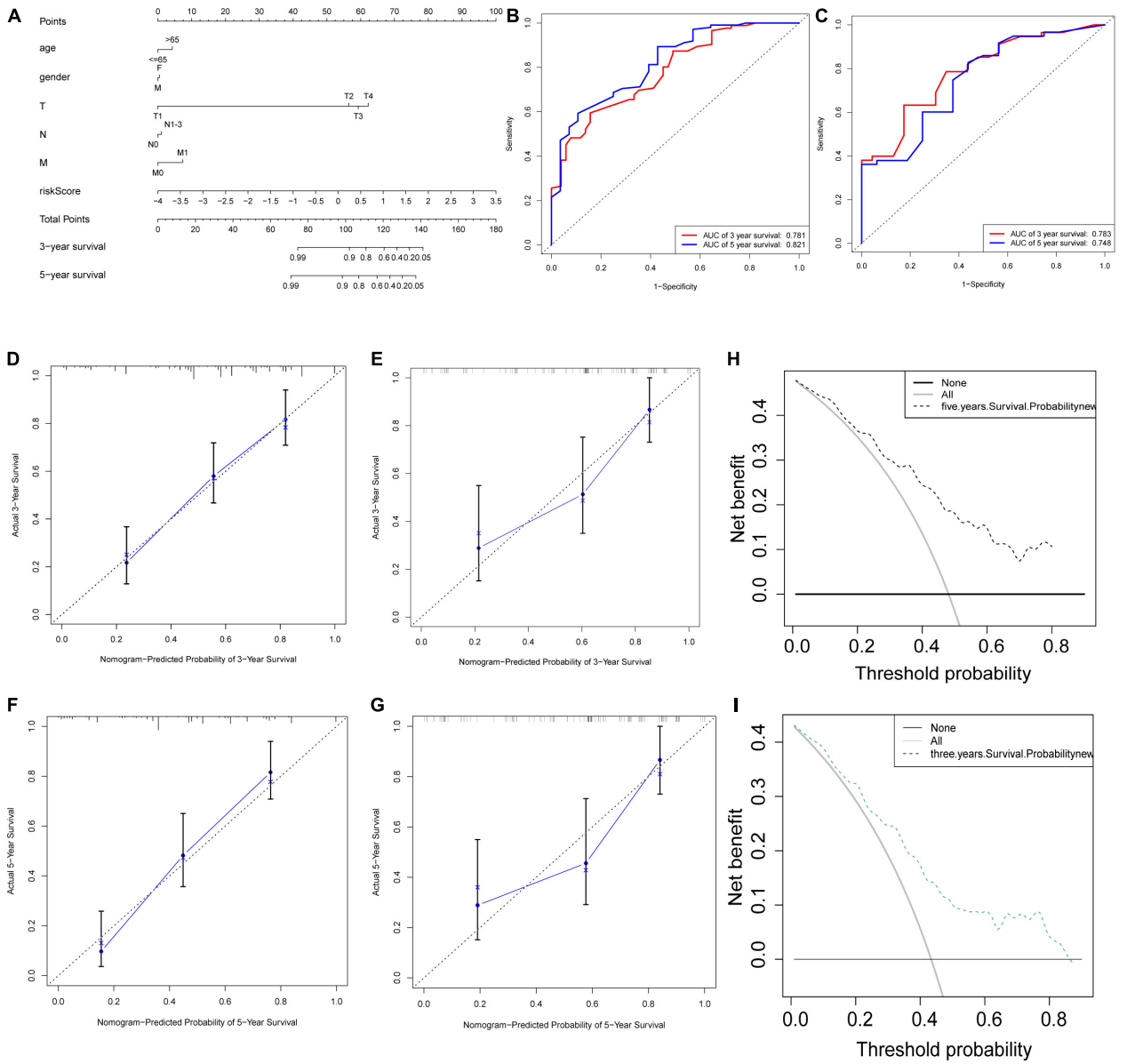


FIGURE 5 | Establishment and evaluation of the nomogram. **(A)** Nomogram for predicting the 3-year and 5-year survival rates of bladder cancer patients. **(B,C)** ROC curve analysis results showed that the AUC of the training set **(B)** for 3-year OS and 5-year OS were 0.781 and 0.821, respectively, and those in the validation set **(C)** were 0.783 and 0.748, respectively. **(D–G)** The calibration curves for the 3-year and 5-year OS of the nomogram in the training set **(D,F)** and validation set **(E,G)**. **(H,I)** The clinical decision curve for the 3-year OS **(H)** and 5-year OS **(I)** of the nomogram.

are a sign of tumorigenesis. Some research has pointed out that the aerobic glycolysis of bladder cancer cells is more active than that of normal urothelial cells, and the increase in aerobic glycolysis is beneficial to the proliferation, invasion, and metastasis of bladder tumor cells (Berrondo et al., 2016). Other studies have found that the activation of the Warburg effect was positively correlated with the degree of tumor malignancy (Agnihotri and Zadeh, 2016). Secondly, previous studies have confirmed that lncRNAs were also involved in glycometabolic reprogramming of tumor cells in BCa. Different lncRNAs play different roles in the glycolysis process of bladder tumor cells. Some lncRNAs accelerate tumor progression by promoting glycolysis, and some of them inhibit glycolysis to suppress cancer. For instance, Li et al. (2014) found that the lncRNA UCA1 could promote glycolysis by upregulation of HK2 through the mTOR-STAT3/miRNA and miR-143 pathways. Hu et al. (2017) found that the lncRNA CASC 8 may act as a tumor suppressor by reducing glycolysis in bladder tumor cells.

In this study, 9 glycolysis-related lncRNAs were screened in relation to the prognosis of BCa patients (AC099850.3, AL589843.1, MAFG-DT, AC011503.2, NR2F1-AS1, AC078778.1, ZNF667-AS1, MNX1-AS1, and AC105942.1). Among them,

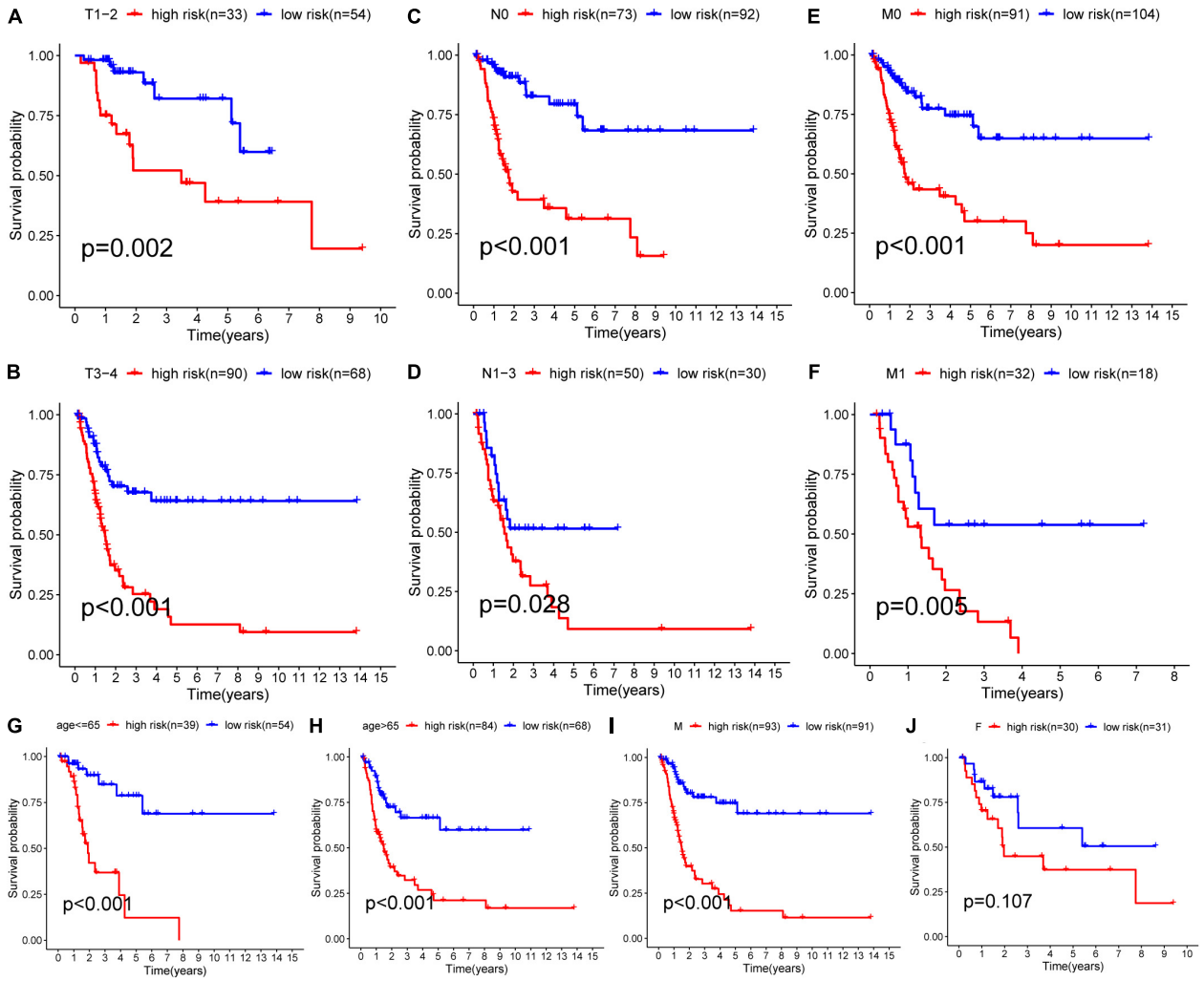


FIGURE 6 | The results of subgroup analysis. Analysis showed that the OS of bladder cancer patients in the high-risk group was significantly lower than that in the low-risk group based on age and TNM stage ($P < 0.05$). **(A–I)** In male patients, the OS in the high-risk group was also significantly shorter than that in the low-risk group ($P < 0.001$), while no significant difference was shown in female patients ($P = 0.107$) **(J)**, possibly due to the relatively small number of cases.

glycolysis-related lncRNA AC011503.2, AC078778.1, and ZNF667-AS1 were protective factors for the prognosis of BCa, and glycolysis-related lncRNAs AC099850.3, AL589843.1, MAFG-DT, NR2F1-AS1, MNX1-AS1, and AC105942.1 were risk factors for the prognosis of bladder cancer. MNX1-AS1 regulates the expression of RAB1A in bladder tumor cells by competitively binding with miR-218-5p, thereby promoting the proliferation, migration, invasion, and epithelial-mesenchymal transition of bladder tumor cells, which contributes to the growth and metastasis of bladder cancer cells (Wang et al., 2020). Multivariate Cox analysis was used to calculate the regression coefficient and construct the prognostic model. According to the risk score, the median was used to divide the bladder cancer patients into high-risk and low-risk groups. Analysis showed that the OS in the low-risk group was longer than that in the high-risk group. According to the results of multivariate Cox regression, a histogram was established, including age, sex, TNM stage, and risk score. The prediction performance of the model was verified by both a calibration diagram and decision curve analysis, and the results demonstrated that the model had good predictive performance.

In terms of GSEA functional enrichment analysis, it was found that these glycolysis related lncRNAs were significantly enriched for the regulation of multiple processes, such as the B cell and T cell receptor signaling pathway, Hedgehog signaling pathway, MAPK signaling pathway, TGF-β signaling pathway, chemokine signaling pathway, EMC receptor interaction, and oxidative phosphorylation processes. Li et al. (2016) found that GALNT1-mediated glycosylation and activation of Sonic Hedgehog signaling were involved in tumor-initiation and self-renewal of bladder cancer stem cells (Li et al., 2016). In another study, Teng Hu et al. found that RSPO3 promoted the aggressiveness of bladder cancer via the Wnt/β-catenin and Hedgehog signaling pathways (Chen et al., 2019b). Transcriptional regulatory factor

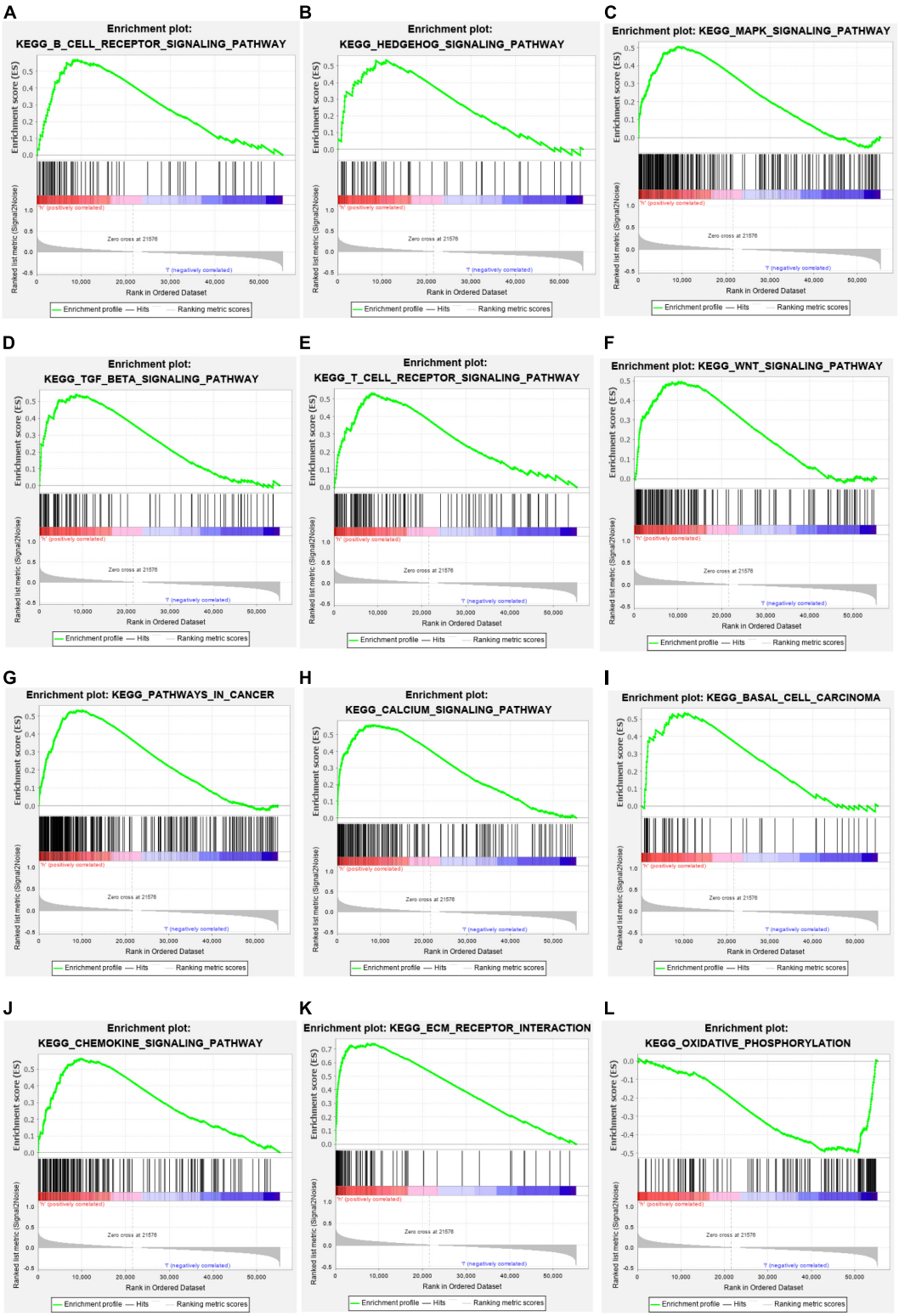


FIGURE 7 | Enrichment analysis. **(A–L)** KEGG pathway analysis indicates that these glycolysis related lncRNAs were involved in the B cell and T cell receptor signaling pathway, Hedgehog signaling pathway, MAPK signaling pathway, TGF-β signaling pathway, chemokine signaling pathway, EMC receptor interaction, and oxidative phosphorylation processes.

YAP was overexpressed in bladder cancer tissue and promoted the spread of bladder cancer by affecting the MAPK pathway (Qiu et al., 2020). All of these prior results further confirmed that the glycolysis associated lncRNAs screened in the present study play an important role in the occurrence and progression of BCa.

It should be pointed out that this study has some limitations: First, this study was a retrospective study of patient data from the TCGA database, which may lead to selection bias. Secondly, the patient information obtained for this study lacks longer follow-up time; therefore, the predictive model still needs to be further verified in large-scale prospective clinical trials.

CONCLUSION

This study established a novel prognostic model of BCa based on 9 glycolysis-related lncRNAs, which can effectively distinguish high-risk and low-risk bladder cancer patients from a new perspective. It can also provide a new point of reference of ideas for clinicians to formulate individualized treatment and review plans for patients with different levels of risk. This study also found that some new glycolysis-related lncRNAs were related to the prognosis of bladder cancer, which points out the direction for the next step of verification experiments and exploration of mechanisms.

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

YY designed the study and took overall control of the manuscript. ZZ analyzed and interpreted the data and drafted the manuscript. CL performed the data analysis and helped write the manuscript. WL collected relevant references, sorted out the data, and helped draft the manuscript. KM and CZ was involved in collecting the data and assisting with manuscript revision. All authors agreed to be accountable for the content of the work.

FUNDING

This work was supported by grants from Guangdong Provincial Clinical Research Center for Urological Diseases (2020B1111170006), from Guangdong Science and Technology Department (2020B1212060018).

ACKNOWLEDGMENTS

We thank all the staff at the Cancer Genome Atlas database and Gene Set Enrichment Analysis official website for managing the database, as well as the patients and researchers involved in the study who provided the data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.720421/full#supplementary-material

REFERENCES

Agnihotri, S., and Zadeh, G. (2016). Metabolic reprogramming in glioblastoma: the influence of cancer metabolism on epigenetics and unanswered questions. *Neuro. Oncol.* 18, 160–172. doi: 10.1093/neuonc/nov125

Berrondo, C., Flax, J., Kucherov, V., Siebert, A., Osinski, T., Rosenberg, A., et al. (2016). Expression of the long non-coding RNA HOTAIR correlates with disease progression in bladder cancer and is contained in bladder cancer patient urinary exosomes. *PloS One* 11:e0147236. doi: 10.1371/journal.pone.0147236

Chen, F., Chen, J., Yang, L., Liu, J., Zhang, X., Zhang, Y., et al. (2019a). Extracellular vesicle-packaged HIF-1α-stabilizing lncRNA from tumour-associated macrophages regulates aerobic glycolysis of breast cancer cells. *Nat. Cell Biol.* 21, 498–510. doi: 10.1038/s41556-019-0299-0

Chen, Z., Zhou, L., Chen, L., Xiong, M., Kazobinka, G., Pang, Z., et al. (2019b). RSPO3 promotes the aggressiveness of bladder cancer via Wnt/β-catenin and Hedgehog signaling pathways. *Carcinogenesis* 40, 360–369. doi: 10.1093/carcin/bgy140

Cumberbatch, M., Jubber, I., Black, P., Esperto, F., Figueroa, J., Kamat, A., et al. (2018). Epidemiology of bladder cancer: a systematic review and contemporary update of risk factors in 2018. *Euro. Urol.* 74, 784–795. doi: 10.1016/j.eururo.2018.09.001

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789. doi: 10.1101/gr.132159.111

Evans, J., Feng, F., and Chinnaiyan, A. (2016). The bright side of dark matter: lncRNAs in cancer. *J. Clin. Invest.* 126, 2775–2782. doi: 10.1172/jci84421

Gui, Y., Guo, G., Huang, Y., Hu, X., Tang, A., Gao, S., et al. (2011). Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. *Nat. Genet.* 43, 875–878. doi: 10.1038/ng.907

Ho, K., Huang, T., Shih, C., Lee, Y., Liu, A., Chen, P., et al. (2021). Glycolysis-associated lncRNAs identify a subgroup of cancer patients with poor prognoses and a high-infiltration immune microenvironment. *BMC Med.* 19:59. doi: 10.1186/s12916-021-01925-6

Hu, R., Zhong, P., Xiong, L., and Duan, L. (2017). Long noncoding RNA cancer susceptibility candidate 8 suppresses the proliferation of bladder cancer cells via regulating glycolysis. *DNA Cell Biol.* 36, 767–774. doi: 10.1089/dna.2017.3785

Knowles, M., and Hurst, C. (2015). Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. *Nat. Rev. Cancer* 15, 25–41. doi: 10.1038/nrc3817

Li, C., Du, Y., Yang, Z., He, L., Wang, Y., Hao, L., et al. (2016). GALNT1-mediated glycosylation and activation of sonic hedgehog signaling maintains the self-renewal and tumor-initiating capacity of bladder cancer stem cells. *Cancer Res.* 76, 1273–1283. doi: 10.1158/0008-5472.Can-15-2309

Li, Z., Li, X., Wu, S., Xue, M., and Chen, W. (2014). Long non-coding RNA UCA1 promotes glycolysis by upregulating hexokinase 2 through the mTOR-STAT3/microRNA143 pathway. *Cancer Sci.* 105, 951–955. doi: 10.1111/cas.12461

Liao, M., Liao, W., Xu, N., Li, B., Liu, F., Zhang, S., et al. (2019). LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. *EBioMedicine* 41, 200–213. doi: 10.1016/j.ebiom.2019.01.035

Logotheti, S., Marquardt, S., Gupta, S., Richter, C., Edelhäuser, B., Engelmann, D., et al. (2020). LncRNA-SLC16A1-AS1 induces metabolic

reprogramming during Bladder Cancer progression as target and co-activator of E2F1. *Theranostics* 10, 9620–9643. doi: 10.7150/thno.44176

Qiu, D., Zhu, Y., and Cong, Z. (2020). YAP Triggers Bladder Cancer Proliferation by Affecting the MAPK Pathway. *Cancer Manag. Res.* 12, 12205–12214. doi: 10.2147/cmar.S273442

Ritterson Lew, C., Guin, S., and Theodorescu, D. (2015). Targeting glycogen metabolism in bladder cancer. *Nat. Rev. Urol.* 12, 383–391. doi: 10.1038/nrurol.2015.111

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

Wang, C., Yang, Y., Zhang, G., Li, J., Wu, X., Ma, X., et al. (2019a). Long noncoding RNA EMS connects c-Myc to cell cycle control and tumorigenesis. *Proc. Natl. Acad. Sci. U. S. A.* 116, 14620–14629. doi: 10.1073/pnas.1903432116

Wang, J., Xing, H., Nikzad, A., Liu, B., Zhang, Y., Li, S., et al. (2020). Long noncoding RNA MNX1 antisense RNA 1 exerts oncogenic functions in bladder cancer by regulating miR-218-5p/RAB1A axis. *J. Pharmacol. Exp. Ther.* 372, 237–247. doi: 10.1124/jpet.119.262949

Wang, Y., Lu, J., Wu, Q., Jin, Y., Wang, D., Chen, Y., et al. (2019b). LncRNA LINRIS stabilizes IGF2BP2 and promotes the aerobic glycolysis in colorectal cancer. *Mol. Cancer* 18:174. doi: 10.1186/s12943-019-1105-0

Warburg, O. (1956). On the origin of cancer cells. *Science* 123, 309–314. doi: 10.1126/science.123.3191.309

Weng, M., Chen, W., Chen, X., Lu, H., Sun, Z., Yu, Q., et al. (2020). Fasting inhibits aerobic glycolysis and proliferation in colorectal cancer via the Fdft1-mediated AKT/mTOR/HIF1α pathway suppression. *Nat. Commun.* 11:1869. doi: 10.1038/s41467-020-15795-8

Yang, F., Zhang, H., Mei, Y., and Wu, M. (2014). Reciprocal regulation of HIF-1α and lincRNA-p21 modulates the Warburg effect. *Mol. Cell* 53, 88–100. doi: 10.1016/j.molcel.2013.11.004

Zhang, L., Fu, Y., and Guo, H. (2019). c-Myc-induced long non-coding RNA small nucleolar RNA host gene 7 regulates glycolysis in breast cancer. *J. Breast Cancer* 22, 533–547. doi: 10.4048/jbc.2019.22.e54

Zhang, Z., Tan, X., Luo, J., Yao, H., Si, Z., and Tong, J. (2020). The miR-30a-5p/CLCF1 axis regulates sorafenib resistance and aerobic glycolysis in hepatocellular carcinoma. *Cell Death Dis.* 11:902. doi: 10.1038/s41419-020-03123-3

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.

The reviewer XZ declared a shared affiliation with the authors to the handling editor at the time of the 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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 31 August 2021
doi: 10.3389/fgene.2021.714915

PLCE1 Polymorphisms Are Associated With Gastric Cancer Risk: The Changes in Protein Spatial Structure May Play a Potential Role

Xi'e Hu[1], Jintong Jia[2], Zhenyu Yang[1], Songhao Chen[1], Jingyi Xue[3], Sensen Duan[1], Ping Yang[1], Shujia Peng[1], Lin Yang[1], Lijuan Yuan[1] and Guoqiang Bao[1]**

[1] Department of General Surgery, The Second Affiliated Hospital of Air Force Medical University, Xi'an, China, [2] Singleron Biotechnologies, Nanjing, China, [3] The Second Clinical Medical College, Shaanxi University of Chinese Medicine, Xianyang, China

OPEN ACCESS

Edited by:
Youping Deng,
Rush University Medical Center,
United States

Reviewed by:
Pedro Dorado,
University of Extremadura, Spain
Lucia Taja-Chayeb,
National Institute of Cancerology (INCAN), Mexico

***Correspondence:**
Lijuan Yuan
564902156@qq.com
Guoqiang Bao
guoqiang@fmmu.edu.cn

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 26 May 2021
Accepted: 04 August 2021
Published: 31 August 2021

Citation:
Hu X, Jia J, Yang Z, Chen S, Xue J, Duan S, Yang P, Peng S, Yang L, Yuan L and Bao G (2021) PLCE1 Polymorphisms Are Associated With Gastric Cancer Risk: The Changes in Protein Spatial Structure May Play a Potential Role. Front. Genet. 12:714915. doi: 10.3389/fgene.2021.714915

Background: Gastric cancer (GC) is one of the most significant health problems worldwide. Some studies have reported associations between Phospholipase C epsilon 1 (*PLCE1*) single-nucleotide polymorphisms (SNPs) and GC susceptibility, but its relationship with GC prognosis lacked exploration, and the specific mechanisms were not elaborated fully yet. This study aimed to further explore the possible mechanism of the association between *PLCE1* polymorphisms and GC.

Materials and Methods: A case-control study, including 588 GC patients and 703 healthy controls among the Chinese Han population, was performed to investigate the association between SNPs of *PLCE1* and GC risk by logistic regression in multiple genetic models. The prognostic value of *PLCE1* in GC was evaluated by the Kaplan-Meier plotter. To explored the potential functions of *PLCE1*, various bioinformatics analyses were conducted. Furthermore, we also constructed the spatial structure of PLCE1 protein using the homology modeling method to analyze its mutations.

Results: Rs3765524 C > T, rs2274223 A > G and rs3781264 T > C in *PLCE1* were associated with the increased risk of GC. The overall survival and progression-free survival of patients with high expression of PLCE1 were significantly lower than those with low expression [HR (95% CI) = 1.38 (1.1–1.63), $P < 0.01$; HR (95% CI) = 1.4 (1.07–1.84), $P = 0.01$]. Bioinformatic analysis revealed that *PLCE1* was associated with protein phosphorylation and played a crucial role in the calcium signal pathway. Two important functional domains, catalytic binding pocket and calcium ion binding pocket, were found by homology modeling of PLCE1 protein; rs3765524 polymorphism could change the efficiency of the former, and rs2274223 polymorphism affected the activity of the latter, which may together play a potentially significant role in the tumorigenesis and prognosis of GC.

Conclusion: Patients with high expression of *PLCE1* had a poor prognosis in GC, and SNPs in *PLCE1* were associated with GC risk, which might be related to the changes in spatial structure of the protein, especially the variation of the efficiency of PLCE1 in the calcium signal pathway.

Keywords: gastric cancer, *PLCE1*, polymorphism, risk, prognosis, bioinformatics, protein structure

INTRODUCTION

Gastric cancer (GC) is becoming a worldwide problem year by year, endangering human life and health severely. It was estimated that over one million new GC cases occurred in 2018 and about 783 000 patients died of that, making GC the fifth most frequently diagnosed cancer and the third deadliest cancer worldwide (Bray et al., 2018). China has a large number of GC patients, with a 5-year overall survival (OS) of less than 25% (Chen et al., 2016; Zeng et al., 2018). The pathogenesis of GC is still unclear till now, but some risk factors have been reported, such as helicobacter pylori (Shimizu et al., 2014; Plummer et al., 2015; Jukic et al., 2021), Epstein-Barr virus infection (Camargo et al., 2014), low consumption of vegetables and fruits, high intake of salts and pickles, smoking and obesity (Lunet et al., 2007; Lin et al., 2014; Li et al., 2019). However, these research results are far from enough for us to understand the oncogenesis and susceptibility mechanism of GC.

In recent years, the genomic analysis of gastric tumors has highlighted the importance of its gene heterogeneity; and differentiations of GC molecular subtypes may be the key to guiding early diagnosis strategies, identifying new therapeutic targets, and predicting the prognosis of patients. In the last decade, single nucleotide polymorphism (SNP) analysis has been extensively used to screen candidate gene and detect various complex human diseases, providing a way to identify genetic loci associated with the heterogeneity of cancers.

Phospholipase C epsilon 1 (*PLCE1*) gene is one of the large-scale candidate genes located at 10q23 and served as a member of the human phosphoinositide-specific phospholipase C family (Song et al., 2001), which exerts an enormous function on growth, differentiation, and oncogenesis (Citro et al., 2007; Bunney and Katan, 2010; Gresset et al., 2012). The most-reported SNPs in *PLCE1* were rs2274223 and rs3765524, which have a significant value in increasing the risk of gastrointestinal tumor progression (Cui et al., 2014a; Mocellin et al., 2015; Mou et al., 2015; Xue et al., 2015; He et al., 2016; Gu et al., 2018). However, relevant studies of the associations between *PLCE1* and GC susceptibility remain inconsistent presently, and the prognostic value of *PLCE1* in GC is unclear; moreover, the specific mechanism between SNPs and GC risk is elusive now. Thus, further studies are still necessary.

This study aimed to analyze the relationship between three SNPs (rs3765524, rs2274223, and rs378126) in the *PLCE1* gene and GC susceptibility by a case-control study in the Chinese Han population firstly; then we explored the prognostic value of *PLCE1* in GC using online databases; finally, we tried to explain the correlation mechanism between the SNPs in *PLCE1* and the risk and prognosis of GC from the perspective of variable bioinformatics and protein spatial structure changes. We hope to make a contribute to the further exploration on the possible mechanism of the association between *PLCE1* polymorphisms and GC.

MATERIALS AND METHODS

Study Population

A case-control study was conducted, including 588 patients with GC (392 males and 196 females) and 703 healthy control subjects (396 males and 307 females). All subjects were genetically related to Chinese Han. Patients with histologically confirmed GC in the Second Affiliated Hospital of Air Force Medical University from January 2015 to January 2019 were enrolled. The exclusion criteria for patients were: Patients who had a family history (three generations) of tumors; Those who had received radiotherapy or chemotherapy before blood sampling collection; Patients with any other digestive diseases or caused by metastasis of other cancer. Additionally, the healthy controls were randomly recruited from the physical examination center of the same hospital during the same period when they visited for an annual health examination. When recruiting healthy participants, we investigated the demographic information by personally interviewing through a structured questionnaire by trained personnel, including age, gender, residential region, ethnicity, and family history of cancer and other diseases. The healthy participants who had a family history of cancer were also excluded from the study. After that, we collected 5 mL peripheral blood of each subject to detect the SNPs of the *PLCE1* gene for our research. All participants were voluntarily recruited and provided written informed consent before taking part in this study. All research analyses were performed following the approved guidelines and regulations. This study was approved by the Research Ethics Committee of the Second Affiliated Hospital of Air Force Medical University (K201501-05) and abided by the Declaration of Helsinki.

Genotyping

Agena MassARRAYAssay Design 4.0 software was used to design the multiplexed SNP Mass EXTEND assay. The *PLCE1* gene rs3765524, rs2274223, and rs3781264 polymorphisms were genotyped on the Agena MassARRAY RS1000 platform according to the standard protocol (Applied Biosystems, Foster City, CA, United States). Then, Agena Typer 4.0 software was applied to analyze and manage our data.

Bioinformatics Analysis

The Prognostic Value of *PLCE1* in GC

The Kaplan Meier (K-M) plotter[1] was used to evaluate the prognostic value of mRNA expression of *PLCE1*in GC patients. They were divided into high- and low-expression groups according to median values of mRNA expression and validated by K-M survival curves, with the hazard ratio (HR) with 95% confidence intervals (CIs) and Logrank *P*-value.

PLCE1 Associated Genes Screening and Enrichment Analysis

STRING database[2] (Szklarczyk et al., 2019) was applied to detect co-expression genes with *PLCE1* in GC, and Cytoscape software (Smoot et al., 2011) was used to explore and construct protein-protein interaction (PPI) network. Gene ontology (GO) enrichment, including biological process (BP), cellular component (CC) and molecular function (MF), and Kyoto Encyclopedia of Genes and Genomes (KEGG), were carried out to annotate PLCE1 functions by the Database for Annotation, Visualization, and Integrated Discovery (DAVID)[3] (Huang da et al., 2009).

Protein Homology-Modeling and Vitalization

The amino acid (aa) sequence of PLCE1 protein was obtained through NCBI.[4] We used SWISS-MODEL[5] to perform PLCE1 protein homology-modeling from its primary sequence (Schwede et al., 2003; Waterhouse et al., 2018). The protein with the highest coverage of the primary sequences was selected as the most homologous protein. We download the files of the constructed protein spatial structures in SWISS-MODEL and then opened them in PyMOL version 2.4[6] for protein visualization to pave the way for PLCE1 protein spatial structure analysis (Arroyuelo et al., 2016; Yuan et al., 2016).

Statistical Analysis

SPSS 26 (IBM SPSS Statistics, RRID:SCR_019096) software was applied to analyze the general characteristics of GC patients and healthy control groups. Welch's *t*-test and the Pearson Chi-square test were applied to analyze differences of the basic characteristics between the two groups. The Pearson Chi-square test was also used to assess deviation from Hardy-Weinberg equilibrium (HWE) to compare the observed and expected genotype frequencies among the control subjects. Allele and genotype frequencies were compared between GC patients and healthy controls using the Pearson Chi-squared test and Fisher's exact test. To evaluate the associations between *PLCE1* SNPs and the risk of GC, we calculated odds ratios (ORs) and 95% confidence intervals (CIs) adjusted by gender and age. Three different genetic models were applied (the codominant model, the dominant model and the recessive model) using PLINK software (PLINK, version 2.0, RRID:SCR_001757). *p*-value < 0.05 was considered statistically significant in all statistical tests in this study.

[1]http://kmplot.com/analysis/

[2]https://www.string-db.org/

[3]https://david.ncifcrf.gov/

[4]https://www.ncbi.nlm.nih.gov

[5]https://swissmodel.expasy.org

[6]http://www.pymol.org/2/

RESULTS

Demographic Characteristics

The primary characteristics of all subjects were shown in **Table 1**. A total of 1,291 participants, including 588 GC patients and 703 healthy controls, were enrolled in this study. The mean age was

TABLE 1 | Basic demographic characteristics of gastric cancer patients and healthy controls.

Characteristics	Patients ($n = 588$) (%)	Controls ($n = 703$) (%)	*P*-value
Age (years), mean ± SD	58.12 ± 11.66	48.57 ± 9.43	**<0.001***
Gender			**<0.001***
Male	392 (66.7)	396 (56.3)	
Female	196 (33.3)	307 (43.7)	
Tumor size (cm)	3.72 ± 2.12		
Location			
Cardia	141 (24.0)		
Body	211 (35.9)		
Pylorus	20 (3.4)		
Antrum	216 (36.7)		
Borrmann			
I	77 (13.1)		
II	56 (9.5)		
III	253 (43.0)		
IV	126 (21.4)		
V	59 (10.0)		
Unknown	17 (2.9)		
Differentiation			
Well	350 (59.5)		
Poor	238 (40.5)		
T stage			
T1	178 (30.3)		
T2	219 (37.2)		
T3	171 (29.1)		
T4	20 (3.4)		
N stage			
N0	294 (50.0)		
N1	141 (24.0)		
N2	75 (12.8)		
N3	78 (13.3)		
M stage			
M0	557 (94.7)		
M1	31 (5.3)		

P-values were calculated using Welch's t-test/Pearson Chi-square test.
SD, standard deviation.
**P < 0.05 indicates statistical significance, which was marked in bold.*

58.12 ± 11.66 years in GC patients and 48.57 ± 9.43 years in healthy controls, which indicated that the patients were elder than the healthy participants ($P < 0.001$). Besides, the scale of males was larger than females in the GC group (male to female is 66.67–33.33%), while the difference between males and females in the control group was minor (male to female is 56.33–43.67%). The difference in the distributions between GC patients and healthy controls suggested that the ORs and *p*-values need to be adjusted according to age and gender in subsequent analysis. Additionally, most of the participants in the study had an adverse family cancer history (cases, 96.3%; controls, 98.0%). Moreover, nearly one-third (30.3%) of patients were at an early stage (the carcinoma was confined to the gastric mucosa and submucosa).

Genotyping Analysis

The detailed information of the three selected SNPs, including roles, MAF, and HWE *P*-values, were listed in **Table 2**. These SNPs were genotyped successfully in further analysis. MAF of all SNPs was greater than 5%, and the observed genotype frequencies of all SNPs in the control groups were in HWE ($P > 0.05$).

TABLE 2 | Basic information of three SNPs in PLCE1 on 10q23.33.

SNP ID	Position	Alleles (A/B)	Role	MAF		O (HET)	E (HET)	*P*-value
				Case	Control			
rs3765524	96,058,298	T/C	Coding exon	0.244	0.214	0.331	0.336	0.654
rs2274223	96,066,341	G/A	Coding exon	0.246	0.213	0.332	0.335	0.822
rs3781264	96,070,375	C/T	Intron	0.190	0.142	0.240	0.243	0.756

Sites with HWE, P < 0.05, are excluded. P-values calculated with two-sided Pearson Chi-square test.
SNP, single nucleotide polymorphism; Alleles (A/B), Alleles (minor/major); MAF, minor allele frequency; O (HET), observed heterozygosity; E (HET), expected heterozygosity; HWE, Hardy-Weinberg equilibrium.
P < 0.05 indicates statistical significance.

TABLE 3 | The relationships of five SNPs with gastric cancer according to the stratification adjusted by gender and age.

SNP ID	Model	Genotype	Case (%)	Control (%)	OR (95%CI)	*P*-value
rs3765524	Allele	T	287 (24.40)	300 (21.34)	1.00	0.067
		C	889 (75.60)	1,104 (78.66)	1.19 (0.99, 1.43)	
	Codominant	T/T	29 (4.93)	34 (4.84)	1.00	0.080
		T/C	229 (50.85)	232 (33.00)	0.86 (0.51–1.47)	
		C/C	330 (56.12)	436 (62.02)	1.62 (0.96–2.72)	
	Dominant	T/T-T/C	258 (43.88)	266 (37.84)	1.00	**0.034***
		C/C	330 (56.12)	436 (62.02)	1.28 (1.03–1.60)	
	Recessive	TT	29 (4.93)	34 (4.84)	1.00	0.955
		T/C-C/C	559 (95.07)	668 (95.02)	1.02 (0.61–1.69)	
rs2274223	Allele	G	289 (24.57)	299 (21.27)	1.00	0.048*
		A	887 (75.43)	1,105 (78.73)	1.20 (1.00, 1.45)	
	Codominant	G/G	29 (4.93)	33 (4.69)	1.00	0.064
		G/A	231 (39.29)	233 (33.14)	0.89 (0.52–1.51)	
		A/A	328 (55.78)	436 (62.02)	1.17 (0.70–1.96)	
	Dominant	G/G-G/A	260 (44.22)	266 (37.84)	1.00	**0.025***
		A/A	328 (55.78)	436 (62.02)	1.30 (1.04–1.62)	
	Recessive	G/G	29 (4.93)	33 (4.69)	1.00	0.950
		G/A-A/A	559 (95.07)	669 (95.16)	1.05 (0.63–1.75)	
rs3781264	Allele	C	224 (19.05)	199 (14.15)	1.00	**0.001***
		T	952 (80.95)	1207 (85.85)	1.43 (1.16, 1.76)	
	Codominant	C/C	17 (2.89)	15 (2.13)	1.00	**0.002***
		C/T	190 (32.31)	169 (24.04)	1.01 (0.49–2.08)	
		T/T	381 (64.80)	519 (73.83)	1.54 (0.76–3.13)	
	Dominant	C/C-C/T	207 (35.20)	184 (26.17)	1.00	**0.001***
		T/T	381 (64.80)	519 (73.83)	1.53 (1.21–1.95)	
	Recessive	C/C	17 (2.89)	15 (2.13)	1	0.489
		C/T-T/T	571 (97.11)	688 (97.87)	1.37 (0.68–2.76)	

ORs and P-values were adjusted by age and gender.
OR, odds ratio; CI, confidence interval.
**P < 0.05 indicates statistical significance, which was marked in bold.*

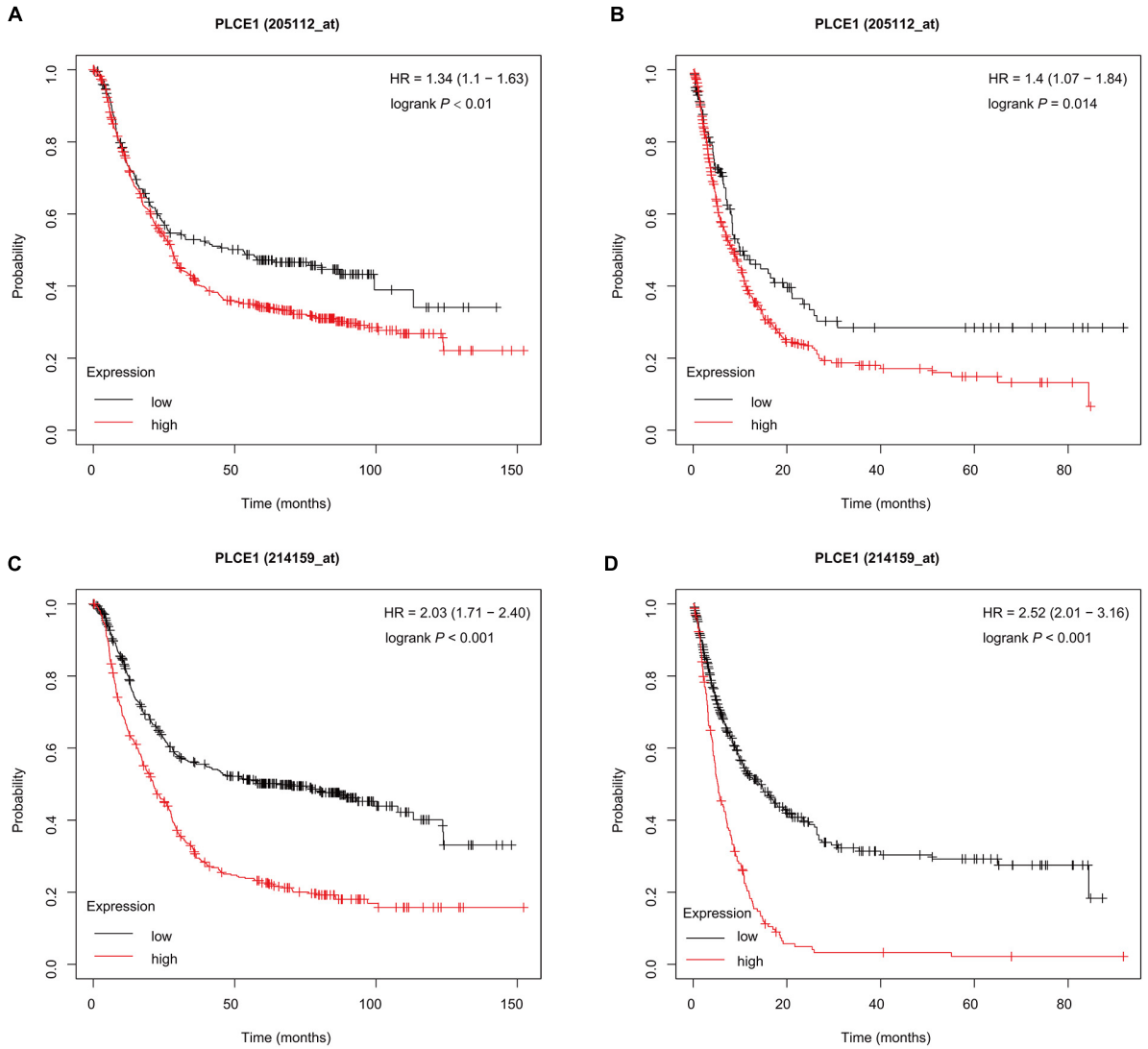


FIGURE 1 | Prognostic characteristic (Kaplan-Meier plotter) of mRNA expression of PLCE1 in gastric cancer patients. The OS curves **(A,C)** and PFS curves **(B,D)** comparing patients with high-expression (red) and low-expression (black) of PLCE1 in gastric cancer by two probes (205112 and 214159) were plotted using the Kaplan-Meier plotter database according to the threshold of *P*-value of < 0.05.

Differences in the frequency distribution of SNPs genotypes and alleles between GC patients and healthy controls were compared by Pearson Chi-squared test and odds ratios (ORs) to evaluate the associations with GC risk, as displayed in **Supplementary Table 1**. The minor allele of each SNP as a risk factor was compared to the wild-type (major) allele. Remarkably, we found that the allele frequency of rs2274223 locating in the exon region was significantly different between GC cases and healthy controls [OR (95% CI) = 1.20 (1.00–1.45), $P = 0.048$]. What's more, the genotype of rs3781264 in the intron region was also significantly different between the two groups [OR (95% CI) = 1.43 (1.16, 1.76), $P = 0.001$].

Then, we analyzed the associations between SNPs with the risk of GC, which was displayed in **Table 3**. Three polymorphisms (rs3765524, rs2274223, rs3781264) in *PLCE1* increased the risk of GC, which were identified through the dominant model (rs3765524, TT-TC vs. CC, OR = 1.28, 95% CI = 1.03–1.60, $P = 0.034$; rs2274223, GG-GA vs. AA, OR = 1.30, 95% CI = 1.04–1.62, $P = 0.025$; rs3781264, CC-CT vs. TT, OR = 1.53, 95% CI = 1.21–1.95, $P = 0.001$), and codominant model [rs3781264, CC vs. CT, OR (95% CI) = 1.01 (0.49–2.08), CC vs. TT, OR (95% CI) = 1.54 (0.76–3.13), $P = 0.002$).

Bioinformatics Analysis of *PLCE1*

The Prognostic Value of *PLCE1* in GC

The K-M plotter (**Figure 1**) showed that patients with *PLCE1* high-expression had lower OS and PFS in both two data sets [data set 205112, OS, HR (95% CI) = 1.34 (1.1–1.63), $P < 0.01$, PFS, HR (95% CI) = 1.4 (1.07–1.84), $P = 0.01$; data set 214159, OS, HR (95% CI) = 2.03 (1.71–2.40), $P < 0.001$, PFS, HR (95% CI) = 2.52

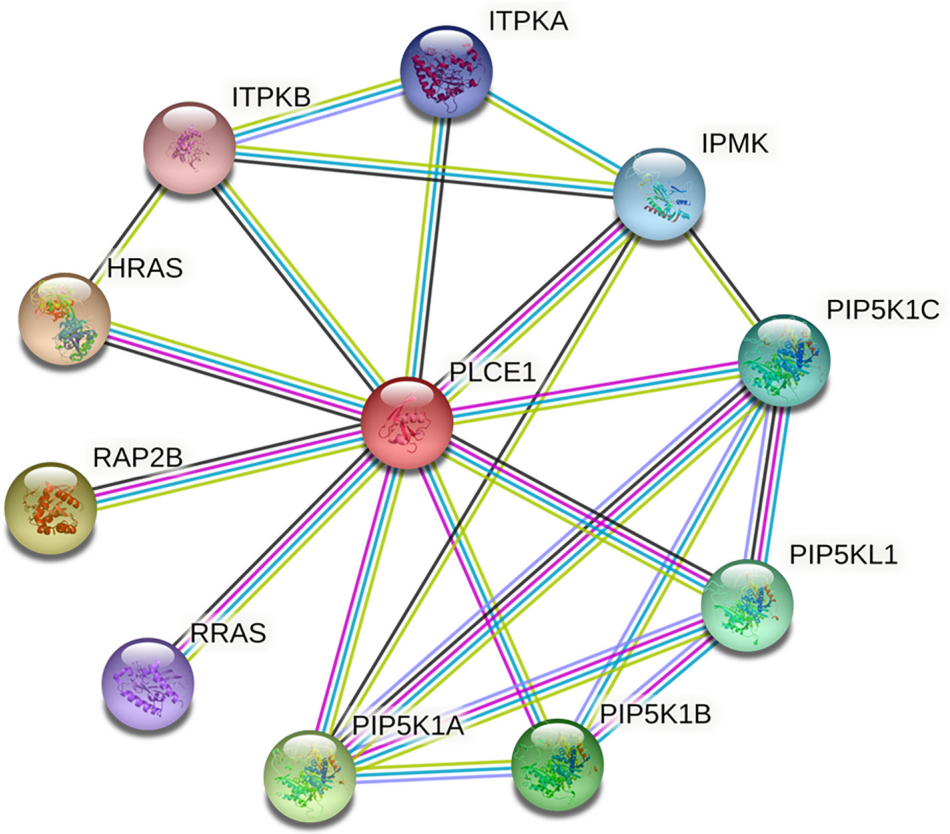


FIGURE 2 | The core protein-protein interaction (PPI) network of PLCE1 constructed by STRING database. It was consisted by 11 nodes and 22 edges with the average node degree of 4 in gastric cancer. The nodes represent proteins; the edges represent interactions between proteins in the network. PLCE1, 1-phosphatidylinositol 4, 5-bisphosphate phosphodiesterase epsilon-1; HRAS, GTPase Hras; ITPKB, inositol-trisphosphate 3-kinase B; ITPKA, inositol-trisphosphate 3-kinase A; IPMK, inositol polyphosphate multikinase; PIP5K1C, phosphatidylinositol 4-phosphate 5-kinase type-1 gamma; PIP5KL1, phosphatidylinositol 4-phosphate 5-kinase-like protein 1; PIP5K1B, phosphatidylinositol 4-phosphate 5-kinase type-1 beta; PIP5K1A, phosphatidylinositol 4-phosphate 5-kinase type-1 alpha; RRAS, Ras-related protein R-Ras; RAP2B, Ras-related protein Rap-2b.

(2.01–3.16), $P < 0.001$], which indicated that *PLCE1* increased the risk of a poor prognosis in GC patients.

PLCE1 PPI Analysis

We investigated the PPI network of PLCE1 by STRING website, and we obtained the core network constructed by 11 nodes and 22 edges with an average node degree of 4 ($P = 0.004$; **Figure 2**). The interaction proteins with PLCE1 were PIP5K1A, PIP5K1B, PIP5K1C, PIP5KL1, IPMK, ITPKA, ITPKB, HRAS, RAP2B, and RRAS.

GO and KEGG Enrichment Analysis

We also analyzed GO and KEGG enrichment by DAVID to explore the potential role of PLCE1. We found that the top five enrichments related to BP were signal transduction (GO:0007165), inositol phosphate metabolic process (GO:0043647), phosphatidylinositol phosphorylation (GO:0046854), phosphatidylinositol biosynthetic process (GO:0006661) and Ras protein signal transduction (GO:0007265) (**Figure 3**). Additionally, we found CC, such as cytosol (GO:0005829), membrane (GO:0016020), uropod (GO:0001931), focal adhesion (GO:0005925) and ruffle membrane (GO:0032587), were significantly associated with PLCE1. Moreover, PLCE1 also affected MF, including ATP binding (GO:0005524), 1-phosphatidylinositol-4-phosphate 5-kinase activity (GO:0016308), 1-phosphatidylinositol-3-phosphate 4-kinase activity (GO:0052811), inositol-1, 4, 5-trisphosphate 3-kinase activity (GO:0008440) and GTP binding (GO:0005525). KEGG enrichment analysis revealed that PLCE1 had a high correlation with inositol phosphate metabolism, phosphatidylinositol signaling system, choline metabolism in cancer, Fc gamma R-mediated phagocytosis, regulation of actin cytoskeleton, endocytosis, calcium signaling pathway, proteoglycans in cancer, rap1 signaling pathway and Ras signaling and metabolic pathways (**Figure 4**). All enrichment results could be attached in **Supplementary Materials**.

PLCE1 Protein Spatial Structure Changes

We modeled the primary PLCE1 protein by SWISS-MODEL. The original (wild-type) model of PLCE1 was shown in **Figure 5A**.

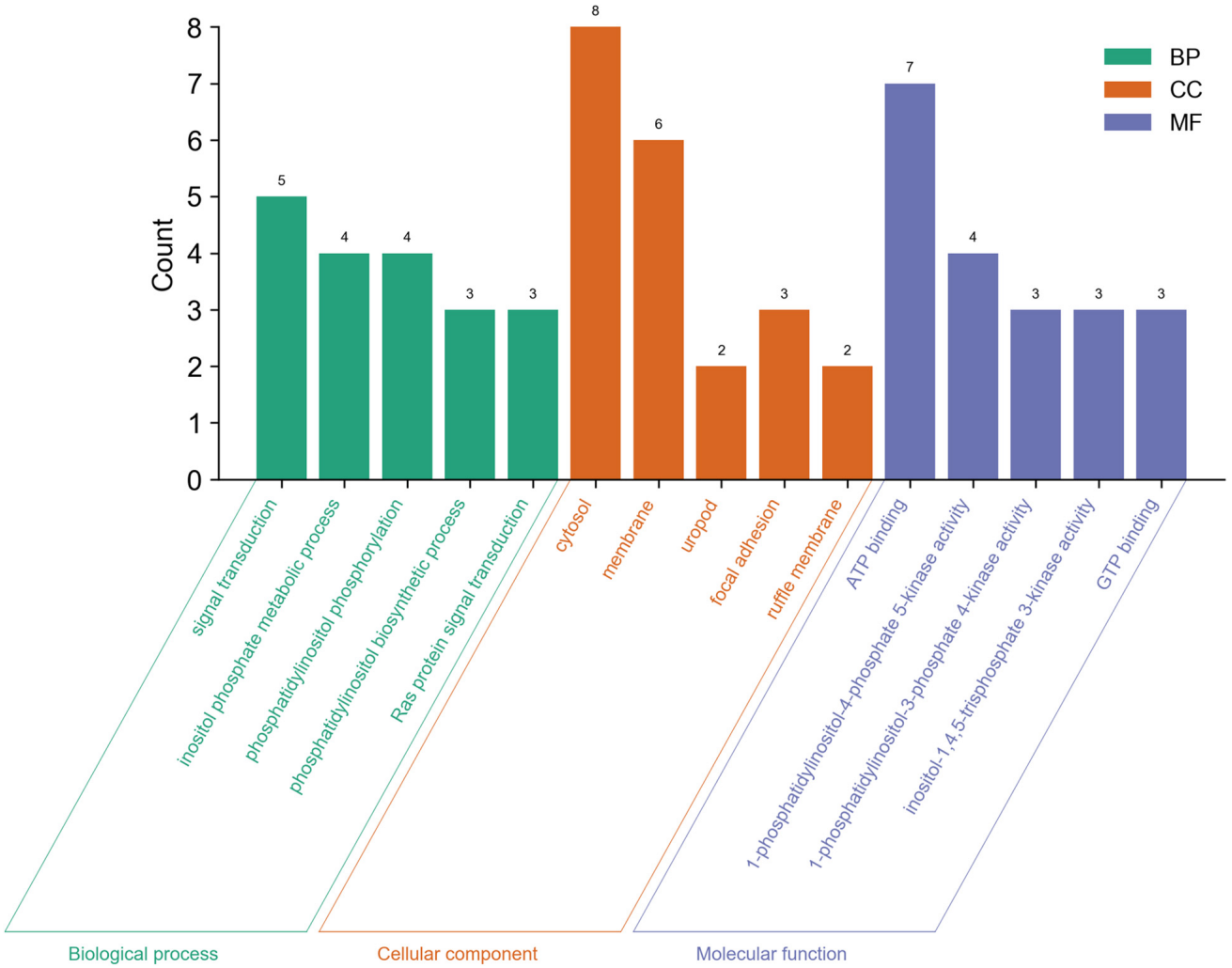


FIGURE 3 | The GO enrichment analysis of PLCE1 and its co-expression genes by DAVID database. BP (biological process) was marked in green; CC (cellular component) was in orange; and MF (molecular function) was in purple.

The protein was colored from blue to red, representing the coiled peptide chain from N- to C-terminal. We found that the PLCE1 protein had two crucial functional domains, namely the calcium ion binding pocket (related to activity), which is composed of 1,873, 1,897, 1,926, 1,928, and 1,933 aa sites (red in **Figure 5B**), and the catalytic binding pocket (related to catalytic efficiency), consisting of 1,391, 1,392, 1,421, 1,423, 1,436, 1,470, 1,637, 1,639, 1,743, 1,770, and 1,772 aa sites (orange in **Figure 5B**).

Hence, the rs2274223 (A > G) changed the aa at the 1927 site, which may affect the activity of the calcium-binding pocket (yellow in **Figure 5C**). Similarly, the mutation of rs3765524 (T > C) enabled the aa at the 1,771 site to change, influencing the catalytic efficiency of the catalytic binding pocket (green in **Figure 5C**).

Interestingly, in further analysis of the impact of the single aa mutation on the protein microenvironment, we found that the ARG1927, in the wild type, formed two ionic bonds with MET1901 and SER1903, respectively (**Figure 5D**), making the interaction force between the two loops extremely tight. However, the mutation (A > G) of rs2274223 resulted in Arg1927His in PLCE1 protein, displayed in **Figure 5E**; although it still formed ionic bonds with these two aa residues after the mutation, one of them was located on the loop of the 1,927 site itself and formed a conjugate bond, causing the attraction between the residues to be stronger than the original one. Consequently, the loop in 1,927 would be tighter than before, and the calcium-binding pocket was more difficult to open after the mutation, leading to the decrease of the protein (PLCE1) activity.

Likewise, in the wild-type, Ile1771 formed two ionic bonds with Gln1687, Val1689 residues, respectively. The interaction force between the ionic bond and the left loop was tight, but no force existed between the loops on the right to "fix" (**Figure 5F**), so it would be easier for the dissociation in a solution or the local changes, facilitating the substrate entered the active center readily. However, rs3765524 (T > C) lead to

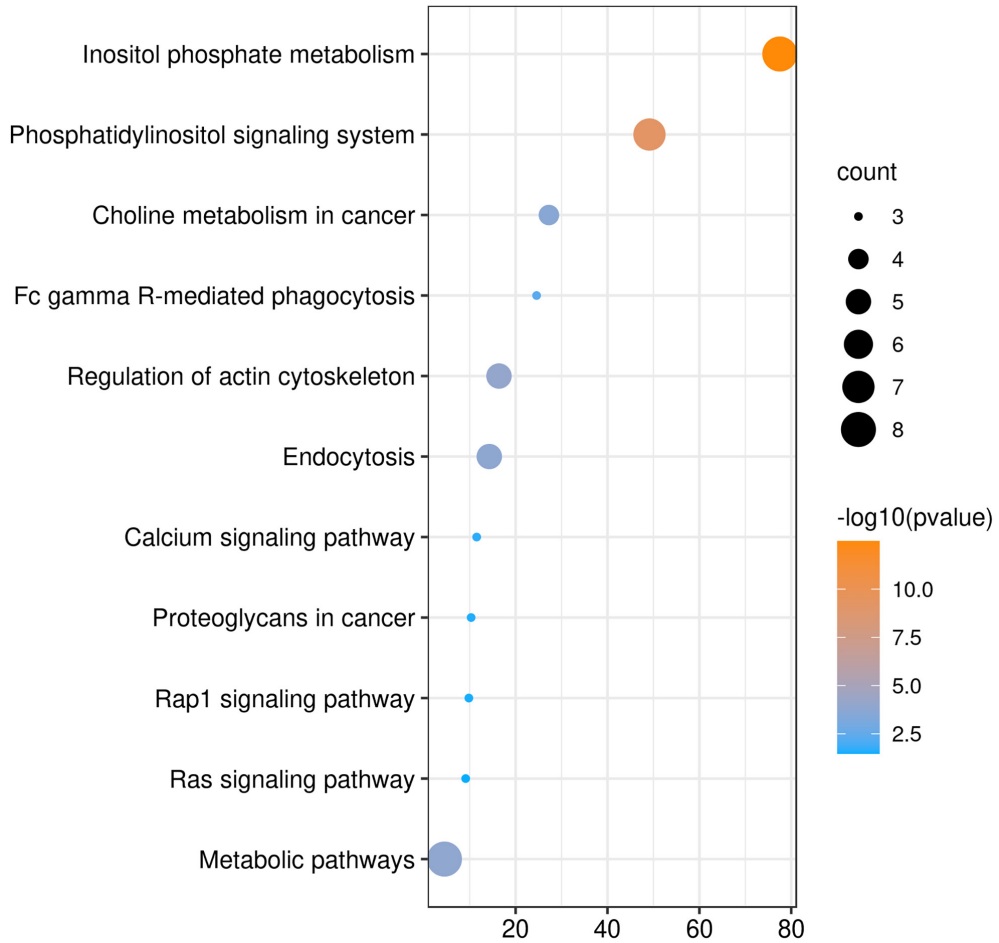


FIGURE 4 | The KEGG enrichment analysis of PLCE1 and its co-expression genes by DAVID database. The size of the circle represents the counts of genes enriched, and the larger the circle, the more genes were enriched. From orange to blue, –10log (*P*-value) gradually decreased.

the Ile1771Thr, which generated four ionic bonds with the four aa residues (Gln1687, Val1689, Ser1772, and Leu1798) in the surrounding space, two of which located on the left loop and the others on the right loop, making the local structure more stable, so the change of the catalytic pocket seemed to be more challenging (**Figure 5G**).

These variations mentioned above combined with the results of bioinformatics analysis indicated that SNPs in *PLCE1* could change the catalytic activity of the protein in Ca^{2+}-related pathways, so more substrates (such as Ca^{2+}) might be required to perform normal functions, which will be verified in our future studies.

DISCUSSION

As a common genetic variation in human genome, SNP is beneficial for understanding the possible relationships between tumors and individuals' biological functions on a genomic scale. It provides a comprehensive tool for identifying candidate genes of cancer, offering fundamental knowledge for clinical diagnosis and revealing drug discovery for relevant genetic diseases; therefore, SNP is considered as a kind of commendable biological marker in diverse tumors (Engle et al., 2006).

Protein is an indispensable carrier of various biological activities and plays a crucial role in the smooth progress of diverse life courses. The primary structure of a protein is aa sequence, which is derived from gene transcription and translation. It is the basis of a high-order structure of a protein and determines the spatial structure and functional properties of a protein. When a SNP is present in a gene, the expressed aa sequence may change, resulting in a change in the spatial structure of the protein. Therefore, it is imperative to study the risk of SNPs and GC from the perspective of protein spatial structure changes, which will contribute to the research on the pathogenesis and prognosis of GC.

In this study, for the first time, we analyzed the correlation between SNPs and GC susceptibility and prognosis in terms of protein spatial structure changes. Firstly, we carried on a case-control study, and by detecting and analyzing the differences on SNPs of *PLCE1* between GC patients and healthy controls, we found that rs3765524 (C > T), rs2274223 (A > G),

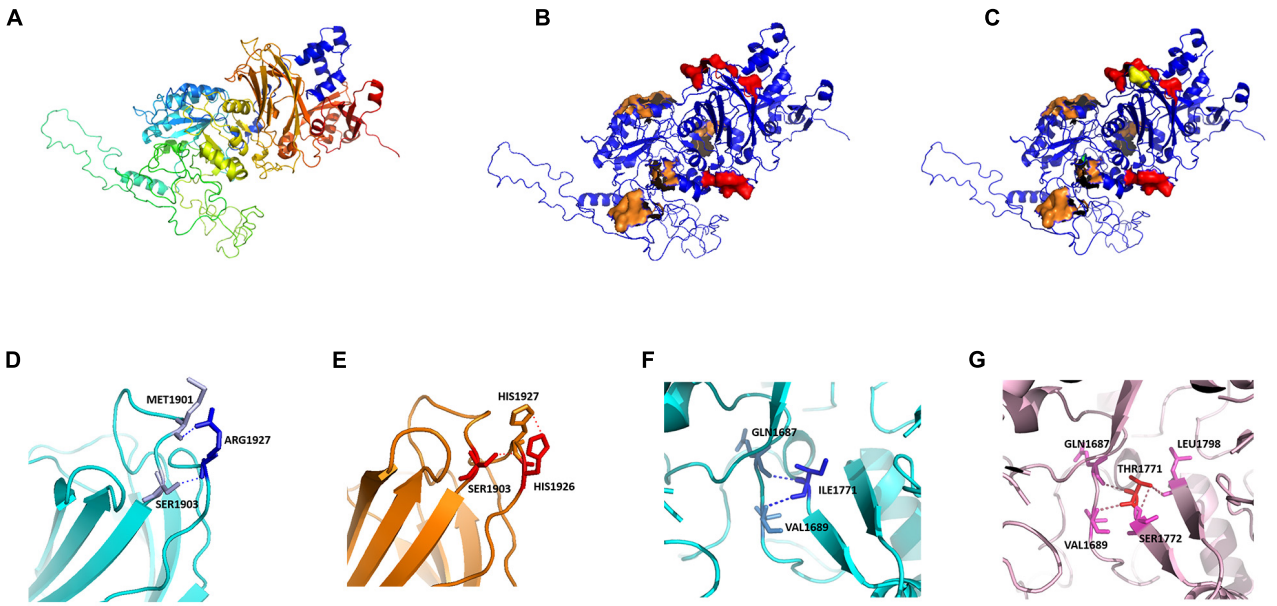


FIGURE 5 | Protein homology modeling of PLCE1 by SWISS-MODEL. **(A)** The spatial protein model of the original (wild-type) PLCE1. From blue to red showed the coiled peptide chain from N- to C- terminal. **(B)** Exploration of crucial functional domains of PLCE1. The calcium ion binding pocket was marked in red, and the catalytic binding pocket was in orange. **(C)** The spatial protein model of the mutant PLCE1 protein. The mutation of rs2274223 (A > G) changed the 1927aa, which may affect the activity of the calcium-binding pocket (yellow). The mutation of rs3765524 (T > C) changed the 1771aa, which may influence the catalytic efficiency of the catalytic binding pocket (green). **(D)** The wild-type protein microenvironment analysis of PLCE1 on the single 1927aa site. ARG1927 formed two ionic bonds with MET1901 and SER1903, respectively, making the force between the two loops very tight. **(E)** The microenvironment analysis of the mutation (rs2274223 A > G) of PLCE1. Arg1927His-mutant of PLCE1 formed ionic bonds with these two aa residues; one of them was located on the loop of the 1927 site itself and formed a conjugate bond, making the attraction between the residues stronger than the wild-type. **(F)** The wild-type protein microenvironment analysis of PLCE1 on the single 1771aa site. Ile1771 formed two ionic bonds with Gln1687 and Val1689 residues, respectively. **(G)** The microenvironment analysis of the mutation (rs3765524 T > C) of PLCE1. Ile1771Thr mutant formed four ionic bonds with Gln1687, Val1689, Ser1772, and Leu1798, two of which located on the left loop and the others on the right loop, making the local structure more stable.

and rs3781264 (T > C) were related to the susceptibility of GC. Then, the K-M plotter demonstrated that high-expression of *PLCE1* was associated with poor survival in GC. To explore the potential function of PLCE1, we used a series of bioinformatics tools, investigating the PPI network, GO and KEGG of PLCE1, and found it played a potential role in the calcium signaling pathway. Furthermore, we constructed the primary and mutant protein spatial structures of PLCE1 by homology modeling method, and interestingly, we found that the changes of the protein spatial structure could reduce the catalytic activity, which might mainly influence its function in Ca^{2+}-related pathways. Combined with the bioinformatic results of *PLCE1*, we speculated that *PLCE1* polymorphisms increase GC susceptibility by changing the spatial structure of PLCE1 protein, affecting its activity and catalytic efficiency in the calcium signaling pathway. This hypothesis will be verified in our future experiments.

As a member of the phospholipase C family of proteins, *PLCE1* encodes a phospholipase C enzyme which mediates the hydrolysis reaction of phosphatidylinositol-4,5-bisphosphate to produce the Ca^{2+}-mobilizing second messenger inositol 1,4,5-triphosphate and the protein kinase C-activating second messenger diacylglycerol. It interacts with the proto-oncogene Ras among other proteins (Bunney et al., 2009). The expression of *PLCE1* was significantly related to tumor differentiation degree, invasion depth, lymph node metastasis and distant metastasis (Cui et al., 2014b; Cheng et al., 2017; Yu et al., 2020).

We confirmed the significance of the two SNPs previously reported, rs3765524 and rs2274223, and revealed another SNP in *PLCE1*, rs3781264, through genotyping and logistic regression in this case-control study was associated with the GC risk. Abnet et al. (2010) firstly used GWAS to identify those variants of *PLCE1* had a significant correlation with GC in the Chinese Han population Until now, an increasing number of studies have identified a shared susceptibility locus in *PLCE1* such as rs2274223 A > G and rs3765524 C > T for gastrointestinal cancer (Abnet et al., 2010; Umar et al., 2013; Cui et al., 2014b; Liu et al., 2014; Malik et al., 2014; Mocellin et al., 2015; He et al., 2016; Gu et al., 2018; Li et al., 2018; Liang et al., 2019; Xie et al., 2020), and the most reported SNP of *PLCE1* was the former, but the conclusions lack consistency. A meta-analysis showed that *PLCE1* rs2274223 polymorphism resulted in susceptibility to esophageal and GC in Asians (Umar et al., 2013). However, another study suggested that an increased association between rs2274223 and GC risk among Asian ethnic groups could only be observed in esophageal cancer rather than GC (Xue et al., 2015). The discrepancy probably

results from considerable heterogeneity in these studies as well as gene-gene interaction and gene-environment interaction. A study (Liang et al., 2019) also confirmed our hypothesis at the protein level by immunohistochemistry (IHC), which confirmed that the PLCE1 protein expression was higher in group of rs3765524 CT/TT than in group of rs3765524 CC. Additionally, our study also showed that rs3781264, located on an intron region, had a potential relationship with GC risk, which was scarcely reported before. Hitherto, most the previous studies focus on the correlation between gene SNPs and cancer susceptibility or risk but never explore its mechanism further.

Currently, the diagnosis, treatment and prognosis of GC are usually based on a risk stratification system. The most efficient curative therapeutic option for GC patients is timely adequate surgical resection (Lutz et al., 2012). Besides, chemotherapy, as a way of second-line treatment, can improve overall survival (Kang et al., 2012). Although we have some understanding of carcinogenesis of GC, early diagnosis and appropriate therapy methods on GC patients still remain a major clinical challenge till now (Choi et al., 2003; Ang and Fock, 2014). It is essential for individuals to identify high-risk GC; thus, more precise gene loci associated with it should be explored. In this study, the K-M plotter analysis was performed in the online bioinformatics database, and both two probes showed that the patients with high mRNA expression of *PLCE1* would have a poorer prognosis. It was suggested that *PLCE1* might have the potential to be a biomarker for the prognosis of GC.

The function of a protein is significantly determined by the spatial structure, which is an indispensable part of protein research. In this study, we analyzed the changes of PLCE1 protein spatial structure after mutations by homology modeling method; and we found it had two important functional domains, calcium-binding pocket related to its protein activity and Ca^{2+} binding pockets associated with the efficiency of Ca^{2+}, which were never reported before. Interestingly, the two SNP sites we focused on, rs2274223 and rs3765524, were located on these important domains. The mutation in rs2274223 affected the Ca^{2+} binding pockets, deregulating its bioactivity efficiency related to Ca^{2+}, and the T > C change in rs3765524 resulted in the efficiency decrease in catalytic activity. All these above together altered the structure, stability, and function of PLCE1 protein. Therefore, by our research, we suppose that SNPs of *PLCE1* may have potential significance in the tumorigenesis and progression of GC, perhaps mainly attributed to the changes of the protein activity, but further studies are needed to confirm.

In summary, this study for the first time analyzes the correlation between SNPs of *PLCE1* and GC in terms of protein spatial structure changes, which has a great significance to the diagnosis and treatment for patients with GC. The more complex connections or the subtle crosstalk will be verified in our future paper, and actually, this experiment is being carried out in full swing.

There were some limitations in this study. Firstly, we selected only three SNPs of *PLCE1*, and more other potentially significant loci were not included in this case-control study. Secondly, the prognostic value of *PLCE1* was investigated in the patients from the online database but not the subjects included in our study, which probably caused background heterogeneity. Thirdly, the mechanism of potential significance in the tumorigenesis and progression of GC was based on the bioinformatic results and the protein homology modeling analysis but lack of experimental verification. Therefore, studies *in vitro* and *in vivo* are needed and will be performed in the future to confirm our results, and we hope to contribute to the era of precise diagnosis and individualized treatment of GC.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of The Second Affiliated Hospital of Air Force Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XH, LYu, and GB designed the research. ZY, SC, JX, and SD performed the study. XH, JJ, SP, PY, LYu, and LYa analyzed the results. XH and JJ edited and commented on the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 81572916 and 81502424).

ACKNOWLEDGMENTS

We would like to acknowledge all the participants involved in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.714915/full#supplementary-material

REFERENCES

Abnet, C. C., Freedman, N. D., Hu, N., Wang, Z., Yu, K., Shu, X. O., et al. (2010). A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat. Genet.* 42, 764–767. doi: 10.1038/ng.649

Ang, T. L., and Fock, K. M. (2014). Clinical epidemiology of gastric cancer. *Singapore Med. J.* 55, 621–628. doi: 10.11622/smedj.2014174

Arroyuelo, A., Vila, J. A., and Martin, O. A. (2016). Azahar: a PyMOL plugin for construction, visualization and analysis of glycan molecules. *J. Comput. Aided Mol. Des.* 30, 619–624. doi: 10.1007/s10822-016-9944-x

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

Bunney, T. D., Baxendale, R. W., and Katan, M. (2009). Regulatory links between PLC enzymes and RAS superfamily GTPASES: signalling via PLCepsilon. *Adv. Enzyme Regul.* 49, 54–58. doi: 10.1016/j.advenzreg.2009.01.004

Bunney, T. D., and Katan, M. (2010). Phosphoinositide signalling in cancer: beyond PI3K and PTEN. *Nat. Rev. Cancer* 10, 342–352. doi: 10.1038/nrc2842

Camargo, M. C., Kim, W. H., Chiaravalli, A. M., Kim, K. M., Corvalan, A. H., Matsuo, K., et al. (2014). Improved survival of gastric cancer with tumour Epstein-Barr virus positivity: an international pooled analysis. *Gut* 63, 236–243. doi: 10.1136/gutjnl-2013-304531

Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., Bray, F., et al. (2016). Cancer statistics in China, 2015. *CA Cancer J. Clin.* 66, 115–132. doi: 10.3322/caac.21338

Cheng, Y., Xing, S. G., Jia, W. D., Huang, M., and Bian, N. N. (2017). Low PLCE1 levels are correlated with poor prognosis in hepatocellular carcinoma. *Onco Targets Ther.* 10, 47–54. doi: 10.2147/OTT.S126401

Choi, N. K., Youn, K. E., Heo, D. S., Lee, S. M., Kim, Y., and Park, B. J. (2003). Stomach cancer incidence, mortality and survival rate in Korean elderly pharmacoepidemiologic cohort (KEPEC) in 1994˜1998. *Cancer Res. Treat.* 35, 383–390. doi: 10.4143/crt.2003.35.5.383

Citro, S., Malik, S., Oestreich, E. A., Radeff-Huang, J., Kelley, G. G., Smrcka, A. V., et al. (2007). Phospholipase CEPSILON is a nexus for Rho and Rap-mediated G protein-coupled receptor-induced astrocyte proliferation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15543–15548. doi: 10.1073/pnas.0702943104

Cui, X. B., Pang, X. L., Li, S., Jin, J., Hu, J. M., Yang, L., et al. (2014a). Elevated expression patterns and tight correlation of the PLCE1 and NF-κB signaling in Kazakh patients with esophageal carcinoma. *Med. Oncol.* 31:791. doi: 10.1007/s12032-013-0791-5

Cui, X. B., Peng, H., Li, S., Li, T. T., Liu, C. X., Zhang, S. M., et al. (2014b). Prognostic value of PLCE1 expression in upper gastrointestinal cancer: a systematic review and meta-analysis. *Asian Pac. J. Cancer Prev.* 15, 9661–9666. doi: 10.7314/apjcp.2014.15.22.9661

Engle, L. J., Simpson, C. L., and Landers, J. E. (2006). Using high-throughput SNP technologies to study cancer. *Oncogene* 25, 1594–1601. doi: 10.1038/sj.onc.1209368

Gresset, A., Sondek, J., and Harden, T. K. (2012). The phospholipase C isozymes and their regulation. *Subcell. Biochem.* 58, 61–94. doi: 10.1007/978-94-007-3012-0_3

Gu, D., Zheng, R., Xin, J., Li, S., Chu, H., Gong, W., et al. (2018). Evaluation of GWAS-identified genetic variants for gastric cancer survival. *EBioMedicine* 33, 82–87. doi: 10.1016/j.ebiom.2018.06.028

He, Y., Wang, C., Wang, Z., and Zhou, Z. (2016). Genetic variant PLCE1 rs2274223 and gastric cancer: more to be explored. *Gut* 65, 359–360. doi: 10.1136/gutjnl-2015-309968

Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Jukic, I., Vukovic, J., Rusic, D., Bozic, J., Bukic, J., Leskur, D., et al. (2021). Adherence to maastricht V/florence consensus report for the management of *Helicobacter* pylori infection among primary care physicians and medical students in Croatia: a cross-sectional study. *Helicobacter* 26:e12775. doi: 10.1111/hel.12775

Kang, J. H., Lee, S. I., Lim, D. H., Park, K. W., Oh, S. Y., Kwon, H. C., et al. (2012). Salvage chemotherapy for pretreated gastric cancer: a randomized phase III trial comparing chemotherapy plus best supportive care with best supportive care alone. *J. Clin. Oncol.* 30, 1513–1518. doi: 10.1200/JCO.2011.39.4585

Li, W. Y., Han, Y., Xu, H. M., Wang, Z. N., Xu, Y. Y., Song, Y. X., et al. (2019). Smoking status and subsequent gastric cancer risk in men compared with women: a meta-analysis of prospective observational studies. *BMC Cancer* 19:377. doi: 10.1186/s12885-019-5601-9

Li, X., Li, X., Jiang, M., Tian, W., and Zhou, B. (2018). Single nucleotide polymorphisms in PLCE1 for cancer risk of different types: a meta-analysis. *Front. Oncol.* 8:613. doi: 10.3389/fonc.2018.00613

Liang, P., Zhang, W., Wang, W., Dai, P., Wang, Q., Yan, W., et al. (2019). PLCE1 polymorphisms and risk of esophageal and gastric cancer in a Northwestern Chinese population. *Biomed. Res. Int.* 2019:9765191. doi: 10.1155/2019/9765191

Lin, X. J., Wang, C. P., Liu, X. D., Yan, K. K., Li, S., Bao, H. H., et al. (2014). Body mass index and risk of gastric cancer: a meta-analysis. *Jpn. J. Clin. Oncol.* 44, 783–791. doi: 10.1093/jjco/hyu082

Liu, X., Zhang, X., Wang, Z., Chang, J., Wu, Z., Zhang, Z., et al. (2014). Genetic polymorphism of the phospholipase C epsilon 1 gene and risk of gastric cancer. *Chin. Med. J.* 127, 2511–2517. doi: 10.3760/cma.j.issn.0366-6999.20133123

Lunet, N., Valbuena, C., Vieira, A. L., Lopes, C., Lopes, C., David, L., et al. (2007). Fruit and vegetable consumption and gastric cancer by location and histological type: case-control and meta-analysis. *Eur. J. Cancer Prev.* 16, 312–327. doi: 10.1097/01.cej.0000236255.95769.22

Lutz, M. P., Zalcberg, J. R., Ducreux, M., Ajani, J. A., Allum, W., Aust, D., et al. (2012). Highlights of the EORTC St. gallen international expert consensus on the primary therapy of gastric, gastroesophageal and oesophageal cancer - differential treatment strategies for subtypes of early gastroesophageal cancer. *Eur. J. Cancer* 48, 2941–2953. doi: 10.1016/j.ejca.2012.07.029

Malik, M. A., Srivastava, P., Zargar, S. A., and Mittal, B. (2014). Phospholipase C epsilon 1 (PLCE1) haplotypes are associated with increased risk of gastric cancer in Kashmir valley. *Saudi J. Gastroenterol.* 20, 371–377. doi: 10.4103/1319-3767.145330

Mocellin, S., Verdi, D., Pooley, K. A., and Nitti, D. (2015). Genetic variation and gastric cancer risk: a field synopsis and meta-analysis. *Gut* 64, 1209–1219. doi: 10.1136/gutjnl-2015-309168

Mou, X., Li, T., Wang, J., Ali, Z., Zhang, Y., Chen, Z., et al. (2015). Genetic variation of BCL2 (rs2279115), neil2 (rs804270), LTA (rs909253), PSCA (rs2294008) and PLCE1 (rs3765524, rs10509670) genes and their correlation to gastric cancer risk based on universal tagged arrays and fe3o4 magnetic nanoparticles. *J. Biomed. Nanotechnol.* 11, 2057–2066. doi: 10.1166/jbn.2015.2113

Plummer, M., Franceschi, S., Vignat, J., Forman, D., and de Martel, C. (2015). Global burden of gastric cancer attributable to *Helicobacter* pylori. *Int. J. Cancer* 136, 487–490. doi: 10.1002/ijc.28999

Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* 31, 3381–3385. doi: 10.1093/nar/gkg520

Shimizu, T., Marusawa, H., Matsumoto, Y., Inuzuka, T., Ikeda, A., Fujii, Y., et al. (2014). Accumulation of somatic mutations in TP53 in gastric epithelium with *Helicobacter* pylori infection. *Gastroenterology* 147, 407–417. doi: 10.1053/j.gastro.2014.04.036

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432. doi: 10.1093/bioinformatics/btq675

Song, C., Hu, C. D., Masago, M., Kariyai, K., Yamawaki-Kataoka, Y., Shibatohge, M., et al. (2001). Regulation of a novel human phospholipase C, PLCepsilon, through membrane targeting by Ras. *J. Biol. Chem.* 276, 2752–2757. doi: 10.1074/jbc.M008324200

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131

Umar, M., Upadhyay, R., and Mittal, B. (2013). PLCE1 rs2274223 A>G polymorphism and cancer risk: a meta-analysis. *Tumour Biol.* 34, 3537–3544. doi: 10.1007/s13277-013-0932-7

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. doi: 10.1093/nar/gky427

Xie, Z., Wang, B., Chai, Y., and Chen, J. (2020). Estimation of associations between 10 common gene polymorphisms and gastric cancer: evidence from a meta-analysis. *J. Clin. Pathol.* 73, 318–321. doi: 10.1136/jclinpath-2019-206189

Xue, W., Zhu, M., Wang, Y., He, J., and Zheng, L. (2015). Association between PLCE1 rs2274223 A > G polymorphism and cancer risk: proof from a meta-analysis. *Sci. Rep.* 5:7986. doi: 10.1038/srep07986

Yu, S., Choi, W. I., Choi, Y. J., Kim, H. Y., Hildebrandt, F., and Gee, H. Y. (2020). PLCE1 regulates the migration, proliferation, and differentiation of podocytes. *Exp. Mol. Med.* 52, 594–603. doi: 10.1038/s12276-020-0410-4

Yuan, S., Chan, H., Filipek, S., and Vogel, H. (2016). PyMOL and INKSCAPE bridge the data and the data visualization. *Structure* 24, 2041–2042. doi: 10.1016/j.str.2016.11.012

Zeng, H., Chen, W., Zheng, R., Zhang, S., Ji, J. S., Zou, X., et al. (2018). Changing cancer survival in China during 2003-15: a pooled analysis of 17 population-based cancer registries. *Lancet Glob. Health* 6, e555–e567. doi: 10.1016/S2214-109X(18)30127-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.

frontiers in Genetics

ORIGINAL RESEARCH
published: 03 September 2021
doi: 10.3389/fgene.2021.714071

Intratumor Epigenetic Heterogeneity—A Panel Gene Methylation Study in Thyroid Cancer

Chaofan Zhu[1,2], Meiying Zhang[2], Qian Wang[2], Jin Jen[3], Baoguo Liu[1] and Mingzhou Guo[2,4]**

[1] Department of Head and Neck Surgery, Peking University Cancer Hospital and Institute, Beijing, China, [2] Department of Gastroenterology and Hepatology, Chinese PLA General Hospital, Beijing, China, [3] Genome Analysis Core, Medical Genome Facility, Center for Individualized Medicine, Mayo Clinic, Rochester, MN, United States, [4] State Key Laboratory of Kidney Diseases, Chinese PLA General Hospital, Beijing, China

Background: Thyroid cancer (TC) is the most common endocrine malignancy, and the incidence is increasing very fast. Surgical resection and radioactive iodine ablation are major therapeutic methods, however, around 10% of differentiated thyroid cancer and all anaplastic thyroid carcinoma (ATC) are failed. Comprehensive understanding the molecular mechanisms may provide new therapeutic strategies for thyroid cancer. Even though genetic heterogeneity is rigorously studied in various cancers, epigenetic heterogeneity in human cancer remains unclear.

Methods: A total of 405 surgical resected thyroid cancer samples were employed (three spatially isolated specimens were obtained from different regions of the same tumor). Twenty-four genes were selected for methylation screening, and frequently methylated genes in thyroid cancer were used for further validation. Methylation specific PCR (MSP) approach was employed to detect the gene promoter region methylation.

Results: Five genes (*AP2*, *CDH1*, *DACT2*, *HIN1,* and *RASSF1A*) are found frequently methylated (>30%) in thyroid cancer. The five genes panel is used for further epigenetic heterogeneity analysis. *AP2* methylation is associated with gender ($P < 0.05$), *DACT2* methylation is associated with age, gender and tumor size (all $P < 0.05$), HIN1 methylation is associated to tumor size ($P < 0.05$) and extra-thyroidal extension ($P < 0.01$). *RASSF1A* methylation is associated with lymph node metastasis ($P < 0.01$). For heterogeneity analysis, *AP2* methylation heterogeneity *is* associated with tumor size ($P < 0.01$), *CDH1* methylation heterogeneity is associated with lymph node metastasis ($P < 0.05$), *DACT2* methylation heterogeneity is associated with tumor size ($P < 0.01$), *HIN1* methylation heterogeneity is associated with tumor size and extra-thyroidal extension (all $P < 0.01$). The multivariable analysis suggested that the risk of lymph node metastasis is 2.5 times in *CDH1* heterogeneous methylation group (*OR* = 2.512, 95% CI 1.135, 5.557, $P = 0.023$). The risk of extra-thyroidal extension is almost 3 times in *HIN1* heterogeneous methylation group (*OR* = 2.607, 95% CI 1.138, 5.971, $P = 0.023$).

OPEN ACCESS

Edited by:
Obul Reddy Bandapalli, Hopp Children's Cancer Center Heidelberg (KiTZ), Germany

Reviewed by:
Jitian Li, Henan Luoyang Orthopedic Hospital, China
Enrique Ambrocio-Ortiz, Instituto Nacional de Enfermedades Respiratorias-México (INER), Mexico

***Correspondence:**
Baoguo Liu
lbg29@163.com
Mingzhou Guo
mzguo@hotmail.com

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *24 May 2021*
Accepted: *16 August 2021*
Published: *03 September 2021*

Citation:
Zhu C, Zhang M, Wang Q, Jen J, Liu B and Guo M (2021) Intratumor Epigenetic Heterogeneity—A Panel Gene Methylation Study in Thyroid Cancer. Front. Genet. 12:714071. doi: 10.3389/fgene.2021.714071

Conclusion: Five of twenty-four genes were found frequently methylated in human thyroid cancer. Based on 5 genes panel analysis, epigenetic heterogeneity is an universal event. Epigenetic heterogeneity is associated with cancer development and progression.

Keywords: epigenetic heterogeneity, DNA methylation, AP2, CDH1, DACT2, HIN1, RASSF1A

INTRODUCTION

Thyroid cancer (TC) is the most common endocrine malignancy and the incidence is 3.4% of all cancers (Seib and Sosa, 2019). Papillary thyroid cancer (PTC) and follicular thyroid cancer (FTC) are two most common thyroid cancer types, account for 80% and 15% of all thyroid cancer cases. Poorly differentiated thyroid carcinoma (PDTC) and anaplastic thyroid carcinoma (ATC) are account for 5% and 1%, respectively. Surgical resection and radioactive iodine ablation are major therapeutic strategies for thyroid cancer. Treatment failure was observed in about 10% of differentiated thyroid cancer and all ATC patients (Laha et al., 2020; Prete et al., 2020). In the last 30 years, genetic study suggests that the frequency of somatic mutations is relatively low in thyroid cancer (Mazzaferri, 1999). The MAPK and PI3K/Akt pathways were reported to be activated by somatic mutations in thyroid cancer (Mercer and Pritchard, 2003; Ball et al., 2007; Leboeuf et al., 2008; Salerno et al., 2009), and targeted therapies were applied to advanced cancers by tyrosine kinase inhibitors and anti-angiogenic drugs (Yamamoto et al., 2014; Valerio et al., 2017). With the understanding of cancer biology and molecular mechanism, precision medicine gets into the main stage of the era. The key to tailor the management of cancer, including thyroid cancer, is based on the better understanding of the molecular pathways. Landscape of cancer genome study has provided a lot of information of genetic alterations and therapeutic targets for cancer therapy. Mutations in different signaling pathways were found by genomic study in thyroid cancer (Vogelstein et al., 2013; Cha and Koo, 2016). DNA methylation was found frequently in human cancers, including esophageal, colorectal, lung, gastric, and hepatic cancers (Esteller et al., 2002; Brock et al., 2003; House et al., 2003; Guo et al., 2006a,b, 2007, 2008; Licchesi et al., 2008; Wang et al., 2009, 2017; Jia et al., 2012, 2013; Hu et al., 2015; Li et al., 2015; Zheng et al., 2017). However, epigenetics was not extensively studied in thyroid cancer. Pan-cancer landscape of aberrant DNA methylation across 24 cancers demonstrated that the methylation frequency was lowest in PTC (Saghafinia et al., 2018). Aberrant methylation of *SOX17* and *DACT2*, the key components of Wnt signaling, were found frequently in thyroid cancer (Li et al., 2012; Zhao et al., 2014). Methylation of *GPX3* was found associated with thyroid cancer metastasis (Zhao et al., 2015). Phenotypic and functional heterogeneity are hallmarks of human cancers (Visvader, 2011). Subpopulations of cancer cells with distinct phenotypic and molecular features within a tumor are called intratumor heterogeneity (ITH) (Bedard et al., 2013). For cancer genetic study, researchers mainly focused on mutational activation of oncogenes or inactivation of tumor-suppressor genes (TSGs). The theory of Darwinian-like clonal evolution of a single tumor was employed to explain the phenomenon of ITH (Nowell, 1976). However, the dominance of gene-centric views is challenged by cancer stem cell hypothesis, and phenotypic variability becomes the major topic of cancer research (Marusyk et al., 2012). The concept of epigenetic silencing being involved in Knudson's two-hit theory was accepted and the causal relevance of epigenetic changes in cancer is being recognized (Nagasaka et al., 2010). The occurrence of abnormal epigenetic change is more frequently than driver mutations in human cancer. In the process of cancer initiation and development, disruption of the "epigenetic machinery" plays an important role. The disruption of "epigenetic machinery" may contribute to tumor phenotype heterogeneity (Guo et al., 2019).

ITH plays an important role in chemotherapeutic resistance (Guo et al., 2019). The best regimen for cancer therapy is to target all different subpopulations of cancer cells at the same time, to avoid chemo-resistance and reduce relapse (Pribluda et al., 2015). Therapeutic failures are often attributed to adaptive responses of cancer stem cells. Environmental and therapeutic pressures may drive transcriptional plasticity through the response of epigenetic regulators to cause durable disease remission in patients for many cancer therapeutic drugs (Dawson, 2017). Epigenetic heterogeneity is more dynamic compare to genetic heterogeneity. Cancer epigenetic heterogeneity is in its infancy and there are very limited studies involved in epigenetic heterogeneity (Dawson, 2017; Guo et al., 2019).

In this study, we evaluated epigenetic heterogeneity by examining promoter region methylation in a panel of genes in primary thyroid cancer. These genes are frequently methylated in thyroid and other cancers, and they are involved in different cancer-related signaling.

MATERIALS AND METHODS

Patients and Specimens

A total of 405 samples from 135 cases of thyroid cancer (46 cases were served as discovery group for methylation screening and all cases were served as validation group) were obtained in Beijing Cancer Hospital from 2018 to 2020 (**Supplementary Table 1**). For heterogeneity analysis, a total of 405 surgical resected thyroid cancer samples were employed (three spatially isolated specimens were obtained from different regions of the same tumor). The median age was 42 years old (range 24–79 years old), including 104 cases of female and 31 cases of male patients. Cancer samples were classified according to the TNM staging system (AJCC2018), including tumor stage I (n = 130), stage II (n = 4), and stage III (n = 1). All of the tissue samples were

immediately snap frozen in liquid nitrogen and preserved at −80°C before analysis. All samples were collected following the guidelines approved by the Institutional Review Board of the Beijing Cancer Hospital.

DNA Extraction, Bisulfite Modification, Methylation Specific PCR, and Screening for Representative Genes

Genomic DNA was extracted from frozen tissues, digested with protease K and then extracted using the standard phenol/chloroform procedure (Cao et al., 2013). Methylation-specific PCR (MSP) primers were designed according to genomic sequences around transcriptional start sites (TSS) and synthesized to detect unmethylated (U) and methylated (M) alleles. Bisulfite treatment was performed as previously described (Herman et al., 1996). MSP primers were listed in **Supplementary Table 2**. MSP amplification conditions were as follows: 95°C 5 min, 1 cycle; 95°C 30 s, 60°C 30 s, 72°C 40 s, 35 cycles; 72°C 5 min, 1 cycle.

To screen representative genes for heterogeneity analysis, twenty-four genes were selected as discovery group, including *MGMT, TIMP3, DAPK, MLH1, TMEM176A, DIRAS1, SFRP1, SFRP2, HIN1, AP2, ER, DACT2, CDH1, RASSF1A, SOX17, GATA4, BCL6B, GPX3, CRBP-1, p16, RUNX3, RARβ, CDX2,* and *WIF1*. All these genes were found frequently methylated in esophageal, colorectal, lung, gastric and hepatic cancers by our previous studies and others, except for *AP2* (an important transcription factor). These genes were well characterized and reported to be involved in different signaling pathways. Among which, *AP2, CDH1, DACT2, HIN1,* and *RASSF1A* genes were found frequently methylated (>30%) and selected for heterogeneity analysis. The workflow is shown in **Supplementary Figure 1**.

Statistical Analysis

The Chi-square or Fisher exact-test was used to analyze the association of gene methylation status and clinical factors. Logistic regression analysis was used to analyze the association of methylation heterogeneity or clinical factors with lymph node Metastasis and Extrathyroidal extension. $P < 0.05^{*}$, $P < 0.01^{**}$, or $P < 0.001^{***}$ was regarded as statistically significant. Data were analyzed by SPSS 22.0 software.

RESULTS

Selection of Frequently Methylated Genes in Human Thyroid Cancer

To explore epigenetic heterogeneity in thyroid cancer, 24 genes, which were found frequently methylated in other cancers, were selected as discovery group to detect 46 cases of thyroid cancer. As shown in **Table 1** and **Figure 1**, the methylation rate is 0% (0/46)–56.52% (26/46). Five genes (*AP2*, *CDH1*, *DACT2*, *HIN1,* and *RASSF1A*) are methylated more than 30% in thyroid cancer. These five genes panel is selected as validation group for methylation heterogeneity analysis.

TABLE 1 | Methylation status of 24 genes in discovery group.

Gene	Thyroid cancer (*n* = 46)
MGMT	0% (0/46)
TIMP3	0% (0/46)
DAPK	2.17% (1/46)
MLH1	2.17% (1/46)
TMEM176A	23.91% (11/46)
DIRAS1	28.26% (13/46)
sFRP1	28.26% (13/46)
sFRP2	10.87% (5/46)
HIN1	32.61% (15/46)
AP2	34.78% (16/46)
ER	4.35% (2/46)
DACT2	45.65% (21/46)
CDH1	47.82% (22/46)
RASSF1A	56.52% (26/46)
SOX17	8.70% (4/46)
GATA4	8.70% (4/46)
BCL6B	28.26% (13/46)
GPX3	17.39% (8/46)
CRBP-1	0% (0/46)
p16	2.17% (1/46)
RUNX3	2.17% (1/46)
RARβ	4.35% (2/46)
CDX2	6.52% (3/46)
WIF1	10.87% (5/46)

Intratumor Epigenetic Heterogeneity in Thyroid Cancer

To evaluate the intratumor epigenetic heterogeneity in thyroid cancer, three different tumor samples are obtain from isolated locations in the same patient. Totally 405 cancer samples are gained from 135 patients. The methylation status of *AP2*, *CDH1*, *DACT2*, *HIN1,* and *RASSF1A* genes is examined by MSP.

The association of promoter region methylation and clinical factors is analyzed by Chi-square tests, including gender, age, tumor size, tumor location, TNM stage, lymph node metastasis, and extra-thyroidal extension. As shown in **Table 2**, *AP2* methylation is associated with gender ($P < 0.05$), while no association were found between *AP2* methylation and age, tumor size, tumor location, TNM stage, lymph node metastasis, and extra-thyroidal extension (all $P > 0.05$). No association are found between *CDH1* methylation and gender, age, tumor size, tumor location, TNM stage, lymph node metastasis, and extra-thyroidal extension (all $P > 0.05$). *DACT2* methylation is associated with age, gender and tumor size (all $P < 0.05$), while no association is found between *DACT2* methylation and tumor location, TNM stage, lymph node metastasis, and extra-thyroidal extension (all $P > 0.05$). *HIN1* methylation is associated with tumor size ($P < 0.05$) and extra-thyroidal extension ($P < 0.01$), while no association is found between *HIN1* methylation and age, tumor location, TNM stage, and lymph node metastasis (all $P > 0.05$). *RASSF1A* methylation is associated with lymph node metastasis ($P < 0.01$), while no association is found between *RASSF1A*

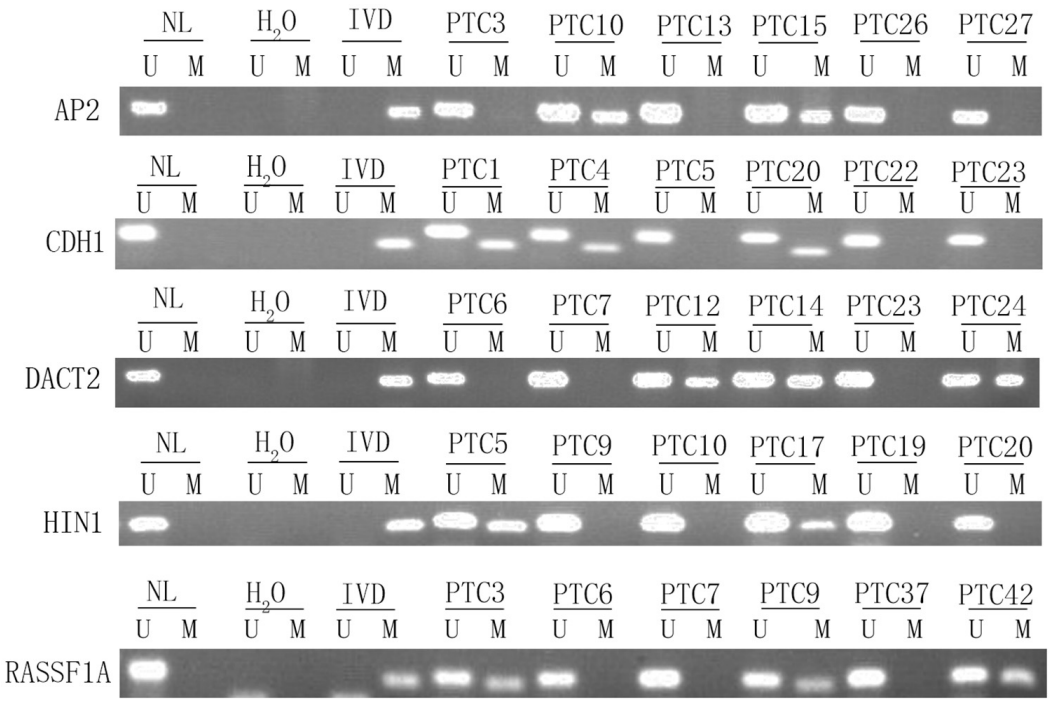


FIGURE 1 | Representative methylation results of *AP2*, *HIN1*, *DACT2, RASSF1A,* and *CDH1* in thyroid cancer (discovery group). IVD, *in vitro*-methylated DNA (methylation control); NL, normal lymphocyte DNA (unmethylation control); H_2O, double distilled water; U, unmethylation; M, methylation; PTC, papillary thyroid cancer.

methylation and age, tumor size, tumor location, TNM stage, and extra-thyroidal extension (all $P > 0.05$).

Unmethylation or methylation in all three samples from the same individual is regarded as homogeneity, while unmethylation or methylation in one or two samples from the same individual is regarded as methylation heterogeneity. As shown in **Figures 2**, **3**, methylation heterogeneity is found in 25.93% (35/135), 37.78% (51/135), 44.44% (60/135), 44.44% (60/135), and 41.48% (56/135) of cases for *RASSF1A*, *CDH1*, *AP2*, *HIN1,* and *DACT2* genes. The results suggest that methylation heterogeneity is a common event in human thyroid cancer.

The association of methylation heterogeneity and clinical factors is further analyzed. As shown in **Table 3**, *AP2* methylation heterogeneity is associated with tumor size ($P < 0.01$), no association are found between *AP2* methylation heterogeneity and gender, age, tumor location, TNM stage, lymph node metastasis and extra-thyroidal extension (all $P > 0.05$). *CDH1* methylation heterogeneity is associated with lymph node metastasis ($P < 0.05$), no association are found between *CDH1* methylation heterogeneity and gender, age, tumor size, tumor location, TNM stage and extra-thyroidal extension (all $P > 0.05$). *DACT2* methylation heterogeneity is associated with tumor size ($P < 0.01$), no association are found between *DACT2* methylation heterogeneity and gender, age, tumor location, TNM stage, lymph node metastasis and extra-thyroidal extension (all $P > 0.05$). *HIN1* methylation heterogeneity is associated with tumor size and extra-thyroidal extension (all $P < 0.01$), no association are found between *HIN1* methylation heterogeneity and gender, age, tumor location, TNM stage and lymph node metastasis (all $P > 0.05$). No association are found between *RASSF1A* methylation heterogeneity and gender, age, tumor size, tumor location, TNM stage, lymph node metastasis and extra-thyroidal extension (all $P > 0.05$). Above results demonstrate that methylation heterogeneity of three genes (*AP2*, *DACT2,* and *HIN1*) is associated with tumor size.

The association of DNA methylation heterogeneity and lymph node metastasis or extra-thyroidal extension is further analyzed by logistic regression model. Univariate logistic analysis indicated that *CDH1* methylation heterogeneity, age and tumor size are associated with lymph node metastasis, independently ($P = 0.026$, $P = 0.000$, $P = 0.040$, **Table 4**). *HIN1* heterogeneous methylation, tumor size and tumor location are associated with extra-thyroidal extension, independently ($P = 0.006$, $P = 0.004$, $P = 0.025$, **Table 5**). Interesting, the multivariable analysis suggested that the risk of lymph node metastasis is 2.5 times in *CDH1* heterogeneous methylation group compare to *CDH1* homogeneous methylation group (*OR* = 2.512, 95% CI 1.135, 5.557, $P = 0.023$, **Table 4**). The risk of extra-thyroidal extension is almost 3 times in *HIN1* heterogeneous methylation group than in *HIN1* homogeneous methylation group (*OR* = 2.607, 95% CI 1.138, 5.971, $P = 0.023$, **Table 5**). The results indicate that with the growing of tumor size, the phenotype of tumor cell becomes diversity. Epigenetic herterogeneity may reflect different phenotypes of population of tumor cells and increased the risk of tumor metastasis significantly. Epigenetic herterogeneity may provide more information for tumor therapeutic strategies.

TABLE 2 | The association of gene methylation and clinical factors in validation group.

		AP2			CDH1			DACT2			HIN1			*RASSF1A*		
	No. ***n* = 135**	***M*** ***n* = 94**	**U** **n = 41**	***p***	**M** **n = 75**	**U** **n = 60**	***p***	**M** **n = 96**	**U** **n = 39**	***p***	**M** **n = 100**	**U** **n = 35**	***p***	**M** **n = 114**	**U** **n = 21**	***p***
Age (year)																
<55	114	77 (67.5%)	37 (32.5%)	0.219	63 (55.3%)	51 (44.7%)	0.873	77 (67.5%)	37 (32.5%)	0.033*	83 (72.8%)	31 (27.2%)	0.434	95 (83.3%)	19 (16.7%)	0.615
≥55	21	17 (81.0%)	4 (19.0%)		12 (57.1%)	9 (42.9%)		19 (90.5%)	2 (9.5%)		17 (81.0%)	4 (19.0%)		19 (90.5%)	2 (9.5%)	
Gender																
Male	31	17 (54.8%)	14 (45.2%)	0.041*	16 (51.6%)	15 (48.4%)	0.615	17 (54.8%)	14 (45.2%)	0.023*	22 (71.0%)	9 (29.0%)	0.653	29 (93.5%)	2 (6.5%)	0.196
Female	104	77 (74.0%)	27 (26.0%)		59 (56.7%)	45 (43.3%)		79 (76.0%)	25 (24.0%)		78 (75.0%)	26 (25.0%)		85 (81.7%)	19 (18.3%)	
Tumor size (cm)																
≤1 cm	103	75 (72.8%)	28 (27.2%)	0.149	57 (55.3%)	46 (44.7%)	0.928	78 (75.7%)	25 (24.3%)	0.034*	81 (78.6%)	22 (21.4%)	0.030*	86 (83.5%)	17 (16.5%)	0.585
>1 cm	32	19 (59.4%)	13 (40.6%)		18 (56.25%)	14 (43.75%)		18 (56.25%)	14 (43.75%)		19 (59.4%)	13 (40.6%)		28 (87.5%)	4 (12.5%)	
Tumor location																
Left lobe	68	46 (67.6%)	22 (32.4%)	0.614	38 (55.9%)	30 (44.1%)	0.939	49 (72.1%)	19 (27.9%)	0.807	53 (77.9%)	15 (22.1%)	0.302	56 (82.4%)	12 (17.6%)	0.499
Right lobe	67	48 (71.6%)	19 (28.4%)		37 (55.2%)	30 (44.8%)		47 (70.1%)	20 (29.9%)		47 (70.1%)	20 (29.9%)		58 (86.6%)	9 (13.4%)	
TNM stage																
I	130	90 (69.2%)	40 (30.8%)	0.985	72 (55.4%)	58 (44.6%)	1.000	91 (70.0%)	39 (30.0%)	0.342	97 (74.6%)	33 (25.4%)	0.832	110 (84.6%)	20 (15.4%)	0.577
II + III	5	4 (80.0%)	1 (20.0%)		3 (60.0%)	2 (40.0%)		5 (100.0%)	0 (0.0%)		3 (60.0%)	2 (40.0%)		4 (80.0%)	1 (20.0%)	
LNM																
N0	76	56 (73.7%)	20 (26.3%)	0.245	45 (59.2%)	31 (40.8%)	0.332	57 (75.0%)	19 (25.0%)	0.258	57 (75.0%)	19 (25.0%)	0.781	70 (92.1%)	6 (7.9%)	0.005**
N1	59	38 (64.4%)	21 (35.6%)		30 (50.8%)	29 (49.2%)		39 (66.1%)	20 (33.9%)		43 (72.9%)	16 (27.1%)		44 (74.6%)	15 (25.4%)	
Extrathyroidal extension																
Negative	91	64 (70.3%)	27 (29.7%)	0.799	51 (56.0%)	40 (44.0%)	0.870	67 (73.6%)	24 (26.4%)	0.354	74 (81.3%)	17 (18.7%)	0.006**	78 (85.7%)	13 (14.3%)	0.558
Positive	44	30 (68.2%)	14 (31.8%)		24 (54.5%)	20 (45.5%)		29 (65.9%)	15 (34.1%)		26 (59.1%)	18 (40.9%)		36 (81.8%)	8 (18.2%)	

*M, methylation; U, unmethylation; LNM, lymph node metastasis; Chi-square test and Fisher exact test, *p < 0.05, **p < 0.01.*

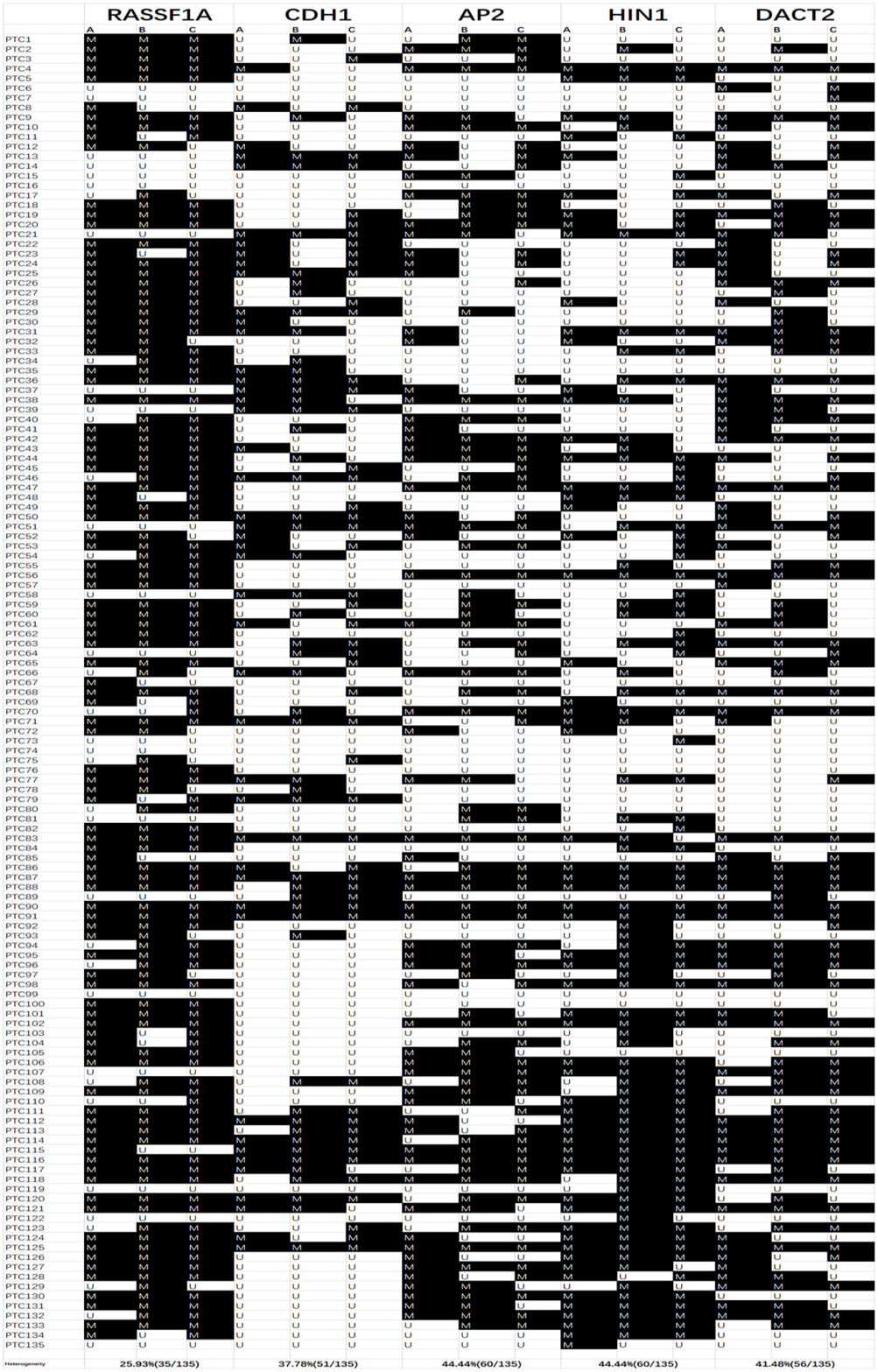


FIGURE 2 | Methylation pattern of 3 different samples for each patient. M, methylation; U, unmethylation; Colum (A–C) are 3 different samples for each case; PTC, papillary thyroid cancer.

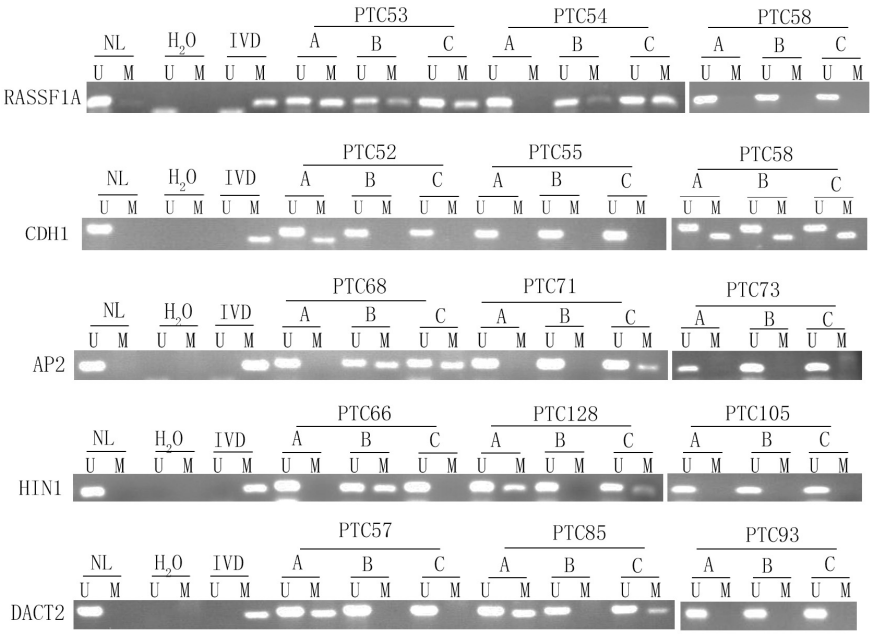


FIGURE 3 | Representative methylation heterogeneity results for each of five genes. IVD, *in vitro*-methylated DNA (methylation control); NL, normal lymphocyte DNA (unmethylation control); H_2O, double distilled water; U, unmethylation; M, methylation; PTC, papillary thyroid cancer.

DISCUSSION

Genomic heterogeneity has been well studied in various cancers (Burrell et al., 2013; Hiley et al., 2016; Maura et al., 2021; Nicholson and Fine, 2021; Wand and Lambert, 2021). However, there are a very few reports about epigenetic heterogeneity in human cancers. Our previous study found that "a field defect of epigenetic changes" was presented in bronchial margins of surgical resected lung cancer samples (Guo et al., 2004). By detecting the methylation status of five genes (*RASSF1A*, *p16*, *DAPK*, *MGMT*, and *Rb*) in 34 tumors (including 15 melanoma primaries, 19 metastases), heterogeneous methylation was found in 70% of the cases (Rastetter et al., 2007). Among 9 MSI-positive primary endometrial cancers lack of MLH1 expression, which was evaluated by immunohistochemistry, Varley et al. (2009) found that 8 cases were methylated in the promoter region. In the 8 cases of methylated patients, four cases were heterogeneously methylated and four tumors were homogeneously methylated. In the discovery group, we select 24 genes, which were well characterized and frequently methylated in various cancers. Five genes are found frequently methylated ($>$ 30%) in 46 cases of human thyroid cancer. In the validation group, total of 405 samples are included (135 cases of patients and 3 samples from different location of each tumor). Among these five genes, *DACT2* methylation is associated with gender, age, and tumor size (all $P < 0.05$), *HIN1* methylation is associated with tumor size ($P < 0.05$) and extra-thyroidal extension ($P < 0.01$), *RASSF1A* methylation is associated with lymph node metastasis ($P < 0.01$). The results suggest that methylation of DACT2, HIN1, and RASSF1A increases the malignance of thyroid cancer. Further analysis find that methylation heterogeneity of *AP2*, *DACT2,* and *HIN1* are associated with tumor size. The results suggest that epigenetic heterogeneity is increasing with tumor growth. It indicates that the phenotype of cancer cells is varied in different tumor stages. Thus, the therapeutic strategies need according to the phenotypes of cancer cells, which are determined by epigenetic changes. The multivariable analysis suggested that *CDH1* methylation heterogeneity is associated with lymph node metastasis and *HIN1* methylation heterogeneity is associated with extra-thyroidal extension. The results suggest that *CDH1* and *HIN1* methylation heterogeneity may increase tumor metastasis. Methylation heterogeneity of these two genes may serve as prognostic marker for thyroid cancer.

The conventional clinical therapeutics is the "one-size-fits-all-approach." However, the ultimate aim of precision medicine is to enable clinicians to accurately and efficiently identify the most effective preventative or therapeutic intervention for a specific patient. Epigenetic switches play important roles in carcinogenesis and tumor progression, and epigenetic switches are reversible (Guo et al., 2019). In this study, we focus mainly on the epigenetic heterogeneity of transcriptional regulators, which were found involved in different cancer-related signaling pathways. Epigenetic heterogeneity is found in four (*CDH1*, *AP2*, *HIN1,* and *DACT2*) of the five detected genes in thyroid cancer. The results suggest that epigenetic heterogeneity is a universal mechanism of cancer development. It is notable that epigenetic heterogeneity is associated with tumor size or tumor metastasis. For the first time, we find that epigenetic heterogeneity is related to cancer development. Based on "BRCAness" principle, our recent study found that methylation of *NRN1* was a novel synthetic lethal marker for PI3K-Akt-mTOR and ATR inhibitors in human esophageal cancer (McLornan et al., 2014;

TABLE 3 | The association of methylation heterogeneity and clinical factors.

		AP2			*CDH1*			*DACT2*			*HIN1*			*RASSF1A*		
	No. n = 135	***HE n = 60***	***HO n = 75***	***p***	***HE n = 51***	***HO n = 84***	***p***	***HE n = 56***	***HO n = 79***	***p***	***HE n = 60***	***HO n = 75***	***p***	***HE n = 35***	***HO n = 100***	***p***
Age (year)																
<55	114	50	64	0.750	43	71	0.974	44	70	0.113	51	63	0.873	31	83	0.434
≥55	21	10	11		8	13		12	9		9	12		4	17	
Gender																
Male	31	13	18	0.749	14	17	0.334	13	18	0.953	13	18	0.749	10	21	0.359
Female	104	47	57		37	67		43	61		47	57		25	79	
Tumor size (cm)																
≤1 cm	103	53	50	0.003**	40	63	0.649	50	53	0.003**	53	50	0.003**	28	75	0.549
>1 cm	32	7	25		11	21		6	26		7	25		7	25	
Tumor location																
Left lobe	68	28	40	0.441	26	42	0.912	29	39	0.782	29	39	0.672	15	53	0.302
Right lobe	67	32	35		25	42		27	40		31	36		20	47	
TNM stage																
I	130	58	72	1.000	50	80	0.651	54	76	1.000	59	71	0.508	33	97	0.604
II + III	5	2	3		1	4		2	3		1	4		2	3	
Lymph node metastasis																
N0	76	32	44	0.535	35	41	0.024*	35	41	0.221	31	45	0.332	20	56	0.907
N1	59	28	31		16	43		21	38		29	30		15	44	
Extrathyroidal extension																
Negative	91	41	50	0.837	35	56	0.814	41	50	0.226	48	43	0.005**	25	66	0.555
Positive	44	19	25		16	28		15	29		12	32		10	34	

*HE, heterogeneity; HO, homogeneity;Chi square test, *p < 0.05, **p < 0.01.*

TABLE 4 | Logistic regression model for lymph node metastasis.

Variable	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*p*	OR (95% CI)	*p*
AP2 (HE vs. HO)	0.805 (0.406, 1.596)	0.535		
CDH1 (HE vs. HO)	2.294 (1.106, 4.761)	0.026*	2.512 (1.135, 5.557)	0.023*
DACT2 (HE vs. HO)	1.545 (0.768, 3.105)	0.222		
HIN1 (HE vs. HO)	0.713 (0.359, 1.414)	0.333		
RASSF1A (HE vs. HO)	1.048 (0.482, 2.279)	0.907		
Age (24–79 years)	0.933 (0.898, 0.969)	< 0.001***	0.927 (0.891, 0.965)	< 0.001***
Gender (female vs. male)	1.080 (0.482, 2.419)	0.852		
Tumor size (0.2–3 cm)	2.190 (1.037, 4.623)	0.040*	2.429 (1.086, 5.436)	0.031*
Tumor location (Right lobe vs. Left lobe)	1.167 (0.591, 2.305)	0.657		
TNM Stage (II + III vs. I)	0.183 (0.020, 1.686)	0.134		
Extrathyroidal extension (positive vs. negative)	1.273 (0.618, 2.626)	0.513		

*HE, heterogeneity methylation; HO, homogeneity methylation; *p < 0.05, ***p < 0.001.*

TABLE 5 | Logistic regression analysis for extrathyroidal extension.

Variable	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*p*	OR (95% CI)	*p*
AP2 (HE vs. HO)	1.079 (0.522, 2.229)	0.837		
CDH1 (HE vs. HO)	1.094 (0.519, 2.305)	0.814		
DACT2 (HE vs. HO)	1.585 (0.751, 3.349)	0.227		
HIN1 (HE vs. HO)	2.977 (1.364, 6.498)	0.006**	2.607 (1.138, 5.971)	0.023*
RASSF1A (HE vs. HO)	1.288 (0.555, 2.989)	0.556		
Age (24–79 years)	1.012 (0.979, 1.047)	0.477		
Gender (female vs. male)	2.031 (0.890, 4.634)	0.092		
Tumor size (0.2–3 cm)	3.249 (1.465, 7.206)	0.004**	2.859 (1.224, 6.679)	0.015*
Tumor location (Right lobe vs. Left lobe)	0.429 (0.204, 0.900)	0.025*	0.373 (0.168, 0.828)	0.015*
TNM stage (II + III vs. I)	0.716 (0.115, 4.448)	0.720		
Lymph node metastasis (positive vs. negative)	1.273 (0.618, 2.626)	0.513		

*HE, heterogeneity methylation; HO, homogeneity methylation; *p < 0.05, **p < 0.01.*

Lord and Ashworth, 2016, 2017; Du et al., 2021). It is reasonable to tailor the regimen for each individual of cancer patients by epigenetic changes or epigenetic heterogeneity.

In conclusion, 5 of 24 genes were found frequently methylated in human thyroid cancer. Based on the 5 genes panel analysis, epigenetic heterogeneity is an universal event. Epigenetic heterogeneity is associated with cancer development and progression.

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

All samples were collected following the guidelines approved by the Institutional Review Board of the Beijing Cancer Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CZ, MZ, and QW designed the project and performed the experiments. MZ and MG wrote the manuscript. MG has made significant contributions to the concept of this study. BL and JJ take part in data interpretation. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Key Research and Development Program of China (Grant Nos. 2018YFA0208902 and 2020YFC2002705), the National Science Foundation of China (NSFC Nos. U1604281 and 81672138), the Beijing Science Foundation of China (BJSFC No. 7171008), and the National Key Scientific Instrument Special Program of China (Grant No. 2011YQ03013405).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.714071/full#supplementary-material

Supplementary Figure 1 | The workflow of epigenetic heterogeneity analysis.

Supplementary Table 1 | Clinical factors of patients.

Supplementary Table 2 | Primers used for methylation analysis.

REFERENCES

Ball, D. W., Jin, N., Rosen, D. M., Dackiw, A., Sidransky, D., Xing, M., et al. (2007). Selective growth inhibition in BRAF mutant thyroid cancer by the mitogen-activated protein kinase kinase 1/2 inhibitor AZD6244. *J. Clin. Endocrinol. Metab.* 92, 4712–4718. doi: 10.1210/jc.2007-1184

Bedard, P. L., Hansen, A. R., Ratain, M. J., and Siu, L. L. (2013). Tumour heterogeneity in the clinic. *Nature* 501, 355–364. doi: 10.1038/nature12627

Brock, M. V., Gou, M., Akiyama, Y., Muller, A., Wu, T. T., Montgomery, E., et al. (2003). Prognostic importance of promoter hypermethylation of multiple genes in esophageal adenocarcinoma. *Clin. Cancer Res.* 9:2912.

Burrell, R. A., McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501, 338–345. doi: 10.1038/nature12625

Cao, B., Yang, Y., Pan, Y., Jia, Y., Brock, M. V., Herman, J. G., et al. (2013). Epigenetic silencing of CXCL14 induced colorectal cancer migration and invasion. *Discov. Med.* 16, 137–147.

Cha, Y. J., and Koo, J. S. (2016). Next-generation sequencing in thyroid cancer. *J. Transl. Med.* 14:322.

Dawson, M. A. (2017). The cancer epigenome: concepts, challenges, and therapeutic opportunities. *Science (New York, NY)* 355, 1147–1152. doi: 10.1126/science.aam7304

Du, W., Gao, A., Herman, J. G., Wang, L., Zhang, L., Jiao, S., et al. (2021). Methylation of NRN1 is a novel synthetic lethal marker of PI3K-Akt-mTOR and ATR inhibitors in esophageal cancer. *Cancer Sci.* 112, 2870–2883. doi: 10.1111/cas.14917

Esteller, M., Guo, M., Moreno, V., Peinado, M. A., Capella, G., Galm, O., et al. (2002). Hypermethylation-associated inactivation of the cellular retinol-binding-protein 1 gene in human cancer. *Cancer Res.* 62, 5902–5905.

Guo, M., House, M. G., Akiyama, Y., Qi, Y., Capagna, D., Harmon, J., et al. (2006a). Hypermethylation of the GATA gene family in esophageal cancer. *Int. J. Cancer* 119, 2078–2083. doi: 10.1002/ijc.22092

Guo, M., House, M. G., Hooker, C., Han, Y., Heath, E., Gabrielson, E., et al. (2004). Promoter hypermethylation of resected bronchial margins: a field defect of changes? *Clin. Cancer Res.* 10, 5131–5136. doi: 10.1158/1078-0432.ccr-03-0763

Guo, M., House, M. G., Suzuki, H., Ye, Y., Brock, M. V., Lu, F., et al. (2007). Epigenetic silencing of CDX2 is a feature of squamous esophageal cancer. *Int. J. Cancer* 121, 1219–1226. doi: 10.1002/ijc.22828

Guo, M., Peng, Y., Gao, A., Du, C., and Herman, J. G. (2019). Epigenetic heterogeneity in cancer. *Biomark. Res.* 7:23.

Guo, M., Ren, J., Brock, M. V., Herman, J. G., and Carraway, H. E. (2008). Promoter methylation of HIN-1 in the progression to esophageal squamous cancer. *Epigenetics* 3, 336–341. doi: 10.4161/epi.3.6.7158

Guo, M., Ren, J., House, M. G., Qi, Y., Brock, M. V., and Herman, J. G. (2006b). Accumulation of promoter methylation suggests epigenetic progression in squamous cell carcinoma of the esophagus. *Clin. Cancer Res.* 12, 4515–4522. doi: 10.1158/1078-0432.ccr-05-2858

Herman, J. G., Graff, J. R., Myöhänen, S., Nelkin, B. D., and Baylin, S. B. (1996). Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9821–9826. doi: 10.1073/pnas.93.18.9821

Hiley, C. T., Le Quesne, J., Santis, G., Sharpe, R., de Castro, D. G., Middleton, G., et al. (2016). Challenges in molecular testing in non-small-cell lung cancer patients with advanced disease. *Lancet (Lond. Engl.)* 388, 1002–1011.

House, M. G., Guo, M., Efron, D. T., Lillemoe, K. D., Cameron, J. L., Syphard, J. E., et al. (2003). Tumor suppressor gene hypermethylation as a predictor of gastric stromal tumor behavior. *J. Gastrointes. Surg.* 7, 1004–1014; discussion 1014.

Hu, S., Cao, B., Zhang, M., Linghu, E., Zhan, Q., Brock, M. V., et al. (2015). Epigenetic silencing BCL6B induced colorectal cancer proliferation and metastasis by inhibiting P53 signaling. *Am. J. Cancer Res.* 5, 651–662.

Jia, Y., Yang, Y., Brock, M. V., Zhan, Q., Herman, J. G., and Guo, M. (2013). Epigenetic regulation of DACT2, a key component of the Wnt signalling pathway in human lung cancer. *J. Pathol.* 230, 194–204. doi: 10.1002/path.4073

Jia, Y., Yang, Y., Zhan, Q., Brock, M. V., Zheng, X., Yu, Y., et al. (2012). Inhibition of SOX17 by microRNA 141 and methylation activates the WNT signaling pathway in esophageal cancer. *J. Mol. Diagn.* 14, 577–585. doi: 10.1016/j.jmoldx.2012.06.004

Laha, D., Nilubol, N., and Boufraqech, M. (2020). New therapies for advanced thyroid cancer. *Front. Endocrinol.* 11:82. doi: 10.3389/fendo.2020.00082

Leboeuf, R., Baumgartner, J. E., Benezra, M., Malaguarnera, R., Solit, D., Pratilas, C. A., et al. (2008). BRAFV600E mutation is associated with preferential sensitivity to mitogen-activated protein kinase kinase inhibition in thyroid cancer cell lines. *J. Clin. Endocrinol. Metab.* 93, 2194–2201. doi: 10.1210/jc.2007-2825

Li, J. Y., Han, C., Zheng, L. L., and Guo, M. Z. (2012). Epigenetic regulation of Wnt signaling pathway gene SRY-related HMG-box 17 in papillary thyroid carcinoma. *Chin. Med. J.* 125, 3526–3531.

Li, Y., Yang, Y., Lu, Y., Herman, J. G., Brock, M. V., Zhao, P., et al. (2015). Predictive value of CHFR and MLH1 methylation in human gastric cancer. *Gastric Cancer* 18, 280–287. doi: 10.1007/s10120-014-0370-2

Licchesi, J. D., Westra, W. H., Hooker, C. M., Machida, E. O., Baylin, S. B., and Herman, J. G. (2008). Epigenetic alteration of Wnt pathway antagonists in progressive glandular neoplasia of the lung. *Carcinogenesis* 29, 895–904. doi: 10.1093/carcin/bgn017

Lord, C. J., and Ashworth, A. (2016). BRCAness revisited. *Nat. Rev. Cancer.* 16, 110–120. doi: 10.1038/nrc.2015.21

Lord, C. J., and Ashworth, A. (2017). PARP inhibitors: synthetic lethality in the clinic. *Science* 355, 1152–1158. doi: 10.1126/science.aam7344

Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer* 12, 323–334. doi: 10.1038/nrc3261

Maura, F., Landgren, O., and Morgan, G. J. (2021). Designing evolutionary-based interception strategies to block the transition from precursor phases to multiple myeloma. *Clin. Cancer Res.* 27, 15–23. doi: 10.1158/1078-0432.ccr-20-1395

Mazzaferri, E. L. (1999). An overview of the management of papillary and follicular thyroid carcinoma. *Thyroid* 9, 421–427. doi: 10.1089/thy.1999.9.421

McLornan, D. P., List, A., and Mufti, G. J. (2014). Applying synthetic lethality for the selective targeting of cancer. *N. Engl. J. Med.* 371, 1725–1735. doi: 10.1056/NEJMra1407390

Mercer, K. E., and Pritchard, C. A. (2003). Raf proteins and cancer: B-Raf is identified as a mutational target. *Biochim. Biophys. Acta* 1653, 25–40. doi: 10.1016/S0304-419X(03)00016-7

Nagasaka, T., Rhees, J., Kloor, M., Gebert, J., Naomoto, Y., Boland, C. R., et al. (2010). Somatic hypermethylation of MSH2 is a frequent event in lynch syndrome colorectal cancers. *Cancer Res.* 70, 3098–3108. doi: 10.1158/0008-5472.CAN-09-3290

Nicholson, J. G., and Fine, H. A. (2021). Diffuse glioma heterogeneity and its therapeutic implications. *Cancer Discov.* 11, 575–590. doi: 10.1158/2159-8290.CD-20-1474

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science (New York, NY)* 194, 23–28. doi: 10.1126/science.959840

Prete, A., Borges de Souza, P., Censi, S., Muzza, M., Nucci, N., and Sponziello, M. (2020). Update on fundamental mechanisms of thyroid cancer. *Front. Endocrinol.* 11:102. doi: 10.3389/fendo.2020.00102

Pribluda, A., de la Cruz, C. C., and Jackson, E. L. (2015). Intratumoral heterogeneity: from diversity comes resistance. *Clin. Cancer Res.* 21, 2916–2923. doi: 10.1158/1078-0432.CCR-14-1213

Rastetter, M., Schagdarsurengin, U., Lahtz, C., Fiedler, E., Marsch, W., Dammann, R., et al. (2007). Frequent intra-tumoural heterogeneity of promoter hypermethylation in malignant melanoma. *Histol. Histopathol.* 22, 1005–1015.

Saghafinia, S., Mina, M., Riggi, N., Hanahan, D., and Ciriello, G. (2018). Pan-cancer landscape of aberrant DNA methylation across human tumors. *Cell Rep.* 25, 1066–1080.e8. doi: 10.1016/j.celrep.2018.09.082

Salerno, P., De Falco, V., Tamburrino, A., Nappi, T. C., Vecchio, G., Schweppe, R. E., et al. (2009). Cytostatic activity of adenosine triphosphate-competitive kinase inhibitors in BRAF mutant thyroid carcinoma cells. *Endocr. Rev.* 30:932. doi: 10.1210/edrv.30.7.9986

Seib, C. D., and Sosa, J. A. (2019). Evolving understanding of the epidemiology of thyroid cancer. *Endocrinol. Metab. Clin. North Am.* 48, 23–35. doi: 10.1016/j.ecl.2018.10.002

Valerio, L., Pieruzzi, L., Giani, C., Agate, L., Bottici, V., Lorusso, L., et al. (2017). Targeted therapy in thyroid cancer: state of the art. *Clin. Oncol. (R. Coll. Radiol.)* 29, 316–324. doi: 10.1016/j.clon.2017.02.009

Varley, K. E., Mutch, D. G., Edmonston, T. B., Goodfellow, P. J., and Mitra, R. D. (2009). Intra-tumor heterogeneity of MLH1 promoter methylation revealed by deep single molecule bisulfite sequencing. *Nucleic Acids Res.* 37, 4603–4612. doi: 10.1093/nar/gkp457

Visvader, J. E. (2011). Cells of origin in cancer. *Nature* 469, 314–322. doi: 10.1038/nature09781

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. Jr., and Kinzler, K. W. (2013). Cancer genome landscapes. *Science (New York, NY)* 339, 1546–1558. doi: 10.1126/science.1235122

Wand, H., and Lambert, S. A. (2021). Improving reporting standards for polygenic scores in risk prediction studies. *Nature* 591, 211–219. doi: 10.1038/s41586-021-03243-6

Wang, J. S., Guo, M., Montgomery, E. A., Thompson, R. E., Cosby, H., Hicks, L., et al. (2009). DNA promoter hypermethylation of p16 and APC predicts neoplastic progression in Barrett's esophagus. *Am. J. Gastroenterol.* 104, 2153–2160. doi: 10.1038/ajg.2009.300

Wang, Y., Zhang, Y., Herman, J. G., Linghu, E., and Guo, M. (2017). Epigenetic silencing of TMEM176A promotes esophageal squamous cell cancer development. *Oncotarget* 8, 70035–70048. doi: 10.18632/oncotarget.19550

Yamamoto, Y., Matsui, J., Matsushima, T., Obaishi, H., Miyazaki, K., Nakamura, K., et al. (2014). Lenvatinib, an angiogenesis inhibitor targeting VEGFR/FGFR, shows broad antitumor activity in human tumor xenograft models associated with microvessel density and pericyte coverage. *Vasc. Cell* 6:18. doi: 10.1186/2045-824X-6-18

Zhao, H., Li, J., Li, X., Han, C., Zhang, Y., Zheng, L., et al. (2015). Silencing GPX3 expression promotes tumor metastasis in human thyroid cancer. *Curr. Protein Pept. Sci.* 16, 316–321. doi: 10.2174/1389203716041504291548 40

Zhao, Z., Herman, J. G., Brock, M. V., Sheng, J., Zhang, M., Liu, B., et al. (2014). Methylation of DACT2 promotes papillary thyroid cancer metastasis by activating Wnt signaling. *PLoS One* 9:e112336. doi: 10.1371/journal.pone.0112336

Zheng, R., Gao, D., He, T., Zhang, M., Zhang, X., Linghu, E., et al. (2017). Methylation of DIRAS1 promotes colorectal cancer progression and may serve as a marker for poor prognosis. *Clin. Epigenetics* 9:50. doi: 10.1186/s13148-017-0348-0

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 Genetics

ORIGINAL RESEARCH
published: 27 September 2021
doi: 10.3389/fgene.2021.732376

CENPN Acts as a Novel Biomarker that Correlates With the Malignant Phenotypes of Glioma Cells

Hailong Wu[1,2†], *Yan Zhou*[1†], *Haiyang Wu*[1†], *Lixia Xu*[3], *Yan Yan*[4], *Xiaoguang Tong*[3,5*] and *Hua Yan*[3*]

[1]Clinical College of Neurology, Neurosurgery and Neurorehabilitation, Tianjin Medical University, Tianjin, China, [2]Department of Neurosurgery, Shijiazhuang Third Hospital, Hebei, China, [3]Tianjin Key Laboratory of Cerebral Vascular and Neurodegenerative Diseases, Tianjin Neurosurgical Institute, Tianjin Huanhu Hospital, Tianjin, China, [4]Clinical Laboratory, Tianjin Huanhu Hospital, Tianjin, China, [5]Department of Neurosurgery, Tianjin Huanhu Hospital, Tianjin, China

Background: Gliomas are the most common intracranial malignant neoplasms and have high recurrence and mortality rates. Recent literatures have reported that centromere protein N (CENPN) participates in tumor development. However, the clinicopathologic significance and biological functions of CENPN in glioma are still unclear.

Methods: Clinicopathologic data and gene expression profiles of glioma cases downloaded from The Cancer Genome Atlas (TCGA) and Chinese Glioma Genome Atlas (CGGA) databases were utilized to determine the associations between the expression of CENPN and clinical features of glioma. Kaplan-Meier and ROC curves were plotted for prognostic analysis. Gene set enrichment analysis (GSEA) and single sample gene set enrichment analysis (ssGSEA) were applied to identify immune-related functions and pathways associated with CENPN' differential expression. *In vitro* experiments were conducted to investigate the impacts of CENPN on human glioma cells.

Results: Elevated CENPN expression was associated with unfavorable clinical variables of glioma patients, which was validated in clinical specimens obtained from our institution by immunohistochemical staining (IHC). The GSEA and ssGSEA results revealed that CENPN expression was strongly correlated with inflammatory activities, immune-related signaling pathways and the infiltration of immune cells. Cell experiments showed that CENPN deficiency impaired cell proliferation, migration and invasion ability and increased glioma apoptosis.

Conclusion: CENPN could be a promising therapeutic target for glioma.

Keywords: CENPN, biomarker, bioinformatics analysis, glioma, immune infiltration

OPEN ACCESS

Edited by:
Jian-Bing Fan,
Illumina, United States

Reviewed by:
Lucia Taja-Chayeb,
National Institute of Cancerology (INCAN), Mexico
Guojun Zhang,
Capital Medical University, China
Jinquan Cai,
Harbin Medical University, China

***Correspondence:**
Hua Yan
yanhua20042007@sina.com
Xiaoguang Tong
tongxg@yahoo.com

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *29 June 2021*
Accepted: *15 September 2021*
Published: *27 September 2021*

Citation:
Wu H, Zhou Y, Wu H, Xu L, Yan Y, Tong X and Yan H (2021) CENPN Acts as a Novel Biomarker that Correlates With the Malignant Phenotypes of Glioma Cells. Front. Genet. 12:732376. doi: 10.3389/fgene.2021.732376

INTRODUCTION

Gliomas are the most common intracranial malignant neoplasms, and they present high recurrence and mortality rates (Ostrom et al., 2018). Although multimodal treatments have been developed, glioma patient's overall survival (OS) is still limited (5 years survival of approximately 5.5%) (OmuroDeAngelis 2013). Molecular markers, such as 1p/19q codeletion status and isocitrate dehydrogenase (IDH) mutations, play valuable roles in tumor formation and progression

(Kristensen et al., 2019; Wang et al., 2020). Individually tailored strategies targeting these biomarkers have been adopted in multiple clinical trials, although few have made breakthroughs. Thus, new targets for glioma therapy are urgently needed.

During mitosis, the assembly of kinetochore proteins at the centromere contributes to the accurate segregation of chromosomes (Kops et al., 2005; McKinleyCheeseman 2016). Centromeric Proteins (CENP), containing 18 subtypes, dynamically associate and dissociate onto centromeric chromatin during mitosis with microtubule regulation (CheesemanDesai 2008; WesthorpeStraight 2014; Hoischen et al., 2018). Among the CENPs, CENPN aggregates toward centromeres in the S phase and eventually dissociates in the G2 phase (Foltz et al., 2006; Carroll et al., 2009; Fang et al., 2015). Recognition of the CENPA nucleosome core region by CENPN is essential for proper chromosome division and kinetochore assembly (Carroll et al., 2009). Recently, the relevance of CENPN to the occurrence and progression of different cancers has been proposed. In oral squamous cell carcinoma, CENPN knockdown arrests the cell cycle in the G1 phase, which leads to the suppression of cellular proliferation (Oka et al., 2019). Based on bioinformatics analyses, Rahman et al. reported that CENPN was a prognostic marker for colorectal cancer (Rahman et al., 2019). Moreover, Sarah An et al. revealed that CENPN expression was involved in the malignant progression of breast cancer patients with a history of smoking (Andres et al., 2015). Most recently, CENPN was shown to promote hepatocellular carcinoma cell proliferation and CENPN deficiency was shown to increase radiotherapy-induced DNA damage (Wang et al., 2021). However, the roles of CENPN in glioma have not been clarified.

In the present study, CENPN served as a prognostic biomarker for glioma based on data obtained from public cancer databases and clinical specimens. Moreover, we investigated the potential functions of CENPN in the immune infiltration and immunoregulation of the glioma microenvironment. *In vitro* cell experiments were conducted to verify the reliability of bioinformatics data and provided a reference for future research.

MATERIALS AND METHODS

Dataset Selection

Next-generation sequencing data (HTSeq-FPKM) and clinical variables were obtained from TCGA (Wang et al., 2016). In addition, data for 1,018 glioma cases generated with the Illumina HiSeq platform were obtained from the CGGA (Zhao et al., 2021). Apart from that, a total of 112 glioma cases enrolled from the Department of Neurosurgery (Huanhu Hospital, Tianjin, China) were used for validation.

Bioinformatics Analysis

In TCGA, the differential expression of CENPN was analyzed using GEPIA, a web database containing abundant normal specimens from GTEx database (Tang et al., 2019). Additionally, GSE16011 (Gravendeel et al., 2009), a dataset from the Gene Expression Omnibus database (GEO), was used for validation. The relationships between CENPN and clinical variables were explored using the beeswarm R package. The survival, survminer and ROC packages were used to plot receiver operating characteristic (ROC) curves and Kaplan–Meier curves.

GSEA was conducted based on the well-established gene sets of c2.cp.kegg.v7.2.symbols.gmt and c5.all.v7.2.symbols.gmt in GSEA 4.0 software. ssGSEA was used to transform the 29 immune signatures into scores for each glioma sample (Subramanian et al., 2005; Bindea et al., 2013; Zhou et al., 2020). Additionally, the CIBERSORT deconvolution algorithm was performed to evaluate the immune infiltration of 22 tumor-infiltrating immune cells (TILs) (Newman et al., 2015). The TIL signatures were downloaded from CIBERSORT (https://cibersortx.stanford.edu/). Relevant R packages were obtained from Bioconductor (http://bioconductor.org/) and CRAN (http://cran.r-project.org/). The bioinformatics analyses were conducted using R software.

Cell Culture and Transfection

The human glioblastoma cell lines LN229 and U251 were purchased from Beijing Beina Chuanglian Biotechnology Institute and cultured in DMEM containing 10% fetal bovine serum (FBS, Gibco, Invitrogen, CA, United States). The siRNAs targeting CENPN mRNA were produced by GenePharma. The sense sequences of these siRNAs were as follows: siCENPN-i, GCGUGCAAGUAUCAGUGAUTT; siCENPN-ii, GACCCU UUGUUACUCAAAUTT; and scramble (siScr), UUCUCC GAACGUGUCACGUTT. siRNA transfections were performed using Lipofectamine 2000 according to the manufacturer's instructions (Thermo Fisher Scientific).

RNA Extraction and Real-Time PCR

TRIzol reagent was used for RNA extraction, and the SureFireRT kit (06-104; Abgen) was applied for RNA reverse transcription. Then, real-time PCR was conducted on a LightCycler 480 II (Roche). The sequences of designed PCR primers were as follows: GAPDH forward, 5′-CAATGACCCCTTCATTGACC-3′ and reverse, 5′-GACAAGCTTCCCGTTCTCAG-3′; CENPN forward, 5′- TGGATGAGACTGTTGCTGA -3′ and reverse, 5′-ACTTGCACGCTTTTCCTC -3′. The PCR conditions were 95°C for 5 min, followed by 40 cycles at 95°C for 10 s and 60°C for 1 min. Relative quantification was calculated by the $2^{-\Delta\Delta Ct}$ method.

Western Blotting and Immunohistochemistry

Cells were lysed using RIPA buffer following the manufacturer's protocols (Solarbio Co., Beijing, China). Protein (50 μg/lane) was added to the 12% SDS-PAGE gel and then transferred to a membrane (PVDF, Merck Millipore). The protein blots on the membrane blocked by BSA were incubated with antibody against CENPN (1:500 dilution; Affinity Biosciences, Jiangsu, China; DF2315) or β-tubulin (1:500 dilution; CST, United States) overnight. After incubation with HRP-linked goat anti-rabbit IgG (CST, United States) for 60 min, the membrane was exposed to enhanced chemiluminescence reagent.

Immunohistochemistry assays were conducted by the pathology department of Huanhu Hospital. Immunostaining of CENPN was performed using a rabbit polyclonal anti-CENPN antibody (1:500 dilution; Affinity Biosciences, Jiangsu, China; DF2315). The positive slides were recognized by integrated scoring, which was performed independently by two experienced pathologists. The glioma tissues were divided into high and low expression groups based on the expression density of CENPN, as previously reported (Xiao et al., 2015).

CCK-8, Colony Formation, Wound Healing, Invasion, Apoptosis and Cell Cycle Analysis

5×10^3 cells per well were seeded in 96-well plates and then transfected with siCENPN. Cells were incubated with CCK-8 reagent (K009-500, ZETA) for 30–60 min at room temperature at 0, 24, 48 and 72 h time points. Eventually, the absorbance changes were observed at 450 nm on a molecular devices microplate reader.

For the wound-healing assay, the bottom of plate was scraped by a pipette tip when 100% cell confluence. Then, wound closure (%) was observed at 0 and 24 h by ImageJ. For the Transwell assay, 5×10^4 cells with or without CENPN knockdown were seeded in the upper chamber (8 μm pore size, Corning, Cambridge, United States) coated with Matrigel (BD Biosciences) while 500 μl of medium containing 10% FBS was placed in the lower chamber. Twenty-four hours later, the upper chamber was immersed in 4% paraformaldehyde for 20 min and then subjected to 20 min crystal violet staining. Eventually, the number of penetrating cells was calculated randomly by three fields of view (FOVs). A colony formation assay was conducted to examine the glioma proliferation ability, with 500 cells per well seeded in six-well plates. Fourteen days later, the cells were fixed with 4% formaldehyde and subjected to crystal violet staining.

Cell apoptosis assay (FITC Annexin V Apoptosis Detection Kit, BD Biosciences, United States) was conducted according to the manufacturer's protocol, as previously described (Zhou et al., 2021). For the cell cycle assay, glioma cells were harvested 48 h following transfection and then fixed with precooled 75% alcohol for 3 h. Subsequently, the cells were incubated with 0.5 ml of PI/RNase reagent (BD Biosciences, United States) at room temperature for 20 min and then analyzed by flow cytometry.

Statistical Analysis

Cases with incomplete information were removed from the corresponding analysis. Student's t-test and one-way analysis of variance (ANOVA) were used to test for significant differences between two groups or multiple groups, respectively. For CENPN protein expression, the chi-square test was employed. Data are presented as the mean ± standard deviation (SD) of at least three independent experiments. $p < 0.05$ was viewed statistically significant. All statistical tests were two-sided. R software v3.6.2. and GraphPad Prism 8 were used for the statistical analyses.

RESULTS

Information of Datasets

In the present study, 1833 glioma cases from the TCGA, CGGA and Huanhu datasets were included. The clinical features of these datasets are summarized in **Table 1**. Pathology reports containing the WHO grade, IDH1 mutation status and other information on the patients in the Huanhu dataset were provided by the Department of Pathology of Huanhu Hospital. The patients in the Huanhu cohort consisted of 44 females and 68 males, and their age ranged from 24 to 82 years old. Detailed information on the patients in the Huanhu cohort is shown in **Supplementary Table S1**.

TABLE 1 | Clinical variables of patients in TCGA, CGGA and Huanhu datasets.

Variables	TCGA	CGGA	Huanhu
Total	703	1,018	112
Age			
≥41	263	191	76
<41	407	558	36
Gender			
Male	651	442	68
Female	460	307	44
Grade			
II	249	218	33
III	265	240	31
IV	596	291	48
Histology			
Astrocytoma	194	150	24
Oligodendroglioma	191	76	41
Oligoastrocytoma	130	232	
Glioblastoma	596	291	47
Mixed glioma	—	—	—
IDH1 mutation			
Yes	91	410	76
No	34	339	36
1p19q codeleted			
Yes	—	155	43
No	—	594	24
KPS			
<80	151	—	—
≥80	584	—	—
Ki-67 index (%)			
<10	—	—	34
≥10	—	—	78
P53 mutation			
Yes	—	—	105
No	—	—	6
MGMT methylation			
Yes	—	—	86
No	—	—	24

mRNA Levels of CENPN in the Databases

The mRNA expression levels of CENPN in glioma were assessed based on GEPIA and GSE16011, a glioma dataset obtained from the Gene Expression Omnibus database. Compared with the values in normal samples, the CENPN mRNA level was markedly higher in the glioma samples (**Figures 1A,B**).

Association Between CENPN and Glioma Clinical Features

To explore the CENPN expression profiles in glioma, the clinical features of glioma patients obtained from TCGA and CGGA were analyzed. As demonstrated in **Figures 1C,E–G**, CENPN was highly expressed in patients older than 41 years and patients

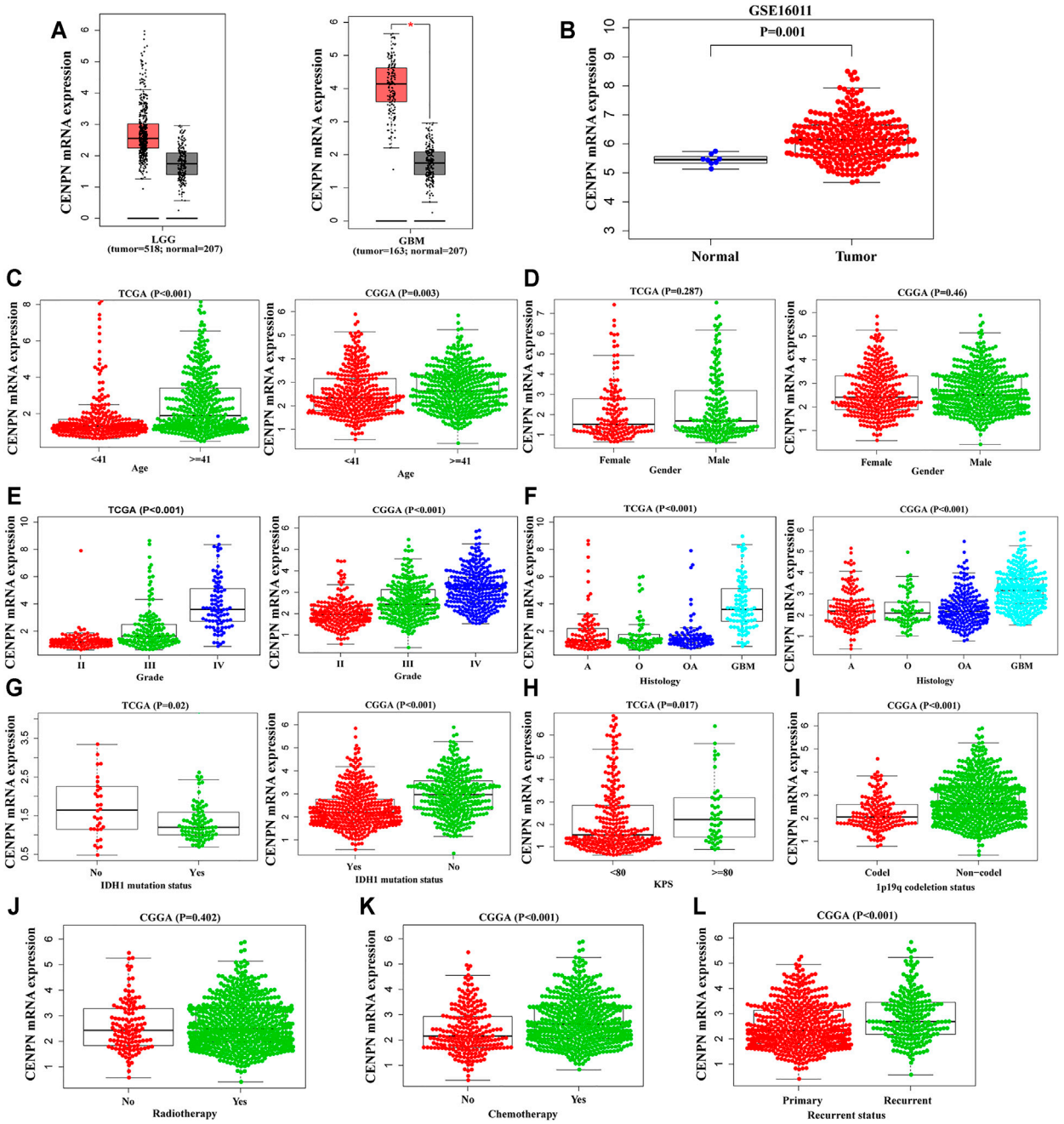


FIGURE 1 | The mRNA expression profiles of CENPN in public databases. **(A,B)** CENPN expression levels in the GEPIA and GSE16011 datasets between tumor and normal tissues. **(C–L)** Associations between CENPN and clinical variables in TCGA and CGGA databases: **(C)** age (year); **(D)** gender; **(E)** WHO grade; **(F)** histological type (A, astrocytoma; O, oligodendroglioma; OA, oligoastrocytoma; GBM, glioblastoma); **(G)** IDH1 mutation status; **(H)** KPS; **(I)** 1p19q codeletion status; **(J)** radiotherapy; **(K)** chemotherapy; and **(L)** recurrence status.

with advanced grades, histopathologies and wild-type IDH1 in both TCGA and CGGA, although no association was found between CENPN expression and gender (**Figure 1D**). In TCGA, CENPN expression increased with higher Karnofsky Performance Status scores (KPS ≥ 80) (**Figure 1H**). In CGGA, some other clinical variables (radiotherapy, 1p19q codeletion, chemotherapy and recurrent status) were also analyzed (**Figure 1I-L**). These results revealed that CENPN upregulation could predict unfavorable glioma.

CENPN Predicts Worse Survival in Glioma

Kaplan-Meier plots were graphed to estimate the prognostic value of CENPN expression. As depicted in **Figure 2A**, CENPN high expression was associated with worse survival of patients in

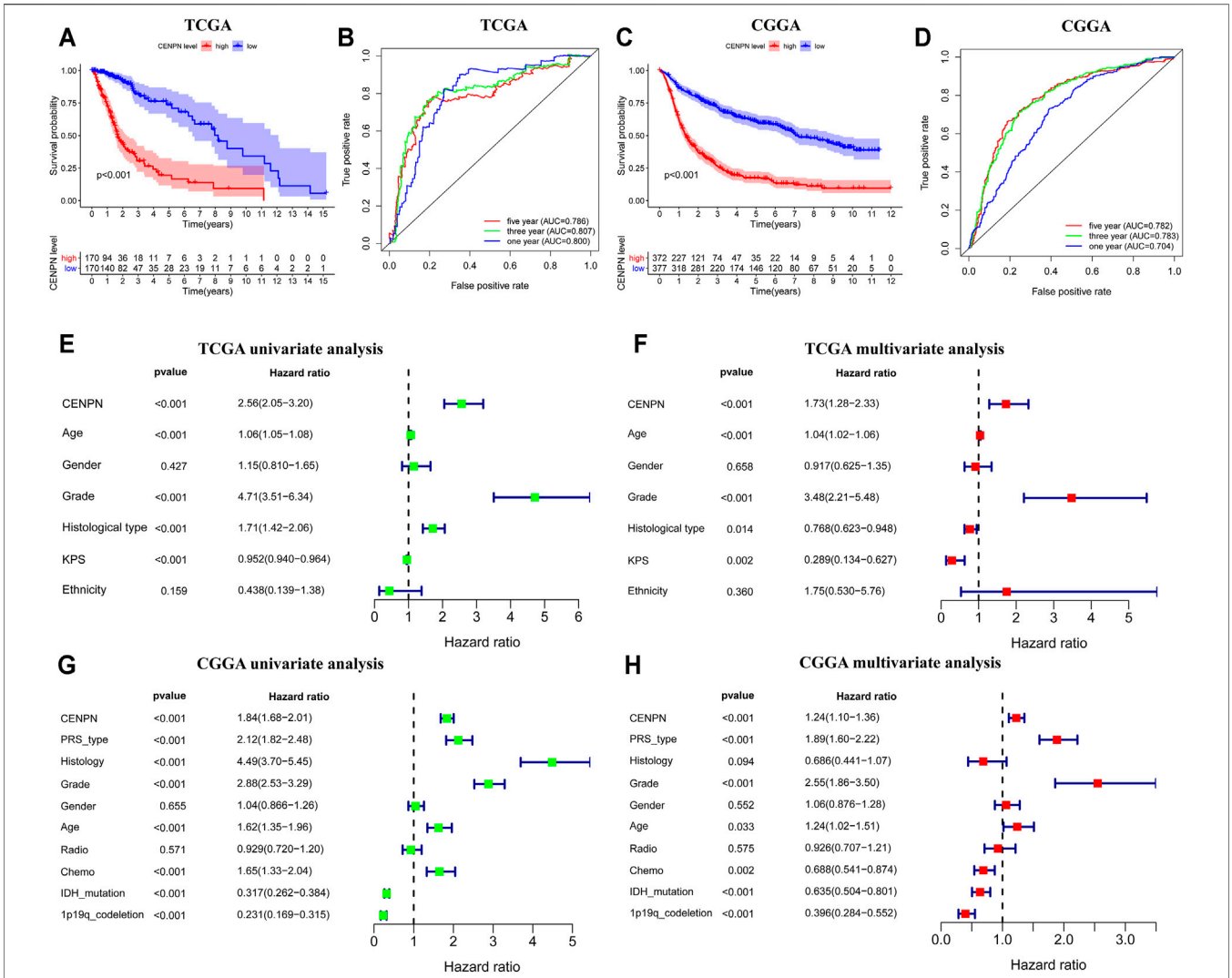


FIGURE 2 | The prognostic values of CENPN in TCGA and CGGA databases. **(A,C)** Kaplan–Meier analysis of OS between groups with different CENPN expression; and **(B,D)** 1, 3- and 5 year ROC curves of OS. Univariate and multivariate Cox regression in TCGA **(E,F)** and CGGA **(G,H)**. Relevant forest plots were plotted.

TCGA. ROC analyses demonstrated that CENPN could be a fine predictor for patient survival: 1 year (AUC = 0.800), 3 years (AUC = 0.807) and 5 years OS (AUC = 0.786) (**Figure 2B**). These results were verified in CGGA dataset (**Figures 2C,D**).

Furthermore, univariate and multivariate Cox regression analyses were employed to assess the independent prognostic roles of CENPN. In TCGA, univariate regression showed that patients with CENPN-high had a shorter overall survival time (hazard ratio (HR): 2.56, 95% CI [2.05–3.20], $p < 0.001$) (**Figure 2E**). The remaining clinical features associated with OS, such as age, grade and histology, are shown in **Figure 2E**. The multivariate regression model demonstrated that CENPN was independently correlated with patient OS (**Figure 2F**, HR = 1.73, 95% CI [1.28–2.33], $p < 0.001$). Moreover, these results were also confirmed in CGGA dataset (**Figures 2G,H**), indicating that CENPN could be an independent risk factor for glioma.

CENPN-Related Biological Processes and Immune Activities

The KEGG analysis based on the GSEA of genes in the high CENPN expression group in TCGA showed that some biological pathways were enriched, including the cell cycle, P53 signaling pathway and apoptosis. Moreover, some pathways related to immunity were also enriched, such as leukocyte transendothelial migration, Fc gamma r-mediated phagocytosis and antigen processing and presentation. The GO analysis indicated that immune-related biological processes, such as immunoglobulin production, leukocyte homeostasis and

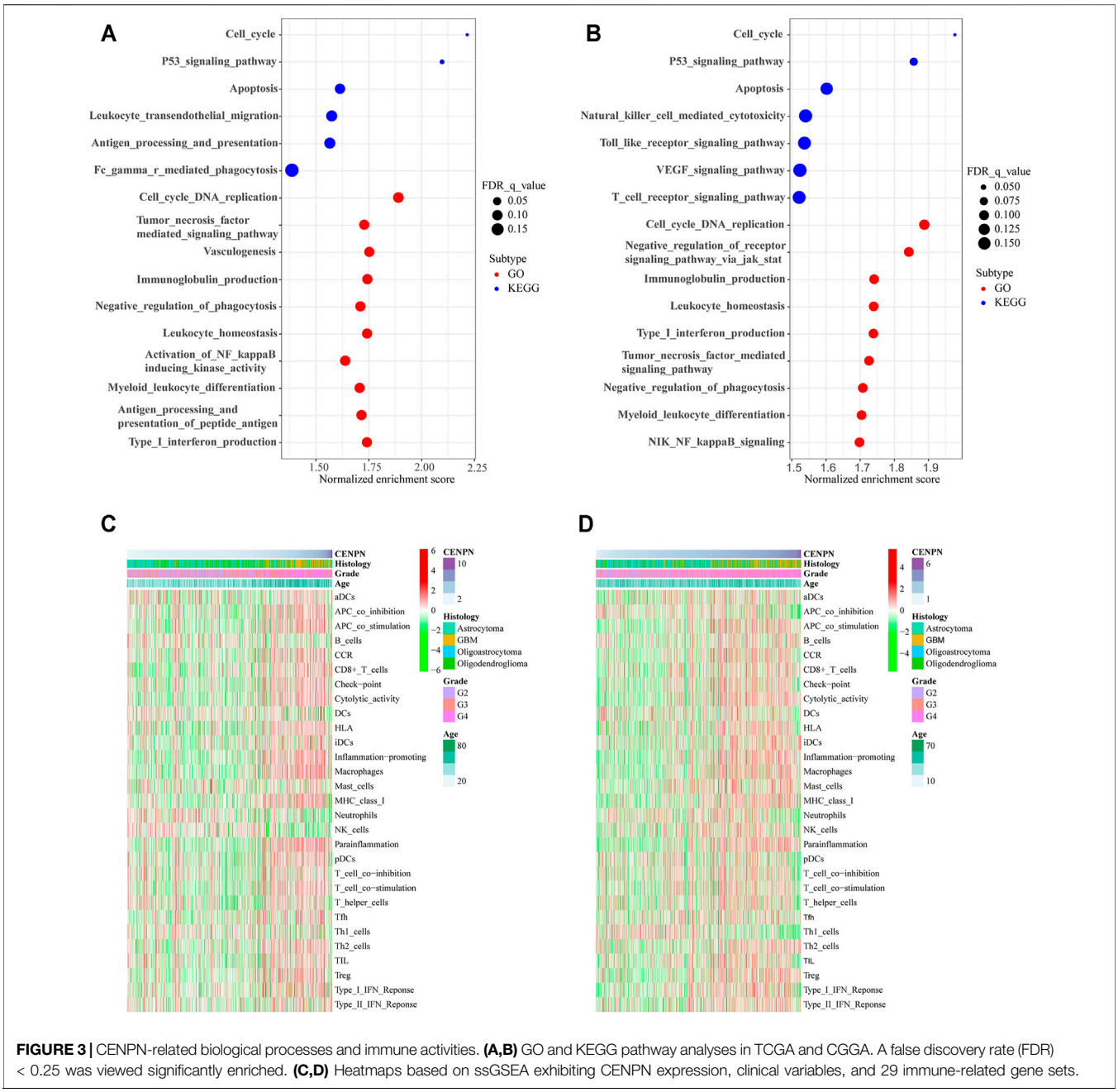


FIGURE 3 | CENPN-related biological processes and immune activities. **(A,B)** GO and KEGG pathway analyses in TCGA and CGGA. A false discovery rate (FDR) < 0.25 was viewed significantly enriched. **(C,D)** Heatmaps based on ssGSEA exhibiting CENPN expression, clinical variables, and 29 immune-related gene sets.

negative phagocytosis regulation, were enriched in the cohort with CENPN-high (**Figure 3A**). Similar results were also verified in the CGGA database (**Figure 3B**). GSEA results suggested that CENPN was associated with multiple molecular mechanisms of glioma development.

To further explore the role of CENPN in the immune microenvironment, ssGSEA was applied to assess the enrichment scores of immune-related pathways, immune cells and immune-related functions grounded on the 29 gene sets derived from previous studies (Subramanian et al., 2005; Bindea et al., 2013; Zhou et al., 2020). In TCGA, the heatmap (**Figure 3C**) shows that some immune-related functions were positively correlated with CENPN expression. APC coinhibition, APC costimulation, chemokine receptor (CCR), cytolytic activity, immune checkpoint, inflammation promoting, human leukocyte antigen (HLA), parainflammation, major histocompatibility complex (MHC) class I, type I interferon (IFN) responses, T cell coinhibition and T cell costimulation were enriched in CENPN-high patients. Additionally, several immune cell types, such as $CD8^+$ T cells, macrophages and regulatory T cells (Tregs), were infiltrated in glioma cases with CENPN-high, while neutrophils, dendritic cells (DCs) and natural killer cells (NK cells) showed opposite trends. Meanwhile, similar results were also obtained in CGGA

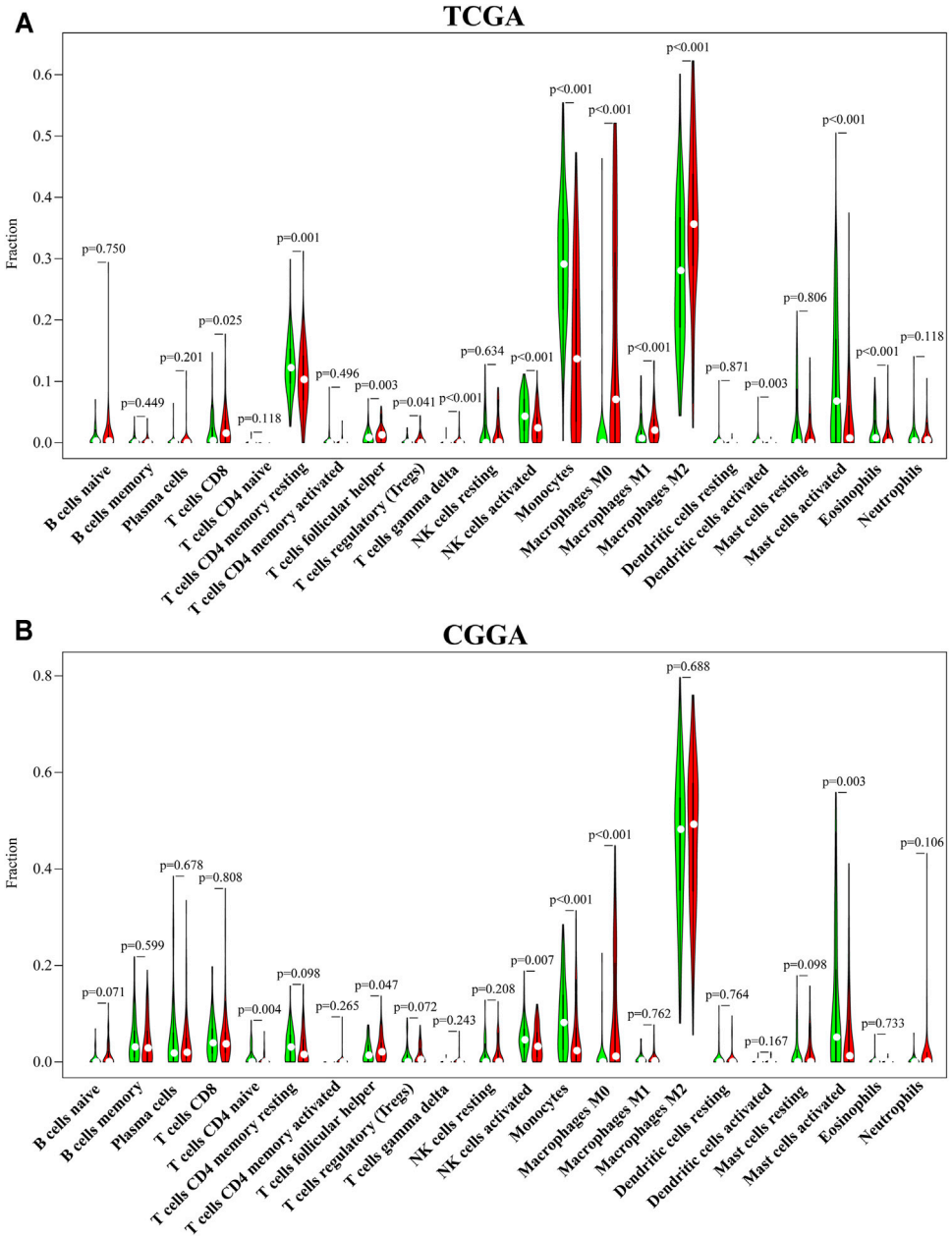


FIGURE 4 | CENPN-related immune infiltration. **(A,B)** The fractions of 22 TILs infiltrating the glioma microenvironment in TCGA and CGGA based on the CIBERSORT algorithm.

database (**Figure 3D**). These results indicated that CENPN has a close relationship with immunomodulation of glioma.

Evaluation of Immune Infiltration

To systematically estimate the relationship between CENPN and tumor-infiltrating lymphocytes (TILs), the CIBERSORT algorithm method was employed to infer the fractions of 22 immune cell types in tumor samples based on gene expression profiles. Consistent with the ssGSEA results, $CD8^+$ T cells, follicular helper T cells, Tregs, gamma delta T cells, M0, M1 and M2 macrophages in TCGA and follicular helper T cells and M0 macrophages in CGGA were enriched in the CENPN-high group. However, memory resting $CD4^+$ T cells, NK cells activated, dendritic cells and mast cells, monocytes and eosinophils in TCGA and CD4 naïve T cells, monocytes, mast cells and NK cells activated in CGGA were enriched in the CENPN-low group (**Figure 4**). These findings indicated an important role of CENPN in the glioma immune microenvironment.

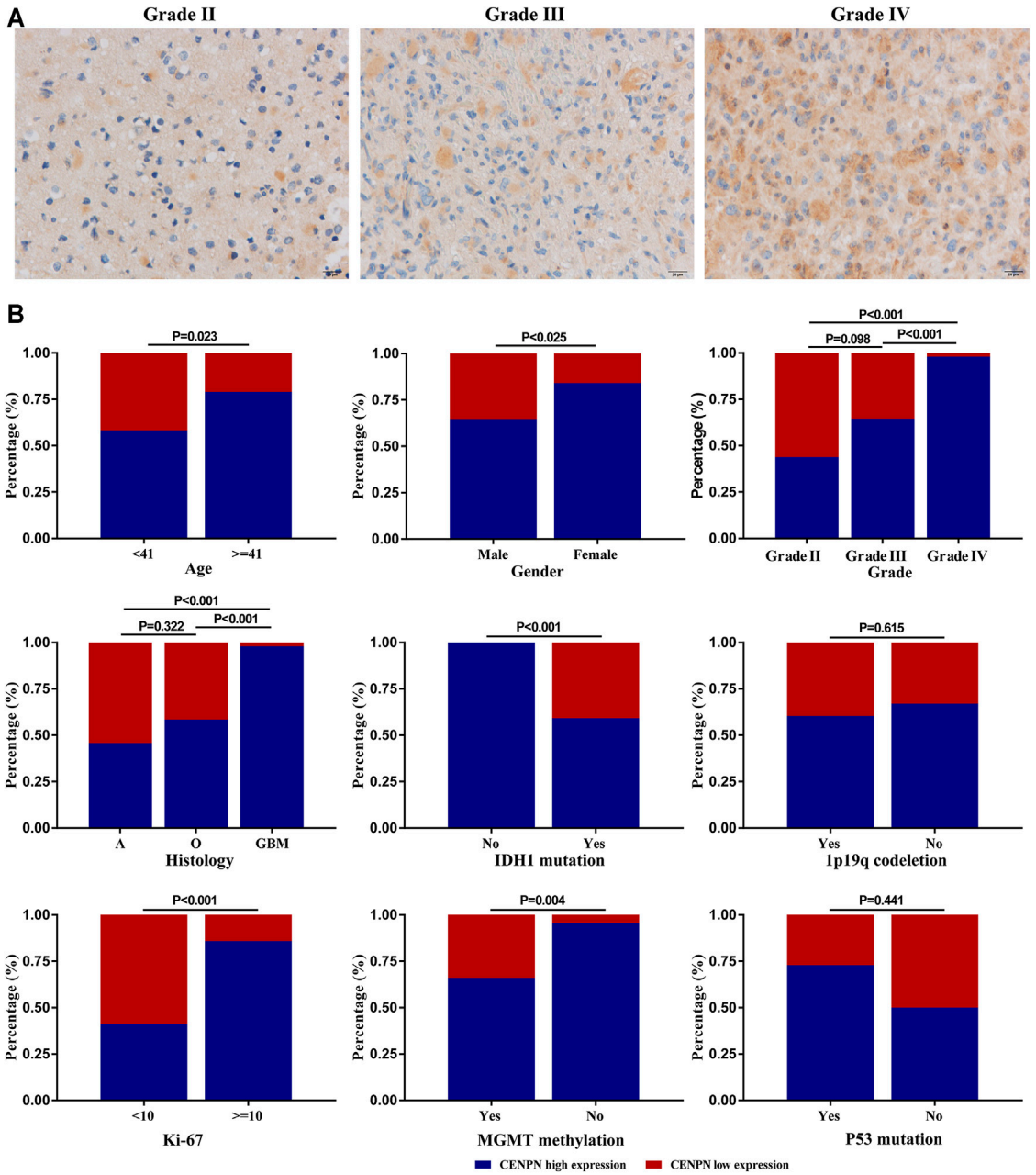


FIGURE 5 | CENPN protein expression in glioma tissues from the Huanhu cohort. **(A)** Representative IHC images of grade II, III and IV samples. **(B)** Associations between CENPN expression and clinical characteristics in the Huanhu cohort.

Protein Expression Level of CENPN in Glioma Tissues

Our study further evaluated the protein level of CENPN in 112 glioma specimens by IHC. We found that CENPN was upregulated in 72.3% (81/112) of glioma tissues (**Supplementary Table S1**). Representative IHC images of samples with grades II, III and IV are displayed in **Figure 5A**. CENPN expression was positively correlated with some adverse outcome clinicopathological variables, such as age, WHO grade and histology (**Figure 5B**). However, no correlation was found between CENPN expression and 1p19q codeletion or P53 mutation.

Biological Functions of CENPN in Human Glioma Cells

To explore the functions of CENPN in glioma, LN229 and U251 cells were transfected with two siRNAs (siCENPN-i and

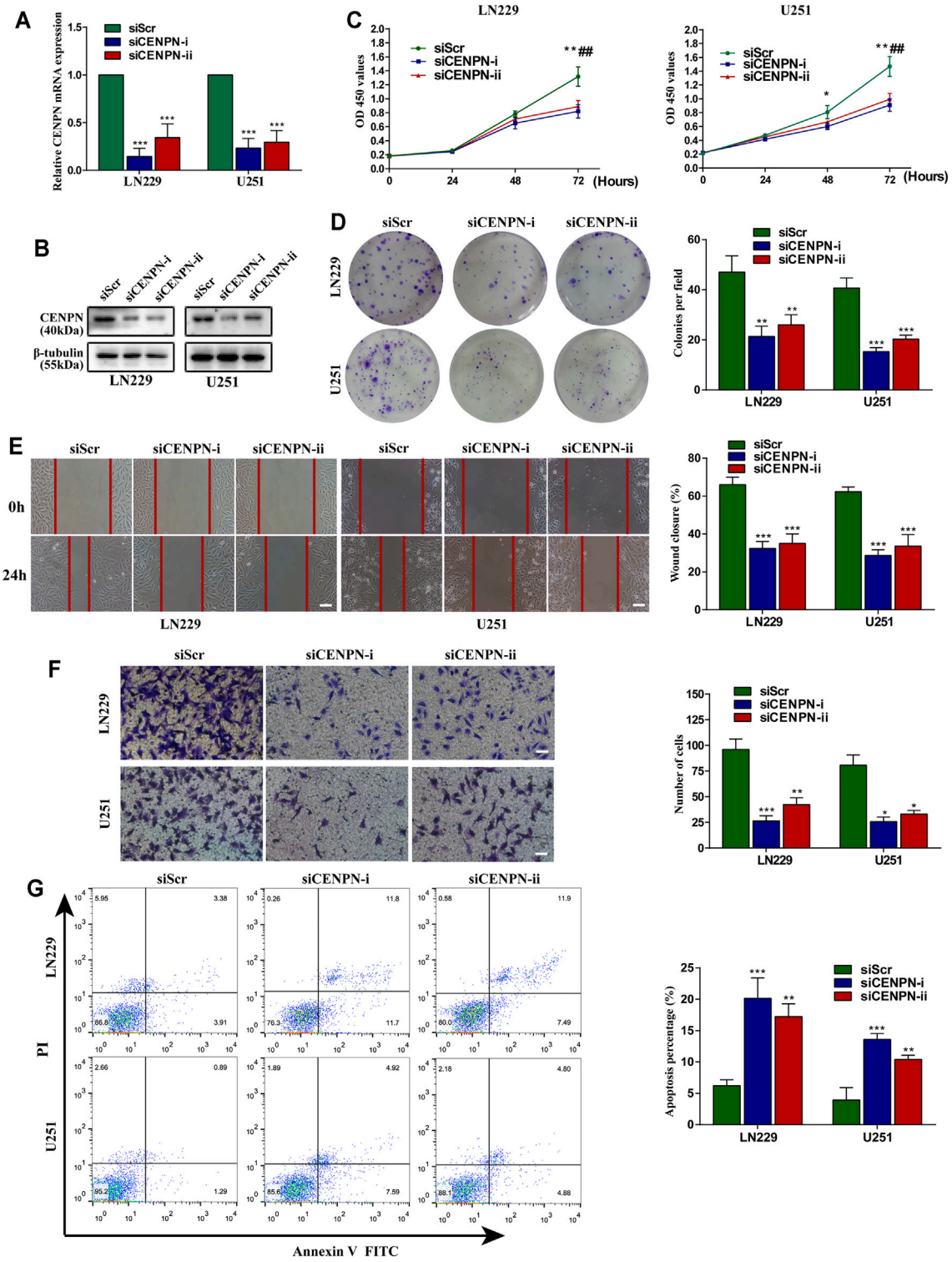


FIGURE 6 | Effects of CENPN knockdown on glioma cells *in vitro*. **(A,B)** Expression of CENPN detected after transfection with siRNA by PCR and western blot. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. **(C)** CCK-8 assay was used to estimate the effect of CENPN knockdown on cell proliferation. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, siCENPN-i vs siScr group; #$p < 0.05$, ##$p < 0.01$, ###$p < 0.001$, siCENPN-ii vs siScr group. **(D)** Cell proliferation was estimated by colony formation assay. **(E)** Wound-healing assay was used to analyze the migration ability of LN229 and U251 cells. Scale bar, 400 μm. **(F,G)** Invasive ability and the degree of apoptosis were estimated by Transwell assay **(F)** and FITC annexin V **(G)**, respectively. Data are the mean ± SD. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, one-way ANOVA. Scale bar, 200 μm.

siCENPN-ii) to silence CENPN transcription. The CENPN siRNAs resulted in significant declines in both mRNA and protein levels (**Figures 6A,B**). The CCK-8 assay showed that transfection with siCENPN significantly reduced the proliferation of the cells (**Figure 6C**). The colony formation assay displayed a similar result (**Figure 6D**). Furthermore, wound healing and Transwell assays showed that downregulation of CENPN suppressed the migration and invasion ability of glioma cells compared with that in the siScr group (**Figures 6E,F**). In addition, we found that higher apoptotic percentages were detected in the siCENPN groups than in the siScr groups (**Figure 6G**). However, the cell cycle assay showed that there were no differences in the fractions of cells in the G0/G1, G2/M, and S phases after CENPN suppression compared with those subjected to siScr (**Supplementary Figure S1**). These results suggest that CENPN could promote malignant phenotypes of glioma cells *in vitro*.

DISCUSSION

The process of tumorigenesis is complicated and involves abnormalities in various genes and signaling pathways (Losurdo et al., 2018; Ronchi et al., 2019; Wang et al., 2019). The lack of timely and effective treatments may lead to the malignant transformation of low-grade glioma, aggravating the prognosis of patients (HanahanWeinberg 2000; FernaldKurokawa 2013; Li et al., 2017; Cai et al., 2018). Therefore, it is essential to clarify the specific mechanisms of glioma occurrence and progression for tumor early diagnosis and treatment.

Taking advantage of the substantial data gained from public databases, we performed comprehensive bioinformatic analyses for CENPN expression in glioma. The results showed that upregulated CENPN expression was correlated with advanced clinicopathological parameters (age, grade, histological type and KPS, etc.) and represented an independent prognostic factor of glioma. These results were validated in clinical samples obtained from our hospital. GSEA and ssGSEA revealed that some immune-related pathways and functions were differentially enriched in the group with CENPN-high. The CIBERSORT algorithm was further used to investigate the influence of CENPN on the infiltration of 22 TILs in the tumor microenvironment. In addition, *in vitro* cell experiments were conducted. Knockdown of CENPN impaired the proliferation, migration and invasion and increased the apoptosis of glioma cell lines. These results demonstrated that CENPN could be a promising target in glioma therapies.

The centromere is a chromosome site that contains chromatin and protein complexes that bind microtubules during mitosis that allows chromosomes to be equally separated (WesthorpeStraight 2014). During the mitosis process, many CENPs associate with and dissociate dynamically from centromeres (CheesemanDesai 2008). In humans, CENPA, also called the histone H3 variant, is always present in the functional centromere. Disruption or mutation of CENPA leads to a complete failure in centromere formation (Palmer et al., 1991; CheesemanDesai 2008). Nevertheless, the biological functions of CENPN are not well clarified. CENP-N interacts directly with the centromere-targeting domain of CENPA to facilitate centromere assembly (Tian et al., 2018), and the depletion of CENPN can induce the loss of many other CENPs (Foltz et al., 2006; Carroll et al., 2009). Therefore, CENPN seems to be essential in centromere assembly, similar to CENPA. However, studies have reported that CENPN is not indispensable in sustaining existing centromeres and recruiting kinetochore proteins because CENPN protein levels decrease before mitosis (McClelland et al., 2007; Hellwig et al., 2011). In cancers, Wang et al. (2021) and Oka et al. (2019) reported that CENPN knockdown could arrest cell cycle at the G1 phase in hepatocellular carcinoma and oral squamous cell carcinoma, respectively. However, in this study, significant changes were not observed in the distribution of cell cycle phases in glioma cells subjected to CENPN knockdown. CENPN may not promote glioma cell progression by influencing the cell cycle; therefore, additional cell experiments are needed to clarify its specific functions in glioma.

In the past decade, immunotherapies have gained great success in multiple tumor treatments (Newick et al., 2017; Ronchi et al., 2020). However, the therapeutic effects of immunotherapies in glioma are less than satisfactory (Lim et al., 2018). As previously reported, the immune infiltration profiles of the tumor microenvironment can directly reflect the immune status and predict the outcomes of cancers (Bindea et al., 2013; Cheng et al., 2016). In our study, patients with different CENPN expression levels had differences in immune infiltration. Based on the CIBERSORT algorithm, five immune cells had the same infiltration trends in both the TCGA and CGGA cohorts. Among these immune cells, the proportions of activated NK cells and monocytes were lower in the CENPN-high group than in the CENPN-low group (**Figure 4**). Activated NK cells are powerful innate immune cells that can help enhance antitumor immunity by releasing perforin and granzyme and secreting various cytokines (Andre et al., 2018). Monocytes have the property of transforming into macrophages and can function as antigen-presenting cells to strengthen adaptive immunity (Jakubzick et al., 2017). Given these findings, CENPN upregulation might promote glioma progression by suppressing the innate immune system. Glioma cells can release multiple immunosuppressive cytokines, such as TGF-β and IL-10, to generate a "cold" tumor microenvironment (Gieryng et al., 2017). Thus, strategies that can reverse the protumor infiltration of immune cells are urgently needed.

Our work presented certain limitations that should be noted. First, the sample size of controls cases in both public databases and clinical cohorts from our institution was small; therefore, more studies with sufficient normal samples must be conducted in the future. Second, CENPN knockdown did not affect the cell cycle distribution, which was inconsistent with previous reports (Oka et al., 2019; Wang et al., 2021). Thus, more glioma cell lines should be adopted for cell biological function assays. At last, functional enrichment analyses revealed that some immune-related functions, immune cells and immune-related pathways were correlated with CENPN expression. The mechanisms by which CENPN may immunomodulate the microenvironment of glioma remain to be understood.

CONCLUSION

This study showed that upregulated CENPN expression was associated with adverse clinical outcomes of glioma patients. CENPN is also associated with multiple immune processes and immune cell infiltration. In addition, *in vitro* experiments showed that CENPN deficiency impaired the proliferation, migration and invasion and increased the apoptosis of glioma cell lines. All in all, these findings indicate that CENPN might be a valid therapeutic target for glioma.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of Huanhu Hospital of Tianjin Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Study design and data collection: HailW, YZ, and HaiyW. Data analysis and interpretation: HaiyW, LX, YY. Writing and review of the manuscript: YZ, HY, HaiyW, and XT

FUNDING

This study was funded by National Natural Science Foundation of China (No. 81972349) and Tianjin Municipal Science and Technology Commission (No. 20JCQNJC00410).

ACKNOWLEDGMENTS

We gratefully acknowledge contributions from Departments of Neurosurgery and Pathology of Huanhu hospital, and TCGA and CGGA databases. We thank American Journal Experts (AJE) for language editing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.732376/full#supplementary-material

REFERENCES

André, P., Denis, C., Soulas, C., Bourbon-Caillet, C., Lopez, J., Arnoux, T., et al. (2018). Anti-NKG2A mAb Is a Checkpoint Inhibitor that Promotes Anti-tumor Immunity by Unleashing Both T and NK Cells. *Cell* 175 (7), 1731–1743. e13. doi:10.1016/j.cell.2018.10.014

Andres, S. A., Bickett, K. E., Alatoum, M. A., Kalbfleisch, T. S., Brock, G. N., and Wittliff, J. L. (2015). Interaction between Smoking History and Gene Expression Levels Impacts Survival of Breast Cancer Patients. *Breast Cancer Res. Treat.* 152 (3), 545–556. doi:10.1007/s10549-015-3507-z

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

Cai, J., Chen, Q., Cui, Y., Dong, J., Chen, M., Wu, P., et al. (2018). Immune Heterogeneity and Clinicopathologic Characterization of IGFBP2 in 2447 Glioma Samples. *Oncoimmunology* 7 (5), e1426516. doi:10.1080/2162402X.2018.1426516

Carroll, C. W., Silva, M. C. C., Godek, K. M., Jansen, L. E. T., and Straight, A. F. (2009). Centromere Assembly Requires the Direct Recognition of CENP-A Nucleosomes by CENP-N. *Nat. Cel Biol* 11 (7), 896–902. doi:10.1038/ncb1899

Cheeseman, I. M., and Desai, A. (2008). Molecular Architecture of the Kinetochore-Microtubule Interface. *Nat. Rev. Mol. Cel Biol* 9 (1), 33–46. doi:10.1038/nrm2310

Cheng, W., Ren, X., Zhang, C., Cai, J., Liu, Y., Han, S., et al. (2016). Bioinformatic Profiling Identifies an Immune-Related Risk Signature for Glioblastoma. *Neurology* 86 (24), 2226–2234. doi:10.1212/WNL.0000000000002770

Fang, J., Liu, Y., Wei, Y., Deng, W., Yu, Z., Huang, L., et al. (2015). Structural Transitions of Centromeric Chromatin Regulate the Cell Cycle-dependent Recruitment of CENP-N. *Genes Dev.* 29 (10), 1058–1073. doi:10.1101/gad.259432.115

Fernald, K., and Kurokawa, M. (2013). Evading Apoptosis in Cancer. *Trends Cel Biol.* 23 (12), 620–633. doi:10.1016/j.tcb.2013.07.006

Foltz, D. R., Jansen, L. E. T., Black, B. E., Bailey, A. O., Cleveland, D. W., and Cleveland, J. R. (2006). The Human CENP-A Centromeric Nucleosome-Associated Complex. *Nat. Cel Biol* 8 (5), 458–469. doi:10.1038/ncb1397

Gieryng, A., Pszczolkowska, D., Walentynowicz, K. A., Rajan, W. D., and Kaminska, B. (2017). Immune Microenvironment of Gliomas. *Lab. Invest.* 97 (5), 498–518. doi:10.1038/labinvest.2017.19

Gravendeel, L. A. M., Kouwenhoven, M. C. M., Gevaert, O., de Rooi, J. J., Stubbs, A. P., Duijm, J. E., et al. (2009). Intrinsic Gene Expression Profiles of Gliomas Are a Better Predictor of Survival Than Histology. *Cancer Res.* 69 (23), 9065–9072. doi:10.1158/0008-5472.CAN-09-2307

Hanahan, D., and Weinberg, R. A. (2000). The Hallmarks of Cancer. *Cell* 100 (1), 57–70. doi:10.1016/s0092-8674(00)81683-9

Hellwig, D., Emmerth, S., Ulbricht, T., Döring, V., Hoischen, C., Martin, R., et al. (2011). Dynamics of CENP-N Kinetochore Binding during the Cell Cycle. *J. Cel Sci* 124 (Pt 22), 3871–3883. doi:10.1242/jcs.088625

Hoischen, C., Yavas, S., Wohland, T., and Diekmann, S. (2018). CENP-C/H/I/K/M/T/W/N/L and hMis12 but Not CENP-S/X Participate in Complex Formation in the Nucleoplasm of Living Human Interphase Cells outside centromeresL and hMis12 but Not CENP-S/X Participate in Complex Formation in the Nucleoplasm of Living Human Interphase Cells outside Centromeres. *PLoS One* 13 (3), e0192572. doi:10.1371/journal.pone.0192572

Jakubzick, C. V., Randolph, G. J., and Henson, P. M. (2017). Monocyte Differentiation and Antigen-Presenting Functions. *Nat. Rev. Immunol.* 17 (6), 349–362. doi:10.1038/nri.2017.28

Kops, G. J. P. L., Weaver, B. A. A., and Cleveland, D. W. (2005). On the Road to Cancer: Aneuploidy and the Mitotic Checkpoint. *Nat. Rev. Cancer* 5 (10), 773–785. doi:10.1038/nrc1714

Kristensen, B. W., Priesterbach-Ackley, L. P., Petersen, J. K., and Wesseling, P. (2019). Molecular Pathology of Tumors of the central Nervous System. *Ann. Oncol.* 30 (8), 1265–1278. doi:10.1093/annonc/mdz164

Li, G., Wang, Z., Zhang, C., Liu, X., Cai, J., Wang, Z., et al. (2017). Molecular and Clinical Characterization of TIM-3 in Glioma through 1,024 Samples. *Oncoimmunology* 6 (8), e1328339. doi:10.1080/2162402X.2017.1328339

Lim, M., Xia, Y., Bettegowda, C., and Weller, M. (2018). Current State of Immunotherapy for Glioblastoma. *Nat. Rev. Clin. Oncol.* 15 (7), 422–442. doi:10.1038/s41571-018-0003-5

Losurdo, G., Principi, M., Girardi, B., Pricci, M., Barone, M., Ierardi, E., et al. (2018). Histamine and Histaminergic Receptors in Colorectal Cancer: From Basic Science to Evidence-Based Medicine. *Acamc* 18 (1), 15–20. doi:10.2174/1871520616666160321115349

McClelland, S. E., Borusu, S., Amaro, A. C., Winter, J. R., Belwal, M., McAinsh, A. D., et al. (2007). The CENP-A NAC/CAD Kinetochore Complex Controls Chromosome Congression and Spindle Bipolarity. *EMBO J.* 26 (24), 5033–5047. doi:10.1038/sj.emboj.7601927

McKinley, K. L., and Cheeseman, I. M. (2016). The Molecular Basis for Centromere Identity and Function. *Nat. Rev. Mol. Cel Biol* 17 (1), 16–29. doi:10.1038/nrm.2015.5

Newick, K., O'Brien, S., Moon, E., and Albelda, S. M. (2017). CAR T Cell Therapy for Solid Tumors. *Annu. Rev. Med.* 68, 139–152. doi:10.1146/annurev-med-062315-120245

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Oka, N., Kasamatsu, A., Endo-Sakamoto, Y., Eizuka, K., Wagai, S., Koide-Ishida, N., et al. (2019). Centromere Protein N Participates in Cellular Proliferation of Human Oral Cancer by Cell-Cycle Enhancement. *J. Cancer* 10 (16), 3728–3734. doi:10.7150/jca.32281

Omuro, A., and DeAngelis, L. M. (2013). Glioblastoma and Other Malignant Gliomas. *JAMA* 310 (17), 1842–1850. doi:10.1001/jama.2013.280319

Ostrom, Q. T., Gittleman, H., Truitt, G., Boscia, A., Kruchko, C., and Barnholtz-Sloan, J. S. (2018). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011-2015, *Neuro Oncol*, 20, Suppl. l_4, iv1–iv86. doi:10.1093/neuonc/noy131

Palmer, D. K., O'Day, K., Trong, H. L., Charbonneau, H., and Margolis, R. L. (1991). Purification of the Centromere-specific Protein CENP-A and Demonstration that it Is a Distinctive Histone. *Proc. Natl. Acad. Sci.* 88 (9), 3734–3738. doi:10.1073/pnas.88.9.3734

Rahman, M. R., Islam, T., Gov, E., Turanli, B., Gulfidan, G., Shahjaman, M., et al. (2019). Identification of Prognostic Biomarker Signatures and Candidate Drugs in Colorectal Cancer: Insights from Systems Biology Analysis. *Medicina* 55 (1), 20. doi:10.3390/medicina55010020

Ronchi, A., Montella, M., Cozzolino, I., Argenziano, G., Moscarella, E., Piccolo, V., et al. (2020). The Potential Diagnostic and Predictive Role of Anaplastic Lymphoma Kinase (ALK) Gene Alterations in Melanocytic Tumors. *Eur. Rev. Med. Pharmacol. Sci.* 24 (7), 3829–3838. doi:10.26355/eurrev_202004_20849

Ronchi, A., Cozzolino, I., Montella, M., Panarese, I., Zito Marino, F., Rossetti, S., et al. (2019). Extragonadal Germ Cell Tumors: Not Just a Matter of Location. A Review about Clinical, Molecular and Pathological Features. *Cancer Med.* 8 (16), 6832–6840. doi:10.1002/cam4.2195

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

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Res.* 47 (W1), W556–W560. doi:10.1093/nar/gkz430

Tian, T., Li, X., Liu, Y., Wang, C., Liu, X., Bi, G., et al. (2018). Molecular Basis for CENP-N Recognition of CENP-A Nucleosome on the Human Kinetochore. *Cell Res* 28 (3), 374–378. doi:10.1038/cr.2018.13

Wang, Q., Yu, X., Zheng, Z., Chen, F., Yang, N., and Zhou, Y. (2021). Centromere Protein N May Be a Novel Malignant Prognostic Biomarker for Hepatocellular Carcinoma. *PeerJ* 9, e11342. doi:10.7717/peerj.11342

Wang, Y., Tang, C., Yao, S., Lai, H., Li, R., Xu, J., et al. (2019). Discovery of a Novel Protein Kinase C Activator from Croton Tiglium for Inhibition of Non-small Cell Lung Cancer. *Phytomedicine* 65, 153100. doi:10.1016/j.phymed.2019.153100

Wang, Y., Tang, S., Lai, H., Jin, R., Long, X., Li, N., et al. (2020). Discovery of Novel IDH1 Inhibitor through Comparative Structure-Based Virtual Screening. *Front. Pharmacol.* 11, 579768. doi:10.3389/fphar.2020.579768

Wang, Z., Jensen, M. A., and Zenklusen, J. C. (2016). A Practical Guide to the Cancer Genome Atlas (TCGA). *Methods Mol. Biol.* 1418, 111–141. doi:10.1007/978-1-4939-3578-9_6

Westhorpe, F. G., and Straight, A. F. (2014). The Centromere: Epigenetic Control of Chromosome Segregation during Mitosis. *Cold Spring Harb Perspect. Biol.* 7 (1), a015818. doi:10.1101/cshperspect.a015818

Xiao, Y., Yuan, Y., Zhang, Y., Li, J., Liu, Z., Zhang, X., et al. (2015). CMTM5 Is Reduced in Prostate Cancer and Inhibits Cancer Cell Growth *In Vitro* and *In Vivo*. *Clin. Transl Oncol.* 17 (6), 431–437. doi:10.1007/s12094-014-1253-z

Zhao, Z., Zhang, K.-N., Wang, Q., Li, G., Zeng, F., Zhang, Y., et al. (2021). Chinese Glioma Genome Atlas (CGGA): A Comprehensive Resource with Functional Genomic Data from Chinese Glioma Patients. *Genomics, Proteomics & Bioinformatics*. doi:10.1016/j.gpb.2020.10.005

Zhou, X., Qiu, S., Nie, L., Jin, D., Jin, K., Zheng, X., et al. (2020). Classification of Muscle-Invasive Bladder Cancer Based on Immunogenomic Profiling. *Front. Oncol.* 10, 1429. doi:10.3389/fonc.2020.01429

Zhou, Y., Wang, X., Lv, P., Yu, H., and Jiang, X. (2021). CDK5 Knockdown Inhibits Proliferation and Induces Apoptosis and Cell Cycle Arrest in Human Glioblastoma. *J. Cancer* 12 (13), 3958–3966. doi:10.7150/jca.53981

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 Genetics

ORIGINAL RESEARCH
published: 25 October 2021
doi: 10.3389/fgene.2021.720849

Prognostic Autophagy-Related Genes of Gastric Cancer Patients on Chemotherapy

Xiaolong Liu[1,2†], Bin Ma[2†], Mali Chen[3], Yaqing Zhang[4], Zhen Ma[1,2] and Hao Chen[1,2*]*

[1]NHC Key Laboratory of Diagnosis and Therapy of Gastrointestinal Tumor, Gansu Provincial Hospital, Lanzhou, China, [2]Department of Surgical Oncology, Lanzhou University Second Hospital, Lanzhou, China, [3]Department of Obstetrics, Gansu Province Maternity and Child-Care Hospital, Lanzhou, China, [4]Department of Gynaecology, Gansu Province Maternity and Child-Care Hospital, Lanzhou, China

Background: Chemotherapy resistance based on fluorouracil and cisplatin is one of the most encountered postoperative clinical problems in patients diagnosed with gastric cancer (GC), resulting in poor prognosis.

Aim of the Study: This study aimed to combine autophagy-related genes (ARGs) to investigate the susceptibility patients with GC to postoperative chemotherapy.

Methods: Based on The Cancer Genome Atlas (TCGA) database, gene expression data for GC patients undergoing chemotherapy were integrated and analyzed. Prognostic genes were screened based on univariate and multivariate analysis regression analysis. Subjects were divided into high-risk and low-risk groups according to the median risk score. Kaplan-Meier method was used to evaluate OS and DFS. The accuracy of the prediction was determined by the subject operating characteristic curve analysis. In addition, stratified analyses based on different clinical variables was performed to assess the correlation between risk scores and clinical variables. Quantitative real-time (qRT) PCR was used to verify the expression of CXCR4 in GC tissues and cell lines.

Results: A total of nine ARGs related to the prognosis of chemotherapy patients were screened out. Compared with normal gastric mucosa cell, CXCR4 showed elevated expression in GC and was significantly associated with survival. Based on GEO and TCGA databases, the model accurately predicted DFS and OS after chemotherapy.

Conclusion: This study established prognostic markers based on nine genes, predicting that ARGs are related to chemotherapy susceptibility of GC patients, which can provide better individualized treatment regimens for clinical practice.

Keywords: gastric cancer, autophagy-related genes, chemotherapy, CXCR4, prognosis

OPEN ACCESS

Edited by:
Jian-Bing Fan,
Illumina, United States

Reviewed by:
Soroush Seifirad,
Harvard Medical School,
United States
Huazhang Wu,
Bengbu Medical College, China
Shihai Liu,
The Affiliated Hospital of Qingdao University, China

***Correspondence:**
Zhen Ma
mazh19@lzu.edu.cn
Hao Chen
ery_chenh@lzu.edu.cn

†These authors share first authorship

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *05 June 2021*
Accepted: *12 October 2021*
Published: *25 October 2021*

Citation:
Liu X, Ma B, Chen M, Zhang Y, Ma Z and Chen H (2021) Prognostic Autophagy-Related Genes of Gastric Cancer Patients on Chemotherapy. Front. Genet. 12:720849. doi: 10.3389/fgene.2021.720849

INTRODUCTION

Gastric cancer (GC) is a major health problem worldwide, which is also a challenge resulted in huge economical burdens. In East Asian countries, especially in China, GC has the highest incidence and mortality rates (GBD 2017 Stomach Cancer Collaborators, 2020). Although overall survival has improved over the past few decades, the prognosis still remains remarkedly poor (Boya et al., 2013a). Drug resistance of chemotherapeutic drugs is the main factor that causes a poor prognosis in GC

patients. Conventional evaluation indexes cannot appropriately evaluate the prognosis of patients with chemotherapy, so it is necessary to have some explicit knowledge and explore victims undergoing chemotherapy.

Autophagy is an important process of eukaryotic transformation of intracellular structures and components (Boya et al., 2013b). In this process, cells wrap their own cytoplasmic proteins or organelles through a single or double membrane to form autophagosomes, which further fuse with lysosomes to form autolysosome, and degrade the contents of the package. According to the different ways of transporting cellular material to lysosomes, autophagy is divided into three types, namely, Macro-autophagy, Micro-autophagy and Chaperon mediated autophagy (CMA) (Kaushik and Cuervo, 2018). What we usually refer to as autophagy is Macro-autophagy. Unlike Macro-autophagy, there is no process of autophagosome formation in Micro-autophagy. The lysosomal membrane itself invades, wrapping and phagocytosing the material to be degraded in the cell, and degrading it. Unlike the former two, CMA is selective in protein degradation. The protein in the cell is restored from the folded state to the unfolded state, and then transferred to the lysosome (Tekirdag and Cuervo, 2018).

Physiological imbalance problems in some processes of autophagy can lead to various diseases and ailments, such as cancer (Levine and Kroemer, 2019). There are some significant pathophysiological processes with autophagy regard to some malignancies (Shen et al., 2008). For instance, Beclin1 gene is associated with autophagy to some extent, which is highly expressed in GC, but not or low expressed in normal tissues (Qu et al., 2017). Glutamine decomposition provides energy for tumor cells, and autophagy activation also contribute to abnormal glutamine decomposition in GC cells, promoting promotion and metastasis (Zhang et al., 2018). LC3 has been widely used as a biomarker for autophagosome, with high expression of LC3 detected in 58% of GC cells, but not in normal gastric epithelial cells (Yoshioka et al., 2008). P62/SQSTM1, a characteristic substrate of ubiquitin-protein in autophagy, which is more significantly up-regulated in GC specimens than in normal gastric mucosa (Kim et al., 2019), while the interpretation of P62/SQSTM1 has some adverse clinical outcomes of the ailment (Masuda et al., 2016). However, whether these autophagy-related genes (ARGs) are correlated with GC patient prognosis remains highly unknown.

Chemotherapy, remains the standard treatment against advanced GC, can exert cytotoxic *via* inducing and enhancing autophagy. It has been reported that autophagy is a survival mechanism that contributes to the development of acquired drug resistance. For instance, autophagy can inhibit the apoptosis of 5-FU-induced MGC803 in GC cells (Korourian et al., 2019). Aquaporin 3(AQP3) promotes the resistance of GC cells to cisplatin *via* autophagy (Dong et al., 2016). Consequently, autophagy might have a fundamental impact on the chemotherapy response of GC. Therefore, it is important to analyze the expression patterns of ARGs in the GC patients on chemotherapy, as well as their prognostic value.

On this basis, our study used bioinformatics methods to predict the prognosis of chemotherapy in GC patients by screening ARGs. This model is helpful for clinicians to develop more individualized chemotherapy regimens and serve patients better and more efficiently. The expression of CXCR4 were verified in GC tissues and cells by qRT.

MATERIALS AND METHODS

Data Collection

ARGs were downloaded and organized from the Human Autophagy Databases (http://autophagy.lu/clustering/index.html). Chemotherapy regimens based on cisplatin and fluorouracil were widely used. Therefore, gene expression data and clinical information were obtained from The Cancer Genome Atlas (TCGA) data portal (https://portal.gdc.cancer.gov/) in 157 patients with GC who received cisplatin or fluorouracil post operatively. The GSE26253 gene expression profile with 432 patients on chemotherapy was downloaded from the Gene Expression Omnibus (GEO) database.

Differential Expression of ARGs and the Enrichment Analysis

The differentially expressed genes (DEGs) of ARGs between chemotherapy group and adjacent nontumorous samples were identified using "limma" R package with a false discovery rate (FDR) <0.05 in the TCGA cohort. To explore the main biological characteristics of ARGs related to chemotherapy, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were performed by the "clusterProfiler" R package.

Construction of Prognostic Gene Signatures

To identify the prognostic value of ARGs with overall survival (OS) and disease-free survival (DFS) in GC chemotherapy group, univariate Cox proportional hazard regression analysis was performed based on TCGA and GEO database. The prognostic model of ARGs was established by multivariate Cox regression analysis. The risk score was calculated based on the expression level of ARGs. Optimal cutoff values were used to divide patients into low-risk and high-risk groups. In addition, Kaplan-Meier method was used to conduct survival analysis based on risk score. To investigate whether the ARGs risk index in the TCGA cohort could be an independent predictor of OS, univariate and multivariate Cox regression analyses were further applied. Risk score, age, sex, tumor subtype, pathological stages, and histological grades were used as covariates. The correlation between risk score and clinicopathological variables was calculated by using the T-test. $p < 0.05$ was considered statistically significant. The Kaplan-Meier plotter database was constructed based on gene chips and RNA-seq data from public databases such as GEO, EGA, and TCGA. We used the Kaplan-Meier plotter database to analyze the relationship between the

expression of CXCR4 and the prognosis of GC, we selected "Pan-cancer RNA-seq" and "Stomach adenocarcinoma."

Gene Set Enrichment Analysis (GSEA)

GSEA was conducted to explore the characteristics of gene Hallmarks in high-risk and low-risk populations. GSEA was performed using GSEA3.0 (http://www.broad.mit.edu/gsea/). Differences for which the nominal $p < 0.05$ and the FDR < 0. 25 were considered statistically enriched.

Tissue Samples

A total of 60 GC cancerous and paracancerous tissue samples were collected in the surgery from May 2010 to December 2018, and the tissues were stored at a −80°C freezer. All patients enrolled in the study signed written informed consent. None of the subjects underwent radiotherapy or chemotherapy prior to the surgery. The tissues were subjected to homogenization, and then total RNA was extracted for RT-PCR. The study was approved by the Clinical Research Ethics Committees.

Cell Culture

Human GC cell lines MKN45, AGS, HGC27, N87 and human normal gastric mucosal epithelial cells GES-1 were purchased from Cell Culture Center of Chinese Academy of Medical Sciences (Beijing, China). All cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (Gibico, United States) and 1% penicillin and streptomycin (Biyuntian, China).

Quantitative Real-Time (qRT) PCR

According to the manufacturer's instructions, the total RNA of the cells was extracted using Trizol reagent (Accurate Biology, China). β-actin as endogenous control, the relative expression of target gene was detected by SYBR Green method on Bio-Rad CFX96GRT-PCR system. The primer sequences were as follows ("F" represents "forward"; "R" represents "reverse"). CXCR4,5′-GGCCCTCAAGACCACAGTC-3′(F), 5′-TTAGCTGGAGTGAAAACTTG-3′(R). Relative quantification of mRNA expression was calculated using the $2^{-\Delta\Delta ct}$ method (Livak and Schmittgen, 2001).

Statistical Analysis

Student's t-test was used to compare gene expression between tumor GC on chemotherapy and normal tissues. Univariate and multivariate cox regression analyses were used to identify independent factors of OS and DFS. Kaplan–Meier curve was implemented to visualize the survival. R software (version 4.0.2) was applied to process and analyze the statistics.

RESULTS

Identification of the Differentially Expressed ARGs in TCGA Cohort

232 ARGs were obtained in our study. A total of 221 ARGs were expressed in TCGA cohort. The results were 157 patients who received chemotherapy and 32 normal samples. The correlated basic clinical characteristics was also compared, as shown in

TABLE 1 | Clinical characteristics of GC patients with chemotherapy in TCGA cohort.

Characteristic	Variables	Total	Percentage (%)
Age	≤65	184	44.9
	>65	226	55.1
Sex	Male	263	64.1
	Female	147	35.9
Grade	G1-2	151	36.8
	G3	251	61.2
	GX	8	2.0
Lauren classification	Intestinal	182	44.4
	Diffuse	74	18.0
	Mixed	154	37.6
Stage	I	57	13.9
	II	129	31.5
	III	181	44.1
	IV	43	10.5
T stage	T1	19	4.6
	T2	84	20.5
	T3	191	46.6
	T4	116	28.3
N stage	N0	130	31.7
	N1	107	26.1
	N2	83	20.2
	N3	90	22.0
M stage	M0	363	88.5
	M1	29	7.1
	Mx	18	4.4

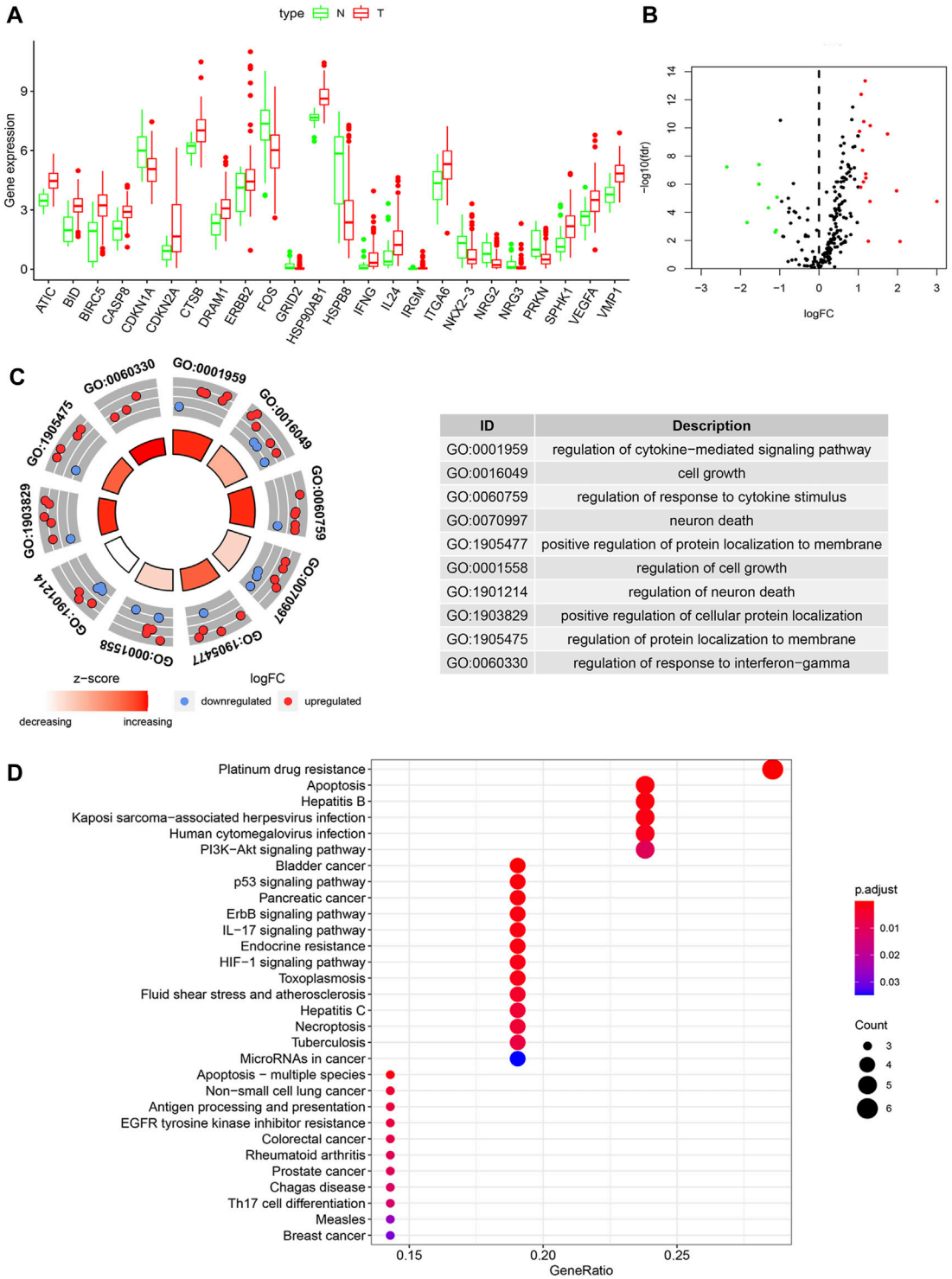
Table 1. With FDR < 0.05 and |log2 FC| > 1 as the screening criteria, 24 DEGs of ARGs were identified (**Figures 1A,B**). The upregulated ATGs were IFNG, ATIC, BIRC5, CASP8, VMP1, IL24, CDKN2A, HSP90AB1, VEGFA, CTSB, and ERBB2. The downregulated ATGs include: PRKN, CDKN1A, GRID2, HSPB8, NRG3, NRG2, FOS, and NKX2-3.

Enrichment Analysis of the Differentially Expressed ARGs

We utilized GO enrichment and KEGG pathway analysis to explore the possible biological functions in GC that may be associated with chemotherapy response. Based on GO analysis, the differentially expressed ARGs were mainly enriched in cell growth, neuron death, positive regulation of protein localization to membrane, autophagy (**Figure 1C**). The KEGG pathways analysis indicated that the DEGs were mainly related to platinum drug resistance, apoptosis, EGFR tyrosine kinase inhibitor resistance and p53 signaling pathway (**Figure 1D**).

The Construction of Prognostic Markers of ARGs for OS in TCGA Cohort

Then 221 ATGs were analyzed by univariate Cox regression analysis. Thirteen ARGs were associated with the prognostic of patients with chemotherapy in TCGA cohort (**Figure 2A**). After multivariate Cox regression analysis, nine ARGs were finally identified to relate to the OS. The coefficients of each gene were shown in **Table 2**.



ID	Description
GO:0001959	regulation of cytokine−mediated signaling pathway
GO:0016049	cell growth
GO:0060759	regulation of response to cytokine stimulus
GO:0070997	neuron death
GO:1905477	positive regulation of protein localization to membrane
GO:0001558	regulation of cell growth
GO:1901214	regulation of neuron death
GO:1903829	positive regulation of cellular protein localization
GO:1905475	regulation of protein localization to membrane
GO:0060330	regulation of response to interferon−gamma

FIGURE 1 | The differentially expressed ARGs and function analysis in a chemotherapeutic group and normal tissues. **(A)** Visualization of the expression levels of the 24 differentially expressed ARGs. **(B)** Volcano plot of 221 ARGs. Red upregulation; Green downregulation. **(C)** GO analysis of 24 differentially expressed ARGs. **(D)** Bubble diagram of KEGG enrichment analysis.

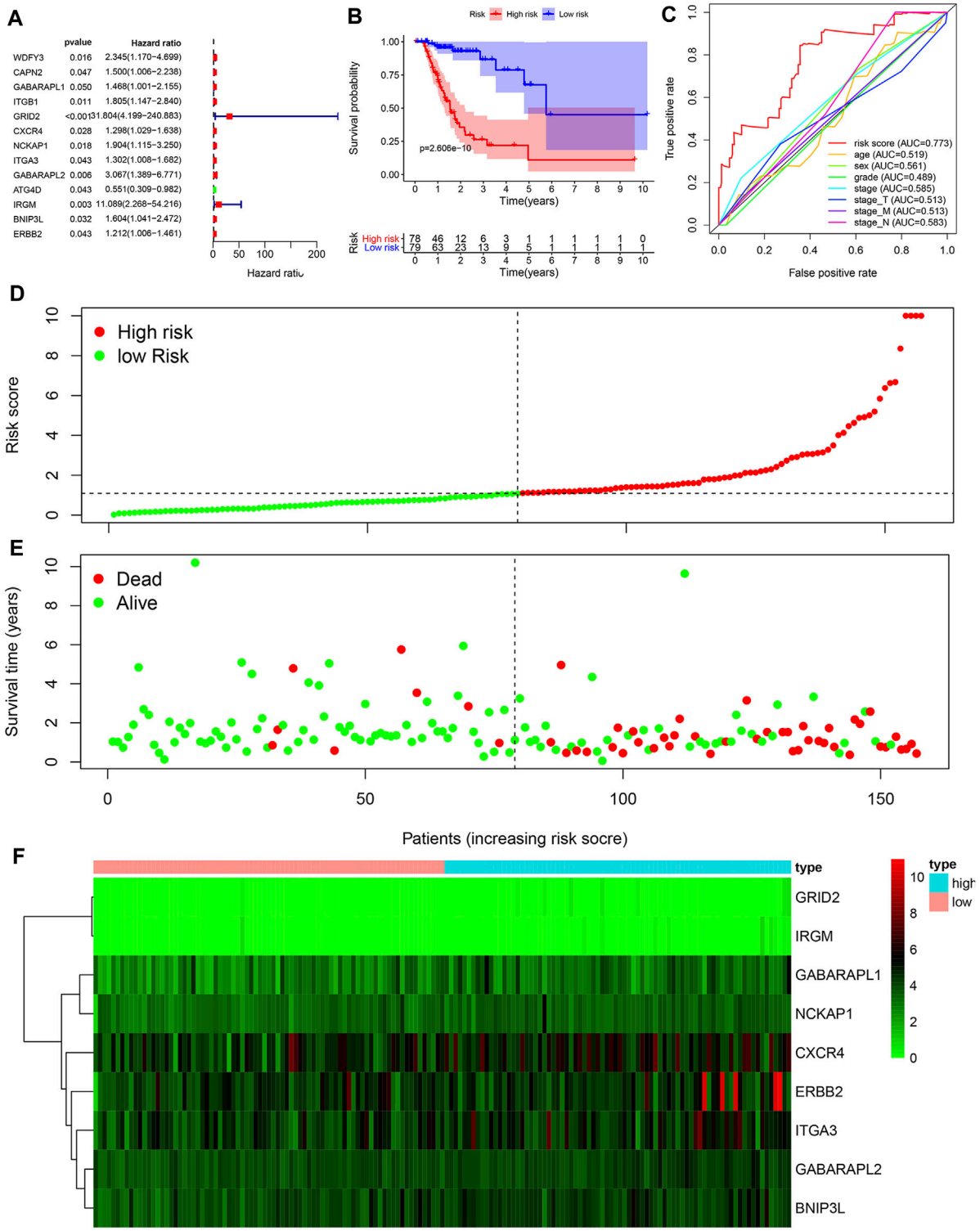


FIGURE 2 | Construction of a prognosis-related risk signature. **(A)** Univariate Cox regression analysis of ARGs related to overall survival of GC patients with chemotherapy. **(B)** Kaplan-Meier OS curves for TCGA GC patients treated with chemotherapy by median risk. **(C)** Multi-index ROC curve of risk score and other indicators. **(D)** Distribution of the risk scores of GC patients. **(E)** The number of survivors and non-survivors with different risk scores. **(F)** The expression of nine ARGs in the **high**- and **low**-risk groups.

TABLE 2 | Multivariate Cox regression analysis of prognostic genes.

Gene	Co-ef	HR	HR.95L	HR.95H
GABARAPL1	0.370661	1.448692	0.912786	2.299233
GRID2	2.358799	10.57824	0.898029	124.6053
CXCR4	0.302963	1.353864	1.034964	1.771025
NCKAP1	0.71455	2.043268	0.967303	4.316067
ITGA3	0.269185	1.308897	0.971892	1.762759
GABARAPL2	1.334027	3.796301	1.55472	9.26977
IRGM	2.963281	19.36138	1.362477	275.1335
BNIP3L	0.592749	1.808954	1.091792	2.997195
ERBB2	0.319098	1.375887	1.105664	1.712152

ARGs as an Independent Prognostic Factor for OS of GC Patients in Chemotherapy Group

Risk scores were calculated based on ATGs mRNA expression levels and risk factors. Patients were divided into high-risk and low-risk groups according to the median risk score. Kaplan-Meier analysis demonstrated that high risk score was associated with poor prognosis, and the 5-year survival rates were 16.5 and 7.7%, respectively, in the high and low risk groups (**Figure 2B**). ROC curves were constructed to determine the ability of ARGs prediction for patients in chemotherapy group (**Figure 2C**). The area under the curve (AUC) of the ARGs for OS was 0.773, which was significantly higher than other indicators. The risk scores of the high-risk and low-risk groups were visualized (**Figure 2D**). As the risk score increased, the number of deaths increased (**Figure 2E**). Heatmaps were constructed for both groups (**Figure 2F**). These results suggested that risk scores accurately reflected patient survival.

To determine whether autophagy-related scoring features were independent prognostic factors in GC patients undergoing chemotherapy, Cox regression analysis was performed. Similarly, the significant correlation between risk scores and OS was achieved by the univariate Cox regression analysis (HR = 1.094, 95%CI = 1.058–1.132, $p < 0.001$)

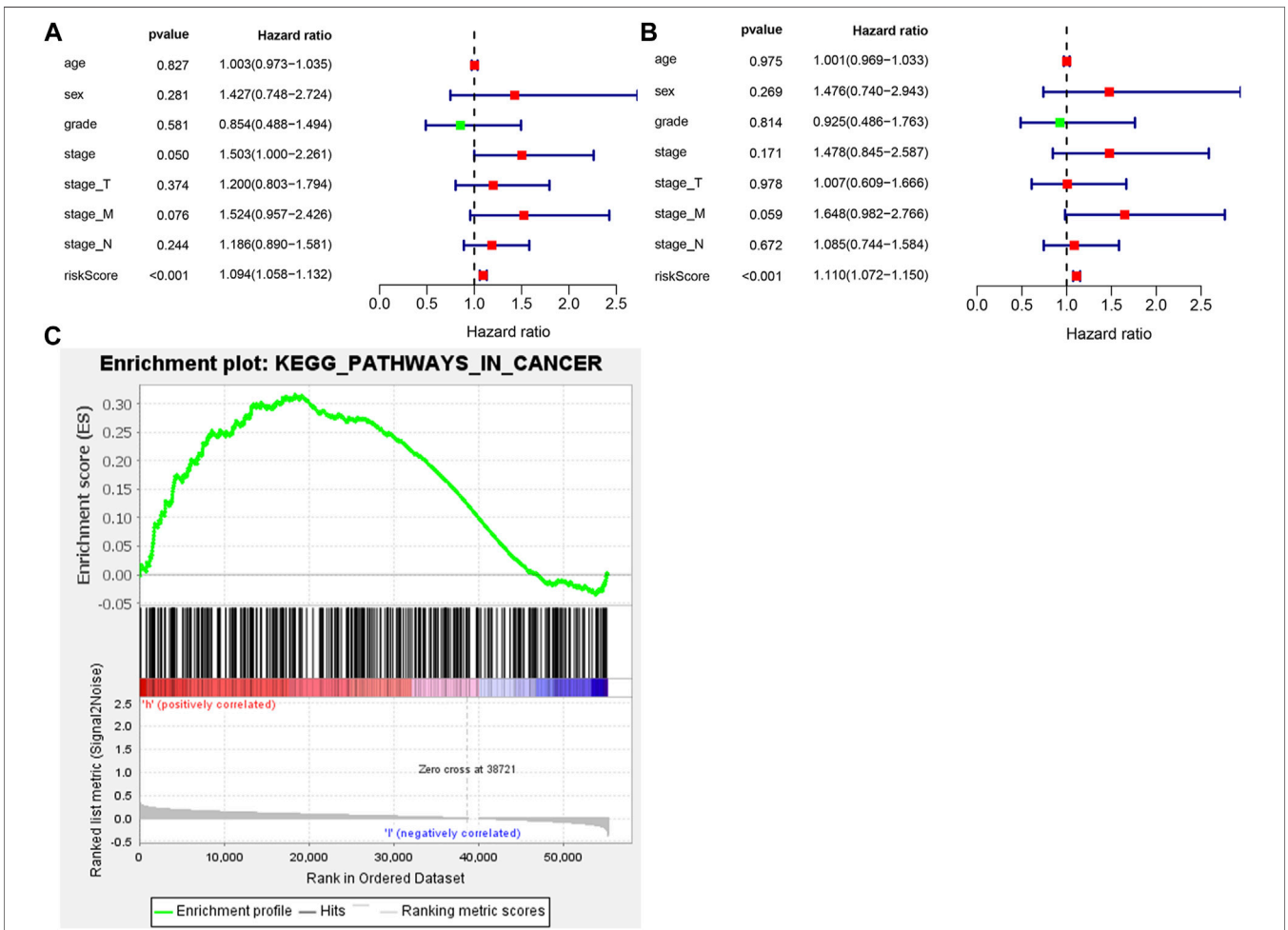


FIGURE 3 | The ATGs for OS is an independent prognostic factor for GC. **(A)** Univariate Cox regression analysis of correlations between the risk score for OS and clinical variables. **(B)** Multivariate Cox regression analysis of correlations between the risk score for OS and clinical variables. **(C)** Gene set enrichment analysis comparing the **high**- and **low**-risk groups.

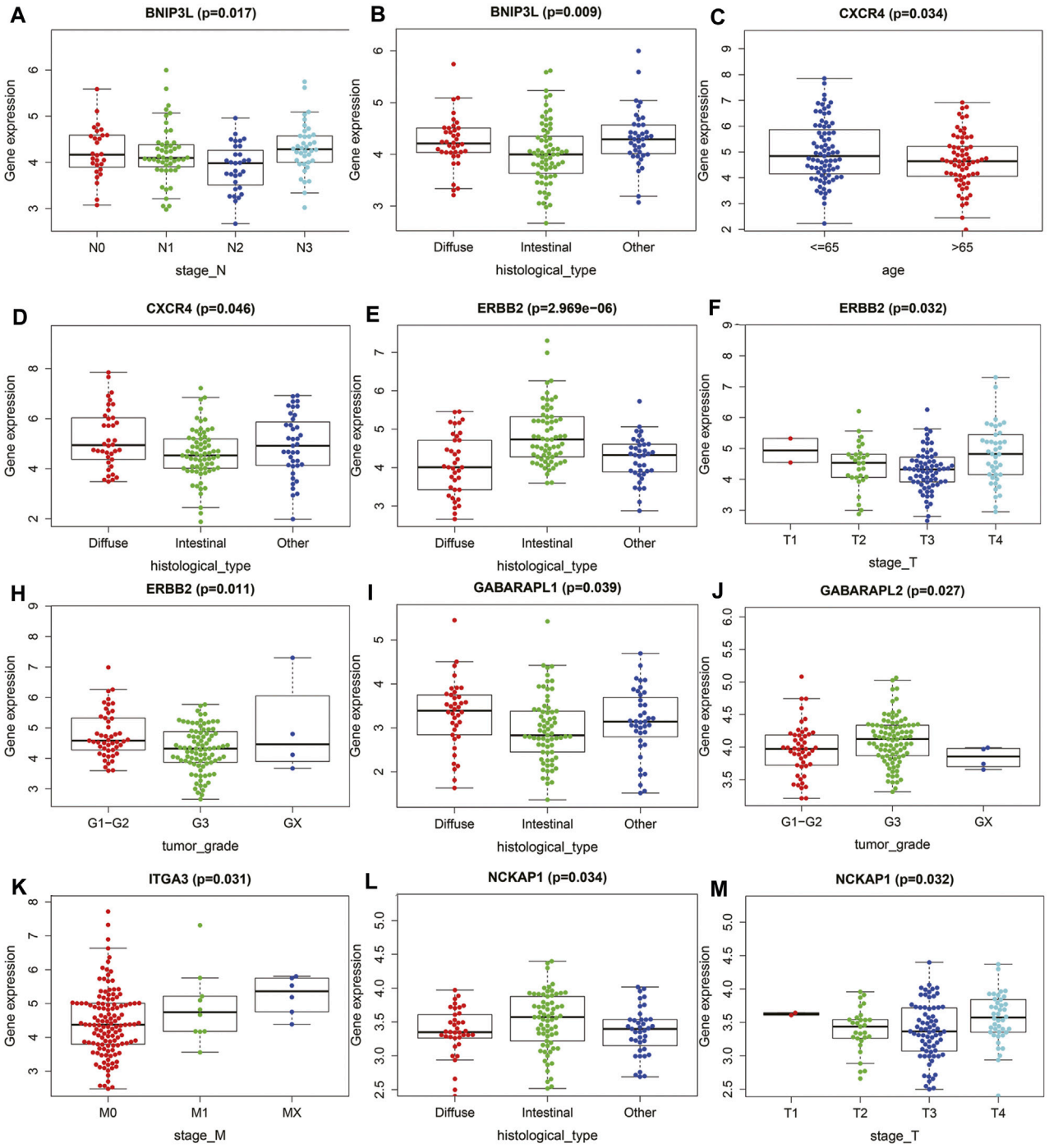


FIGURE 4 | The relationships between the ATGs and clinicopathological variables. **(A,B)** BNIP3L. **(C,D)** CXCR4. **(E,F,H)** ERBB2. **(I,J)** GABRAPL. **(K)** ITGA3. **(L,M)** NCKAP1.

(**Figure 3A**). Multivariate Cox regression analysis showed that risk score was an independent factor affecting the prognosis of CG patients undergoing chemotherapy (HR = 1.110, 95%CI = 1.072–1.150, $p < 0.001$) (**Figure 3B**). Considering the survival differences between the high-risk and low-risk groups, we conducted GSEA to investigate the functional differences between the two groups (**Figure 3C**). Cancer pathways were enriched, suggesting that autophagy is involved in the regulation of chemotherapy for high-risk GC patients. Furthermore, the expression of CXCR4 in GC cell lines was detected by qRT PCR. Compared with normal gastric mucosa epithelial cells, the expression of CXCR4 was significantly increased in GC cell lines.

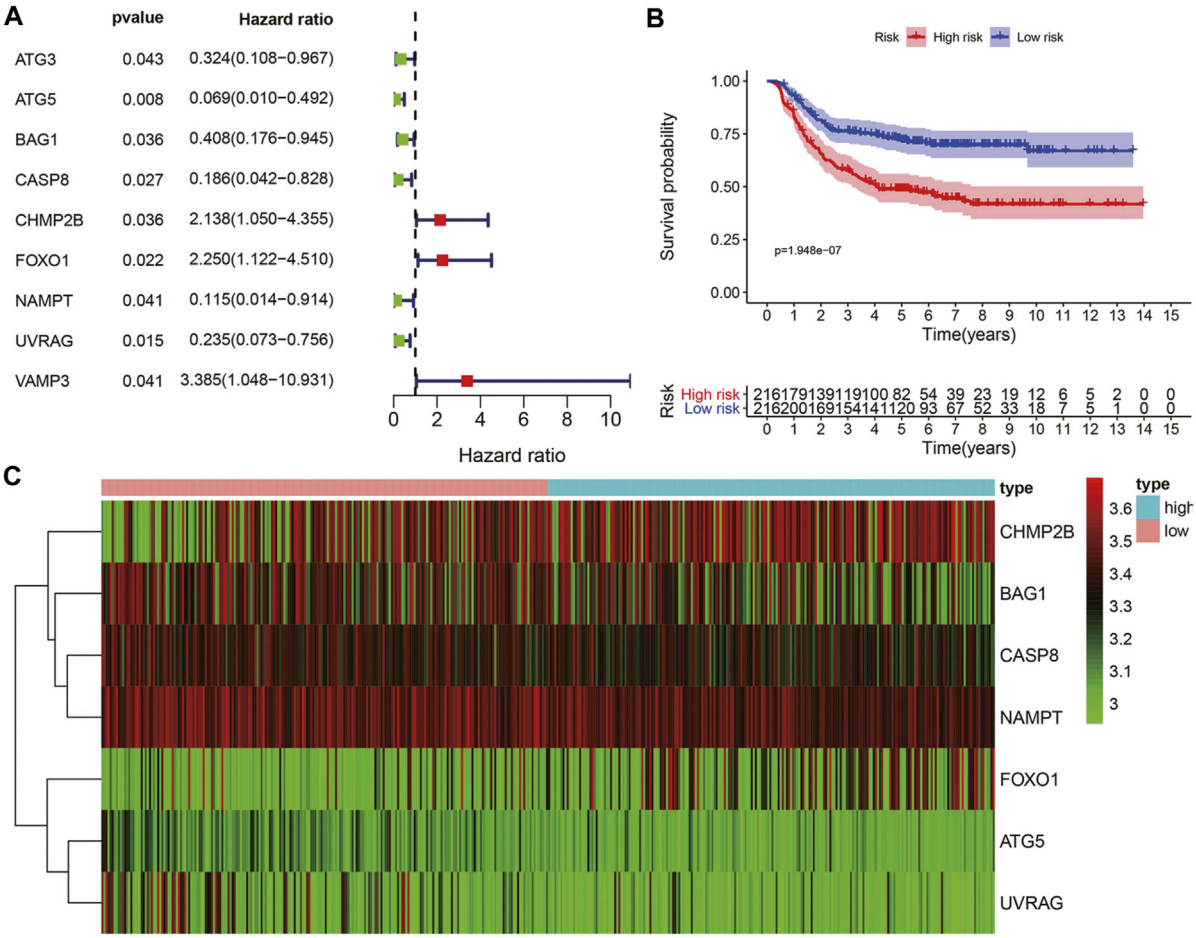


FIGURE 5 | The ATGs for DFS is an independent prognostic factor for GC. **(A)** Univariate Cox regression analysis of autophagy genes related to DFS of GC patients with chemotherapy. **(B)** Kaplan-Meier DFS curves for **high** and **low**-risk groups. **(C)** The expression of seven autophagy-related genes in the **high** and **low**-risk groups.

Relationship Between the Prognostic ARGs for OS and Clinicopathological Variables

To determine whether ARGs affect the progression of gastric cancer, we analyzed the correlation between OS autophagy-related genes and clinicopathological variables. **Figure 4** showed that BNIP3L, CXCR4, ERBB2, GABRAPL, ITGA3, and NCKAP1 were significantly correlated with the pathological classification of GC. On the one hand, BNIP3L, CXCR4, ERBB2, GABRAPL, and NCKAP1 were significantly correlated with Lauren typing. ERBB2 and GABRAPL were also significantly correlated with tumor grade. BNIP3L, ERBB2, ITGA3, and NCKAP1 were significantly correlated with TNM staging.

Prognostic ARGs for DFS of GC Patients in Chemotherapy Group

Considering the significance of DFS in the prognosis of GC patients undergoing chemotherapy, we also established a prognostic marker for DFS. GSE26253 dataset was incorporated. According to univariate Cox regression analysis, there was a certain significant correlation among the nine ARGs (**Figure 5A**). After multivariate Cox regression analysis, we finally obtained seven ARGs and divided the patients in the whole data set into high-risk group and low-risk group according to the median of risk scores. Kaplan-Meier analysis revealed that the DFS in the high-risk group was significantly shorter than that in the low-risk group ($p < 0.001$, **Figure 5B**). Heatmaps were developed for both groups (**Figure 5C**). These results suggest that the prognostic marker of DFS can also predict the prognosis of GC patients undergoing chemotherapy.

Validation of CXCR4 in Independent GC Cohorts

We used qRT-PCR to characterize the expression of CXCR4 in GC tissues and paracancerous tissues. CXCR4 expression in GC tissues was observably higher than that in paracancerous tissues (**Figure 6A**). The expression of CXCR4 was elevated in GC cell lines (**Figure 6B**). Then, we explored the effects of CXCR4 on prognosis of GC patients and found that higher expression of

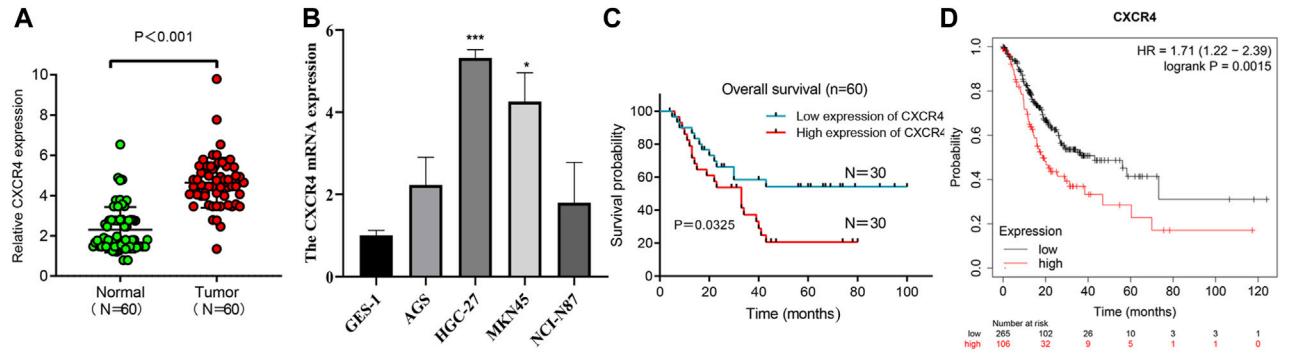


FIGURE 6 | CXCR4 expression in GC. **(A)** The expression of CXCR4 in GC tissues and paracancerous tissues. **(B)** CXCR4 expression in GC cell lines. **(C)** Survival analysis of CXCR4 in GC patients. **(D)** survival curve of CXCR4 based on Kaplan Meier Plotter.

CXCR4 was significantly associated with poor survival (**Figure 6C**). The survival curve based on Kaplan Meier Plotter showed that high CXCR4 expression was closely associated with poor prognosis in GC patients (**Figure 6D**).

DISCUSSION

The treatment cost of advanced gastric cancer is very high, and the prognosis is quite poor, so it has caused huge economic challenges on a global scale. The drug resistance of gastric cancer patients is the main reason for this phenomenon (Tsai et al., 2020). Autophagy is a process by which the body itself regulates cellular mechanisms and homeostasis (Hou et al., 2020). Some recent studies have shown that autophagy may be closely related to the resistance of chemotherapy drugs in GC patients (Peng et al., 2020; Zhao et al., 2020). Studies have found that the autophagy of GC cells with enhanced chemotherapy-drug resistance, and inhibition of autophagy can eliminate chemotherapy-resistance (Xu et al., 2018; Guo et al., 2019). Considering the importance of autophagy in chemotherapy resistance of GC, we further explored the prognostic value of autophagy in the treatment of GC. In recent years, with the advancement of genome sequencing, biochips, and high-throughput sequencing technologies, more and more studies have applied bioinformatics methods to the analysis of chip data sets, providing an effective method for the diagnosis, treatment and prognosis of gastric cancer. In this study, we combined TCGA and GEO databases to accomplish our work. The prognosis of GC patients receiving postoperative chemotherapy was analyzed. We also studied the biological function and role of ARGs in GC.

First and foremost, the differentially expressed ARGs between GC chemotherapy group and normal stomach were identified in our study. Furthermore, GO and KEGG analysis showed that the differentially expressed ARGs were mainly enriched in platinum resistance (Wu et al., 2018; Su et al., 2019; Herhaus et al., 2020). Research has demonstrated that a combination of inhibitors in GC can improve cisplatin resistance, which is consistent and concurs with our results. ARGs can promote progress in GC disease progress through platinum resistance. Moreover, there were 13 genes associated with prognosis in the GC chemotherapy group. We used multivariate Cox regression to construct and compute data set for nine genes.

Among the nine genes in the prognostic model that we constructed, GABARAPL1 knockdown has been shown to inhibit the growth of prostate cancer cells *in vitro* or *in vivo* (Keulers et al., 2015). In head and neck squamous cell carcinoma, the high expression level of GABARAPL1 is associated with the poor prognosis of patients (Liu et al., 2014). However, in certain cancers, high levels of GABARAPL1 expression are associated with better results, such as hepatocellular carcinoma (HCC) (Berthier et al., 2010) and node-positive breast cancer (Zhang et al., 2018). GC cells activate autophagy through GABARAPL1 to supplement glutamine breakdown and promote the growth and metastasis of GC cells (Ali et al., 2017). The GluD2 protein encoded by GRID2 is a member of the ionotropic glutamate receptor family that mediates excitatory synaptic transmission (Ngollo et al., 2017). Ngollo et al. showed that GRID2 is significantly overexpressed in prostate cancer (Zhong et al., 2019). NCKAP1 is abnormally expressed in HCC and used as an independent prognostic factor for patients (Teng et al., 2016). High expression level of NCKAP1 is associated with poorer survival in breast cancer patients (Jiao et al., 2019). ITGA3 has been confirmed to be associated with poor prognosis in a variety of cancers (Miao et al., 2010; Wang et al., 2019; Tian et al., 2020). Miao et al. found that GABARAP is overexpressed in colorectal cancer, and patients with high GABARAP expression have a shorter survival time (Song et al., 2015). IRGM has been shown to be dysregulated in GC and affect the occurrence and development of GC (Burada et al., 2012). BNIP3L has different expressions in a variety of cancers. It is highly expressed in HCC (Chen et al., 2020), breast cancer (Real et al., 2005) and ovarian cancer (Jia et al., 2020). However, in colorectal and pancreatic cancers BNIP3 is frequently epigenetically silenced (Mellor and Harris, 2007).

CXCR4 is a chemokine receptor, which is highly expressed in breast cancer patients, and high expression indicates a poor prognosis. We also got the same result in ovarian cancer (Jiang et al., 2006; Mirisola et al., 2009). These findings suggest that CXCR4 is a promising prognostic factor. In

addition, CXCR4 also plays an important role in the chemotherapy resistance of a variety of malignant tumors. Gemcitabine is a chemotherapeutic agent for the treatment of advanced and metastatic pancreatic cancer. However, chemotherapy resistance is a critical factor affecting the clinical prognosis of pancreatic cancer. Studies have shown that activation of Akt and ERK signaling pathways mediate the resistance of pancreatic cancer to gemcitabine. Blocking CXCR4 can effectively eliminate these survival signals and restore the sensitivity of pancreatic cancer cells to gemcitabine (Singh et al., 2010). Another study confirmed that targeting CXCR4 can inhibit the growth of pancreatic cancer cells and increase the sensitivity of pancreatic cancer cells to gemcitabine (Khan et al., 2020). Similar results have been observed in colorectal cancer, miR-193a-5p reduces the chemotherapy resistance of colorectal cancer to 5-FU and oxaliplatin by targeting CXCR4 (Azar et al., 2021). An analysis based on clinical samples showed that ovarian cancer patients with high expression of CXCR4 were significantly less sensitive to chemotherapy and had a poor prognosis, suggesting that CXCR4 is the key molecules for chemotherapy resistance (Li et al., 2014). In acute myeloid leukemia, targeting CXCR4 has been proven to be one of the potential treatment methods to overcome chemotherapy resistance (Cho et al., 2015). In view of the correlation between CXCR4 and chemotherapy resistance of various tumors, we investigated the role of CXCR4 gene in GC. We have verified the expression of CXCR4 in gastric cancer cells. Our results show that the expression of gastric cancer cells is higher than that of normal gastric epithelial cells, and the expression level of gastric cancer tissues is also high. Beside the cancer, the prognosis of patients with high expression of CXCR4 is also worse than that of patients with low expression of CXCR4. Our results indicate that CXCR4 can be used as a prognostic indicator for patients with gastric cancer. GSEA results showed that autophagy regulation was mainly concentrated in the high-risk group, suggesting that autophagy in the high-risk group may regulate the tolerance of GC patients to chemotherapy and thus lead to poor prognosis (Samiei et al., 2020; Zhang et al., 2020). This is consistent with previous research.

CONCLUSION

In conclusion, we constructed autophagy related genes for OS and DFS in patients with GC undergoing chemotherapy. It may perovide alternative choices for treatment strategies of GC patients with chemotherapy resistant. At the same time, CXCR4 may be used as a promising prognostic indicator for gastric cancer. However, our study still has considerable limitations. First, due to insufficient data, it is not possible to evaluate the prognostic capacity of autophagy related genes in other independent GC data sets, only a strong prospective cohort can actually evaluate predictability of the provided prognostic markers accurately. In addition, there are other prognostic factors that affect patients receiving chemotherapy, such as tumor immune microenvironment, which requires further research. Secondly, further *in vivo* and *in vitro* experiments need to be carried out to explore the expression of genes other than CXCR4 in GC tissues and the potential mechanism. Although previous studies have constructed GC-related prognostic models based on ARGs, there is no model to evaluate the prognosis of patients receiving chemotherapy.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

XL and BM conceived of the study and participated in design and coordination, MC and YZ drafted the manuscript. ZM collected and analyzed immune related information. HC revised manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2019PT320005), Key Talents Project of Gansu Province (No. 2019RCXM020) and Key Project of Science and Technology in Gansu province (19ZD2WA001).

REFERENCES

Ali, Z., Zulfiqar, S., Klar, J., Wikström, J., Ullah, F., Khan, A., et al. (2017). Homozygous GRID2 Missense Mutation Predicts a Shift in the D-Serine Binding Domain of GluD2 in a Case with Generalized Brain Atrophy and Unusual Clinical Features. *BMC Med. Genet.* 18 (1), 144. doi:10.1186/s12881-017-0504-6

Azar, M. R. M. H., Aghazadeh, H., Mohammed, H. N., Sara, M. R. S., Hosseini, A., Shomali, N., et al. (2021). miR-193a-5p as a Promising Therapeutic Candidate in Colorectal Cancer by Reducing 5-FU and Oxaliplatin Chemoresistance by Targeting CXCR4. *Int. immunopharmacology* 92, 107355. doi:10.1016/j.intimp.2020.107355

Berthier, A., Seguin, S., Sasco, A. J., Bobin, J. Y., De Laroche, G., Datchary, J., et al. (2010). High Expression of Gabarapl1 Is Associated with a Better Outcome for

Patients with Lymph Node-Positive Breast Cancer. *Br. J. Cancer* 102 (6), 1024–1031. doi:10.1038/sj.bjc.6605568

Boya, P., Reggiori, F., and Codogno, P. (2013). Emerging Regulation and Functions of Autophagy. *Nat. Cel Biol* 15 (7), 713–720. doi:10.1038/ncb2788

Boya, P., Reggiori, F., and Codogno, P. (2013). Emerging Regulation and Functions of Autophagy. *Nat. Cel Biol* 15 (7), 713–720. doi:10.1038/ncb2788

Burada, F., Plantinga, T. S., Ioana, M., Rosentul, D., Angelescu, C., Joosten, L. A., et al. (2012). IRGM Gene Polymorphisms and Risk of Gastric Cancer. *J. Dig. Dis.* 13 (7), 360–365. doi:10.1111/j.1751-2980.2012.00602.x

Chen, Y.-Y., Wang, W.-H., Che, L., Lan, Y., Zhang, L.-Y., Zhan, D.-L., et al. (2020). BNIP3L-Dependent Mitophagy Promotes HBx-Induced Cancer Stemness of Hepatocellular Carcinoma Cells via Glycolysis Metabolism Reprogramming. *Cancers* 12 (3), 655. doi:10.3390/cancers12030655

Cho, B.-S., Zeng, Z., Mu, H., Wang, Z., Konoplev, S., McQueen, T., et al. (2015). Antileukemia Activity of the Novel Peptidic CXCR4 Antagonist LY2510924 as Monotherapy and in Combination with Chemotherapy. *Blood* 126 (2), 222–232. doi:10.1182/blood-2015-02-628677

Dong, X., Wang, Y., Zhou, Y., Wen, J., Wang, S., and Shen, L. (2016). Aquaporin 3 Facilitates Chemoresistance in Gastric Cancer Cells to Cisplatin *via* Autophagy. *Cel Death Discov.* 2, 16087. doi:10.1038/cddiscovery.2016.87

GBD 2017 Stomach Cancer Collaborators (2020). 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.* 5 (1), 42–54. doi:10.1016/S2468-1253(19)30328-0

Guo, Q., Jing, F.-J., Xu, W., Li, X., Li, X., Sun, J.-L., et al. (2019). Ubenimex Induces Autophagy Inhibition and EMT Suppression to Overcome Cisplatin Resistance in GC Cells by Perturbing the CD13/EMP3/PI3K/AKT/NF-κB axis. *Aging* 12 (1), 80–105. doi:10.18632/aging.102598

Herhaus, L., Bhaskara, R. M., Lystad, A. H., Gestal-Mato, U., Covarrubias-Pinto, A., Bonn, F., et al. (2020). TBK1-mediated Phosphorylation of LC3C and GABARAP-L2 Controls Autophagosome Shedding by ATG4 Protease. *EMBO Rep.* 21 (1), e48317. doi:10.15252/embr.201948317

Hou, J., Tan, Y., Su, C., Wang, T., Gao, Z., Song, D., et al. (2020). Inhibition of Protein FAK Enhances 5-FU Chemosensitivity to Gastric Carcinoma *via* P53 Signaling Pathways. *Comput. Struct. Biotechnol. J.* 18, 125–136. doi:10.1016/j.csbj.2019.12.010

Jia, J., Yang, X., Zhao, Q., Ying, F., Cai, E., Sun, S., et al. (2020). BNIP3 Contributes to Cisplatin-induced Apoptosis in Ovarian Cancer Cells. *FEBS open bio* 10 (8), 1463–1473. doi:10.1002/2211-5463.12881

Jiang, Y.-p., Wu, X.-h., Shi, B., Wu, W.-x., and Yin, G.-r. (2006). Expression of Chemokine CXCL12 and its Receptor CXCR4 in Human Epithelial Ovarian Cancer: an Independent Prognostic Factor for Tumor Progression. *Gynecol. Oncol.* 103 (1), 226–233. doi:10.1016/j.ygyno.2006.02.036

Jiao, Y., Li, Y., Liu, S., Chen, Q., and Liu, Y. (2019). ITGA3 Serves as a Diagnostic and Prognostic Biomarker for Pancreatic Cancer. *Ott* Vol. 12, 4141–4152. doi:10.2147/OTT.S201675

Kaushik, S., and Cuervo, A. M. (2018). The Coming of Age of Chaperone-Mediated Autophagy. *Nat. Rev. Mol. Cel. Biol.* 19 (6), 365–381. doi:10.1038/s41580-018-0001-6

Keulers, T. G., Schaaf, M. B. E., Peeters, H. J. M., Savelkouls, K. G. M., Vooijs, M. A., Bussink, J., et al. (2015). GABARAPL1 Is Required for Increased EGFR Membrane Expression during Hypoxia. *Radiother. Oncol.* 116 (3), 417–422. doi:10.1016/j.radonc.2015.06.023

Khan, M. A., Srivastava, S. K., Zubair, H., Patel, G. K., Arora, S., Khushman, M. d., et al. (2020). Co-targeting of CXCR4 and Hedgehog Pathways Disrupts Tumor-Stromal Crosstalk and Improves Chemotherapeutic Efficacy in Pancreatic Cancer. *J. Biol. Chem.* 295 (25), 8413–8424. doi:10.1074/jbc.RA119.011748

Kim, J. S., Bae, G. E., Kim, K.-H., Lee, S.-I., Chung, C., Lee, D., et al. (2019). Prognostic Significance of LC3B and p62/SQSTM1 Expression in Gastric Adenocarcinoma. *Anticancer Res.* 39 (12), 6711–6722. doi:10.21873/anticanres.13886

Korourian, A., Madjd, Z., Roudi, R., Shariftabrizi, A., and Soleimani, M. (2019). Induction of miR-31 Causes Increased Sensitivity to 5-FU and Decreased Migration and Cell Invasion in Gastric Adenocarcinoma. *Bll* 120 (1), 35–39. doi:10.4149/BLL_2019_005

Levine, B., and Kroemer, G. (2019). Biological Functions of Autophagy Genes: A Disease Perspective. *Cell* 176 (1-2), 11–42. doi:10.1016/j.cell.2018.09.048

Li, J., Jiang, K., Qiu, X., Li, M., Hao, Q., Wei, L., et al. (2014). Overexpression of CXCR4 Is Significantly Associated with Cisplatin-Based Chemotherapy Resistance and Can Be a Prognostic Factor in Epithelial Ovarian Cancer. *BMB Rep.* 47 (1), 33–38. doi:10.5483/bmbrep.2014.47.1.069

Liu, C., Xia, Y., Jiang, W., Liu, Y., and Yu, L. (2014). Low Expression of GABARAPL1 Is Associated with a Poor Outcome for Patients with Hepatocellular Carcinoma. *Oncol. Rep.* 31 (5), 2043–2048. doi:10.3892/or.2014.3096

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* 25 (4), 402–408. doi:10.1006/meth.2001.1262

Masuda, G. O., Yashiro, M., Kitayama, K., Miki, Y., Kasashima, H., Kinoshita, H., et al. (2016). Clinicopathological Correlations of Autophagy-Related Proteins LC3, Beclin 1 and P62 in Gastric Cancer. *Anticancer Res.* 36 (1), 129–136. PMID: 26722036.

Mellor, H. R., and Harris, A. L. (2007). The Role of the Hypoxia-Inducible BH3-Only Proteins BNIP3 and BNIP3L in Cancer. *Cancer Metastasis Rev.* 26 (3-4), 553–566. doi:10.1007/s10555-007-9080-0

Miao, Y., Zhang, Y., Chen, Y., Chen, L., and Wang, F. (2010). GABARAP Is Overexpressed in Colorectal Carcinoma and Correlates with Shortened Patient Survival. *Hepatogastroenterology* 57 (98), 257–261. PMID: 20583424.

Mirisola, V., Zuccarino, A., Bachmeier, B. E., Sormani, M. P., Falter, J., Nerlich, A., et al. (2009). CXCL12/SDF1 Expression by Breast Cancers Is an Independent Prognostic Marker of Disease-free and Overall Survival. *Eur. J. Cancer* 45 (14), 2579–2587. doi:10.1016/j.ejca.2009.06.026

Ngollo, M., Lebert, A., Daures, M., Judes, G., Rifai, K., Dubois, L., et al. (2017). Global Analysis of H3K27me3 as an Epigenetic Marker in Prostate Cancer Progression. *BMC cancer* 17 (1), 261. doi:10.1186/s12885-017-3256-y

Peng, L., Li, Y., Wei, S., Li, X., Dang, Y., Zhang, W., et al. (2020). LAMA4 Activated by Androgen Receptor Induces the Cisplatin Resistance in Gastric Cancer. *Biomed. Pharmacother.* 124, 109667. doi:10.1016/j.biopha.2019.109667

Qu, B., Yao, L., Ma, H.-l., Chen, H.-l., Zhang, Z., and Xie, J. (2017). Prognostic significance of autophagy-related proteins expression in resected human gastric adenocarcinomaMedical sciences = Hua zhong ke ji da xue xue bao. Yi xue Ying De wen ban = Huazhong keji daxue xuebao. *J. Huazhong Univ. Sci. Technol. [Med. Sci.]Yixue Yingdewen Ban* 37 (1), 37–43. doi:10.1007/s11596-017-1691-2

Real, P. J., Benito, A., Cuevas, J., Berciano, M. T., de Juan, A., Coffer, P., et al. (2005). Blockade of Epidermal Growth Factor Receptors Chemosensitizes Breast Cancer Cells through Up-Regulation of Bnip3L. *Cancer Res.* 65 (18), 8151–8157. doi:10.1158/0008-5472.CAN-05-1134

Samiei, E., Seyfoori, A., Toyota, B., Ghavami, S., and Akbari, M. (2020). Investigating Programmed Cell Death and Tumor Invasion in a Three-Dimensional (3D) Microfluidic Model of Glioblastoma. *Ijms* 21 (9), 3162. doi:10.3390/ijms21093162

Shen, Y., Li, D.-D., Wang, L.-L., Deng, R., and Zhu, X.-F. (2008). Decreased Expression of Autophagy-Related Proteins in Malignant Epithelial Ovarian Cancer. *Autophagy* 4 (8), 1067–1068. doi:10.4161/auto.6827

Singh, S., Srivastava, S. K., Bhardwaj, A., Owen, L. B., and Singh, A. P. (2010). CXCL12-CXCR4 Signalling axis Confers Gemcitabine Resistance to Pancreatic Cancer Cells: a Novel Target for Therapy. *Br. J. Cancer* 103 (11), 1671–1679. doi:10.1038/sj.bjc.6605968

Song, Z., Guo, C., Zhu, L., Shen, P., Wang, H., Guo, C., et al. (2015). Elevated Expression of Immunity-Related GTPase Family M in Gastric Cancer. *Tumor Biol.* 36 (7), 5591–5596. PMID: 25707354. doi:10.1007/s13277-015-3229-1

Su, B., Zhang, L., Liu, S., Chen, X., and Zhang, W. (2019). GABARAPL1 Promotes AR+ Prostate Cancer Growth by Increasing FL-AR/AR-V Transcription Activity and Nuclear Translocation. *Front. Oncol.* 9, 1254. doi:10.3389/fonc.2019.01254

Tekirdag, K., and Cuervo, A. M. (2018). Chaperone-mediated Autophagy and Endosomal Microautophagy: Jointed by a Chaperone. *J. Biol. Chem.* 293 (15), 5414–5424. doi:10.1074/jbc.R117.818237

Teng, Y., Qin, H., Bahassan, A., Bendzunas, N. G., Kennedy, E. J., and Cowell, J. K. (2016). The WASF3-NCKAP1-CYFIP1 Complex Is Essential for Breast Cancer Metastasis. *Cancer Res.* 76 (17), 5133–5142. doi:10.1158/0008-5472.CAN-16-0562

Tian, L., Chen, M., He, Q., Yan, Q., and Zhai, C. (2020). MicroRNA-199a-5p S-uppresses C-ell P-roliferation, M-igration and I-nvasion by T-argeting

ITGA3 in C-olorectal C-ancer. *Mol. Med. Rep.* 22 (3), 2307–2317. doi:10.3892/mmr.2020.11323

Tsai, C. Y., Lin, T. A., Huang, S. C., Hsu, J. T., Yeh, C. N., Chen, T. C., et al. (2020). Is Adjuvant Chemotherapy Necessary for Patients with Deficient Mismatch Repair Gastric Cancer?-Autophagy Inhibition Matches the Mismatched. *Oncol.* 25 (7), e1021–e1030. doi:10.1634/theoncologist.2019-0419

Wang, J.-R., Liu, B., Zhou, L., and Huang, Y.-X. (2019). MicroRNA-124-3p Suppresses Cell Migration and Invasion by Targeting ITGA3 Signaling in Bladder Cancer. *Cbm* 24 (2), 159–172. doi:10.3233/CBM-182000

Wu, J., Chen, S., Liu, H., Zhang, Z., Ni, Z., Chen, J., et al. (2018). Tunicamycin Specifically Aggravates ER Stress and Overcomes Chemoresistance in Multidrug-Resistant Gastric Cancer Cells by Inhibiting N-Glycosylation. *J. Exp. Clin. Cancer Res.* 37 (1), 272. doi:10.1186/s13046-018-0935-8

Xu, Z., Chen, L., Xiao, Z., Zhu, Y., Jiang, H., Jin, Y., et al. (2018). Potentiation of the Anticancer Effect of Doxorubicinin Drug-Resistant Gastric Cancer Cells by Tanshinone IIA. *Phytomedicine* 51, 58–67. doi:10.1016/j.phymed.2018.05.012

Yoshioka, A., Miyata, H., Doki, Y., Yamasaki, M., Sohma, I., Gotoh, K., et al. (2008). LC3, an Autophagosome Marker, Is Highly Expressed in Gastrointestinal Cancers. *Int. J. Oncol.* 33 (3), 461–468. PMID: 18695874.

Zhang, X., Deibert, C. P., Kim, W.-J., Jaman, E., Rao, A. V., Lotze, M. T., et al. (2020). Autophagy Inhibition Is the Next Step in the Treatment of Glioblastoma Patients Following the Stupp Era. *Cancer Gene Ther.* 28, 971–983. doi:10.1038/s41417-020-0205-8

Zhang, X., Li, Z., Xuan, Z., Xu, P., Wang, W., Chen, Z., et al. (2018). Novel Role of miR-133a-3p in Repressing Gastric Cancer Growth and Metastasis via Blocking Autophagy-Mediated Glutaminolysis. *J. Exp. Clin. Cancer Res.* 37 (1), 320. doi:10.1186/s13046-018-0993-y

Zhao, R., Zhang, X., Zhang, Y., Zhang, Y., Yang, Y., Sun, Y., et al. (2020). HOTTIP Predicts Poor Survival in Gastric Cancer Patients and Contributes to Cisplatin Resistance by Sponging miR-216a-5p. *Front. Cel Dev. Biol.* 8, 348. doi:10.3389/fcell.2020.00348

Zhong, X.-p., Kan, A., Ling, Y.-h., Lu, L.-h., Mei, J., Wei, W., et al. (2019). NCKAP1 Improves Patient Outcome and Inhibits Cell Growth by Enhancing Rb1/p53 Activation in Hepatocellular Carcinoma. *Cell Death Dis* 10 (5), 369. doi:10.1038/s41419-019-1603-4

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 Genetics

ORIGINAL RESEARCH
published: 08 November 2021
doi: 10.3389/fgene.2021.736766

Inflammation-Related Long Non-Coding RNA Signature Predicts the Prognosis of Gastric Carcinoma

ShuQiao Zhang[1†], *XinYu Li*[2†], *ChunZhi Tang*[2*] *and WeiHong Kuang*[3*]

[1]First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine, Guangzhou, China, [2]Medical College of Acupuncture-Moxibustion and Rehabilitation, Guangzhou University of Chinese Medicine, Guangzhou, China, [3]Guangdong Key Laboratory for Research and Development of Natural Drugs, School of Pharmacy, Guangdong Medical University, Dongguan, China

Background: Gastric carcinoma (GC) is a molecularly and phenotypically highly heterogeneous disease, making the prognostic prediction challenging. On the other hand, Inflammation as part of the active cross-talk between the tumor and the host in the tumor or its microenvironment could affect prognosis.

Method: We established a prognostic multi lncRNAs signature that could better predict the prognosis of GC patients based on inflammation-related differentially expressed lncRNAs in GC.

Results: We identified 10 differently expressed lncRNAs related to inflammation associated with GC prognosis. Kaplan-Meier survival analysis demonstrated that high-risk inflammation-related lncRNAs signature was related to poor prognosis of GC. Moreover, the inflammation-related lncRNAs signature had an AUC of 0.788, proving their utility in predicting GC prognosis. Indeed, our risk signature is more precise in predicting the prognosis of GC patients than traditional clinicopathological manifestations. Immune and tumor-related pathways for individuals in the low and high-risk groups were further revealed by GSEA. Moreover, TCGA based analysis revealed significant differences in HLA, MHC class-I, cytolytic activity, parainflammation, co-stimulation of APC, type II INF response, and type I INF response between the two risk groups. Immune checkpoints revealed CD86, TNFSF18, CD200, and LAIR1 were differently expressed between lowand high-risk groups.

Conclusion: A novel inflammation-related lncRNAs (AC015660.1, LINC01094, AL512506.1, AC124067.2, AC016737.1, AL136115.1, AP000695.1, AC104695.3, LINC00449, AC090772.1) signature may provide insight into the new therapies and prognosis prediction for GC patients.

Keywords: gastric carcinoma (GC), inflammation, lncRNAs, long non-coding RNAs, prognosis (carcinoma), immune infiltration, signature

Abbreviation: GC, Gastric carcinoma; OS, overall survival (OS); DEGs, differentially expressed genes; TCGA, The Cancer Genome Atlas; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, Gene set enrichment analyses; FDR, false discovery rate; ssGSEA, single-sample gene set enrichment analysis (ssGSEA).

OPEN ACCESS

Edited by:
Jian-Bing Fan,
Illumina, United States

Reviewed by:
Zhipeng Liu,
Purdue University, United States
Chao Di,
Children's Hospital of Philadelphia, United States

***Correspondence:**
ChunZhi Tang
jordan664@163.com
WeiHong Kuang
Kuangwh@gdmu.edu.cn

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *05 July 2021*
Accepted: *05 October 2021*
Published: *08 November 2021*

Citation:
Zhang S, Li X, Tang C and Kuang W (2021) Inflammation-Related Long Non-Coding RNA Signature Predicts the Prognosis of Gastric Carcinoma. Front. Genet. 12:736766. doi: 10.3389/fgene.2021.736766

INTRODUCTION

Gastric carcinoma (GC) is a highly lethal and aggressive cancer, the third most common cause of cancer death globally, a disease with high molecular and phenotypic heterogeneity (Smyth et al., 2020), with over 1 million estimated new cases annually (Hacker and Lordick, 2015). *Helicobacter pylori* (HP) infection, age, high salt intake, and diets low in fruit and vegetables are significant risk factors for GC progression (Smyth et al., 2020). Although combined treatment, including surgery, chemo-radiotherapy, and chemotherapy, have shown remarkable improvements, the prognosis of GC remains undesirable (Shah and Kelsen, 2010; Cristescu et al., 2015). Moreover, the current TNM staging system of the American Joint Committee on Cancer (AJCC) or the Joint International Committee on Cancer (UICC) has shown valuable but insufficient prognostication for the prognosis and estimation of subsets of GC patients (Marano et al., 2015; Kim et al., 2016). Therefore, new biomarkers are needed to discriminate high-risk patients with GC to improve personalized cancer treatment.

The study of tumor-associated inflammation has increased rapidly in the past few decades, and it has also been described as a hallmark of cancer (Hanahan and Weinberg, 2011). Many cancers arise from irritation, chronic infection, and Inflammation, and the tumor microenvironment is mainly coordinated by inflammatory cells, which are indispensable players in fostering proliferation, neoplastic process, survival, and migration. Inflammation as an opportunity for anti-cancer therapies is on the rise. At the same time, long non-coding RNAs (lncRNAs) are a subset of non-coding RNA molecules with about 200 nt, which regulates gene expression and participates in various biological regulatory processes, including the regulatory process related to tumor genesis, progression, and metastasis (Gupta et al., 2010). At present, functional lncRNAs are considered critical roles in several biological regulatory processes, such as cell growth, development, angiogenesis, and inflammation (Ng et al., 2013; Li et al., 2014; Michalik et al., 2014; Ghazal et al., 2015). One recent study revealed that lincRNA-Cox2 and lncRNA NEAT1 promotes IL-6, an inflammatory cytokine, expression via distinct mechanisms. LncRNA NEAT1 acts further upstream and potentiates IL-6 expression by promoting the JNK1/2 and ERK1/2 signaling cascades (Ma et al., 2020). In a related study, Zhou et al. (2016) found that lncRNA ANRIL is involved in TNF-α-NF-κB signaling to regulate the inflammatory response in endothelial cells. Nevertheless, to date, serial studies that systematically evaluate inflammation-related lncRNA prognostic signature and its correlation with GC patients remain scarce.

In our study, a prognostic multi lncRNA signature of inflammation-related differentially expressed lncRNAs based on the Cancer Genome Atlas (TCGA) data was established for the first time. Then, we investigated the role of inflammation-related mRNA, immune responses, and N6-methylated-adenosine (m6A) mRNA status in GC prognosis.

METHODS

Data Collection

RNA-sequence (32 normal and 375 tumor) data of 443 patients were extracted from the TCGA-STAD database. Clinical characteristics of the GC patients used in this study are shown in **Table 1**. The corresponding inflammation-related genes were downloaded from The Molecular Signatures database (MSigDB, http://www.gsea-msigdb.org/gsea/login.jsp) (Liberzon et al., 2015), a collection of annotated gene sets for use with Gene Set Enrichment Analysis (GSEA) software, which provided comprehensive and gene sets for inflammatory markers. Overall, we identified 561 inflammation-related genes (**Supplementary Table S1**). The association between lncRNAs in TCGA-LIHC dataset and inflammation-related genes from MSigDB was assessed by Pearson correlation analysis. If the correlation coefficient $|R^2|$ was greater than 0.6 and $p < 0.001$, the correlation is considered significant, and then inflammation-related lncRNAs were selected. The clinicopathological data of GC patients were collected, including grade, age, stage, TMN, gender, survival time, and survival status. False discovery rate (FDR) < 0.05 and $|\log_2 FC| \geq 1$ were set as the significant differential expression of lncRNAs related to Inflammation. Firstly, we explored the functions of up and downregulated Inflammation related differentially expressed genes (DEGs). Then the biological processes, cellular components, and molecular functions associated with the DEGs were then evaluated by Gene Ontology (GO). Based on data from Kyoto Encyclopedia of Genes and Genomes (KEGG), the functions of biological pathways by different expressions of Inflammation-related lncRNAs were further analyzed in R software.

TABLE 1 | The clinical characteristics of patients in the TCGA dataset.

Variable	Number of samples
Gender	
Male/Female	285/158
Age at diagnosis	
≤65/>65/NA	197/241/5
Grade	
G1/G2/G3/NA	12/159/263/9
Stage	
I/II/III/IV/NA	59/130/183/44/27
T	
T1/T2/T3/T4/NA	23/93/198/119/10
M	
M0/M1/NA	391/30/22
N	
N0/N1/N2/N3/NA	132/119/85/88/19

Construction of the Inflammation-Related lncRNAs Prognostic Signature

To construct a robust and stable prognostic, predictive signature, we first used univariate Cox regression and lasso regression analysis, and finally used multivariate Cox regression analysis to construct an inflammation-related lncRNA signature, and stratified them based on the risk score (βlncRNA1 × ExpressionlncRNA1 + βlncRNA2 × ExpressionlncRNA2 + βlncRNA3×ExpressionlncRNA3 +. . .+ βlncRNAn × ExpressionlncRNAn). The relevant risk score of each GC patient was also evaluated. The lncRNAs were divided into low-risk (<median number) and high-risk (≥median number) groups based on median scores.

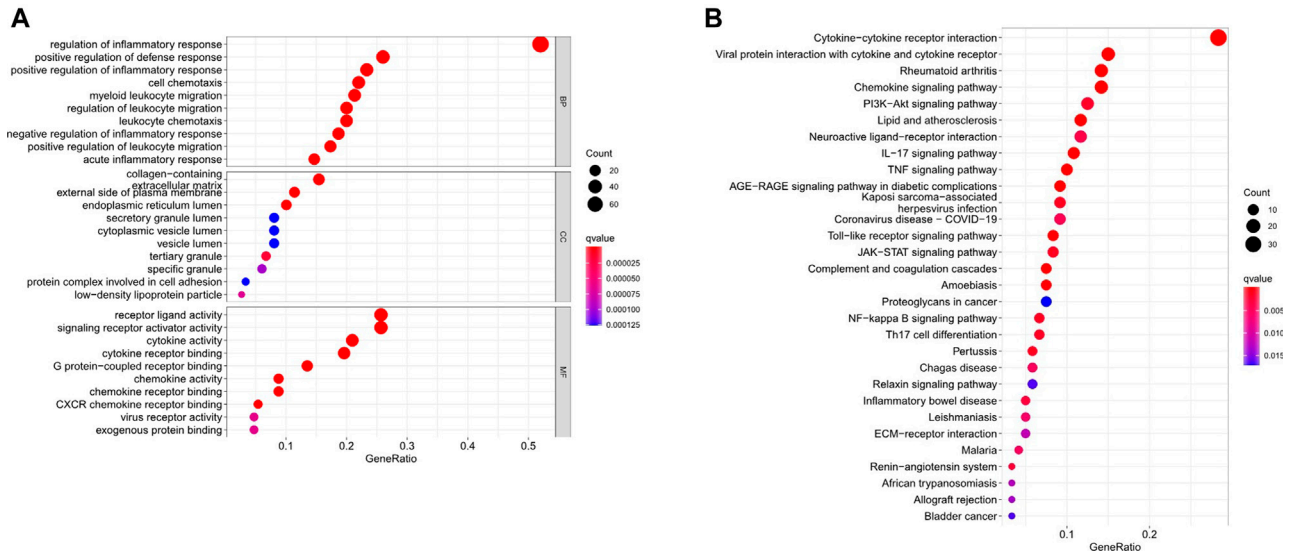


FIGURE 1 | Go and KEGG analysis of differentially expressed inflammation-related genes. **(A)** Biological processes, molecular functions, and cellular components analysis of GO **(B)** Pathway enrichment analysis of KEGG.

The Predictive Nomogram

To establish a clinically feasible method to predict the survival rate of GC patients, we set a prediction model using nomograms and considered the clinicopathological manifestations of GC patients. The nomogram predicted 1 -, 2 -, and 3-years overall survival (OS) of GC patients, and the statistical significance was based on multivariate analysis of OS rate. The predictive factors included risk model and grade, age, stage, TMN, gender.

Immunity Analysis and Related Gene Expression

Meanwhile, the relative immune cell infiltration level of individual samples was quantified by single-sample gene set enrichment analysis (ssGSEA) (Yi et al., 2020) comparing the CIBERSORT (Newman et al., 2015; Charoentong et al., 2017), CIBERSORT–ABS (Wang et al., 2020), QUANTISEQ (Plattner et al., 2020), MCPCOUNTER (Shi et al., 2020a), XCELL (Aran et al., 2017), EPIC (Racle et al., 2017) and TIMER (Li et al., 2017a)algorithms to evaluate the cellular component or cellular immune response lncRNAs signature between the high-risk and low-risk groups based on inflammation correlation. Then, the heatmap demonstrates the differences in immune responses under different algorithms. In addition, we used the ssGSEA method to quantitatively analyze the tumor-infiltrating immune cell subsets in the high-risk and low-risk group and evaluate their immune functions.

Statistical Analysis

All the statistical analyses of data were performed in R software (version 4.0.3) with Bioconductor packages including "limma," "survival," "survminer." Non paired *t*-test and Wilcoxon test were used to analyze the normal distribution and non-normal distribution variables. Based on FDR, the different expression of lncRNAs was corrected by the Benjamin Hochberg method to control the elevated false-positive rate. To analyze the GC DEGs associated immune status in each sample in the TCGA cohort, the relative infiltration of 20 immune cell types in the tumor microenvironment was calculated via ssGSEA with the application of the"GSVA" package in R. The prognostic performance of the GC predictive signature compared with other clinicopathological features was measured using the time-dependent receiver operating characteristic (ROC) and the decision curve analysis (DCA) (Vickers et al., 2008). Logistic regression analysis and a heatmap graph were used to analyze the association between inflammation-related lncRNAs and clinicopathological manifestations. Kaplan Meier survival analysis was used to evaluate the OS of GC patients based on Inflammation-related lncRNA signature. For each analysis, $p < 0.05$ in the results was considered statistically significant.

RESULTS

Enrichment Analysis of Inflammation-Related Genes

We uncovered 150 inflammation-related DEGs (114 downregulated and 36 upregulated; **Supplementary Table S2**). Biological processes participated in the inflammatory response, defense response, and immune system process, among others. Molecular functions mainly regulated receptor ligand and signaling receptor activator activity, cytokine activity, and chemokine activity. Cellular components were regulated primarily on the collagen-containing extracellular matrix, external side of the plasma membrane, and endoplasmic reticulum lumen. Furthermore, KEGG analysis showed that the differentailly expressed genes were significantly enriched in Cytokine-cytokine receptor interaction, IL-17 signaling pathway, PI3K-AKT signaling

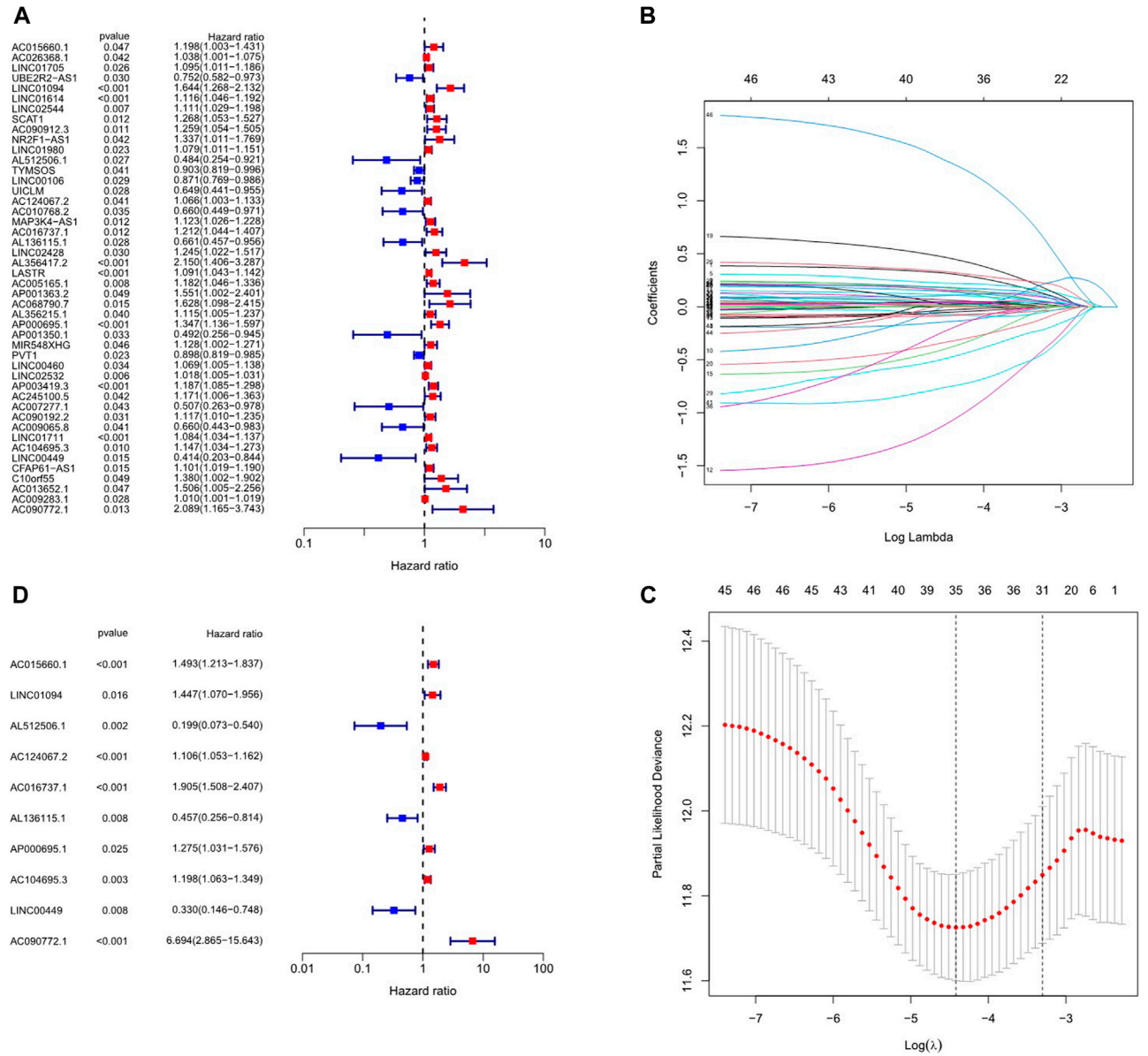


FIGURE 2 | Construction of inflammation-related lncRNAs prognostic signature. **(A)** Univariate cox regression illustrated 46 inflammation-related lncRNAs associated with prognosis. **(B, C)** Inflammation-related lncRNAs with $p < 0.05$ in univariate Cox analysis were penalized by LASSO regression analysis. **(D)** Multivariate cox regression analysis revealed 10 inflammation-related lncRNAs prognostic signature of GC.

pathway, TNF signaling pathway, JAK-STAT signaling pathway, Toll-like receptor signaling pathway, Chemokine signaling pathway, and NF-kappa B signaling pathway (**Figure 1**).

The Inflammation-Based lncRNAs Prognostic Signature

In the preliminary screening, 46 inflammation-related lncRNAs were obtained to be associated with OS by univariate Cox analysis (**Figure 2A**). Next, inflammation-related lncRNAs with $p < 0.05$ in univariate Cox analysis were penalized by Lasso regression (**Figures 2B,C**). Finally, 10 inflammation-related lncRNAs signature were found to be independent prognostic factors of GC patients by multivariate Cox regression analysis. (**Figure 2D**). A novel risk score was calculated by multiplying the lncRNA expression of each lncRNA and its corresponding coefficient, which was obtained by multivariate Cox regression analysis. The formula of the risk score was as follows: risk score = (Coefficient lncRNA1 × expression of lncRNA1) + (Coefficient lncRNA2 × expression of lncRNA2) + + (Coefficient lncRNAn × expression lncRNAn). Overall, we evaluated

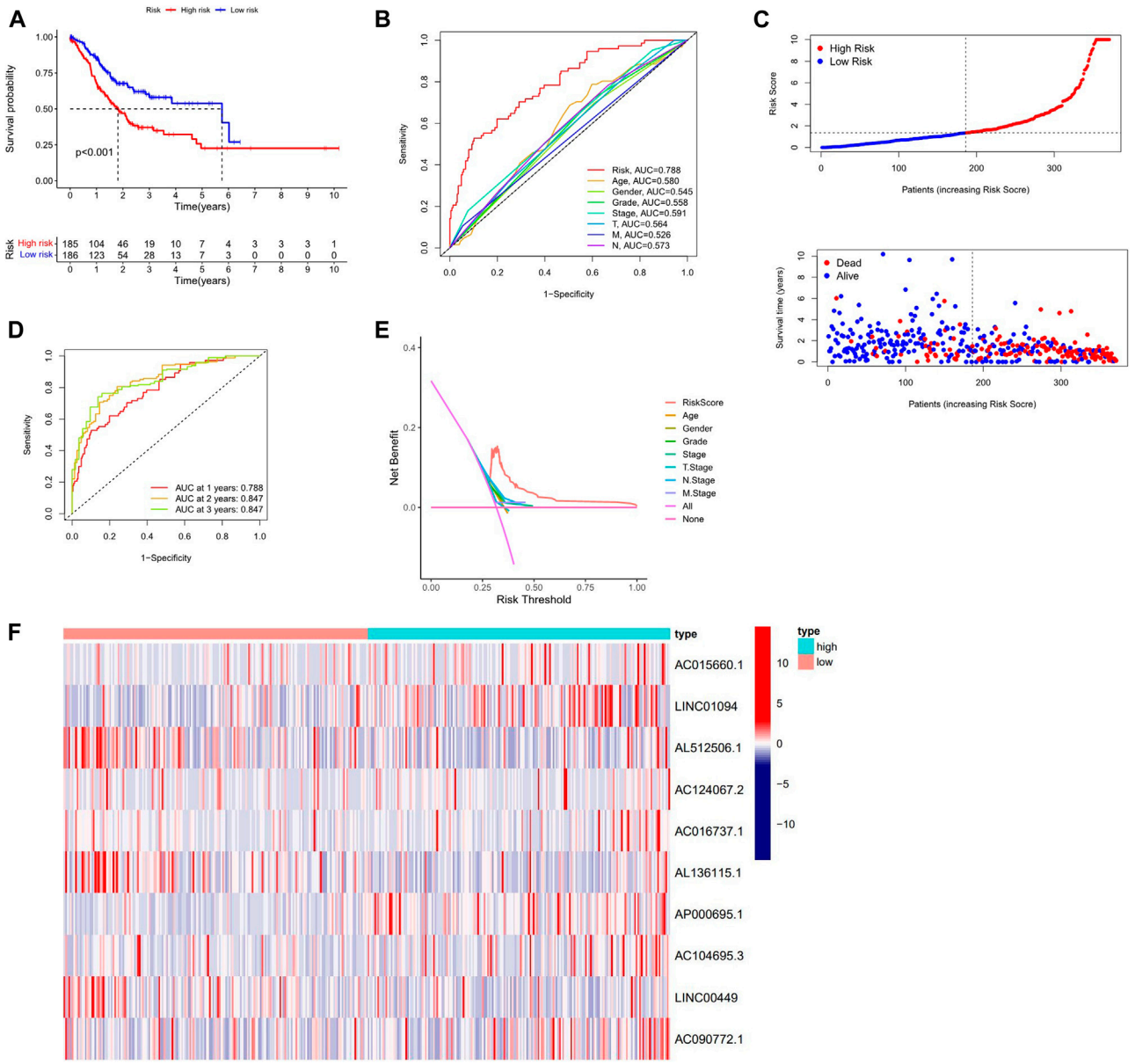


FIGURE 3 | Validation of the Inflammation-related lncRNAs signature. **(A)** Survival curves result, **(B)**. The AUC values for forecasting OS based on risk factors, **(C)**. Risk score distribution and survival status of GC patients, **(D)** AUC of time-dependent ROC curves represented of 1, 2, 3-years survival rates of GC and risk score **(E)** The DCA plot. **(F)** Expression profile heatmap of the 10 inflammation-related lncRNAs signature.

the risk score and identified 10 differently expressed inflammation-related lncRNAs as independent prognosis predictors of GC (**Supplementary Table S3**).

Survival Results and Multivariate Examination

Kaplan-Meier analysis revealed that GC patients in the high-risk group had poorer survival rates than those in the low-risk group (**Figure 3A**). Meanwhile, the signature lncRNAs, with an AUC of 0.788, exhibited superior performance in predicting GC prognosis than traditional clinicopathological features (**Figures 3B,E**). We used the patients' risk survival status plots and found that the patients' risk scores were inversely related to the GC patients' survival. The heatmap revealed that the novel inflammation-related lncRNAs identified in our study were positively correlated with the risk model. (**Figure 3C**). The AUC for ROC analysis was 0.788, 0.847, and 0.847 for the predictive value of the 10 inflammation-related lncRNA signatures of GC patients for 1 -, 2 -, 3-years survival, respectively, (**Figure 3D**). Besides, the DCA

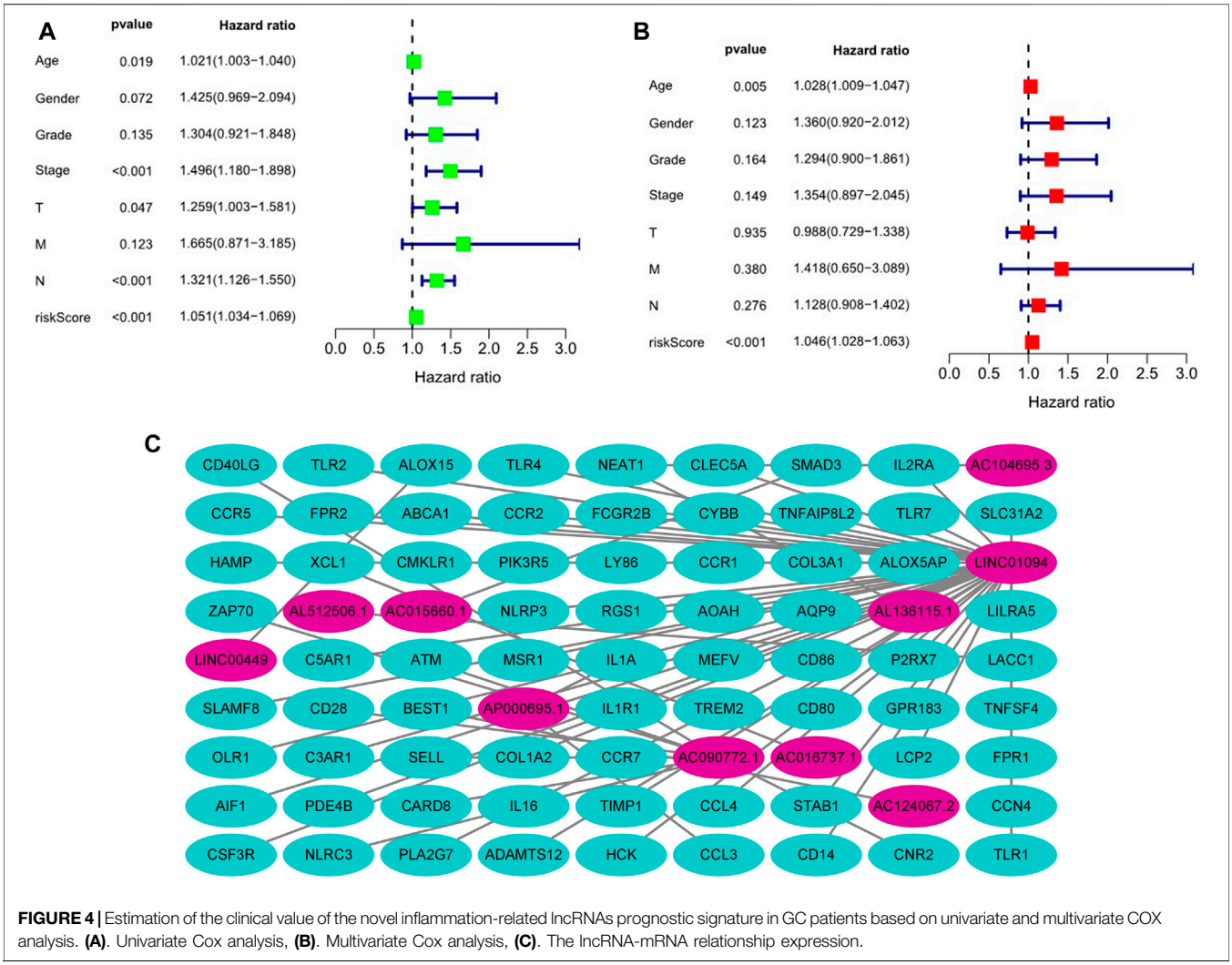


FIGURE 4 | Estimation of the clinical value of the novel inflammation-related lncRNAs prognostic signature in GC patients based on univariate and multivariate COX analysis. **(A)**. Univariate Cox analysis, **(B)**. Multivariate Cox analysis, **(C)**. The lncRNA-mRNA relationship expression.

plot indicated a robust and stable prognostic-predictive ability of the inflammation-related lncRNAs signature (**Figure 3E**). The heatmap demonstrated the difference in the expression of ten inflammation-related lncRNAs in the two risk population groups (**Figure 3F**). Univariate and multivariate Cox analysis revealed that the novel lncRNAs signature (HR: 1.05, 95CI: 1.03–1.07) was an independent prognosis predictor of GC patients (**Figures 4A,B**). The lncRNA-mRNA relationship was shown in the network (**Figure 4C**). Also, the heatmap analyzed the correlation between the novel prognostic signature and clinical-pathological features (**Figure 5**). The nomogram, which combined both the novel prognostic signature and the clinicopathological characteristics (**Figure 6**), was accurate and stable and thus applied to the clinical management for GC patients.

Gene Set Enrichment Analyses

To investigate the underlying pathways and biological processes, GSEA was used to reveal the significant pathways by which novel inflammation-related lncRNA signatures regulate tumor-associated and immune, such as natural killer cell-mediated cytotoxicity, primary immunodeficiency, T cell receptor signaling pathway, Toll-like receptor signaling pathway, MAPK signaling pathway, Notch signaling pathway, JAK-STAT signaling pathway and pathways in cancer (**Figure 7**; **Supplementary Table S4**).

Immunological Reaction and Related Gene Expression

The heatmap showed the immunological responses based on QUANTISEQ, CIBERSORT, CIBERSORT-ABS, MCPCOUNTER, XCELL, TIMER, and EPIC algorithms. (**Figure 8**). ssGSEA correlation analysis of immune cell subsets with relevant functions based on TCGA-STAD data showed cytolytic activity, MHC class-I, HLA, type I INF response, type II INF response, parainflammation, and co-stimulation of APC differed significantly between the high and low-risk groups (**Figure 9A**). Given the significance of checkpoint inhibitor-based immunotherapy, differences in immune checkpoint expression among high-risk and low-risk groups were further explored. Among two risk groups, significant differences were discovered in gene expression, including CD86,

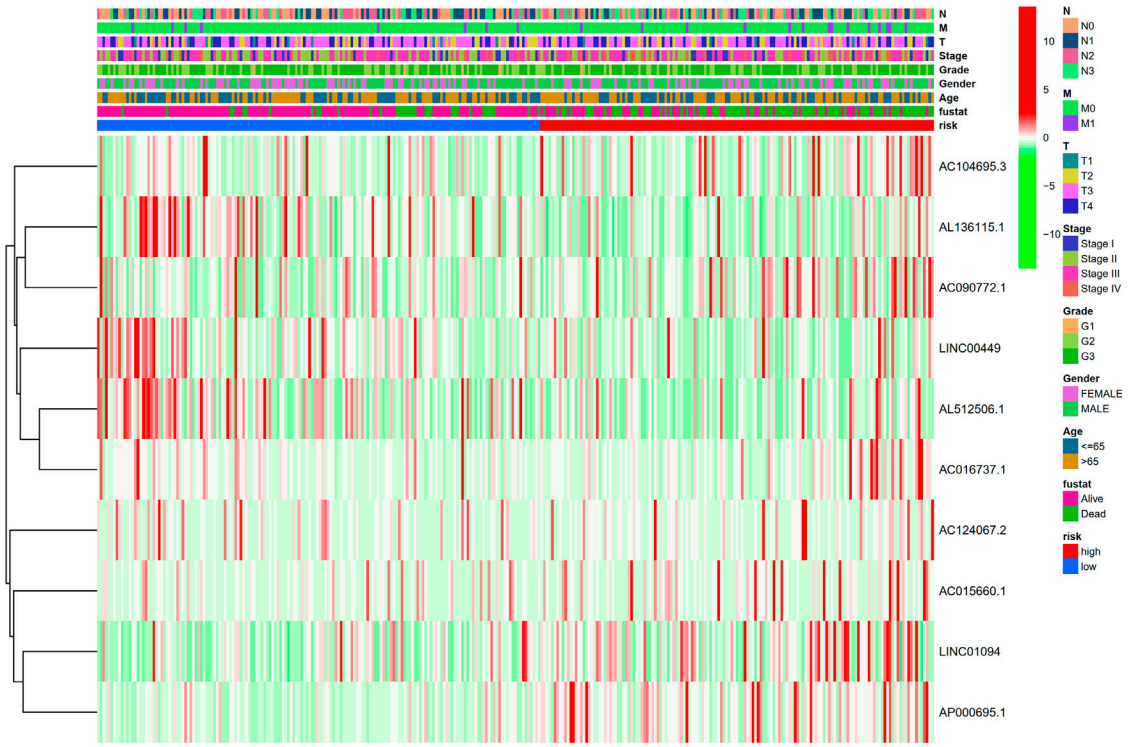


FIGURE 5 | Heatmap for clinical pathology manifestations and inflammation-related lncRNAs prognostic signature.

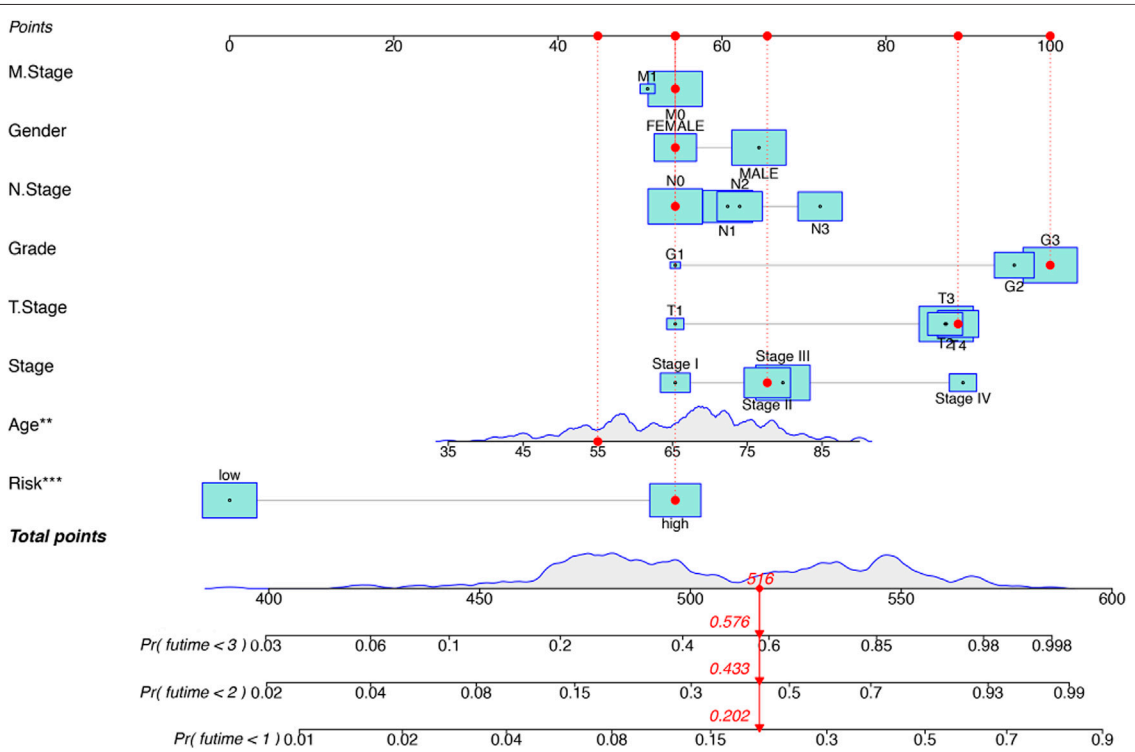


FIGURE 6 | Construction of a hybrid nomogram based on both prognostic inflammation-related lncRNAs and prognostic-related clinicopathological factors.

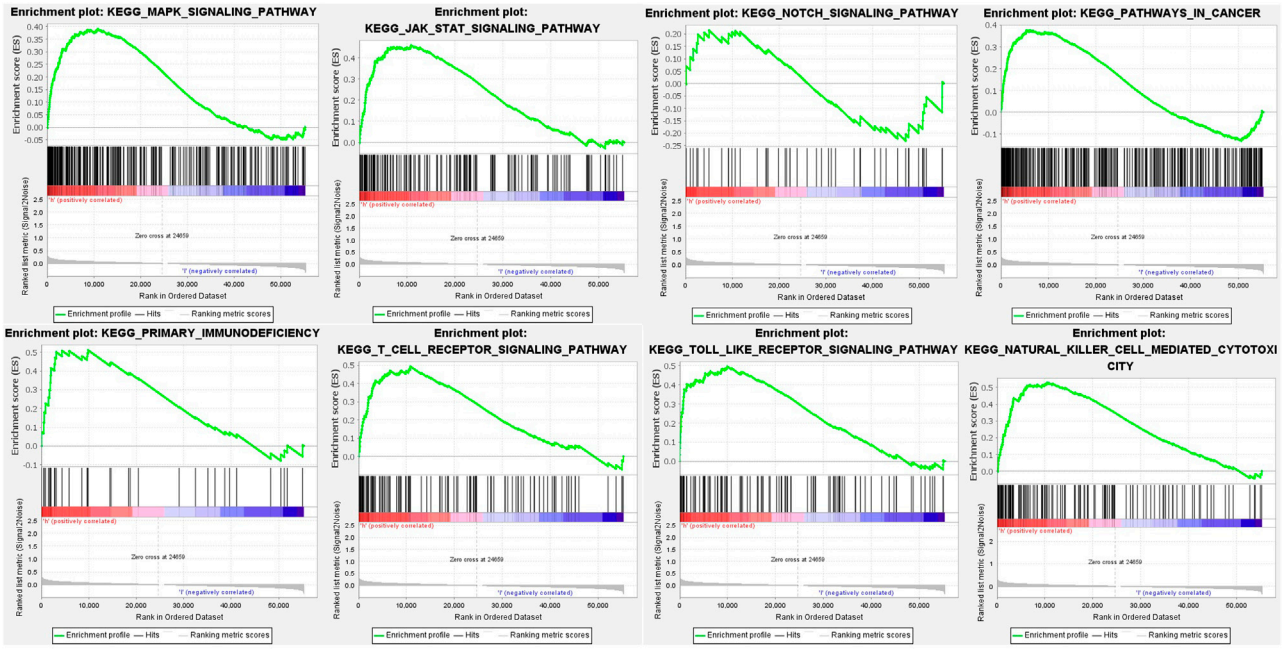


FIGURE 7 | GSEA for inflammation-related lncRNAs signature.

TNFSF18, CD200, and LAIR1 (**Figure 9B**). Comparison of m6A-related mRNA expression among high- and low-risk groups revealed the significant expression of HNRNPC, RBM15, YTHDC1, YTHDF2, FTO, ZC3H13, and METTL3 (**Figure 10**).

DISCUSSION

Inflammation is of essentiality in the development of cancer because chronic Inflammation has been shown to increase cancer risk by causing tumor initiation, promotion, and metastatic progression (So et al., 2020). Meanwhile, one key hallmark of cancer is the ability of cancer cells to evade immune destruction (Hanahan and Weinberg, 2011). Thus, it has the potential to become a novel tumor therapy. In our study, a novel inflammation-related prognostic lncRNA signature was first constructed based on the existing clinical characteristic of GC patients from the TCGA dataset. Then, we investigated the potential roles of immune infiltrating cells and immune checkpoint inhibitors in the tumor microenvironment in GC prognosis. The findings of our study revealed potential biomarkers and therapeutic targets in the inflammation signaling pathways. Overall, our analyses uncovered 150 inflammation-related DEGs. KEGG analysis further indicated that these genes were mainly involved in the IL-17 signaling pathway, PI3K-Akt signaling pathway, Toll-like receptor signaling pathway, TNF signaling pathway, and JAK-STAT signaling pathway, and so forth. Some recent studies found that Toll-like receptors play a vital role in the innate immune system, particularly in the inflammatory response (Takeda and Akira, 2004), and Dectin-1 could suppress Toll-like receptor four signals to protect against chronic liver disease in hepatic inflammatory cells (Seifert et al., 2015). Balkwill (2006) reported that one of the vital chemical mediators implicated in inflammation-related cancer is TNFα-α; it is involved in the promotion and progression of experimental and human cancers, and its pathways lead to the activation of NF-κKey to the B, and AP-1 transcription factor complex is the intracellular connection.

Blocking the link of inflammation to cancer might be potentially a new way for tumor treatment. Overall, in our study, the ten differently expressed inflammation-related lncRNA signature found was an independent prognostic predictor of GC. A recent study found LINC01094 could promote the proliferation, migration, invasion and epithelial-mesenchymal transition of ovarian cancer cells by adsorbing miR-577 (Xu et al., 2020). Besides, LINC01094 could promote clear cell renal cell carcinoma development through the miR-224-5p/CHSY1 regulatory axis (Jiang et al., 2020). Linc01094 promotes proliferation, migration and invasion of glioma cells by adsorbing miR-330-3p and upregulating MSI1 expression (Zhu et al., 2020). In a related study, overexpression of LINC00449 could inhibit acute myeloid leukemia cell invasion and cell proliferation *in vitro* and *in vivo*, and the LINC00449/miR-150/FOXD3 signaling pathway may be a novel prognostic biomarker or therapeutic target for the treatment of acute myeloid leukemia (Shi et al., 2020b). Thus, it is of importance to establish a novel prognostic multiinflammation-related lncRNAs signature for GC patients. Such findings may provide valuable insights for cancer control in the future.

The Inflammation-related lncRNAs with different expressions were divided into low-risk and high-risk groups to investigate their potential roles in GC. The Inflammation-related lncRNA signature

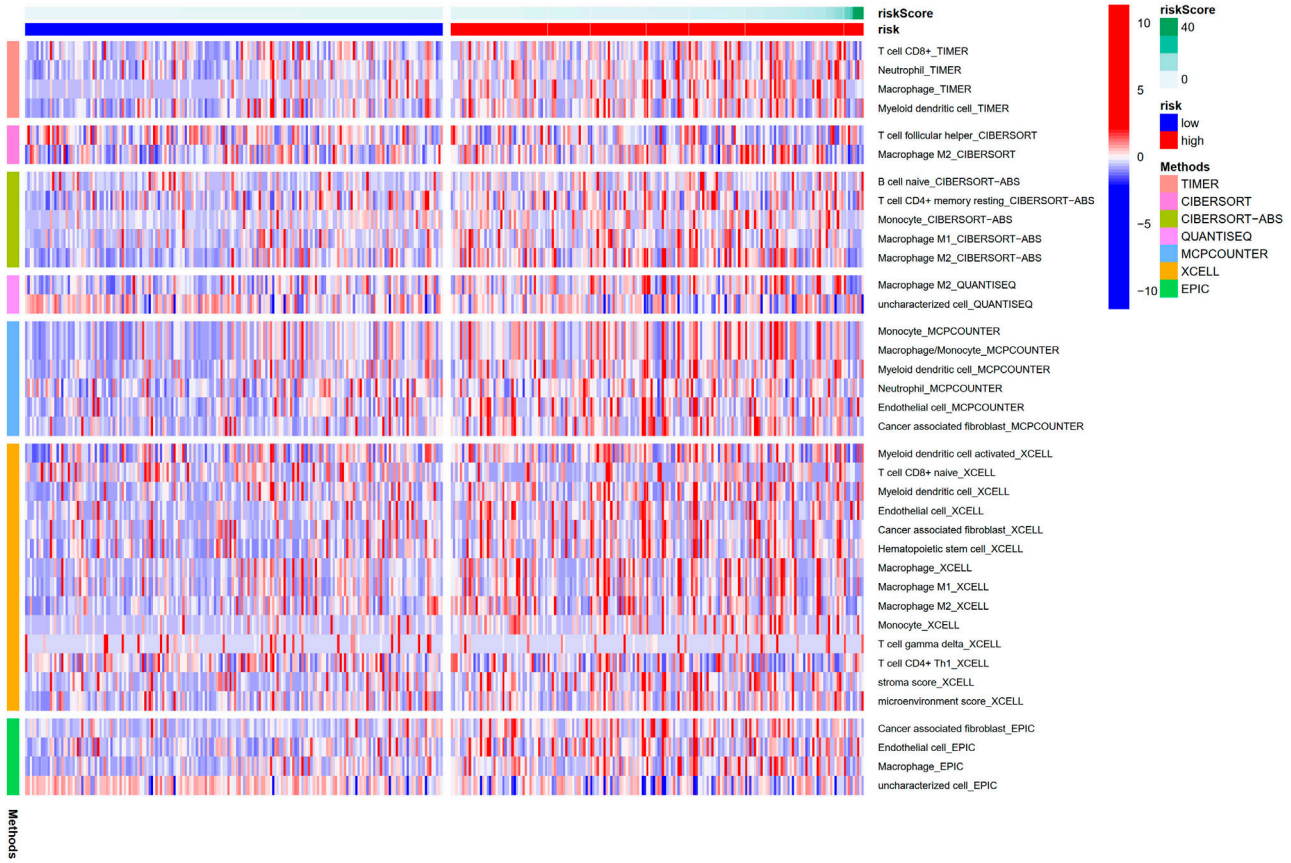


FIGURE 8 | The immune responses among the high and low-risk group.

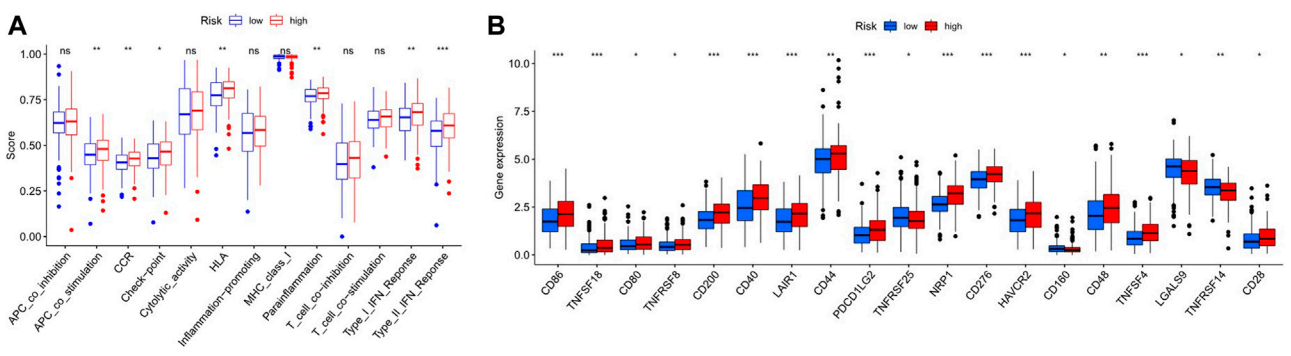


FIGURE 9 | (A). Correlation of immune cell subsets with related functions **(B)**. Comparison of immune checkpoints among the two GC risk groups.

of the high-risk group was correlated with poor prognosis of GC patients in the study results. Moreover, the inflammation-related lncRNA signature had an AUC of 0.788 but performed well in other validations such as DCA, demonstrating their utility in predicting GC prognosis and indicating that our risk signature also outperforms traditional clinic pathology characteristics.

The direct correlations between lncRNAs and cancer-derived Inflammation, epithelial-to-mesenchymal transition, metastasis, and other hallmarks of cancer indicate their potential as cancer biomarkers and targets (Tamang et al., 2019). LncRNA H19 could be induced by *helicobacter pylori* infection to promote gastric cancer cell growth via strengthening NF-κB-induced inflammation (Zhang et al., 2019). Besides, CCAT1 lncRNA Promotes Inflammatory bowel disease malignancy by destroying Intestinal Barrier via downregulating miR-185-3p (Ma et al., 2019). Increasing evidence suggests that lncRNA is critical in mediating

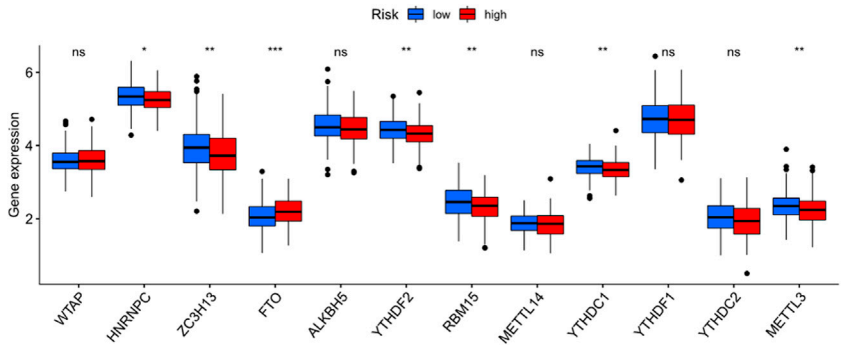


FIGURE 10 | Comparison of m6A-related genes expression among the two GC risk group.

both inflammation and cancer (Barsyte-Lovejoy et al., 2006; Guo et al., 2019; Song et al., 2020; Sun et al., 2020).

Our study, GSEA revealed immune and tumor-related pathways for individuals in the low- and high-risk groups. TCGA revealed significant differences in HLA, MHC class-I, cytolytic activity, parainflammation, co-stimulation of APC, type II INF response, and type I INF response between the low and high-risk groups. Moreover, Immune checkpoints such as CD86, TNFSF18, CD200, and LAIR1 were differently expressed between high and low-risk groups. Li et al. speculated that Inflammation triggered by lncRNA CRNDE could regulate tumorigenesis and development through the TLR pathway (Li et al., 2017b). In recent years, lncRNAs have gained attention as essential regulators of gene expression acting through versatile interactions with DNA, RNA, or proteins. Intriguingly, lncRNAs play crucial roles in modulating innate immune cell development and inflammatory gene expression (Elling et al., 2016). Activating inflammatory pathways, particularly the IFN response, can sensitize animal cancer models and cancer patients to immune checkpoint inhibitors and positively contribute to anti-tumor activity (Zemek et al., 2019). At present, few studies have explored the relationship among inflammation and immune checkpoint inhibitors. Therefore, lncRNAs might be critical players in the transformation of inflammation to cancer through the immune microenvironment.

Although we identified a predictive risk model with ten inflammation-related lncRNAs and confirmed that the risk model was significantly associated with Inflammation, this work has limitations. The study conducted with bioinformatics analysis was not robust enough and needed to be confirmed via experimental verification. Hence, further laboratory experiments, including a multicenter study with larger sample sizes, are required. In summary, in this study, we developed a robust prognostic risk model with ten inflammation-related lncRNAs. Compared to other clinical parameters, the risk score is an independent prognostic index. Therefore, this risk model may serve as a prognostic signature and provide clues for individualized immunotherapy for GC patients.

CONCLUSION

Specific Inflammation-related lncRNAs can provide individualized prediction of GC patients' 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

SZ designed and analyzed the research study; SZ wrote and revised the article, XL collected and analyzed the data and all authors have read and approved the article.

FUNDING

Discipline construction project of Guangdong Medical University (No. 4SG21009G).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.736766/full#supplementary-material.

REFERENCES

Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Genome Biol.* 18, 220. doi:10.1186/s13059-017-1349-1

Balkwill, F. (2006). TNF-α in Promotion and Progression of Cancer. *Cancer Metastasis Rev.* 25, 409–416. doi:10.1007/s10555-006-9005-3

Barsyte-Lovejoy, D., Lau, S. K., Boutros, P. C., Khosravi, F., Jurisica, I., Andrulis, I. L., et al. (2006). The C-Myc Oncogene Directly Induces the H19 Noncoding RNA by Allele-specific Binding to Potentiate Tumorigenesis. *Cancer Res.* 66, 5330–5337. doi:10.1158/0008-5472.can-06-0037

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cel Rep.* 18, 248–262. doi:10.1016/j.celrep.2016.12.019

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

Elling, R., Chan, J., and Fitzgerald, K. A. (2016). Emerging Role of Long Noncoding RNAs as Regulators of Innate Immune Cell Development and Inflammatory Gene Expression. *Eur. J. Immunol.* 46, 504–512. doi:10.1002/eji.201444558

Ghazal, S., McKinnon, B., Zhou, J., Mueller, M., Men, Y., Yang, L., et al. (2015). H19 Lnc RNA Alters Stromal Cell Growth via IGF Signaling in the Endometrium of Women with Endometriosis. *EMBO Mol. Med.* 7, 996–1003. doi:10.15252/emmm.201505245

Guo, L., Li, L., Zhang, Y., Fu, S., Zhang, J., Wang, X., et al. (2019). Long Non-coding RNA Profiling in LPS-Induced Intestinal Inflammation Model: New Insight into Pathogenesis. *Innate Immun.* 25, 491–502. doi:10.1177/1753425919872812

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

Hacker, U., and Lordick, F. (2015). Aktuelle Standards in der Therapie des Magenkarzinoms. *Dtsch Med. Wochenschr* 140, 1417. doi:10.1055/s-0041-106868

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of Cancer: the Next Generation. *Cell* 144, 646–674. doi:10.1016/j.cell.2011.02.013

Jiang, Y., Zhang, H., Li, W., Yan, Y., Yao, X., and Gu, W. (2020). FOXM1-Activated LINC01094 Promotes Clear Cell Renal Cell Carcinoma Development via MicroRNA 224-5p/CHSY1. *Mol. Cel Biol* 40. doi:10.1128/MCB.00357-19

Kim, B. S., Park, Y. S., Yook, J. H., and Kim, B. S. (2016). Comparison of the Prognostic Values of the 2010 WHO Classification, AJCC 7th Edition, and ENETS Classification of Gastric Neuroendocrine Tumors. *Medicine (Baltimore)* 95, e3977. doi:10.1097/md.0000000000003977

Li, H., Li, Q., Guo, T., He, W., Dong, C., and Wang, Y. (2017). Lncrna Crnde Triggers Inflammation Through The Tlr3-Nf-Kb-Cytokine Signaling Pathway. *Tumour Biol.* 39, 1010428317703821. doi:10.1177/1010428317703821

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* 77, e108–e110. doi:10.1158/0008-5472.can-17-0307

Li, Z., Chao, T.-C., Chang, K.-Y., Lin, N., Patil, V. S., Shimizu, C., et al. (2014). The Long Noncoding Rna Thril Regulates Tnf Expression Through Its Interaction With Hnrnpl. *Proc. Natl. Acad. Sci.* 111, 1002–1007. doi:10.1073/pnas.1313768111

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. *Cel Syst.* 1, 417–425. doi:10.1016/j.cels.2015.12.004

Ma, D., Cao, Y., Wang, Z., He, J., Chen, H., Xiong, H., et al. (2019). CCAT1 lncRNA Promotes Inflammatory Bowel Disease Malignancy by Destroying Intestinal Barrier via Downregulating miR-185-3p. *Inflamm. Bowel Dis.* 25, 862–874. doi:10.1093/ibd/izy381

Ma, S., Long, T., and Huang, W. J. M. (2020). Noncoding RNAs in Inflammation and Colorectal Cancer. *RNA Biol.* 17, 1628–1635. doi:10.1080/15476286.2019.1705610

Marano, L., Boccardi, V., Braccio, B., Esposito, G., Grassia, M., Petrillo, M., et al. (2015). Comparison Of The 6th And 7th Editions Of The Ajcc/Uicc Tnm Staging System For Gastric Cancer Focusing On The "N" Parameter-Related Survival: The Monoinstitutional Nodus Italian Study. *World J. Surg. Onc* 13, 215. doi:10.1186/s12957-015-0633-3

Michalik, K. M., You, X., Manavski, Y., Doddaballapur, A., Zörnig, M., Braun, T., et al. (2014). Long Noncoding RNA MALAT1 Regulates Endothelial Cell Function and Vessel Growth. *Circ. Res.* 114, 1389–1397. doi:10.1161/circresaha.114.303265

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

Ng, S.-Y., Lin, L., Soh, B. S., and Stanton, L. W. (2013). Long Noncoding RNAs in Development and Disease of the central Nervous System. *Trends Genet.* 29, 461–468. doi:10.1016/j.tig.2013.03.002

Plattner, C., Finotello, F., and Rieder, D. (2020). Deconvoluting Tumor-Infiltrating Immune Cells from RNA-Seq Data Using quanTIseq. *Methods Enzymol.* 636, 261–285. doi:10.1016/bs.mie.2019.05.056

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

Seifert, L., Deutsch, M., Alothman, S., Alqunaibit, D., Werba, G., Pansari, M., et al. (2015). Dectin-1 Regulates Hepatic Fibrosis and Hepatocarcinogenesis by Suppressing TLR4 Signaling Pathways. *Cel Rep.* 13, 1909–1921. doi:10.1016/j.celrep.2015.10.058

Shah, M. A., and Kelsen, D. P. (2010). Gastric Cancer: a Primer on the Epidemiology and Biology of the Disease and an Overview of the Medical Management of Advanced Disease. *J. Natl. Compr. Canc Netw.* 8, 437–447. doi:10.6004/jnccn.2010.0033

Shi, J., Jiang, D., Yang, S., Zhang, X., Wang, J., Liu, Y., et al. (2020). LPAR1, Correlated with Immune Infiltrates, Is a Potential Prognostic Biomarker in Prostate Cancer. *Front. Oncol.* 10, 846. doi:10.3389/fonc.2020.00846

Shi, Y., Zhu, Y., Zheng, X., and Zheng, Z. (2020). LINC00449 Regulates the Proliferation and Invasion of Acute Monocytic Leukemia and Predicts Favorable Prognosis. *J. Cel Physiol* 235, 6536–6547. doi:10.1002/jcp.29487

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

So, A. R., Si, J. M., Lopez, D., and Pellegrini, M. (2020). Molecular Signatures for Inflammation Vary across Cancer Types and Correlate Significantly with Tumor Stage, Sex and Vital Status of Patients. *PLoS One* 15, e0221545. doi:10.1371/journal.pone.0221545

Song, B., Ye, L., Wu, S., and Jing, Z. (2020). Long Non-coding RNA MEG3 Regulates CSE-Induced Apoptosis and Inflammation via Regulating miR-218 in 16HBE Cells. *Biochem. Biophysical Res. Commun.* 521, 368–374. doi:10.1016/j.bbrc.2019.10.135

Sun, H., Ke, C., Zhang, L., Tian, C., Zhang, Z., and Wu, S. (2020). Long Non-coding RNA (LncRNA)-ATB Promotes Inflammation, Cell Apoptosis and Senescence in Transforming Growth Factor-Beta1 (TGF-Beta1) Induced Human Kidney 2 (HK-2) Cells via TGFbeta/SMAD2/3 Signaling Pathway. *Med. Sci. Monit.* 26, e922029. doi:10.12659/msm.922029

Takeda, K., and Akira, S. (2004). TLR Signaling Pathways. *Semin. Immunol.* 16, 3–9. doi:10.1016/j.smim.2003.10.003

Tamang, S., Acharya, V., Roy, D., Sharma, R., Aryaa, A., Sharma, U., et al. (2019). SNHG12: An LncRNA as a Potential Therapeutic Target and Biomarker for Human Cancer. *Front. Oncol.* 9, 901. doi:10.3389/fonc.2019.00901

Vickers, A. J., Cronin, A. M., Elkin, E. B., and Gonen, M. (2008). Extensions to Decision Curve Analysis, a Novel Method for Evaluating Diagnostic Tests, Prediction Models and Molecular Markers. *BMC Med. Inform. Decis. Mak* 8, 53. doi:10.1186/1472-6947-8-53

Wang, L., Sebra, R. P., Sfakianos, J. P., Allette, K., Wang, W., Yoo, S., et al. (2020). A Reference Profile-free Deconvolution Method to Infer Cancer Cell-Intrinsic Subtypes and Tumor-type-specific Stromal Profiles. *Genome Med.* 12, 24. doi:10.1186/s13073-020-0720-0

Xu, J., Zhang, P., Sun, H., and Liu, Y. (2020). LINC01094/miR-577 axis Regulates the Progression of Ovarian Cancer. *J. Ovarian Res.* 13, 122. doi:10.1186/s13048-020-00721-9

Yi, M., Nissley, D. V., McCormick, F., and Stephens, R. M. (2020). ssGSEA Score-Based Ras Dependency Indexes Derived from Gene Expression Data Reveal Potential Ras Addiction Mechanisms with Possible Clinical Implications. *Sci. Rep.* 10, 10258. doi:10.1038/s41598-020-66986-8

Zemek, R. M., De Jong, E., Chin, W. L., Schuster, I. S., Fear, V. S., Casey, T. H., et al. (2019). Sensitization to Immune Checkpoint Blockade through Activation of a STAT1/NK axis in the Tumor Microenvironment. *Sci. Transl Med.* 11. doi:10.1126/scitranslmed.aav7816

Zhang, Y., Yan, J., Li, C., Wang, X., Dong, Y., Shen, X., et al. (2019). LncRNA H19 Induced by helicobacter Pylori Infection Promotes Gastric Cancer Cell Growth via Enhancing NF-Kb-Induced Inflammation. *J. Inflamm.* 16, 23. doi:10.1186/s12950-019-0226-y

Zhou, X., Han, X., Wittfeldt, A., Sun, J., Liu, C., Wang, X., et al. (2016). Long Non-Coding Rna Anril Regulates Inflammatory Responses As A Novel Component Of Nf-Kb Pathway. *RNA Biol.* 13, 98–108. doi:10.1080/15476286.2015.1122164

Zhu, B., Liu, W., Liu, H., Xu, Q., and Xu, W. (2020). LINC01094 Down-Regulates miR-330-3p and Enhances the Expression of MSI1 to Promote the Progression of Glioma. *Cmar* 12, 6511–6521. doi:10.2147/cmar.s254630

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 Genetics

ORIGINAL RESEARCH
published: 04 January 2022
doi: 10.3389/fgene.2021.732822

Comprehensive Analysis to Identify *SPP1* as a Prognostic Biomarker in Cervical Cancer

*Kaidi Zhao, Zhou Ma and Wei Zhang**

Department of Obstetrics and Gynecology, Zhongnan Hospital of Wuhan University, Wuhan, China

Background: *SPP1*, secreted phosphoprotein 1, is a member of the small integrin-binding ligand N-linked glycoprotein (SIBLING) family. Previous studies have proven *SPP1* overexpressed in a variety of cancers and can be identified as a prognostic factor, while no study has explored the function and carcinogenic mechanism of *SPP1* in cervical cancer.

Methods: We aimed to demonstrate the relationship between *SPP1* expression and pan-cancer using The Cancer Genome Atlas (TCGA) database. Next, we validated *SPP1* expression of cervical cancer in the Gene Expression Omnibus (GEO) database, including GSE7803, GSE63514, and GSE9750. The receiver operating characteristic (ROC) curve was used to evaluate the feasibility of *SPP1* as a differentiating factor by the area under curve (AUC) score. Cox regression and logistic regression were performed to evaluate factors associated with prognosis. The *SPP1*-binding protein network was built by the STRING tool. Enrichment analysis by the R package clusterProfiler was used to explore potential function of *SPP1*. The single-sample GSEA (ssGSEA) method from the R package GSVA and TIMER database were used to investigate the association between the immune infiltration level and *SPP1* expression in cervical cancer.

Results: Pan-cancer data analysis showed that *SPP1* expression was higher in most cancer types, including cervical cancer, and we got the same result in the GEO database. The ROC curve suggested that *SPP1* could be a potential diagnostic biomarker (AUC = 0.877). High *SPP1* expression was associated with poorer overall survival (OS) ($P = 0.032$). Further enrichment and immune infiltration analysis revealed that high *SPP1* expression was correlated with regulating the infiltration level of neutrophil cells and some immune cell types, including macrophage and DC.

Conclusion: *SPP1* expression was higher in cervical cancer tissues than in normal cervical epithelial tissues. It was significantly associated with poor prognosis and immune cell infiltration. Thus, *SPP1* may become a promising prognostic biomarker for cervical cancer patients.

Keywords: SPP1, biomarker, cervical cancer, prognosis, immune infiltration

OPEN ACCESS

Edited by:
Jian-Bing Fan,
Illumina, United States

Reviewed by:
Lucia Tata-Chayeb,
National Institute of Cancerology (INCAN), Mexico
Shanqiang Qu,
Southern Medical University, China

****Correspondence:***
Wei Zhang
zn002646@whu.edu.cn

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: *29 June 2021*
Accepted: *03 December 2021*
Published: *04 January 2022*

Citation:
Zhao K, Ma Z and Zhang W (2022) Comprehensive Analysis to Identify SPP1 *as a Prognostic Biomarker in Cervical Cancer. Front. Genet. 12:732822. doi: 10.3389/fgene.2021.732822*

1 INTRODUCTION

Cervical cancer remains the fourth most common cancer among women and accounts for 527,624 new diagnosed cases and 265,672 deaths in 2018 (Bray et al. (2018)). Cervical cancer continues to be the first or second leading cause of cancer-related death among women for many low- and middle-income countries (LMICs) (Wang et al. (2018)). Persistent HPV infection, especially types 16 and 18, is a high-risk factor but not the only one for cervical cancer (Revathidevi et al. (2020)). Host genetic factors may also be involved in tumor development. The major treatments for cervical cancer patients include surgery, chemotherapy, and radiotherapy. For patients with early-stage cervical cancer, 5-year survival is up to 91.5%, while the treatment of advanced cervical cancer is not ideal (Luan and Wang (2018)). The median survival time of metastatic cervical cancer patients is about 8–13 months, and the 5-year overall

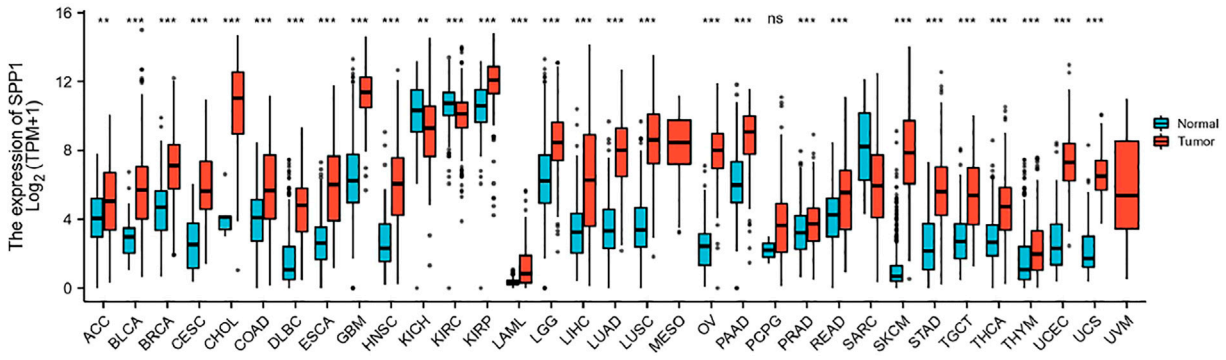


FIGURE 1 | *SPP1* expression in normal and tumor tissues in TCGA and GTEx databases.

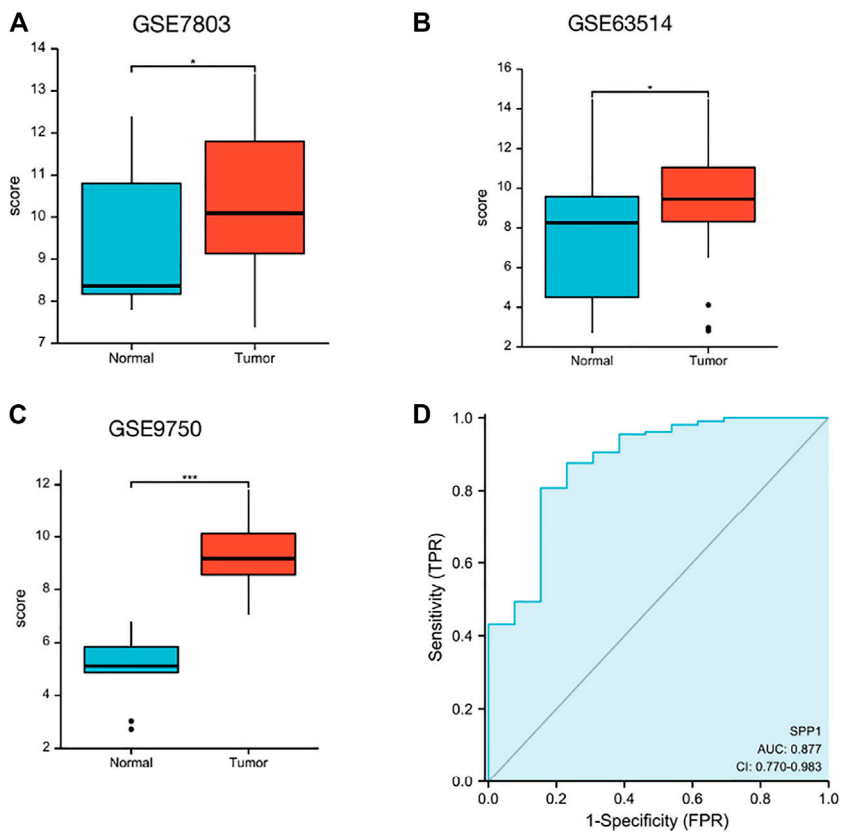


FIGURE 2 | *SPP1* expression in the GEO database. **(A)** *SPP1* expression in normal and tumor tissues in cervical cancer from GSE7803. **(B)** *SPP1* expression in normal cervical epithelial and cervical cancer tissues from GSE63514. **(C)** *SPP1* expression in normal cervical tissues and cervical cancer epithelial component from GSE9750. **(D)** ROC curve of *SPP1* in cervical cancer. *X*-axis represents false-positive rates, and *Y*-axis represents true-positive rates.

TABLE 1 | Correlation analyzed between *SPP1* expression and clinicopathologic characteristics in cervical cancer based on TCGA database.

Characteristic	Low expression of SPP1	High expression of SPP1	p value
N	153	153	
T stage, n (%)			0.020
T1	82 (33.7%)	58 (23.9%)	
T2	31 (12.8%)	41 (16.9%)	
T3	6 (2.5%)	15 (6.2%)	
T4	4 (1.6%)	6 (2.5%)	
N stage, n (%)			0.243
N0	73 (37.4%)	61 (31.3%)	
N1	27 (13.8%)	34 (17.4%)	
M stage, n (%)			0.699
M0	55 (43.3%)	61 (48%)	
M1	4 (3.1%)	7 (5.5%)	
Clinical stage, n (%)			0.020
Stage I	95 (31.8%)	67 (22.4%)	
Stage II	30 (10%)	39 (13%)	
Stage III	17 (5.7%)	29 (9.7%)	
Stage IV	9 (3%)	13 (4.3%)	
Radiation therapy, n (%)			0.726
No	63 (20.6%)	59 (19.3%)	
Yes	90 (29.4%)	94 (30.7%)	
Primary therapy outcome, n (%)			0.106
PD	7 (3.2%)	16 (7.3%)	
SD	2 (0.9%)	4 (1.8%)	
PR	4 (1.8%)	4 (1.8%)	
CR	101 (46.1%)	81 (37%)	
Race, n (%)			0.444
Asian	12 (4.6%)	8 (3.1%)	
Black or African American	13 (5%)	18 (6.9%)	
White	106 (40.6%)	104 (39.8%)	
Histologic type, n (%)			<0.001
Adenosquamous	40 (13.1%)	13 (4.2%)	
Squamous cell carcinoma	113 (36.9%)	140 (45.8%)	
Histologic grade, n (%)			0.954
G1	10 (3.6%)	9 (3.3%)	
G2	69 (25.2%)	66 (24.1%)	
G3	62 (22.6%)	57 (20.8%)	
G4	0 (0%)	1 (0.4%)	
Age (years), median (IQR)	45 (37, 54)	49 (40, 60)	0.038

survival rate is only around 16.5% (Ferlay et al. (2013); van Meir et al. (2014)). Therefore, it is urgent to find more accurate biomarkers for early detection of cervical cancer and monitoring the disease progression.

Secreted phosphoprotein 1 (*SPP1*) is a secreted multifunctional phosphoprotein located in 4q13 with seven exons and six introns. *SPP1*, also known as osteopontin-like protein or early T-lymphocyte activation 1 protein, is a member of the small integrin-binding ligand N-linked glycoprotein (SIBLING) family which can specifically bind and activate matrix metalloproteinases (MMPs) in cancer (Su et al. (2020)). Its main biological functions are involved in immune response, biomineralization, and tissue remodeling. *SPP1* is also related to the growth, proliferation, migration, apoptosis, and chemotaxis of cells. Previous studies have proven that *SPP1* is overexpressed in a variety of cancers and can be used to predict the adverse consequences, including ovarian cancer (Zeng et al. (2018)), glioblastoma (Kijewska et al. (2017)), hepatocellular carcinoma (Wang et al. (2019)), and gastric cancer (Song et al. (2019)). Recently, the relationship between the expression of *SPP1* and chemotherapy resistance, such as prostate cancer and hepatocellular carcinoma, has also attracted the attention of researchers (Liu et al. (2016); Pang et al. (2019)), while no study has explored the correlation between *SPP1* and cervical cancer. Therefore, our study aimed to explore the expression of *SPP1* in cervical cancer tissues and its potential clinical values.

In our research, we utilized the cervical cancer RNA-seq data from The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), and Genotype-Tissue Expression databases to compare the differential expression of *SPP1* between normal cervical tissues and cervical cancer samples. Next, we investigated the relationship between *SPP1* expression levels and clinical pathological features of cervical cancer. Furthermore, we explored the prognostic value of *SPP1* in cervical cancer. Besides, we performed gene enrichment analysis to reveal its potential functions. Finally, we analyzed the relationship between *SPP1* expression and immune infiltration and comprehensively

TABLE 2 | *SPP1* expression associated with clinicopathologic characteristics by logistic regression.

Characteristic	Total (N)	Odds ratio (OR)	*p* value
T stage (T2 and T3 and T4 vs. T1)	243	2.138 (1.278–3.609)	0.004
N stage (N1 vs. N0)	195	1.507 (0.821–2.786)	0.187
M stage (M1 vs. M0)	127	1.578 (0.451–6.294)	0.485
Clinical stage (Stage II and Stage III and Stage IV vs. Stage I)	299	2.051 (1.295–3.269)	0.002
Primary therapy outcome (SD and PR and CR vs. PD)	219	0.364 (0.135–0.893)	0.033
Histologic type (squamous cell carcinoma vs. adenosquamous)	306	3.812 (1.993–7.732)	<0.001
Age (>50 vs. ≤50 years)	306	1.743 (1.097–2.787)	0.019
Radiation therapy (yes vs. no)	306	1.115 (0.706–1.765)	0.641
Histologic grade (G2 and G3 and G4 vs. G1)	274	1.052 (0.411–2.731)	0.916

TABLE 3 | Univariate and multivariate Cox analyses of prognostic factors in cervical cancer.

Characteristic	Total (N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p* value	Hazard ratio (95% CI)	*p* value
T stage (T2 and T3 and T4 vs. T1)	243	1.906 (1.085–3.348)	**0.025**	1.193 (0.419–3.395)	0.741
N stage (N1 vs. N0)	195	2.844 (1.446–5.593)	**0.002**	3.117 (1.517–6.403)	**0.002**
M stage (M1 vs. M0)	127	3.555 (1.187–10.641)	**0.023**		
TP53 (high vs. low)	306	0.854 (0.537–1.356)	0.503		
Clinical stage (Stage II and Stage III and Stage IV vs. Stage I)	299	1.462 (0.920–2.324)	0.108	0.464 (0.160–1.345)	0.157
Radiation therapy (yes vs. no)	306	1.172 (0.694–1.981)	0.553		
Race (Black or African American and White vs. Asian)	261	1.537 (0.374–6.317)	0.552		
Age (>50 vs. ≤50 years)	306	1.289 (0.810–2.050)	0.284	0.658 (0.298–1.452)	0.299
Histologic type (squamous cell carcinoma vs. adenosquamous)	306	1.033 (0.543–1.969)	0.920		
Histologic grade (G2 and G3 vs. G1)	273	1.212 (0.378–3.882)	0.746		
SPP1 (high vs. low)	306	1.686 (1.046–2.719)	**0.032**	2.207 (1.019–4.777)	**0.045**

The value in bold indicates that p is less than 0.05, which is meaningful.

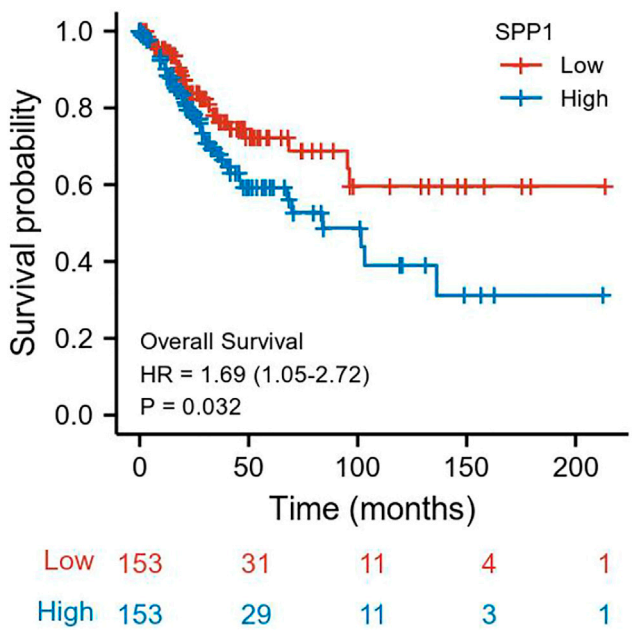


FIGURE 3 | Association between *SPP1* expression and OS in cervical cancer patients.

explored its mechanism in inducing and promoting cervical cancer.

2 MATERIALS AND METHODS

2.1 RNA Sequencing Data Collection and Analysis

To evaluate the *SPP1* expression level in pan-cancer, we downloaded data from the UCSC Xena (https://xenabrowser.net/datapages/). We selected samples from the TCGA database for the analysis of *SPP1* expression in tumor tissues, while the combined analysis of TCGA and Genotype-Tissue Expression (GTEx) databases was used for the normal tissue samples. GSE7803 (Platform: GPL96), GSE63514 (Platform: GPL570), and GSE9750 (Platform: GPL96) downloaded from GEO were used to obtain cervical cancer microarray data.

2.2 Correlation and Gene Set Enrichment Analysis

We used data collected from TCGA to perform correlation analysis between *SPP1* and other mRNAs in cervical cancer. To demonstrate the biological function of *SPP1*, we selected

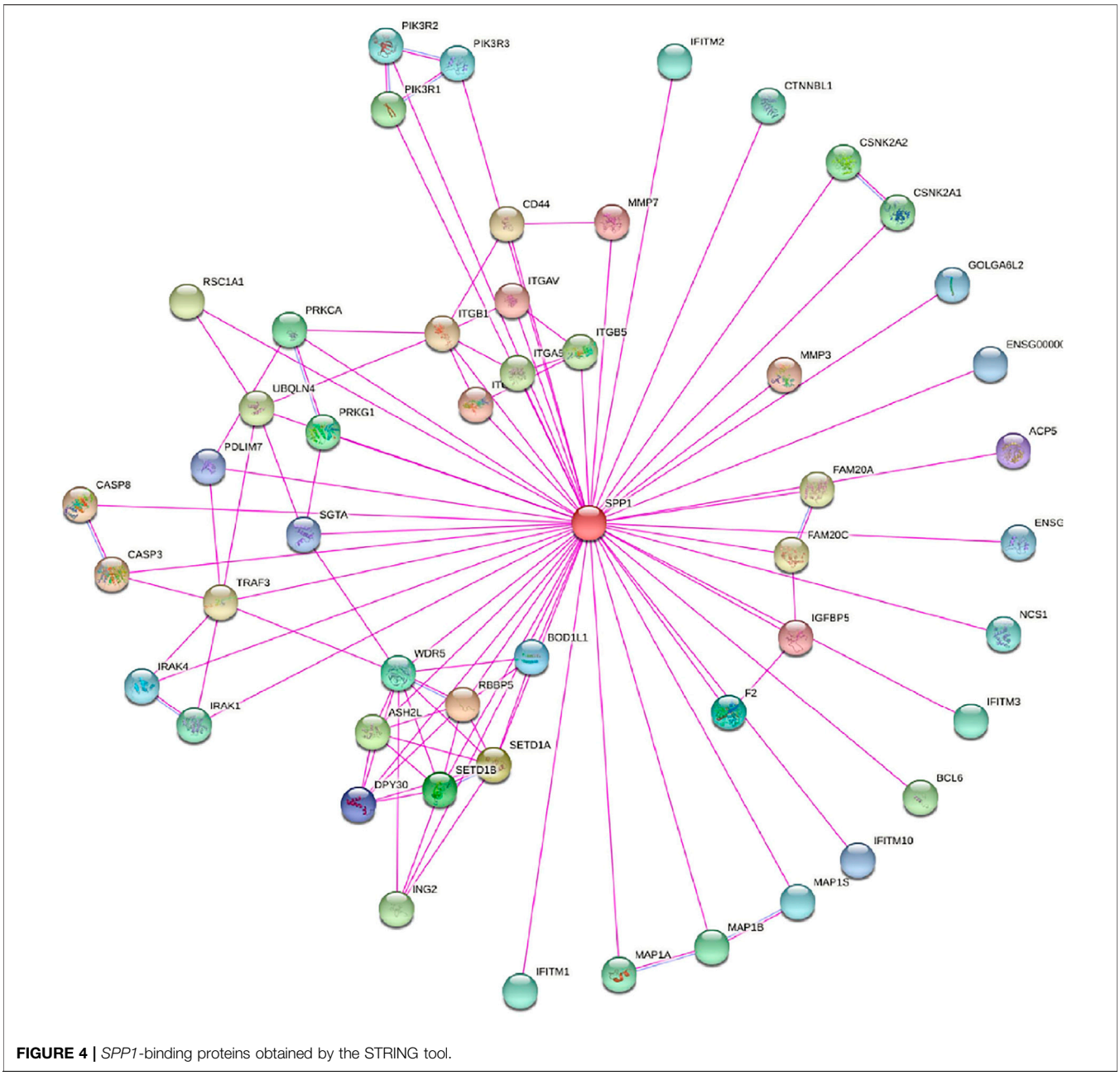


FIGURE 4 | *SPP1*-binding proteins obtained by the STRING tool.

the top 100 genes most positively correlated with *SPP1* for enrichment analysis. EnrichGO function in the R package "clusterProfiler" was used to perform gene ontology (GO) enrichment, including BP, CC, and MF. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was performed using the EnrichKEGG function of the R package "clusterProfiler."

2.3 Survival Prognosis Analysis

We used the R package "survival" (version 3.6) to obtain the overall survival (OS) survival plots of *SPP1*. Selecting the cutoff value of 50% as the dividing threshold, the cohorts were divided into high-expression and low-expression groups. To evaluate the value of *SPP1* in predicting the prognosis of cervical cancer patients, we used the R package (version 3.6.3) "ROC" for analysis and "ggplot2" for visual.

2.4 Immune Cell Infiltration Analysis

We used the single-sample GSEA (ssGSEA) method from the R package GSVA (version 3.6) and Tumor Immune Estimation Resource (TIMER) database (http://timer.cistrome.org/) to comprehensively investigate molecular characterization of tumor–immune interactions in cervical cancer. In the literature, we examined the impact of *SPP1* expression on immune cell infiltration using gene expression profiling data. To investigate the correlation between *SPP1* expression and the abundances of tumor-infiltrating immune cells, *p*-values were

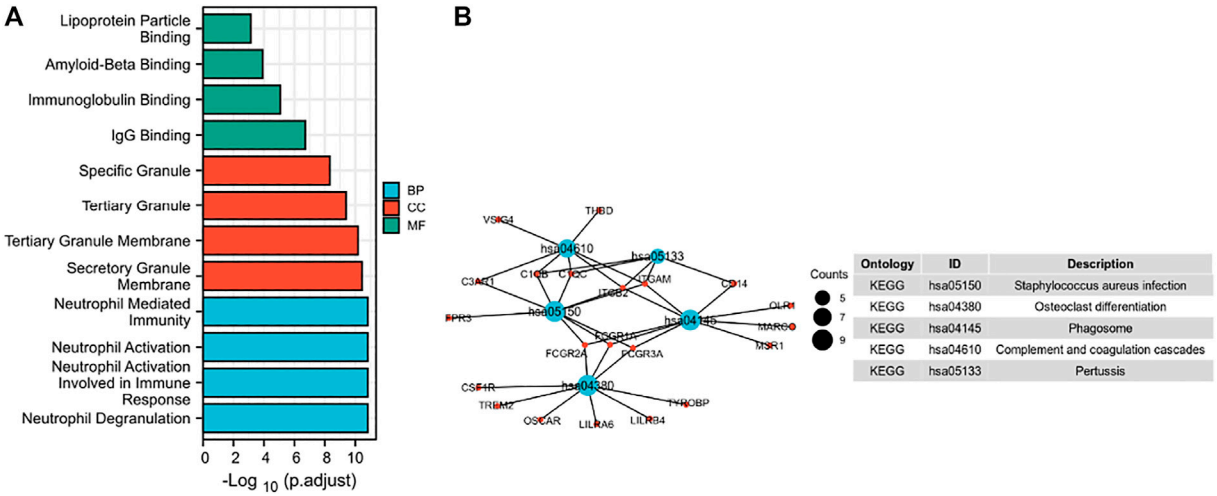


FIGURE 5 | Function and pathway enrichment analysis of *SPP1* in cervical cancer. **(A)** Significant Gene Ontology terms (including BP, MF, and CC) of the top 100 genes most positively associated with *SPP1*. **(B)** Significant KEGG pathway of the top 100 genes most positively associated with *SPP1*.

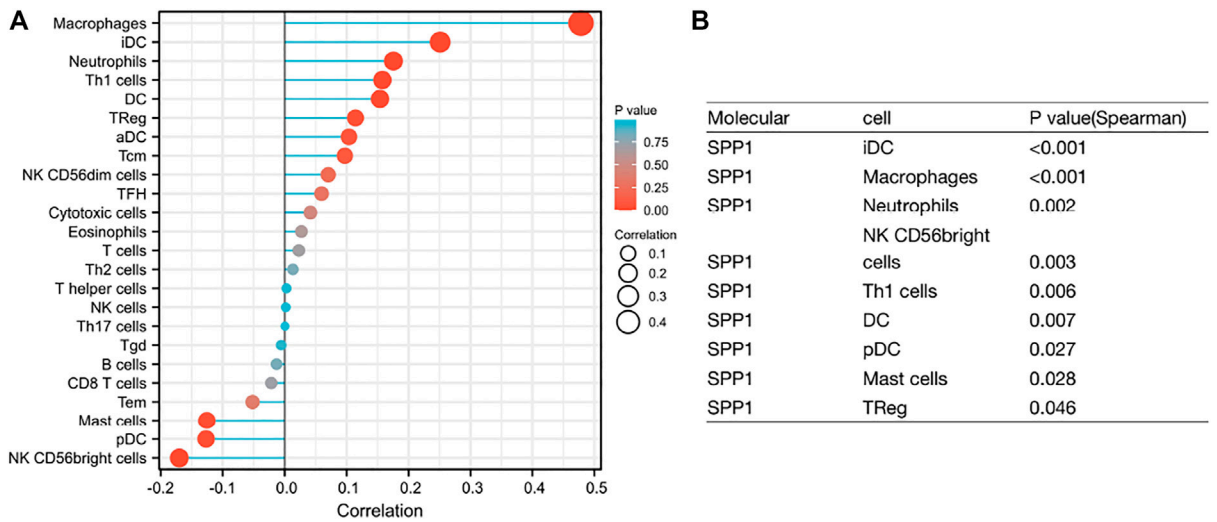


FIGURE 6 | (A) Lollipop chart of *SPP1* expression level in 24 immune cells. **(B)** The immune cell infiltration associated with *SPP1* expression, $P < 0.05$, represents a significant result.

calculated using the Wilcoxon rank-sum and Spearman's rank correlation tests.

3 RESULTS

3.1 The mRNA Expression Analysis of *SPP1* in Pan-Cancer

Data downloaded from TCGA and GTEx were used to analyze *SPP1* expression in 33 types of cancer. The result revealed that *SPP1* was overexpressed in most cancers, including ACC, BLCA, BRCA, CESC, CHOL, COAD, DLBC, ESCA, GBM, HNSC, KIRP, LAML, LGG, LIHC, LUAD, LUSC, OV, PAAD, PRAD, READ, SKCM, STAD, TGCT, THCA, THYM, UCEC, and UCS. However, the expression of *SPP1* was low in KICH and KIRC (**Figure 1**). Furthermore, we assessed *SPP1* expression in cervical cancer in the GEO database, including GSE7803 (Platform: GPL96), GSE63514 (Platform: GPL570), and GSE9750, and the results confirmed that *SPP1* was overexpressed in cervical cancer tissues (**Figures 2A–C**). Additionally, we performed the receiver operating characteristic (ROC) curve to evaluate the feasibility of the *SPP1* expression level to distinguish cervical cancer tissues from normal cervical tissues. The

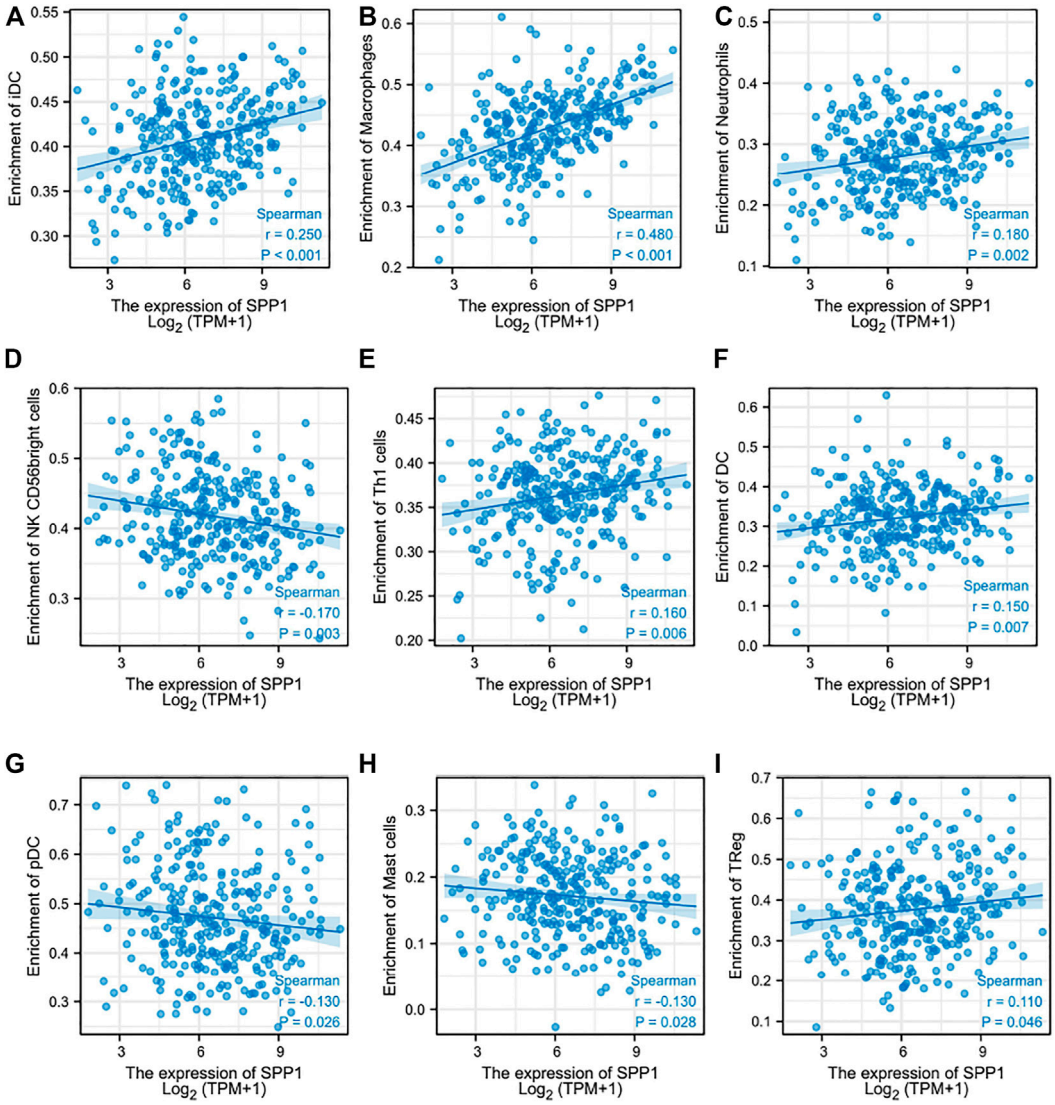


FIGURE 7 | Correlation between *SPP1* expression and immune cell infiltration. **(A–I)** Correlation between *SPP1* expression and iDC, macrophages, neutrophils, NK CD56 bright cells, Th1 cells, DC, pDC, mast cells, and Treg cells.

area under the ROC curve (AUC) was 0.877, representing the quality of the test.

3.2 Clinical Relevance of the *SPP1* Expression in Cervical Cancer Patients

The characteristics of 306 primary cervical cancer patients with both clinical and gene expression data were downloaded from TCGA database. With the cutoff value of 50% as the dividing threshold, the patients were divided into a high–*SPP1* expression group ($n = 153$) and a low–*SPP1* expression group ($n = 153$). The correlation of the *SPP1* expression level and patients' clinicopathologic characteristics was explored. We found that *SPP1* expression was significantly associated with T stage ($P = 0.02$), clinical stage ($P = 0.02$), and histologic type ($P < 0.001$) by using the chi-square test or Fisher's exact test. The Wilcoxon rank-sum test revealed that *SPP1* expression was associated with age ($P = 0.038$) (**Table 1**).

We conducted the logistic regression method to further analyze the relationship between the *SPP1* expression level and the clinicopathologic characteristics of cervical cancer. The results showed that the expression level of *SPP1* was significantly associated with T stage ($P = 0.004$), clinical stage ($P = 0.002$), primary therapy outcome ($P = 0.033$), histologic type ($P < 0.001$), and age ($P = 0.019$) (**Table 2**).

Association Between *SPP1* Expression and Cancer Patient Survival Prognosis

We performed univariate and multivariate Cox analyses of overall survival (OS) in cervical cancer patients, and results are shown in **Table 3**. In univariate Cox analysis of *SPP1*, T stage

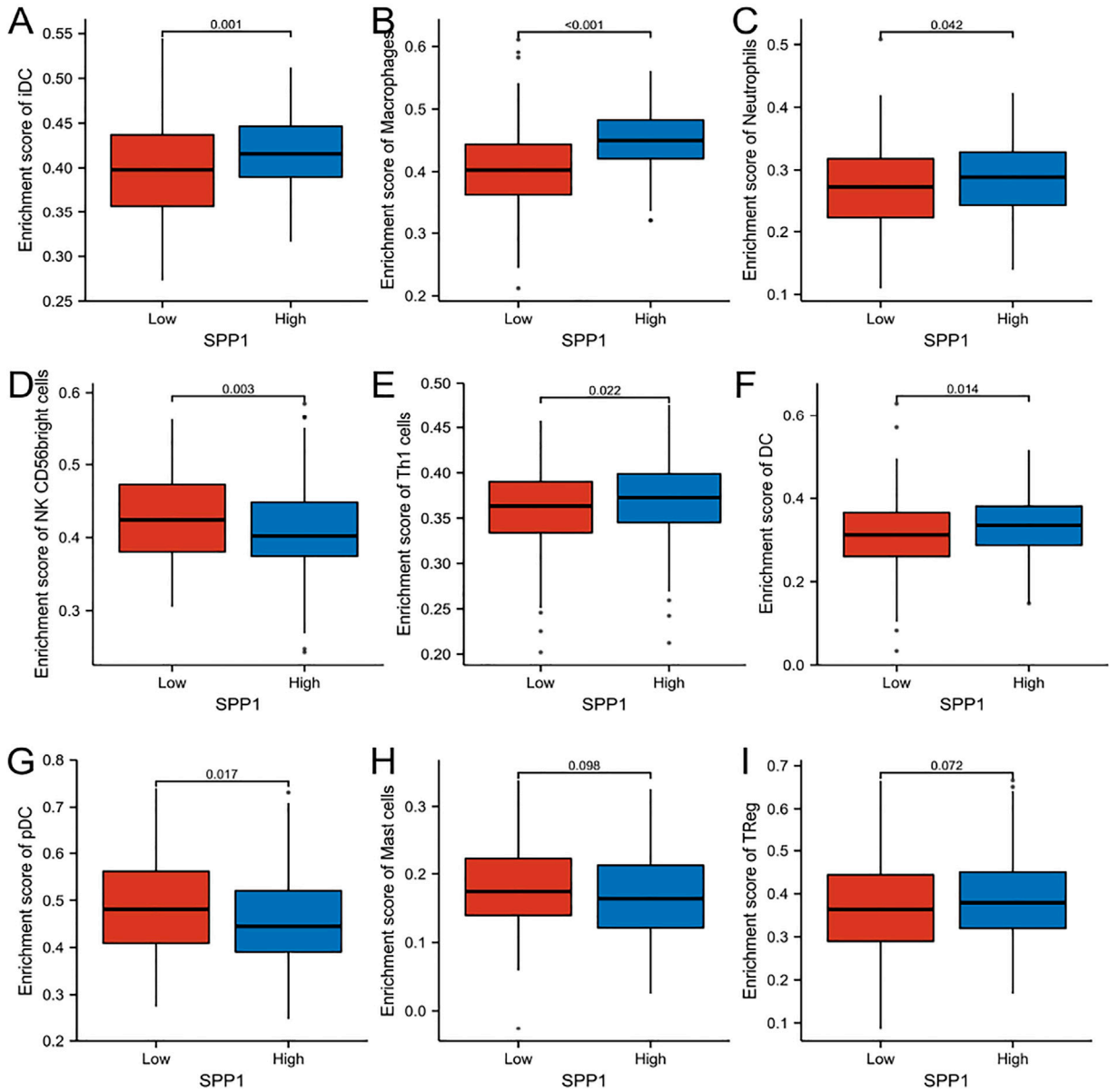


FIGURE 8 | Comparison of immune cells between high– and low–*SPP1* expression groups. **(A–I)** Histogram showing the difference of iDC, macrophages, neutrophils, NK CD56 bright cells, Th1 cells, DC, pDC, mast cells, and Treg cell infiltration level between high–and low–*SPP1* expression groups.

($P = 0.025$), N stage ($P = 0.002$), M stage ($P = 0.023$), and *SPP1* expression ($P = 0.032$) were associated with overall survival (OS) in cervical cancer patients. In the multivariate Cox model, we found that N stage ($P = 0.002$) and *SPP1* expression ($P = 0.045$) were still relevant to worse prognosis. Furthermore, we investigated the relationship between *SPP1* expression and overall survival (OS) of cervical cancer patients. According to the KM plot, patients with higher *SPP1* mRNA expression showed poorer prognosis than the lower group (HR = 1.69, 95% CI: 1.05–2.72, $P = 0.032$) (**Figure 3**). Thus, *SPP1* may become a promising prognostic biomarker for cervical cancer patients.

3.4 Correlation and *SPP1*-Related Gene Enrichment Analysis

In this study, we only considered physically binding protein interactions and obtained 50 experimental supported *SPP1*-binding proteins from the STRING network (**Figure 4**). We downloaded data from TCGA database to further investigate

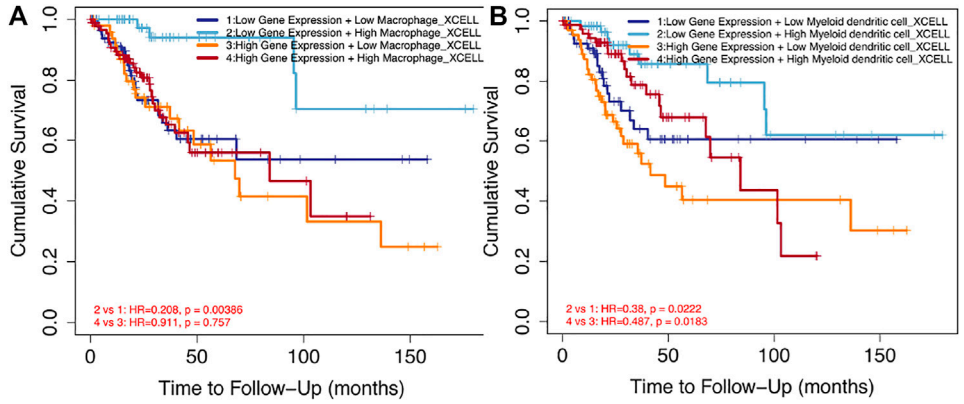


FIGURE 9 | Impact of immune cell infiltration on prognosis in cervical cancer patients. **(A)** Clinical survival outcome of cervical cancer patients in the high-macrophage group. **(B)** Clinical survival outcome of cervical cancer patients in the high–DC cell group.

the function of *SPP1* and search *SPP1* expression–correlated genes for related pathway analysis. We obtained the top 100 most positively correlated genes with *SPP1* for GO and KEGG enrichment analysis by the "clusterProfile" R package. The GO analysis data showed that most of the genes were associated with neutrophil degranulation, neutrophil activation involved in immune response, neutrophil activation, and neutrophil-mediated immunity (**Figure 5A**). The KEGG data suggested that the "phagosome" may be related to the carcinogenic mechanism of *SPP1* (**Figure 5B**).

3.5 Relationship Between *SPP1* Expression and Immune Cell Infiltration

Through the previous enrichment analysis, we found that *SPP1* was mainly related to neutrophils and phagosomes. We hypothesized that there might be some relationship between *SPP1* and immune cells. Thus, we further assessed whether the *SPP1* expression level was associated with immune cell infiltration. We used ssGSEA from the R package with Spearman's r to investigate the potential association between the *SPP1* expression level and 24 types of immune cells. The result revealed that *SPP1* expression had significant correlation with iDC, macrophages, neutrophils, NK CD56 bright cells, Th1 cells, DC, pDC, mast cells, and Treg cells (**Figure 6**). Further research showed that *SPP1* expression was positively correlated with infiltration levels of iDC (**Figure 7A**) ($r = 0.250$, $P < 0.001$), macrophages (**Figure 7B**) ($r = 0.480$, $P < 0.001$), neutrophils (**Figure 7C**) ($r = 0.180$, $P = 0.002$), Th1 cells (**Figure 7E**) ($r = 0.160$, $P = 0.006$), DC (**Figure 7F**) ($r = 0.150$, $P = 0.007$), and Treg cells (**Figure 7I**) ($r = 0.110$, $P = 0.046$). In contrast, *SPP1* expression was negatively correlated with that of NK CD56 bright cells (**Figure 7D**) ($r = -0.170$, $P = 0.003$), pDC (**Figure 7G**) ($r = -0.130$, $P = 0.026$) and mast cells (**Figure 7H**) ($r = -0.130$, $P = 0.028$). This prompted us to examine the relationship between the *SPP1* expression level and immune infiltration. Surprisingly, we found significant differences in infiltrating immune cell levels, including iDC, macrophages, neutrophils, NK CD56 bright cells, Th1 cells, DC, and pDC ($P < 0.05$), when *SPP1* expression was categorized into high and low groups (**Figures 8A–G**), while no significant difference in mast cells and Treg cells was noted (**Figures 8H,I**). Finally, we assessed the impact of immune cell infiltration on clinical survival outcome of cervical cancer patients by TIMER (http://timer.cistrome.org/). We found that high levels of macrophages and DC cells were associated with poor prognosis of cervical cancer patients ($P < 0.05$) (**Figures 9A,B**).

4 DISCUSSION

Invasive cervical cancer remains the leading cause of cancer death among women worldwide (Shen et al. (2020)). Thus, it is necessary to find more accurate biomarkers to detect at an early stage and monitor disease progression. According to the previous studies, *SPP1* is overexpressed in various cancer types (Xu et al. (2017); Choe et al. (2018); Zhang et al. (2020)) and identified as a prognostic factor (Li et al. (2018); Chen J et al. (2019); Guo et al. (2020)), while to our knowledge, no study has explored the relationship of *SPP1* expression and cervical cancer. In our study, we attempted to explore the potential mechanism of *SPP1* in promoting cervical cancer and its feasibility as a molecular biomarker.

In pan-cancer analysis, we found that *SPP1* was upregulated in most cancer types. Further exploration revealed that higher *SPP1* expression was associated with reduced overall survival (OS) in cervical cancer patients. We performed logistic regression to evaluate the relationship between the *SPP1* expression level and the clinicopathologic characteristics of cervical cancer. The result showed that *SPP1* was significantly correlated with clinical stages. In addition, univariate and multivariate Cox analyses indicated that *SPP1* was an independent factor to predict prognosis of patients. All these aforementioned results and ROC analysis suggest that *SPP1* may be a promising prognostic biomarker for cervical cancer patients.

The tumor microenvironment (TME), composed of various types of immune cells, played an important role in tumor progression, metastasis, and treatment resistance (Usui et al. (2016)). The composition of tumor-infiltrating immune cells strongly influenced the tumor microenvironment and the

behavior of the tumor. Our gene enrichment analysis revealed that the main biological function of *SPP1* was mainly involved in immune response. We next confirmed that *SPP1* expression correlated with immune cell infiltration. Hence, we hypothesized that *SPP1* may affect the tumor microenvironment by changing proportions of specific immune cell types, thereby promoting tumor progression and metastasis. It was, indeed, the case that *SPP1* had recently been shown to be an important component in maintaining the tumor microenvironment in AML (Ruvolo et al. (2019)). Our research demonstrated the significant positive correlation between macrophages and the expression of *SPP1*. Macrophages are important components of the tumor microenvironment, and tumor-associated macrophages play complex roles in cancer pathophysiology (Gibson et al. (2019)). A previous study found that *SPP1* was involved in the function, migration, and differentiation of macrophages (Zhang et al. (2017); Wei et al. (2019); Jaitin et al. (2019); Srirussamee et al. (2019)). A recent study also showed that *SPP1* was essential for M2-like macrophage, the tumor-associated macrophage, and promoted tumor growth (Chen P et al. (2019)). Furthermore, we found that the increased level of macrophages and DC infiltration were correlated with poor prognosis. Our results were supported by the findings of similar studies about this topic (Long et al. (2016); Ndiaye et al. (2019)). Certainly, the tumor microenvironment had a high level of complexity in its regulation; other immune cell types in the tumor microenvironment may also influence tumor cell survival, including iDC, neutrophils, NK CD56 bright cells, Th1 cells, DC, and pDC. Future studies were needed to further explore the relationship between *SPP1* expression and these cells.

In conclusion, we demonstrated that *SPP1* expression was upregulated in cervical cancer and significantly related to poor survival outcome. In addition to this, *SPP1* might participate in the occurrence and development of cervical cancer by influencing the infiltration level of immune cells. Therefore, our study revealed the role of *SPP1* in cervical cancer and identified a promising prognostic biomarker.

Although our study is the first work to explore the relationship between *SPP1* expression and cervical cancer, it also has some limitations. First, all of the data analyzed by bioinformatics methods in this study were downloaded directly from public databases, so it requires further validation by experimental investigations; second, the number of normal samples used as controls was considerably different from that of patients with tumor in the TCGA database; therefore, further studies based on an equal balance of sample size are necessary. Third, further validation studies with a long-term follow-up and larger cohorts of patients are needed to definitely validate *SPP1* as an OS predictor. Last but not least, our study laid the foundation for detailed studies of the correlation between *SPP1* and the tumor-associated immune microenvironment. However, more studies are required to explore the hypothesis in depth.

STATEMENT

The cervical cancer cell lines (Siha and Hela) present in this study were obtained from the Scientific Research Center of Zhongnan Hospital of Wuhan University. And normal cervical epithelial cell (END1) was donated by Wuhan University Basic Medical College.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found freely from TCGA data portal (https://portal.gdc.cancer.gov/) and GEO database (https://www.ncbi.nlm.nih.gov/geo/).

AUTHOR CONTRIBUTIONS

KZ and WZ contributed to the study conception and design. Material preparation, data collection, and analysis were performed by KZ and ZM. KZ contributed to the literature search. The first draft of the manuscript was written by KZ, and all authors commented on previous versions of the manuscript. WZ reviewed the article and gave suggestions on the revision of the article. All authors read and approved the final manuscript.

FUNDING

Our research was supported by the project of improving the ability of diagnosis and treatment of difficult diseases in Zhongnan Hospital of Wuhan University. The project number is ZLYNXM202019.

REFERENCES

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: Globocan Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* 68, 394–424. doi:10.3322/caac.21492

Chen J, J., Hou, C., Zheng, Z., Lin, H., Lv, G., and Zhou, D. (2019). Identification of Secreted Phosphoprotein 1 (Spp1) as a Prognostic Factor in Lower-Grade Gliomas. *World Neurosurg.* 130, e775. doi:10.1016/j.wneu.2019.06.219

Chen P, P., Zhao, D., Li, J., Liang, X., Li, J., Chang, A., et al. (2019). Symbiotic Macrophage-Glioma Cell Interactions Reveal Synthetic Lethality in Pten-Null Glioma. *Cancer cell* 35, 868–884. doi:10.1016/j.ccell.2019.05.003

Choe, E. K., Yi, J. W., Chai, Y. J., and Park, K. J. (2018). Upregulation of the Adipokine Genes Adipor1 and Spp1 Is Related to Poor Survival Outcomes in Colorectal Cancer. *J. Surg. Oncol.* 117, 1833–1840. doi:10.1002/jso.25078

Ferlay, J., Steliarova-Foucher, E., Lortet-Tieulent, J., Rosso, S., Coebergh, J. W. W., Comber, H., et al. (2013). Cancer Incidence and Mortality Patterns in Europe: Estimates for 40 Countries in 2012. *Eur. J. Cancer* 49 (6), 1374–1403. doi:10.1016/j.ejca.2012.12.027

Gibson, E. M., Nagaraja, S., Ocampo, A., Tam, L. T., Wood, L. S., Pallegar, P. N., et al. (2019). Methotrexate Chemotherapy Induces Persistent Tri-glial Dysregulation that Underlies Chemotherapy-Related Cognitive Impairment. *Cell* 176, 43–55. doi:10.1016/j.cell.2018.10.049

Guo, Z., Huang, J., Wang, Y., Liu, X. P., Li, W., Yao, J., et al. (2020). Analysis of Expression and its Clinical Significance of the Secreted Phosphoprotein 1 in Lung Adenocarcinoma. *Front. Genet.* 11, 547. doi:10.3389/fgene.2020.00547

Jaitin, D. A., Adlung, L., Thaiss, C. A., Weiner, A., Li, B., Descamps, H., et al. (2019). Lipid-associated Macrophages Control Metabolic Homeostasis in a Trem2-dependent Manner. *Cell* 178, 686–698. doi:10.1016/j.cell.2019.05.054

Kijewska, M., Kocyk, M., Kloss, M., Stepniak, K., Korwek, Z., Polakowska, R., et al. (2017). The Embryonic Type of Spp1 Transcriptional Regulation Is Re-activated in Glioblastoma. *Oncotarget* 8, 16340–16355. doi:10.18632/oncotarget.14092

Li, S., Yang, R., Sun, X., Miao, S., Lu, T., Wang, Y., et al. (2018). Identification of Spp1 as a Promising Biomarker to Predict Clinical Outcome of Lung Adenocarcinoma Individuals. *Gene* 679, 398–404. doi:10.1016/j.gene.2018.09.030

Liu, G., Fan, X., Tang, M., Chen, R., Wang, H., Jia, R., et al. (2016). Osteopontin Induces Autophagy to Promote Chemo-Resistance in Human Hepatocellular Carcinoma Cells. *Cancer Lett.* 383 (2), 171–182. doi:10.1016/j.canlet.2016.09.033

Long, K. B., Gladney, W. L., Tooker, G. M., Graham, K., Fraietta, J. A., and Beatty, G. L. (2016). IFNγ and CCL2 Cooperate to Redirect Tumor-Infiltrating Monocytes to Degrade Fibrosis and Enhance Chemotherapy Efficacy in Pancreatic Carcinoma. *Cancer Discov.* 6 (4), 400–413. doi:10.1158/2159-8290.cd-15-1032

Luan, X., and Wang, Y. (2018). Lncrna Xloc_006390 Facilitates Cervical Cancer Tumorigenesis and Metastasis as a Cerna against Mir-331-3p and Mir-338-3p. *J. Gynecol. Oncol.* 29, e95. doi:10.3802/jgo.2018.29.e95

Ndiaye, P. D., Dufies, M., Giuliano, S., Douguet, L., Grépin, R., Durivault, J., et al. (2019). Vegfc Acts as a Double-Edged Sword in Renal Cell Carcinoma Aggressiveness. *Theranostics* 9, 661–675. doi:10.7150/thno.27794

Pang, X., Xie, R., Zhang, Z., Liu, Q., Wu, S., and Cui, Y. (2019). Identification of Spp1 as an Extracellular Matrix Signature for Metastatic Castration-Resistant Prostate Cancer. *Front. Oncol.* 9, 924. doi:10.3389/fonc.2019.00924

Revathidevi, S., Murugan, A. K., Nakaoka, H., Inoue, I., and Munirajan, A. K. (2020). Apobec: A Molecular Driver in Cervical Cancer Pathogenesis. *Cancer Lett.* 496, 104–116. doi:10.1016/j.canlet.2020.10.004

Ruvolo, P. P., Hu, C. W., Qiu, Y., Ruvolo, V. R., Go, R. L., Hubner, S. E., et al. (2019). Lgals3 Is Connected to Cd74 in a Previously Unknown Protein Network that Is Associated with Poor Survival in Patients with Aml. *EBioMedicine* 44, 126–137. doi:10.1016/j.ebiom.2019.05.025

Shen, S., Zhang, S., Liu, P., Wang, J., and Du, H. (2020). Potential Role of Micrornas in the Treatment and Diagnosis of Cervical Cancer. *Cancer Genet.* 248-249, 25–30. doi:10.1016/j.cancergen.2020.09.003

Song, S. Z., Lin, S., Liu, J. N., Zhang, M. B., Du, Y. T., Zhang, D. D., et al. (2019). Retracted : Targeting of SPP1 by microRNA-340 Inhibits Gastric Cancer Cell Epithelial-Mesenchymal Transition through Inhibition of the PI3K/AKT Signaling Pathway. *J. Cel Physiol.* 234, 18587–18601. doi:10.1002/jcp.28497

Srirussamee, K., Mobini, S., Cassidy, N. J., and Cartmell, S. H. (2019). Direct Electrical Stimulation Enhances Osteogenesis by Inducing Bmp2 and Spp1 Expressions from Macrophages and Preosteoblasts. *Biotechnol. Bioeng.* 116, 3421–3432. doi:10.1002/bit.27142

Su, X., Xu, B., Zhou, D.-L., Ye, Z., He, H., Yang, X. H., et al. (2020). Polymorphisms in Matricellular Spp1 and Sparc Contribute to Susceptibility to Papillary Thyroid Cancer. *Genomics* 112, 4959. doi:10.1016/j.ygeno.2020.09.018

Usui, T., Sakurai, M., Enjoji, S., Kawasaki, H., Umata, K., Ohama, T., et al. (2016). Establishment of a Novel Model for Anticancer Drug Resistance in Three-Dimensional Primary Culture of Tumor Microenvironment. *Stem Cell Int.* 2016, 7053872. doi:10.1155/2016/7053872

van Meir, H., Kenter, G., Burggraaf, J., Kroep, J., Welters, M., Melief, C., et al. (2014). The Need for Improvement of the Treatment of Advanced and Metastatic Cervical Cancer, the Rationale for Combined Chemo-Immunotherapy. *Anticancer Agents Med. Chem.* 14 (2), 190–203. doi:10.2174/18715206113136660372

Wang, L., Zhao, Y., Wang, Y., and Wu, X. (2018). The Role of Galectins in Cervical Cancer Biology and Progression. *Biomed. Res. Int.* 2018, 2175927. doi:10.1155/2018/2175927

Wang, J., Hao, F., Fei, X., and Chen, Y. (2019). SPP1 Functions as an Enhancer of Cell Growth in Hepatocellular Carcinoma Targeted by Mir-181c. *Am. J. Transl. Res.* 11 (11), 6924–6937.

Wei, J., Marisetty, A., Schrand, B., Gabrusiewicz, K., Hashimoto, Y., Ott, M., et al. (2019). Osteopontin Mediates Glioblastoma-Associated Macrophage Infiltration and Is a Potential Therapeutic Target. *J. Clin. Invest.* 129, 137–149. doi:10.1172/JCI121266

Xu, C., Sun, L., Jiang, C., Zhou, H., Gu, L., Liu, Y., et al. (2017). Spp1, Analyzed by Bioinformatics Methods, Promotes the Metastasis in Colorectal Cancer by Activating Emt Pathway. *Biomed. Pharmacother.* 91, 1167–1177. doi:10.1016/j.biopha.2017.05.056

Zeng, B., Zhou, M., Wu, H., and Xiong, Z. (2018). Spp1 Promotes Ovarian Cancer Progression via Integrin β1/fak/akt Signaling Pathway. *Onco. Targets Ther.* 11, 1333–1343. doi:10.2147/ott.s154215

Zhang, Y., Du, W., Chen, Z., and Xiang, C. (2017). Upregulation of Pd-L1 by Spp1 Mediates Macrophage Polarization and Facilitates Immune Escape in Lung Adenocarcinoma. *Exp. Cel Res.* 359, 449–457. doi:10.1016/j.yexcr.2017.08.028

Zhang, Q., Li, L., Lai, Y., and Zhao, T. (2020). Silencing of Spp1 Suppresses Progression of Tongue Cancer by Mediating the Pi3k/akt Signaling Pathway. *Technol. Cancer Res. Treat.* 19, 1533033820971306. doi:10.1177/1533033820971306

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.

GLOSSARY

aDC activated DC
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
GEO Gene Expression Omnibus
GO Gene Ontology
HNSC head and neck squamous cell carcinoma
iDC immature DC
KICH kidney chromophobe
KIRC kidney renal clear cell carcinoma
KIRP kidney renal papillary cell carcinoma
KEGG Kyoto Encyclopedia of Genes and Genomes
LAML acute myeloid leukemia
LGG lower grade glioma
LIHC liver hepatocellular carcinoma
LUAD lung adenocarcinoma
LUSC lung squamous cell carcinoma
OS overall survival
OV ovarian serous cystadenocarcinoma
pDC plasmacytoid DC
PAAD pancreatic adenocarcinoma
PRAD prostate adenocarcinoma
READ rectum adenocarcinoma
SKCM skin cutaneous melanoma
STAD stomach adenocarcinoma
SPP1 secreted phosphoprotein 1
Tcm T central memory
Tem T effector memory
Tfh T follicular helper
Tgd T gamma delta.
TCGA The Cancer Genome Atlas
TGCT testicular germ cell tumor
THCA thyroid carcinoma
THYM thymoma
UCEC uterine corpus endometrial carcinoma
UCS uterine carcinosarcoma

frontiers in Genetics

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fgene.2021.730847

Roles and Clinical Significances of ATF6, EMC6, and APAF1 in Prognosis of Pancreatic Cancer

Wang Xiao[1†], Rong-Chang Cao[1†], Wan-Jun Yang[1†], Jie-Hui Tan[1], Ruo-Qi Liu[1], He-Ping Kan[1], Lei Zhou[2], Na Zhang[3], Zhi-Ye Chen[3], Xue-Mei Chen[4], Jia Xu[5], Guo-Wei Zhang[1] and Peng Shen[6]**

[1]Division of Hepatobiliopancreatic Surgery, Department of General Surgery, Nanfang Hospital, Southern Medical University, Guangzhou, China, [2]Department of Hepoctobiliary Pancreatic Surgery, The Eighth Affiliated Hospital, Sun Yat-sen University, Shenzhen, China, [3]Department of Anesthesiology, Nanfang Hospital, Southern Medical University, Guangzhou, China, [4]Department of Occupational Health and Medicine, Guangdong Provincial Key Laboratory of Tropical Disease Research, School of Public Health, Southern Medical University, Guangzhou, China, [5]Department of Pathophysiology, Southern Medical University, Guangzhou, China, [6]Department of Oncology, Nanfang Hospital, Southern Medical University, Guangzhou, China

OPEN ACCESS

Edited by:
Youping Deng,
Rush University Medical Center,
United States

Reviewed by:
Shuangyu Lv,
Henan University, China
Muhammad Khan,
University of the Punjab, Pakistan

***Correspondence:**
Guo-Wei Zhang
guoweizhang77@163.com
Peng Shen
shenbo20110311@163.com

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

Specialty section:
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

Received: 25 June 2021
Accepted: 14 December 2021
Published: 11 February 2022

Citation:
Xiao W, Cao R-C, Yang W-J, Tan J-H, Liu R-Q, Kan H-P, Zhou L, Zhang N, Chen Z-Y, Chen X-M, Xu J, Zhang G-W and Shen P (2022) Roles and Clinical Significances of ATF6, EMC6, and APAF1 in Prognosis of Pancreatic Cancer. Front. Genet. 12:730847. doi: 10.3389/fgene.2021.730847

Background: Pancreatic cancer (PC) is prevalent among malignant tumors with poor prognosis and lacks efficient therapeutic strategies. Endoplasmic reticulum (ER) stress and apoptosis are associated with chronic inflammation and cancer progression. However, the prognostic value of ER stress-related, and apoptosis-related genes in PC remains to be further elucidated. Our study aimed at confirming the prognostic values of the ER stress-related genes, ATF6, EMC6, XBP1, and CHOP, and the apoptosis-related gene, APAF1, in PC patients.

Methods: Gene Expression Profiling Interactive Analysis 2 (GEPIA2) was used to evaluate prognosis value of ATF6, EMC6, XBP1, CHOP, and APAF1 in PC. Clinical data from 69 PC patients were retrospectively analyzed. Immunohistochemistry, Western blotting, and qRT-PCR were used for the assessment of gene or protein expression. The cell counting kit-8 (CCK-8) and the Transwell invasion assays were, respectively, used for the assessment of the proliferative and invasive abilities of PC cells. The prognostic values of ATF6, XBP1, CHOP, EMC6, and APAF1 in PC patients were evaluated using Kaplan–Meier and Cox regression analyses.

Results: XBP1 and CHOP expressions were not associated with PC recurrence-free survival (RFS), overall survival (OS) and disease-specific survival (DSS). ATF6 upregulation and EMC6 and APAF1 downregulations significantly correlated with the poor RFS, OS, and DSS of PC patients. ATF6 promoted PC cell proliferation and invasion, while EMC6 and APAF1 inhibited these events.

Conclusion: ATF6 upregulation and EMC6 and APAF1 downregulations may be valid indicators of poor prognosis of PC patients. Moreover, ATF6, EMC6, and APAF1 may constitute potential therapeutic targets in PC patients.

Keywords: pancreatic cancer, endoplasmic reticulum stress (ER stress), apoptosis, ATF6, EMC6, Apaf1, prognosis

INTRODUCTION

Pancreatic cancer (PC) is the fourth most common cause of cancer-related death in the United States with a 5-years survival rate of 10% (Siegel et al., 2021). The main treatment options include surgery, chemotherapy, radiotherapy, targeted therapy, supportive care, and their combination, however, surgical resection is the only curative therapy (Mizrahi et al., 2020). However, post-surgical resection recurrence is observed in approximatively 80% of PC patients (Groot et al., 2019; Kim et al., 2019). Despite remarkable improvements in surgical techniques in recent years, surgically resected patients are susceptible to death from their disease due to the high rate of recurrence (Sakamoto et al., 2020). Thus, the evaluation of the prognosis of PC patients and the development of new therapeutic methods are urgently needed to improve PC prognostic efficiency.

The endoplasmic reticulum (ER) stress is induced by various physiological or pathological strains on the cell, such as glucose deprivation, hypoxia, or chemotherapeutics, that subsequently activate unfolded protein response (UPR) as an adaptive response for cell recovery from stress (Dauer et al., 2019). Protein kinase R-like ER kinase (PERK), the transcription factor 6 (ATF6), and the inositol requiring enzyme 1α (IRE1α) constitute the three branches of the UPR signaling pathway (Lin et al., 2019). X-box-binding protein 1 (XBP1) is generated through the activation of the IRE1α-mediated cleavage of XBP1 mRNA cleavage (Dauer et al., 2019; Barez et al., 2020), which expression is associated with poor prognosis in cancer patients (Bagratuni et al., 2010; Chen et al., 2014; Kwon et al., 2018). The C/EBP homologous protein (CHOP) is a downstream factor of severe ER stress (Cao et al., 2019), which is upregulated in response to dysregulated UPR and which is used in the stratification of mesothelioma patients (Dalton et al., 2013). ATF6 is a crucial regulator of the UPR pathway that is involved in coagulation (Zheng et al., 2019), and that has been identified as a poor prognosis factor in biliopancreatic carcinoma (Martinez-Useros et al., 2015) and colon cancer (Liu et al., 2018). Although ATF6, XBP1, and CHOP are involved in the prognosis of multiple diseases, their roles in PC remain not well-known.

ER membrane protein complex subunit 6 (EMC6) is a novel positive regulator of autophagy regulator in human cells (Shen et al., 2016) that has been demonstrated to influence the development of ER stress (Chitwood and Hegde, 2019), and to induce apoptosis in gastric cancer cells (Wang et al., 2017). Apoptotic protease-activating factor 1 (APAF1) is a crucial factor in the mitochondria-dependent death pathway, which also plays a significant role in ER stress-induced apoptosis (Shiraishi et al., 2006). In our previous study, we found that ATF6/XBP1/CHOP axis could promote the progression of chronic pancreatitis (CP) (Zhou et al., 2019), and EMC6 could upregulate the expression of APAF1 to promote pancreatic acinar apoptosis and inflammatory injury of CP (Tan et al., 2020). Given that CP was regarded as a high risk factor for PC, in the present study, we aimed at characterizing the expression of ER stress-related proteins ATF6/XBP1/CHOP/EMC6 and apoptosis-related protein APAF1 in PC, and at analyzing the relationship between their expression, the clinico-pathological variables, and prognosis of surgically resected PC patients.

MATERIALS AND METHODS

Survival Analysis Based on Gene Expression Profiling Interactive Analysis 2

Different expressions of ATF6, EMC6, XBP1, CHOP, and APAF1 in PC and normal tissues were analyzed in Gene Expression Profiling Interactive Analysis 2 (GEPIA2; http://gepia2.cancer-pku.cn/#index). GEPIA2 is an interactive web server for analyzing the expression data of RNA from 9,736 tumors and 8,587 normal samples from the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) datasets (Tang et al., 2019). The *p*-value cutoff of the expression of gene for analysis was 0.01. The |Log2FC| cutoff of the expression of gene for analysis was 1, and we used log2 (TPM + 1) for log scale.

Survival analysis for overall survival (OS) and disease-free survival (DFS) in GEPIA2 was also used to estimate the relationship between PC prognostic value and the expression of genes ATF6, EMC6, XBP1, CHOP, and APAF1. The meaning of DFS was similar to recurrence-free survival (RFS) for this study. Hazards ratio (HR) was calculated based on Cox PH Model. The 95% confidence interval (CI) was shown by dotted line. The expression median value of gene for analyzing was identified as group cutoff to distinguish the high-expression group and the low-expression group.

Patients and Tissue Samples

PC and adjacent normal pancreatic tissue samples were retrospectively collected from 69 PC patients, including 39 males and 30 females, with a mean age of about 57 years ranging from 34 to 79 years. These patients underwent surgical resections in Nanfang Hospital, Southern Medical University, between October 2010 and April 2019. A more detailed information about the patients is shown in **Supplementary Table S1**. Before the experiments, each patient provided a written informed consent. After resection, each sample was frozen at −80°C until analysis and all patients were followed up until May 2019. Complete clinical and pathological data and follow-up documentations were recorded and analyzed for all patients in the study, who had never received preoperative chemotherapy or radiotherapy. The tumor stages in the study were classified according to the Union for International Cancer Control (UICC). RFS was defined as the period from the date of pancreatic resection until the date of recurrence diagnosis. OS was defined as the period from the date of surgical resection until the date of death or last follow-up. Disease-specific survival (DSS) was defined as the period from the date of surgical resection until the date of death due to PC. The collection and analysis of tissue and data were approved by the Ethics Committee of the Southern Medical University.

Hematoxylin and Eosin Staining and Immunohistochemistry

Pancreatic tissues were fixed in 4% neutral phosphate-buffered formalin, embedded in paraffin, and cut into 5-μm thick

sections. Hematoxylin and eosin (H&E) staining were performed by experienced pathologists and followed by double-blinded histological evaluations. For the immunohistochemical detection, the sections were sequentially incubated overnight at 4°C with anti-ATF6 (Bioss, diluted 1:100), anti-XBP1 (Bioss, diluted 1:200), anti-CHOP (Bioss, diluted 1:100), anti-EMC6 (Proteintech, diluted 1:100) and anti-APAF1 (Abcam, diluted 1:100) antibodies. After 30 min of incubation with secondary antibodies at room temperature, the sections were counterstained with DAB solution and hematoxylin. Positively stained cells were evaluated by two experienced pathologists according to a previously described protocol (Zhou et al., 2020). Five high magnification areas were evaluated from each sample.

The immunohistochemical scores were assessed based on the intensity of staining and the proportion of stained cells. The scores of 0, 1, 2, and 3, respectively, corresponded to negative, weak, moderate, and strong staining intensities. The proportion of positively stained cells for each intensity was scored as follows: 0 (0%–5% positive cells), 1 (5%–25% positive cells), 2 (26%–50% positive cells), 3 (51%–75% positive cells), and 4 (76%–100% positive cells). The IHC scores given by each pathologist were calculated by multiplying the proportion of positively stained cells by the staining intensity scores. The final IHC scores were the mean value of scores from two pathologists and divided into low expression (0–7) and high expression (8–12) groups.

Cell Culture and Transfection

The human PC cell lines, SW1990, HUPT4, PATU8988, PANC1, and ASPC1, were obtained from American Type Culture Collection (ATCC, Rockville, MD, USA). All the cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco). All mediums were supplemented with 10% fetal bovine serum (FBS) and maintained in a 37°C and 5% CO_2 atmosphere.

ATF6, EMC6, and APAF1 expression and functions were investigated by Western blotting, qRT-PCR, CCK8 assay, and Transwell assay. Si-ATF6, Si-EMC6, and Si-APAF1 were, respectively, used to inhibit the expression of ATF6, EMC6, and APAF1. The constructs OE-ATF6, OE-EMC6, and OE-APAF1 were used to, respectively, overexpress ATF6, EMC6, and APAF1. The negative control (NC) and the vector were designed and synthesized by RiboBio Co., Ltd. (Guangzhou, China). PC cells were separately seeded in 24-well plates at a density of 5×10^5 cells and transfected with Si-ATF6, Si-EMC6, Si-APAF1, OE-ATF6, OE-EMC6, OE-APAF1, NC, and the vector using riboFECT mRNA Transfection Reagent (RiboBio Co., Ltd. Guangzhou, China) according to the recommendations of the manufacturer. After 48 h of transfection and incubation in a 37°C and 5% CO_2 atmosphere, the cells were harvested for subsequent experiments.

Quantitative Real-Time PCR

Trizol Reagent (Merck, Germany) was used to extract RNA from cultured cells according to the instructions of the manufacturer. QRT-PCR was performed using the SYBR Premix Ex Tag kit (Takara Biotechnology Co., Ltd.) and the Applied Biosystems 7500 Real-Time PCR system (Thermo Fisher Scientific Inc., UK). The primers were designed and synthesized by RiboBio Co., Ltd. (Guangzhou, China). ATF6 forward, 5′-CGC CTT TTA GTC CGG TTC TT-3′ and reverse, 5′-CCA GTT GGT AAC AAT GCC ATG T-3′; EMC6 forward, 5′-GTC GCC AAG ATT TGC TCC CT-3′ and reverse, 5′-AAA CAC ACA ATG CCG GTA CAC-3′; APAF1 forward, 5′-GAT CCA CAC AGG CCA TCA CA-3′ and reverse, 5′-GGC GGG AGT CTA TGT TCC AC-3′. GAPDH forward, 5′-ATC ATC AGC AAT GCC TCC TG-3′ and reverse, 5′-ATG GAC TGT GGT CAT GAG TC-3′. ATF6, EMC6, and APAF1 expressions were normalized by GAPDH. The $2^{-\triangle\triangle Ct}$ method was used to calculate the relative expression levels. The expression levels of the genes were measured by qRT-PCR.

Western Blot Assay

RiPA buffer (GenStar, China) was used to extract the proteins from the transfected cell lines. The proteins were loaded and separated on SDS-PAGE, and transferred onto PVDF membranes (Millipore, USA). After blocking with 5% non-fat milk, the membranes were incubated with anti-ATF6 (Bioss, China), anti-EMC6 (Proteintech, USA), anti-APAF1 (Abcam, UK), and anti-GAPDH (Fude Biological Technology Co., Ltd. China) antibodies overnight at 4°C. Following this step, the membranes were incubated with horseradish peroxidase-coupled secondary antibodies for 1 h, and finally, the expression of the proteins was revealed using the enhanced chemiluminescence solution (ECL, PerkinElmer, USA).

Cell Counting Kit-8 Assay

The cell counting kit-8 assay was performed to assess the viability of the transfected cell lines. The transfected cells were seeded in a 96-well plate and cultured for 72 h. Then, 10 μl of CCK8 solution (Dojindo, Japan) was added to each well and incubated for 2 h. The absorbance at 450 nm was measured using a microplate reader.

Transwell Invasion Assay

The Transwell invasion assay was performed to measure the invasion ability of the transfected cells using the Transwell chambers. The transfected cells were suspended in serum-free medium and seeded in the upper chambers that were precoated with Matrigel. The medium containing 10% FBS was placed in the lower chamber. After 24 h of incubation at 37°C, the cells on the bottom chamber were stained with 0.1% crystal violet for 10 min at 37°C. The number of stained cells in the lower chambers was calculated using a microscope.

Statistical Analysis

The SPSS software version 26.0 (SPSS, Chicago, IL) and the GraphPad Prism software version 8.2 (San Diego, CA, USA) were used for statistical analyses. The data were reported as mean ± standard deviation. The paired Student's t-test or the chi-square test were used to analyze the expression differences of genes among normal, high, and low expression groups. The

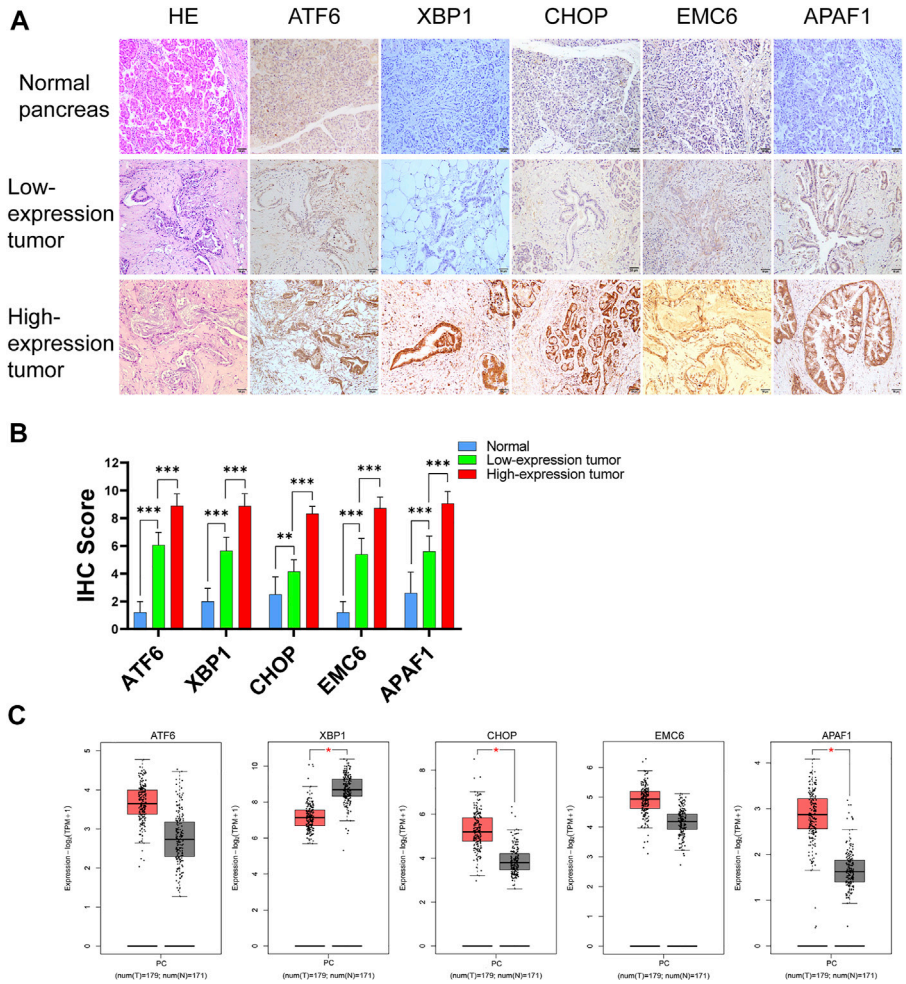


FIGURE 1 | Analysis of ATF6, XBP1, CHOP, EMC6, and APAF1 expression in tumor and adjacent normal tissue samples of human. **(A)** H&E staining and immunohistochemical detection of ATF6, XBP1, CHOP, EMC6, and APAF1 protein expression in pancreatic tissue from normal and PC patients. **(B)** The expression of ATF6, XBP1, CHOP, EMC6, and APAF1 in tumor and normal samples. **(C)** The expression of ATF6, XBP1, CHOP, EMC6, and APAF1 in tumor (red) and normal samples (gray) *via* GEPIA2. T, tumors; N, normal tissues; $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$. Scale bars = 50 μm. ATF6, transcription factor 6; XBP1, X-box-binding protein 1; CHOP, C/EBP homologous protein; EMC6, ER membrane protein complex subunit 6; APAF1, apoptotic protease-activating factor 1; H&E, hematoxylin and eosin; PC, pancreatic cancer; GEPIA2, gene expression profiling interactive analysis.

Kaplan–Meier method was performed for survival curves between low-expression and high-expression groups using the log-rank test. To identify the factors involved in PC, the Cox proportional hazards regression method was applied using univariate and multivariate analyses. Differences were considered significant if $p < 0.05$.

RESULTS

Elevated Expression of ATF6 and Reduced Expression of EMC6 and APAF1 Associated with Worse Prognosis in PC According to the GEPIA2 Database

GEPIA2 was used to evaluate the relationship between the expression of ATF6, EMC6, XBP1, CHOP, APAF1, and prognosis value of PC. There was an upregulated trend of ATF6, EMC6, APAF1, and CHOP expression in PC compared with that in normal pancreatic tissues, while the result of XBP1 was opposite in this event (**Figure 1C**). Noticeably, the elevated expression of ATF6 and reduced expression of EMC6 and APAF1 showed a statistically significant association with poor OS for PC, but not with DFS, while the expression level of CHOP and XBP1 had no correlation with OS and DFS of PC (**Figure 2**).

Clinicopathological Characteristics of Patients With Pancreatic Cancer

The clinicopathological features of the 69 patients are described in **Table 1**. Briefly, the median age of the patients at diagnosis was 58 years (age range: 34–79 years), and 39.1% of patients were

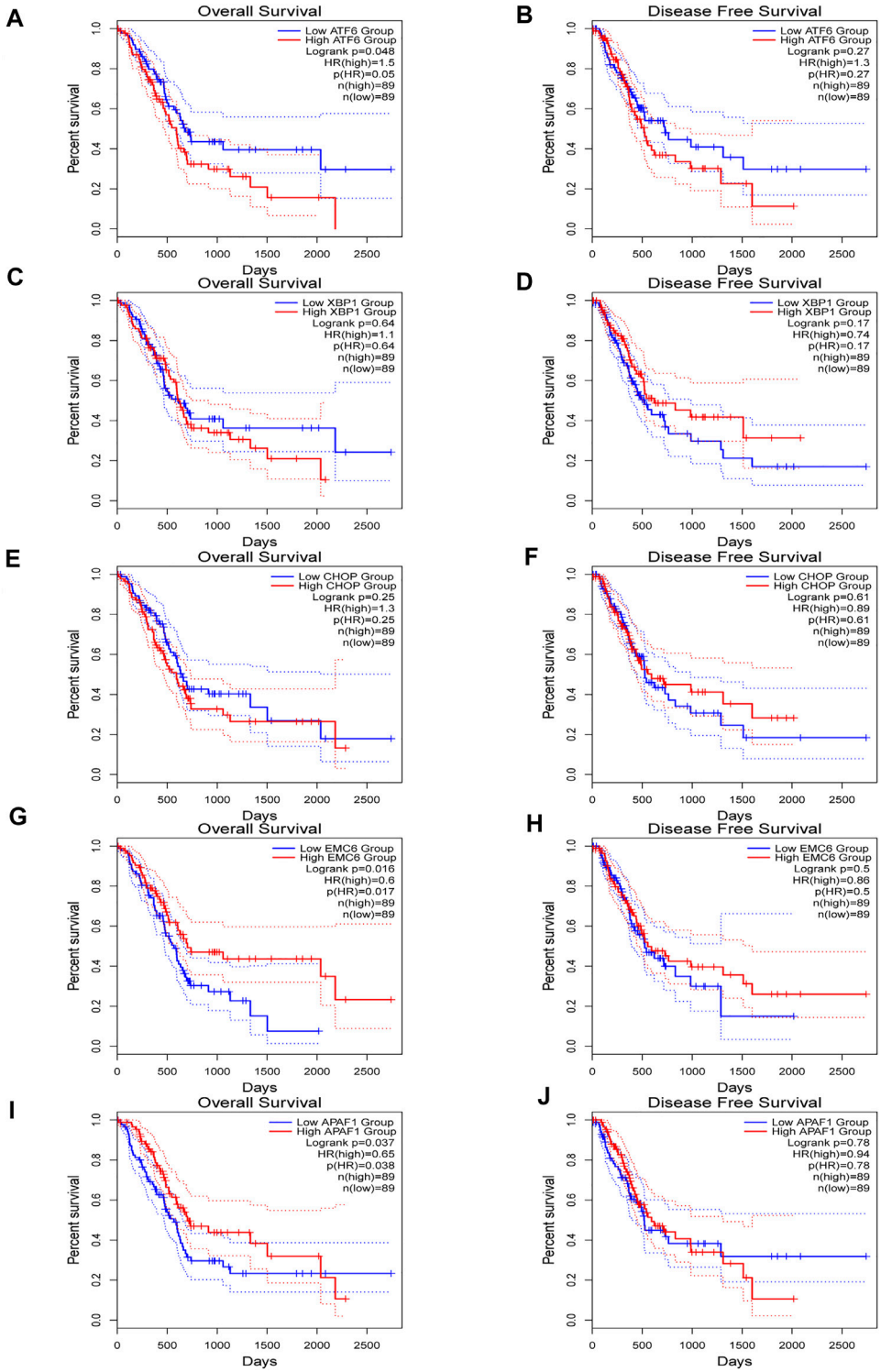


FIGURE 2 | Kaplan–Meier plotter for overall survival (OS) **(A,C,E,G,I)** and disease-free survival (DFS) **(B,D,F,H,J)** based ATF6, XBP1, CHOP, EMC6, or APAF1 expression *via* survival analysis in GEPIA2 database. The two-sided log-rank test was performed to compare differences by *p*-values.

>60 years old and 63.8% had tumors with sizes>3 cm. Patients were male in 56.5%, and the proportion of smokers was 23.2%. Adenocarcinoma was the most common histological type that was observed in 95.7% of patients. Stage I (42.0%) and II (44.9%) were common, with involvement of lymph nodes, and vascular and neural invasions observed in 31.9%, 13.0%, and 26.1%, respectively.

TABLE 1 | Clinicopathological characteristics of patients with pancreatic cancer (PC).

Variable	Number (%)
Age (years)	
≤60	42 (60.9%)
>60	27 (39.1%)
Gender	
Male	39 (56.5%)
Female	30 (43.5%)
Smoking	
Yes	16 (23.2%)
No	53 (76.8%)
Tumor size	
≤3 cm	25 (36.2%)
>3 cm	44 (63.8%)
Histology	
Adenocarcinoma	66 (95.7%)
Others	3 (4.3%)
Stage	
I	29 (42.0%)
II	31 (44.9%)
III	2 (2.9%)
IV	7 (10.1%)
Lymph nodes involved	
Yes	22 (31.9%)
No	47 (68.1%)
Vascular invasion	
Yes	9 (13.0%)
No	60 (87.0%)
Neural invasion	
Yes	18 (26.1%)
Noo	51 (73.9%)

Elevated Expression of Endoplasmic Reticulum Stress and Apoptosis-Related Proteins in Pancreatic Cancer

To determine the expression levels of ER stress-related and apoptosis-related proteins in human PC, we collected PC and normal pancreatic tissues for immunohistochemistry analysis. The expression of the ER stress-related proteins, ATF6, XBP1, CHOP, and EMC6, and the apoptosis-related protein, APAF1, were significantly higher in PC tissues compared with those in normal pancreatic tissues (**Figures 1A,B**). Compared with these results to GEPIA2, the reason for the different results in XBP1 may be the differences of sources of data analyzed. The RNA-seq datasets were used by GEPIA2, instead of the IHC data. These results demonstrate that the expressions of ATF6, XBP1, CHOP, EMC6, and APAF1 were upregulated in PC.

Prognostic Values of Transcription Factor 6, ER Membrane Protein Complex Subunit 6, and Apoptotic Protease-Activating Factor 1 in Pancreatic Cancer Patients

To validate the associations between clinicopathological and molecular variables, and prognostic values in the PC patients, survival analysis was performed using the Kaplan–Meier method and the significance was tested with the log rank test. Univariate Cox regression analysis revealed that the TNM stage, lymph node involvement, and the expression levels of ATF6, EMC6, and APAF1, significantly correlate with RFS, OS, and DSS, while no significant differences were observed between the two groups based on age, sex, smoking, tumor size, neural and vascular invasions, and XBP1 and CHOP expression (**Tables 2–4**). Notably, the multivariate Cox regression analysis revealed that the TNM stage [HR = 3.578; $p = 0.027$], ATF6 expression [HR = 0.220; $p < 0.001$], EMC6 expression [HR = 2.571; $p = 0.020$], and APAF1 expression [HR = 2.426; $p = 0.026$] were also independent prognostic factors for RFS (**Table 2**). Simultaneously, the TNM stage [HR = 4.064; $p = 0.014$], lymph node involvement [HR = 0.380; $p = 0.034$], ATF6 expression [HR = 0.229; $p = 0.001$], EMC6 expression [HR = 2.956; $p = 0.010$], and APAF1 expression [HR = 2.369; $p = 0.034$] were also found to be significant prognostic factors for OS (**Table 3**). Moreover, the TNM stage [HR = 4.073; $p = 0.017$], lymph node involvement [HR = 0.396; $p = 0.046$], ATF6 expression [HR = 0.183; $p < 0.001$], EMC6 expression [HR = 3.275; $p = 0.015$], and APAF1 expression [HR = 2.887; $p = 0.029$] were also prognostic factors for DSS (**Table 4**). Kaplan–Meier survival plots indicated significantly higher survival rates at each time point for the $ATF6^{low}$, $EMC6^{high}$, and $APAF1^{high}$ groups compared with those in the $ATF6^{high}$, $EMC6^{low}$ and $APAF1^{low}$ groups ($p < 0.05$, **Figures 3A,D,E**). There was no statistical significance for XBP1 and CHOP ($p > 0.05$, **Figures 3B,C**). Discrepancy of results in ATF6, EMC6, and APAF1 from GEPIA2 survival analysis and our experiment may be due to different sources of data analyzed. Another reason probably lies in the different group cutoff in distinguishing the high-expression group and the low-expression group. The samples were divided into two groups by expression median of genes in GEPIA2 database, instead of defining "high expression" as the IHC scores ≥8 in our study. Nevertheless, both GEPIA2 survival analysis and our study identified that the expression of ATF6, EMC6, and APAF1 was related to PC patients' survival. The results demonstrate that among the genes that are related to ER stress and apoptosis in this research, only ATF6, EMC6, and APAF1 were associated with PC patients' survival, which would be used in further studies.

The Expression of Transcription Factor 6, ER Membrane Protein Complex Subunit 6, and Apoptotic Protease-Activating Factor 1 in Pancreatic Cancer Cell Lines

To evaluate the expression of ATF6, EMC6, and APAF1 in pancreatic carcinoma cells by qRT-PCR, the PC cell lines, SW1990, HUPT4, PATU8988, PANC1, and ASPC1, were used. Noticeably, the highest and the lowest level of expression of ATF6 was observed in the cell lines, ASPC1 and SW1990 (**Figure 4A**). Therefore, SW1990 and ASPC1 were used as representative PC cell lines for ATF6 subsequent studies. Similarly, PATU8988 and SW1990 were selected for EMC6, PANC1, and ASPC1 was selected for APAF1 experiments (**Figures 4B,C**).

TABLE 2 | Univariate and multivariate Cox regression analysis for recurrence-free survival.

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*p*-Value	HR (95% CI)	*p*-Value
Age (≤60 vs. >60 years)	1.751 (0.964–3.180)	0.066		
Gender (male vs. female)	0.833 (0.464–1.496)	0.541		
Smoking (yes vs. no)	1.184 (0.571–2.457)	0.650		
Size (>3 vs. ≤3 cm)	0.581 (0.301–1.124)	0.107		
Stage (I, II, III vs. IV)	5.109 (2.121–12.305)	0.000[a]	3.578 (1.154–11.099)	0.027[a]
Lymph nodes involved (yes vs. no)	0.435 (0.240–0.788)	0.006[a]		
Neural invasion (yes vs. no)	1.044 (0.484–2.253)	0.912		
Vascular invasion (yes vs. no)	0.572 (0.264–1.240)	0.157		
ATF6 (high vs. low)	0.392 (0.213–0.720)	0.003[a]	0.220 (0.095–0.510)	0.000[a]
XBP1 (high vs. low)	1.027 (0.573–1.840)	0.929		
CHOP (high vs. low)	0.948 (0.530–1.694)	0.856		
EMC6 (high vs. low)	2.056 (1.091–3.874)	0.026[a]	2.571 (1.160–5.700)	0.020[a]
APAF1 (high vs. low)	2.017 (1.040–3.911)	0.038[a]	2.426 (1.114–5.281)	0.026[a]

Note. HR, hazard ratio.
95% CI, 95% confidence interval.
[a]*Statistically significant results (p < 0.05).*

TABLE 3 | Univariate and multivariate Cox regression analysis for overall survival.

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*p*-Value	HR (95% CI)	*p*-Value
Age (≤60 vs. >60 years)	1.880 (1.034–3.419)	0.039[a]		
Gender (male vs. female)	0.823 (0.458–1.478)	0.514		
Smoking (yes vs. no)	1.137 (0.547–2.362)	0.731		
Size (>3 vs. ≤3 cm)	0.599 (0.310–1.158)	0.128		
Stage (I, II, III vs. IV)	6.648 (2.653–16.657)	0.000[a]	4.064 (1.325–12.465)	0.014[a]
Lymph nodes involved (yes vs. no)	0.426 (0.236–0.769)	0.005[a]	0.380 (0.155–0.932)	0.034[a]
Neural invasion (yes vs. no)	0.958 (0.422–2.074)	0.913		
Vascular invasion (yes vs. no)	0.545 (0.252–1.181)	0.124		
ATF6 (high vs. low)	0.378 (0.206–0.696)	0.002[a]	0.229 (0.099–0.530)	0.001[a]
XBP1 (high vs. low)	1.080 (0.603–1.936)	0.795		
CHOP (high vs. low)	0.981 (0.549–1.753)	0.948		
EMC6 (high vs. low)	2.082 (1.105–3.924)	0.023[a]	2.956 (1.290–6.772)	0.010[a]
APAF1 (high vs. low)	2.117 (1.092–4.103)	0.026[a]	2.369 (1.069–5.249)	0.034[a]

Note. HR, hazard ratio.
95% CI, 95% confidence interval.
[a]*Statistically significant results (p < 0.05).*

High Expression of Transcription Factor 6 and Low Expression of ER Membrane Protein Complex Subunit 6, or Apoptotic Protease-Activating Factor 1 Promote Proliferative and Invasive Abilities of Pancreatic Cancer Cells

To further explore the function of ATF6, EMC6, and APAF1 in PC, we determined the effect of ATF6, EMC6, APAF1 on the proliferation and invasion abilities of PC cells using the CCK8 and Transwell assays on ATF6 transfected PC cell lines, SW1990 and ASPC1, EMC6 transfected PC cell lines, PATU8988 and SW 1990, APAF1 transfected cell lines, PANC1 and ASPC1. For this, qRT-PCR and Western blotting were used and revealed that the expression of ATF6, EMC6, and APAF1 were markedly increased in PC cells that were transfected with OE-ATF6, EMC6, and APAF1, and markedly inhibited in PC cells transfected with Si-ATF6, EMC6, and APAF1 when compared with the control (**Figures 4D–L**).

The result of the CCK8 assay showed that ATF6 overexpression and EMC6 or APAF1 knockdown enhances the growth of PC cells (**Figures 5B,D,G**), whereas ATF6 knockdown and EMC6 or APAF1 overexpression decreases the proliferation of PC cells compared with the control (**Figures 5A,E,H**). These results indicate that ATF6 promotes the viability of PC cells, while EMC6 and APAF1 have an inhibitory role.

The Transwell assay showed that the invasiveness of PC cells was markedly increased when ATF6 expression level is elevated and when EMC6 and APAF1 expression levels are reduced. However, ATF6 knockdown and EMC6 or APAF1 overexpression had the opposite effect (**Figures 5C,F,I**). Therefore, ATF6 promotes the invasion ability of PC cells, while EMC6 and APAF1 impair it.

TABLE 4 | Univariate and multivariate Cox regression analysis for disease-specific survival.

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*p*-Value	HR (95% CI)	*p*-Value
Age (≤60 vs. >60 years)	2.029 (1.077–3.823)	0.029[a]		
Gender (male vs. female)	0.865 (0.467–1.602)	0.645		
Smoking (yes vs. no)	1.068 (0.508–2.245)	0.861		
Size (>3 vs. ≤3 cm)	0.592 (0.296–1.184)	0.138		
Stage (I, II, III vs. IV)	7.494 (2.865–19.604)	0.000[a]	4.073 (1.284–12.922)	0.017[a]
Lymph nodes involved (yes vs. no)	0.396 (0.213–0.739)	0.004[a]	0.396 (0.159–0.983)	0.046[a]
Neural invasion (yes vs. no)	1.016 (0.445–2.319)	0.969		
Vascular invasion (yes vs. no)	0.546 (0.239–1.246)	0.150		
ATF6 (high vs. low)	0.327 (0.170–0.631)	0.001[a]	0.183 (0.077–0.438)	0.000[a]
XBP1 (high vs. low)	1.073 (0.579–1.988)	0.824		
CHOP (high vs. low)	1.072 (0.579–1.983)	0.826		
EMC6 (high vs. low)	2.028 (1.044–3.937)	0.037[a]	3.275 (1.255–8.550)	0.015[a]
APAF1 (high vs. low)	2.397 (1.170–4.909)	0.017[a]	2.887 (1.112–7.496)	0.029[a]

Note. HR, hazard ratio.
95% CI, 95% confidence interval.
[a]*Statistically significant results (p < 0.05).*

DISCUSSION

PC has always been one of the greatest challenges of human health, especially in East Asia, where 458,918 newly diagnosed cases and approximately 432,242 death cases were recorded in 2018 (Bray et al., 2018). Despite continuing advancements in surgery, chemotherapy, radiotherapy, immunotherapy, and targeted therapy of PC, the prognosis of PC patients is still poor (Mizrahi et al., 2020; Thakur et al., 2021). Unfortunately, adjuvant and neoadjuvant treatments, which were used for PC therapy, had a scarce benefit on PC prognosis (O'Reilly and Ferrone, 2020). Thus, the evaluation of the prognosis of PC patients and the development of new therapeutic methods are in great need to be improved .

ATF6, XBP1, and CHOP were identified as core proteins in UPR signaling, which contribute to various physiological processes and cancer development (Hetz et al., 2020). The ER stress was linked to a variety of cancers and was associated with their prognosis (Nikesitch et al., 2016). Furthermore, blocking moderate ER stress and UPR could lead to tumoricidal effects (Mohamed et al., 2017). In the present study, we showed that ATF6, EMC6, XBP1, and CHOP expression are significantly higher in PC tissues compared with those in adjacent normal pancreatic tissues, indicating the potential involvement of ER stress-related proteins in PC progression. Thus, ER stress and UPR could be potential regulators of treatment effects in PC.

ATF6 is a UPR sensor that is located in the ER membrane, and that is associated with poor prognosis in Biliopancreatic and colon cancers (Martinez-Useros et al., 2015; Liu et al., 2018). Mutations in p53 result in tumor-cell differentiation and transition to malignant lesions in human PC (Morris et al., 2019). It was reported that p53 mutants enhance tumor aggressiveness by promoting cell invasion, metastasis, and chemoresistance through their interactions with ATF6 (Sicari et al., 2019). We supposed that p53 may be an ATF6 potential downstream molecule associated with a poor PC prognosis. Additionally, some studies reported that OTUB1 promotes the progression of bladder cancer through its interaction with ATF6 (Zhang et al., 2021), and that ATF6 could facilitate cervical cancer cell growth and migration through the MAPK pathway (Liu et al., 2020), resulting in poor cancer prognosis. As observed in other cancers, our study indicated that of ATF6 increased expression correlates with poor prognosis of PC. We found that the survival rate of patients was lower, and that the proliferation and aggressiveness of tumor cells were stronger in PC when ATF6 expression is elevated. Further studies are required to reveal the exact mechanism of ATF6 in cancer. These can provide opportunities for the development of new targeting therapies for PC.

EMC6 is an autophagy-related protein that is overexpressed in U2OS osteosarcoma and HCT116 colon carcinoma cells, and that participates in the formation of autophagosomes and in accelerating the degradation of autophagic substrates in lysosomes (Li et al., 2019). Meanwhile, EMC6 functions as a tumor suppressor and its overexpression induces apoptosis and cell cycle arrest in gastric cancer cells (Wang et al., 2017). Similarly, in our study, EMC6 was expressed at low levels in PC tissues from better surviving patients. EMC6 protein reduced PC cells viability and invasion, and cancer patients with high EMC6 expression had longer OS and RFS. EMC6 participates in cell autophagy through its interaction with RAB5A, and its deficiency induces the impairment of autophagy (Li et al., 2013). Recent studies showed that autophagy plays a dual role in PC progression, which suggest its potential therapeutic targeting. Facilitating and inhibiting autophagy were both effective therapeutic methods (Li et al., 2020; Piffoux et al., 2021). However, the specific mechanism of the dual effects of autophagy in PC progression is unclear, and therefore, further studies are required. In this study, EMC6 is a regulator of autophagy that exhibited an inhibitory effect on PC cells. Thus, the regulation of EMC6 expression could offer a novel direction for PC treatment.

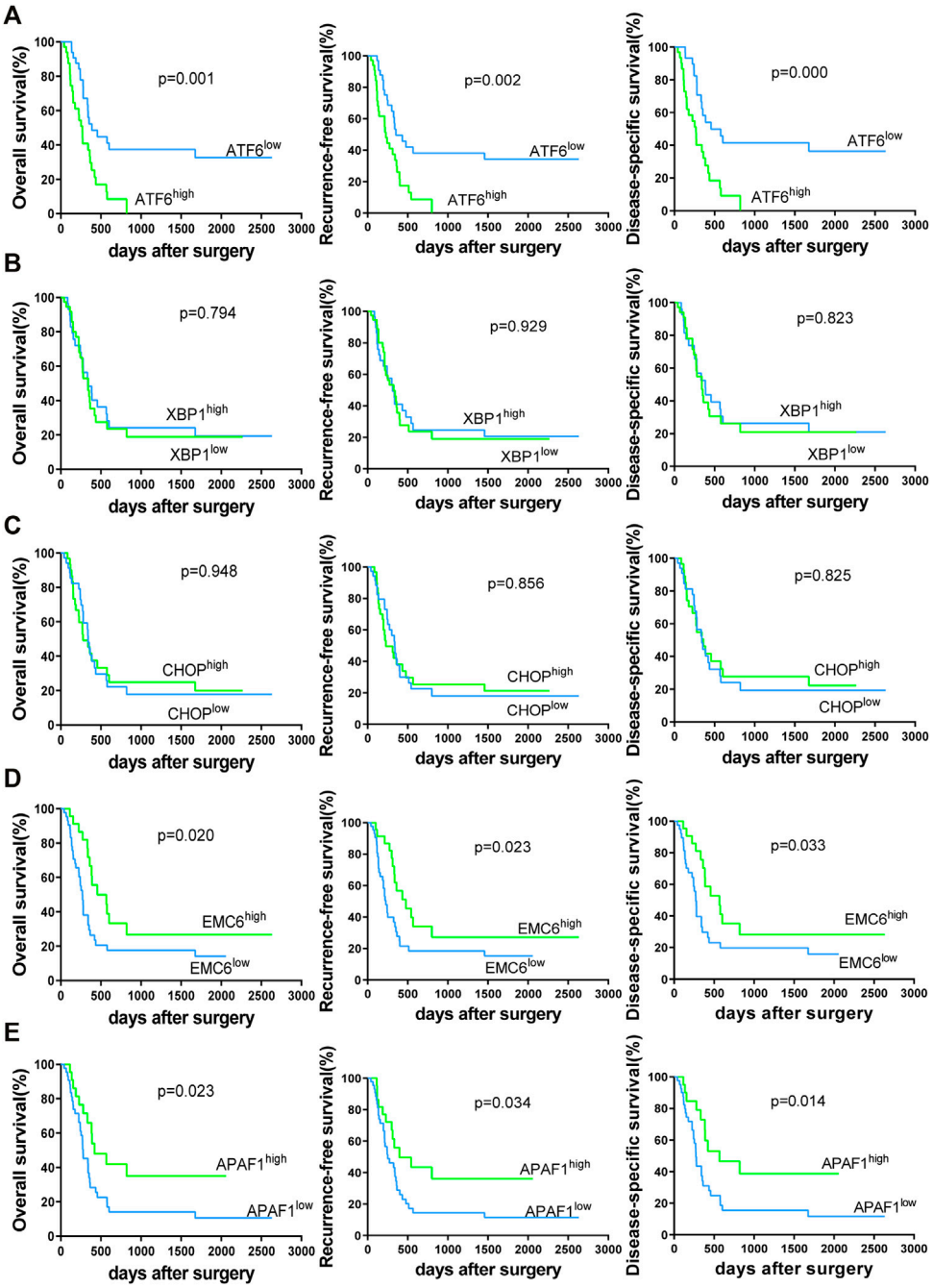


FIGURE 3 | Kaplan–Meier survival curves of 69 PC patients for overall survival (OS), recurrence-free survival (RFS) and disease-specific survival (DSS)-based ATF6 **(A)**, XBP1 **(B)**, CHOP **(C)**, EMC6 **(D)**, or APAF1 **(E)** expression. The two-sided log-rank test was performed to compare differences by *p*-values.

XBP1 is considered as a biomarker of poor clinical outcomes in patients with pulmonary adenocarcinoma (Kwon et al., 2018), breast cancers (Chen et al., 2014), and multiple myeloma (Bagratuni et al., 2010). CHOP serves as an apoptosis specific transcription factor (Zhang et al., 2012), which expression is related to mesothelioma stratification of patients (Dalton et al., 2013) and cancer staging (Lee et al., 2013). In this study, we demonstrated that XBP1 and CHOP high expression occur in PC tissues, however, the correlation between XBP1 and CHOP

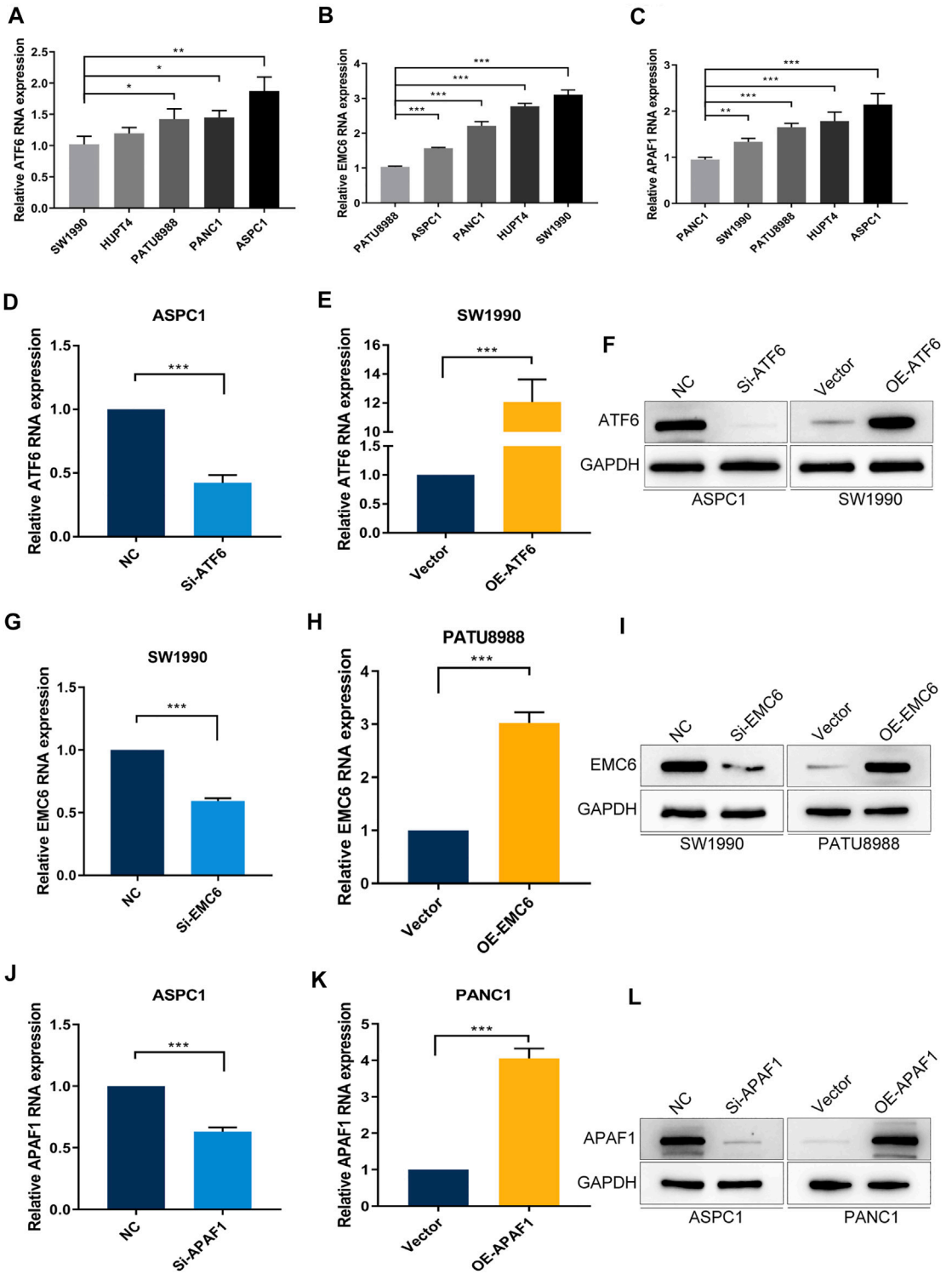


FIGURE 4 | PC cell lines with different expression levels of ATF6, EMC6, and APAF1. The expression of **(A)** ATF6, **(B)** EMC6, and **(C)** APAF1 in different PC cell lines were detected by qRT-PCR. ATF6 expression in ASPC1 and SW1990 cell lines that were transfected with Si-ATF6 and OE-ATF6, respectively, were detected by qRT-PCR **(D,E)** and Western blot **(F)**. The expression of EMC6 in SW1990 and PATU8988 cell lines that were transfected with Si-EMC6 and OE-EMC6, respectively, were measured by qRT-PCR **(G,H)** and Western blot **(I)**. APAF1 expression in ASPC1 and PANC1 cell lines that were transfected with Si-APAF1 and OE-APAF1, respectively, were evaluated by qRT-PCR **(J,K)** and Western blot **(L)**. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$.

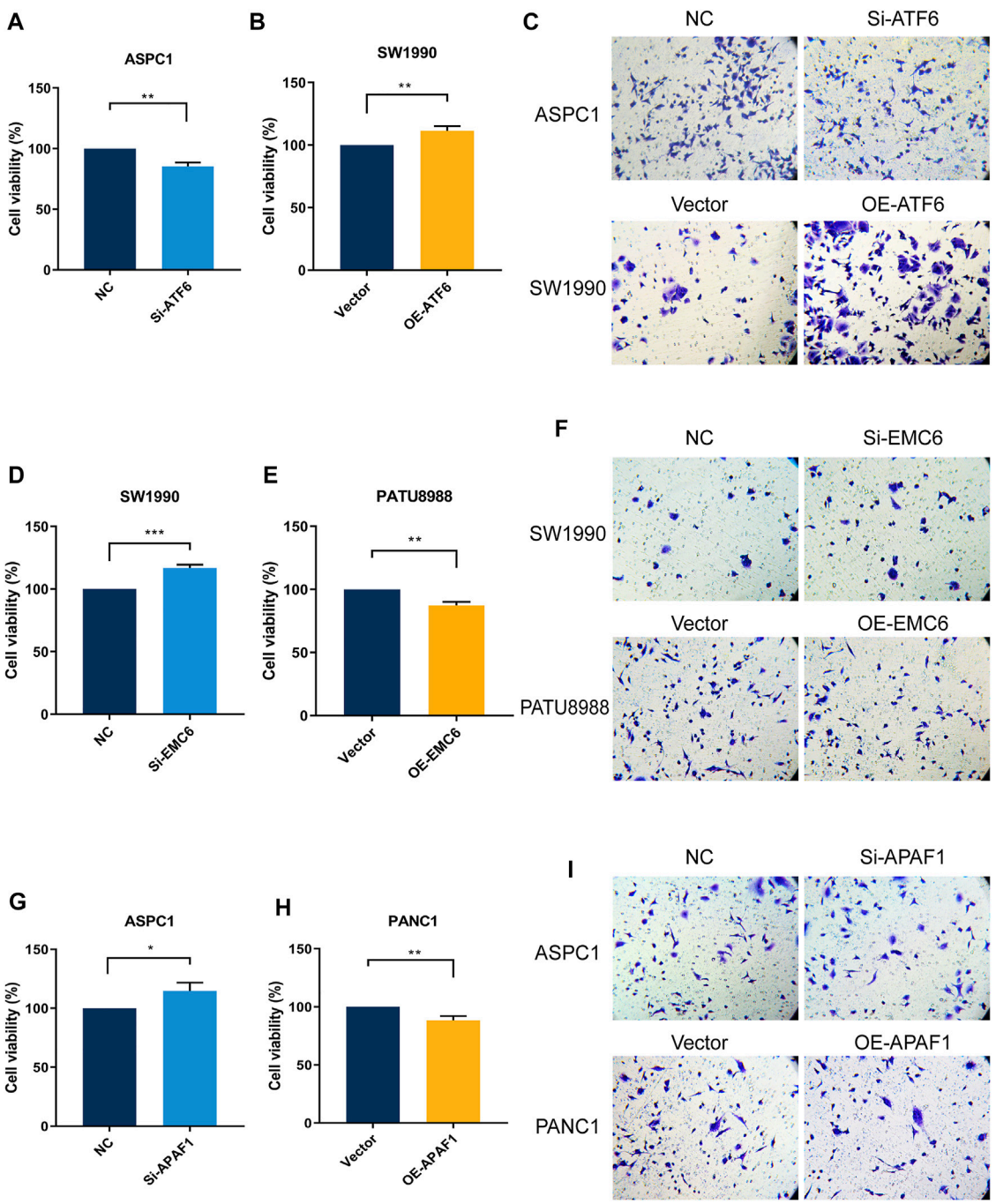


FIGURE 5 | ATF6 promoted PC cell viability and invasion, while EMC6 and APAF1 inhibited these events. ATF6 effects on cell viability and invasion of Si-ATF6 transfected ASPC1 cells and OE-ATF6 transfected SW1990 cells were determined using **(A,B)** Cell-counting kit-8 (CCK-8) assay and **(C)** Transwell assay, respectively. **(D,E)** CCK8 and **(F)** Transwell assays were, respectively, performed to measure cell viability and invasion of Si-EMC6 transfected SW1990 cells and OE-EMC6 transfected PATU-8988 cells. Cell viability and invasion of Si-APAF1 transfected ASPC1 cells and OE-APAF1 transfected PANC1 were separately evaluated using **(G,H)** the CCK8 assay and **(I)** the Transwell assay. $^{*}p \leq 0.05$, $^{**}p \leq 0.01$, $^{***}p \leq 0.001$.

expression and the survival of PC patients was not statistically significant.

APAF1 is as a key regulator of cell death and cell recovery pathways, and therefore, the dysregulation of apoptosis is at the root of various diseases (Gortat et al., 2015). Moreover, APAF1 expression was significantly suppressed by miR-23a in PC cells, which promotes PC cell proliferation and represses apoptosis (Liu et al., 2015). Our results revealed that APAF1 was overexpressed in PC tissues and inhibited the proliferation and invasion of PC cells, that contributed to a promising prognosis in PC patients. Shiraishi et al.

reported that APAF1 plays a crucial role in ER stress-induced apoptosis (Shiraishi et al., 2006), and EMC6 has been demonstrated to influence the development of ER stress (Chitwood and Hegde, 2019). In this study, we show that APAF1 also inhibits the viability and invasion of PC cells. The exact mechanism of the interaction between ER stress and EMC6 or APAF1 on the prognosis of PC is worthy of further investigation.

CONCLUSION

The expressions of ATF6, CHOP, XBP1, EMC6, and APAF1 are significantly involved in PC progression, and ATF6 overexpression and the inhibition of EMC6 or APAF1 expression are associated with poor clinical outcome in PC patients. These results suggest the potential use of these biomarkers as prognostic predictors for PC patients following surgery.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Southern Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

G-WZ conceived and designed the study. WX, R-CC, W-JY, and R-QL contributed in carrying out the cell experiments. WX, R-CC, and W-JY collected and organized the clinical samples and clinical information. J-HT, LZ, NZ, Z-YC, X-MC, and JX participated in the data analysis. WX drafted the manuscript. G-WZ, H-PK, and PS revised the manuscript. All authors read and approved the final manuscript.

FUNDING

The article is supported by the National Natural Science Foundation of China (82170655), the Guangdong Science and Technology Planning Project (2019A030317018), the Scientific Research Startup Program of Southern Medical University by High-level University Construction Funding of Guangdong Provincial Department of Education (CX2018N012), the Clinical Research Program of Nanfang Hospital, Southern Medical University (2018CR046), and the Clinical Research Startup Program of Southern Medical University by High-Level University Construction Funding of Guangdong Provincial Department of Education (LC2016PY011).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.730847/full#supplementary-material

REFERENCES

Bagratuni, T., Wu, P., Gonzalez de Castro, D., Davenport, E. L., Dickens, N. J., Walker, B. A., et al. (2010). XBP1s Levels Are Implicated in the Biology and Outcome of Myeloma Mediating Different Clinical Outcomes to Thalidomide-Based Treatments. *Blood* 116, 250–253. doi:10.1182/blood-2010-01-263236

Barez, S. R., Atar, A. M., and Aghaei, M. (2020). Mechanism of Inositol-Requiring Enzyme 1-alpha Inhibition in Endoplasmic Reticulum Stress and Apoptosis in Ovarian Cancer Cells. *J. Cel. Commun. Signal.* 14, 403–415. doi:10.1007/s12079-020-00562-7

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J. Clinicians* 68, 394–424. doi:10.3322/caac.21492

Cao, Y., Trillo-Tinoco, J., Sierra, R. A., Anadon, C., Dai, W., Mohamed, E., et al. (2019). ER Stress-Induced Mediator C/EBP Homologous Protein Thwarts Effector T Cell Activity in Tumors through T-Bet Repression. *Nat. Commun.* 10, 1280. doi:10.1038/s41467-019-09263-1

Chen, X., Iliopoulos, D., Zhang, Q., Tang, Q., Greenblatt, M. B., Hatziapostolou, M., et al. (2014). XBP1 Promotes Triple-Negative Breast Cancer by Controlling the HIF1α Pathway. *Nature* 508, 103–107. doi:10.1038/nature13119

Chitwood, P. J., and Hegde, R. S. (2019). The Role of EMC during Membrane Protein Biogenesis. *Trends Cel. Biol.* 29, 371–384. doi:10.1016/j.tcb.2019.01.007

Dalton, L. E., Clarke, H. J., Knight, J., Lawson, M. H., Wason, J., Lomas, D. A., et al. (2013). The Endoplasmic Reticulum Stress Marker CHOP Predicts Survival in Malignant Mesothelioma. *Br. J. Cancer* 108, 1340–1347. doi:10.1038/bjc.2013.66

Dauer, P., Sharma, N. S., Gupta, V. K., Durden, B., Hadad, R., Banerjee, S., et al. (2019). ER Stress Sensor, Glucose Regulatory Protein 78 (GRP78) Regulates Redox Status in Pancreatic Cancer Thereby Maintaining "stemness". *Cell Death Dis.* 10, 132. doi:10.1038/s41419-019-1408-5

Gortat, A., Sancho, M., Mondragón, L., Messeguer, À., Pérez-Payá, E., and Orzáez, M. (2015). Apaf1 Inhibition Promotes Cell Recovery from Apoptosis. *Protein Cell* 6, 833–843. doi:10.1007/s13238-015-0200-2

Groot, V. P., Gemenetzis, G., Blair, A. B., Rivero-Soto, R. J., Yu, J., Javed, A. A., et al. (2019). Defining and Predicting Early Recurrence in 957 Patients with Resected Pancreatic Ductal Adenocarcinoma. *Ann. Surg.* 269, 1154–1162. doi:10.1097/SLA.0000000000002734

Hetz, C., Zhang, K., and Kaufman, R. J. (2020). Mechanisms, Regulation and Functions of the Unfolded Protein Response. *Nat. Rev. Mol. Cel. Biol.* 21, 421–438. doi:10.1038/s41580-020-0250-z

Kim, Y. I., Song, K. B., Lee, Y.-J., Park, K.-M., Hwang, D. W., Lee, J. H., et al. (2019). Management of Isolated Recurrence after Surgery for Pancreatic Adenocarcinoma. *Br. J. Surg.* 106, 898–909. doi:10.1002/bjs.11144

Kwon, D., Koh, J., Kim, S., Go, H., Min, H. S., Kim, Y. A., et al. (2018). Overexpression of Endoplasmic Reticulum Stress-Related Proteins, XBP1s and GRP78, Predicts Poor Prognosis in Pulmonary Adenocarcinoma. *Lung Cancer* 122, 131–137. doi:10.1016/j.lungcan.2018.06.005

Lee, Y. J., Ha, Y. J., Na Kang, Y., Kang, K. J., Hwang, J. S., Chung, W. J., et al. (2013). The Autophagy-Related Marker LC3 Can Predict Prognosis in Human Hepatocellular Carcinoma. *PLoS One* 8, e81540. doi:10.1371/journal.pone.0081540

Li, J., Chen, X., Kang, R., Zeh, H., Klionsky, D. J., and Tang, D. (2020). Regulation and Function of Autophagy in Pancreatic Cancer. *Autophagy* 17, 3275–3296. doi:10.1080/15548627.2020.1847462

Li, R., Wang, X., Zhang, X., Yu, J., Feng, J., Lv, P., et al. (2019). Ad5-EMC6 Mediates Antitumor Activity in Gastric Cancer Cells through the Mitochondrial Apoptosis Pathway. *Biochem. Biophysical Res. Commun.* 513, 663–668. doi:10.1016/j.bbrc.2019.04.023

Li, Y., Zhao, Y., Hu, J., Xiao, J., Qu, L., Wang, Z., et al. (2013). A Novel ER-Localized Transmembrane Protein, EMC6, Interacts with RAB5A and Regulates Cell Autophagy. *Autophagy* 9, 150–163. doi:10.4161/auto.22742

Lin, Y., Jiang, M., Chen, W., Zhao, T., and Wei, Y. (2019). Cancer and ER Stress: Mutual Crosstalk between Autophagy, Oxidative Stress and Inflammatory Response. *Biomed. Pharmacother.* 118, 109249. doi:10.1016/j.biopha.2019.109249

Liu, C. Y., Hsu, C. C., Huang, T. T., Lee, C. H., Chen, J. L., Yang, S. H., et al. (2018). ER Stress-Related ATF 6 Upregulates CIP 2A and Contributes to Poor Prognosis of colon Cancer. *Mol. Oncol.* 12, 1706–1717. doi:10.1002/1878-0261.12365

Liu, F., Chang, L., and Hu, J. (2020). Activating Transcription Factor 6 Regulated Cell Growth, Migration and Inhibiteds Cell Apoptosis and Autophagy via MAPK Pathway in Cervical Cancer. *J. Reprod. Immunol.* 139, 103120. doi:10.1016/j.jri.2020.103120

Liu, N., Sun, Y.-Y., Zhang, X.-W., Chen, S., Wang, Y., Zhang, Z.-X., et al. (2015). Oncogenic miR-23a in Pancreatic Ductal Adenocarcinogenesis via Inhibiting APAF1. *Dig. Dis. Sci.* 60, 2000–2008. doi:10.1007/s10620-015-3588-x

Martinez-Useros, J., Georgiev-Hristov, T., Borrero-Palacios, A., Fernandez-Aceñero, M. J., Rodríguez-Remírez, M., Del Puerto-Nevado, L., et al. (2015). Identification of Poor-Outcome Biliopancreatic Carcinoma Patients with Two-Marker Signature Based on ATF6α and P-P38 "STARD Compliant". *Medicine* 94, e1972. doi:10.1097/MD.0000000000001972

Mizrahi, J. D., Surana, R., Valle, J. W., and Shroff, R. T. (2020). Pancreatic Cancer. *The Lancet* 395, 2008–2020. doi:10.1016/S0140-6736(20)30974-0

Mohamed, E., Cao, Y., and Rodriguez, P. C. (2017). Endoplasmic Reticulum Stress Regulates Tumor Growth and Anti-tumor Immunity: a Promising Opportunity for Cancer Immunotherapy. *Cancer Immunol. Immunother.* 66, 1069–1078. doi:10.1007/s00262-017-2019-6

Morris, J. P., YashinskieYashinskie, J. J., Koche, R., Chandwani, R., Tian, S., Chen, C.-C., et al. (2019). α-Ketoglutarate Links P53 to Cell Fate during Tumour Suppression. *Nature* 573, 595–599. doi:10.1038/s41586-019-1577-5

Nikesitch, N., Tao, C., Lai, K., Killingsworth, M., Bae, S., Wang, M., et al. (2016). Predicting the Response of Multiple Myeloma to the Proteasome Inhibitor Bortezomib by Evaluation of the Unfolded Protein Response. *Blood Cancer J.* 6, e432. doi:10.1038/bcj.2016.40

O'Reilly, E. M., and Ferrone, C. (2020). Neoadjuvant or Adjuvant Therapy for Resectable or Borderline Resectable Pancreatic Cancer: Which Is Preferred. *J. Clin. Oncol.* 38, 1757–1759. doi:10.1200/JCO.19.03318

Piffoux, M., Eriau, E., and Cassier, P. A. (2021). Autophagy as a Therapeutic Target in Pancreatic Cancer. *Br. J. Cancer* 124, 333–344. doi:10.1038/s41416-020-01039-5

Sakamoto, H., Attiyeh, M. A., Gerold, J. M., Makohon-Moore, A. P., Hayashi, A., Hong, J., et al. (2020). The Evolutionary Origins of Recurrent Pancreatic Cancer. *Cancer Discov.* 10, 792–805. doi:10.1158/2159-8290.CD-19-1508

Shen, X., Kan, S., Hu, J., Li, M., Lu, G., Zhang, M., et al. (2016). EMC6/TMEM93 Suppresses Glioblastoma Proliferation by Modulating Autophagy. *Cel. Death Dis.* 7–e2043. doi:10.1038/cddis.2015.408

Shiraishi, H., Okamoto, H., Yoshimura, A., and Yoshida, H. (2006). ER Stress-Induced Apoptosis and Caspase-12 Activation Occurs Downstream of Mitochondrial Apoptosis Involving Apaf-1. *J. Cel. Sci.* 119 (Pt 19), 3958–3966. doi:10.1242/jcs.03160

Sicari, D., Fantuz, M., Bellazzo, A., Valentino, E., Apollonio, M., Pontisso, I., et al. (2019). Mutant P53 Improves Cancer Cells' Resistance to Endoplasmic Reticulum Stress by Sustaining Activation of the UPR Regulator ATF6. *Oncogene* 38, 6184–6195. doi:10.1038/s41388-019-0878-3

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

Tan, J.-h., Cao, R.-c., Zhou, L., Zhou, Z.-t., Chen, H.-j., Xu, J., et al. (2020). EMC6 Regulates Acinar Apoptosis via APAF1 in Acute and Chronic Pancreatitis. *Cel. Death Dis.* 11 (11), 966. doi:10.1038/s41419-020-03177-3

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Res.* 47 (W1), W556–W560. doi:10.1093/nar/gkz430

Thakur, G., Kumar, R., Kim, S.-B., Lee, S.-Y., Lee, S.-L., and Rho, G.-J. (2021). Therapeutic Status and Available Strategies in Pancreatic Ductal Adenocarcinoma. *Biomedicines* 9, 178. doi:10.3390/biomedicines9020178

Wang, X., Xia, Y., Xu, C., Lin, X., Xue, P., Zhu, S., et al. (2017). ER Membrane Protein Complex Subunit 6 (EMC6) Is a Novel Tumor Suppressor in Gastric Cancer. *BMB Rep.* 50, 411–416. doi:10.5483/bmbrep.2017.50.8.065

Zhang, H. H., Li, C., Ren, J. W., Liu, L., Du, X. H., Gao, J., et al. (2021). OTUB1 Facilitates Bladder Cancer Progression by Stabilizing ATF6 in Response to Endoplasmic Reticulum Stress. *Cancer Sci.* 112, 2199–2209. doi:10.1111/cas.14876

Zhang, R., Piao, M. J., Kim, K. C., Kim, A. D., Choi, J.-Y., Choi, J., et al. (2012). Endoplasmic Reticulum Stress Signaling Is Involved in Silver Nanoparticles-Induced Apoptosis. *Int. J. Biochem. Cel. Biol.* 44, 224–232. doi:10.1016/j.biocel.2011.10.019

Zheng, Z., Nayak, L., Wang, W., Yurdagul, A., Jr, Wang, X., Cai, B., et al. (2019). An ATF6-tPA Pathway in Hepatocytes Contributes to Systemic Fibrinolysis and Is Repressed by DACH1. *Blood* 133, 743–753. doi:10.1182/blood-2018-07-864843

Zhou, L., Tan, J.-h., Cao, R.-c., Xu, J., Chen, X.-m., Qi, Z.-c., et al. (2019). ATF6 Regulates the Development of Chronic Pancreatitis by Inducing P53-Mediated Apoptosis. *Cel. Death Dis.* 10 (9), 662. doi:10.1038/s41419-019-1919-0

Zhou, L., Tan, J.-h., Zhou, W.-y., Xu, J., Ren, S.-j., Lin, Z.-y., et al. (2020). P53 Activated by ER Stress Aggravates Caerulein-Induced Acute Pancreatitis Progression by Inducing Acinar Cell Apoptosis. *Dig. Dis. Sci.* 65, 3211–3222. doi:10.1007/s10620-020-06052-5

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership