frontiers | Research Topics

Cancer systems biology

Edited by
Jesús Espinal-Enríquez, Claudia Rangel-Escareño and Frank Emmert-Streib

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-83250-934-0
DOI 10.3389/978-2-83250-934-0

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

Cancer systems biology

Topic editors

Jesús Espinal-Enríquez — Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Claudia Rangel-Escareño — Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Frank Emmert-Streib — Tampere University, Finland

Citation

Espinal-Enríquez, J., Rangel-Escareño, C., Emmert-Streib, F., eds. (2022). *Cancer systems biology*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83250-934-0

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.

Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 18 November 2022
DOI 10.3389/fgene.2022.1083902

OPEN ACCESS

EDITED AND REVIEWED BY
Jared C. Roach,
Institute for Systems Biology (ISB),
United States

*CORRESPONDENCE
Jesús Espinal-Enríquez,
jespinal@inmegen.gob.mx
Claudia Rangel-Escareño,
crangel@inmegen.gob.mx
Frank Emmert-Streib,
frank.emmert-streib@tuni.fi

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

RECEIVED 29 October 2022
ACCEPTED 04 November 2022
PUBLISHED 18 November 2022

CITATION
Espinal-Enríquez J, Rangel-Escareño C and Emmert-Streib F (2022), Editorial: Cancer systems biology.
Front. Genet. 13:1083902.
doi: 10.3389/fgene.2022.1083902

Editorial: Cancer systems biology

Jesús Espinal-Enríquez[1]*, Claudia Rangel-Escareño[1,2]* and Frank Emmert-Streib[3]*

[1]Computational Genomics Division, National Institute of Genomic Medicine, Mexico City, Mexico, [2]School on Engineering and Sciences, Tecnologico de Monterrey Epigmenio González, Santiago de Querétaro, Mexico, [3]Predictive Society and Data Analytics Lab, Tampere University, Tampere, Finalnd

KEYWORDS

cancer systems biology, cancer Metabolism, biomarkers in cancer, cancer immunobiology, non-coding regulation in cancer

Editorial on the Research Topic
Cancer systems biology

Introduction

Cancer has been a paradigmatic example of a complex disease. It is the second leading cause of death worldwide. Knowledge acquired on the mechanisms and potential causes underlying the origins and progression of cancer from single-scale studies have led to the development of personalized treatments, which in turn have decreased mortality rates in several types of cancer. However, because it is not a single and isolated disease, but rather a heterogeneous set of multi-scale alterations, current research requires computational and systems biology approaches that deal with complex systems on gene regulatory networks. Multi-omics is one of these approaches. Aimed at developing multi-scale mathematical modeling, it allows us to integrate the dynamics of biological perturbations for different types of cancer. Moreover, if tissue/organ-specific path is studied we are able to create mechanistic models for a particular type of the disease. That, combined with current low-cost high throughput technologies have made possible massive cohorts of -omics and clinical data to be publicly available. This immeasurable amount of information promotes harnessing data science approaches for the analysis, integration, and mining of multi-omics data. Large and multidisciplinary groups can now focus on creating models and provide theoretical frameworks that describe cancer in a variety of ways. Besides cancer cell transformation or cancer evolution, currently single cell genomics and spatial single cell transcriptomics is moving forward with deeper understanding of the disease. Also, cancer driver mutations and mutational processes, development of target-specific treatments have been leading to a prediction of drug responses. Notwithstanding, several challenges that remain to be undertaken such as mechanisms of metastasis, resistance to treatment, intra-tumoral heterogeneity, molecular, cellular, and metabolic changes during progression stages, epigenetic modifications, to mention a few. This cross-disciplinary interaction at these levels of description have contributed to develop more efficient models that improve the predictive capacity and ultimately help clinicians and medical scientists in the treatment and therapies.

The aim of this Research Topic is to discuss and explore state-of-the-art research on cancer systems biology. Throughout this Research Topic of articles, several aspects of cancer systems biology are covered. On the one hand, the cancer-specific point of view, which tries to identify molecules, pathways, mechanisms that generate or alter a concrete type of cancer, in one tissue. On the other hand, the multi-cancer point of view, which tries to observe generalities, common aspects, or the widely mentioned hallmarks of cancer. Also, several bioinformatics tools were implemented to dissect and discriminate between groups and to observe specific patterns. Among these computational tools we can find language-specific or web-based bioinformatic packages such as TIMER, GEPIA2, LinkedOmics, GSCALite, GSEA, CIBERSORT, ESTIMATE, WGCNA ARACNe, or Infotheo, among others. Tissue-specific approach was widely used in mainly four applications: cancer metabolism, biomarkers and prognosis, immune system-related research and, non-coding regulation. A highlight of this Research Topic is that the majority of manuscripts focus on establishing a better prognosis or a more accurate classification of cancer groups. We summarize the 15 articles accepted for publication in the systems biology Research Topic.

Cancer metabolism

Zhu et al. studied the relationships between glycolysis, tumor microenvironment, and therapeutic response in esophageal adenocarcinoma (EAC). By classifying into low and high risk groups and based on a glycolysis-related genes signature, these authors established a correlation between glycolytic and ATP/ADP metabolic pathways and a poor overall survival. Wang et al. systematically analyzed the expression patterns of lysyl oxidase family genes in gastric cancer. They found that LOX and LOX2 may have a role in tumor prognosis and therefore, in gastric cancer therapy. Sun et al. investigated the association between COX-2 −1195G/A single nucleotide polymorphism and lung cancer susceptibility in a Japanese population. By evaluating the genotype distribution of COX-2 −1195G/A with a PCR-restriction fragment length polymorphism assay for 330 lung cancer patients and 162 healthy controls, they show that COX-2 −1195G/A does not have a relationship with the risk of developing lung cancer. However, homozygous COX-2–1195A genotype confers an increased risk for lung squamous cell carcinoma in Japanese individuals and could be used as a predictive factor for early detection of lung squamous cell carcinoma.

Biomarkers and prognosis

Regarding biomarkers and prognosis, Chen et al. evaluated the correlation between NCKAP1 expression and clinical features of clear cell renal cell carcinoma. These authors found that overexpression of NCKAP1 in ACHN cells reduced proliferation, invasion and migration capacity *in vitro* and inhibited tumor growth *in vivo*. Hu et al. explored the molecular mechanism of LYPD3 in the regulation during transformation and throughout the development of acute myeloid leukemia providing a research basis for the screening of markers related to the treatment and prognosis. They specifically suggest, by means of a dataset analysis and gene knockdown mediated by small interfering RNA (siRNA), that LYPD3 participates in the development of AML through the p53 signaling pathway or/and PI3K/AKT signaling pathway. Zhou et al. assessed the diagnostic and prognostic significance of ATP binding cassette subfamily C (ABCC) genes in hepatocellular carcinoma (HCC). ABCC1,4,5 were found to be positively associated with infiltration of multiple immune cells, while ABCC6 was found to be the opposite. In conclusion, they suggest that ABCC1, ABCC4, ABCC5, and ABCC6 might be prognostic biomarkers in HCC. Zhang et al. explored the expression and carcinogenic effect of keratin 17 (KRT17) in human tumors. They show that KRT17 was highly expressed in most tumors (such as esophageal cancer, lung cancer, cervical cancer, *etc.*), and the high expression level correlated with tumor stage and prognosis. Zheng et al. evaluated the relationship between EFNA3 and gastric cancer (GC) prognosis and tumor-infiltrating lymphocytes. The authors suggest based on bioinformatics analyses, that EFNA3 participates in changes in GC immune checkpoint markers in a co-linear manner. Turns out that EFNA3 expression in HGC-27, AGS, MKN45, and NCI-N87 cell lines was higher than that in GES-1 and observed that patients with high expression of EFNA3 had a worse prognosis.

Cancer and immune system

Among the immune-system related works, Zhang S. et al. constructed a prognostic model for the response to immunotherapy in thyroid carcinoma. The authors show that patients with high tumor mutational burden and low PD-L1 expression levels might respond poorly to immunotherapy. Yang S. et al. identified three different immune cell infiltration signatures. The one with the highest risk was characterized by an enhanced activation of the immune system as well as a significantly high tumor mutational burden. Xie et al. evaluated the thymidine kinase 1 (TK1) role in prostate cancer databases with results validated for *in vitro* and *in vivo* models. They showed that TK1 is a prognostic predictor correlated with poor outcomes of PCa patients. Moreover, TK1 inactivation can significantly restrain tumor growth. Zhang Y. et al. constructed a prognostic model based on pyroptosis-related genes to provide new insights into the prognosis of osteosarcoma patients. Based on a pyroptosis-related signature score, they were able to differentiate patients with high and low risk of metastasis.

Non-coding regulation

In terms of gene expression correlating with other omics, Hernández-Gómez et al. evaluated the role that CNVs may play in shaping gene co-expression patterns in luminal B breast cancer. The authors construct a conditional mutual-information-based network for gene expression and copy number alterations. They report that, for luminal B breast tumors, the co-expression program is not necessarily determined by its CNV structure. Additionally, by analyzing the network topology, they suggested that MAF1 and POLR3D may constitute an axis of regulation of gene transcription, in particular for non-coding RNA species. Methylation is a widely known mechanism involved in cancer development. Shi et al. analyzed the m6A RNA methylation in several types of cancer. The results revealed that the m6A regulators exhibited widespread dysregulation, genetic alteration, and the modulation of oncogenic pathways across cancer types. In addition, most of the m6A regulators were relevant for prognosis in many cancer types. It has been recently shown that long non-coding RNAs play a crucial role in the development of several cancer types. Yang X. et al. developed a pyroptosis-related lncRNA signature to evaluate and predict overall survival in breast cancer. The risk model comprised 10 pyroptosis-related lncRNAs, and was identified as an independent predictor for overall survival (OS). The low-risk group had a higher expression of immune checkpoint markers and exhibited higher fractions of activated immune cells, while the high-risk group had a higher percentage of TMB. A validation on a separate cohort of breast cancer samples found that RP11-459E5.1 was significantly upregulated, while RP11-1070N10.3 and RP11-817J15.3 were downregulated and significantly associated with worse OS.

Altogether, this Research Topic of 15 articles provides a broad overview of the current status of cancer systems biology, showing promising advances of cancer research through the application of computational approaches.

Author contributions

JE-E drafted the first vesion of the manuscript. JE-E, CR-E, and FE-S made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was supported by the National Institute of Genomic Medicine, México (JE-E and CR-E).

Acknowledgments

The guest editors acknowledge the peer reviewers who made this Research Topic possible with their timely, insightful, and critical comments for improving the manuscripts in this Research Topic.

Conflict of interest

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

Publisher's note

frontiers in Genetics

ORIGINAL RESEARCH
published: 25 November 2021
doi: 10.3389/fgene.2021.777931

Identification of the Immune Cell Infiltration Landscape in Hepatocellular Carcinoma to Predict Prognosis and Guide Immunotherapy

Shiyan Yang[1,2†], Yajun Cheng[3†], Xiaolong Wang[4†], Ping Wei[5], Hui Wang[6] and Shanzhong Tan[1]*

[1]Department of Integrated TCM and Western Medicine, Nanjing Hospital Affiliated to Nanjing University of Chinese Medicine, Nanjing, China, [2]Department of Gastroenterology, Huaian Hospital Affiliated to Xuzhou Medical University, Huaian, China, [3]Department of Gastroenterology, People's Hospital of Lianshui, Huaian, China, [4]The Department of General Surgery, Tumor Hospital of Huaian, Huaian, China, [5]The Department of Ultrasound, Huaian Hospital Affiliated to Xuzhou Medical University, Huaian, China, [6]The Department of Rehabilitation Medicine, Huaian Hospital Affiliated to Xuzhou Medical University, Huaian, China

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez, Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Zaoqu Liu, First Affiliated Hospital of Zhengzhou University, China
Michael Poidinger, Royal Children's Hospital, Australia

***Correspondence:**
Shanzhong Tan
fsyy01455@njucm.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: 16 September 2021
Accepted: 01 November 2021
Published: 25 November 2021

Citation:
Yang S, Cheng Y, Wang X, Wei P, Wang H and Tan S (2021) Identification of the Immune Cell Infiltration Landscape in Hepatocellular Carcinoma to Predict Prognosis and Guide Immunotherapy. Front. Genet. 12:777931. doi: 10.3389/fgene.2021.777931

Background: Globally, hepatocellular carcinoma (HCC) is the sixth most frequent malignancy with a high incidence and a poor prognosis. Immune cell infiltration (ICI) underlies both the carcinogenesis and immunogenicity of tumors. However, a comprehensive classification system based on the immune features for HCC remains unknown.

Methods: The HCC dataset from The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC) cohorts was used in this study. The ICI patterns of 571 patients were characterized using two algorithms: the patterns were determined based on the ICI using the ConsensusClusterPlus package, and principal component analysis (PCA) established the ICI scores. Differences in the immune landscape, biological function, and somatic mutations across ICI scores were evaluated and compared, followed by a predictive efficacy evaluation of ICI scores for immunotherapy by the two algorithms and validation using an external immunotherapy cohort.

Results: Based on the ICI profile of the HCC patients, three ICI patterns were identified, including three subtypes having different immunological features. Individual ICI scores were determined; the high ICI score subtype was characterized by enhanced activation of immune-related signaling pathways and a significantly high tumor mutation burden (TMB); concomitantly, diminished immunocompetence and enrichment of pathways associated with cell cycle and RNA degradation were found in the low ICI score subtype. Taken

Abbreviations: CDF: Cumulative distribution function; CR: Complete response; DEGs: differentially expressed genes; FDR: false discovery rate; FKPM: Fragments Per Kilobase per Million; GO: Gene Ontology; GSEA: Gene set enrichment analysis; HCC: Hepatocellular carcinoma; ICGC: International Cancer Genome Consortium; ICI: Immune Cell Infiltration; KEGG: Kyoto Encyclopedia of Genes and Genomes; PD: Progressive disease; PR: Partial response; SD: Static disease; TME: Tumor Microenvironment; TCGA: The Cancer Genome Atlas; TPM: Transcripts Per Million; TMB: Tumor Mutation Burden; TIDE: Tumor Immune Dysfunction and Exclusion.

together, our results contribute to a better understanding of an active tumor and plausible reasons for its poor prognosis.

Conclusion: The present study reveals that ICI scores may serve as valid prognostic biomarkers for immunotherapy in HCC.

Keywords: carcinoma, immune cell infiltration landscape, immunotherapy, ICI scores, prognosis

INTRODUCTION

HCC is an aggressive malignancy that frequently develops and progresses in the setting of chronic liver disease or cirrhosis (Park et al., 2015). Statistics from 2018 indicate that HCC is the sixth most frequently occurring malignancy and the fourth highest cause of cancer-related deaths (Kulik and El-Serag, 2019). To date, approximately 841,000 new cases are registered and more than 782,000 HCC-related deaths are recorded (Singal et al., 2020). Alcohol consumption, obesity, fatty liver, and hepatitis infection are some of the important risk factors for HCC (Caruso et al., 2021). Current advances in HCC-diagnosis, surgical treatment, transplantation, chemotherapy, radiotherapy, and targeted molecular therapies, to some extent, have improved the prognosis of HCC patients (Fan, 2012), but the majority of the diagnosed patients are already at an advanced stage and have only limited conservative treatment options. The rate of cure in HCC remains low due to its high malignancy, recurrence rate, increased metastasis, and adverse response to chemotherapy (Pillai et al., 2020; Farzaneh et al., 2021).

As a treatment for HCC, despite its limited efficacy, immunotherapy has yielded promising results (Silva et al., 2020). However, the benefits of immunotherapy are largely limited to only a small number of HCC patients. Existing studies have shown that immune-associated genes and lymphocytes infiltrating tumors play a key role in tumor oncogenesis and its progression (Wang et al., 2020); the dynamic interactions between immune cells infiltration into the tumor microenvironment, cytokines secreted by immune cell types, and cancerous cells are involved in HCC tumor progression (Choi and Park, 2017; Sachdeva and Arora, 2020). A clearer understanding of these specific dynamical patterns may be beneficial for immunotherapy. Therefore, detailed investigations of the immune landscape of the tumor microenvironment (TME) and identification of ideal HCC subgroups for immunotherapy are important to improve the immunotherapeutic responses and prognostic prediction (Hosseinzadeh et al., 2018; Robert et al., 2020).

Extensive studies on the TME indicate the critical functions of infiltrating immune cells in tumor dissemination, recurrence, metastatic activity, and immunotherapeutic responses (Jiang et al., 2018a; Zeng et al., 2018). As an example, $CD8^+$ T cells are potent regulators of adaptive immunity as they can eliminate pathogen-infected and tumor cells (Stairiker et al., 2020), and thus, critically affect tumor immunity (Han et al., 2020). Tumor-associated macrophages (TAM) exert multiple tumor-beneficial effects through the secretion of immunosuppressive cytokines, associated with unfavorable prognoses (De Palma and Lewis, 2013; Noy and Pollard, 2014). Through their inhibitory activity, M2-type macrophages critically regulate the tumor microenvironment (Mehla and Singh, 2019). Taken together, these studies suggest that immune cell interactions in TME may provide new insights for cancer therapy. However, a comprehensive and clear understanding of immune landscape complexity in HCC is still lacking.

Here, we evaluated the immune landscape of HCC using the CIBERSOFT algorithm. Based on their ICI features, the HCC patients were classified into four subtypes. Subsequently, based on immune subtypes, ICI scores were established to further assess the immune landscape of HCC, for accurate prognostic prediction of the patients and their immunotherapeutic responses.

MATERIALS AND METHODS

Hepatocellular Carcinoma Sample and Data Acquisition

Patients with complete clinical information (Stage, Follow-up Information, Age and Gender) were selected in this study, after removing patients who did not meet the criteria. RNA-Seq data of 340 HCC patients and their complete corresponding clinical information were acquired from The Cancer Genome Atlas (TCGA) using the GDC API; for the training cohort, expression data in FPKM (Fragments Per Kilobase per Million) were subsequently TPM-normalized (Transcripts Per Kilobase per Million). In addition, RNA-Seq data of 231 HCC patients and their complete corresponding clinical information were obtained from the International Cancer Genome Consortium (ICGC) database (Fujimoto et al., 2016). Similarly, for the validation set, the raw sequencing data were TPM normalized.

Evaluation of Immune Infiltration Levels and Consensus Clustering

The level of infiltration of the 22 immune cells was quantified for each sample of the HCC-TCGA cohort using the "CIBERSORT" R package with the LM22 signature (Newman et al., 2015). Next, the ESTIMATE algorithm was used to compute the scores for immune and stromal characteristics for each patient (Yoshihara et al., 2013). Hierarchical consensus clustering for HCC was performed for each sample based on the individual pattern of ICI. In this analysis, the PAM unsupervised clustering based on Pearson's correlation and Ward's linkage based on the "ConsensusClusterPlus" R package, were used (Yu et al., 2012)

and repeated 1,000 times to reduce sampling errors and ensure a stable classification. Consensus clustering is a popular bioinformatics algorithm, which was extensively utilized in cancer-related studies (Liu et al., 2021a; Liu et al., 2021b; Liu et al., 2021c; Liu et al., 2021d).

DEGs Identified Based on ICI Phenotype

Patients were subdivided based on ICI and were referred to as the ICI subgroups. Subsequently, differentially expressed genes between subgroups were analyzed using the "limma" package, and genes associated with the ICI patterns were identified. Significance criteria of p. adjust < 0.05 and | Log2FC | > 1 were set to identify the significant DEGs among the different ICI subgroups.

Dimensionality Reduction and the Construction of ICI Scores

The ICI scores were constructed following the work of Zhang et al. (2020). First, to classify the patients in the training set based on DEGs, an unsupervised clustering method was used; the positively and negatively correlated DEGs with the clustering features were called ICI gene signatures A and B, respectively. Second, the dimensionality reduction of gene signatures A and B based on ICI was performed using the Boruta algorithm, followed by subsequent extraction of the signature score (corresponding to principal component 1) using the PCA algorithm. Finally, the computation of ICI scores for each patient was according to the following equation:

$$\textbf{ICI score} = \sum \mathbf{PC1_A} - \sum \mathbf{PC1_B}$$

Somatic Mutations in the The Cancer Genome Atlas Cohort

The corresponding data for the patient mutations in the HCC-TCGA cohort were collated on the Mutect2 platform and were downloaded using the "TCGAbiolinks" package (Colaprico et al., 2016). The total number of nonsynonymous mutations in the samples was calculated to compare the differences in the mutation burdens between the two ICI score-based subgroups. Subsequently, using the "maftool" in R, the top 25 driver genes having the highest mutation frequency were identified and the mutation differences in the driver genes between the high- and low-score subgroups were compared (Mayakonda et al., 2018).

Immunotherapeutic Responses of ICI Subgroup

Since different ICI subgroups may have different sensitivities to immunotherapy, the TIDE (http://tide.dfci.harvard.edu/) algorithm was used to predict the anti-PD1 and anti-CTLA4 treatment responses of patients in the TCGA and ICGC cohorts (Jiang et al., 2018b; Fu et al., 2020). Subsequently, with the aid of unsupervised subclass mapping (https://cloud.genepattern.org/gp/) (Hoshida et al., 2007), data from the high- and low-score subgroups were compared to a published dataset consisting of 47 patients' responses to anti-PD1 and anti-CTLA4 treatments (Roh et al., 2017). This analysis predicted the immunotherapeutic responses of the high- and low-subgroups; FDR < 0.05 was set as the threshold for a significant response to anti-PD1 and anti-CTLA4 treatment. Additionally, the independent dataset IMvigor210 was used to analyze the predictive efficacy of ICI scores. The IMvigor210 dataset consisting of 298 cases of uroepithelial carcinoma samples and their corresponding clinical information, were obtained from the freely available, fully documented software and data packages under the Creative Commons 3.0 Attribution License, available at http://research-pub.gene.com/IMvigor210CoreBiologies.

Statistical Analysis

All statistical analyses and plotting were performed using R software (version 4.04). For comparisons of more than two groups, the Kruskal-Wallis test was used, else, we used the Wilcoxon test. For the subgroups in each data set, the Kaplan-Meier plotter generated the survival curves, and the log-rank tests were determined any statistically significant differences. The correlations between ICI score for the subgroups and associated somatic mutation frequencies were evaluated and analyzed by the chi-square test. Unless stated, $p < 0.05$ (two-tailed) was considered to be statistically significant.

RESULTS

Immune cell infiltrationICI Landscape in the TCGA Cohort

Supplementary Figure S1 displayed a brief flow chart of this study. The execution of the CIBERSORT algorithm quantified the activity or enrichment of immune cells in the HCC tumor tissues (**Figure 1A**, **Supplementary Table S1**). Based on the 340 tumor samples and their corresponding ICI features in the training set, the ConsensusClusterPlus package of R software executed the unsupervised clustering method. Thus, we classified the HCC patients into three different ICI subtypes.

There were significant survival differences among the subtypes (log-rank test, $p < 0.0001$; **Supplementary Figure S2A–E**; **Figures 1B,C**); ICI cluster 1 was associated with a good prognosis while ICI cluster 3 had the worst prognosis. Additionally, to assess the intrinsic differences among the biological parameters underlying the different clinical phenotypes, ICI differences were compared between the three subgroups. ICI cluster 1 showed the highest infiltration of activated B cells, monocytes, and resting memory CD4 T cells, and the lowest infiltration of regulatory T cells. A more favorable prognosis of patients in ICI cluster 2 may be attributed to the high degree of infiltration of plasma cells, activated memory CD4T cells, M1 macrophages, and CD8T cells. However, in ICI cluster 3, higher infiltration of regulatory T cells, MO macrophages may underlie the poorest prognosis due to suppressed tumor immunity responses (**Figure 1D**). We also plotted the correlation heat maps to depict the interactions between immune cells in TME (**Figure 1E**). The expression

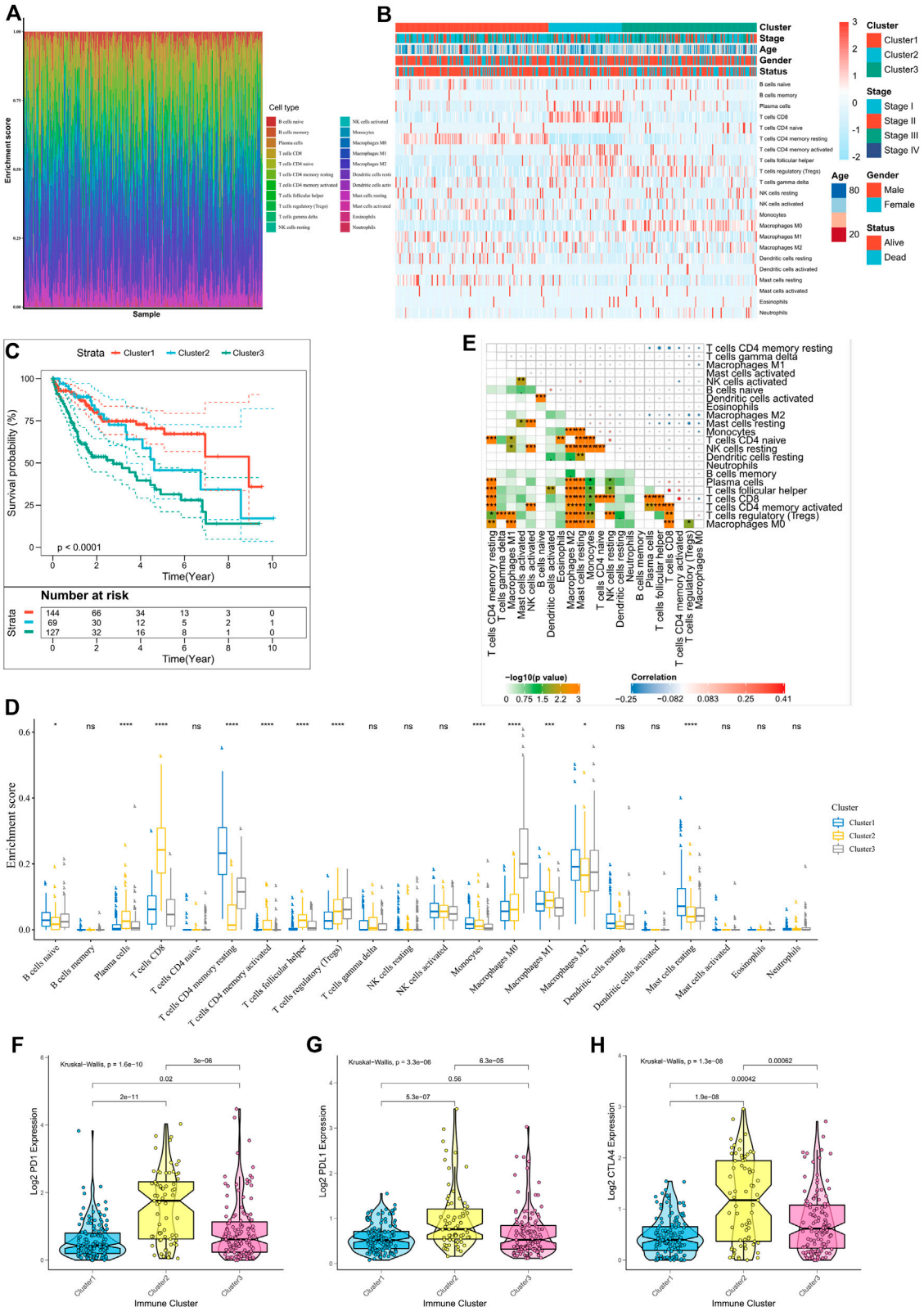


FIGURE 1 | Immune landscape in the TCGA cohort. **(A)** The immune landscape of 22 ICIs in HCC patients; **(B)** Unsupervised clustering of tumor-infiltrating immune cells in the TCGA cohort, where rows represent tumor-infiltrating immune cells and columns represent samples; **(C)** Kaplan-Meier curves for overall survival (OS) of patients with different ICI clusters, where log-rank $p = 0.018$; **(D)** Proportion of tumor-infiltrating immune cells in the three ICI clusters, where Kruskal-Wallis was used to test and compare the statistical differences of the three ICI clusters. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; **(E)** Cell interactions of tumor-infiltrating immune cell types. **(F–H)** Expression differences in PD-L1 **(F)**, PD1 **(G)**, and CTLA4 **(H)** between different ICI clusters (Kruskal-Wallis test, $p < 0.0001$).

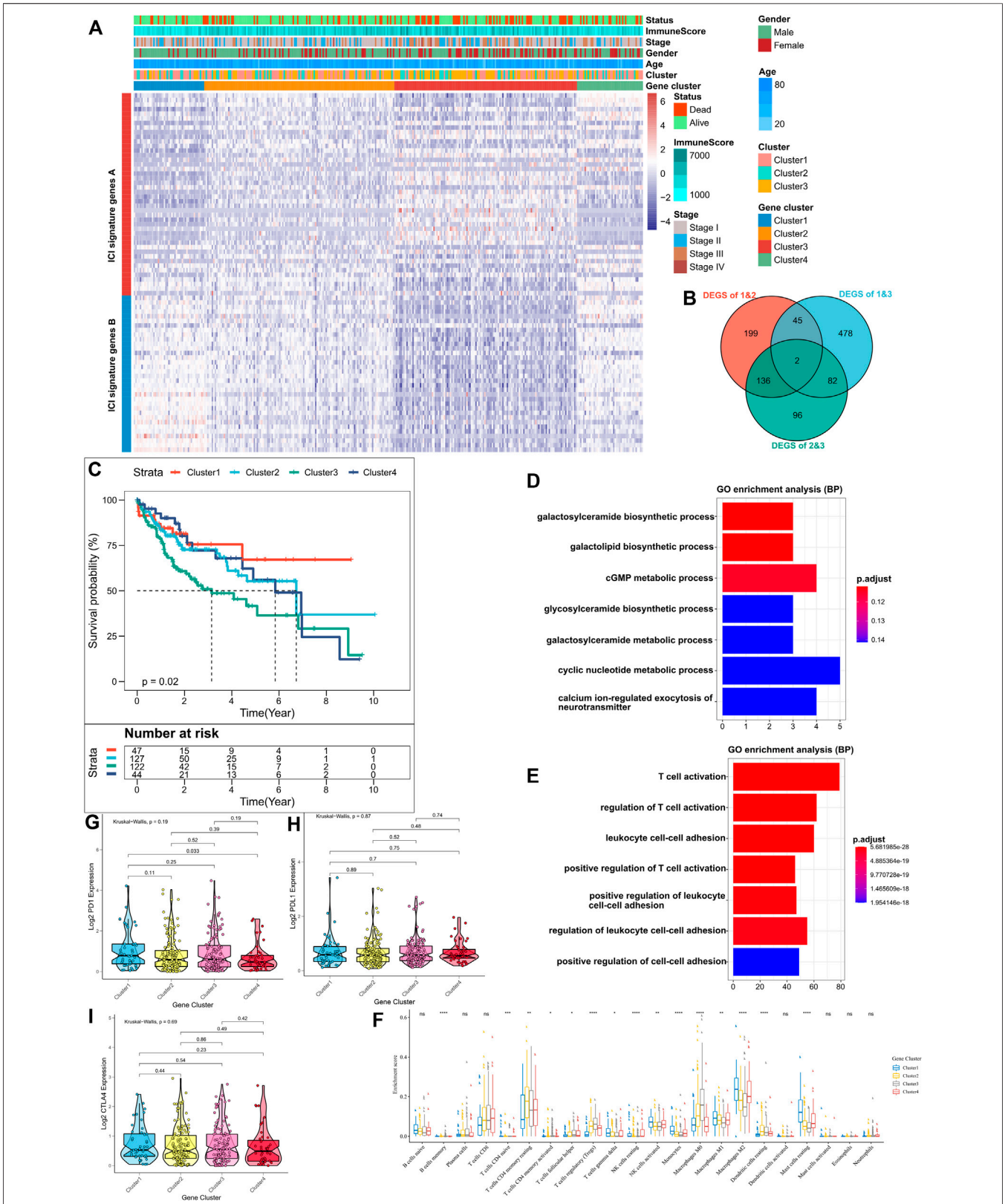


FIGURE 2 | Identification of Immune-Related Gene Subtypes. **(A)** Unsupervised clustering of common DEGs in the three ICI subgroups, dividing patients into four groups; **(B)** Number of DEGs among subgroups as shown by Venn diagram; **(C)** Kaplan-Meier curves for the three subgroups of patients, where the log-rank test shows

(*Continued*)

FIGURE 2 | an overall $p = 0.02$; **(D,E)** Gene ontology (GO) enrichment analysis of ICI-associated signature genes: ICI signature genes A **(D)** and B **(E)**, where x-axis indicates the number of genes in each GO term; **(F)** Proportion of tumor-infiltrating immune cells in the three gene clusters, where Kruskal-Wallis was used to test and compare the statistical differences of the three ICI clusters. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; **(G–I)** Expression differences of PD-L1 **(G)**, PD1 **(H)** and CTLA4 **(I)** using Kruskal-Wallis test.

differences in the different ICI subtypes, for three important immune checkpoints, PD1, PDL1, and CTLA4, were also analyzed. ICI cluster 2 had the highest levels of expression of immune checkpoint genes, while these were lowest in ICI cluster 1 (**Figures 1F–H**).

Identification of Immunogenic Subtypes

To better understand the underlying biological features of different immunophenotypes, differentially expressed genes (DEGs) between these subtypes were identified using the "limma" package of the R software. A total of 1,038 DEGs were identified (**Supplementary Table S2**), and their intersections are shown in the Venn diagram (**Figure 2B**). Subsequently, based on DEGs, the ConsensusClusterPlus package was executed for unsupervised clustering analysis; thus, the TCGA cohort was divided into four gene clusters (**Supplementary Figure S3A–F**); Positively associated 318 DEGs in the gene clusters were defined as ICI gene feature A, and the remaining were defined as ICI gene feature B (**Supplementary Table S3**). Moreover, to attenuate noise and gene redundancy, dimensionality reduction of ICI gene features A and B was performed using the Boruta algorithm. The transcriptional profiles of the 78 signature DEGs identified after dimensionality reduction are shown in the heat map (**Figure 2A**). The significantly enhanced biological processes among the DEGs are shown in **Figures 2D,E** and **Supplementary Table S4**. Kaplan-Meier analysis showed a significant difference in survival outcomes among the four subgroups ($p = 0.02$, **Figure 2C**). Patients in clusters 1 and 2 had a better prognosis as compared to those in cluster 3. The presence of higher infiltration levels of M1 macrophages, monocytes, gamma delta T cell, and lower infiltration levels of regulatory T cell in clusters 1 and 2, indicated that patients in these two clusters may have a stronger anti-tumor immune response (Biswas and Mantovani, 2010; Chen and Mellman, 2017). In contrast, the highest levels of infiltration of regulatory T cell and M0 macrophage, and lowest

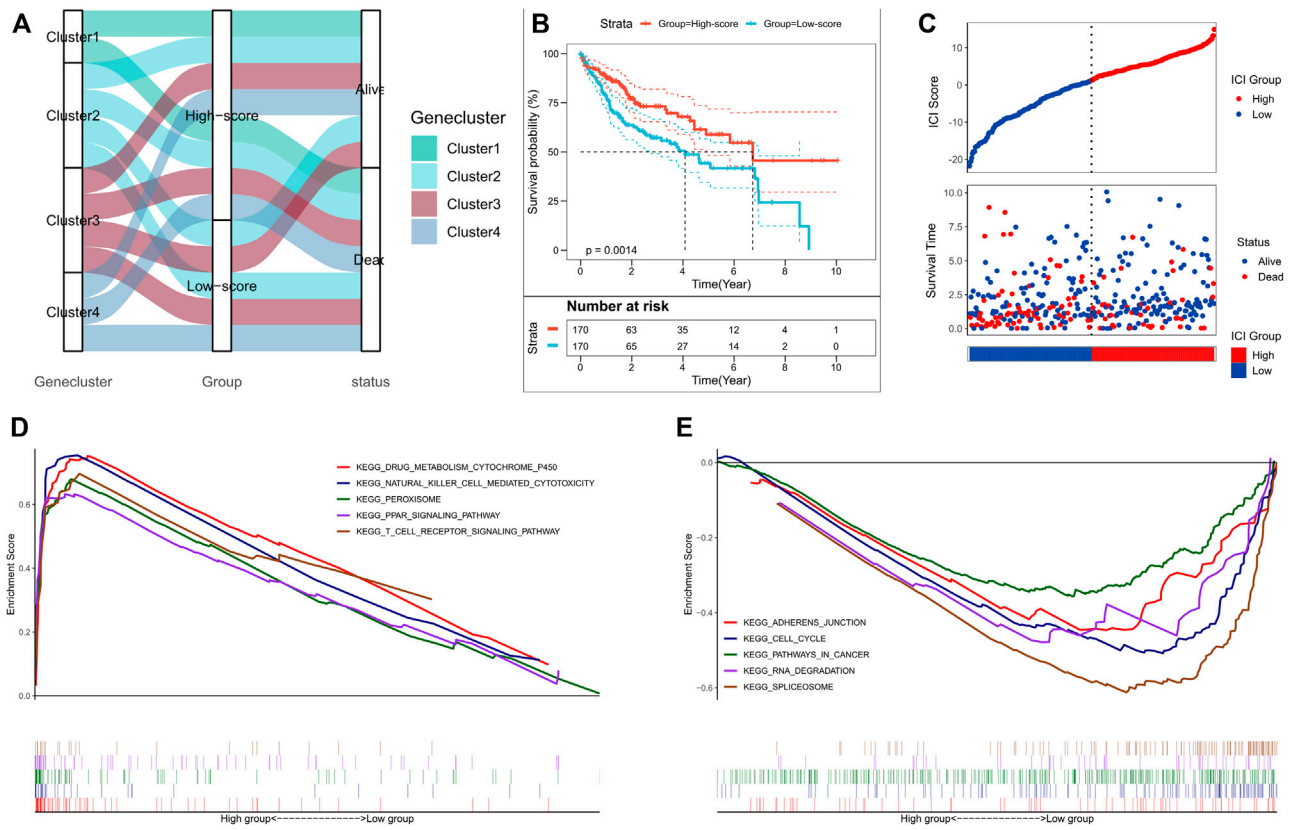


FIGURE 3 | Construction of the ICI scores. **(A)** Alluvial plot of the ICI gene cluster distribution in subgroups with different ICI clusters, ICI scores, and survival outcomes; **(B)** Kaplan-Meier curves for the high and low ICI score subgroups in the TCGA cohort, where $p = 0.0014$ for the log-rank test; **(C)** Survival status of patients in the high and low ICI score subgroups in the TCGA cohort; **(D)** GSEA enrichment results for the high ICI score subgroup; **(E)** GSEA enrichment results for the low ICI score subgroup.

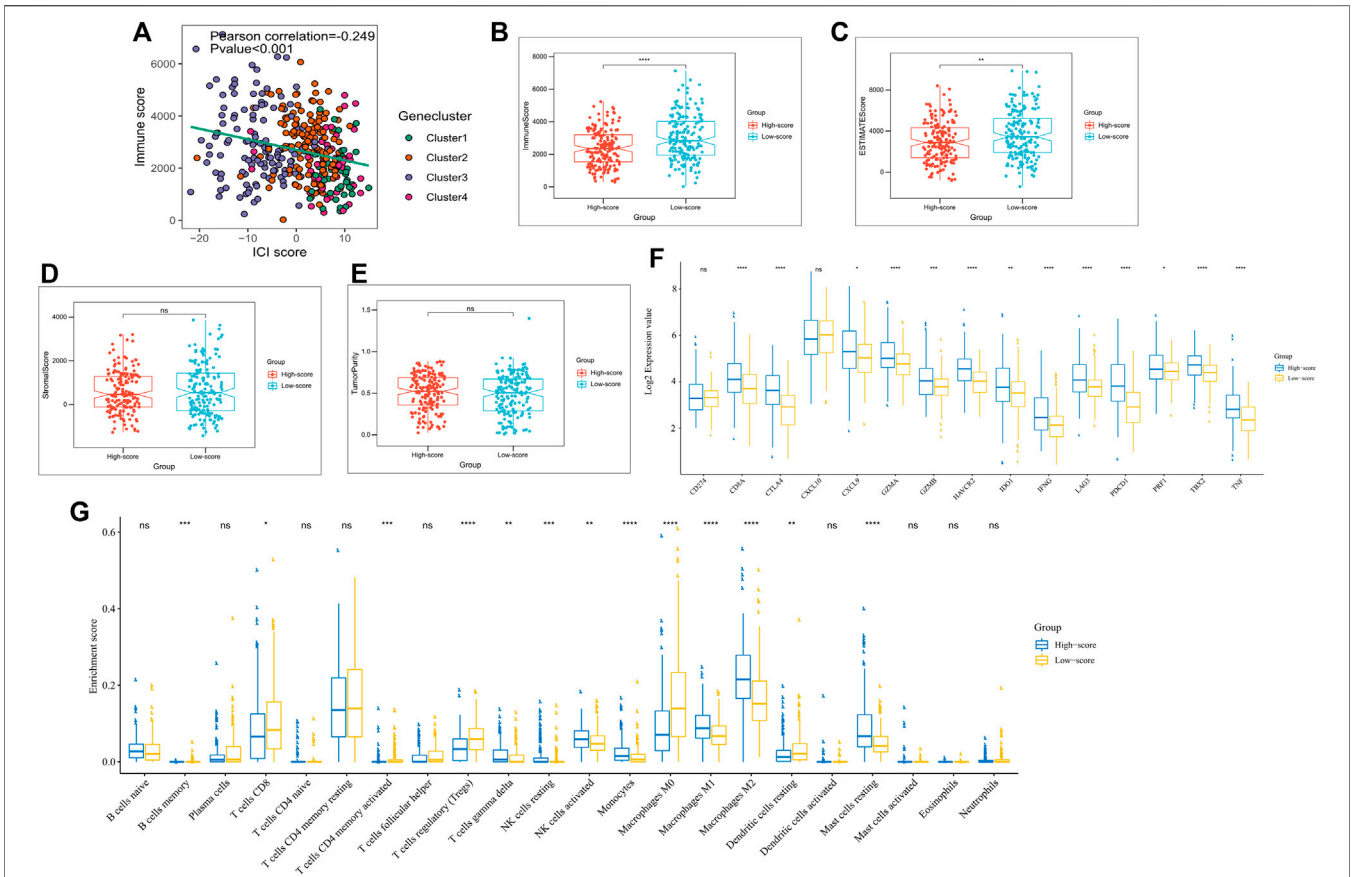


FIGURE 4 | Immune landscape of subgroups with different ICI scores. **(A)** Scatter plots depicting a negative correlation between ICI scores and immune scores in the TCGA cohort, showing a Pearson correlation between ICI scores and immune scores; **(B)** Expression of immune checkpoint-related genes (IDO1, CD274, HAVCR2, PDCD1, CTLA4, and LAG3) and immunoreactive-related genes (CD8A, CXCL10, CXCL9, GZMA, GZMB, PRF1, IFNG, TBX2, and TNF) in high and low ICI score subgroups; **(C)** ICI Proportion in different ICI score subgroups; **(D)** Immune scores of the high and low ICI score groups; **(E)** ESTIMATEscores of the high and low ICI score groups; **(F)** Stromal scores in the high and low ICI score subgroups; **(G)** Tumor purity in the high and low ICI score subgroups. *$p < 0.05$; **$p < 0.01$; ****$p < 0.0001$.

levels of infiltration of all other immune cells in cluster 3, suggested that this may be an immune desert subtype (Biswas and Mantovani, 2010; Chen and Mellman, 2017). The concordance between the immune profiles and the prognosis using different gene clusters suggested that our classification strategy is scientifically sound and reasonably good. The levels of PD1, PDL1, and CTLA4 expression among the four clusters, however, were not significantly different (**Figures 2F–H**).

Construction of the Immune Cell Infiltration Scores

PCA analysis was used to quantify the ICI status of HCC patients. We calculated the sum of ICI scores A from ICI signature gene A minus the sum of ICI score B from ICI signature gene B. Thus, the prognostic feature scores defined as ICI scores were obtained. Additionally, ICI scores for the validation cohort, from ICGC were calculated using the same gene signatures A and B and the algorithm as described above. Patients were divided into high- or low-score subgroups based on median ICI score, and the distribution of patients in the four clusters is shown in **Figure 3A**. Kaplan-Meier analysis showed a significant difference in the prognoses between the two groups; the high-score subgroup had the best prognosis ($p = 0.0014$, **Figure Figure3B**). The prognostic efficacy of ICI scores was also validated in the ICGC cohort ($p < 0.001$, **Supplementary Figure S4A**); the high-score subgroup patients had better survival outcomes in both the TCGA and ICGC cohorts (**Figure 3C**; **Supplementary Figure S4B**). GSEA analysis showed that NK cell-mediated cytotoxicity, T cell receptor signaling, and peroxisome-related pathways were substantially enriched in the high-score subgroup, while cancer-related, cell cycle, and RNA degradation pathways were substantially enriched in the low-score subgroup (**Figures 3D,E**; **Supplementary Table S5**).

Correlation of Immune Cell Infiltration Scores With Immune Landscape

The immunocompetence and stromal content of the TCGA cohort were quantified using the ESTIMATE algorithm. ICI scores and immune scores were negatively correlated (Pearson correlation:

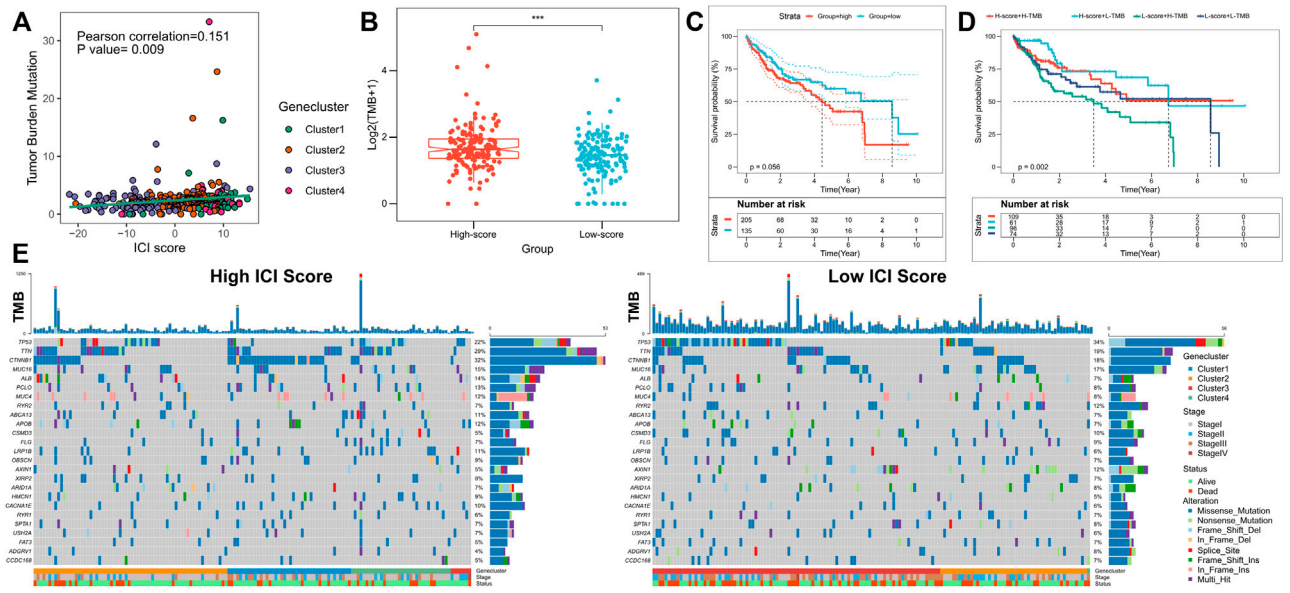


FIGURE 5 | Correlation between ICI scores and somatic cell variation. **(A)** Difference in TMB between the high and low ICI score subgroups, where $p < 0.001$ for Wilcoxon test; **(B)** Scatter plot depicting a positive correlation between ICI scores and the mutational burden in the TCGA cohort, showing the Pearson correlation between ICI scores and the mutational burden; **(C)** Kaplan-Meier curves for the high and low TMB score subgroups in the TCGA cohort, where $p = 0.056$ for the log-rank test; **(D)** Kaplan-Meier curves for patients in the TCGA cohort stratified by TMB and ICI scores, where $p = 0.002$ for the log-rank test; **(E)** "oncoplot"showing the high and low ICI score subgroups for the top 25 mutant genes, with each column representing one patient.

TABLE 1 | The association of ICI scores with somatic cell variation, where chi-square tests were used to compare statistical differences between high and low ICI score subgroups.

Gene symbol	High ICI score (%)	Low ICI score (%)	*p* Value
CTNNB1	53 (32)	30 (18)	0.0054
TP53	37 (22)	56 (34)	0.0147
AXIN1	8 (5)	19 (12)	0.0270
TTN	49 (29)	31 (19)	0.0298
ALB	23 (14)	12 (7)	0.0734
CSMD3	9 (5)	17 (10)	0.1039
LRP1B	18 (11)	9 (6)	0.1074
RYR2	11 (7)	19 (12)	0.1272
ADGRV1	7 (4)	13 (8)	0.1713
APOB	20 (12)	12 (7)	0.1934
HMCN1	15 (9)	8 (5)	0.1945
PCLO	21 (13)	13 (8)	0.2057
CACNA1E	16 (10)	9 (6)	0.2122
ABCA13	18 (11)	11 (7)	0.2441
MUC4	20 (12)	13 (8)	0.2720
FAT3	8 (5)	12 (7)	0.3635
CCDC168	9 (5)	11 (7)	0.6503
MUC16	25 (15)	28 (17)	0.6536
RYR1	10 (6)	12 (7)	0.6638
ARID1A	11 (7)	13 (8)	0.6757
SPTA1	11 (7)	13 (8)	0.6757
FLG	12 (7)	14 (9)	0.6863
OBSCN	15 (9)	12 (7)	0.6892
XIRP2	15 (9)	12 (7)	0.6892
USH2A	11 (7)	10 (6)	1.0000

R = −0.249, $p < 0.001$; **Figure 4A**). Box plots exhibited lower immune scores and ESTIMATE scores for the high-score subgroups ($p < 0.05$; **Figures 4B,C**), while stromal scores and tumor purity scores did not differ significantly between the two subgroups (**Figures 4D,E**). To assess immunocompetence among subgroups, CD274, CTLA4, HAVCR2, IDO1, LAG3, and PDCD1 were selected as immune checkpoint-related features, while CD8A, CXCL10, CXCL9, GZMA, GZMB, IFNG, PRF1, TBX2, and TNF were selected as immunocompetence features (Hugo et al., 2016; Ayers et al., 2017). Our results showed that almost all, immune checkpoint-related and immunocompetence-related genes (except CD274 and CXCL10), had a significant overexpression in the high ICI score subgroup (**Figure 4F**). Additionally, higher infiltration levels of NK cells, gamma delta T cells, monocytes, and M1 macrophage and lower infiltration levels of regulatory T cells were observed in the high-score subgroups (**Figure 4G**), which was also consistent in the immune landscape of the ICGC cohort (**Supplementary Figures S4C,D**).

Association Between Immune Cell Infiltration Scores and Somatic Cell Variation

Previous investigations have revealed that increased infiltration of CD8T cells in high mutation burden-associated tumor tissues (nonsynonymous variants) can identify and eliminate these cancers (McGranahan et al., 2016). Higher tumor mutation burden (TMB) and somatic mutation rates are associated with stronger anti-cancer immunity (Rizvi et al., 2015; Rooney et al., 2015). The KEYNOTE 012 clinical trial showed that TMB increase was associated with improved PD-1 inhibitors and prolonged progression-free survival of patients (Seiwert et al., 2016; Cristescu et al., 2018). Because of the clinical significance of TMB, the correlation between TMB and ICI scores was analyzed

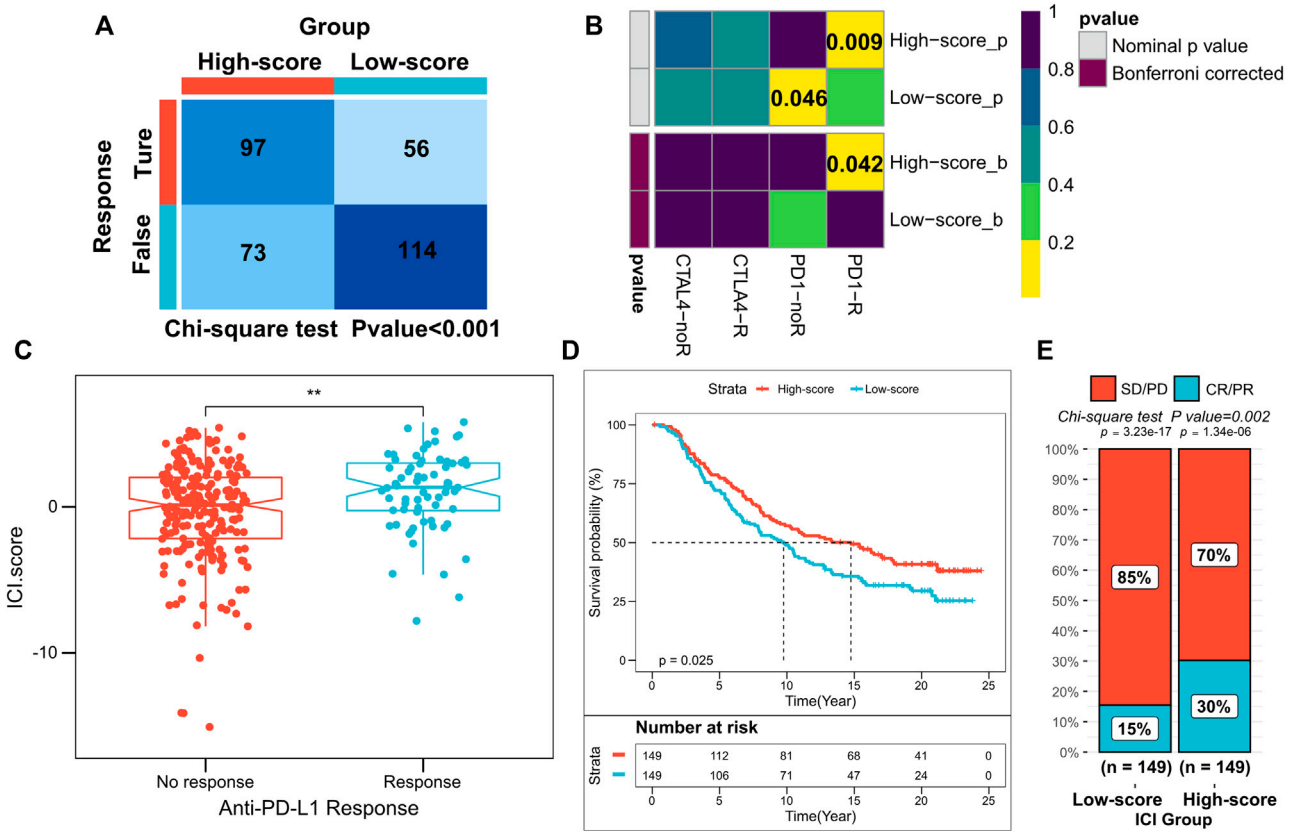


FIGURE 6 | Role of ICI scores in predicting the efficacy of immunotherapy. **(A)** TIDE algorithm showing higher immunotherapy response in patients with high ICI scores, where $p < 0.001$ for the chi-square test; **(B)** Subclass mapping showing higher sensitivity to anti-PD1 treatment in patients with high-score subgroups (FDR = 0.042); **(C)** ICI scores for subgroups with different anti-PD-1 clinical response status, where $p < 0.01$ for the Wilcoxon test; **(D)** Kaplan-Meier curves for patients with high and low ICI scores in the IMvigor210 cohort, where $p = 0.025$ for the log-rank test; (d) Clinical response rates to anti-PD-L1 immunotherapy in the high or low ICI score subgroups in the IMvigor210 cohort [complete response (CR)/partial response (PR) and stable disease (SD)/progressive disease (PD)].

in detail. For this purpose, first, the TMB comparison between patients in the high- and low-score subgroups were analyzed; ICI score and TMB were positively correlated (Pearson correlation: R = 0.151, $p = 0.009$; **Figure 5A**). TMB was significantly higher in the high-score subgroup (Wilcoxon test $p < 0.001$; **Figure 5B**). Patients were divided into high- and low-TMB score subgroups based on the optimal cut-off value of TMB calculated from the "survminer" package; patients with high TMB scores exhibited poorer OS ($p = 0.056$; **Figure 5C**). Due to the opposing predictions of OS by ICI and TMB scores, the combined effect of these scores in the prognostic stratification of HCC was subsequently evaluated. Stratified survival analyses showed that TMB did not affect ICI score-based prediction; significant survival differences for ICI score subtypes were obtained between the two TMB-based score subgroups (log-rank test, $p = 0.002$; High TMB & High ICI score (HH) versus High TMB & Low ICI score (HL), $p = 0.011$; Low TMB & Low ICI score (LH) versus Low TMB & Low ICI score (LL), $p = 0.047$; **Figure 5D**). Overall, our findings suggested that ICI scores may have implications as an independent predictor of TMB and could be a reliable parameter for patient prognosis. In addition, differences in somatic variant driver genes between the low and high ICI score subgroups were evaluated. The driver genes for HCC were obtained using "maftools"; among them the most frequently altered top 25 genes were further analyzed (**Figure 5E**). The frequencies of CTNNB1, TP53, AXIN1, and TTN were significantly altered between the high- and low-score subgroups (chi-square test; **Table 1**). Taken together, these results may provide new insights for future investigations on the constituents of tumor ICI and the mechanisms of gene mutations in immune checkpoint blockade therapy.

Predictive Efficacy of Immune Cell Infiltration Scores for Immunotherapy

Novel immune checkpoint inhibition has shown promising results in both preclinical trials and real clinic settings. However, only a small proportion of patients respond to these therapies (Curran et al., 2010; Grosso and Jure-Kunkel, 2013; Larkin et al., 2015). Our subsequent analyses assessed the utility of scores based on ICI in predicting the efficacy of immunotherapy in HCC. Differences in response to anti-PD1 and anti-CTLA4 therapy between the high- and low-score subgroups of patients in the TCGA and ICGC cohorts were

evaluated using the TIDE algorithm. In the high-score subgroup, the patients had a higher immunotherapy response rate (chi-square test $p < 0.001$; **Figure 6A**; **Supplementary Figure S4E**). Subclass mapping analysis predicted the immunotherapy responses of both subgroups to PD1 and CTLA4 inhibitors. The high-score subgroups in both the TCGA and ICGC cohorts were found to be more sensitive to anti-PD1 treatment (TCGA: FDR = 0.042; ICGC: FDR = 0.022, respectively) (**Figure 6B**; **Supplementary Figure S4F**). In addition, patients in the IMvigor210 cohort administered with anti-PD-L1 immunotherapy were also assigned the corresponding ICI scores (high or low). Notably, patients with high ICI scores in the IMvigor210 cohort survived longer as compared to those with low ICI scores (log-rank test, $p = 0.0017$; **Figure 6C**). In the IMvigor210 cohort, anti-PD-L1 therapy's objective remission rate was significantly higher in the high ICI score subgroup (Chi-square test, $p = 0.002$; **Figure 6D**). Our results also indicated that higher ICI scores in the IMvigor210 cohort were associated with objective responses to anti-PD-L1 therapy (Wilcoxon test, $p < 0.01$; **Figure 6E**). Overall, these findings suggest a possible association between ICI scores and immunotherapeutic responses.

DISCUSSION

HCC is an aggressive tumor with a high degree of malignancy, and most patients are diagnosed initially at an advanced stage (Zhou et al., 2020). High recurrence and metastasis rates of advanced HCC to a low possibility for surgical resection (Zhou et al., 2020; Feng et al., 2021). Within the local area, the complex genomic alterations, differences in biological behaviors, and heterogeneity of the tumor microenvironment resulting in a complex HCC process. Currently, immunotherapy is a promising treatment strategy available for HCC (Huang et al., 2020). Due to the limitations of surgical resection, chemotherapy and immunotherapy have received increasing attention in the treatment of advanced HCC (Brown et al., 2019). However, immunotherapeutic response rates are highly heterogeneous and remain considerably low (Feng et al., 2021). Thus, in HCC, immune-related classification criteria may provide new insights to assess the efficacy of immunotherapy and predict the patient prognosis.

The high genomic heterogeneity of HCC results in the complexity of the immune microenvironment (Dal Bo et al., 2020). Therefore, the identification of novel signatures in HCC based on immune-related genes provides a new direction for assessing the efficacy of immunotherapy. Further assessment of these classifications based on gene signatures may help in developing immunotherapy strategies with improved sensitivity for different subtypes of HCC. Zhang et al. characterized the ICI dynamics in HCC by single-cell sequencing, and thus provided a new basis for investigations of the immune landscape (Zhang et al., 2019a). Sia et al. identified active or depleted immune subtypes in HCC based on immune gene transcriptional profiling. This suggests that active immune subtypes may be more sensitive to immunosuppressant therapy (Sia et al., 2017). Zhang et al. integrated multi-omics data and show new immunophenotypic classifications in HCC which may be useful for prognostic prediction and potentially supporting new treatment targets (Zhang et al., 2019b). Indeed, these studies have their unique strengths and potential and complement each other. Therefore, investigations of HCC immune subtypes from different perspectives hold great promise for research, and a better classification of immune features would enhance the overall understanding of HCC immunotherapy.

In the present study, we analyzed the classical HCC dataset from the TCGA and ICGC cohorts and divided the patients into three different immune subtypes. Our results suggested that high infiltration levels of CD4 T cells, CD8 T cells, and M1 macrophage and low infiltration levels of regulatory T cells were associated with good prognosis, consistent with previous studies (Rooney et al., 2015; He et al., 2018). Due to the heterogeneity of immune landscape and prognosis among the three immune subtypes, we speculated that an integrated ICI profile analysis and evaluation of immune-based gene expression patterns would be a new approach to develop patient-customized and tailored treatment strategies. Four distinct gene clusters were obtained based on differentially expressed genes between the subtypes; clusters 1 and 2 exhibited a more favorable immune activation phenotype, exhibited higher infiltration of M1 macrophages, monocytes, gamma delta T cells, and lower infiltration levels of regulatory T cells (Biswas and Mantovani, 2010; Chen and Mellman, 2017); in contrast, the highest infiltration levels of regulatory T-cells and M0 macrophage and lowest infiltration levels of other cell types were found in cluster 3, which suggested a general immune failure phenotype (Biswas and Mantovani, 2010; Chen and Mellman, 2017). TME impact on patient's OS has been well documented in previous studies; ICI differences resulted in cluster 1 and cluster 2 patients having a good prognosis, while patients in cluster 3 had the worst prognosis, consistent with previous studies (Chen et al., 2019; Li et al., 2019). These findings suggested that the gene clusters in this study may have a potential role in more accurate predictions of patient outcomes.

Given the differences in patient prognosis and immune landscape between gene clusters, it was imperative to quantify the individual patient ICI patterns for improved outcome prediction. Individual models based on tumor subtype-specific biomarkers show good efficacy for HCC (Sia et al., 2017; Kurebayashi et al., 2018). In this study, potential "subtype biomarkers" were obtained using the Boruta algorithm and ICI scores were calculated to quantify ICI patterns. GSEA showed that cancer-related pathways including cell cycle pathways and RNA degradation pathways were significantly enriched in the low ICI score group. Recently, preclinical trial reports show the correlation of gene mutations with tolerance or immunotherapeutic responses (Rizvi et al., 2015; Rooney et al., 2015). Several genes with significant differences in mutation frequencies exist between the high and low ICI score subgroups. All of these play an important role in cancer progression (Mazzoni and Fearon, 2014; Mantovani et al., 2019; Wen et al., 2019; Yang et al., 2020). Moreover, the highly immunotherapy-sensitive, TMB, was significantly lowered in patients with lower ICI scores (correlation 0.151). The stratified analysis could independently predict the prognostic value of ICI scores for TMB. These results implied that ICI scores and TMB represented different aspects of tumor immunobiology and ICI scores could indeed predict patient responses to immunotherapy in conditions independent of TMB.

The efficacy of ICI scores in predicting immunotherapeutic responses was further evaluated by multiple methods; TIDE and subclass mapping analyses showed that patients with higher ICI scores were more sensitive to anti-PD1 therapy. After evaluating patients in the anti-PD1 immunotherapy regime in the IMvigor210 cohort, a significant increase in ICI scores was found which validated the predictive value of patients' response. These results indicated that mono-immunotherapy may benefit patients with high ICI scores.

However, the present study has some limitations. The current results need to be validated for their efficiency in immunotherapy clinical trials with larger HCC cohorts. This would confirm the utility of classification for clinical evaluation and decision-making. Additionally, transcriptomic information was obtained from post-surgical liver tissues. Thus, the model may not accurately predict outcomes prior to the onset of HCC. Therefore, a better understanding of circulating biomarkers released into the bloodstream from tumor cells and tumor-associated immune cells is important. Further *in vivo* and *in vitro* experiments should investigate the potential functional and mechanical differences between the subtypes. Finally, the findings of this study and ICI scores may apply to other cancers, and these require further studies.

In conclusion, a comprehensive analysis of ICI patterns in HCC provides a foundational basis for the regulation of anti-tumor/tumor-promoting immune responses in HCC. These suggested that differences in ICI patterns correlated with tumor heterogeneity and therapeutic complexity. Based on this, a practical model for quantifying individual ICI patterns was proposed, which could predict the prognosis of HCC patients and identify potential candidates for developing immunotherapy regimens.

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

ST conceived and designed the whole project and drafted the manuscript. SY, YC, and XW analyzed the data and wrote the manuscript. PW carried out data interpretations and helped data discussion. HW provided specialized expertise and collaboration in data analysis. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Leading Talent Project of Jiangsu Province Traditional Chinese Medicine (No. SLJ0216), the National Natural Science Foundation of China (Nos 81870423, 82073914,82000572), the Major Project of the Natural Science Research of Jiangsu Higher Education Institutions (No. 19KJA310005), the Postgraduate Research and Practice Innovation Program of Jiangsu Province (No. KYCX201493), the Joint Project of Jiangsu Key Laboratory for Pharmacology and Safety Evaluation of Chinese Materia Medica and Yangtze River Pharmaceutical (No. JKLPSE202005), and the Natural Science Foundation of Jiangsu Province (No. BK20200056).

ACKNOWLEDGMENTS

The authors hereby express their gratitude to all participants who supported the study, especially the TCGA and ICGC database providers who provided the data for the analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.777931/full#supplementary-material

REFERENCES

Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D. R., et al. (2017). IFN-γ-related mRNA Profile Predicts Clinical Response to PD-1 Blockade. *J. Clin. Invest.* 127 (8), 2930–2940. doi:10.1172/jci91190

Biswas, S. K., and Mantovani, A. (2010). Macrophage Plasticity and Interaction with Lymphocyte Subsets: Cancer as a Paradigm. *Nat. Immunol.* 11 (10), 889–896. doi:10.1038/ni.1937

Brown, Z. J., Greten, T. F., and Heinrich, B. (2019). Adjuvant Treatment of Hepatocellular Carcinoma: Prospect of Immunotherapy. *Hepatology* 70 (4), 1437–1442. doi:10.1002/hep.30633

Caruso, S., O'Brien, D. R., Cleary, S. P., Roberts, L. R., and Zucman-Rossi, J. (2021). Genetics of Hepatocellular Carcinoma: Approaches to Explore Molecular Diversity. *Hepatology* 73 (Suppl. 1), 14–26. doi:10.1002/hep.31394

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

Chen, Y.-P., Wang, Y.-Q., Lv, J.-W., Li, Y.-Q., Chua, M. L. K., Le, Q.-T., et al. (2019). Identification and Validation of Novel Microenvironment-Based Immune Molecular Subgroups of Head and Neck Squamous Cell Carcinoma: Implications for Immunotherapy. *Ann. Oncol.* 30 (1), 68–75. doi:10.1093/annonc/mdy470

Choi, S. H., and Park, J. Y. (2017). Regulation of the Hypoxic Tumor Environment in Hepatocellular Carcinoma Using RNA Interference. *Cancer Cell Int* 17, 3. doi:10.1186/s12935-016-0374-6

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAbiolinks: an R/Bioconductor Package for Integrative Analysis of TCGA Data. *Nucleic Acids Res.* 44 (8), e71. doi:10.1093/nar/gkv1507

Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., et al. (2018). Pan-tumor Genomic Biomarkers for PD-1 Checkpoint Blockade-Based Immunotherapy. *Science* 362, eaar3593. doi:10.1126/science.aar3593

Curran, M. A., Montalvo, W., Yagita, H., and Allison, J. P. (2010). PD-1 and CTLA-4 Combination Blockade Expands Infiltrating T Cells and Reduces Regulatory T and Myeloid Cells within B16 Melanoma Tumors. *Proc. Natl. Acad. Sci.* 107 (9), 4275–4280. doi:10.1073/pnas.0915174107

Dal Bo, M., De Mattia, E., Baboci, L., Mezzalira, S., Cecchin, E., Assaraf, Y. G., et al. (2020). New Insights into the Pharmacological, Immunological, and CAR-T-Cell Approaches in the Treatment of Hepatocellular Carcinoma. *Drug Resist. Updates* 51, 100702. doi:10.1016/j.drup.2020.100702

De Palma, M., and Lewis, C. E. (2013). Macrophage Regulation of Tumor Responses to Anticancer Therapies. *Cancer cell* 23 (3), 277–286. doi:10.1016/j.ccr.2013.02.013

Fan, S. T. (2012). Hepatocellular Carcinoma-Resection or Transplant? *Nat. Rev. Gastroenterol. Hepatol.* 9 (12), 732–737. doi:10.1038/nrgastro.2012.158

Farzaneh, Z., Vosough, M., Agarwal, T., and Farzaneh, M. (2021). Critical Signaling Pathways Governing Hepatocellular Carcinoma Behavior; Small Molecule-Based Approaches. *Cancer Cell Int* 21 (1), 208. doi:10.1186/s12935-021-01924-w

Feng, G. S., Hanley, K. L., Liang, Y., and Lin, X. (2021). Improving the Efficacy of Liver Cancer Immunotherapy: The Power of Combined Preclinical and Clinical Studies. *Hepatology* 73 (Suppl. 1), 104–114. doi:10.1002/hep.31479

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale Public Data Reuse to Model Immunotherapy Response and Resistance. *Genome Med.* 12 (1), 21. doi:10.1186/s13073-020-0721-z

Fujimoto, A., Furuta, M., Totoki, Y., Tsunoda, T., Kato, M., Shiraishi, Y., et al. (2016). Whole-genome Mutational Landscape and Characterization of Noncoding and Structural Mutations in Liver Cancer. *Nat. Genet.* 48 (5), 500–509. doi:10.1038/ng.3547

Grosso, J. F., and Jure-Kunkel, M. N. (2013). CTLA-4 Blockade in Tumor Models: an Overview of Preclinical and Translational Research. *Cancer Immun.* 13, 5.

Han, J., Khatwani, N., Searles, T. G., Turk, M. J., and Angeles, C. V. (2020). Memory CD8+ T Cell Responses to Cancer. *Semin. Immunol.* 49, 101435. doi:10.1016/j.smim.2020.101435

He, Y., Jiang, Z., Chen, C., and Wang, X. (2018). Classification of Triple-Negative Breast Cancers Based on Immunogenomic Profiling. *J. Exp. Clin. Cancer Res.* 37 (1), 327. doi:10.1186/s13046-018-1002-1

Hoshida, Y., Brunet, J.-P., Tamayo, P., Golub, T. R., and Mesirov, J. P. (2007). Subclass Mapping: Identifying Common Subtypes in Independent Disease Data Sets. *PloS one* 2 (11), e1195. doi:10.1371/journal.pone.0001195

Hosseinzadeh, F., Verdi, J., Ai, J., Hajighasemlou, S., Seyhoun, I., Parvizpour, F., et al. (2018). Combinational Immune-Cell Therapy of Natural Killer Cells and Sorafenib for Advanced Hepatocellular Carcinoma: a Review. *Cancer Cell Int* 18, 133. doi:10.1186/s12935-018-0624-x

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

Hugo, W., Zaretsky, J. M., Sun, L., Song, C., Moreno, B. H., Hu-Lieskovan, S., et al. (2016). Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* 165 (1), 35–44. doi:10.1016/j.cell.2016.02.065

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

Jiang, Y., Zhang, Q., Hu, Y., Li, T., Yu, J., Zhao, L., et al. (2018). ImmunoScore Signature. *Ann. Surg.* 267 (3), 504–513. doi:10.1097/sla.0000000000002116

Kulik, L., and El-Serag, H. B. (2019). Epidemiology and Management of Hepatocellular Carcinoma. *Gastroenterology* 156 (2), 477–491. doi:10.1053/j.gastro.2018.08.065

Kurebayashi, Y., Ojima, H., Tsujikawa, H., Kubota, N., Maehara, J., Abe, Y., et al. (2018). Landscape of Immune Microenvironment in Hepatocellular Carcinoma and its Additional Impact on Histological and Molecular Classification. *Hepatology* 68 (3), 1025–1041. doi:10.1002/hep.29904

Larkin, J., Hodi, F. S., and Wolchok, J. D. (2015). Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N. Engl. J. Med.* 373 (13), 1270–1271. doi:10.1056/NEJMc1509660

Li, B., Cui, Y., Nambiar, D. K., Sunwoo, J. B., and Li, R. (2019). The Immune Subtypes and Landscape of Squamous Cell Carcinoma. *Clin. Cancer Res.* 25 (12), 3528–3537. doi:10.1158/1078-0432.Ccr-18-4085

Liu, Z., Liu, L., Lu, T., Wang, L., Li, Z., Jiao, D., et al. (2021). Hypoxia Molecular Characterization in Hepatocellular Carcinoma Identifies One Risk Signature and Two Nomograms for Clinical Management. *J. Oncol.* 2021, 1–20. doi:10.1155/2021/6664386

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

Liu, Z., Zhang, Y., Dang, Q., Wu, K., Jiao, D., Li, Z., et al. (2021). Genomic Alteration Characterization in Colorectal Cancer Identifies a Prognostic and Metastasis Biomarker: FAM83A|Ido1. *Front. Oncol.* 11, 632430. doi:10.3389/fonc.2021.632430

Liu, Z., Zhang, Y., Shi, C., Zhou, X., Xu, K., Jiao, D., et al. (2021). A Novel Immune Classification Reveals Distinct Immune Escape Mechanism and Genomic Alterations: Implications for Immunotherapy in Hepatocellular Carcinoma. *J. Transl Med.* 19 (1), 5. doi:10.1186/s12967-020-02697-y

Mantovani, F., Collavin, L., and Del Sal, G. (2019). Mutant P53 as a Guardian of the Cancer Cell. *Cell Death Differ* 26 (2), 199–212. doi:10.1038/s41418-018-0246-9

Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

Mazzoni, S. M., and Fearon, E. R. (2014). AXIN1 and AXIN2 Variants in Gastrointestinal Cancers. *Cancer Lett.* 355 (1), 1–8. doi:10.1016/j.canlet.2014.09.018

McGranahan, N., Furness, A. J. S., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S. K., et al. (2016). Clonal Neoantigens Elicit T Cell Immunoreactivity and Sensitivity to Immune Checkpoint Blockade. *Science* 351 (6280), 1463–1469. doi:10.1126/science.aaf1490

Mehla, K., and Singh, P. K. (2019). Metabolic Regulation of Macrophage Polarization in Cancer. *Trends Cancer* 5 (12), 822–834. doi:10.1016/j.trecan.2019.10.007

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

Noy, R., and Pollard, J. W. (2014). Tumor-associated Macrophages: from Mechanisms to Therapy. *Immunity* 41 (1), 49–61. doi:10.1016/j.immuni.2014.06.010

Park, J. W., Chen, M., Colombo, M., Roberts, L. R., Schwartz, M., Chen, P. J., et al. (2015). Global Patterns of Hepatocellular Carcinoma Management from Diagnosis to Death: the BRIDGE Study. *Liver Int.* 35 (9), 2155–2166. doi:10.1111/liv.12818

Pillai, A., Ahn, J., and Kulik, L. (2020). Integrating Genomics into Clinical Practice in Hepatocellular Carcinoma: The Challenges Ahead. *Am. J. Gastroenterol.* 115 (12), 1960–1969. doi:10.14309/ajg.0000000000000843

Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., et al. (2015). Mutational Landscape Determines Sensitivity to PD-1 Blockade in Non-small Cell Lung Cancer. *Science* 348 (6230), 124–128. doi:10.1126/science.aaa1348

Robert, C., Marabelle, A., Herrscher, H., Caramella, C., Rouby, P., Fizazi, K., et al. (2020). Immunotherapy Discontinuation - How, and when? Data from Melanoma as a Paradigm. *Nat. Rev. Clin. Oncol.* 17 (11), 707–715. doi:10.1038/s41571-020-0399-6

Roh, W., Chen, P.-L., Reuben, A., Spencer, C. N., Prieto, P. A., Miller, J. P., et al. (2017). Integrated Molecular Analysis of Tumor Biopsies on Sequential CTLA-4 and PD-1 Blockade Reveals Markers of Response and Resistance. *Sci. Transl. Med.* 9 (379). doi:10.1126/scitranslmed.aah3560

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. *Cell* 160 (1-2), 48–61. doi:10.1016/j.cell.2014.12.033

Sachdeva, M., and Arora, S. K. (2020). Prognostic Role of Immune Cells in Hepatocellular Carcinoma. *EXCLI J.* 19, 718–733. doi:10.17179/excli2020-1455

Seiwert, T. Y., Burtness, B., Mehra, R., Weiss, J., Berger, R., Eder, J. P., et al. (2016). Safety and Clinical Activity of Pembrolizumab for Treatment of Recurrent or Metastatic Squamous Cell Carcinoma of the Head and Neck (KEYNOTE-012): an Open-Label, Multicentre, Phase 1b Trial. *Lancet Oncol.* 17 (7), 956–965. doi:10.1016/s1470-2045(16)30066-3

Sia, D., Jiao, Y., Martinez-Quetglas, I., Kuchuk, O., Villacorta-Martin, C., Castro de Moura, M., et al. (2017). Identification of an Immune-specific Class of Hepatocellular Carcinoma, Based on Molecular Features. *Gastroenterology* 153 (3), 812–826. doi:10.1053/j.gastro.2017.06.007

Silva, L., Egea, J., Villanueva, L., Ruiz, M., Llopiz, D., Repáraz, D., et al. (2020). Cold-Inducible RNA Binding Protein as a Vaccination Platform to Enhance Immunotherapeutic Responses against Hepatocellular Carcinoma. *Cancers* 12 (11), 3397. doi:10.3390/cancers12113397

Singal, A. G., Lampertico, P., and Nahon, P. (2020). Epidemiology and Surveillance for Hepatocellular Carcinoma: New Trends. *J. Hepatol.* 72 (2), 250–261. doi:10.1016/j.jhep.2019.08.025

Stairiker, C. J., Thomas, G. D., and Salek-Ardakani, S. (2020). EZH2 as a Regulator of CD8+ T Cell Fate and Function. *Front. Immunol.* 11, 593203. doi:10.3389/fimmu.2020.593203

Wang, W.-J., Wang, H., Hua, T.-Y., Song, W., Zhu, J., Wang, J.-J., et al. (2020). Establishment of a Prognostic Model Using Immune-Related Genes in Patients

with Hepatocellular Carcinoma. *Front. Genet.* 11, 55. doi:10.3389/fgene.2020.00055

Wen, J., Min, X., Shen, M., Hua, Q., Han, Y., Zhao, L., et al. (2019). ACLY Facilitates colon Cancer Cell Metastasis by CTNNB1. *J. Exp. Clin. Cancer Res.* 38 (1), 401. doi:10.1186/s13046-019-1391-9

Yang, Y., Zhang, J., Chen, Y., Xu, R., Zhao, Q., and Guo, W. (2020). MUC4 , MUC16 , and TTN Genes Mutation Correlated with Prognosis, and Predicted Tumor Mutation burden and Immunotherapy Efficacy in Gastric Cancer and pan-cancer. *Clin. translational Med.* 10 (4), e155. doi:10.1002/ctm2.155

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

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zeng, D., Zhou, R., Yu, Y., Luo, Y., Zhang, J., Sun, H., et al. (2018). Gene Expression Profiles for a Prognostic Immunoscore in Gastric Cancer. *Br. J. Surg.* 105 (10), 1338–1348. doi:10.1002/bjs.10871

Zhang, Q., He, Y., Luo, N., Patel, S. J., Han, Y., Gao, R., et al. (2019). Landscape and Dynamics of Single Immune Cells in Hepatocellular Carcinoma. *Cell* 179 (4), 829–845. e20. doi:10.1016/j.cell.2019.10.003

Zhang, Q., Lou, Y., Yang, J., Wang, J., Feng, J., Zhao, Y., et al. (2019). Integrated Multiomic Analysis Reveals Comprehensive Tumour Heterogeneity and Novel Immunophenotypic Classification in Hepatocellular Carcinomas. *Gut* 68 (11), 2019–2031. doi:10.1136/gutjnl-2019-318912

Zhang, X., Shi, M., Chen, T., and Zhang, B. (2020). Characterization of the Immune Cell Infiltration Landscape in Head and Neck Squamous Cell Carcinoma to Aid Immunotherapy. *Mol. Ther. - Nucleic Acids* 22, 298–309. doi:10.1016/j.omtn.2020.08.030

Zhou, T., Liang, X., Wang, P., Hu, Y., Qi, Y., Jin, Y., et al. (2020). A Hepatocellular Carcinoma Targeting Nanostrategy with Hypoxia-Ameliorating and Photothermal Abilities that, Combined with Immunotherapy, Inhibits Metastasis and Recurrence. *ACS nano* 14 (10), 12679–12696. doi:10.1021/acsnano.0c01453

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: 26 November 2021
doi: 10.3389/fgene.2021.780780

A Novel Pyroptosis-Related Signature for Predicting Prognosis and Indicating Immune Microenvironment Features in Osteosarcoma

Yiming Zhang[1†], Rong He[2†], Xuan Lei[3], Lianghao Mao[1], Pan Jiang[1,4], Chenlie Ni[1], Zhengyu Yin[1], Xinyu Zhong[1], Chen Chen[5], Qiping Zheng[5,6] and Dapeng Li[1]**

[1]Department of Orthopedics, Affiliated Hospital of Jiangsu University, Zhenjiang, China, [2]Cancer Institute, The Affiliated People's Hospital of Jiangsu University, Zhenjiang, China, [3]Department of Burn and Plastic Surgery, Affiliated Hospital of Jiangsu University, Zhenjiang, China, [4]Guizhou Orthopedics Hospital, Guiyang, China, [5]Department of Hematological Laboratory Science, Jiangsu Key Laboratory of Medical Science and Laboratory Medicine, School of Medicine, Jiangsu University Zhenjiang, Guiyang, China, [6]Shenzhen Academy of Peptide Targeting Technology at Pingshan, and Shenzhen Tyercan Bio-Pharm Co., Ltd., Shenzhen, China

Osteosarcoma is a common malignant bone tumor with a propensity for drug resistance, recurrence, and metastasis. A growing number of studies have elucidated the dual role of pyroptosis in the development of cancer, which is a gasdermin-regulated novel inflammatory programmed cell death. However, the interaction between pyroptosis and the overall survival (OS) of osteosarcoma patients is poorly understood. This study aimed to construct a prognostic model based on pyroptosis-related genes to provide new insights into the prognosis of osteosarcoma patients. We identified 46 differentially expressed pyroptosis-associated genes between osteosarcoma tissues and normal control tissues. A total of six risk genes affecting the prognosis of osteosarcoma patients were screened to form a pyroptosis-related signature by univariate and LASSO regression analysis and verified using GSE21257 as a validation cohort. Combined with other clinical characteristics, including age, gender, and metastatic status, we found that the pyroptosis-related signature score, which we named "PRS-score," was an independent prognostic factor for patients with osteosarcoma and that a low PRS-score indicated better OS and a lower risk of metastasis. The result of ssGSEA and ESTIMATE algorithms showed that a lower PRS-score indicated higher immune scores, higher levels of tumor infiltration by immune cells, more active immune function, and lower tumor purity. In summary, we developed and validated a pyroptosis-related signature for predicting the prognosis of osteosarcoma, which may contribute to early diagnosis and immunotherapy of osteosarcoma.

Keywords: osteosarcoma, pyroptosis, prognosis, immunotherapy, survival analysis

OPEN ACCESS

Edited by:
Frank Emmert-Streib, Tampere University, Finland

Reviewed by:
Junming Ren, Stanford University, United States
Emil Bulatov, Kazan Federal University, Russia

****Correspondence:***
Dapeng Li
lidapeng706@hotmail.com
Qiping Zheng
qp_zheng@hotmail.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: *21 September 2021*
Accepted: *08 November 2021*
Published: *26 November 2021*

Citation:
Zhang Y, He R, Lei X, Mao L, Jiang P, Ni C, Yin Z, Zhong X, Chen C, Zheng Q and Li D (2021) A Novel Pyroptosis-Related Signature for Predicting Prognosis and Indicating Immune Microenvironment Features in Osteosarcoma. Front. Genet. 12:780780. doi: 10.3389/fgene.2021.780780

INTRODUCTION

Osteosarcoma is the most common primary aggressive malignancies of the skeleton, and it occurs mainly in children and adolescents, in which distant metastasis still leads to a poor prognosis (Chow et al., 2020; Rojas et al., 2021). With a combination of neoadjuvant chemotherapy, surgery, chemotherapy, and biological therapy in the last few years, the 5-year survival rate for osteosarcoma patients has improved significantly, from 20% to 65–70% (Yao and Chen, 2020; Gazouli et al., 2021). However, due to the limited efficacy of current treatment strategies, nearly 30% of osteosarcoma patients are prone to metastasis or recurrence, with poor prognosis and low 5-year survival rates (Fan et al., 2021). Recently, immunotherapy has undergone a dramatic transformation, demonstrating superior anticancer efficacy in many tumors and being recognized as a more potent and antigen-specific form of antitumor therapy (Constantinidou et al., 2019; Shin et al., 2021). For example, adoptive cellular immunotherapy is a promising option for tumors resistant to current conventional therapy, and chimeric antigen receptor T-cell therapy has been shown to cure 25–50% of patients with previously incurable B-cell malignancies, revolutionizing the treatment of drug-resistant hematologic malignancies (Titov et al., 2021). In addition, specific immune checkpoint inhibitors are being explored as new immunotherapeutic strategies for osteosarcoma, such as CTLA-4, LAG3, TIGIT, and PD-1/L1 (Wang S.-D. et al., 2016; Hashimoto et al., 2020; Judge et al., 2020; Park and Cheung, 2020; Ligon et al., 2021). However, cancer immunotherapies, including checkpoint inhibitors, have varying response rates due to multiple primary and acquired resistance mechanisms (Bashash et al., 2021). In order to improve the early diagnosis and treatment of osteosarcoma, novel biomarkers and therapeutic targets are needed.

Pyroptosis is a newly discovered form of programmed cell death that is morphologically distinct from apoptosis and necrosis while releasing inflammatory mediators in the process (Wu et al., 2021). Pyroptosis is mediated by pore-forming proteins, such as the gasdermin family, of which gasdermin D (GSDMD) is a primary substrate for the caspase family (Li L. et al., 2021). After cleavage by activated caspases, the N-terminal fragment of GSDMD oligomerizes in the membrane to form pores, leading to pyroptosis (Lu et al., 2021). Pyroptosis acts as a double-edged sword in cancer. On the one hand, pyroptosis can create a tumor-promoting environment by releasing inflammatory factors; on the other hand, pyroptosis can inhibit tumor occurrence and progression as a form of programmed death (Xia et al., 2019). As research progresses, the impact of pyroptosis-related genes on the proliferation, migration, and invasion of tumor cells becomes increasingly prominent and is strongly associated with cancer prognosis (Ju A. et al., 2021; Lin W. et al., 2021; Shao et al., 2021; Ye et al., 2021). For instance, Tang et al. (2020) reported that pyroptosis inhibited metastasis of colorectal cancer cells through activation of NLRP3-ASC-Caspase-1 signaling by FL118. In another study by Wang Y. et al. (2016), it was found that the NLRP3 inflammasome can promote the proliferation and migration of A549 lung cancer cells via the caspase-1-IL-1β/IL-18 signaling pathway. Studies have shown that GSDMD was notably upregulated in osteosarcoma compared to normal skeletal tissue as well as associated with drug resistance and prognosis for patients with osteosarcoma (Lin R. et al., 2020). Alternatively, GSDMD expression was significantly downregulated in gastric cancer tissues, which may contribute to the development of gastric cancer through the regulation of cell cycle transition (Wang et al., 2018). However, the mechanism of pyroptosis-related genes in osteosarcoma is still not fully elucidated.

Recently, high-throughput sequencing technologies and bioinformatics analysis have enabled the exploration of genetic alterations in osteosarcoma and provided an effective way to identify potentially beneficial markers and the most appropriate treatment strategies for other cancer types (Li M. et al., 2021; Na et al., 2021; Pan et al., 2021). According to Zhang et al. (Xing et al., 2021), TIMELESS was the most significantly upregulated gene within the 16 clock-related genes by analyzing The Cancer Genome Atlas (TCGA) database and promoted cancer cell proliferation and migration via increasing macrophage infiltration in ovarian cancer. An analysis of the relationship between osteosarcoma development and KIF21B using bioinformatics analysis showed that knockdown of KIF21B inhibited cell proliferation and reduced tumor formation *in vivo* by modulating the PI3K/AKT pathway and that KIF21B was an independent prognostic factor in osteosarcoma patients (Ni et al., 2020). The previous success of projects to identify prognostic target genes suggests that it may be possible to uncover more molecular mechanisms in osteosarcoma.

We used microarray data from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) and Genotype-Tissue Expression (GTEx) database for differential expression analysis and identified 46 differentially expressed pyroptosis-related genes (DEPRGs) in osteosarcoma and normal muscle tissues. We then constructed a six-gene signature (that could determine the PRS-score) based on DEPRGs to predict osteosarcoma outcomes. We validated the signature by evaluating the association between the PRS-scores and clinical characteristics and immune microenvironment features in osteosarcoma tumors. The differential genes among the PRS-score-based subgroups are also enriched for immunological functions and may be involved in regulating the composition of the immune microenvironment. These results reveal that the pyroptosis-related prognostic signature may provide new insights into osteosarcoma diagnosis and prognosis prediction.

MATERIALS AND METHODS

Data Acquisition

The workflow chart of this study is shown in **Figure 1**. We extracted the RNA sequencing (RNA-seq) data and the corresponding clinical information of 88 osteosarcoma patients from the TARGET database (https://ocg.cancer.gov/programs/target). The RNA-seq data of 396 normal human muscle tissue samples were obtained from the GTEx database (https://

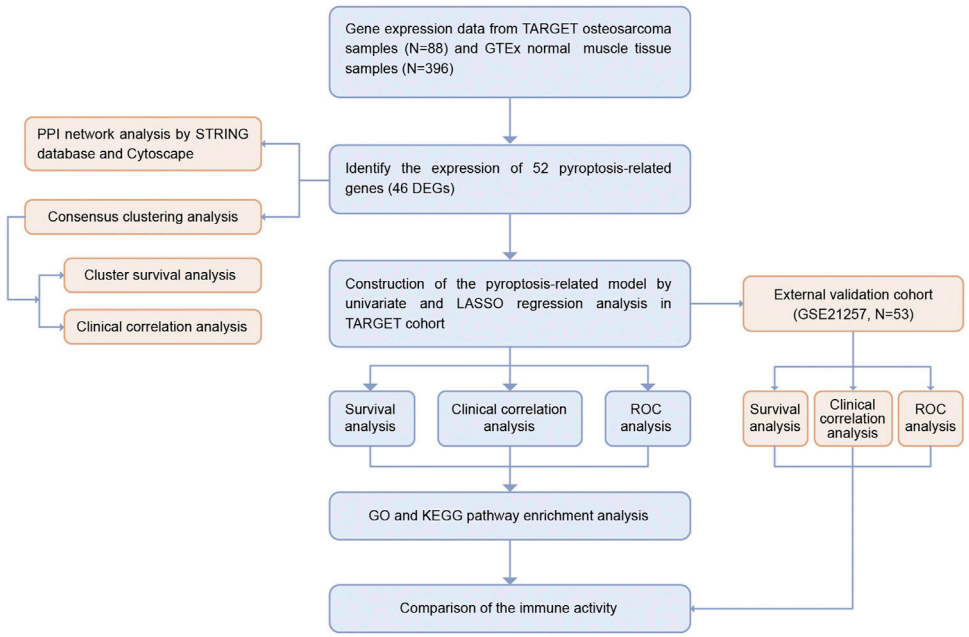


FIGURE 1 | Flow chart of the study.

xenabrowser.net/datapages/). Both data types were HTseq-FPKM, and all gene expression levels were processed with $\log^2$ (FPKM + 1). The independent cohort GSE21257, which contained 53 osteosarcoma samples, was downloaded from Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE21257).

Identification of DEPRGs

We obtained 52 pyroptosis-related genes (PRGs) from prior reviews (Xia et al., 2019; Zhou and Fang, 2019; Li et al., 2020; Ju X. et al., 2021; Wu et al., 2021; Ye et al., 2021) and MSigDB database v7.4 (Subramanian et al., 2005) (listed in **Supplementary Table S1**). We identified DEPRGs between tumor and normal tissues using the "limma" package, with a p-value < 0.05. A protein-protein interaction (PPI) network of all DEPRGs was obtained by STRING database (http://www.string-db.org/). We used Molecular Complex Detection (MCODE), a plugin for Cytoscape, to cluster the genes and find a densely connected area based on the following criteria: degree cut-off = 2, haircut on, node score cut-off = 0.2, Max depth = 100, k-score = 2, score ≥ 5, and node ≥ 10.

Consensus Clustering Analysis

We downloaded all clinical data from the TARGET dataset and further analyzed a total of 85 patients with survival time and status. We performed consensus clustering analysis based on the clinical characteristics of osteosarcoma patients in the TARGET dataset using the "ConsensusClusterPlus" package. The clustering index "k" was increased from 2 to 10 to identify the clustering index with the minor interference and the greatest difference between clusters.

Construction of a Pyroptosis-Related Scoring Signature

We conducted univariate Cox analysis with the "survival" package to screen for prognosis-related DEPRGs and set 0.1 as the threshold p-value for omission prevention (Ye et al., 2021). We then conducted the LASSO Cox regression analysis to narrow the risk of overfitting to develop a prognostic signature using "glmnet" package. The TARGET osteosarcoma patients were divided into low and high PRS-score groups based on the median PRS-score, and the PRS-score formula was as follows: PRS-score = Σ ($\beta i \times Expi$) (β: coefficients, Exp: gene expression level). We created a Kaplan–Meier survival curve using the R "survival" and "survminer" packages to determine the OS time between the two subgroups. The principal component analysis (PCA) based on the signature was performed using the R package "Rtsne" and "ggplot2". The specificity and sensitivity of this prognostic signature were determined by the receiver operating characteristic (ROC) curve constructed with the "SurvivalROC" package. In addition, we identified copy-number alterations and performed mutation analysis of the risk genes in sarcomas using the cBioportal database (http://www.cbioportal.org/). Additionally, 53 osteosarcoma patient samples from the GSE21257 dataset were used to verify the reliability of the prognostic model.

Independent Prognostic Analysis and Clinical Correlation Analysis

We extracted clinical information (gender, age, and metastasis status) of patients in the TARGET cohort. We implemented the "survival" package to conduct both univariate and multivariate

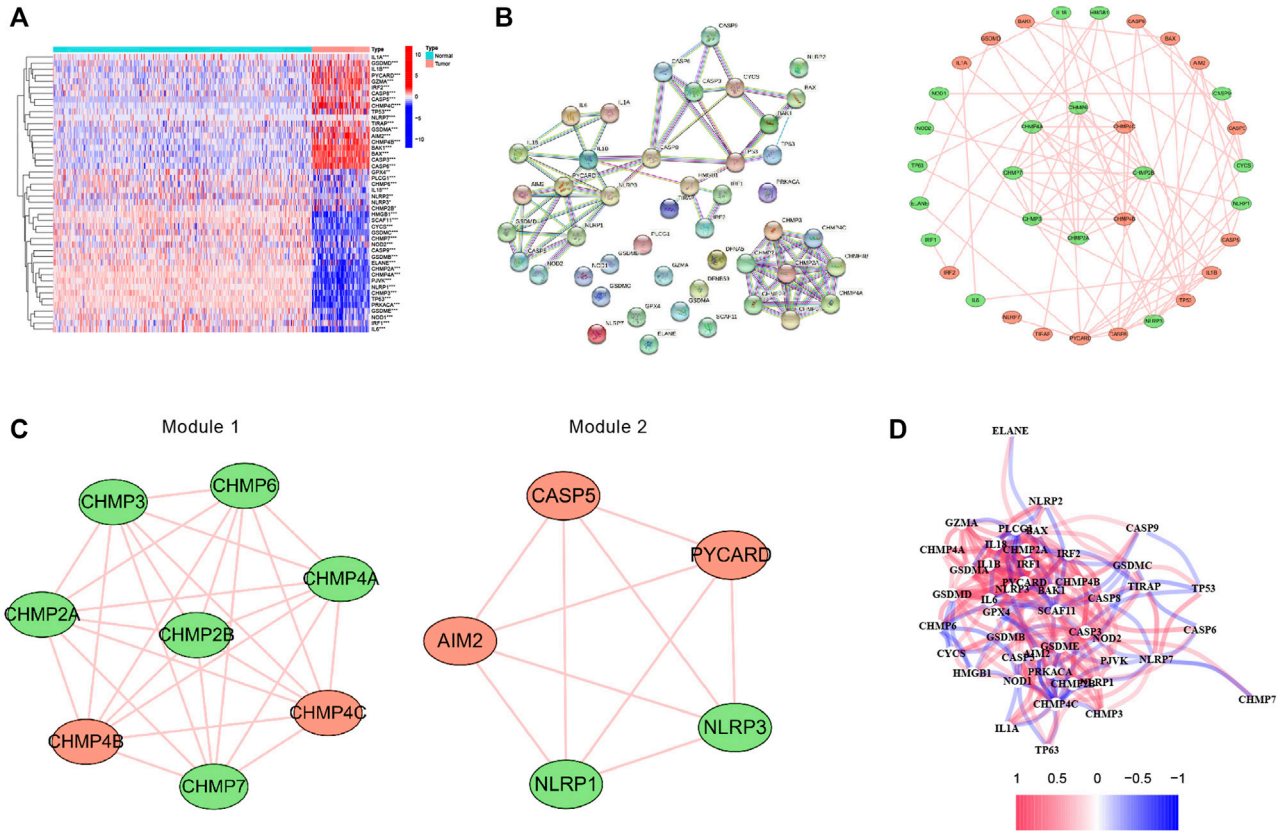


FIGURE 2 | Expression and interconnectedness of the pyroptosis-related genes in osteosarcoma. **(A)** The heatmap showed the differential expressed PRGs between human osteosarcoma samples and normal muscle tissues (red: high expression level, blue: low expression level). **(B)** PPI network of differentially expressed PRGs (The red nodes indicate upregulated PRGs and the green nodes indicate downregulated PRGs) **(C)** Critical modules from the PPI network. **(D)** The correlation network of the differential expressed PRGs (red lines indicate positive correlation and blue lines indicate negative correlation). PRGs, pyroptosis-related genes.

Cox regression analysis to assess the independence of the PRS-score from other clinical variables. The R "RMS" package was then used to generate nomograms to predict survival in patients with osteosarcoma over the course of 1, 3, and 5 years. Additionally, osteosarcoma patients were divided into two subgroups according to age (≤ 18 or > 18 years old), gender (female or male), and metastasis status (M0 and M1). The R "Beeswarm," "limma," and "pheatmap" package was used to assess the correlation between the PRGs involved in the prognostic signature and clinical parameters mentioned above.

Functional Enrichment Analyses

We applied the "limma" R package to identify differentially expressed genes (DEGs) in the PRS-score-classified subgroups, with a false discovery rate (FDR) < 0.05 and absolute value of the log2 fold change (|log2FC|) ≥ 1 as a threshold. We implemented the "clusterProfiler" package to conduct the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis based on the DEGs between different PRS-score subgroups, with an adjusted *p*-value (adj. P) < 0.05. Subsequently, the "GSVA" package was used to conduct the single sample Gene Set Enrichment Analysis (ssGSEA) to calculate the enrichment scores of immunological cells and functions.

Analysis of the Immune Microenvironment Features and Immune Response

Immunoscore and stromal scores for each osteosarcoma patient were obtained using the "estimate" and "limma" packages and were used to derive tumor purity. Using the "ggpubr" and "limma" packages, we assessed the differential expression of immune checkpoints (CTLA4, PDL1, LAG3, TIGIT, TIM3, PDCD1, IDO1, and TDO2) between subgroups to estimate the predictive power of the signature for immunotherapy response.

Statistical Analysis

We executed all statistical analyses with R software (v4.0.5). The threshold for statistical significance was taken as $p < 0.05$ if it was not explicitly stated.

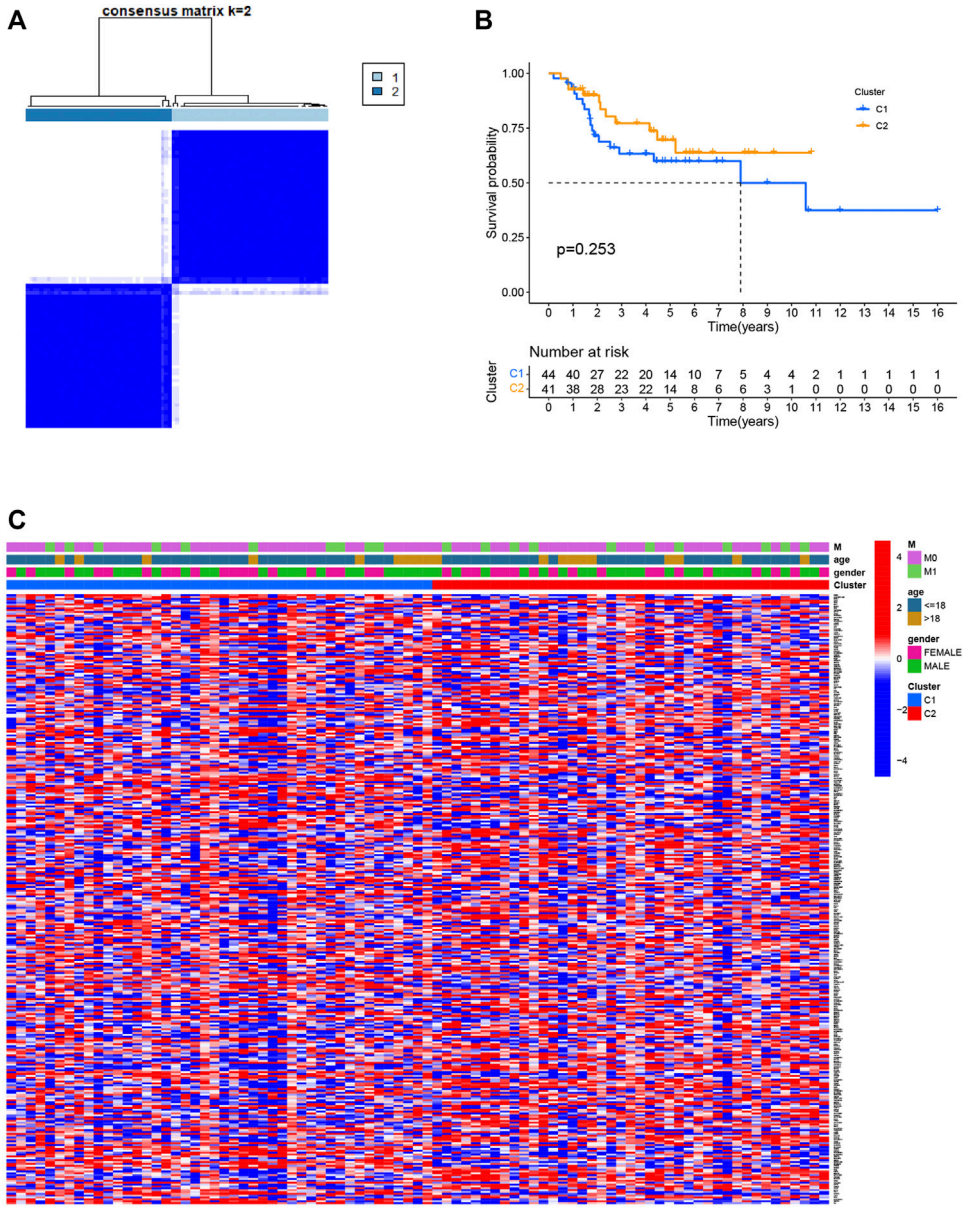


FIGURE 3 | Classification of osteosarcoma patients based on pyroptosis-related regulators. **(A)** Consensus clustering of osteosarcoma patients for k = 2. **(B)** The prognostic analysis between the two pyroptosis-related clusters. **(C)** Heatmap of the differentially expressed genes and clinical characteristics between the two pyroptosis-related clusters.

RESULTS

DEPRGs in Human Osteosarcoma and Normal Tissues

The expression levels of 52 PRGs were compared in the human osteosarcoma samples and normal muscle tissues, and we detected 19 DEPRGs that were upregulated and 27 DEPRGs that were down-regulated using our threshold criteria (p value < 0.05) (**Figure 2A**). The PPI network of DEPRGs created with the minimum required interaction score > 0.9 is presented in **Figure 2B**. We then screened out the two most crucial network modules using MCODE (**Figure 2C**) and drew the correlation network of the differentially expressed PRGs (**Figure 2D**).

Identification of Subgroups Based on PRGs by Consensus Clustering

Consensus clustering was used to separate all 85 osteosarcoma patients into subgroups according to the expression of PRGs. By increasing the clustering index "k" from 2 to 10, we found that k = 2 seems to be the optimal point to identify the smallest

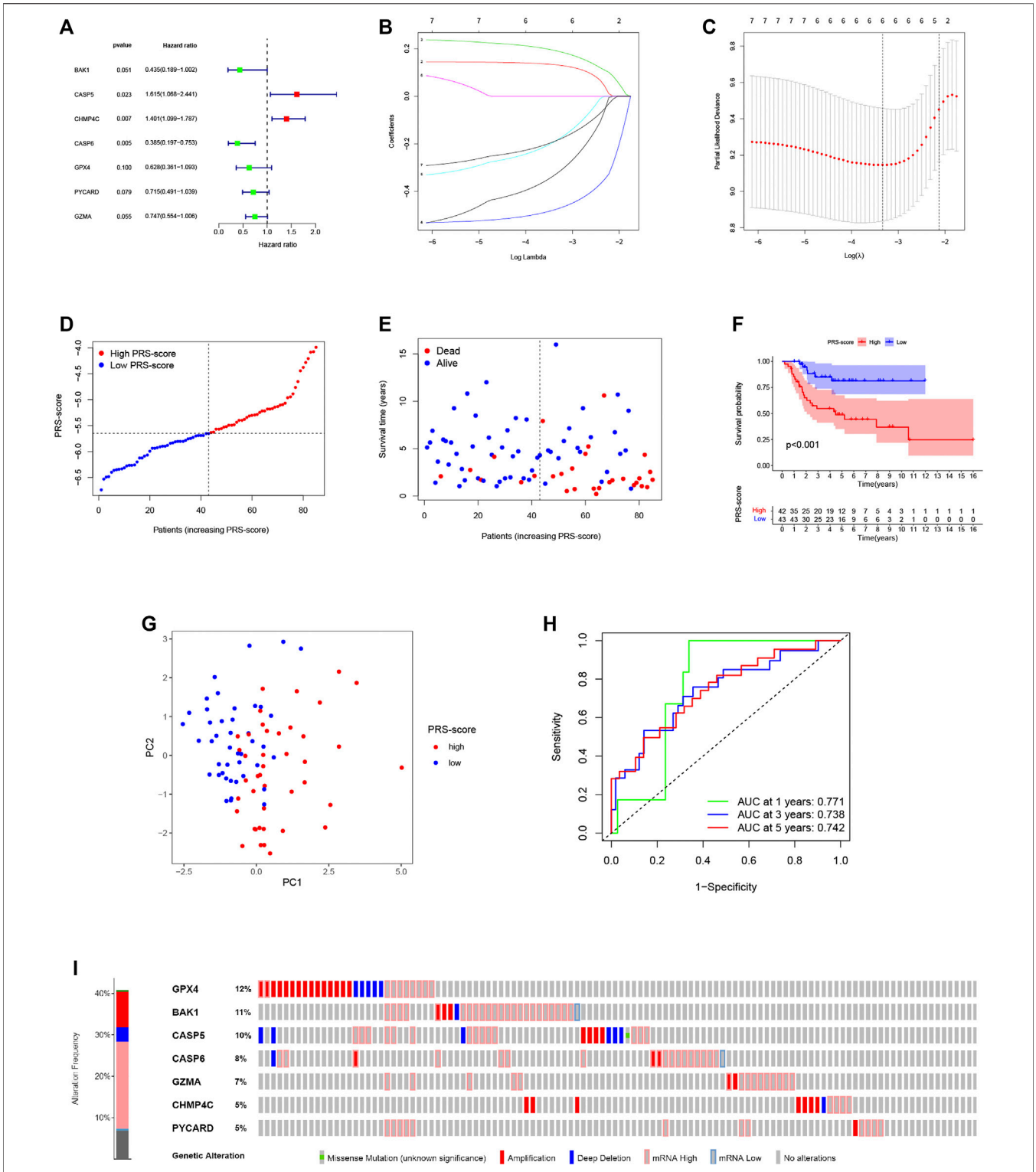


FIGURE 4 | Construction of the pyroptosis-related prognostic signature for osteosarcoma. **(A,B)** Cox regression analysis of pyroptosis-related genes. **(A)** Univariate Cox regression analysis. **(B)** LASSO Cox regression analysis. **(C)** Selection of the optimal penalty parameter for LASSO regression. **(D)** The PRS-score distribution of the patients with osteosarcoma in the TARGET cohort **(E)** The survival status and survival time distribution of the patients with osteosarcoma in the TARGET cohort. **(F)** Kaplan–Meier curves of the high and low PRS-score subgroups in the TARGET cohort. **(G)** PCA plot based on the PRS-scores in the TARGET cohort. **(H)** Time-dependent ROC curve for predicting the 1-, 3-, and 5-year overall survival in the TARGET cohort. **(I)** Genomic alterations of hub genes.

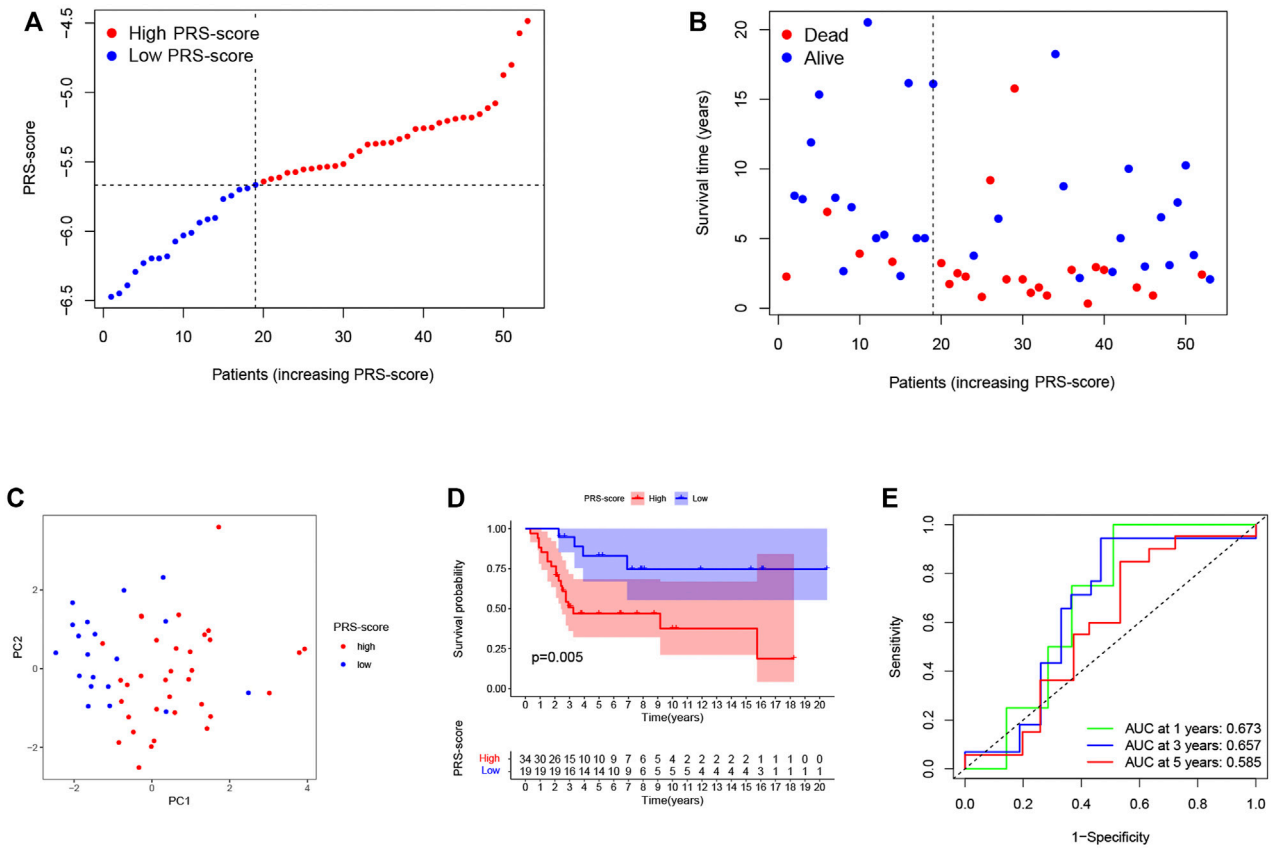


FIGURE 5 | Validation of the prognostic signature in the GEO cohort. **(A)** The PRS-score distribution of the patients with osteosarcoma in the GEO cohort. **(B)** The survival status and survival time distribution of the patients with osteosarcoma in the GEO cohort. **(C)** PCA plot based on the PRS-scores in the GEO cohort. **(D)** Kaplan–Meier curves of the high and low PRS-score subgroups in the GEO cohort. **(E)** Time-dependent ROC curve for predicting 1-, 3-, and 5-year overall survival in the GEO cohort.

interferences and the most significant differences between clusters (**Figure 3A**). Consequently, patients with osteosarcoma in the training group were classified into two clusters. However, a comparison of overall survival between the two clusters revealed no significant difference ($p = 0.253$, **Figure 3B**). We also plotted a heatmap to express the differences in gene expression and clinical characteristics, including age (≤ 18 or > 18 years old), gender (male or female), and metastasis status (metastatic, non-metastatic) between the clusters, but we found there are little differences (**Figure 3C**).

Construction of the PRG-Based Prognostic Signature

To construct a pyroptosis-related prognostic model, we further screened seven-candidate prognostic PRGs by univariate Cox regression analysis (**Figure 4A**). Of the seven prognostic PRGs, CASP5 and CHMP4C were regarded as high-risk genes based on their HRs, whereas BAK1, CASP6, GPX4, PYCARD, and GZMA were regarded as low-risk genes. Subsequently, LASSO Cox regression analysis was performed to construct a 6-gene signature according to the optimum penalty parameter (λ) value (**Figures 4B,C**). We then divided the patients in the TARGET cohort into high and low scoring subgroups based on a composite signature score termed the "PRS-score" (PRS-score = [BAK1 expression × (−0.325)] + [CASP5 expression × (0.132)] + [CHMP4C expression × (0.191)] + [CASP6 expression × (−0.475)] + [GPX4 expression × (−0.185)] + [GZMA expression × (−0.185)]). The PRS-scores, survival status, and survival time in the two groups of patients are shown in **Figures 4D,E**. The results showed that patients with higher PRS-scores had worse prognoses than patients with lower PRS-scores. Kaplan-Meier curves showed that the patients in the high PRS-score group had worse OS than the patients in the low PRS-score group ($p < 0.001$; **Figure 4F**). Analyses of PCA revealed that high and low PRS-score patients were separated into two clusters (**Figure 4G**). To assess the accuracy of the signature, we then constructed a time-dependent ROC curve. We found the area under the ROC curve (AUC) was 0.771 for 1-year OS, 0.738 for 3-year OS, and 0.742 for 5-year OS, providing evidence that this six-gene prognostic model performed well as a predictor of OS (**Figure 4H**). Mutations and copy number alterations of the six hub genes (BAK1, CASP6, GPX4, PYCARD, GZMA, CASP5, and CHMP4C) were analyzed together using the cBioportal database.

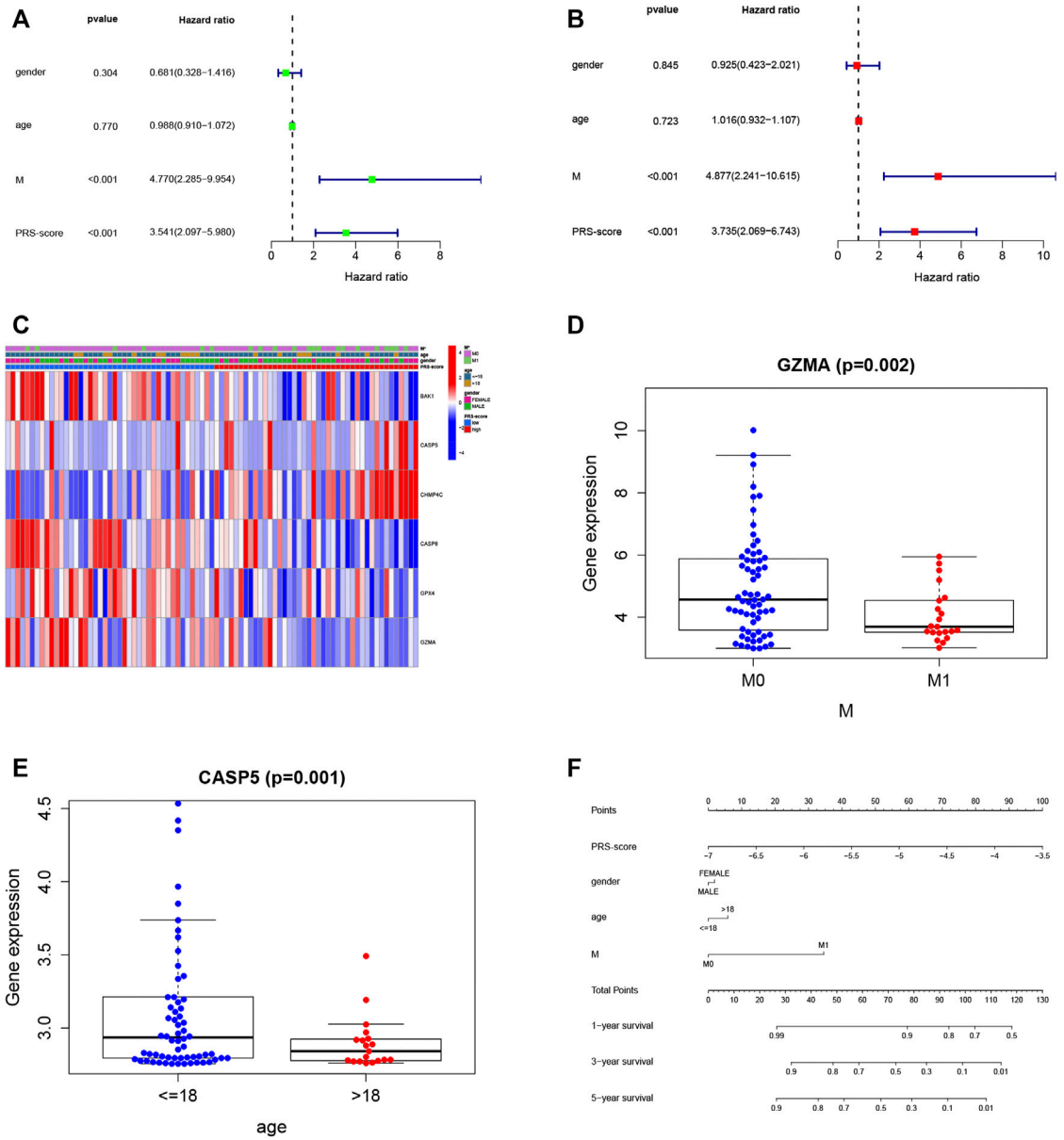


FIGURE 6 | Independent prognosis analysis and clinical utility. **(A,B)** Cox regression analysis of pyroptosis-related genes **(A)** Univariate Cox regression analysis **(B)** Multivariate Cox regression analysis **(C)** Heatmap (blue: low expression level; red: high expression level) of the correlation between clinical features and the risk groups (*$p < 0.05$) **(D)** Relationship between GZMA and metastasis. **(E)** Relationship between CASP5 and age category. **(F)** A prognostic nomogram based on the PRG-related model for prediction of 1-, 3-, and 5-year survival rates.

These six hub genes were altered in 99 of 241 samples (41%) (**Figure 4I**). Since the frequency of mutations in GPX4 and BAK1 exceeded 10%, we hypothesized that these two genes might be key therapeutic targets (**Figure 4I**).

Validation of the PRG-Based Prognostic Signature

To reliability of this six-gene prognostic signature, a total of 53 patients from GSE21257 were used as the test set. Based on the median cut-off of the PRS-score in the TARGET cohort, patients with osteosarcoma in the GEO cohort were separated into high (n = 34) and low (n = 19) scoring groups (**Figure 5A**). The survival time and survival status distribution showed that patients in the low PRS-score subgroup had a higher possibility of surviving (**Figure 5B**). The PCA of the two subgroups showed a clear separation (**Figure 5C**). Furthermore, Kaplan-Meier analysis revealed that osteosarcoma patients with high PRS-scores had a significantly poorer prognosis than those with low PRS-

TABLE 1 | The relationship between PRS-scores and clinical characteristics.

Id	Gender (female, male) t (p)	Age (≤ 18, > 18) t (p)	M stage (M0, M1) t (p)
BAK1	−0.182 (0.856)	1.324 (0.197)	0.091 (0.928)
CASP5	1.257 (0.213)	**3.287(0.001)**	−0.08 (0.936)
CHMP4C	0.84 (0.404)	1.609 (0.115)	−0.568 (0.574)
CASP6	−0.724 (0.472)	0.128 (0.899)	1.965 (0.056)
GPX4	0.062 (0.951)	0.32 (0.751)	1.664 (0.107)
GZMA	0.341 (0.734)	1.434 (0.160)	**3.293(0.002)**
PRS-scores	0.742 (0.461)	−0.093 (0.926)	**−2.58(0.015)**

t, t value from Student's t test; p: p-value from Student's t test. Bold indicates statistical significance, $p < 0.05$.

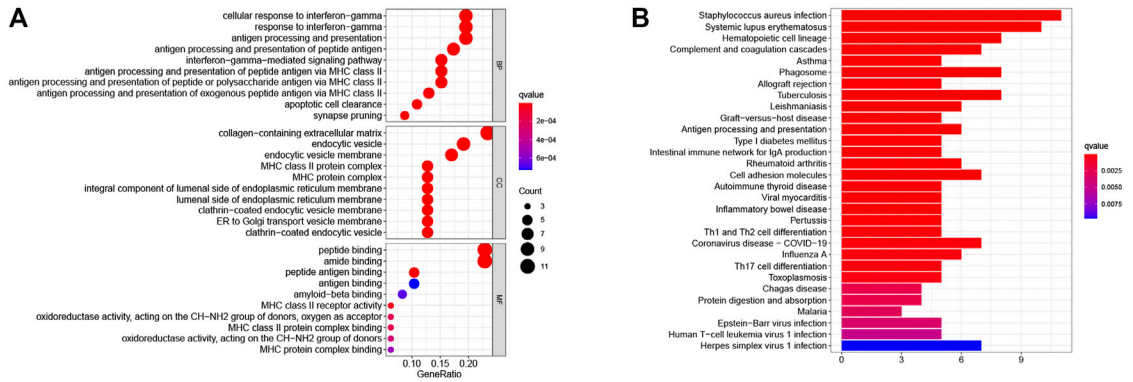


FIGURE 7 | Functional enrichment analysis of DEGs between the two pyroptosis-related subgroups. **(A)** GO enrichment analysis of DEGs based on PRS-score, including BP, CC, and MF. **(B)** KEGG pathway enrichment analysis of DEGs based on PRS-score. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological processes; CC, cell component; MF, molecular function.

scores (**Figure 5D**), with AUC = 0.673, 0.657, and 0.585 for 1, 3, and 5 years survival, respectively (**Figure 5E**).

Independent Prognostic Value and Clinical Utility of the Prognostic Signature

We then utilized univariate and multivariable Cox regression analyses to evaluate the independent prognostic value of the model with other clinical features. The univariate Cox analysis indicated that the PRS-score (HR = 3.541, 95% CI = 2.097–5.980, $p < 0.001$) and M-stage (HR = 4.770, 95% CI = 2.285–9.954, $p < 0.001$) were significantly associated with OS (**Figure 6A**). The multivariate Cox analysis confirmed that the PRS-score (HR = 3.735, 95% CI = 2.069–6.743, $p < 0.001$) and M-stage (HR = 4.877, 95% CI = 2.241–10.615, $p < 0.001$) were independent factors affecting the prognosis of osteosarcoma patients (**Figure 6B**). We then plotted a clinical information-related heatmap for the TARGET cohort and found significant differences in M-stage distribution between low- and high-scoring subgroups (**Figure 6C**). The results of clinical correlation analysis showed that the M stage of osteosarcoma patients decreased with increasing GZMA expression (**Figure 6D**, $p < 0.01$), while osteosarcoma patients with high CASP5 expression were younger (**Figure 6E**, $p < 0.01$), and all results are shown in **Table 1**. Additionally, a pyroptosis-related signature-based nomogram showed that the OS of patients at 1, 3, and 5 years decreased with increasing PRS-score (**Figure 6F**).

Functional Analysis of DEGs Based on PRS-Score

To further investigate differences in PRS-score-classified subgroups, we identified 34 genes that were down-regulated and 14 genes that were up-regulated in the high PRS-score subgroup compared with the low PRS-score subgroup in the TARGET group (**Supplementary Table S2**). GO analysis revealed that the 48 DEGs were mainly involved in the cellular response to interferon-gamma, MHC class II protein complex, peptide binding, and amide binding (**Figure 7A**). According to the KEGG pathway analysis, these DEGs were primarily associated with staphylococcus aureus infection, systemic lupus erythematosus, hematopoietic cell lineage, and complement and coagulation cascades (**Figure 7B**).

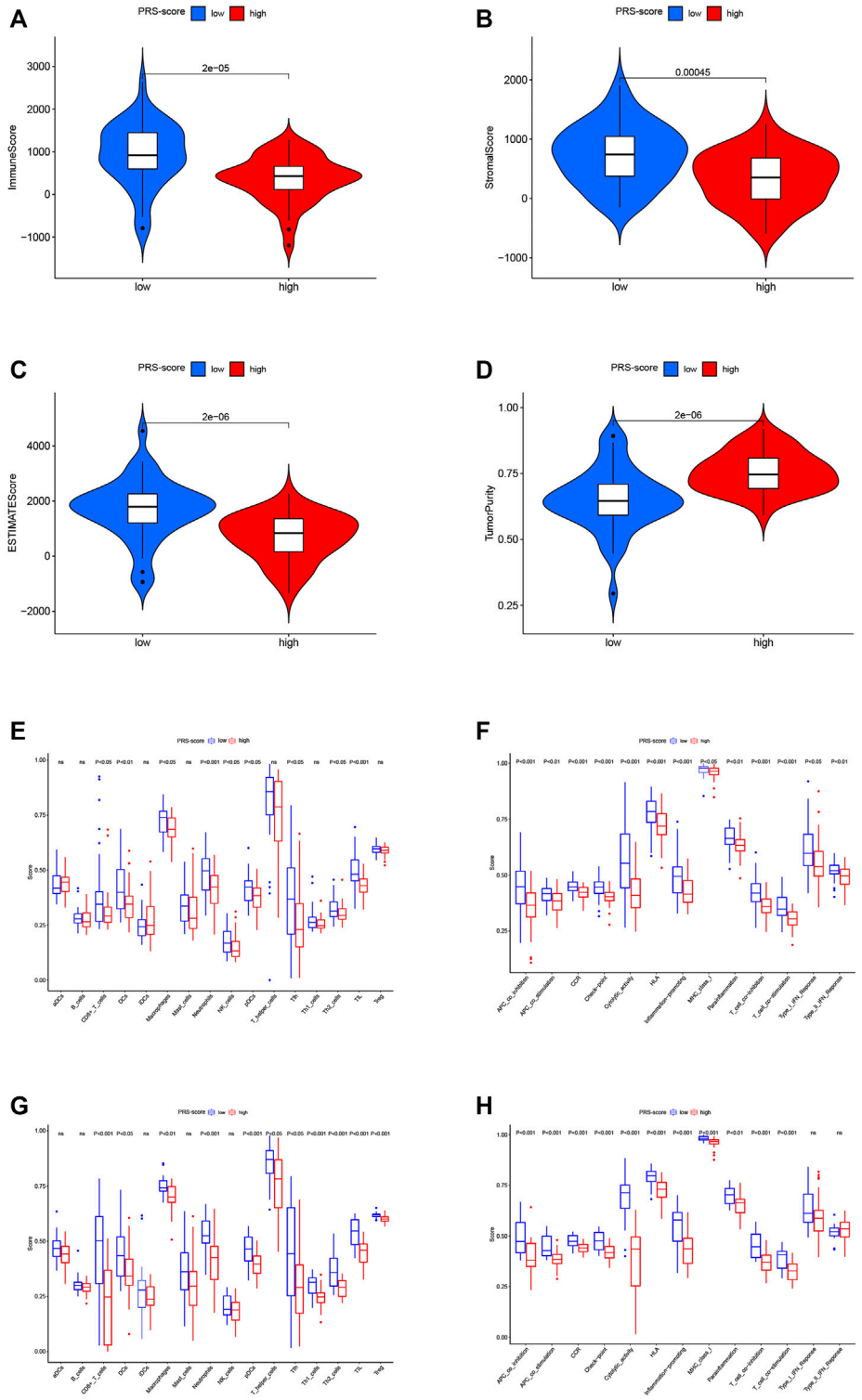


FIGURE 8 | Immune characteristics analysis of the prognostic signature. **(A)** Immune scores between high and low PRS-score groups**. (B)** Stromal scores between high and low PRS-score groups **(C)** ESTIMATE scores between high and low PRS-score groups. **(D)** Tumor purity between high and low PRS-score groups **(E)** Comparisons of the level of immune cell infiltration between high and low PRS-score groups in the TARGET cohort. **(F)** Comparisons of immune functions between high and low PRS-score groups in the TARGET cohort. **(G)** Comparisons of the level of immune cell infiltration between high and low PRS-score groups in the GEO cohort. **(H)** Comparisons of immune functions between high and low PRS-score groups in the GEO cohort.

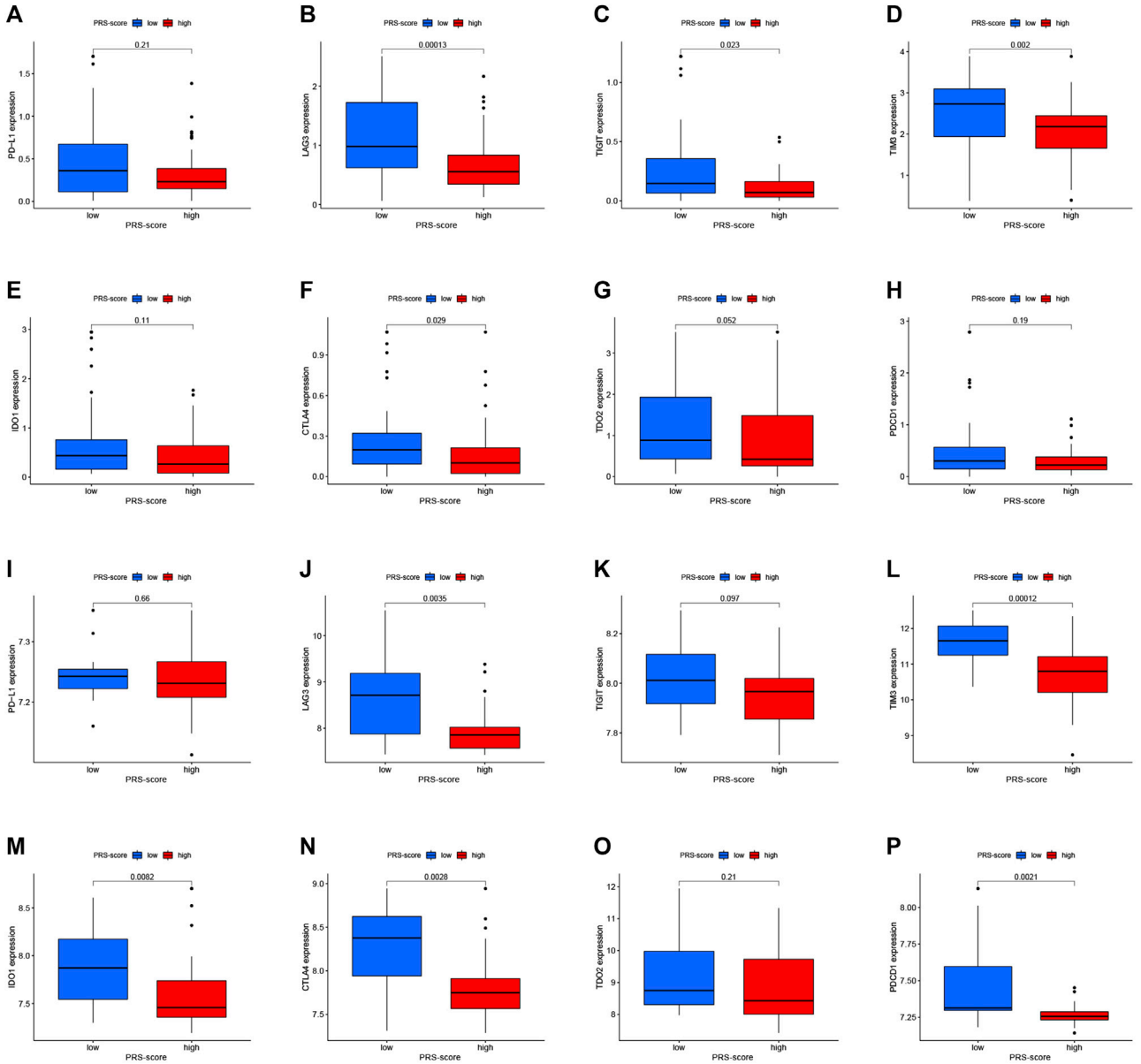


FIGURE 9 | Immune checkpoint molecules expression analysis. **(A–H)** The expression levels of Immune checkpoint molecules, including PD-L1 **(A)**, LAG-3 **(B)**, TIGIT **(C)**, TIM-3 **(D)**, IDO1 **(E)**, CTLA-4 **(F)**, TDO2 **(G)**, and PDCD1 **(H)** between high and low PRS-score groups in the TARGET cohort. **(I–P)** The expression levels of Immune checkpoint molecules, including PD-L1 **(I)**, LAG-3 **(J)**, TIGIT **(K)**, TIM-3 **(L)**, IDO1 **(M)**, CTLA-4 **(N)**, TDO2 **(O)**, and PDCD1 **(P)** between high and low PRS-score groups in the GEO cohort.

Analysis of Immune Microenvironment Characteristics Between Subgroups

Several studies have shown that the tumor immune microenvironment correlates strongly with malignant behavior; thus, we investigated the unique features of the tumor microenvironment (TME) to distinguish between the two subgroups of patients. Based on the ESTIMATE algorithm, the overall level of immune cell infiltration and tumor purity were examined. As shown in **Figures 8A–D**, the immune score, stromal score, and ESTIMATE score were significantly higher in the low scoring group than in the high scoring group, while the tumor purity was lower. We then explored the distribution patterns of infiltrating immune cells in different subgroups using the ssGSEA algorithm. In the TARGET cohort, the patients in the high PRS-score group had lower levels of tumor infiltration by $CD8^+$ T cells, dendritic cells (DCs), macrophages, neutrophils, natural killer cells, plasmacytoid dendritic cells (pDCs), Th2 cells, Tfh cells, and tumor-infiltrating lymphocytes (TILs) compared with the patients in

the low PRS-score group (**Figure 8E**). All 13 immune functions were down-regulated in the patients in the high PRS-score group in comparison with the patients in the low PRS-score group (**Figure 8F**). In the GEO cohort, compared with the patients in the low PRS-score group, the patients in the high PRS-score group had lower levels of tumor infiltration by immune cells, including $CD8^+$ T cells, DCs, macrophages, neutrophils, pDCs, TILs, T regulatory, Tfh, Th1, and Th2 cells (**Figure 8G**). Moreover, in contrast to the type-1 and type-2 interferon response pathways, the other 11 immune pathways had lower activity in the high PRS-score group than in the low PRS-score group (**Figure 8H**). Our investigation showed that PRS-scores were associated with immune characteristics and that elevated immune activity in the low-scoring samples may contribute to the antitumor effect in osteosarcoma.

In addition, we analyzed the changes in immune checkpoint expression between the high and low PRS-score groups. **Figures 9A–H** shows that in the TARGET cohort, LAG3 (p = 1.3e-04), TIGIT (p = 0.023), TIM3 (p = 0.002), and CTLA4 (p = 0.029) expressions were down-regulated in the high-scoring group in comparison to the low-scoring group. On the other hand, as the PRS-score increased, the expression of LAG3 (p = 0.0035), TIM3 (p = 1.2e-04), IDO1 (p = 0.0082), CTLA4 (p = 0.0028), and PDCD1 (p = 0.0021) in patients with osteosarcoma also decreased in the GEO cohort (**Figures 9I–P**).

DISCUSSION

Pyroptosis, a form of programmed cell death, was found to play a dual role in both promoting and inhibiting the growth of different tumor cells (Loveless et al., 2021). Several recent studies have highlighted the relevance of pyroptosis-related genes as candidate biomarkers for prognosis and therapeutic response in patients with different cancer types (Ju A. et al., 2021; Lin W. et al., 2021; Shao et al., 2021; Ye et al., 2021). In the current study, we identified the mRNA levels of 52 pyroptosis-related genes in osteosarcoma and normal tissues based on public databases and found that most of these genes were differentially expressed. However, DEPRGs-based consensus clustering analysis produced two clusters that showed no significant differences in clinical characteristics. Subsequently, we performed univariate and LASSO Cox regression analyses to further identify six prognosis-related RPGs. To further explore their biological function and clinical significance, we also performed survival and ROC analyses to develop an accurate pyroptosis-related prognostic signature in osteosarcoma. Subsequently, ssGSEA found that the high-scoring group had lower levels of immune infiltration and fewer immune-related pathways than the low-scoring group. These results suggest that the novel pyroptosis-related genes signature has the potential to predict prognosis accurately and could provide new diagnostic biomarkers and therapeutic targets for patients with osteosarcoma.

As a result of the present study, we constructed a 6-gene pyroptosis-related signature, including BAK1, CASP5, CASP6, GPX4, GZMA, and CHMP4C. Notably, six genes involved in this signature have been implicated in apoptotic pathways as well (Li et al., 2015; Skotte et al., 2017; Wu et al., 2019; Zhou et al., 2020; Darweesh et al., 2021; Ding et al., 2021). Following apoptotic signals, caspase 8 and caspase 3 initiate pyroptosis by processing GSDMC and GSDME, respectively (Liu et al., 2021). The close relationship between pyroptosis and apoptosis may explain the dual role of these genes. Caspase 5 is an essential player in canonical or noncanonical inflammasome-induced pyroptosis. Upon activation, caspase-5 can act on the GSDMD, leading to the formation of cell membrane pores. Activated caspase-5 can also interact with caspase-1 to promote its activation, and the latter cleaves the precursors of IL-1+ and IL-18 to form active IL-1+ and IL-18, which are released through the channels formed by GSDMD-cNT and lead to pyroptosis (Kayagaki et al., 2015; Xia et al., 2019). Studies have reported that caspase-5 is associated with various malignancies, including gastric cancer, cervical cancer, lung cancer, and human glioblastoma (Babas et al., 2010; Zhou et al., 2018; Wang et al., 2019). Caspase-6 plays a vital role in promoting cell death, ZBP1-mediated inflammasome activation, and host defense during IAV infection (Zheng et al., 2020; Zheng and Kanneganti, 2020). In addition, caspase-6 can also be involved in cancer progression by regulating tumor apoptosis and metastasis (Capo-Chichi et al., 2018). GPX4 was found to negatively regulate Gasdermin D-mediated pyroptosis in lethal polymicrobial sepsis by reducing lipid peroxidation; in contrast, conditional GPX4 knockdown in myeloid cells triggers macrophage pyroptosis with caspase-1/caspase-11-GSDMD-phospholipase C gamma 1 axis. (Kang et al., 2018). Chen et al. (2021a) found that circKIF4A promoted papillary thyroid tumors by sponging miR-1231 and upregulating GPX4 expression. GPX4 is also a ferroptosis-related factor playing an essential role in iron-dependent oxidative cell death driven by lipid peroxidation (Chen X. et al., 2021). Lin H. et al. (2021) discovered that upregulation of HMOX1 to inhibit GPX4 expression induced ferroptosis in osteosarcoma cells by increasing reactive oxygen species levels, malondialdehyde levels, and intracellular ferric ion level. GZMA from cytotoxic lymphocytes enhances antitumor immunity and promotes tumor clearance by cleavage of GSDMB triggering pyroptosis (Zhou et al., 2020). On the other hand, GZMA acts as a pro-inflammatory cytokine to promote cancer development (van Daalen et al., 2020); for instance, GZMA deficiency inhibits colon cancer development and inflammatory response in colon tissue through the NF-κB-IL-6-pSTAT3 axis (Santiago et al., 2020). The polymorphism of CHMP4C increased the cancer susceptibility and was imbalanced in many cancers, including lung, ovarian, prostate, and cervical cancers (Lin S. L. et al., 2020). Another study showed that CHMP4C is also an autophagy-related gene, and its participation in the construction of risk models could effectively predict the prognosis of cervical cancer patients and help develop precise treatment strategies (Shi et al., 2020). Notably, similar to CHMP4C, BAK1 was found to be an apoptosis and pyroptosis-related gene (Cowan et al., 2020; Deo et al., 2020). BAK1 is a member of the Bcl-2 family and can induce mitochondria-mediated apoptosis via regulating the release of cytochrome c (Vervliet et al., 2016). Recent studies have shown that miR-125b, miR-410, and miR-103a-3p could all directly target BAK1 to inhibit apoptosis, and upregulation of BAK1

may contribute to the treatment of cisplatin-resistant non-small cell lung cancer (Wen et al., 2018; Wang H. et al., 2021; Zhang et al., 2021). A prognostic signature based on 14 genes, including BAK1, was able to predict the survival outcome for patients with osteosarcoma (Qi et al., 2021). We also found that PYCARD, although not included in the construction of the model, was also associated with patient outcomes. PYCARD is an adaptor protein that helps form inflammasomes, which contribute to inflammation by promoting the release of the active IL-1β and IL-18 (Hoffman and Wanderer, 2010; Protti and De Monte, 2020). Inflammation is commonly thought to contribute to driving tumor growth, metastasis, and immune escape; for example, IL-1 promotes tumor angiogenesis, recruitment of myeloid cells and contributes to tumor metastasis by recognizing endothelial cell adhesion molecules (Mantovani et al., 2018; Karin and Shalapour, 2021). On the other hand, PYCARD was found to be silenced by promoter methylation in various cancer cells, suggesting its anti-tumor role as a pro-apoptotic factor (Agrawal and Jha, 2020). These studies further confirmed the potential prognostic value of the identified pyroptosis-related genes in osteosarcoma. However, the exact mechanism of their involvement in pyroptosis in osteosarcoma needs to be verified by further *in vivo* and *in vitro* experiments.

The enrichment analysis results showed that the DEGs between high and low PRS-score subgroups were mainly enriched in interferon-gamma mediated signaling pathways, antigen processing, and peptide antigens presented via MHC class II, peptide binding. The MHC-II is the critical component of adaptive anti-tumor immunity, and its upregulation is closely associated with increased levels of interferon-gamma in tumors (Dubrot et al., 2014; Cook et al., 2021). During inflammation, epithelial cells could act as accessory antigen-presenting cells along with the expression of MHC-II (Ghasemi et al., 2020). Tumor-specific MHC-II expression is associated with better prognosis, T-cell infiltration, higher levels of Th1 cytokines, and sensitivity to anti-PD-1 therapies (Johnson et al., 2020). Lu et al. (2017), used the adoptive transfer of MHC-II-restricted tumor-reactive T cells in patients with metastatic cancer (which contained patients with osteosarcoma) and achieved different degrees of tumor regressions in these patients. Coincidentally, the ssGSEA results indicated lower levels of principal anti-tumor infiltrating immune cells in the high PRS-score group, providing further evidence that these genes may play a role in anti-tumor immunity. Studies have shown that chimeric antigen receptor T-cell immunotherapy, a potent option for drug-resistant tumors, has transformed the treatment of drug-resistant hematologic malignancies yet remains largely ineffective against solid tumors, which may be related to the tumor immune microenvironment, the stromal barrier, and the lack of surface tumor-specific targets (Titov et al., 2021). Therefore, we used the ESTIMATE algorithm to examine the distribution of immune scores, stromal scores, and tumor purity in osteosarcoma patients in high and low PRS-score groups. We found that the low-scoring group showed higher immune scores, ESTIMATE scores. Consistent with these results, the high-scoring group had high tumor purity. The PRS-score may help assess the immune microenvironment features of patients and thus predict their sensitivity to immunotherapy, which will help to guide individualized anti-tumor treatment strategies. Finally, we evaluated the differences in immune checkpoint expression between the two subgroups to determine whether patients would benefit from immune checkpoint inhibitor therapy.

In previous studies, several prognostic signatures have been constructed from different perspectives to predict the prognosis of oeosarcoma. Jiang et al. (2021) created a hypoxia gene-based signature to predict the survival in childhood osteosarcoma. Wang et al. developed a new classification system of osteosarcoma based on immune features and identified TYROBP as a key immune regulatory gene (Wang X. et al., 2021). Qi et al. identified a prognostic signature of osteosarcoma based on 14 autophagy-related genes that can guide clinical decisions in treating osteosarcoma (Qi et al., 2021). Nonetheless, no research has so far concentrated on PRGs-related models, and the current study was designed to fill the vacancy in PRGs-based models for predicting outcomes. Of course, there are inevitably some limitations to this study. Firstly, the verification cohort has a relatively small sample size due to the inherent property of osteosarcoma. Secondly, it lacks experimental work, and the molecular mechanisms of its specific involvement still need further study.

In conclusion, we have developed a novel prognostic model based on six pyroptosis-related genes through comprehensive and systematic bioinformatics analysis, providing an essential foundation for future studies of the association between pyroptosis-related genes and immunity in osteosarcoma.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

YZ and RH conceived of the research. XL, LM, PJ, and CN collected the data. YZ, RH, XL, XZ, ZY, and CC interpreted the data. YZ and RH drafted the manuscript. DL and QZ critically revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 81601931; 81672229), the Natural Science Foundation of Jiangsu Province (BK20150475), the Youth Medical Key Talent Project of Jiangsu (QNRC2016844), “Six One Projects” for high-level health professionals in Jiangsu Province Top Talent Project (LGY2019089), Jiangsu Provincial key research and development program (BE2020679), and the Shenzhen Science and Technology Program (No. KQTD20170810154011370)

ACKNOWLEDGMENTS

The authors would like to give their sincere appreciation to the reviewers for their helpful comments on this article and research groups for the GTEx, TARGET, and CEO, which provided data for this collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.780780/full#supplementary-material

REFERENCES

Agrawal, I., and Jha, S. (2020). Comprehensive Review of ASC Structure and Function in Immune Homeostasis and Disease. *Mol. Biol. Rep.* 47 (4), 3077–3096. doi:10.1007/s11033-020-05345-2

Babas, E., Ekonomopoulou, M. T., Karapidaki, I., Doxakis, A., Betsas, G., and Iakovidou-Kritsi, Z. (2010). Indication of Participation of Caspase-2 and Caspase-5 in Mechanisms of Human Cervical Malignancy. *Int. J. Gynecol. Cancer* 20 (8), 1381–1385. doi:10.1111/IGC.0b013e3181ed7896

Bashash, D., Zandi, Z., Kashani, B., Pourbagheri-Sigaroodi, A., Salari, S., and Ghaffari, S. H. (2021). Resistance to Immunotherapy in Human Malignancies: Mechanisms, Research Progresses, Challenges, and Opportunities. *J. Cel Physiol.*, [Epub ahead of print]. doi:10.1002/jcp.30575

Capo-Chichi, C. D., Cai, K. Q., and Xu, X.-X. (2018). Overexpression and Cytoplasmic Localization of Caspase-6 Is Associated with Lamin A Degradation in Set of Ovarian Cancers. *Biomark Res.* 6, 30. doi:10.1186/s40364-018-0144-9

Chen, W., Fu, J., Chen, Y., Li, Y., Ning, L., Huang, D., et al. (2021a). Circular RNA circKIF4A Facilitates the Malignant Progression and Suppresses Ferroptosis by Sponging miR-1231 and Upregulating GPX4 in Papillary Thyroid Cancer. *Aging* 13 (12), 16500–16512. doi:10.18632/aging.203172

Chen, X., Yu, C., Kang, R., Kroemer, G., and Tang, D. (2021b). Cellular Degradation Systems in Ferroptosis. *Cell Death Differ* 28 (4), 1135–1148. doi:10.1038/s41418-020-00728-1

Chow, T., Wutami, I., Lucarelli, E., Choong, P. F., Duchi, S., and Di Bella, C. (2021). Creating *In Vitro* Three-Dimensional Tumor Models: A Guide for the Biofabrication of a Primary Osteosarcoma Model. *Tissue Eng. B: Rev.* 27, 514–529. doi:10.1089/ten.TEB.2020.0254

Constantinidou, A., Alifieris, C., and Trafalis, D. T. (2019). Targeting Programmed Cell Death -1 (PD-1) and Ligand (PD-L1): A new era in Cancer Active Immunotherapy. *Pharmacol. Ther.* 194, 84–106. doi:10.1016/j.pharmthera.2018.09.008

Cook, K. W., Xue, W., Symonds, P., Daniels, I., Gijon, M., Boocock, D., et al. (2021). Homocitrullination of Lysine Residues Mediated by Myeloid-Derived Suppressor Cells in the Tumor Environment Is a Target for Cancer Immunotherapy. *J. Immunother. Cancer* 9 (7), e001910. doi:10.1136/jitc-2020-001910

Cowan, A. D., Smith, N. A., Sandow, J. J., Kapp, E. A., Rustam, Y. H., Murphy, J. M., et al. (2020). BAK Core Dimers Bind Lipids and Can Be Bridged by Them. *Nat. Struct. Mol. Biol.* 27 (11), 1024–1031. doi:10.1038/s41594-020-0494-5

Darweesh, O., Al-Shehri, E., Falquez, H., Lauterwasser, J., Edlich, F., and Patel, R. (2021). Identification of a Novel Bax-Cdk1 Signalling Complex that Links Activation of the Mitotic Checkpoint to Apoptosis. *J. Cel Sci* 134 (8), jcs244152. doi:10.1242/jcs.244152

Deo, P., Chow, S. H., Han, M.-L., Speir, M., Huang, C., Schittenhelm, R. B., et al. (2020). Mitochondrial Dysfunction Caused by Outer Membrane Vesicles from Gram-Negative Bacteria Activates Intrinsic Apoptosis and Inflammation. *Nat. Microbiol.* 5 (11), 1418–1427. doi:10.1038/s41564-020-0773-2

Ding, Y., Chen, X., Liu, C., Ge, W., Wang, Q., Hao, X., et al. (2021). Identification of a Small Molecule as Inducer of Ferroptosis and Apoptosis through Ubiquitination of GPX4 in Triple Negative Breast Cancer Cells. *J. Hematol. Oncol.* 14 (1), 19. doi:10.1186/s13045-020-01016-8

Dubrot, J., Duraes, F. V., Potin, L., Capotosti, F., Brighouse, D., Suter, T., et al. (2014). Lymph Node Stromal Cells Acquire Peptide-MHCII Complexes from Dendritic Cells and Induce Antigen-specific CD4+ T Cell Tolerance. *J. Exp. Med.* 211 (6), 1153–1166. doi:10.1084/jem.20132000

Fan, Z., Huang, G., Zhao, J., Li, W., Lin, T., Su, Q., et al. (2021). Establishment and Characterization of a Highly Metastatic Human Osteosarcoma Cell Line from Osteosarcoma Lung Metastases. *J. Bone Oncol.* 29, 100378. doi:10.1016/j.jbo.2021.100378

Gazouli, I., Kyriazoglou, A., Kotsantis, I., Anastasiou, M., Pantazopoulos, A., Prevezanou, M., et al. (2021). Systematic Review of Recurrent Osteosarcoma Systemic Therapy. *Cancers* 13 (8), 1757. doi:10.3390/cancers13081757

Ghasemi, F., Tessier, T. M., Gameiro, S. F., Maciver, A. H., Cecchini, M. J., and Mymryk, J. S. (2020). High MHC-II Expression in Epstein-Barr Virus-Associated Gastric Cancers Suggests that Tumor Cells Serve an Important Role in Antigen Presentation. *Sci. Rep.* 10 (1), 14786. doi:10.1038/s41598-020-71775-4

Hashimoto, K., Nishimura, S., and Akagi, M. (2020). Characterization of PD-1/pd-L1 Immune Checkpoint Expression in Osteosarcoma. *Diagnostics* 10 (8), 528. doi:10.3390/diagnostics10080528

Hoffman, H. M., and Wanderer, A. A. (2010). Inflammasome and IL-1β-Mediated Disorders. *Curr. Allergy Asthma Rep.* 10 (4), 229–235. doi:10.1007/s11882-010-0109-z

Jiang, F., Miao, X.-L., Zhang, X.-T., Yan, F., Mao, Y., Wu, C.-Y., et al. (2021). A Hypoxia Gene-Based Signature to Predict the Survival and Affect the Tumor Immune Microenvironment of Osteosarcoma in Children. *J. Immunol. Res.* 2021, 1–13. doi:10.1155/2021/5523832

Johnson, A. M., Bullock, B. L., Neuwelt, A. J., Poczobutt, J. M., Kaspar, R. E., Li, H. Y., et al. (2020). Cancer Cell-Intrinsic Expression of MHC Class II Regulates the Immune Microenvironment and Response to Anti-PD-1 Therapy in Lung Adenocarcinoma. *J. Immunol.* 204 (8), 2295–2307. doi:10.4049/jimmunol.1900778

Ju, A., Tang, J., Chen, S., Fu, Y., and Luo, Y. (2021a). Pyroptosis-Related Gene Signatures Can Robustly Diagnose Skin Cutaneous Melanoma and Predict the Prognosis. *Front. Oncol.* 11, 709077. doi:10.3389/fonc.2021.709077

Ju, X., Yang, Z., Zhang, H., and Wang, Q. (2021b). Role of Pyroptosis in Cancer Cells and Clinical Applications. *Biochimie* 185, 78–86. doi:10.1016/j.biochi.2021.03.007

Judge, S. J., Darrow, M. A., Thorpe, S. W., Gingrich, A. A., O'Donnell, E. F., Bellini, A. R., et al. (2020). Analysis of Tumor-Infiltrating NK and T Cells Highlights IL-15 Stimulation and TIGIT Blockade as a Combination Immunotherapy Strategy for Soft Tissue Sarcomas. *J. Immunother. Cancer* 8 (2), e001355. doi:10.1136/jitc-2020-001355

Kang, R., Zeng, L., Zhu, S., Xie, Y., Liu, J., Wen, Q., et al. (2018). Lipid Peroxidation Drives Gasdermin D-Mediated Pyroptosis in Lethal Polymicrobial Sepsis. *Cell Host & Microbe* 24 (1), 97–108.e4. doi:10.1016/j.chom.2018.05.009

Karin, M., and Shalapour, S. (2021). Regulation of Antitumor Immunity by Inflammation-Induced Epigenetic Alterations. *Cell Mol Immunol.*, [Epub ahead of print]. doi:10.1038/s41423-021-00756-y

Kayagaki, N., Stowe, I. B., Lee, B. L., O'Rourke, K., Anderson, K., Warming, S., et al. (2015). Caspase-11 Cleaves Gasdermin D for Non-canonical Inflammasome Signalling. *Nature* 526 (7575), 666–671. doi:10.1038/nature15541

Li, K., Liu, J., Tian, M., Gao, G., Qi, X., Pan, Y., et al. (2015). CHMP4C Disruption Sensitizes the Human Lung Cancer Cells to Irradiation. *Int. J. Mol Sci.* 17 (1), 18. doi:10.3390/ijms17010018

Li, L., Jiang, M., Qi, L., Wu, Y., Song, D., Gan, J., et al. (2021a). Pyroptosis, a New Bridge to Tumor Immunity. *Cancer Sci.* 112, 3979–3994. doi:10.1111/cas.15059

Li, M., Wu, W., Deng, S., Shao, Z., and Jin, X. (2021b). TRAIP Modulates the IGFBP3/AKT Pathway to Enhance the Invasion and Proliferation of Osteosarcoma by Promoting KANK1 Degradation. *Cell Death Dis* 12 (8), 767. doi:10.1038/s41419-021-04057-0

Li, L., Li, Y., and Bai, Y. (2020). Role of GSDMB in Pyroptosis and Cancer. *Cancer Manag. Res.* 12, 3033–3043. doi:10.2147/cmar.S246948

Ligon, J. A., Choi, W., Cojocaru, G., Fu, W., Hsiue, E. H.-C., Oke, T. F., et al. (2021). Pathways of Immune Exclusion in Metastatic Osteosarcoma Are Associated with Inferior Patient Outcomes. *J. Immunother. Cancer* 9 (5), e001772. doi:10.1136/jitc-2020-001772

Lin, H., Chen, X., Zhang, C., Yang, T., Deng, Z., Song, Y., et al. (2021a). EF24 Induces Ferroptosis in Osteosarcoma Cells through HMOX1. *Biomed. Pharmacother.* 136, 111202. doi:10.1016/j.biopha.2020.111202

Lin, W., Chen, Y., Wu, B., Chen, Y., and Li, Z. (2021b). Identification of the Pyroptosis-related P-rognostic G-ene S-ignature and the A-ssociated R-egulation axis in L-ung A-denocarcinoma. *Cell Death Discov.* 7 (1), 161. doi:10.1038/s41420-021-00557-2

Lin, R., Wei, H., Wang, S., Huang, Z., Chen, H., Zhang, S., et al. (2020a). Gasdermin D Expression and Clinicopathologic Outcome in Primary Osteosarcoma Patients. *Int. J. Clin. Exp. Pathol.* 13 (12), 3149–3157.

Lin, S. L., Wang, M., Cao, Q. Q., and Li, Q. (2020b). Chromatin Modified Protein 4C (CHMP4C) Facilitates the Malignant Development of Cervical Cancer Cells. *FEBS Open Bio* 10 (7), 1295–1303. doi:10.1002/2211-5463.12880

Liu, X., Xia, S., Zhang, Z., Wu, H., and Lieberman, J. (2021). Channelling Inflammation: Gasdermins in Physiology and Disease. *Nat. Rev. Drug Discov.* 20 (5), 384–405. doi:10.1038/s41573-021-00154-z

Loveless, R., Bloomquist, R., and Teng, Y. (2021). Pyroptosis at the Forefront of Anticancer Immunity. *J. Exp. Clin. Cancer Res.* 40 (1), 264. doi:10.1186/s13046-021-02065-8

Lu, X., Guo, T., and Zhang, X. (2021). Pyroptosis in Cancer: Friend or Foe? *Cancers* 13 (14), 3620. doi:10.3390/cancers13143620

Lu, Y.-C., Parker, L. L., Lu, T., Zheng, Z., Toomey, M. A., White, D. E., et al. (2017). Treatment of Patients with Metastatic Cancer Using a Major Histocompatibility Complex Class II-Restricted T-Cell Receptor Targeting the Cancer Germline Antigen MAGE-A3. *J. Clin. Oncol.* 35 (29), 3322–3329. doi:10.1200/jco.2017.74.5463

Mantovani, A., Barajon, I., and Garlanda, C. (2018). IL-1 and IL-1 Regulatory Pathways in Cancer Progression and Therapy. *Immunol. Rev.* 281 (1), 57–61. doi:10.1111/imr.12614

Na, Z., Fan, L., and Wang, X. (2021). Gene Signatures and Prognostic Values of N6-Methyladenosine Related Genes in Ovarian Cancer. *Front. Genet.* 12, 542457. doi:10.3389/fgene.2021.542457

Ni, S., Li, J., Qiu, S., Xie, Y., Gong, K., and Duan, Y. (2020). KIF21B Expression in Osteosarcoma and its Regulatory Effect on Osteosarcoma Cell Proliferation and Apoptosis through the PI3K/AKT Pathway. *Front. Oncol.* 10, 606765. doi:10.3389/fonc.2020.606765

Pan, J., Zhang, X., Fang, X., and Xin, Z. (2021). Construction on of a Ferroptosis-Related lncRNA-Based Model to Improve the Prognostic Evaluation of Gastric Cancer Patients Based on Bioinformatics. *Front. Genet.* 12, 739470. doi:10.3389/fgene.2021.739470

Park, J. A., and Cheung, N.-K. V. (2020). GD2 or HER2 Targeting T Cell Engaging Bispecific Antibodies to Treat Osteosarcoma. *J. Hematol. Oncol.* 13 (1), 172. doi:10.1186/s13045-020-01012-y

Protti, M. P., and De Monte, L. (2020). Dual Role of Inflammasome Adaptor ASC in Cancer. *Front. Cel Dev. Biol.* 8, 40. doi:10.3389/fcell.2020.00040

Qi, W., Yan, Q., Lv, M., Song, D., Wang, X., and Tian, K. (2021). Prognostic Signature of Osteosarcoma Based on 14 Autophagy-Related Genes. *Pathol. Oncol. Res.* 27, 1609782. doi:10.3389/pore.2021.1609782

Rojas, G. A., Hubbard, A. K., Diessner, B. J., Ribeiro, K. B., and Spector, L. G. (2021). International Trends in Incidence of Osteosarcoma (1988-2012). *Int. J. Cancer* 149 (5), 1044–1053. doi:10.1002/ijc.33673

Santiago, L., Castro, M., Sanz-Pamplona, R., Garzón, M., Ramirez-Labrada, A., Tapia, E., et al. (2020). Extracellular Granzyme A Promotes Colorectal Cancer Development by Enhancing Gut Inflammation. *Cel Rep.* 32 (1), 107847. doi:10.1016/j.celrep.2020.107847

Shao, W., Yang, Z., Fu, Y., Zheng, L., Liu, F., Chai, L., et al. (2021). The Pyroptosis-Related Signature Predicts Prognosis and Indicates Immune Microenvironment Infiltration in Gastric Cancer. *Front. Cel Dev. Biol.* 9, 676485. doi:10.3389/fcell.2021.676485

Shi, H., Zhong, F., Yi, X., Shi, Z., Ou, F., Xu, Z., et al. (2020). Application of an Autophagy-Related Gene Prognostic Risk Model Based on TCGA Database in Cervical Cancer. *Front. Genet.* 11, 616998. doi:10.3389/fgene.2020.616998

Shin, C., Kim, S. S., and Jo, Y. H. (2021). Extending Traditional Antibody Therapies: Novel Discoveries in Immunotherapy and Clinical Applications. *Mol. Ther. - Oncolytics* 22, 166–179. doi:10.1016/j.omto.2021.08.005

Skotte, N. H., Sanders, S. S., Singaraja, R. R., Ehrnhoefer, D. E., Vaid, K., Qiu, X., et al. (2017). Palmitoylation of Caspase-6 by HIP14 Regulates its Activation. *Cel Death Differ* 24 (3), 433–444. doi:10.1038/cdd.2016.139

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., Ji, L., Han, M., Xie, J., Zhong, F., Zhang, X., et al. (2020). Pyroptosis Is Involved in the Inhibitory Effect of FL118 on Growth and Metastasis in Colorectal Cancer. *Life Sci.* 257, 118065. doi:10.1016/j.lfs.2020.118065

Titov, A., Zmievskaya, E., Ganeeva, I., Valiullina, A., Petukhov, A., Rakhmatullina, A., et al. (2021). Adoptive Immunotherapy beyond CAR T-Cells. *Cancers* 13 (4), 743. doi:10.3390/cancers13040743

van Daalen, K. R., Reijneveld, J. F., and Bovenschen, N. (2020). Modulation of Inflammation by Extracellular Granzyme A. *Front. Immunol.* 11, 931. doi:10.3389/fimmu.2020.00931

Vervliet, T., Parys, J. B., and Bultynck, G. (2016). Bcl-2 Proteins and Calcium Signaling: Complexity beneath the Surface. *Oncogene* 35 (39), 5079–5092. doi:10.1038/onc.2016.31

Wang, H., Huang, H., Wang, L., Liu, Y., Wang, M., Zhao, S., et al. (2021a). Cancer-associated Fibroblasts Secreted miR-103a-3p Suppresses Apoptosis and Promotes Cisplatin Resistance in Non-small Cell Lung Cancer. *Aging* 13 (10), 14456–14468. doi:10.18632/aging.103556

Wang, X., Wang, L., Xu, W., Wang, X., Ke, D., Lin, J., et al. (2021b). Classification of Osteosarcoma Based on Immunogenomic Profiling. *Front. Cel Dev. Biol.* 9, 696878. doi:10.3389/fcell.2021.696878

Wang, S.-D., Li, H.-Y., Li, B.-H., Xie, T., Zhu, T., Sun, L.-L., et al. (2016a). The Role of CTLA-4 and PD-1 in Anti-tumor Immune Response and Their Potential Efficacy against Osteosarcoma. *Int. Immunopharmacol.* 38, 81–89. doi:10.1016/j.intimp.2016.05.016

Wang, Y., Kong, H., Zeng, X., Liu, W., Wang, Z., Yan, X., et al. (2016b). Activation of NLRP3 Inflammasome Enhances the Proliferation and Migration of A549 Lung Cancer Cells. *Oncol. Rep.* 35 (4), 2053–2064. doi:10.3892/or.2016.4569

Wang, W. J., Chen, D., Jiang, M. Z., Xu, B., Li, X. W., Chu, Y., et al. (2018). Downregulation of Gasdermin D Promotes Gastric Cancer Proliferation by Regulating Cell Cycle-Related Proteins. *J. Dig. Dis.* 19 (2), 74–83. doi:10.1111/1751-2980.12576

Wang, Z., Ni, F., Yu, F., Cui, Z., Zhu, X., and Chen, J. (2019). Prognostic Significance of mRNA Expression of CASPs in Gastric Cancer. *Oncol. Lett.* 18 (5), 4535–4554. doi:10.3892/ol.2019.10816

Wen, R., Umeano, A. C., Essegian, D. J., Sabitaliyevich, U. Y., Wang, K., and Farooqi, A. A. (2018). Role of microRNA-410 in Molecular Oncology: A Double Edged Sword. *J. Cel Biochem* 119 (11), 8737–8742. doi:10.1002/jcb.27251

Wu, D., Wei, C., Li, Y., Yang, X., and Zhou, S. (2021). Pyroptosis, a New Breakthrough in Cancer Treatment. *Front. Oncol.* 11, 698811. doi:10.3389/fonc.2021.698811

Wu, Y., Zhuang, J., Zhao, D., and Xu, C. (2019). Interaction between Caspase-3 and Caspase-5 in the Stretch-induced Programmed Cell Death in the Human Periodontal Ligament Cells. *J. Cel Physiol* 234 (8), 13571–13581. doi:10.1002/jcp.28035

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

Xing, X., Gu, F., Hua, L., Cui, X., Li, D., Wu, Z., et al. (2021). TIMELESS Promotes Tumor Progression by Enhancing Macrophages Recruitment in Ovarian Cancer. *Front. Oncol.* 11, 732058. doi:10.3389/fonc.2021.732058

Yao, Q., and Chen, T. (2020). LINC01128 Regulates the Development of Osteosarcoma by Sponging miR-299-3p to Mediate MMP2 Expression and Activating Wnt/β-catenin Signalling Pathway. *J. Cel. Mol. Med.* 24 (24), 14293–14305. doi:10.1111/jcmm.16046

Ye, Y., Dai, Q., and Qi, H. (2021). A Novel Defined Pyroptosis-Related Gene Signature for Predicting the Prognosis of Ovarian Cancer. *Cel Death Discov.* 7 (1), 71. doi:10.1038/s41420-021-00451-x

Zhang, B., Mao, S., Liu, X., Li, S., Zhou, H., Gu, Y., et al. (2021). MiR-125b Inhibits Cardiomyocyte Apoptosis by Targeting BAK1 in Heart Failure. *Mol. Med.* 27 (1), 72. doi:10.1186/s10020-021-00328-w

Zheng, M., and Kanneganti, T. D. (2020). The Regulation of the ZBP1-NLRP3 Inflammasome and its Implications in Pyroptosis, Apoptosis, and Necroptosis (PANoptosis). *Immunol. Rev.* 297 (1), 26–38. doi:10.1111/imr.12909

Zheng, M., Karki, R., Vogel, P., and Kanneganti, T.-D. (2020). Caspase-6 Is a Key Regulator of Innate Immunity, Inflammasome Activation, and Host Defense. *Cell* 181 (3), 674–687.e13. doi:10.1016/j.cell.2020.03.040

Zhou, C.-B., and Fang, J.-Y. (2019). The Role of Pyroptosis in Gastrointestinal Cancer and Immune Responses to Intestinal Microbial Infection. *Biochim. Biophys. Acta (Bba) - Rev. Cancer* 1872 (1), 1–10. doi:10.1016/j.bbcan.2019.05.001

Zhou, Y., Dai, W., Wang, H., Pan, H., and Wang, Q. (2018). Long Non-coding RNA CASP5 Promotes the Malignant Phenotypes of Human Glioblastoma Multiforme. *Biochem. Biophysical Res. Commun.* 500 (4), 966–972. doi:10.1016/j.bbrc.2018.04.217

Zhou, Z., He, H., Wang, K., Shi, X., Wang, Y., Su, Y., et al. (2020). Granzyme A from Cytotoxic Lymphocytes Cleaves GSDMB to Trigger Pyroptosis in Target Cells. *Science* 368 (6494), eaaz7548. doi:10.1126/science.aaz7548

Conflict of Interest: Author QZ was employed by the company Shenzhen Tyercan Bio-Pharm Co., Ltd.

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

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 08 December 2021
doi: 10.3389/fgene.2021.743133

Glycolysis Changes the Microenvironment and Therapeutic Response Under the Driver of Gene Mutation in Esophageal Adenocarcinoma

Lei Zhu[1,2†], *Fugui Yang*[2†], *Xinrui Li*[3†], *Qinchuan Li*[2*] *and Chunlong Zhong*[1*]

[1]Department of Neurosurgery, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai, China, [2]Department of Thoracic Surgery, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai, China, [3]Department of Neurology, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai, China

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez, Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Divya Mundackal Sivaraman, Sree Chitra Tirunal Institute for Medical Sciences and Technology (SCTIMST), India
Espiridión Ramos-Martínez, Universidad Nacional Autónoma de México, Mexico

***Correspondence:**
Qinchuan Li
qinchuanli@tongji.edu.cn
Chunlong Zhong
drchunlongzhong@126.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: 17 July 2021
Accepted: 25 October 2021
Published: 08 December 2021

Citation:
Zhu L, Yang F, Li X, Li Q and Zhong C (2021) Glycolysis Changes the Microenvironment and Therapeutic Response Under the Driver of Gene Mutation in Esophageal Adenocarcinoma. Front. Genet. 12:743133. doi: 10.3389/fgene.2021.743133

Background: Esophageal cancer is one of the most leading and lethal malignancies. Glycolysis and the tumor microenvironment (TME) are responsible for cancer progressions. We aimed to study the relationships between glycolysis, TME, and therapeutic response in esophageal adenocarcinoma (EAC).

Materials and Methods: We used the ESTIMATE algorithm to divide EAC patients into ESTIMATE^{high} and ESTIMATE^{low} groups based on the gene expression data downloaded from TCGA. Weighted gene co-expression network analysis (WGCNA) and Gene Set Enrichment Analysis (GSEA) were performed to identify different glycolytic genes in the TME between the two groups. The prognostic gene signature for overall survival (OS) was established through Cox regression analysis. Impacts of glycolytic genes on immune cells were assessed and validated. Next, we conducted the glycolytic gene mutation analysis and drug therapeutic response analysis between the two groups. Finally, the GEO database was employed to validate the impact of glycolysis on TME in patients with EAC.

Results: A total of 78 EAC patients with gene expression profiles and clinical information were included for analysis. Functional enrichment results showed that the genes between ESTIMATE^{high} and ESTIMATE^{low} groups (N = 39, respectively) were strongly related with glycolytic and ATP/ADP metabolic pathways. Patients in the low-risk group had probabilities to survive longer than those in the high-risk group ($p < 0.001$). Glycolytic genes had significant impacts on the components of immune cells in TME, especially on the T-cells and dendritic cells. In the high-risk group, the most common mutant genes were

Abbreviations: EAC, esophageal adenocarcinoma; TME, tumor microenvironment; TCGA, the cancer genome atlas; IICs, infiltrating immune cells; DEGs, differently expressed genes; FDR, false discovery rate; log2FC, log2 fold change; GSEA, gene set enrichment analysis; NOM, nominal; NES: normalized enrichment score; IGs, intersection genes; GO: gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; BP, biological process; CC, cellular component; MF, molecular function; OS, overall survival; SNV, single nucleotide variant; TMB, tumor mutation burden; K-M, Kaplan–Meier; GDSC, genomics of drug sensitivity in cancer; TIDE, tumor immune dysfunction and exclusion; GEO, gene expression omnibus; VEGF, vascular endothelial growth factor; NPCs, nuclear pore complexes; HR, hazard ratio; SNP, single-nucleotide polymorphism; NPCs, nuclear pore complexes.

TP53 and TTN, and the most frequent mutation type was missense mutation. Glycolysis significantly influenced drug sensitivity, and high tumor mutation burden (TMB) was associated with better immunotherapeutic response. GEO results confirmed that glycolysis had significant impacts on immune cell contents in TME.

Conclusion: We performed a comprehensive study of glycolysis and TME and demonstrated that glycolysis could influence the microenvironment and drug therapeutic response in EAC. Evaluation of the glycolysis pattern could help identify the individualized therapeutic regime.

Keywords: glycolysis, tumor microenvironment, immunotherapy, drug response, esophageal adenocarcinoma

INTRODUCTION

Esophageal cancer is the eighth most common malignancy and the sixth cause of cancer death globally, which accounts for more than 570,000 new cases and 500,000 deaths annually (Bray et al., 2018; Wang et al., 2018). Esophageal adenocarcinoma (EAC) is the predominant pathological type in western countries, with an increasing proportion from 35 to 61% over the past 30 years (Alsop and Sharma, 2016). The global incidence rate of EAC is approximately 0.7/100,000 person-years, and the 5-year survival rate is merely less than 20%, although multidisciplinary treatments have been applied, including esophagectomy, radiation, and chemotherapy (Arnold et al., 2015; Markar et al., 2017; Smyth et al., 2017; Zhao et al., 2019). Considering the chemotherapeutic resistance, several targeted agents have been applied in patients with EAC, such as imatinib (Mayr et al., 2012). Unfortunately, the efficacy is still not satisfactory. Recently, immunotherapy-targeting PD-1 has revolutionized the therapy in cancer patients. However, not all patients with EAC respond to immunotherapy (Däster et al., 2020). Therefore, there is an urgent necessity to better understand the molecular characteristics and genetic features that could help predict accurate survival and identify suitable patients who will benefit from immunotherapy.

The tumorigenesis and development of EAC is a highly complex biology, involving the tumor cell-intrinsic and cell-extrinsic factors (Quante et al., 2018; Talukdar et al., 2018). Genetic alterations are the primary mechanisms that drive the initiation and progression of EAC, not only conferring tumor cells infinite proliferative abilities but also reprogramming metabolic pathways to adapt to the hostile environment, such as aerobic glycolysis (Hochwald and Zhang, 2017; Talukdar et al., 2018). The seminal discovery of tumor glycolysis has been considered a hallmark of cancer, proposed by Otto Warburg in 1923 (Warburg and Minami, 1923). The glycolytic phenotype renders cancer cells selective advantages by unlimited growth and attenuated apoptosis (Xu et al., 2017). In addition, it is gradually evident that elevated glycolysis is closely related to the immune escape by changing the microenvironment and inhibiting the functions of immune cells (Jiang et al., 2019a). Mounting evidence from cell-based assays has linked glycolysis to TME, and preclinical investigations have demonstrated the effectiveness targeting glycolysis in some cancers (Lim et al., 2017; Kornberg et al., 2018; Jiang et al., 2019a; Jiang et al., 2019b; Kang et al., 2020).

Genetic mutation results in the rewire of the glucose metabolism decreased cancer cell apoptosis and immortal growth. Consequently, these events bring about the component reconstruction in the tumor microenvironment (TME), thus changing the purity of the tumor. Reciprocally, the intimate interactions between glycolytic cells and the extracellular matrix further exacerbate the remodeling of TME, including the stromal and immune cells. It is well accepted that the tumor is highly dependent on TME, which is preponderant on prognosis and impacts the therapeutic efficacy profoundly, such as the immune checkpoint therapy (Wu and Dai, 2017; Alsina et al., 2018; Taube et al., 2018; Hinshaw and Shevde, 2019; Li et al., 2020). However, there are few studies exploring the associations between glycolysis and TME in EAC, and far less is known about how genetic mutations orchestrate the glycolysis under these aberrant TME conditions. Herein, we investigated the effects of glycolysis on immune cells and revealed the genetic mutation diversity. Our study unraveled that glycolysis could influence TME under the driver of genetic mutation and could serve as prognostic biomarkers. Moreover, we constructed the risk score system to predict drug sensitivity and immunotherapeutic response. The results hold great promise in targeting glycolysis and utilizing TME to improve the treatment in patients with EAC.

MATERIALS AND METHODS

Data Acquisition

Gene expression data and clinical information were downloaded from the Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). The mRNA expression profiles were log2 normalized for further analysis. Clinical information included gender, age, stage, survival status, and follow-up time.

Tumor Microenvironment and Glycolysis

TME is composed of resident stromal cells and infiltrating immune cells (IICs), reflecting tumor purity. With the increase of stromal cells and IICs, the tumor purity becomes lower. The stromal score, immune score, and ESTIMATE score were calculated by applying the ESTIMATE algorithm (Chakraborty

and Hossain, 2018). The ESTIMATE score is the comprehensive parameter of the stromal and immune scores. Then, patients with EAC were classified into ESTIMATE ^{high} and ESTIMATE ^{low} groups according to the median of the ESTIMATE score. The differently expressed genes (DEGs) were screened by the weighted gene co-expression network analysis (WGCNA) with the false discovery rate (FDR) ≤0.05 and log2 fold change (log2FC)| >2.

To explore whether glycolysis affects tumor purity, we performed gene set enrichment analysis (GSEA) between the ESTIMATE ^{high} and ESTIMATE ^{low} groups. Five glycolysis-related gene sets, namely, Hallmark, BioCarta, KEGG, GO, and Reactome, were downloaded from the Molecular Signatures Database (http://www.gsea-msigdb.org/gsea/msigdb) and analyzed using the GSEA software (version 4.1.0). The permutation number was set as 1,000 for every phenotype. The gene sets were considered statistically significant when the nominal (NOM) p-value ≤0.05, FDR ≤0.05, and normalized enrichment score |(NES)| >1. Finally, the intersection genes (IGs) from the WGCNA and GSEA gene sets were identified for further analysis.

Gene Interaction Analysis and Enrichment Analysis

Gene interaction analysis was performed through the "corrplot" package in R software (version 4.1.0). Hub genes were screened with "cytoHubba" in Cytoscape software.

The IGs based on TME and glycolysis were analyzed for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) by the "clusterProfiler" R package. GO analysis has three functional parts, including the biology process (BP), cellular component (CC), and molecular function (MF).

Next, we performed functional similarity analysis, which was measured through the "GOSemSim" R package (Wang et al., 2007). Functional similarity could be used for the purpose of assessing the intimacy and relationship between each gene and its partners by evaluating the function and location.

Establishment of Prognostic Signatures

First, univariate Cox regression analysis was used to identify IGs which were related to patients' overall survival (OS). Then, statistically significant IGs ($p < 0.05$) were enrolled into the multivariate Cox regression. Finally, patients were divided into high- and low-risk groups according to the median of the risk score, in which the risk score was calculated as follows: risk score = $\sum_{n=1}^{j} Coef\,j * Xj$, with Coef j indicating the coefficient and Xj representing the relative expression levels of each IG standardized by the z-score.

TME and Gene Mutation

Next, we selected the prognostic glycolysis-related genes and hub genes to investigate their relations with IICs between the two groups through single-sample gene set enrichment analysis (ssGSEA) using the "GSVA" R package. The effects of OGG on immune cells were assessed using the linear regression.

To analyze why glycolysis affects TME, we calculated the glycolytic gene mutation frequency, variant classification, variant type, and single nucleotide variants (SNVs) between the ESTIMATE ^{high} and ESTIMATE ^{low} groups. Additionally, to fully understand the role of gene mutation in TME, we performed tumor mutation burden (TMB) analyses and explored their relationships with IICs through the simple nucleotide variation data from TCGA and cBioPortal online databases (http://www.cbioportal.org/).

Drug Sensitivity Analysis and Immunotherapy Response

The drug sensitivity of each patient with EAC was predicted by the Genomics of Drug Sensitivity in Cancer database (GDSC; https://www.cancerrxgene.org/). The half-maximal inhibitory concentration (IC50) was calculated through the "pRRophetic" R package, and the IC50 differences between the high- and low-risk groups were compared (Geeleher et al., 2014).

The response to immunotherapy was estimated using the Tumor Immune Dysfunction and Exclusion website (TIDE; http://tide.dfci.harvard.edu/login/). The TIDE and PDL-1 scores were compared between the high- and low-risk groups.

External Cohort Validation

The impacts of glycolysis on TME in EAC were validated through the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds/). The study was considered eligible for external cohort validation according to the following criteria: 1) studies with *Homo sapiens* samples and 2) studies with a sample number more than 50, and 3) studies with detailed experiment information and complete expression profiles. The primary goal of validation was to confirm whether the ESTIMATE algorithm method is suitable for patients with EAC, and the secondary goals were to determine whether glycolysis could influence the components of the microenvironment and affect drug sensitivity. The overall design of this study is shown in **Figure 1**.

Statistical Analysis

All statistical analyses were performed by the R software (version 4.1.0). The DEG analysis between the ESTIMATE ^{high} and ESTIMATE ^{low} groups was carried out by applying the unpaired t-test. Cox regression analysis was used to determine the prognostic factors. Kaplan–Meier (K-M) curves and log-rank tests were utilized to assess the prognostic outcome. The Mann–Whitney U test was used to compare the immune score, immune cell infiltrations, and immune signatures. Spearman's correlation analysis was used to evaluate the interactions. $p < 0.05$ was considered significant.

RESULTS

Identification of Tumor Purity and Glycolysis

A total of 87 EAC samples and gene expression data were available from the TCGA database, including 9 normal and 78

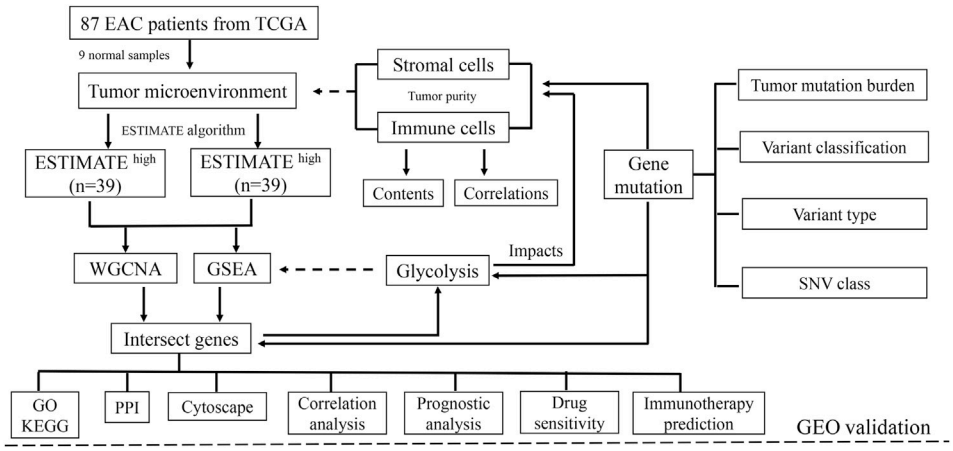


FIGURE 1 | Schematic of the study design. A total of 78 EAC patients were recruited for further study. TME is mainly composed of tumor cells, stromal cells, and immune cells. With higher stromal and immune cells, the tumor purity is low.

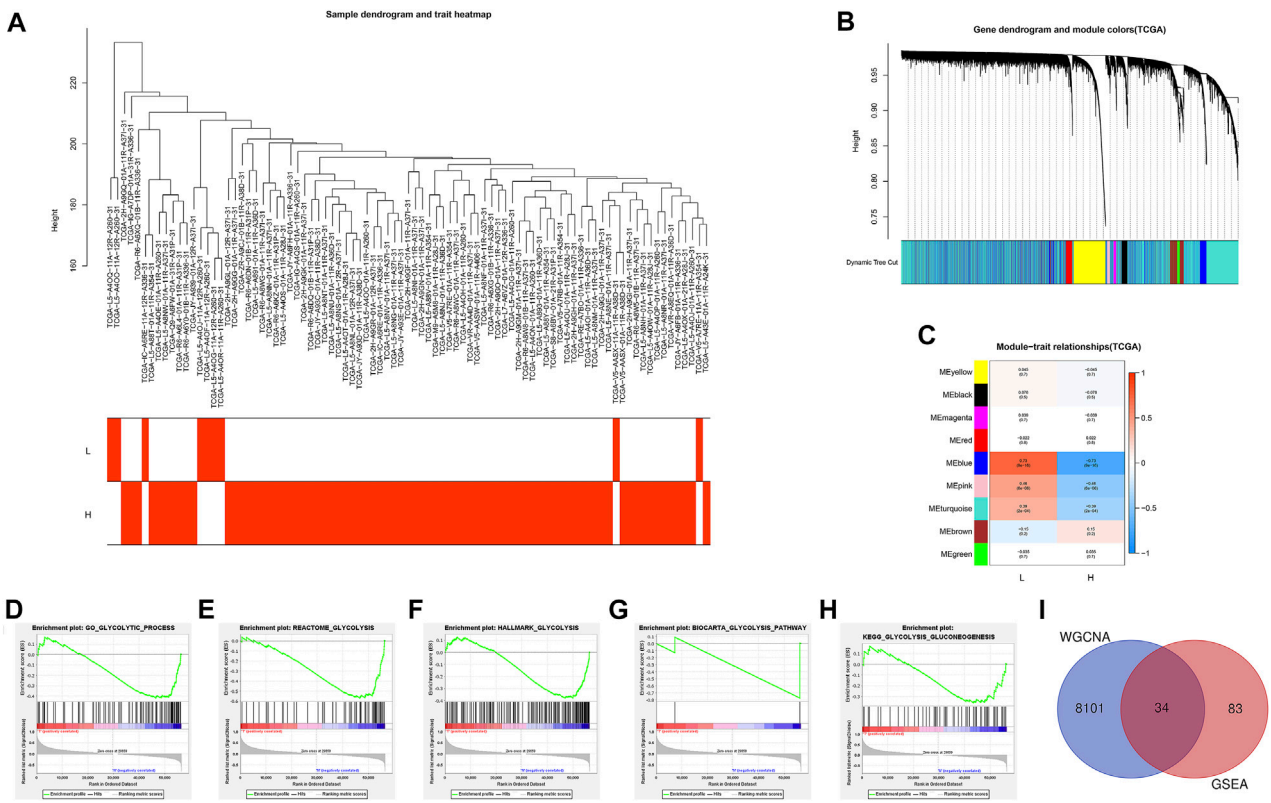


FIGURE 2 | Identification of the intersection genes. **(A)** 78 EAC sample clustering. All samples were clustered, and difference analysis was performed. **(B)** Gene dendrogram and dynamic tree cuts. Each color represents a module, and genes with similar expression patterns will be classified into the same module. **(C)** Module-trait relationships. Every row represents a module eigengene (ME). The red module corresponds to the significant positive correlation, and blue corresponds to the significant negative correlation. **(D)** GO glycolytic process gene set. **(E)** Reactome glycolysis gene set. **(F)** Hallmark glycolysis gene set. **(G)** Biocarta glycolysis pathway. **(H)** KEGG glycolysis gluconeogenesis. **(I)** Venn diagram shows the 34 IGs (intersection genes from WGCNA and GSEA). L: ESTIMATE ^{low}; H: ESTIMATE ^{high}.

TABLE 1 | Significant IG expression levels in the ESTIMATE ^{low} and ESTIMATE ^{high} tissues.

Gene	ESTIMATE ^{low}	ESTIMATE ^{high}	logFC	*p*	FDR
GAPDH	693.986	509.158	−0.447	0.002	0.011
NUP133	9.617	8.402	−0.195	0.035	0.118
GCK	0.077	0.232	1.595	0.000	0.001
HTR2A	0.037	0.323	3.116	0.000	0.000
NUP205	15.387	12.149	−0.341	0.037	0.122
NUP43	11.641	9.217	−0.337	0.000	0.004
SEH1L	7.027	5.367	−0.389	0.023	0.086
INSR	11.440	15.775	0.464	0.003	0.017
BPGM	5.508	7.941	0.528	0.000	0.000
PRKACB	4.543	6.572	0.533	0.010	0.045
CBFA2T3	0.390	1.505	1.947	0.000	0.000
NUP214	9.564	8.187	−0.224	0.015	0.062
NDC1	16.439	12.273	−0.422	0.001	0.006
HK3	0.409	1.216	1.572	0.000	0.000
NUP62	14.909	12.765	−0.224	0.027	0.096
EIF6	105.279	83.551	−0.333	0.022	0.084
PRXL2C	4.076	5.085	0.319	0.020	0.080
DDIT4L	0.098	0.215	1.134	0.000	0.001
ZBTB20	1.195	1.939	0.699	0.001	0.006
ENO1	227.228	180.617	−0.331	0.006	0.031
MLXIPL	7.779	4.578	−0.765	0.001	0.005
PRKAG1	9.508	8.284	−0.199	0.022	0.084
P2RX7	0.417	1.300	1.641	0.000	0.000
RAE1	10.815	8.043	−0.427	0.002	0.012
GPI	48.943	35.859	−0.449	0.005	0.026
NUP188	19.729	17.185	−0.199	0.037	0.122
HIF1A	49.095	66.493	0.438	0.002	0.014
HKDC1	18.783	12.578	−0.579	0.039	0.127
NUP37	6.737	5.076	−0.408	0.002	0.013
HK1	16.837	20.341	0.273	0.045	0.140
IGF1	0.090	0.414	2.196	0.000	0.000
TPI1	142.420	116.178	−0.294	0.025	0.091
ENO3	1.130	0.929	−0.281	0.022	0.084
NUP88	10.051	8.744	−0.201	0.045	0.140

LogFC: log fold change; FDR: false discovery rate.

EAC cases. Based on the ESTIMATE algorithm, the stromal score ranged from −2,315.387 to 1903.167 and the immune score ranged from −1,224.491 to 3,362.338. The range of the comprehensive ESTIMATE score was from −3,375.446 to 5,265.505 (**Supplementary Table S1**). According to the median of the ESTIMATE score, 78 patients with EAC were categorized into the ESTIMATE ^{high} and ESTIMATE ^{low} groups (39 cases, respectively). There are 8,135 DEGs between two groups according to WGCNA results (**Figures 2A–C**) (**Supplementary Table S2**).

Then, GSEA was conducted to assess the glycolytic differences between the ESTIMATE ^{high} and ESTIMATE ^{low} groups. The results showed that the GO glycolytic process (NES = −1.54, NOM p = 0.050, FDR = 0.050) and Reactome glycolysis (NES = −1.83, NOM p = 0.008, FDR = 0.008) were significantly enriched in ESTIMATE ^{high} group patients (**Figures 2D,E**). There were no significant enrichments in the Hallmark (NES = −1.44, NOM p = 0.067, FDR = 0.067), BioCarta (NES = −1.23, NOM p = 0.227, FDR = 0.227), and KEGG (NES = −1.20, NOM p = 0.228, FDR = 0.228) pathways (**Figures 2F–H**). There were 117 DEGs between GO and Reactome glycolysis gene sets. After screening, a total of 34 IGs were selected from WGCNA and GSEA for further analysis (**Table 1**) (**Figure 2I**). The details about IGs in each EAC sample are shown in **Supplementary Table S3**.

Gene Interaction Networks and Functional Enrichment Analysis

To explore the correlation between the IGs, we calculated their coefficients. Gene interaction analysis showed that GAPDH and TPI1 had the strongest positive correlation (coef = 0.81), whereas PRKACB and RAE1 had the strongest negative correlation (coef = −0.48) (**Figure 3A**). To explore the IG functions, we performed GO and KEGG enrichment analyses using R packages. GO results showed that IGs were significantly enriched in the glycolytic and ATP/ADP metabolic pathways. In addition, nuclear-, glucose-, and carbohydrate-related activities were closely associated with CC and MF terms (**Figure 3B**). KEGG results demonstrated that carbon, gluconeogenesis, and HIF−1 signaling pathways were enriched (**Figure 3C**).

Based on the GO analysis and semantic similarities, we ranked the genes by average functional similarities between IGs and their partners, with the cutoff value at 0.75. The box plots and raincloud plots are demonstrated in **Figures 3D,E**. From the pictures, we can clearly see that NUP43, NUP37, and DDIT4 had strong similarities and weak correlation with BPGM.

Prognostic IG Signatures

Univariate Cox regression analysis revealed that NUP88, RAE1, SEH1L, NUP37, and NUP43 were significantly associated with patients' OS (all $p < 0.05$) (**Figure 4A**). After multivariate Cox regression analysis, three IGs (NUP88, SEH1L, and NUP37) were used to develop the risk score based on the following formula: risk score = 0.637 * expression of NUP88 + 0.494 * expression of SEH1L + 0.657 * expression of NUP37. Also, the three genes were all risk genes with hazard ratio (HR) > 1. A total of 78 patients with EAC were classified into low- and high-risk groups according to the median risk score (n = 39). The K-M survival plot showed that patients in the low-risk group had significant probabilities to survive longer than those in the high-risk group (median time = 1.75 vs 0.745 years, $p < 0.001$) (**Figure 4B**).

In order to evaluate the prognostic values of clinical information in OS, we integrated the patients' clinical features with IGs. Univariate Cox regression analysis showed that the tumor stage (HR = 3.308, $p < 0.001$) and risk score (HR = 1.954, p = 0.013) were significantly associated with OS (**Figure 4C**). Multivariate Cox regression analysis results demonstrated that tumor stage (HR = 7.971, $p < 0.001$), metastasis (HR = 0.167, p = 0.033), and risk score (HR = 2.822, p = 0.002) were independent risk factors for OS (**Figure 4D**). In addition, the distributions of each patient and their survival statuses are shown in **Figures 4E–G**. We can clearly see that patients in the low-risk group had a better prognosis than those in the high-risk group.

Effect of Glycolytic Genes on IICs

We selected three prognostic genes and five hub genes (NUP88, SEH1L, NUP37, GCK, NUP62, NUP155, NUP205, and NUP214) to assess whether glycolysis affects the IICs in TME. The results demonstrated that NUP62, NUP155, NUP205, and SEH1L had

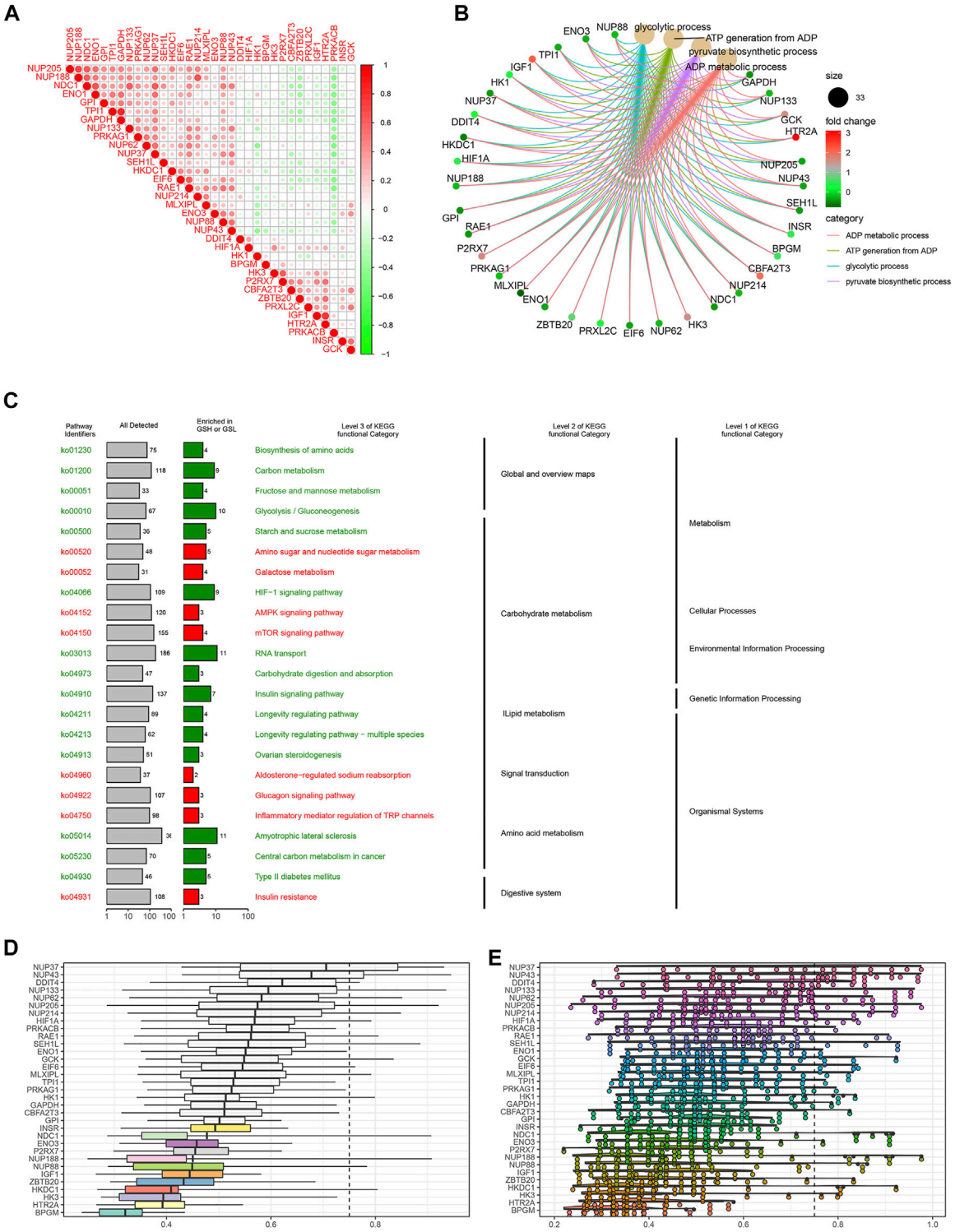


FIGURE 3 | Gene interaction analysis and functional enrichment. **(A)** Gene interaction network. Red represents positive correlation, while negative correlation is represented in green. **(B)** GO enrichment analysis, including the biological process (BP), cellular component (CC), and molecular function (MF). Every term shows top 10 pathways. **(C)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in the IGs. IG network analysis. **(D)** Summary of OGG similarities. The boxes indicated the middle 50% of the similarities, and the upper and lower boundaries show the 75th and 25th percentiles, respectively. **(E)** Raincloud plots of OGG. Data are expressed as the mean and standard error. Each dot represents the single gene. The dashed line represents the cutoff value (0.75).

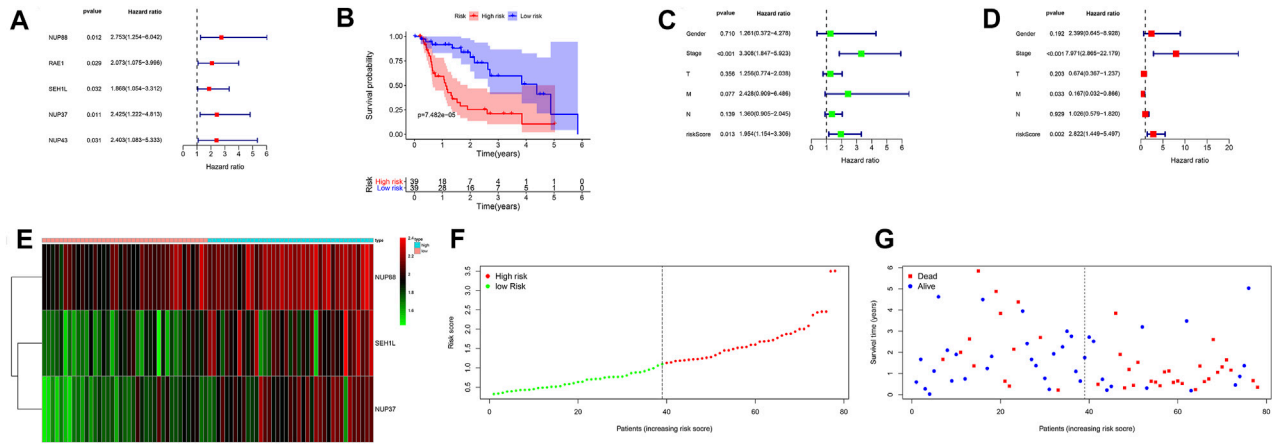


FIGURE 4 | Prognostic signatures of OS in EAC patients. **(A)** Forest plot of univariate Cox regression analysis based on IGs. **(B)** K-M survival plot of high- and low-risk score patients. **(C)** Forest plot of univariate Cox regression analysis based on the clinical information and risk score. **(D)** Forest plot of multivariate Cox regression analysis based on the clinical information and risk score. **(E)** Heatmap of NUP88, NUP37, and SEH1L in EAC patients. Green represents low gene expression, and red represents high expression. **(F)** Risk score curve of high- and low-risk score patients. The dotted line represents every individual, and patients are categorized into low-risk and high-risk groups at the inflection point. **(G)** Survival status and time distributed by the risk score. The red dot represents the dead, and blue represents the living. With an increase in time, more and more patients died. T: tumor; M: metastasis; N: lymph node.

significant impacts on the IIC expression level, especially on the T-cells and mast cells (all $p < 0.05$). The details are shown in **Figure 5A**.

To fully explore the relationships between these genes and IICs, we performed Spearman correlation analysis by using the "Limma" package. The results showed that NUP62 was strongly associated with T-cells, dendritic cells, and antigen-presenting cells (all $p < 0.05$) (**Figure 5B**). NUP155 had a close relationship with T-cells, mast cells, and MHC class I activity (all $p < 0.05$) (**Figure 5C**). NUP205 showed close relationships with T-cells, mast cells, and type I IFN response (all $p < 0.05$) (**Figure 5D**). SEH1L exhibited significant associations with dendritic cells, antigen-presenting cells, and chemokine receptors (CCR) (all $p < 0.05$) (**Figure 5E**). Collectively, these findings suggested that the IGs had profound effects on immune cells and immunological functions.

Genetic Mutation and the Tumor Microenvironment

Gene mutations were analyzed for further investigation into the mechanisms that TME is affected by gene alteration. The genetic mutations are significantly different between the ESTIMATE ^{high} and ESTIMATE ^{low} groups. In the ESTIMATE ^{high} group, the five most common mutant genes were TP53, TTN, HMCN1, DNAH5, and SYNE1, and the most common mutational type was missense mutation (**Figures 6A,C**). In the ESTIMATE ^{low} group, the most common mutant genes were TP53, TTN, MUC16, SYNE1, and PCLO. The most common mutational type was also missense mutation (**Figures 6B,F**). The frequencies of missense mutation were significantly lower than that in the ESTIMATE ^{high} group, indicating that the glycolytic level and tumor purity were different from those of the ESTIMATE ^{high} group (**Figures 6A,B**). Single-nucleotide polymorphism (SNP) had the highest frequency in the variant type (**Figures 6D,G**). G > A was the most frequent type in the SNV class (**Figures 6E,H**). The results imply that these mutant genes drive a higher glycolytic level, consequently changing the tumor purity in the microenvironment.

To broaden the understanding of the glycolytic gene mutations, the cBioPortal database was applied to validate these findings. We selected the most common glycolytic genes (TP53, TTN, HMCN1, DNAH5, SYNE1, MUC16, and PCLO) for further verification. Consistent with the above findings, the results from the cBioPortal database showed that TP53, TTN, and MUC16 possessed the highest mutation frequencies (87, 42, and 26%, respectively), and missense mutation was the commonest type (**Figure 7**).

Tumor Mutation Burden Analysis and Immunotherapy Response

TMB refers to the total number of mutations per mega base in tumor tissues. By analyzing the SNP data downloaded from TCGA, we calculated the TMB frequency in each patient with EAC. The range of TMB is from 0.053 to 41.053 in EAC. Moreover, we further analyzed the effect of TMB on survival in patients with EAC. A total of 78 patients were classified into high- and low-TMB groups according to the median TMB. As shown in the K-M curves, patients in the high-TMB group had significantly higher mortality than those in the low-TMB group ($p = 0.05$) (**Figure 8A**). To exhibit the relationships between TMB, glycolysis, and TME (ESTIMATE score), we applied the Sankey diagram to visualize their correlations with the "ggalluvial" package in R. The result is shown in **Figure 8B**.

Low TMB usually implies poor response for immunotherapy (Chan et al., 2019). To explore whether TMB will influence the immunotherapy response, we compared the PDL-1 and TIDE

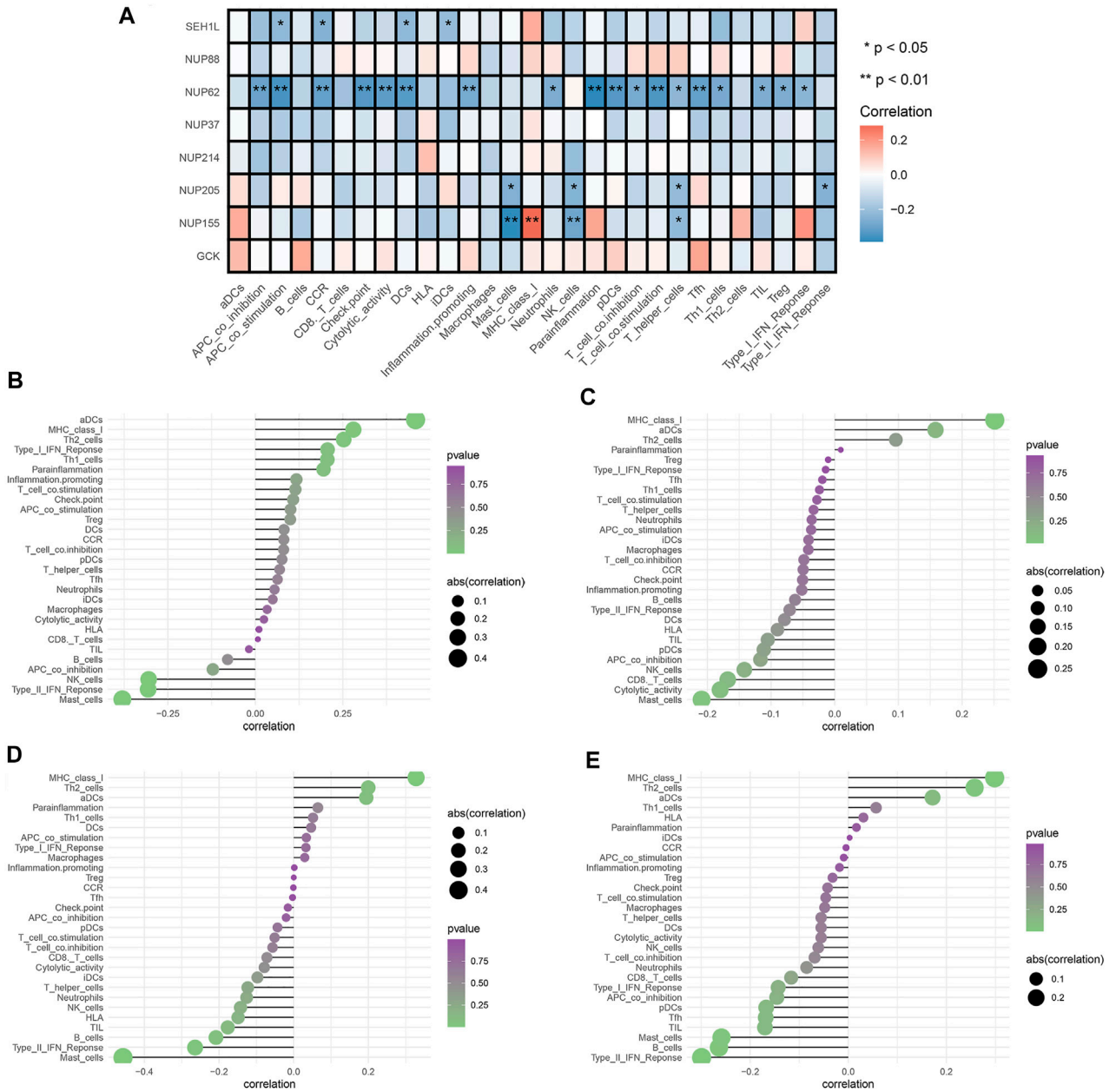


FIGURE 5 | Analysis of glycolytic gene effects on immune signatures. **(A)** Heatmap demonstrating the correlation between 8 genes and the ssGSEA scores of 29 immune signatures. **(B)** NUP62 and immune signatures. **(C)** NUP155 and immune signatures. **(D)** NUP205 and immune signatures. **(E)** SEH1L and immune signatures. Spearman's correlation analysis was used to evaluate the relations with $p < 0.05$.

scores between high- and low-TMB groups. As a result, patients in the high-TMB group significantly responded to anti–PDL-1 therapy ($p = 0.040$) (**Figure 8C**). In **Figure 8D**, patients with high TMB had a higher TIDE score than those with low TMB ($p = 0.000$). In addition, the correlations between the immune checkpoints are also explored in **Figure 8E**.

Drug Sensitivity Analysis

We compared the IC50 differences of chemotherapeutic and targeted drugs between high- and low-risk score groups, including bexarotene (**Figure 9A**), camptothecin (**Figure 9B**), gemcitabine (**Figure 9C**), imatinib (**Figure 9D**), methotrexate (**Figure 9E**), and vorinostat (**Figure 9F**). The results demonstrated that there were higher IC50 levels of bexarotene and imatinib in the high-risk score group, which indicated that patients with a low-risk score were more sensitive to the two drugs. Oppositely, the IC50 levels of camptothecin, gemcitabine, methotrexate, and vorinostat were higher in the low-risk score group, implying that patients in the high-risk score group were more sensitive to the four drugs.

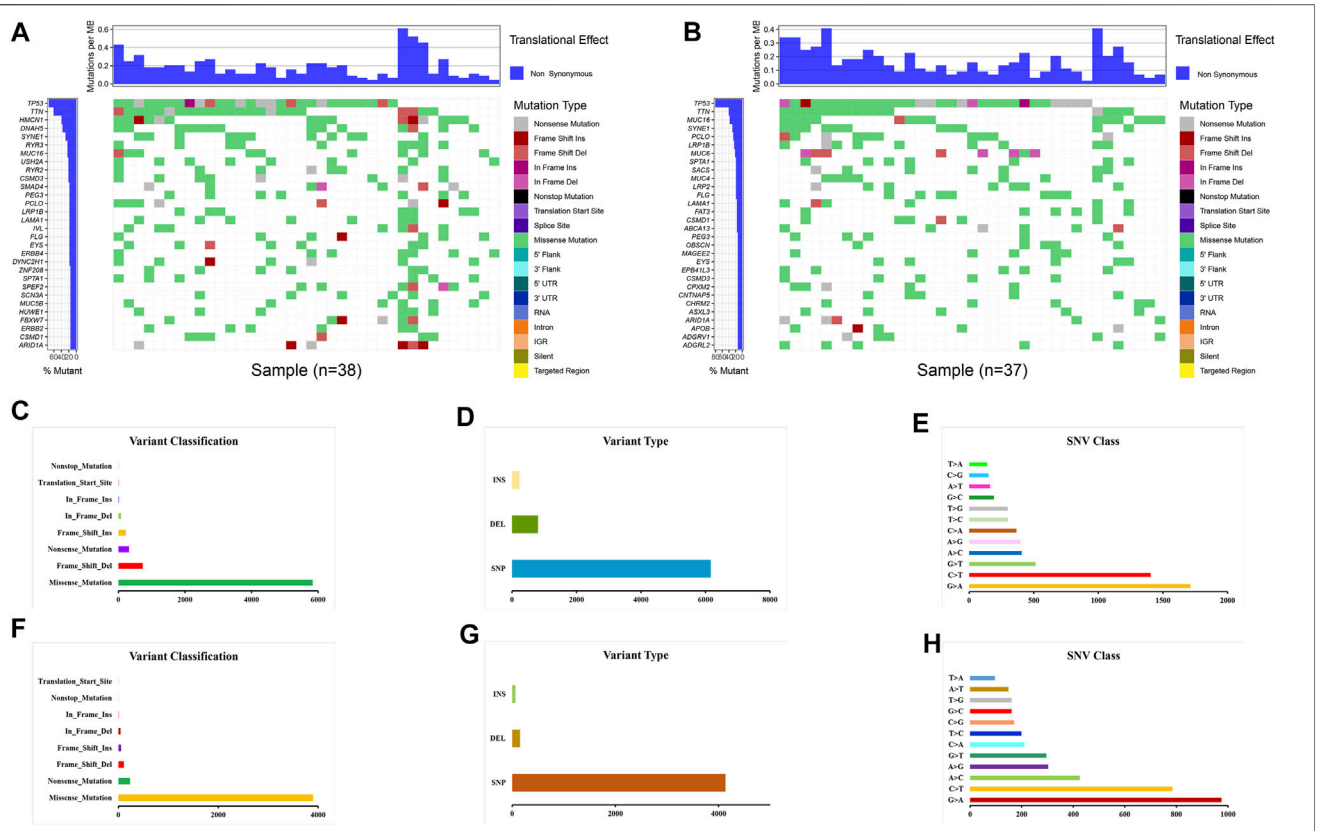


FIGURE 6 | Landscape of gene mutations in EAC patients. **(A)** Waterfall plot of mutational genes in the ESTIMATE ^{high} group. **(B)** Waterfall plot of mutational genes in the ESTIMATE ^{low} group. The left panel shows the gene mutational frequencies, and the right panel represents the mutational type. **(C)** Variant classification and frequency of gene mutations in the ESTIMATE ^{high} group. **(D)** Variant type in the ESTIMATE ^{high} group. **(E)** Frequency of SNV classes in the ESTIMATE ^{high} group. **(F)** Variant classification and frequency of gene mutations in the ESTIMATE ^{low} group. **(G)** Variant type in the ESTIMATE ^{low} group. **(H)** Frequency of SNV classes in the ESTIMATE ^{low} group.

Validation of Glycolysis and Its Impact on TME

For external cohort validation, GSE12898 was employed, which consisted of 75 EAC samples. Consistent with the classification result of the ESTIMATE algorithm in TCGA, this method successfully divided 75 patients with EAC into the ESTIMATE ^{high} and ESTIMATE ^{low} groups (n = 38 and 37, respectively). There were 3,428 genes between the two groups, including 52 significantly different glycolysis-related genes (**Supplementary Table S4**). The correlation analysis showed that 28 glycolytic genes had significant impacts on immune cells in the microenvironment, and the majority were B-cell and T-cell subtypes. The details are shown in **Table 2**.

Furthermore, the drug sensitivities were also analyzed between the ESTIMATE ^{high} and ESTIMATE ^{low} groups. The results showed that the IC50 level of bexarotene (**Figure 10A**) was higher in the ESTIMATE ^{high} group. However, the IC50 levels of camptothecin (**Figure 10B**), gemcitabine (**Figure 10C**), and vorinostat (**Figure 10F**) were lower in the ESTIMATE ^{high} group, implying that the patients in the ESTIMATE ^{high} group were more sensitive to the four drugs. There were no significant differences regarding IC50 in imatinib (**Figure 10D**) and methotrexate (**Figure 10E**). Taken together, glycolysis directly changed TME and indirectly influenced drug sensitivity.

DISCUSSION

The abilities of cancer cells to switch metabolisms and evade the immunity system in TME are well-documented characteristics in tumors. Elevated glycolysis is commonly observed in cancer progression and is associated with significant disruptions of a previous finely tuned microenvironment (Chang et al., 2015; Jiang et al., 2019a). As the tumors develop, they constantly interact with neighboring cells, such as stromal cells and immune cells, under the driver of genetic mutations, thus altering their phenotypes and functions (Butturini et al., 2019). In the context of these intricate crosstalks between malignant cells and non-malignant cells, the impacts of increased glycolysis and a dysregulated TME on immune response and effective therapy are of vital importance (Roma-Rodrigues et al., 2019). Although the research studies focusing on the tumor glycolysis and TME have exploded exponentially in recent years, the underlying mechanisms of how they act both independently and

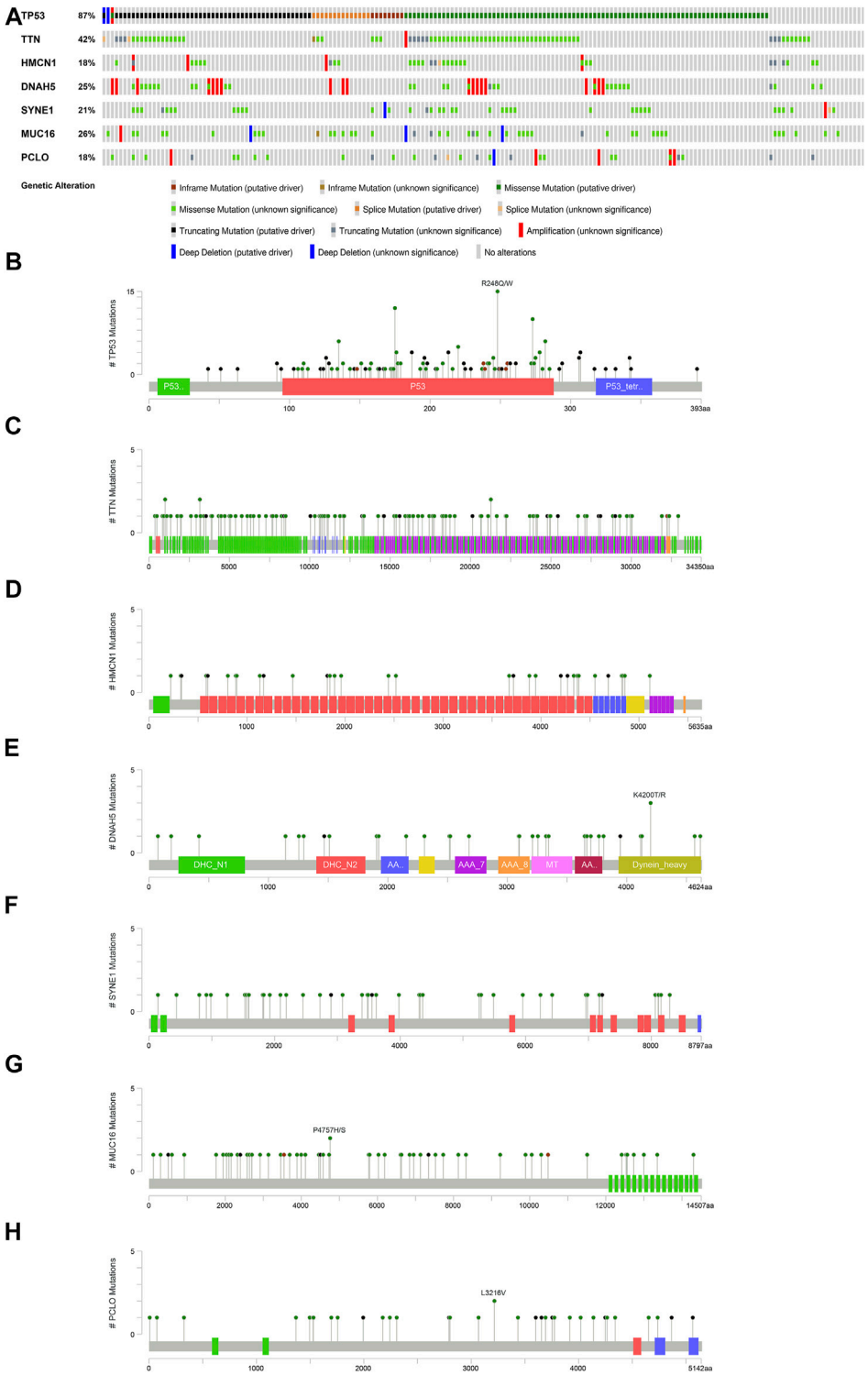


FIGURE 7 | Overview of the seven most common mutant genes in EAC patients. **(A)** Proportion and mutation type of genes. Different color bars represent different mutational types. **(B)** TP53-specific mutation site. **(C)** TTN-specific mutation site. **(D)** HMCN1-specific mutation site. **(E)** DNAH5-specific mutation site. **(F)** SYNE1-specific mutation site. **(G)** MUC16-specific mutation site. **(H)** PCLO-specific mutation site.

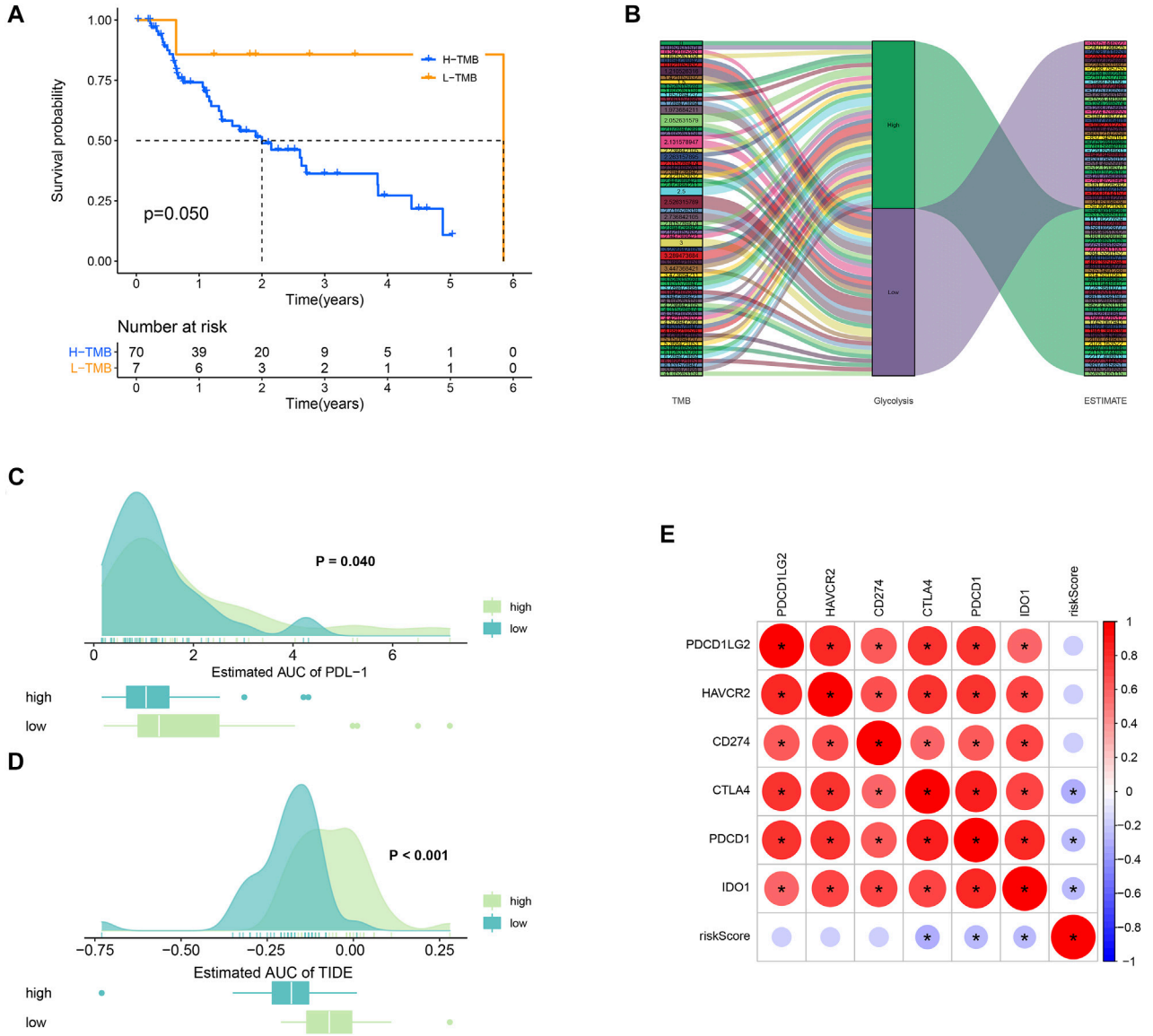


FIGURE 8 | TMB and immunotherapy response. **(A)** K-M survival plot. Patients with low TMB have significant probabilities to survive longer than those with high TMB. **(B)** Sankey diagram showing the relations between TMB, glycolysis, and the ESTIMATE score. **(C)** PD-L1 score comparison between high- and low-TMB groups. **(D)** TIDE algorithm analysis showed that patients with high TMB had better immunotherapeutic response than those with low TMB. **(E)** Correlations plots between the immune checkpoints.

synergistically still remain elusive. In this study, we explored systematically the links between glycolysis and TME in patients with EAC as well as the relations with genetic mutations. In the present study, we found that the glycolytic level was higher in the ESTIMATE ^{high} group, reasonably reflecting the fact that glycolysis may change the tumor purity. By constructing the predictive survival model based on IG signatures, we discovered that three genes (NUP88, SEH1L, and NUP37) may serve as independent prognostic biomarkers for OS in EAC. In addition, we revealed that gene mutation types and frequencies were distinct between the different ESTIMATE score groups, lending us a hypothesis that genetic alteration may drive TME changes. In addition, our data suggested that glycolysis could influence drug sensitivity and immunotherapeutic response.

Several studies have confirmed the close relationships between glycolysis and EAC (Lynam-Lennon et al., 2014; Harada et al., 2020; Kang et al., 2020). Consistent with these findings, our functional enrichment analysis results showed that IGs were strongly enriched in the glycolytic processes, ATP generation, the HIF−1 signal pathway, RNA activities, and so on. The presence of aerobic glycolysis under normal conditions efficiently promotes tumor cell growth by the following mechanisms: 1) the rate of glycolysis is accelerated at 100 times compared to oxidative phosphorylation to

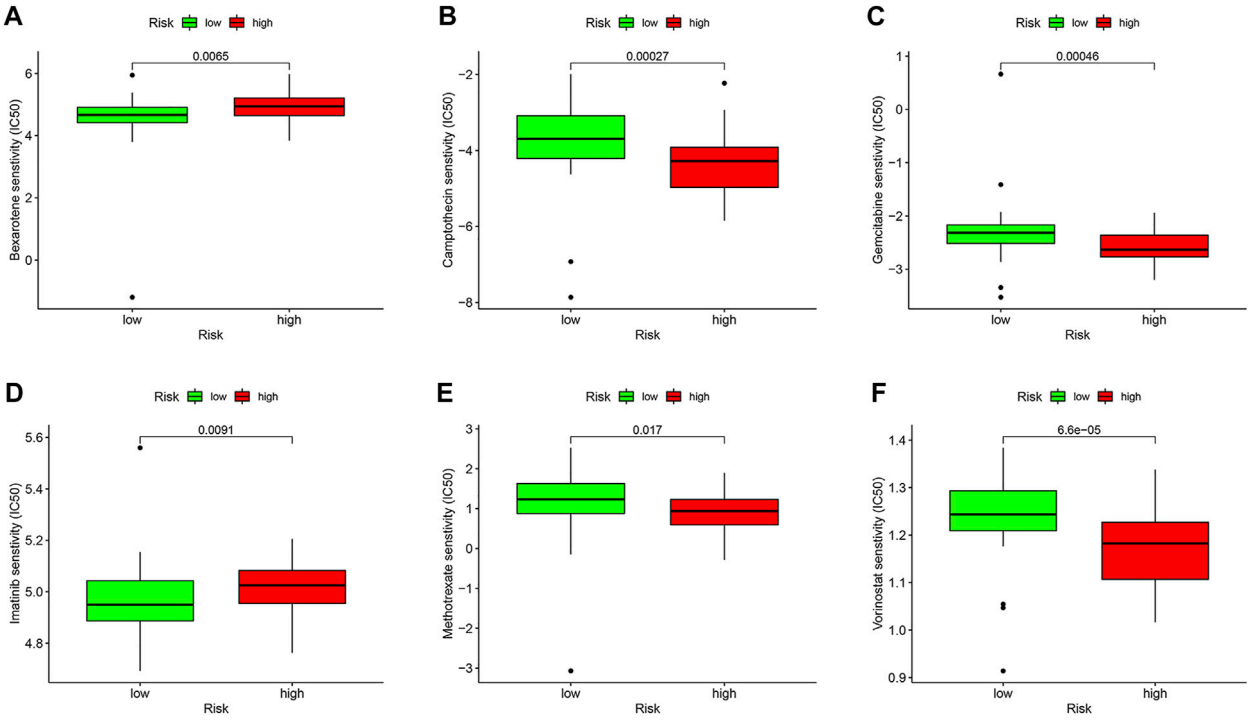


FIGURE 9 | Drug sensitivity analysis in the TCGA database. Box plots demonstrate the estimated IC50 values of bexarotene **(A)**, camptothecin **(B)**, gemcitabine **(C)**, imatinib **(D)**, methotrexate **(E),** and vorinostat **(F)**. The lower the IC50 value, the higher is the sensitivity to the drug.

compensate for the mathematical disadvantage in terms of net ATP production (glucose oxidative phosphorylation although mitochondria generate 18 times ATP compared to glycolysis) (Pfeiffer et al., 2001); 2) glycolysis provides sufficient and essential intermediates, such as NADPH and ribose-5-phosphate, which are indispensable for biosynthesis to meet avid proliferative requirements (Boroughs and DeBerardinis, 2015); 3) lactate, the obligatory product derived from glycolysis, could activate the HIF−1 signal pathway to induce vascular endothelial growth factor (VEGF) expression to stimulate angiogenesis in the microenvironment (Sonveaux et al., 2012). In addition, lactate is crucial to reorganize the tumor physical architectures in TME, and the accumulation of extracellular lactate is detrimental for normal healthy cells, such as immune cells (Romero-Garcia et al., 2016; Cassim and Pouyssegur, 2019). The study about how RNA interacts precisely with glycolysis is still in its infancy. However, Hua Q et al. gave us a hint that long non-coding RNA may promote glycolysis by sponging miRNA (Hua et al., 2019). Hence, it is reasonable to speculate that certain mutational genes regulated and modified by RNA may be involved in the glycolysis in EAC (Hochwald and Zhang, 2017).

The prognostic signature of glycolysis in EAC was established based on three genes (NUP88, SEH1L, and NUP37). NUP88, located at chromosome 17p13, encodes Nup88 protein (Zhao et al., 2012). Nup88 is a nucleoporin comprising nuclear pore complexes (NPCs) and plays critical roles in maintaining the spindle stability and preventing aneuploidy formation during mitosis (Hashizume et al., 2010). Unanimously, the GO results in our study also illustrated that glycolytic genes had intimacy with the nuclear pore, emphasizing the importance of nuclear proteins in glycolysis. The Nup88 overexpression is highly associated with tumor development and decreased survival, suggesting that NUP88 acts as an oncogene (Martínez et al., 1999; Naylor et al., 2016). In line with these studies, our results also proved that NUP88 was a risk factor for OS (HR > 1). Another prognostic gene is NUP37, which shares similarities with NUP88 and is also a member of NPC. It exerts primary functions of sustaining NPC integrity and modulates the cell cycle (Chen et al., 2019). Previous studies showed that the elevated expression of NUP37 is associated with worsened survival rates in liver cancer (Uhlen et al., 2017). In addition, a latest study by Huang L et al. demonstrates that NUP37 silencing induces inhibition of lung cancer cell proliferation (Huang et al., 2020). These findings were in agreement with our results, pointing out that NUP37 played an oncogenic role in OS (HR > 1). Nonetheless, the function of NUP37 in EAC has never been explored and needs further experiments to confirm *in vitro* and *in vivo*. SEH1L, also known as Seh1, is a part of NPC as well. However, the field is still in its infancy, and only a handful of animal models have been developed to investigate the role of SEH1L. Studies have shown that Seh1 could promote oligodendrocyte differentiation (Liu et al., 2019; Raices and D'Angelo, 2019). However, little is known about how it works in EAC. Undeniably, more research

TABLE 2 | Impacts of glycolytic genes on immune cells in TME.

Type	B n	B m	M2	Mono	CD4 a	CD4 r	CD4 n	Tfh	NK r	Neu	DC a	γδT
ADPGK	0.016*	0.490	0.037*	0.199	0.482	0.728	0.922	0.881	0.336	0.934	0.892	0.768
ALG1	0.037*	0.119	0.254	0.723	0.238	0.376	0.718	0.385	0.762	0.269	0.765	0.670
CACNA1H	0.457	0.535	0.031*	0.805	0.157	0.166	0.372	0.673	0.846	0.676	0.366	0.409
CHST12	0.230	0.739	0.201	0.010*	0.567	0.367	0.713	0.927	0.250	0.159	0.390	0.713
COL5A1	0.080	0.637	0.024*	0.018*	0.129	0.437	0.718	0.653	0.680	0.472	0.514	0.158
CXCR4	0.067	0.856	0.013*	0.339	0.298	0.769	0.718	0.881	0.260	0.678	0.828	0.768
DCN	0.037*	0.446	0.330	0.076	0.312	0.538 0.870	0.976	0.196	0.507	0.532	0.450	
DPYSL4	0.259	0.488	0.556	0.527	0.009*	0.727	0.137	0.976	0.361	0.738	0.702	0.120
DSC2	0.123	0.913	0.889	0.178	0.734	0.767	0.197	0.047*	0.719	0.330	0.460	0.155
FUT8	0.007*	0.490	0.126	0.287	0.555	0.320	0.922	0.590	0.196	0.934	0.765	0.974
GALK2	0.012*	0.075	0.303	0.891	0.216	0.689	0.922	0.590	0.492	0.580	0.807	0.870
GPC3	0.039*	0.053	0.300	0.094	0.713	0.267	0.869	0.470	0.455	0.559	1.000	0.120
GPC4	0.016*	0.119	0.277	0.643	0.312	0.347	0.718	0.385	0.237	0.934	0.496	0.577
GUSB	0.132	0.537	0.079	0.643	0.238	0.470	0.158	0.905	0.045*	0.619	0.683	0.224
HSPA5	0.037*	0.637	0.070	0.219	0.978	0.728	0.718	1.000	0.336	0.580	0.807	0.870
KDELR3	0.007*	0.490	0.126	0.287	0.555	0.320	0.922	0.590	0.196	0.934	0.765	0.974
NDUFV3	0.642	0.537	0.277	0.339	0.030*	0.574	0.158	0.255	0.051	0.333	0.978	0.818
NT5E	0.396	0.690	0.015*	0.076	0.555	0.689	0.533	0.491	0.309	0.176	0.957	0.718
NUP210	0.659	0.742	0.110	0.051	0.224	0.979	0.197	0.195	0.489	0.019*	0.956	0.974
PFKFB3	0.508	0.116	0.123	0.564	1.000	0.025*	0.447	0.880	0.934	1.000	0.035*	0.716
PLOD1	0.539	0.970	0.094	0.522	0.197	0.892	0.366	0.329	0.978	0.866	0.868	0.048*
PLOD2	0.870	0.046*	0.597	0.429	0.757	0.437	0.039*	0.072	0.309	0.825	0.870	0.158
SDC2	0.013*	0.361	0.153	0.197	0.224	0.936	0.372	0.833	0.978	0.867	0.622	0.120
STC1	0.338	0.690	0.079	0.604	0.933	0.503	0.224	0.454	0.022*	0.868	0.663	0.670
TGFBI	0.639	0.535	0.053	0.045*	0.045*	0.851	0.372	0.292	0.306	0.303	0.239	0.530
TPST1	0.934	0.635	0.077	0.011*	1.000	0.687	0.574	0.382	0.391	0.182	0.239	0.221
VCAN	0.218	0.856	0.021*	0.067	0.072	0.376	0.718	0.436	0.284	0.376	0.913	0.251
ZBTB20	0.119	0.537	0.172	0.643	0.086	0.810	0.158	0.811	0.039	0.619	0.978	0.577

B n: naive B-cells; B m: memory B-cells; M2: M2 macrophages; Mono: monocytes; CD4 a: memory-activated CD4 T-cells; CD4 r: memory resting CD4 T-cells; CD4 n: naive CD4 T-cells; Tfh: follicular helper T-cells; NK r: resting NK cells; Neu: neutrophils; DC a: activated dendritic cells; γδT: gamma delta T-cells.
**: p<0.05.*

studies are warranted to understand the specific effects of SEH1L on EAC.

The notion that glycolysis has profound impacts on immune cells in TME is well recognized. The investigation linking metabolic demands and immune cells was first documented in neutrophils and macrophages (Alonso and Nungester, 1956; Newsholme et al., 1986). Immune cells hold a resemblance with tumor cells to engage glycolysis, which require rapid energy sources to produce immune-mediators for migration and phagocytosis. Our study demonstrated that NUP88 had significant correlation with B-cells and mast cells (all $p < 0.05$) and NUP37 had positive correlation with T-cells and negative correlation with dendritic cells (DCs) and macrophages (all $p < 0.05$). These findings enforced the concept that glycolysis contributes to dramatic alterations of immune cells in TME, hence influencing the immune response and immune-based treatments. Glycolytic tumor cells compete with T-cells for glucose and impair T-cell activation (Peng et al., 2016). In addition, IFN-γ secreted by Th1 cells is sensitive and is unstable to lactate. Noteworthily, this could polarize the T-cells toward differentiating into the Th2 subpopulation that favors tumor progression by inhibiting the antitumor effect of M1 macrophages (Kareva and Hahnfeldt, 2013). Paradoxically, direct evidence from *in vivo* experiments supports that glycolysis could promote Th1 cell differentiation through an epigenetic mechanism (Peng et al., 2016). This calls for further studies aiming to shed light on the characteristics of glycolysis on T-cells. The divergent effects of glycolysis on DC are also observed. Glycolysis produces excessive lactate and lowers the pH in TME. Lactate together with decreased pH suppresses DC differentiation and consequently abolishes the antigen-presenting functions to T-cells (Harmon et al., 2020). However, inhibition of glycolysis also blocks DC maturation through HIF−1α. This highlights the multiple roles of DC in glycolysis and needs to be carefully interpreted in a context-dependent manner.

Gene mutations facilitate tumor cell metabolic plasticity to create favorable microenvironments beneficial to uncontrolled proliferation. Deepening our knowledge about the differences of gene alterations between different TMEs may allow for the development of therapeutic strategies. Spurred by this promising target, we explored upon this issue and showed that ESTIMATE ^{high} and ESTIMATE ^{low} groups manifested different gene mutation profiles, in which TP53 and TTN were the most prevalent mutant genes. Moreover, the most common type is missense mutation. Mutant TP53 endows tumor cells with adaptabilities to cope with the harsh microenvironment by providing adequate nutrients, thereby escaping from antitumor immune attack. Experimental mice models in the study by Basu S et al. testified that mutant TP53 could rewire the tumor glycolytic metabolism and enhance metastasis in TME (Basu et al., 2018). Mutant TP53 has additional impacts on TME beyond changing

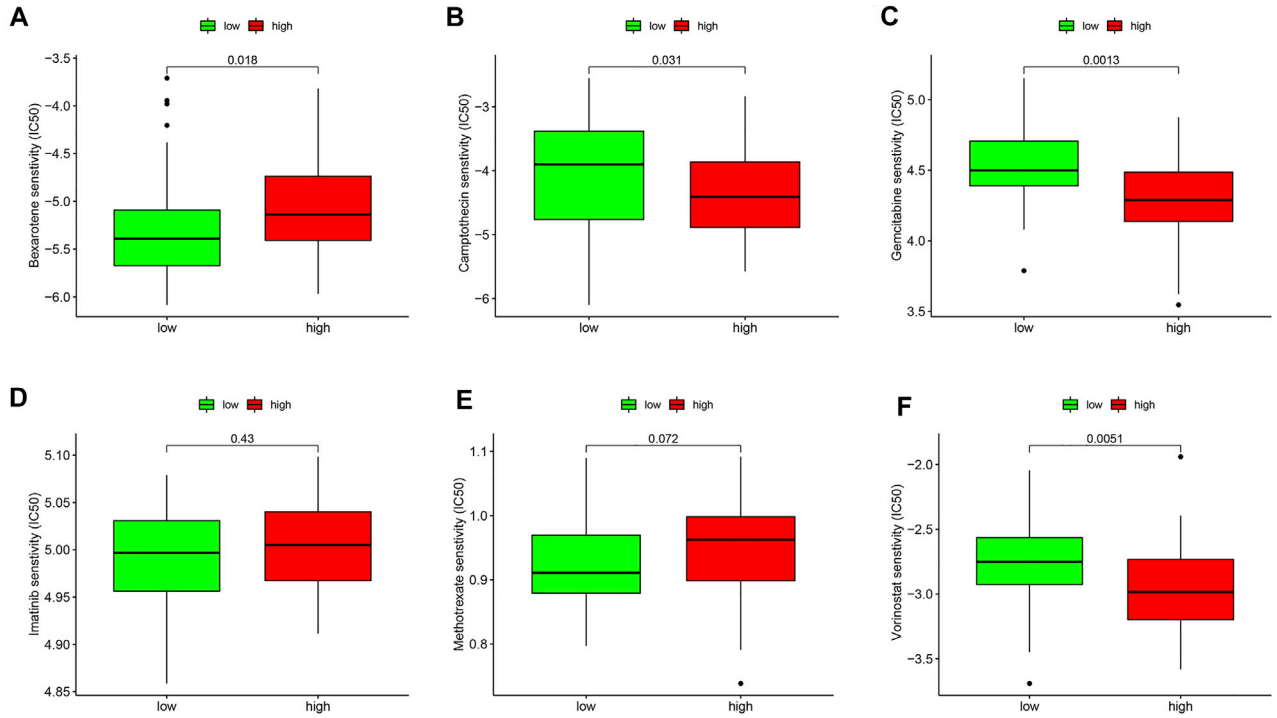


FIGURE 10 | Validation of drug sensitivity analysis in the GEO database. Box plots demonstrate the estimated IC50 values of bexarotene **(A)**, camptothecin **(B)**, gemcitabine **(C)**, imatinib **(D)**, methotrexate **(E),** and vorinostat **(F)**. The lower the IC50 value, the higher is the sensitivity to the drug.

tumor metabolic phenotypes. Myriad studies have indicated that mutant TP53 can remodel TME by several mechanisms. First, mutant TP53 could induce neo-angiogenesis by stimulating VEGF secretion (Kieser et al., 1994). Second, mutant TP53 could regulate chemokines that are involved in the homeostatic microenvironment (Yeudall et al., 2012). Last but not the least, mutant TP53 could reprogram the infiltrating immune cells and reshape the microenvironment (Cooks et al., 2018). TTN is located on chromosome 2q31, consisting of 364 exons, and it is the longest described coding gene (Chauveau et al., 2014). TTN ranks the third most abundant protein in both cardiac and skeletal muscle tissues, followed by actin and myosin (Chauveau et al., 2014). However, its mutation is not a rare entity in various cancers. Consistent with prior investigation by Cheng X et al., we also found that TTN missense mutation was a rather frequent type (Cheng et al., 2019). In addition, TTN and TP53 co-mutation is often accompanied during tumorigenesis and may serve as a prognostic biomarker either alone or in combination [61-63].

Current guidelines recommend adjuvant chemotherapy for patients at an advanced stage. However, how to select suitable patients who will benefit more from the chemotherapeutic regime is the prior concern. Our data demonstrated that patients with a high-risk score were more sensitive to camptothecin, gemcitabine, methotrexate, and vorinostat, suggesting that targeting glycolysis may alleviate the chemotherapeutic resistance. Despite immunotherapy bringing about a breakthrough for cancer patients, only a minority of patients could reap survival benefits actually. The TMB analysis in our study will accurately and effectively identify which patients will respond to immunotherapy in patients with EAC. Collectively, the proposed risk score system in our study has potency to help clinicians devise an individualized treatment strategy.

The strength of the present study is such that we performed a systematic analysis about glycolysis and TME in EAC for the first time based on the National Public Database, which provides robust data and statistical support. This study draws a close link between tumor glycolysis and the microenvironment and tentatively explains the mechanisms from the viewpoint of genetic alteration. Meanwhile, there are several limitations. First of all, TME is a complex mixture of parenchymal cells, the extracellular matrix, and numerous cytokines except for tumor and immune cells. These are not available from the public database and may greatly affect the analysis. Second, the results are not validated in *in vitro* and *in vivo* experiments. Last, the methods proposed in this study may not be applicable to all tumors as a result of heterogeneity. Notwithstanding its limitations, our study does provide an overview of glycolysis and TME in EAC, and this lays the foundation for further basic research in the area of metabolism and the microenvironment.

In summary, we found that glycolysis could change the microenvironment under the driver of genetic mutation and influence the immunotherapy in EAC. New efforts target that

EAC should incorporate the idea that the glycolytic metabolism could reshape TME. Further studies are necessary to confirm our conclusion.

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

LZ contributed to conception, design, and data acquisition of the study. FY contributed to interpretation and data analysis. XL contributed to data collection and interpreted the results. QL conceived and designed the study protocol. CZ reviewed and approved the final version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Natural Science Foundation of China (NSFC, No. 81773266) and the Key Discipline Group Construction Project of Pudong Health Bureau of Shanghai, China (No. PWZxq 2017-13) to QL.

ACKNOWLEDGMENTS

We would like to thank all the authors listed for their contribution to the present study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at https://www.frontiersin.org/articles/10.3389/fgene.2021.743133/full#supplementary-material

REFERENCES

Alonso, D., and Nungester, W. J. (1956). Comparative Study of Host Resistance of Guinea Pigs and Rats V. The Effect of Pneumococcal Products on Glycolysis and Oxygen Uptake by Polymorphonuclear Leucocytes. *J. Infect. Dis.* 99 (2), 174–181. doi:10.1093/infdis/99.2.174

Alsina, M., Moehler, M., and Lorenzen, S. (2018). Immunotherapy of Esophageal Cancer: Current Status, Many Trials and Innovative Strategies. *Oncol. Res. Treat.* 41 (5), 266–271. doi:10.1159/000488120

Alsop, B. R., and Sharma, P. (2016). Esophageal Cancer. *Gastroenterol. Clin. North America* 45 (3), 399–412. doi:10.1016/j.gtc.2016.04.001

Arnold, M., Soerjomataram, I., Ferlay, J., and Forman, D. (2015). Global Incidence of Oesophageal Cancer by Histological Subtype in 2012. *Gut* 64 (3), 381–387. doi:10.1136/gutjnl-2014-308124

Basu, S., Gnanapradeepan, K., Barnoud, T., Kung, C.-P., Tavecchio, M., Scott, J., et al. (2018). Mutant P53 Controls Tumor Metabolism and Metastasis by Regulating PGC-1α. *Genes Dev.* 32 (3-4), 230–243. doi:10.1101/gad.309062.117

Boroughs, L. K., and DeBerardinis, R. J. (2015). Metabolic Pathways Promoting Cancer Cell Survival and Growth. *Nat. Cell Biol* 17 (4), 351–359. doi:10.1038/ncb3124

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J. Clinicians* 68 (6), 394–424. doi:10.3322/caac.21492

Butturini, E., Carcereri de Prati, A., Boriero, D., and Mariotto, S. (2019). Tumor Dormancy and Interplay with Hypoxic Tumor Microenvironment. *Ijms* 20 (17), 4305. doi:10.3390/ijms20174305

Cassim, S., and Pouyssegur, J. (2019). Tumor Microenvironment: A Metabolic Player that Shapes the Immune Response. *Ijms* 21 (1), 157. doi:10.3390/ijms21010157

Chakraborty, H., and Hossain, A. (2018). R Package to Estimate Intracluster Correlation Coefficient with Confidence Interval for Binary Data. *Comp. Methods Programs Biomed.* 155, 85–92. doi:10.1016/j.cmpb.2017.10.023

Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., et al. (2019). Development of Tumor Mutation burden as an Immunotherapy Biomarker: Utility for the Oncology Clinic. *Ann. Oncol.* 30 (1), 44–56. doi:10.1093/annonc/mdy495

Chang, C.-H., Qiu, J., O'Sullivan, D., Buck, M. D., Noguchi, T., Curtis, J. D., et al. (2015). Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell* 162 (6), 1229–1241. doi:10.1016/j.cell.2015.08.016

Chauveau, C., Rowell, J., and Ferreiro, A. (2014). A Rising Titan:TTNReview and Mutation Update. *Hum. Mutat.* 35 (9), 1046–1059. doi:10.1002/humu.22611

Chen, J., Wo, D., Ma, E., Yan, H., Peng, J., Zhu, W., et al. (2019). Deletion of Low-Density Lipoprotein-Related Receptor 5 Inhibits Liver Cancer Cell Proliferation via Destabilizing Nucleoporin 37. *Cell Commun Signal* 17 (1), 174. doi:10.1186/s12964-019-0495-3

Cheng, X., Yin, H., Fu, J., Chen, C., An, J., Guan, J., et al. (2019). Aggregate Analysis Based on TCGA: TTN Missense Mutation Correlates with Favorable Prognosis in Lung Squamous Cell Carcinoma. *J. Cancer Res. Clin. Oncol.* 145 (4), 1027–1035. doi:10.1007/s00432-019-02861-y

Cooks, T., Pateras, I. S., Jenkins, L. M., Patel, K. M., Robles, A. I., Morris, J., et al. (2018). Mutant P53 Cancers Reprogram Macrophages to Tumor Supporting Macrophages via Exosomal miR-1246. *Nat. Commun.* 9 (1), 771. doi:10.1038/s41467-018-03224-w

Däster, S., Eppenberger-Castori, S., Mele, V., Schäfer, H. M., Schmid, L., Weixler, B., et al. (2020). Low Expression of Programmed Death 1 (PD-1), PD-1 Ligand 1 (PD-L1), and Low CD8+ T Lymphocyte Infiltration Identify a Subgroup of Patients with Gastric and Esophageal Adenocarcinoma with Severe Prognosis. *Front. Med.* 7, 144. doi:10.3389/fmed.2020.00144

Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R Package for Prediction of Clinical Chemotherapeutic Response from Tumor Gene Expression Levels. *PLoS One* 9 (9), e107468. doi:10.1371/journal.pone.0107468

Harada, K., Wu, C. C., Wang, X., Mizrak Kaya, D., Amlashi, F. G., Iwatsuki, M., et al. (2020). Total Lesion Glycolysis Assessment Identifies a Patient Fraction with a High Cure Rate Among Esophageal Adenocarcinoma Patients Treated with Definitive Chemoradiation. *Ann. Surg.* 272 (2), 311–318. doi:10.1097/SLA.0000000000003228

Harmon, C., O'Farrelly, C., and Robinson, M. W. (2020). The Immune Consequences of Lactate in the Tumor Microenvironment. *Adv. Exp. Med. Biol.* 1259, 113–124. doi:10.1007/978-3-030-43093-1_7

Hashizume, C., Nakano, H., Yoshida, K., and Wong, R. W. (2010). Characterization of the Role of the Tumor Marker Nup88 in Mitosis. *Mol. Cancer* 9, 119. doi:10.1186/1476-4598-9-119

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

Hochwald, J., and Zhang, J. (2017). Glucose Oncometabolism of Esophageal Cancer. *Acamc* 17 (3), 385–394. doi:10.2174/1871520616666160627092716

Hua, Q., Jin, M., Mi, B., Xu, F., Li, T., Zhao, L., et al. (2019). LINC01123, a C-Myc-Activated Long Non-coding RNA, Promotes Proliferation and Aerobic Glycolysis of Non-small Cell Lung Cancer through miR-199a-5p/c-Myc axis. *J. Hematol. Oncol.* 12 (1), 91. doi:10.1186/s13045-019-0773-y

Huang, L., Wang, T., Wang, F., Hu, X., Zhan, G., Jin, X., et al. (2020). NUP37 Silencing Induces Inhibition of Cell Proliferation, G1 Phase Cell Cycle Arrest and Apoptosis in Non-small Cell Lung Cancer Cells. *Pathol. - Res. Pract.* 216 (3), 152836. doi:10.1016/j.prp.2020.152836

Jiang, L., Zhao, L., Bi, J., Guan, Q., Qi, A., Wei, Q., et al. (2019). Glycolysis Gene Expression Profilings Screen for Prognostic Risk Signature of Hepatocellular Carcinoma. *Aging* 11 (23), 10861–10882. doi:10.18632/aging.102489

Jiang, Z., Liu, Z., Li, M., Chen, C., and Wang, X. (2019). Increased Glycolysis Correlates with Elevated Immune Activity in Tumor Immune Microenvironment. *EBioMedicine* 42, 431–442. doi:10.1016/j.ebiom.2019.03.068

Kang, H., Wang, N., Wang, X., Zhang, Y., Lin, S., Mao, G., et al. (2020). A Glycolysis-Related Gene Signature Predicts Prognosis of Patients with Esophageal Adenocarcinoma. *Aging* 12 (24), 25828–25844. doi:10.18632/aging.104206

Kareva, I., and Hahnfeldt, P. (2013). The Emerging "hallmarks" of Metabolic Reprogramming and Immune Evasion: Distinct or Linked? *Cancer Res.* 73 (9), 2737–2742. doi:10.1158/0008-5472.CAN-12-3696

Kieser, A., Weich, H. A., Brandner, G., Marmé, D., and Kolch, W. (1994). Mutant P53 Potentiates Protein Kinase C Induction of Vascular Endothelial Growth Factor Expression. *Oncogene* 9 (3), 963–969.

Kornberg, M. D., Bhargava, P., Kim, P. M., Putluri, V., Snowman, A. M., Putluri, N., et al. (2018). Dimethyl Fumarate Targets GAPDH and Aerobic Glycolysis to Modulate Immunity. *Science* 360 (6387), 449–453. doi:10.1126/science.aan4665

Li, D., Jiao, W., Liang, Z., Wang, L., Chen, Y., Wang, Y., et al. (2020). Variation in Energy Metabolism Arising from the Effect of the Tumor Microenvironment on Cell Biological Behaviors of Bladder Cancer Cells and Endothelial Cells. *Biofactors* 46 (1), 64–75. doi:10.1002/biof.1568

Lim, Y., Bang, J.-I., Han, S.-W., Paeng, J. C., Lee, K.-H., Kim, J. H., et al. (2017). Total Lesion Glycolysis (TLG) as an Imaging Biomarker in Metastatic Colorectal Cancer Patients Treated with Regorafenib. *Eur. J. Nucl. Med. Mol. Imaging* 44 (5), 757–764. doi:10.1007/s00259-016-3577-0

Liu, Z., Yan, M., Liang, Y., Liu, M., Zhang, K., Shao, D., et al. (2019). Nucleoporin Seh1 Interacts with Olig2/Brd7 to Promote Oligodendrocyte Differentiation and Myelination. *Neuron* 102 (3), 587–601. e7. doi:10.1016/j.neuron.2019.02.018

Lynam-Lennon, N., Connaughton, R., Carr, E., Mongan, A.-M., O'Farrell, N. J., Porter, R. K., et al. (2014). Excess Visceral Adiposity Induces Alterations in Mitochondrial Function and Energy Metabolism in Esophageal Adenocarcinoma. *BMC Cancer* 14, 907. doi:10.1186/1471-2407-14-907

Markar, S. R., Noordman, B. J., Mackenzie, H., Findlay, J. M., Boshier, P. R., Ni, M., et al. (2017). Multimodality Treatment for Esophageal Adenocarcinoma: Multi-center Propensity-Score Matched Study. *Ann. Oncol.* 28 (3), 519–527. doi:10.1093/annonc/mdw560

Martínez, N., Alonso, A., Moragues, M. D., Pontón, J., and Schneider, J. (1999). The Nuclear Pore Complex Protein Nup88 Is Overexpressed in Tumor Cells. *Cancer Res.* 59 (21), 5408–5411. doi:10.1172/JCI82277

Mayr, M., Becker, K., Schulte, N., Belle, S., Hofheinz, R., Krause, A., et al. (2012). Phase I Study of Imatinib, Cisplatin and 5-fluoruracil or Capecitabine in Advanced Esophageal and Gastric Adenocarcinoma. *BMC Cancer* 12, 587. doi:10.1186/1471-2407-12-587

Naylor, R. M., Jeganathan, K. B., Cao, X., and van Deursen, J. M. (2016). Nuclear Pore Protein NUP88 Activates Anaphase-Promoting Complex to Promote Aneuploidy. *J. Clin. Invest.* 126 (2), 543–559. doi:10.1172/JCI82277

Newsholme, P., Curi, R., Gordon, S., and Newsholme, E. A. (1986). Metabolism of Glucose, Glutamine, Long-Chain Fatty Acids and Ketone Bodies by Murine Macrophages. *Biochem. J.* 239 (1), 121–125. doi:10.1042/bj2390121

Peng, M., Yin, N., Chhangawala, S., Xu, K., Leslie, C. S., and Li, M. O. (2016). Aerobic Glycolysis Promotes T Helper 1 Cell Differentiation through an Epigenetic Mechanism. *Science* 354 (6311), 481–484. doi:10.1126/science.aaf6284

Pfeiffer, T., Schuster, S., and Bonhoeffer, S. (2001). Cooperation and Competition in the Evolution of ATP-Producing Pathways. *Science* 292 (5516), 504–507. doi:10.1126/science.1058079

Quante, M., Graham, T. A., and Jansen, M. (2018). Insights into the Pathophysiology of Esophageal Adenocarcinoma. *Gastroenterology* 154 (2), 406–420. doi:10.1053/j.gastro.2017.09.046

Raices, M., and D'Angelo, M. A. (2019). Nuclear Pore Complexes Are Key Regulators of Oligodendrocyte Differentiation and Function. *Neuron* 102 (3), 509–511. doi:10.1016/j.neuron.2019.04.025

Roma-Rodrigues, C., Mendes, R., Baptista, P., and Fernandes, A. (2019). Targeting Tumor Microenvironment for Cancer Therapy. *Ijms* 20 (4), 840. doi:10.3390/ijms20040840

Romero-Garcia, S., Moreno-Altamirano, M. M. B., Prado-Garcia, H., and Sánchez-García, F. J. (2016). Lactate Contribution to the Tumor Microenvironment: Mechanisms, Effects on Immune Cells and Therapeutic Relevance. *Front. Immunol.* 7, 52. doi:10.3389/fimmu.2016.00052

Smyth, E. C., Lagergren, J., Fitzgerald, R. C., Lordick, F., Shah, M. A., Lagergren, P., et al. (2017). Oesophageal Cancer. *Nat. Rev. Dis. Primers* 3, 17048. doi:10.1038/nrdp.2017.48

Sonveaux, P., Copetti, T., De Saedeleer, C. J., Végran, F., Verrax, J., Kennedy, K. M., et al. (2012). Targeting the Lactate Transporter MCT1 in Endothelial Cells Inhibits Lactate-Induced HIF-1 Activation and Tumor Angiogenesis. *PLoS One* 7 (3), e33418. doi:10.1371/journal.pone.0033418

Talukdar, F. R., di Pietro, M., Secrier, M., Moehler, M., Goepfert, K., Lima, S. S. C., et al. (2018). Molecular Landscape of Esophageal Cancer: Implications for Early Detection and Personalized Therapy. *Ann. N.Y. Acad. Sci.* 1434 (1), 342–359. doi:10.1111/nyas.13876

Taube, J. M., Galon, J., Sholl, L. M., Rodig, S. J., Cottrell, T. R., Giraldo, N. A., et al. (2018). Implications of the Tumor Immune Microenvironment for Staging and Therapeutics. *Mod. Pathol.* 31 (2), 214–234. doi:10.1038/modpathol.2017.156

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., et al. (2017). A Pathology Atlas of the Human Cancer Transcriptome. *Science* 357 (6352), eaan2507. doi:10.1126/science.aan2507

Wang, J. Z., Du, Z., Payattakool, R., Yu, P. S., and Chen, C.-F. (2007). A New Method to Measure the Semantic Similarity of GO Terms. *Bioinformatics* 23 (10), 1274–1281. doi:10.1093/bioinformatics/btm087

Wang, Y., Cheng, J., Xie, D., Ding, X., Hou, H., Chen, X., et al. (2018). NS1-binding Protein Radiosensitizes Esophageal Squamous Cell Carcinoma by Transcriptionally Suppressing C-Myc. *Cancer Commun.* 38 (1), 33. doi:10.1186/s40880-018-0307-y

Warburg, O., and Minami, S. (1923). Versuche an Überlebendem Carcinom-Gewebe. *Klin Wochenschr* 2 (17), 776–777. doi:10.1007/BF01712130

Wu, T., and Dai, Y. (2017). Tumor Microenvironment and Therapeutic Response. *Cancer Lett.* 387, 61–68. doi:10.1016/j.canlet.2016.01.043

Xu, D., Jin, J., Yu, H., Zhao, Z., Ma, D., Zhang, C., et al. (2017). Chrysin Inhibited Tumor Glycolysis and Induced Apoptosis in Hepatocellular Carcinoma by Targeting Hexokinase-2. *J. Exp. Clin. Cancer Res.* 36 (1), 44. doi:10.1186/s13046-017-0514-4

Yeudall, W. A., Vaughan, C. A., Miyazaki, H., Ramamoorthy, M., Choi, M.-Y., Chapman, C. G., et al. (2012). Gain-of-function Mutant P53 Upregulates CXC Chemokines and Enhances Cell Migration. *Carcinogenesis* 33 (2), 442–451. doi:10.1093/carcin/bgr270

Zhao, Q., Yu, J., and Meng, X. (2019). A Good Start of Immunotherapy in Esophageal Cancer. *Cancer Med.* 8 (10), 4519–4526. doi:10.1002/cam4.2336

Zhao, Z.-R., Zhang, L.-J., Wang, Y. Y., Li, F., Wang, M.-W., and Sun, X.-F. (2012). Increased Serum Level of Nup88 Protein Is Associated with the Development of Colorectal Cancer. *Med. Oncol.* 29 (3), 1789–1795. doi:10.1007/s12032-011-0047-1

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

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

APPENDIX A:

TABLE A1 | R Glycolytic Effects in Other Cancers

Glycolysis factors	Tumor	Glycolytic effects
GRG	Bladder cancer	Influences cell proliferation and cycle
GRG	Breast cancer	Pro-tumor immunity
glycolysis	Cervical cancer	Indirect effect
SPAG4, ENO3, et al	Colon adenocarcinoma	Double effects on prognosis
GRG	Colorectal cancer	Double effects on prognosis
STC1, CLDN9, et al	Gastric cancer	Influence the microenvironment and prognosis
GGESS	Glioblastoma	Poor prognosis and poor chemotherapy
GRGPs	Hepatocellular carcinoma	Double effects on prognosis and treatments
GRG	HNSCC	Double effects on prognosis and treatments
HMMR, GPC1, et al	Lung cancer	Poor prognosis and metastasis
GRG	Ovarian cancer	Influence the microenvironment and therapy
glycolysis	Pancreatic cancer	Promote progression and reduce drug sensitivity
GRG	Prostate cancer	Cell migration and invasion inhibition
CD44, PLOD1, et al	Renal cell carcinoma	Tumor-promoting
GRG	Uveal melanoma	Influence the microenvironment, prognosis, and therapy

GRG: glycolysis-related gene; GGESS: glycolytic gene expression signature score; GRGPs: glycolysis-related gene pairs; HNSCC: head and neck squamous cell carcinoma.

frontiers in Genetics

ORIGINAL RESEARCH
published: 17 December 2021
doi: 10.3389/fgene.2021.744304

Association of a Novel Prognosis Model with Tumor Mutation Burden and Tumor-Infiltrating Immune Cells in Thyroid Carcinoma

Siqin Zhang[1,2†], *Shaoyong Chen*[1,2†], *Yuchen Wang*[1,2], *Yuxiang Zhan*[1,2], *Jiarui Li*[1,2], *Xiaolin Nong*[1,2*] *and Biyun Gao*[1,2*]

[1]College of Stomatology, Guangxi Medical University, Nanning, China, [2]Guangxi Key Laboratory of Oral and Maxillofacial Surgery Disease Treatment, Nanning, China

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez,
Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Burcu Bakir-Gungor,
Abdullah Gül University, Turkey
Nazanin Hosseinkhan,
Iran University of Medical Science, Iran

***Correspondence:**
Xiaolin Nong
xnong@gxmu.edu.cn
Biyun Gao
joely918@163.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: 20 July 2021
Accepted: 25 November 2021
Published: 17 December 2021

Citation:
Zhang S, Chen S, Wang Y, Zhan Y, Li J, Nong X and Gao B (2021) Association of a Novel Prognosis Model with Tumor Mutation Burden and Tumor-Infiltrating Immune Cells in Thyroid Carcinoma. Front. Genet. 12:744304. doi: 10.3389/fgene.2021.744304

Although immunotherapy has recently demonstrated a substantial promise in treating advanced thyroid carcinoma (THCA), it is not appropriate for all THCA patients. As a result, this study aims to identify biomarkers for predicting immunotherapy efficacy and prognosis in THCA patients based on a constructed prognostic model. The transcriptomic and corresponding clinical data of THCA patients were obtained from the Cancer Genome Atlas (TCGA) database. We identified differentially expressed genes (DEGs) between THCA and normal samples and performed an intersection analysis of DEGs with immune-related genes (IRGs) downloaded from the ImmPort database. Functional enrichment analysis was performed on the chosen immune-related DEGs. Subsequently, Cox and LASSO regression analyses were conducted to obtain three hub immune-related DEGs, including PPBP, SEMA6B, and GCGR. Following that, a prognostic risk model was established and validated based on PPBP, SEMA6B, and GCGR genes to predict immunotherapy efficacy and THCA prognosis. Finally, we investigated the association between the constructed risk model and tumor mutational burden (TMB), abundance of tumor-infiltrating immune cells (TICs) as well as immunotherapeutic targets (PDL-1, PD-1, and CTLA4) in THCA. THCA patients in the high-risk score (RS) group showed higher TMB levels and worse prognosis than the low RS group. Patients in the high-RS group had higher proportions of monocytes, M2 macrophages, and activated dendritic cells, whereas those in the low-RS group exhibited higher numbers of M1 macrophages and dendritic resting cells. Our data implied that the constructed THCA prognostic model was sound and we concluded that the THCA patients having high TMB and low PD-L1 expression levels might respond poorly to immunotherapy. Taken together, we constructed a novel prognostic model for THCA patients to predict their prognosis and immunotherapy efficacy, providing a viable option for the future management of THCA patients in the clinic.

Keywords: tumor mutational burden, tumor-infiltrating immune cells, immunotherapy, prognosis, thyroid carcinoma

INTRODUCTION

Thyroid carcinoma (THCA) is the fifth most prevalent malignancy affecting women (Zhang et al., 2019b) and a major cause of annual endocrine malignancies death (Aboelnaga and Ahmed., 2015). Over the past decades, the incidence of THCA has been escalating worldwide (Cabanillas et al., 2016). THCA is generally classified into four pathological subtypes: papillary thyroid cancer (PTC, 80–85%), follicular thyroid cancer (FTC, 10–15%), medullary thyroid carcinoma (MTC, less than 2%) and anaplastic thyroid carcinoma (ATC, less than 2%) (Laha et al., 2020). Although most THCA is well-differentiated PTC with a 10-years survival rate of over 95%, some variants of PTC may demonstrate increased aggressive behavior, particularly in older patients, contributing to significant mortality (Nath and Erickson., 2018). Therefore, surgical resection and radioiodine (RAI) therapy were considered standard treatments for most THCA patients. Nonetheless, survival rates for advanced thyroid cancer patients remain low. For these reasons, novel therapeutic strategies, such as immunotherapies and targeted molecular therapy, are under investigation for treating advanced or metastatic THCA patients.

In recent years, immune checkpoint inhibitor (ICI) has achieved remarkable progress in treating breast cancer (Santa-Maria and Nanda., 2018), lung cancer (Anagnostou et al., 2017), melanoma (Amaria et al., 2018; Barrios et al., 2020) and squamous cell carcinoma of the head and neck (HNSCC) (Ferris et al., 2016; Ferris et al., 2018). It has been shown that numerous immune cells and their mediators exist in the tumor microenvironment (TME) of THCA with various interactions between them (Varricchi et al., 2019). In addition, it was demonstrated that increased frequency of regulatory and PD-1^+ T cells was associated with the recurrent or aggressive PTC (French et al., 2012), and that the frequency of PD-1^+ T cells was higher in patients with extranodal invasion than in those without lymph nodes metastases, indicating that it may be associated with the THCA prognosis. Furthermore, Gunda and his colleagues have recently revealed that anti-PD-1/PD-L1 therapy could beneficially modulate the immune microenvironment in orthotopic ATC murine model while simultaneously enhancing the efficacy of lenvatinib, a multi-targeted tyrosine kinase inhibitor (Gunda et al., 2019). The efficacy of ICIs has also yielded encouraging results in clinical trials, manifesting potent antitumor effects and improved tolerability in advanced FTC patients (Mehnert et al., 2019). However, the expensive cost of ICIs, which averages around $150,000 per year (Oiseth and Aziz., 2017), and their unpredictable efficacies (Giannone et al., 2020), prelude their widespread clinical use. Considering these factors, screening for appropriate molecular markers to improve treatment precision is highly demanding.

Tumor mutational burden (TMB), or the number of nonsynonymous mutations in a genomic region of somatic cells, is a biomarker employed to predict immunotherapy efficacy in various cancers (Samstein et al., 2019). As is widely known, tumor-infiltrating immune cells (TICs), an essential component of TME, are critical for in tumor initiation and progression (Bissell and Hines., 2011). Furthermore, gene locus mutations are observed in many histological subtypes of THCA, such as anaplastic lymphoma kinase (ALK), neurotrophic receptor tyrosine kinase 1 (NTRK1) genes rearrangements, BRAF and GTPase RAS family genes (Bos, 1989; Nikiforov and Nikiforova., 2011; Cancer Genome Atlas Research Network, 2014; Arndt et al., 2018; Varricchi et al., 2019). As high-throughput sequencing advances, large-scale acquisition of relevant cancer genomic data has become possible. However, few studies have conducted in-depth investigations to evaluate immunotherapy efficacy and prognosis in THCA.

Herein, based on Cancer Genome Atlas (TCGA) and ImmPort databases, we screened immune-related differentially expressed genes (IRDEGs), explored their functional enrichment, and constructed a prognostic prediction model. Additionally, we further investigated the association of TMB with immune infiltration and prognosis in THCA.

MATERIALS AND METHODS

Flow Diagram of Analysis

We designed a flow chart of analysis for the construction, validation and evaluation of the prognostic model in THCA. The analysis process was performed strictly according to the flow chart (**Figure 1**).

Data Downloading and Processing

We downloaded transcriptomic data with HTSeq-FPKM workflow type from TCGA of THCA project database (https://portal.gdc.cancer.gov/), including 510 tumor and 58 normal samples. Besides, masked somatic mutation data processed by VarScan software was also acquired from the GDC (Genomic Data Commons) data portal of TCGA database. Next, we downloaded clinical data of corresponding patients, including their age, gender, tumor grade, tumor stage, survival time and survival status.

Screening of Immune-Related Differentially Expressed Genes

Firstly, differentially expressed genes (DEGs) between normal and tumor tissues of THCA were analyzed using R software "limma" package[4], and the screening criteria were false discovery rate (FDR) < 0.05 and $|\log2FC| > 1$. Immune-related genes (IRGs) were obtained from ImmPort database (http://www.immport.org/) and then performed an intersection analysis on DEGs and IRGs to identify IRDEGs. R "venn" package was used to visualize to result of intersection analysis and R "pheatmap" package was performed to manifest the expression levels of IRDEGs of the THCA samples in the TCGA database.

Functional Enrichment Analysis

We used R package "org.Hs.eg.db" to obtain Entrez-ID of each IRDEG, and then performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses on IRDEGs using "clusterProfiler," "enrichplot," and "ggplot2" R

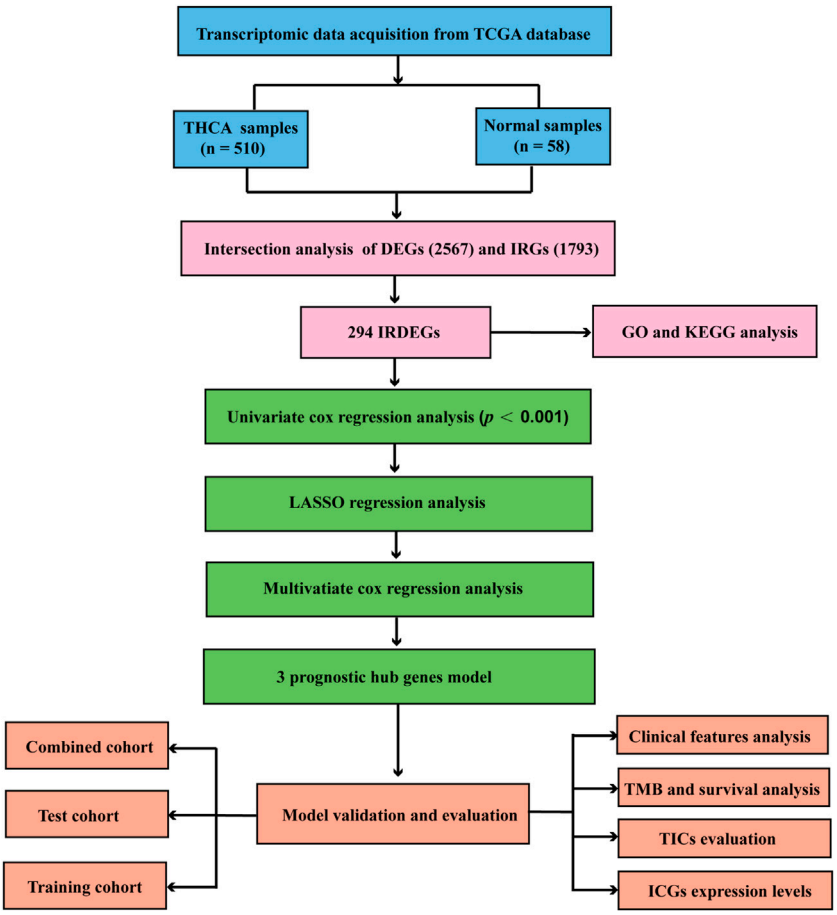


FIGURE 1 | Flow chart for construction and validation the THCA prognostic model. THCA, thyroid carcinoma.

packages. $p < 0.05$ and $q < 0.05$ were utilized as statistical significance thresholds.

Construction and Validation of a Prediction Model

For cross-validation, we randomly divided all samples into two groups with a 2:1 ratio using R package "caret", referred to as the training cohort and the test cohort, containing 2/3 and 1/3 of THCA cases, respectively. First, we determined IRDEGs associated with THCA survival in the training cohort using univariate Cox regression analysis with a threshold value of $p < 0.001$ via R package "survival". Then, key IRDEGs were identified using LASSO regression using R package "glmnet", and multivariate Cox regression analyses, thus constructing a THCA prediction model. Following that, we divided all THCA samples into high- and low-risk score (RS) groups according to their median RS. For high- and low-RS groups, Kaplan-Meier survival analysis and time-related receiver operating characteristic (ROC) curve analysis through R package "timeROC", were employed to cross-validate the predictive power of the model in the training test and the combined cohorts (containing all THCA samples).

Analysis of Tumor-Infiltrating Immune Cells

We utilized CIBERSORT algorithm (R script v 1.03), a deconvolution algorithm to quantify immune cell proportions based on transcriptomic expression profiles (Newman et al., 2015; Yao et al., 2020), to calculate the relative frequencies of 22 tumor-infiltrating immune cells (TIC) subtypes in tumor samples with high and low RS groups. $p < 0.05$ was utilized as statistical significance thresholds, and Wilcoxon test was performed to compare differences in the relative frequencies of various type of immune cells.

Analysis of High- and Low-RS Groups in Terms of their Clinical Information

After establishing the prediction model, we conducted the difference analysis to present the diversity of the clinical information between the high- and low-RS groups through R packages "ggpubr" and "limma". Next, Kaplan-Meier survival analysis was used to compare the survival outcome of patients between the two groups, and performed the difference analysis of TMB levels and immune checkpoint genes (ICGs) expression levels between high- and low-RS groups to predict their treatment responses when receiving immunotherapy.

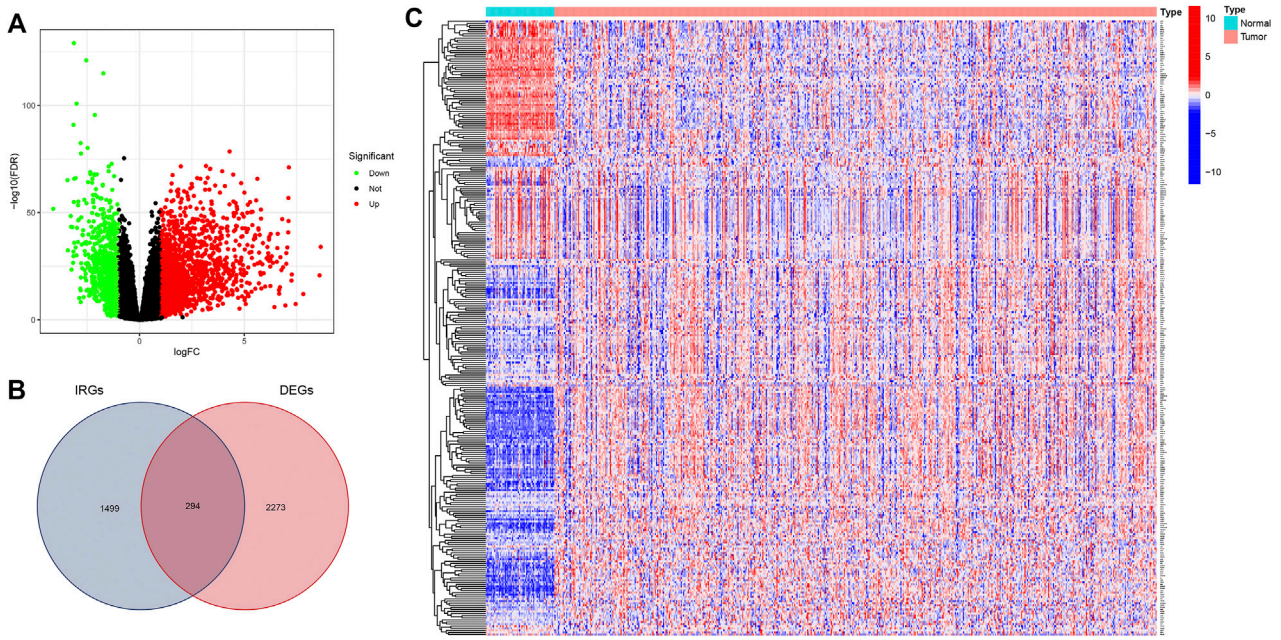


FIGURE 2 | Identification of immune-related DEGs in THCA. **(A)** Volcano plot for DEGs between THCA and normal samples from TCGA database. **(B)** Intersection of 2567 DEGs and 1793 IRGs in Venn plot **(C)** Heatmap showing immune-related IRDEGs. DEGs, differentially expressed genes; IRGs, immune-related genes.

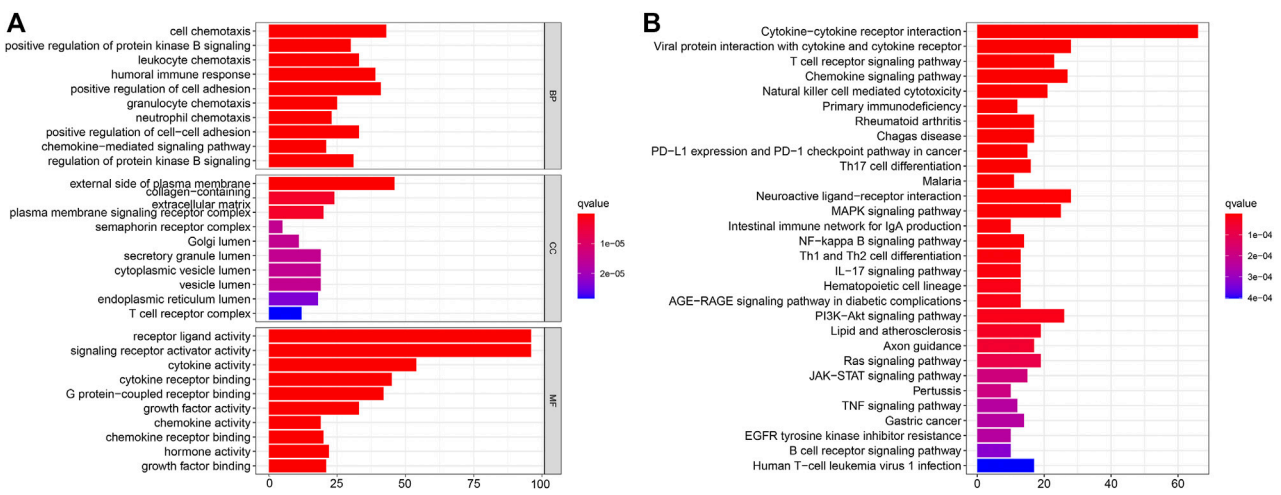


FIGURE 3 | Enrichment analysis of IRDEGs. **(A)** GO enrichment analysis. **(B)** KEGG pathway enrichment analysis. IRDEGs, immune-related differentially expressed gene; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological process; CC, cellular component; MF, molecular function.

Statistical Analysis

Statistical analyses were performed using R language (version 4.0.3). Kaplan-Meier survival analysis was used to assess the differences of the survival outcome between different groups. Univariate and multivariate Cox regression analyses were used to identify independent prognostic factors. p value < 0.05 was considered statistically significant in all tests.

RESULTS

Identification of DEGs and IRDEGs

We first used R software "limma" package[4] with false discovery rate (FDR) < 0.05 and |log2FC| >1 to identify DEGs between normal and THCA tissues, resulting in 2567 DEGs, as displayed in the volcano plot (**Figure 2A**). Then 294 IRDEGs were obtained

TABLE 1 | Cox regression analysis for screening of IRDEGs influencing the THCA patients' prognosis.

Gene ID	Univariate cox analysis		Multivariate cox analysis	
	HR (95%CI)	p-value	HR (95%CI)	p-value
PPBP	2.49 (1.50,4.13)	4.18E-04	3.21 (1.78,5.79)	1.02E-04
RBP4	3.20 (1.76,5.84)	1.42E-04		
SEMA6B	2.61 (1.49,4.62)	8.76E-04	2.46 (1.01,6.00)	4.74E-02
VGF	1.56 (1.22,2.00)	3.94E-04		
GCGR	1.75 (1.29,2.39)	3.72E-04	1.60 (1.05,2.44)	2.71E-02

IRDEGs, immune-related differentially expressed genes; THCA, thyroid carcinoma.

from the intersection analysis of 567 DEGs and 1793 IRGs from ImmPort database and displayed in a Venn diagram (**Figure 2B**). expression levels of 294 IRDEGs between tumor and normal samples were visualized via a heatmap (**Figure 2C**).

Functional Enrichment Analysis for IRDEGs

To further investigate relevant pathway, biological process, cellular component, and molecular function of 294 IRDEGs, we performed GO and KEGG pathway enrichment analysis for these IRDEGs, as shown in **Figure 3**. GO enrichment analysis indicated that these IRDEGs were mostly involved in chemotaxis-related activities, cell adhesion regulation and other immune-related responses (**Figure 3A**). Besides, KEGG pathway analysis revealed that these IRDEGs were chiefly enriched in immune-related pathways, such as cytokine–cytokine receptor interaction, Viral protein with cytokine receptor, T cell receptor signaling pathway, Chemokine signaling pathway, Natural killer cell mediated cytotoxicity, PD–L1 expression and PD–1 checkpoint pathway in cancer (**Figure 3B**).

Construction and Validation of THCA Prognosis Model

We first performed univariate Cox regression analysis on 294 IRDEGs to find the key IRDEGs that affect patients' survival outcome with a threshold value of $p < 0.001$, which five genes (PPBP, RBP4, SEMA6B, VGF and GCGR) were selected as being significantly associated with THCA patients' prognosis in this step (**Table 1**). Subsequently, LASSO regression was used for the following analyses (**Figures 4A–B**), thereby obtaining four candidate genes (PPBP, RBP4, SEMA6B, and GCGR). Eventually, three prognostic hub genes (PPBP, SEMA6B and GCGR) were screened from five these candidate genes by multivariate Cox regression analysis (**Table 1**). Next, we randomly assigned all THCA cases into training and test cohorts at a 2: 1 ratio for cross-validation, referred to as the training cohort and the test cohort, containing 2/3 and 1/3 of THCA cases, respectively. Additionally, no significant difference was observed in the clinical characteristics between the training and test cohorts ($p > 0.01$, **Table 2**). The THCA prognostic model was constructed based on the three hub genes and the RS of each patient in the prognostic model was calculated using the following formula: RS = (1.167391 × expression of PPBP) + (0.900831 × expression of SEMA6B) + (0.471683 × expression of GCGR), thereby dividing the patients into low- and high-RS groups according to their median value of RS. The heatmap for THCA tissues in the combined, test, and training three sets (**Figures 5A–C**) revealed that expression levels of three prognostic hub genes were downregulated in the low-RS group. RS distribution (**Figures 5D–F**) and RS-related survival status among

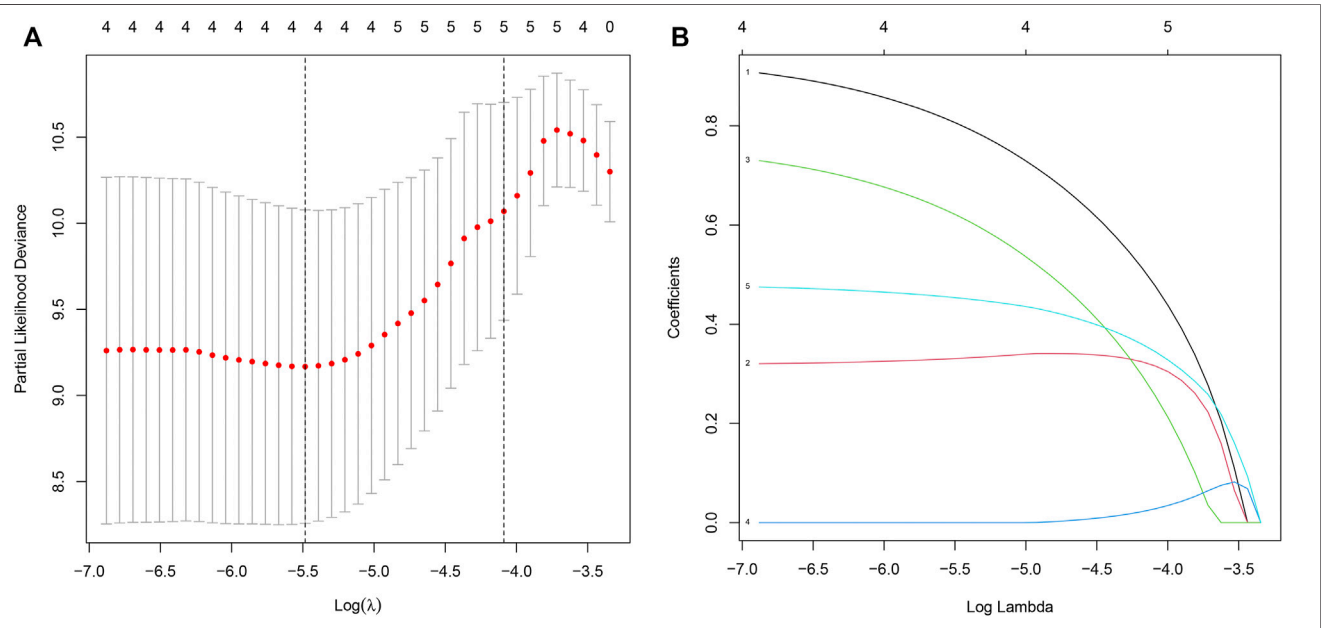


FIGURE 4 | Identification of IRDEGs associated with THCA prognosis. **(A–B)** LASSO coefficient profiles of prognostic-related IRDEGs. LASSO, least absolute shrinkage and selection operator. IRDEGs, immune-related differentially expressed genes.

TABLE 2 | Clinical characteristics between test and training cohorts.

Clinical characteristics		Number	Test (%)	Training (%)	*p*-value
Age	≤65	429 (85.97)	137 (83.54)	292 (87.16)	0.3376
	>65	70 (14.03)	27 (16.46)	43 (12.84)	
Gender	FEMALE	364 (72.95)	111 (67.68)	253 (75.52)	0.0811
	MALE	135 (27.05)	53 (32.32)	82 (24.48)	
Stage	Stage I-II	332 (66.53)	100 (60.98)	232 (69.25)	0.0889
	Stage III-IV	165 (33.07)	63 (38.41)	102 (30.45)	
	unknow	2 (0.4)	1 (0.61)	1 (0.3)	
T	T1-2	305 (61.12)	99 (60.37)	206 (61.49)	0.8227
	T3-4	192 (38.48)	65 (39.63)	127 (37.91)	
	TX	2 (0.4)	0 (0)	2 (0.6)	
M	M0	282 (56.51)	93 (56.71)	189 (56.42)	1
	M1	9 (1.8)	3 (1.83)	6 (1.79)	
	MX	208 (41.68)	68 (41.46)	140 (41.79)	
N	N0	229 (45.89)	70 (42.68)	159 (47.46)	0.2707
	N1	220 (44.09)	79 (48.17)	141 (42.09)	
	NX	50 (10.02)	15 (9.15)	35 (10.45)	

TX, unknown T stage; MX, unknown M stage; NX, unknown N stage.

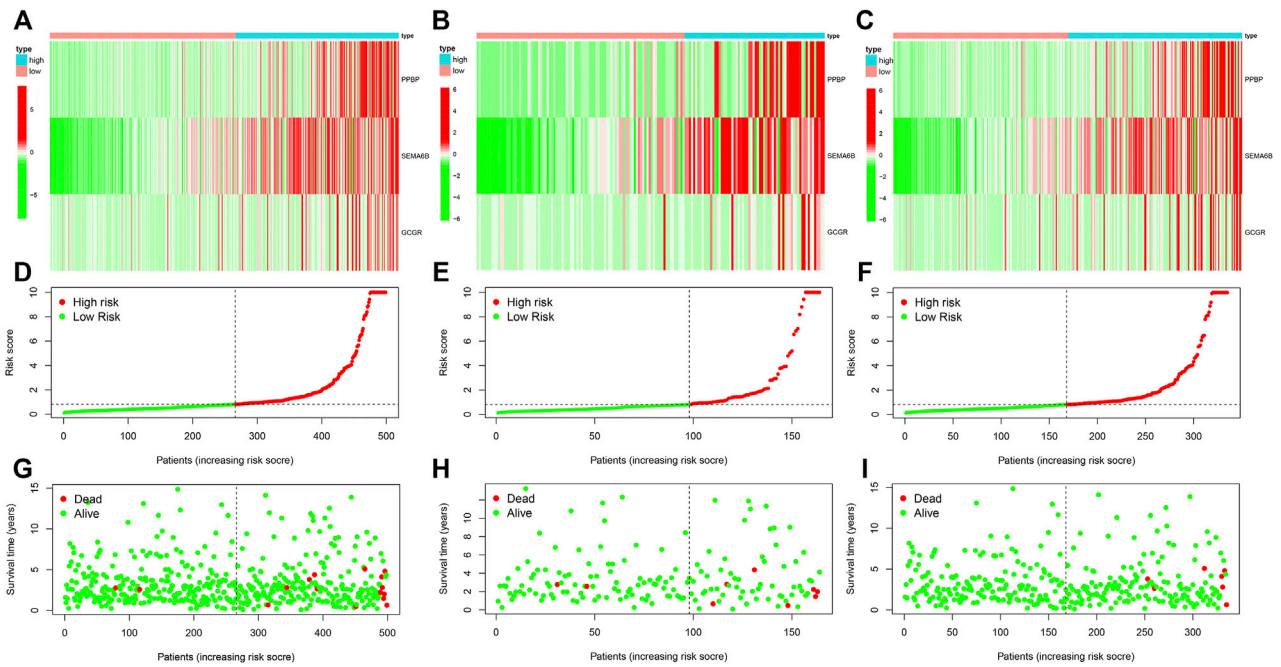


FIGURE 5 | Construction of THCA prognostic model **(A–C)** Expression levels of three prognostic hub IRDEGs among THCA patients in the combined test, and training cohorts, respectively. **(D–F)** The distribution of RS among THCA patients in three cohorts. **(G–I)** The survival status correlated with RS among THCA patients. THCA, thyroid carcinoma; RS, risk score.

patients (**Figures 5G–I**) indicated that higher RS corresponded to higher mortality risk in THCA patients.

To further validate the reliability of THCA prognostic model, we further performed survival and ROC curves analyses. Kaplan–Meier survival plots for THCA patients in the combined, test and training three sets indicated that patients in the high-RS group had a substantially worse survival outcome than those in the low-RS group (**Figures 6A–C**), with $p < 0.001$, $p = 0.023$, and $p = 0.006$, respectively. Besides, the combined set showed that area under the curve (AUC) values for 1-, 3-, 5-yearsurvival were 0.847, 0.722, and 0.781, respectively (**Figure 6D**). Similarly, each AUC for 1-, 3-, 5-years survival was more than 0.6 in the test and training sets (**Figures 6E, F**). In brief, these findings

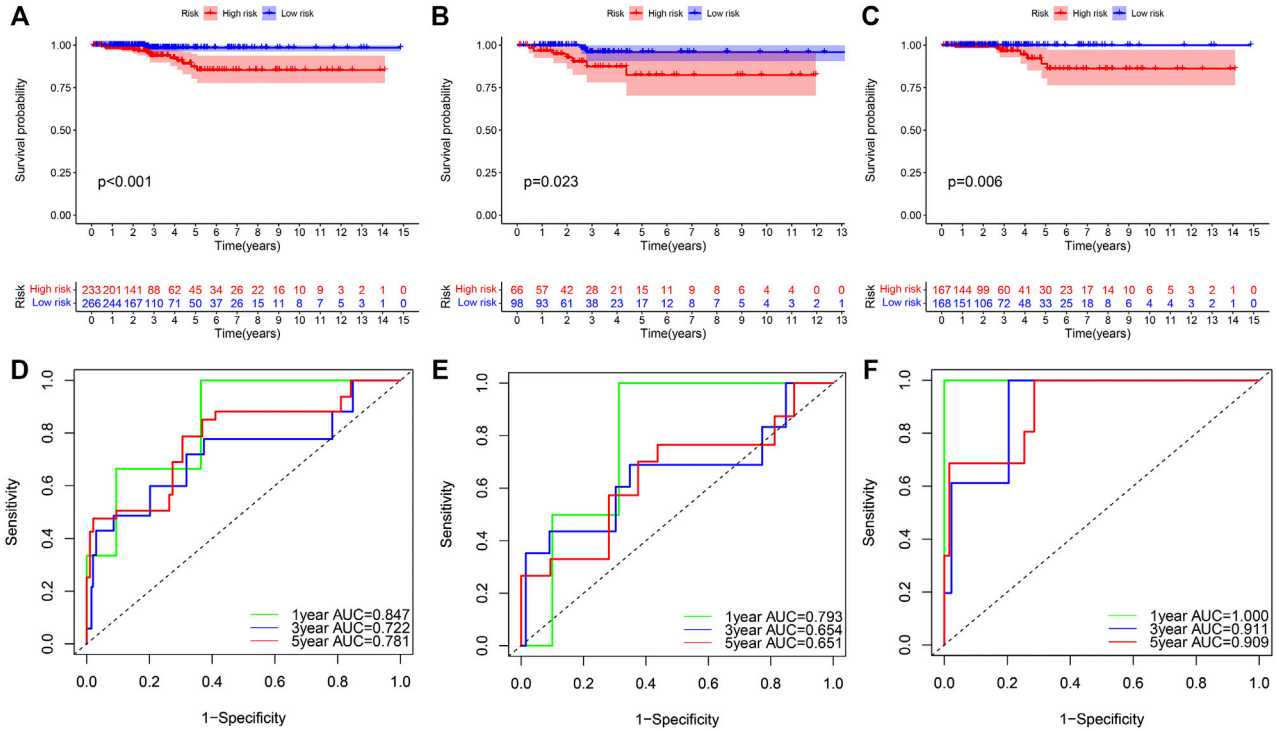


FIGURE 6 | Cross-validation of a prognostic model in THCA. **(A–C)** OS analysis of THCA patients with high- and low-RS in the combined test, and training cohorts, with $p < 0.001$, $p = 0.023$, and $p = 0.006$, respectively. **(D–F)** ROC curve analysis for 1–, 3–, 5-years validation of RS model in three different cohorts. THCA, thyroid carcinoma; OS, overall survival; RS, risk score; ROC, receiver operating characteristic.

TABLE 3 | Cox regression analysis for clinical characteristics and RS influencing prognosis in THCA patients.

Variable	Univariate cox analysis		Multivariate cox analysis	
	HR (95%CI)	*p*-value	HR (95%CI)	*p*-value
Age	1.16 (1.10,1.22)	2.27E-08	1.16 (1.09,1.22)	2.89E-07
Gender	1.92 (0.69,5.30)	2.09E-01		
Stage	2.42 (1.53,3.80)	1.39E-04	1.32 (0.74,2.37)	3.38E-01
Risk Score	1.01 (1.00,1.01)	5.88E-06	1.00 (1.00,1.01)	3.22E-03

RS, risk score; THCA, thyroid carcinoma.

demonstrated predictive accuracy and stability of the constructed model.

Identification of Independent Risk Factors for Prognosis in THCA Patients

We conducted univariate and multivariate Cox regression analyses to determine the correlation between RS and clinical features in THCA patients. The results revealed that prognosis of THCA patients was linked to three significant risk factors, including the age at diagnosis, pathological stage and RS (**Table 3**). Boxplots of the relationship between patients' clinical characteristics and RS revealed that patients aged over 65 years old (**Figure 7A**), had advanced pathological stages (**Figure 7C**), or had a higher T stage (**Figure 7D**) showed a higher RS, with $p = 0.003$, $p = 0.014$, and $p = 0.002$, respectively. However, no significant difference was observed in the correlations of RS with gender (**Figure 7B**), N (**Figure 7E**) and M (**Figure 7F**)stages, with $p > 0.05$. These results revealed that RS was associated with progression and development of THCA patients.

Associations of TMB With Prognosis and RS

To examine the value of TMB in THCA prognosis, we analyzed the associations between TMB and overall survival (OS) time or RS. The Kaplan–Meier survival plot revealed that THCA patients with high TMB exhibited shorter OS time than those with low TMB (**Figure 8A**, $p = 0.033$). Besides, THCA patients in the high-RS group presented higher TMB levels than those in the low-RS group (**Figure 8B**, $p = 0.0041$). Such results indicated a negative relationship between THCA prognosis model and TMB. Waterfall plots displayed a landscape of the top 20 somatic gene mutations in 475 THCA samples from TCGA database, with different colors signifying different mutation types (**Figures 8C,D**).

Evaluation of TICs and Immune Checkpoint Genes Expression

To investigate the effect of constructed prognostic model on TICs we used CIBERSORT algorithm to calculate the distribution of 22 TIC subtypes in THCA samples. The Wilcoxon rank-sum test was utilized to compare the proportions of different TIC subtypes in

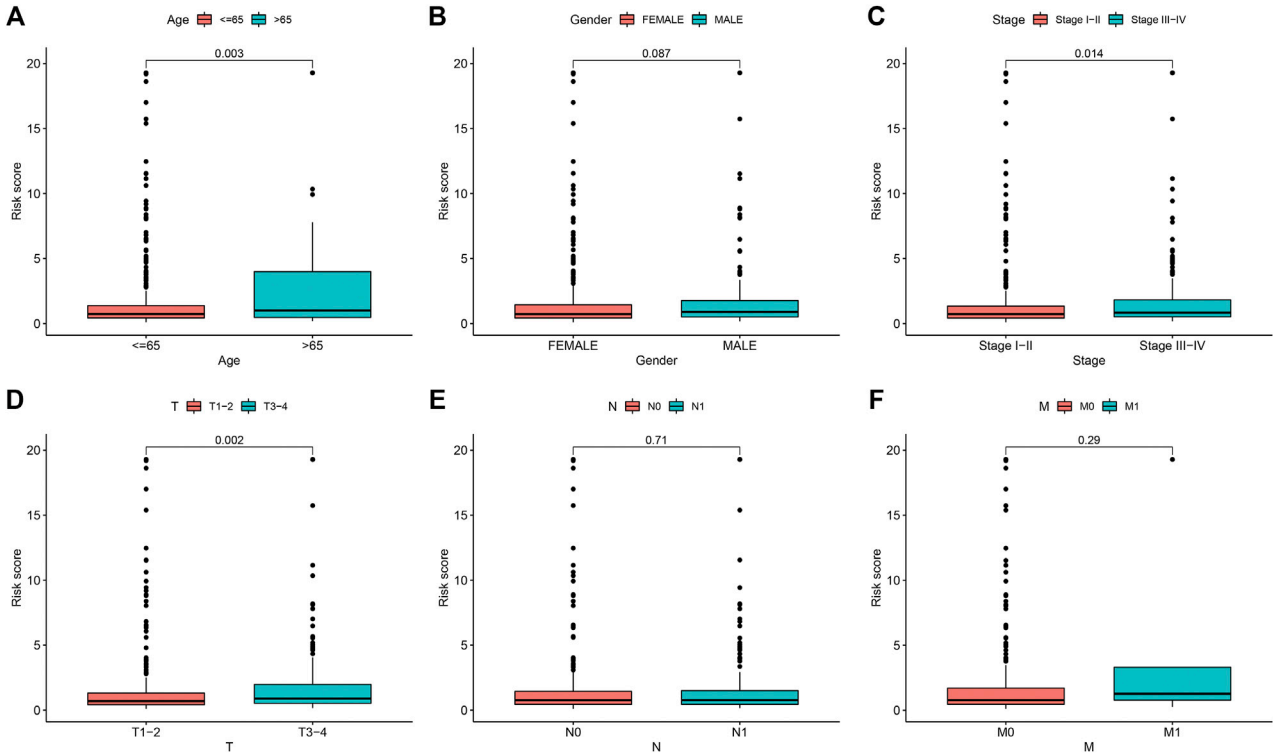


FIGURE 7 | Correlation between RS and clinical characteristics in THCA. **(A)** age > 65 years old and < 65 years old **(B)** female and male **(C)** Stage I-II and stage III–IV, **(D)** T stage **(E)** Nstage **(F)** Mstage. Age over 65 years old, higher pathological stage, and higher T stage correlated with higher RS, with $p = 0.003$, $p = 0.014$, and $p = 0.002$, respectively. RS, risk score; THCA, thyroid carcinoma.

high- and low-RS groups, and the findings were then visualized using a violin plot (**Figure 9A**). The results revealed that the relative abundance of monocytes, M2 macrophages, activated dendritic cells was significantly higher in high-RS group than those in low-RS group ($p = 0.017$, $p = 0.042$ and $p < 0.001$, respectively), whereas M1 macrophages and dendritic resting cells in high-RS group exhibited lower relative abundance ($p = 0.036$ and $p = 0.002$, respectively). According to their median TICs values, we classified THCA patients into high- and low- RS groups and then performed Kaplan–Meier survival analysis. The results displayed that high memory B cells was associated with poor OS ($p = 0.002$, **Figure 9B**).

To further evaluate the effectiveness of immunotherapy on THCA, we analyzed expression levels of immunotherapy targets (PD-L1, PD-1 and CTLA4) in the high- and low-RS groups (**Figure 10**). The results revealed that the low-risk group significantly increased PD–L1 expression levels compared with the high-risk group ($p = 0.044$, **Figure 10A**), implying that low-RS THCA patients might be more susceptible to immunotherapy.

DISCUSSION

THCA is the fifth most prevalent cancer in the United States, with an annual incidence increase of ~5% (Varricchi et al., 2019). Although most THCA patients have a favorable prognosis, approximately 15–20% of DTC patients, most ATC patients display resistance to standard treatment methods, such as RAI. Sorafenib and lenvatinib, two recently approved multikinase inhibitors (MKIs), have demonstrated encouraging outcomes in progressive, RAI-refractory DTC. However, adverse consequences have been identified, restricting their use (Cabanillas and Habra., 2016; Krajewska et al., 2017). Notably, the development of ICIs, such as anti–CTLA-4 and anti–PD-1 agents, have revolutionized THCA treatment (Antonelli et al., 2018; Varricchi et al., 2019). However, not all patients benefit from them due to individual variances. In the current study, we constructed a prognostic prediction model by screening IRDEGs to predict immunotherapy efficacy and survival outcome of THCA.

In the present study, we first identified 294 IRDEGs between THCA and normal samples and then explored potential functional enrichment pathways of these IRDEGs via KEGG and GO functional enrichment analyses. We discovered that the IRDEGs were primally enriched in some immune-related pathways, suggesting that IRDEGs might play a critical role in alteration of THCA immune microenvironment. Subsequently, three hub IRDEGs associated with THCA prognosis were identified, including PPBP, SEMA6B and GCGR. PPBP, alternatively known as C-X-C chemokine ligand 7, stimulates various cellular processes, such as DNA synthesis, glucose metabolism, intracellular cAMP accumulation and fibrinogen

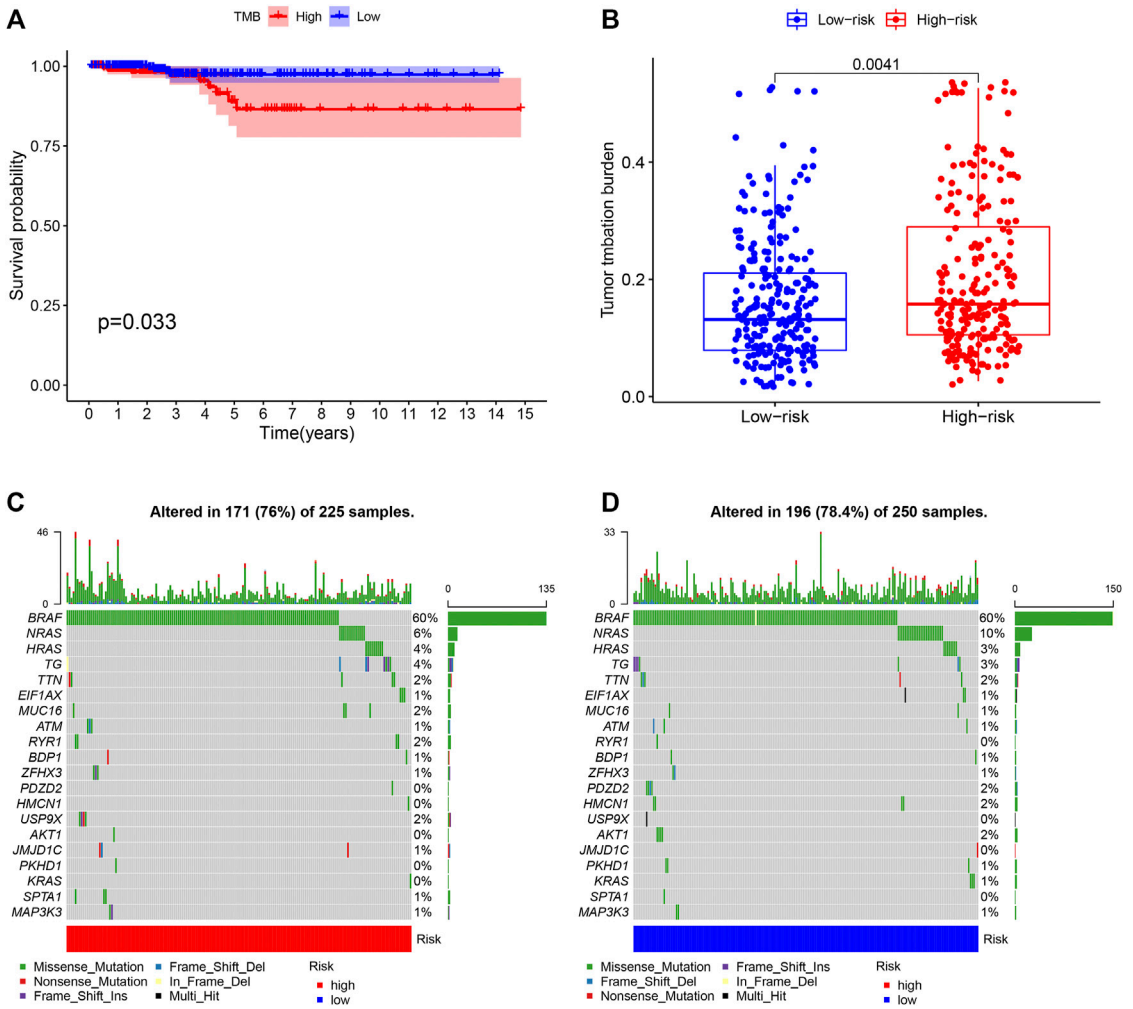


FIGURE 8 | Association of TMB with OS and RS **(A)** Survival analysis of THCA patients with high- and low-TMB **(B)** Comparisons of TMB in low- and high-RS groups **(C) (D)** Waterfall plot for mutation profiles of the top 20 genes in THCA samples of high- and low-RS groups, respectively. Annotations with different colors at the bottom referred to the various mutation types and bar chart above presented mutation burden. The right showed name of mutated genes and the right displayed percent of gene mutation. TMB, tumor mutation burden; OS, overall survival; RS, risk score.

activator synthesis. Some studies have indicated that PPBP and its encoded proteins might be linked to the progression of Wilms tumor (Guo et al., 2017) and gastric cancer (Yamamoto et al., 2019). SEMA6B a member of the semaphorin family, is mainly involved in developing peripheral and central nervous systems (Andermatt et al., 2014). SEMA6B has been demonstrated to have a critical role in the prognosis of various tumors, such as gliomas (Sun et al., 2019), breast cancer (Murad et al., 2006) and testicular cancer (Ji et al., 2020). A recent study has shown that SEMA6B promoted occurrence and development of THCA via regulating Notch signaling pathway (Lv et al., 2021). GCGR, a member of G protein-coupled receptor family, is critical in regulating glucose homeostasis. Previous studies indicated that GCGR aberrant expression might lead to adverse survival in endometrial stromal sarcoma (Davidson et al., 2013), renal cell carcinoma (Schrödter et al., 2016), and non-small cell lung cancer (NSCLC) (Li et al., 2020). In addition, it has been revealed that GCGR overexpression results in poor survival of PTC by activating epithelial-mesenchymal transition (EMT) and P38/ERK pathway (Jiang et al., 2020). Although we speculated that these genes might be a potential therapeutic target and/or prognostic biomarker for treating THCA patients, this hypothesis still requires additional validation in future studies.

Next, we firstly built a THCA prognostic model using the three screened hub genes (PPBP, SEMA6B and GCGR), and validated the reliability of the prognostic model in THCA. The results in the combined, test, and training sets revealed that THCA patients in the high-RS group exhibited poor survival outcomes than those in the low-RS group, and AUC values were over 0.6, implying that the constructed prognostic model accurately predicted OS in THCA. Besides, we found that older patients in the more advanced stage had significantly greater RS levels than younger ones in the earlier stage, consistent with the conclusion drawn by Ibrahimpasic et al. (Ibrahimpasic et al., 2019).

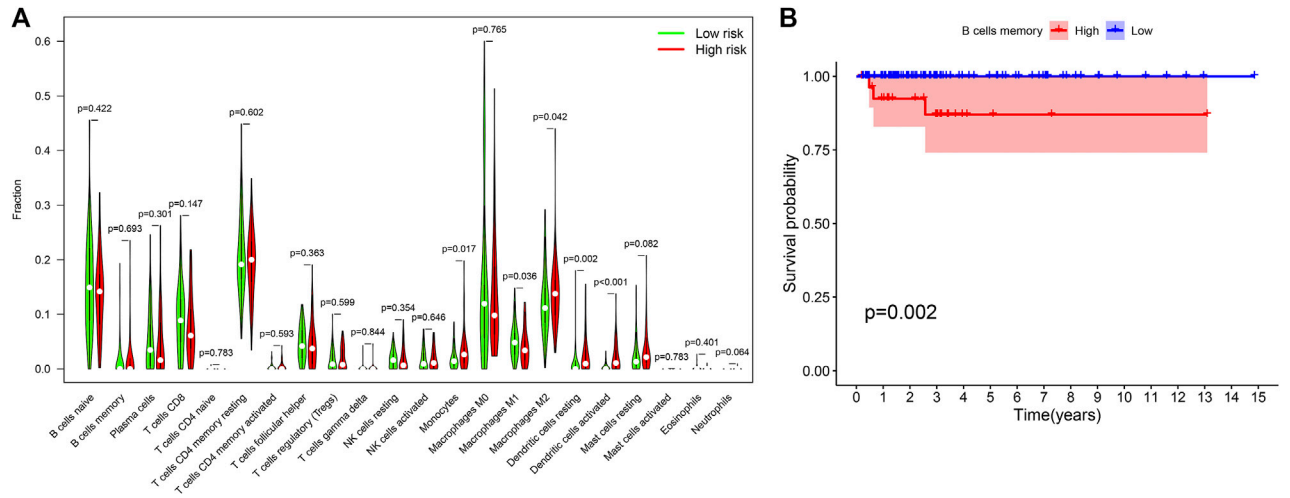


FIGURE 9 | TICs profile of THCA cases **(A)** Violin plot for quantification of 22 TIC subtypes between low- and high- RS groups. **(B)** Memory B cells influencing survival outcome of THCA patients. TIC, tumor-infiltrating immune cell; THCA, thyroid carcinoma; RS, risk score.

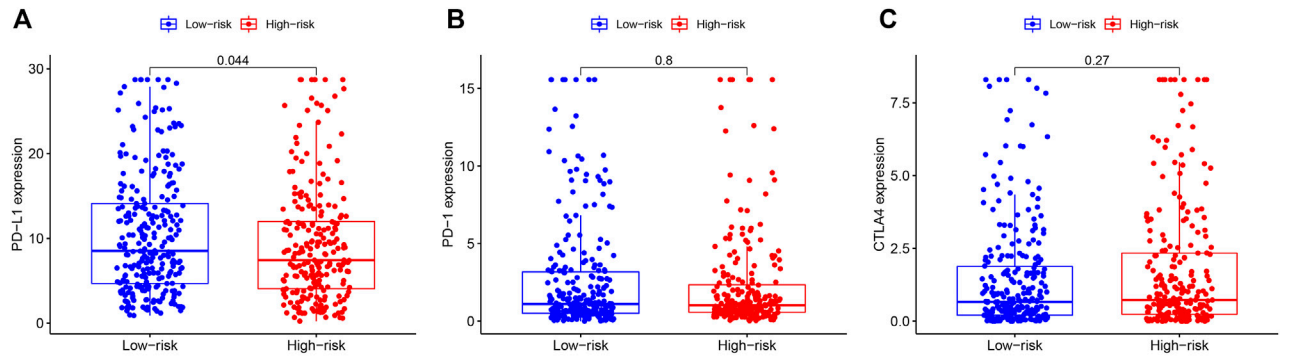


FIGURE 10 | Expression levels of ICGs in high- and low-RS groups **(A)** PD-L1 expression of THCA patients in high- and low-RS groups. **(B)** PD-1 expression of THCA patients in high- and low-RS groups. **(C)** CTLA4 expression of THCA patients in high- and low-RS groups. RS, risk score; PD-L1, programmed cell death ligand 1; THCA, thyroid carcinoma; PD-1, programmed death 1; CTLA4, cytotoxic T lymphocyte-associated protein 4.

TMB, as a novel biomarker, has been recently known to forecast the clinical efficacy of immunotherapy in many cancers (Samstein et al., 2019), since it is closely associated with tumor immune infiltration and microenvironment alteration (Kang et al., 2020). Zhou et al. (Zhou et al., 2021) discovered that patients in the low-TMB group exhibited a worse survival outcome and immune response than those in the high-TMB group in. However, few studies focused on the predictive value of TMB for immunotherapy in THCA patients. In this current study we revealed that THCA patients in the high-RS group possessed higher TMB than those in the low-RS group. Herein, we speculated that higher TMB might be correlated with a worse prognosis in THCA, consistent with the conclusion drawn by Zhang et al. (Zhang et al., 2019a) in their work on clear cell renal cell carcinoma. Besides considering the important roles of TICs in TME on prognosis of numerous malignancies (Zhang et al., 2003; Tran Janco et al., 2015), we further assessed the distribution of 22 TIC subtypes in THCA samples from high- and low-RS groups and the relationship with survival outcome. The analysis results revealed that patients in the high-RS group chiefly possessed a higher level of monocytes, M2 macrophages and activated dendritic cells (DCs). In contrast, those in the low-RS group exhibited higher proportions of M1 macrophages and resting DCs, manifesting that tumor-associated macrophages were associated with tumor progression (Jackaman et al., 2017). Nevertheless, the underlying mechanism remains yet unclear. Similarly, Travers et al. (Travers et al., 2019) found that increasing tumor-killing M1 macrophages and decreasing M2 macrophages in TME contributed to reduced TMB and improved survival in mice with ovarian cancer. Additionally, a study indicated a strong correlation between DCs and advanced patients with PTC (Bergdorf et al., 2019), which might explain why THCA samples in the high-RS group exhibited advanced pathological stages. Besides, it has been reported that tumor-infiltrating memory B cells are linked to

superior clinical outcomes in breast cancer (Garaud et al., 2019). In contrast, in the present study, we revealed that a high percent of memory B cells was highly correlated with poor survival in THCA.

Immune checkpoint molecules, such as PD-L1, PD-1 and CTLA4, have been demonstrated to connect with the efficacy of immunotherapy (Kalbasi and Ribas., 2020) Hence, we further investigated these biomarkers, including PD-L1, PD-1 and CTLA4, expression levels of THCA patients between in the high- and low-RS group. Results showed that PD-L1 expression was significantly upregulated in the low-RS group compared to the high-RS group, suggesting that this prognostic model might have ability to determine THCA patients' response to immunotherapy. Based on these findings, we supposed that high TMB and low PD-L1 expression in THCA patients might respond poorly to immunotherapy.

CONCLUSIONS

In conclusion, this study constructed and validated a THCA prognostic prediction model based on TCGA database, displaying good predictability and reliability for THCA prognosis. To our knowledge, our group was the first to screen out three potential therapeutic target genes and elucidate the association of TICs with RS and OS in THCA. Additionally, we figured out that THCA patients in the high-RS group had high TMB and low PD-L1 expression, establishing a baseline and reference for predicting THCA immunotherapy efficacy in clinical trials. However, future investigations require relevant basic experiments and large sample clinical trials.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

SZ and SC wrote the manuscript. SC, YW, YZ, and JL analyzed data. BG and XN were responsible for data acquisition and interpretation SZ and SC revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the project of improving scientific ability of young college teachers in Guangxi (2021KY0110), Nanning Qingxiu District Science and Technology project (2020027), and Guangxi Health Commission project (Z20201043).

REFERENCES

Aboelnaga, E. M., and Ahmed, R. A. (2015). Difference between Papillary and Follicular Thyroid Carcinoma Outcomes: an Experience from Egyptian Institution. *Cancer Biol. Med.* 12 (1), 53–59. doi:10.7497/j.issn.2095-3941.2015.0005

Amaria, R. N., Reddy, S. M., Tawbi, H. A., Davies, M. A., Ross, M. I., Glitza, I. C., et al. (2018). Neoadjuvant Immune Checkpoint Blockade in High-Risk Resectable Melanoma. *Nat. Med.* 24 (11), 1649–1654. doi:10.1038/s41591-018-0197-1

Anagnostou, V., Smith, K. N., Forde, P. M., Niknafs, N., Bhattacharya, R., White, J., et al. (2017). Evolution of Neoantigen Landscape during Immune Checkpoint Blockade in Non-small Cell Lung Cancer. *Cancer Discov.* 7 (3), 264–276. doi:10.1158/2159-8290.cd-16-0828

Andermatt, I., Wilson, N. H., Bergmann, T., Mauti, O., Gesemann, M., Sockanathan, S., et al. (2014). Semaphorin 6B Acts as a Receptor in post-crossing Commissural Axon Guidance. *Development* 141 (19), 3709–3720. doi:10.1242/dev.112185

Antonelli, A., Ferrari, S. M., and Fallahi, P. (2018). Current and Future Immunotherapies for Thyroid Cancer. *Expert Rev. Anticancer Ther.* 18 (2), 149–159. doi:10.1080/14737140.2018.1417845

Arndt, A., Steinestel, K., Rump, A., Sroya, M., Bogdanova, T., Kovgan, L., et al. (2018). Anaplastic Lymphoma Kinase (ALK) Gene Rearrangements in Radiation-Related Human Papillary Thyroid Carcinoma after the Chernobyl Accident. *J. Path: Clin. Res.* 4 (3), 175–183. doi:10.1002/cjp2.102

Barrios, D. M., Do, M. H., Phillips, G. S., Postow, M. A., Akaike, T., Nghiem, P., et al. (2020). Immune Checkpoint Inhibitors to Treat Cutaneous Malignancies. *J. Am. Acad. Dermatol.* 83 (5), 1239–1253. doi:10.1016/j.jaad.2020.03.131

Bergdorf, K., Ferguson, D. C., Mehrad, M., Ely, K., Stricker, T., and Weiss, V. L. (2019). Papillary Thyroid Carcinoma Behavior: Clues in the Tumor Microenvironment. *Endocrine-Related Cancer* 26 (6), 601–614. doi:10.1530/erc-19-0074

Bissell, M. J., and Hines, W. C. (2011). Why Don't We Get More Cancer? A Proposed Role of the Microenvironment in Restraining Cancer Progression. *Nat. Med.* 17 (3), 320–329. doi:10.1038/nm.2328

Bos, J. L. (1989). Ras Oncogenes in Human Cancer: a Review. *Cancer Res.* 49 (17), 4682–4689.

Cabanillas, M. E., and Habra, M. A. (2016). Lenvatinib: Role in Thyroid Cancer and Other Solid Tumors. *Cancer Treat. Rev.* 42, 47–55. doi:10.1016/j.ctrv.2015.11.003

Cabanillas, M. E., McFadden, D. G., and Durante, C. (2016). Thyroid Cancer. *The Lancet* 388 (10061), 2783–2795. doi:10.1016/s0140-6736(16)30172-6

Cancer Genome Atlas Research Network (2014). Integrated Genomic Characterization of Papillary Thyroid Carcinoma. *Cell* 159 (3), 676–690. doi:10.1016/j.cell.2014.09.050

Davidson, B., Abeler, V. M., Hellesylt, E., Holth, A., Shih, I.-M., Skeie-Jensen, T., et al. (2013). Gene Expression Signatures Differentiate Uterine Endometrial Stromal Sarcoma from Leiomyosarcoma. *Gynecol. Oncol.* 128 (2), 349–355. doi:10.1016/j.ygyno.2012.11.021

Ferris, R. L., Blumenschein, G., Jr., Fayette, J., Guigay, J., Colevas, A. D., Licitra, L., et al. (2018). Nivolumab vs Investigator's Choice in Recurrent or Metastatic Squamous Cell Carcinoma of the Head and Neck: 2-year Long-Term Survival Update of CheckMate 141 with Analyses by Tumor PD-L1 Expression. *Oral Oncol.* 81, 45–51. doi:10.1016/j.oraloncology.2018.04.008

Ferris, R. L., Blumenschein, G., Jr., Fayette, J., Guigay, J., Colevas, A. D., Licitra, L., et al. (2016). Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *N. Engl. J. Med.* 375 (19), 1856–1867. doi:10.1056/NEJMoa1602252

French, J. D., Kotnis, G. R., Said, S., Raeburn, C. D., McIntyre, R. C., Jr., Klopper, J. P., et al. (2012). Programmed Death-1+ T Cells and Regulatory T Cells Are Enriched in Tumor-Involved Lymph Nodes and Associated with Aggressive

Features in Papillary Thyroid Cancer. *J. Clin. Endocrinol. Metab.* 97 (6), E934–E943. doi:10.1210/jc.2011-3428

Garaud, S., Buisseret, L., Solinas, C., Gu-Trantien, C., de Wind, A., Van den Eynden, G., et al. (2019). Tumor-infiltrating B Cells Signal Functional Humoral Immune Responses in Breast Cancer. *JCI Insight* 4 (18). doi:10.1172/jci.insight.129641

Giannone, G., Ghisoni, E., Genta, S., Scotto, G., Tuninetti, V., Turinetto, M., et al. (2020). Immuno-Metabolism and Microenvironment in Cancer: Key Players for Immunotherapy. *Ijms* 21 (12), 4414. doi:10.3390/ijms21124414

Gunda, V., Gigliotti, B., Ashry, T., Ndishabandi, D., McCarthy, M., Zhou, Z., et al. (2019). Anti-PD-1/PD-L1 Therapy Augments Lenvatinib's Efficacy by Favorably Altering the Immune Microenvironment of Murine Anaplastic Thyroid Cancer. *Int. J. Cancer* 144 (9), 2266–2278. doi:10.1002/ijc.32041

Guo, F., Zhang, J., Wang, L., Zhao, W., Yu, J., Zheng, S., et al. (2017). Identification of Differentially Expressed Inflammatory Factors in Wilms Tumors and Their Association with Patient Outcomes. *Oncol. Lett.* 14 (1), 687–694. doi:10.3892/ol.2017.6261

Ibrahimpasic, T., Ghossein, R., Shah, J. P., and Ganly, I. (2019). Poorly Differentiated Carcinoma of the Thyroid Gland: Current Status and Future Prospects. *Thyroid* 29 (3), 311–321. doi:10.1089/thy.2018.0509

Jackaman, C., Tomay, F., Duong, L., Razak, N. B. A., Pixley, F. J., Metharom, P., et al. (2017). Aging and Cancer: The Role of Macrophages and Neutrophils. *Ageing Res. Rev.* 36, 105–116. doi:10.1016/j.arr.2017.03.008

Ji, C., Wang, Y., Wang, Y., Luan, J., Yao, L., Wang, Y., et al. (2020). Immune-related Genes Play an Important Role in the Prognosis of Patients with Testicular Germ Cell Tumor. *Ann. Transl Med.* 8 (14), 866. doi:10.21037/atm-20-654

Jiang, H.-C., Chen, X.-R., Sun, H.-F., and Nie, Y.-W. (2020). Tumor Promoting Effects of Glucagon Receptor: a Promising Biomarker of Papillary Thyroid Carcinoma via Regulating EMT and P38/ERK Pathways. *Hum. Cell* 33 (1), 175–184. doi:10.1007/s13577-019-00284-y

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

Kang, K., Xie, F., Mao, J., Bai, Y., and Wang, X. (2020). Significance of Tumor Mutation Burden in Immune Infiltration and Prognosis in Cutaneous Melanoma. *Front. Oncol.* 10, 573141. doi:10.3389/fonc.2020.573141

Krajewska, J., Gawlik, T., and Jarzab, B. (2017). Advances in Small Molecule Therapy for Treating Metastatic Thyroid Cancer. *Expert Opin. Pharmacother.* 18 (11), 1049–1060. doi:10.1080/14656566.2017.1340939

Laha, D., Nilubol, N., and Boufraqech, M. (2020). New Therapies for Advanced Thyroid Cancer. *Front. Endocrinol.* 11, 82. doi:10.3389/fendo.2020.00082

Li, R., Liu, X., Zhou, X. J., Chen, X., Li, J. P., Yin, Y. H., et al. (2020). Identification and Validation of the Prognostic Value of Immune-Related Genes in Non-small Cell Lung Cancer. *Am. J. Transl Res.* 12 (9), 5844–5865.

Lv, X. J., Chen, X., Wang, Y., Yu, S., Pang, L., and Huang, C. (2021). Aberrant Expression of Semaphorin 6B Affects Cell Phenotypes in Thyroid Carcinoma by Activating the Notch Signalling Pathway. *Endokrynol Pol.* 72 (1), 29–36. doi:10.5603/EP.a2020.0072

Mehnert, J. M., Varga, A., Brose, M. S., Aggarwal, R. R., Lin, C.-C., Prawira, A., et al. (2019). Safety and Antitumor Activity of the Anti-PD-1 Antibody Pembrolizumab in Patients with Advanced, PD-L1-Positive Papillary or Follicular Thyroid Cancer. *BMC Cancer* 19 (1), 196. doi:10.1186/s12885-019-5380-3

Murad, H., Collet, P., Huin-Schohn, C., Al-Makdissy, N., Kerjan, G., Chedotal, A., et al. (2006). Effects of PPAR and RXR Ligands in Semaphorin 6B Gene Expression of Human MCF-7 Breast Cancer Cells. *Int. J. Oncol.* 28 (4), 977–984. doi:10.3892/ijo.28.4.977

Nath, M. C., and Erickson, L. A. (2018). Aggressive Variants of Papillary Thyroid Carcinoma: Hobnail, Tall Cell, Columnar, and Solid. *Adv. Anat. Pathol.* 25 (3), 172–179. doi:10.1097/pap.0000000000000184

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

Nikiforov, Y. E., and Nikiforova, M. N. (2011). Molecular Genetics and Diagnosis of Thyroid Cancer. *Nat. Rev. Endocrinol.* 7 (10), 569–580. doi:10.1038/nrendo.2011.142

Oiseth, S. J., and Aziz, M. S. (2017). Cancer Immunotherapy: a Brief Review of the History, Possibilities, and Challenges Ahead. *Jcmt* 3, 250–261. doi:10.20517/2394-4722.2017.41

Samstein, R. M., Lee, C.-H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor Mutational Load Predicts Survival after Immunotherapy across Multiple Cancer Types. *Nat. Genet.* 51 (2), 202–206. doi:10.1038/s41588-018-0312-8

Santa-Maria, C. A., and Nanda, R. (2018). Immune Checkpoint Inhibitor Therapy in Breast Cancer. *J. Natl. Compr. Canc Netw.* 16 (10), 1259–1268. doi:10.6004/jnccn.2018.7046

Schrödter, S., Braun, M., Syring, I., Klümper, N., Deng, M., Schmidt, D., et al. (2016). Identification of the Dopamine Transporter SLC6A3 as a Biomarker for Patients with Renal Cell Carcinoma. *Mol. Cancer* 15, 10. doi:10.1186/s12943-016-0495-5

Sun, X., Liu, X., Xia, M., Shao, Y., and Zhang, X. D. (2019). Multicellular Gene Network Analysis Identifies a Macrophage-Related Gene Signature Predictive of Therapeutic Response and Prognosis of Gliomas. *J. Transl Med.* 17 (1), 159. doi:10.1186/s12967-019-1908-1

Tran Janco, J. M., Lamichhane, P., Karyampudi, L., and Knutson, K. L. (2015). Tumor-infiltrating Dendritic Cells in Cancer Pathogenesis. *J.I.* 194 (7), 2985–2991. doi:10.4049/jimmunol.1403134

Travers, M., Brown, S. M., Dunworth, M., Holbert, C. E., Wiehagen, K. R., Bachman, K. E., et al. (2019). DFMO and 5-Azacytidine Increase M1 Macrophages in the Tumor Microenvironment of Murine Ovarian Cancer. *Cancer Res.* 79 (13), 3445–3454. doi:10.1158/0008-5472.can-18-4018

Varricchi, G., Loffredo, S., Marone, G., Modestino, L., Fallahi, P., Ferrari, S. M., et al. (2019). The Immune Landscape of Thyroid Cancer in the Context of Immune Checkpoint Inhibition. *Ijms* 20 (16), 3934. doi:10.3390/ijms20163934

Yamamoto, Y., Kuroda, K., Sera, T., Sugimoto, A., Kushiyama, S., Nishimura, S., et al. (2019). The Clinicopathological Significance of the CXCR2 Ligands, CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, CXCL7, and CXCL8 in Gastric Cancer. *Anticancer Res.* 39 (12), 6645–6652. doi:10.21873/anticanres.13879

Yao, Y., Yan, Z., Lian, S., Wei, L., Zhou, C., Feng, D., et al. (2020). Prognostic Value of Novel Immune-Related Genomic Biomarkers Identified in Head and Neck Squamous Cell Carcinoma. *J. Immunother. Cancer* 8 (2), e000444. doi:10.1136/jitc-2019-000444

Zhang, C., Li, Z., Qi, F., Hu, X., and Luo, J. (2019a). Exploration of the Relationships between Tumor Mutation burden with Immune Infiltrates in clear Cell Renal Cell Carcinoma. *Ann. Transl. Med.* 7 (22), 648. doi:10.21037/atm.2019.10.84

Zhang, L., Conejo-Garcia, J. R., Katsaros, D., Gimotty, P. A., Massobrio, M., Regnani, G., et al. (2003). Intratumoral T Cells, Recurrence, and Survival in Epithelial Ovarian Cancer. *N. Engl. J. Med.* 348 (3), 203–213. doi:10.1056/NEJMoa020177

Zhang, W., Sun, W., Qin, Y., Wu, C., He, L., Zhang, T., et al. (2019b). Knockdown of KDM1A Suppresses Tumour Migration and Invasion by Epigenetically Regulating the TIMP1/MMP9 Pathway in Papillary Thyroid Cancer. *J. Cell Mol Med* 23 (8), 4933–4944. doi:10.1111/jcmm.14311

Zhou, H., Chen, L., Lei, Y., Li, T., Li, H., and Cheng, X. (2021). Integrated Analysis of Tumor Mutation burden and Immune Infiltrates in Endometrial Cancer. *Curr. Probl. Cancer* 45 (2), 100660. doi:10.1016/j.currproblcancer.2020.100660

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: 23 December 2021
doi: 10.3389/fgene.2021.771853

Comprehensive Analyses of the Expression, Genetic Alteration, Prognosis Significance, and Interaction Networks of m^{6}A Regulators Across Human Cancers

Xiujuan Shi[1,2†], Jieping Zhang[1,2†], Yuxiong Jiang[2], Chen Zhang[2], Xiaoli Luo[1], Jiawen Wu[1] and Jue Li[1,2,3]*

[1]Clinical Research Center for Mental Disorders, Shanghai Pudong New Area Mental Health Center, School of Medicine, Tongji University, Shanghai, China, [2]School of Medicine, Tongji University, Shanghai, China, [3]Shanghai East Hospital, Tongji University School of Medicine, Shanghai, China

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez, Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Qianqian Song, Wake Forest School of Medicine, United States
Xinyue Song, China Medical University, China

***Correspondence:**
Jue Li
jueli@tongji.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: *07 September 2021*
Accepted: *23 November 2021*
Published: *23 December 2021*

Citation:
Shi X, Zhang J, Jiang Y, Zhang C, Luo X, Wu J and Li J (2021) Comprehensive Analyses of the Expression, Genetic Alteration, Prognosis Significance, and Interaction Networks of m^{6}A Regulators Across Human Cancers. Front. Genet. 12:771853. doi: 10.3389/fgene.2021.771853

Accumulating lines of evidence indicate that the deregulation of m^{6}A is involved in various cancer types. The m^{6}A RNA methylation is modulated by m^{6}A methyltransferases, demethylases, and reader proteins. Although the aberrant expression of m^{6}A RNA methylation contributes to the development and progression of multiple cancer types, the roles of m^{6}A regulators across numerous types of cancers remain largely unknown. Here, we comprehensively investigated the expression, genetic alteration, and prognosis significance of 20 commonly studied m^{6}A regulators across diverse cancer types using TCGA datasets *via* bioinformatic analyses. The results revealed that the m^{6}A regulators exhibited widespread dysregulation, genetic alteration, and the modulation of oncogenic pathways across TCGA cancer types. In addition, most of the m^{6}A regulators were closely relevant with significant prognosis in many cancer types. Furthermore, we also constructed the protein–protein interacting network of the 20 m^{6}A regulators, and a more complex interacting regulatory network including m^{6}A regulators and their corresponding interacting factors. Besides, the networks between m^{6}A regulators and their upstream regulators such as miRNAs or transcriptional factors were further constructed in this study. Finally, the possible chemicals targeting each m^{6}A regulator were obtained by bioinformatics analysis and the m^{6}A regulators–potential drugs network was further constructed. Taken together, the comprehensive analyses of m^{6}A regulators might provide novel insights into the m^{6}A regulators' roles across cancer types and shed light on their potential molecular mechanisms as well as help develop new therapy approaches for cancers.

Keywords: m^{6}A methylation, bioinformatics, TCGA, cancer, comprehensive analyses

INTRODUCTION

A variety of biological processes are orchestrated by post-transcriptional modifications including RNA modifications (Zhao et al., 2017). As the most common type of RNA methylation modifications, N6-methyladenosine (m^{6}A), first unraveled in 1970s, modulates the corresponding target RNAs *via* influencing RNA translation, degradation, splicing, folding, or stability (He and He, 2021). Although studies had revealed that one to two m^{6}A residues were found in an average of one thousand nucleotides, nearby the 3′ untranslated region (UTR), stop codon as well as long internal exon might exhibit the relatively richer m^{6}A in mRNAs (Meyer et al., 2012). In addition to mRNAs, m^{6}A RNA methylation was also found to be distributed in other RNAs, such as ribosomal RNA (rRNA) and RNAs of bacteria and viruses (Ma et al., 2019).

The regulators involved in modulating m^{6}A methylation include three types of proteins called "writer", "eraser", and "reader", respectively (Zaccara et al., 2019). The writers consisting of m^{6}A methyltransferases such as METTL3, METTL14, and their corresponding cofactors like RBM15 and WTAP exhibit in cellular nuclei and increase the m^{6}A levels (Meyer and Jaffrey, 2017). On the contrary, the erasers, also being discovered in the cellular nuclei, are m^{6}A demethylase enzymes such as FTO and ALKBH5, which remove the m^{6}A and thus result in reducing the m^{6}A levels (Zhang et al., 2017). Moreover, the readers such as IGF2BP1 and RBMX, distributing in both cellular nuclei and cytoplasm, can decode the m^{6}A methylation information *via* binding to the m^{6}A sites and further initiate the different downstream signals (Sun et al., 2019). The processes of m^{6}A methylation are reversible and dynamic, which are homeostatically modulated by these writers, erasers, and readers (Chen et al., 2019).

Since the m^{6}A regulators acted as crucial roles in a variety of biological processes, the abnormalities of m^{6}A methylation might lead to multiple kinds of diseases including neuronal diseases, diabetes, immunological disorders, liver metabolic disorders, and numerus cancer types (He et al., 2019). For example, recent studies had demonstrated that the decreased RNA methylation of critical genes in β-cell markedly contributed to the pathophysiology of human T2D (De Jesus et al., 2019). Additionally, METTL3 was found to have dramatical overexpression in hepatocellular carcinoma (HCC), and the depletion of METTL3 contributed to the significant suppression of the HCC growth and metastasis (Chen et al., 2018). Besides, findings had uncovered that YTHDF2, an m^{6}A reader, was markedly upregulated in human acute myeloid leukemia (AML), and targeting YTHDF2 might compromise the cancer stem cells in AML (Paris et al., 2019).

Although the m^{6}A methylation has been identified as the most abundant modification of RNAs, and served as crucial regulators in diverse biological processes and diseases including numerous types of cancers, the relevant factors involved in that modification are still not completely discovered, and their associated molecular mechanisms, expression, and interacting networks remain unclear. Therefore, in the present study, we comprehensively investigated the expression, genetic alteration, and prognosis significance of 20 commonly studied m^{6}A regulators across diverse cancer types using TCGA *via* bioinformatic analyses. In addition, we also constructed the networks between m^{6}A regulators and potential chemical drugs, miRNAs, or upstream transcriptional factors. These comprehensive analyses of m^{6}A regulators might provide novel understanding of these m^{6}A regulators' roles across cancer types and shed light on their potential molecular mechanisms in cancers as well as helping developing new therapy approaches for cancers.

MATERIALS AND METHODS

The Gene Expression and Methylation Analyses of m^{6}A Regulators

The expression of the gene set (the 20 m^{6}A methylation regulators) and m^{6}A regulators' interacting proteins across diverse cancer types based on TCGA data was analyzed using the GSCALite database (Liu et al., 2018). Besides, we also analyzed the expression of 20 m^{6}A methylation regulators through R software package using microarray data (GSE11969, GSE63898, GSE37182, GSE22820, GSE54129, GSE53757, GSE23036, GSE33630, and GSE11024) from the Gene Expression Omnibus (GEO) datasets. The heatmaps of these GEO data were displayed by the R software package pheatmap. The expression of IGF2BP1, IGF2BP2, IGF2BP3, SP1, ELK1, and EGR1 across diverse TCGA cancer types was analyzed using the UALCAN database (Chandrashekar et al., 2017). In addition, the methylation of the gene set, and the correlation between the methylation and m^{6}A methylation regulators' gene expression were also analyzed using GSCALite database.

The Genetic Alteration Analyses of the m^{6}A Regulators

The single nucleotide variations (SNVs) and copy number variations (CNVs) of the m^{6}A regulators across cancer types were analyzed by the GSCALite database using TCGA data. The SNV-oncoplot and CNV-percent-profile (CNV pie plots) were also generated by GSCALite database. In addition, the bubble plots describing the correlation between CNV and m^{6}A methylation regulators' mRNA expression were generated by GSCALite database based on TCGA data. Besides, the genetic alterations of the 20 m^{6}A regulators were also analyzed by cBioportal database (Cerami et al., 2012).

The Oncogenic Pathway Analyses of the m^{6}A Regulators and Protein–Protein Interaction Network Construction

The m^{6}A methylation regulators-related oncogenic pathways were analyzed by GSCALite database. The pathway activity pie plots, and the interaction map of genes and pathways were also generated using the GSCALite database. The protein–protein interaction (PPI) networks were generated by STRING database (Szklarczyk et al., 2019).

The Overall Survivals Analyses

The overall survivals of the m^{6}A methylation regulators across cancer types were analyzed by the GSCALite database. The overall survivals

of the m^{6}A regulators in kidney renal clear cell carcinoma (KIRC) were evaluated by the GEPIA database based on TCGA data.

The m^{6}A Regulators–Drug Interacting Network Construction

The potential chemicals targeting each m^{6}A regulator were obtained by applying the comparative toxicogenomics database (CTD) (Davis et al., 2021). Thereafter, the chemicals and their corresponding m^{6}A regulator were inputted into Cytoscape software (Shannon et al., 2003) to generate the m^{6}A regulators–drug interacting network.

The Generation of the MiRNAs–m^{6}A Regulators Network

First, the potential miRNAs targeting each m^{6}A regulator were predicted by miRDB (Chen and Wang, 2020), targetScan (Agarwal et al., 2015), and starbase (Li et al., 2014). Subsequently, the overlapping miRNAs (commonly expressed in the prediction of miRDB, targetScan, and starbase) were obtained using the VENNY 2.1 database (https://bioinfogp.cnb.csic.es/tools/venny/index.html). Then, the Cytoscape software was utilized for generating the miRNAs–m^{6}A regulators networks.

Constructing the Transcription Factors–m^{6}A Regulators Network

The potential TFs targeting each group of m^{6}A regulators (writers, erases, and readers) were obtained by KnockTF database (Feng et al., 2020). The erasers–TFs interacting network, writers–TFs interacting network, and readers–TFs interacting network were also generated by KnockTF database. In addition, the top 25 potential TFs were achieved exhibited by doughnut plots using FunRich software (Pathan et al., 2015).

The Interacting Networks Construction and Gene Ontology and Biological Pathway Analyses

The m^{6}A regulators' interacting proteins were obtained by FunRich software, and the interacting regulatory network including m^{6}A regulators and their corresponding interacting proteins was then constructed by FunRich software. The FunRich software was also utilized for investigating the GO analyses and biological pathways of the m^{6}A regulators interacting proteins. The column diagrams (exhibiting CC: cellular component; MF: molecular function; BP: biological process) and doughnut plots (exhibiting biological pathways) were also generated using FunRich software.

RESULTS

The Expression of m^{6}A Methylation Regulators Across Cancer Types

The m^{6}A methylation regulators could be clarified into three types: writers, erasers, and readers (**Figure 1A**). The reports relevant with m^{6}A methylation regulators in recent years were reviewed and a total of 20 genes (writers: 7; erasers: 2; readers: 11) were included in this study (**Figure 1B**). Next, we sought to evaluate the expressing levels of these 20 m^{6}A methylation regulators across diverse cancer types using TCGA datasets. By searching the GSCALite database, we found that the expressions of many m^{6}A methylation regulators (especially IGF2BP1, IGF2BP2, and IGF2BP3) were changed across multiple cancer types (**Figure 1C**). Furthermore, the detail expressing situations of IGF2BP1, IGF2BP2, and IGF2BP3 were evaluated by applying UALCAN algorithm. The results revealed that the levels of IGF2BP1, particularly IGF2BP2 and IGF2BP3, were remarkably upregulated in many cancer types such as bladder urothelial carcinoma (BLCA), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), and lung squamous cell carcinoma (LUSC) (**Figure 1D**). In addition, to further verify the above results using TCGA data, we also analyzed the expression of 20 m^{6}A regulators using microarray data from GEO datasets in many cancer types including LUAD, LUSC, LIHC, colon adenocarcinoma (COAD), breast invasive carcinoma (BRCA), stomach adenocarcinoma (STAD), kidney renal clear cell carcinoma (KIRC), head and neck squamous cell carcinoma (HNSC), thyroid carcinoma (THCA), and kidney chromophobe (KICH). The results revealed that the expression of many m^{6}A regulators including IGF2BP1, IGF2BP2, and IGF2BP3 were notably changed, which was consistent with the results from TCGA analysis (**Supplementary Figure S1**). Given that the methylation of genes was able to influence genes expression, we next attempted to investigate the methylation of the 20 m^{6}A methylation regulators across cancer types. Through searching GSCALite database, we found that the majority of the 20 m^{6}A regulators' methylation was lower in the tumor samples than that of the normal control samples in prostate adenocarcinoma (PRAD), LIHC, LUSC, KIRC, BLCA, and THCA (**Figure 1E**). Besides, the correlation between the methylation and m^{6}A methylation regulators' gene expression was further analyzed using the GSCALite database. According to the data, we found that most of these m^{6}A regulators' expression was negatively correlated with the methylation across diverse cancer types, which was consistent with the above finding that the majority of the 20 m^{6}A regulators' methylation was downregulated in the tumor samples of many cancer types (**Figure 1F**). Overall, our data suggested that the expression and methylation of the 20 m^{6}A regulators were remarkable dysregulation across many cancer types.

Genetic Alterations of m^{6}A Regulators Across Cancer Types

Considering that the alteration of the genome might always affect the gene expression, we next attempted to explore the genetic alterations including single nucleotide variations (SNVs) and copy number variations (CNVs) of the m^{6}A regulators across cancer types. First, the GSCALite database was utilized for analyzing the SNVs of the 20 m^{6}A regulators. The results suggested that the SNVs of the 20 m^{6}A regulators altered in 74.8% TCGA samples across cancer types, and the waterfall plots

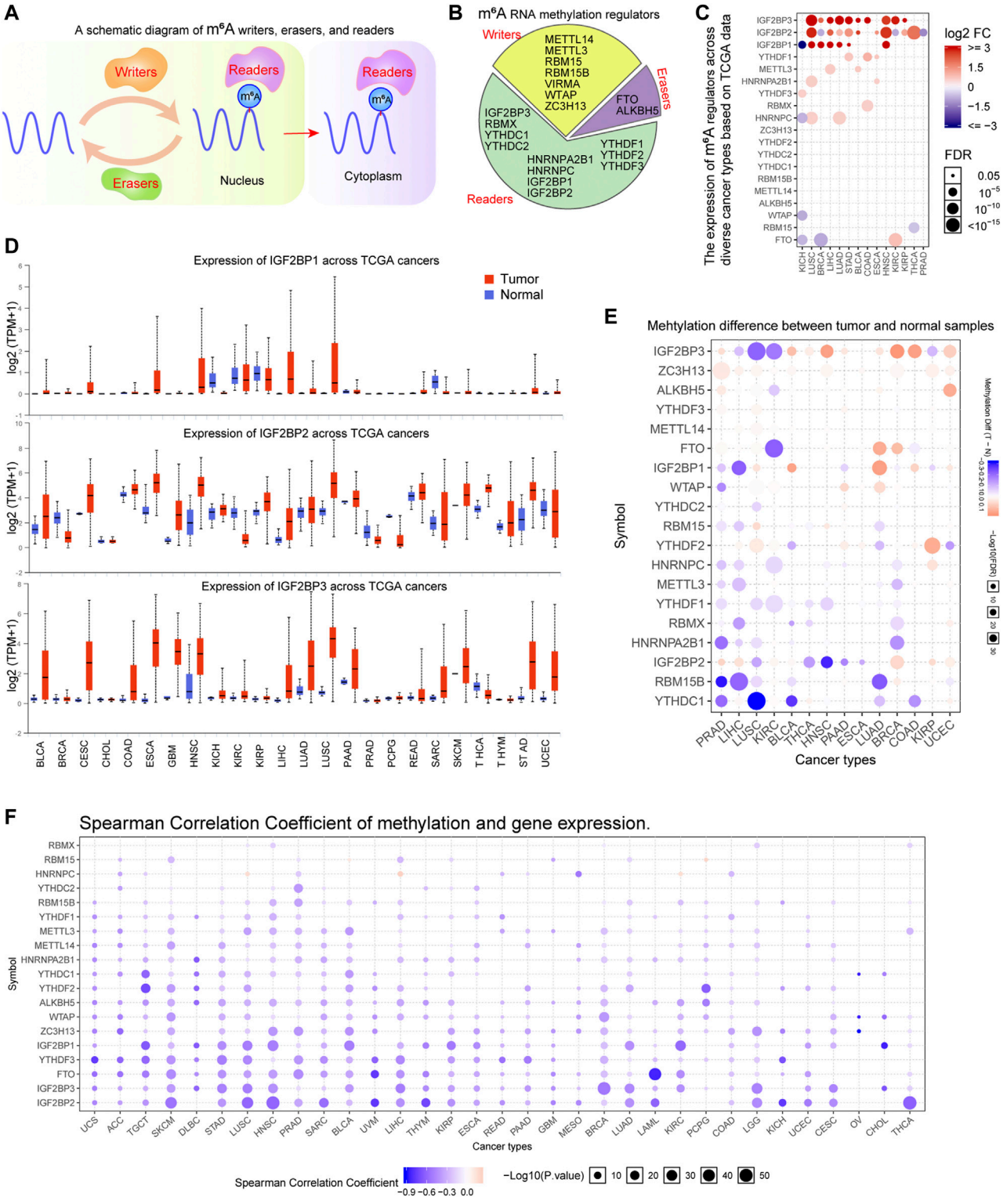


FIGURE 1 | The expression analyses of m^{6}A regulators in TCGA cancer types. **(A)** A schematic diagram of m^{6}A regulators in cells. **(B)** The m^{6}A regulators are divided into writers, erasers, and readers. **(C)** The GSCALite database revealed the expressing levels of the 20 m^{6}A regulators across diverse cancer types. **(D)** Details expressing situations of IGF2BP1, IGF2BP2, and IGF2BP3 were evaluated by applying the UALCAN algorithm. **(E)** The methylation difference analyzed by GSCALite between tumor and normal samples of m^{6}A regulators across TCGA cancer types. **(F)** The correlation between the methylation and m^{6}A methylation regulators' gene expression were analyzed using GSCALite database.

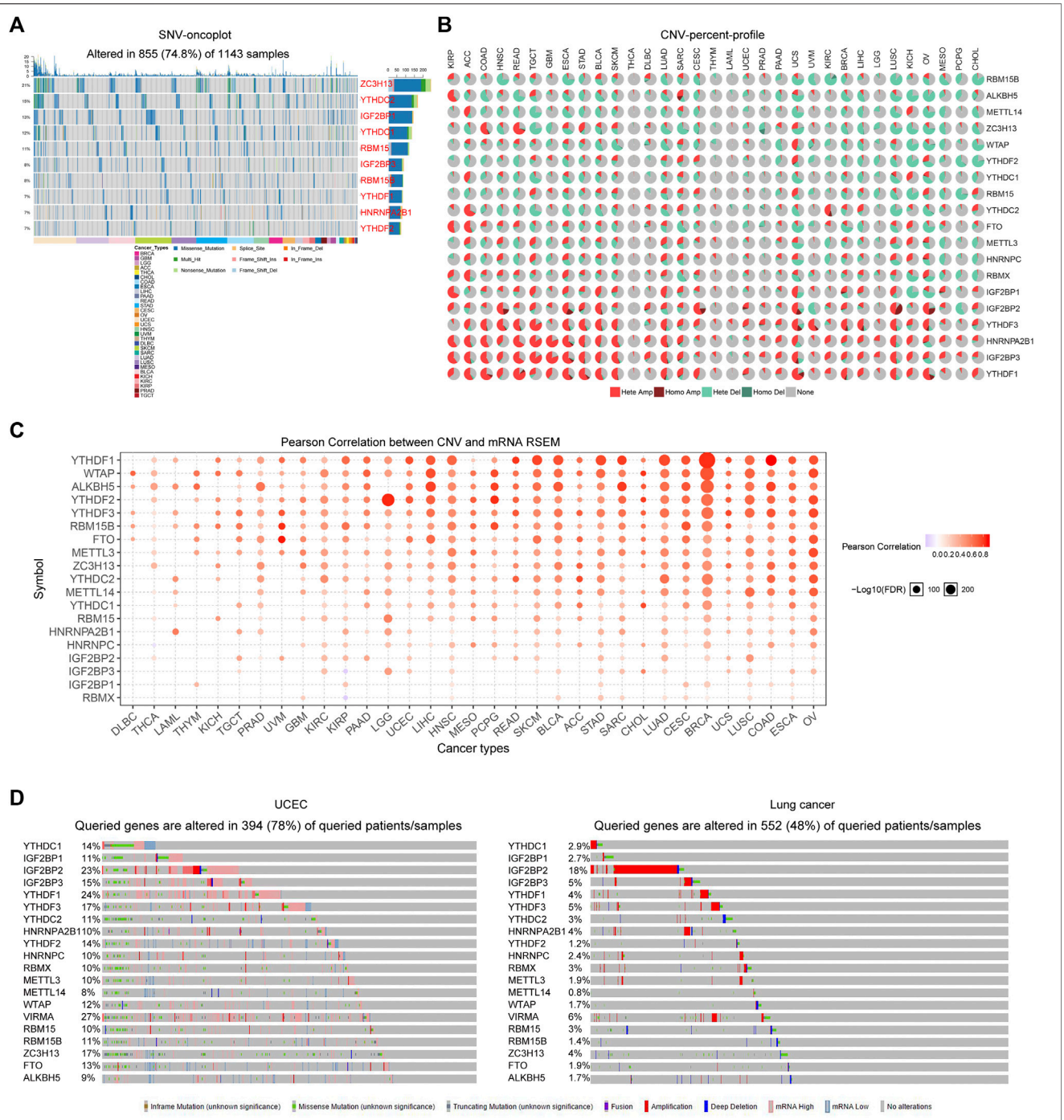


FIGURE 2 | The genetic alterations of m6A regulators. **(A)** The SNVs of the m^{6}A regulators in TCGA samples across cancer types. The waterfall plots presented the top ten SNVs-changed genes. **(B)** The CNV pie plots revealed the CNVs alteration frequency of the 20 m^{6}A regulators across diverse cancer types. **(C)** The bubble plots from GSCALite database showed that the mRNA expression of the m^{6}A regulators was positively correlated with their corresponding CNVs across most cancer types. **(D)** The genetic alterations of the 20 m^{6}A regulators in UCEC and lung cancer using cBioportal analyses.

presented the top ten SNVs-changed genes, such ZC3H13, YTHDC2, and IGF2BP1 (**Figure 2A**). Thereafter, we sought to investigate the CNVs alteration frequency for the 20 m^{6}A regulators using the GSCALite database. The CNV pie plots revealed that several readers (YTHDF1, YTHDF3, IGF2BP1, IGF2BP2, IGF2BP3, and HNRNPA2B1) exhibited very high percentages of heterozygous CNVs, particularly amplification (Hete Amp) in multiple cancer types, while genes like RBM15B, ALKBH5, METTL14, ZC3H13, and WTAP had high percentages of heterozygous CNVs with depletion (Hete Del)

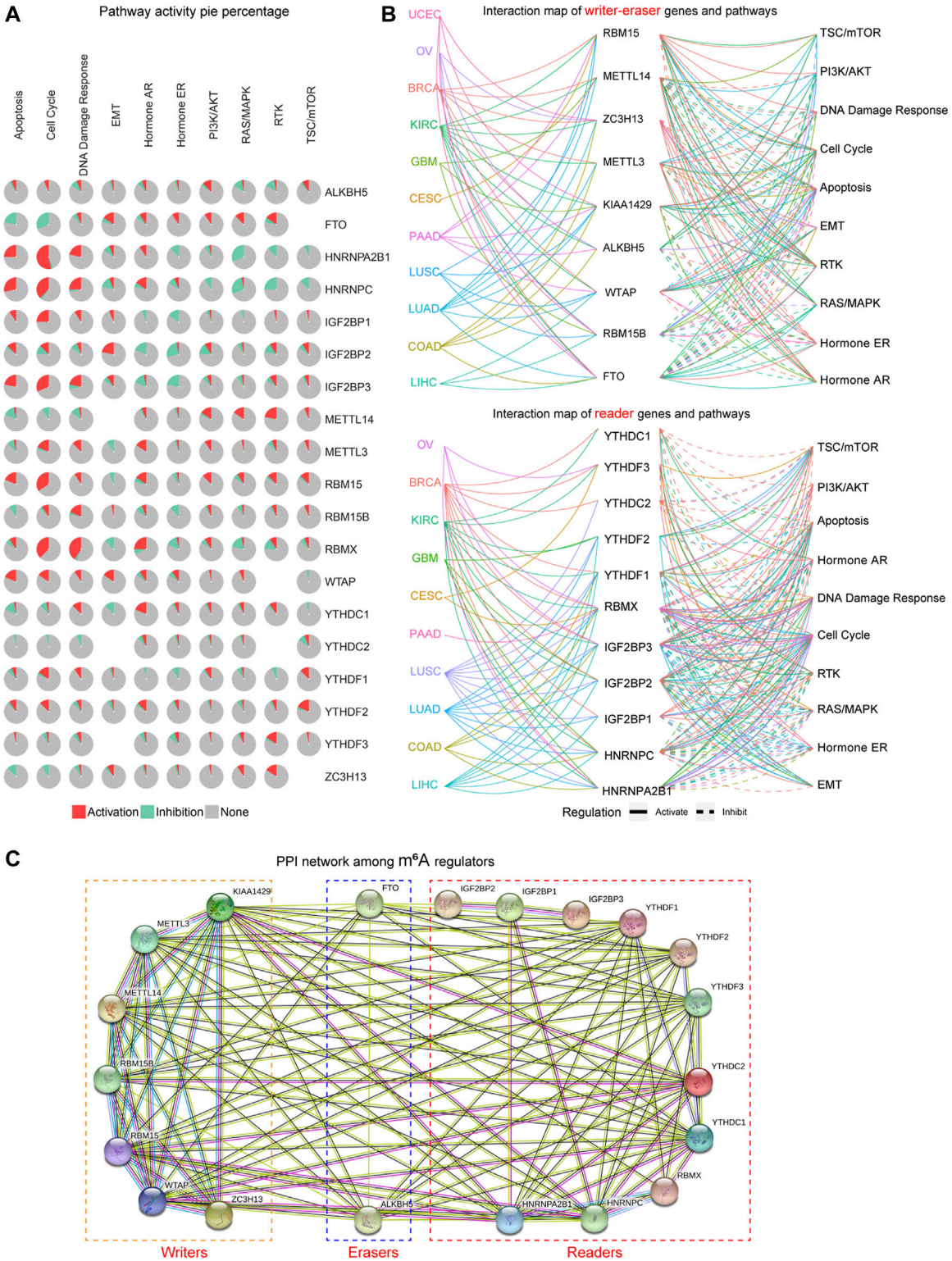


FIGURE 3 | The oncogenic pathways related to the m^{6}A regulators. **(A)** The pathway pie plots analysis of m^{6}A regulators from the GSCALite database. **(B)** The interaction map of m^{6}A regulators and pathways in numerous cancer types using GSCALite database. **(C)** The protein–protein interaction (PPI) network of the 20 m^{6}A regulators was constructed by STRING database.

(**Figure 2B**). Afterwards, we attempted to explore whether these m^{6}A regulators' CNVs alterations were able to influence the expression of their mRNA expression. The bubble plots from GSCALite database demonstrated that the mRNA expression of the majority of the m^{6}A regulators was positively correlated with their corresponding CNVs across most cancer types, which indicated that CNVs alterations could remarkably promote m^{6}A regulators' expression (**Figure 2C**). In addition, the genetic alterations of the 20 m^{6}A regulators in TCGA cancer types were also investigated using cBioportal database. The data suggested that the genetic alterations of the m^{6}A regulators were remarkably high in tumor specimens of many cancer types, particularly uterine corpus endometrial carcinoma (UCEC; with 78% genetic alterations) and lung cancer (with 48% genetic alterations) (**Figure 2D**). Taken together, these results revealed that the m^{6}A regulators exhibited widespread genetic alterations across cancer types, and these genetic alterations could significantly affect their expression.

The Analyses of the Oncogenic Pathways Relevant With the m^{6}A Regulators

Next, we attempted to investigate whether these m^{6}A regulators were associated with oncogenic pathways. According to the results of the pathway pie plots from GSCALite database, we found that HNRNPA2B1, HNRNPC, IGF2BP1, IGF2BP3, RBM15, and RBMX were markedly related with the activation of the cell cycle (**Figure 3A**). In addition, the pathway pie plots also showed that FTO was relevant with the inhibition of apoptosis and cell cycle; HNRNPA2B1 and HNRNPC were significantly correlated with the inhibition of the RAS/MAPK pathway; RBMX was also related with the activation of DNA damage response pathway (**Figure 3A**). Besides, the 20 m^{6}A regulators were divided into two groups (writer–eraser genes and reader genes) to respectively construct the interaction map of genes and pathways using the GSCALite database, and the results further confirmed the above findings that many m^{6}A regulators were associated with the activation or inhibition of these famous cancer-related pathways across TCGA cancer types (**Figure 3B**). Furthermore, considering that genes always exerted their functions *via* interacting with other genes, we thereby next sought to investigate the interaction among these writers, erasers, and readers. The protein–protein interaction (PPI) network of the 20 m^{6}A regulators was constructed by the STRING database, and the PPI network demonstrated that the m^{6}A regulators interacted with each other with very high frequency, which indicated that the m^{6}A methylation in cancers might be regulated by collaboration among writers, erasers, and readers (**Figure 3C**). Collectively, these data validated that the m^{6}A regulators could modulate the oncogenic pathways *via* collaboration.

Prognosis Significance of the m^{6}A Regulators Across Cancer Types

Since the m^{6}A regulators were dysregulated in many cancer types and several of them were closely relevant with oncogenic pathways, we next sought to explore whether the aberrant expression of the m^{6}A regulators was associated with prognosis significance. After inputting the gene set of the m^{6}A regulators into the GSCALite database, we found that most of the m^{6}A regulators were associated with overall survivals across TCGA cancer types (**Figure 4A**). Particularly, more than half of the 20 m^{6}A regulators were notably correlated with poor or good prognosis in multiple cancer types, such as KIRC, brain lower grade glioma (LGG), adrenocortical carcinoma (ACC), breast invasive carcinoma (BRCA), and sarcoma (SARC). Therefore, we next attempted to investigate the detailed overall survivals of the m^{6}A regulators (high or low expression) in KIRC. The overall survivals of the 20 m^{6}A regulators in KIRC were analyzed by applying the GEPIA database. The results validated that 19 of the 20 m^{6}A regulators were dramatically correlated with significantly good or poor prognosis (**Figure 4B**). Especially, high expression of all the erasers (FTO and ALKBH5) and most writers (METTL14, RBM15, RBM15B, WTAP, and ZC3H13) and readers (YTHDF1, YTHDF2, YTHDF3, HNRNPA2B1, HNRNPC, YTHDC1, YTHDC2, and RBMX) was significantly with poor prognosis, while the high expression of VIRMA (also named KIAA1429; a writer) and IGF2BPs (IGF2BP1, IGF2BP2, and IGF2BP3; readers) predicted good prognosis (**Figure 4B**). Therefore, these data revealed that the dysregulation of the m^{6}A regulators was remarkably associated with significant prognosis in many cancer types, especially KIRC, which indicated that the aberrant expression of the m^{6}A regulators might be a prognostic marker in cancers including KIRC.

The Construction of the m^{6}A Regulators–Potential Drugs Network

Since the above findings revealed that the dysregulation of the m^{6}A regulators might be correlated with tumor progression, we next thought to investigate whether there were some potential chemicals that could increase or decrease the expression of the m^{6}A regulators. First, the comparative toxicogenomics database (CTD) was utilized for analyzing the possible chemicals targeting each m^{6}A regulator. Afterwards, the 20 m^{6}A regulators were divided into three groups (writers, erasers, and readers) and we subsequently drew three sub gene–drug interaction networks (writers–drugs interaction network, erasers–drugs interaction network, and readers–drugs interaction network), using Cytoscape software. The m^{6}A regulators–potential drugs network is presented in **Figure 5**, and these chemicals were able to increase or decrease the expression of the m^{6}A regulators (**Supplementary Table S1**). Therefore, the m^{6}A regulators–potential drugs network provided benefits for potential drugs discovery to target specific m^{6}A regulators.

The Upstream MiRNAs–m^{6}A Regulators Network

Although our above findings demonstrated that the methylation and genetic alterations were capable to affect the expression of the m^{6}A regulators across TCGA cancer types, there might be other factors such as miRNAs that could also contribute to the dysregulation of the m^{6}A regulators in cancers. Therefore, we next sought to investigate the potential upstream miRNAs, which

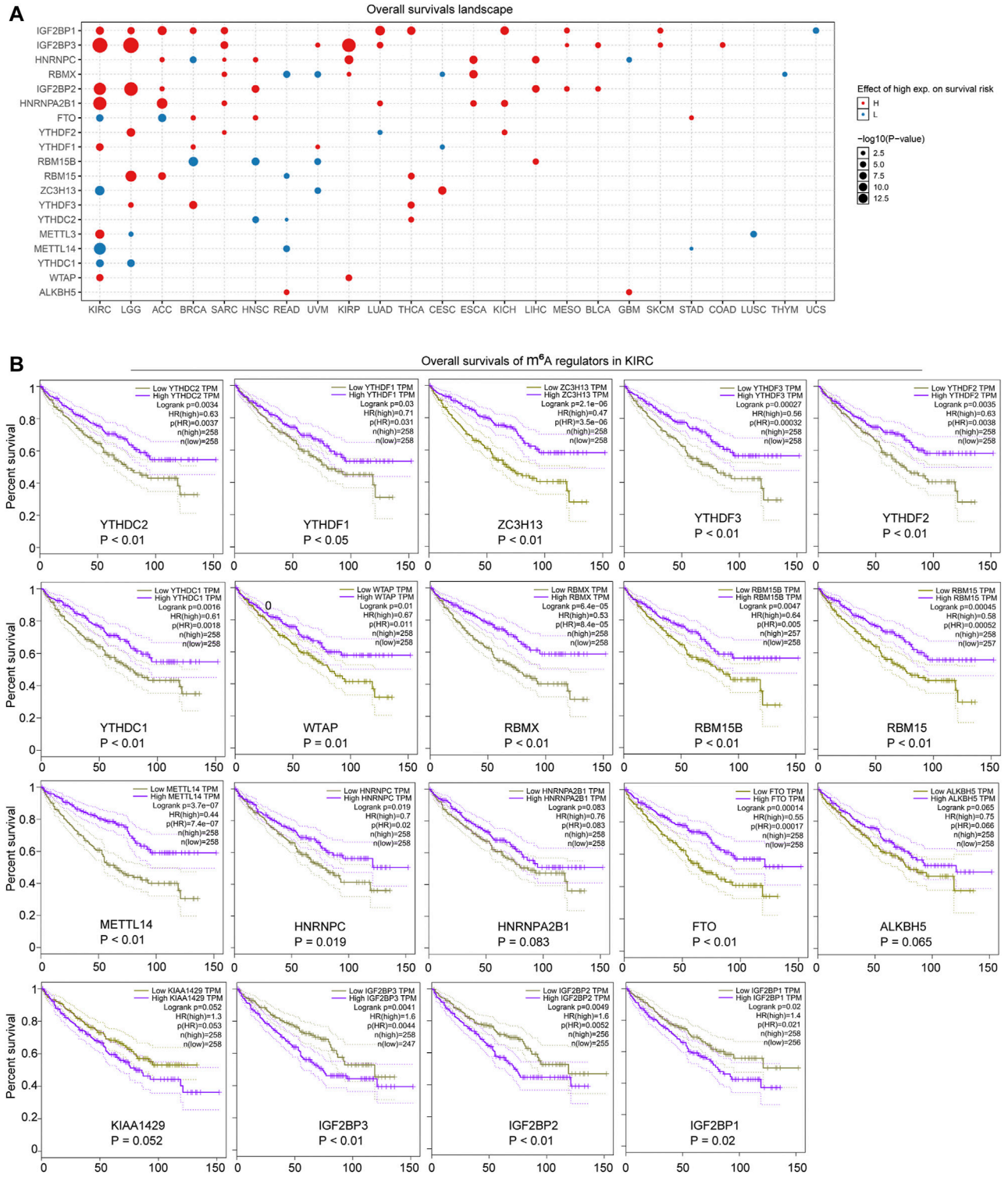


FIGURE 4 | Prognosis significance of the m^{6}A regulators across cancer types. **(A)** Overall survivals landscape of the m^{6}A regulators across cancer types was generated by GSCALite database. **(B)** GEPIA analyzed the overall survivals of the m^{6}A regulators in KIRC.

were able to target these m^{6}A regulators. First, we employed three classical miRNAs predicting databases—miRDB, targetScan, and starbase—to predict the possible miRNAs that could target each m^{6}A regulator. Thereafter, the common miRNAs targeting each m^{6}A regulator in the three databases were selected. Subsequently, we applied the Cytoscape software to generate the miRNAs–m^{6}A

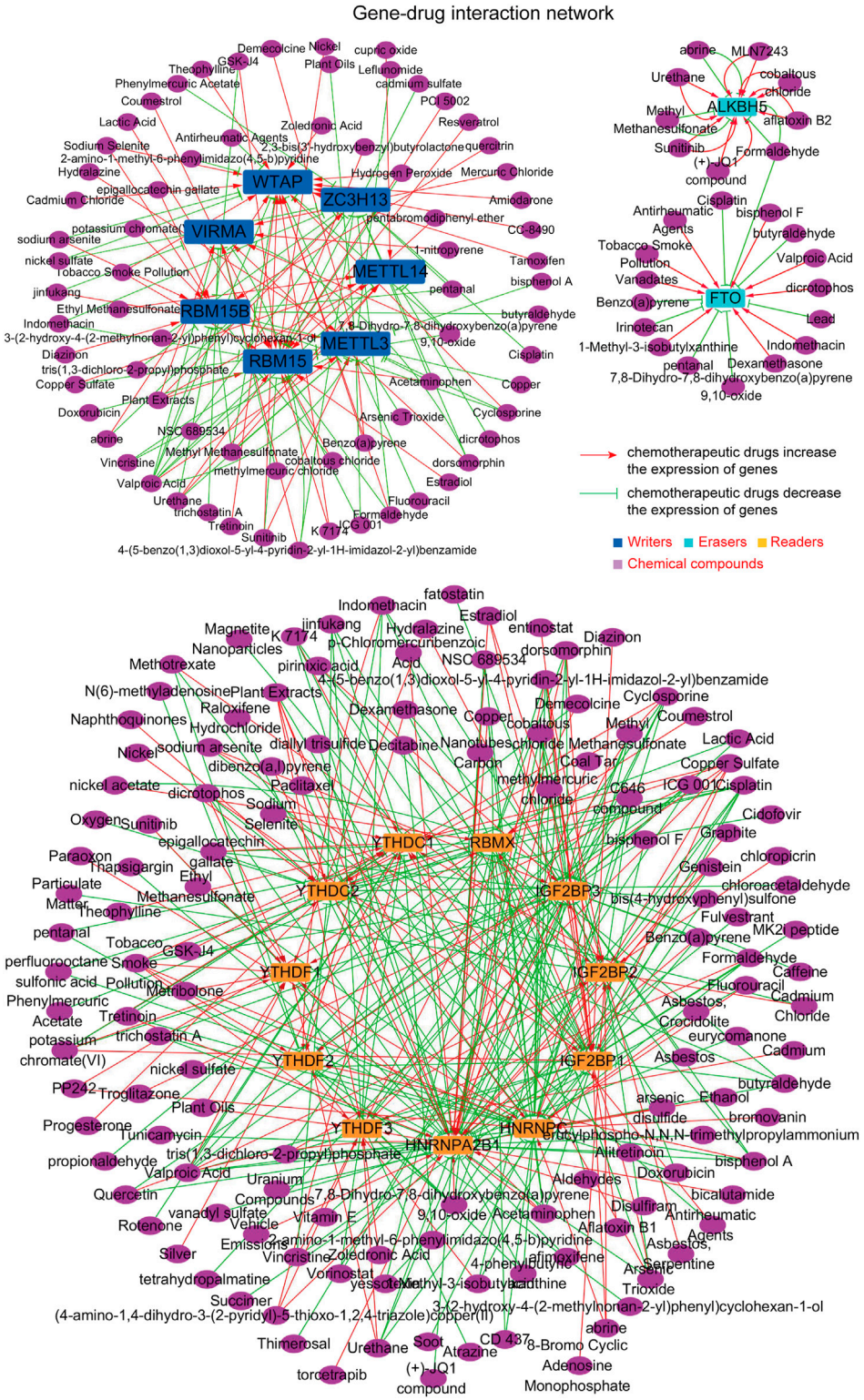


FIGURE 5 | The construction of the m^{6}A regulators–potential drugs network. The potential chemicals targeting m^{6}A regulators were obtained by the comparative toxicogenomics database (CTD), and the network was generated by Cytoscape software.

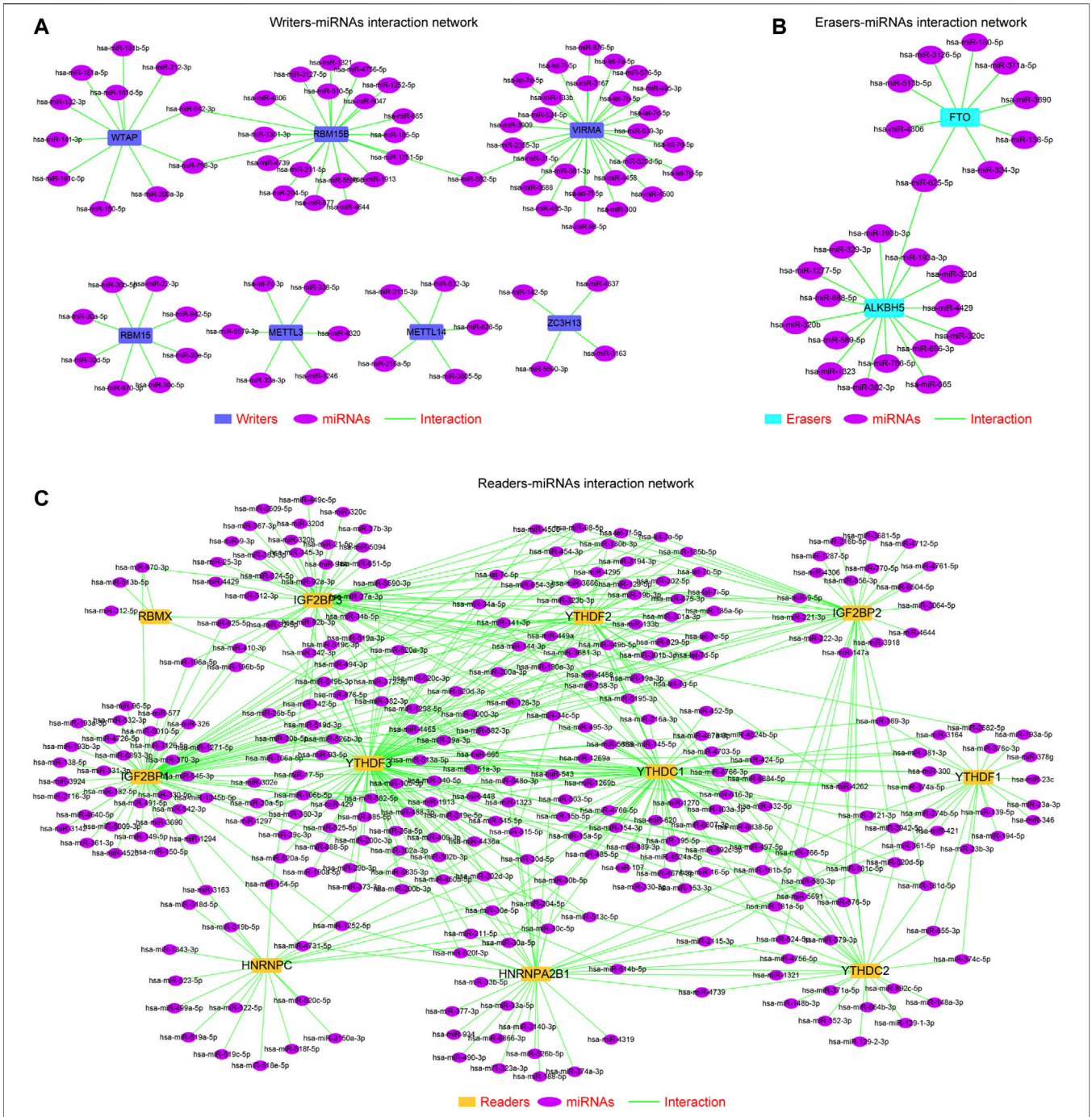


FIGURE 6 | MiRNAs–m^{6}A regulators network construction. **(A)** Writers–miRNAs interaction network. **(B)** Erasers–miRNAs interaction network. **(C)** Readers–miRNAs interaction network.

regulators networks, including the writers–miRNAs interaction network (**Figure 6A**), erasers–miRNAs interaction network (**Figure 6B**), and readers-miRNAs interaction network (**Figure 6C**). These miRNAs–m^{6}A regulators networks provided new supplementary knowledge about the modulation of the m^{6}A regulators' dysregulation across cancer types.

The Upstream TFs–m^{6}A Regulators Networks

Besides, the transcription factors (TFs) could also contribute to the dysregulation of m^{6}A regulators. Hence, we next sought to uncover the potential TFs that were capable to modulate the expression of the m^{6}A regulators. To achieve that, we first utilized an online database,

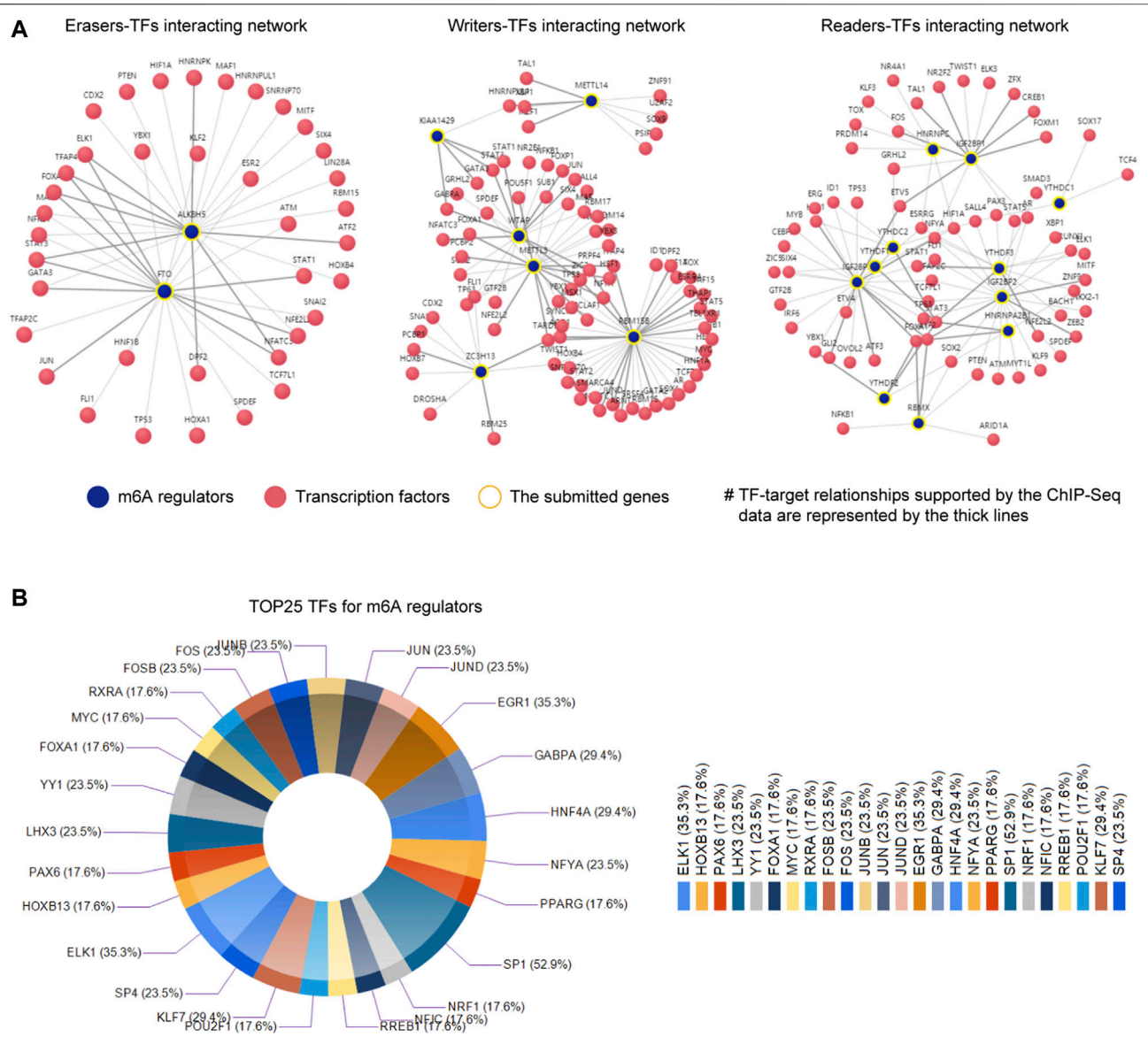


FIGURE 7 | The construction of TFs–m^{6}A regulators networks. **(A)** The TFs–m^{6}A regulators networks (including three sub-networks: erasers–TFs interacting network, writers–TFs interacting network, and readers–TFs interacting network) were generated by KnockTF software. **(B)** The top 25 potential TFs (ranked by targeting percentages) were achieved and exhibited by doughnut plots using FunRich software.

KnockTF, to analyze the possible TFs of the writers, erases, and readers, respectively. We generated the TFs–m^{6}A regulators networks (including three sub-networks: erasers–TFs interacting network, writers–TFs interacting network, and readers–TFs interacting network), and they are presented in **Figure 7A**. The TFs–m^{6}A regulators networks indicated that a plethora of TFs might regulate the m^{6}A regulators. For example, there were more than 20 TFs that were possibly able to target the promoter of RBM15B. In addition, another software, FunRich, was also utilized for calculating the potential TFs targeting the m^{6}A regulators. The top 25 potential TFs (ranked by targeting percentages) were achieved and exhibited by doughnut plots (**Figure 7B**). The data demonstrated that SP1 were possibly able to modulate more than half of the 20 m^{6}A regulators (52.9%), ELK1 regulated 35.3% of the 20 m^{6}A regulators, and EGR1 also orchestrated 35.3% of the 20 m^{6}A regulators. Indeed, analyses from the UALCAN database revealed that the expressions of SP1, ELK1, and EGR1 were remarkably aberrant in many TCGA cancer types (**Supplementary Figure S2**). Collectively, these data provided novel insights into the possible molecular mechanisms of the m^{6}A regulators' dysregulation in TCGA cancer types.

The Interacting Regulatory Network of Each m^{6}A Regulators

Although we had generated the PPI network of the 20 m^{6}A regulators in the above findings, which demonstrated that

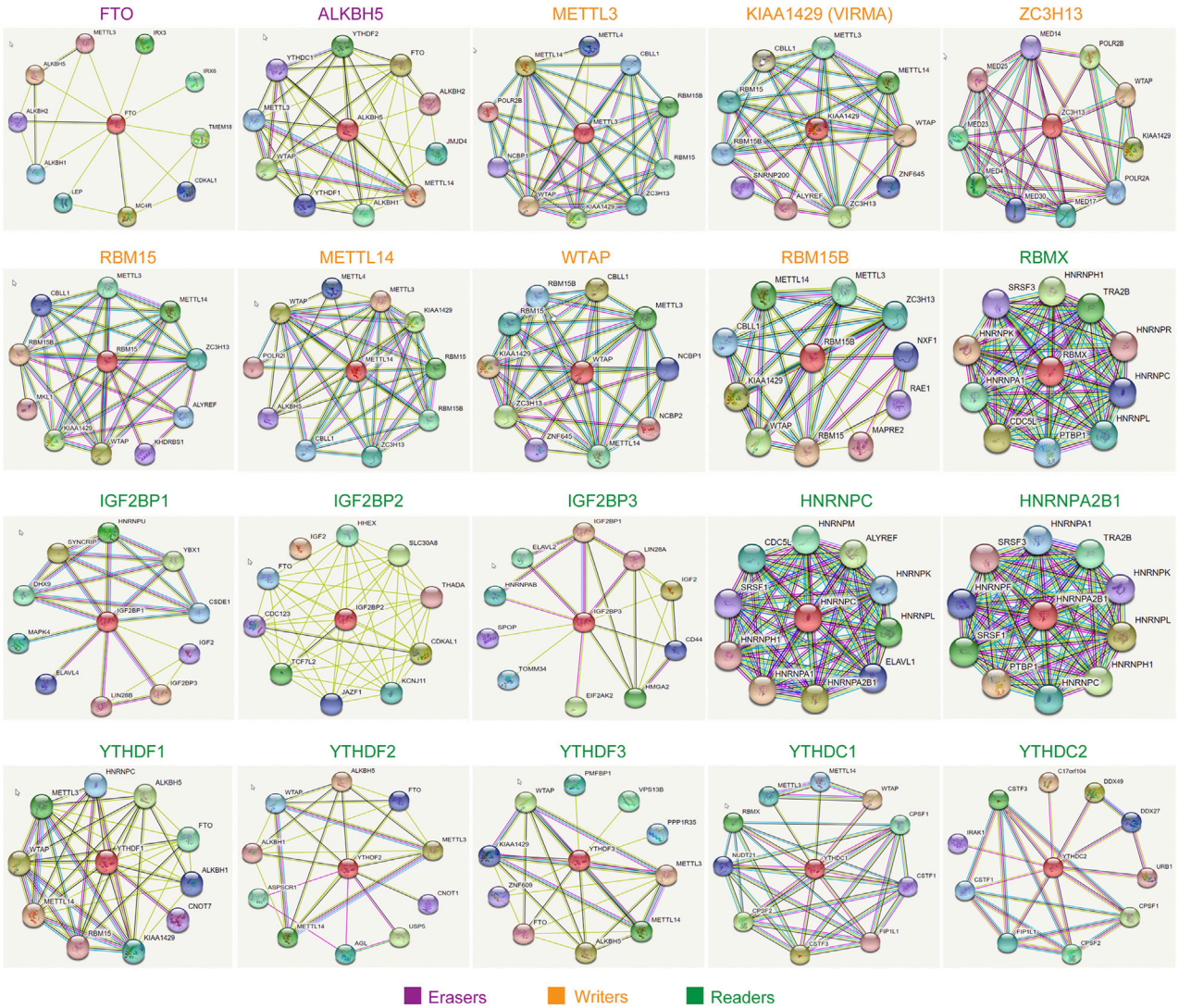


FIGURE 8 | The interacting regulatory network of each m^{6}A regulator. The interacting proteins of each eraser, writer, and reader were obtained by using STRING database.

these 20 m^{6}A regulators interacted with each other with very high frequency, the m^{6}A regulators exerted their functions and might also collaborate with other factors. Therefore, we next attempted to construct the interacting regulatory network of each m^{6}A regulator alone. The interacting proteins of each eraser, writer, and reader were obtained by using the STRING database, and their corresponding interacting networks are presented in **Figure 8**. The results indicated that every m^{6}A regulator had a complex interacting regulatory network, and these interacting networks might provide new insights into how the m^{6}A regulators exerted their modulatory functions.

Furthermore, a more complex interacting regulatory network including m^{6}A regulators and their corresponding interacting factors was constructed by applying FunRich software. This network not only validated that m^{6}A regulators interacted with other relevant factors, but also revealed that the erasers, writers, and readers interacted with each other frequently (**Figure 9A**). In addition, we then attempted to explore the expressions of the m^{6}A regulators-related genes in that network using GSCALite database, because these genes might be the downstream targets of the m^{6}A regulators or even the modulators of the m^{6}A regulators, and their dysregulation should be critical in m^{6}A regulation. According to the data, numerous genes (such as H2AFX, CDKN2A, TTF2, IKBKE, and UBE2I) were markedly upregulated in many cancer types, while some genes (such as ARRB1, LMO3, KHDRBS2, CIRBP, and RALYL) were remarkably downregulated in multiple cancer types (**Figure 9B**). Therefore, these data suggested that the majority of the m^{6}A regulators-related genes were dysregulated across many cancer types.

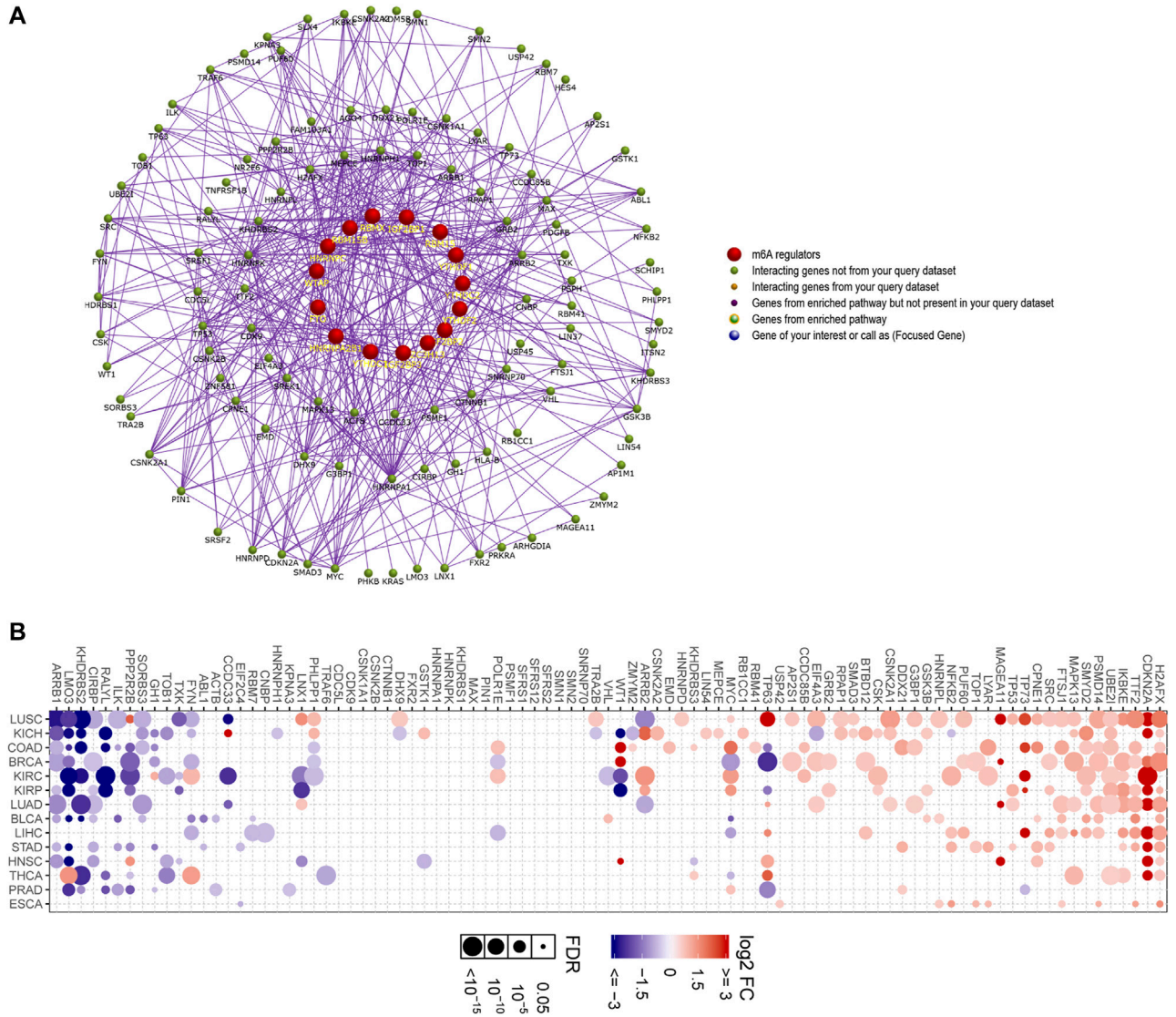


FIGURE 9 | Constructing the interacting regulatory network of m^{6}A regulators and their corresponding interacting factors. **(A)** The m^{6}A regulators and their corresponding interacting factors were constructed into a complex network by applying FunRich software. **(B)** The expressions of the m^{6}A regulators-related genes in diverse cancer types were analyzed using GSCALite database.

Next, we aimed to investigate the gene ontology analyses (GO analyses; including three sub-analyses: CC: cellular component; MF: molecular function; BP: biological process) and biological pathways of the m^{6}A regulators-related genes. To achieve that, the FunRich software was utilized. The gene ontology analyses revealed that the m^{6}A regulators-related genes were significantly associated with cytoplasm (CC), nucleus (CC), RNA binding (MF), regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism (BP) (**Figure 10A**). Moreover, the top 20 biological pathways for m^{6}A regulators-related genes (ranked by percentages) were exhibited by doughnut plots and the data showed that these genes were dramatically correlated with many tumorigenesis-relevant pathways, such as TRAIL signaling pathway, S1P pathway, and mTOR signaling (**Figure 10B**). Therefore, these results indicated that m^{6}A regulators as well as their relevant genes were potentially associated with cancers.

DISCUSSION

Cancers account for the major public health problems, and it leads to the second cause of death, ranking behind cardiovascular diseases, in most countries (Siegel et al., 2019). Therefore, seeking novel approaches for cancer therapy is urgent. Based on deeply understanding the molecular mechanisms, several effectively new methods for treating cancers such as cellular immuno-therapy and PD1/PDL1 antibodies therapy had emerged currently (Sharpe and Pauken, 2018). Moreover, epigenetics including

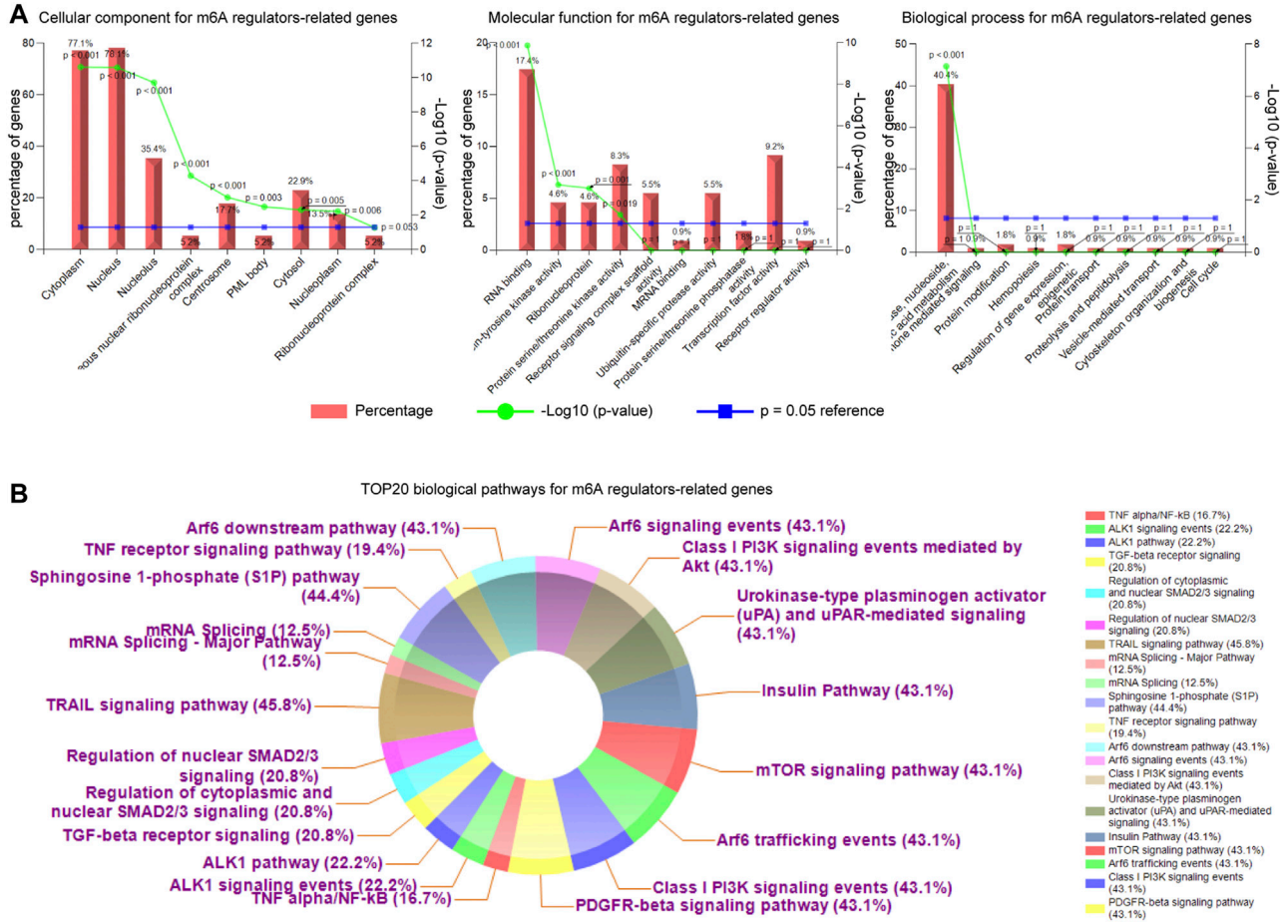


FIGURE 10 | The analyses of gene ontology and biological pathways of the m^{6}A regulators-related genes. **(A)** The cellular component (CC), molecular function (MF), and biological process (BP) analyses for the m^{6}A regulators-related genes. **(B)** The top 20 biological pathways for m^{6}A regulators-related genes (ranked by percentages) were exhibited by doughnut plots using FunRich software.

m^{6}A methylation, as a popular field of cancer research, might also emerge as a new approach for cancer treatment if its detailed molecular mechanisms in tumorigenesis were deeply unraveled (Zhao et al., 2020). Hence, it is necessary to investigate the effects of m^{6}A RNA methylation regulators on multiple cancer types. In the present study, we applied bioinformatics analyses to explore the expression, genetic alterations, prognosis significance, the networks between m^{6}A regulators, and potential chemical drugs, miRNAs, or upstream transcriptional factors in multiple cancer types, which deeply uncovered the critical roles and molecular mechanisms of m^{6}A regulators in cancers.

Emerging lines of evidence had indicated that the m^{6}A regulators served as critical roles in regulating numerous biological processes, diseases, and especially tumor development. For example, m^{6}A regulator YTHDF1 was recently identified as a novel prognostic marker and potential target for HCC (Bian et al., 2020). Moreover, m^{6}A regulator HNRNPA2B1 was found to function as an oncogenic factor to accelerate esophageal cancer (ESCA) progression, and it might be a promising prognostic biomarker for ESCA (Guo et al., 2020). In addition, the low expression of METTL3, an important m^{6}A writer, was found to be correlated with the poor prognosis of triple-negative breast cancer (TNBC), and METTL3 might serve as a novel therapeutic target in TNBC metastasis (Shi et al., 2020). In the present study, our bioinformatics analysis also revealed that YTHDF1, HNRNPA2B1, and METTL3 were highly expressed in colon adenocarcinoma (COAD), lung squamous cell carcinoma (LUSC), and liver hepatocellular carcinoma (LIHC), respectively, and the expression of many other m^{6}A regulators such as IGF2BP1, IGF2BP2, and IGF2BP3 was remarkably dysregulated across multiple cancer types.

Deeply understanding cancer hallmarks requires the detailed information of molecular alterations at multiple dimensions such as gene expression, genetic alteration, epigenomics, clinical information, and metabolome. Therefore, the multi-omics analysis approaches were particularly valuable to deeply discover the molecular alterations in pan-cancer. For example, a multi-omics approach was applied to characterize brain metastasis, and the findings revealed that two molecular subtypes showed notably differential prognosis irrespective of brain tumor subtype (Su et al., 2020). Besides, single-cell sequencing was also an important aspect of multi-omics

analysis. Recently, several single-cell datasets, including CancerSEA and scLM, were developed to facilitate the mechanism discovery and understanding of complex biosystems such as in cancers (Yuan et al., 2019; Song et al., 2020). In the present study, although lacking single-cell sequencing data, we also applied the multi-omics analysis to uncover the molecular mechanisms of m^{6}A regulators in pan-cancer at levels of gene expression, genetic alteration, epigenomics, and clinical information, which might help to facilitate the deep understanding of the modulating mechanisms of m^{6}A regulators in pan-cancer.

Up to now, the great majority of the studies focused on researching one or several m^{6}A RNA methylation regulators in one or several cancer types (Barbieri et al., 2017; Huang et al., 2019). However, the m^{6}A regulators exerted their functions in tumor development and might also collaborate with each other or other factors, and accumulating lines of evidence had indicated that m^{6}A regulators might play a dual role as tumor promoters or tumor suppressors in variously different cancer types, implying that the levels or functions of m^{6}A RNA methylation were determined by the collaboration of m^{6}A regulators in certain conditions (Roundtree et al., 2017; Panneerdoss et al., 2018). Therefore, the comprehensive analyses of all the m^{6}A regulators but not several of them across all the cancer types might help supply unique insights into the molecular mechanisms of m^{6}A RNA methylation in many cancer types. In the present study, the landscapes of the gene expression, genetic alterations, the prognosis significance, and interacting networks of the 20 m^{6}A regulators across dozens of cancer types were revealed by integrative bioinformatics analyses. These results provided new supplementary knowledge about the modulation of the m^{6}A regulators' dysregulation across cancer types and novel insights into the possible molecular mechanisms of the m^{6}A regulators' dysregulation in TCGA cancer types.

The expression alteration of m^{6}A regulators in various cancer types might provide novel insight into the molecular mechanisms of tumorigenesis and new therapy approaches (Chen and Wong, 2020). In addition, many aspects such as genetic alternations, epigenetics, and transcriptional factors could contribute to the dysregulation of the m^{6}A regulators in cancers (Li et al., 2019). For example, the m^{6}A levels were increased through miR-145 targeting YTHDF2, which caused the suppression of cancer cell proliferation in HCC (Yang et al., 2017). Another study demonstrated that SPI1, as a transcriptional factor in hematopoietic cancer cells, could directly suppress the expression of METTL14 (Weng et al., 2018). Therefore, in this study, we not only explored that the methylation and genetic alterations were capable of affecting the expression of the m^{6}A regulators across TCGA cancer types, but also investigated the potential upstream miRNAs and transcriptional factors that were able to target these m^{6}A regulators. Our results including miRNAs or TF–m^{6}A regulators networks provided new supplementary knowledge about the modulation of the m^{6}A regulators' dysregulation across cancer types.

The dysregulation of m^{6}A regulators was involved in the procedures of cancer development (Huang et al., 2020). Hence, discovering novel drugs targeting these m^{6}A regulators was critical for cancer therapy. For example, a chemical compound, MA2, as an inhibitor of FTO, could effectively suppress the tumor progression of glioblastoma (Cui et al., 2017). Besides, FB23-2 was also capable of inhibiting FTO expression to suppress the proliferation of AML cells (Huang et al., 2019). In the present study, we thereby investigated whether there were some potential chemicals that could increase or decrease the expression of the m^{6}A regulators. By analyzing the chemical database, the m^{6}A regulators–potential drugs network was constructed and it might provide benefits for potential drugs discovery to target specific m^{6}A regulators.

CONCLUSION

In summary, our results not only systematically analyze the expression, genetic alterations, oncogenic pathways, and prognosis significance of m^{6}A regulators across multiple cancer types, but also constructed the networks between m^{6}A regulators and potential chemical drugs, miRNAs, or upstream transcriptional factors. These comprehensive analyses might provide novel understanding of these m^{6}A regulators' roles and shed light on their potential molecular mechanisms in cancers as well as help develop new therapy approaches for cancers.

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

Conceptualization, XS and JL; methodology, XS and JZ; software, XS; validation, CZ, XL, and JW; formal analysis, YJ; investigation, XS; resources, JZ; data curation, CZ; writing—original draft preparation, XS; writing—review and editing, XS and JZ; visualization, XS; supervision, YJ; project administration, XS; funding acquisition, JL. All authors have read and approved the final manuscript.

FUNDING

This work was supported by the Science and Technology Commission of Shanghai Municipality (No. 19DZ2201000) and the Outstanding Clinical Discipline Project of Shanghai Pudong (No. PWYgy2018-10).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.771853/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, e05005. doi:10.7554/eLife.05005

Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 Maintains Myeloid Leukaemia by m6A-dependent Translation Control. *Nature* 552 (7683), 126–131. doi:10.1038/nature24678

Bian, S., Ni, W., Zhu, M., Song, Q., Zhang, J., Ni, R., et al. (2020). Identification and Validation of the N6-Methyladenosine RNA Methylation Regulator YTHDF1 as a Novel Prognostic Marker and Potential Target for Hepatocellular Carcinoma. *Front. Mol. Biosci.* 7, 604766. doi:10.3389/fmolb.2020.604766

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discov.* 2 (5), 401–404. doi:10.1158/2159-8290.CD-12-0095

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V. S. K., et al. (2017). UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Chen, M., Wei, L., Law, C.-T., Tsang, F. H.-C., Shen, J., Cheng, C. L.-H., et al. (2018). RNA N6-Methyladenosine Methyltransferase-like 3 Promotes Liver Cancer Progression through YTHDF2-dependent Posttranscriptional Silencing of SOCS2. *Hepatology* 67 (6), 2254–2270. doi:10.1002/hep.29683

Chen, M., and Wong, C.-M. (2020). The Emerging Roles of N6-Methyladenosine (m6A) Deregulation in Liver Carcinogenesis. *Mol. Cancer* 19 (1), 44. doi:10.1186/s12943-020-01172-y

Chen, X.-Y., Zhang, J., and Zhu, J.-S. (2019). The Role of m6A RNA Methylation in Human Cancer. *Mol. Cancer* 18 (1), 103. doi:10.1186/s12943-019-1033-z

Chen, Y., and Wang, X. (2020). miRDB: an Online Database for Prediction of Functional microRNA Targets. *Nucleic Acids Res.* 48 (D1), D127–D131. doi:10.1093/nar/gkz757

Cui, Q., Shi, H., Ye, P., Li, L., Qu, Q., Sun, G., et al. (2017). m 6 A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cel Rep.* 18 (11), 2622–2634. doi:10.1016/j.celrep.2017.02.059

Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., Wiegers, J., Wiegers, T. C., et al. (2021). Comparative Toxicogenomics Database (CTD): Update 2021. *Nucleic Acids Res.* 49 (D1), D1138–D1143. doi:10.1093/nar/gkaa891

De Jesus, D. F., Zhang, Z., Kahraman, S., Brown, N. K., Chen, M., Hu, J., et al. (2019). m6A mRNA Methylation Regulates Human β-cell Biology in Physiological States and in Type 2 Diabetes. *Nat. Metab.* 1 (8), 765–774. doi:10.1038/s42255-019-0089-9

Feng, C., Song, C., Liu, Y., Qian, F., Gao, Y., Ning, Z., et al. (2020). KnockTF: a Comprehensive Human Gene Expression Profile Database with Knockdown/knockout of Transcription Factors. *Nucleic Acids Res.* 48 (D1), D93–D100. doi:10.1093/nar/gkz881

Guo, H., Wang, B., Xu, K., Nie, L., Fu, Y., Wang, Z., et al. (2020). m6A Reader HNRNPA2B1 Promotes Esophageal Cancer Progression via Up-Regulation of ACLY and ACC1. *Front. Oncol.* 10, 553045. doi:10.3389/fonc.2020.553045

He, L., Li, H., Wu, A., Peng, Y., Shu, G., and Yin, G. (2019). Functions of N6-Methyladenosine and its Role in Cancer. *Mol. Cancer* 18 (1), 176. doi:10.1186/s12943-019-1109-9

He, P. C., and He, C. (2021). m 6 A RNA Methylation: from Mechanisms to Therapeutic Potential. *EMBO J.* 40 (3), e105977. doi:10.15252/embj.2020105977

Huang, H., Weng, H., and Chen, J. (2020). m6A Modification in Coding and Non-coding RNAs: Roles and Therapeutic Implications in Cancer. *Cancer Cell* 37 (3), 270–288. doi:10.1016/j.ccell.2020.02.004

Huang, Y., Su, R., Sheng, Y., Dong, L., Dong, Z., Xu, H., et al. (2019). Small-Molecule Targeting of Oncogenic FTO Demethylase in Acute Myeloid Leukemia. *Cancer Cell* 35 (4), 677–691. doi:10.1016/j.ccell.2019.03.006

Li, J.-H., Liu, S., Zhou, H., Qu, L.-H., and Yang, J.-H. (2014). starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and Protein-RNA Interaction Networks from Large-Scale CLIP-Seq Data. *Nucl. Acids Res.* 42, D92–D97. doi:10.1093/nar/gkt1248

Li, Y., Xiao, J., Bai, J., Tian, Y., Qu, Y., Chen, X., et al. (2019). Molecular Characterization and Clinical Relevance of m6A Regulators across 33 Cancer Types. *Mol. Cancer* 18 (1), 137. doi:10.1186/s12943-019-1066-3

Liu, C.-J., Hu, F.-F., Xia, M.-X., Han, L., Zhang, Q., and Guo, A.-Y. (2018). GSCALite: a Web Server for Gene Set Cancer Analysis. *Bioinformatics* 34 (21), 3771–3772. doi:10.1093/bioinformatics/bty411

Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S. K., Lv, R., et al. (2019). N6-Methyladenosine Methyltransferase ZCCHC4 Mediates Ribosomal RNA Methylation. *Nat. Chem. Biol.* 15 (1), 88–94. doi:10.1038/s41589-018-0184-3

Meyer, K. D., and Jaffrey, S. R. (2017). Rethinking m6A Readers, Writers, and Erasers. *Annu. Rev. Cel Dev. Biol.* 33, 319–342. doi:10.1146/annurev-cellbio-100616-060758

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3′ UTRs and Near Stop Codons. *Cell* 149 (7), 1635–1646. doi:10.1016/j.cell.2012.05.003

Panneerdoss, S., Eedunuri, V. K., Yadav, P., Timilsina, S., Rajamanickam, S., Viswanadhapalli, S., et al. (2018). Cross-talk Among Writers, Readers, and Erasers of M 6 A Regulates Cancer Growth and Progression. *Sci. Adv.* 4 (10), eaar8263. doi:10.1126/sciadv.aar8263

Paris, J., Morgan, M., Campos, J., Spencer, G. J., Shmakova, A., Ivanova, I., et al. (2019). Targeting the RNA m6A Reader YTHDF2 Selectively Compromises Cancer Stem Cells in Acute Myeloid Leukemia. *Cell Stem Cell* 25 (1), 137–148. doi:10.1016/j.stem.2019.03.021

Pathan, M., Keerthikumar, S., Ang, C.-S., Gangoda, L., Quek, C. Y. J., Williamson, N. A., et al. (2015). FunRich: An Open Access Standalone Functional Enrichment and Interaction Network Analysis Tool. *Proteomics* 15 (15), 2597–2601. doi:10.1002/pmic.201400515

Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. *Cell* 169 (7), 1187–1200. doi:10.1016/j.cell.2017.05.045

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Sharpe, A. H., and Pauken, K. E. (2018). The Diverse Functions of the PD1 Inhibitory Pathway. *Nat. Rev. Immunol.* 18 (3), 153–167. doi:10.1038/nri.2017.108

Shi, Y., Zheng, C., Jin, Y., Bao, B., Wang, D., Hou, K., et al. (2020). Reduced Expression of METTL3 Promotes Metastasis of Triple-Negative Breast Cancer by m6A Methylation-Mediated COL3A1 Up-Regulation. *Front. Oncol.* 10, 1126. doi:10.3389/fonc.2020.01126

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer Statistics, 2019. *CA A. Cancer J. Clin.* 69 (1), 7–34. doi:10.3322/caac.21551

Song, Q., Su, J., Miller, L. D., and Zhang, W. (2021). scLM: Automatic Detection of Consensus Gene Clusters across Multiple Single-Cell Datasets. *Genomics, Proteomics & Bioinformatics* 19, 330–341. doi:10.1016/j.gpb.2020.09.002

Su, J., Song, Q., Qasem, S., O'Neill, S., Lee, J., Furdui, C. M., et al. (2020). Multi-Omics Analysis of Brain Metastasis Outcomes Following Craniotomy. *Front. Oncol.* 10, 615472. doi:10.3389/fonc.2020.615472

Sun, T., Wu, R., and Ming, L. (2019). The Role of m6A RNA Methylation in Cancer. *Biomed. Pharmacother.* 112, 108613. doi:10.1016/j.biopha.2019.108613

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 (D1), D607–D613. doi:10.1093/nar/gky1131

Weng, H., Huang, H., Wu, H., Qin, X., Zhao, B. S., Dong, L., et al. (2018). METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m6A Modification. *Cell Stem Cell* 22 (2), 191–205. doi:10.1016/j.stem.2017.11.016

Yang, Z., Li, J., Feng, G., Gao, S., Wang, Y., Zhang, S., et al. (2017). MicroRNA-145 Modulates N6-Methyladenosine Levels by Targeting the 3′-Untranslated mRNA Region of the N6-Methyladenosine Binding YTH Domain Family 2 Protein. *J. Biol. Chem.* 292 (9), 3614–3623. doi:10.1074/jbc.M116.749689

Yuan, H., Yan, M., Zhang, G., Liu, W., Deng, C., Liao, G., et al. (2019). CancerSEA: a Cancer Single-Cell State Atlas. *Nucleic Acids Res.* 47 (D1), D900–D908. doi:10.1093/nar/gky939

Zaccara, S., Ries, R. J., and Jaffrey, S. R. (2019). Reading, Writing and Erasing mRNA Methylation. *Nat. Rev. Mol. Cel Biol* 20 (10), 608–624. doi:10.1038/s41580-019-0168-5

Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., et al. (2017). m 6 A Demethylase ALKBH5 Maintains Tumorigenicity of Glioblastoma Stem-like Cells by Sustaining FOXM1 Expression and Cell Proliferation Program. *Cancer Cell* 31 (4), 591–606. doi:10.1016/j.ccell.2017.02.013

Zhao, B. S., Roundtree, I. A., and He, C. (2017). Post-transcriptional Gene Regulation by mRNA Modifications. *Nat. Rev. Mol. Cel Biol* 18 (1), 31–42. doi:10.1038/nrm.2016.132

Zhao, W., Qi, X., Liu, L., Ma, S., Liu, J., and Wu, J. (2020). Epigenetic Regulation of m6A Modifications in Human Cancer. *Mol. Ther. - Nucleic Acids* 19, 405–412. doi:10.1016/j.omtn.2019.11.022

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: 19 January 2022
doi: 10.3389/fgene.2021.796592

EFNA3 Is a Prognostic Biomarker Correlated With Immune Cell Infiltration and Immune Checkpoints in Gastric Cancer

Peng Zheng[1,2†], *XiaoLong Liu*[1†], *Haiyuan Li*[1], *Lei Gao*[1], *Yang Yu*[1], *Na Wang*[1] *and Hao Chen*[1*]

[1]*The Second Clinical Medical College of Lanzhou University, Lanzhou, China,* [2]*Abdominal Department III, Gansu Provincial Tumor Hospital, Lanzhou, China*

Background: Ephrin A3 (EFNA3), like most genes in the ephrin family, plays a central role in embryonic development and can be dysregulated in a variety of tumors. However, the relationship between EFNA3 and gastric cancer (GC) prognosis and tumor-infiltrating lymphocytes remains unclear.

Methods: Tumor Immune Estimation Resource (TIMER) and Gene Expression Profiling Interactive Analysis 2 (GEPIA2) were used to analyze the expression of EFNA3. Kaplan-Meier plots and GEPIA2 were used to evaluate the relationship between EFNA3 expression and GC prognosis. Univariable survival and multivariate Cox analyses were used to compare various clinical characteristics with survival. LinkedOmics database was used for gene set enrichment analysis (GSEA). TIMER database and CIBERSORT algorithm were used to examine the relationship between EFNA3 expression and immune infiltration in GC and to explore cumulative survival in GC. The relationship between EFNA3 and immune checkpoints was examined using cBioPortal genomics analysis. Finally, EFNA3 expression in GC cells and tissues was assayed using quantitative real-time polymerase chain reaction.

Results: EFNA3 expression differs in a variety of cancers, and EFNA3 expression was higher in GC tissue than normal gastric tissue. GC patients with high expression of EFNA3 had worse overall survival, disease-free survival, and first progression. Multivariate analysis identified EFNA3 as an independent prognostic factor for GC. GSEA identified ribosome, cell cycle, ribosome biogenesis in eukaryotes, and aminoacyl-tRNA biosynthesis pathways as differentially enriched in patients with high EFNA3 expression. B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells were significantly negatively correlated with a variety of immune markers. EFNA3 participates in changes in GC immune checkpoint markers in a collinear manner. EFNA3 expression in HGC-27, AGS, MKN45, and NCI-N87 was cell lines higher than that in GES-1, and patients with high expression of EFNA3 had a worse prognosis.

Conclusion: EFNA3 can be used as a prognostic and immune infiltration and checkpoint marker in GC patients.

Keywords: gastric cancer, EFNA3, prognosis biomarkers, immune infiltrates, immune checkpoints

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez, Instituto Nacional de Medicina Genómica, Mexico

Reviewed by:
Sohini Chakraborty, NYU Grossman School of Medicine, United States
Yandong Li, Tongji University, China
Renata Dos Santos Almeida, Aggeu Magalhães Institute (IAM), Brazil

****Correspondence:***
Hao Chen
ery_chenh@lzu.edu.cn

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

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

Received: *17 October 2021*
Accepted: *21 December 2021*
Published: *19 January 2022*

Citation:
Zheng P, Liu X, Li H, Gao L, Yu Y, Wang N and Chen H (2022) EFNA3 Is a Prognostic Biomarker Correlated With Immune Cell Infiltration and Immune Checkpoints in Gastric Cancer. Front. Genet. 12:796592. doi: 10.3389/fgene.2021.796592

INTRODUCTION

Gastric cancer (GC) is one of the most common cancers worldwide, and the mortality rate ranks third among all cancers (Smyth et al., 2020). Surgery is the only cure for GC, but even if tumors are surgically removed, recurrence is common. Radiotherapy, chemotherapy, targeted therapy, and immunotherapy for GC are advancing rapidly, but the prognosis of patients with advanced GC remains poor. Therefore, it is very important to identify effective early diagnostic and prognostic biomarkers.

Ephrin is a general term for a class of cell surface ligands. Ephrin binds to members of the Eph tyrosine kinase receptor family and thus plays an essential role in the migration, rejection, and adhesion of neurons, blood vessels, and epithelial cells during development (Nievergall et al., 2012). Eph receptors and ephrins are signaling molecules involved in axon guidance. Recent studies have shown that they play a critical role in cancer proliferation, invasion, metastasis, and angiogenesis (Chen, 2012). Therefore, many members of the ephrin family are abnormally expressed in cancer cells, and changes in ephrin genes are often associated with greater likelihood of invasion and metastasis and worse prognosis (Kou and Kandpal, 20182018). We found that the tumor microenvironment (TME) promotes tumor growth and suppresses anti-tumor immunity *via* complex signaling pathways. Ephrins expressed in the TME play roles in tumor invasion, metastasis, and angiogenesis (Janes et al., 2021). Modulating the expression of ephrins may affect the TME and ultimately the tumor itself. In the past decade, tremendous advances in immune-related treatments and technologies have occurred. Considerable progress has been made in the development of both treatment methods and treatment techniques (Pulendran and Davis, 2020), particularly those related to immune checkpoints. Considerable research has also focused on the relationship between ephrins and immunity. Ephrin expression has been detected on both human B cells and T cells (Alonso-C et al., 2009; Luo et al., 2016), suggesting that these proteins are involved in immunity.

Our current research primarily focuses on the relationship between the expression of ephrin family proteins and various malignant tumors. For example, ephrin-A1 is highly expressed in hepatocellular carcinoma and associated with poor prognosis (Wada et al., 2014). We also found the same relationship in GC and colorectal cancer (Yuan et al., 2009; Yamamoto et al., 2013). The expression of ephrin-B1 is higher in bladder cancer tissues than normal urothelial tissue, suggesting that ephrin-B1 can be used as a biomarker of bladder cancer aggressiveness (Mencucci et al., 2020). Ephrin-B2 is also highly expressed in endometrial cancer, and patients with low ephrin-B2 expression have a better prognosis (Alam et al., 2007). As a member of the Ephrin family, EFNA3 also plays an important role in the occurrence and development of tumors. EFNA3 promotes the occurrence and development of oral tumors as well as the formation of blood vessels in oral cancer (Wang et al., 2020). EFNA3 also inhibits the proliferation and invasion of Malignant peripheral nerve sheath tumor(15). Sheath tumor cells (Wang et al., 2015). The different roles of EFNA3 in different tumors suggests the protein has diverse functions. To the best of our knowledge, only two studies examining EFNA3 in relation to GC have been published, but these studies did not examine the relationship between expression level and prognosis (Yu et al., 2020; Pei et al., 2021). In view of the role of EFNA3 in other tumors, the relationship between EFNA3 and GC requires further study.

Based on the involvement of ephrins such as EFNA3 in a variety of tumors, we hypothesized that EFNA3 would be a useful diagnostic and prognostic marker in GC patients. Although ephrins play a role in anti-cancer immunity, very little research has focused on this relationship. We therefore examined the relationship between EFNA3 expression and the immune microenvironment and immune checkpoints due to the potential usefulness of monitoring EFNA3 in clinical treatment.

In this study, we used the online tools Tumor Immune Estimation Resource (TIMER) and Gene Expression Profiling Interactive Analysis 2 (GEPIA2) to analyze the expression of EFNA3 in GC tissues. Kaplan-Meier plots and GEPIA2 were employed to explore the relationship between EFNA3 expression and GC prognosis as well as the relationship between EFNA3 and immune cell infiltration and immune checkpoints. To examine the relationship between immune checkpoints and EFNA3, gene set enrichment analysis (GSEA) was used to identify pathways enriched in GC patients with high or low expression of EFNA3. The expression of EFNA3 in GC cells and tissues as it relates to prognosis was evaluated using quantitative real-time polymerase chain reaction (qRT-PCR). The results of our research indicate that EFNA3 plays an important role in GC and clarify the relationship between EFNA3 and GC immunity.

MATERIALS AND METHODS

Tissue Samples

A total of 50 cancerous and paracancerous tissue samples were collected from GC patients during surgery in Gansu provincial Tumor Hospital, and the tissues were stored at −80°C until analysis. Prior to analysis, tissues were homogenized, and total RNA was extracted for qRT-PCR. The study was approved by our institutional Clinical Research Ethics Committee.

qRT-PCR

Cells were collected using a cell scraper and washed twice with cold phosphate-buffered saline. The cells were then lysed with TRIzol RNA extraction reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. RNA was reversely transcribed into cDNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo-Fisher Scientific, Waltham, MA, USA). Subsequently, with cDNA as the template, SYBR Premix Ex Taq™ (TaKaRa, Otsu, Shiga, Japan) was utilized for qRT-PCR. SYBR Green qPCR was used to evaluate the mRNA levels of indicated genes. Expression of target genes was normalized to that of GAPDH, and the data were analyzed according to the $2^{-\Delta\Delta CT}$ method. Primers used for qRT-PCR were as follows: GAPDH, 5′-AGAAGGCTGGGGCTCATTTG-3′ (F), 5′AGGGGCCATCCACAGTCTTC-3′ (R); EFNA3, TACTACTACATCTCCACGCCC

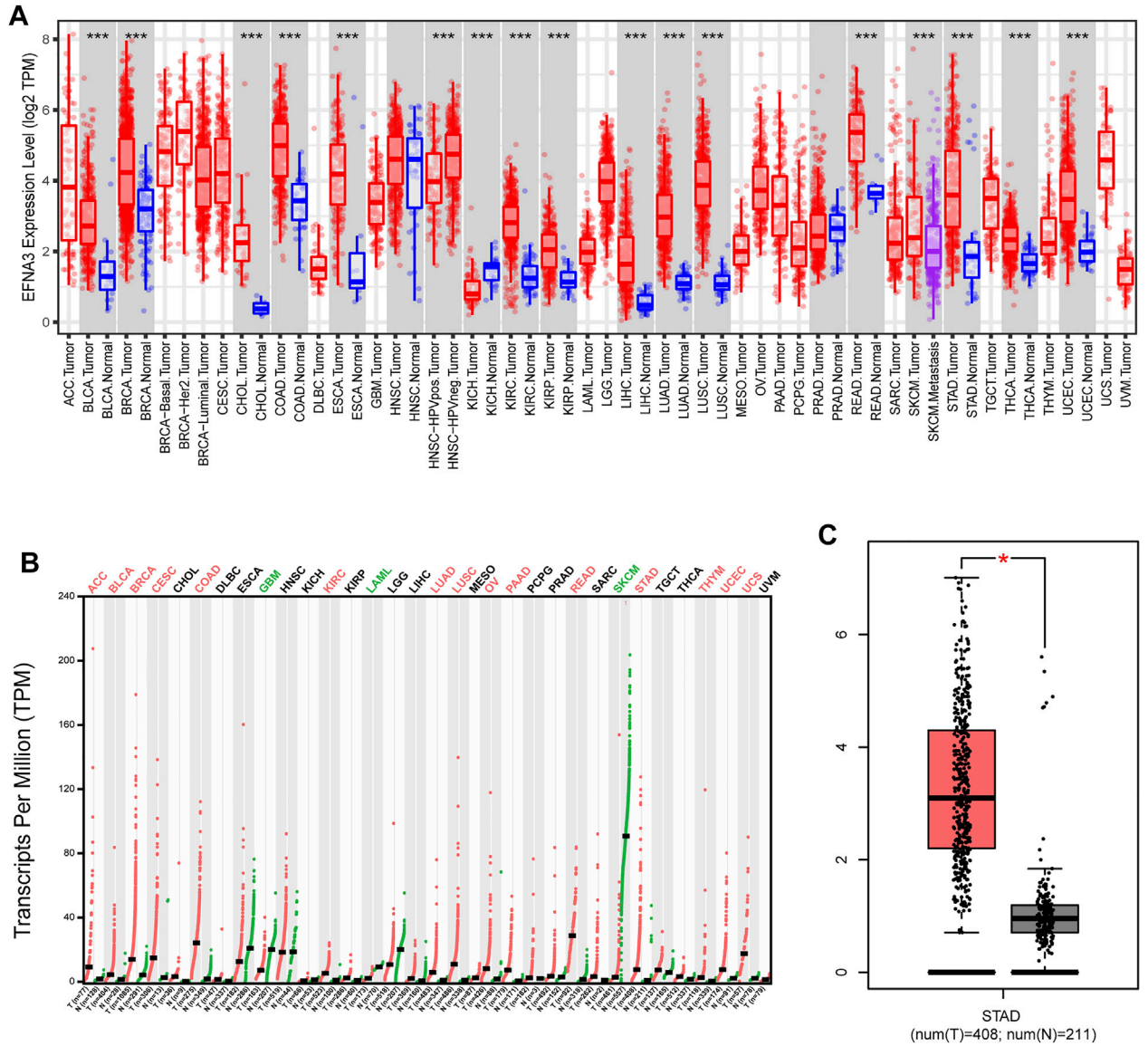


FIGURE 1 | EFNA3 expression levels in different types of human cancers. **(A)** Expression of EFNA3 in different tumor types from TIMER. **(B)** Expression of EFNA3 in different tumor types in GEPIA2. **(C)** GEPIA2 generates box plots for comparing EFNA3 expression in GC and paired normal tissues (TCGA tumor versus TCGA normal + GTEx normal). (*P< 0.05, **P< 0.01, ***P< 0.001).

ACTC-3′ (F), 5′-TCCCGCTGATGCTCTTCTCAA-3′(R) (Zhu et al., 2021). Based on the results of qPCR, we sorted the expression levels of EFNA3 in all patients from high to low. Those with higher than the median value were the high-expression groups, and those below the median were the low-expression groups.

Cell Culture

GES-1 gastric epithelial cells and the GC cell lines HGC-27, AGS, MKN45, and NCI-N87 were purchased from the Cell Resource Center, Peking Union Medical CollegePMUC (Beijing, China). All cells were maintained in Dulbecco's modified Eagle's medium (RPMI-1640; Gibco, USA) supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin and streptomycin (Gibco). All cells were cultured in a 5% CO^2 humidified atmosphere at 37°C (Baust et al., 2017).

GEPIA2 Database Analysis

GEPIA2 (http://gepia2.cancer-pku.cn/) is a newly developed bioinformatics platform for the analysis and processing of transcriptome data from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx)databases.

Survival Analysis and Prognosis Evaluation

Kaplan–Meier plots (http://kmplot.com/analysis/) were generated for prognostic analysis. Based on the median

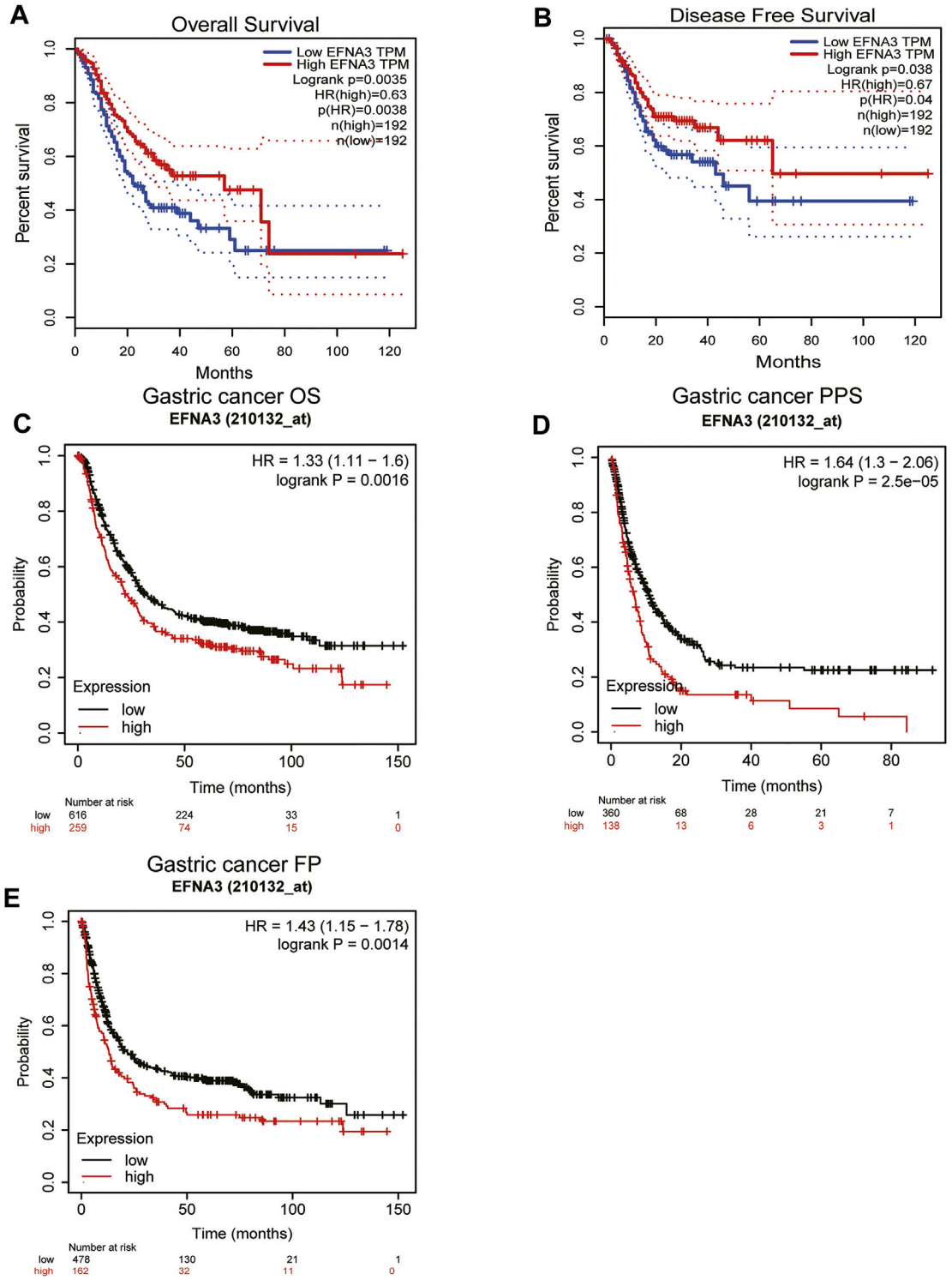


FIGURE 2 | Correlation of EFNA3 expression with prognostic value in GC. Survival curves of differential EFNA3 expression were analyzed using GEPIA2 **(A, B)**. Correlation between EFNA3 and prognosis of STAD in the Kaplan-Meier plot database **(C–E)**. OS, overall survival; DFS, disease-free survival; FP, fast progression; PPS, post-progression survival.

TABLE 1 | Correlation of EFNA3 mRNA expression and clinical prognosis in gastric cancer with different clinicopathological factors by Kaplan-Meier plotter Clinicopathological characteristics. Bold values indicate $p < 0.05$.

	Overall survival (n = 875)			First progression (n = 640)		
	N	**Hazard ratio**	***p*-value**	**N**	**Hazard ratio**	***p*-value**
SEX	—	—	—	—	—	—
—	236	1.58 (1.09–2.28)	**0.015**	201	1.58 (1.05–2.37)	**0.025**
—	534	1.35 (1.09–1.67)	**0.006**	437	1.36 (1.06–1.76)	**0.017**
STAGE	—	—	—	—	—	—
1	67	0.63 (0.23–1.74)	0.37	60	0.63 (0.19–2.06)	0.44
2	140	0.58 (0.3–1.1)	0.091	131	1.69 (0.78–3.66)	0.18
3	305	1.5 (1.09–2.08)	**0.013**	186	1.99 (1.28–3.08)	**0.002**
4	148	0.68 (0.46–1)	**0.049**	141	0.62 (0.41–0.94)	**0.024**
STAGE T	—	—	—	—	—	—
2	241	0.73 (0.47–1.14)	0.16	239	0.71 (0.42–1.17)	0.17
3	204	1.52 (1.07–2.16)	**0.019**	204	1.35 (0.96–1.89	0.085
4	38	0.44 (0.19–1.02)	0.05	39	0.52 (0.22–1.21)	0.12
STAGE N	—	—	—	—	—	—
0	74	0.54 (0.23–1.26)	0.15	72	0.6 (0.26–1.39)	0.23
1	225	1.98 (1.18–3.31	**0.009**	222	1.98 (1.18–3.27)	**0.008**
2	121	2.25 (1.14–3.56)	**0.001**	125	1.79 (1.15–2.78)	**0.009**
3	76	0.62 (0.36–1.07)	0.083	76	0.61 (0.35–1.04)	0.068
1 + 2 + 3	442	1.26 (0.96–1.65)	0.096	423	1.29 (0.97–1.72)	0.084
STAGE M	—	—	—	—	—	—
0	444	1.29 (0.97–1.71)	0.08	443	1.25 (0.95–1.65)	0.11
1	241	1.83 (0.98–3.4)	0.054	56	0.7 (0.38–1.31)	0.26
LAUREN CLASSIFICATION	—	—	—	—	—	—
Intestinal	320	1.67 (1.22–2.3)	**0.001**	263	1.53 (1.08–2.18)	**0.017**
Diffuse	241	1.38 (0.97–1.98)	0.075	231	1.39 (0.97–1.99)	0.068
Mixed	32	0.42 (0.15–1.19)	0.095	28	0.2 (0.06–0.69)	0.0057
DIFFERENTIATION	—	—	—	—	—	—
Poor	165	0.58 (0.38–0.88)	**0.01**	121	0.59 (0.37–0.96)	**0.03**
Moderate	67	2.5 (1.17–5.35)	**0.014**	67	2.27 (1.1–4.68)	**0.023**
Well	32	0.35 (0.13–0.97)	**0.034**	—	—	—
—	First Progression (n = 640)	—	—	—	—	—
—	N	Hazard ratio	*p*-value	—	—	—
SEX	—	—	—	—	—	—
—	149	1.63 (1.04–2.53))	**0.03**	—	—	—
—	348	1.67 (1.28–2.17)	**0.001**	—	—	—
STAGE	—	—	—	—	—	—
2	105	0.56 (0.28–1.09)	**0.001**	—	—	—
3	142	2.77 (1.79–4.27)	**0.002**	—	—	—
4	104	0.69 (0.43–1.08)	0.1	—	—	—
STAGE T	—	—	—	—	—	—
2	196	0.71 (0.42–1.17)	0.17	—	—	—
3	150	1.35 (0.96–1.89	0.085	—	—	—
4	—	—	—	—	—	—
STAGE N	—	—	—	—	—	—
1	169	1.89 (1.11–3.21)	**0.017**	—	—	—
2	105	2.64 (1.62–4.32)	**0.009**	—	—	—
3	63	0.57 (0.31–1.06)	**0.001**	—	—	—
1 + 2 + 3	337	1.44 (1.07–1.94)	**0.014**	—	—	—
STAGE M	—	—	—	—	—	—
0	342	1.47 (1.08–2)	**0.013**	—	—	—
LAUREN CLASSIFICATION	—	—	—	—	—	—
Intestinal	192	2.39 (1.56–3.66)	**0.001**	—	—	—
Diffuse	176	1.34 (0.91–1.98)	0.14	—	—	—

expression of EFNA3, patient samples were divided into two groups for analysis with respect to overall survival (OS), fast progression (FP), and post-progression survival (PPS). The GEPIA2 database was used to determine the prognostic value of EFNA3 expression in relation to the OS and disease-free survival (DFS) of GC patients.

TIMER Database Analysis

TIMER (https://cistrome.shinyapps.io/timer/) is a web server for the comprehensive analysis of tumor-infiltrating immune cells and comprehensive analysis of tumor immunity. We verified the differential expression of EFNA3 between GC samples and samples of adjacent tissues (Ma et al., 2020). We also used the

TABLE 2 | Univariate COX regression analysis and Multivariate COX regression analysis for EFNA3.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	*p* Value	Hazard ratio (95% CI)	*p* Value
Age (>65 *vs*. ≤65)	367	1.620 (1.154–2.276)	**0.005**	1.974 (1.364–2.857)	**<0.001**
Gender (Male *vs*. Female)	370	1.267 (0.891–1.804)	0.188	—	—
T stage (T3&T4 *vs*. T1&T2)	362	1.719 (1.131–2.612)	**0.011**	1.446 (0.919–2.276)	**0.111**
N stage (N2&N3 *vs*. N0&N1)	352	1.650 (1.182–2.302)	**0.003**	1.542 (1.086–2.190)	**0.015**
M stage (M1 *vs*. M0)	352	2.254 (1.295–3.924)	**0.004**	2.860 (1.593–5.133)	**<0.001**
EFNA3 (High *vs*. Low)	370	0.703 (0.506–0.978)	**0.036**	0.625 (0.439–0.889)	**0.009**

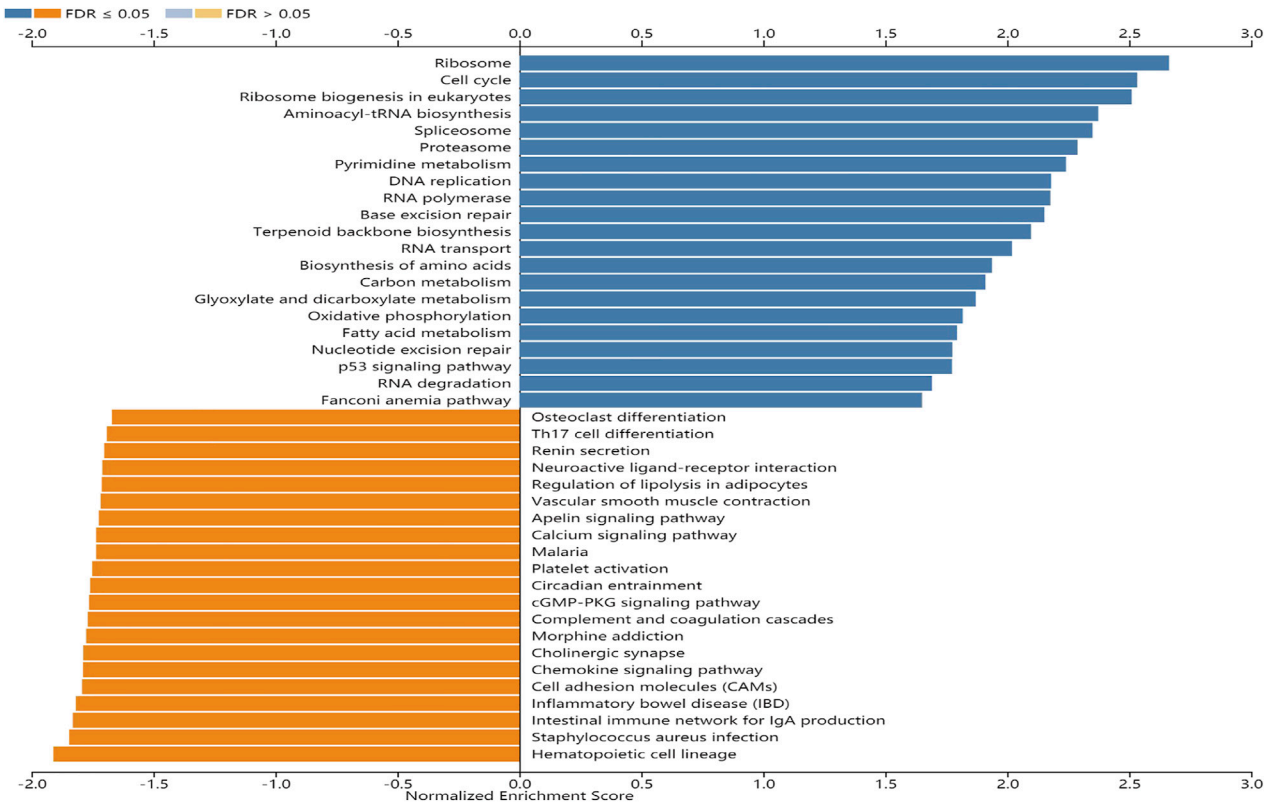


FIGURE 3 | Enrichment plots from GSEA. KEGG pathways of EFNA3 in STAD cohort. FDR, false-discovery rate.

database to analyze EFNA3 expression in Stomach adenocarcinoma (STAD) and the correlation between EFNA3 expression and the abundance of immune infiltrating cells, including B cells, $CD4^+$ T cells, $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells.

Gauging the Immune Response of 22 Tumor-Infiltrating Immune Cells in GC

CIBERSORT (Gentles et al., 2015) (http://cibersort.stanford.edu/) is a deconvolution algorithm based on gene expression that can be used to evaluate changes in the expression of a set of genes relative to all other genes in the sample. We used the CIBERSORT algorithm to examine the responses of 22 TIICs(B cells naïve,B cells memory,Plasma cells, T cells CD8,T cells CD4 naïve,T cells CD4 memory resting,T cells CD4 memory activated,T cells follicular helper,T cells regulatory(Tregs),T cells gamma delta,NK cells resting,NK cells activated, Monocytes,Macrophages M0, Macrophages M1, Macrophages M2,Dendritic cells resting, Dendritic cells activated, Mast cells resting, Mast cells activated, Eosiniphils and Neutrophils) in GC in order to assess the correlations with survival and molecular subgroups.

cBioPortal Analysis

cBioPortal (http://cbioportal.org) provides web resources for exploring, visualizing, and analyzing multi-dimensional cancer genome data (Gao et al., 2013). We used cBioPortal to visualize and compare genetic changes in the following immune checkpoint molecules: PD-L1 (CD274), PD-L2 (PDCD1LG2), CD80, CD86, VTCN1, VSIR, HHLA2, TNFRSF14, PVR, CD112 (NECTIN2), CD200, LGALS9, ICOSLG, TNFSF9, TNFSF4, CD70, TNFSF18, and CD48.

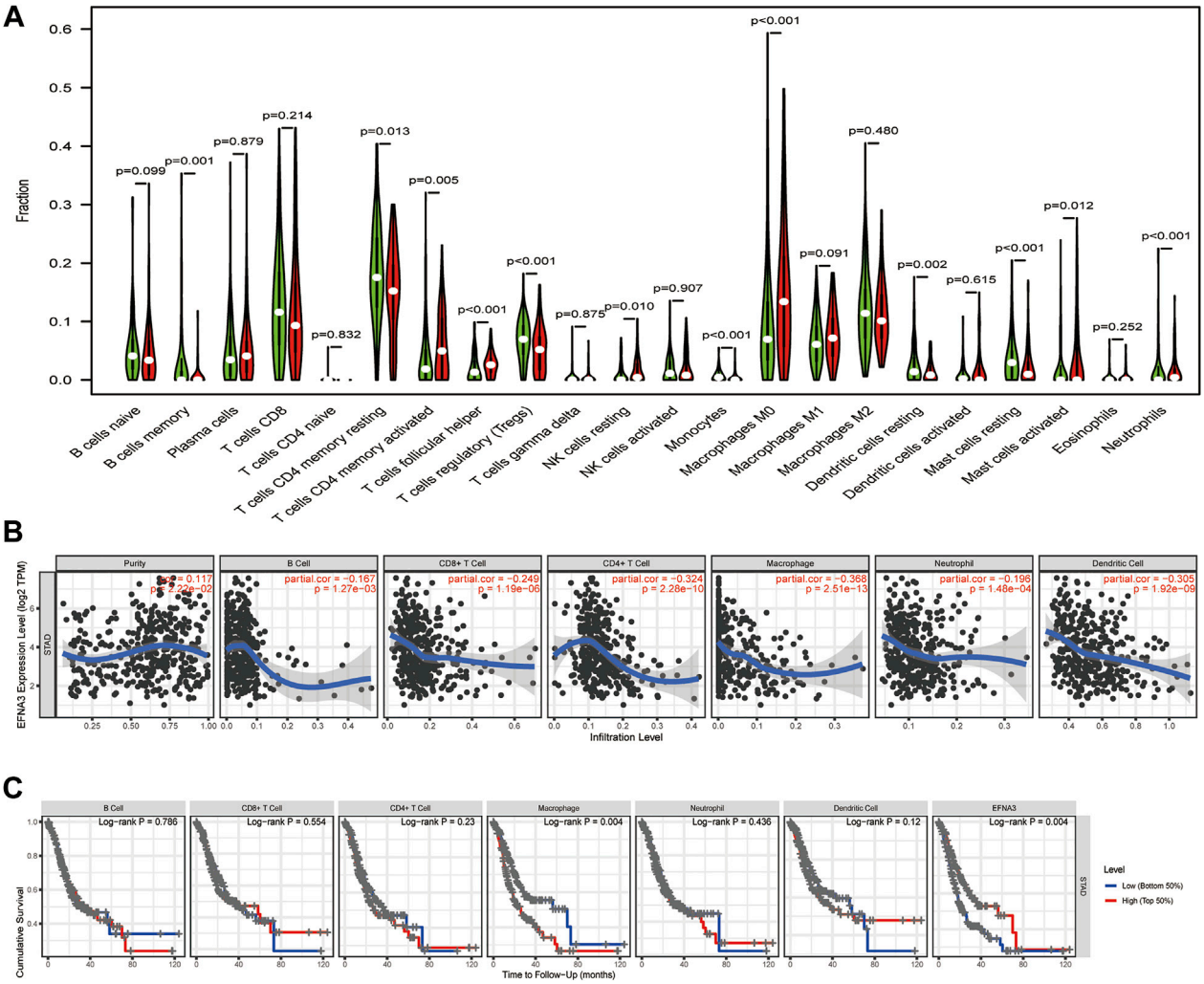


FIGURE 4 | (A) Proportion of 22 subpopulations of immune cells. **(B)** EFNA3 expression level was significantly negatively correlated with levels of infiltrating B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and DCs in GC. **(C)** Cumulative survival was related to macrophages and EFNA3 in GC.

LinkedOmics Database Analysis of EFNA3-Related Pathways

The LinkedOmics (http://www.linkedomics.org) database includes 32 cancer types from TCGA project and a total of 11,158 patients with multiple omics and clinical data. It is also the first multi-omics database that integrates mass spectrometry–based global proteomics data generated by the Clinical Proteomics Cancer Analysis Alliance on selected TCGA tumor samples. We use the LinkedOmics database to view these pathways. We also conducted GSEA in LinkInterpreter and KEGG Pathways Enrichment Analysis of EFNA3-related pathways.

Univariate and Multivariate Cox Regression Analyses

Univariate and multivariate Cox regression were used to analyze survival. Multivariate Cox analysis was used to compare the effects of EFNA3 expression and other clinical characteristics on survival. Patients were divided into high– and low–EFNA3 expression groups. The statistical significance level for the two-tailed test was set to 0.05.

Statistical Analysis

Survival curves were generated using GEPIA2 and Kaplan-Meier plots. The results are displayed with hazard ratio (HR) and P or Cox *p*-values from log-rank tests.

RESULTS

mRNA Expression Levels of EFNA3 in Different Types of Human Cancers

To evaluate differences in EFNA3 expression in tumor and normal tissues, the EFNA3 mRNA levels in tumor and normal tissues of patients with multiple types of cancer were analyzed using the TIMER database. EFNA3 expression was higher in BLCA, CHOL,

TABLE 3 | Correlation analysis between EFNA3 and relate genes and markers of immune cells in TIMER.

Description	Gene markers	None		Purity	
		Cor	P	Cor	P
CD8$^+$ T cell	—	—	—	—	—
—	CD8A	−0.207315046	***	−0.20242845	***
—	CD8B	−0.149493123	**	−0.143242331	**
T cell (general)	—	—	—	—	—
—	CD3D	−0.263887993	***	−0.249219653	***
—	CD3E	−0.267123175	***	−0.26059211	***
—	CD2	−0.238296665	***	−0.224033483	***
B cell	—	—	—	—	—
—	CD19	−0.257111288	***	−0.242635831	***
—	CD79A	−0.304089267	***	−0.295011059	***
Monocyte	—	—	—	—	—
—	CD86	−0.174594623	***	−0.159553219	**
—	CD115 (CSF1R)	−0.232143121	***	−0.217217322	***
TAM	—	—	—	—	—
—	CCL2	−0.170170145	**	-0.150541669	**
—	CD68	−0.110379646	*	-0.10019684	0.051284861
—	IL10	−0.151439718	*	−0.11591185	*
M1 Macrophage	—	—	—	—	—
—	INOS (NOS2)	0.01731189	0.725113784	0.033387321	0.516974952
—	IRF5	−0.058798236	0.231890806	−0.051855302	0.314003055
—	COX2(PTGS2)	0.038192905	0.437599325	0.04611645	0.370625499
M2 Macrophage	—	—	—	—	—
—	CD163	−0.117692005	*	−0.100753601	*
—	VSIG4	−0.128470571	**	−0.108138566	*
—	MS4A4A	−0.245048757	***	−0.228561813	***
Neutrophils	—	—	—	—	—
—	CD66b (CEACAM8)	−0.038892665	0.429403341	−0.045277856	0.379398546
—	CD11b (ITGAM)	−0.237196109	***	−0.222545419	***
—	CCR7	−0.328484339	***	−0.303657644	***
Natural killer cell	—	—	—	—	—
—	KIR2DL1	−0.054005227	0.272356068	−0.046098831	0.370808521
—	KIR2DL3	−0.074236678	0.131084612	−0.048572102	0.345664963
—	KIR2DL4	0.052175156	0.288960296	0.080487011	0.117752329
—	KIR3DL1	−0.047502648	0.33438018	−0.044153189	0.391362199
—	KIR3DL2	−0.092480122	0.059793067	−0.081231199	0.11438445
—	KIR3DL3	0.002491516	0.959641852	−0.001757506	0.972795818
—	KIR2DS4	−0.015480848	0.753189563	−0.036168244	0.482663905
Dendritic cell	—	—	—	—	—
—	HLA-DPB1	−0.25863437	***	−0.235387886	***
—	HLA-DQB1	−0.141898195	*	−0.111787485	*
—	HLA-DRA	−0.155704255	**	−0.124386463	*
—	HLA-DPA1	−0.188483029	**	−0.158691397	*
—	BDCA-1(CD1C)	−0.419921274	***	−0.395322677	***
—	BDCA-4(NRP1)	-0.256360056	***	-0.251782132	***
—	CD11c (ITGAX)	−0.164137654	**	−0.145959864	*
Th1	—	—	—	—	—
—	T-bet (TBX21)	−0.211737847	***	−0.191068168	**
—	STAT4	−0.280788203	***	−0.28766748	***
—	STAT1	0.104556017	*	0.107212799	*
—	IFN-γ (IFNG)	0.036859801	0.453928996	0.055368894	0.28229336
—	TNF-α (TNF)	−0.013194081	0.788624466	0.030997142	0.547439608
Th2	—	—	—	—	—
—	GATA3	−0.268551295	***	−0.27844804	***
—	STAT6	−0.1308076	*	−0.135425002	*
—	STAT5A	−0.10826853	*	−0.101518944	*
—	IL13	−0.050161416	0.30800105	−0.042617284	0.408064011
Tfh	—	—	—	—	—
—	BCL6	−0.057315888	0.243998933	−0.049118066	0.340263959
—	IL21	−0.006365356	0.897134245	0.018712321	0.716516372
Th17	—	—	—	—	—
—	STAT3	−0.061558105	0.210680659	−0.057538398	0.263832341
—	IL17A	−0.019216313	0.696298946	−0.000146833	0.997726758
Treg	—	—	—		—

(Continued on following page)

COAD, KIRC, ESCA, NISC, KIRC, KIRP, LIHC, LUAD, LUSC, READ, SKCM, STAD, THCA, and UCEC compared with normal tissues. In addition, lower expression was observed in KICH (**Figure 1A**). In GEPIA2, we observed high expression in ACC, BLCA, RBCA, COAD, KIRC, LUAD, LUSC, OV, PAAD, READ, STAD, THYM, UCEC, and UCS and low expression in GBM, LAML, and SKCM (**Figure 1B**). EFNA3 was highly expressed in STAD (**Figure 1C**).

Prognostic Utility of EFNA3 in GC

The prognostic value of EFNA3 expression in GC was evaluated using Kaplan-Meier plots and GEPIA2. Expression of EFNA3 was significantly associated with the prognosis of GC patients. The results of GEPIA2 analysis showed that high expression of EFNA3 was associated with longer OS (HR = 0.63, p = 0.0038) and DFS (HR = 0.67, p = 0.04) of GC patients (**Figures 2A,B**) compared with GC patients with low EFNA3 expression. In analyses using the KM web tool, the OS (HR = 1.33 [1.11–1.6], p = 0.0016), PPS (HR = 1.64 [1.3–2.06], p = 2.5e-05), and FP (HR = 1.43 [1.15–1.78], p = 0.0014) of GC patients with high EFNA3 expression (**Figures 2C–E**) values were significantly lower than those of patients with low EFNA3 expression.

In order to better understand the relationship between the expression of EFNA3 and GC, we examined the expression of EFNA3 in relation to various clinical characteristics in GC patients using the KM web tool. High expression of EFNA3 in males and females with stage 3 disease of intestinal type was associated with poor OS, PPS, and FP. In terms of differentiation, high expression of EFNA3 was associated with poor OS, regardless of high, medium, or low differentiation (**Table 1**). Finally, we downloaded GC-related information from TCGA, remove the missing information. Univariate and multivariate Cox analysis identified EFNA3 expression (HR = 0.701 [0.504–0.974], p = 0.034) as an independent prognostic factor in patients with GC, (**Table 2**).

Identification of EFNA3-Related Signaling Pathways Using GSEA

GSEA was performed to identify signaling pathways that are activated in GC. Ribosome, cell cycle, ribosome biogenesis in eukaryotes, and aminoacyl-tRNA biosynthesis pathways were differentially enriched and positively correlated with EFNA3 mRNA expression phenotype. In contrast, hematopoietic cell lineage, *Staphylococcus aureus* infection, intestinal immune network for IgA production, and inflammatory bowel disease pathways were negatively correlated with EFNA3 mRNA expression (**Figure 3**).

Relationship Between EFNA3 Expression and TIICs

We also evaluated whether the expression of EFNA3 is related to immune cell infiltration in GC using data downloaded from TCGA. Tumor specimens were divided into groups based on high and low EFNA3 expression. We used CIBERSORT to calculate and download the gene expression profiles of the samples to infer the immune infiltration of 22 immune cells. The results showed that memory B cells, memory resting CD4 T cells, follicular T helper cells, regulatory T cells, resting NK cells, monocytes, M0 macrophages, resting dendritic cells, resting mast cells, activated mast cells, and neutrophils were the primary immune cells affected by the expression of EFNA3 (**Figure 4A**).

TIMER Analysis of Correlation Between EFNA3 Expression and Immune Cell Infiltration Level and Cumulative Survival in GC

As TIICs are independent predictors of cancer prognosis, it is very important to study the relationship between the expression of EFNA3 and the level of immune cell infiltration. Using the TIMER database, we found that EFNA3 expression was significantly negatively correlated with the infiltration of B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells (**Figure 4B**). Macrophage infiltration and EFNA3 expression were related to the cumulative survival rate of GC patients over time (**Figure 4C**).

In order to further characterize the role of EFNA3 expression and TIICs, we analyzed the relationship between the expression of EFNA3 and immune marker genes in different types of immune cells, including $CD8^+$ T cells, T cells (general), B cells, monocytes, tumor-associated macrophages (TAMs), M1 and M2 macrophages, neutrophils, NK cells, DCs, Th1 cells, Th2 cells, follicular T helper cells, Th17 cells, Tregs, and T cell exhaustion. We found that EFNA3 expression was also related to several immune markers of B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells. These results were consistent with our previous results. Interestingly, the expression of EFNA3 was not related to M1 macrophages but closely related to M2 macrophages. In addition, the expression levels of most marker sets of monocytes and TAMs were closely related to the expression of EFNA3 (**Table 3**).

EFNA3 and Immune Checkpoints

We also explored the genetic changes in the EFNA3 gene and the immune checkpoints we mentioned earlier in GC. The general landscape of EFNA3 and immune checkpoint alteration in GC was compactly visualized, including fusion, amplification, deep deletion, truncating, and missense mutations (**Figure 5**). Genetic alterations in EFNA3 in GC reached as high as 3%, a level higher than that of other immune checkpoint changes (**Figure 5**).

Next, we examined the relationship between EFNA3 and each representative immune checkpoint separately. Mutations in EFNA3 exhibited statistically significant co-occurrences rather than mutual exclusivity with a variety of immune checkpoints, such as CD48, TNFSF4, TNFSF18, PVR, NECTIN2, CD274, and TNFRSF14 (**Table 4**).

Elevated EFNA3 Expression in GC Cell Lines and Tissues

In order to characterize EFNA3 expression in GC tissues and cell lines, qRT-PCR was performed, and the results showed that EFNA3

TABLE 3 | (*Continued*) Correlation analysis between EFNA3 and relate genes and markers of immune cells in TIMER.

Description	Gene markers	None		Purity	
		Cor	P	Cor	P
—	FOXP3	−0.119564378	*	−0.107616351	*
—	CCR8	−0.129723551	*	−0.115791686	*
—	STAT5B	−0.204293596	***	−0.202750648	***
—	TGFβ (TGFB1)	−0.125655895	*	−0.125169215	*
T cell exhaustion	—	—	—	—	—
—	PD-1 (PDCD1)	−0.104825489	*	−0.084697986	0.099680031
—	CTLA4	−0.016864448	0.731835488	0.011626566	0.821508615
—	LAG3	−0.065025822	0.186062406	−0.053968224	0.294665055
—	TIM-3 (HAVCR2)	−0.126420397	*	−0.106746361	*
—	GZMB	0.096918275	*	0.122835389	*

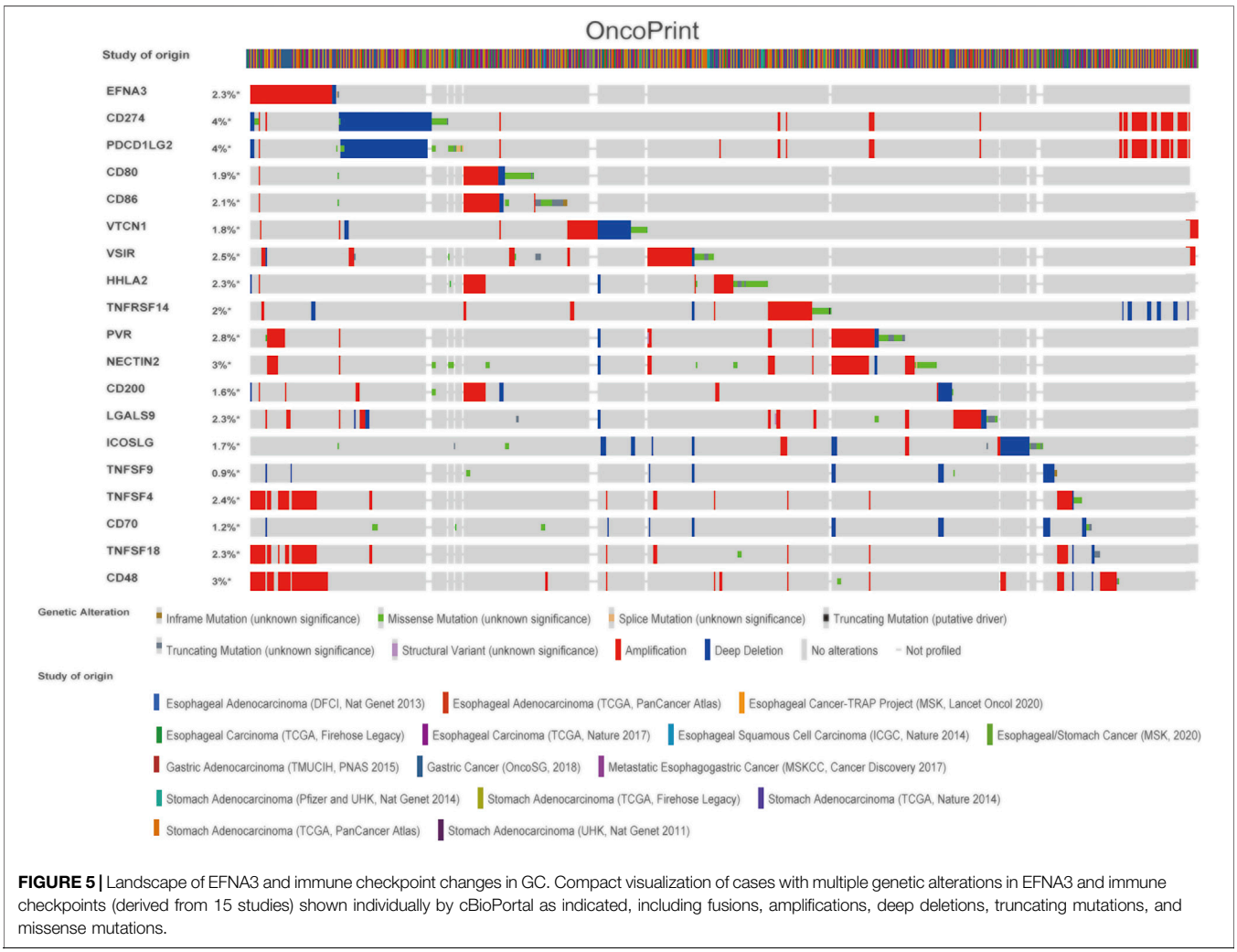


FIGURE 5 | Landscape of EFNA3 and immune checkpoint changes in GC. Compact visualization of cases with multiple genetic alterations in EFNA3 and immune checkpoints (derived from 15 studies) shown individually by cBioPortal as indicated, including fusions, amplifications, deep deletions, truncating mutations, and missense mutations.

expression was significantly higher in GC tissues than adjacent non-cancerous tissues (**Figure 6B**). In GC cell lines, the expression of EFNA3 was significantly higher than in GES-1 cells (**Figure 6A**). In addition, after grouping patients based on high versus low EFNA3 expression, log-rank tests showed that high EFNA3 expression was associated with poor prognosis (**Figure 6C**).

DISCUSSION

Cancer cells at different stages of transformation and metastasis rely on signal transduction between cells. Ephs/Ephrins act as repulsive and attractive signaling molecules between cells and can bind to Eph receptors on neighboring cells, resulting in contact-

TABLE 4 | Mutual-exclusivity analysis between EFNA3 and multiple-immune checkpoints in gastric cancer.

A	B	Neither	A not B	B not A	Both	Log2 odds ratio	*p*-Value	q-Value	Tendency	Significant
EFNA3	CD48	2,374	10	34	51	>3	<0.001	<0.001	Co-occurrence	Yes
EFNA3	TNFSF4	2,384	21	24	40	>3	<0.001	<0.001	Co-occurrence	Yes
EFNA3	TNFSF18	2,384	25	24	36	>3	<0.001	<0.001	Co-occurrence	Yes
EFNA3	PVR	2,347	47	61	14	>3	<0.001	<0.001	Co-occurrence	Yes
EFNA3	NECTIN2	2,335	53	73	8	2.271	<0.001	0.004	Co-occurrence	Yes
EFNA3	CD274	2,290	53	118	8	1.551	0.011	0.042	Co-occurrence	Yes
EFNA3	TNFRSF14	2,346	56	62	5	1.756	0.023	0.071	Co-occurrence	Yes
EFNA3	LGALS9	2,353	57	55	4	1.586	0.056	0.145	Co-occurrence	No
EFNA3	CD200	2,373	58	35	3	1.81	0.066	0.158	Co-occurrence	No
EFNA3	VSIR	2,349	57	59	4	1.482	0.068	0.162	Co-occurrence	No
EFNA3	TNFSF9	2,385	59	23	2	1.814	0.125	0.249	Co-occurrence	No
EFNA3	CD80	2,360	59	48	2	0.737	0.352	0.503	Co-occurrence	No
EFNA3	PDCD1LG2	2,290	57	118	4	0.446	0.356	0.503	Co-occurrence	No
EFNA3	CD86	2,357	59	51	2	0.648	0.379	0.512	Co-occurrence	No
EFNA3	HHLA2	2,351	59	57	2	0.484	0.432	0.535	Co-occurrence	No
EFNA3	CD70	2,378	60	30	1	0.402	0.542	0.59	Co-occurrence	No
EFNA3	VTCN1	2,355	60	53	1	−0.433	0.612	0.642	Mutual exclusivity	No
EFNA3	ICOSLG	2,361	60	47	1	−0.256	0.666	0.674	Mutual exclusivity	No

dependent bidirectional signal transduction between neighboring cells (Kandouz, 2012). Abnormal signal transduction leads to the occurrence and development of tumors. Our research focuses on the ephrin family member EFNA3. Although EFNA3 has not been extensively studied, available research in sheath tumor and oral cancers is gratifying (Yuan et al., 2009; Yamamoto et al., 2013). Here, we studied the expression of EFNA3 in different cancers, focusing on the high expression of EFNA3 in GC and its relationship to poor prognosis. In addition, the expression level of EFNA3 in GC is related to the levels of immune cell infiltration and different immune markers.

The online data results of our study show that the expression of EFNA3 in many types of cancers differs from that in normal tissues. EFNA3 is highly expressed in GC, hepatocellular carcinoma, and other cancers. We found that KICH expression was low in the TIMER database, but there was no difference in the GEPIA2 data. By comparison, SKCM expression was high in TIMER but low in GEIPA2. Differences in expression in the same cancer noted in different databases may be related to differences in data collection methods, statistical analyses, and biological characteristics. In both databases we examined, EFNA3 was highly expressed in GC, consistent with our qRT-PCR results. Kaplan-Meier plot analyses of OS between the databases showed that GC patients with high expression of EFNA3 had a poor prognosis, which was also closely related to gender and classification, stages 3 and 4, stage T3, stages N1 and 2, classification of intestinal, whereas the prognosis of GC patients with high expression of EFNA3 in GEPIA2 was good. The prognostic differences between databases may be related to the study subject inclusion and rejection criteria, the amount of specimens analyzed, as well as other human or random factors. Therefore, we grouped GC patients based on EFNA3 expression from the results of qRT-PCR analyses. Our results show that GC patients with high expression of EFNA3 have a significantly worse prognosis than GC patients with low EFNA3 expression ($p = 0.0376$). The prognostic utility of EFNA3 for GC patients was further evaluated using univariate and multivariate Cox analyses, which indicated that EFNA3 is a useful independent prognostic factor for GC. These results strongly indicate that EFNA3 is a promising prognostic biomarker for GC.

Based on our initial results, we sought to identify signaling pathways that are enriched in GC patients with high expression of EFNA3, because these patients are at higher risk of poor outcome. The results of GSEA showed that high expression of EFNA3 was associated primarily with enrichment of six pathways. The most markedly enriched pathway was the ribosome pathway. In cancer cells, increased ribosome synthesis leads to a corresponding increase in protein synthesis, which plays an important role in the development of most tumors. Inhibition of ribosome biosynthesis has become a new target in cancer treatment (Pelletier et al., 2018; Catez et al., 2019). The change in EFNA3 expression leads to enrichment of the ribosome pathway, indicating that EFNA3 expression is closely related to ribosome biosynthesis in GC cells. Further interactions need to be verified by related experiments; however, our present research still provides new insights regarding the treatment of GC.

Studies of immune cell infiltration have shown that immune cells in the TME play an important role in the progression of cancer (Lei et al., 2020). A deeper understanding of immune cell infiltration in the immune microenvironment could facilitate the development of new strategies for cancer immunotherapy. Our results show that the expression of EFNA3 is negatively correlated with the infiltration of a variety of immune cells, with the highest correlation with macrophages (Cor = −0.368, $p = 2.51e\text{-}13$). Based on that result, we explored tumor-associated macrophages (TAMs) and genetic markers of M1 and M2 macrophages. Interestingly, the three genetic markers of M1 macrophages were not correlated with the expression of EFNA3, whereas the expression of the three genetic markers of M2

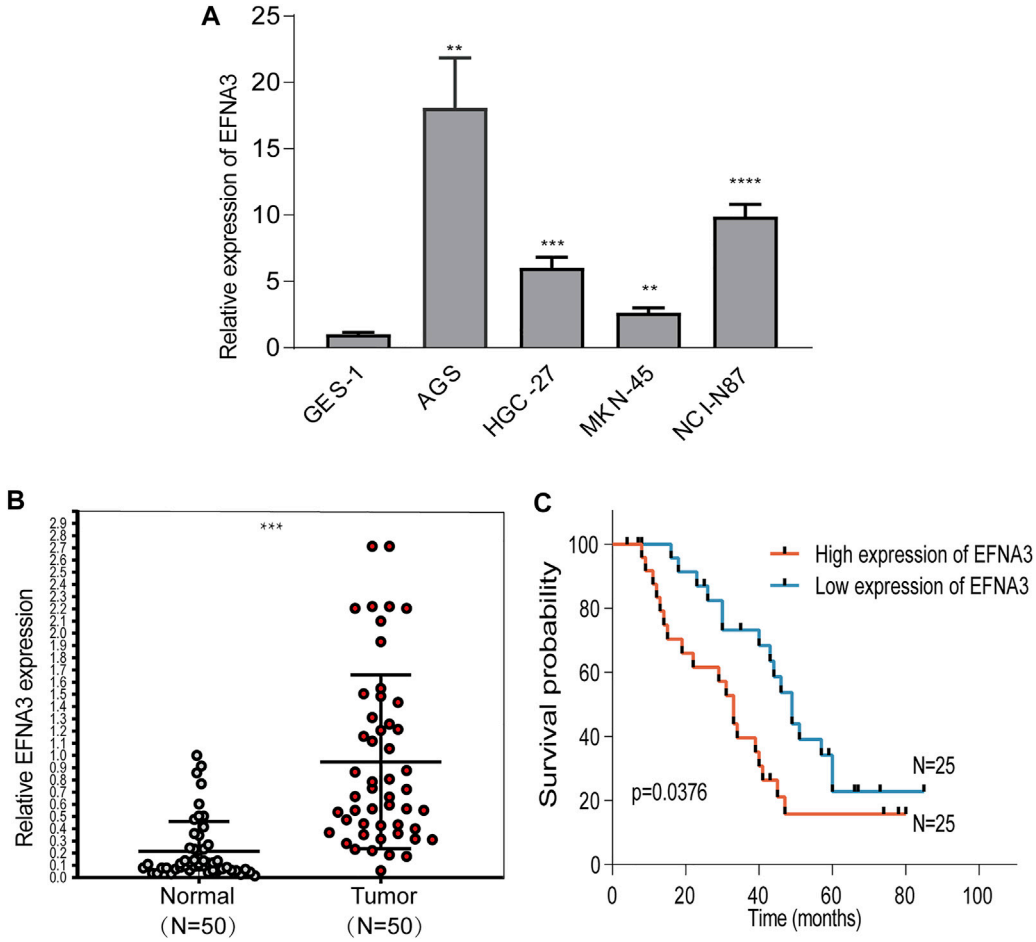


FIGURE 6 | EFNA3 expression was significantly up-regulated in cell lines and GC tissues. qRT-PCR showed that EFNA3 expression was up-regulated in GC cell lines **(A)** and GC tissues **(B)**. High EFNA3 expression was related to shorter OS time **(C)**. (**P< 0.01, ***P< 0.001, ****P< 0.0001).

macrophages examined were closely related to the expression of EFNA3. M1 macrophages mainly participate in positive immune responses such as immune surveillance and inhibition of tumor growth, whereas M2 macrophages mainly secrete inhibitory cytokines (such as IL-10 and TGF-β) to down-regulate the immune response, thereby promoting tumor growth (Cortese et al., 2020). Therefore, the close relationship between EFNA3 expression and M2 macrophages may be related to its M2 macrophages down-regulation of the immune response. We know that TAMs are not identical to the M1 and M2 macrophage subtypes, but TAMs are similar to M2 macrophages and promote tumor growth by inducing immunosuppression (Mehla and Singh, 2019). M2 macrophages also cooperate with Th2 and Treg cells to affect multiple steps of tumor development (Najafi et al., 2019). The expression of markers of Th2 cells (GATA3, STAT6, STAT5A) and Tregs (FOXP3, CCR8, STAT5B, TGF-β [TGFB1]) differed significantly. These results may indicate that EFNA3 has the potential to regulate TAMs. We therefore examined the relationship between high and low expression of EFNA3 and 22 types of immune cells and found that high EFNA3 expression is correlated is correlated with M2 macrophages (p = 0.034), consistent with the TIMER results. In our research, we found many articles related to ephrins and T cells. For example, in GC, EFNB1 inhibits T cells *via* follicular T helper cells (Lu et al., 2017). In experimental autoimmune encephalomyelitis and multiple sclerosis, the expression of EFNB1 and EFNB2 was found to be related to the migration of T cells (Luo et al., 2016). Based on this observation, we explored TIICs because the analysis of TIICs in human tumors usually focuses on T cells. We found that EFNA3 is highly expressed primarily by infiltrating activated CD4 memory T cells and follicular T helper cells, whereas low EFNA3 expression is primarily associated with infiltrating resting CD4 memory T cells and Tregs.

The blocking of immune checkpoints is increasingly considered a primary future method for cancer immunotherapy. However, at least in current clinical practice, the treatment of GC is focused primarily on surgery and radiotherapy (Zhao et al., 2019). Genomic investigations showed that EFNA3 actually participates in the changes in

immune checkpoints. Changes of EFNA3 in expression co-occur with changes in a wide range of immune checkpoints (CD48, TNFSF4, TNFSF18, PVR, NECTIN2, CD274, and TNFRSF14), which strongly suggest that EFNA3 is a co-regulator of immune checkpoints in GC.

Although the present study further elucidated the relationship between EFNA3 expression and prognosis in GC through analyses involving multiple databases and experiments, the pathogenic mechanism of EFNA3 in GC was only examined to a limited degree using GSEA. Further studies are needed to verify our present results. In addition, the reasons for the prognostic differences between the different databases could not be conclusively determined. In order to eliminate potentially interfering factors, studies with larger sample sizes will be needed to minimize potential errors. Our study was limited by the small sample size and therefore could not fully elucidate the relationship between EFNA3 expression and GC prognosis. Finally, although we studied the relationship between EFNA3 expression and immune checkpoints using online databases, clearly determining this relationship requires further confirmation. Although these problems will likely be solved in the future, our research clearly shows that GC tissues express significantly higher levels of EFNA3, and high expression of EFNA3 is associated with a worse outcome in GC, as it is closely related to immune cell infiltration and regulation of immune checkpoints. In short, EFNA3 appears to hold tremendous promise as both a target in GC immunotherapy and a promising prognostic indicator of GC.

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 Research Ethics Committee of Gansu provincial Tumor Hospital (China). Informed consent was obtained from all individual participants included in the study. All methods or experimental protocols were approved by Gansu provincial Tumor Hospital (No. A202106110018) and carried out in accordance with relevant guidelines and regulations.

AUTHOR CONTRIBUTIONS

PZ and HC designed the study. XL analyzed the data and wrote the manuscript. HL and LG collected the data. YY and NW analyzed the data. All authors revised the manuscript. All authors approved the final version for submission.

FUNDING

This research was supported by Key Talents Project of Gansu Province (No. 2019RCXM020), Key Project of Science and Technology in Gansu province (19ZD2WA001), Science and technology project of Chengguan District of Lanzhou City (2019RCCX0034), Cuiying Scientific and Technological Innovation Program of Lanzhou University Second Hospital (No. CY2017-ZD01) and Science and technology project of Chengguan District of Lanzhou City (2020SHFZ0039).

REFERENCES

Alam, S. M., Fujimoto, J., Jahan, I., Sato, E., and Tamaya, T. (2007). Overexpression of EphrinB2 and EphB4 in Tumor Advancement of Uterine Endometrial Cancers. *Ann. Oncol.* 18 (3), 485–490. doi:10.1093/annonc/mdl414

Alonso-C, L. M., Trinidad, E. M., de Garcillan, B., Ballesteros, M., Castellanos, M., Cotillo, I., et al. (2009). Expression Profile of Eph Receptors and Ephrin Ligands in Healthy Human B Lymphocytes and Chronic Lymphocytic Leukemia B-Cells. *Leuk. Res.* 33 (3), 395–406. doi:10.1016/j.leukres.2008.08.010

Baust, J. M., Buehring, G. C., Campbell, L., Elmore, E., Harbell, J. W., Nims, R. W., et al. (2017). Best Practices in Cell Culture: an Overview. In Vitro *Cell.Dev.Biol.-Animal* 53 (8), 669–672. doi:10.1007/s11626-017-0177-7

Catez, F., Dalla Venezia, N., Marcel, V., Zorbas, C., Lafontaine, D. L. J., and Diaz, J.-J. (2019). Ribosome Biogenesis: An Emerging Druggable Pathway for Cancer Therapeutics. *Biochem. Pharmacol.* 159, 74–81. doi:10.1016/j.bcp.2018.11.014

Chen, J. (2012). Regulation of Tumor Initiation and Metastatic Progression by Eph Receptor Tyrosine Kinases. *Adv. Cancer Res.* 114, 1–20. doi:10.1016/B978-0-12-386503-8.00001-6

Cortese, N., Carriero, R., Laghi, L., Mantovani, A., and Marchesi, F. (2020). Prognostic Significance of Tumor-Associated Macrophages: Past, Present and Future. *Semin. Immunol.* 48, 101408. doi:10.1016/j.smim.2020.101408

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* 6 (269), l1. doi:10.1126/scisignal.2004088

Gentles, A. J., Newman, A. M., Liu, C. L., Bratman, S. V., Feng, W., Kim, D., et al. (2015). The Prognostic Landscape of Genes and Infiltrating Immune Cells across Human Cancers. *Nat. Med.* 21 (8), 938–945. doi:10.1038/nm.3909

Janes, P. W., Vail, M. E., Ernst, M., and Scott, A. M. (2021). Eph Receptors in the Immunosuppressive Tumor Microenvironment. *Cancer Res.* 81 (4), 801–805. doi:10.1158/0008-547210.1158/0008-5472.can-20-3047

Kandouz, M. (2012). The Eph/Ephrin Family in Cancer Metastasis: Communication at the Service of Invasion. *Cancer Metastasis Rev.* 31 (1-2), 353–373. doi:10.1007/s10555-012-9352-1

Kou, C.-T. J., and Kandpal, R. P. (20182018). Differential Expression Patterns of Eph Receptors and Ephrin Ligands in Human Cancers. *Biomed. Res. Int.* 2018, 1–23. doi:10.1155/2018/7390104

Lei, X., Lei, Y., Li, J.-K., Du, W.-X., Li, R.-G., Yang, J., et al. (2020). Immune Cells within the Tumor Microenvironment: Biological Functions and Roles in Cancer Immunotherapy. *Cancer Lett.* 470, 126–133. doi:10.1016/j.canlet.2019.11.009

Lu, P., Shih, C., and Qi, H. (2017). Ephrin B1-Mediated Repulsion and Signaling Control Germinal Center T Cell Territoriality and Function. *Science* 356 (6339), eaai9264. doi:10.1126/science.aai9264

Luo, H., Broux, B., Wang, X., Hu, Y., Ghannam, S., Jin, W., et al. (2016). EphrinB1 and EphrinB2 Regulate T Cell Chemotaxis and Migration in Experimental Autoimmune Encephalomyelitis and Multiple Sclerosis. *Neurobiol. Dis.* 91, 292–306. doi:10.1016/j.nbd.2016.03.013

Ma, C., Luo, H., Cao, J., Gao, C., Fa, X., and Wang, G. (2020). Independent Prognostic Implications of *RRM2* in Lung Adenocarcinoma. *J. Cancer* 11 (23), 7009–7022. doi:10.7150/jca.47895

Mehla, K., and Singh, P. K. (2019). Metabolic Regulation of Macrophage Polarization in Cancer. *Trends Cancer* 5 (12), 822–834. doi:10.1016/j.trecan.2019.10.007

Mencucci, M. V., Lapyckyj, L., Rosso, M., Besso, M. J., Belgorosky, D., Isola, M., et al. (2020). Ephrin-B1 Is a Novel Biomarker of Bladder Cancer Aggressiveness. Studies in Murine Models and in Human Samples. *Front. Oncol.* 10, 283. doi:10.3389/fonc.2020.00283

Najafi, M., Hashemi Goradel, N., Farhood, B., Salehi, E., Nashtaei, M. S., Khanlarkhani, N., et al. (2019). Macrophage Polarity in Cancer: A Review. *J. Cel. Biochem.* 120 (3), 2756–2765. doi:10.1002/jcb.27646

Nievergall, E., Lackmann, M., and Janes, P. W. (2012). Eph-Dependent Cell-Cell Adhesion and Segregation in Development and Cancer. *Cell. Mol. Life Sci.* 69 (11), 1813–1842. doi:10.1007/s00018-011-0900-6

Pei, J.-P., Zhang, C.-D., Yusupu, M., Zhang, C., and Dai, D.-Q. (2021). Screening and Validation of the Hypoxia-Related Signature of Evaluating Tumor Immune Microenvironment and Predicting Prognosis in Gastric Cancer. *Front. Immunol.* 12, 705511. doi:10.3389/fimmu.2021.705511

Pelletier, J., Thomas, G., and Volarević, S. (2018). Ribosome Biogenesis in Cancer: New Players and Therapeutic Avenues. *Nat. Rev. Cancer* 18 (1), 51–63. doi:10.1038/nrc.2017.104

Pulendran, B., and Davis, M. M. (2020). The Science and Medicine of Human Immunology. *Science* 369 (6511), eaay4014. doi:10.1126/science.aay4014

Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C., and Lordick, F. (2020). Gastric Cancer. *The Lancet* 396 (10251), 635–648. doi:10.1016/S0140-6736(20)31288-5

Wada, H., Yamamoto, H., Kim, C., Uemura, M., Akita, H., Tomimaru, Y., et al. (2014). Association between Ephrin-A1 mRNA Expression and Poor Prognosis after Hepatectomy to Treat Hepatocellular Carcinoma. *Int. J. Oncol.* 45 (3), 1051–1058. doi:10.3892/ijo.2014.2519

Wang, H., Wang, L., Zhou, X., Luo, X., Liu, K., Jiang, E., et al. (2020). OSCC Exosomes Regulate miR-210-3p Targeting EFNA3 to Promote Oral Cancer Angiogenesis through the PI3K/AKT Pathway. *Biomed. Res. Int.* 2020, 1–13. doi:10.1155/2020/2125656

Wang, Z., Liu, Z., Liu, B., Liu, G., and Wu, S. (2015). Dissecting the Roles of Ephrin-A3 in Malignant Peripheral Nerve Sheath Tumor by TALENs. *Oncol. Rep.* 34 (1), 391–398. doi:10.3892/or.2015.3966

Yamamoto, H., Tei, M., Uemura, M., Takemasa, I., Uemura, Y., Murata, K., et al. (2013). Ephrin-A1 mRNA Is Associated with Poor Prognosis of Colorectal Cancer. *Int. J. Oncol.* 42 (2), 549–555. doi:10.3892/ijo.2012.1750

Yu, S., Hu, C., Cai, L., Du, X., Lin, F., Yu, Q., et al. (2020). Seven-Gene Signature Based on Glycolysis Is Closely Related to the Prognosis and Tumor Immune Infiltration of Patients with Gastric Cancer. *Front. Oncol.* 10, 1778. doi:10.3389/fonc.2020.01778

Yuan, W.-J., Ge, J., Chen, Z.-K., Wu, S.-B., Shen, H., Yang, P., et al. (2009). Over-expression of EphA2 and EphrinA-1 in Human Gastric Adenocarcinoma and its Prognostic Value for Postoperative Patients. *Dig. Dis. Sci.* 54 (11), 2410–2417. doi:10.1007/s10620-008-0649-4

Zhao, Q., Cao, L., Guan, L., Bie, L., Wang, S., Xie, B., et al. (2019). Immunotherapy for Gastric Cancer: Dilemmas and Prospect. *Brief. Funct. Genom* 18 (2), 107–112. doi:10.1093/bfgp/ely019

Zhu, X., Jiang, S., Wu, Z., Liu, T., Zhang, W., Wu, L., et al. (2021). Long Non-coding RNA TTN Antisense RNA 1 Facilitates Hepatocellular Carcinoma Progression via Regulating miR-139-5p/SPOCK1 Axis. *Bioengineered* 12 (1), 578–588. doi:10.1080/21655979.2021.1882133

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

frontiers in Genetics

ORIGINAL RESEARCH
published: 20 January 2022
doi: 10.3389/fgene.2021.760534

Systematic Analysis of Expression and Prognostic Values of Lysyl Oxidase Family in Gastric Cancer

Li Wang[1†], *Shan Cao*[2†], *Rujun Zhai*[1], *Yang Zhao*[3] *and Guodong Song*[1*]

[1]Department of Gastrointestinal Surgery, The Second Hospital of Tianjin Medical University, Tianjin, China, [2]Department of Respiratory, The Second Hospital of Tianjin Medical University, Tianjin, China, [3]Radiology Department, The Second Hospital of Tianjin Medical University, Tianjin, China

OPEN ACCESS

Edited by:
Frank Emmert-Streib,
Tampere University, Finland

Reviewed by:
Yuen Ting Lau,
Queen Mary Hospital, Hong Kong SAR, China
Chenyu Sun,
AMITA Health St Joseph Hospital, United States
Shihori Tanabe,
National Institute of Health Sciences (NIHS), Japan
Shan Yu,
Fudan University, China
Shihai Liu,
The Affiliated Hospital of Qingdao University, China

***Correspondence:**
Guodong Song
tijdoctor@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: 18 August 2021
Accepted: 16 December 2021
Published: 20 January 2022

Citation:
Wang L, Cao S, Zhai R, Zhao Y and Song G (2022) Systematic Analysis of Expression and Prognostic Values of Lysyl Oxidase Family in Gastric Cancer. Front. Genet. 12:760534. doi: 10.3389/fgene.2021.760534

Background: Gastric cancer (GC) remains the fifth most commonly diagnosed malignancy worldwide, with a poor prognosis. The lysyl oxidase (LOX) family, a type of secreted copper-dependent amine oxidases, is comprised of LOX and four LOX-like (LOXL) 1–4 isoforms and has been reported to be dysregulated in a number of different type cancers. However, the diverse expression patterns and prognostic values of LOX family in GC have yet to be systematically analyzed.

Methods: ONCOMINE, GEPIA, UALCAN, Kaplan–Meier Plotter, LOGpc, cBioPortal, GeneMANIA and Metascape databases were utilized in this study to analyze the expression, prognostic values, mutations and functional networks of LOX family in GC.

Results: The mRNA expression levels of LOX, LOXL1 and LOXL2 were significantly higher in GC, the expression level of LOXL3 was contrary in different databases, while the expression level of LOXL4 made no difference; the expression levels of LOX, LOXL1 and LOXL3 were higher in stages 2–4 than that of normal tissues and stage 1, while the mRNA level of LOXL2 in stage 1–4 was higher than normal tissues; patients with high expression of LOX and LOXL 2-4 had poor OS; the genes correlated with LOX and LOXL2 were enriched in extracellular matrix organization, vasculature development and skeletal system development.

Conclusion: Our results indicated that the LOX family, especially LOX and LOXL2, might play an important role in GC oncogenesis, and they may become biomarkers for predicting tumor prognosis and potential targets for tumor therapy.

Keywords: gastric cancer, LOX family, bioinformatics analysis, biomarker, prognosis

INTRODUCTION

Even though declines in (GC)incidence and mortality rates have been observed consistently across world regions, GC remains the fifth most commonly diagnosed malignancy worldwide, with over 1 million estimated new cases in 2018 (Bray et al., 2018). However, due to its advanced-stage diagnosis, excess mortality from this cancer is high, making GC the third most common cause of cancer related death with 784,000 deaths globally (Bray et al., 2018). Despite major advances in understanding the epidemiology, pathology, and molecular mechanisms of GC and in implementing emerging therapeutic options such as targeted and immune-based therapies, not all patients respond

to existing molecularly targeted agents developed for certain acknowledged biomarkers (Chau, 2017). As a result, it is of great importance to further identify novel diagnostic and prognostic biomarkers in terms of the biological complexity, poor prognosis and high reoccurrence of GC (Kang et al., 2018).

The extracellular matrix (ECM) is a complex network of secreted molecules, the function of which is to program cell behavior including supporting cell adhesion, survival and migration. The remodeling of the ECM in cancer plays an important role in controlling the progression of disease and influences cell growth, motility and survival (Yamauchi et al., 2018a). The lysyl oxidase (LOX) family, which are well known as ECM-modifying proteins, they participate in the crosslinking of collagens and elastin in the ECM, promoting its maturation (Yamauchi et al., 2018b; Chitty et al., 2019). The LOX family, a type of secreted copper-dependent amine oxidases, is comprised of five homologous members: LOX and lysyl oxidase-like proteins 1–4 (LOXL1, LOXL2, LOXL3 and LOXL4) (Molnar et al., 2003). Structurally, these members are all characterized by a highly conserved C-terminal domain and a variable N-terminal domain. The composition of C-terminal domain contains copper binding domain, amino acid residues forming lysine tryosylquinone (LTQ), cofactor formation, and a cytokine receptor-like (CRL) domain (Wang et al., 2016). The pro-domains are existed in the N-terminal region of LOX and LOXL1, whereas four scavenger-receptor cysteine-rich (SRCR) domains in the N-terminal are observed in LOXL2-4 (Xiao and Ge, 2012). Mature active forms of LOX and LOXL1 are obtained through a specific cleavage process induced by bone morphogenetic protein 1 (BMP-1), whereas LOXL2, LOXL3, and LOXL4 do not require this cleavage process to mature. In particular, a pre-pro-LOX protein is encoded by LOX mRNA and converted to the inactive LOX preprotein (pro-LOX) in the cytoplasm. The Pro-LOX protein is further cleaved by BMP-1 to form an active LOX with the LOX propeptide (LOX-PP) to perform its function (Wang et al., 2016). In healthy tissue, the synthesis of the LOX family is tightly regulated to control the amount of active LOX family members present. While the LOX family has been reported to be dysregulated in a number of different type cancers (Li et al., 2015; Salvador et al., 2017; Shao et al., 2019; Zeltz et al., 2019; Hu L. et al., 2020). The changes in LOX family member regulation, expression and subsequently enzymatic activity are therefore important factors in cancer progression (Setargew et al., 2021). The LOX family of enzymes may be favorable targets for anti-stromal therapeutics due to their importance in cancer development and progression when compared to healthy state ECM (Setargew et al., 2021). Additionally, highly LOX family expressing tumors have increased LOX family levels detectable in plasma (Rachman-Tzemah et al., 2017), and thus indicate the potential to be used as tumor serum markers.

The dysregulated expression level of LOX family and their relationship with clinicopathological features and prognosis have been partly reported in human GC. With the revolutionized development of microarray and bioinformatic technology, we conducted this systematical study using the data from The Cancer Genome Atlas (TCGA) and other versatile public databases to analyze the expression levels, mutations, functional networks and prognostic values of different LOX in GC, so as to reveal potential diagnostic, therapeutic, and prognostic targets for GC, and the results in different databases were verified with each other to make the results more convincible.

MATERIALS AND METHODS

Oncomine Database

The mRNA expression levels of LOX family in various cancers and their normal tissue counterparts were analyzed using the Oncomine database (http://www.oncomine.org/) (Rhodes et al., 2007). A p-value of 0.001, a fold change of 2, and a gene rank in the top 10% were set as the significance thresholds. The p value was calculated using the Student's t-test.

GEPIA Database

GEPIA (http://gepia2.cancer-pku.cn/) is a gene expression analysis web which contains 9,736 tumors and 8,587 normal samples from the TCGA and the Genotype-Tissue Expression (GTEx) project (Tang et al., 2019). Here we used GEPIA to compare the expression levels between TCGA cancer and matched TCGA normal and GTEx normal. The results were expressed as boxplots, and the cutoff criteria were set as $p < 0.01$ and $|\text{Log2FC}| > 1$.

UALCAN

UALCAN (http://ualcan.path.uab.edu/) is a comprehensive, user-friendly, and interactive web resource for analyzing cancer OMICS data (Chandrashekar et al., 2017). In this study, we used UALCAN to compare the expression levels of LOX family and their relationship with tumor stages. Student's t-test was used to generate a p-value and the p-value cutoff was 0.05.

Survival Analysis

We used The Kaplan Meier plotter (http://kmplot.com/analysis/) and LOGpc (http://bioinfo.henu.edu.cn/DatabaseList.jsp) to evaluate the prognostic value of LOX family mRNA expression in which cancer patients were split into high and low expression group based on median values of mRNA expression and validated by K-M survival curves. The Kaplan Meier plotter is capable to assess the effect of 54 k genes (mRNA, miRNA, protein) on survival in 21 cancer types including breast, ovarian, lung and GC (Szász et al., 2016). LOGpc (Long-term Outcome and Gene Expression Profiling Database of pan-cancers) encompasses 209 expression datasets, provides 13 types of survival terms for 31,310 patients of 27 distinct malignancies (Liu et al., 2018). The log-rank test was used for computing p-value, with the hazard ratio (HR) and 95% confidence intervals (CI), and $p < 0.05$ was regarded as significant.

cBioPortal

The cBioPortal (https://www.cbioportal.org/) for Cancer Genomics provides visualization, analysis and download of large-scale cancer genomics data sets (Cerami et al., 2012).

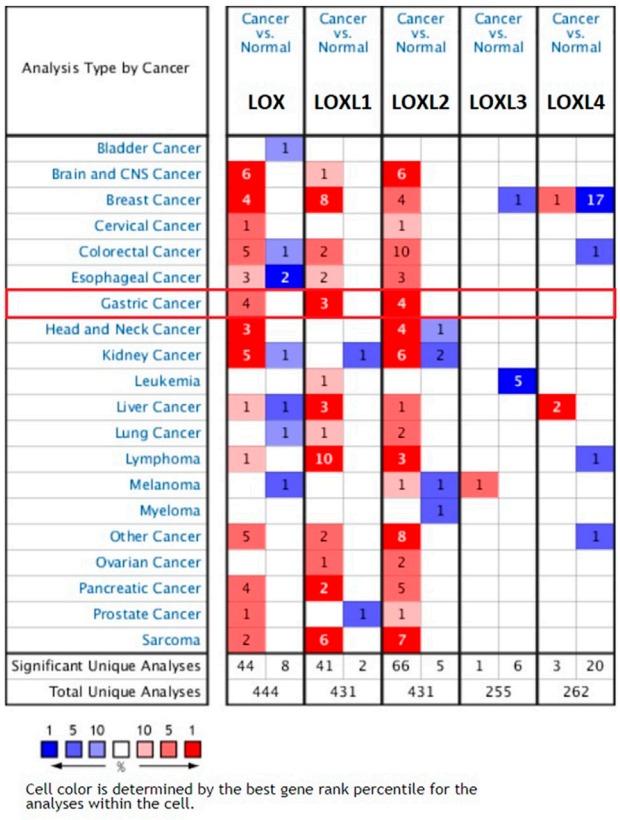


FIGURE 1 | Transcription levels of LOX family in different cancer types (Oncomine).

The stomach adenocarcinoma (TCGA, Firehose Legacy) dataset was selected to figure out the alterations of LOX family. We also estimated the correlations of each LOX family by analyzing their mRNA expression (RNA Seq V2 RSEM), and then the Spearman correlation coefficient was put into Microsoft Excel 2016 to draw the heat maps. Besides, genes with the highest expression correlation with each LOX protein were generated by cBioPortal, and the top 50 co-expressed genes with highest Spearman correlation score were included in the following functional enrichment analysis.

GeneMANIA Database

GeneMANIA (https://genemania.org/) is actively developed at the University of Toronto, in the Donnelly Centre for Cellular and Biomolecular Research, in the labs of Gary Bader and Quaid Morris, and it can find other genes that are related to a set of input genes, using a very large set of functional association data (Warde-Farley et al., 2010). GeneMANIA constructed protein–protein interaction (PPI) networks in terms of physical interaction, coexpression, predicted, colocalization, common pathway, genetic interaction, and shared protein domains. In this study, GeneMANIA was used to generate and analyze gene co-expression network.

Metascape

We used the Metascape web (http://metascape.org) (Zhou et al., 2019) to perform functional enrichment analysis by using the top 50 co-expressed genes of LOX family. The functional process and pathway, following the default, included Canonical Pathway (MSigDB), Hallmark Gene Sets (MSigDB), Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway and Gene Oncology (GO).

RESULTS

Transcriptional Levels of LOX Family in GC and Other Cancers

The transcription level differences of LOX family between tumor and normal tissues were analyzed in multiple cancer types using the Oncomine database. As shown in **Figure 1**, the expressions of LOX, LOXL1 and LOXL2 have been up-regulated in most of the studied tumors, while LOXL3 and LOXL4 only have differences in expression in a small number of tumors.

The mRNA expression level of LOX was significantly up-regulated in patients with GC in 4 analyses out of 23 in 2 datasets out of 7. Chen gastric statistics (Chen et al., 2003) indicated that LOX is overexpressed in gastric intestinal adenocarcinoma compared with gastric normal tissue with a fold change of 2.312, diffuse gastric adenocarcinoma with a fold change of 2.004, and gastric mixed adenocarcinoma with a fold change of 3.232 (**Table 1**). Wang gastric analysis (Wang et al., 2012) revealed that LOX is upregulated in GCwith a fold change of 2.287 (**Table 1**). The transcriptional level of LOXL1 was significantly up-regulated in GC in 3 analyses out of 23 in 3 datasets out of 7. In Chen's (Chen et al., 2003) dataset, the expression of LOXL1 was 2.077 times higher in gastric mixed adenocarcinoma than normal tissues (**Table 1**). In Wang's dataset (Wang et al., 2012), the expression of LOXL1 was 2.083 times higher in GC tissues than normal tissues (**Table 1**). In DErrico's dataset (D'Errico et al., 2009), the expression of LOXL1 was 2.192 times higher in gastric mixed adenocarcinoma than normal tissues (**Table 1**). Upregulation of LOXL2 was observed in 4 analyses in GC tissues compared with normal tissues, with a fold change of 2.118 in Cui's dataset (Cui et al., 2011), a fold change of 2.424 in Wang's dataset, a fold change of 2.039 in Chen's (Chen et al., 2003) dataset, and a fold change of 2.681 in DErrico's dataset (D'Errico et al., 2009) respectively (**Table 1**). The analyses of Oncomine database showed no difference in transcriptional levels of LOXL3 and LOXL4 in GC (**Figure 1**).

Furthermore, we used GEPIA to compare the mRNA expression of LOX family between 408 TCGA GC and 211 matched TCGA normal and GTEx normal, used UALCAN to compare the expression levels of LOX family between 415 TCGA GC and 34 TCGA normal. The GEPIA analyses showed that LOX/LOXL1/LOXL2 were higher in GC than in normal tissues (**Figure 2A**). The results of UALCAN indicated that LOX/LOXL1/LOXL2/LOXL3 were over-expressed in GC (**Figure 2B**).

TABLE1 | Remarkable changes of LOX family expression in transcription level between GC and normal gastric tissues (ONCOMINE).

Genes	Type of GC versus normal	Fold change	*p* value	t value	References
LOX	Gastric intestinal type adenocarcinoma	2.312	8.32E-14	9.498	Chen gastric statistics
	Diffuse gastric adenocarcinoma	2.004	1.56E-5	5.536	Chen gastric statistics
	Gastric mixed adenocarcinoma	3.232	1.79E-4	4.679	Chen gastric statistics
	Gastric cancer	2.287	3.13E-4	4.231	Wang gastric statistics
LOXL1	Gastric mixed adenocarcinoma	2.077	5.58E-6	7.664	Chen gastric statistics
	Gastric cancer	2.083	3.47E-5	4.767	Wang gastric statistics
	Gastric mixed adenocarcinoma	2.192	3.75E-5	7.558	DErrico gastric statistics
LOXL2	Gastric cancer	2.118	7.91E-14	8.084	Cui gastric statistics
	Gastric cancer	2.424	2.93E-5	4.838	Wang gastric statistics
	Diffuse gastric adenocarcinoma	2.039	6.41E-5	4.880	Chen gastric statistics
	Gastric intestinal type adenocarcinoma	2.681	2.19E-9	7.165	DErrico gastric statistics

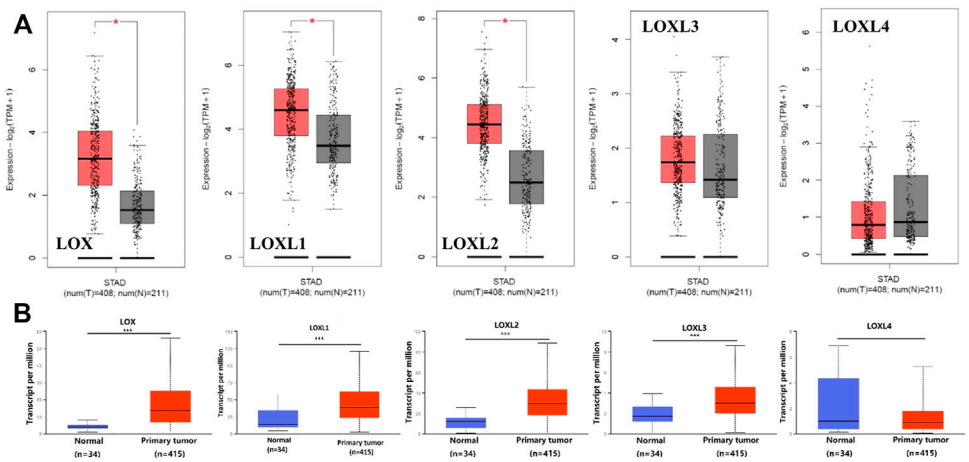


FIGURE 2 | Boxplot showing the expression levels of LOX family inGC. **(A)** GEPIA analysis. The number of normal samples are 211 (grey box), and number of primary tumor samples are 408 (red box), red star means $p < 0.01$; **(B)** UALCAN analysis. The number of normal samples are 34 (blue box), and number of primary tumor samples are 415 (red box) $^{***}p < 0.001$.

Relationship Between the mRNA Levels of LOX Family and the Cancer Stages of GC

Next, we analyzed the relationship between the mRNA expression of different LOX family members with patients' individual cancer stages of GC patients by using UALCAN. LOX, LOXL1, LOXL2 and LOXL3 groups significantly varied, whereas LOXL4 groups did not significantly differ (**Figure 3**). According to clinical stages, the mRNA level of LOX in stage 2–4 was higher than normal tissues and stage 1, while there is no difference between stage 2–4 (**Figure 3A**). The similar result was found in expression of LOXL1 and LOXL3 (**Figures 3B,D**). The mRNA level of LOXL2 in stage 1–4 was higher than normal tissues and highest expression was found in stage 2 tissues (**Figure 3C**). There was no difference in expression of LOXL4 among different stages (**Figure 3E**).

Prognostic Values of LOX Family in GC

We further explored the prognostic values of LOX family in patients with GC by using the Kaplan–Meier Plotter database and LOGpc database. We separated all GC patients into two groups (high vs. low) based on median expression values for each LOX protein across all GC samples and compared overall survival (OS) between the two groups. The Kaplan-Meier curve and log rank test analyses revealed that all of LOX family members were significantly associated with the OS ($p < 0.05$) in patients with GC (**Figure 4A**). Meanwhile, the results of LOGpc analyses indicated that the increased LOX, LOXL2-4 mRNA expression were associated with low OS ($p < 0.05$) in patients with GC, but the expression of LOXL1 had no correction with prognosis ($p > 0.05$) in patients with GC (**Figure 4B**).

Genetic Mutations and PPI Network of LOX Family

We analyzed the types and frequency of LOX Family alterations in a cohort of GC patients using cBioPortal. The LOX family were altered in 55 (14%) samples of 393 patients with stomach

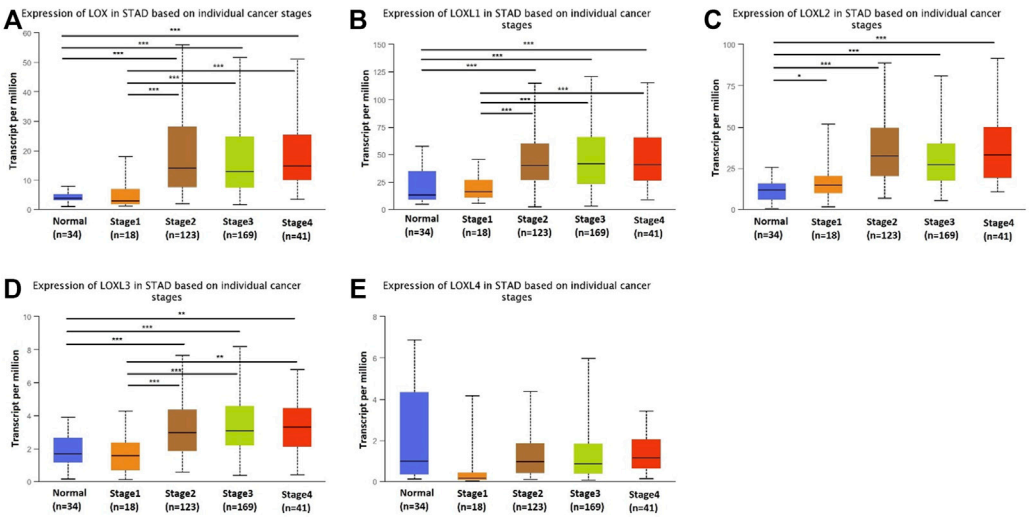


FIGURE 3 | Transcription levels of LOX family in different stage of patients with GC. LOX **(A)**, LOXL1 **(B)**, LOXL2 **(C)**, and LOXL3 **(D)** groups significantly varied, whereas LOXL4 **(E)** groups did not significantly differ. (UALCAN, ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

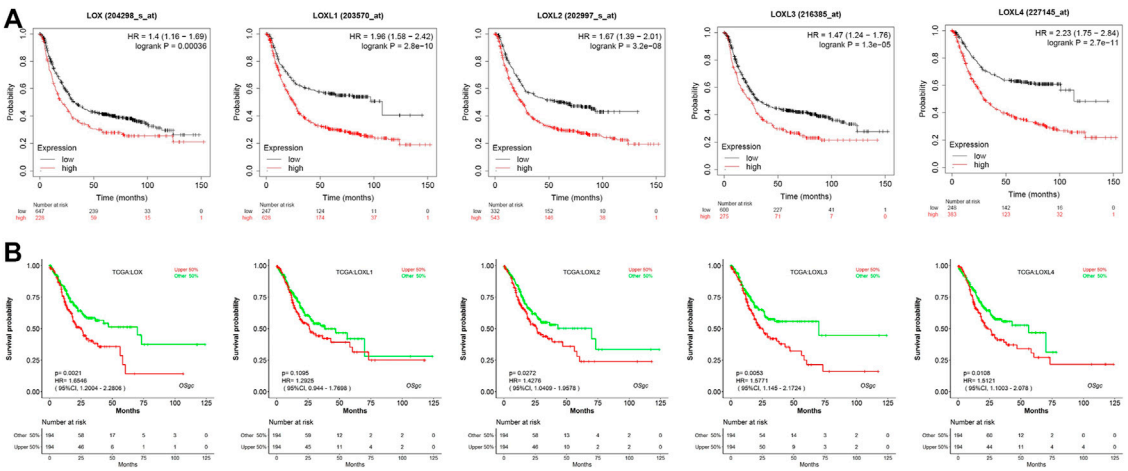


FIGURE 4 | Prognostic feature of mRNA expression of distinct LOX family members in GC patients. **(A)** OS of Kaplan–Meier plotter revealed that all of LOX family members were significantly associated with the OS ($p < 0.05$); **(B)** OS of LOGpc indicated that LOX, LOXL2-4 mRNA expression were associated with OS ($p < 0.05$).

adenocarcinoma (**Figure 5A**). We also calculated the correlations of LOX family with each other by analyzing their mRNA expressions (RNA sequencing (RNA seq V2 RSEM)) via the cBioPortal online tool for stomach adenocarcinoma (TCGA, Firehose Legacy) and Pearson's correction was included. The results indicated significant and positive correlations in the following pairs: LOX and LOXL2, LOX and LOXL3, LOXL1 and LOXL4 (**Figure 5B**). The mutation rates of LOX, LOXL1, LOXL2, LOXL3 and LOXL4 were 4, 2.5, 5, 2.5 and 2.8%, respectively (**Figure 5C**).

Next, we used Gene-MANIA to construct a PPI network for LOX family, and the result is shown in **Figure 5D**. The most top 20 related genes are as follows: TLL1, BMP1, EGFL7, EFEMP2, SNAI1, COL1A2, COL1A1, MFAP4, FBLN5, COL3A1, ELN, JTB, DHDDS, SEPHS2, CD5, SCART1, CD163L1, CD163, SSC4D and SSC5D.

Functional Enrichment Analysis of Genes Co-expressed With LOX/LOXL2

Considering the expression level of LOX family in GC tumor tissues and their prognostic values in GC, LOX and LOXL2 were taken into next functional enrichment analysis. The top 50 genes, which had the most significant correlation with LOX/LOXL2 generated by cBioPortal, were included in the following functional enrichment analysis using Metascape.

The results shown in **Figure 6** indicated that the most top 5 significant biological process with LOX and its co-expressed

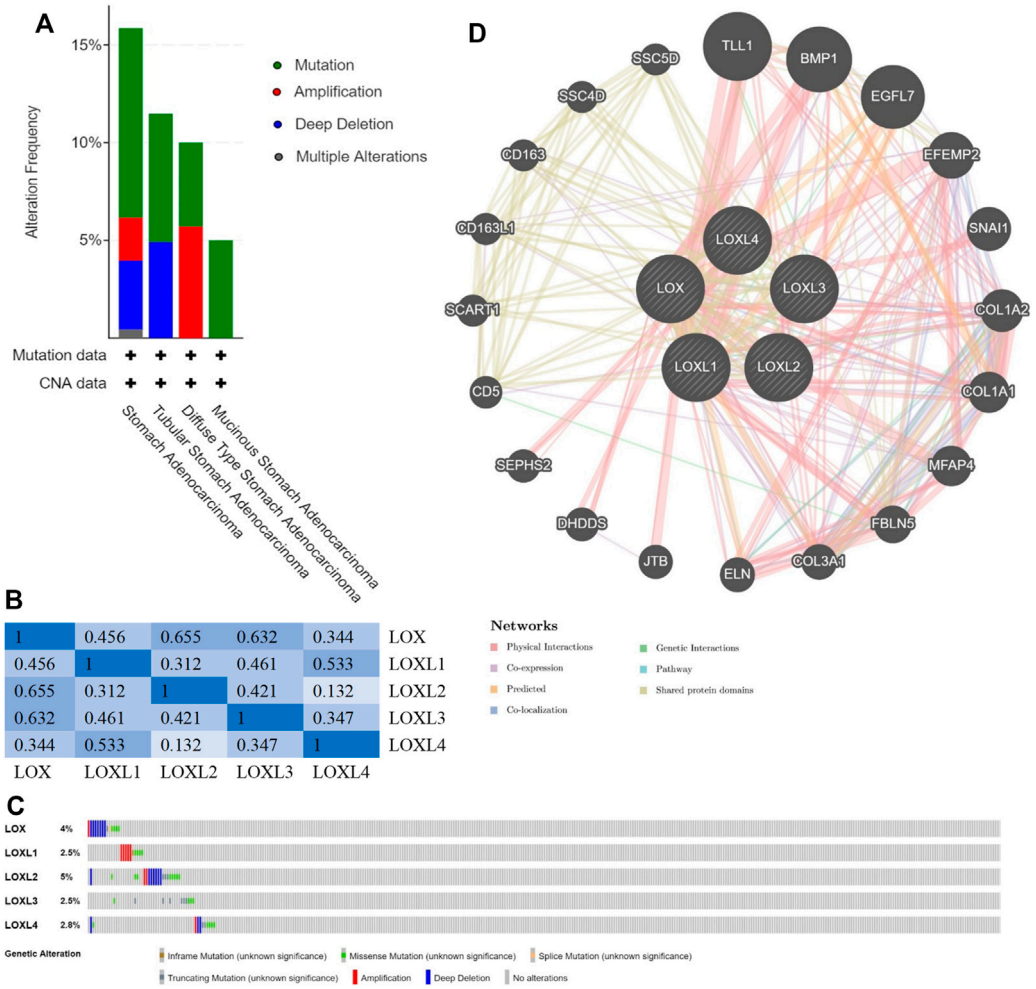


FIGURE 5 | Genomic alterations (cBioPortal) and network (GeneMANIA) of LOX family inGC. **(A)** Distribution of LOX family genomic alterations according to cancer type. **(B)** Correlations of different LOX family members with each other inGC. **(C)** OncoPrint of LOX family alterations inGC. **(D)** Network of the 20 most frequently altered neighboring genes of LOX family.

genes were GO: 0030198 (extracellular matrix organization), M5884 (NABA CORE MATRISOME), M18 (PID INTEGRINL PATHWAY), GP: 0001944 (vasculature development), and GO: 0001501 (skeletal system development) (**Figure 6A**); while the most top 5 with LOXL2 were GO: 0030198 (extracellular matrix organization), R-HAS-1474290 (Collagen formation), GO: 0035987 (endodermal cell differentiation), GP: 0001944 (vasculature development), and GO: 0001501 (skeletal system development) (**Figure 6C**). **Figures 6B,D** were networks that exhibited the interactions among cluster of genes enriched in biological processes and pathways mentioned above.

DISCUSSION

The LOX family has been reported to be dysregulated in a number of cancers (Li et al., 2015; Salvador et al., 2017; Shao et al., 2019; Zeltz et al., 2019; Hu L. et al., 2020). Although the role of LOX family in tumorigenesis and prognosis of several cancers has been partially confirmed, further bioinformatics analysis of GChas yet to be performed. In this study, we used multitalented public databases to reveal the dysregulated expression of the LOX family and their relations with tumor stage and prognosis. We mainly found that the mRNA expression levels of LOX, LOXL1 and LOXL2 were significantly higher in GC, the expression level of LOXL3 was contrary in different databases, while the expression level of LOXL4 made no difference; the expression levels of LOX, LOXL1 and LOXL3 were higher in stages 2–4 than that of normal tissues and stage 1, while the mRNA level of LOXL2 in stage 1–4 was higher than normal tissues; patients with high expression of LOX and LOXL 2-4 had poor OS; the genes correlated with LOXL2/4 were enriched in extracellular matrix organization, vasculature development and skeletal system development.

LOX is a secreted extracellular matrix protein that plays an important role in remodeling the extracellular matrix and

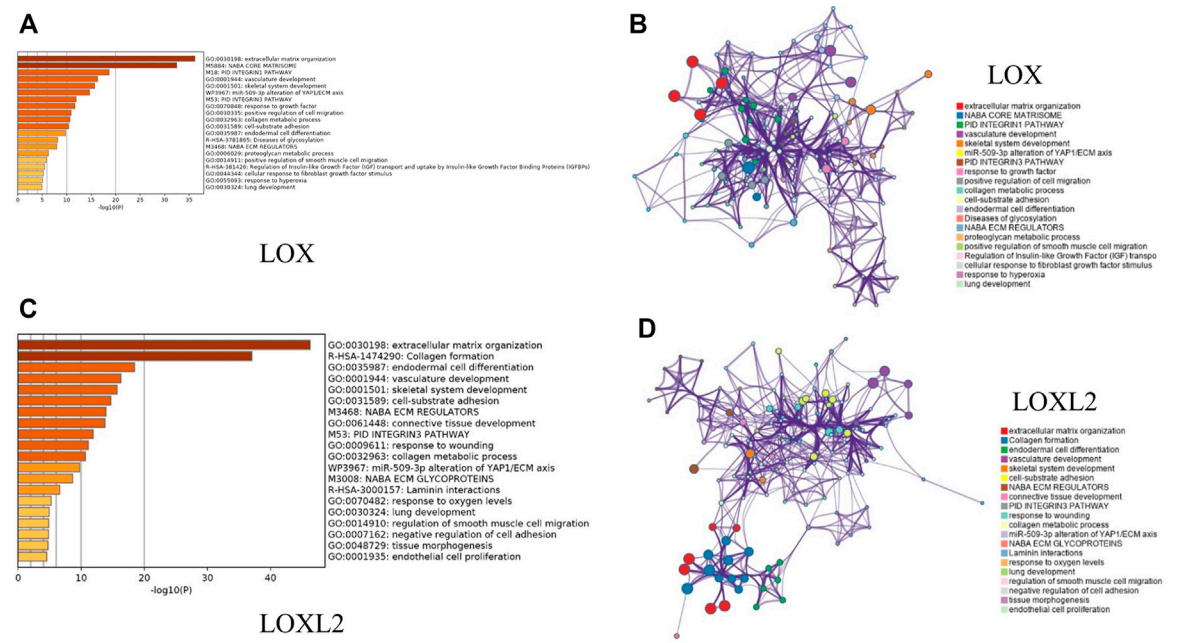


FIGURE 6 | Functional enrichment analysis of genes co-expressed with LOX and LOXL2 using Metascape. **(A,C)** Heatmaps of the molecular functions, biological processes, or pathways enriched with LOX and LOXL2 co-expressed genes. The bar color shade was decided by the *p* value, the deeper the shade, the less the *p* value. **(B,D)** Networks exhibiting interactions among the clusters of genes enriched in the molecular functions, biological processes, or pathways presented in the heatmaps. The points in different colors represented clusters of genes enriched in different molecular functions, biological processes, and pathways.

promoting tumor progression. Higher LOX mRNA expression was detected in GC tissues than that in adjacent normal gastric tissues, and was significantly correlated with the invasion depth, tumor differentiation, lymph node metastasis, lymphatic invasion, venous invasion, and peritoneal metastasis in GC patients, predicting a poor prognosis of GC patients with high expression of LOX (Zhang et al., 2013). Besides, the mRNA and protein levels of LOX in GC cells and tissues were higher under hypoxia condition than that under normoxia condition (Kasashima et al., 2016). The number of migrating and invading GC cells in hypoxia was significantly decreased after knockdown of LOX (Kasashima et al., 2016). The mechanism involved in LOX-mediated proliferation facilitation in GC is enhancement of Warburg effect through regulation of HIF-1α and c-Myc (Li et al., 2019). The risk of macrophages high density, high microvessel density (MVD), low neomicrovessel maturation, MMP-9 expression and low type IV collagen was elevated after LOX overexpression, suggesting that LOX activated cancer stromal cells and facilitated the progression of GC (Peng et al., 2018). Combining LOX with CEA, CA724, CA199, and CA125 could increase the sensitivity of predicting lymph nodes metastasis and peritoneal metastasis in GC (Lai et al., 2014). Similar results have also been confirmed in our research, however, one study indicated that LOX expression was downregulated in GC, and LOX functioned as a tumor suppressor (Kaneda et al., 2004). Therefore, LOX function in GC needs to be further explored.

Relatively, few data are available on the role of LOXL1 in tumorigenesis. LOXL1 was overexpressed in GC cells, and high LOXL1 expression is a poor prognostic factor in GC patients (Kasashima et al., 2018). Moreover, LOXL1 is associated with peritoneal dissemination, potentially via promoting EMT in GC cells, and high LOXL1 expression was associated with poorly differentiated histological type, lymph node metastasis, and poor prognosis in GC (Hu Q. et al., 2020). Our study also revealed that LOXL1 is highly expressed in GC and may be related to the prognosis, although the results of the two survival databases are inconsistent.

It is first reported that LOXL2 promotes tumor progression and is associated with poor prognosis in breast cancer (Akiri et al., 2003). LOXL2 was overexpressed in GC versus normal tissues, and overexpression of LOXL2 was associated with depth of tumor invasion, lymph node metastasis and poorer overall survival (Peng et al., 2009). Furthermore, secreted LOXL2 promotes GC metastasis via Src kinase/Focal adhesion kinase (Src/FAK) pathway (Peng et al., 2009). LOXL2 expression in stromal cells was significantly associated with tumor invasion depth, lymph node metastasis, lymphatic invasion, venous invasion, peritoneal dissemination, and survival in GC patients (Kasashima et al., 2014). In our report, we illustrated that the expression of LOXL2 in GC tissues was higher than that in normal tissues, and this expression was markedly correlated with tumor stage and poor OS in patients with GC, which was consistent with reports above.

Even though LOXL3 expression was also detected in some kinds of tumors, studies have been conducted on LOXL3 were fewer. The expression of LOXL3 was detected mainly in the nucleus, and the expression of LOXL3 was correlated with tumor invasion, lymph node metastasis, and poorer prognosis of

patients (Kasashima et al., 2018). Additionally, TGF-induced LOXL3 upregulation in GC cells, suggesting that LOXL3 was downstream from the TGF-signaling pathway (Kasashima et al., 2018). Our research showed that high expression of LOXL3 was confirmed in GC of TCGA data by using UALCAN database, and the expression was correlated with tumor stage, while there was no difference in the results of Oncomine and GEPIA. The survival analysis also verified that high expression was related to poor prognosis, we speculate that the prognosis of tumors are related to a variety of factors and this may be related to the target of certain drugs in the treatment of gastric cancer. Recently, more studies about LOXL3 have been published, the roles attributed to LOXL3 should be further determined.

LOXL4 was significantly up-regulated in gastric carcinoma tissues, and this over-expression is significantly correlated with tumor size, depth of tumor invasion, lymph node metastasis, higher TNM stages and poor prognosis (Li et al., 2015). LOXL4 may promote proliferation and metastasis via regulate FAK/Src pathway in GC cells (Li et al., 2015). The results about LOXL4 conducted by Kasashima et al. were similar with LOXL3 (Kasashima et al., 2018). In contrast, comprehensive bioinformatics analysis of multiple databases in our study did not yield positive results in expression of LOXL4. In other aspects, survival analysis showed that LOXL4 is associated with poor prognosis, suggesting that LOXL4 is implicated in the progression of GC. This prognostic-related reason may be the same as LOXL3.

In the functional enrichment analysis of genes co-expressed with LOX/LOXL2, the most significant biological process of LOX/LOXL2 and their co-expressed genes is ECM. ECM plays a key role in the occurrence and metastasis of gastric cancer. The destruction of the tightly coordinated ECM tissue will damage the structure and function of the gastric tissue, eventually leading to the progression of gastric cancer (Moreira et al., 2020). We speculate that LOX and LOXL2 affect the occurrence and development of gastric cancer by participating in the regulation of extracellular matrix, but further research is still needed.

Due to the secreted nature of the LOX family members, their detectable presence in the blood, and the well-established correlation between LOX family enzyme expression and prognosis in many cancers, the LOX family offers promise as an inexpensive and non-invasive companion biomarker for cancers (Setargew et al., 2021). The LOX family of enzymes are favorable targets for anti-stromal therapeutics because of their importance in cancer development and progression. A number of studies have examined the use of LOX family inhibitors in cancer therapy (Jiang et al., 2020; Smithen et al., 2020). CCT365623 is a LOX inhibitor based on methylaminothiophene. It has shown that its inhibitory effect can lead to delayed tumor development and reduced lung metastasis in mouse breast cancer models. But it has not yet been tested in a clinical (Smithen et al., 2020). XS-5382A, an oral LOXL2 inhibitor, has been shown to slow tumor growth and reduce collagen accumulation in LY2 oral cancer models and is currently being investigated in Phase 1 clinical trials in healthy adults (Clinical trial identifier: NCT04183517) (Mahjour et al., 2019). Although no inhibitors of the LOX family have currently been approved for routine clinical practice, the developing LOX family inhibitors have shown high specificity and low toxicity.

However, there are limitations in our research. Bioinformatics analysis alone cannot determine the specific mechanism of LOX family in GC. The role of the LOX family in GC might be complex, and more clinical studies and in-depth experiments are needed to verify the diagnostic value of these LOX family and explore the potential mechanism of LOX family affecting the development of GC.

CONCLUSION

In this study, through systematically analyzing the expression and prognostic value of LOX family in GC, we indicated that the LOX family, especially LOX and LOXL2, might play an important role in GC oncogenesis, and they may become biomarkers for predicting tumor prognosis and potential targets for tumor therapy.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

LW and GS developed the idea, designed the research, and performed the data analysis work. SC and RZ reorganized the data and edited the manuscript and GS and YZ reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by National Natural Science Foundation of China (Grants 91959114, 81872106), Natural Science Foundation of Tianjin (19JCZDJC33900).

REFERENCES

Akiri, G., Sabo, E., Dafni, H., Vadasz, Z., Kartvelishvily, Y., Gan, N., et al. (2003). Lysyl Oxidase-Related Protein-1 Promotes Tumor Fibrosis and Tumor Progression *In Vivo*. *Cancer Res.* 63 (7), 1657–1666.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer J. Clinicians* 68 (6), 394–424. doi:10.3322/caac.21492

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring

Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discov.* 2 (5), 401–404. doi:10.1158/2159-8290.CD-12-0095

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V. S. K., et al. (2017). UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Chau, I. (2017). Checkpoint Inhibition: an ATTRACTION in Advanced Gastric Cancer? *The Lancet* 390 (10111), 2418–2419. doi:10.1016/S0140-6736(17)32131-1

Chen, X., Leung, S. Y., Yuen, S. T., Chu, K.-M., Ji, J., Li, R., et al. (2003). Variation in Gene Expression Patterns in Human Gastric Cancers. *MBoC* 14 (8), 3208–3215. doi:10.1091/mbc.e02-12-0833

Chitty, J. L., Setargew, Y. F. I., and Cox, T. R. (2019). Targeting the Lysyl Oxidases in Tumour Desmoplasia. *Biochem. Soc. Trans.* 47 (6), 1661–1678. doi:10.1042/BST20190098

Cui, J., Chen, Y., Chou, W.-C., Sun, L., Chen, L., Suo, J., et al. (2011). An Integrated Transcriptomic and Computational Analysis for Biomarker Identification in Gastric Cancer. *Nucleic Acids Res.* 39 (4), 1197–1207. doi:10.1093/nar/gkq960

D'Errico, M., Rinaldis, E. d., Blasi, M. F., Viti, V., Falchetti, M., Calcagnile, A., et al. (2009). Genome-wide Expression Profile of Sporadic Gastric Cancers with Microsatellite Instability. *Eur. J. Cancer* 45 (3), 461–469. doi:10.1016/j.ejca.2008.10.032

Hu, L., Wang, J., Wang, Y., Wu, L., Wu, C., Mao, B., et al. (2020). LOXL1 Modulates the Malignant Progression of Colorectal Cancer by Inhibiting the Transcriptional Activity of YAP. *Cell Commun Signal* 18 (1), 148. doi:10.1186/s12964-020-00639-1

Hu, Q., Masuda, T., Kuramitsu, S., Tobo, T., Sato, K., Kidogami, S., et al. (2020). Potential Association of LOXL1 with Peritoneal Dissemination in Gastric Cancer Possibly via Promotion of EMT. *PLoS One* 15 (10), e0241140. doi:10.1371/journal.pone.0241140

Jiang, H., Torphy, R. J., Steiger, K., Hongo, H., Ritchie, A. J., Kriegsmann, M., et al. (2020). Pancreatic Ductal Adenocarcinoma Progression Is Restrained by Stromal Matrix. *J. Clin. Invest.* 130 (9), 4704–4709. doi:10.1172/JCI136760

Kaneda, A., Wakazono, K., Tsukamoto, T., Watanabe, N., Yagi, Y., Tatematsu, M., et al. (2004). Lysyl Oxidase Is a Tumor Suppressor Gene Inactivated by Methylation and Loss of Heterozygosity in Human Gastric Cancers. *Cancer Res.* 64 (18), 6410–6415. doi:10.1158/0008-5472.CAN-04-1543

Kang, M.-H., Choi, H., Oshima, M., Cheong, J.-H., Kim, S., Lee, J. H., et al. (2018). Estrogen-related Receptor Gamma Functions as a Tumor Suppressor in Gastric Cancer. *Nat. Commun.* 9 (1), 1920. doi:10.1038/s41467-018-04244-2

Kasashima, H., Yashiro, M., Kinoshita, H., Fukuoka, T., Morisaki, T., Masuda, G., et al. (2016). Lysyl Oxidase Is Associated with the Epithelial-Mesenchymal Transition of Gastric Cancer Cells in Hypoxia. *Gastric Cancer* 19 (2), 431–442. doi:10.1007/s10120-015-0510-3

Kasashima, H., Yashiro, M., Kinoshita, H., Fukuoka, T., Morisaki, T., Masuda, G., et al. (2014). Lysyl Oxidase-like 2 (LOXL2) from Stromal Fibroblasts Stimulates the Progression of Gastric Cancer. *Cancer Lett.* 354 (2), 438–446. doi:10.1016/j.canlet.201410.1016/j.canlet.2014.08.014

Kasashima, H., Yashiro, M., Okuno, T., Miki, Y., Kitayama, K., Masuda, G., et al. (2018). Significance of the Lysyl Oxidase Members Lysyl Oxidase like 1, 3, and 4 in Gastric Cancer. *Digestion* 98 (4), 238–248. doi:10.1159/000489558

Lai, H., Jin, Q., Lin, Y., Mo, X., Li, B., He, K., et al. (2014). Combined Use of Lysyl Oxidase, Carcino-Embryonic Antigen, and Carbohydrate Antigens Improves the Sensitivity of Biomarkers in Predicting Lymph Node Metastasis and Peritoneal Metastasis in Gastric Cancer. *Tumor Biol.* 35 (10), 10547–10554. doi:10.1007/s13277-014-2355-5

Li, Q., Zhu, C.-C., Ni, B., Zhang, Z.-Z., Jiang, S.-H., Hu, L.-P., et al. (2019). Lysyl Oxidase Promotes Liver Metastasis of Gastric Cancer via Facilitating the Reciprocal Interactions between Tumor Cells and Cancer Associated Fibroblasts. *EBioMedicine* 49, 157–171. doi:10.1016/j.ebiom.2019.10.037

Li, R.-k., Zhao, W.-y., Fang, F., Zhuang, C., Zhang, X.-x., Yang, X.-m., et al. (2015). Lysyl Oxidase-like 4 (LOXL4) Promotes Proliferation and Metastasis of Gastric Cancer via FAK/Src Pathway. *J. Cancer Res. Clin. Oncol.* 141 (2), 269–281. doi:10.1007/s00432-014-1823-z

Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* 173 (2), 400–416.e11. doi:10.1016/j.cell.2018.02.052

Mahjour, F., Dambal, V., Shrestha, N., Singh, V., Noonan, V., Kantarci, A., et al. (2019). Mechanism for Oral Tumor Cell Lysyl Oxidase Like-2 in Cancer Development: Synergy with PDGF-AB. *Oncogenesis* 8 (5), 34. doi:10.1038/s41389-019-0144-0

Molnar, J., Fong, K. S. K., He, Q. P., Hayashi, K., Kim, Y., Fong, S. F. T., et al. (2003). Structural and Functional Diversity of Lysyl Oxidase and the LOX-like Proteins. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1647, 220–224. doi:10.1016/s1570-9639(03)00053-0

Moreira, A. M., Pereira, J., Melo, S., Fernandes, M. S., Carneiro, P., Seruca, R., et al. (2020). The Extracellular Matrix: An Accomplice in Gastric Cancer Development and Progression. *Cells* 9 (2), 394. doi:10.3390/cells9020394

Peng, C., Liu, J., Yang, G., and Li, Y. (2018). Lysyl Oxidase Activates Cancer Stromal Cells and Promotes Gastric Cancer Progression: Quantum Dot-Based Identification of Biomarkers in Cancer Stromal Cells. *Int. J. Nanomedicine* 13, 161–174. doi:10.2147/IJN.S143871

Peng, L., Ran, Y.-L., Hu, H., Yu, L., Liu, Q., Zhou, Z., et al. (2009). Secreted LOXL2 Is a Novel Therapeutic Target that Promotes Gastric Cancer Metastasis via the Src/FAK Pathway. *Carcinogenesis* 30 (10), 1660–1669. doi:10.1093/carcin/bgp178

Rachman-Tzemah, C., Zaffryar-Eilot, S., Grossman, M., Ribero, D., Timaner, M., Mäki, J. M., et al. (2017). Blocking Surgically Induced Lysyl Oxidase Activity Reduces the Risk of Lung Metastases. *Cel. Rep.* 19 (4), 774–784. doi:10.1016/j.celrep.2017.04.005

Rhodes, D. R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J., Briggs, B. B., et al. (2007). Oncomine 3.0: Genes, Pathways, and Networks in a Collection of 18,000 Cancer Gene Expression Profiles. *Neoplasia* 9 (2), 166–180. doi:10.1593/neo.07112

Salvador, F., Martin, A., López-Menéndez, C., Moreno-Bueno, G., Santos, V., Vázquez-Naharro, A., et al. (2017). Lysyl Oxidase-like Protein LOXL2 Promotes Lung Metastasis of Breast Cancer. *Cancer Res.* 77 (21), 5846–5859. doi:10.1158/0008-5472.CAN-16-3152

Setargew, Y. F. I., Wyllie, K., Grant, R. D., Chitty, J. L., and Cox, T. R. (2021). Targeting Lysyl Oxidase Family Meditated Matrix Cross-Linking as an Anti-stromal Therapy in Solid Tumours. *Cancers* 13 (3), 491. doi:10.3390/cancers13030491

Shao, J., Lu, J., Zhu, W., Yu, H., Jing, X., Wang, Y.-L., et al. (2019). Derepression of LOXL4 Inhibits Liver Cancer Growth by Reactivating Compromised P53. *Cell Death Differ* 26 (11), 2237–2252. doi:10.1038/s41418-019-0293-x

Smithen, D. A., Leung, L. M. H., Challinor, M., Lawrence, R., Tang, H., Niculescu-Duvaz, D., et al. (2020). 2-Aminomethylene-5-sulfonylthiazole Inhibitors of Lysyl Oxidase (LOX) and LOXL2 Show Significant Efficacy in Delaying Tumor Growth. *J. Med. Chem.* 63 (5), 2308–2324. doi:10.1021/acs.jmedchem.9b01112

Szász, A. M., Lánczky, A., Nagy, Á., Förster, S., Hark, K., Green, J. E., et al. (2016). Cross-validation of Survival Associated Biomarkers in Gastric Cancer Using Transcriptomic Data of 1,065 Patients. *Oncotarget* 7 (31), 49322–49333. doi:10.18632/oncotarget.10337

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

Wang, Q., Wen, Y.-G., Li, D.-P., Xia, J., Zhou, C.-Z., Yan, D.-W., et al. (2012). Upregulated INHBA Expression Is Associated with Poor Survival in Gastric Cancer. *Med. Oncol.* 29 (1), 77–83. doi:10.1007/s12032-010-9766-y

Wang, T.-H., Hsia, S.-M., and Shieh, T.-M. (2016). Lysyl Oxidase and the Tumor Microenvironment. *Ijms* 18 (1), 62. doi:10.3390/ijms18010062

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

Xiao, Q., and Ge, G. (2012). Lysyl Oxidase, Extracellular Matrix Remodeling and Cancer Metastasis. *Cancer Microenvironment* 5 (3), 261–273. doi:10.1007/s12307-012-0105-z

Yamauchi, M., Barker, T. H., Gibbons, D. L., and Kurie, J. M. (2018a). The Fibrotic Tumor Stroma. *J. Clin. Invest.* 128 (1), 16–25. doi:10.1172/JCI93554

Yamauchi, M., Taga, Y., Hattori, S., Shiiba, M., and Terajima, M. (2018b). Analysis of Collagen and Elastin Cross-Links. *Methods Cel. Biol.* 143, 115–132. doi:10.1016/bs.mcb.2017.08.006

Zeltz, C., Pasko, E., Cox, T. R., Navab, R., and Tsao, M.-S. (2019). LOXL1 Is Regulated by Integrin α11 and Promotes Non-small Cell Lung Cancer Tumorigenicity. *Cancers* 11 (5), 705. doi:10.3390/cancers11050705

Zhang, Q., Jin, X.-S., Yang, Z.-Y., Wei, M., Zhu, X.-C., Wang, P., et al. (2013). Upregulated Expression of LOX Is a Novel Independent Prognostic Marker of Worse Outcome in Gastric Cancer Patients after Curative Surgery. *Oncol. Lett.* 5 (3), 896–902. doi:10.3892/ol.2012.1092

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6

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: 07 March 2022
doi: 10.3389/fgene.2022.805961

Clinical Significance and Potential Mechanisms of ATP Binding Cassette Subfamily C Genes in Hepatocellular Carcinoma

Xin Zhou[1,2†], *Jia-mi Huang*[1,2†], *Tian-man Li*[1,3], *Jun-qi Liu*[1,2], *Zhong-liu Wei*[1,2], *Chen-lu Lan*[1,2], *Guang-zhi Zhu*[1,2], *Xi-wen Liao*[1,2], *Xin-ping Ye*[1,2] *and Tao Peng*[1,2*]

[1]*Department of Hepatobiliary Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, China,* [2]*Key Laboratory of Early Prevention and Treatment for Regional High Frequency Tumor (Guangxi Medical University), Ministry of Education, Nanning, China,* [3]*Department of Hepatobiliary Surgery, The Sixth Affiliated Hospital of Guangxi Medical University, Yulin, China*

OPEN ACCESS

Edited by:
Frank Emmert-Streib,
Tampere University, Finland

Reviewed by:
Asli Suner,
Ege University, Turkey
Jiyou Han,
Hyupsung University, South Korea
Yi Bai,
Tianjin First Central Hospital, China

***Correspondence:**
Tao Peng
pengtaogmu@163.com
orcid.org/0000-0001-6133-7078

†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: *31 October 2021*
Accepted: *15 February 2022*
Published: *07 March 2022*

Citation:
Zhou X, Huang J-m, Li T-m, Liu J-q, Wei Z-l, Lan C-l, Zhu G-z, Liao X-w, Ye X-p and Peng T (2022) Clinical Significance and Potential Mechanisms of ATP Binding Cassette Subfamily C Genes in Hepatocellular Carcinoma. Front. Genet. 13:805961. doi: 10.3389/fgene.2022.805961

The purpose of this investigation was to assess the diagnostic and prognostic significance of *ATP binding cassette subfamily C (ABCC)* genes in hepatocellular carcinoma (HCC). The Student t-test was used to compare the expression level of *ABCCs* between HCC and paraneoplastic tissues. Receiver operating characteristic curve (ROC) analysis was applied for diagnostic efficiency assessment. The Kaplan–Meier method and Cox proportional hazards model were respectively applied for survival analysis. Genes with prognostic significance were subsequently used to construct prognostic models. From the perspective of genome-wide enrichment analysis, the mechanisms of prognosis-related ABCC genes were attempted to be elaborated by gene set enrichment analysis (GSEA). It was observed in the TCGA database that *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC10* were significantly upregulated in tumor tissues, while *ABCC6* and *ABCC7* were downregulated in HCC tissues. Receiver operating characteristic analysis revealed that *ABCC7* might be a potential diagnostic biomarker in HCC. *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* were significantly related to the prognosis of HCC in the TCGA database. The prognostic significance of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* was also observed in the Guangxi cohort. In the Guangxi cohort, both polymerase chain reaction and IHC (immunohistochemical) assays demonstrated higher expression of ABCC1, ABCC4, and ABCC5 in HCC compared to liver tissues, while the opposite was true for ABCC6. GSEA analysis indicated that *ABCC1* was associated with tumor differentiation, nod-like receptor signal pathway, and so forth. It also revealed that *ABCC4* might play a role in HCC by regulating epithelial-mesenchymal transition, cytidine analog pathway, met pathway, and so forth. *ABCC5* might be associated with the fatty acid metabolism and KRT19 in HCC. *ABCC6* might impact the cell cycle in HCC by regulating E2F1 and myc. The relationship between ABCC genes and immune infiltration was explored, and ABCC1,4,5 were found to be positively associated with infiltration of multiple immune cells, while ABCC6 was found to be

Abbreviations: ABCC, ATP binding cassette subfamily C; HCC, hepatocellular carcinoma; ROC, receiver operating characteristic curve; HBV, hepatitis B virus; HCV, hepatitis C virus; AUC, area under curve; GSEA, gene set enrichment analysis; IHC, Immunohistochemical.

the opposite. In conclusion, *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* might be prognostic biomarkers in HCC. The prognostic models constructed with *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* had satisfactory efficacy.

Keywords: HCC (hepatic cellular carcinoma), ABCC gene family, prognosis (carcinoma), nomogram, GSEA (gene set enrichment analysis)

BACKGROUND

Hepatocellular carcinoma (HCC) generally followed cirrhosis given rise by metabolic disorder (Yang et al., 2019), chronic ethanol intake (Llovet et al., 2016), and hepatitis virus infection (Fujiwara et al., 2018). The leading metabolic risk factor for HCC is non-alcoholic fatty liver disease (NAFLD) (Zhang, 2018), which is mainly related to obesity and type 2 diabetes. Currently, it is acknowledged that hepatitis B virus (HBV) and hepatitis C virus (HCV) were the main infectious etiologies for cirrhosis and HCC. It is worth mentioning that viral hepatitis could skip cirrhosis and induced HCC directly and independently (El-Serag, 2012; Levrero and Zucman-Rossi, 2016). Besides the aforementioned factors, the intake of Aflatoxin B1 (AFB1) was also demonstrated to be related to HCC (Rushing and Selim, 2019). People in specific regions entailing relatively high exposure to Aflatoxin B1 were accompanied by high incidence and mortality of HCC (Long et al., 2008; Wogan et al., 2012; Zhang W et al., 2017). More than 8 million new cases of liver cancer occurred worldwide each year, which directly or indirectly gave rise to more than 4 million deaths worldwide each year (Torre et al., 2015; Bray et al., 2018). Asia has the highest incidence of liver cancer in the world, particularly in China, which accounts for almost half the global cases (Akinyemiju et al., 2017). At the same time, Asia is the high-incidence area of HBV and HCV (Gower et al., 2014; Polaris Observatory Collaborators, 2018). The main treatment methods of liver cancer mainly include surgical resection, transcatheter arterial chemoembolization (TACE), ablation, liver transplantation, radiotherapy, and so forth (Fattovich et al., 2004). Sorafenib, the multi-kinase inhibitor, is one of first-line drugs approved for the treatment of advanced HCC. Although it can improve survival, the long-term survival of HCC patients is limited due to the drug resistance. Hence, the discovery of new hub genes for developing HCC-targeted drugs and specific genes that improve and maintain drug susceptibility might be hopeful for advanced-stage HCC patients.

The ATP binding cassette subfamily C (ABCC) subfamily includes 13 members whose protein products take effect in transporters with different functional profiles, including ion transport, cell surface receptor, and toxin secretion activity (Childs and Ling, 1994; Dean and Allikmets, 2001; Robey et al., 2018; Yamada et al., 2018). The ATP-binding domain of the ABCC product possesses distinctive conserved motifs (Walker A and B motifs), which are separated by an uncertain sequence of around 100 amino acids (Dean et al., 2001). The distinctive interval and conserved motifs distinguish ABCC members from other ATP-binding proteins (Higgins et al., 1986). Genetic variations in these genes are substantiated in numerous research studies to be the cause or contributor to a variety of complex human diseases, including cystic fibrosis, neurological diseases, defects in cholesterol and bile transport, and drug responses. The ABCC subfamily plays an important role in the pharmacokinetics of endogenous and exogenous compounds. Studies have shown that the members of the ABCC family could transport drugs to the extracellular substances by virtue of ATP energy (Chen and Tiwari, 2011; Keppler, 2011; Leslie, 2012).

METHODS

Data Acquisition and Specimen Collection

RNA-Seq data (FPKM) of 412 samples, 362 tumors, and 50 paraneoplastic tissues were acquired from the TCGA database (https://portal.gdc.cancer.gov/, accessed on 22 December 2019). The *limma* package was employed for normalization of this RNA-Seq data in R. Matched prognostic/clinicopathologic data of these 362 patients were acquired from UCSC Xena (http://xena.ucsc.edu/, accessed on 23 December 2019).

The HCC tissues and matched paracancer tissues of 102 patients hospitalized in the first affiliated hospital of Guangxi Medical University from September 2016 to December 2018 were collected after informed consent was obtained. Among them, excised tissues during surgery of 72 patients were well preserved in the Department of Pathology. Tissue slices of these patients were obtained from the Department of Pathology.

Expression Difference and Diagnostic Efficiency Analysis of ABCC Genes

The expression levels of *ABCCs* in HCC and paraneoplastic tissues were extracted from the RNA-Seq Chip matrix in the TCGA database. The normality test was assessed using the Kolmogorov–Smirnov normality test. Student's t-test was used to assess the statistical significance of *ABCC* genes' expression between HCC and paraneoplastic tissues. The area under the curve (AUC) of the receiver operating characteristic curve (ROC) was used to access the diagnostic efficiency of each *ABCC* gene in HCC. AUC > 0.8 with $p < 0.05$ was considered as satisfactory diagnostic performance (Hosmer et al., 2013).

Immunohistochemistry

Tissue sections were sequentially placed in xylene and graded concentrations of ethanol to achieve hydration. Antigens were repaired with a pH 6.0 citrate repair solution (ZSGB-BIO, Beijing, China). Subsequent antigen–antibody reactions and color development reactions were performed with the help of a universal two-step detection kit (Mouse/Rabbit Enhanced Polymer Detection System). Immunohistochemical scores were assessed by two experienced pathologists. Antibodies for ABCC1, ABCC4,

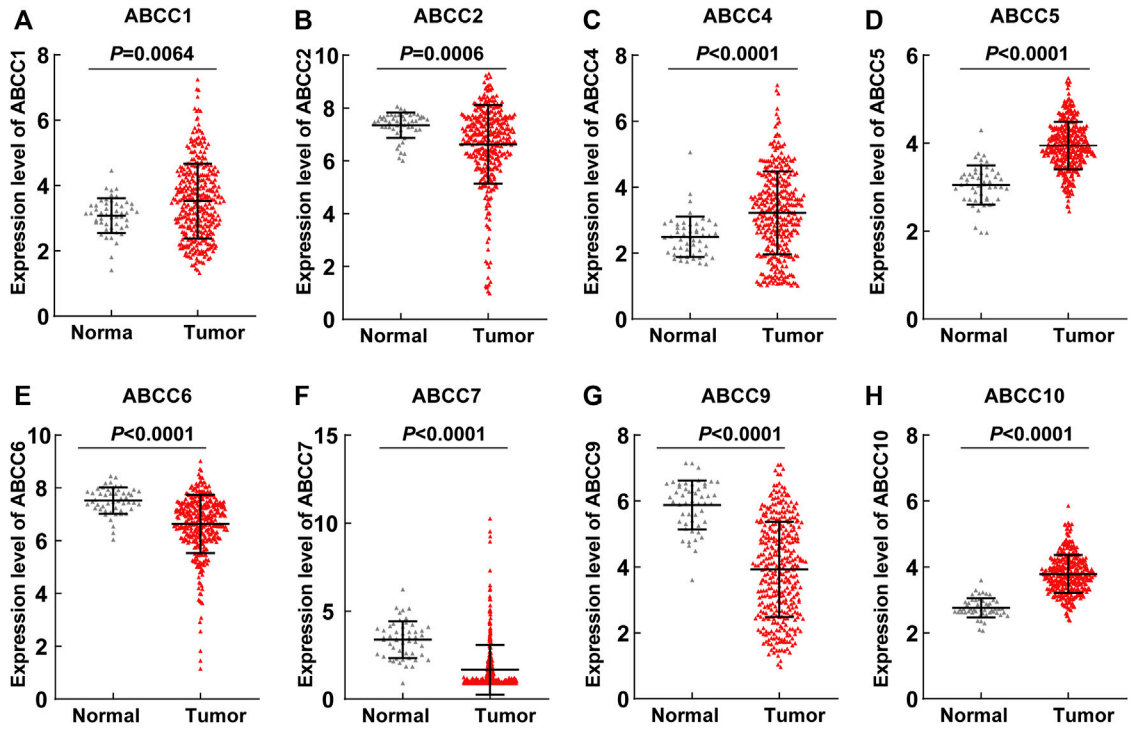


FIGURE 1 | *ABCC1*, *ABCC2*, *ABCC4*, *ABCC5*, *ABCC6*, *ABCC7*, *ABCC9*, and *ABCC10* were differently expressed between the HCC tissues and paraneoplastic tissues based on the RNA-seq data of the TCGA database: **(A)** *ABCC1*, **(B)** *ABCC2*, **(C)** *ABCC4*, **(D)** *ABCC5*, **(E)** *ABCC6*, **(F)** *ABCC7*, **(G)** *ABCC9*, and **(H)** *ABCC10*.

ABCC5, and ABCC6 were diluted according to the recommended concentrations of the manufacturer (Proteintech, Wuhan, China).

Prognostic Significance Assessment of ABCC Genes

The patients in the TCGA database were divided into two groups in terms of the median value of each *ABCC* gene expression for survival analysis. The Kaplan–Meier method with a log-rank test was applied to assess the prognostic significance of each *ABCC* gene. The Cox proportional hazards model was applied to adjust the bias caused by prognosis-related clinicopathologic factors.

In terms of survival analysis results in the TCGA database, the prognostic significance of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* was further validated in the Guangxi cohort.

For better predicting the prognosis and evaluating the combined effect of *ABCCs*, prognosis-related *ABCCs* (*ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6*) were integrated in pairs into combined effect survival analysis. The patients were divided into four groups in terms of the expression level of *ABCCs* with details displayed in **Table 2**. The Kaplan–Meier method with the log-rank test and Cox proportional hazards model were applied to assess the prognostic significance.

Nomogram

Independent prognostic factors, including *ABCCs* and clinicopathologic features, were integrated to construct the nomogram in R with the *rms* package (Iasonos et al., 2008). In the nomogram, the risk degree of each variable in the nomogram was displayed by the integration line, and the total risk score is obtained by adding up the risk value of each variable (Zhang Z et al., 2017). The model was validated for calibration and discrimination using the *bootstrap* method (Wang et al., 2013).

Prognostic Signature Construction

The Cox proportional hazards model was used to assess the risk coefficient of *ABCCs* in overall survival. Then, the prognostic signature was constructed in terms of the expression of *ABCCs* and the corresponding risk coefficient. The formula of prognostic signature construction is as follows: Risk score = $\sum_{1}^{N}(ExpVluei*\beta i)$ (Chen M et al., 2017). N is the number of prognostic genes. ExpVluei is the expression value of each *ABCC* gene. βi is the risk coefficient of the corresponding *ABCC* gene. A time-dependent ROC curve was constructed in R (version 3.6.2; www.r-project.org) with the *survivalROC* package to evaluate the availability of this prognostic signature (Chen M et al., 2017).

Biological Functional Exploration of ABCC Genes

The Gene Ontology (GO) database, the integrated database of calculable information about the functions of genes, was comprehensively used for identifying unique biological properties of high-throughput transcriptome or genome data (The Gene Ontology Consortium, 2017; Chen L et al., 2017). KEGG is a collection of databases dealing with genomes, diseases,

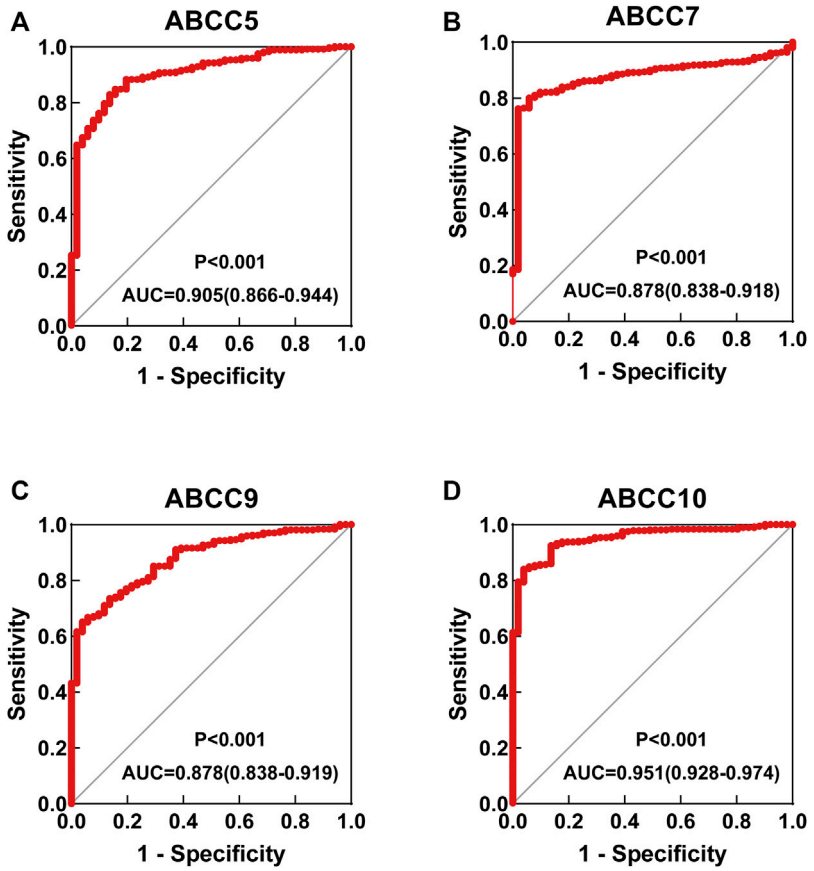


FIGURE 2 | ROC curve for ABCC genes in the TCGA database: **(A)** *ABCC5*, **(B)** *ABCC7*, **(C)** *ABCC9*, and **(D)** *ABCC10*.

biological pathways, drugs, and chemical materials (Kanehisa et al., 2017). DAVID (The Database for Annotation, Visualization, and Integrated Discovery, https://david.ncifcrf.gov/) is an online bioinformatics tool to access the GO database and the KEGG database (Long et al., 2008). DAVID was used to access the enrichment of biological functions and pathways of ABCC genes in this investigation. Then, the enrichment biological functions and pathways were visualized in R Studio (Version 1. 2.5033) with packages *Goplot*, *Hmisc*, and *ggplot2* (Nolan et al., 2013; Ito and Murphy, 2013). The Biological Networks Gene Ontology tool (BiNGO) is an open-source online database, which was employed to determine the significantly overrepresented GO terms of ABCC genes (Maere et al., 2005). Functions and interactions of ABCC genes were performed in Genemania (http://genemania.org/. accessed on 11 August 2020) and STRING (https://string-db.org/. accessed on 11 August 2020), respectively (Szklarczyk et al., 2015; Luo et al., 2020).

Gene Set Enrichment Analysis (GSEA)

GSEA is software with additional resources for analyzing, annotating, and interpreting standardized chip matrices. In this investigation, GSEA enrichment was used to analyze the enriched biological pathways of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* in the TCGA database. The Oncogenic Signatures c2.all. v7.1.symbols.gmt data set was adopted as the reference data set. The biological pathways exported from GSEA with $p < 0.05$ and FDR < 0.25 were considered as significant results.

Correlation Analysis of Tumor-Infiltrating Immune Cells and ABCC Gene Expression

TIMER (http://timer.cistrome.org/) is a comprehensive resource for the systematical analysis of immune infiltrates across diverse cancer types, which provides immune infiltrates abundances estimated by multiple immune deconvolution methods. In this investigation, TIMER was accessed to explore the correlation between infiltrating immune cells and *ABCC* expression in HCC.

RESULTS

Expression and Diagnostic Efficiency of ABCC Genes in HCC

Several *ABCC* genes were discovered to be differentially expressed in HCC and paraneoplastic tissues based on the RNA-seq data of the TCGA database. *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC10* (**Figures 1A,C,D,H**) were significantly higher expressed in HCC tissues, but *ABCC2*, *ABCC6*, *ABCC7*, and *ABCC9* (**Figures**

TABLE 1 | Survival analysis results of ABCC genes in the TCGA database.

Gene expression	Patients (*n* = 362)	Overall survival				
		Number of events	Crude HR (95% CI)	Crude P	Adjusted HR (95% CI)	Adjusted P §
ABCC1						
Low	181	51	1		1	
High	181	78	1.759 (0.235–2.504)	0.002	1.656 (1.137–2.410)	0.008
ABCC2						
Low	181	64	1		1	
High	181	65	1.079 (0.761–1.529)	0.670	1.210 (0.833–1.758)	0.317
ABCC3						
Low	181	64	1		1	
High	181	65	1.018 (0.718–1.443)	0.919	0.863 (0.595–1.422)	0.438
ABCC4						
Low	181	57	1		1	
High	181	72	1.489 (1.046–2.121)	0.026	1.479 (1.021–2.142)	0.038
ABCC5						
Low	181	50	1		1	
High	181	79	1.759 (0.234–2.508)	0.002	1.928 (1.318–2.820)	0.001
ABCC6						
Low	181	81	1		1	
High	181	48	0.495 (0.346–0.708)	<0.001	0.534 (0.366–0.778)	0.001
ABCC7						
Low	181	63	1		1	
High	181	66	1.185 (0.836–1.680)	0.340	1.077 (0.743–1.562)	0.695
ABCC8						
Low	181	57	1		1	
High	181	72	1.306 (0.920–1.853)	0.134	1.227 (0.843–1.785)	0.286
ABCC9						
Low	181	71	1		1	
High	181	58	0.757 (0.535–1.072)	0.116	0.794 (0.549–1.149)	0.221
ABCC10						
Low	181	60	1		1	
High	181	69	1.283 (0.907–1.815)	0.157	1.275 (0.883–1.841)	0.195
ABCC11						
Low	181	67	1		1	
High	181	62	0.893 (0.632–1.263)	0.523	0.832 (0.575–1.205)	0.331
ABCC12						
Low	181	58	1		1	
High	181	71	1.308 (0.923–1.852)	0.130	1.287 (0.891–1.859)	0.179
ABCC13						
Low	181	60	1		1	
High	181	69	1.313 (0.928–1.857)	0.112	1.261 (0.871–1.823)	0.220

Notes: § Adjusted for tumor stage. HR, hazard ratio; ABCC, ATP binding cassette subfamily C.

1B,E–G) were significantly lower expressed in HCC tissues. No significant differences in the expression of ABCC3, ABCC8, ABCC11, ABCC12, and ABCC13 were observed between liver and HCC tissues (**Supplementary Figures S1A–E**). The diagnostic efficacy of the genes differentially expressed between HCC and paraneoplastic tissues was subsequently evaluated using ROC curve analysis. Among them (*ABCC1*, *ABCC2*, *ABCC4*, *ABCC5*, *ABCC6*, *ABCC7*, *ABCC9*, and *ABCC10*), high diagnostic efficiencies of *ABCC5* (AUC = 0.905, $p < 0.001$), *ABCC7* (AUC = 0.878, $p < 0.001$), *ABCC9* (AUC = 0.878, $p < 0.001$), and *ABCC10* (AUC = 0.951, $p < 0.001$) (**Figures 2A–D**) were observed in HCC. The results of the diagnostic efficacy analysis of other ABCC genes are shown in **Supplementary Figures S2A–I**.

Prognostic Significance of ABCC Genes

Subsequently, the prognostic significance of *ABCC* genes was systematically discussed. The clinicopathologic characteristics of 362 HCC tissues in the TCGA database are displayed in **Supplementary Table S1**. The expression levels of *ABCC1* (log-rank $p = 0.002$, adjusted $p = 0.008$, adjusted HR = 1.656), *ABCC4* (log-rank $p = 0.026$, adjusted $p = 0.038$, adjusted HR = 1.479), *ABCC5* (log-rank $p = 0.002$, adjusted $p = 0.001$, adjusted HR = 1.928), and *ABCC6* (log-rank $p < 0.001$, adjusted $p = 0.001$, adjusted HR = 0.534) were significantly associated with the overall survival of HCC patients in univariate and multivariate survival analysis (**Table 1**; **Figures 3A–D**). In terms of the prognostic value of a single ABCC gene, patients with high expression of *ABCC1*, *ABCC4*, or *ABCC5* tend to be with a shorter median survival time, while high-expression *ABCC6* was associated with longer survival. The results of survival analysis of other ABCC genes are shown in **Supplementary Figures S3A–I**.

To more accurately predict the prognosis of HCC patients, multivariate survival analysis was integrated into the combined

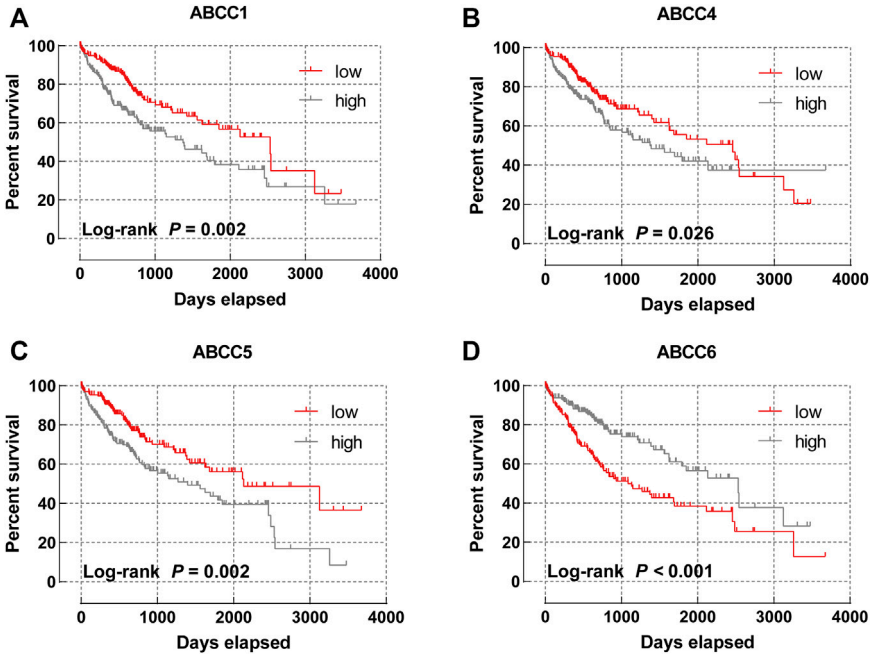


FIGURE 3 | Survival curve of ABCC genes in the TCGA database: **(A)** *ABCC1*, **(B)** *ABCC4*, **(C)** *ABCC5*, and **(D)** *ABCC6*.

TABLE 2 | Combined effect survival analysis of ABCC1, ABCC4, ABCC5, and ABCC6.

Group	ABCC1	ABCC4	ABCC5	ABCC6	Patients	No. of events	MST (days)	Crude HR (95%CI)	Crude P	Adjusted HR (95%CI)	Adjusted P δ
A	low	low			106	27	2,532	1		1	
B	low	high			150	54	2,116	1.474 (0.928–2.342)		1.500 (0.925–2.432)	
	high	low									
C	high	high			106	48	1,135	2.322 (1.445–3.731)	0.001	2.191 (1.328–3.614)	0.002
1	low		low		117	27	3,125	1		1	
2	low		high		128	47	1,685	1.564 (0.973–2.513)		1.637 (0.985–2.719)	
	high		low								
3	high		high		117	55	1,135	2.487 (1.568–3.945)	<0.001	2.572 (1.565–4.227)	<0.001
a	low			high	116	29	2,532	1		1	
b	low			low	130	41	1791	1.698 (1.053–2.739)		1.813 (1.098–2.993)	
	high			high							
c	high			low	116	59	931	2.574 (1.649–4.018)	<0.001	2.315 (1.447–3.704)	<0.001
I		low	low		100	23	3,125	1		1	
II		low	high		162	61	1852	1.767 (1.093–2.856)		1.897 (1.138–3.162)	
		high	low								
III		high	high		100	45	1,149	2.528 (1.522–4.201)	0.001	2.790 (1.618–4.813)	<0.001
i		low		high	97	21	2,542	1		1	
ii		low		low	168	63	1791	2.009 (1.221–3.307)		2.322 (1.379–3.910)	
		high		high							
iii		high		low	97	45	837	2.988 (1.771–5.042)	<0.001	2.792 (1.595–4.887)	<0.001
α			low	high	111	25	3,125	1		1	
β			low	low	140	48	1791	1.628 (1.004–2.641)		1.641 (0.983–2.739)	
			high	high							
γ			high	low	111	56	802	2.850 (1.777–4.571)	<0.001	2.939 (1.772–4.874)	<0.001

Notes: δ Adjusted for tumor stage. MST, median survival time; No. of events, number of events; HR, hazard ratio; ABCC, ATP binding cassette subfamily C.

effect survival analysis. In the combined effect survival analysis, it was observed that there was an even bigger prognosis difference among groups in combined effect survival analysis (**Table 2**). Group C with high expression of *ABCC1* and *ABCC4* was significantly correlated to bad outcome (p = 0.001, **Figure 4A**), so were group 3 with high expression of *ABCC1* and *ABCC5* (p <

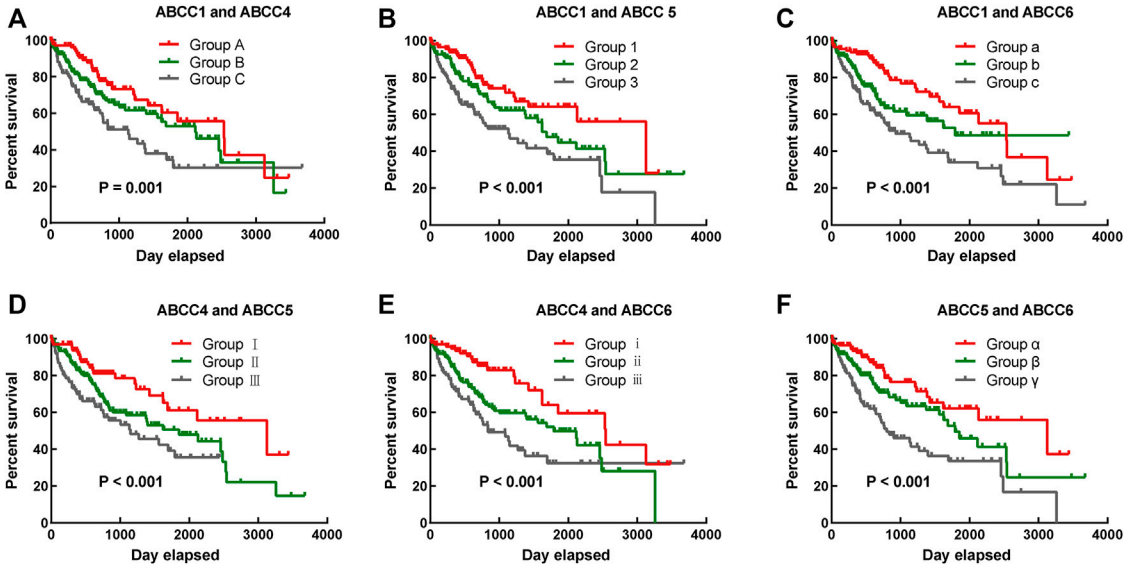


FIGURE 4 | Combined effect survival analysis of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* in the TCGA database: **(A)** combined effect survival analysis of *ABCC1* and *ABCC4*, **(B)** combined effect survival analysis of *ABCC1* and *ABCC5*, **(C)** combined effect survival analysis of *ABCC1* and *ABCC6*, **(D)** combined effect survival analysis of *ABCC4* and *ABCC5*, **(E)** combined effect survival analysis of *ABCC4* and *ABCC6*, and **(F)** combined effect survival analysis of *ABCC5* and *ABCC6*. The details about the groups are displayed in **Table 2**.

0.001, **Figure 4B**), group c with high expression of *ABCC1* and low expression of *ABCC4* ($p < 0.001$, **Figure 4C**), group Ⅲ with high expression of *ABCC4* and *ABCC5* ($p = 0.001$, **Figure 4D**), group ⅲ with high expression of *ABCC4* and low expression of *ABCC6* ($p < 0.001$, **Figure 4E**), and group γ with high expression of *ABCC5* and low expression of *ABCC6* ($p < 0.001$, **Figure 4F**).

Nomogram Based on *ABCC1*, *4*, *5*, and *6* and Tumor Stage

In the survival analysis, we found that ABCC1.4.5.6 was strongly associated with the prognosis of HCC. In addition, the clinical factor tumor stage could also partially distinguish patients with good and bad prognoses. Thus, a nomogram integrating clinical elements and ABCC gene expression was constructed in terms of the COX proportional hazards model. In the nomogram, the contribution of *ABCC1*, *ABCC4*, *ABCC5*, *ABCC6*, and clinicopathologic features to the overall survival of HCC patients was displayed by virtue of the length of the scales (**Figure 5A**). The calibration plot for 1-, 3-, and 5-year survival after surgery revealed a satisfactory overlap between calculation and reality (**Figures 5B–D**).

Prognostic Signature Based on the TCGA Database

In terms of the expressions of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6*, the prognostic signature for HCC patients was built in the TCGA database and Guangxi cohort. Each HCC patient was assigned with a risk score in terms of the expression of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6*. In the prognostic signature built for the TCGA database, the risk score for each patient was displayed in the upper scatter plot, and the patients were divided into two groups based on the median value (**Figure 5E**). The survival time and survival status of specific patients can be observed from the middle scatter plot, which showed that the dots representing patients in the high-risk group tended to cluster lower (**Figure 5F**). The expression levels of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* in patients were presented in the form of heat maps (**Figure 5G**). A significant difference in overall survival was observed between the high-risk and low-risk groups (**Figure 5H**, $p = 0.003$). The AUC value of the prognostic signature for 1-year, 3-year, and 5-year overall survival prediction was 0.689, 0.619, and 0.598, respectively (**Figure 5I**).

Validation in the Guanxi HCC Cohort

A total of 102patients who were hospitalized in the first affiliated hospital of Guangxi Medical University from September 2016 to December 2018 were taken into the group for validation. The baseline information for these patients is presented in **Table 3**. The expressions of ABCC in HCC tissues and in paraneoplastic tissues were detected by immunohistochemical (IHC) and polymerase chain reaction (PCR) assays, respectively. In the IHC assay, the expressions of ABCC1, ABCC4, and ABCC5 in HCC tissues were significantly higher than that of paraneoplastic tissues, while ABCC6 was higher expressed in paraneoplastic tissues (**Figure 6A**). The same expression trends of ABCC1, ABCC4, ABCC5, and ABCC6 were observed at the mRNA level (**Figures 6B–E**). The prognostic significance of ABCC1, ABCC4, ABCC5, and ABCC6 was

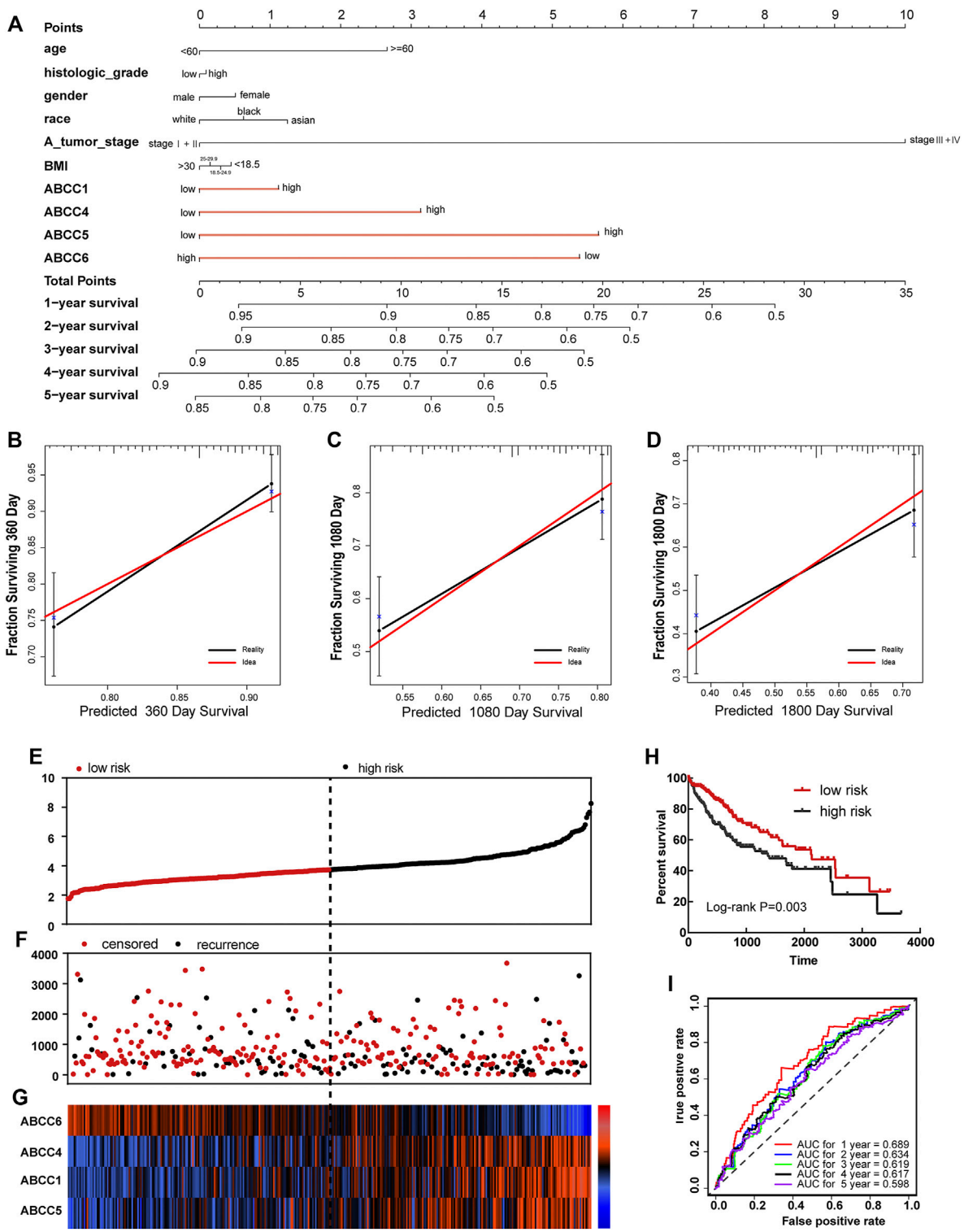


FIGURE 5 | Nomogram and prognostic signature constructed in terms of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* in the TCGA database: **(A)** Nomogram based on expression of *ABCC* genes and clinicopathologic features; **(B)** internal validation for 1-year survival; **(C)** internal validation for 3-year survival; **(D)** internal validation for 5-year survival; **(E)**, scatter plot for risk score; **(F)** scatter plot for survival time (days); **(G)**, heat map corresponding to the expression of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6*; **(H)** survival analysis for high- and low-risk score groups; and **(I)** AUC for inspecting the efficiency of the prognostic signature for predicting long-term prognosis.

TABLE 3 | Clinical characteristics of patients in HCC from Guangxi China.

Variables	Patients	Overall survival			
	(*n* = 102)	No. of events	MST (days)	HR (95% CI)	P
Age					
<60	76	37	23.9		
60	26	12	NA	0.956 (0.498–1.834)	0.891
Gender					
Female	14	5	NA		
Male	88	44	35	0.625 (0.247–1.583)	0.314
BMI					
<24.9	81	38	45		
>24.9	21	11	40	1.054 (0.539–2.063)	0.876
Alcohol					
No	65	34	33		
Yes	37	15	NA	1.595 (0.862–2.950)	0.131
Cirrhosis					
No	9	3	NA		
Yes	93	46	40	1.175 (0.544–5.637)	0.337
Child					
No	100	48	45		
Yes	2	1	3	6.122 (0.796–47.096)	0.045
BCLC					
A	67	32	45	1	
B	28	19	NA	3.757 (1.267–4.144)	
C	7	5	30	6.677 (5.878–8.201)	0.032
Missing	2				
AFP					
<200	50	18	NA	1	
>200	51	31	30	2.038 (1.139–3.648)	0.030
Missing	1				
Radical resection					
No	30	16	33	1	
Yes	70	32	45	0.807 (0.442–1.473)	0.480
Missing	2				
Histological					
Low	5	1	48	1	
Middle	65	32	35	2.978 (0.407–21.813)	
High	22	11	33	3.255 (0.419–25.275)	0.448

Notes: HCC, hepatocellular carcinoma; MST, median survival time; OS, overall survival; HR, hazard ratio; CI, confidence interval.

also observed in the Guangxi HCC cohort (**Figures 6F–I**; **Table 4**).

In the prognostic signature built for the Guangxi cohort, patients were divided into two groups in terms of the risk score (**Figure 6J**). The same as above-mentioned, the dots representing patients in the high-risk group also tended to cluster lower (**Figure 6K**). The expression levels of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* in patients were presented in the form of heat maps (**Figure 6L**). The prognosis of the high-risk group was significantly worse than that of the low-expression group (**Figure 6M**, $p = 0.001$). The AUC value of the prognostic signature for 1-year, 2-year, 3-year, and 4-year overall survival prediction was 0.787, 0.772, 0.838, and 0.788, respectively (**Figure 6N**). In the nomogram, the contribution of *ABCC1*, *ABCC4*, *ABCC5*, *ABCC6*, and clinicopathologic features to overall survival was displayed by the length of the corresponding scales (**Supplementary Figure S4A**). The calibration plot for 1-, 2- and 3-year survival after the surgery revealed a satisfactory overlap between calculation and reality (**Supplementary Figures S4B–D**).

Biological Functional Exploration of *ABCC*s

The enrichment analysis of the *ABCC* gene by setting Homo sapiens as the background was performed on the DAVID online database for obtaining enrichment information about GO terms. The corresponding relationship between ABCCs and GO terms is displayed in **Figure 7A**. The enrichment analysis of GO showed that *ABCCs* were mainly related to ATP binding, ATP activity, transmembrane, and other biological functions (**Figure 7B**). The bubble color from red to green represents the biological function of the –log (*p*-value) from high to low. The network diagram of the relationship between enriched GO terms is shown in **Figures 7C,D**. Interactions of *ABCCs* which were analyzed from STRING and Genemania are respectively displayed in **Figures 7E,F**.

GSEA

The GSEA results revealed that the expression of *ABCC1* was associated with tumor differentiation, nod-like receptor signal pathway, resistance to the bcl2 inhibitor up, and so on (**Figures 8A–F**). The pathways that *ABCC4* might regulate are shown in **Figures 8G–L**. *ABCC5* might impact HCC by regulating the fatty

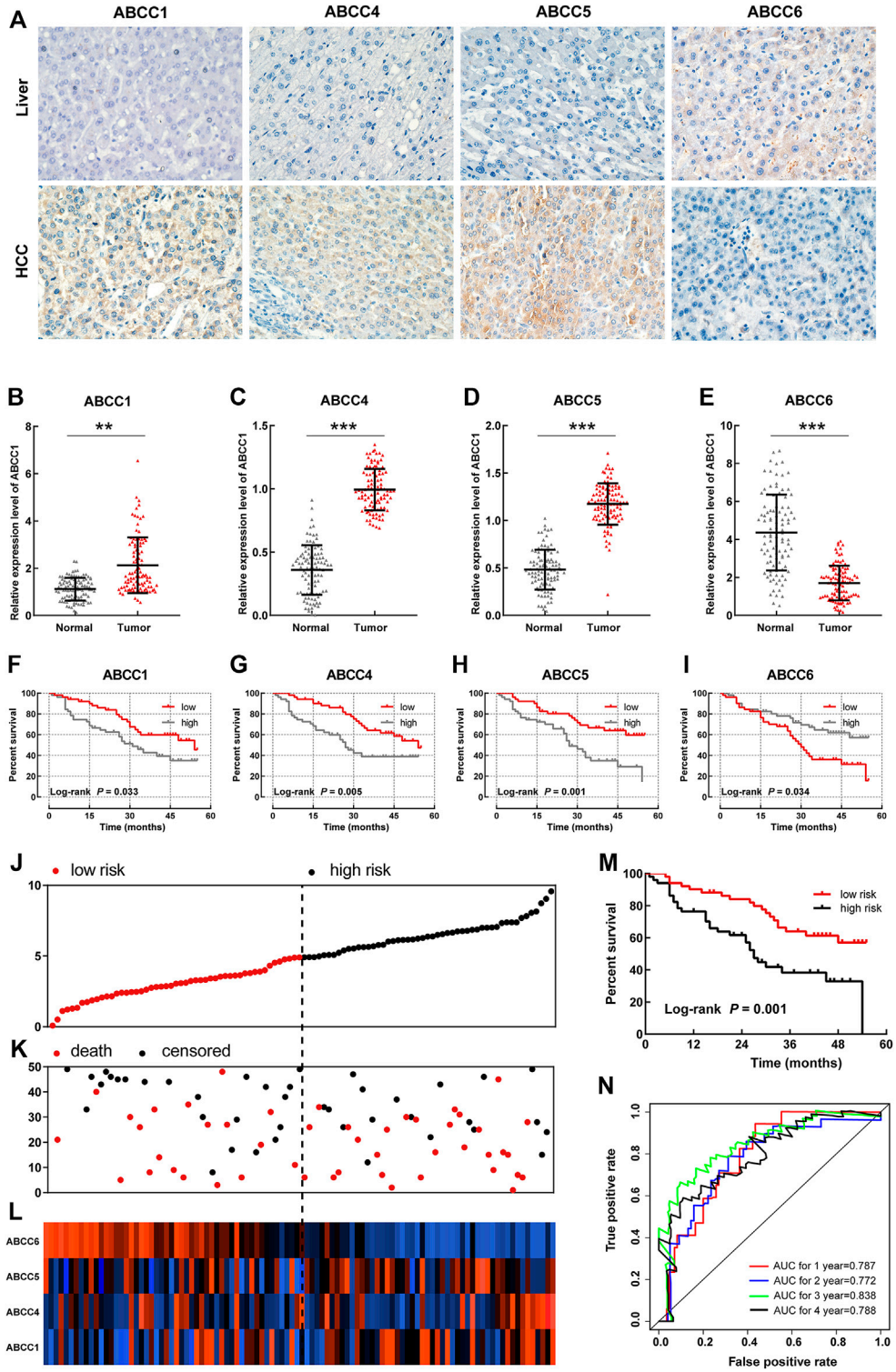


FIGURE 6 | Validation for the prognostic significance of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* in the Guangxi HCC cohort: **(A)** expression of ABCC1, ABCC4, ABCC5, and ABCC6 in HCC tissues and paraneoplastic tissues assessed by IHC assays; **(B–E)** histogram showing ABCC1, ABCC4, ABCC5, and ABCC6 expression levels in HCC tissues and paraneoplastic tissues assessed by PCR assays; **(F–I)** survival curve of ABCC1, ABCC4, ABCC5, and ABCC6 in the Guangxi HCC cohort; the patients were grouped based on median expression; **(J)** Scatter plot for risk score; **(K)** scatter plot for survival time (months); **(L)** heat map corresponding to the expression of *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6*; **(M)** survival analysis for high- and low-risk score groups; and **(N)** AUC for inspecting the efficiency of the prognostic signature for predicting long-term prognosis.

TABLE 4 | Survival analysis results of ABCC genes in the Guangxi cohort.

Gene expression	Patients (*n* = 102)	Overall survival					
		No. of event	MST (months)	Crude HR (95% CI)	Crude P	Adjusted HR (95% CI)	Adjusted P ζ
ABCC1							
Low	51	20	54				
High	51	29	31	1.835 (1.036–3.251)	0.033	1.81 (0.998–3.283)	0.034
ABCC4							
Low	51	22	54				
High	51	27	27	1.991 (1.124–3.557)	0.005	1.912 (1.063–3.437)	0.03
ABCC5							
Low	51	17	NA				
High	51	32	27	2.895 (1.594–5.258)	0.001	2.750 (1.509–5.010)	0.001
ABCC6							
Low	51	28	31				
High	51	21	NA	0.582 (0.329–1.029)	0.034	0.065 (0.303–1.038)	0.046

Notes: ζ Adjusted for child pugh stage, BCLC, stage and AFP; NA, not available; MST, median survival time; HR, hazard ratio.

acid metabolism and the expression of kt19 and myc (**Figures 8M–O**). The result of GSEA revealed that high expression of *ABCC6* was accompanied with lower HCC late recurrence (**Figure 8P**). It also illustrated that *ABCC6* might impact HCC by regulating E2F1 and myc (**Figures 8Q,R**).

Correlation Analysis of ABCC Gene Expression and Tumor-Infiltrating Immune Cells

The estimation of the abundance of immune cell infiltration showed that *ABCC1*, *ABCC4*, and *ABCC5* were significantly positively associated with infiltration of immune cells, which include B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells (**Figures 9A–C**). However, *ABCC6* was negatively associated with the infiltration of immune cells (**Figure 9D**).

DISCUSSION

ABCC expressions were analyzed in two data sets, and consistent results were obtained. Compared with normal tissues, they revealed that *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC10* were significantly upregulated in HCC tissues, while *ABCC6* and *ABCC7* were significantly downregulated in HCC tissues. In the TCGA database, *ABCC5*, *ABCC7*, *ABCC9*, and *ABCC10* were equipped with high diagnostic efficacy for HCC (AUC > 0.8). In GSE76427, the good diagnostic efficacy for HCC was only discovered in *ABCC7*. Combining the results of the two data sets, we consider ABCC7 as a potential diagnostic marker for HCC.

In the TCGA database, *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* were found to be associated with the prognosis of HCC, while further verification in GSE14250 indicated that only *ABCC6* was significantly correlated to the prognosis. The results of survival analysis in the two data sets were very similar, although not identical. We observed that the expression of *ABCC1* and *ABCC5* was associated with the prognosis of liver cancer in both data sets. The reason for the different conclusions may lie in the population difference and inconformity in the causes of neoplasm. HCC patients in GSE76427 were mainly in the Asian population, and the proportion of hepatitis B virus infection was high. However, the majority of HCC patients in the TCGA database were Caucasian and the proportion of hepatitis B virus infection was low. The sample size of both databases is relatively large, and the follow-up data were also of high quality. The results from both databases should be reliable but may apply to different populations. Both clinicopathologic features and biomarker expression were included in the nomogram as prognostic dependent variables, with the length of each variable clearly reflecting its contribution to the prognosis of liver cancer.

Based on the four prognostic biomarkers obtained from the survival analysis, we further performed combined effect survival analysis, nomogram, and prognostic signature based on biomarker expression. The combined survival analysis had obvious advantages, and the prognostic difference between groups was more remarkable. The length of each variable in the nomogram clearly reflects its contribution to the prognosis of liver cancer.

ABCC1 transports drugs to the extracellular substances, thereby reducing the drug concentration and generating drug resistance in cancer (Wlcek and Stieger, 2014). In the liver, ABCC1 undertakes excretion of the drugs into the bile (Zhou, 2008). The ontogeny, localization, expression, and function of ABCC1 in HCC were reported in several research studies, and the previous reports mainly focused on the role of ABCC1 in HCC drug resistance (Flens et al., 1996; Nies et al., 2001; Vander Borght et al., 2005). It was reported that ABCC1 was significantly upregulated in the tissues of oxaliplatin-resistant, 5-fluorouracil-resistant, and sorafenib-resistant HCC patients (Ding et al., 2017; Huang et al., 2018; Ding et al., 2019). In HCC, increased ABCC1 expression was related to increasing dedifferentiation, tumor size, and microvascular invasion (Vander Borght et al., 2008; Zhou, 2008).

Located on the inner surface of the basal side of the liver cells, ABCC4 undertakes bile salt transport (Borst et al., 2007). Previous studies have shown that ABCC4 expression is extremely low in the normal adult liver and fetal liver (Sharma

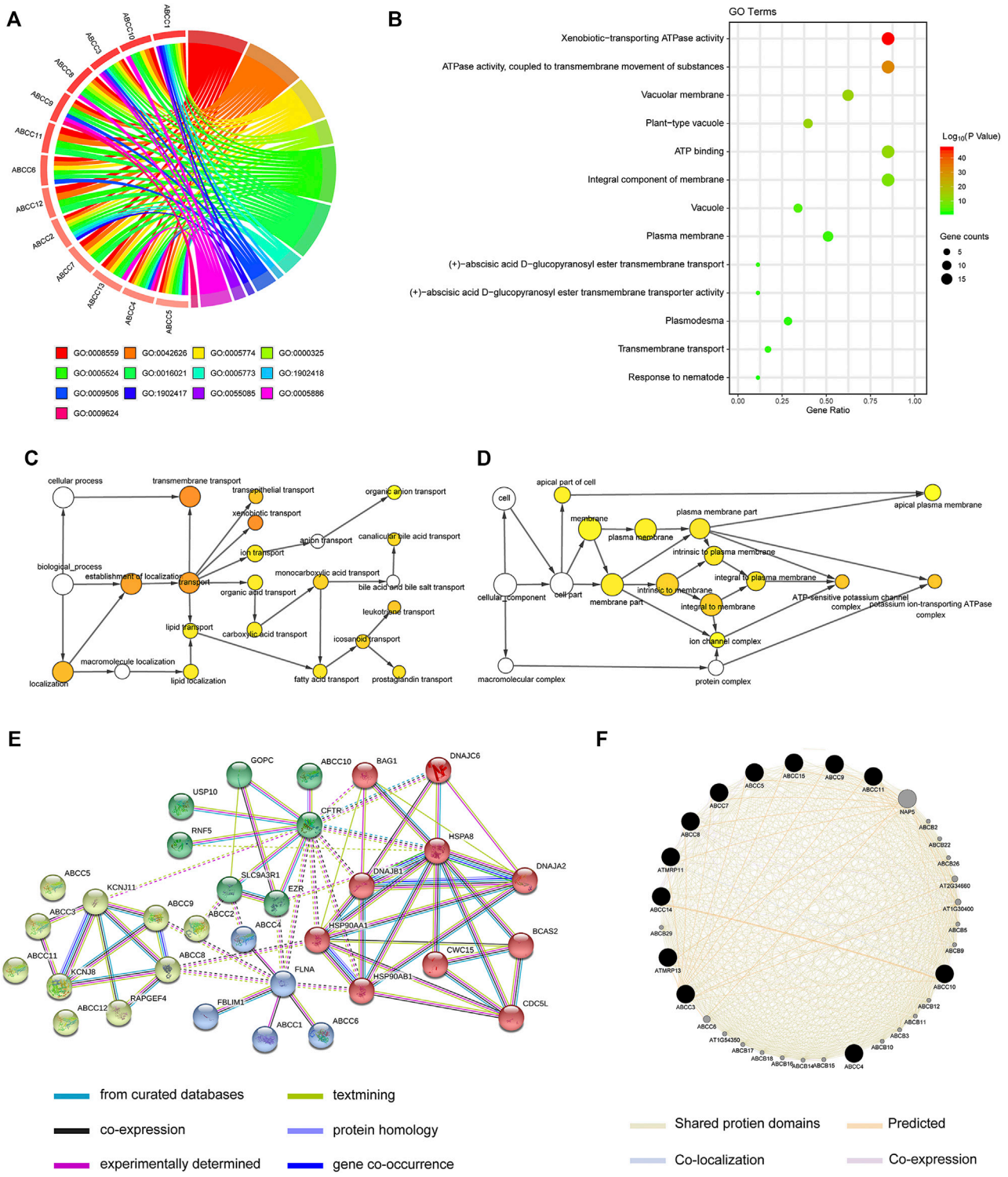


FIGURE 7 | Biological function analysis for ABCC genes by bioinformatics: **(A)** circle plot for displaying the relationship between gene and GO term, **(B)** bubble plot for GO terms, **(C-D)** the network diagram of the relationship between enriched GO terms, **(E-F)** interactions of *ABCCs* that were analyzed from STRING and Genemania, respectively.

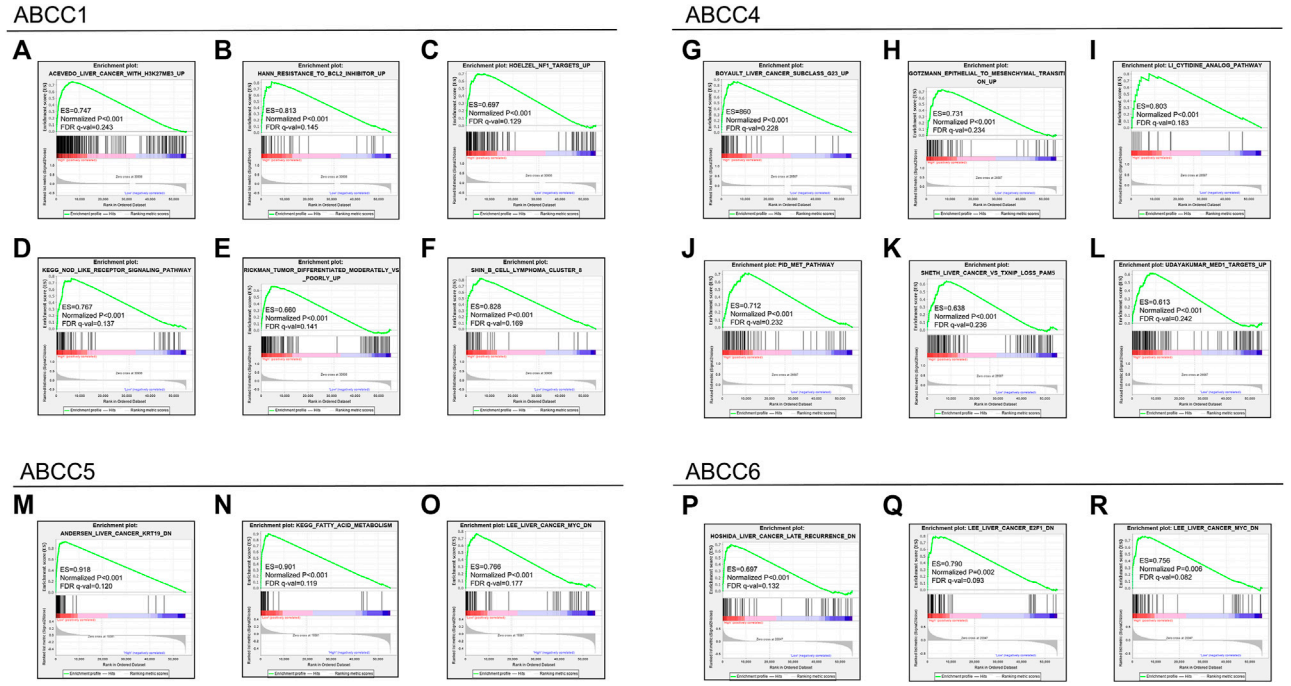


FIGURE 8 | GSEA results for the *ABCC* genes in the TCGA database: **(A–F)** GSEA results for *ABCC1* in HCC, **(G–L)** GSEA results for *ABCC4* in HCC, **(M–O)** GSEA results for *ABCC5* in HCC, and **(P–R)** GSEA results for *ABCC6* in HCC.

et al., 2013), and *ABCC4* expression is significantly increased in cholestatic hepatocyte cell membranes (Gradhand et al., 2008; Sharma et al., 2013). Studies have shown that *ABCC4* is highly expressed in HCC tissues (Sekine et al., 2011; Borel et al., 2012; Luo et al., 2020). Recently, ABCC4 was found to play an important role in HCC oncogenesis and development promoted by decreasing the haploid of p53 (Luo et al., 2020). In addition, ABCC4 could specifically and independently distinguish the aggressive subtypes of HCC (Gradhand et al., 2008).

Here are a few reports on ABCC5 in HCC, with the existing relevant study indicating that ABCC5 is highly expressed in the liver cancer tissues. Our findings in this investigation also confirm this conclusion.

T lymphocytes are known as the main cells of the tumor immunity. Cytotoxic CD8+T cells play a particularly vital role in anticancer immune response (Vesely et al., 2011; Raskov et al., 2021). Once successfully activated, CD8+T cells secreted death-inducing granules to enhance the killing effect of target cells (Basu et al., 2016). Accumulating evidence indicates that TRM (tissue-resident $CD8^+$ memory T cells) is essential for suppressing cancer growth. In a mouse model, whether generated during tumorigenesis or prior to tumor challenge, antitumor TRM cells revealed suppression in cancer growth (Park et al., 2019). Regulatory T cells inhibit anticancer immunity by preventing the protective immunosurveillance of neoplasia and hindering antitumor immune responses in tumor-bearing hosts, thereby promoting the tumor progression (Sakaguchi et al., 2010; Wing and Sakaguchi, 2010; Togashi and Nishikawa, 2017).

B cells have a crucial part in the regulation of T cell response against tumors (Olkhanud et al., 2011; Tadmor et al., 2011). There is a crosstalk between the B and T lymphocytes in antitumor immunity (Blair et al., 2010; DiLillo et al., 2010). Natural killer cells (NK cells) in cancer are involved in priming a multilayered immune response to achieving long-lasting immunity against tumors, in which T cells are involved (Morandi et al., 2012; Ferlazzo and Moretta, 2014). Moreover, NK cells generate cytokines and chemokines that regulate immune responses. The function of non-NK ILCs (innate lymphoid cells, ILCs) in cancer remains unclear.

Combining this investigation and previous research studies, we could preliminarily conclude that *ABCC1*, *ABCC4*, and *ABCC5* reduce drug sensitivity by influencing drug transport out of cells, thus resulting in a poor prognosis in these patients with HCC. In this study, we also found a significant positive correlation between *ABCC1*, *ABCC4*, and *ABCC5* expression and immune cell infiltration.

There were no reports on *ABCC6* in HCC before. The role of *ABCC6* in HCC is completely opposite to that of *ABCC1*, *ABCC4*, and *ABCC5*. We found that *ABCC6* expression was decreased in the liver cancer tissues, and the patients with low *ABCC6* expression had a better prognosis. We speculate that *ABCC6* may function through a completely different mechanism, and the specific findings need to be further studied.

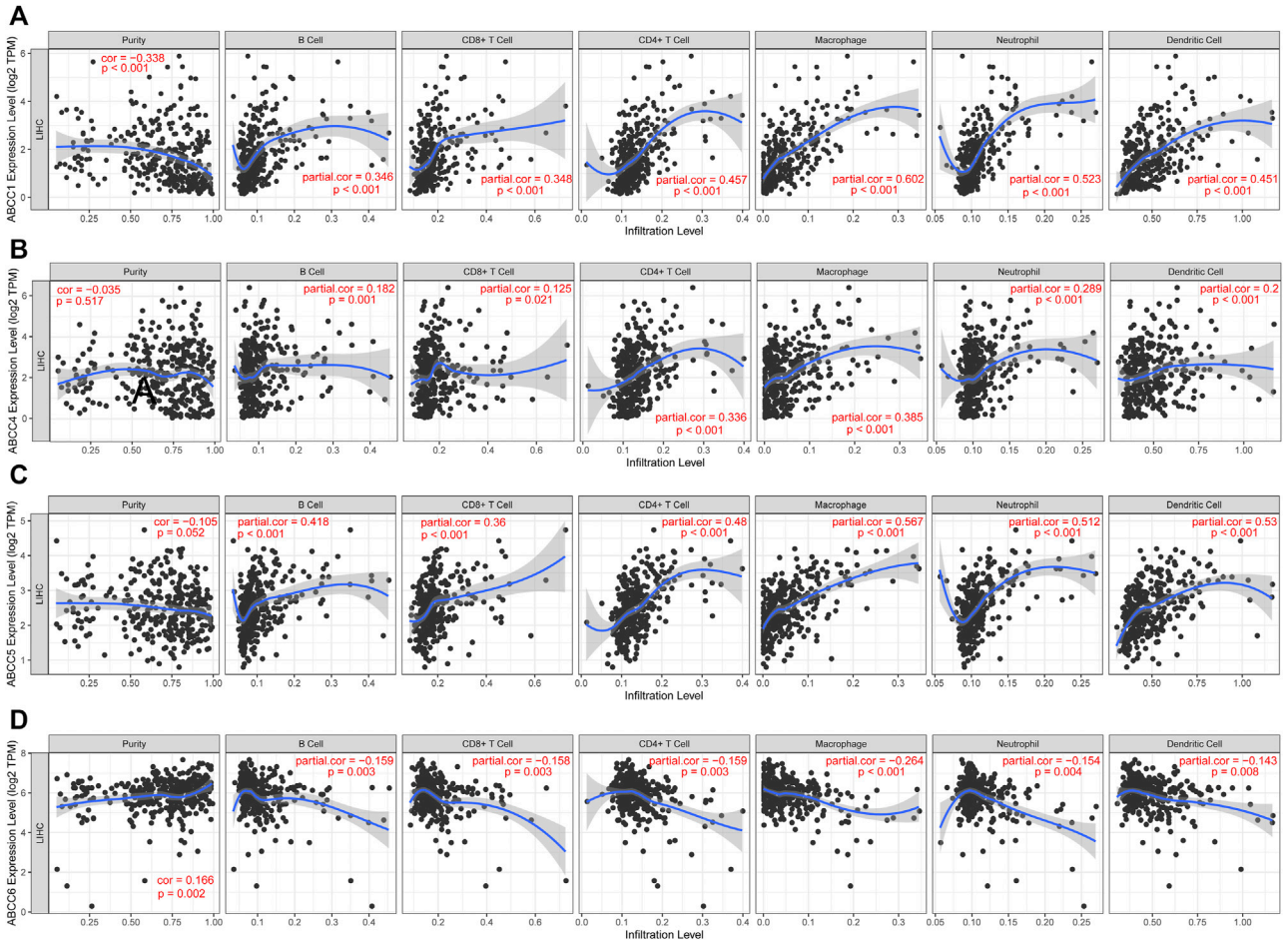


FIGURE 9 | Correlation between ABCC gene expression and tumor-infiltrating immune cells: **(A)** scatter plot corresponding to ABCC1 expression and tumor-infiltrating immune cells, **(B)** scatter plot corresponding to ABCC4 expression and tumor-infiltrating immune cells, **(C)** scatter plot corresponding to ABCC5 expression and tumor-infiltrating immune cells, and **(D)** scatter plot corresponding to ABCC6 expression and tumor-infiltrating immune cells.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

XZ and J-mH formulated the research direction and inclusion criteria for this study. G-zZ and XZ were mainly responsible for data extraction and the elimination of cases that did not meet the inclusion criteria. X-pY and TP analyzed the data, drew charts, and wrote the manuscript. All authors read and approved the final manuscript. All authors made a significant contribution to the work reported, whether this is in the conception, study design, execution, acquisition of data, analysis, and interpretation or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

FUNDING

This work was supported by the Key Laboratory of High-Incidence-Tumor Prevention and Treatment (Guangxi Medical University), Ministry of Education (grant nos. GKE 2018-01, GKE 2019-11, and GKE-ZZ202009); the Guangxi Key Laboratory for the Prevention and Control of Viral Hepatitis (No. GXCDCKL201902); and the Natural Science Foundation of Guangxi Province of China (grant no. 2020GXNSFAA159127).

ACKNOWLEDGMENTS

The authors thank the contributors TCGA database (https://cancergenome.nih.gov/) GSE14520 (https://www.ncbi.nlm.nih.gov/gds/?term=GSE145020) and GSE76427 (https://www.ncbi.

nlm.nih.gov/gds/?term=GSE76427) for sharing the RNA-Sequencing of patients with pancreatic cancer on open access.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.805961/full#supplementary-material

Supplementary Figure 1 | ABCC genes that were not found to be differentially expressed between HCC tissues and paraneoplastic tissues: **(A)** ABCC3, **(B)** ABCC8, **(C)** ABCC11, **(D)**, ABCC12, and **(E)** ABCC13.

Supplementary Figure 2 | ROC curve for ABCC genes with average performance in diagnostic efficacy assessment: **(A)** ABCC1, **(B)** ABCC2, **(C)** ABCC3, **(D)** ABCC4, **(E)** ABCC6, **(F)** ABCC8, **(G)** ABCC11, **(H)** ABCC12, and **(I)** ABCC13.

Supplementary Figure 3 | Survival curves of ABCC genes not found to be associated with HCC prognosis: **(A)** ABCC2, **(B)** ABCC3, **(C)** ABCC7, **(D)** ABCC8, **(E)** ABCC9, **(F)** ABCC10, **(G)** ABCC11, **(H)** ABCC12, and **(I)** ABCC13.

Supplementary Figure 4 | Nomogram constructed based on *ABCC1*, *ABCC4*, *ABCC5*, and *ABCC6* in the Guangxi cohort. **(A)** Nomogram based on expression of *ABCC* genes and clinicopathologic features; **(B)** internal validation for 1-year survival; **(C)** internal validation for 2-year survival; **(D)** internal validation for 3-year survival.

Supplementary Figure 5 | Workflow to summary manuscript.

REFERENCES

Akinyemiju, T., Akinyemiju, T., Abera, S., Ahmed, M., Alam, N., Alemayohu, M. A., et al. (2017). The Burden of Primary Liver Cancer and Underlying Etiologies from 1990 to 2015 at the Global, Regional, and National Level: Results from the Global Burden of Disease Study 2015. *JAMA Oncol.* 3, 1683–1691. doi:10.1001/jamaoncol.2017.3055

Basu, R., Whitlock, B. M., Husson, J., Le Floc'h, A., Jin, W., Oyler-Yaniv, A., et al. (2016). Cytotoxic T Cells Use Mechanical Force to Potentiate Target Cell Killing. *Cell* 165, 100–110. doi:10.1016/j.cell.2016.01.021

Blair, P. A., Noreña, L. Y., Flores-Borja, F., Rawlings, D. J., Isenberg, D. A., Ehrenstein, M. R., et al. (2010). CD19(+)CD24(hi)CD38(hi) B Cells Exhibit Regulatory Capacity in Healthy Individuals but Are Functionally Impaired in Systemic Lupus Erythematosus Patients. *Immunity* 32, 129–140. doi:10.1016/j.immuni.2009.11.009

Borel, F., Han, R., Visser, A., Petry, H., van Deventer, S. J. H., Jansen, P. L. M., et al. (2012). Adenosine Triphosphate-Binding Cassette Transporter Genes Up-Regulation in Untreated Hepatocellular Carcinoma Is Mediated by Cellular microRNAs. *Hepatology* 55, 821–832. doi:10.1002/hep.24682

Borst, P., de Wolf, C., and van de Wetering, K. (2007). Multidrug Resistance-Associated Proteins 3, 4, and 5. *Pflugers Arch. - Eur. J. Physiol.* 453, 661–673. doi:10.1007/s00424-006-0054-9

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

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

Chen, M., Liu, B., Xiao, J., Yang, Y., and Zhang, Y. (2017). A Novel Seven-Long Non-Coding RNA Signature Predicts Survival in Early Stage Lung Adenocarcinoma. *Oncotarget* 8, 14876–14886. doi:10.18632/oncotarget.14781

Chen, Z.-S., and Tiwari, A. K. (2011). Multidrug Resistance Proteins (MRPs/ABCCs) in Cancer Chemotherapy and Genetic Diseases. *FEBS J.* 278, 3226–3245. doi:10.1111/j.1742-4658.2011.08235.x

Childs, S., and Ling, V. (1994). The MDR Superfamily of Genes and its Biological Implications. *Important Adv. Oncol.*, 21–36.

Dean, M., and Allikmets, R. (2001). Complete Characterization of the Human ABC Gene Family. *J. Bioenerg. Biomembr* 33, 475–479. doi:10.1023/a:1012823120935

Dean, M., Rzhetsky, A., and Allikmets, R. (2001). The Human ATP-Binding Cassette (ABC) Transporter Superfamily. *Genome Res.* 11, 1156–1166. doi:10.1101/gr.184901

DiLillo, D. J., Yanaba, K., and Tedder, T. F. (2010). B Cells Are Required for Optimal CD4+ and CD8+ T Cell Tumor Immunity: Therapeutic B Cell Depletion Enhances B16 Melanoma Growth in Mice. *J. Immunol.* 184, 4006–4016. doi:10.4049/jimmunol.0903009

Ding, J., Zhou, X.-T., Zou, H.-Y., and Wu, J. (2017). Hedgehog Signaling Pathway Affects the Sensitivity of Hepatoma Cells to Drug Therapy through the ABCC1 Transporter. *Lab. Invest.* 97, 819–832. doi:10.1038/labinvest.2017.34

Ding, Y., Li, S., Ge, W., Liu, Z., Zhang, X., Wang, M., et al. (2019). Design and Synthesis of Parthenolide and 5-Fluorouracil Conjugates as Potential Anticancer Agents against Drug Resistant Hepatocellular Carcinoma. *Eur. J. Med. Chem.* 183, 111706. doi:10.1016/j.ejmech.2019.111706

El-Serag, H. B. (2012). Epidemiology of Viral Hepatitis and Hepatocellular Carcinoma. *Gastroenterology* 142, 1264. doi:10.1053/j.gastro.2011.12.061

Fattovich, G., Stroffolini, T., Zagni, I., and Donato, F. (2004). Hepatocellular Carcinoma in Cirrhosis: Incidence and Risk Factors. *Gastroenterology* 127, S35–S50. doi:10.1053/j.gastro.2004.09.014

Ferlazzo, G., and Moretta, L. (2014). Dendritic Cell Editing by Natural Killer Cells. *Crit. Rev. Oncog* 19, 67–75. doi:10.1615/critrevoncog.2014010827

Flens, M. J., Zaman, G. J., van der Valk, P., Izquierdo, M. A., Schroeijers, A. B., Scheffer, G. L., et al. (1996). Tissue Distribution of the Multidrug Resistance Protein. *Am. J. Pathol.* 148, 1237–1247.

Fujiwara, N., Friedman, S. L., Goossens, N., and Hoshida, Y. (2018). Risk Factors and Prevention of Hepatocellular Carcinoma in the Era of Precision Medicine. *J. Hepatol.* 68, 526–549. doi:10.1016/j.jhep.2017.09.016

Gower, E., Estes, C., Blach, S., Razavi-Shearer, K., and Razavi, H. (2014). Global Epidemiology and Genotype Distribution of the Hepatitis C Virus Infection. *J. Hepatol.* 61, S45–S57. doi:10.1016/j.jhep.2014.07.027

Gradhand, U., Lang, T., Schaeffeler, E., Glaeser, H., Tegude, H., Klein, K., et al. (2008). Variability in Human Hepatic MRP4 Expression: Influence of Cholestasis and Genotype. *Pharmacogenomics J.* 8, 42–52. doi:10.1038/sj.tpj.6500451

Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., et al. (1986). A Family of Related ATP-Binding Subunits Coupled to many Distinct Biological Processes in Bacteria. *Nature* 323, 448–450. doi:10.1038/323448a0

Hosmer, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Assessing the Fit of the Model, Applied Logistic Regression*, 153–225.

Huang, H., Chen, J., Ding, C. M., Jin, X., Jia, Z. M., and Peng, J. (2018). Lnc RNA NR 2F1- AS 1 Regulates Hepatocellular Carcinoma Oxaliplatin Resistance by Targeting ABCC 1 via miR-363. *J. Cell. Mol. Medi* 22, 3238–3245. doi:10.1111/jcmm.13605

Iasonos, A., Schrag, D., Raj, G. V., and Panageas, K. S. (2008). How to Build and Interpret a Nomogram for Cancer Prognosis. *J. Clin. Oncol.* 26, 1364–1370. doi:10.1200/jco.2007.12.9791

Ito, K., and Murphy, D. (2013). Application of Ggplot2 to Pharmacometric Graphics. *CPT Pharmacometrics Syst. Pharmacol.* 2, e79. doi:10.1038/psp.2013.56

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* 45, D353–D361. doi:10.1093/nar/gkw1092

Keppler, D. (2011). Multidrug Resistance Proteins (MRPs, ABCCs): Importance for Pathophysiology and Drug Therapy. *Handb Exp. Pharmacol.*, 299–323. doi:10.1007/978-3-642-14541-4_8

Leslie, E. M. (2012). Arsenic-Glutathione Conjugate Transport by the Human Multidrug Resistance Proteins (MRPs/ABCCs). *J. Inorg. Biochem.* 108, 141–149. doi:10.1016/j.jinorgbio.2011.11.009

Levrero, M., and Zucman-Rossi, J. (2016). Mechanisms of HBV-Induced Hepatocellular Carcinoma. *J. Hepatol.* 64, S84–S101. doi:10.1016/j.jhep.2016.02.021

Llovet, J. M., Zucman-Rossi, J., Pikarsky, E., Sangro, B., Schwartz, M., Sherman, M., et al. (2016). Hepatocellular Carcinoma. *Nat. Rev. Dis. Primers* 2, 16018. doi:10.1038/nrdp.2016.18

Long, X. D., Ma, Y., Qu, D. Y., Liu, Y. G., Huang, Z. Q., Huang, Y. Z., et al. (2008). The Polymorphism of XRCC3 Codon 241 and AFB1-Related Hepatocellular Carcinoma in Guangxi Population, China. *Ann. Epidemiol.* 18, 572–578. doi:10.1016/j.annepidem.2008.03.003

Luo, Y.-D., Fang, L., Yu, H.-Q., Zhang, J., Lin, X.-T., Liu, X.-Y., et al. (2020). p53 Haploinsufficiency and Increased mTOR Signaling Define a Subset of Aggressive Hepatocellular Carcinoma. *J. Hepatol.* 74 (1), 96–108. doi:10.1016/j.jhep.2020.07.036

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: A Cytoscape Plugin to Assess Overrepresentation of Gene Ontology Categories in Biological Networks. *Bioinformatics* 21, 3448–3449. doi:10.1093/bioinformatics/bti551

Morandi, B., Mortara, L., Chiossone, L., Accolla, R. S., Mingari, M. C., Moretta, L., et al. (2012). Dendritic Cell Editing by Activated Natural Killer Cells Results in a More Protective Cancer-Specific Immune Response. *PloS one* 7, e39170. doi:10.1371/journal.pone.0039170

Nies, A. T., König, J., Pfannschmidt, M., Klar, E., Hofmann, W. J., and Keppler, D. (2001). Expression of the Multidrug Resistance Proteins MRP2 and MRP3 in Human Hepatocellular Carcinoma. *Int. J. Cancer* 94, 492–499. doi:10.1002/ijc.1498

Nolan, D. J., Ginsberg, M., Israely, E., Palikuqi, B., Poulos, M. G., James, D., et al. (2013). Molecular Signatures of Tissue-Specific Microvascular Endothelial Cell Heterogeneity in Organ Maintenance and Regeneration. *Develop. Cel* 26, 204–219. doi:10.1016/j.devcel.2013.06.017

Olkhanud, P. B., Damdinsuren, B., Bodogai, M., Gress, R. E., Sen, R., Wejksza, K., et al. (2011). Tumor-Evoked Regulatory B Cells Promote Breast Cancer Metastasis by Converting Resting CD4+ T Cells to T-Regulatory Cells. *Cancer Res.* 71, 3505–3515. doi:10.1158/0008-5472.can-10-4316

Park, S. L., Buzzai, A., Rautela, J., Hor, J. L., Hochheiser, K., Effern, M., et al. (2019). Tissue-resident Memory CD8+ T Cells Promote Melanoma-Immune Equilibrium in Skin. *Nature* 565, 366–371. doi:10.1038/s41586-018-0812-9

Polaris Observatory Collaborators (2018). Global Prevalence, Treatment, and Prevention of Hepatitis B Virus Infection in 2016: A Modelling Study. *Lancet Gastroenterol. Hepatol.* 3, 383–403. doi:10.1016/S2468-1253(18)30056-6

Raskov, H., Orhan, A., Christensen, J. P., and Gögenur, I. (2021). Cytotoxic CD8(+) T Cells in Cancer and Cancer Immunotherapy. *Br. J. Cancer* 124, 359–367. doi:10.1038/s41416-020-01048-4

Robey, R. W., Pluchino, K. M., Hall, M. D., Fojo, A. T., Bates, S. E., and Gottesman, M. M. (2018). Revisiting the Role of ABC Transporters in Multidrug-Resistant Cancer. *Nat. Rev. Cancer* 18, 452–464. doi:10.1038/s41568-018-0005-8

Rushing, B. R., and Selim, M. I. (2019). Aflatoxin B1: A Review on Metabolism, Toxicity, Occurrence in Food, Occupational Exposure, and Detoxification Methods. *Food Chem. Toxicol.* 124, 81–100. doi:10.1016/j.fct.2018.11.047

Sakaguchi, S., Miyara, M., Costantino, C. M., and Hafler, D. A. (2010). FOXP3+ Regulatory T Cells in the Human Immune System. *Nat. Rev. Immunol.* 10, 490–500. doi:10.1038/nri2785

Sekine, S., Ogawa, R., Ojima, H., and Kanai, Y. (2011). Expression of SLCO1B3 Is Associated with Intratumoral Cholestasis and CTNNB1 Mutations in Hepatocellular Carcinoma. *Cancer Sci.* 102, 1742–1747. doi:10.1111/j.1349-7006.2011.01990.x

Sharma, S., Ellis, E. C. S., Gramignoli, R., Dorko, K., Tahan, V., Hansel, M., et al. (2013). Hepatobiliary Disposition of 17-OHPC and Taurocholate in Fetal Human Hepatocytes: A Comparison with Adult Human Hepatocytes. *Drug Metab. Dispos* 41, 296–304. doi:10.1124/dmd.112.044891

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2015). STRING V10: Protein-Protein Interaction Networks, Integrated over the Tree of Life. *Nucleic Acids Res.* 43, D447–D452. doi:10.1093/nar/gku1003

Tadmor, T., Zhang, Y., Cho, H.-M., Podack, E. R., and Rosenblatt, J. D. (2011). The Absence of B Lymphocytes Reduces the Number and Function of T-Regulatory Cells and Enhances the Anti-Tumor Response in a Murine Tumor Model. *Cancer Immunol. Immunother.* 60, 609–619. doi:10.1007/s00262-011-0972-z

The Gene Ontology Consortium (2017). Expansion of the Gene Ontology Knowledgebase and Resources. *Nucleic Acids Res.* 45, D331–D338. doi:10.1093/nar/gkw1108

Togashi, Y., and Nishikawa, H. (2017). Regulatory T Cells: Molecular and Cellular Basis for Immunoregulation. *Curr. Top. Microbiol. Immunol.* 410, 3–27. doi:10.1007/82_2017_58

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global Cancer Statistics, 2012. *CA: A Cancer J. Clinicians* 65, 87–108. doi:10.3322/caac.21262

Vander Borght, S., Komuta, M., Libbrecht, L., Katoonizadeh, A., Aerts, R., Dymarkowski, S., et al. (2008). Expression of Multidrug Resistance-Associated Protein 1 in Hepatocellular Carcinoma Is Associated with a More Aggressive Tumour Phenotype and May Reflect a Progenitor Cell Origin. *Liver Int.* 28, 1370–1380. doi:10.1111/j.1478-3231.2008.01889.x

Vander Borght, S., Libbrecht, L., Blokzijl, H., Nico Faber, K., Moshage, H., Aerts, R., et al. (2005). Diagnostic and Pathogenetic Implications of the Expression of Hepatic Transporters in Focal Lesions Occurring in normal Liver. *J. Pathol.* 207, 471–482. doi:10.1002/path.1852

Vesely, M. D., Kershaw, M. H., Schreiber, R. D., and Smyth, M. J. (2011). Natural Innate and Adaptive Immunity to Cancer. *Annu. Rev. Immunol.* 29, 235–271. doi:10.1146/annurev-immunol-031210-101324

Wang, Y., Li, J., Xia, Y., Gong, R., Wang, K., Yan, Z., et al. (2013). Prognostic Nomogram for Intrahepatic Cholangiocarcinoma after Partial Hepatectomy. *J. Clin. Oncol.* 31, 1188–1195. doi:10.1200/jco.2012.41.5984

Wing, K., and Sakaguchi, S. (2010). Regulatory T Cells Exert Checks and Balances on Self Tolerance and Autoimmunity. *Nat. Immunol.* 11, 7–13. doi:10.1038/ni.1818

Wlcek, K., and Stieger, B. (2014). ATP-Binding Cassette Transporters in Liver. *Biofactors* 40, 188–198. doi:10.1002/biof.1136

Wogan, G. N., Kensler, T. W., and Groopman, J. D. (2012). Present and Future Directions of Translational Research on Aflatoxin and Hepatocellular Carcinoma. A Review. *Food Addit. Contam. A* 29, 249–257. doi:10.1080/19440049.2011.563370

Yamada, A., Nagahashi, M., Aoyagi, T., Huang, W.-C., Lima, S., Hait, N. C., et al. (2018). ABCC1-Exported Sphingosine-1-Phosphate, Produced by Sphingosine Kinase 1, Shortens Survival of Mice and Patients with Breast Cancer. *Mol. Cancer Res.* 16, 1059–1070. doi:10.1158/1541-7786.mcr-17-0353

Yang, J. D., Hainaut, P., Gores, G. J., Amadou, A., Plymoth, A., and Roberts, L. R. (2019). A Global View of Hepatocellular Carcinoma: Trends, Risk, Prevention and Management. *Nat. Rev. Gastroenterol. Hepatol.* 16, 589–604. doi:10.1038/s41575-019-0186-y

Zhang, W., He, H., Zang, M., Wu, Q., Zhao, H., Lu, L.-L., et al. (2017). Genetic Features of Aflatoxin-Associated Hepatocellular Carcinoma. *Gastroenterology* 153, 249–262. doi:10.1053/j.gastro.2017.03.024

Zhang, X. (2018). NAFLD Related-HCC: The Relationship with Metabolic Disorders. *Adv. Exp. Med. Biol.* 1061, 55–62. doi:10.1007/978-981-10-8684-7_5

Zhang, Z., Geskus, R. B., Kattan, M. W., Zhang, H., and Liu, T. (2017). Nomogram for Survival Analysis in the Presence of Competing Risks. *Ann. Translational Med.* 5, 403. doi:10.21037/atm.2017.07.27

Zhou, S.-F. (2008). Structure, Function and Regulation of P-Glycoprotein and its Clinical Relevance in Drug Disposition. *Xenobiotica* 38, 802–832. doi:10.1080/00498250701867889

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: 11 March 2022
doi: 10.3389/fgene.2022.795820

LYPD3, a New Biomarker and Therapeutic Target for Acute Myelogenous Leukemia

Tingting Hu[1], *Yingjie Zhang*[2], *Tianqing Yang*[1], *Qingnan He*[1]* *and Mingyi Zhao*[1]*

[1]*Department of Pediatrics, Third Xiangya Hospital, Central South University, Changsha, China,* [2]*College of Biology, Hunan University, Changsha, China*

Background: Acute myelogenous leukemia (AML) is nosocomial with the highest pediatric mortality rates and a relatively poor prognosis. *C4.4A*(*LYPD3*) is a tumorigenic and high-glycosylated cell surface protein that has been proven to be linked with the carcinogenic effects in solid tumors, but no hematologic tumors have been reported. We focus on exploring the molecular mechanism of *LYPD3* in the regulation of the occurrence and development of AML to provide a research basis for the screening of markers related to the treatment and prognosis.

Methods: Datasets on RNA Sequencing (RNA-seq) and mRNA expression profiles of 510 samples were obtained from The Cancer Genome Atlas Program/The Genotype-Tissue Expression (Tcga-gtex) on 10 March 2021, which included the information on 173 AML tumorous tissue samples and 337 normal blood samples. The differential expression, identification of prognostic genes based on the COX regression model, and LASSO regression were analyzed. In order to better verify, experiments including gene knockdown mediated by small interfering RNA (siRNA), cell proliferation assays, and Western blot were prefomed. We studied the possible associated pathways through which *LYPD3* may have an impact on the pathogenesis and prognosis of AML by gene set enrichment analysis (GSEA).

Results: A total of 11,490 differential expression genes (DEGs) were identified. Among them, 4,164 genes were upregulated, and 7,756 genes were downregulated. The univariate Cox regression analysis and LASSO regression analysis found that 28 genes including *LYPD3*, *DNAJC8*, and other genes were associated with overall survival (OS). After multivariate Cox analysis, a total of 10 genes were considered significantly correlated with OS in AML including *LYPD3*, which had a poor impact on AML (p <0.05). The experiment results also supported the above conclusion. We identified 25 pathways, including the E2F signaling pathway, *p53* signaling pathway, and *PI3K_AKT* signaling pathway, that were significantly upregulated in AML samples with high *LYPD3* expression (p < 0.05) by GSEA. Further, the results of the experiment suggested that *LYPD3* participates in the development of AML through the *p53* signaling pathway or/and *PI3K/AKT* signaling pathway.

Conclusion: This study first proved that the expression of *LYPD3* was elevated in AML, which was correlated with poor clinical characteristics and prognosis. In addition, *LYPD3* participates in the development of AML through *p53 or/and* the *PI3K/AKT* signaling pathway.

Keywords: acute myelogenous leukemia, COX, LASSO, LYPD3, poor prognosis, p53, PI3K/AKT, GSEA

OPEN ACCESS

Edited by:
Frank Emmert-Streib, Tampere University, Finland

Reviewed by:
Raquel Duarte, University of the Witwatersrand, South Africa
Enrique Ambrocio-Ortiz, Instituto Nacional de Enfermedades Respiratorias-México (INER), Mexico

***Correspondence:**
Qingnan He
heqn2629@csu.edu.cn
Mingyi Zhao
zhao_mingyi@csu.edu.cn

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

Received: *15 October 2021*
Accepted: *15 February 2022*
Published: *11 March 2022*

Citation:
Hu T, Zhang Y, Yang T, He Q and Zhao M (2022) LYPD3, a New Biomarker and Therapeutic Target for Acute Myelogenous Leukemia. Front. Genet. 13:795820. doi: 10.3389/fgene.2022.795820

INTRODUCTION

Acute leukemia (AL) is reported to be the 10th most common cancer all over the world with over 350,000 new cases diagnosed per year (Voelker, 2019). AL can be roughly divided into acute lymphoblastic leukemia (ALL) and acute non-lymphoblastic leukemia (ANLL), and acute myelogenous leukemia (AML) is a subtype of ANLL (Castagnola et al., 2010). AML is the deadliest form of hemal tumor worldwide, and according to the reports from the Global Burden of Disease (GBD), there were 147,000 deaths in 2015 (Mortality, 2016; Fitzmaurice et al., 2017). AML is more likely to occur in children and adolescents, making up 15–20% of AL in children and about 33% in adolescents (Creutzig et al., 2018). There is convincing evidence of a link between abnormal expression, factors of the oxidative stress, activation of various cytokines, signaling pathways, and proliferation and metastasis of tumor cells (Milkovic et al., 2014; Wee and Wang, 2017; Iqbal et al., 2019).

It has been reported that high-glycosylated cell surface proteins are overexpressed or abnormally expressed in several kinds of cancers, including breast cancer, pancreatic cancer, colon cancer, and carcinoma of the ovary. Notably, it is well documented that the abnormal expressions are related to the poor prognosis (Hollingsworth and Swanson, 2004; Vincent et al., 2008; Kitamoto et al., 2011; Gupta et al., 2012; Yang et al., 2015; van Putten and Strijbis, 2017). Therefore, what do high-glycosylated cell surface proteins have to do with cancers? According to the universally accepted views, high-glycosylated cell surface proteins promote oncogenic effects by their glycosylated extracellular domains, which might protect cancer cells under harmful conditions, and by the intracellular domain that is associated with pathways that regulate inflammation, apoptosis, and cell differentiation (van Putten and Strijbis, 2017). In addition, cancer cells appear to utilize the high-glycosylated cell surface proteins to regulate detachment and reattachment during metastasis (Maher et al., 2011). Furthermore, there is more evidence that high-glycosylated cell surface proteins are associated with cellular growth, differentiation, transformation, adhesion, invasion, and immune surveillance (Hollingsworth and Swanson, 2004; Chauhan et al., 2006; Ohtsubo and Marth, 2006; Acar et al., 2008; Maher et al., 2011). However, the molecular biomarkers which are available for early diagnosing and prognosis prediction are still required to be further studied and developed.

C4.4A (*Ly6/PLAUR domain-containing protein 3*, *LYPD3*), first reported in 1998, is a tumorigenic and high-glycosylated cell surface protein that has been proven to be linked with the carcinogenic effects in different solid tumors (Rösel et al., 1998; Hansen et al., 2004; Hansen et al., 2008; Jacobsen et al., 2012; Görtz et al., 2017; De Loma et al., 2020; Yue et al., 2020). The elevated expression of *LYPD3* is not only demonstrated to be associated with lung adenocarcinoma carcinogenesis and poor prognosis (Jacobsen et al., 2014; Cohen et al., 2017; Hu et al., 2020) but also there is evidence that *LYPD3* can lead to the initiation and development of cancers and the chemoresistance of metastatic cancers by impacting the proliferation and apoptosis of the tumor, which are involved in many important regulatory mechanisms of cancers (Lamouille et al., 2014; Fischer et al., 2015). Furthermore, the relationship between the molecule and AML remains unclear. Also, *LYPD3* has not been reported to be found in normal blood, and in our previous bioanalysis, we found that *LYPD3* expression was increased in AML, suggesting that it may be an emerging marker. *LYPD3* can be the ideal target for the therapy method and early detection of AML.

In our study, we analyzed the relationship between the expression of *LYPD3* and the clinical variables of AML according to the data obtained from the public database, and they were experimentally validated. Based on these studies, we aimed to determine the clinical value and prognostic significance of *LYPD3* in patients with AML based on our study so as to provide a theoretical basis and molecular basis for future clinical and basic research.

MATERIALS AND METHODS

Data Source

Datasets on RNA Sequencing (RNA-seq) and mRNA expression profiles of 510 samples were obtained from TCGA-GTEx on 10 March 2021, which included the information on 173 AML tumorous tissue samples and 337 normal blood samples. Among the AML tumorous tissue samples, the survival time of 12 samples was 0, and 12 samples had no survival information, which have been all removed. Finally, a total of 149 AML samples were included in this study. Also, genes with an expression value of 0, which occupied more than 60% of all the samples, were excluded. RGene expression data and phenotype data of 337 normal whole blood samples were downloaded from the GTEx (https://www.gtexportal.org/home/) database. The abovementioned databases were obtained from UCSC XENA (https://xenabrowser.net/) (**Figure 1A**).

Differential Expression Analysis

The mRNA expression levels between tumor and normal samples were analyzed by using the DESeq2 package in R software. First, genes with an expression value of 0 in more than half of the samples were excluded. Differentially expressed genes were screened with the cutoff value of adjusted *p-value* < 0.05 and | log2-fold change [FC] | > 2.

Identification of Prognostic Genes Based on the COX Regression Model and LASSO Regression Analysis

To figure out the prognostic genes of the AML samples, the related hazard ratios (HRs), 95% confidence intervals (CIs) of the HRs, and *p*-values were analyzed using univariate and multivariate Cox regression. Least absolute shrinkage and selection operator (LASSO) Cox regression analysis, which could reduce the dimensionality and select the most robust markers to predict prognosis, was used after the univariate Cox analysis to further identify the candidate genes that could be used to construct the multivariate Cox regression model. Survival analysis was conducted using the R Bioconductor. The survival package was used for univariate and multivariate Cox regression analyses, while the glmnet package was used for the LASSO regression analysis.

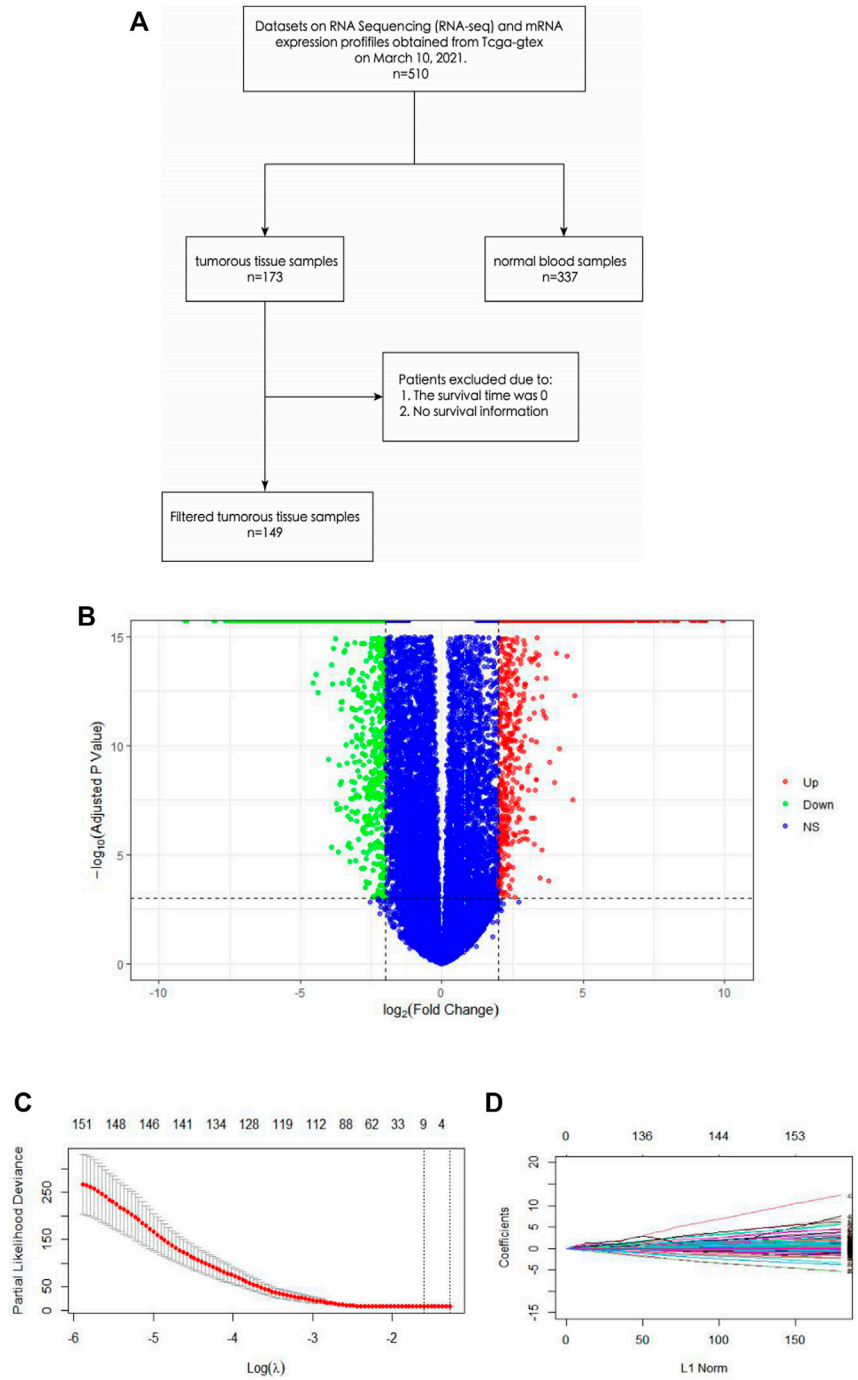


Figure 1 (A) The entire screening process of Patients cohort ;
(B) Volcane plot of Differential expression genes (DEGs) (4164 genes were up-regulated, and 7756 genes were down-regulated);
(C) The univariate Cox regression analysis of AML;
(D) LASSO regression analysis of AML (C and D showed that 28 genes were associated with OS);

FIGURE 1 | (A) Entire screening process of the patient cohort; **(B)** volcano plot of DEGs (4164 genes were upregulated, and 7756 genes were downregulated); **(C)** univariate Cox regression analysis of AML; and **(D)** LASSO regression analysis of AML. **(C,D)** show that 28 genes were associated with OS.

Gene Enrichment Analysis

Genes in each module were subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes Genomes (KEGG) pathway analysis to understand their biological function better. The GO enrichment analysis and KEGG signal pathway analysis of differential expression genes (DEGs) were carried out by clusterProfiler in R software, and the results were visualized. For this study, we analyzed the gene set named h.all.v6.2 symbols.gmt with the cutoff value of normal $p < 0.05$.

Single Gene Enrichment Analysis

The Gene Set Enrichment Analysis (GSEA) software was used for single-gene enrichment analysis aiming at *LYPD3*. Based on LASSO regression and Cox regression, survival analysis was performed to screen out the prognostic genes. Finally, the receiver operating characteristic (ROC) curve was plotted, and the 1-, 3, and 5-year survival rates were calculated.

Cell Culture

HL-60 cells (non-adherent human acute myelogenous leukemia cells), A549 cells (adherent human lung cancer cells), HCT 116 cells (adherent human colon cancer cells), EC cells (human umbilical vein endothelial cells), and CEM cells (non-adherent human lymphoma cells) were all obtained from the ScienCell (San Diego, California, United States). They were all incubated with RPMI Medium Modified medium (Cytiva, Dana Hector, United States) containing 10% fetal bovine serum (Gibco, Thermo Fisher Scientific, Waltham, MA, United States). The cells were cultured using a previously described method (Chen et al., 2018).

LYPD3 Gene Knockdown Mediated by Small Interfering RNA (siRNA)

We obtained *LYPD3* siRNA and scrambled negative control siRNA for transfection from Shanghai GenePharma (China). The siRNA sequences were *LYPD3* siRNA (sense: 5′- *GCU GUAACUCUGACCUCCGCAACAA*-3′; antisense: 5′ *UUGUUG CGGAGGUCAGAGUUACA GC*-3′) and scrambled negative control siRNA (sense, 5′-*UUCUCCGAACGUGUCACGUTT*-3′; antisense, 5′-*ACGUGACACGUUCGGAGAATT*-3′). HL-60 cells (0.4×10^6 cells/well) were seeded in six-well plates and grown to 30% confluence. The HL-60 cells were divided into three groups: a control group, a blank group (transfected with negative control siRNA), and an Si-*LYPD3* group (transfected with *LYPD3* siRNA). Lip2000 (Invitrogen, China) was used to improve the transfection efficiency. The transfection mixture was replaced after 24 h with 1640 with 10% fetal bovine serum (FBS). Then, the cells were incubated for another 24 h and subjected to Western blot analysis. Also, *LYPD3* knockdown was confirmed by Western blot.

Cell Proliferation Assays

The effect of *LYPD3* on HL-60 proliferation was assayed using the Cell Counting Kit-8 (CCK8, Beyotime, China). In summary, for the CCK8 assay, cultured HL-60 were suspended in a culture medium with 0.1% FBS and inoculated in a 96-well plate (1×10^4cells/well) along with 0, 5 μM siRNA(Shanghai GenePharma, China). After 0, 24, 48, and 72 h incubation, 10 μL of CCK8 solution was added to each well. The plate was incubated for an additional 1.5 h before measuring the absorbance at a 450 nm wavelength using a microplate reader (Thermo MK3, United States).

Western Blot

For the analysis of HO-1, Nrf2, Cas-3, Cas-1, PARP-1, akt-1, P-akt, P53, and β-actin at the protein level, HL-60 cells were seeded in six-well plates with a density of 7.35×10^5cells per well. After 24 h, the medium was changed. Each sample was homogenized in 3 mL of lysis buffer [50 mm Tris (pH 8.0), 150 mm NaCl, and 0.5% NP40] with protease inhibitors (1 mm phenylmethylsulfonyl fluoride, 10 μg/ml aprotinin, and 10 μg/ml leupeptin), followed by incubation on ice for 30 min. After centrifugation at 15,000g for 10 min at 4°C, the protein content of the cell lysates was determined by the Bio-Rad protein assay (Hercules, CA). Twenty-five micrograms of protein per lane (unless otherwise noted) was resolved by 7.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to the nitrocellulose membrane. The membranes were blocked with 1% non-fat dried milk in 50 mM Tris (pH 7.5) with 150 mM NaCl and 0.05% Tween-20 and sequentially incubated with antibodies against *LYPD3* (EPR9107, #ab1517-9, ABCAM, MA, United States), β-actin (SANYING, Wuhan, China), Cas-1(PROTEINTECH, United States), Cas-3(PROTEINTECH, United States), PARP-1(PROTEINTECH, United States), P-akt(PROTEINTECH, United States), akt-1(PROTEINTECH, United States), P53(PROTEINTECH, United States), and the appropriate horseradish peroxidase-conjugated, secondary anti-mouse (for ERα and β-actin), or anti-rabbit (for ERβ) antibodies (Amersham Biosciences, Piscataway, NJ). Blots were visualized using enhanced chemiluminescence (ECL Plus) reagents as recommended by the manufacturer (Amersham Biosciences). Densitometric analysis of band intensities was conducted using optical scanning and qualification with ImageJ. After the analysis was completed, protein expression was normalized to β-actin and compared to the corresponding vehicle controls.

Statistical Analysis

All experiment data are expressed as mean ± standard deviation (SD). Graph Prism 8 software was used for statistical analysis. Mean values of the experimental groups were compared by the *t*-test and Chi-square test, and *p-value* < 0.05 was accepted as statistically significant. Western blot results were analyzed with ImageJ software. Experiments were repeated in triplicate, with similar results each time, and the figures show representative experimental results.

RESULTS

Differentially Expressed Genes in the AML Sample and Normal Sample

The DEGs of 149 AML samples obtained from the TCGA database and 337 normal whole blood controls from the GTEx

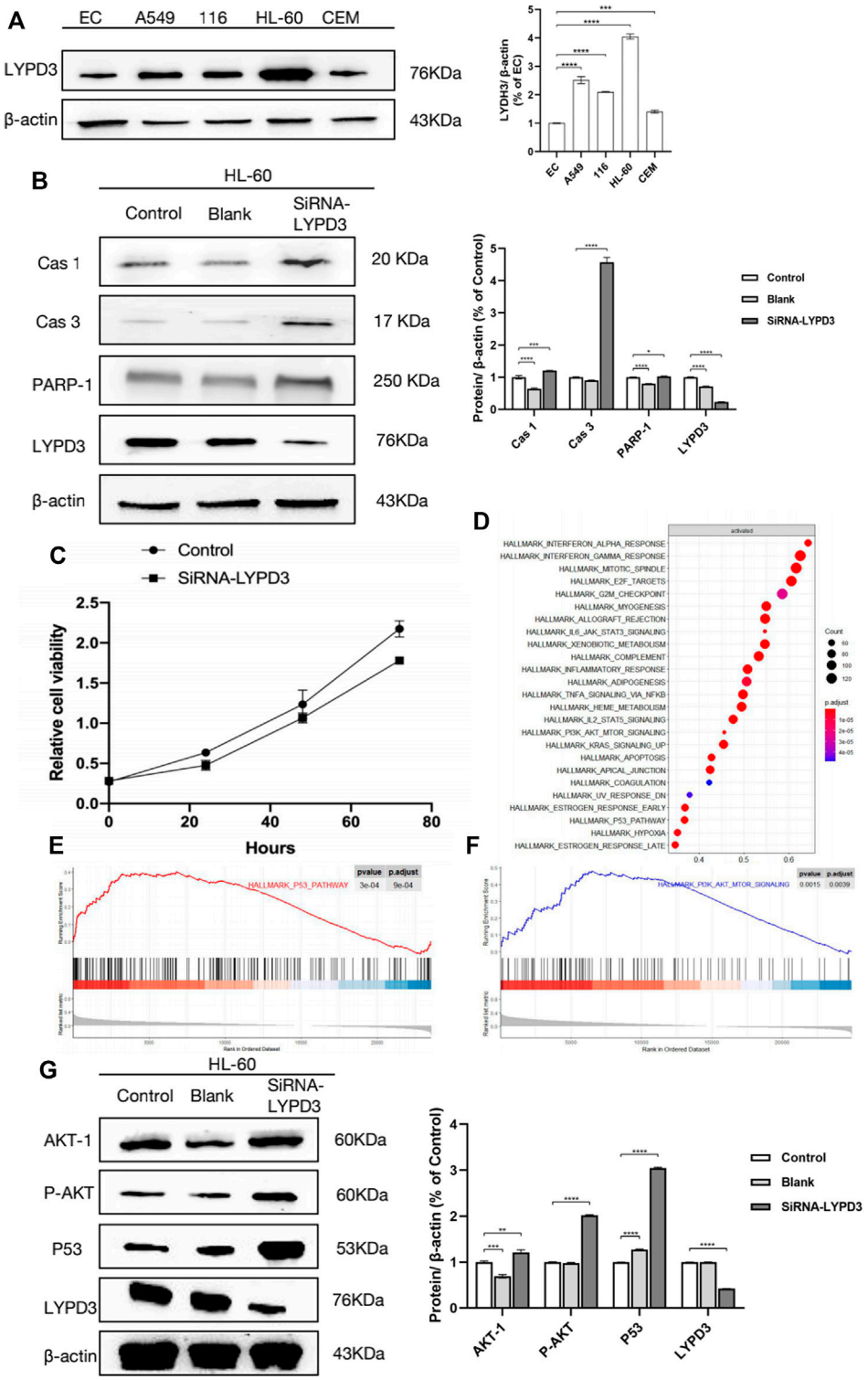


FIGURE 2 | (A) Expression of *LYPD3* in cancer lines (the expression of *LYPD3* is the highest in HL-60 cells); **(B)** *LYPD3* gene knockdown mediated by siRNA-induced apoptosis in AML cells (HL-60 cells); **(C)** *LYPD3* gene knockdown mediated by siRNA-suppressed proliferation in AML cells (HL-60 cells); **(D)** significantly enriched pathways in AML samples with high *LYPD3* expression; **(E)** significantly enriched pathways (the P53 signaling pathways); **(F)** significantly enriched pathways (PI3K_AKT signaling pathway); **(G)** relationship between *LYPD3* and the molecules Akt and P53 (the expression of the *LYPD3* gene knockdown-mediated SiRNA group was obviously increased in p53 and PI3K_AKT signaling).

database were identified. To sum up, a total of 11,490 DEGs were identified. Among them, 4164 genes were upregulated and 7756 genes were downregulated (**Figure 1B**).

Identification of AML-Related Genes Associated With OS

The univariate Cox regression analysis and LASSO regression analysis showed that 28 genes were associated with OS (**Figures 1C,D**). In order to identify the prognosis genes of AML, the multivariate Cox regression model was performed. Also, multivariate Cox analysis further narrowed these candidates into 10 genes, including *AC108479.3*, *FAM207A*, *GS1-304P7.1*, *LINC00540*, *LYPD3*, *NDST3*, *RP11-379P1.4*, *RP11-615I2.1*, *RP11-672L10.6*, *and TREML2*, which were significantly correlated with OS in AML; *AC108479.3*, *NDST3*, *RP11-379P1.4*, and *RP11-672L10.6* benefited AML, while *FAM207A*, *GS1-304P7.1*, *LINC00540*, *LYPD3*, *RP11-615I2.1*, *and TREML2* had a poor impact on AML (**Supplementary Table S1**). Surprisingly, *LYPD3* stands out from these 10 genes and shows a potentially negative impact on AML (p <0.05).

Expression of *LYPD3* in Cancer Lines

To validate the expression of the *LYPD3* protein with the *LYPD3* GPI Ab in cancer lines (see **Figure 2A**), a Western blot analysis was performed. This antibody gave prominent bands around 76 kDa in the A549, hct 116, and HL-60 cells, while no bands were shown in EC and CEM cells, and the results of these are consistent with the study reported by Ping Hu et al. in 2020, the research reported by Willuda J et al. in 2017, and the study reported by Wang L in 2017 (Wang et al., 2017; Willuda et al., 2017; Hu et al., 2020). As shown in **Figure 2A**, the color of HL-60 cell bands was darker than that of the other two bands (that is, HCT116 cell bands and A549 cell bands), indicating that the expression of *LYPD3* was the highest in the HL-60 cells. It is well documented that increased *LYPD3* expression is related to lung adenocarcinoma and colon cancer, which is the same as our experimental results mentioned above. In general, the results are expected to provide a research basis for the screening of markers related to the treatment and prognosis of AML.

LYPD3 Gene Knockdown Mediated by Small Interfering RNA (siRNA) Suppressed Proliferation and Induced Apoptosis in AML Cells (HL-60 Cells)

To investigate the function of *LYPD3*, proliferation and apoptosis assays were carried out in HL-60 transfected with siRNA-*LYPD3*. As shown in **Figure 2C**, *LYPD3* expression was significantly suppressed in HL-60 cells by siRNA-*LYPD3*. It is significant because it has to do with one of the mechanisms of leukemia, abnormal cell proliferation, which might provide a molecular basis for future clinical and basic research. After that, one is that the proliferation of HL-60 cells was evaluated more distinctly compared with the HL-60 cells by siRNA-*LYPD3*; that is to say, *LYPD3* gene knockdown mediated by siRNA-*LYPD3* suppressed proliferation. Furthermore, the expression of apoptotic markers including *CAS-1*, *CAS-3*, and *PARP-1* was significantly upregulated, indicating that obvious inducing of cell apoptosis occurred in HL-60 cells (**Figure 2B**). This result demonstrated that knocking down *LYPD3* induced multiple cell death modes including apoptosis, pyroptosis, and parthanatos, which suppressed the growth of HL-60 cells.

Pathway Analysis of the Effect of *LYPD3* on AML

According to the above results, we have definitely confirmed that *LYPD3* should play an important role in AML. Thus, we further explored the possible related pathways through which *LYPD3* affected the pathogenesis and prognosis of AML. As shown in **Figure 2D**, the significantly enriched pathways in AML samples with high *LYPD3* expression were analyzed by GSEA. We identified 25 pathways that were significantly upregulated in AML samples with high *LYPD3* expression (**Figure 2D**, normal $p < 0.05$), including the E2F signaling pathway, the p53 signaling pathway, the *PI3K_AKT* signaling pathway, and so forth. Among them, the two most significantly enriched pathways were the *P53* signaling pathway and the *PI3K_AKT* signaling pathway (**Figures 2E,F**). In conclusion, we thought that it was doubtful that *LYPD3* might affect the clinical features of AML by regulating the *P53* and/or *PI3K_AKT* signaling pathway. To further explore our conjecture, Western blot was performed (**Figure 2G**). As shown, the expression of *LYPD3* gene knockdown mediated by the siRNA group was obviously increased in *p53* and *PI3K_AKT* signaling, which was consistent with our predictions. However, their specific mechanisms and the upstream and downstream molecules that regulate them need to be further studied.

DISCUSSION

LYPD3, "C4.4A", a membrane protein, partially anchored to the cell surface by glycosylphosphatidylinositol (GPI), which showed predicted structural homology to other members of the *Ly6/uPAR (LU)* protein family (Hansen et al., 2004; Korkmaz et al., 2008; Fujihara et al., 2013; Gårdsvoll et al., 2013). The genes that encode these proteins are clustered in a tiny region on chromosome 19q13. After post-translation processing, *LYPD3* was composed of 278 amino acids distributed in a C-terminal region rich in serine and two N-terminal Lu domains of threonine (Jacobsen et al., 2014). In addition, *LYPD3* was reported to be normally mostly expressed in meningioma tissues according to the study by Mette C in 2011 (Kriegbaum et al., 2011). Furthermore, *LYPD3* has been demonstrated to be highly expressed in several human malignancies, such as breast cancer, colorectal cancer, esophageal cancer, renal cell carcinomas, and so forth (Fletcher et al., 2003; Hansen et al., 2004; Hansen et al., 2008; Miyake et al., 2015; Cohen et al., 2017; Hu et al., 2020; Monteiro et al., 2020). It was found that tumor cell expression of *LYPD3* correlates with poor prognosis in non-small cell lung cancer (NSCLC), esophageal cancer, and renal cell carcinomas. Thus, the association between *LYPD3* and cancer

development is receiving increasing scientific attention and is well worth investigating. Also, studies have found that *LYPD3* is consistently associated with tumor progression and wound healing (Jacobsen and Ploug, 2008). It is also well documented that *LYPD3* can specifically be involved in tumor cell invasion through its interaction with the extracellular matrix (Paret et al., 2007). However, the role of *LYPD3* expression in the occurrence and development of AML remains unclear.

In our study, we found that a total of 11,490 DEGs were identified with 4,164 genes upregulated and 7,756 genes downregulated. Also, in order to identify the prognosis genes of AML, the multivariate Cox regression model was performed. After multivariate Cox analysis, a total of 10 genes including *LYPD3* were significantly correlated with OS in AML, and the results of univariate Cox regression analysis and LASSO regression analysis also indicated that *LYPD3* is associated with AML and poor prognosis (p = 0,01), which supported the theory that the expression of *LYPD3* would be closely correlated with the development of AML and might function as an oncogene for AML.

To validate the expression of the *LYPD3* protein with *LYPD3* GPI Ab in cancer lines (see **Figure 2A**), a Western blot analysis was performed, and the results also support the above-mentioned thought. Then, we further explored the possible related pathways through which *LYPD3* affected the pathogenesis and prognosis of AML. GSEA demonstrated that the *P53* signaling pathway and/or *PI3K_AKT* were significantly enriched in the high-*LYPD3* expression group. The incidence of AML is generally believed that the reasons for the proliferation and apoptosis of leukemia cells were inhibited (Heinrich, 2004; Kornblau et al., 2010; Shih et al., 2013; Elgarten and Aplenc, 2020). In addition, since there is no evidence that *LYPD3* plays a role in proliferative activity or resistance to apoptosis and the two above-mentioned characteristics are thought to be associated with AML pathogenesis according to the above universally recognized views, the most likely explanation may be that *LYPD3* acts as a coactivator. Moreover, it has been reported that a train of AML has intact, unaltered *P53* alleles (Ley et al., 2013; Papaemmanuil et al., 2016). Furthermore, the conundrum of infrequent *P53* mutations in AML is emphasized by the evidence that inactivation of P53 potently promotes AML (Barbosa et al., 2019). Indeed, the study also found that *P53* is one of the most powerful independent indicators of poor outcomes in AML. Thus, we proved that *LYPD3* can be involved in regulating the occurrence, invasion, and metastasis of AML. As shown in **Figure 2B**, the expression of apoptotic markers including *CAS-1*, *CAS-3*, *and PARP-1* was significantly upregulated, which indicated that obvious inducing of cell apoptosis occurred in HL-60 cells. The view that was widely approved was that apoptosis, the most widely studied cell death program, which can be retained as a capacity to undergo, as will be discussed as follows, contributes to both carcinogenesis and anticancer processes (Gregory and Paterson, 2018). Studies have shown that in cancer, separation from neighbors or the substrate triggers a type of spontaneous apoptotic suicide called nest-loss apoptosis. In part, nest-loss apoptosis occurs because cells are deprived of essential integrins and cadherin-mediated survival signals. However, recent studies have shown that interference with the intracellular cytoskeleton caused by detachment can directly trigger apoptosis through the release of pro-apoptotic *BH3* proteins (Evan and Vousden, 2001). In addition, *Akt* mutations are activated in the apoptotic survival signaling pathway in tumors (Russo et al., 2020). *Akt* is a serine/threonine kinase that induces strong survival signaling that is related to the loss of the inhibitor of *Akt* function *PTEN*, which is consistent with the conclusion that the expression of Akt is increased, as shown in **Figure 2G**. In summary, the relationship between *LYPD3*, apoptosis, and leukemia remains complex and unclear, and the specific mechanisms and the upstream and downstream molecules that regulate them need to be further studied.

CONCLUSION

For the first time, we identified that *LYPD3* may promote AML progress through the *PI3K/AKT* and *p53* pathway, which provided a brand new potential biomarker and target for the clinical test and therapy of AML.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

TH and TY analyzed the data and finished the experiment, TH and YZ drafted the manuscript, QH generated the figure, and MZ edited the manuscript. All authors have read and approved the content of the manuscript.

FUNDING

This work was supported by the Hunan innovative province construction project (Grant no. 2019SK2211) and Hunan innovative province construction project (Grant no. 2019SK2211).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.795820/full#supplementary-material

REFERENCES

Acar, M., Jafar-Nejad, H., Takeuchi, H., Rajan, A., Ibrani, D., Rana, N. A., et al. (2008). Rumi Is a CAP10 Domain Glycosyltransferase that Modifies Notch and Is Required for Notch Signaling. *Cell* 132 (2), 247–258. doi:10.1016/j.cell.2007.12.016

Barbosa, K., Li, S., Adams, P. D., and Deshpande, A. J. (2019). The Role of TP53 in Acute Myeloid Leukemia: Challenges and Opportunities. *Genes Chromosomes Cancer* 58 (12), 875–888. doi:10.1002/gcc.22796

Cancer Genome Atlas Research NetworkLey, T. J., Miller, C., Ding, L., Raphael, B. J., Mungall, A. J., Robertson, A., et al. (2013). Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N. Engl. J. Med.* 368 (22), 2059–2074. doi:10.1056/NEJMoa1301689

Castagnola, E., Rossi, M. R., Cesaro, S., Livadiotti, S., Giacchino, M., Zanazzo, G., et al. (2010). Incidence of Bacteremias and Invasive Mycoses in Children with Acute Non-lymphoblastic Leukemia: Results from a Multi-center Italian Study. *Pediatr. Blood Cancer* 55 (6), 1103–1107. doi:10.1002/pbc.22750

Chauhan, S. C., Singh, A. P., Ruiz, F., Johansson, S. L., Jain, M., Smith, L. M., et al. (2006). Aberrant Expression of MUC4 in Ovarian Carcinoma: Diagnostic Significance Alone and in Combination with MUC1 and MUC16 (CA125). *Mod. Pathol.* 19 (10), 1386–1394. doi:10.1038/modpathol.3800646

Chen, X., Cai, X., Le, R., Zhang, M., Gu, X., Shen, F., et al. (2018). Isoliquiritigenin Protects against Sepsis-Induced Lung and Liver Injury by Reducing Inflammatory Responses. *Biochem. Biophys. Res. Commun.* 496 (2), 245–252. doi:10.1016/j.bbrc.2017.11.159

Cohen, A. S., Khalil, F. K., Welsh, E. A., Schabath, M. B., Enkemann, S. A., Davis, A., et al. (2017). Cell-surface Marker Discovery for Lung Cancer. *Oncotarget* 8 (69), 113373–113402. doi:10.18632/oncotarget.23009

Creutzig, U., Kutny, M. A., Barr, R., Schlenk, R. F., and Ribeiro, R. C. (2018). Acute Myelogenous Leukemia in Adolescents and Young Adults. *Pediatr. Blood Cancer* 65 (9), e27089. doi:10.1002/pbc.27089

De Loma, J., Gliga, A. R., Levi, M., Ascui, F., Gardon, J., Tirado, N., et al. (2020). Arsenic Exposure and Cancer-Related Proteins in Urine of Indigenous Bolivian Women. *Front. Public Health* 8, 605123. doi:10.3389/fpubh.2020.605123

Elgarten, C. W., and Aplenc, R. (2020). Pediatric Acute Myeloid Leukemia: Updates on Biology, Risk Stratification, and Therapy. *Curr. Opin. Pediatr.* 32 (1), 57–66. doi:10.1097/mop.0000000000000855

Evan, G. I., and Vousden, K. H. (2001). Proliferation, Cell Cycle and Apoptosis in Cancer. *Nature* 411 (6835), 342–348. doi:10.1038/35077213

Fischer, K. R., Durrans, A., Lee, S., Sheng, J., Li, F., Wong, S. T. C., et al. (2015). Epithelial-to-mesenchymal Transition Is Not Required for Lung Metastasis but Contributes to Chemoresistance. *Nature* 527 (7579), 472–476. doi:10.1038/nature15748

Fletcher, G. C., Patel, S., Tyson, K., Adam, P. J., Schenker, M., Loader, J. A., et al. (2003). hAG-2 and hAG-3, Human Homologues of Genes Involved in Differentiation, Are Associated with Oestrogen Receptor-Positive Breast Tumours and Interact with Metastasis Gene C4.4a and Dystroglycan. *Br. J. Cancer* 88 (4), 579–585. doi:10.1038/sj.bjc.6600740

Fujihara, Y., Tokuhiro, K., Muro, Y., Kondoh, G., Araki, Y., Ikawa, M., et al. (2013). Expression of TEX101, Regulated by ACE, Is Essential for the Production of fertile Mouse Spermatozoa. *Proc. Natl. Acad. Sci.* 110 (20), 8111–8116. doi:10.1073/pnas.1222166110

Gårdsvoll, H., Kriegbaum, M. C., Hertz, E. P., Alpízar-Alpízar, W., and Ploug, M. (2013). The Urokinase Receptor Homolog Haldisin Is a Novel Differentiation Marker of Stratum Granulosum in Squamous Epithelia. *J. Histochem. Cytochem.* 61 (11), 802–813. doi:10.1369/0022155413501879

Global Burden of Disease Cancer CollaborationFitzmaurice, C., Allen, C., Barber, R. M., Barregard, L., Bhutta, Z. A., Brenner, H., et al. (2017). Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived with Disability, and Disability-Adjusted Life-Years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol.* 3 (4), 524–548. doi:10.1001/jamaoncol.2016.5688

Görtz, M., Galli, U., Longerich, T., Zöller, M., Erb, U., and Schemmer, P. (2017). De Novo synthesis of C4.4A in Hepatocellular Carcinoma Promotes Migration and Invasion of Tumor Cells. *Oncol. Rep.* 38 (5), 2697–2704. doi:10.3892/or.2017.5980

Gregory, C. D., and Paterson, M. (2018). An Apoptosis-Driven 'onco-Regenerative Niche': Roles of Tumour-Associated Macrophages and Extracellular Vesicles. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373 (1737), 20170003. doi:10.1098/rstb.2017.0003

Gupta, B. K., Maher, D. M., Ebeling, M. C., Sundram, V., Koch, M. D., Lynch, D. W., et al. (2012). Increased Expression and Aberrant Localization of Mucin 13 in Metastatic colon Cancer. *J. Histochem. Cytochem.* 60 (11), 822–831. doi:10.1369/0022155412460678

Hansen, L. V., Gårdsvoll, H., Nielsen, B. S., Lund, L. R., Danø, K., Jensen, O. N., et al. (2004). Structural Analysis and Tissue Localization of Human C4.4A: a Protein Homologue of the Urokinase Receptor. *Biochem. J.* 380 (Pt 3), 845–857. doi:10.1042/BJ20031478

Hansen, L. V., Lærum, O. D., Illemann, M., Nielsen, B. S., and Ploug, M. (2008). Altered Expression of the Urokinase Receptor Homologue, C4.4A, in Invasive Areas of Human Esophageal Squamous Cell Carcinoma. *Int. J. Cancer* 122 (4), 734–741. doi:10.1002/ijc.23082

Heinrich, M. (2004). Targeting FLT3 Kinase in Acute Myelogenous Leukemia: Progress, Perils, and Prospects. *Mini Rev. Med. Chem.* 4 (3), 255–271. doi:10.2174/1389557043487394

Hollingsworth, M. A., and Swanson, B. J. (2004). Mucins in Cancer: protection and Control of the Cell Surface. *Nat. Rev. Cancer* 4 (1), 45–60. doi:10.1038/nrc1251

Hu, P., Huang, Y., Gao, Y., Yan, H., Li, X., Zhang, J., et al. (2020). Elevated Expression of *LYPD3* Is Associated with Lung Adenocarcinoma Carcinogenesis and Poor Prognosis. *DNA Cel Biol.* 39 (4), 522–532. doi:10.1089/dna.2019.5116

Iqbal, B., Ali, J., Ganguli, M., Mishra, S., and Baboota, S. (2019). Silymarin-loaded Nanostructured Lipid Carrier Gel for the Treatment of Skin Cancer. *Nanomedicine* 14 (9), 1077–1093. doi:10.2217/nnm-2018-0235

Jacobsen, B., Kriegbaum, M. C., Santoni-Rugiu, E., and Ploug, M. (2014). C4.4A as a Biomarker in Pulmonary Adenocarcinoma and Squamous Cell Carcinoma. *World J. Clin. Oncol.* 5 (4), 621–632. doi:10.5306/wjco.v5.i4.621

Jacobsen, B., and Ploug, M. (2008). The Urokinase Receptor and its Structural Homologue C4.4A in Human Cancer: Expression, Prognosis and Pharmacological Inhibition. *Curr. Med. Chem.* 15 (25), 2559–2573. doi:10.2174/092986708785909012

Jacobsen, B., Santoni-Rugiu, E., Illemann, M., Kriegbaum, M. C., Laerum, O. D., and Ploug, M. (2012). Expression of C4.4A in Precursor Lesions of Pulmonary Adenocarcinoma and Squamous Cell Carcinoma. *Int. J. Cancer* 130 (11), 2734–2739. doi:10.1002/ijc.26305

Kitamoto, S., Yamada, N., Yokoyama, S., Houjou, I., Higashi, M., Goto, M., et al. (2011). DNA Methylation and Histone H3-K9 Modifications Contribute to MUC17 Expression. *Glycobiology* 21 (2), 247–256. doi:10.1093/glycob/cwq155

Korkmaz, B., Kuhl, A., Bayat, B., Santoso, S., and Jenne, D. E. (2008). A Hydrophobic Patch on Proteinase 3, the Target of Autoantibodies in Wegener Granulomatosis, Mediates Membrane Binding via NB1 Receptors. *J. Biol. Chem.* 283 (51), 35976–35982. doi:10.1074/jbc.m806754200

Kornblau, S. M., McCue, D., Singh, N., Chen, W., Estrov, Z., and Coombes, K. R. (2010). Recurrent Expression Signatures of Cytokines and Chemokines Are Present and Are Independently Prognostic in Acute Myelogenous Leukemia and Myelodysplasia. *Blood* 116 (20), 4251–4261. doi:10.1182/blood-2010-01-262071

Kriegbaum, M. C., Jacobsen, B., Hald, A., and Ploug, M. (2011). Expression of C4.4A, a Structural uPAR Homolog, Reflects Squamous Epithelial Differentiation in the Adult Mouse and during Embryogenesis. *J. Histochem. Cytochem.* 59 (2), 188–201. doi:10.1369/0022155410394859

Lamouille, S., Xu, J., and Derynck, R. (2014). Molecular Mechanisms of Epithelial-Mesenchymal Transition. *Nat. Rev. Mol. Cel Biol* 15 (3), 178–196. doi:10.1038/nrm3758

Maher, D. M., Gupta, B. K., Nagata, S., Jaggi, M., and Chauhan, S. C. (2011). Mucin 13: Structure, Function, and Potential Roles in Cancer Pathogenesis. *Mol. Cancer Res.* 9 (5), 531–537. doi:10.1158/1541-7786.mcr-10-0443

Milkovic, L., Siems, W., Siems, R., and Zarkovic, N. (2014). Oxidative Stress and Antioxidants in Carcinogenesis and Integrative Therapy of Cancer. *Curr. Pharm. Des.* 20 (42), 6529–6542. doi:10.2174/1381612820666140826152822

Miyake, T., Ito, T., Yanai, A., Inoue, N., Miyagawa, Y., Murase, K., et al. (2015). C4.4A Highly Expressed in HER2-Positive Human Breast Cancers May Indicate a Good Prognosis. *Breast Cancer* 22 (4), 366–373. doi:10.1007/s12282-013-0487-x

Monteiro, M. B., Pelaes, T. S., Santos-Bezerra, D. P., Thieme, K., Lerario, A. M., Oba-Shinjo, S. M., et al. (2020). Urinary Sediment Transcriptomic and

Longitudinal Data to Investigate Renal Function Decline in Type 1 Diabetes. *Front. Endocrinol.* 11, 238. doi:10.3389/fendo.2020.00238

Mortality, G. B. D. (2016). Global, Regional, and National Life Expectancy, All-Cause Mortality, and Cause-specific Mortality for 249 Causes of Death, 1980-2015: a Systematic Analysis for the Global Burden of Disease Study 2015. *Lancet* 388 (10053), 1459–1544. doi:10.1016/S0140-6736(16)31012-1

Ohtsubo, K., and Marth, J. D. (2006). Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* 126 (5), 855–867. doi:10.1016/j.cell.2006.08.019

Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., et al. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N. Engl. J. Med.* 374 (23), 2209–2221. doi:10.1056/nejmoa1516192

Paret, C., Hildebrand, D., Weitz, J., Kopp-Schneider, A., Kuhn, A., Beer, A., et al. (2007). C4.4A as a Candidate Marker in the Diagnosis of Colorectal Cancer. *Br. J. Cancer* 97 (8), 1146–1156. doi:10.1038/sj.bjc.6604012

Rösel, M., Claas, C., Seiter, S., Herlevsen, M., and Zöller, M. (1998). Cloning and Functional Characterization of a New Phosphatidyl-Inositol Anchored Molecule of a Metastasizing Rat Pancreatic Tumor. *Oncogene* 17 (15), 1989–2002. doi:10.1038/sj.onc.1202079

Russo, M., Newell, J. M., Budurlean, L., Houser, K. R., Sheldon, K., Kesterson, J., et al. (2020). Mutational Profile of Endometrial Hyperplasia and Risk of Progression to Endometrioid Adenocarcinoma. *Cancer* 126 (12), 2775–2783. doi:10.1002/cncr.32822

Shih, A. H., Chung, S. S., Dolezal, E. K., Zhang, S.-J., Abdel-Wahab, O. I., Park, C. Y., et al. (2013). Mutational Analysis of Therapy-Related Myelodysplastic Syndromes and Acute Myelogenous Leukemia. *Haematologica* 98 (6), 908–912. doi:10.3324/haematol.2012.076729

van Putten, J. P. M., and Strijbis, K. (2017). Transmembrane Mucins: Signaling Receptors at the Intersection of Inflammation and Cancer. *J. Innate Immun.* 9 (3), 281–299. doi:10.1159/000453594

Vincent, A., Ducourouble, M. P., and Van Seuningen, I. (2008). Epigenetic Regulation of the Human Mucin Gene MUC4 in Epithelial Cancer Cell Lines Involves Both DNA Methylation and Histone Modifications Mediated by DNA Methyltransferases and Histone Deacetylases. *FASEB j.* 22 (8), 3035–3045. doi:10.1096/fj.07-103390

Voelker, R. (2019). New Acute Myeloid Leukemia Therapy. *JAMA* 321 (1), 23. doi:10.1001/jama.2018.20416

Wang, L., Hirohashi, Y., Ogawa, T., Shen, M., Takeda, R., Murai, A., et al. (2017). LY6/PLAUR Domain Containing 3 Has a Role in the Maintenance of Colorectal Cancer Stem-like Cells. *Biochem. Biophys. Res. Commun.* 486 (2), 232–238. doi:10.1016/j.bbrc.2017.02.112

Wee, P., and Wang, Z. (2017). Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers (Basel)* 9 (5), 52. doi:10.3390/cancers9050052

Willuda, J., Linden, L., Lerchen, H.-G., Kopitz, C., Stelte-Ludwig, B., Pena, C., et al. (2017). Preclinical Antitumor Efficacy of BAY 1129980-a Novel Auristatin-Based Anti-C4.4A (*LYPD3*) Antibody-Drug Conjugate for the Treatment of Non-small Cell Lung Cancer. *Mol. Cancer Ther.* 16 (5), 893–904. doi:10.1158/1535-7163.mct-16-0474

Yang, C.-W., Chang, C. Y.-Y., Lai, M.-T., Chang, H.-W., Lu, C.-C., Chen, Y., et al. (2015). Genetic Variations of MUC17 Are Associated with Endometriosis Development and Related Infertility. *BMC Med. Genet.* 16, 60. doi:10.1186/s12881-015-0209-7

Yue, J., Zhu, T., Yang, J., Si, Y., Xu, X., Fang, Y., et al. (2020). CircCBFB-mediated miR-28-5p Facilitates Abdominal Aortic Aneurysm via *LYPD3* and GRIA4. *Life Sci.* 253, 117533. doi:10.1016/j.lfs.2020.117533

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: 14 March 2022
doi: 10.3389/fgene.2021.792106

Pyroptosis-Related lncRNAs Predict the Prognosis and Immune Response in Patients With Breast Cancer

Xia Yang[1], *Xin Weng*[2], *Yajie Yang*[2] *and ZhiNong Jiang*[1]*

[1]*Department of Pathology, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China,* [2]*Department of Pathology, Shenzhen Second People's Hospital, Shenzhen, China*

Background: Breast cancer (BC) is the most common malignant tumor and the leading cause of cancer-related death in women worldwide. Pyroptosis and long noncoding RNAs (lncRNAs) have been demonstrated to play vital roles in the tumorigenesis and development of BC. However, the clinical significance of pyroptosis-related lncRNAs in BC remains unclear.

Methods: Using the mRNA and lncRNA profiles of BC obtained from TCGA dataset, a risk model based on the pyroptosis-related lncRNAs for prognosis was constructed using univariate and multivariate Cox regression model, and least absolute shrinkage and selection operator. Patients were divided into high- and low-risk groups based on the risk model, and the prognosis value and immune response in different risk groups were analyzed. Furthermore, functional enrichment annotation, therapeutic signature, and tumor mutation burden were performed to evaluate the risk model we established. Moreover, the expression level and clinical significance of the selected pyroptosis-related lncRNAs were further validated in BC samples.

Results: 3,364 pyroptosis-related lncRNAs were identified using Pearson's correlation analysis. The risk model we constructed comprised 10 pyroptosis-related lncRNAs, which was identified as an independent predictor of overall survival (OS) in BC. The nomogram we constructed based on the clinicopathologic features and risk model yielded favorable performance for prognosis prediction in BC. In terms of immune response and mutation status, patients in the low-risk group had a higher expression of immune checkpoint markers and exhibited higher fractions of activated immune cells, while the high-risk group had a highly percentage of TMB. Further analyses in our cohort BC samples found that RP11-459E5.1 was significantly upregulated, while RP11-1070N10.3 and RP11-817J15.3 were downregulated and significantly associated with worse OS.

Conclusion: The risk model based on the pyroptosis-related lncRNAs we established may be a promising tool for predicting the prognosis and personalized therapeutic response in BC patients.

Keywords: breast cancer, pyroptosis, lncRNA, prognosis, immune response

OPEN ACCESS

Edited by:
Frank Emmert-Streib, Tampere University, Finland

Reviewed by:
Taobo Hu, Peking University People's Hospital, China
Jing Dong, Medical College of Wisconsin, United States

***Correspondence:**
ZhiNong Jiang
3200039@zju.edu.cn

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

Received: *09 October 2021*
Accepted: *21 December 2021*
Published: *14 March 2022*

Citation:
Yang X, Weng X, Yang Y and Jiang Z (2022) Pyroptosis-Related lncRNAs Predict the Prognosis and Immune Response in Patients With Breast Cancer. Front. Genet. 12:792106. doi: 10.3389/fgene.2021.792106

INTRODUCTION

Breast cancer (BC) is the most common malignant tumor and the leading cause of cancer-related death in women worldwide (Siegel et al., 2020). Owing to the development of local control and systematic treatment, including surgical resection, radiotherapy, systemic chemotherapy, in combination with targeted therapy, hormone replacement therapy, and other novel therapeutic methods, the survival outcome of BCs has significantly improved in the past decade (Harbeck and Gnant, 2017). Despite these efforts, the curative rate and prognosis of BC patients are still unsatisfactory, and up to 15% of cancer-related deaths occurred after treatment according to GLOBOCAN 2018 (Bray et al., 2018). In addition, the heterogeneity of BC results in the diversity of tumor evolution scenarios and traditional therapeutic response (Holm et al., 2017). Thus, there is an urgent requirement to identify novel sensitive biomarkers for predicting prognosis and developing targeted therapeutic agents in BC patients.

Pyroptosis is a novel programmed cell death mediated by the gasdermin family, accompanied by inflammatory and immune responses (Vande Walle and Lamkanfi, 2016). The relationship between pyroptosis and tumor remains mysterious, and the role of pyroptosis in cancer vary in different tissues and genetic

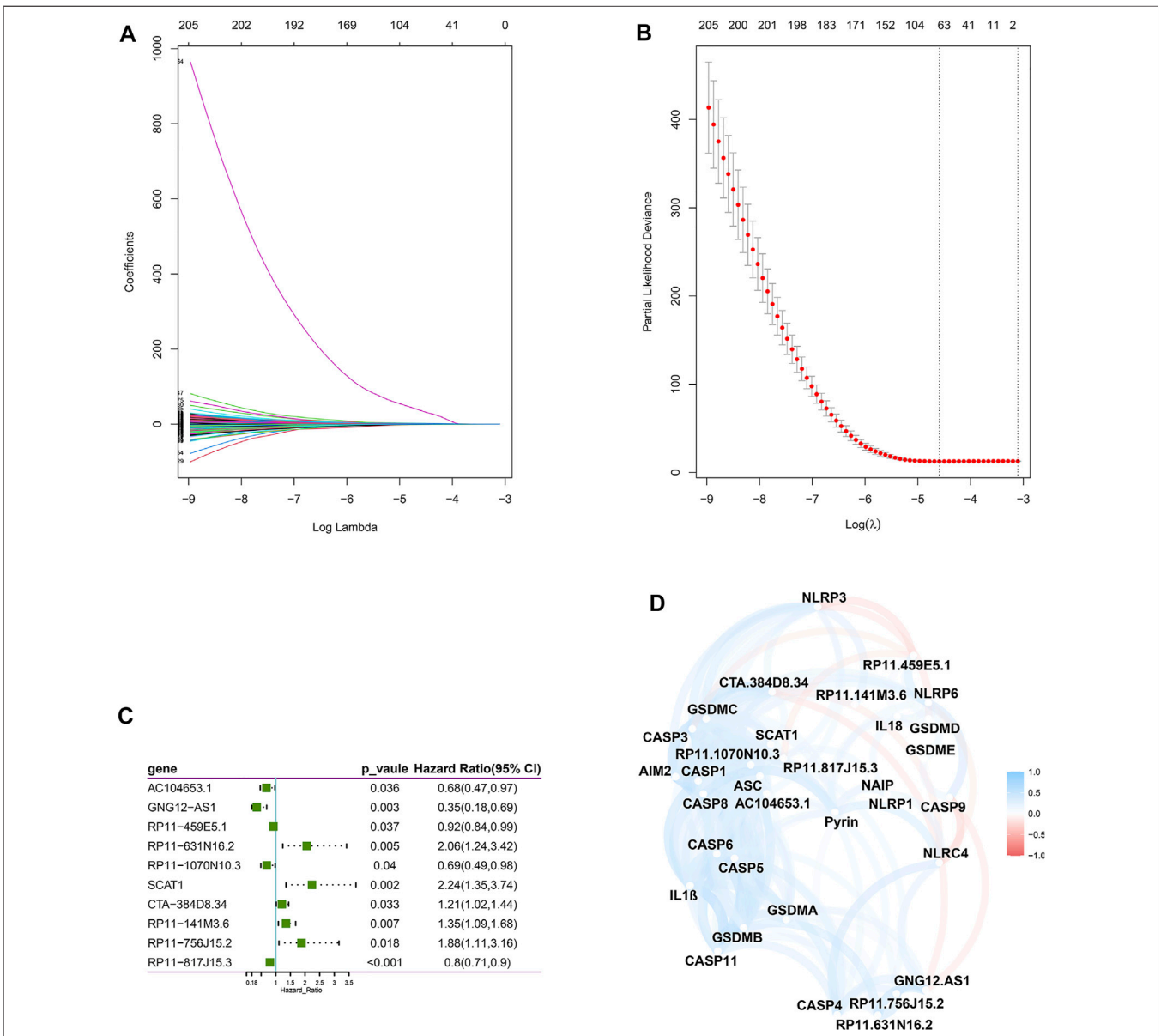


FIGURE 1 | Construction of a risk model for BC patients based on the pyroptosis-related lncRNAs. **(A,B)** The LASSO coefficient profile was constructed from 64 prognostic pyroptosis-related lncRNAs based on the minimum criteria for OS with 10-fold cross-validation. **(C)** Multivariate Cox regression analysis showed 10 independent prognostic pyroptosis-related lncRNAs. **(D)** Correlation network between the 23 pyroptosis genes and the 10 prognostic pyroptosis-related lncRNAs.

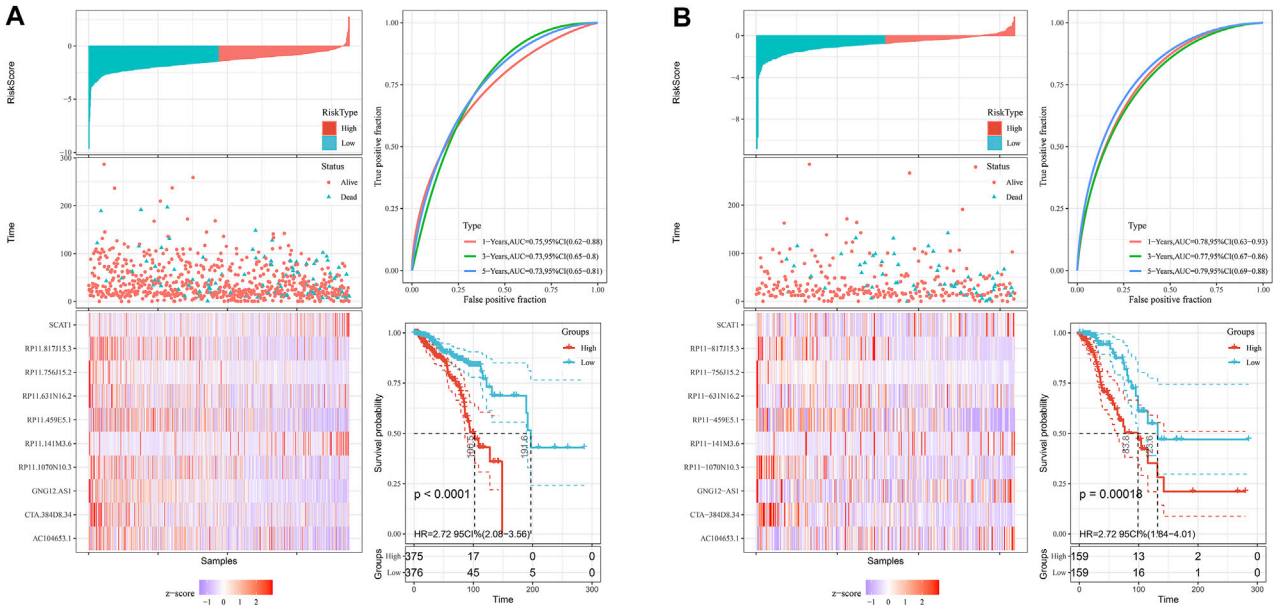


FIGURE 2 | The distribution of risk score, survival analysis, and the expression level of the 10 pyroptosis-related lncRNAs in the training **(A)** and testing sets **(B)**.

backgrounds (Xia et al., 2019; Lu et al., 2021). On the one hand, pyroptosis as a type of cell death can inhibit the pathogenesis and development of tumor (Nagarajan et al., 2019; Tan et al., 2021); on the other hand, pyroptosis can form a suitable microenvironment for tumorigenesis and chemotherapeutic resistance by releasing multiple inflammatory mediators (Lee et al., 2019; Zhou and Fang, 2019; Li et al., 2021).

Recently, solid evidence suggested that pyroptosis plays a vital role in various tumors by regulating tumor cell proliferation, invasion, and migration. Lu et al. found that GSDME-mediated pyroptosis contributed to the drug response in a subset of lung cancer models, including KRAS-mutant, EGFR-altered, and ALK-rearranged adenocarcinomas (Lu et al., 2018). IL18, a pyroptosis-related inflammatory mediator, was observed to exert inflammation-dependent tumor-suppressive effects by promoting the differentiation, activity, and survival of tumor-infiltrating T cells in hepatocellular carcinoma (Markowitz et al., 2016). In terms of the innate immune microenvironment in tumor, Storr et al. found that macrophage-derived IL-1β promoted the migration of breast cancer cell and lymphatic endothelial cell adhesion, then contributed to the tumorigenesis and metastasis of BC (Storr et al., 2017; Tulotta et al., 2019). Moreover, IL-1β activated IRAK4 in cancer-associated fibroblasts, and then drove tumor fibrosis, chemoresistance, and poor prognosis in pancreatic cancer (Zhang et al., 2018). These results pinpoint pyroptosis as an unrecognized mechanism involved in the tumorigenesis and development in various tumors, which may have important implications for the clinical development and optimal application of anticancer therapeutics.

Long non-coding RNAs (lncRNAs), a cluster of RNAs that have no protein-encoding ability, have been widely accepted to play an important role in the pathogenesis and development of cancer (Fatica and Bozzoni, 2014; Bhan et al., 2017; Lin and Yang, 2018). Several studies proved the association between lncRNAs and the progression of breast cancer. Wang et al. found that H19 induced autophagy activation *via* the H19/SAHH/DNMT3B axis, which contributed to tamoxifen resistance in BC (Wang et al., 2019). Using lncRNA/mRNA microarray assays, Qin et al. observed that lnc030 maintained breast cancer stem cell stemness and tumorigenesis by stabilizing SQLE mRNA and increasing cholesterol synthesis (Qin et al., 2021). A study conducted by Xiu et al. showed that LINC02273 drove breast cancer metastasis by epigenetically upregulated AGR2 (Xiu et al., 2019). However, few efforts have been devoted to the role of pyroptosis regulators in the dysregulation of lncRNAs in breast cancer. Thus, by performing a comprehensive bioinformatics analysis, we aimed to construct and validate a risk model based on the pyroptosis-related lncRNAs to predict prognosis and immune response in breast cancer.

MATERIALS AND METHODS

Data Acquisition and Preparation

The RNA sequence transcriptome profiling data and mutation data of BC patients with clinical features and survival information were downloaded from The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/repository). Then, using the Ensembl human genome browser (http://asia.ensembl.org/info/data/index.html) by the Perl program, the data were collated and annotated to protein-coding genes and lncRNAs. In total, 48,608 genes were annotated, in which 15,058 lncRNAs were identified.

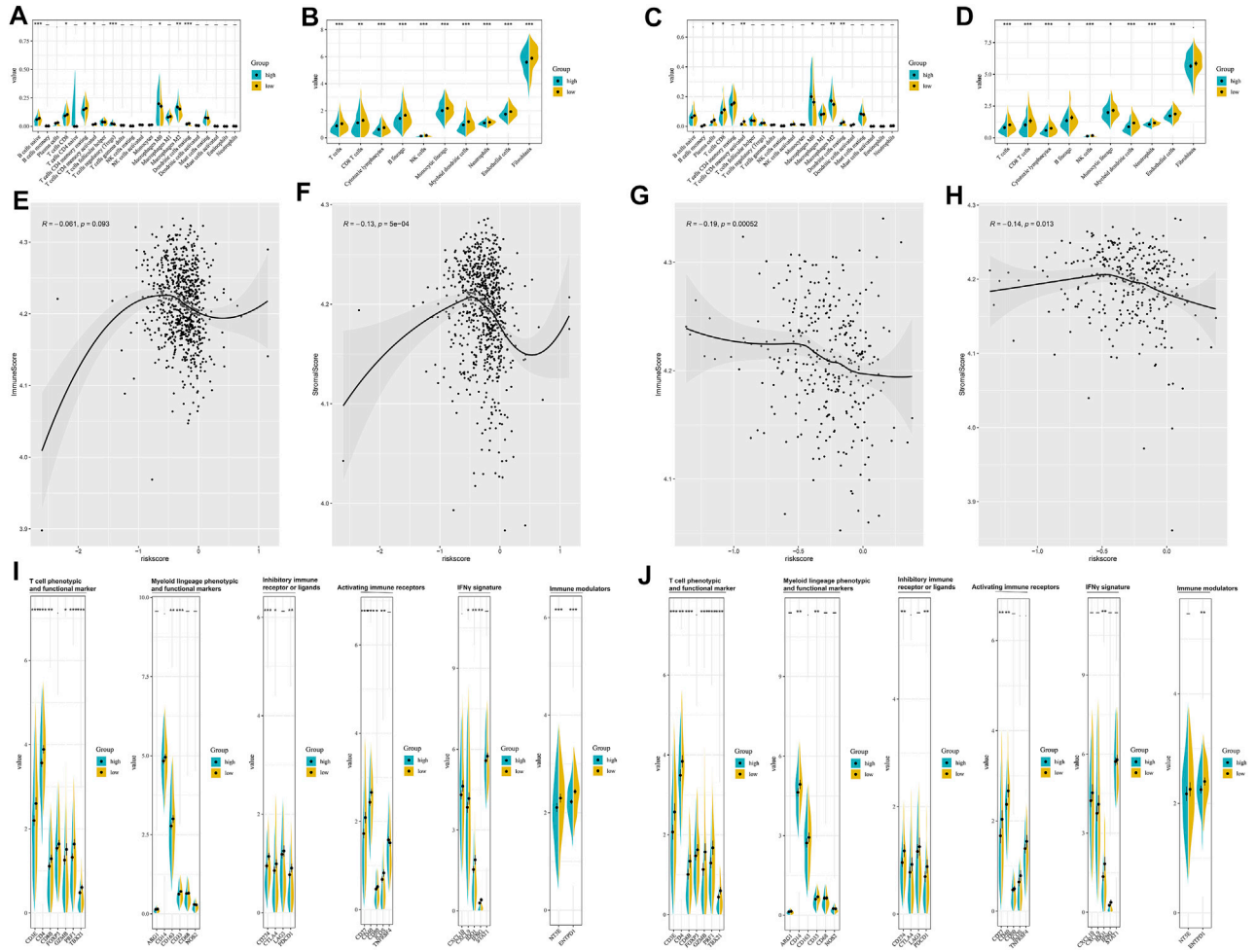


FIGURE 3 | Estimation of tumor immune microenvironment in the pyroptosis-related lncRNA-based risk groups. CIBERSORT analyses between different risk groups in the training set **(A)** and testing set **(B)**. MCP-counter analyses between different risk groups in the training set **(C)** and testing set **(D)**. ESTIMATE analyses between different risk groups in the training set **(E,F)** and testing set **(G,H)**. The expression of immune profiles between different risk groups in the training set **(I)** and testing set **(J)**.

Selection of Pyroptosis Genes and Pyroptosis-Related lncRNAs

Based on previous review (Liu et al., 2021; Yu et al., 2021a), 23 genes (CASP1, CASP3, CASP4, CASP5, CASP6, CASP8, CASP9, CASP11, GSDMA, GSDMB, GSMDC, GSDMD, GSDME, NAIP, NLRC4, NLRP1, NLRP3, NLRP6, Pyrin, ASC, IL1β, IL18, AIM2) were defined as pyroptosis-related regulators. Pyroptosis-related lncRNAs were defined as lncRNAs that were significantly related to one or more of the 23 pyroptosis genes (|Pearson R| > 0.3 and $p < 0.001$) (Tang et al., 2021; Xu et al., 2021). Finally, 3,364 lncRNAs were identified as pyroptosis-related lncRNAs.

Establishment and Verification of a Pyroptosis-Related lncRNA Signature

The entire TCGA set was randomized as a training set and a testing set with a ratio of 7:3. The training set (N = 764) was utilized to construct a risk model based on the pyroptosis-related lncRNAs, and the testing set (N = 327) was applied to validate this established model. Combined with BC survival information in TCGA, we screened the prognosis value of 3,364 pyroptosis-related lncRNAs in the training dataset by univariate Cox regression using the R package "survival". Then using the R package "glmnet" (Tibshirani, 1997), we selected the most robust prognostic pyroptosis-related lncRNAs in LASSO Cox regression. Subsequently, multivariable Cox regression was applied to find the independent prognostic lncRNAs for overall survival (OS). Finally, a 10-pyroptosis-related lncRNAs risk signature was ultimately established (Yu et al., 2021b).

The risk score formula was calculated as follows:

$$\text{Risk score} = \Sigma i\, Coefficient\,(lncRNA) \times Expression\,(lncRNA)^{[29]}$$

where coefficient (lncRNA) was the coefficient of lncRNAs correlated with survival and expression (lncRNA) was the

expression of lncRNAs. According to the median risk score, patients were divided into low- and high-risk groups.

Evaluation of the Immune Landscape

The abundance of 22 immunocytes between the low-risk and high-risk groups were calculated through CIBERSORT algorithm (Newman et al., 2015). The immune-related efficiency was estimated using the "MCPcounter" package (Becht et al., 2016). Immune and stromal scores of BC patients were estimated applying the "estimate" package (Yoshihara et al., 2013). Besides, the expression of key immune profiles between the low-risk and high-risk groups was compared using the Wilcoxon test.

Exploration of the Overall Gene Mutation and Tumor Mutation Burden in Different Risk Groups

Using the R package maftools (Mayakonda et al., 2018), the overall gene mutation status was analyzed and summarized in the high- and low-risk groups. Then, TMB scores based on the TGCA somatic mutation data were calculated to evaluate the mutation status between different risk groups.

Exploration of Potential Compounds Targeting the Pyroptosis-Related lncRNA Model

To identify potential drugs targeting the lncRNA-based risk model for treating BC patients, we estimated the therapeutic response based on the half-maximal inhibitory concentration (IC_{50}) of various molecular data available in the CellMiner database for each sample (Reinhold et al., 2012).

Functional Enrichment Analysis

Gene set enrichment analysis (GSEA) was performed to investigate the potential biological process and cellular pathway between the low- and high-risk groups through the Clusterprofile package (Subramanian et al., 2005). The FDR q <0.25 and p <0.05 were considered statistically significant.

Independence of the Pyroptosis-Related lncRNA Model

The Kaplan–Meier survival curve was performed to compare the survival diversities of the high-risk and low-risk groups. Univariable and multivariable Cox regression analyses were conducted to test whether the risk model we constructed was an independent risk factor for survival in BC patients.

Establishment and Validation of a Prognostic Nomogram

To predict the prognosis of BC patients, a nomogram based on risk model and other clinicopathologic features were constructed to predict the 1-, 3-, and 5-year OS using the R package "rms" (Zhang and Kattan, 2017). The concordance index (C-index) and calibration plots were applied to reflect the predictive accuracy of the prognostic nomogram we constructed. The area under the time-dependent receiver operating characteristic curve (AUC) were performed to evaluate the sensitivity and specificity of the prognostic nomogram in both the training and validation sets.

Validation of the Bioinformatics Results Using RT-qPCR Assay

A total of 133 paired BC tissues (T) and adjacent normal tissues (N) were obtained from Sir Run Run Shaw Hospital of Zhejiang University from 2014 to 2017. Total RNA was extracted using TRIzol reagent (Invitrogen, USA). Reverse transcription was conducted using PrimeScript RT MasterMix (Takara, China). qRT-PCR was performed using SYBR Green PCR MasterMix (Takara, China). The qRT-PCR primers are listed in **Supplementary Table S1.1**; target lncRNA expression was normalized to those of GAPDH. The workflow of this study is shown in **Supplementary Figure S1**.

RESULTS

Identification of Pyroptosis-Related lncRNAs

The matrix expression of 23 pyroptosis genes and 15,058 lncRNAs were abstracted from the TCGA database. Then, applying Pearson's correlation analysis with a criteria of |Pearson R| >0.3 and $p < 0.001$, 3,364 lncRNAs were identified closely related to the 23 pyroptosis-related regulators, and these lncRNAs were defined as pyroptosis-related lncRNAs.

Construction and Validation of a Risk Model Based on the Pyroptosis-Related lncRNAs

Next, we screened prognostic-related lncRNAs from 3,364 pyroptosis-related lncRNAs in the training set using univariate Cox regression analysis. In total, 381 pyroptosis-related lncRNAs were significantly correlated with OS. Sixty-four pyroptosis-related lncRNAs were selected by performing the LASSO Cox analysis (**Figures 1A,B**). Next, 10 pyroptosis-related lncRNAs were found as independent predictors for OS in the training set using multivariate Cox ratio hazard regression analysis (**Figure 1C**). Then, a risk model was constructed and patients were clustered into low- and high-risk groups based on their risk scores (**Supplementary Table S1.2**). In addition, the correlation between the pyroptosis-related lncRNAs and pyroptosis genes was analyzed, which is shown in **Figure 1D**. The distribution of risk score, survival analysis, and the expression level of the 10 pyroptosis-related lncRNAs in the training and testing sets are shown in **Figures 2A,B**.

Estimation of the Tumor Immune Microenvironment

The enrichment level and activity of infiltrating immune cells between the different risk groups were further analyzed. CIBERSORT algorithm (**Figures 3A,B**) confirmed that patients

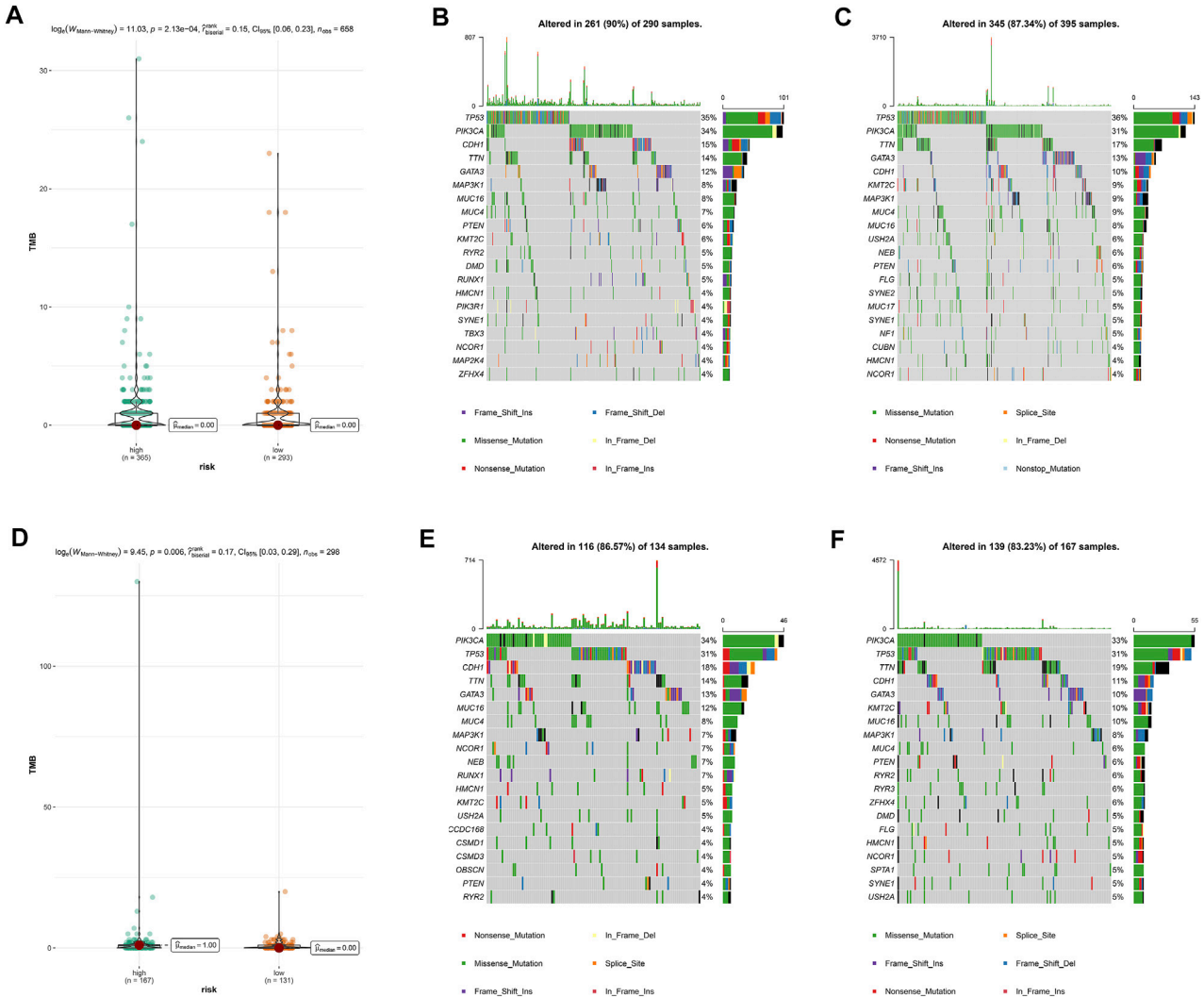


FIGURE 4 | Overall gene mutation and tumor mutation burden analysis based on the pyroptosis-related lncRNAs model. The difference of tumor mutational burden between the high- and low-risk subtypes in the training set **(A)** and testing set **(D)**. Waterfall plots of top 20 mutated genes in the low-/high-risk groups in the training set **(B,C)** and testing set **(E,F)**.

in the low-risk group were qualified with more antitumoral immune cells ($CD8^+$ T cells, activated memory $CD4^+$ T cells), while patients in the high-risk group were characteristics of more regulatory T cells and M2 macrophages. MCP-counter (**Figures 3C,D**) suggested that patients in the low-risk group had a higher level of activated immune cells ($CD8^+$ T cells, cytotoxic lymphocytes, B cells, and NK cells). ESTIMATE algorithm indicated that the risk model was negatively correlated with the immune score and the stromal score (**Figures 3E–H**). Besides, the low-risk and high-risk groups showed prominent differences in the expression of immune profiles (**Figures 3I,J**), and a higher expression of immune-related profiles, like T cell phenotypic and functional marker (CD3E, CD4, CD8B, FOXP3, GZMB, PRF1, and TBX21), activating immune receptors (CD27, CD40, CD80, ICOS, and TNFRSF4), IFNγ signature (CXCL9, CXCL10, IDO1, IFNG, and STAT1), and the immune checkpoint markers (CTLA4, CD274, PDCD1), was observed in the low-risk group, which indicated that the pyroptosis-based lncRNA signature might serve as an effective indicator for immunotherapeutic response.

Overall Gene Mutation and Tumor Mutation Burden Analysis

Then, we calculated TMB scores according to tumor-specific mutated genes. Patients in the high-risk group showed a significantly higher TMB than their counterpart in the low-risk group (**Figures 4A,D**). Using the R package maftools, we analyzed the overall gene mutation in different risk groups. The top 20 driver genes with the highest alteration frequency were depicted by waterfall plots in the low-/high-risk groups in the training set (**Figures 4B,C**) and testing set (**Figures 4E,F**).

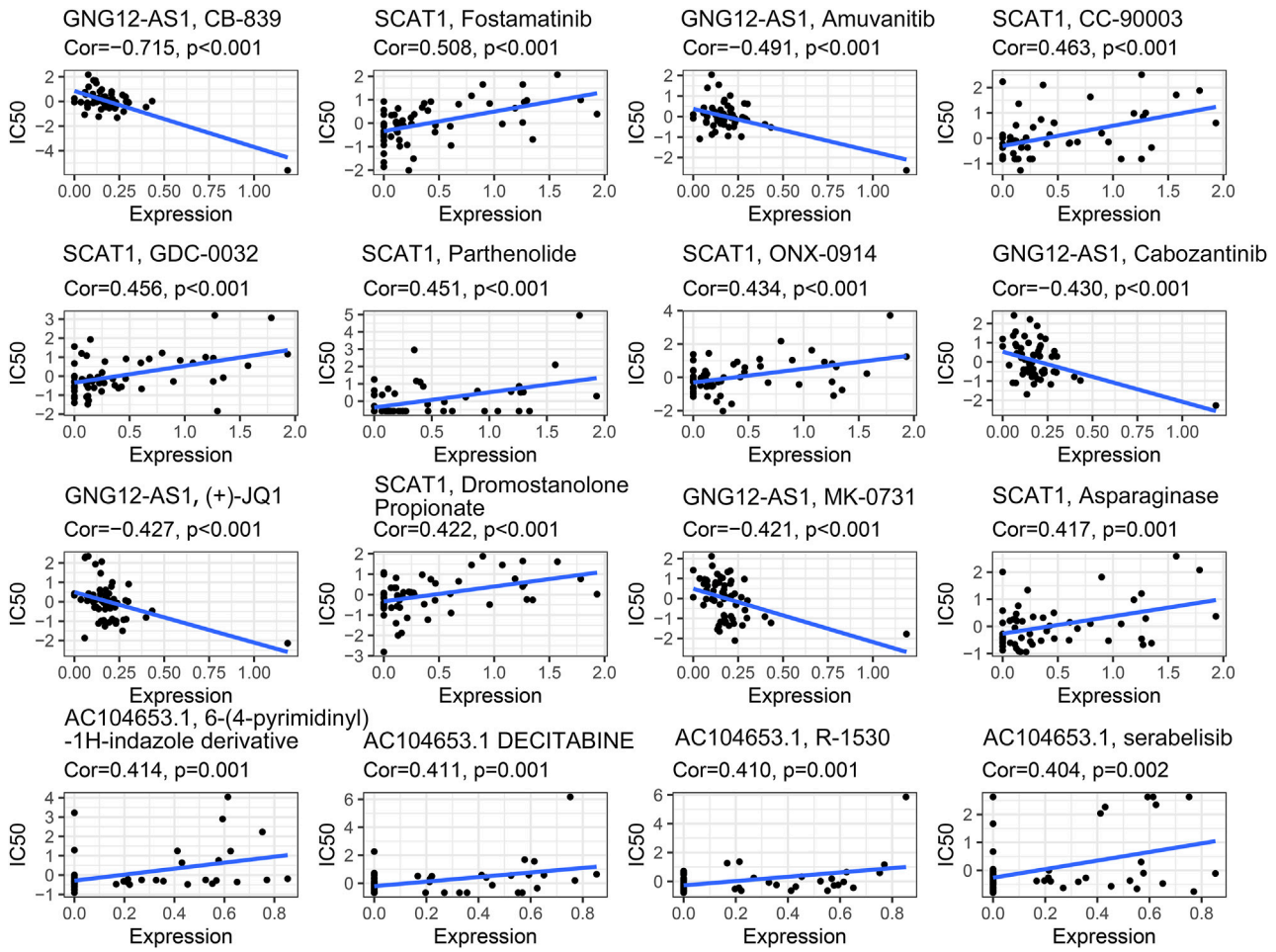


FIGURE 5 | Identification of novel candidate compounds targeting the pyroptosis-related lncRNA model.

Identification of Novel Candidate Compounds Targeting the Pyroptosis-Related lncRNA Model

To identify potential compounds for the treatment of BC, we calculated the IC_{50} of compounds obtained from the CellMiner database. Robust negative correlation has been found between the expression level of GNG12-AS1 with IC_{50} of CB-839, amuvanitib, cabozantinib, (+)-JQ1, and MK-0731 (all $p < 0.001$). A significant positive correlation was observed between the expression of SCAT1 and IC_{50} of fostamatinib, CC-90003, GDC-0032, parthenolide, ONX-0914, asparaginase, and dromostanolone propionate (all $p < 0.005$). The expression level of AC104653.1 was significantly positively associated with the IC_{50} of 6-(4-pyrimidinyl)-1H-indazole derivative, decitabine, R-1530, and serabelisib (all $p < 0.005$), which is shown in **Figure 5**.

Functional Enrichment Analysis

Gene set enrichment analysis (GSEA) revealed that apoptosis and JAK-STAT signaling pathway were significantly enriched in the high-risk group in both training and testing sets. In terms of gene annotation (GO) analysis, protein transmembrane import into intracellular organelle, mitochondrial membrane part, and plasma membrane receptor complex were the most relevant biological process (BP), cellular component (CC), and molecular function (MF) of pyroptosis-related lncRNAs, respectively. Pertaining to cancer hallmark, IL6-JAK-STAT3 signaling pathway and inflammatory response were the most relevant cancer hallmarks (**Figure 6**).

Evaluation of the Prognostic Value of the Risk Model

To further evaluate the prognostic value of the risk model, univariate and multivariate Cox regression models that contained the clinicopathologic features and the risk model were analyzed in the training and testing sets. As shown in **Supplementary Table S2**, the risk model was an independent risk factor for OS in BC patients. Moreover, the expression level and prognostic value of these pyroptosis-related lncRNAs were further analyzed using GEPIA dataset (http://gepia.cancer-pku.cn/). Kaplan–Meier survival analyses showed the survival

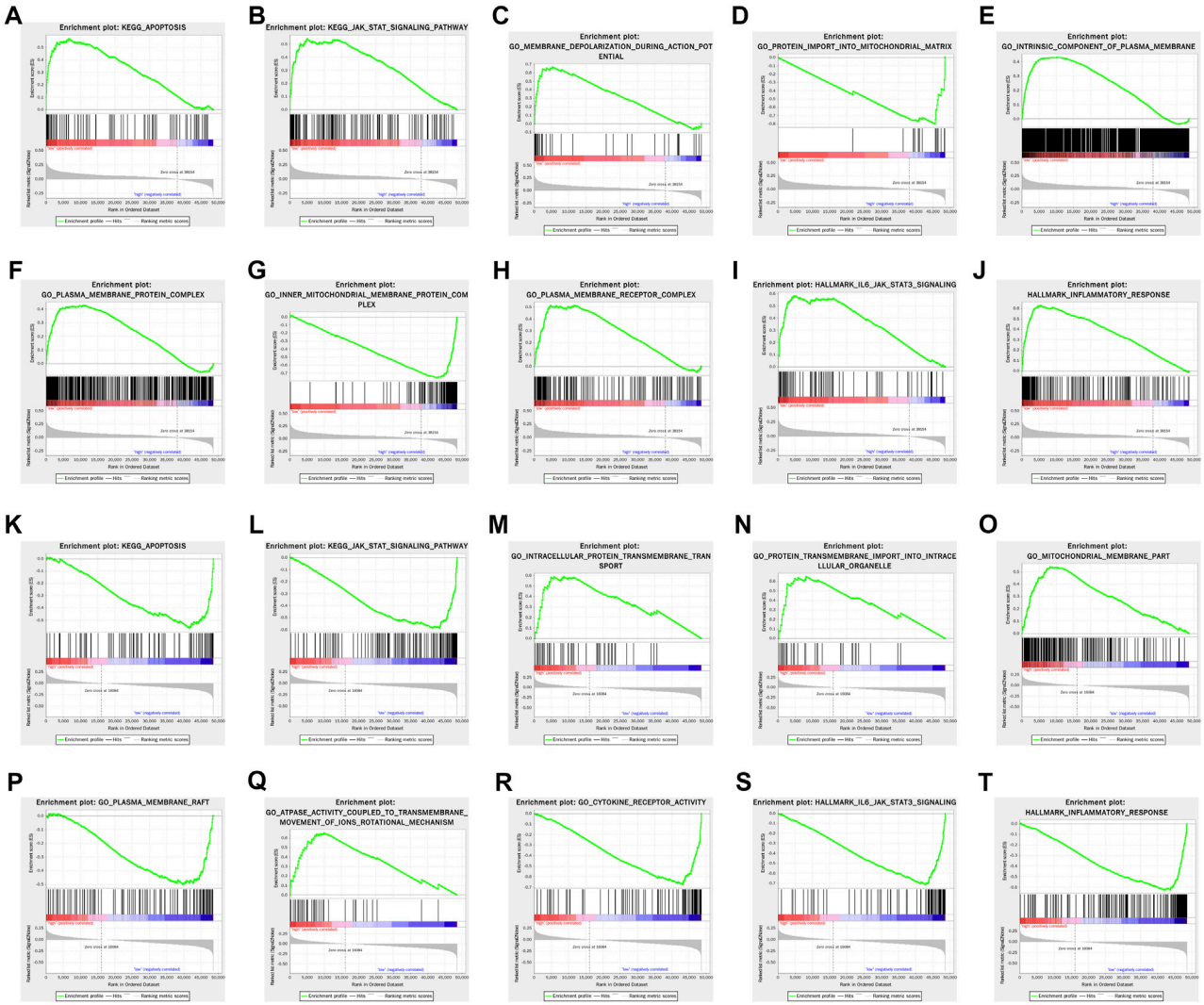


FIGURE 6 | GSEA of enriched potential biological process and cellular pathway between the high- and low-risk groups in the training set **(A–J)** and testing set **(K–T)**.

diversity of the pyroptosis-related lncRNAs (**Supplementary Figure S2**). The expression of these pyroptosis-related lncRNA signatures is depicted in **Supplementary Figure S2**. Results showed that RP11-756J15.2, RP11-1070N10.3, RP11-817J15.3, RP11-459E5.1, and RP11-141M3.6 were lowly expressed in tumor tissues and positively associated with OS in BC patients.

Construction and Evaluation of the Pyroptosis-Related lncRNA-Based Nomogram

A nomogram comprising the risk model and clinicopathologic features was established to predict the 1-, 3-, and 5-year OS. By comparison with other clinicopathologic parameters, the risk model showed predominant predictive ability in the nomogram (**Figure 7A**). The AUC value (**Figures 7B,C**) and calibration plots (**Figures 7D–I**) showed excellent consistency between the actual and nomogram-predicted survival probabilities for 1-, 3-, and 5-year OS in the training and testing sets.

The Expression Levels and Prognostic Value of 10 Selected Pyroptosis-Related lncRNAs in Our Cohort

The expression levels of 10 selected pyroptosis-related lncRNAs (AC104653.1, GNG12-AS1, RP11-141M3.6, RP11-631N16.2, RP11-459E5.1, RP11-756J15.2, RP11-817J15.3, RP11-1070N10.3, CTA-384D8.34, and CTD-2357A8.3) were examined by qRT-PCR. In detail, RP11-141M3.6, RP11-1070N10.3, and RP11-817J15.3 were downregulated, while RP11-459E5.1 was significantly upregulated in tumor tissues compared with that in the paired normal tissues (**Figure 8A**). The four differentially expressed lncRNAs were further analyzed

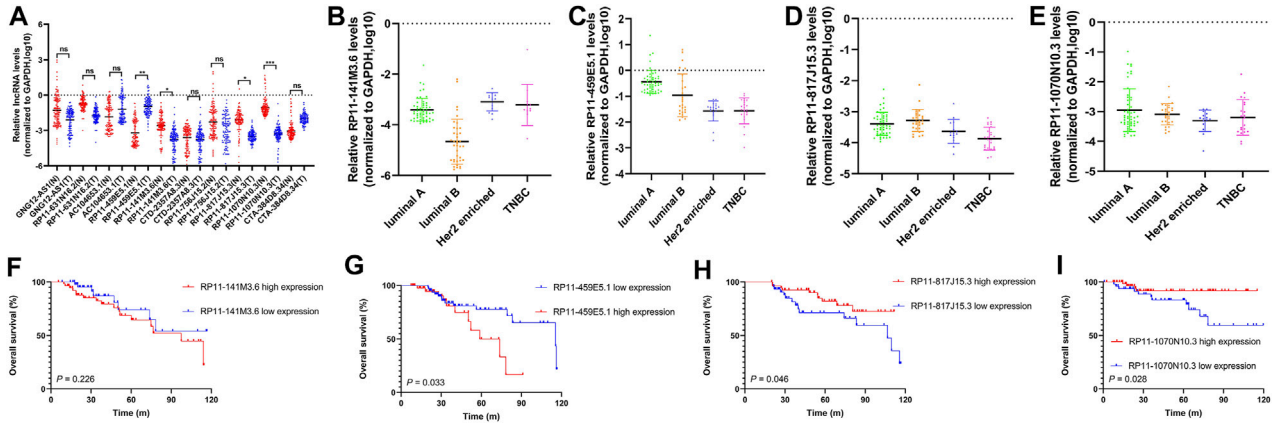


FIGURE 7 | The expression levels and prognostic value of 10 selected pyroptosis-related lncRNAs in our cohort. **(A)** Comparison of expression levels of 10 selected pyroptosis-related lncRNAs in tumor tissues (T) and adjacent normal tissues (N) by RT-qPCR assay. The expression level of RP11-141M3.6, RP11-459E5.1, RP11-817J15.3, and RP11-1070N10.3 in different breast cancer subtypes, respectively **(B–E)**. Kaplan–Meier curve shows the survival diversity between different expressions of RP11-141M3.6 **(F)**, RP11-459E5.1 **(G)**, RP11-817J15.3 **(H)**, and RP11-1070N10.3 **(I)** in our cohort. Non-significant (ns) $p > 0.05$, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

in different breast cancer subtypes (**Figures 8B–E**). Importantly, Kaplan–Meier survival analysis showed that high expression of RP11-459E5.1 was significantly associated with worse OS (**Figure 8G**), while RP11-1070N10.3 and RP11-817J15.3 high expression (**Figures 8H,I**) were correlated with better survival in breast cancer, which is in accord with the bioinformatics results.

DISCUSSION

Breast cancer is the most frequent malignant tumor and the leading cause of cancer-related death in women globally. Although dramatic improvements in early diagnosis and effective treatment have been made for this malignant tumor, a considerable proportion of patients still succumb to metastasis and recurrence due to therapeutic failure. It is well established that breast cancer displays extreme heterogeneity in histology and molecular analysis, which contributes to significant diversities in incidence, malignant procession, treatment response, and prognosis (Holm et al., 2017; Yeo and Guan, 2017). Thus, more insights are required to identify critical signaling molecules that contribute to the tumorigenesis and malignant procedure of breast cancer.

Pyroptosis is a novel type of programmed cell death characterized by gasdermin-mediated and proinflammatory factor release (Shi et al., 2015). Increasing evidence suggested that pyroptosis may play a critical role in the pathogenesis and development of various tumors, including BC. However, studies on the pathological role of pyroptosis in BC progression remain limited. Similarly, lncRNAs, the largest class of noncoding RNA, modulate chromatin functions through interactions with DNA, RNA, and proteins (Qian et al., 2019; Statello et al., 2021). Numerous studies have explored the correlation of lncRNA with various cancers, including BC. However, studies on the biological mechanisms and prognostic biomarkers of BC concerning pyroptosis-related lncRNAs are still lacking.

In the present study, we were inspired by the biological function of pyroptosis and lncRNAs in BC; thus, we attempted to construct a risk model based on pyroptosis-related lncRNAs for predicting prognosis and immune response in BC patients. First, we identified 3,364 pyroptosis-related lncRNAs in the TCGA cohort. Then, using univariate and multivariate Cox regression analyses, and LASSO cox regression analysis, a risk model based on 10 pyroptosis-related lncRNAs were constructed and patients were classified into high- and low-risk groups based on the median risk score. Results showed that the high-risk group had apparently poorer OS than the low-risk group. Multivariate Cox regression analysis showed that the pyroptosis-related lncRNA model was an independent risk factor of OS. Then, a nomogram was established to predict the 1-year, 3-year, and 5-year OS of BC patients. The calibration plots and AUC value showed excellent consistency between the actual and nomogram-predicted survival probabilities for 1-, 3-, and 5-year OS in the training and testing sets. Moreover, the expression level and clinical significance of the selected pyroptosis-related lncRNAs were further validated in our cohort, which is in accord with the bioinformatics results.

Next, we explored the relationship of the risk model and tumor microenvironment in BC. Results showed that patients in the low-risk group had a higher expression of immune-checkpoint markers (like CTLA4, CD274, and PDCD1) as compared with the high-risk group. Similarly, we observed that the low-risk group had higher infiltration levels of activated immune cells than their counterpart ($CD8^{+}$ T cells, B cells, NK cells, etc.). The results were consistent with previous findings that immune checkpoints executed a vital role on tumorigenesis and development in tumors by inducing tumor immune-suppressive activities, and patients with high PD-L1 expression in tumor cells and stromal immune cells are more likely to respond to chemotherapy and immunotherapy (Gibson, 2015; Kwapisz, 2021). Liu et al. reported that the aberrant expression of CTLA4 and PDCD1 was associated with tumorigenesis and immunocyte infiltration

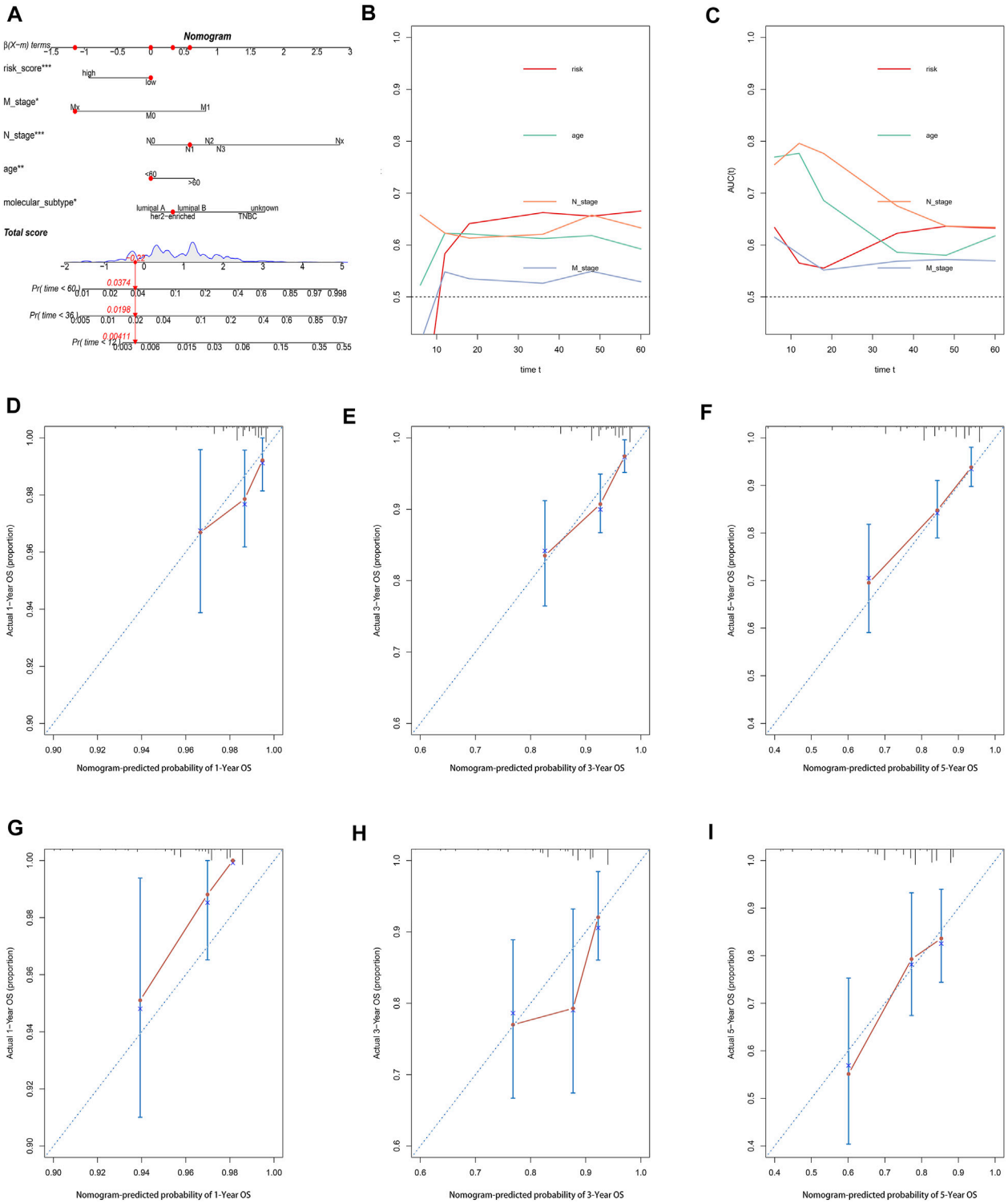


FIGURE 8 | Construction and evaluation of a pyroptosis-related lncRNA-based nomogram. **(A)** The nomogram predicts the probability of 1-, 3-, and 5-year OS for individual patients. AUC values of the risk score and clinical characteristics in the training set **(B)** and testing set **(C)**. Calibration plots evaluate the prediction accuracy of the nomogram in the training set **(D–F)** and testing set **(G–I)**.

in pan-cancer, including breast cancer (Liu et al., 2020). Another study conducted by Park et al. showed that the expression level of CD274 was associated with prognosis in breast cancer patients who received neoadjuvant chemotherapy (Park et al., 2020). Moreover, Tekpli et al. (2019) found an immune infiltration–based subtype of breast cancer to predict therapeutic response and prognosis in breast cancer patients.

Furthermore, we estimated the TMB and overall gene mutations in different risk groups. TMB is defined as the total number of somatic mutations per megabase of interrogated genomic sequence, which related to the emergence of neoantigens that trigger antitumor immunity (Sha et al., 2020). Recent studies revealed that breast cancer patients with high TMB were more likely to benefit from PD-L1 inhibitors (Thomas et al., 2018; Barroso-Sousa et al., 2020; Chumsri et al., 2020). The results of our study showed that the difference in the amount of overall gene mutations between the high- and low-risk groups, and the TMB in the low-risk group exceeded that in high-risk group. In terms of chemotherapy response, the expression of GNG12-AS1 was negatively associated with IC_{50} of CB-839, amuvanitib, cabozantinib, (+)-JQ1, and MK-0731. Also, significant positive correlations were observed between the expression of SCAT1(CTD-2357A8.3) and AC104653.1 with IC_{50} of other agents, which may be potential compounds for the treatment of BC by targeting these specific pyroptosis-related lncRNAs.

In addition, as revealed in the GSEA results, the tumor functional patterns including apoptosis and JAK-STAT signaling pathways were enriched in the high-risk group. Pertaining to cancer hallmark, IL6-JAK-STAT3 signaling pathway and inflammatory response were the most relevant cancer hallmark. It is well established that JAK-STAT signaling is involved in breast cancer cell proliferation, metastasis, and chemotherapeutic sensitiveness. For instance, Wang et al. reported that CircNOL10 suppressed breast cancer progression by sponging miR-767-5p to regulate SOCS2/JAK/STAT signaling (Wang et al., 2021). Several studies found that JAK-STAT signaling was involved in the chemotherapeutic and endocrine therapeutic resistance in breast cancer (Lui et al., 2017; Zhu et al., 2020). Considering that, the 10 pyroptosis-related lncRNAs and their relative pathways may be involved in the tumorigenesis and development of breast cancer.

Among the lncRNA signatures, GNG12-AS1 has been found to coregulate with DIRAS3, and then inhibit cell cycle progression and migration in various tissues (Stojic et al., 2016). Moreover, GNG12-AS1 caused allele-specific splicing in breast cancer, which may contribute to the tumorigenesis and development of breast cancer (Niemczyk et al., 2013). Xiang et al. observed that GNG12-AS1 induced glioma cell proliferation and migration through AKT/GSK-3β/β-catenin signaling (Xiang et al., 2020). Zheng et al. found that SCAT1(CTD-2357A8.3) served as a predictive biomarker for pathologic complete response of chemotherapy in esophageal squamous cell carcinoma (Zhang et al., 2020). A previous study by Fan et al. found that SCAT1(CTD-2357A8.3) could be a novel prognostic biomarker for esophageal cancer (Fan and Liu, 2016). Similarly, Lei et al. established a risk model, which contained AC104653.1 to predict the prognosis of glioblastoma, and results showed that the model was a powerful tool for survival prediction in this malignant tumor (Lei et al., 2018). Furthermore, pan-cancer analysis of S-phase enriched lncRNAs identified that SCAT1(CTD-2357A8.3) was differentially expressed in several cancers, SCAT1(CTD-2357A8.3) induced cell proliferation and correlated with poor prognosis in lung cancer (Ali et al., 2018), and these publications provided a novel biological function and mechanism of these lncRNAs in tumors.

It is an undeniable fact that several limitations existed in this study. Due to lack of available data about lncRNAs in other databases, like the Gene Expression Omnibus (GEO) database, we could not validate the results of our study in other public datasets. In this background, we collected BC samples in our cohort to further explore the expression level and clinical significance of these pyroptosis-related lncRNAs, which validated the clinical significance of the selected pyroptosis-related lncRNAs. However, further experimental studies are needed to elucidate the underlying biological function and mechanism of these pyroptosis-related lncRNAs in breast cancer.

In conclusion, the established pyroptosis-related lncRNA model provides a new method for prognostic prediction in BC patients, and may help elucidate the important role of these pyroptosis-related lncRNAs in the tumorigenesis and development of breast cancer. In addition, our study provides new insight in identifying BC patients who may benefit from immunotherapy.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

XY designed the study, analyzed data, and wrote the article. ZJ supervised the research. All authors read and approved the final submitted article.

ACKNOWLEDGMENTS

We greatly acknowledge TCGA database for providing their platforms and contributors for uploading their meaningful datasets.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.792106/full#supplementary-material

Supplementary Figure S1 | Kaplan–Meier curves to depict the survival diversity of the pyroptosis-related lncRNAs.

Supplementary Figure S2 | The expression of pyroptosis-related lncRNAs signature in the TCGA dataset.

REFERENCES

Ali, M. M., Akhade, V. S., Kosalai, S. T., Subhash, S., Statello, L., Meryet-Figuiere, M., et al. (2018). PAN-Cancer Analysis of S-Phase Enriched lncRNAs Identifies Oncogenic Drivers and Biomarkers. *Nat. Commun.* 9, 883. doi:10.1038/s41467-018-03265-1

Barroso-Sousa, R., Jain, E., Cohen, O., Kim, D., Buendia-Buendia, J., Winer, E., et al. (2020). Prevalence and Mutational Determinants of High Tumor Mutation Burden in Breast Cancer. *Ann. Oncol.* 31, 387–394. doi:10.1016/j.annonc.2019.11.010

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. *Genome Biol.* 17, 218. doi:10.1186/s13059-016-1070-5

Bhan, A., Soleimani, M., and Mandal, S. S. (2017). Long Noncoding RNA and Cancer: A New Paradigm. *Cancer Res.* 77, 3965–3981. doi:10.1158/0008-5472.can-16-2634

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

Chumsri, S., Sokol, E. S., Soyano-Muller, A. E., Parrondo, R. D., Reynolds, G. A., Nassar, A., et al. (2020). Durable Complete Response with Immune Checkpoint Inhibitor in Breast Cancer with High Tumor Mutational Burden and APOBEC Signature. *J. Natl. Compr. Canc Netw.* 18, 517–521. doi:10.6004/jnccn.2020.7543

Fan, Q., and Liu, B. (2016). Identification of a RNA-Seq Based 8-Long Non-Coding RNA Signature Predicting Survival in Esophageal Cancer. *Med. Sci. Monit.* 22, 5163–5172. doi:10.12659/msm.902615

Fatica, A., and Bozzoni, I. (2014). Long Non-coding RNAs: New Players in Cell Differentiation and Development. *Nat. Rev. Genet.* 15, 7–21. doi:10.1038/nrg3606

Gibson, J. (2015). Anti-PD-L1 for Metastatic Triple-Negative Breast Cancer. *Lancet Oncol.* 16, e264. doi:10.1016/s1470-2045(15)70208-1

Harbeck, N., and Gnant, M. (2017). Breast Cancer. *The Lancet* 389, 1134–1150. doi:10.1016/s0140-6736(16)31891-8

Holm, J., Eriksson, L., Ploner, A., Eriksson, M., Rantalainen, M., Li, J., et al. (2017). Assessment of Breast Cancer Risk Factors Reveals Subtype Heterogeneity. *Cancer Res.* 77, 3708–3717. doi:10.1158/0008-5472.can-16-2574

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

Lee, C., Do, H. T. T., Her, J., Kim, Y., Seo, D., and Rhee, I. (2019). Inflammasome as a Promising Therapeutic Target for Cancer. *Life Sci.* 231, 116593. doi:10.1016/j.lfs.2019.116593

Lei, B., Yu, L., Jung, T., Deng, Y., Xiang, W., Liu, Y., et al. (2018). Prospective Series of Nine Long Noncoding RNAs Associated with Survival of Patients with Glioblastoma. *J. Neurol. Surg. A. Cent. Eur. Neurosurg.* 79, 471–478. doi:10.1055/s-0038-1655549

Li, L., Jiang, M., Qi, L., Wu, Y., Song, D., Gan, J., et al. (2021). Pyroptosis, a New Bridge to Tumor Immunity. *Cancer Sci.* 112 (10), 3979–3994. doi:10.1111/cas.15059

Lin, C., and Yang, L. (2018). Long Noncoding RNA in Cancer: Wiring Signaling Circuitry. *Trends Cel. Biol.* 28, 287–301. doi:10.1016/j.tcb.2017.11.008

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

Liu, X., Xia, S., Zhang, Z., Wu, H., and Lieberman, J. (2021). Channelling Inflammation: Gasdermins in Physiology and Disease. *Nat. Rev. Drug Discov.* 20, 384–405. doi:10.1038/s41573-021-00154-z

Lu, H., Zhang, S., Wu, J., Chen, M., Cai, M.-C., Fu, Y., et al. (2018). Molecular Targeted Therapies Elicit Concurrent Apoptotic and GSDME-Dependent Pyroptotic Tumor Cell Death. *Clin. Cancer Res.* 24, 6066–6077. doi:10.1158/1078-0432.ccr-18-1478

Lu, X., Guo, T., and Zhang, X. (2021). Pyroptosis in Cancer: Friend or Foe? *Cancers (Basel)* 13, 3620. doi:10.3390/cancers13143620

Lui, A. J., Geanes, E. S., Ogony, J., Behbod, F., Marquess, J., Valdez, K., et al. (2017). IFITM1 Suppression Blocks Proliferation and Invasion of Aromatase Inhibitor-Resistant Breast Cancer *In Vivo* by JAK/STAT-Mediated Induction of P21. *Cancer Lett.* 399, 29–43. doi:10.1016/j.canlet.2017.04.005

Markowitz, G. J., Yang, P., Fu, J., Michelotti, G. A., Chen, R., Sui, J., et al. (2016). Inflammation-Dependent IL18 Signaling Restricts Hepatocellular Carcinoma Growth by Enhancing the Accumulation and Activity of Tumor-Infiltrating Lymphocytes. *Cancer Res.* 76, 2394–2405. doi:10.1158/0008-5472.can-15-1548

Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res.* 28, 1747–1756. doi:10.1101/gr.239244.118

Nagarajan, K., Soundarapandian, K., Thorne, R. F., Li, D., and Li, D. (2019). Activation of Pyroptotic Cell Death Pathways in Cancer: An Alternative Therapeutic Approach. *Translational Oncol.* 12, 925–931. doi:10.1016/j.tranon.2019.04.010

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

Niemczyk, M., Ito, Y., Huddleston, J., Git, A., Abu-Amero, S., Caldas, C., et al. (2013). Imprinted Chromatin Around DIRAS3 Regulates Alternative Splicing of GNG12-AS1, a Long Noncoding RNA. *Am. J. Hum. Genet.* 93, 224–235. doi:10.1016/j.ajhg.2013.06.010

Park, Y. H., Lal, S., Lee, J. E., Choi, Y.-L., Wen, J., Ram, S., et al. (2020). Chemotherapy Induces Dynamic Immune Responses in Breast Cancers that Impact Treatment Outcome. *Nat. Commun.* 11, 6175. doi:10.1038/s41467-020-19933-0

Qian, X., Zhao, J., Yeung, P. Y., Zhang, Q. C., and Kwok, C. K. (2019). Revealing lncRNA Structures and Interactions by Sequencing-Based Approaches. *Trends Biochem. Sci.* 44, 33–52. doi:10.1016/j.tibs.2018.09.012

Qin, Y., Hou, Y., Liu, S., Zhu, P., Wan, X., Zhao, M., et al. (2021). A Novel Long Non-Coding RNA Lnc030 Maintains Breast Cancer Stem Cell Stemness by Stabilizing SQLE mRNA and Increasing Cholesterol Synthesis. *Adv. Sci.* 8, 2002232. doi:10.1002/advs.202002232

Reinhold, W. C., Sunshine, M., Liu, H., Varma, S., Kohn, K. W., Morris, J., et al. (2012). CellMiner: A Web-Based Suite of Genomic and Pharmacologic Tools to Explore Transcript and Drug Patterns in the NCI-60 Cell Line Set. *Cancer Res.* 72, 3499–3511. doi:10.1158/0008-5472.can-12-1370

Sha, D., Jin, Z., Budczies, J., Kluck, K., Stenzinger, A., and Sinicrope, F. A. (2020). Tumor Mutational Burden as a Predictive Biomarker in Solid Tumors. *Cancer Discov.* 10, 1808–1825. doi:10.1158/2159-8290.cd-20-0522

Shi, J., Zhao, Y., Wang, K., Shi, X., Wang, Y., Huang, H., et al. (2015). Cleavage of GSDMD by Inflammatory Caspases Determines Pyroptotic Cell Death. *Nature* 526, 660–665. doi:10.1038/nature15514

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

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

Stojic, L., Niemczyk, M., Orjalo, A., Ito, Y., Ruijter, A. E. M., Uribe-Lewis, S., et al. (2016). Transcriptional Silencing of Long Noncoding RNA GNG12-AS1 Uncouples its Transcriptional and Product-Related Functions. *Nat. Commun.* 7, 10406. doi:10.1038/ncomms10406

Storr, S. J., Safuan, S., Ahmad, N., El-Refaee, M., Jackson, A. M., and Martin, S. G. (2017). Macrophage-Derived Interleukin-1beta Promotes Human Breast Cancer Cell Migration and Lymphatic Adhesion *In Vitro*. *Cancer Immunol. Immunother.* 66, 1287–1294. doi:10.1007/s00262-017-2020-0

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

Tan, Y., Chen, Q., Li, X., Zeng, Z., Xiong, W., Li, G., et al. (2021). Pyroptosis: A New Paradigm of Cell Death for Fighting against Cancer. *J. Exp. Clin. Cancer Res.* 40, 153. doi:10.1186/s13046-021-01959-x

Tang, Y., Li, C., Zhang, Y.-J., and Wu, Z.-H. (2021). Ferroptosis-Related Long Non-Coding RNA Signature Predicts the Prognosis of Head and Neck Squamous Cell Carcinoma. *Int. J. Biol. Sci.* 17, 702–711. doi:10.7150/ijbs.55552

Tekpli, X., Lien, T., Lien, T., Røssevold, A. H., Nebdal, D., Borgen, E., et al. (2019). An Independent Poor-Prognosis Subtype of Breast Cancer Defined by a

Distinct Tumor Immune Microenvironment. *Nat. Commun.* 10, 5499. doi:10.1038/s41467-019-13329-5

Thomas, A., Routh, E. D., Pullikuth, A., Jin, G., Su, J., Chou, J. W., et al. (2018). Tumor Mutational burden Is a Determinant of Immune-Mediated Survival in Breast Cancer. *Oncoimmunology* 7, e1490854. doi:10.1080/2162402x.2018.1490854

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

Tulotta, C., Lefley, D. V., Freeman, K., Gregory, W. M., Hanby, A. M., Heath, P. R., et al. (2019). Endogenous Production of IL1B by Breast Cancer Cells Drives Metastasis and Colonization of the Bone Microenvironment. *Clin. Cancer Res.* 25, 2769–2782. doi:10.1158/1078-0432.ccr-18-2202

Vande Walle, L., and Lamkanfi, M. (2016). Pyroptosis. *Curr. Biol.* 26, R568–R572. doi:10.1016/j.cub.2016.02.019

Wang, J., Xie, S., Yang, J., Xiong, H., Jia, Y., Zhou, Y., et al. (2019). The Long Noncoding RNA H19 Promotes Tamoxifen Resistance in Breast Cancer via Autophagy. *J. Hematol. Oncol.* 12, 81. doi:10.1186/s13045-019-0747-0

Wang, F., Wang, X., Li, J., Lv, P., Han, M., Li, L., et al. (2021). CircNOL10 Suppresses Breast Cancer Progression by Sponging miR-767-5p to Regulate SOCS2/JAK/STAT Signaling. *J. Biomed. Sci.* 28, 4. doi:10.1186/s12929-020-00697-0

Xia, X., Wang, X., Cheng, Z., Qin, W., Lei, L., Jiang, J., et al. (2019). The Role of Pyroptosis in Cancer: Pro-Cancer or Pro-"Host"? *Cell Death Dis.* 10, 650. doi:10.1038/s41419-019-1883-8

Xiang, Z., Lv, Q., Chen, X., Zhu, X., Liu, S., Li, D., et al. (2020). Lnc GNG12-AS1 Knockdown Suppresses Glioma Progression through the AKT/GSK-3β/β-catenin Pathway. *Biosci. Rep.* 40 (8), BSR20201578. doi:10.1042/BSR20201578

Xiu, B., Chi, Y., Liu, L., Chi, W., Zhang, Q., Chen, J., et al. (2019). LINC02273 Drives Breast Cancer Metastasis by Epigenetically Increasing AGR2 Transcription. *Mol. Cancer* 18, 187. doi:10.1186/s12943-019-1115-y

Xu, F., Huang, X., Li, Y., Chen, Y., and Lin, L. (2021). m6A-Related lncRNAs Are Potential Biomarkers for Predicting Prognoses and Immune Responses in Patients with LUAD. *Mol. Ther. - Nucleic Acids* 24, 780–791. doi:10.1016/j.omtn.2021.04.003

Yeo, S. K., and Guan, J.-L. (2017). Breast Cancer: Multiple Subtypes within a Tumor? *Trends Cancer* 3, 753–760. doi:10.1016/j.trecan.2017.09.001

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

Yu, P., Zhang, X., Liu, N., Tang, L., Peng, C., and Chen, X. (2021). Pyroptosis: Mechanisms and Diseases. *Sig. Transduct. Target. Ther.* 6, 128. doi:10.1038/s41392-021-00507-5

Yu, J., Mao, W., Sun, S., Hu, Q., Wang, C., Xu, Z., et al. (2021). Identification of an m6A-Related lncRNA Signature for Predicting the Prognosis in Patients with Kidney Renal Clear Cell Carcinoma. *Front. Oncol.* 11, 663263. doi:10.3389/fonc.2021.663263

Zhang, Z., and Kattan, M. W. (2017). Drawing Nomograms with R: Applications to Categorical Outcome and Survival Data. *Ann. Transl. Med.* 5, 211. doi:10.21037/atm.2017.04.01

Zhang, D., Li, L., Jiang, H., Li, Q., Wang-Gillam, A., Yu, J., et al. (2018). Tumor-Stroma IL1β-IRAK4 Feedforward Circuitry Drives Tumor Fibrosis, Chemoresistance, and Poor Prognosis in Pancreatic Cancer. *Cancer Res.* 78, 1700–1712. doi:10.1158/0008-5472.can-17-1366

Zhang, C., Zhang, Z., Zhang, G., Xue, L., Yang, H., Luo, Y., et al. (2020). A Three-lncRNA Signature of Pretreatment Biopsies Predicts Pathological Response and Outcome in Esophageal Squamous Cell Carcinoma with Neoadjuvant Chemoradiotherapy. *Clin. Transl. Med.* 10, e156. doi:10.1002/ctm2.156

Zhou, C.-B., and Fang, J.-Y. (2019). The Role of Pyroptosis in Gastrointestinal Cancer and Immune Responses to Intestinal Microbial Infection. *Biochim. Biophys. Acta (Bba) - Rev. Cancer* 1872, 1–10. doi:10.1016/j.bbcan.2019.05.001

Zhu, N., Zhang, J., Du, Y., Qin, X., Miao, R., Nan, J., et al. (2020). Loss of ZIP Facilitates JAK2-STAT3 Activation in Tamoxifen-Resistant Breast Cancer. *Proc. Natl. Acad. Sci. USA* 117, 15047–15054. doi:10.1073/pnas.1910278117

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 | Frontiers in Genetics

ORIGINAL RESEARCH
published: 01 April 2022
doi: 10.3389/fgene.2022.806607

The Role of Copy Number Variants in Gene Co-Expression Patterns for Luminal B Breast Tumors

Candelario Hernández-Gómez[1], *Enrique Hernández-Lemus*[1,2]* and *Jesús Espinal-Enríquez*[1,2]*

[1]Computational Genomics Division, National Institute of Genomic Medicine, Mexico City, Mexico, [2]Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Mexico City, Mexico

Gene co-expression networks have become a usual approach to integrate the vast amounts of information coming from gene expression studies in cancer cohorts. The reprogramming of the gene regulatory control and the molecular pathways depending on such control are central to the characterization of the disease, aiming to unveil the consequences for cancer prognosis and therapeutics. There is, however, a multitude of factors which have been associated with anomalous control of gene expression in cancer. In the particular case of co-expression patterns, we have previously documented a phenomenon of loss of long distance co-expression in several cancer types, including breast cancer. Of the many potential factors that may contribute to this phenomenology, copy number variants (CNVs) have been often discussed. However, no systematic assessment of the role that CNVs may play in shaping gene co-expression patterns in breast cancer has been performed to date. For this reason we have decided to develop such analysis. In this study, we focus on using probabilistic modeling techniques to evaluate to what extent CNVs affect the phenomenon of long/short range co-expression in Luminal B breast tumors. We analyzed the co-expression patterns in chromosome 8, since it is known to be affected by amplifications/deletions during cancer development. We found that the CNVs pattern in chromosome 8 of Luminal B network does not alter the co-expression patterns significantly, which means that the co-expression program in this cancer phenotype is not determined by CNV structure. Additionally, we found that region 8q24.3 is highly dense in interactions, as well as region p21.3. The most connected genes in this network belong to those cytobands and are associated with several manifestations of cancer in different tissues. Interestingly, among the most connected genes, we found MAF1 and POLR3D, which may constitute an axis of regulation of gene transcription, in particular for non-coding RNA species. We believe that by advancing on our knowledge of the molecular mechanisms behind gene regulation in cancer, we will be better equipped, not only to understand tumor biology, but also to broaden the scope of diagnostic, prognostic and therapeutic interventions to ultimately benefit oncologic patients.

Keywords: breast cancer, co-expression networks, conditional mutual information, copy number variants, luminal B breast cancer

OPEN ACCESS

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

Reviewed by:
Ying Cai,
Albert Einstein College of Medicine, United States
Prashanth N. Suravajhala,
Amrita Vishwa Vidyapeetham University, India

***Correspondence:**
Jesús Espinal-Enríquez
jespinal@inmegen.gob.mx
Enrique Hernández-Lemus
ehernandez@inmegen.gob.mx

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

Received: *01 November 2021*
Accepted: *03 March 2022*
Published: *01 April 2022*

Citation:
Hernández-Gómez C, Hernández-Lemus E and Espinal-Enríquez J (2022) The Role of Copy Number Variants in Gene Co-Expression Patterns for Luminal B Breast Tumors. Front. Genet. 13:806607. doi: 10.3389/fgene.2022.806607

1 INTRODUCTION

It is difficult to exaggerate the negative impact that cancer has, not only as a global health burden, but also as a phenomenon with enormous societal and economic consequences. Last year, cancer was the cause of death for 10 million people, mainly in poor and developing countries. Of all tumor types, breast cancer is the malignant neoplasm with the largest incidence worldwide (Siegel et al., 2020). Many survivors of the disease (five or more years after diagnosis) still have to live with physical and psychological problems that persist over time (Stein et al., 2008) and drastically reduce their quality of life and productivity.

Breast cancer is also a highly complex and heterogeneous disease, both from the molecular and from the clinical/phenotypic standpoints. It is known that breast cancer diagnosis, response to treatment, relapse, and outcome are largely dependent on the molecular features that have been associated with the so-called breast cancer subtypes (Liu et al., 2007; Kittaneh et al., 2013; de Anda-Jáuregui et al., 2016): Luminal A, the most common, estrogen and progesterone receptor positive but, in general, epidermal growth factor 2 receptor negative; Luminal B subtype is positive for estrogen receptor and epidermal growth and negative for progesterone; HER2+, these tumors are negative for estrogen and progesterone and positive for epidermal growth factor 2; finally the Basal subtype tumors, which are mostly (around 80%) *triple negative*, i.e., negative for estrogen, progesterone, and epidermal growth factor 2. These classifications are useful to determine the origin, evolution and treatment to be followed in each case, although each patient is unique and each subtype has peculiarities.

Of the estrogen-positive subtypes, Luminal B is often the most aggressive and although luminal breast tumors are susceptible to be treated with targeted therapy (a fact that commonly, but not always, is associated with better outcomes), in some cases present the worst prognosis for patients. This is so, due to several molecular and functional features that have been related to higher proliferation rates and pharmacological resistance (Cheang et al., 2009; Tran and Bedard, 2011; Creighton, 2012; Ades et al., 2014), or even due to metabolic alterations (Serrano-Carbajal et al., 2020).

Luminal B breast cancer subtype is hormone-receptor positive (estrogen-receptor and/or progesterone-receptor positive), and either HER2 positive or HER2 negative. This subtype presents high levels of Ki-67. In general, Luminal B tumors grow slightly faster than those from Luminal A subtype. Additionally, the prognosis is commonly worse (Li et al., 2016).

Luminal B breast tumors are characterized by a lower expression of estrogen receptor, and low expression of progesterone receptor (Harbeck et al., 2013). It is also defined by aggressive clinical behavior; its prognosis is similar to that of non-luminal cancers (Tran and Bedard, 2011). Bone metastasis appears more often in Luminal B patients than in non-luminal ones. However, recurrence or metastasis in this subtype have a better prognosis after treatment than non-luminal tumors. It has also been shown that Luminal B subtype presents high metabolic deregulation (Li et al., 2016; Serrano-Carbajal et al., 2020). Luminal B subtype tumors accounted for nearly 40% of all breast cancers (Metzger-Filho et al., 2013). Therefore, understanding the molecular basis of the luminal B subtype is a matter of current concern.

One of the most actively investigated genomic regions in the manifestation of breast cancer is cytoband 8q24.3 (Wokolorczyk et al., 2008; Dorantes-Gilardi et al., 2021). This region results particularly relevant to study in the present context, since 8q24 has been repeatedly reported to harbor multiple variants associated with the incidence of breast cancer and other type of neoplasms (Tong et al., 2020). Indeed, genomic variants in 8q24 have been ascertained to be associated with risk of breast cancer on systematic reviews and meta-analyses (Wang et al., 2020). Particularly interesting is the fact that larger genomic alterations (including copy number variants, CNVs) have been linked to breast cancer onset (Jia et al., 2019), often via disruption of healthy breast cells transcriptional programmes (Ibragimova et al., 2020). Indeed, such effects have been actually linked via clinical and pathological features to breast tumors of the Luminal B subtype (or related: ER+, PR- and HER2-) giving rise to basal-like and endocrine resistant phenotypes (Liu et al., 2018).

It has been argued that some of these genomic alterations have relevant consequences for transcriptional regulation anomalies associated with cancer. Gene regulatory programmes are known to be altered, a fact that has been linked with the onset and development of tumor phenotypes (de Anda-Jáuregui et al., 2016; Hernández-Lemus et al., 2019). In this regard, our group has thoroughly described how the more relevant gene-gene co-expression interactions in several cancer types often occur between genes from the same chromosome (*cis*-), even in proximal chromosomal locations. Conversely, inter-chromosome (*trans*-) interactions (or even long distance intra-chromosome co-expression interactions) are comparatively less abundant and have lower values of diverse statistical dependency measures (Espinal-Enríquez et al., 2017; de Anda-Jáuregui et al., 2019; Dorantes-Gilardi et al., 2020; García-Cortés et al., 2020; Zamora-Fuentes et al., 2020; Andonegui-Elguera et al., 2021; Dorantes-Gilardi et al., 2021; García-Cortés et al., 2021).

Since a large number of molecular players and processes are known to be involved in (normal and) anomalous transcriptional regulatory patterns, establishing which, among the multitude of potential causes of this phenomenon in tumors are actually more relevant given the available experimental evidence becomes desirable. Particularly important has been the discussion on the effects that CNVs have on gene expression patterns in several diseases (Stranger et al., 2007; Henrichsen et al., 2009), including breast cancer (Kumaran et al., 2017; Ohshima et al., 2017; Safonov et al., 2017; Sun et al., 2018; Shao et al., 2019).

With this in mind, we have decided to investigate the effect that CNVs may have on the phenomenon of distance-associated gene co-expression in breast cancer. For the reasons already discussed we have chosen to perform a detailed study centered in the chromosome 8 and the 8q24 region in Luminal B breast tumors. In this study, we implemented a probabilistic modeling approach to evaluate how CNVs affect long and short range co-expression profiles in Luminal B breast tumors. We analyzed the co-expression patterns in chromosome 8, since it is known that

amplifications and deletions in that chromosome may influence cancer development. We found that CNV signatures in chromosome 8 of Luminal B network do not significantly alter the co-expression patterns, hence the co-expression program in this tumor phenotype is not determined by CNV structure. Additionally, we found that the 8q24.3 region is highly dense in co-expression interactions, as well as p21.3. The most connected genes in this network belong to those cytobands and have been associated with several tumor types. Interestingly, among the most connected genes, we found MAF1 and POLR3D, which may constitute an axis of regulation of gene transcription, in particular for non-coding RNA species. We believe that by advancing our knowledge of the molecular mechanisms behind gene regulation in cancer, we will be improving, not only to understand tumor biology, but also the scope of diagnostic, prognostic and therapeutic interventions to ultimately benefit oncologic patients.

This article is organized as follows: **Section 2** explains what conditional mutual information (CMI) is and why it is appropriate in the present investigation; the Kolmogorov-Smirnov method is also presented to quantify the difference between two probability distributions. In **section 3**, the method used to build the networks is explained. It is also established (via conditional mutual information distributions), that CNVs are not significantly associated with the structural features (in particular distance-dependent co-expression patterns) of the analyzed co-expression networks. The results are analyzed in **section 4**. There, the genes with the highest connectivity as well as their possible relationship in the development of cancer are identified and discussed. Finally, in **section 5** the conclusions and some consequences of this work are presented.

2 METHODS

2.1 Data Acquisition

The complete collection of The Cancer Genome Atlas (TCGA) breast RNA-Seq samples was downloaded in January 2019 from the GDC repository https://portal.gdc.cancer.gov/repository. This collection included 113 solid tissue, normal samples and 1,102 primary tumor samples. From these samples, 192 corresponded to Luminal B tumors. Data acquisition was carried out by using the TCGABiolinks R package (Colaprico et al., 2016).

2.2 Data Integration

An integrity check was carried out in raw expression files using gene annotations from BioMart. Only protein coding genes belonging to conventional chromosomes (1, 2, . . . , 22, X and Y) were kept. The CNVs of the micro-RNAs were *masked* from the 8q24.3 region and excluded from the analysis. **Supplementary Materials S1, S2** contains gene expression and genetic information of chromosome 8 genes. Pre-processing and quality control of the gene expression samples were performed as in (Espinal-Enríquez et al., 2017). In brief, we used for NOIseq R library for quality control (Tarazona et al., 2011, 2015). For batch effect removal, normalization, transcript length and GC content correction, EDASeq library was implemented (Risso et al., 2011). Finally, for multi-dimensional noise reduction we used ARSyN R library (Nueda et al., 2012).

2.3 Conditional Mutual Information Measures

We considered information theoretical measures of statistical dependency as follows: Let X, Y and Z denote discrete random variables having the following features:

1 Finite alphabets $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$, respectively
2 Joint probability mass distributions $p\,(X, Y, Z)$, and partial-joint probability mass distributions $p\,(X, Y)$, $p\,(X, Z)$, etc.,
3 Marginal probability mass distributions p (X), p (Y) and p (Z)

Let also $\hat{X}$, $\hat{Y}$ and $\hat{Z}$ denote additional discrete random variables defined on X, Y and Z respectively, the associated probability mass distributions will be $\widehat{p(X)}$, $\widehat{p(Y)}$ and $\widehat{p(Z)}$, their joint probability mass distribution $\widehat{p(X,Y,Z)}$ defined on $\mathcal{J}$, the joint probability sampling space; $\mathcal{J} = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. For particular realizations, we have $p(x) = P\,(X = x)$, $\widehat{p(y)} = P(\hat{Y} = y)$, etc., It is possible to define the Conditional Mutual Information (CMI) function $I\,(X; Y|Z)$ as follows:

$$I\,(X;Y|Z) = \sum_{z\in\mathcal{Z}} \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} p_{X,Y,Z}\,(x,y,z) \log \frac{p_Z\,(z)\,p_{X,Y,Z}\,(x,y,z)}{p_{X,Z}\,(x,z)\,p_{Y,Z}\,(y,z)} \tag{1}$$

Formally $I\,(X; Y|Z)$ is a measure representing the expected value of the mutual information of two random variables X and Y given the value of a third random value Z. Thus $I\,(X; Y|Z)$ represents the expected value (w.r.t. Z) of the Kullback-Liebler divergence from the conditional joint distribution $P\,(X, Y|Z)$ to the product of the conditional marginals $P\,(X|Z)$ and $P\,(Y|Z)$.

CMI calculations were performed with R infotheo library (Meyer and Meyer, 2009).

2.4 Assessment of the Impact of Copy Number Variant in Gene Co-Expression Programs

In order to ascertain to what extent the CNVs are able to influence the gene co-expression programs, we performed Kolmogorov-Smirnov (KS) tests to evaluate the differences among the diverse CMI distributions. The KS statistic between 2 distributions is defined as:

$$D_{n,m} = \sup_x |F_{1,n}\,(x) - F_{2,m}\,(x)| \tag{2}$$

here, $F_{1,n}\,(x)$ and $F_{2,m}\,(x)$ are the empirical distribution functions of the first and the second sets. Statistical significance of the KS tests is asymptotically given as follows.

The null hypothesis is rejected (at significance level α), whenever

$$D_{n,m} > c(\alpha)\sqrt{\frac{n+m}{n \cdot m}} \quad (3)$$

with $c(\alpha) = \sqrt{-\ln\left(\frac{\alpha}{2}\right) \cdot \frac{1}{2}}$

KS tests were performed with the ks. test library in R.

Circos plots visualizations were made using the BioCircos JavaScript library (Cui et al., 2016).

3 IMPLEMENTATION

It has been argued that mutual information (MI) is a reliable measure to establish links between genes in co-expression networks (Margolin et al., 2006a,b; Hernández-Lemus and Siqueiros-García, 2013; Lachmann et al., 2016). Given that there is generally a sufficient amount of data available to reconstruct the probability distributions associated with the expression of genes, it is possible to use it without major restrictions when measuring the statistical dependency structures between them (Hernández-Lemus and Rangel-Escareño, 2011). It is also a measure that takes into account the non-linear contributions of interdependence in the series, which makes it more appropriate in the context of the complex regulatory patterns of gene expression.

The adoption of MI-based network deconvolution methods has opened the entrance of the tenets of information theory in the analysis of biomolecular networks. However, when, as in the present work, it is necessary to evaluate the mutual information between two random variables (corresponding to each of the individual gene expression profiles in the present case), given a third, potentially influential feature, one relevant alternative is the evaluation of the Conditional Mutual Information measures. Indeed CMI has already been used in the construction of regulatory networks in a different but related context to the one presented here (Liang and Wang, 2008; Zhang et al., 2012).

An underlying problem when applying MI and CMI to establish the dependency between variables is choosing the method for reconstruction of the associated probability distributions. There are two main ways to do this, using the k-nearest neighbor non-parametric method (Kraskov et al., 2004) and through kernel density estimation (Terrell and Scott, 1992). In this work we have chosen the second method, in particular with a Gaussian kernel estimation. The implementation used is the one corresponding to the *R* package Infotheo.

The role of CNVs has been pointed out as a possible element causing the loss of co-expression with the distance that is observed in cancer and that has been reported previously (García-Cortés et al., 2021). Ideally, a non-cancerous person has two copies of each gene, however many disorders of genetic origin are caused by the deletion, repetition or insertion of DNA segments, sometimes as long as the arm of a chromosome or a complete chromosome. Given that in the past, they have been found to be responsible for genetic diseases, to think that they may also have a relevant role in the loss of co-expression is an appealing idea.

To systematically evaluate the contribution of CNVs in gene co-expression in cancer, it is convenient to use an area of the genome that is particularly active in its manifestation and to calculate the influence of CNVs both in that region and in the surrounding *areas*. As previously mentioned, 8q24.3 has been identified as particularly active and connected in co-expression networks, mainly in breast cancer tissues. Thus, we chose the Luminal B breast cancer subtype, given the poor prognosis compared with the Luminal A subtype. Furthermore, gene expression patterns have shown altered metabolic pathways even more evident than in Basal-like or HER2+ phenotypes (Serrano-Carbajal et al., 2020). We took the CNV values of 183 coding genes from the 8q24.3 region (**Supplementary Materials S3, S4**), we take them one by one as conditional and calculate the CMI of 442 genes, including not only those of this region but of the entire chromosome 8. **Supplementary Materials S3–S6** show the copy number alteration map of all samples for chromosome 8.

We repeated the above calculation for the same genes with control tissue samples assuming a number of CNVs equal to two in all samples as conditional. From the obtained values we construct co-expression networks. For each of the 183 conditional layers there are 97,461 links. The question behind this analysis is: How important is the difference obtained between the different conditional layers? For control tissue networks the question becomes trivial, since all the conditionals are essentially the same, but for those corresponding to Luminal B this difference may be relevant due to the already documented anomalous CNV structure in the region.

To quantify the differences between each of the 183 conditional layers obtained, we calculated the difference between the distributions of CMI values for all possible pairs of them (16,653 comparisons) using the Kolmogorov-Smirnov test. In essence, this test compares the cumulative functions of two distributions and it establishes as a measure, the D statistic, i.e. the maximum vertical difference between them (**Supplementary material S5**). In **Figure 1** we show graphically the workflow for this project.

4 RESULTS AND DISCUSSION

4.1 Contribution of Copy Number Variants to the Co-Expression Program is Marginal

Figure 2 shows a heatmap with the values of the D statistic of the Kolmogorov-Smirnov test for each different pair of the 183 distributions of CMI values obtained. The largest of them, which occurs between the layers corresponding to the CNVs of the COL22A1 gene (ENSG00000169436) and (ENSG0000016943), is less than 0.06. This result indicates that the contribution of 8q24.3 CNVs in the co-expression networks of genes on chromosome 8 in the luminal B subtype of cancer is marginal.

This result acquires relevance since copy number alterations are a mesoscopic phenomena, which can affect large parts of the genome. On the other hand, changes in the co-expression landscape are considered a microscopic event, since those changes may affect specific genes and their regulatory relationships. Hence, with this result we provide evidence that

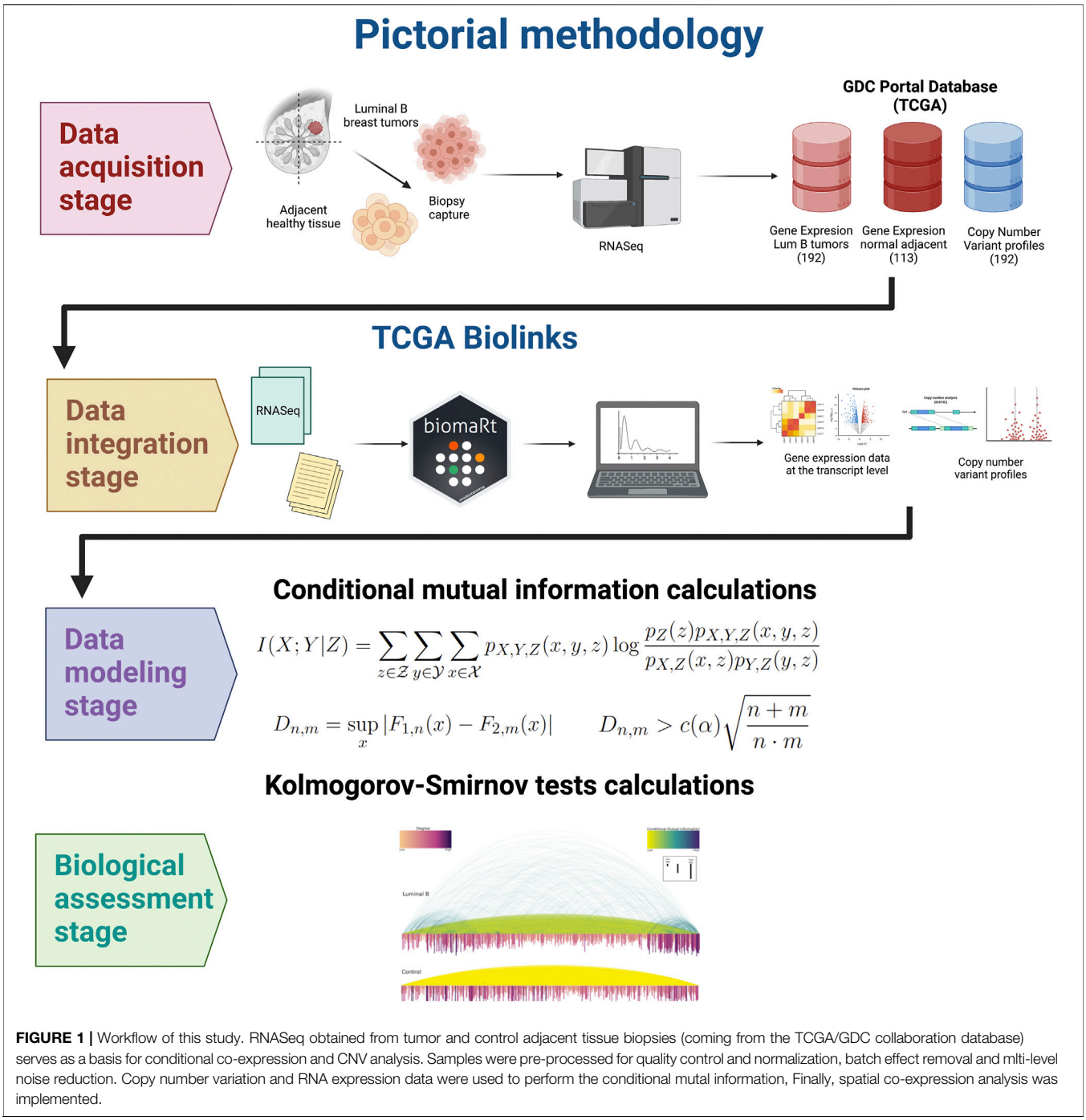


FIGURE 1 | Workflow of this study. RNASeq obtained from tumor and control adjacent tissue biopsies (coming from the TCGA/GDC collaboration database) serves as a basis for conditional co-expression and CNV analysis. Samples were pre-processed for quality control and normalization, batch effect removal and mlti-level noise reduction. Copy number variation and RNA expression data were used to perform the conditional mutal information, Finally, spatial co-expression analysis was implemented.

a large-scale event such as amplifications/deletions of large portions of chromosome 8 do not significantly alter the gene co-expression program in cancer.

4.2 Conditional Mutual Information is Higher in Luminal B Than in Controls

Since all the layers corresponding to Luminal B samples have essentially the same distribution of CMI, we take the first, whose conditionals are the CNVs of the COL22A1 gene, as representative of the entire family of networks and proceed to analyze it. **Figure 3** shows the difference between CMI from the Luminal B subtype network and the one from the control network. As it can be clearly observed, the control CMI values are lower than the cancer counterpart.

In (Espinal-Enríquez et al., 2017), we observed that mutual information values from control gene co-expression networks are higher than in the cancer-derived network. However, the calculation was taken over all the gene-gene co-expression interactions, i.e. intra and inter-chromosomal gene

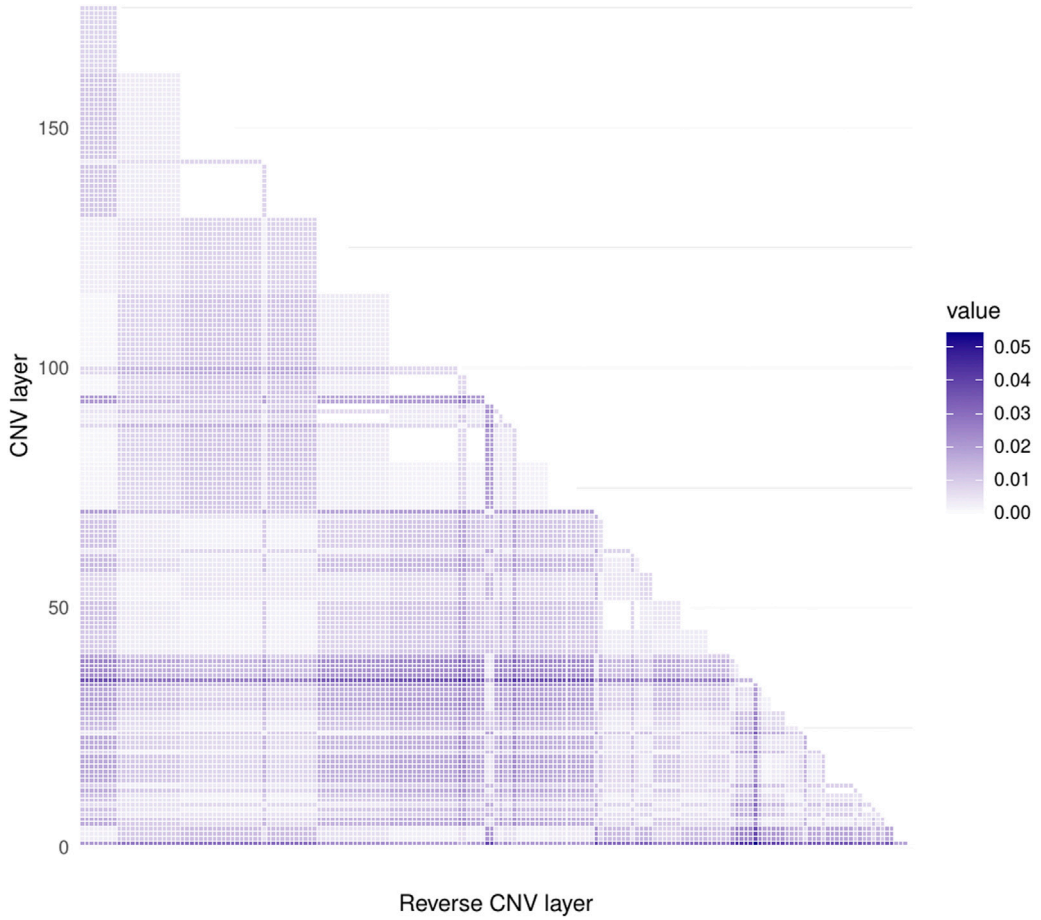


FIGURE 2 | Heatmap corresponding to the values of the D statistic for the Kolmogorov-Smirnov test. Both axes represent the same CNV layers. As it can be observe from the color code at the right part of the figure (very low values of D), none of the CNV layers present a larger KS statistics, which reflects that copy number alterations do not significantly change gene co-expression values.

correlations. In this case, the CMI is taken between chromosome 8 genes only, i.e., just intra-chromosome interactions.

Taking into account that the loss of long-range co-expression phenomenon has been previously reported in networks form different cancer types (Zamora-Fuentes et al., 2020; Andonegui-Elguera et al., 2021), and also for breast carcinoma (Espinal-Enríquez et al., 2017; Alcalá-Corona et al., 2018; de Anda-Jáuregui et al., 2019; Dorantes-Gilardi et al., 2020), and in particular for breast cancer subtype networks (de Anda-Jáuregui et al., 2019; García-Cortés et al., 2020; Dorantes-Gilardi et al., 2021), the finding of a higher average co-expression in intra-chromosome 8 of Luminal B breast cancer compared with the control one, reinforces the fact that this phenomenon is a common trait in several types of cancer.

4.3 Cancer Network *Grows* in Small Local Regions, but not in Control

Figure 4 corresponds to Luminal B subtype interactions. Top-left to bottom-right shows the Top-500, 1,000, 2,000, and 3,000 interactions in chromosome 8. As it can be appreciated, interactions appear local in the first place, markedly in cytobands q24.3 (yellow arcs) and p11.23. Subsequently, inter-cytoband and inter-arm interactions occur, but in a small fraction (top-right). For the top-2,000 edges, it is clear that several interactions appear from q24.3. Finally, for the top-3,000 edges, 8q24.3 region is strongly connected with the rest of chromosome 8, and more inter-arm and inter-cytoband edges appear.

On the other hand, the **Figure 5** shows the existence of a non-localized connectivity for the control network; since for the first 500 highest connections, interactions occur between genes from different cytobands, or even from different arms, much more often than in the cancer phenotype. The same effect is shown in the top-1,000 and 2,000 interaction circos plots. It is worth noticing the small number of intra-cytoband interactions, even in the top-3,000 edges (830 out of 3,000, dark blue edges).

We believe that this pattern where the highest correlations appear between physically close genes, is in agreement with other phenomena observed in breast and other cancers in which the loss of long-distance co-expression is evident, such as kidney (Zamora-Fuentes et al., 2020) or lung (Andonegui-Elguera et al., 2021).

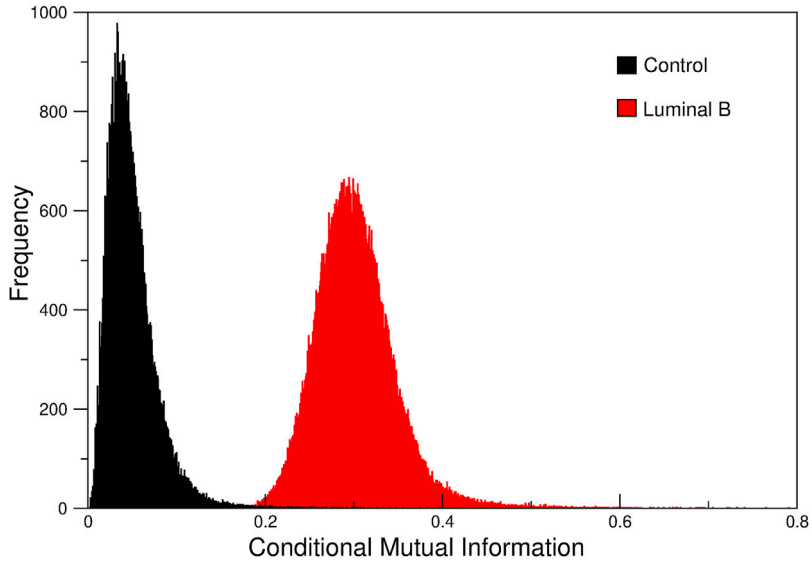


FIGURE 3 | Conditional Mutual Information for normal adjacent tissue (black) and Luminal B tumors (red). We can notice that Luminal B tumors present distinctively higher values of CMI than Controls.

4.4 Extreme Regions Exhibit a High Connectivity Pattern in Luminal B Network

We decided to analyze the topology of the resulting networks, but we only kept the highest interactions for this purpose. We set 0.35 as a threshold for CMI in the Luminal B network. This threshold resulted in a network with 11,449 edges and 420 genes. For comparison, we conserved the same number of edges in the control network. Taking into account the highest CMI values in the Luminal B phenotype, a very intriguing phenomenon appears: the first (8p11.21–23) and last (8q24.3) codifying regions of the chromosome 8 exhibit strong connections. Additionally, this strength decrease towards the centromere. Conversely, in the control network, the strength of interactions does not depend on the location of genes. This remarkable difference between the breast cancer and the non-tumor adjacent phenotype can be appreciated in **Figure 6**

Regarding the Luminal B chromosome 8 network, the genes of the extreme regions also constitute the most connected nodes of the network. The case of cytoband q24.3 is the most emblematic one. Cytoband q24.3 is the one with more intra-cytoband edges (1,640 out of 11,449). It is also the cytoband with more genes in the network (79 out of 420). Regarding the inter-cytoband edges, the large majority of interactions from any cytoband also correspond to those from q24.3. That is the reason for which the average degree of region q24.3 is the highest of the network (52). The non-extreme region p11.23 is an exception that has strong interactions.

It is important to mention that in region p11.23, there are some crucial genes in terms of cell maintenance, as well as for cancer development. For instance, four genes, namely BAG4, LSM1, ASH2L and BRF2 serve as housekeeping genes. On the other hand, FGFR, a keystone gene in cancer development, has been observed to be amplified in luminal B breast tumors (Erber et al., 2020; Voutsadakis, 2020; Amina et al., 2021).

4.5 Highly Connected Genes and Their Possible Relationship With Cancer

In **Table 1** we show the most connected genes in the Luminal B network. As it can be observed, most of the genes belong to region q24.3. However, three genes, KCTD9, POL3RD, and ATP6V1B2 belong to the region p21.3. The gene with the highest Betweenness Centrality (BC) is KCTD9. Below we will provide a brief summary of what it is known of those genes in Luminal B breast cancer or other carcinomas.

ZNF7 (Zinc Finger Protein 7) has been indicated as a biomarker of survival in glioblastoma Esteve-Codina et al. (2021) and Burkitt's lymphoma (Gallego and Lazo, 1994). ZNF7 is the most connected gene in the Luminal B breast cancer co-expression network, and it is slighlty overexpressed (**Table 1**).

Many genes on chromosome 8 have been associated with mental illnesses, mainly schizophrenia, and their mutations are presumed to be involved in the development of our mental abilities. This is the case of the coding genes for the KCTD (Potassium Channel Tetramerization Domain) proteins. However, it has recently been indicated (Angrisani et al., 2021) that the 25 members of this family are potentially involved in a second fundamental activity: 13 of them have an exclusive pro-tumor function, 5 an exclusive anti-tumor function, 5 a pro/anti-tumor role, and 2 with a function not yet determined. KCTD9 has a pro-tumor function not yet reported in the literature, but inferred through databases. However, in this case, KCTD9 is underexpressed (−1.16) which may implicate a dual role in the phenotype.

MAF1 gene is known for its regulatory effect on the polymerase III, although it has also been associated with cancer, given that it activates the expression of the PTEN protein, which is an important tumor suppressor (Zhang et al.,

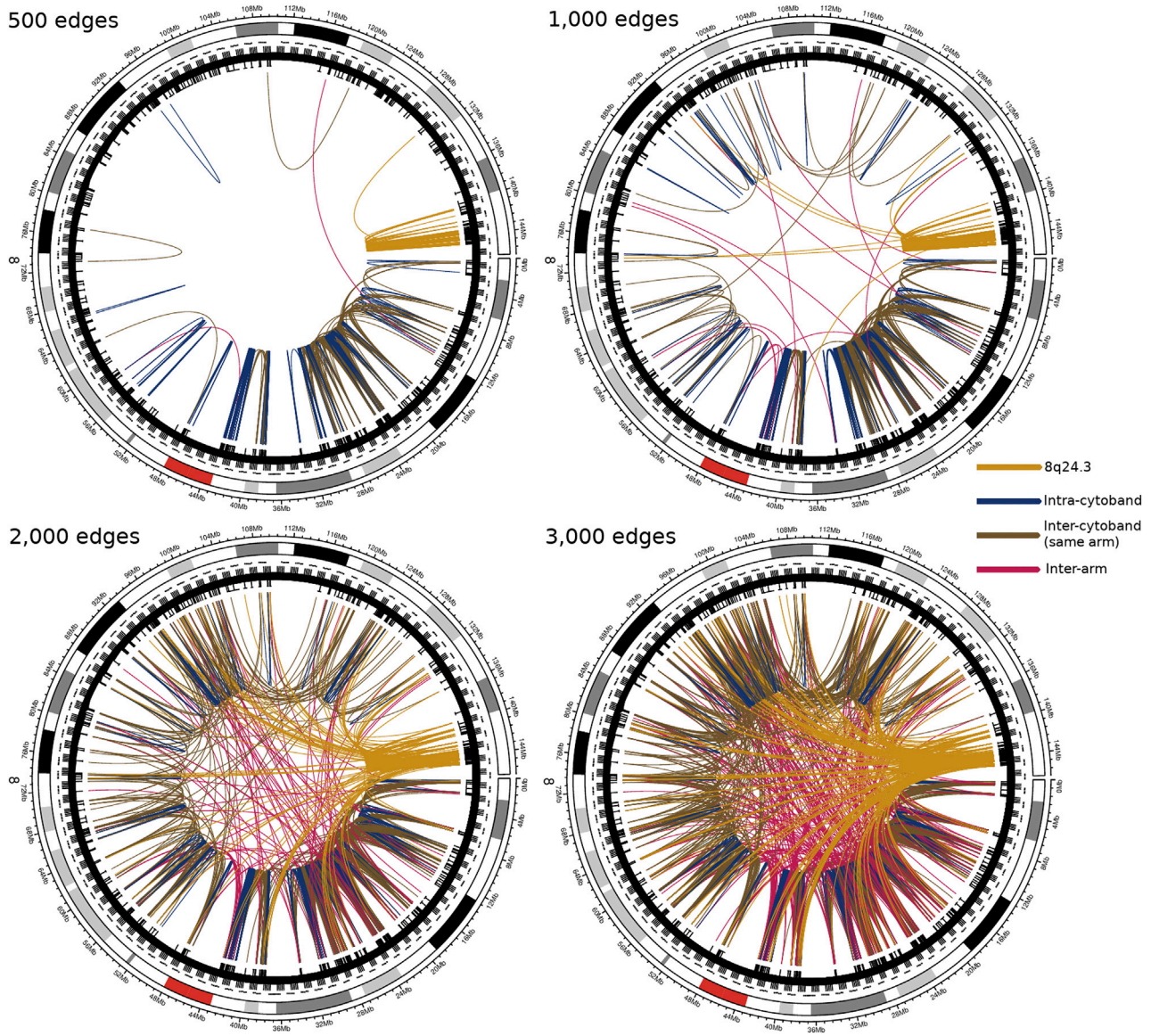


FIGURE 4 | Top interactions of Luminal B CMI network at different cut-offs (500, 1,000, 2,000, and 3,000 edges). At first, the intra-cytoband interactions dominate, mainly in q24.3, p21.3, p11.21 and p11.23; secondly, the inter-cytoband interactions (particularly in p-arm), and finally, the inter-arm edges. Red arc at the external circle represents the centromere of Chr8. Color code of the co-expression interactions are also described. The take-home message is that for the top interactions, the location of participating genes is very close between them, in particular, at cytoband q24.3.

2018). Interestingly POLR3D is also one of the most connected genes in the Luminal B breast cancer network. POLR3D gene encodes an RNA polymerase 3 subunit D, and synthesizes small RNAs, such as 5S rRNA and tRNAs, which, when inhibited by TRIPLIDE (TPL) influence the control of colorectal cancer (Liang et al., 2019). POLR3D is also inhibited by miR-320 (Ramassone et al., 2018).

The fact that MAF1 and POLR3D genes were highly connected could be an indicator of alterations in the transcriptional regulation program. Since MAF1 regulates polymerase III, and POLR3D encodes an RNA polymerase III subunit, the emergence of an axis of transcriptional regulation of non-coding RNA species results appealing. The latter may imply that in the Luminal B subtype, transcriptional regulation mediated by non-coding RNAs could affect the gene regulatory program. The latter coincides with the fact that POLR3D abnormal activity is characteristic of cancer cells (White, 2004).

ADCK5 has been indicated as an intermediary in the growth and metastasis of lung cancer. This gene promotes invasion and migration of lung cancer cells through the ADCK5-SOX9-PTTG1 pathway Qiu et al. (2020).

Regarding PTK2, this gene in association with KCNMA1 gene has been reported as a tumor suppressor in gastric cancer Ma

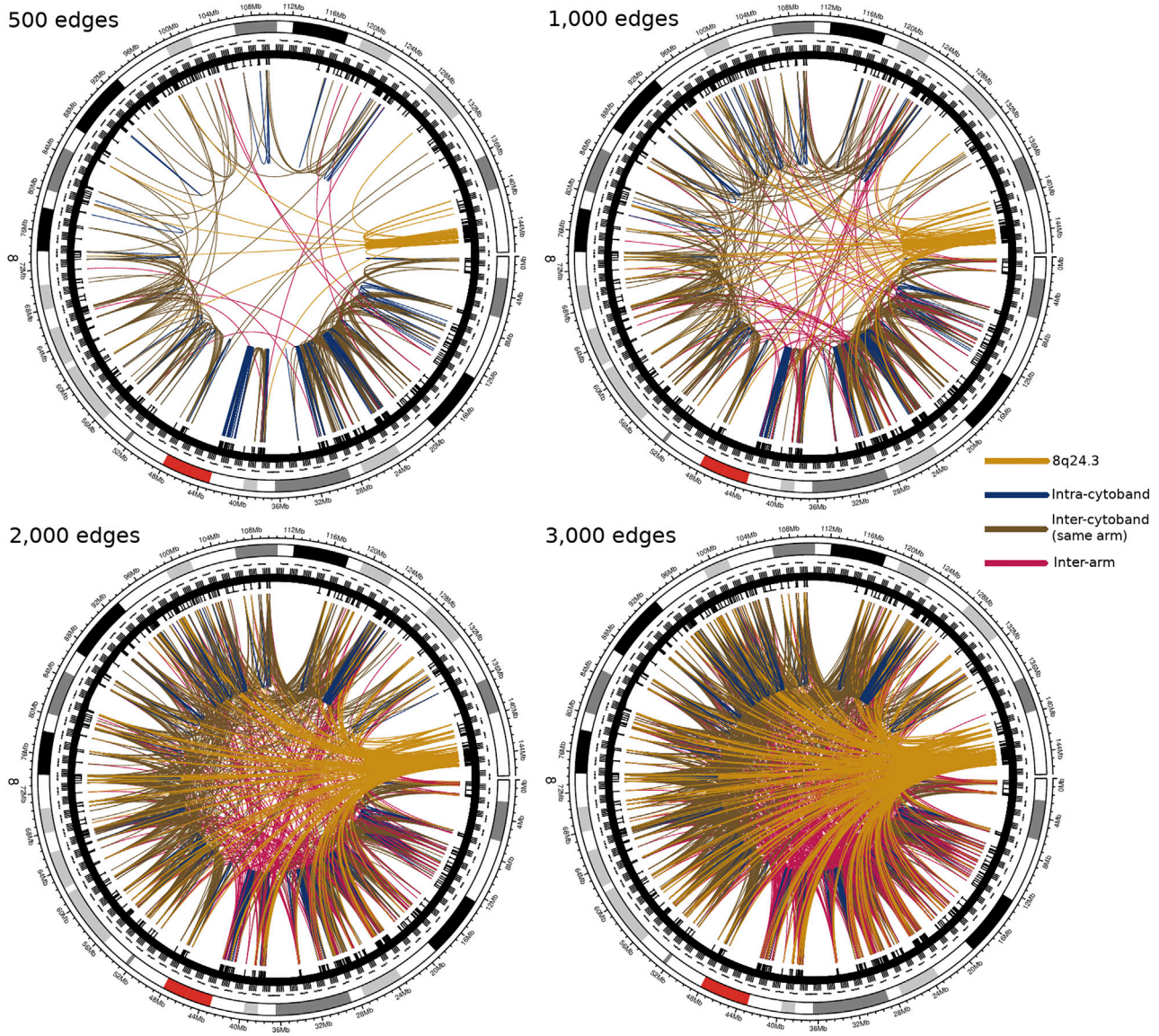


FIGURE 5 | Top interactions of control tissue CMI network at different cut-offs. In this figure, analogue to **Figure 4**, although both, expression and co-expression are not uniform, many of the strongest links correspond to inter-arm interactions. The color code is the same than in **Figure 4**. It is clear the difference between the area covered by top-3,000 interactions in control circos-plot and the one observed in the Luminal B case. The density of links in the control network is larger than in **Figure 4**.

et al. (2017). Additionally, the SMARCE1 gene regulates metastasis in breast cancer through its interaction with HIF1A and PTK2 Sethuraman et al. (2016). It has also been associated with hepatocarcinoma (Okamoto et al., 2003). Additionally, PTK2/FAK is considered a driver of radio-resistance in HPV-negative head and neck cancer (Skinner et al., 2016).

Recently, the FAM83H gene (also located on chromosome 8, specifically in the 8q24.3 region) has been related to Zinc Finger Proteins (ZNFs) genes, specifically ZNF16. Gallbladder cancer is highly associated with the expression of these two genes Ahn et al. (2020).

ATP6V1B2 gene has been identified as a possible biomarker (from controlled-to-aggressive growth with invasion of muscle tissue) for bladder cancer. Specifically, it is underexpressed in the early stages of the disease and over-expressed in the advanced stages Fang et al. (2013). Additionally, it has been associated with follicular lymphoma, activating autophagic flux and mTOR pathway (Wang et al., 2019).

ZNF707 has recently been Kim et al. (2020) identified as highly overexpressed in the Japanese population. High expression has also been shown in kidney and luminal B cancer patients Machnik et al. (2019) which suggests that ZNF707 could be involved in the development of cancer in general, regardless of the tissue.

In brief, we can establish that the most connected genes in the chromosome 8 network for Luminal B breast cancer are

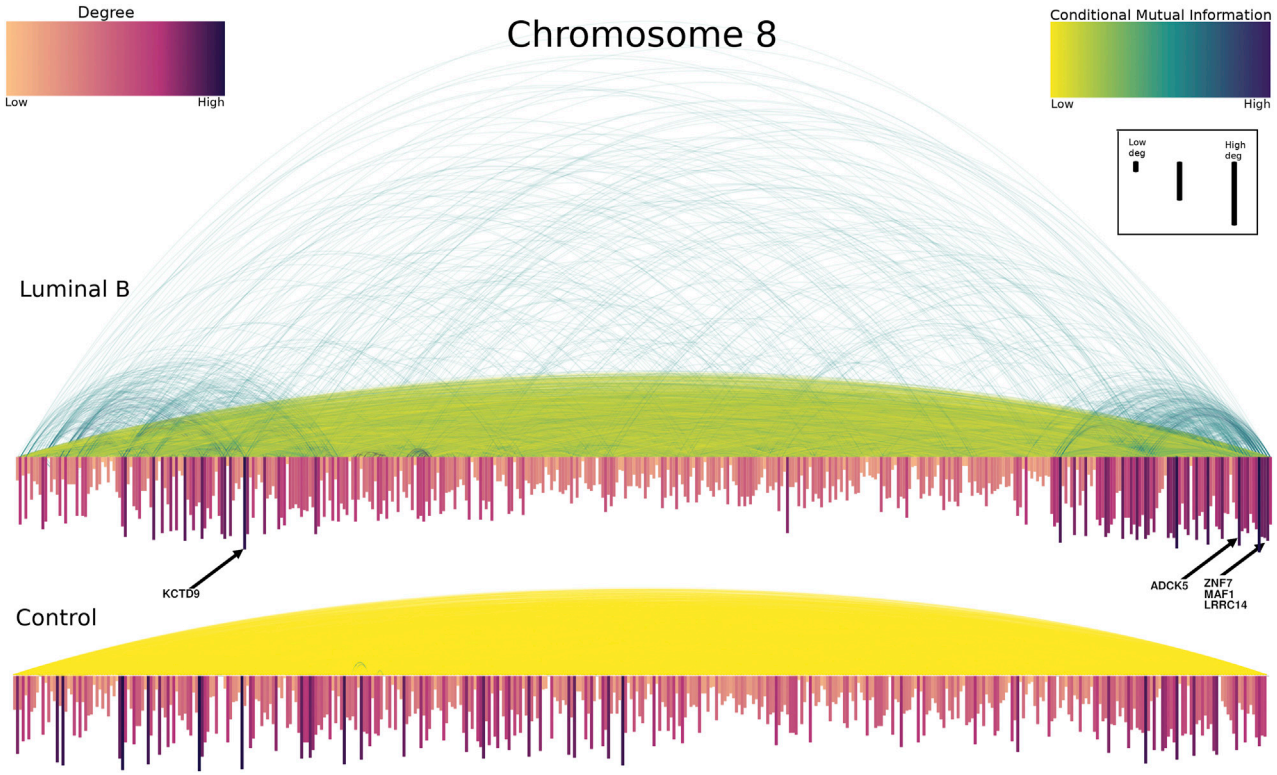


FIGURE 6 | Top CMI values between genes of Chromosome 8 for Luminal B (upper) and control (lower) networks. In this figure, the genes are placed according to its start position. Yellow-to-blue arcs represent co-expression interactions between the connecting genes. The size and color of the genes are proportional to the degree. As it can be observed, for the case of Luminal B, the extreme left and right sides contain the majority of highest CMI interactions (dark arcs), in particular, cytobands q24.3 and p11.21–23. Conversely, in the control network, interactions are not particularly biased to a specific region. Top-degree genes are indicated with black arrows. ZNF7, MAF1 and LRRC14 genes are overlapped in the figure.

TABLE 1 | The most connected genes belong to q24.3 and p21.2.

Gene	Band	Degree	BC	Log_2FC
ZNF7	q24.3	109	0.00544	0.47
KCTD9	p21.2	106	0.00656	−1.16
MAF1	q24.3	105	0.00562	0.07
LRRC14	q24.3	102	0.00512	0.34
ADCK5	q24.3	101	0.00523	0.93
PTK2	q24.3	99	0.00516	−0.03
POLR3D	p21.3	97	0.00508	−0.36
ZNF16	q24.3	97	0.00518	0.38
ATP6V1B2	p21.3	96	0.00579	−0.15
ZNF707	q24.3	96	0.00410	0.99

significantly related to the oncogenic phenotype. Particularly interesting is the case of MAF1, which in turn regulates POL3RD expression, suggesting an axis of non-coding RNA regulation.

As it can be also observed in **Table 1**, the differential expression in the large majority of hub genes is not significant. However, the co-expression patterns of these genes are importantly different than in the control case. We have observed previously, in clear cell renal carcinoma progression, that the gene expression patterns do not change significantly over progression stages, however, the co-expression networks are clearly different between those stages (Zamora-Fuentes et al., 2020). Therefore, slightly different gene expression patterns may implicate a dramatical alteration in the co-expression landscape. The differential expression of all Chr8 genes is provided in **Supplementary Material S6**.

5 CONCLUSION

The understanding of the intricate relationship between copy number alterations, which can be seen as a mesoscopic dimension, with the regulation of the gene co-expression program, which can be understood as a microscopic phenomenon, is a highly promissory pathway for intense research in the near future. In this work, we have shown, for the case of Chromosome 8 in Luminal B breast cancer, that the Copy Number Alteration scenario does not influence, in a relevant manner, the Conditional Mutual Information program. That is independent on which region is analyzed. As a summary of findings, we can establish the following:

- CNVs do not influence Conditional Mutual Information (as observed in **Figure 2**).
- Co-expression program in the chromosome 8 for Luminal B breast cancer shows localized hotspots regions in certain cytobands.
- The large majority of Chromosome 8 gene co-expression interactions shows low CMI values, meanwhile the extreme parts of the chromosome show higher values.
- Cytoband q24.3 has the highest values of MI, is the most dense in terms of interactions, and its genes have the highest degree.
- In the control phenotype there is a homogeneous CMI distribution regarding the location of genes in the top interactions, contrary to the case of Luminal B network.
- Taking into account a *growing* of the networks from highest to lowest CMI values, in the case of Luminal B network, the top CMI values appear between intra-cytoband genes; after that, between inter-cytoband and same-arm genes; finally between inter-arm genes. In the case of the control network, there is no clear localization pattern.
- Genes such as ZNF7, KCTD9, MAF1, or POLR3D have the highest degree centralities. Those genes have been reported to have influence in cancer.
- MAF1 and POLR3D could form an axis of non-coding RNA regulation, which can be a possible complex for future research.

Further steps towards a whole understanding of how copy number alterations may affect the co-expression program in breast cancer must include the analysis of the conditional mutual information of all chromosomes in Luminal B breast cancer. A similar study in the other molecular subtypes is also needed. The influence of the progression stage must be also taken into account. Finally, this analysis over other cancer tissues will provide a solid and robust landscape of the role of copy number alterations in the rise and development of cancer.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

CH-G organized data, developed code, performed calculations, analyzed data, drafted the manuscript; EH-L co-devised the project, designed the methodological approach, analyzed data, discussed results, co-supervised the project. JE-E co-devised and coordinated the project, contributed to the methodological strategy, analyzed data and integrated the biological results, discussed results, co-supervised the project; All authors read and approved the final manuscript.

FUNDING

This work was supported by the Consejo Nacional de Ciencia y Tecnología (Estancia Posdoctoral de Incidencia, CHG), and the National Institute of Genomic Medicine, México. Additional support has been granted by the Laboratorio Nacional de Ciencias de la Complejidad, from the Universidad Nacional Autónoma de México. EHL is recipient of the 2016 Marcos Moshinsky Fellowship in the Physical Sciences.

ACKNOWLEDGMENTS

Authors want to thank Diana García-Cortés for her support in the figure design.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.806607/full#supplementary-material

Supplementary Material S1 | Gene expression data. This table includes the normalized RNA-seq expression of all Luminal B samples.

Supplementary Material S2 | Gene information. Name, ENSEMBL-id, GC content, biotype, gene description, cytoband location, base pair and base end of each gene of chromosome 8 is provided.

Supplementary Material S3 | CNV data. This table contains the copy number alteration information for all Luminal B breast cancer samples.

Supplementary Material S4 | Gene information data from copy number regions.

Supplementary Material S5–S8 | Chromosome 8 copy number alteration map of 140 Luminal B samples. For these figures we used the IGV platform (https://software.broadinstitute.org/software/igv/).

Supplementary Material S9 | D-statistic heatmap and p-values. In the first figure, the layers with the largest discrepancies, all less than 0.05, appear in red. The corresponding p-values appear in the second heatmap; these should be read in conjunction with the D-values. Most of these numbers are low and appear in shades of white, the blue dots correspond to *p*-values equal to one, indicating that the data are the same in both distributions. Like adjacency matrices, these matrices are symmetric.

Supplementary Material S10 | Differential gene expression of Chr8. Log2 fold change and *p*-values are provided.

REFERENCES

Ades, F., Zardavas, D., Bozovic-Spasojevic, I., Pugliano, L., Fumagalli, D., De Azambuja, E., et al. (2014). Luminal B Breast Cancer: Molecular Characterization, Clinical Management, and Future Perspectives. *J. Clin. Oncol.* 32, 2794–2803. doi:10.1200/jco.2013.54.1870

Ahn, S. W., Ahn, A.-R., Ha, S. H., Hussein, U. K., Do Yang, J., Kim, K. M., et al. (2020). Expression of Fam83h and Znf16 Are Associated with Shorter Survival of Patients with Gallbladder Carcinoma. *Diagn. Pathol.* 15, 1–14. doi:10.1186/s13000-020-00985-1

Alcalá-Corona, S. A., de Anda-Jáuregui, G., Espinal-Enriquez, J., Tovar, H., and Hernández-Lemus, E. (2018). "Network Modularity and Hierarchical Structure in Breast Cancer Molecular Subtypes," in International Conference on Complex Systems, Cambridge, MA (Springer), 352–358. doi:10.1007/978-3-319-96661-8_36

Amina, B., Lynda, A. K., Sonia, S., Adel, B., Jelloul, B. H., Miloud, M., et al. (2021). Fibroblast Growth Factor Receptor 1 Protein (Fgfr1) as Potential Prognostic and Predictive Marker in Patients with Luminal B Breast Cancers Overexpressing Human Epidermal Receptor 2 Protein (Her2). *Indian J. Pathol. Microbiol.* 64, 254. doi:10.4103/IJPM.IJPM_87_20

Andonegui-Elguera, S. D., Zamora-Fuentes, J. M., Espinal-Enríquez, J., and Hernández-Lemus, E. (2021). Loss of Long Distance Co-expression in Lung Cancer. *Front. Genet.* 12, 625741. doi:10.3389/fgene.2021.625741

Angrisani, A., Di Fiore, A., De Smaele, E., and Moretti, M. (2021). The Emerging Role of the Kctd Proteins in Cancer. *Cell Commun. Signal.* 19, 1–17. doi:10.1186/s12964-021-00737-8

Cheang, M. C., Chia, S. K., Voduc, D., Gao, D., Leung, S., Snider, J., et al. (2009). Ki67 index, Her2 Status, and Prognosis of Patients with Luminal B Breast Cancer. *J. Natl. Cancer Inst.* 101, 736–750. doi:10.1093/jnci/djp082

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). Tcgabiolinks: an R/bioconductor Package for Integrative Analysis of Tcga Data. *Nucleic Acids Res.* 44, e71. doi:10.1093/nar/gkv1507

Creighton, C. J. (2012). The Molecular Profile of Luminal B Breast Cancer. *Biol. Targets Ther.* 6, 289. doi:10.2147/btt.s29923

Cui, Y., Chen, X., Luo, H., Fan, Z., Luo, J., He, S., et al. (2016). Biocircos. Js: an Interactive Circos Javascript Library for Biological Data Visualization on Web Applications. *Bioinformatics* 32, 1740–1742. doi:10.1093/bioinformatics/btw041

de Anda-Jáuregui, G., Velázquez-Caldelas, T. E., Espinal-Enríquez, J., and Hernández-Lemus, E. (2016). Transcriptional Network Architecture of Breast Cancer Molecular Subtypes. *Front. Physiol.* 7, 568. doi:10.3389/fphys.2016.00568

de Anda-Jáuregui, G., Espinal-Enriquez, J., and Hernández-Lemus, E. (2019). Spatial Organization of the Gene Regulatory Program: an Information Theoretical Approach to Breast Cancer Transcriptomics. *Entropy* 21, 195. doi:10.3390/e21020195

Dorantes-Gilardi, R., García-Cortés, D., Hernández-Lemus, E., and Espinal-Enríquez, J. (2020). Multilayer Approach Reveals Organizational Principles Disrupted in Breast Cancer Co-expression Networks. *Appl. Netw. Sci.* 5, 1–23. doi:10.1007/s41109-020-00291-1

Dorantes-Gilardi, R., García-Cortés, D., Hernández-Lemus, E., and Espinal-Enríquez, J. (2021). K-Core Genes Underpin Structural Features of Breast Cancer. *Sci. Rep.* 11, 1–17. doi:10.1038/s41598-021-95313-y

Erber, R., Rübner, M., Davenport, S., Hauke, S., Beckmann, M. W., Hartmann, A., et al. (2020). Impact of Fibroblast Growth Factor Receptor 1 (Fgfr1) Amplification on the Prognosis of Breast Cancer Patients. *Breast Cancer Res. Treat.* 184, 311–324. doi:10.1007/s10549-020-05865-2

Espinal-Enríquez, J., Fresno, C., Anda-Jáuregui, G., and Hernández-Lemus, E. (2017). Rna-seq Based Genome-wide Analysis Reveals Loss of Inter-chromosomal Regulation in Breast Cancer. *Sci. Rep.* 7, 1–19. doi:10.1038/s41598-017-01314-1

Esteve-Codina, A., Alameda, F., Carrato, C., Pineda, E., Arpí, O., Martinez-García, M., et al. (2021). Rna Sequencing and Immunohistochemistry Reveal Zfn7 as a Stronger Marker of Survival Than Molecular Subtypes in G-Cimp–Negative Glioblastoma. *Clin. Cancer Res.* 27, 645–655. doi:10.1158/1078-0432.ccr-20-2141

Fang, Z., Zang, W., Chen, R., Ye, B., Wang, X., Yi, S., et al. (2013). Gene Expression Profile and Enrichment Pathways in Different Stages of Bladder Cancer. *Genet. Mol. Res.* 12, 1479–1489. doi:10.4238/2013.may.6.1

Gallego, M. I., and Lazo, P. A. (1994). The Human Zinc-finger Protein-7 Gene Is Located 90 Kb 3 Prime of Myc and Is Not Expressed in Burkitt Lymphoma Cell Lines. *Int. J. Cancer* 58, 855–859. doi:10.1002/ijc.2910580618

García-Cortés, D., de Anda-Jáuregui, G., Fresno, C., Hernández-Lemus, E., and Espinal-Enríquez, J. (2020). Gene Co-expression Is Distance-dependent in Breast Cancer. *Front. Oncol.* 10, 1232. doi:10.3389/fonc.2020.01232

García-Cortés, D., Hernández-Lemus, E., and Espinal-Enríquez, J. (2021). Luminal a Breast Cancer Co-expression Network: Structural and Functional Alterations. *Front. Genet.* 12, 514. doi:10.3389/fgene.2021.629475

Harbeck, N., Thomssen, C., and Gnant, M. (2013). St. Gallen 2013: Brief Preliminary Summary of the Consensus Discussion. *Breast care* 8, 102–109. doi:10.1159/000351193

Henrichsen, C. N., Chaignat, E., and Reymond, A. (2009). Copy Number Variants, Diseases and Gene Expression. *Hum. Mol. Genet.* 18, R1–R8. doi:10.1093/hmg/ddp011

Hernández-Lemus, E., and Rangel-Escareño, C. (2011). The Role of Information Theory in Gene Regulatory Network Inference. *Inf. Theor. New Res.*, 109–144.

Hernández-Lemus, E., and Siqueiros-García, J. M. (2013). Information Theoretical Methods for Complex Network Structure Reconstruction. *Complex Adaptive Syst. Model.* 1, 1–22. doi:10.1186/2194-3206-1-8

Hernández-Lemus, E., Reyes-Gopar, H., Espinal-Enríquez, J., and Ochoa, S. (2019). The many Faces of Gene Regulation in Cancer: A Computational Oncogenomics Outlook. *Genes* 10, 865. doi:10.3390/genes10110865

Ibragimova, M., Tsyganov, M., and Litviakov, N. (2020). Transcriptome of Breast Tumor with Different Amplification Statuses of Long Arm of Chromosome 8. *Med. Genet.* 19, 31–32. doi:10.20538/1682-0363-2020-3-22-28

Jia, H., Li, S., Han, Y., Liu, F., Li, W., and Fu, L. (2019). Research Advances of Chromosomal 8q24 Aberrance in Breast Cancer. *Chin. J. Clin. Oncol.* 46, 150–153. doi:10.3969/j.issn.1000-8179.2019.03.157

Kim, J., Chung, J.-Y., Hwang, J. R., Lee, Y.-Y., Kim, T.-J., Lee, J.-W., et al. (2020). Identification of Candidate Genes Associated with Susceptibility to Ovarian clear Cell Adenocarcinoma Using Cis-Eqtl Analysis. *J. Clin. Med.* 9, 1137. doi:10.3390/jcm9041137

Kittaneh, M., Montero, A. J., and Glück, S. (2013). Molecular Profiling for Breast Cancer: a Comprehensive Review. *Biomarkers Cancer* 5, BIC.S9455. doi:10.4137/bic.s9455

Kraskov, A., Stögbauer, H., and Grassberger, P. (2004). Estimating Mutual Information. *Phys. Rev. E* 69, 066138. doi:10.1103/PhysRevE.69.066138

Kumaran, M., Cass, C. E., Graham, K., Mackey, J. R., Hubaux, R., Lam, W., et al. (2017). Germline Copy Number Variations Are Associated with Breast Cancer Risk and Prognosis. *Sci. Rep.* 7, 1–15. doi:10.1038/s41598-017-14799-7

Lachmann, A., Giorgi, F. M., Lopez, G., and Califano, A. (2016). Aracne-ap: Gene Network Reverse Engineering through Adaptive Partitioning Inference of Mutual Information. *Bioinformatics* 32, 2233–2235. doi:10.1093/bioinformatics/btw216

Li, Z.-h., Hu, P.-h., Tu, J.-h., and Yu, N.-s. (2016). Luminal B Breast Cancer: Patterns of Recurrence and Clinical Outcome. *Oncotarget* 7, 65024. doi:10.18632/oncotarget.11344

Liang, K.-C., and Wang, X. (2008). Gene Regulatory Network Reconstruction Using Conditional Mutual Information. *EURASIP J. Bioinformatics Syst. Biol.* 2008, 1–14. doi:10.1155/2008/253894

Liang, X., Xie, R., Su, J., Ye, B., Wei, S., Liang, Z., et al. (2019). Inhibition of Rna Polymerase Iii Transcription by Triptolide Attenuates Colorectal Tumorigenesis. *J. Exp. Clin. Cancer Res.* 38, 1–13. doi:10.1186/s13046-019-1232-x

Liu, R., Wang, X., Chen, G. Y., Dalerba, P., Gurney, A., Hoey, T., et al. (2007). The Prognostic Role of a Gene Signature from Tumorigenic Breast-Cancer Cells. *New Engl. J. Med.* 356, 217–226. doi:10.1056/nejmoa063994

Liu, X., Ma, D., Jiang, Y., Liu, Y., Yu, K., and Shao, Z. (2018). Abstract P2-06-03: Clinicopathologic Characteristics and Genomic Essence Analyses Revealed Estrogen Receptor Positive, Progesterone Receptor Negative and Human Epidermal Growth Factor Receptor 2 Negative Breast Cancer to Be More Basal-like and Endocrine Resistant. *Cancer Res.* 78, 6711–6910. doi:10.1158/1538-7445.sabcs17-p2-06-03

Ma, G., Liu, H., Hua, Q., Wang, M., Du, M., Lin, Y., et al. (2017). Kcnma1 Cooperating with Ptk2 Is a Novel Tumor Suppressor in Gastric Cancer and Is Associated with Disease Outcome. *Mol. Cancer* 16, 1–10. doi:10.1186/s12943-017-0613-z

Machnik, M., Cylwa, R., Kiełczewski, K., Biecek, P., Liloglou, T., Mackiewicz, A., et al. (2019). The Expression Signature of Cancer-Associated Krab-Znf Factors Identified in Tcga Pan-Cancer Transcriptomic Data. *Mol. Oncol.* 13, 701–724. doi:10.1002/1878-0261.12407

Margolin, A. A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla Favera, R., et al. (2006a). Aracne: an Algorithm for the Reconstruction of Gene Regulatory Networks in a Mammalian Cellular Context. *BMC Bioinformatics* 7, 1–15. Springer. doi:10.1186/1471-2105-7-S1-S7

Margolin, A. A., Wang, K., Lim, W. K., Kustagi, M., Nemenman, I., and Califano, A. (2006b). Reverse Engineering Cellular Networks. *Nat. Protoc.* 1, 662–671. doi:10.1038/nprot.2006.106

Metzger-Filho, O., Sun, Z., Viale, G., Price, K. N., Crivellari, D., Snyder, R. D., et al. (2013). Patterns of Recurrence and Outcome According to Breast Cancer Subtypes in Lymph Node–Negative Disease: Results from International Breast Cancer Study Group Trials Viii and Ix. *J. Clin. Oncol.* 31, 3083. doi:10.1200/jco.2012.46.1574

Meyer, P. E., and Meyer, M. P. E. (2009). *Package 'infotheo'*. Princeton, NJ, USA: Citeseer. R Package Version.

Nueda, M. J., Ferrer, A., and Conesa, A. (2012). Arsyn: a Method for the Identification and Removal of Systematic Noise in Multifactorial Time Course Microarray Experiments. *Biostatistics* 13, 553–566. doi:10.1093/biostatistics/kxr042

Ohshima, K., Hatakeyama, K., Nagashima, T., Watanabe, Y., Kanto, K., Doi, Y., et al. (2017). Integrated Analysis of Gene Expression and Copy Number Identified Potential Cancer Driver Genes with Amplification-dependent Overexpression in 1,454 Solid Tumors. *Sci. Rep.* 7, 1–13. doi:10.1038/s41598-017-00219-3

Okamoto, H., Yasui, K., Zhao, C., Arii, S., and Inazawa, J. (2003). Ptk2 and Eif3s3 Genes May Be Amplification Targets at 8q23–Q24 and Are Associated with Large Hepatocellular Carcinomas. *Hepatology* 38, 1242–1249. doi:10.1053/jhep.2003.50457

Qiu, M., Li, G., Wang, P., Li, X., Lai, F., Luo, R., et al. (2020). Aarf Domain Containing Kinase 5 Gene Promotes Invasion and Migration of Lung Cancer Cells through Adck5-Sox9-Pttg1 Pathway. *Exp. Cel. Res.* 392, 112002. doi:10.1016/j.yexcr.2020.112002

Ramassone, A., Pagotto, S., Veronese, A., and Visone, R. (2018). Epigenetics and Microrna s in Cancer. *Int. J. Mol. Sci.* 19, 459. doi:10.3390/ijms19020459

Risso, D., Schwartz, K., Sherlock, G., and Dudoit, S. (2011). Gc-content Normalization for Rna-Seq Data. *BMC Bioinformatics* 12, 1–17. doi:10.1186/1471-2105-12-480

Safonov, A., Jiang, T., Bianchini, G., Győrffy, B., Karn, T., Hatzis, C., et al. (2017). Immune Gene Expression Is Associated with Genomic Aberrations in Breast Cancer. *Cancer Res.* 77, 3317–3324. doi:10.1158/0008-5472.can-16-3478

Serrano-Carbajal, E. A., Espinal-Enríquez, J., and Hernández-Lemus, E. (2020). Targeting Metabolic Deregulation Landscapes in Breast Cancer Subtypes. *Front. Oncol.* 10, 97. doi:10.3389/fonc.2020.00097

Sethuraman, A., Brown, M., Seagroves, T. N., Wu, Z.-H., Pfeffer, L. M., and Fan, M. (2016). Smarce1 Regulates Metastatic Potential of Breast Cancer Cells through the Hif1a/ptk2 Pathway. *Breast Cancer Res.* 18, 1–15. doi:10.1186/s13058-016-0738-9

Shao, X., Lv, N., Liao, J., Long, J., Xue, R., Ai, N., et al. (2019). Copy Number Variation Is Highly Correlated with Differential Gene Expression: a Pan-Cancer Study. *BMC Med. Genet.* 20, 1–14. doi:10.1186/s12881-019-0909-5

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

Skinner, H. D., Giri, U., Yang, L., Woo, S. H., Story, M. D., Pickering, C. R., et al. (2016). Proteomic Profiling Identifies Ptk2/fak as a Driver of Radioresistance in Hpv-Negative Head and Neck Cancer. *Clin. Cancer Res.* 22, 4643–4650. doi:10.1158/1078-0432.ccr-15-2785

Stein, K. D., Syrjala, K. L., and Andrykowski, M. A. (2008). Physical and Psychological Long-Term and Late Effects of Cancer. *Cancer* 112, 2577–2592. doi:10.1002/cncr.23448

Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N., et al. (2007). Relative Impact of Nucleotide and Copy Number Variation on Gene Expression Phenotypes. *Science* 315, 848–853. doi:10.1126/science.1136678

Sun, W., Bunn, P., Jin, C., Little, P., Zhabotynsky, V., Perou, C. M., et al. (2018). The Association between Copy Number Aberration, Dna Methylation and Gene Expression in Tumor Samples. *Nucleic Acids Res.* 46, 3009–3018. doi:10.1093/nar/gky131

Tarazona, S., García, F., Ferrer, A., Dopazo, J., and Conesa, A. (2011). Noiseq: a Rna-Seq Differential Expression Method Robust for Sequencing Depth Biases. *Embnet. J.* 17, 18–19. doi:10.14806/ej.17.b.265

Tarazona, S., Furió-Tarí, P., Turrà, D., Pietro, A. D., Nueda, M. J., Ferrer, A., et al. (2015). Data Quality Aware Analysis of Differential Expression in Rna-Seq with Noiseq R/bioc Package. *Nucleic Acids Res.* 43, e140. doi:10.1093/nar/gkv711

Terrell, G. R., and Scott, D. W. (1992). Variable Kernel Density Estimation. *Ann. Stat.* 20, 1236–1265. doi:10.1214/aos/1176348768

Tong, Y., Tang, Y., Li, S., Zhao, F., Ying, J., Qu, Y., et al. (2020). Cumulative Evidence of Relationships between Multiple Variants in 8q24 Region and Cancer Incidence. *Medicine* 99, e20716. doi:10.1097/MD.0000000000020716

Tran, B., and Bedard, P. L. (2011). Luminal-b Breast Cancer and Novel Therapeutic Targets. *Breast Cancer Res.* 13, 1–10. doi:10.1186/bcr2904

Voutsadakis, I. A. (2020). 8p11. 23 Amplification in Breast Cancer: Molecular Characteristics, Prognosis and Targeted Therapy. *J. Clin. Med.* 9, 3079. doi:10.3390/jcm9103079

Wang, F., Gatica, D., Ying, Z. X., Peterson, L. F., Kim, P., Bernard, D., et al. (2019). Follicular Lymphoma–Associated Mutations in Vacuolar Atpase Atp6v1b2 Activate Autophagic Flux and Mtor. *J. Clin. Invest.* 129, 1626–1640. doi:10.1172/jci98288

Wang, X., He, X., Guo, H., and Tong, Y. (2020). Variants in the 8q24 Region Associated with Risk of Breast Cancer: Systematic Research Synopsis and Meta-Analysis. *Medicine* 99, e19217. doi:10.1097/MD.0000000000019217

White, R. J. (2004). Rna Polymerase Iii Transcription and Cancer. *oncogene* 23, 3208–3216. doi:10.1038/sj.onc.1207547

Wokolorczyk, D., Gliniewicz, B., Sikorski, A., Zlowocka, E., Masojc, B., Dkebniak, T., et al. (2008). A Range of Cancers Is Associated with the Rs6983267 Marker on Chromosome 8. *Cancer Res.* 68, 9982–9986. doi:10.1158/0008-5472.CAN-08-1838

Zamora-Fuentes, J. M., Hernández-Lemus, E., and Espinal-Enríquez, J. (2020). Gene Expression and Co-expression Networks Are Strongly Altered through Stages in clear Cell Renal Carcinoma. *Front. Genet.* 11, 578679. doi:10.3389/fgene.2020.578679

Zhang, X., Zhao, X.-M., He, K., Lu, L., Cao, Y., Liu, J., et al. (2012). Inferring Gene Regulatory Networks from Gene Expression Data by Path Consistency Algorithm Based on Conditional Mutual Information. *Bioinformatics* 28, 98–104. doi:10.1093/bioinformatics/btr626

Zhang, S., Li, X., Wang, H.-Y., and Zheng, X. S. (2018). Beyond Regulation of Pol Iii: Role of Maf1 in Growth, Metabolism, Aging and Cancer. *Biochim. Biophys. Acta (BBA)-Gene Regul. Mech.* 1861, 338–343. doi:10.1016/j.bbagrm.2018.01.019

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 | Frontiers in Genetics

ORIGINAL RESEARCH
published: 05 April 2022
doi: 10.3389/fgene.2022.796444

The Association Between Cyclooxygenase-2 –1195G/A (rs689466) Gene Polymorphism and the Clinicopathology of Lung Cancer in the Japanese Population: A Case-Controlled Study

*Rong Sun [1], Ryosuke Tanino [1], Xuexia Tong [2], Minoru Isomura [3], Li-Jun Chen [4], Takamasa Hotta [1], Tamio Okimoto [1], Megumi Hamaguchi [1], Shunichi Hamaguchi [1], Yasuyuki Taooka [5], Takeshi Isobe [1] and Yukari Tsubata [1]**

[1]Department of Internal Medicine, Division of Medical Oncology and Respiratory Medicine, Faculty of Medicine, Shimane University, Shimane, Japan, [2]Department of Respiratory and Critical Care Medicine, General Hospital of Ningxia Medical University, Yinchuan, China, [3]Department of Pathology, Shimane University Faculty of Medicine, Shimane University, Shimane, Japan, [4]Department of Respiratory Medicine, Second Affiliated Hospital of Ningxia Medical University, Yinchuan, China, [5]Division of Internal Medicine, Department of Respiratory Medicine, Medical Corporation JR Hiroshima Hospital, Hiroshima, Japan

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez, Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Benjamin Rybicki, Henry Ford Health System, United States
Katarzyna Bogunia-Kubik, Polish Academy of Sciences, Poland

***Correspondence:**
Yukari Tsubata
ytsubata@med.shimane-u.ac.jp

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

Received: 16 October 2021
Accepted: 17 February 2022
Published: 05 April 2022

Citation:
Sun R, Tanino R, Tong X, Isomura M, Chen L-J, Hotta T, Okimoto T, Hamaguchi M, Hamaguchi S, Taooka Y, Isobe T and Tsubata Y (2022) The Association Between Cyclooxygenase-2 –1195G/A (rs689466) Gene Polymorphism and the Clinicopathology of Lung Cancer in the Japanese Population: A Case-Controlled Study. Front. Genet. 13:796444. doi: 10.3389/fgene.2022.796444

The single nucleotide polymorphisms of COX-2 gene, also known as *PTGS2*, which encodes a pro-inflammatory factor cyclooxygenase-2, alter the risk of developing multiple tumors, but these findings are not consistent for lung cancer. We previously reported that the homozygous COX-2 –1195A genotype is associated with an increased risk for chronic obstructive pulmonary disease (COPD) in Japanese individuals. COPD is a significant risk factor for lung cancer due to genetic susceptibility to cigarette smoke. In this study, we investigated the association between COX-2 –1195G/A polymorphism and lung cancer susceptibility in the Japanese population. We evaluated the genotype distribution of COX-2 –1195G/A using a polymerase chain reaction-restriction fragment length polymorphism assay for 330 newly diagnosed patients with lung cancer and 162 healthy controls. Our results show that no relationship exists between the COX-2 –1195G/A polymorphism and the risk of developing lung cancer. However, compared to the control group, the homozygous COX-2 –1195A genotype increased the risk for lung squamous cell carcinoma (odds ratio = 2.902; 95% confidence interval, 1.171–7.195; p = 0.021), whereas no association is observed with the risk for adenocarcinoma. In addition, Kaplan-Meier analysis shows that the genotype distribution of homozygous COX-2 –1195A does not correlate with the overall survival of patients with lung squamous cell carcinoma. Thus, we conclude that the homozygous COX-2 –1195A genotype confers an increased risk for lung squamous cell carcinoma in Japanese individuals and could be used as a predictive factor for early detection of lung squamous cell carcinoma.

Keywords: cyclooxygenase-2, single nucleotide polymorphism, promoter region, lung cancer risk, squamous cell carcinoma, Japanese

1 INTRODUCTION

Lung cancer has the highest morbidity and mortality among cancers worldwide (Bray et al., 2018). There are many environmental risk factors for lung cancer and the most common is tobacco smoking (de Groot et al., 2018). However, exposure to environmental risk factors and genetic susceptibility are necessary for the development of lung carcinogenesis (Barta et al., 2019). Therefore, identifying genetic variants is vital for early prevention and screening of lung cancer.

Chronic inflammation plays an important role in carcinogenesis. Airway injuries related to inflammation caused by tobacco smoke or other environmental exposures increase the risk of developing lung cancer (Hecht, 2008). Moreover, inflammation-related genetic variants are highly associated with carcinogenesis for various cancers, including lung cancer (Rudnicka et al., 2019; Tan et al., 2019; Xiong et al., 2019).

The expression level of cyclooxygenase-2, a pro-inflammatory factor encoded by the COX-2 gene, is negligible in normal cells (Gurram et al., 2018). However, it is commonly overexpressed in various types of cancers and is implicated in the tumorigenesis, proliferation, metastasis, prognosis, and treatment of cancers (Tomozawa et al., 2000; Liu et al., 2001; Patti et al., 2002; Li et al., 2020). COX-2 is also involved in the molecular pathogenesis of chronic lung diseases (Park and Christman, 2006). Polymorphisms in the COX-2 gene alter the risk for chronic lung disease (Szczeklik et al., 2004; Xaubet et al., 2010). Three single nucleotide polymorphisms (SNPs) of the potential function, −1290G/A (rs689465), −1195G/A (rs689466), and −765G/C (rs20417), were identified in the COX-2 gene in esophageal cancer (Zhang et al., 2005). We previously reported that the homozygous COX-2 −1195A genotype is associated with an increased risk for chronic obstructive pulmonary disease (COPD) in the Japanese population (Chen et al., 2013). COPD is the single risk factor identified for the development of lung cancer after smoking exposure (Young et al., 2009). Chronic inflammation increases the risk for lung cancer by 2- to 3- fold in patients with COPD (Koshiol et al., 2009; Schwartz et al., 2016). Meta-analysis suggests that the emphysema detected visually on chest computed tomography (CT) and reduced forced expiratory volume in 1 s (FEV1) have strong effect on the increased odds of developing lung cancer (Wasswa-Kintu et al., 2005; Smith et al., 2012). Genetic analysis suggests that the genetic risk factors predisposing smokers to COPD and lung cancer may overlap (Young and Hopkins, 2011), and the key inflammatory-related genes and pathways impact the risk for lung cancer in a COPD-dependent manner (Watza et al., 2020). COX-2 is reported to be one of the candidate susceptibility genes related to inflammation involved in both COPD and lung cancer (Sekine et al., 2012).

The COX-2 −1195G/A gene polymorphism is functional and associated with an increased risk for various human cancers; however, the results are controversial in lung cancer (Zhang et al., 2005; Dong et al., 2010; Coskunpinar et al., 2011; Tang et al., 2011; Moraes et al., 2017). Therefore, this case-controlled study aimed to investigate the association between the COX-2 −1195G/A gene polymorphism and lung cancer susceptibility in the Japanese population.

2 MATERIALS AND METHODS

2.1 Study Design and Participants

This study included 492 participants from the Japanese population. The enrolled 330 patients with lung cancer were newly diagnosed at the Shimane University Hospital or Higashi Hiroshima Medical Center between 2009 and 2012. The lung cancer cases consisted of 221 patients with lung adenocarcinoma, 85 patients with lung squamous cell carcinoma, 9 patients with small cell lung cancer, and 15 patients with the other types. The 162 healthy controls were randomly selected from participants who received an annual health screening at the Shimane Institute of Health Science between 2009 and 2012. Those who were diagnosed with any cancer or any respiratory disease should be excluded from the controls. Ethical approval was obtained from the Institutional Review Board at the Shimane University Faculty of Medicine and the Higashi Hiroshima Medical Center (approved number 1022). Each enrolled participant signed an informed consent form.

2.2 DNA Preparation and Genotype Determination

DNA from enrolled participants was isolated from whole blood. A polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) assay was used for COX-2 polymorphism determination. The PCR reactions were performed in a reaction mixture system volume of 50 μl that contained 2.5 U Taq and 1 μl template DNA at a concentration of 50–150 ng/ml. The genotype of COX-2 −1195G/A was determined using the following specific primers: 5′- CCC TGA GCA CTA CCC ATG AT -3′ (forward) and 5′- GCC CTT CAT AGG AGA TAC TGG -3′ (reverse). The PCR cycling program was as follows: incubation at 96°C for 5 min followed by 32 cycles of 96°C for 20 s, 52°C for 30 s, and 72°C for 30 s with a final extension at 72°C for 6 min. The COX-2 273 bp PCR product was digested into 220 bp and 53 bp fragments with PvuII for the −1195G allele (restriction products: AA, 273 bp; GG, 220 bp + 53 bp; GA, 220 bp + 53 bp + 273 bp). The digested products were observed in a 2% agarose gel stained with ethidium bromide, and images were obtained under ultraviolet light.

2.3 Statistical Analysis

Data are presented as the number (%) of participants. All statistical analyses were performed using SPSS Statistics version 27.0 (IBM, Armonk, NY, United States). Demographic characteristics were analyzed using a Mann-Whitney *U*-test or a Chi-squared test. The distribution of genotypes was assessed using the Hardy-Weinberg equilibrium. Differences in the genotype distribution of COX-2 gene were analyzed using a Chi-squared test. The association between the genotype distribution of COX-2 −1195G/A and lung cancer risk was estimated using odds ratios (ORs) and a 95% confidence interval (95% CI) that were computed using logistic regression analysis. Survival distributions were estimated using Kaplan-Meier analysis and compared using the log-rank test. $P < 0.05$ was considered statistically significant.

TABLE 1 | Demographic characteristics of the study participants.

	Lung cancer (n = 330)	Control (n = 162)	*p*-value
Age			
Median	71	51	
Range	35–96	18–86	< 0.001
Sex			
Male (%)	255 (77.3)	99 (61.1)	
Female (%)	75 (22.7)	63 (38.9)	< 0.001
Smoker			
No (%)	82 (24.8)	85 (52.5)	
Yes (%)	248 (75.2)	77 (47.5)	< 0.001

Note: p-*values are presented for the comparison between the lung cancer group and the control group.*

TABLE 2 | Baseline clinicopathology and stage characteristics for patients with lung cancer.

	Number (%)
Histology	
Adenocarcinoma	221 (67.0)
Squamous cell carcinoma	85 (25.8)
Small cell carcinoma	9 (2.7)
Others	15 (4.5)
Stage	
I	9 (2.7)
II	4 (1.2)
IIIA	38 (11.5)
IIIB	78 (23.6)
IV	201 (60.9)

Note: values represent the number (%) of participants.

3 RESULTS

The demographic characteristics of all study participants are summarized in **Table 1**, which includes age, sex, and smoking history. We observed the lung cancer group had a higher median age than that of the control group ($p < 0.001$). Moreover, there was a higher prevalence of men and smoking in the lung cancer group than in the control group (all values of $p < 0.001$). The clinicopathology and the disease stages for the patients with lung cancer are listed in **Table 2**. Adenocarcinoma, squamous cell carcinoma, small cell carcinoma, and the others represented 67, 25.8, 2.7, and 4.5% of all patients with lung cancer, respectively. Stage I + II, stage III, and stage IV represented 3.9, 35.1, and 60.9% of all patients with lung cancer, respectively. The 330 patients with lung cancer and 162 healthy controls were genotyped for the COX-2 −1195G/A polymorphism (**Table 3**). The genotype distribution in the control group was consistent with the Hardy-Weinberg equilibrium. Moreover, for the genotype distribution of COX-2 −1195G/A, there is no significant difference between lung cancer patients and the controls ($p = 0.17$), and also no significant difference was obsearved between the adenocarcinoma group and the squamous cell carcinoma group ($p = 0.158$).

4 OUTCOME

No association on the risk of developing lung cancer with the genotype distribution of COX-2 −1195G/A was observed between the lung cancer and the control groups in both unadjusted analyses and adjusted analyses with age, sex, and smoking status (**Table 3**). Even though stratified by smoking status or sex in this study, we did not observe any association on the genotype distribution of COX-2 −1195G/A with the risk for lung cancer in both unadjusted and adjusted analyses with their respective factors (**Table 4** and **Table 5**). In addition, no significant difference of the genotype distribution of COX-2 −1195G/A was found among the disease stages and the controls (**Table 6**). In the subgroup analyses, no increased risk was obsearved on the genotype distribution of the homozygous −1195A compared to the homozygous −1195G in the patients with stage IV (OR = 0.664; 95% CI, 0.346–1.272; $p = 0.271$). However, the multinomial logistic regression analysis represented that the genotype distribution of COX-2 −1195G/A was associated with the increasing risk for squamous cell carcinoma while not associated with adenocarcinoma (**Table 7**). The homozygous −1195A genotype increased the risk of 2.902 times for squamous cell carcinoma than the homozygous −1195G genotype (OR = 2.902; 95% CI, 1.171–7.195; $p = 0.021$). While no significant difference was observed from the heterozygous −1195G/A genotype for developing squamous cell carcinoma (OR = 1.618; 95% CI, 0.681–3.843; $p = 0.275$). To assess whether the prognosis was affected by the genotype of COX-2 −1195G/A, we compared the overall survival (OS). The median OS did not correlate with the genotype distribution of COX-2 −1195G/A among patients with

TABLE 3 | The genotype distribution of COX-2 −1195G/A in patients with lung cancer and control participants.

	Lung cancer (n = 330)	Control (n = 162)	Unadjusted OR (95% CI)	*p*-value	Adjusted OR (95% CI)	*p*-value
Overall lung cancer						
Homozygous G	52 (15.8)	28 (17.3)	1		1	
Heterozygous G/A	167 (50.6)	93 (57.4)	0.967 (0.572–1.634)	0.9	0.949 (0.469–1.921)	0.884
Homozygous A	111 (33.6)	41 (25.3)	1.458 (0.814–2.610)	0.2	1.316 (0.608–2.845)	0.486

Note: OR: odds ratio; 95% CI: 95% confidence interval. The model is adjusted for the distributions of age, sex and smoking status in all the participants.

TABLE 4 | The genotype distribution of COX-2 −1195G/A gene stratified by smoking status.

	Lung cancer (n = 330)	Control (n = 162)	Unadjusted OR (95% CI)	*p*-value	Adjusted OR (95% CI)	*p*-value
Non-smokers						
Homozygous G	18 (22.0)	18 (21.2)	1		1	
Heterozygous G/A	45 (54.9)	45 (52.9)	1.000 (0.462–2.166)	1.0	0.975 (0.346–2.749)	0.962
Homozygous A	19 (23.1)	22 (25.9)	0.864 (0.352–2.117)	0.749	1.039 (0.329–3.279)	0.948
Smokers						
Homozygous G	34 (13.7)	10 (13.0)	1		1	
Heterozygous G/A	122 (49.2)	48 (62.3)	0.748 (0.343–1.631)	0.465	0.983 (0.373–2.590)	0.972
Homozygous A	92 (37.1)	19 (24.7)	1.424 (0.602–3.368)	0.421	1.522 (0.533–4.343)	0.433

Note: OR: odds ratio; 95% CI: 95% confidence interval. The model is adjusted for the distributions of age and sex in the respective participants.

TABLE 5 | The genotype distribution of COX-2 −1195G/A gene stratified by sex.

	Lung cancer (n = 330)	Control (n = 162)	Unadjusted OR (95% CI)	*p*-value	Adjusted OR (95% CI)	*p*-value
Male						
Homozygous G	40 (15.7)	18 (18.2)	1		1	
Heterozygous G/A	122 (47.8)	56 (56.6)	0.980 0.517–1.859)	0.952	1.253 (0.526–2.985)	0.611
Homozygous A	93 (36.5)	25 (25.3)	1.674 (0.823–3.406)	0.155	1.802 (0.707–4.588)	0.217
Female						
Homozygous G	12 (16.0)	10 (15.9)	1		1	
Heterozygous G/A	45 (60.0)	37 (58.7)	1.014 (0.394–2.608)	0.978	0.620 (0.181–2.127)	0.447
Homozygous A	18 (24.0)	16 (25.4)	0.938 (0.320–2.750)	0.906	0.676 (0.171–2.669)	0.576

Note: OR: odds ratio; 95% CI: 95% confidence interval. The model is adjusted for the distributions of age and smoking status in the respective participants.

TABLE 6 | The comparison on the genotype distribution of COX-2 −1195G/A gene among lung cancer patients with different disease stages and control participants.

	Stage IV (n = 201)	Stage IIIB (n = 78)	Stage IIIA (n = 38)	Stage I + II (n = 13)	Control (n = 162)	*p*-value
Homozygous G	29 (14.4)	17 (21.8)	4 (10.5)	2 (15.4)	28 (17.3)	
Heterozygous G/A	108 (53.7)	35 (44.9)	19 (50.0)	5 (38.5)	93 (57.4)	0.371
Homozygous A	64 (31.8)	26 (33.3)	15 (39.5)	6 (46.2)	41 (25.3)	

Note: p-*values are presented for comparison among lung cancer patients with different disease stages and control participants.*

TABLE 7 | The genotype distribution of COX-2 −1195G/A in adenocarcinoma and squamous cell carcinoma.

	Lung cancer (n = 330)	Control (n = 162)	OR (95%CI)	*p*-value
Adenocarcinoma				
Homozygous G	39 (17.6)	28 (17.3)	1	
Heterozygous G/A	110 (49.8)	93 (57.4)	0.849 (0.486–1.484)	0.566
Homozygous A	72 (32.6)	41 (25.3)	1.261 (0.679–2.341)	0.463
Squamous cell carcinoma				
Homozygous G	8 (9.4)	28 (17.3)	1	
Heterozygous G/A	43 (50.6)	93 (57.4)	1.618 (0.681–3.843)	0.275
Homozygous A	34 (40.0)	41 (25.3)	2.902 (1.171–7.195)	0.021

Note: OR: odds ratio; 95% CI: 95% confidence interval.

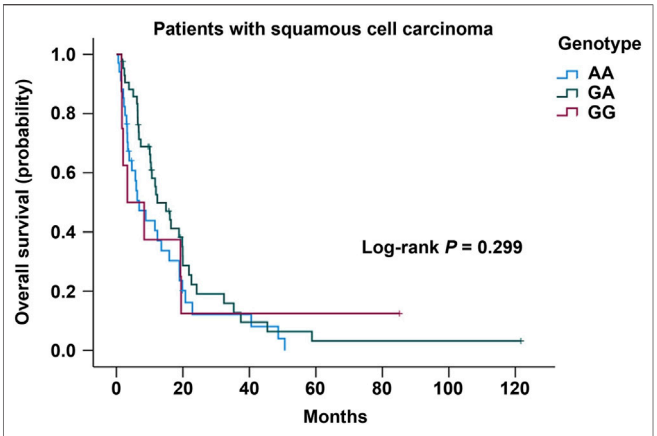


FIGURE 1 | Kaplan-Meier analysis of overall survival in patients with lung squamous cell carcinoma stratified by the genotype of COX-2 –1195G/A.

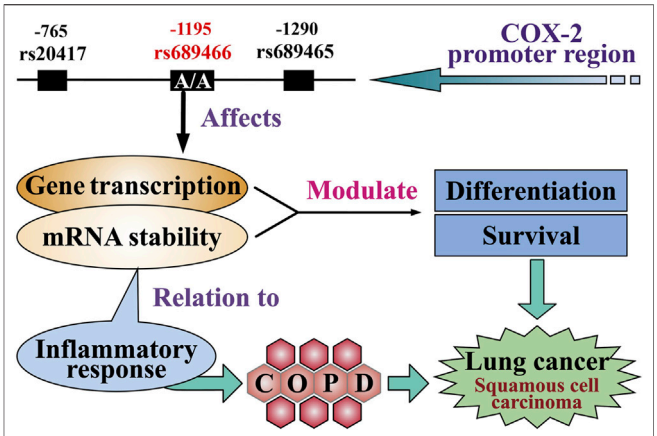


FIGURE 2 | Potential function of the COX-2 single nucleotide polymorphisms in lung cancer. The genotype of homozygous COX-2 –1195A in the COX-2 promoter region affects gene transcription, thereby increasing the expression of COX-2 in lung cells. The inflammatory response is related to mRNA stability. Both factors enhance the level of differentiation of lung cells and promote the development of lung squamous cell carcinoma. The inflammatory response also leads to chronic obstructive pulmonary disease and lung cancer.

squamous cell carcinoma (log-rank: p = 0.299) (**Figure 1**). In addition, the clinicopathology and sex were not related to the OS in patients with lung cancer (**Supplementary Figure S1** and **Supplementary Figure S2**). Furthermore, there is no significant difference in the OS between the genotype of COX-2 –1195G/A stratified by sex (**Supplementary Figure S3** and **Supplementary Figure S4**). Based on the above outcome, we summarized that the homozygous COX-2 –1195A genotype might increase the risk for lung squamous cell carcinoma in the Japanese population but no effect on the prognosis of squamous cell carcinoma.

5 DISCUSSION

This study analyzed the association between genotypes of the COX-2 –1195G/A polymorphism and different clinicopathology of lung cancer, and the results demonstrate that the homozygous COX-2 –1195A genotype was associated with the increased risk of developing lung squamous cell carcinoma. To date, the related studies have inconsistent results. The homozygous COX-2 –1195A genotype increased the risk of lung cancer development in the Turkish population, wherein patients with lung squamous cell carcinoma represented 53.2% (Coskunpinar et al., 2011). By contrast, a study from the Brazilian population concluded that the COX-2 –1195G/A polymorphism was not associated with the risk for lung cancer (Moraes et al., 2017), which is in agreement with the results from our present study. On the other hand, the number of patients diagnosed with lung squamous cell carcinoma observed in the Brazilian study is 39.4% (Moraes et al., 2017), and that of our study is 25.8%. Based on these results, we hypothesized that the discrepancy observed in the number of patients with homozygous COX-2 –1195A might be due to the different distribution of lung cancer clinicopathology. Indeed, our results reveal that the homozygous COX-2 –1195A genotype increases the risk for lung squamous cell carcinoma.

Increased levels of COX-2 expression were observed in bronchial precursors of squamous cell carcinoma using immunohistochemistry (Petkova et al., 2004; Mascaux et al., 2005). The substitution of –1195G > A creates a binding site for a transcription factor c-MYB in the COX-2 promoter region, which regulates the balance among cell division, differentiation, and survival, resulting in facilitation of COX-2 transcription (Ramsay et al., 2003; Zhang et al., 2005). Moreover, the homozygous COX-2 –1195A genotype exhibit a significant increase in the mRNA level of COX-2 expression than the genotypes of homozygous –1195G and/or the heterozygous –1195 GA in esophageal tissue (Zhang et al., 2005). In lung cancer, a Brazilian study demonstrated that the homozygous –1195A did not increase the mRNA expression of COX-2 compared to the other genotypes (Moraes et al., 2017). However, the 34 lung tumor specimens comprised 19 (55.9%) adenocarcinoma cases and 15 (44.1%) squamous cell carcinoma cases in the Brazilian study. The association between COX-2 -1195G/A polymorphism and the risk for lung cancer may be pathologically and ethnically dependent. Further studies with larger sample sizes that include populations of different races and analyses stratified by histology classifications are necessary to investigate the controversial results.

We previously demonstrated that the homozygous COX-2 –1195A genotype increased the risk for COPD in Japanese individuals (Chen et al., 2013). A Swedish study showed that the association with a lower FEV1 was higher for patients with lung squamous cell carcinoma than those with lung adenocarcinoma (Purdue et al., 2007). Further, the presence of emphysema, a typical manifestation of COPD on a chest CT scan, is associated with significantly increased odds of developing squamous cell carcinoma (Wang et al., 2018). Moreover, smoking is a major risk factor in the pathogenesis of lung squamous cell carcinoma and COPD, and it upregulates inflammation-related genes, including COX-2, in tracheal smooth muscle cells (Yang et al., 2009). Therefore, a potential

link might exist between the functional COX-2 SNPs, COPD, and lung cancer, particularly for lung squamous cell carcinoma (**Figure 2**) (Lee et al., 2009). Considering the possible relationship between pulmonary function, emphysema CT scan parameters, smoking status, and COX-2 –1195A homozygosity, further studies are required.

We did not observe any association between the risk for lung cancer and the COX-2 –1195G/A polymorphism stratified by smoking status. By contrast, a previous study from the Taiwanese population reported that the enrolled patients who smoked and carried the A allele of rs2066826 in the COX-2 intron 6 had an increased risk of 2.21 for lung cancer. (Liu et al., 2010). Further studies are needed to comprehensively analyze the functional COX-2 polymorphisms in addition to geographic populations.

The relationship between COX-2, COPD, and lung cancer is complicated. Epithelial-to-mesenchymal transition (EMT) is critical for lung carcinogenesis and observation of a malignant phenotype, and inhibition of COX-2 reverses EMT-induced changes in lung cancer patients (Dohadwala et al., 2006; Peebles et al., 2007). EMT in COPD and the resultant association with the risk for lung cancer have not been completely elucidated. Fundamental research is necessary to identify the molecular mechanisms linking these diseases.

Genotyping patients and identifying those with homozygous COX-2 –1195A could be combined with identifying emphysema using chest CT scans to serve as predictive markers for the early prevention and screening of lung squamous cell carcinoma. The increased odds of developing lung cancer in the presence of emphysema on CT may prove to be useful in targeting resources for the prevention and screening of lung squamous cell carcinoma. In addition, our findings suggest that either shared host susceptibility or an uncharacterized novel mechanism promotes the pathogenesis of both COPD and lung squamous cell carcinoma. It is necessary to further explore the benefit of clinical interventions to prevent or detect lung cancer after a patient is diagnosed with emphysema.

On the other hand, it has been reported that high levels of COX-2 mRNA transcription are associated with a more aggressive phenotype and poor prognosis for patients with non-small cell lung cancer (NSCLC) (Brabender et al., 2002). The homozygous COX-2 –1195A genotype is associated with poor overall survival in Chinese patients with NSCLC treated with chemoradiotherapy or radiotherapy alone (Bi et al., 2010). Although the homozygous COX-2 –1195A increased the risk for lung squamous cell carcinoma, this genotype did not correlate with poor prognosis in our study when evaluated using median overall survival. One reason for this discrepancy might be the fact that certain genetic markers are ethnicity-specific; another reason might be that different treatment regimens play a role in the prognosis of lung cancer and influence the effects of the COX-2 genotypes.

The three limitations of this study are listed as follows: the number of enrolled participants was low; only patients from the Japanese population were included; it was an imbalance of the baseline characteristics between the patients with lung cancer and the control participants.

In conclusion, the homozygous COX-2 –1195A increased the risk of developing lung squamous cell carcinoma and might be used as a predictive marker for early detection and screening of lung squamous cell carcinoma in Japanese individuals, but not as a predictive marker for the prognosis.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The data that support the findings of this study are available from Shimane university hospital but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Requests to access these datasets should be directed to Yukari Tsubata, ytsubata@med.shimane-u.ac.jp.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board at the Shimane University Faculty of Medicine and the Higashi Hiroshima Medical Center (approved number 1022). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RS, RT, TI, and YT: conception, design, supervision, resources, manuscript writing, revising, and editing. MI and LC: experiments conduction and data analysis. RS, MI, RT, XT, TH, TO, MH, SH, and YT: data visualization and statistical analysis. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

RS would like to thank the Otsuka Toshimi Scholarship Foundation for its support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.796444/full#supplementary-material

REFERENCES

Barta, J. A., Powell, C. A., and Wisnivesky, J. P. (2019). Global Epidemiology of Lung Cancer. *Ann. Glob. Health* 85. doi:10.5334/aogh.2419

Bi, N., Yang, M., Zhang, L., Chen, X., Ji, W., Ou, G., et al. (2010). Cyclooxygenase-2 Genetic Variants Are Associated with Survival in Unresectable Locally Advanced Non-small Cell Lung Cancer. *Clin. Cancer Res.* 16, 2383–2390. doi:10.1158/1078-0432.ccr-09-2793

Brabender, J., Park, J., Metzger, R., Schneider, P. M., Lord, R. V., Hölscher, A. H., et al. (2002). Prognostic Significance of Cyclooxygenase 2 mRNA Expression in Non-small Cell Lung Cancer. *Ann. Surg.* 235, 440–443. doi:10.1097/00000658-200203000-00017

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

Chen, L. J., Xu, W., Taooka, Y., Ohe, M., Takahashi, H., Sutani, A., et al. (2013). Cyclooxygenase 2 1195G > A Polymorphism Is Associated with Chronic Obstructive Pulmonary Disease in Japanese and Chinese Patients$\\gtrbin $. *Chin. Med. J. (Engl)* 126, 2215–2221.

Coskunpinar, E., Eraltan, I. Y., Turna, A., and Agachan, B. (2011). Cyclooxygenase-2 Gene and Lung Carcinoma Risk. *Med. Oncol.* 28, 1436–1440. doi:10.1007/s12032-010-9627-8

de Groot, P. M., Wu, C. C., Carter, B. W., and Munden, R. F. (2018). The Epidemiology of Lung Cancer. *Transl. Lung Cancer Res.* 7, 220–233. doi:10.21037/tlcr.2018.05.06

Dohadwala, M., Yang, S.-C., Luo, J., Sharma, S., Batra, R. K., Huang, M., et al. (2006). Cyclooxygenase-2-Dependent Regulation of E-Cadherin: Prostaglandin E2Induces Transcriptional Repressors ZEB1 and Snail in Non-small Cell Lung Cancer. *Cancer Res.* 66, 5338–5345. doi:10.1158/0008-5472.can-05-3635

Dong, J., Dai, J., Zhang, M., Hu, Z., and Shen, H. (2010). Potentially functionalCOX-2–1195G>A Polymorphism Increases the Risk of Digestive System Cancers: A Meta-Analysis$\\gtrbin $A Polymorphism Increases the Risk of Digestive System Cancers: A Meta-Analysis. *J. Gastroenterol. Hepatol.* 25, 1042–1050. doi:10.1111/j.1440-1746.2010.06293.x

Gurram, B., Zhang, S., Li, M., Li, H., Xie, Y., Cui, H., et al. (2018). Celecoxib Conjugated Fluorescent Probe for Identification and Discrimination of Cyclooxygenase-2 Enzyme in Cancer Cells. *Anal. Chem.* 90, 5187–5193. doi:10.1021/acs.analchem.7b05337

Hecht, S. S. (2008). Progress and Challenges in Selected Areas of Tobacco Carcinogenesis. *Chem. Res. Toxicol.* 21, 160–171. doi:10.1021/tx7002068

Koshiol, J., Rotunno, M., Consonni, D., Pesatori, A. C., De Matteis, S., Goldstein, A. M., et al. (2009). Chronic Obstructive Pulmonary Disease and Altered Risk of Lung Cancer in a Population-Based Case-Control Study. *PloS one* 4, e7380. doi:10.1371/journal.pone.0007380

Lee, G., Walser, T. C., and Dubinett, S. M. (2009). Chronic Inflammation, Chronic Obstructive Pulmonary Disease, and Lung Cancer. *Curr. Opin. Pulm. Med.* 15, 303–307. doi:10.1097/mcp.0b013e32832c975a

Li, M., Li, M., Wei, Y., and Xu, H. (2020). Prognostic and Clinical Significance of Cyclooxygenase-2 Overexpression in Endometrial Cancer: a Meta-Analysis. *Front. Oncol.* 10, 1202. doi:10.3389/fonc.2020.01202

Liu, C. H., Chang, S.-H., Narko, K., Trifan, O. C., Wu, M.-T., Smith, E., et al. (2001). Overexpression of Cyclooxygenase-2 Is Sufficient to Induce Tumorigenesis in Transgenic Mice. *J. Biol. Chem.* 276, 18563–18569. doi:10.1074/jbc.m010787200

Liu, C. J., Hsia, T. C., Wang, R. F., Tsai, C. W., Chu, C. C., Hang, L. W., et al. (2010). Interaction of Cyclooxygenase 2 Genotype and Smoking Habit in Taiwanese Lung Cancer Patients. *Anticancer Res.* 30, 1195–1199.

Mascaux, C., Martin, B., Verdebout, J.-M., Ninane, V., and Sculier, J.-P. (2005). COX-2 Expression during Early Lung Squamous Cell Carcinoma Oncogenesis. *Eur. Respir. J.* 26, 198–203. doi:10.1183/09031936.05.00001405

Moraes, J. L., Moraes, A. B., Aran, V., Alves, M. R., Schluckbier, L., Duarte, M., et al. (2017). Functional Analysis of Polymorphisms in the COX-2 Gene and Risk of Lung Cancer. *Mol. Clin. Oncol.* 6, 494–502. doi:10.3892/mco.2017.1167

Park, G. Y., and Christman, J. W. (2006). Involvement of Cyclooxygenase-2 and Prostaglandins in the Molecular Pathogenesis of Inflammatory Lung Diseases. *Am. J. Physiology-Lung Cell Mol. Physiol.* 290, L797–L805. doi:10.1152/ajplung.00513.2005

Patti, R., Gumired, K., Reddanna, P., Sutton, L. N., Phillips, P. C., and Reddy, C. D. (2002). Overexpression of Cyclooxygenase-2 (COX-2) in Human Primitive Neuroectodermal Tumors: Effect of Celecoxib and Rofecoxib. *Cancer Lett.* 180, 13–21. doi:10.1016/s0304-3835(02)00003-4

Peebles, K. A., Lee, J. M., Mao, J. T., Hazra, S., Reckamp, K. L., Krysan, K., et al. (2007). Inflammation and Lung Carcinogenesis: Applying Findings in Prevention and Treatment. *Expert Rev. anticancer Ther.* 7, 1405–1421. doi:10.1586/14737140.7.10.1405

Petkova, D. K., Clelland, C., Ronan, J., Pang, L., Coulson, J. M., Lewis, S., et al. (2004). Overexpression of Cyclooxygenase-2 in Non-small Cell Lung Cancer. *Respir. Med.* 98, 164–172. doi:10.1016/j.rmed.2003.09.006

Purdue, M. P., Gold, L., Jarvholm, B., Alavanja, M. C. R., Ward, M. H., and Vermeulen, R. (2007). Impaired Lung Function and Lung Cancer Incidence in a Cohort of Swedish Construction Workers. *Thorax* 62, 51–56. doi:10.1136/thx.2006.064196

Ramsay, R. G., Barton, A. L., and Gonda, T. J. (2003). Targeting C-Myb Expression in Human Disease. *Expert Opin. Ther. Targets* 7, 235–248. doi:10.1517/14728222.7.2.235

Rudnicka, K., Backert, S., and Chmiela, M. (2019). Genetic Polymorphisms in Inflammatory and Other Regulators in Gastric Cancer: Risks and Clinical Consequences. *Mol. Mech. Inflamm. induction, resolution escape by* Helicobacter pylori, 53–76. doi:10.1007/978-3-030-15138-6_3

Schwartz, A. G., Lusk, C. M., Wenzlaff, A. S., Watza, D., Pandolfi, S., Mantha, L., et al. (2016). Risk of Lung Cancer Associated with COPD Phenotype Based on Quantitative Image Analysis. *Cancer Epidemiol. Biomarkers Prev.* 25, 1341–1347. doi:10.1158/1055-9965.epi-16-0176

Sekine, Y., Katsura, H., Koh, E., Hiroshima, K., and Fujisawa, T. (2012). Early Detection of COPD Is Important for Lung Cancer Surveillance. *Eur. Respir. J.* 39, 1230–1240. doi:10.1183/09031936.00126011

Smith, B. M., Pinto, L., Ezer, N., Sverzellati, N., Muro, S., and Schwartzman, K. (2012). Emphysema Detected on Computed Tomography and Risk of Lung Cancer: a Systematic Review and Meta-Analysis. *Lung cancer* 77, 58–63. doi:10.1016/j.lungcan.2012.02.019

Szczeklik, W., Sanak, M., and Szczeklik, A. (2004). Functional Effects and Gender Association of COX-2 Gene Polymorphism G-765C in Bronchial Asthma. *J. Allergy Clin. Immunol.* 114, 248–253. doi:10.1016/j.jaci.2004.05.030

Tan, N., Song, J., Yan, M., Wu, J., Sun, Y., Xiong, Z., et al. (2019). Association between IL-4 Tagging Single Nucleotide Polymorphisms and the Risk of Lung Cancer in China. *Mol. Genet. Genomic Med.* 7, e00585. doi:10.1002/mgg3.585

Tang, Z., Nie, Z.-L., Pan, Y., Zhang, L., Gao, L., Zhang, Q., et al. (2011). The Cox-2 -1195 G > A Polymorphism and Cancer Risk: a Meta-Analysis of 25 Case-Control Studies$\\gtrbin $A Polymorphism and Cancer Risk: a Meta-Analysis of 25 Case-Control Studies. *Mutagenesis* 26, 729–734. doi:10.1093/mutage/ger040

Tomozawa, S., Tsuno, N. H., Sunami, E., Hatano, K., Kitayama, J., Osada, T., et al. (2000). Cyclooxygenase-2 Overexpression Correlates with Tumour Recurrence, Especially Haematogenous Metastasis, of Colorectal Cancer. *Br. J. Cancer* 83, 324–328. doi:10.1054/bjoc.2000.1270

Wang, W., Xie, M., Dou, S., Cui, L., Zheng, C., and Xiao, W. (2018). The Link between Chronic Obstructive Pulmonary Disease Phenotypes and Histological Subtypes of Lung Cancer: a Case–Control Study. *Copd* Vol. 13, 1167–1175. doi:10.2147/copd.s158818

Wasswa-Kintu, S., Gan, W., Man, S., Pare, P., and Sin, D. (2005). Relationship between Reduced Forced Expiratory Volume in One Second and the Risk of Lung Cancer: a Systematic Review and Meta-Analysis. *Thorax* 60, 570–575. doi:10.1136/thx.2004.037135

Watza, D., Lusk, C. M., Dyson, G., Purrington, K. S., Wenzlaff, A. S., Neslund-Dudas, C., et al. (2020). COPD-dependent Effects of Genetic Variation in Key Inflammation Pathway Genes on Lung Cancer Risk. *Int. J. Cancer* 147, 747–756. doi:10.1002/ijc.32780

Xaubet, A., Fu, W. J., Li, M., Serrano-Mollar, A., Ancochea, J., Molina-Molina, M., et al. (2010). A Haplotype of Cyclooxygenase-2 Gene Is Associated with Idiopathic Pulmonary Fibrosis. *Sarcoidosis Vasc. Diffuse Lung Dis.* 27, 121–130.

Xiong, Z., Sun, Y., Wu, J., Niu, F., Jin, T., and Li, B. (2019). Genetic Polymorphisms in IL1R1 and IL1R2 Are Associated with Susceptibility to Thyroid Cancer in the Chinese Han Population. *J. Gene Med.* 21, e3093. doi:10.1002/jgm.3093

Yang, C.-M., Lee, I.-T., Lin, C.-C., Yang, Y.-L., Luo, S.-F., Kou, Y. R., et al. (2009). Cigarette Smoke Extract Induces COX-2 Expression via a PKCα/c-Src/EGFR,

PDGFR/PI3K/Akt/NF-κB Pathway and P300 in Tracheal Smooth Muscle Cells. *Am. J. Physiology-Lung Cell Mol. Physiol.* 297, L892–L902. doi:10.1152/ajplung.00151.2009

Young, R. P., Hopkins, R. J., Christmas, T., Black, P. N., Metcalf, P., and Gamble, G. D. (2009). COPD Prevalence Is Increased in Lung Cancer, Independent of Age, Sex and Smoking History. *Eur. Respir. J.* 34, 380–386. doi:10.1183/09031936.00144208

Young, R. P., and Hopkins, R. J. (2011). How the Genetics of Lung Cancer May Overlap with COPD. *Respirology* 16, 1047–1055. doi:10.1111/j.1440-1843.2011.02019.x

Zhang, X., Miao, X., Tan, W., Ning, B., Liu, Z., Hong, Y., et al. (2005). Identification of Functional Genetic Variants in and Their Association with Risk of Esophageal Cancer. *Gastroenterology* 129, 565–576. doi:10.1016/j.gastro.2005.05.003

Conflict of Interest: TI reports personal fees from Boehringer Ingelheim, Pfizer, AstraZeneca and Daiichi-Sankyo outside the submitted work. YT reports personal fees from AstraZeneca, Daiichi-Sankyo, Chugai Pharmaceutical Co., Ltd. and Ono Pharmaceutical Co., Ltd. outside the submitted work. YT also reports scholarship grants from Pfizer Health Research Foundation outside the submitted work.

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 26 April 2022
doi: 10.3389/fgene.2022.778850

Assessing the Potential Prognostic and Immunological Role of TK1 in Prostate Cancer

Hui Xie[1†], Linpei Guo[2†], Zhun Wang[1], Shuanghe Peng[3], Qianwang Ma[1], Zhao Yang[1], Zhiqun Shang[1] and Yuanjie Niu[1]**

[1]Department of Urology, Tianjin Institute of Urology, the Second Hospital of Tianjin Medical University, Tianjin, China, [2]Department of Urology, the Affiliated Wuxi No. 2 People's Hospital of Nanjing Medical University, Wuxi, China, [3]Department of Pathology, Tianjin Institute of Urology, the Second Hospital of Tianjin Medical University, Tianjin, China

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez,
Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Zhitong Bing,
Institute of Modern Physics (CAS), China
Elena Pudova,
Engelhardt Institute of Molecular Biology (RAS), Russia

***Correspondence:**
Zhiqun Shang
zhiqun_shang@tmu.edu.cn
Yuanjie Niu
niuyuanjie68@163.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: 17 September 2021
Accepted: 22 February 2022
Published: 26 April 2022

Citation:
Xie H, Guo L, Wang Z, Peng S, Ma Q, Yang Z, Shang Z and Niu Y (2022) Assessing the Potential Prognostic and Immunological Role of TK1 in Prostate Cancer. Front. Genet. 13:778850. doi: 10.3389/fgene.2022.778850

Background: It has been reported that thymidine kinase 1 (TK1) was up-regulated in multiple malignancies and participated in the regulation of tumor malignant behavior. However, its specific role in prostate cancer (PCa) remains unclear.

Methods: TK1 expression in PCa patients and cell lines was identified via crossover analysis of the public datasets. A series of *in vitro* experiments and *in vivo* models was applied to investigate the function of TK1 in PCa. Functional enrichment analyses were further conducted to explore the underlying mechanism. Additionally, TISIDB was applied to explore the correlation between TK1 expression and tumor-infiltrating lymphocytes, immune subtypes, and immune regulatory factors.

Results: TK1 expression was significantly up-regulated in PCa patients and cell lines. TK1 ablation inhibited tumor cell proliferation and migration potential, and *in vivo* experiments showed that TK1 inactivation can significantly restrain tumor growth. Functional enrichment analysis revealed TK1-related hub genes (AURKB, CCNB2, CDC20, CDCA5, CDK1, CENPA, CENPM, KIF2C, NDC80, NUF2, PLK1, SKA1, SPC25, ZWINT), and found that TK1 was closely involved in the regulation of cell cycle. Moreover, elevated mRNA expression of TK1 was related with higher Gleason score, higher clinical stage, higher pathological stage, higher lymph node stage, shorter overall survival, and DFS in PCa patients. Particularly, TK1 represented attenuated expression in C3 PCa and was related with infiltration of $CD4^+$, $CD8^+$ T cells, and dendritic cells as well as immunomodulator expression.

Conclusion: Our study indicates that TK1 is a prognostic predictor correlated with poor outcomes of PCa patients, and for the first time represented that TK1 can promote the progression of PCa. Therefore, TK1 may be a potential diagnostic and prognostic biomarker, as well as a therapeutic target for PCa.

Keywords: thymidine kinase 1, bioinformatics, prognostic biomarker, tumor immunity, prostate cancer

INTRODUCTION

According to the latest American Cancer Society's statistics, prostate cancer (PCa) ranks first among estimated new cases and second in the number of estimated deaths (Siegel et al., 2020). Furthermore, with an estimated nearly 1.4 million new cases and 375,000 deaths worldwide, PCa is the second most frequent cancer and the fifth leading cause of cancer death among men in 2020 (Sung et al., 2021). At present, fully curative treatment still has not been found for the terminal stage of PCa, castration resistant prostate cancer (CRPC) (Wong et al., 2014). Meanwhile, numerous microarray and next-generation sequencing technologies have been applied to explore the etiology of PCa and to find the specific drug targets (Barbieri et al., 2017). Although important insights have been gained through the efforts, the underlying mechanisms are still not fully clarified. Cumulative evidence suggested that the carcinogenesis and development of PCa is a process involving multiple genes and signaling pathways (Taylor et al., 2010; Grasso et al., 2012; Latonen et al., 2018). Therefore, it is urgent to determine effective molecules to better perform PCa management.

Thymidine kinase 1 (TK1) is a cytosolic enzyme involved in pyrimidine metabolism that catalyzes the addition of a gamma-phosphate group to thymidine and in regenerating thymidine for DNA synthesis and DNA damage (Malvi et al., 2019; Bitter et al., 2020). Among the four deoxyribonucleoside-specific kinases in mammalian cells, TK1 is the only one with the most restricted substrates specificity (Eriksson et al., 2002). Its expression is S-phase dependent and elevated expression of TK1 has been noted in cell proliferation. Since Ki67 is present in all phases of the cell cycle and PCNA is mainly present in later G1, TK1 is more informative because it peaks in S phase expression, closely mimicking the rate of DNA synthesis (Bitter et al., 2020). Recently, it has been applied as an important biomarker for the diagnosis of various cancers, including breast cancer, esophageal cancer, and lung cancer (Li et al., 2005; He et al., 2010; Nisman et al., 2014; Jagarlamudi et al., 2015; Weagel et al., 2018; Malvi et al., 2019). TK1 upregulation was indicated as an early event in a study of breast cancer and further studies demonstrated a positive correlation between TK1 and cancer stage (He et al., 2010; Alegre et al., 2012). Subsequent studies support the potential of utilizing TK1 clinically to identify treatment effectiveness, cancer stage, and prognoses (Nisman et al., 2014; McCartney et al., 2019). Nisman et al. demonstrated that increased serum TK1 levels after chemotherapy for NSCLS indicate treatment failure and poor overall survival (Nisman et al., 2014). As for PCa, a few studies reported that TK1 can be used as a diagnostic biomarker through bioinformatic analysis and serological TK1 may be a potential proliferating biomarker for early detection (Li et al., 2018; Wang et al., 2018; Jagarlamudi et al., 2019; Song et al., 2019; Wang et al., 2020). Wang et al. identified TK1 as a core gene directly related to the recurrence and prognosis of PCa via bioinformatics analysis in multiple databases (Wang et al., 2020). Jagarlamudi et al. found that serum TK1 protein was significantly higher in patients with PCa than in patients with benign urological conditions and that TK1 protein determinations together with PHI or PSAD could be a valuable tool in PCa management (Jagarlamudi et al., 2019). In addition, Li et al. found that serum TK1 levels correlated with Gleason scores of prostate cancer patients whereas PSA levels did not (Li et al., 2018). However, the specific function of TK1 in PCa and the underlying mechanism are still lacking experimental verification.

In the present research, we first systematically investigated the function of TK1 in PCa *via in vivo* and *in vitro* experiments. Cox regression model analysis revealed that the expression of TK1 is significantly correlated with the pathology of PCa and associated with poor survival. Our study revealed that TK1 may be applied as a potential biomarker for PCa.

MATERIALS AND METHODS

Bioinformatic Analysis

The mRNA expression profiles and clinical data were obtained from the cancer genome atlas (TCGA), Gene Expression Omnibus (GEO), Prostate Cancer Transcriptome Atlas (PCTA), and PRAD-TCGA datasets (Rotinen et al., 2018; Cancer Genome Atlas Research, 2015). The PCTA dataset included 1321 clinical specimens. The PRAD dataset refers to the Prostate Adenocarcinoma (TCGA, TCGA Provisional) dataset and contains 497 PCa samples with fully collected data. GEPIA2 (http://gepia2.cancer-pku.cn/) was used to analyze data from the TCGA dataset (Tang et al., 2019). Most gene expression and clinical data were downloaded from cBioPortal (http://cbioportal.org). Also, two PCa microarray datasets were obtained from NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/) (Edgar et al., 2002): GSE70769 (Ross-Adams et al., 2015) and GSE21032 (Taylor et al., 2010). The status of neoadjuvant therapy was not considered as a criterion when selecting samples for analysis. For the PCa specimen shown in the figures, TK1 antibody (Atlas Antibodies, Cat# CAB004683) and AURKB antibody (Atlas Antibodies, Cat#CAB005862) were applied. Immunohistochemical staining of PCa specimens represented moderate cytoplasmic and nuclear positivity in the Human Protein Atlas database[1] (Uhlen et al., 2010; Uhlén et al., 2015).

Since co-expressed genes may act synergistically with TK1 to play a similar biological function in PCa, we screened the co-expressed genes via Spearman correlation analysis in the PRAD dataset from the cBioPortal[2] (Cerami et al., 2012). Then Metascape (https://metascape.org)[3] (Zhou et al., 2019) was applied to conduct further gene enrichment analysis using positively co-expressed genes ($r \geq 0.7$, $p < 0.01$, $q < 0.01$) and TK1. The protein-protein interaction (PPI) enrichment analyses were explored via The Molecular Complex Detection (MCODE) algorithm.

To investigate the correlation between TK1 expression and gene-level copy number variation, the PRAD dataset from TCGA was obtained from cBioPortal online dataset. TIMER was used to analyze the association between TK1 and tumor

[1]https://www.proteinatlas.org/
[2]http://cbioportal.org
[3]https://metascape.org

immune infiltration, immune subtype of PCa[4] (Li et al., 2017). TISIDB was used to investigate the expression of TK1 in PCa patients with different immune subtypes, as well as the correlation between tumor immune infiltration and TK1[5] (Ru et al., 2019).

Cell Culture and Transfection

7PCa cells applied in all experiments including BPH-1, LNCaP, C4-2, 22RV1, and DU145 were all derived from ATCC and cultured in RPMI 1640 (Gibco, United States) with 10% fetal bovine serum (FBS, Gibco, United States) in 5% CO_2 at 37°C. TK1 shRNA was used to target TK1 mRNA region (GCACAGAGU UGAUGAGACG) following the manufacturer's instructions.

RNA Isolation, Reverse Transcription, and Quantitative RT-PCR

TRIzol reagent (Sigma, United States) was applied to conduct RNA extraction. RNA reverse transcription was conducted following the protocol by using RevertAid First Strand cDNA Synthesis Kit (ThermoFisher, United States). Quantitative RT-PCR was performed using Fast SYBR Green Master Mix (Roche, United States) on LightCycler 480 System (Roche). Gene expression levels were identified via the Ct method and further normalized to GAPDH levels. The primer sequences are listed in **Supplementary Table S1**.

Western Blot

RIPA buffer was applied to extract total cellular protein. The concentration of the protein was quantified by BCA analysis. Then sodium dodecyl sulfate-polyacrylamide sodium gel electrophoresis (SDS-PAGE) and PVDF membrane (Millipore, Bedford, MA) were used to separate the protein. The PVDF membrane was blocked with 5% skim milk for 1 h and incubated overnight with anti-GAPDH (1:2000, ab8245, Abcam) and anti-TK1 (1:1000, ab76495, Abcam) antibody at 4°C. The next day, the membrane was washed and incubated with HRP-conjugated goat anti-rabbit IgG antibody at room temperature for 1 h. Visualization and photography were performed using immobilon western chemilum hrp substrate (WBKLS0100, Millipore).

Cell Growth Assay

Cell Counting Kit-8 (CCK8, Dojindo, Japan) was applied to analyze cell viability following the corresponding protocols.

Transwell

For migration assessment, standard transwell chambers (Corning, United States) were used. There were 1.5×10^4 cells with RPMI 1640 medium added to the upper chamber and the lower chamber and was supplemented with 10% fetal bovine serum medium of a 24-well plate. After incubating for 2 d in 37°C, cells were washed with cool PBS twice, fixed with methanol for 30 min at room temperature, stained with 0.2% crystal violet for 20 min, and observed under microscope. Each experiment was conducted in triplicate and repeated three times.

Flow Cytometry

The effects of TK1 ablation on PCa cell cycle were explored via flow cytometry (FC5000, BD, United States). There was 1 ug/ml propidium iodide (BD Biosciences, Germany) used to stain cancer cells.

Colony Formation Assay

First, about 1000 cells were seeded in a 6-well plate. After 10 d incubation, the cells were washed with cold PBS, fixed with 4% paraformaldehyde, and stained with 1% crystal violet solution for 10 min. Then the colonies were counted under an optional microscope.

Tumor Xenograft

All animal experiments were approved by the ethics committee of the Second Hospital of Tianjin Medical University. Eight-week-old nude mice were obtained from Beijing HFK Bioscience Co. Ltd. (Beijing, China). In brief, a total of 10 mice were randomly allocated to 2 groups, and 2×10^6 TK1 knockdown and control cells were suspended in 0.1 ml PBS and injected subcutaneously into the right groin of nude mice. Then the speed of tumor growth was measured every other day.

Statistics

All statistical analysis was conducted via R-4.0.0 and SPSS 22.0. The following R package were used: edgeR, WGCNA, survival, and ggplot2. Independent Student t-test and ANOVA were both applied for comparison. The Cox regression was used to explore the prognostic value of TK1 expression for OS, as well as DFS. Survival analysis was calculated and carried out by Kaplan-Meier method, and log-rank test was used to determine the distinctions. The data was demonstrated as mean ± standard deviation (SD). A p value < 0.05 was considered statistically significant.

RESULTS

Elevated Expression of TK1 in Human Prostate Cancer and Cancer Cells

Previously studies have reported that TK1 took a key role in tumor initiation and progression. We first explored the expression pattern of TK1 in certain tumors using the TCGA dataset. We found that TK1 was upregulated in most human cancers, including PCa (**Figures 1A,B**, $p < 0.05$). We also identified elevated expression of TK1 in the Chinese cohort population (**Figure 1C**) (Ren et al., 2018). Moreover, the expression of TK1 was also elevated in mCRPC patients comparing with primary PCa (**Figure 1D**). To further assess the expression pattern of TK1 expression in PCa, the correlation between tumor Gleason score and TK1 expression was also explored. As depicted in **Figures 1E,F**, the expression of TK1 increased with the increase of tumor Gleason score ($p < 0.001$). Next, we explored the TK1 protein expression via The Human

[4]https://cistrome.shinyapps.io/timer/
[5]http://cis.hku.hk/TISIDB/index.php

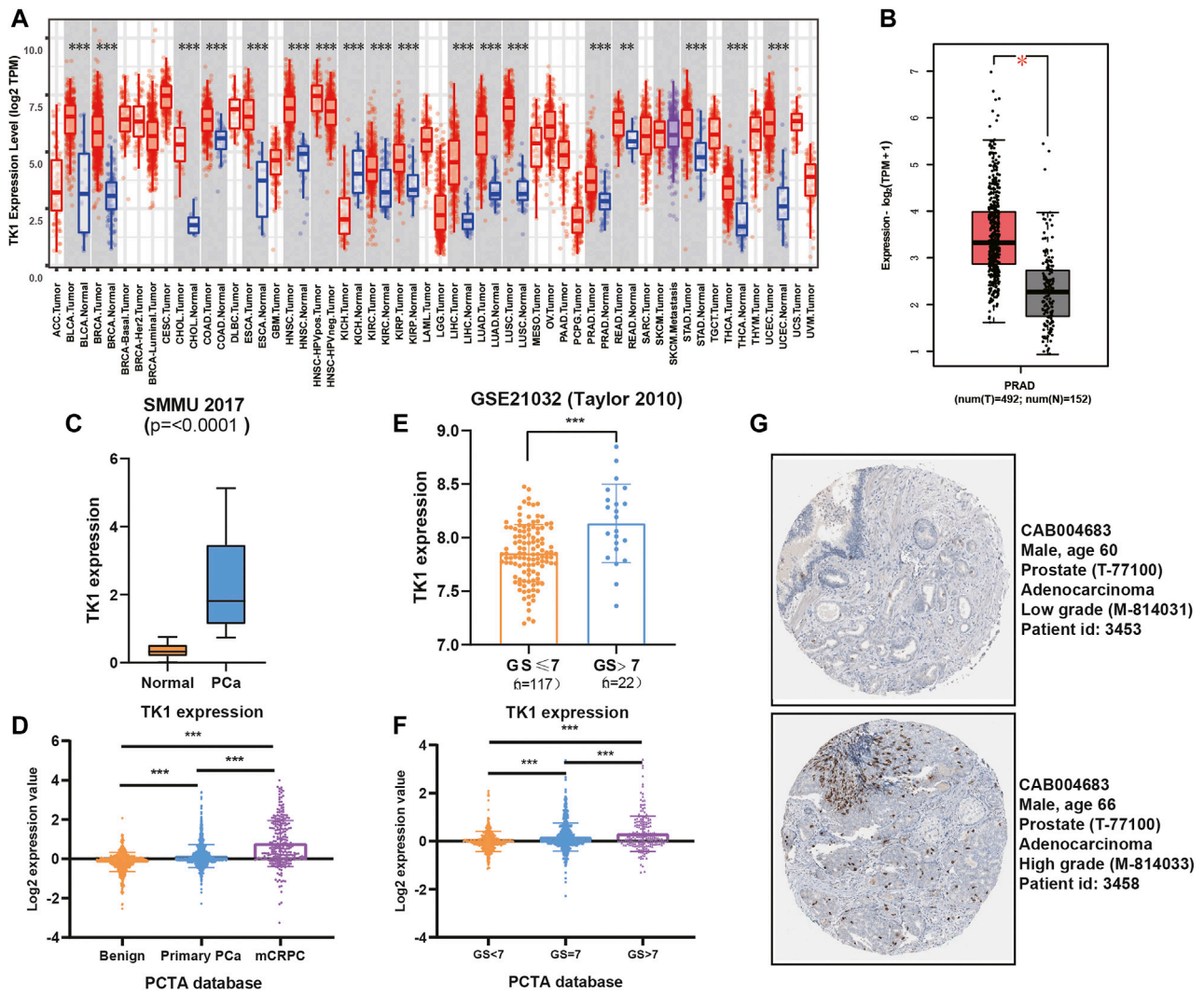


FIGURE 1 | TK1 expression in PCa patients. **(A)** TK1 expression in various cancer tissues and normal tissues. **(B)** TK1 expression in TCGA PRAD cohort. **(C,E)** TK1 expression across several independent clinical studies. **(D,F)** TK1 mRNA expression in the PCTA dataset. **(G)** TK1 protein expression showed by immunohistochemical staining in high-grade and low-grade patient. The pictures were taken from the Human Protein Atlas dataset. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; GS, Gleason score; mCRPC, metastatic castration resistant prostate cancer.

Protein Atlas. As showed in **Figure 1G**, a high-grade PCa patient (ID:3458) showed significantly higher intensity level of TK1 protein expression relative to a low-grade PCa patient (ID:3453).

To validate the findings in the above datasets, we evaluated the expression of TK1 among multiple human prostate cancer cell lines using the CCLE dataset and quantitative RT-PCR analysis. The data from the CCLE dataset exhibited a certain amount of TK1 expression in PCa cells, and PCR verification demonstrated that its expression was dramatically elevated compared with BPH1 (**Figures 2A,B**).

TK1 Inactivation in Prostate Cancer Cells Inhibits Tumor Malignant Behavior

To explore the specific function of TK1 in PCa cells, shRNA-mediated assay was applied to ablate TK1 function. We used shRNA-containing lentiviruses to target TK1 (PC-3 and C4-2) and the knockdown efficacy was verified via qPCR and Western blot (**Figures 2C,D**). Then all cell lines were tested for their tumor malignant behavior including proliferation, migration, and invasion. As **Figures 2E,F** show, CCK8 assays were performed to determine the cell proliferation viability, and cells from the shTK1 group grew significantly slower than the control group. Moreover, TK1 ablation also significantly inhibited colony formation and brought about a dramatic reduction in the rate of colony formation (**Figure 2G**). In addition, transwell assay further revealed the potential stimulative role of TK1 on tumor cell mobility in C42 and PC-3 cells. As depicted in **Figure 2H**, cells that knocked down TK1 failed to cross over the chambers because of their impaired migration capability. Furthermore, xenograft model assay suggested that knockdown TK1 in PC-3 cells significantly inhibited tumor growth compared with

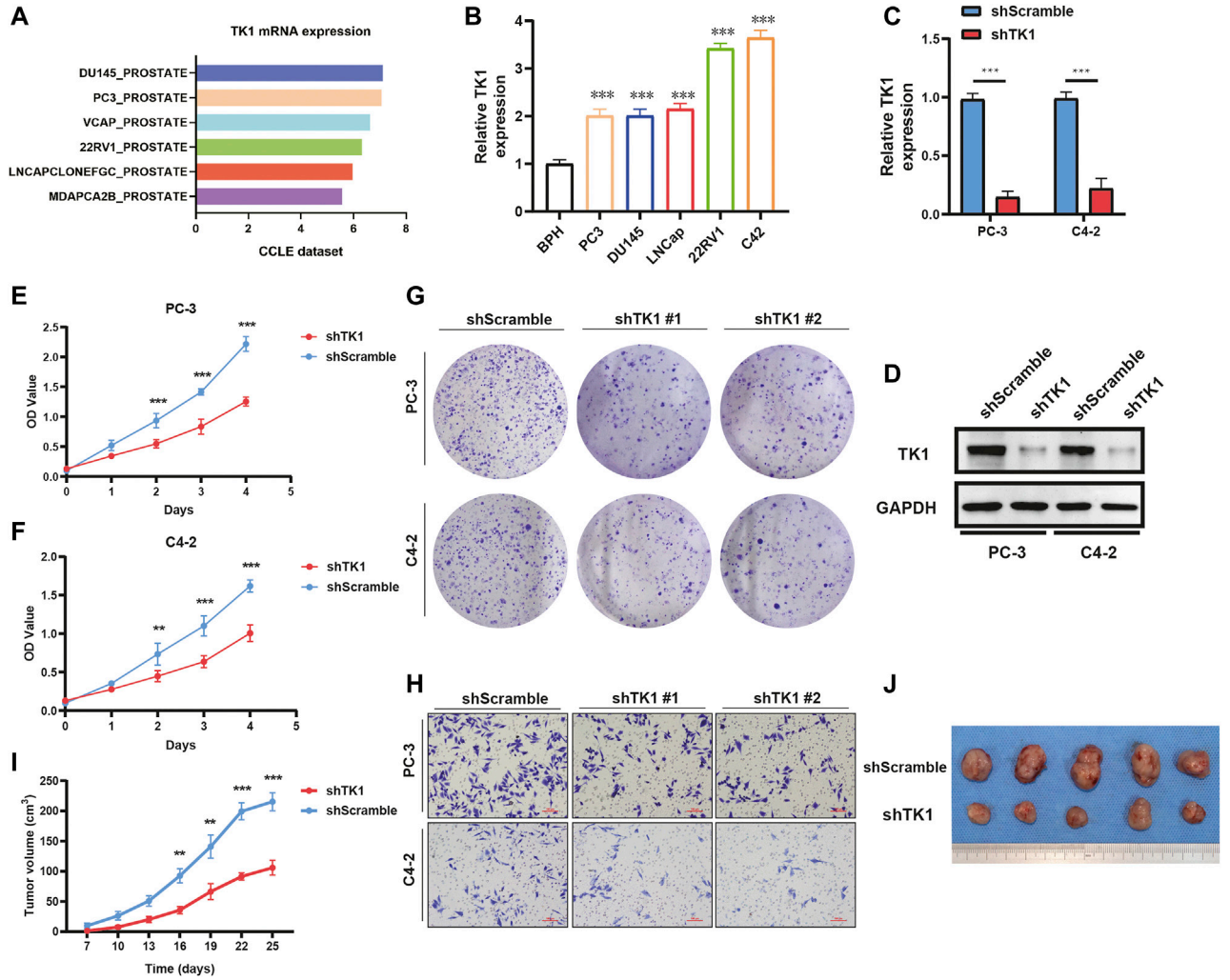


FIGURE 2 | TK1 ablation inhibits tumor cell growth both *in vitro* and *in vivo*. **(A)** TK1 mRNA expression of prostate cancer cell lines in the CCLE dataset. **(B)** TK1 mRNA expression in prostate cancer lines validated by qPCR. **(C,D)** TK1 knockdown efficacy validated by qPCR **(C)** and Western blot **(D)**. **(E,F)** The cell proliferation capacity in shTK1 cells is significantly suppressed compared to control cells. Both PC-3 and C4-2 cell lines were applied. **(G)** TK1 silencing dramatically inhibits the colony formation of prostate cancer cells. **(H)** The migration ability in shTK1 cells is significantly inhibited compared to control cells. **(I)** Tumor growth curves of the TK1-silenced and control groups. **(J)** The photograph of tumors implanted with TK1-silenced PC-3 cells and control tumors from nude mice. $^{**}p < 0.01$, $^{***}p < 0.001$.

scramble cells (**Figures 2I,J**). All results indicated that TK1 is closely involved in the malignant behavior of PCa cells.

Enrichment Analysis and PPI

To explore the potential biological significance and underlying mechanism of TK1 in PCa, gene co-expression analysis was performed via cBioPortal dataset. Biological functions and related signaling pathways were determined using the top 50 co-expressed genes (r > 0.75, *p*-value <0.05, **Table 1**). As demonstrated in **Figure 3A** and **Table 2**, pathway enrichment analysis revealed the 18 most statistically significant clusters (*p*-value <0.05 and enrichment factor >1.5). Cell cycle, cell division, and chromosome segregation were the top 3 clusters with the most enrichment. Meanwhile, the top-level Gene Ontology biological processes were also demonstrated (**Figure 3B**). To prove the results of enrichment analysis and further explore the function of TK1 in PCa, cell cycle distributions were identified via flow cytometry. The results indicated that TK1 ablation in prostate cancer cells leads to cell arrest in G2/M phase compared to control cells (**Figures 3C,D**).

PPI enrichment analyses were also carried out with the MCODE algorithm to determine densely connected network components. As depicted in **Figures 4A,B**, the MCODE results were gathered and demonstrated. Fourteen hub genes (AURKB, CCNB2, CDCA5, CDK1, CENPA, CENPM, KIF2C, NDC80, CDC20, NUF2, PLK1, SKA1, SPC25, ZWINT) constituted the MCODE-1 component. The expression

TABLE 1 | Gene positively correlated with TK1 mRNA expression in the PRAD dataset (Top 50 ranked by Spearman's correlation coefficient).

Correlated gene	Spearman's correlation	*p*-Value	q-Value
MCM2	0.857687	3.20E-109	6.42E-105
GINS1	0.837886	1.40E-99	1.41E-95
CDCA5	0.836338	6.99E-99	4.68E-95
KIF2C	0.833796	9.46E-98	4.75E-94
TROAP	0.82942	7.57E-96	3.04E-92
CDC20	0.829119	1.02E-95	3.41E-92
CDC45	0.828222	2.46E-95	7.06E-92
RAD54L	0.826461	1.37E-94	3.43E-91
CHAF1B	0.823953	1.52E-93	3.40E-90
SPC25	0.823389	2.60E-93	5.23E-90
CDT1	0.82312	3.36E-93	6.14E-90
ZWINT	0.822677	5.11E-93	8.56E-90
KIFC1	0.822012	9.58E-93	1.48E-89
NCAPG	0.821661	1.33E-92	1.91E-89
FANCG	0.820582	3.66E-92	4.90E-89
OIP5	0.818721	2.06E-91	2.58E-88
RAD51	0.81848	2.57E-91	3.04E-88
FEN1	0.818403	2.76E-91	3.08E-88
EXO1	0.816745	1.26E-90	1.29E-87
KIF4A	0.816725	1.28E-90	1.29E-87
CDC6	0.816016	2.44E-90	2.34E-87
KIF18B	0.815072	5.74E-90	5.24E-87
CCNB2	0.814834	7.11E-90	6.21E-87
NDC80	0.814465	9.92E-90	8.30E-87
CENPM	0.813052	3.51E-89	2.82E-86
TPX2	0.8123	6.86E-89	5.30E-86
HJURP	0.81152	1.37E-88	1.02E-85
MYBL2	0.810855	2.46E-88	1.76E-85
E2F1	0.810435	3.55E-88	2.46E-85
SKA1	0.810176	4.46E-88	2.92E-85
FANCI	0.810163	4.51E-88	2.92E-85
NUF2	0.808864	1.40E-87	8.79E-85
CENPA	0.807952	3.09E-87	1.88E-84
SKA3	0.807729	3.74E-87	2.21E-84
CDCA3	0.807196	5.93E-87	3.40E-84
FANCD2	0.807027	6.86E-87	3.83E-84
DTL	0.805966	1.70E-86	9.24E-84
MCM10	0.805741	2.06E-86	1.09E-83
TEDC2	0.805711	2.12E-86	1.09E-83
CDK1	0.805485	2.57E-86	1.29E-83
CCNF	0.804815	4.54E-86	2.22E-83
MCM7	0.804701	5.00E-86	2.39E-83
ORC1	0.804614	5.38E-86	2.51E-83
ASF1B	0.802442	3.35E-85	1.53E-82
FAM72B	0.801841	5.54E-85	2.47E-82
PLK1	0.801381	8.13E-85	3.55E-82
PTTG1	0.799776	3.07E-84	1.31E-81
AURKB	0.79901	5.77E-84	2.42E-81
CDC25C	0.798915	6.24E-84	2.56E-81

relationship between TK1 and all the hub genes was also shown in **Supplementary Figure S1**, respectively. Moreover, we confirmed that the expression of some hub genes was suppressed in shTK1 cells via RT-PCR (**Figure 4C**), and the protein co-expression of TK1 and AURKB in a same patient sample in the Human Protein Atlas was depicted in **Figure 4D**. All the above results indicated that the function of TK1 in PCa was closely involved in cell cycle regulation, which was also in accordance with the phenotypic results characterized previously.

TK1 Is Correlated With Clinical Features of PCa and Elevated Expression of TK1 Represents a Prognostic Factor for PCa

Given the crucial capacity of TK1 in PCa, we examined the potential relationship between TK1 expression and clinical features, including multiple clinic-pathological characteristics and survival of PCa patients. Data from the TCGA dataset showed that patients with elder age (>60 years; $p = 0.003$), higher Gleason score (>7; $p < 0.005$), higher clinical stage (≥T3a; $p < 0.005$), higher pathological stage (≥T3a; $p < 0.001$), lymph node metastasis ($p < 0.005$), shorter OS ($p < 0.005$), and shorter DFS ($p < 0.005$) had higher levels of TK1 expression (**Table 3**). Moreover, the Kaplan-Meier curve method was conducted to determine the correlation between TK1 expression level and OS and DFS (**Figures 5A,B**). The quartile TK1 mRNA expression level was used as the cutoff point to divided patients into the low TK1 ($n = 246$, TCGA dataset) and high TK1 ($n = 246$, TCGA dataset) group and conducted statistically significant validation of survival analyses in both groups. As **Figures 3A,B** show, patients in the high TK1 class had a shorter probability of OS ($p = 0.017$) and DFS ($p < 0.001$) compared to the low TK1 group. Moreover, we also investigated the prognostic role of TK1 across several independent clinical data sets (Taylor et al., 2010; Ross-Adams et al., 2015). As depicted in **Figures 5C,D**, the time to biochemical relapse was significantly shorter in the group of PCa patients with higher TK1 expression.

To explore the prognostic significance of TK1 in PCa, the Cox regression method was applied. As demonstrated in **Table 4**, clinical stage ($p < 0.005$), Gleason score ($p < 0.001$), pathological stage ($p < 0.005$), lymph node stage ($p = 0.014$), and TK1 mRNA expression ($p < 0.001$) were suitable to be regarded as prognostic factors for DFS by univariate analysis. In addition, we also found that Gleason score ($p = 0.007$), clinical stage ($p = 0.01$), and TK1 mRNA expression ($p < 0.001$) could be taken as prognostic factors for OS. Furthermore, multi-variate analyses suggested that Gleason score was an independent factor predicting the shortened survival of DFS ($p < 0.001$) and OS ($p < 0.05$), and the clinical stage predicted shorter DFS ($p = 0.003$). Perhaps because of the finite number of deceased in the PRAD dataset, TK1 mRNA expression showed limited prognostic value for survival via multi-variate analysis.

Immune Analysis of TK1 in Prostate Cancer

Next, the correlation between tumor immune infiltration and TK1 expression was analyzed. The results demonstrated that TK1 expression was closely correlated to immune subtypes of PCa, and TK1 was dramatically downregulated in the C3 subtype of PCa (**Figure 6A**). We further explored the genetic variations of TK1 in 497 cases of PCa in PRAD datasets via cBioPortal. As depicted in **Figure 6B**, amplification, deletion, and mRNA high were the main genetic variation types in TK1 in all samples. The overall variation rates of TK1 were also represented. In addition, **Figure 6C** presented the correlation of TK1 mRNA expression and the copy number in PCa. Using TIMER, the correlation between the TK1 copy number and tumor-infiltrating lymphocytes (TILs) was investigated. As shown in **Figure 6D**, high amplification of TK1 significantly decreased the TILs in PCa ($p < 0.05$). TK1 expression of immune cells in normal tissues and prostate tumor was also

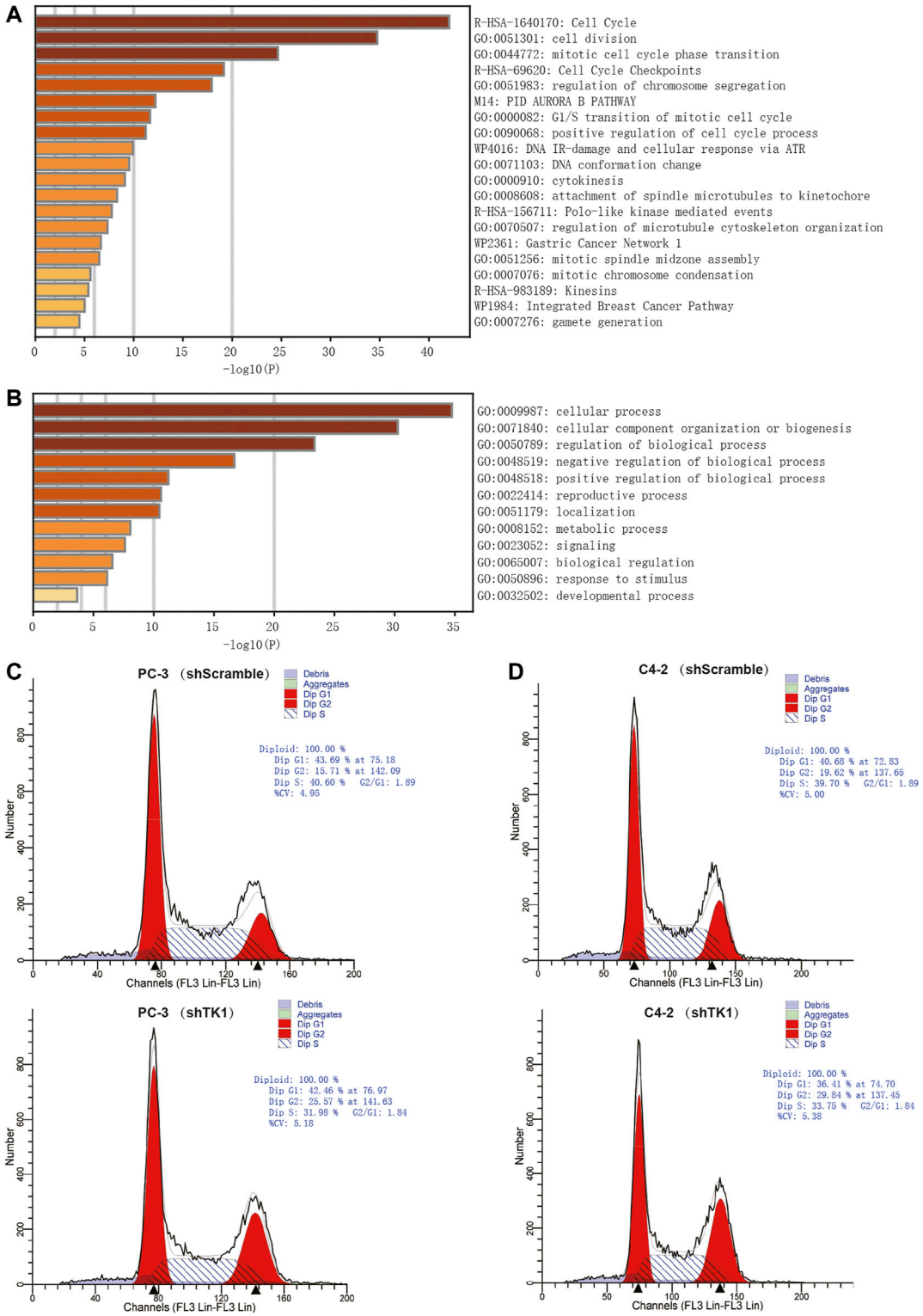


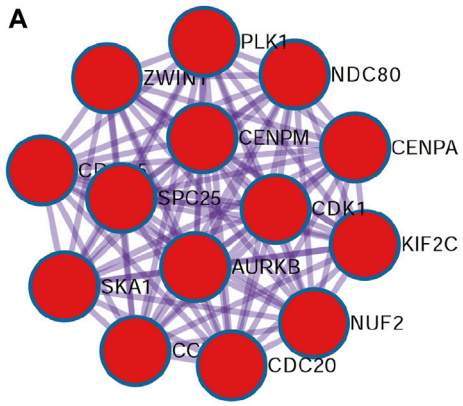
FIGURE 3 | Enrichment analysis and verification of the co-expressed genes. **(A,B)** Bar graph of enriched pathways **(A)** and top-level Gene-Ontology biological processes **(B)** cross the co-expressed genes. **(C,D)** TK1 silencing increases the percentage of cells in the G2/M phase. Cell cycle distributions were investigated by flow cytometry.

TABLE 2 | Top 18 clusters with their representative enriched terms by Metascape.

Category	Description	LogP	Log (q-value)	Symbols
Reactome Gene Sets	Cell cycle	41.0644	−36.706	CDK1, CDC6, CDC20, CDC25C, CENPA, E2F1, FEN1, MCM2, MCM7, MYBL2, ORC1, PLK1, RAD51, CDC45, CCNB2, EXO1, AURKB, PTTG1, GINS1, NDC80, KIF2C, ZWINT, OIP5, TPX2, HJURP, MCM10, SPC25, NCAPG, CENPM, CDT1, NUF2, CDCA5, SKA1, KIFC1
GO Biological Processes	Cell division	30.1137	−26.233	CCNF, CDK1, CDC6, CDC20, CDC25C, CENPA, KIFC1, PLK1, CCNB2, AURKB, PTTG1, NDC80, KIF2C, ZWINT, OIP5, TPX2, KIF4A, SPC25, NCAPG, CDT1, CDCA3, NUF2, CDCA5, KIF18B, SKA1, SKA3, MYBL2
GO Biological Processes	Chromosome segregation	28.9194	−25.164	CDC6, CDC20, FANCD2, FEN1, KIFC1, PLK1, AURKB, PTTG1, NDC80, KIF2C, ZWINT, OIP5, KIF4A, HJURP, SPC25, NCAPG, CDT1, NUF2, CDCA5, KIF18B, SKA1, SKA3, CDC25C, MYBL2, RAD51, RAD54L, CCNB2, TPX2, CCNF, CDK1, E2F1, MCM2, MCM7, ORC1, CDC45, DTL, MCM10, FANCI
WikiPathways	DNA IR-damage and cellular response via ATR	17.8918	−14.813	CDK1, CDC25C, E2F1, FANCD2, FEN1, MCM2, PLK1, RAD51, CDC45, EXO1, FANCI, CDC6, ORC1, DTL, CDT1, MCM7, CCNB2, CENPA, AURKB, MYBL2, TPX2, NCAPG, CCNF
WikiPathways	Cell cycle	17.7783	−14.721	CDK1, CDC6, CDC20, CDC25C, E2F1, MCM2, MCM7, ORC1, PLK1, CDC45, CCNB2, PTTG1, FEN1, RAD51, CHAF1B, EXO1, GINS1, DTL, MCM10, CDT1, FANCG, KIF4A, MYBL2, CDCA5, AURKB
GO Biological Processes	DNA repair	13.5907	−10.876	CDK1, FANCD2, FANCG, FEN1, MCM2, MCM7, RAD51, CHAF1B, CDC45, RAD54L, EXO1, PTTG1, DTL, FANCI, CDCA5, AURKB, E2F1, PLK1
GO Biological Processes	DNA conformation change	12.5687	−9.967	CENPA, MCM2, MCM7, RAD51, CHAF1B, RAD54L, OIP5, HJURP, ASF1B, NCAPG, CENPM, CDCA5, CDC45, CDT1, FEN1
GO Biological Processes	Meiotic cell cycle	10.7912	−8.368	CDC20, CDC25C, FANCD2, PLK1, RAD51, RAD54L, CCNB2, EXO1, PTTG1, NUF2
GO Biological Processes	Positive regulation of cell cycle process	−9.9885	−7.656	CDK1, CDC6, CDC25C, E2F1, FEN1, AURKB, NDC80, DTL, CDT1, CDCA5, ORC1, PLK1, CCNB2, TPX2, KIF18B, CDC20
Canonical Pathways	PID PLK1 pathway	9.74329	−7.423	CDK1, CDC20, CDC25C, PLK1, NDC80, TPX2, CENPA, AURKB, CDT1, CDCA5, CDC6, KIF4A, CCNF, E2F1, KIF2C
Reactome Gene Sets	DNA strand elongation	8.59771	−6.337	FEN1, MCM2, MCM7, CDC45, GINS1, RAD51, FANCD2, RAD54L, CDCA5, EXO1, MCM10
Reactome Gene Sets	Transcriptional regulation by TP53	6.67593	−4.538	CDK1, CDC25C, E2F1, FANCD2, EXO1, AURKB, TPX2, FANCI, CENPA, KIF2C
GO Biological Processes	Positive regulation of chromosome segregation	6.38871	−4.266	CDC6, FEN1, AURKB, CDT1, E2F1, PLK1, HJURP, CDCA5, RAD51
GO Biological Processes	Microtubule polymerization or depolymerization	−5.7257	−3.647	KIF2C, TPX2, KIF18B, SKA1, SKA3, KIFC1, KIF4A, CDK1, PLK1, CCNB2
GO Biological Processes	Gamete generation	5.55847	−3.493	CDC25C, E2F1, FANCD2, FANCG, KIFC1, PLK1, CCNB2, PTTG1, ASF1B
GO Biological Processes	Regulation of microtubule cytoskeleton organization	4.73284	−2.715	CCNF, PLK1, TPX2, SKA1, SKA3
KEGG Pathway	HTLV-I infection	4.08352	−2.118	CDC20, E2F1, MYBL2, CCNB2, PTTG1
GO Biological Processes	Telomere maintenance	3.74684	−1.825	FEN1, RAD51, EXO1, AURKB, DTL

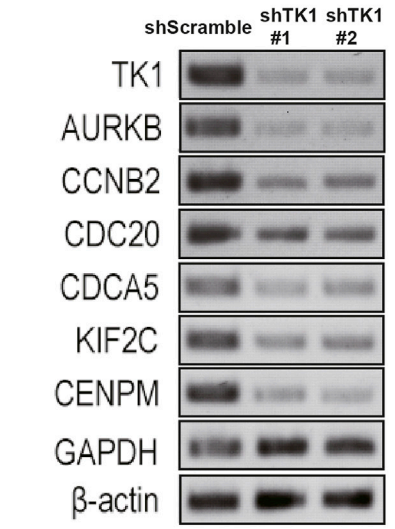
shown in **Figure 6E**. The expression of TK1 in prostate tumors was significantly elevated in Treg cells and decreased in B cells and activated dendritic cells compared with normal tissues (all $p < 0.05$).

We further determined the correlation between TK1 expression, immunomodulators (immunostimulators and immunoinhinitors), and TILs via TISIDB. **Figures 6F–H** respectively showed the top three TILs and immunomodulators with a Spearman's correlation coefficient greater than 0.2 with TK1 expression. Activated $CD4^+$ (r = 0.307, p = 3.4e-12) and $CD8^+$ (r = 0.208, p = 2.97e-06) T cells depicted the densest association with TK1 (**Figure 6F**). As depicted in **Figures 6G,H**, the greatest related immunostimulators with TK1 expression in PCa were interleukin 6 receptor (IL6R, r = −0.414, $p <$ 2.2e-16), 5′-nucleotidase ecto (NT5E, r = −0.372, $p <$ 2.2e-16), and TNF receptor superfamily member 18 (TNFRSF18, r = 0.348, p = 1.75e-15) and the most relevant immunoinhibitors correlated with TK1 expression in PCa were CD274 (r = −0.279, p = 2.93e-10),



B

Color	MCODE	GO	Description	Log10(P)
	MCODE_1	R-HSA-2500257	Resolution of Sister Chromatid Cohesion	-33.2
	MCODE_1	R-HSA-68877	Mitotic Prometaphase	-30.2
	MCODE_1	R-HSA-68882	Mitotic Anaphase	-29.2



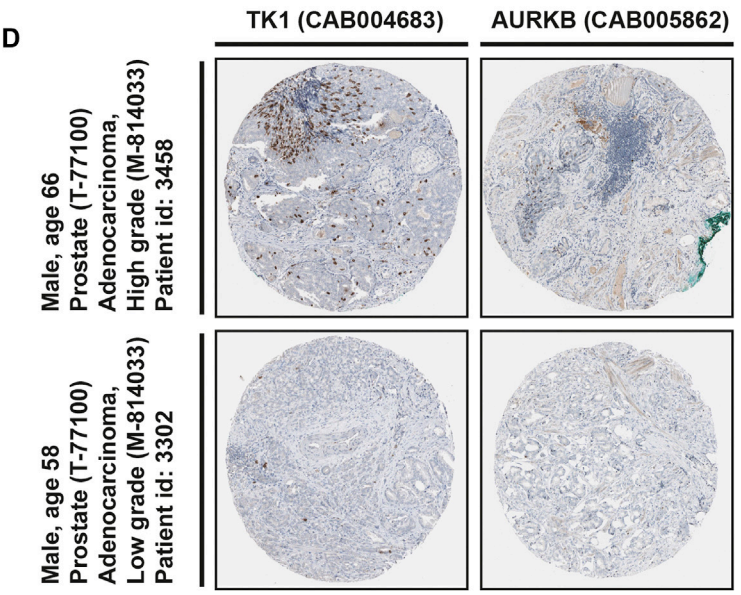


FIGURE 4 | Protein networks and the correlation between TK1 and the hub genes. **(A,B)** Molecular Complex Detection (MCODE) components of the hub genes. **(C)** The expression of serval hub genes was down-regulated in the TK1-silencing cells verified by RT-PCR. **(D)** TK1 and AURKB protein expression showed by immunohistochemical staining in the same high-grade and low-grade patient. The pictures were taken from the Human Protein Atlas dataset.

kinase insert domain receptor (KDR, r = −0.278, p = 3.49e-10), and adenosine a2a receptor (ADORA2A, r = 0.222, p = 5.89e-07).

DISCUSSION

In this research, we expanded the capacity of TK1, and explored the specific function of TK1 in PCa, as well as its underlying mechanism for the first time. Moreover, we also found that the functions of TK1 were strongly associated with related signaling pathways, including cell cycle, cell division, and mitotic cell cycle phase transition, thereby promoting tumor malignant behavior.

TK1 is a cytosolic enzyme involved in salvage pathway and plays a vital role in pyrimidine deoxynucleotide synthesis during the cell cycle. Thymidine is transferred from the extracellular space to the cell membrane by facilitated diffusion and is converted to the monophosphate form (dTMP) by TK1 at the cell membrane (Bello, 1974; Johnson et al., 1982). In addition to DNA synthesis, TK1 is also essential for cell repair following DNA damage due to its vital role in nucleotides formation beyond the S phase (Chen et al., 2010; Jagarlamudi and Shaw, 2018). The expression level of TK1 increases significantly after cellular damage caused by radiation or chemotherapeutic agents, and depletion of TK1 in cells exposed to DNA damage can lead to cell death (Chen et al., 2010; Fischer et al., 2016; Jagarlamudi and Shaw, 2018). Multiple studies have reported that regulation of cell cycle factors, including TK1, is critical for cell homeostasis and that mutations or dysregulation of cell cycle proteins is a major cause of tumorigenesis (Collins et al., 1997; Levine and Holland, 2018; Wenzel and Singh, 2018). Moreover, TK1 has been identified as a malignant biomarker in multiple malignancies due to its close correlation to cell proliferation, including lung, breast, and colorectal (Li et al., 2005; He et al., 2010; Nisman et al., 2014; Jagarlamudi et al., 2015; Weagel et al., 2018; McCartney et al., 2020). As for PCa, only several studies speculated that TK1 can be used as a diagnostic biomarker via bioinformatics analysis (Song et al., 2019; Wang et al., 2020). Song et al. integrated 10

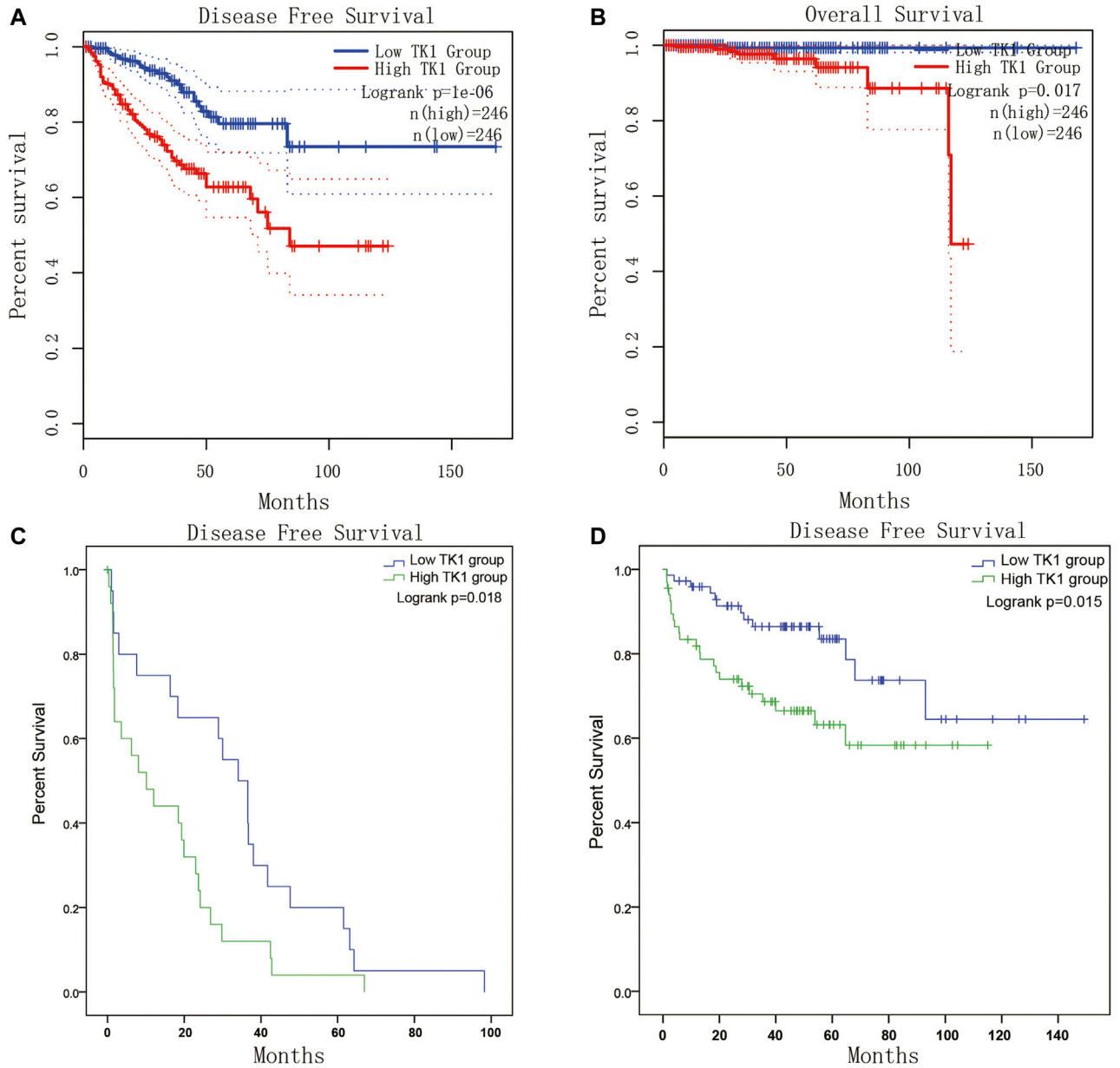


FIGURE 5 | Survival analysis of TK1 expression in PCa. **(A,B)** The TK1 mRNA expression level represented a prognostic value in OS **(A)** and in DFS **(B)** in the PRAD dataset. **(C,D)** Kaplan-Meier plots of the risk of biochemical recurrence in PCa patients with high or low expression of TK1 in several cohorts of human prostate tumors.

eligible PCa microarray datasets via the Robust Rank Aggregation method and identified four candidate biomarkers, including TK1, for the diagnosis and prognosis of PCa (Song et al., 2019). Similarly, by informatic analysis of four PCa microarray datasets, Tian et al. identified six core genes including TK1 directly involved in the recurrence and prognosis of PCa (Wang et al., 2020). Although the above research has noted that TK1 is involved in PCa progression, experimental verification and potential mechanisms are still limited. In this study, the expression profile of TK1 was examined and results suggested that TK1 was up-regulated in PCa patients and cell lines, especially those with higher Gleason scores (> 7). We also identified the role of TK1 in PCa proliferation and migration via a series of experiments. In addition, 14 hub genes were identified via enrichment analysis and PPI network analysis, and their functions indicated that TK1 was closely involved in cell cycle-related signaling pathways, which was in accordance with the phenotypic results characterized previously. Moreover, we conducted the Kaplan-Meier survival analysis and Cox regression model and found that elevated TK1 expression was dramatically correlated with worse clinical survival.

TABLE 3 | The correlation between clinicopathological characteristics and TK1 expression in the PRAD dataset.

Characteristics	N	TK1 expression (mean ± SD)	P
Age			
≤60y	224	296.5 ± 306.8	0.003
>60y	275	384.6 ± 345.2	
Clinical stage			
<T3a	352	328.5 ± 296.3	<0.001
≥T3a	55	513.4 ± 557.4	
Pathological stage			
<T3a	188	253.4 ± 189.5	<0.001
≥T3a	304	394.0 ± 340.0	
Gleason score			
≤7	293	263.7 ± 187.7	<0.001
>7	206	460.8 ± 439.6	
Lymph node stage			
N0	346	324.6 ± 282.4	<0.001
N1	80	462.9 ± 379.4	
Overall survival			
Alive	489	337.4 ± 298.7	<0.001
Decease	10	717.8 ± 1034.2	
Disease-free survival			
Disease-free	401	317.2 ± 286.9	0.001
Recurred/progressed	92	435.9 ± 333.6	

TABLE 4 | Prognostic value of TK1 mRNA expression level for the disease-free survival (DFS) and overall survival (OS) via Cox proportional model.

	DFS		OS	
	Hazard ratio (95% CI)	P	Hazard ratio (95% CI)	P
Univariate analysis				
Age	1.027 (0.996–1.060)	0.09	1.053 (0.955–1.160)	**0.032**
TK1 mRNA	1.001 (1.000–1.001)	**<0.001**	1.002 (1.001–1.003)	**<0.001**
Clinical stage	1.437 (1.263–1.635)	**<0.001**	1.666 (1.130–2.459)	**0.01**
Pathological stage	1.801 (1.437–2.259)	**<0.001**	1.630 (0.766–3.467)	0.205
Gleason score	2.227 (1.794–2.764)	**<0.001**	2.981 (1.346–6.601)	**0.007**
Lymph node stage	1.831 (1.130–2.969)	**0.014**	3.523 (0.778–15.942)	0.102
Multivariate analysis				
Age	0.997 (0.962–1.034)	0.885	1.041 (0.931–1.163)	0.480
TK1 mRNA	1.000 (1.000–1.001)	0.373	0.999 (0.997–1.002)	0.673
Clinical stage	1.255 (1.079–1.459)	**0.003**	1.278 (0.761–2.145)	0.353
Pathological stage	1.117 (0.794–1.572)	0.525	0.772 (0.255–2.335)	0.647
Gleason score	1.801 (1.310–2.474)	**<0.001**	3.489 (1.035–11.758)	**0.044**
Lymph node stage	0.994 (0.558–1.769)	0.983	2.537 (0.447–14.391)	0.293

*The bold values in **Table 4** represent values less than 0.05 and are statistically significant.*

The management of PCa still imposes an urgent challenge on society. Prostate specific antigen (PSA) screening has been performed for PCa diagnosis and relapse monitoring. But this could also lead to a series of problems such as overdiagnosis and overtreatment due to lack of specificity and poor indication of aggressiveness (Diamandis, 1998; Hayes and Barry, 2014). Therefore, new prognostic factor identification for biomedical recurrence and overall survival of PCa patients is crucial and urgent. Recently, many efforts have been made to find better biomarkers for PCa. Prostate-specific membrane antigen (PSMA), a type II transmembrane protein, has been found to be significantly overexpressed on prostatic cancer cells, including advanced-stage prostate carcinomas, but a low expression in normal tissues. It can be considered as ideal for developing small and low-molecular-weight targeted radiopharmaceuticals for

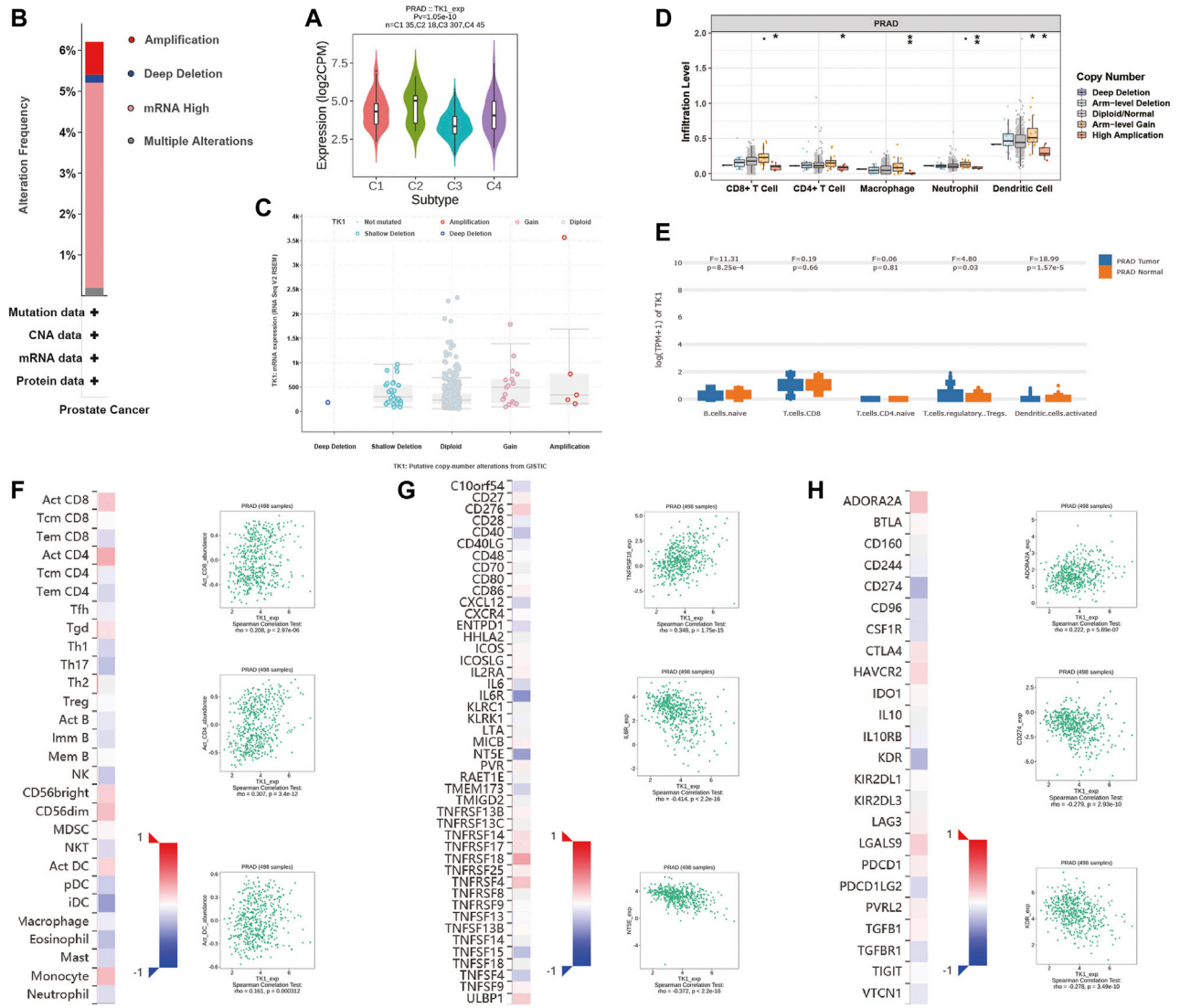


FIGURE 6 | Immune analysis of TK1 in PCa. **(A)** Relationships between TK1 expression and immune subtype in TCGA prostate cancer dataset. **(B)** The mutation types and mutation frequencies of TK1 in PCa. **(C)** Correlation between mRNA expression of TK1 and the copy number in PCa. **(D)** Correlation between TK1 copy number and tumor-infiltrating lymphocytes (TILs). **(E)** TK1 expression of immune cells in the prostate tumor and normal tissues. **(F)** Correlation between TK1 expression and TILs (TISIDB). **(G,H)** Correlation between TK1 expression and immunostimulators **(G)** and immunoinhibitors **(H)**. In the heatmaps of **(F-H)**, the red and blue squares represent positive and negative correlations, respectively. The scatter plots show TILs or immunomodulators with the strongest correlation with TK1 expression.

diagnosis and treatment in imaging (Haberkorn et al., 2016). Therefore, it is more of a diagnostic and therapeutic target for imaging rather than a prognostic biomarker. Combined RankProd with genetic algorithm optimized artificial neural network (GA-ANN), Hou et al. identified a 15-gene signature that exhibited a great capacity for diagnosis and prognosis of PCa and found that C1QTNF3 was a good predictor for PCa diagnosis (Hou et al., 2018). However, the underlying mechanism lacks experimental validation, and more studies are warranted. Herein, we systemically demonstrated the function of TK1 in PCa and found that it can be applied as a prognostic biomarker. Similar to our results, much research has investigated the clinical value of serological TK1 in the diagnosis of PCa. Wang et al. determined the mean values and the concentration distribution of serological TK1 protein in a cohort of 56,178 persons consisting of people with different disease stages, and found that serological TK1 was a proliferating biomarker for early discovery of malignancy in the prostate (Wang et al., 2018). Jagarlamudi et al. demonstrated that there were inconsistencies in the particular activities as well as the subunit compositions of serological TK1 in different cancers. Meanwhile, serological

TK1 protein assays can distinguish early-phase cancer formation in prostate and breast cancer more usefully than serological TK1 activity assays (Jagarlamudi et al., 2015). Furthermore, by collecting and analyzing serum samples from 140 patients, they also demonstrated that TK1 protein determinations together with Prostate Health Index (PHI) or PSA density (PSAD) can be worthy additional tools for PCa treatment (Jagarlamudi et al., 2019). However, the present study did not further determine the function of serological TK1 protein in PCa.

Tumor immune response plays a vital role in cancer formation and development. Though increasing evidence has proved the non-negligible role of immune system in PCa management, few approved immunotherapy exists (Bilusic et al., 2017; Cha et al., 2020). Using the TCGA database, Vesteinn Thorsson et al. classified tumors into six immune subtypes (Thorsson et al., 2019). Dramatic dissimilarities in lymphocyte infiltration, prognosis, and immune regulation gene expression existed among distinct subtypes. The present studies indicated that TK1 expression was dramatically decreased in C3 subtype of PCa, which had the best prognosis. This suggested that TK1 can be applied for immunophenotyping and prognosis prediction. The data from TISIDB also revealed that TK1 was significantly related with TILs and immunomodulators. Since the accumulation of TILs and immunomodulators expression in PCa was associated with patient prognosis, TK1 may be involved in immune tolerance via interacting with TILs and immunomodulatory molecules, and can be used as a potential marker for prostate immunotherapy (Steele et al., 2018; Pérez-Ruiz et al., 2020; Yang et al., 2021).

In conclusion, our research systematically explored the capacity of TK1 in PCa for the first time. Elevated expression of TK1 in PCa patients can be applied as a valuable prognostic biomarker for predicting poor survival (both DFS and OS). TK1 ablation inhibits tumor malignant behavior and may serve as a therapeutic target for PCa.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The animal study was reviewed and approved by the Affidavit of Approval of Animal Ethical and Welfare of Tianjin Medical University.

AUTHOR CONTRIBUTIONS

HX and LG performed most of the experiments and interpretation of the data. ZW, SP, QM, and ZY contributed to the analysis of the data and conducted parts of the experiment. HX and LG wrote the manuscript. YN and ZS critically revised the manuscript and contributed to the conception and design. All authors read and approved the final article.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 91959114, 81872106, and 81802573).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.778850/full#supplementary-material

REFERENCES

Alegre, M. M., Robison, R. A., and O'Neill, K. L. (2012). Thymidine Kinase 1 Upregulation Is an Early Event in Breast Tumor Formation. *J. Oncol.* 2012, 1–5. Epub 2012/07/11PubMed PMID: 22778736; PubMed Central PMCID: PMCPMC3388419. doi:10.1155/2012/575647

Barbieri, C. E., Chinnaiyan, A. M., Lerner, S. P., Swanton, C., and Rubin, M. A. (2017). The Emergence of Precision Urologic Oncology: A Collaborative Review on Biomarker-Driven Therapeutics. *Eur. Urol.* 71 (2), 237–246. Epub 2016/08/28PubMed PMID: 27567210; PubMed Central PMCID: PMCPMC5195855. doi:10.1016/j.eururo.2016.08.024

Bello, L. (1974). Regulation of Thymidine Kinase Synthesis in Human Cells*1. *Exp. Cel Res.* 89 (2), 263–274. Epub 1974/12/01PubMed PMID: 4457349. doi:10.1016/0014-4827(74)90790-3

[1]https://www.proteinatlas.org/

[2]http://cbioportal.org

[3]https://metascape.org

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

[5]http://cis.hku.hk/TISIDB/index.php

Bilusic, M., Madan, R. A., and Gulley, J. L. (2017). Immunotherapy of Prostate Cancer: Facts and Hopes. *Clin. Cancer Res.* 23 (22), 6764–6770. Epub 2017/07/01PubMed PMID: 28663235; PubMed Central PMCID: PMCPMC5690854. doi:10.1158/1078-0432.CCR-17-0019

Bitter, E. E., Townsend, M. H., Erickson, R., Allen, C., and O'Neill, K. L. (2020). Thymidine Kinase 1 through the Ages: a Comprehensive Review. *Cell Biosci* 10 (1), 138. Epub 2020/12/10PubMed PMID: 33292474; PubMed Central PMCID: PMCPMC7694900. doi:10.1186/s13578-020-00493-1

Cancer Genome Atlas Research, N. (2015). The Molecular Taxonomy of Primary Prostate Cancer. *Cell* 163 (4), 1011–1025. Epub 2015/11/07PubMed PMID: 26544944; PubMed Central PMCID: PMCPMC4695400. doi:10.1016/j.cell.2015.10.025

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data: Figure 1. *Cancer Discov.* 2 (5), 401–404. Epub 2012/05/17PubMed PMID: 22588877; PubMed Central PMCID: PMCPMC3956037. doi:10.1158/2159-8290.CD-12-0095

Cha, H.-R., Lee, J. H., and Ponnazhagan, S. (2020). Revisiting Immunotherapy: A Focus on Prostate Cancer. *Cancer Res.* 80 (8), 1615–1623. Epub 2020/02/19PubMed PMID: 32066566; PubMed Central PMCID: PMCPMC7641094. doi:10.1158/0008-5472.CAN-19-2948

Chen, Y.-L., Eriksson, S., and Chang, Z.-F. (2010). Regulation and Functional Contribution of Thymidine Kinase 1 in Repair of DNA Damage. *J. Biol. Chem.*

285 (35), 27327–27335. Epub 2010/06/18PubMed PMID: 20554529; PubMed Central PMCID: PMCPMC2930731. doi:10.1074/jbc.M110.137042

Collins, K., Jacks, T., and Pavletich, N. P. (1997). The Cell Cycle and Cancer. *Proc. Natl. Acad. Sci.* 94 (7), 2776–2778. Epub 1997/04/01PubMed PMID: 9096291; PubMed Central PMCID: PMCPMC34145. doi:10.1073/pnas.94.7.2776

Diamandis, E. P. (1998). Prostate-specific Antigen: Its Usefulness in Clinical Medicine. *Trends Endocrinol. Metab.* 9 (8), 310–316. Epub 2008/04/15PubMed PMID: 18406295. doi:10.1016/s1043-2760(98)00082-4

Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene Expression Omnibus: NCBI Gene Expression and Hybridization Array Data Repository. *Nucleic Acids Res.* 30 (1), 207–210. Epub 2001/12/26PubMed PMID: 11752295; PubMed Central PMCID: PMCPMC99122. doi:10.1093/nar/30.1.207

Eriksson, S., Munch-Petersen, B., Johansson, K., and Ecklund, H. (2002). Structure and Function of Cellular Deoxyribonucleoside Kinases. *Cell Mol. Life Sci. (Cmls)* 59 (8), 1327–1346. Epub 2002/10/05PubMed PMID: 12363036. doi:10.1007/s00018-002-8511-x

Fischer, M., Quaas, M., Steiner, L., and Engeland, K. (2016). The P53-P21-DREAM-CDE/CHR Pathway Regulates G2/M Cell Cycle Genes. *Nucleic Acids Res.* 44 (1), 164–174. Epub 2015/09/20PubMed PMID: 26384566; PubMed Central PMCID: PMCPMC4705690. doi:10.1093/nar/gkv927

Grasso, C. S., Wu, Y.-M., Robinson, D. R., Cao, X., Dhanasekaran, S. M., Khan, A. P., et al. (2012). The Mutational Landscape of Lethal Castration-Resistant Prostate Cancer. *Nature* 487 (7406), 239–243. Epub 2012/06/23PubMed PMID: 22722839; PubMed Central PMCID: PMCPMC3396711. doi:10.1038/nature11125

Haberkorn, U., Eder, M., Kopka, K., Babich, J. W., and Eisenhut, M. (2016). New Strategies in Prostate Cancer: Prostate-specific Membrane Antigen (PSMA) Ligands for Diagnosis and Therapy. *Clin. Cancer Res.* 22 (1), 9–15. Epub 2016/01/06PubMed PMID: 26728408. doi:10.1158/1078-0432.CCR-15-0820

Hayes, J. H., and Barry, M. J. (2014). Screening for Prostate Cancer with the Prostate-specific Antigen Test. *JAMA* 311 (11), 1143–1149. Epub 2014/03/20PubMed PMID: 24643604. doi:10.1001/jama.2014.2085

He, E., Xu, X. H., Guan, H., Chen, Y., Chen, Z. H., Pan, Z. L., et al. (2010). Thymidine Kinase 1 Is a Potential Marker for Prognosis and Monitoring the Response to Treatment of Patients with Breast, Lung, and Esophageal Cancer and Non-hodgkin's Lymphoma. *Nucleosides, Nucleotides and Nucleic Acids* 29 (4-6), 352–358. Epub 2010/06/15PubMed PMID: 20544519. doi:10.1080/15257771003738535

Hou, Q., Bing, Z.-T., Hu, C., Li, M.-Y., Yang, K.-H., Mo, Z., et al. (2018). RankProd Combined with Genetic Algorithm Optimized Artificial Neural Network Establishes a Diagnostic and Prognostic Prediction Model that Revealed C1QTNF3 as a Biomarker for Prostate Cancer. *EBioMedicine* 32, 234–244. Epub 2018/06/05PubMed PMID: 29861410; PubMed Central PMCID: PMCPMC6021271. doi:10.1016/j.ebiom.2018.05.010

Jagarlamudi, K. K., Hansson, L. O., and Eriksson, S. (2015). Breast and Prostate Cancer Patients Differ Significantly in Their Serum Thymidine Kinase 1 (TK1) Specific Activities Compared with Those Hematological Malignancies and Blood Donors: Implications of Using Serum TK1 as a Biomarker. *BMC Cancer* 15, 66. Epub 2015/04/17PubMed PMID: 25881026; PubMed Central PMCID: PMCPMC4336758. doi:10.1186/s12885-015-1073-8

Jagarlamudi, K. K., and Shaw, M. (2018). Thymidine Kinase 1 as a Tumor Biomarker: Technical Advances Offer New Potential to an Old Biomarker. *Biomarkers Med.* 12 (9), 1035–1048. Epub 2018/07/25PubMed PMID: 30039979. doi:10.2217/bmm-2018-0157

Jagarlamudi, K. K., Zupan, M., Kumer, K., Fabjan, T., Hlebič, G., Eriksson, S., et al. (2019). The Combination of AroCell TK 210 ELISA with Prostate Health Index or Prostate-specific Antigen Density Can Improve the Ability to Differentiate Prostate Cancer from Noncancerous Conditions. *Prostate* 79 (8), 856–863. Epub 2019/03/20PubMed PMID: 30889628. doi:10.1002/pros.23791

Johnson, L., Rao, L. G., and Muench, A. J. (1982). Regulation of Thymidine Kinase Enzyme Level in Serum-Stimulated Mouse 3T6 Fibroblasts*1. *Exp. Cel Res.* 138 (1), 79–85. Epub 1982/03/01PubMed PMID: 7067741. doi:10.1016/0014-4827(82)90093-3

Latonen, L., Afyounian, E., Jylhä, A., Nättinen, J., Aapola, U., Annala, M., et al. (2018). Integrative Proteomics in Prostate Cancer Uncovers Robustness against Genomic and Transcriptomic Aberrations during Disease Progression. *Nat. Commun.* 9 (1), 1176. Epub 2018/03/23PubMed PMID: 29563510; PubMed Central PMCID: PMCPMC5862881. doi:10.1038/s41467-018-03573-6

Levine, M. S., and Holland, A. J. (2018). The Impact of Mitotic Errors on Cell Proliferation and Tumorigenesis. *Genes Dev.* 32 (9-10), 620–638. Epub 2018/05/29PubMed PMID: 29802124; PubMed Central PMCID: PMCPMC6004076. doi:10.1101/gad.314351.118

Li, H., Lei, D., Wang, X., Skog, S., and He, Q. (2005). Serum Thymidine Kinase 1 Is a Prognostic and Monitoring Factor in Patients with Non-small Cell Lung Cancer. *Oncol. Rep.* 13 (1), 145–149. Epub 2004/12/08. PubMed PMID: 15583816. doi:10.3892/or.13.1.145

Li, S., Zhou, J., Wang, Y., Zhang, K., Yang, J., Zhang, X., et al. (2018). Serum Thymidine Kinase 1 Is Associated with Gleason Score of Patients with Prostate Carcinoma. *Oncol. Lett.* 16 (5), 6171–6180. Epub 2018/10/20PubMed PMID: 30333882; PubMed Central PMCID: PMCPMC6176382. doi:10.3892/ol.2018.9345

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* 77 (21), e108–e110. Epub 2017/11/03PubMed PMID: 29092952; PubMed Central PMCID: PMCPMC6042652. doi:10.1158/0008-5472.CAN-17-0307

Malvi, P., Janostiak, R., Nagarajan, A., Cai, G., and Wajapeyee, N. (2019). Loss of Thymidine Kinase 1 Inhibits Lung Cancer Growth and Metastatic Attributes by Reducing GDF15 Expression. *Plos Genet.* 15 (10), e1008439. Epub 2019/10/08PubMed PMID: 31589613; PubMed Central PMCID: PMCPMC6797230. doi:10.1371/journal.pgen.1008439

McCartney, A., Biagioni, C., Schiavon, G., Bergqvist, M., Mattsson, K., Migliaccio, I., et al. (2019). Prognostic Role of Serum Thymidine Kinase 1 Activity in Patients with Hormone Receptor-Positive Metastatic Breast Cancer: Analysis of the Randomised Phase III Evaluation of Faslodex versus Exemestane Clinical Trial (EFECT). *Eur. J. Cancer* 114, 55–66. Epub 2019/05/07PubMed PMID: 31059974. doi:10.1016/j.ejca.2019.04.002

McCartney, A., Bonechi, M., De Luca, F., Biagioni, C., Curigliano, G., Moretti, E., et al. (2020). Plasma Thymidine Kinase Activity as a Biomarker in Patients with Luminal Metastatic Breast Cancer Treated with Palbociclib within the TREnd Trial. *Clin. Cancer Res.* 26 (9), 2131–2139. Epub 2020/01/16PubMed PMID: 31937617. doi:10.1158/1078-0432.CCR-19-3271

Nisman, B., Nechushtan, H., Biran, H., Gantz-Sorotsky, H., Peled, N., Gronowitz, S., et al. (2014). Serum Thymidine Kinase 1 Activity in the Prognosis and Monitoring of Chemotherapy in Lung Cancer Patients: a Brief Report. *J. Thorac. Oncol.* 9 (10), 1568–1572. Epub 2014/12/19PubMed PMID: 25521401. doi:10.1097/JTO.0000000000000276

Pérez-Ruiz, E., Melero, I., Kopecka, J., Sarmento-Ribeiro, A. B., García-Aranda, M., and De Las Rivas, J. (2020). Cancer Immunotherapy Resistance Based on Immune Checkpoints Inhibitors: Targets, Biomarkers, and Remedies. *Drug Resist. Updates* 53, 100718. Epub 2020/08/01PubMed PMID: 32736034. doi:10.1016/j.drup.2020.100718

Ren, S., Wei, G.-H., Liu, D., Wang, L., Hou, Y., Zhu, S., et al. (2018). Whole-genome and Transcriptome Sequencing of Prostate Cancer Identify New Genetic Alterations Driving Disease Progression. *Eur. Urol.* 73 (3), 322–339. PubMed PMID: WOS:000425085700021. doi:10.1016/j.eururo.2017.08.027

Ross-Adams, H., Lamb, A. D., Dunning, M. J., Halim, S., Lindberg, J., Massie, C. M., et al. (2015). Integration of Copy Number and Transcriptomics Provides Risk Stratification in Prostate Cancer: A Discovery and Validation Cohort Study. *EBioMedicine* 2 (9), 1133–1144. Epub 2015/10/27PubMed PMID: 26501111; PubMed Central PMCID: PMCPMC4588396. doi:10.1016/j.ebiom.2015.07.017

Rotinen, M., You, S., Yang, J., Coetzee, S. G., Reis-Sobreiro, M., Huang, W.-C., et al. (2018). ONECUT2 Is a Targetable Master Regulator of Lethal Prostate Cancer that Suppresses the Androgen axis. *Nat. Med.* 24 (12), 1887–1898. Epub 2018/11/28PubMed PMID: 30478421; PubMed Central PMCID: PMCPMC6614557. doi:10.1038/s41591-018-0241-1

Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., et al. (2019). TISIDB: an Integrated Repository portal for Tumor-Immune System Interactions. *Bioinformatics* 35 (20), 4200–4202. Epub 2019/03/25PubMed PMID: 30903160. doi:10.1093/bioinformatics/btz210

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer Statistics, 2020. *CA A. Cancer J. Clin.* 70 (1), 7–30. Epub 2020/01/09PubMed PMID: 31912902. doi:10.3322/caac.21590

Song, Z.-y., Chao, F., Zhuo, Z., Ma, Z., Li, W., and Chen, G. (2019). Identification of Hub Genes in Prostate Cancer Using Robust Rank Aggregation and Weighted Gene Co-expression Network Analysis. *Aging* 11 (13), 4736–4756. Epub 2019/07/16PubMed PMID: 31306099; PubMed Central PMCID: PMCPMC6660050. doi:10.18632/aging.102087

Steele, K. E., Tan, T. H., Korn, R., Dacosta, K., Brown, C., Kuziora, M., et al. (2018). Measuring Multiple Parameters of CD8+ Tumor-Infiltrating Lymphocytes in Human Cancers by Image Analysis. *J. Immunotherapy Cancer* 6 (1), 20. Epub 2018/03/08PubMed PMID: 29510739; PubMed Central PMCID: PMCPMC5839005. doi:10.1186/s40425-018-0326-x

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71 (3), 209–249. Epub 2021/02/05PubMed PMID: 33538338. doi:10.3322/caac.21660

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. Epub 2019/05/23PubMed PMID: 31114875; PubMed Central PMCID: PMCPMC6602440. doi:10.1093/nar/gkz430

Taylor, B. S., Schultz, N., Hieronymus, H., Gopalan, A., Xiao, Y., Carver, B. S., et al. (2010). Integrative Genomic Profiling of Human Prostate Cancer. *Cancer Cell* 18 (1), 11–22. Epub 2010/06/29PubMed PMID: 20579941; PubMed Central PMCID: PMCPMC3198787. doi:10.1016/j.ccr.2010.05.026

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2019). The Immune Landscape of Cancer. *Immunity* 51 (2), 411–412. Epub 2019/08/23PubMed PMID: 31433971. doi:10.1016/j.immuni.2019.08.004

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Tissue-based Map of the Human Proteome. *Science* 347 (6220), 1260419. Epub 2015/01/24PubMed PMID: 25613900. doi:10.1126/science.1260419

Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., et al. (2010). Towards a Knowledge-Based Human Protein Atlas. *Nat. Biotechnol.* 28 (12), 1248–1250. Epub 2010/12/09PubMed PMID: 21139605. doi:10.1038/nbt1210-1248

Wang, Y., Jiang, X., Wang, S., Yu, H., Zhang, T., Xu, S., et al. (2018). Serological TK1 Predict Pre-cancer in Routine Health Screenings of 56,178 People. *Cbm* 22 (2), 237–247. Epub 2018/04/25PubMed PMID: 29689706. doi:10.3233/CBM-170846

Wang, Y., Wang, J., Yan, K., Lin, J., Zheng, Z., and Bi, J. (2020). Identification of Core Genes Associated with Prostate Cancer Progression and Outcome via Bioinformatics Analysis in Multiple Databases. *PeerJ* 8, e8786. Epub 2020/04/09PubMed PMID: 32266115; PubMed Central PMCID: PMCPMC7120053. doi:10.7717/peerj.8786

Weagel, E. G., Burrup, W., Kovtun, R., Velazquez, E. J., Felsted, A. M., Townsend, M. H., et al. (2018). Membrane Expression of Thymidine Kinase 1 and Potential Clinical Relevance in Lung, Breast, and Colorectal Malignancies. *Cancer Cel Int* 18, 135. Epub 2018/09/15PubMed PMID: 30214377; PubMed Central PMCID: PMCPMC6131957. doi:10.1186/s12935-018-0633-9

Wenzel, E. S., and Singh, A. T. K. (2018). Cell-cycle Checkpoints and Aneuploidy on the Path to Cancer. *Iv* 32 (1), 1–5. Epub 2017/12/25PubMed PMID: 29275292; PubMed Central PMCID: PMCPMC5892633. doi:10.21873/invivo.11197

Wong, Y. N. S., Ferraldeschi, R., Attard, G., and de Bono, J. (2014). Evolution of Androgen Receptor Targeted Therapy for Advanced Prostate Cancer. *Nat. Rev. Clin. Oncol.* 11 (6), 365–376. Epub 2014/05/21PubMed PMID: 24840076. doi:10.1038/nrclinonc.2014.72

Yang, Y., Attwood, K., Bshara, W., Mohler, J. L., Guru, K., Xu, B., et al. (2021). High Intratumoral CD8 + T-cell Infiltration Is Associated with Improved Survival in Prostate Cancer Patients Undergoing Radical Prostatectomy. *Prostate* 81 (1), 20–28. Epub 2020/10/22PubMed PMID: 33085799. doi:10.1002/pros.24068

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10 (1), 1523. Epub 2019/04/05PubMed PMID: 30944313; PubMed Central PMCID: PMCPMC6447622. doi:10.1038/s41467-019-09234-6

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

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

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 12 May 2022
doi: 10.3389/fgene.2022.801698

Analysis of the Expression and Role of Keratin 17 in Human Tumors

Hanqun Zhang[1†], Yun Zhang[2†], Zhiyu Feng[1†], Liang Lu[1], Yong Li[1], Yuncong Liu[1] and Yanping Chen[1]**

[1]Department of Oncology, Guizhou Provincial People's Hospital, Guiyang, China, [2]Department of Pathology, Guizhou Provincial People's Hospital, Guiyang, China

Objective: We aimed to explore the expression and carcinogenic effect of KRT17 in human tumors and provide useful information for the study of KRT17.

Methods: We used databases including the Cancer Genome Atlas, Gene Expression Omnibus, GTEx, and GEPIA2 to analyze the expression, mutation, and prognosis of KRT17 in human tumors. Through webservers, including UALCAN, TIMER2.0, and STRING, we learned about the genetic variation, immune cell penetration, and enrichment analysis of KRT17-related genes.

Results: KRT17 was highly expressed in most tumors (such as esophageal cancer, lung cancer, cervical cancer, etc.), and the high expression level correlated with tumor stage and prognosis. In addition, amplification was the main type of KRT17 tumor variation, with an amplification rate of about 9%, followed by mutation, with a mutation rate of 4%. Moreover, KRT17 was strongly associated with tumor-infiltrating immune cells (such as macrophages, CD8+T, Tregs, and cancer-associated fibroblasts). KEGG analysis suggested that KRT17 may play a role in tumor pathogenesis following human papillomavirus infection, and the gene ontology enrichment analysis indicated that the carcinogenicity of KRT17 can be attributed to cadherin binding, intermediate fibrocytoskeleton and epidermal development.

Conclusion: KRT17 may play an important role in the occurrence, development, and prognosis of malignant tumors. We provided a relatively comprehensive description of the carcinogenic role of KRT17 in different tumors for the first time.

Keywords: Krt17, carcinogenesis, mechanism, prognosis, cancer

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez, Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Chih-Yang Wang, Taipei Medical University, Taiwan
Liang Shao, Wuhan University, China

****Correspondence:***
Yuncong Liu
sonia_lyc@foxmail.com
Yanping Chen
zenghao1234hui@sina.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 October 2021*
Accepted: *25 March 2022*
Published: *12 May 2022*

Citation:
Zhang H, Zhang Y, Feng Z, Lu L, Li Y, Liu Y and Chen Y (2022) Analysis of the Expression and Role of Keratin 17 in Human Tumors. Front. Genet. 13:801698. doi: 10.3389/fgene.2022.801698

INTRODUCTION

Malignant tumors are a serious threat to human health and one of the major causes of death worldwide. In recent years, the morbidity and mortality rates of malignant tumors have increased significantly (Yao et al., 2019). Moreover, with the increase of population and poor lifestyle choices, the number of new cases and deaths related to malignant tumors is expected to increase rapidly (Torre et al., 2016). Early detection, early diagnosis, and treatment have become the goals for prevention and treatment of malignant tumors (Jemal et al., 2010). With the development of science and technology and the arrival of the era of precision medicine, searching for sensitive biomarkers and prognostic indicators of malignant tumors and exploring their molecular mechanisms are important for prevention and treatment. Given the complexity of malignancies, it is important to

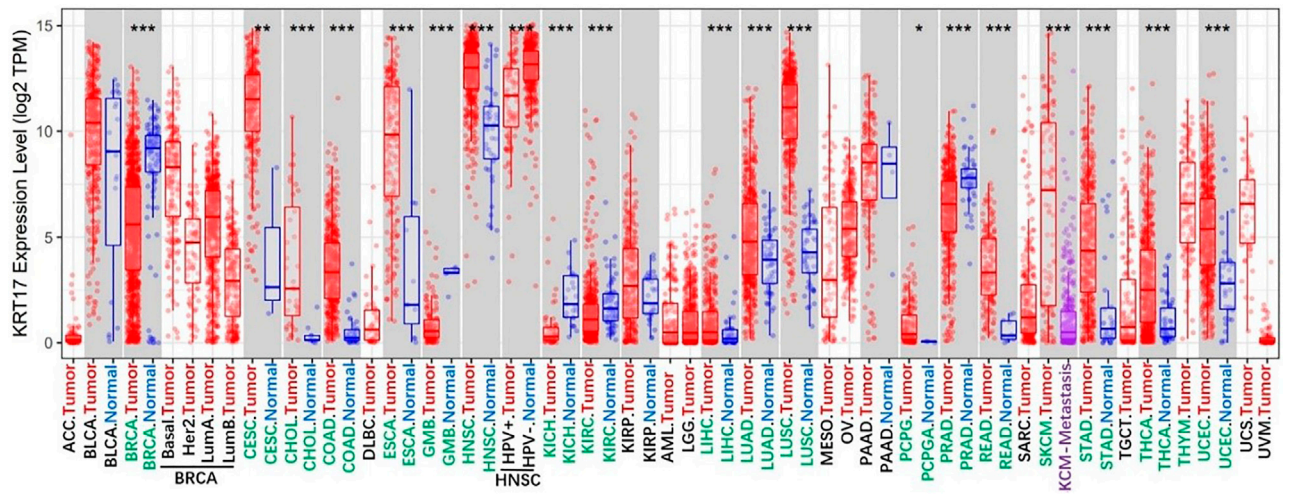


FIGURE 1 | Expression status of KRT17 in various malignant tumors was analyzed through TIMER2; the figure showed that KRT17 expression in most malignant tumor tissues was higher than that in normal tissues, and it was statistically significant (*p < 0.05; **p < 0.01; ***p < 0.001) (TCGA dataset).

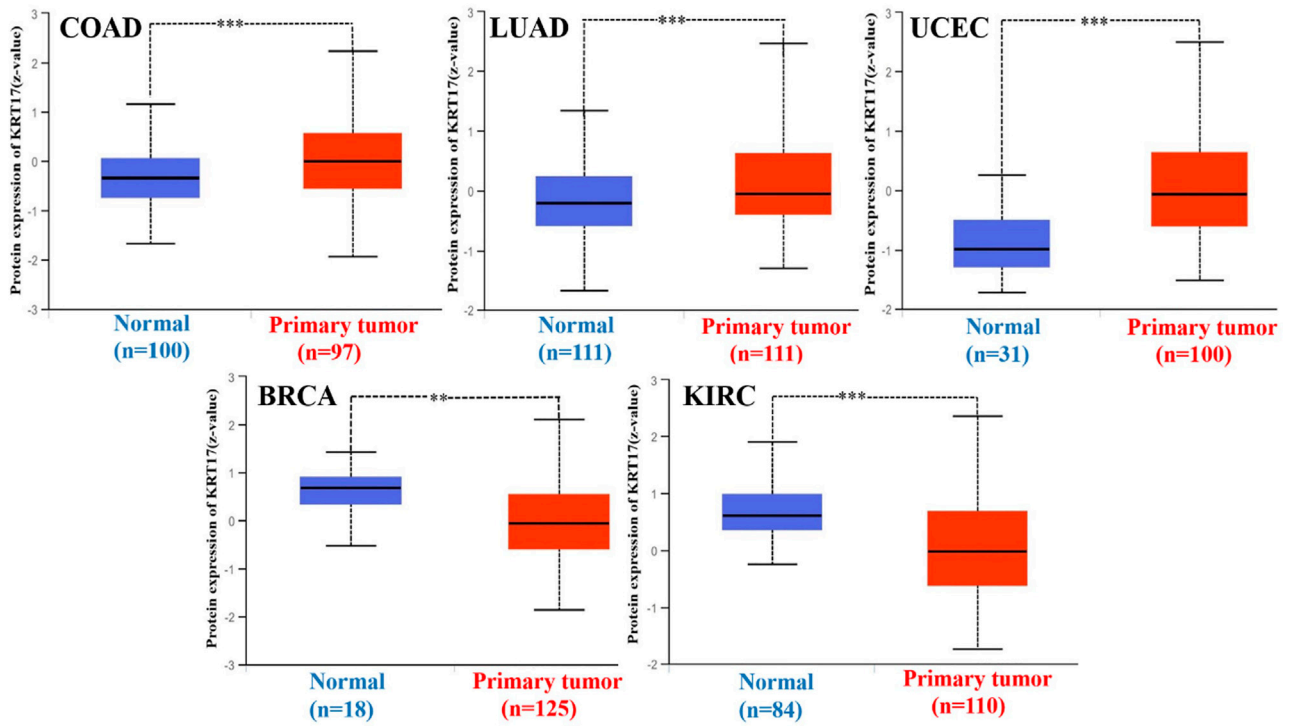


FIGURE 2 | Expression of KRT17 in total protein of COAD, LUAD, UCEC, BRCA, and KIRC. The expression of KRT17 in COAD, LUAD and UCEC tumor tissues was higher than that in normal tissues, while in BRCA and KIRC, the expression of KRT17 in normal tissues was higher than that in tumor tissues (*p < 0.05; **p < 0.01; ***p < 0.001) (CPTAC dataset).

conduct pan-cancer expression analysis of any gene of interest and assess its correlation with clinical prognosis and its potential molecular mechanisms. With the rapid development of genomics, transcriptomics and proteomics, and the establishment of databases [The Gene Expression Omnibus (GEO), The Cancer Genome Atlas (TCGA), and The Human Protein Atlas (HPA)], the data can be conveniently accessed, allowing us to perform a pan-cancer analysis.

KRT17 is a triplet structure protein comprising 432 amino acids: a non-helical head (1–83), an α helical rod (84–392), and a

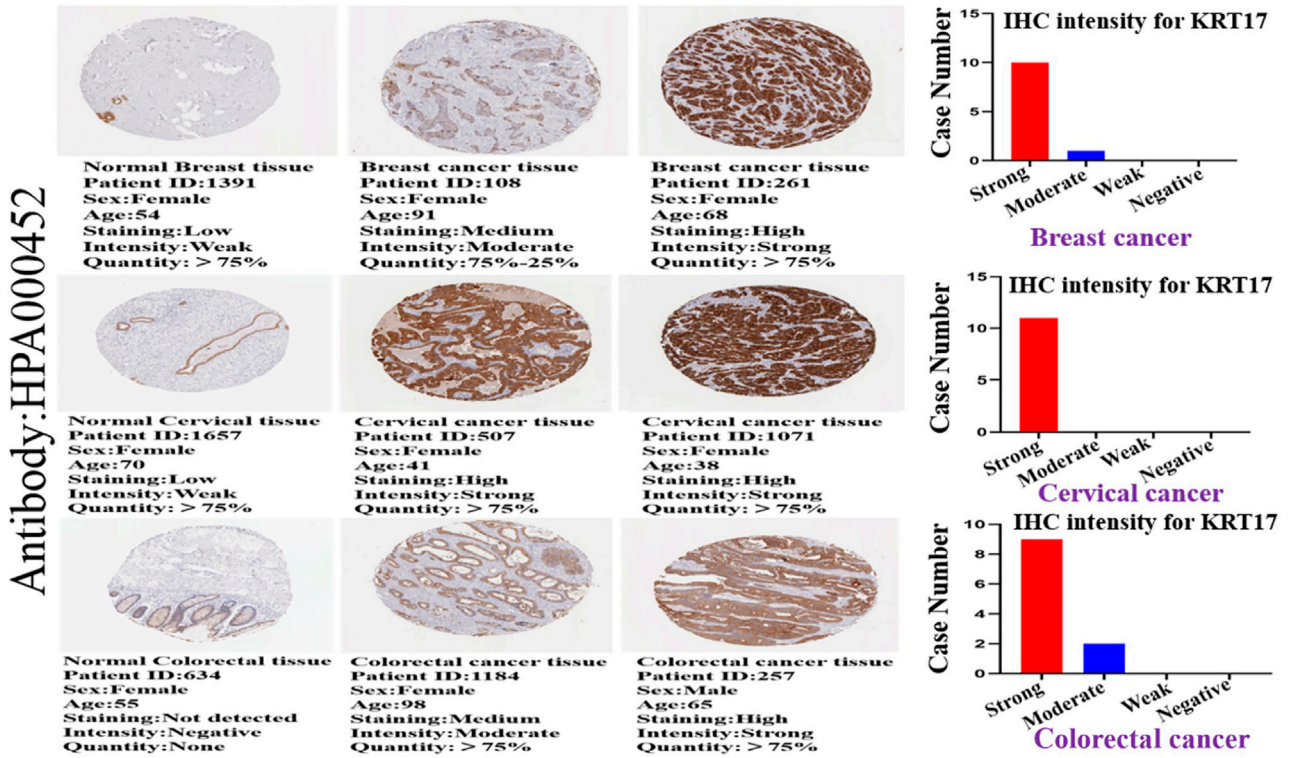


FIGURE 3 | Immunohistochemical (IHC) images of normal and tumor tissues of KRT17 from patients with breast cancer, cervical cancer, and colorectal cancer, and the intensity of the IHC of KRT17. The bar graph shows the IHC intensity of KRT17 (breast cancer: 11 patients, cervical cancer: 11 patients, and colorectal cancer: 11 patients). All IHC images and patient information were derived from the HPA.

non-helical tail domain (393–432) (Yang et al., 2019). The KRT17 gene is located on chromosome 17q21.2 (Kurokawa et al., 2011). The KRT17 is a multifunctional protein that regulates numerous cellular processes, including cell proliferation and growth (Depianto et al., 2010; Mikami et al., 2015). In addition, KRT17 can promote the release of inflammatory cytokines and promote the occurrence and development of tumors (Lo et al., 2010; Chung et al., 2015). Our team found that KRT17 expression was significantly different before and after cervical cancer radiotherapy when screening radiotherapy sensitivity genes (GSE6213). Considering that KRT17 may be a gene related to the radiotherapy sensitivity of cervical cancer, we assessed KRT17 and found that it was abnormally expressed in a variety of tumors after reviewing the literature. This abnormal expression is related to the occurrence, development, treatment, and prognosis of tumors (Ide et al., 2012; Ujiie et al., 2020). Moreover, despite the large number of clinical data, there is no pan-cancer evidence of a relationship between KRT17 and various tumor types. Therefore, we used databases or webservers such as TCGA, Tumor Immune Estimation Resource 2.0 (TIMER2.0), GEO, and Gene Expression Profiling Interactive Analysis 2 (GEPIA2) to conduct pan-cancer analysis of KRT17 and explore the potential molecular mechanisms by which it relates to the occurrence, development, and clinical prognosis of different cancer types.

MATERIALS AND METHODS

Gene Expression Analysis

We searched the TIMER2.0 (http://timer.cistrome.org/) webservers in the Gene_DE KRT17 module input and found differences in KRT17 expression in tumor and normal tissues in TCGA database. There were no matched normal tissues in TCGA database. We obtained tumor tissues from the GEPIA2 (http://gepia2.cancer-pku.cn/) database and normal tissues from the Genotype-Tissue Expression (GTEx) database to assess the expression differences between the two tissue types. The *p* value cut-off was below 0.01, and the log fold change (logFC) cut-off was equal to 1. We selected "match TCGA normal and GTEx data" in the field of "match normal data." With UALCAN tools (http://ualcan.path.uab.edu/index.html) and TCGA data analysis, we obtained KRT17 expression profiles for different tumor stages. UALCAN protein expression analysis (http://ualcan.path.uab.edu/home), *via* the "CPTAC analysis" module, was used to obtain the gene and protein expression profiles of KRT17 in tumor tissue and normal tissue.

Analysis of Protein Expression

We entered "KRT17" into the "search" module of the HPA (https://www.proteinatlas.org/) network database, clicked the "search" button, and then selected the "tissue" and "pathology" modules

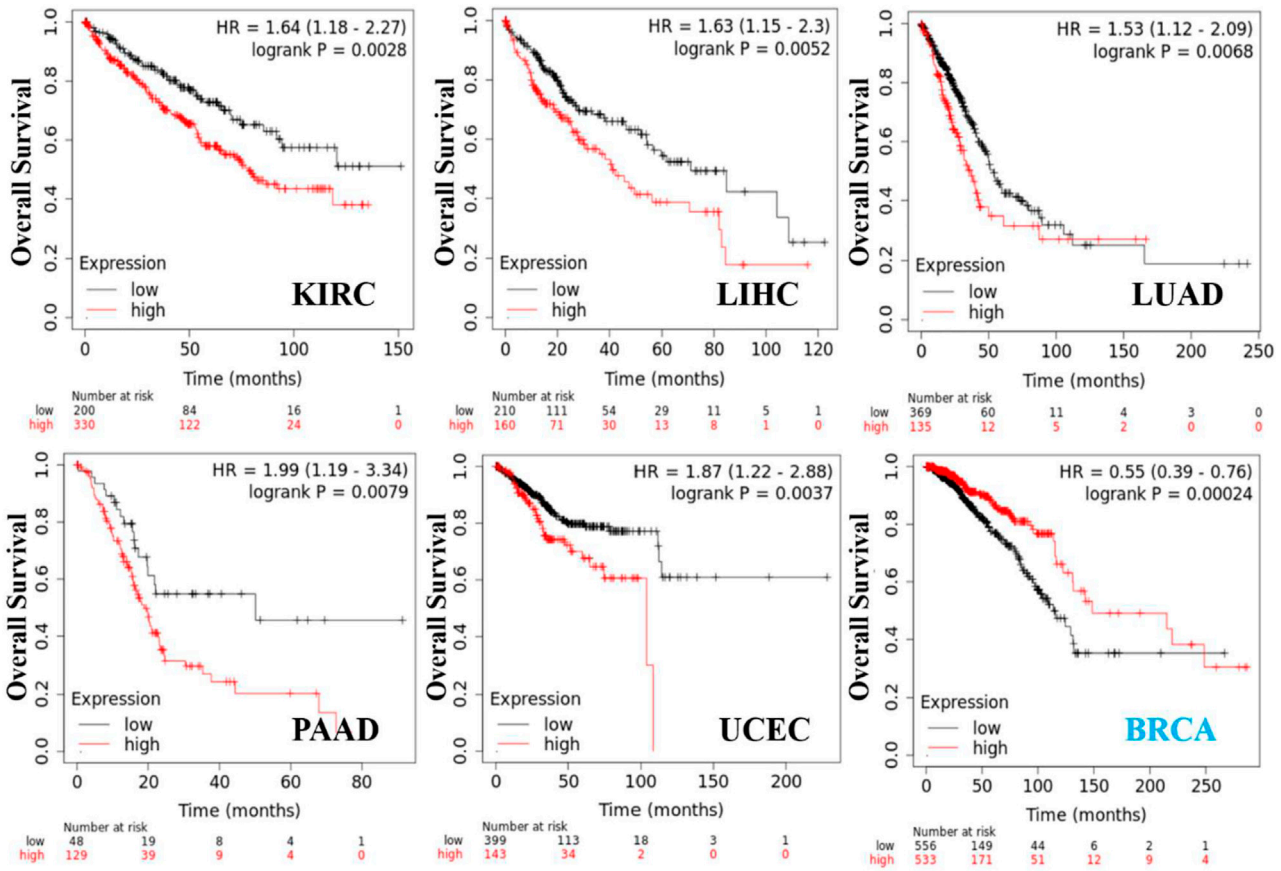


FIGURE 4 | KRT17 expression in OS in KIRC, LIHC, LUAD, PAAD, UCEC, and BRCA. We obtained a relationship between KRT17 and survival prognosis of cancer from Kaplan–Meier Plotter. Kaplan–Meier curves were all positive. The solid red line represents the high expression of KRT17 in tumor tissues, and the solid black line represents the low expression of KRT17 in tumor tissues.

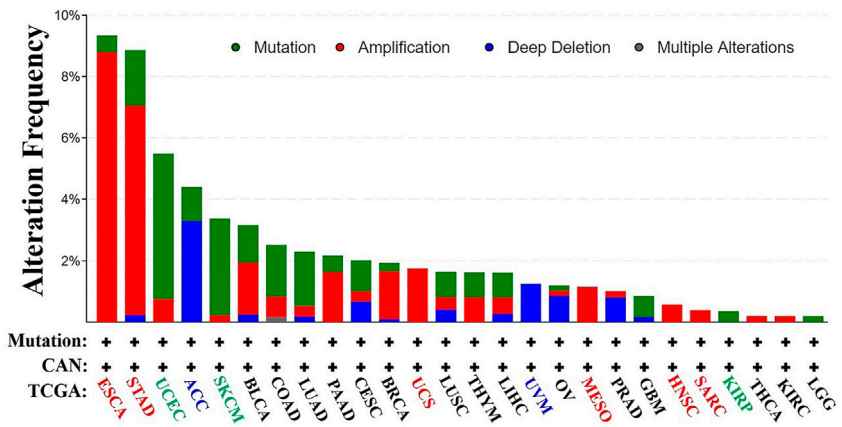


FIGURE 5 | Alteration frequency of KRT17 in a variety of malignancies. The figure showed that KRT17 was the main type of amplification in esophageal cancer, the main type of mutation in UCEC, and the main type of deep deletion in ACC (TCGA dataset).

to obtain KRT17 expression data for human tumor tissue samples and normal tissue samples and evaluate the protein expression of KRT17 on clinical specimen images. Staining reports included information of intensity, subcellular localization, single-cell variability, and antibodies. The staining intensity was classified into four categories (strong, moderate, weak, and negative) by

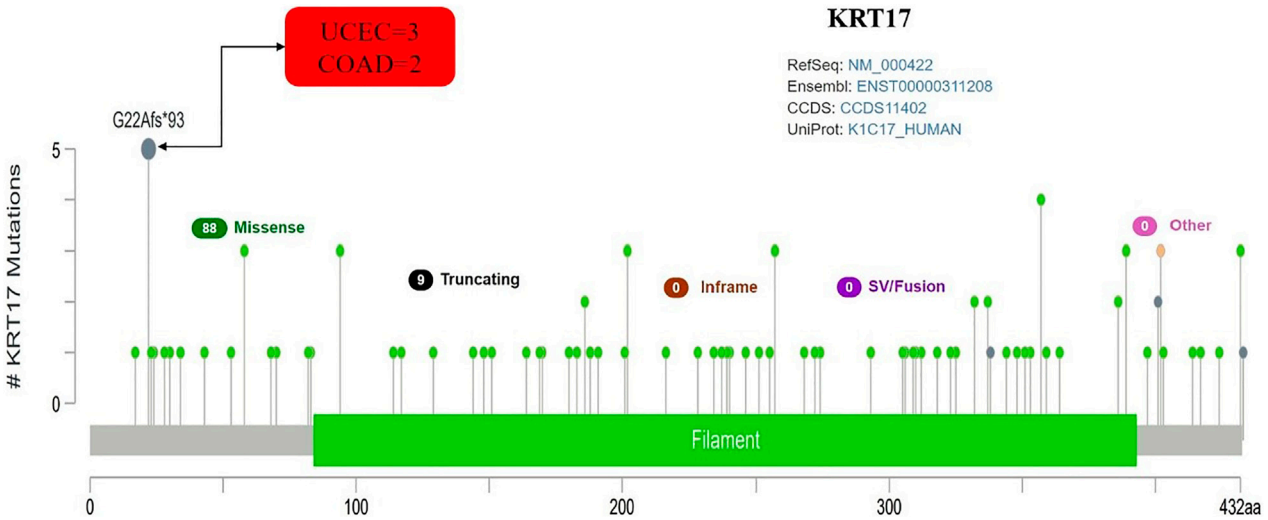


FIGURE 6 | Mutation site and number of cases of KRT17. G22Afs*93 was the site with the highest mutation frequency. There were three cases in UCEC and two cases in COAD (TCGA dataset).

using image capture and visualization techniques. The protein expression score was determined by the staining intensity from the immunohistochemistry (IHC) data and the proportion of stained cells as follows: negative–not detected; weak <25%–not detected; weak combined with either 25–75% or 75%–low; moderate <25%–low; moderate combined with either 25–75% or 75%–medium; strong <25%–medium; and strong combined with either 25–75% or 75%–high (Ta et al., 2021).

Survival Prognosis Analysis

In the "start KM plotter for pan-cancer" list of the Kaplan–Meier Plotter (https://kmplot.com/analysis/), we obtained the overall survival (OS) and relapse-free survival (RFS) data of patients with various human malignancies sorted according to KRT17 expression. The median value was set as the cut-off. A Cox proportional hazards (PH) model was used to calculate the risk ratio. The log-rank sum test was used for hypothesis testing, and the OS plots and RFS plots were obtained *via* Kaplan–Meier Plotter survival analysis.

Genetic Variation Analysis

In the cBioPortal (https://www.cbioportal.org/) dataset, we chose "TGCA extensive cancer atlas research" and searched for "KRT17" genetic variation characteristics. All the changes, mutational results, and copy number changes in TGCA tumors were reviewed in the cancer type summary. We also used the comparison/survival module to assess the differences in overall survival and disease-free, progression-free, and disease-free survival for cancer patients from TCGA database. Kaplan–Meier survival plots were generated by the log-rank sum test, and $p < 0.05$ was considered to indicate significance.

Immune Infiltration Analysis

On the TIMER (http://timer.cistrome.org/) website "immune" template "gene expression" input ("KRT17") and "immune infiltrates" ["CD8 + T cells", "regulatory T cells" (Tregs), and "cancer-associated fibroblasts"] were selected to determine the relationships between KRT17 and tumor immune infiltration. The TIMER, CIBERSORT, CIBERSORT-abs, Quantiseq, Xcell, MCPCounter, and EPIC algorithms were used to estimate immune infiltration. The *p* values and bias correlation values were obtained by Spearman's rank correlation test with purity adjustment. The data are presented as a scatter plot.

Enrichment Analysis

First, the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) (https://string-db.org/) was used to screen 50 proteins that are experimentally verified to bind to KRT17. In the STRING webserver, we selected the column of "protein name", entered "KRT17", and selected "*Homo sapiens*"; the parameters were set as follows: "full network", "evidence", "experiments", "low confidence (0.150)", and "no more than 50 interactors in the first shell". After parameter setting, we continued to follow the instructions for the next step to obtain the binding proteins of KRT17. GEPIA2 was used for similar gene detection, and the first 200 similar genes were obtained. In addition, we used Jvenn, a Venn diagram viewer (http://bioinformatics.psb.ugent.be/webtools/Venn/), for cross analysis of KRT17 and its interacting genes, GEPIA2 was used for correlation analysis, Pearson's correlation analysis was used for paired genes, and log2 transcripts per million (TPM) was applied to the dot plot to obtain the *p* value and correlation coefficient. In the Gene_cor module of the TIMER2.0 website, KRT17 was input as the gene of interest, and KRT5, KRT6a, KRT6b, KRT6c, and SFN gene expressions were analyzed. After execution, we generated a heatmap. The data included *p* values and bias correlation values obtained by purity-adjusted Spearman's rank correlation test. For the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, we uploaded the data to the Database for Annotation, Visualization, and Discovery (DAVID) and then selected the settings "official_gene_symbol" and "*Homo sapiens*" to obtain the functionally annotated

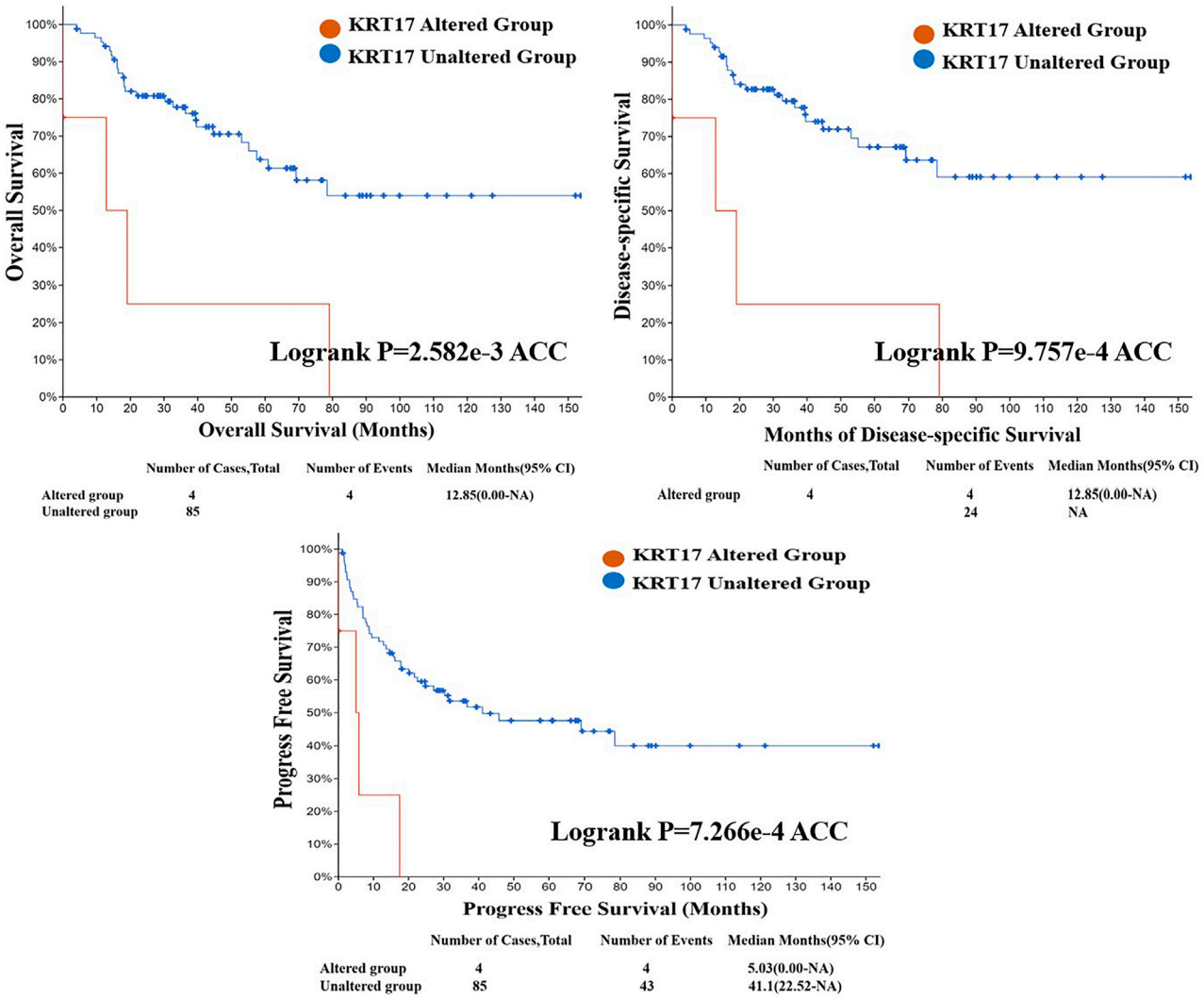


FIGURE 7 | Correlation between KRT17 mutation status and overall survival (OS), disease-specific survival (DSS), and progression-free survival (PFS) of ACC. The solid red line represents the KRT17-altered group in tumor tissues, and the dash-dotted blue line represents KRT17-unaltered group in tumor tissues (TCGA dataset).

map data. The enriched pathways were analyzed using the "Tidyr" (https://cran.r-project.org/web/packages/tidyr/index.html) and "ggplot2" (https://Cran.r-project.org/web/packages/ggplot2/index.html) R language packages (version: version number: 3.6.2). In addition, we used the " clusterProfiler" (http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) R language package for the Gene Ontology (GO) enrichment analysis. $p < 0.05$ was considered to indicate statistical significance.

RESULTS

Gene Expression Results

We determined the difference in the expression of KRT17 between various cancer types in TCGA database *via* the TIMER2.0 webserver (**Figure 1**). KRT17 expression was higher in cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), kidney renal clear cell carcinoma (KIRC), liver hepatocellular carcinoma (LIHC), lung squamous cell carcinoma (LUSC), lung adenocarcinoma (LUAD), prostate adenocarcinoma (PRAD), rectum adenocarcinoma (READ), skin cutaneous melanoma (SKCM), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), uterine corpus endometrial carcinoma (UCEC) ($p < 0.001$), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) ($p < 0.01$), and pheochromocytoma and paraganglioma (PCPG) ($p < 0.05$) tissues than in normal tissues. In breast invasive carcinoma (BRCA) and kidney chromophobe (KICH), the expression of KRT17 was higher in normal tissues than in tumor tissues ($p < 0.001$). There was no difference in KRT17 expression between

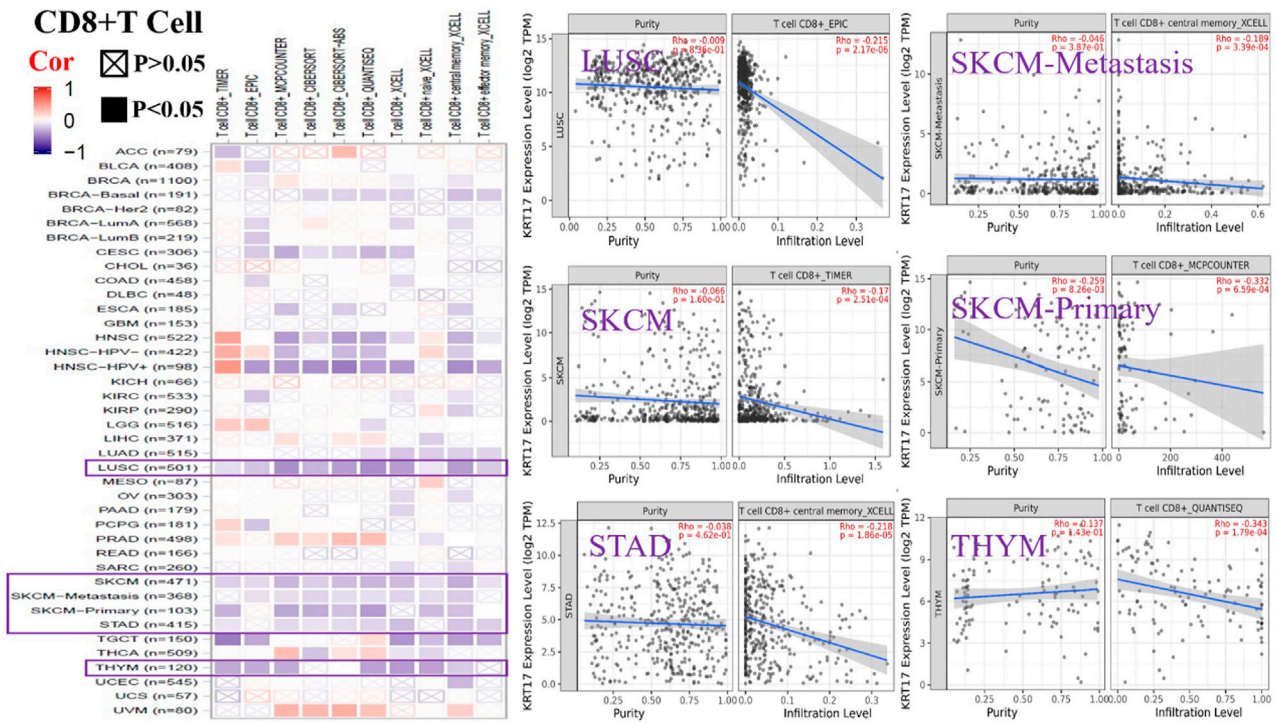


FIGURE 8 | Correlation between KRT17 expression and immune infiltration of CD8+T cells. In LUSC, SKCM, SKCM-metastasis, SKCM-primary, STAD and THYM and KRT17 expression was negatively correlated with CD8+T cell expression (TCGA dataset).

tumor and normal tissues in bladder urothelial carcinoma (BLCA), kidney renal papillary cell carcinoma (KIRP), and pancreatic adenocarcinoma (PAAD) ($p > 0.05$). TCGA, adrenocortical carcinoma (ACC), lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), brain lower grade glioma (LGG), acute myeloid leukemia (AML), ovarian serous cystadenocarcinoma (OV), sarcoma (SARC), mesothelioma (MESO), testicular germ cell tumors (TGCT), thymoma (THYM), uterine carcinosarcoma (UCS), and uveal melanoma (UVM) datasets did not include matched normal tissues. Therefore, we looked for matched normal tissues in the GTEx database and used them as controls for tumor tissues in TCGA. We analyzed the previously mentioned tumors, and the expression of KRT17 in OV, THYM, and UCS tissues was higher than that in normal tissues ($p < 0.05$), while in LGG and TGCT, the expression of KRT17 in tumor tissues was lower than that in normal tissues ($p < 0.05$) (**Supplementary Figure S1A**). There was no significant difference in expression between ACC, DLBC, AML, and SARC tumor tissues and normal tissues. Unfortunately, there were no matched normal tissues for MESO and UVM.

Next, we obtained the total protein expression data for KRT17 in BRCA, LUAD, COAD, UCEC, OV, and KIRC in tumor tissues and normal tissues in the CPTAC dataset, and the total protein expression data for the other cancers were not included in the CPTAC dataset. The total protein levels in COAD, LUAD, and UCEC tissues were higher than those in normal tissues, and the difference was statistically significant ($p < 0.05$); the levels in normal tissues were higher than those in tumor tissues in BRCA and KIRC, and the difference was statistically significant ($p < 0.05$); no significant difference was seen in OV (**Figure 2**). Since there were only six tumors with available total protein expression data in the CPTAC database, we obtained the expression data for BLCA, CESC, ESCA, HNSC, leukemia, LUSC, PAAD, and STAD from the oncomine database, and the differences in expression between these tumor tissues and their corresponding normal tissues were statistically significant (**Supplementary Figure S1B**).

We then analyzed the correlation between KRT17 expression in normal tissues and that in tumor tissues at different stages. The expression of KRT17 in CESC, HNSC, LUAD, LUSC, READ, and UCEC tissues of all stages was higher than that in normal tissues, and the differences were statistically significant ($p < 0.05$) (**Supplementary Figure S1C**); differences were also observed for BLCA (normal vs. Stage 1 and 2), COAD (normal vs. Stage 1, 3, and 4), ESCA (normal vs. Stage 2 and3), KIRP (normal vs. Stage 1 and 3) (LIHC (normal vs. Stage 1 and 2), and STAD (normal vs. Stage 2, 3, and 4), and THCA) (Normal vs. Stage 1 and 4). The expression of KRT17 in some stages was higher than that in normal tissues (**Supplementary Figure S1D**), while in BRCA, KICH and KIRC, the expression of KRT17 in each stage was lower than that in normal tissues, and the difference was statistically significant ($p < 0.05$) (**Supplementary Figure S1E**); no significant differences were found in other tumors (some of which did not have available data for matched normal tissue comparison; these included ACC, DLBC, MESO, UVM, OV, TGCT, and UCS) (**Supplementary Figures S1F,H**).

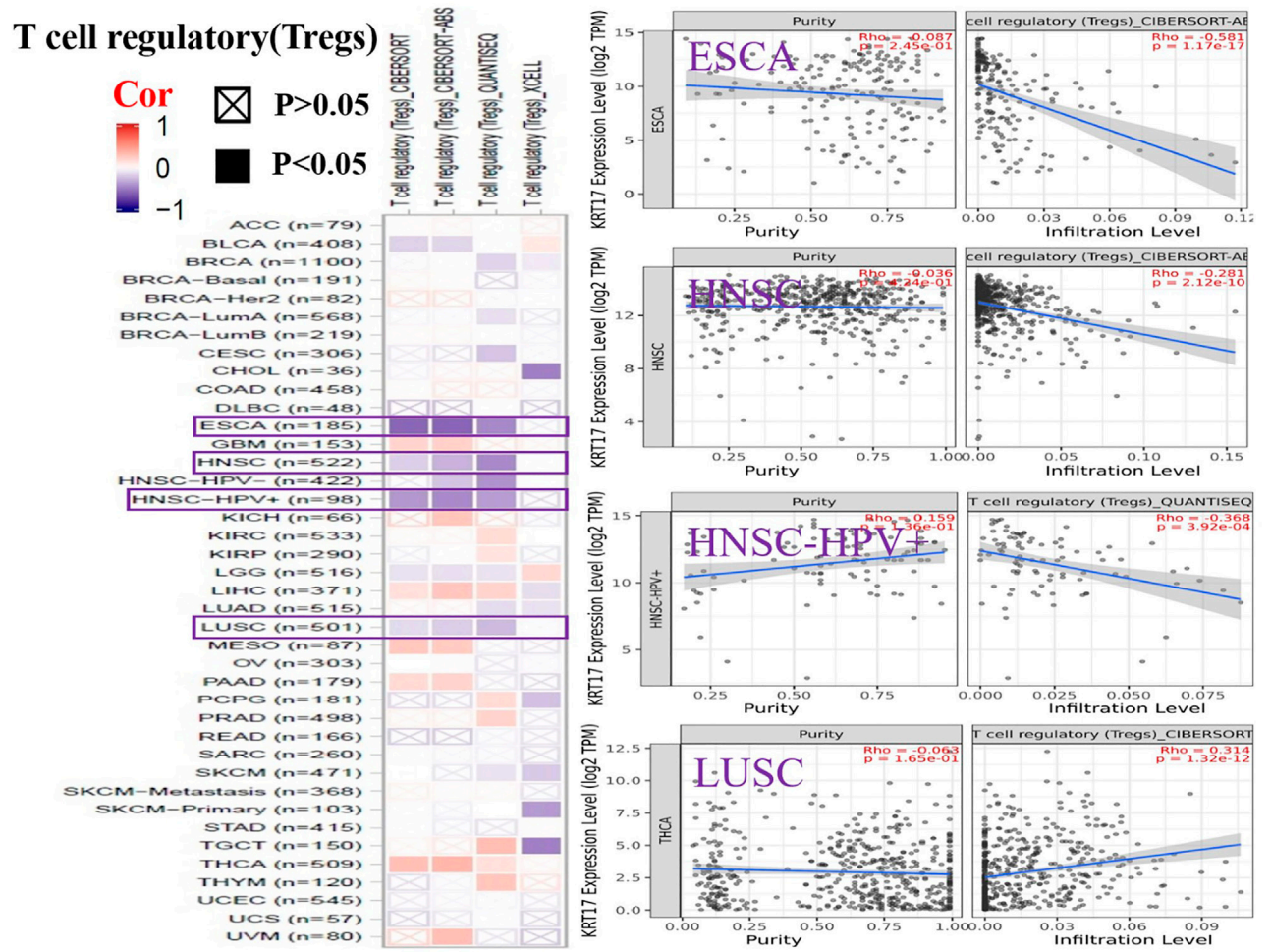


FIGURE 9 | Correlation between KRT17 expression and immune infiltration of Tregs cells. In ESCA, HNSC, and HNSC-HPV+ and LUSC, KRT17 expression was negatively correlated with Treg expression (TCGA dataset).

Protein Expression Outcomes of KRT17 in Human Clinical Specimen

We investigated the protein expression of KRT17 in the HPA database, and we obtained IHC images of 19 types of cancer tissues and corresponding normal tissues; we also obtained corresponding clinicopathological parameters, such as patient ID, sex, age, and antibody. We found that KRT17 was overexpressed in BRCA, CESC, and colorectal cancer tissues versus normal tissues, and the difference was statistically significant (**Figure 3**). Moreover, KRT17 was overexpressed in BLCA, but there was no significant difference in the expression between cancer and normal tissues. In addition, the expression of KRT17 was low in glioma, LIHC, renal cancer, testicular cancer, and melanoma tissues, but significant differences were still observed between tumor and normal tissues. KRT17 was not expressed in lymphomas or normal lymph nodes. Other tumors showed moderate expression (**Supplementary Figures S2A,B,C**; **Supplementary Table S1**).

Survival Outcomes Related to KRT17 Expression in Tumors

We obtained the survival prognosis information of patients sorted according to KRT17 expression level (high or low) for various cancer types through the Kaplan–Meier plotter. The correlation between KRT17 expression levels and survival prognosis of patients with different tumors was studied (**Figure 4**). The OS of KIRC, LIHC, LUAD, PAAD, and UCEC patients with high expression of KRT17 was lower than that of those with low expression of KRT17 ($p < 0.05$); however, the OS rate in BRCA was higher for patients with high expression of KRT17 than for those with low expression of KRT17 ($p < 0.05$). High expression of KRT17 was associated with favorable RFS in THCA, KRIC, and UCEC, while unfavorable RFS was associated with BLCA and PAAD ($p < 0.05$) (**Supplementary Figure S3**).

Genetic Variation Results for KRT17

From TCGA database, we learned the expression state of KRT17 gene genetic variation in different tumors (**Figure 5**). Among all

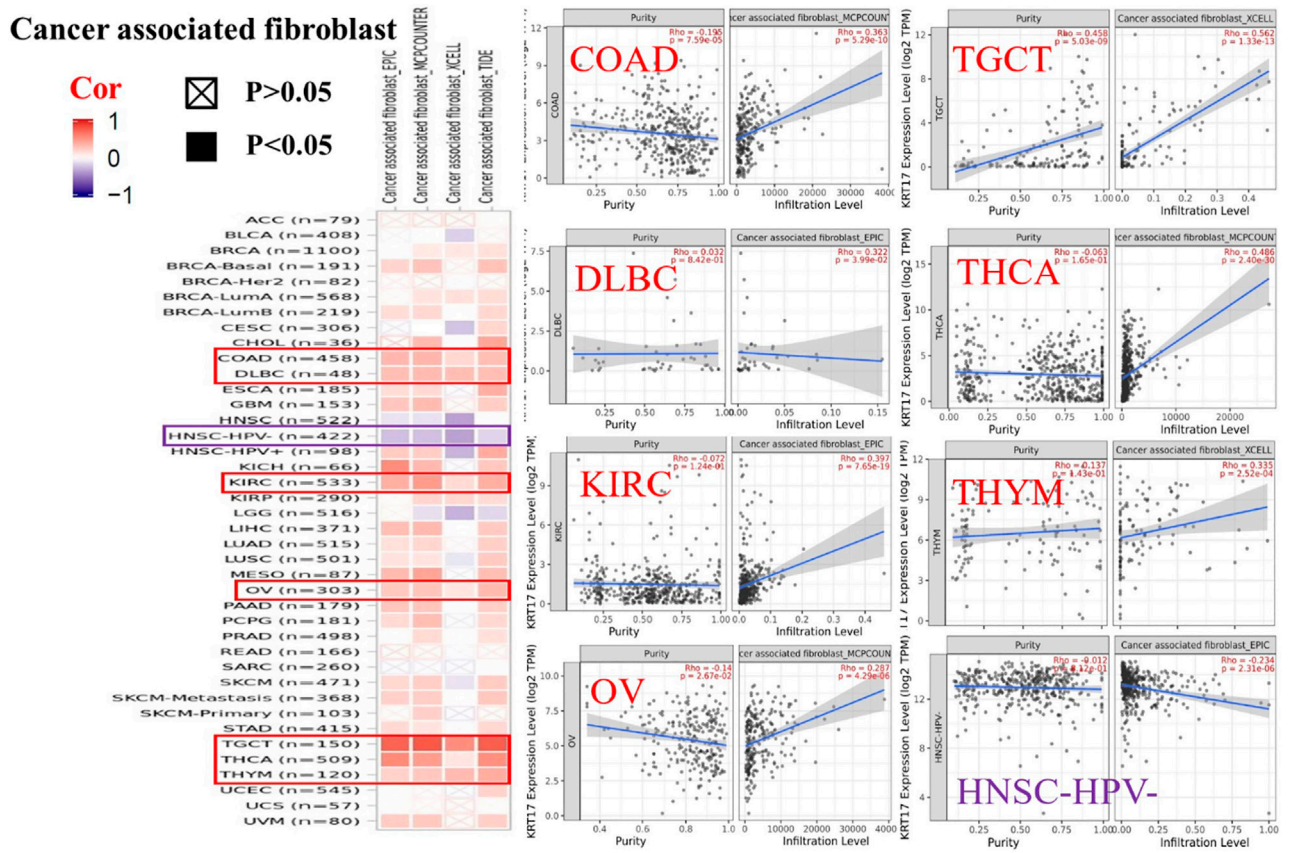


FIGURE 10 | Correlation between KRT17 expression and immune infiltration of cancer-associated fibroblasts. In THYM, TGCT, KIRC, COAD, THCA, OV, and DLBC, KRT17 expression was positively correlated with cancer-associated fibroblasts expression; however, in HNSC-HPV-, KRT17 expression was negatively correlated with cancer-associated fibroblast cells (TCGA dataset).

the variation expression states, the amplification type was associated with the highest expression, and the amplification rate was approximately 9% and was the highest in ESCA and STAD. Mutation was the dominant type of genetic variation in UCEC, with a mutation frequency of approximately 4%. In addition, KRT17 gene copy deletion was found in both ACC and MESO. Furthermore, we learned about the type, locus, and number of cases of genetic variation in KRT17. The main type of genetic variation in KRT17 was frameshift mutation. In addition, we noted that the KRT17 protein exhibited a change from glycine (G) to alanine (A) at site 22 in three UCEC cases and two COAD cases (**Figure 6**). We also explored the correlation between KRT17 gene mutation and survival prognosis of patients with different tumors. Compared with patients without KRT17 mutations, the overall survival (P = 2.582e-3), disease-specific (P = 9.757e-4), and progression-free (P = 7.266e-4) rates for ACC were lower for the KRT17 mutant group than for the non-mutant group (**Figure 7**). In addition, KRT17 mutations were found to be associated with survival in CESC (progression-free survival), THCA (progression-free survival and disease-free), and THYM (overall survival and survival). However, in SKCM (progression-free survival), the prognosis of the KRT17 mutant group was superior to that of the KRT17 mutant group (**Supplementary Figure S4**).

Results of Immune Cell Infiltration Analysis

To further clarify the relationship between KRT17 and tumor-infiltrating immune cells, we used the TIMER, CIBERSORT, CIBERSORT-abs, Quantiseq, Xcell, MCPCounter, and EPIC methods to investigate the potential relationship between the level of infiltration of different immune cells and the expression of the KRT17 gene in different types of cancer in TCGA database. We found that KRT17 expression and macrophage infiltration were positively correlated in LIHC, THCA, THYM, and UVM (**Supplementary Figure S5A**). KRT17 expression and CD8+T cell infiltration were negatively correlated in LUSC, SKCM, SKCM-metatasis, SKCM-primary, STAD, and THYM (**Figure 8**). KRT17 expression was negatively associated with Treg infiltration in ESCA, HNSC, HNSC- human papilloma virus positive (HNSC-HPV+), and LUSC (**Figure 9**) but positively associated with Treg infiltration in THCA (**Supplementary Figure S5B**). In addition, we found that KRT17 expression was positively correlated with cancer-associated fibroblast infiltration in COAD, DLBC, KIRC, OV, TGCT, THCA, and

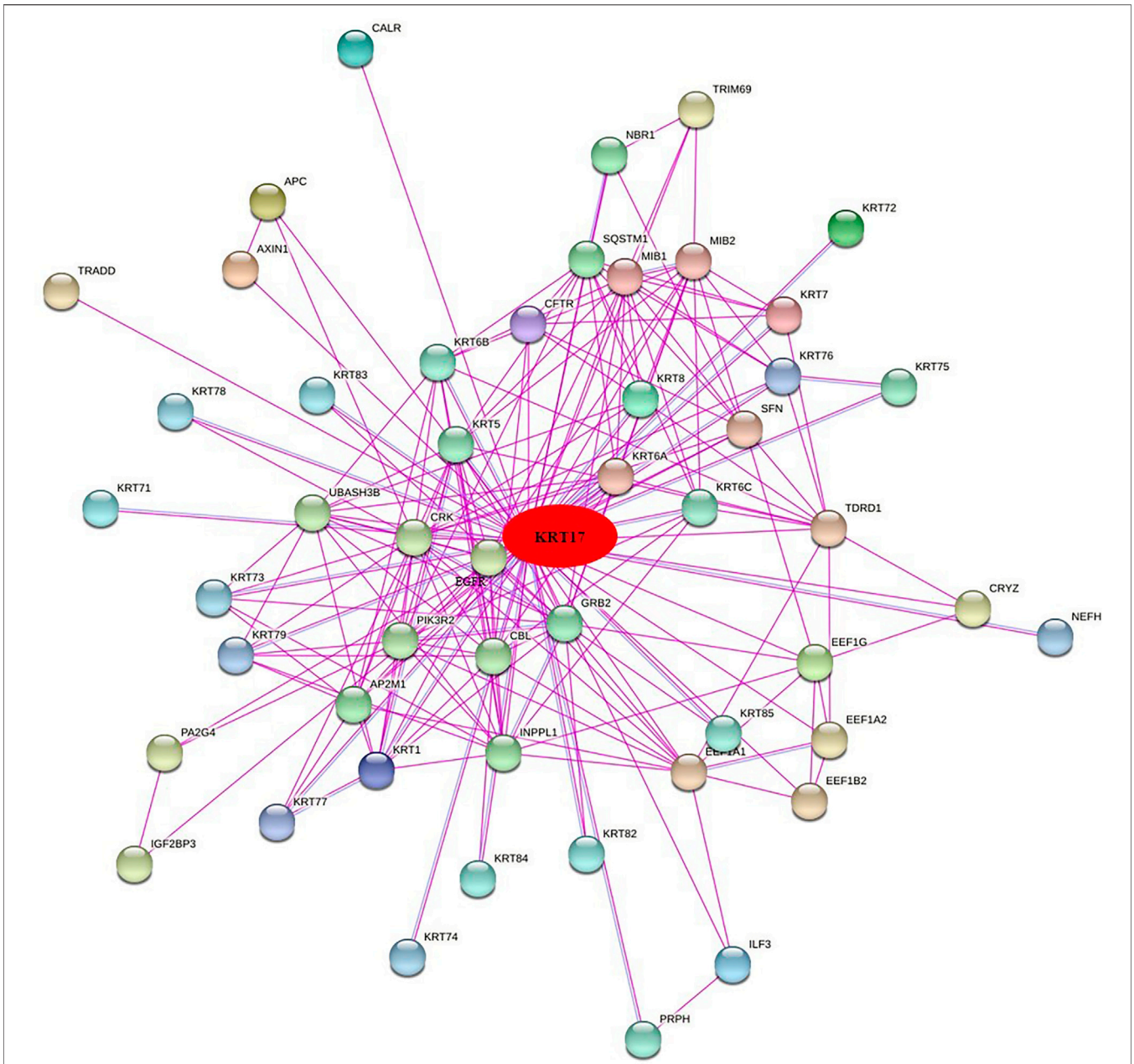


FIGURE 11 | Fifty experiments confirmed the binding proteins interacting with KRT17 and their interaction networks.

THYM and negatively correlated with cancer-associated fibroblast infiltration in HNSC (**Figure 10**).

Results of Enrichment Analysis

To further clarify the mechanism by which KRT17 mediated the pathogenesis of tumors, we screened the binding proteins interacting with KRT17 and the genes related to KRT17 expression. Using the STRING tool, we obtained 50 experimentally confirmed KRT17 binding proteins and generated an interaction network of these proteins (**Figure 11**). We combined the tumor expression data of TCGA and GTEx with the GEPIA2 tool to obtain the top 200 genes related to KRT17 expression. To further screen genes, we analyzed the intersection between 50 binding proteins interacting with KRT17 and the top 200 genes related to KRT17 expression. Five genes (KRT5, KRT6A, KRT6B, KRT6C, and SFN) were obtained (**Figure 12**). Through the GEPIA2 tool, we obtained the correlation between the expression of KRT17 and that of KRT5, KRT6a, KRT6b, KRT6c, and SFN, and the results showed that the expression of KRT17 was positively correlated with that of these

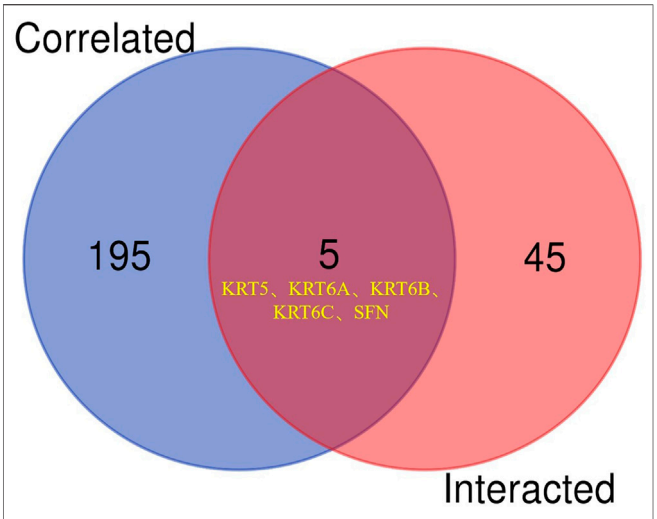


FIGURE 12 | Cross analysis of KRT17 binding protein and related genes. An intersection analysis of the abovementioned two groups showed five common members, namely, KRT5, KRT6A, KRT6B, KRT6C and SFN.

five genes (**Figure 13**). In addition, we used TIMER2.0 to obtain a heatmap between KRT17 and KRT5, KRT6A, KRT6B, KRT6C, and SFN, and the corresponding heatmap results also showed that KRT17 was positively correlated with the abovementioned five genes in most tumors (**Figure 14**).

We combined 50 binding proteins interacting with KRT17 with two datasets of the top 200 genes related to KRT17 expression and carried out KEGG and GO enrichment analyses. Through KEGG analysis, we learned that KRT17 may play a role in the pathogenesis of tumors mainly through HPV infection (**Figure 15**). GO enrichment analysis showed that these two sets of genes were mainly involved in cadherin binding, intermediate cytoskeleton filament formation, and epidermal development (**Supplementary Figures S6A–C**).

DISCUSSION

Keratin is a member of the intermediate filament superfamily that makes up the cytoskeleton and is encoded by 54 evolutionarily conserved genes. According to gene substructure and nucleotide sequence homology, keratin can be divided into two types: 28 type I acidic proteins and 26 type II basic proteins (Jacob et al., 2018). Keratin 17 (KRT17) belongs to the type I intermediate family and is an intermediate filament of the cytoskeleton, involved in structural support, metabolism, and various developmental processes (McGowan and Coulombe, 1998; Pan et al., 2013). KRT17 is mainly found in epithelial appendages, such as hair follicles, sebaceous glands, and other glands (Kurokawa et al., 2011). KRT17 is not expressed in the epidermis of normal skin, but its expression can be induced under stress conditions, such as skin injury (Kim et al., 2006) and virus infection (Proby et al., 1993). KRT17, as an intermediate filament, has long been thought to play a role in the cytoplasm. However, recent studies have shown that due to the presence of nuclear localization signals and nuclear output signals, KRT17 can move both inside and outside the nucleus, suggesting that KRT17 may regulate additional cellular processes (Hobbs et al., 2016a). In recent years, there have been an increasing number of reports on the relationship between KRT17 and malignant tumors, especially the functional association between tumors (Chivu-Economescu et al., 2017; Liu et al., 2018; Chen et al., 2020; Yan et al., 2020). However, it is still unclear whether KRT17 has a common molecular mechanism that plays a role in the occurrence and development of different tumors. To date, there have been no reports of KRT17 in pan-cancer studies. Therefore, we analyzed KRT17 in 33 different tumors by employing TCGA, TIMER2.0, GEO, GEPIA2, and HPA databases or websites and analyzed and summarized its molecular characteristics such as gene expression and gene mutation and its associations with clinical prognosis and immune infiltration. KRT17 is highly expressed in most malignant tumors, suggesting that KRT17 may play a role as an oncogene in cancers. Among the total proteins, KRT17 expression was higher in COAD, LUAD, and UCEC than in normal tissues, while KRT17 expression was higher in BRCA and KIRC normal tissues than in tumor tissues. In addition, we analyzed the expression of KRT17 protein in human tumor tissues and normal tissues in HPA and found that KRT17 was overexpressed in breast cancer, cervical cancer, colorectal cancer, and bladder cancer, but there was no statistical significance between bladder cancer tissues and normal tissues. The mRNA and total protein expressions of KRT17 in normal tissues were higher than those in tumor tissues, while the immunohistochemical results showed that KRT17 was highly expressed in tumor tissues, which may be related to the selected specimens, tumor heterogeneity, and detection technology, so it is necessary to continue to expand clinical samples for research. In glioma, liver cancer, kidney cancer, testicular cancer, and melanoma, KRT17 expression is low but higher than that in corresponding normal tissues. However, KRT17 is moderately expressed in thyroid cancer, lung cancer, gastric cancer, prostate cancer, ovarian cancer, head and neck cancer, and endometrial cancer. These results were consistent with the mRNA expression results, suggesting that KRT17 may play an oncogenic role in the occurrence and development of malignant tumors (except for BRCA). The results of the *in vitro* experiments were consistent with the previously mentioned results. KRT17 was found to be highly expressed in cervical cancer, esophageal cancer, lung cancer, gastric cancer, and colorectal cancer cell lines. KRT17 knockout was also found to inhibit cell proliferation and migration and increase sensitivity to cisplatin chemotherapy in cervical cancer cells (Escobar-Hoyos et al., 2015). Similarly, in esophageal cancer, lung cancer, gastric cancer, and colorectal cancer, KRT17 knockout can inhibit cancer cell proliferation, migration, invasion, and colony formation and induce apoptosis (Chivu-Economescu et al., 2017; Wang et al., 2019; Liu et al., 2020; Ujiie et al., 2020). The previously mentioned results confirmed that the knockout of KRT17 inhibited the growth, migration, and invasion of tumor cells, while the overexpression of KRT17 had the opposite effect, suggesting that KRT17 may be involved in the occurrence and development of malignant tumors and play an oncogenic role in malignant tumors.

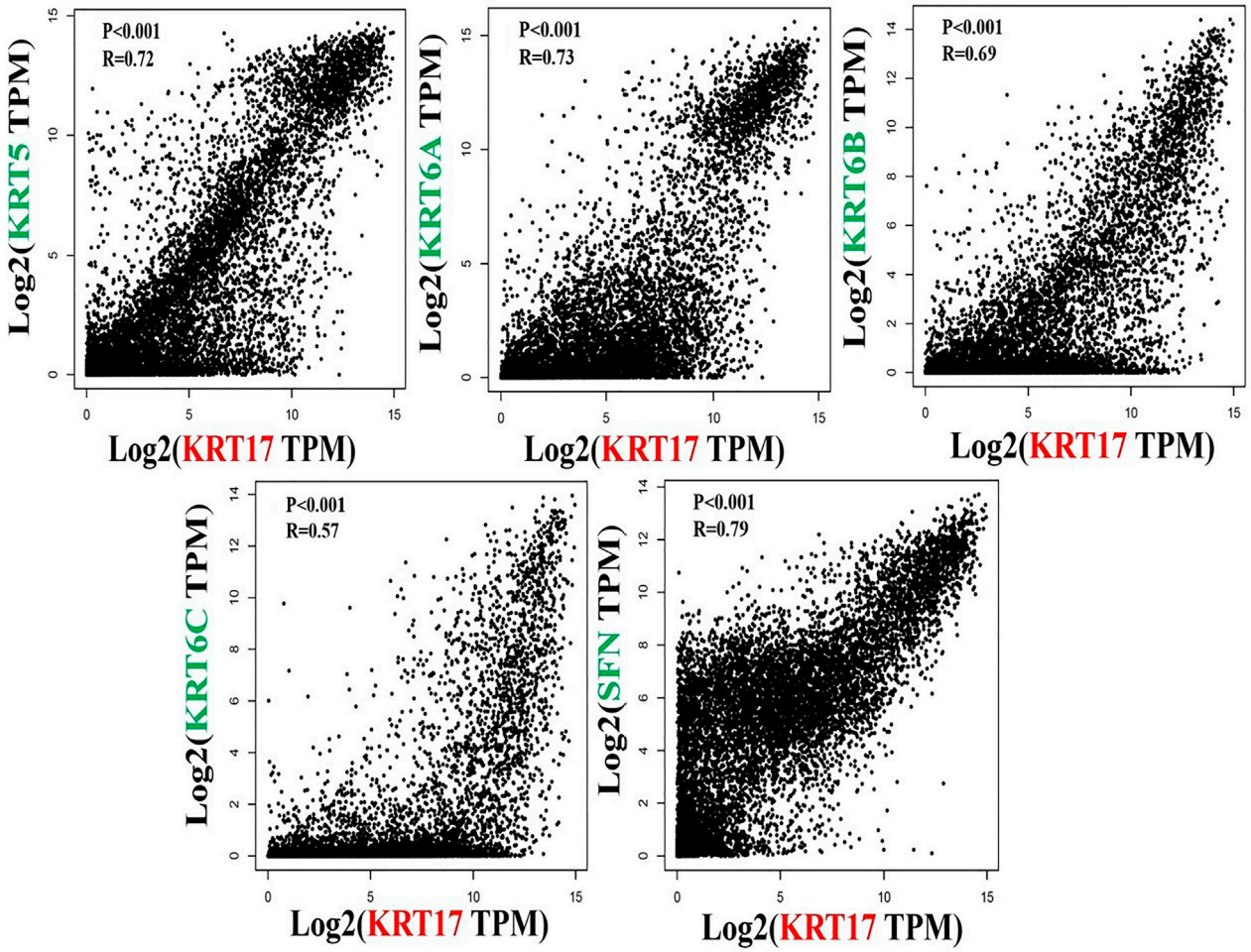


FIGURE 13 | Correlation between KRT17 and KRT5, KRT6a, KRT6b, KRT6c and SFN. The results showed that the expression of KRT17 was positively correlated with the expression of KRT5, KRT6a, KRT6b, KRT6c and SFN.

Subsequently, the expression of KRT17 at various tumor pathological stages was analyzed, and the expression of KRT17 in most tumors at various pathological stages was higher than that in normal tissues, while the expression in normal tissues was higher than that in BRCA, KICH, and KIRC tissues at different stages. According to GEPIA2, high expression of KRT17 in most tumors is associated with poor prognosis, while high expression of KRT17 is associated with a favorable prognosis in BRCA. These results indicate that KRT17 may play a role as a tumor suppressor gene in BRCA; this finding could also be explained by other reasons. The prognosis of BRCA is related not only to stage and treatment but also to molecular type. In BRCA, the prognosis of triple-negative breast cancer is worse than that of other types of breast cancer. Merkin et al. studied tissues from 164 breast cancer patients, among which 82% (28/34) of triple-negative [estrogen receptor [ER]/progesterone receptor/human epidermal growth factor receptor-2 (HER2) negative] breast cancers showed positive KRT17 expression. The positive expression rate of KRT17 in non-triple-negative breast cancer was 46% (52/112). High expression of KRT17 was associated with reduced 5-year DFS in patients with advanced cancer. Studies have shown that high KRT17 expression is associated with triple-negative status and reduced survival (Merkin et al., 2017). However, the Kaplan–Meier plotter indicated that high expression of KRT17 has good prognostic implications in breast cancer. Considering that the proportion of triple-negative breast cancer cases included in TCGA is small (approximately 14% of all breast cancer cases), high expression of KRT17 seems to indicate a good prognosis. Similarly, in lung cancer, high expression of KRT17 is not associated with OS in LUSC (Wang et al., 2019), while in LUAD, high expression of KRT17 is associated with survival (Liu et al., 2018), indicating that the expression status and clinical prognosis implications of KRT17 may be related to pathological classification. However, only 239 patients were included in the two studies, so a larger sample size is needed to confirm the role of KRT17 in survival outcomes in patients with different types of lung cancer. In addition, in KIRC and UCEC, high expression of KRT17 was negatively correlated with OS, while in RFS, high expression of KRT17 was positively correlated with RFS. Considering that KRT17 may play different roles in the

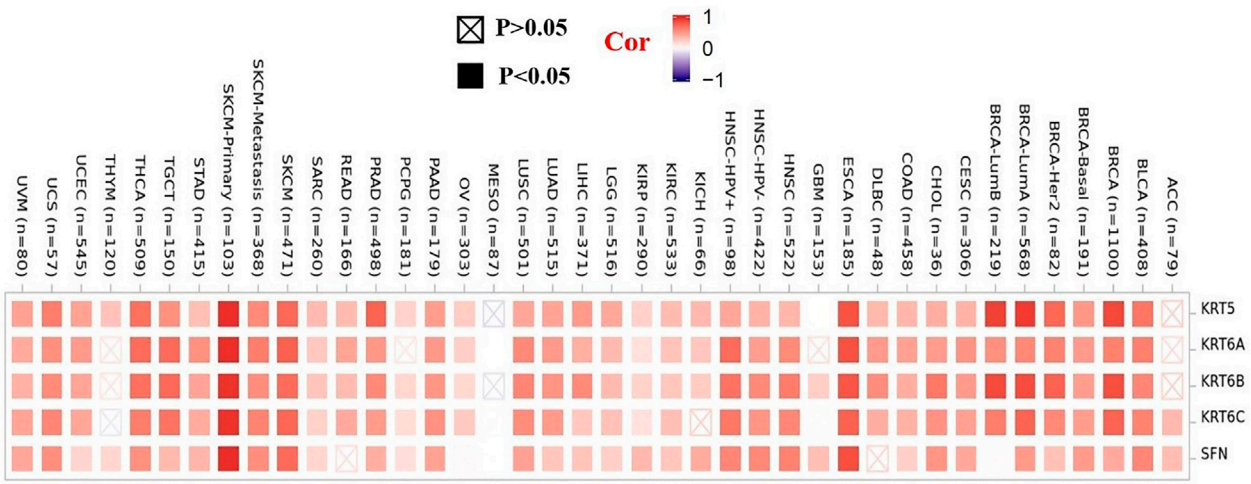


FIGURE 14 | Heatmap data between KRT17 and KRT5, KRT6A, KRT6B, KRT6C, and SFN. The heatmap results showed that the expression of KRT17 was positively correlated with the expression of KRT5, KRT6a, KRT6b, KRT6c, and SFN.

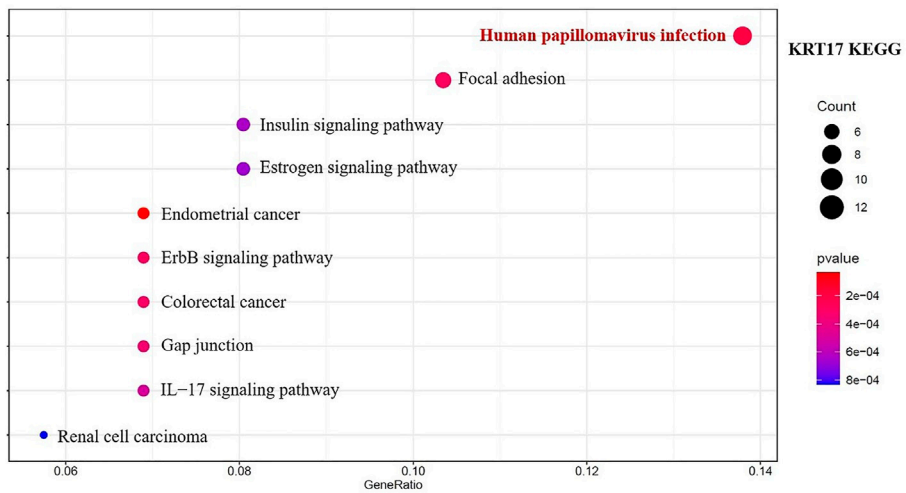


FIGURE 15 | Results of KEGG pathway analysis of KRT17 interaction binding and expression-related genes.

development of tumor diseases, how it affects the survival and mechanism of the disease still needs to be further explored.

Our study also showed that the expression state of KRT17 genetic variations in various tumors, with gene amplification, gene mutation, and gene deletion as the main variation type. Analysis of TCGA–ACC dataset revealed that the KRT17 mutation status of ACC was associated with OS, DSS, and PFS, In ESCA, STAD, and UCEC, a correlation was not found, indicating that tumor development and clinical outcomes are not simply influenced by mutation status but rather by a complex genetic process. Among human diseases, mutations in KRT17 have been reported to be associated with congenital thyroid disease. Among KRT17 mutations, c.275A > G missense mutations that cause asparagine to be replaced by serine (Asn92Ser) are the most common (Cogulu et al., 2009; Ofaiche et al., 2014). In UCEC and COAD, we found that KRT17 translation from G (glycine) to A (alanine) (Gly22Ala) at site 22 is the most common; the change at this point may be the main reason for the occurrence and development of malignant tumors caused by KRT17, but the specific molecular mechanism is still not clear and needs further study. Cancer-related immunology is complex and poorly understood, partly reflected in the diversity of immune responses and the spatial and temporal heterogeneity of developing tumors (Melero et al., 2014). Studies have shown that during the occurrence of cervical cancer, the expression of some inflammatory cytokines and immune cytokines in tumors is significantly dependent on KRT17 (Hobbs et al., 2016b). In addition, KRT17 can activate different macrophage

populations and subtypes (such as M1 and M2) through interferon γ, tumor necrosis factor-α and interleukin 10, which play an important role in tumor proliferation and differentiation (Sica and Mantovani, 2012). Therefore, we used the Quantiseq, Xcell, Mpcounter, and EPIC methods to analyze the relationship between KRT17 expression and the infiltration of macrophages, CD8+T cells, Tregs, and cancer-related fibroblasts. There was a positive correlation between KRT17 expression and macrophage infiltration in LIHC, THCA, THYM, and UVM, but no positive correlation was found in cervical cancer. Considering that the expression of interferon γ, tumor necrosis factor-α and interleukin 10 may be different in different patients with cervical cancer, KRT17 may have little effect on stimulating the production of macrophages through the abovementioned factors. KRT17 expression was negatively associated with Treg infiltration T-regulatory cells in ESCA, HNSC, HNSC- human papilloma virus positive (HNSC-HPV+), and LUSC but positively associated with Treg infiltration in THCA. THCA is an endocrine gland. Considering that the high expression of Tregs may be related to hormones secreted by endocrine glands, studies have shown that Tregs infiltration is greater in THCA and the increase of Tregs tissue infiltration is positively correlated with advanced disease (Gogali et al., 2012). Therefore, the expression of Tregs may be used as a biomarker and prognostic indicator of THCA. This also indicates that there are multiple ways in which KRT17 acts in malignant tumors. In addition, we proposed the relationship between KRT17 and CD8+T cells, Tregs, and cancer-related fibroblasts for the first time, suggesting that KRT17 and immune cells are involved in the formation and development of tumors, However, the specific mechanism of action is still unclear and needs further study.

Next, GO enrichment analysis was performed on the binding proteins interacting with KRT17 and the genes related to KRT17 expression. Through the analysis, we learned that the expression of KRT17 was positively correlated with the expression of KRT5, KRT6a, KRT6b, KRT6c, and SFN. We also found that HPV infection, the estrogen signaling pathway, cadherin binding, and intermediate cytoskeleton filament formation may affect KRT17-mediated tumor pathogenesis and development. Recent studies have shown that KRT17 has a variety of mechanisms of action in malignant tumors and can inhibit tumor cell proliferation, migration, and invasion by regulating the Akt/mTOR pathway, glucose uptake (Khanom et al., 2016), Wnt signaling pathway, epithelial–mesenchymal transition (EMT) (Wang et al., 2019), and mTOR/S6K1 signaling pathway (Li et al., 2020). This suggests that the mechanism of action of KRT17 in malignant tumors is complex and diverse, that there may be multiple mechanisms in one tumor, and that multiple tumors may share similar mechanisms (Li et al., 2019; Liu et al., 2020; Zeng et al., 2020).

To summarize, the correlation between KRT17 expression and clinical prognosis, genetic variation and immune cell infiltration were analyzed for the first time. These results are helpful for understanding the role of KRT17 in tumorigenesis and development and exploring its potential clinical application value, providing useful information for KRT17 research and drug development.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

HQ-Z: Methodology, Software, Formal analysis, Writing–Original Draft, Funding acquisition. Y-Z: Methodology, Visualization, Writing–Original Draft. ZY-F: Software, Formal analysis, Writing–Original Draft. L-L: Formal analysis, Visualization. Y-L: Methodology, Software. YP-C: Methodology, Formal analysis, Review and Editing. YC-L: Conceptualization, Writing–Review and Editing, Supervision, Project administration.

FUNDING

This article was supported by the Science and Technology Foundation Project of Guizhou Health Commission (award number: No. gzwjkj2020-1-031).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.801698/full#supplementary-material

REFERENCES

Chen, P., Shen, Z., Fang, X., Wang, G., Wang, X., Wang, J., et al. (2020). Silencing of Keratin 17 by Lentivirus Mediated Short Hairpin RNA Inhibits the Proliferation of PANC1 Human Pancreatic Cancer Cells. *Oncol. Lett.* 19, 3531–3541. doi:10.3892/ol.2020.11469

Chivu-Economescu, M., Dragu, D. L., Necula, L. G., Matei, L., Enciu, A. M., Bleotu, C., et al. (2017). Knockdown of KRT17 by siRNA Induces Antitumoral Effects on Gastric Cancer Cells. *Gastric Cancer* 20, 948–959. doi:10.1007/s10120-017-0712-y

Chung, B. M., Arutyunov, A., Ilagan, E., Yao, N., Wills-Karp, M., and Coulombe, P. A. (2015). Regulation of C-X-C Chemokine Gene Expression by Keratin 17 and hnRNP K in Skin Tumor Keratinocytes. *J. Cel Biol.* 208, 613–627. doi:10.1083/jcb.201408026

Cogulu, O., Onay, H., Aykut, A., Wilson, N. J., Smith, F. J. D., Dereli, T., et al. (2009). Pachyonychia Congenita Type 2, N92S Mutation of Keratin 17 Gene: Clinical Features, Mutation Analysis and Pathological View. *Eur. J. Pediatr.* 168, 1269–1272. doi:10.1007/s00431-008-0908-6

Depianto, D., Kerns, M. L., Dlugosz, A. A., and Coulombe, P. A. (2010). Keratin 17 Promotes Epithelial Proliferation and Tumor Growth by Polarizing the Immune Response in Skin. *Nat. Genet.* 42, 910–914. doi:10.1038/ng.665

Escobar-Hoyos, L. F., Shah, R., Roa-Peña, L., Vanner, E. A., Najafian, N., Banach, A., et al. (2015). Keratin-17 Promotes p27KIP1 Nuclear Export and

Degradation and Offers Potential Prognostic Utility. *Cancer Res.* 75, 3650–3662. doi:10.1158/0008-5472.CAN-15-0293

Gogali, F., Paterakis, G., Rassidakis, G. Z., Kaltsas, G., Liakou, C. I., Gousis, P., et al. (2012). Phenotypical Analysis of Lymphocytes with Suppressive and Regulatory Properties (Tregs) and NK Cells in the Papillary Carcinoma of Thyroid. *J. Clin. Endocrinol. Metab.* 97, 1474–1482. doi:10.1210/jc.2011-1838

Hobbs, R. P., Jacob, J. T., and Coulombe, P. A. (2016a). Keratins Are Going Nuclear. *Develop. Cel* 38, 227–233. doi:10.1016/j.devcel.2016.07.022

Hobbs, R. P., Batazzi, A. S., Han, M. C., and Coulombe, P. A. (2016b). Loss of Keratin 17 Induces Tissue-specific Cytokine Polarization and Cellular Differentiation in HPV16-Driven Cervical Tumorigenesis *In Vivo*. *Oncogene* 35 (43), 5653–5662. doi:10.1038/onc.2016.102

Ide, M., Kato, T., Ogata, K., Mochiki, E., Kuwano, H., and Oyama, T. (2012). Keratin 17 Expression Correlates with Tumor Progression and Poor Prognosis in Gastric Adenocarcinoma. *Ann. Surg. Oncol.* 19, 3506–3514. doi:10.1245/s10434-012-2437-9

Jacob, J. T., Coulombe, P. A., Kwan, R., and Omary, M. B. (2018). Types I and II Keratin Intermediate Filaments. *Cold Spring Harb. Perspect. Biol.* 10, a018275. doi:10.1101/cshperspect.a018275

Jemal, A., Center, M. M., DeSantis, C., and Ward, E. M. (2010). Global Patterns of Cancer Incidence and Mortality Rates and Trends. *Cancer Epidemiol. Biomarkers Prev.* 19 (8), 1893–1907. doi:10.1158/1055-9965.EPI-10-0437

Khanom, R., Nguyen, C. T. K., Kayamori, K., Zhao, X., Morita, K., Miki, Y., et al. (2016). Keratin 17 Is Induced in Oral Cancer and Facilitates Tumor Growth. *PLoS One* 11 (8), e0161163. doi:10.1371/journal.pone.0161163

Kim, S., Wong, P., and Coulombe, P. A. (2006). A Keratin Cytoskeletal Protein Regulates Protein Synthesis and Epithelial Cell Growth. *Nature* 441, 362–365. doi:10.1038/nature04659

Kurokawa, I., Takahashi, K., Moll, I., and Moll, R. (2011). Expression of Keratins in Cutaneous Epithelial Tumors and Related Disorders - Distribution and Clinical Significance. *Exp. Dermatol.* 20 (3), 217–228. doi:10.1111/j.1600-0625.2009.01006.x

Li, J., Chen, Q., Deng, Z., Chen, X., Liu, H., Tao, Y., et al. (2019). KRT17 Confers Paclitaxel-Induced Resistance and Migration to Cervical Cancer Cells. *Life Sci.* 224, 255–262. doi:10.1016/j.lfs.2019.03.065

Li, D., Ni, X.-F., Tang, H., Zhang, J., Zheng, C., Lin, J., et al. (2020). KRT17 Functions as a Tumor Promoter and Regulates Proliferation, Migration and Invasion in Pancreatic Cancer via mTOR/S6k1 Pathway. *Cancer Manag. Res.* 12, 2087–2095. doi:10.2147/CMAR.S243129

Liu, J., Liu, L., Cao, L., and Wen, Q. (2018). Keratin 17 Promotes Lung Adenocarcinoma Progression by Enhancing Cell Proliferation and Invasion. *Med. Sci. Monit.* 24, 4782–4790. doi:10.12659/MSM.909350

Liu, Z., Yu, S., Ye, S., Shen, Z., Gao, L., Han, Z., et al. (2020). Keratin 17 Activates AKT Signalling and Induces Epithelial-Mesenchymal Transition in Oesophageal Squamous Cell Carcinoma. *J. Proteomics* 211, 103557. doi:10.1016/j.jprot.2019.103557

Lo, B. K. K., Yu, M., Zloty, D., Cowan, B., Shapiro, J., and McElwee, K. M. (2010). CXCR3/ligands Are Significantly Involved in the Tumorigenesis of Basal Cell Carcinomas. *Am. J. Pathol.* 176, 2435–2446. doi:10.2353/ajpath.2010.081059

McGowan, K. M., and Coulombe, P. A. (1998). Onset of Keratin 17 Expression Coincides with the Definition of Major Epithelial Lineages during Skin Development. *J. Cel Biol.* 143, 469–486. doi:10.1083/jcb.143.2.469

Melero, I., Gaudernack, G., Gerritsen, W., Huber, C., Parmiani, G., Scholl, S., et al. (2014). Therapeutic Vaccines for Cancer: an Overview of Clinical Trials. *Nat. Rev. Clin. Oncol.* 11, 509–524. doi:10.1038/nrclinonc.2014.111

Merkin, R. D., Vanner, E. A., Romeiser, J. L., Shroyer, A. L. W., Escobar-Hoyos, L. F., Li, J., et al. (2017). Keratin 17 Is Overexpressed and Predicts Poor Survival in Estrogen Receptor-Negative/human Epidermal Growth Factor Receptor-2-Negative Breast Cancer. *Hum. Pathol.* 62, 23–32. doi:10.1016/j.humpath.2016.10.006

Mikami, T., Maruyama, S., Abé, T., Kobayashi, T., Yamazaki, M., Funayama, A., et al. (2015). Keratin 17 Is Co-expressed with 14-3-3 Sigma in Oral Carcinoma *In Situ* and Squamous Cell Carcinoma and Modulates Cell Proliferation and Size but Not Cell Migration. *Virchows Arch.* 466, 559–569. doi:10.1007/s00428-015-1735-6

Ofaiche, J., Duchatelet, S., Fraitag, S., Nassif, A., Nougué, J., and Hovnanian, A. (2014). Familial Pachyonychia Congenita with Steatocystoma Multiplex and Multiple Abscesses of the Scalp Due to the p.Asn92Ser Mutation in Keratin 17. *Br. J. Dermatol.* 171, 1565–1567. doi:10.1111/bjd.13123

Pan, X., Hobbs, R. P., and Coulombe, P. A. (2013). The Expanding Significance of Keratin Intermediate Filaments in normal and Diseased Epithelia. *Curr. Opin. Cel Biol.* 25, 47–56. doi:10.1016/j.ceb.2012.10.018

Proby, C. M., Churchill, L., Purkis, P. E., Glover, M. T., Sexton, C. J., and Leigh, I. M. (1993). Keratin 17 Expression as a Marker for Epithelial Transformation in Viral Warts. *Am. J. Pathol.* 143, 1667–1678.

Sica, A., and Mantovani, A. (2012). Macrophage Plasticity and Polarization: *In Vivo* Veritas. *J. Clin. Invest.* 122, 787–795. doi:10.1172/JCI59643

Ta, H. D. K., Wang, W.-J., Phan, N. N., An Ton, N. T., Anuraga, G., Ku, S.-C., et al. (2021). Potential Therapeutic and Prognostic Values of LSM Family Genes in Breast Cancer. *Cancers* 13, 4902. doi:10.3390/cancers13194902

Torre, L. A., Siegel, R. L., Ward, E. M., and Jemal, A. (2016). Global Cancer Incidence and Mortality Rates and Trends-An Update. *Cancer Epidemiol. Biomarkers Prev.* 25 (1), 16–27. doi:10.1158/1055-9965.EPI-15-0578

Ujiie, D., Okayama, H., Saito, K., Ashizawa, M., Thar Min, A. K., Endo, E., et al. (2020). KRT17 as a Prognostic Biomarker for Stage II Colorectal Cancer. *Carcinogenesis* 41, 591–599. doi:10.1093/carcin/bgz192

Wang, Z., Yang, M.-Q., Lei, L., Fei, L.-R., Zheng, Y.-W., Huang, W.-J., et al. (2019). Overexpression of KRT17 Promotes Proliferation and Invasion of Non-small Cell Lung Cancer and Indicates Poor Prognosis. *Cancer Manag. Res.* 11, 7485–7497. doi:10.2147/CMAR.S218926

Yan, X., Yang, C., Hu, W., Chen, T., Wang, Q., Pan, F., et al. (2020). Knockdown of KRT17 Decreases Osteosarcoma Cell Proliferation and the Warburg Effect via the AKT/mTOR/HIF1α Pathway. *Oncol. Rep.* 44, 103–114. doi:10.3892/or.2020.7611

Yang, L., Zhang, S., and Wang, G. (2019). Keratin 17 in Disease Pathogenesis: from Cancer to Dermatoses. *J. Pathol.* 247 (2), 158–165. doi:10.1002/path.5178

Yao, Z., Zhang, Y., Xu, D., Zhou, X., Peng, P., Pan, Z., et al. (2019). Research Progress on Long Non-Coding RNA and Radiotherapy. *Med. Sci. Monit.* 25, 5757–5770. doi:10.12659/MSM.915647

Zeng, Y., Zou, M., Liu, Y., Que, K., Wang, Y., Liu, C., et al. (2020). Keratin 17 Suppresses Cell Proliferation and Epithelial-Mesenchymal Transition in Pancreatic Cancer. *Front. Med.* 7 (7), 572494. doi:10.3389/fmed.2020.572494

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 | Frontiers in Genetics

ORIGINAL RESEARCH
published: 26 July 2022
doi: 10.3389/fgene.2022.764957

NCKAP1 is a Prognostic Biomarker for Inhibition of Cell Growth in Clear Cell Renal Cell Carcinoma

Jiasheng Chen[1], *Jianzhang Ge*[2], *Wancong Zhang*[1], *Xuqi Xie*[1], *Xiaoping Zhong*[1]* and *Shijie Tang*[1]*

[1]Department of Burns and Plastic Surgery, The Second Affiliated Hospital of Shantou University Medical College, Shantou, China, [2]Department of Urology, Changsha Central Hospital, Changsha, China

Background: Clear cell renal cell carcinoma (ccRCC) is the most frequent type of kidney cancer. Nck-associated protein 1 (NCKAP1) is associated with poor prognosis and tumor progression in several cancer types, but the function and prognostic value of NCKAP1 in ccRCC remain poorly understood.

Methods: Using the Ualcan database, we evaluated the correlation between NCKAP1 expression and clinical features of ccRCC. These data were validated by immunohistochemical staining for NCKAP1 in a cohort of ccRCC patients. We assessed the prognostic value of NCKAP1 using GEPIA2 survival analysis. NCKAP1 function was characterized *in vitro* and *in vivo* using NCKAP1-overexpression ACHN cell lines. The LinkedOmics and GSCALite databases were used to investigate identify potential NCKAP1-targeted medicines that may play a role in the treatment of ccRCC. The impact of NCKAP1 expression on immune infiltration was also evaluated.

Results: NCKAP1 was significantly downregulated in ccRCC and correlated with advanced clinicopathological features and poor prognosis. Overexpression of NCKAP1 in ACHN cells reduced proliferation, invasion and migration capacity *in vitro* and inhibited tumor growth *in vivo*. According to the LinkedOmics, GSCALite and TIMER databases, NCKAP1 and related genes function primarily in ribosomal signaling, oxidative phosphorylation, TGF-β, and EMT-related signaling pathways. NCKAP1 was also shown to positively correlate with immune cell types, biomarkers, and immune checkpoints in ccRCCs.

Conclusions: NCKAP1 may play a vital tumor-suppressive role in ccRCC and is potentially a useful prognostic biomarker.

Keywords: clear cell renal cell carcinoma, Nckap1, biomarkers, prognosis, progression

OPEN ACCESS

Edited by:
Jesús Espinal-Enríquez, Instituto Nacional de Medicina Genómica (INMEGEN), Mexico

Reviewed by:
Subhradip Karmakar, All India Institute of Medical Sciences, India
Xiaofu Qiu, Guangdong Second Provincial General Hospital, China

***Correspondence:**
Xiaoping Zhong
zhongxiaoping6@foxmail.com
Shijie Tang
sjtang3@stu.edu.cn

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

Received: *26 August 2021*
Accepted: *07 June 2022*
Published: *26 July 2022*

Citation:
Chen J, Ge J, Zhang W, Xie X, Zhong X and Tang S (2022) NCKAP1 is a Prognostic Biomarker for Inhibition of Cell Growth in Clear Cell Renal Cell Carcinoma. Front. Genet. 13:764957. doi: 10.3389/fgene.2022.764957

INTRODUCTION

Renal cell carcinoma (RCC) is a malignant tumor of the urinary system and accounts for about 3% of cancers worldwide (Kotecha et al., 2019). Based on the World Health Organization (WHO) classification system, RCC in adults is classified into four types: clear cell, papillary, pigmented, and collecting duct type, of which clear cell RCC (ccRCC) is the most common type worldwide (Patel et al., 2012). Over the past few decades, the incidence of RCC has increased by 2% per year, due in

part to the difficulty of early diagnosis, and approximately 25% of patients present with metastatic disease (Perazella et al., 2018; Capitanio et al., 2019). There is a tremendous unmet need for developing novel diagnostic biomarkers and therapeutic targets that have the potential to improve the prognosis of RCC patients.

NCK-associated protein 1 (NCKAP1) is a protein found in sporadic Alzheimer's disease (AD) as part of the WAVE complex along with ABI1-2, BRK1, CYFIP1-2, and WASF1-2 proteins (Suzuki et al., 2000; Innocenti et al., 2004). NCKAP1 regulates various intracellular processes such as apoptosis, migration, and invasion, and plays an essential role in disease pathogenesis (Whitelaw et al., 2020). NCKAP1 expression is highly tissue-specific, and its expression has been found in colon, breast, and lung cancers (Teng et al., 2016; Xiong et al., 2019a; Rai et al., 2020; Zhang et al., 2020). On the other hand, we previously showed that downregulation of NCKAP1 in liver cancer patients is associated with poor prognosis (Zhong et al., 2019). These data suggest that NCKAP1 may have tumor-promoting or suppressive effects on certain types of cancer. However, the clinicopathological characteristics and function of NCKAP1 in ccRCC have not yet been confirmed.

In this study, we aimed to determine the function of NCKAP1 in ccRCC using bioinformatics analysis portal tools and immunohistochemical validation, to examine the relationship between NCKAP1 expression and clinicopathological features of ccRCC, and to determine the *in vitro* and *in vivo* NCKAP1 expression was measured to characterize the clinicopathological features of ccRCC. We also determined the *in vitro* and *in vivo* role of NCKAP1 in a related cell line (ACHN). In addition, the predicted functions of NCKAP1 and tumor immune infiltrating cells were discussed.

MATERIALS AND METHODS

Patients and Samples

Shanghai Xinchao Biotechnology (Shanghai, China) was the commercial source of renal cell carcinoma tissue microarray (TMA). The staging of tumors was done employing the 2010 revised TNM system and the TMA consisted of stages I-II disease ($n = 52$) and stage III-IV ($n = 23$) disease. The WHO criteria were utilized to specify the histological grades of tumors as mentioned, low grade (Grade I and II; $n = 55$) and high grade (Grade III and IV; $n = 20$).

Bioinformatics Analysis

NCKAP1 mRNA, protein expression and the associated clinical features were examined in ccRCC using the UALCAN database (Chandrashekar et al., 2017) (http://ualcan.path.uab.edu/). Gene Expression Profiling Interactive Analysis (GEPIA) (Tang et al., 2017) (http://gepia.cancer-pku.cn/index.html) was used to analyze the survival information between NCKAP1 and ccRCC. The LinkedOmics database (Vasaikar et al., 2018) (www.linkedomics.org) was used to analyze the genes that significantly correlated with NCKAP1, GO enrichment, KEGG pathways, kinase targets and miRNA targets in ccRCC. GSCALite (www.bioinfo.life.hust.edu.cn/web/GSCALite/) was used to analyze and visualize the gene sets correction with pathway activity in our study with TCGA ccRCC sample. TIMER (Li et al., 2017) (www.cistrome.shinyapps.io/timer/) was employed to probe components of tumor immune cell characteristics. Immune Checkpoints, TMB, MSI R package were implemented by R foundation for statistical computing (2020) version 4.0.3 and software packages ggplot2 and pheatmap (Thorsson et al., 2019).

Immunohistochemical Analysis

Immunohistochemical (IHC) staining of NCKAP1,Ki-67 and E-cad were carried out as previously reported (Zhong et al., 2019). The TMA and xenograft tumor sections were incubated with the NCKAP1, Ki-67, E-cad antibody (1:100; Proteintech, China). Stained sections were then independently assessed by two pathologists.

Cell Culture and Transfections

The American Type Culture Collection (ATCC, Manassas, VA, United States) was the source of human RCC cell lines ACHN, 786-O, and 769-P. The OS-RC-2 cell line was received as a gift from the Cancer Research Center of Shantou University Medical College (Shantou, China). Culture of the cell lines was done in Roswell Park Memorial Institute medium (RPMI-1640, Gibco, Gaithersburg, MD) supplemented with 10% fetal bovine serum (FBS; Gibco, United States) and maintained in an atmosphere of 5% CO_2 atmosphere at 37°C.

Hanbio Biotechnology Co., Ltd. (Shanghai, China) was the commercial source of alentiviral NCKAP1 overexpression vector and an empty vector. ACHN cells were transfected using Lipofectamine 2000 and Opti-MEMI (Gibco, United States) in accordance with the prescribed protocol of the manufacturer.

qRT-PCR Analysis

Trizol reagent (Tiangen Biotech, China) was employed to extract the total RNA and cDNA synthesis performed using a Revert Aid First Strand cDNA Synthesis Kit (Thermo Scientific, United States). qRT-PCR analysis for the expression of NCKAP1 was performed in triplicate using SYBR Green I (Tiangen Biotech, China) in accordance with the prescribed protocol of the manufacturer. The internal control was GAPDH. The primer sequences for NCKAP1 and GAPDH were as follows:

NCKAP1
5′-TCCTAAATACTGACGCTACAGCA-3′(forward)
5′-GCCTCCTTGCATTCTCTTATGTC-3′(reverse)
GAPDH
5′-GTCTCCTCTGACTTCAACAGCG-3′(forward)
5 ′-ACCACCCTGTTGCTGTAGCCAA-3′(reverse)

Western Blotting

Proteins were isolated using a whole-cell lysis assay (Beyotime Biotechnology, Jiangsu, China). Following the resolution of the protein samples by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) transfer to polyvinylidene difluoride (PVDF) membranes was done for western blotting.

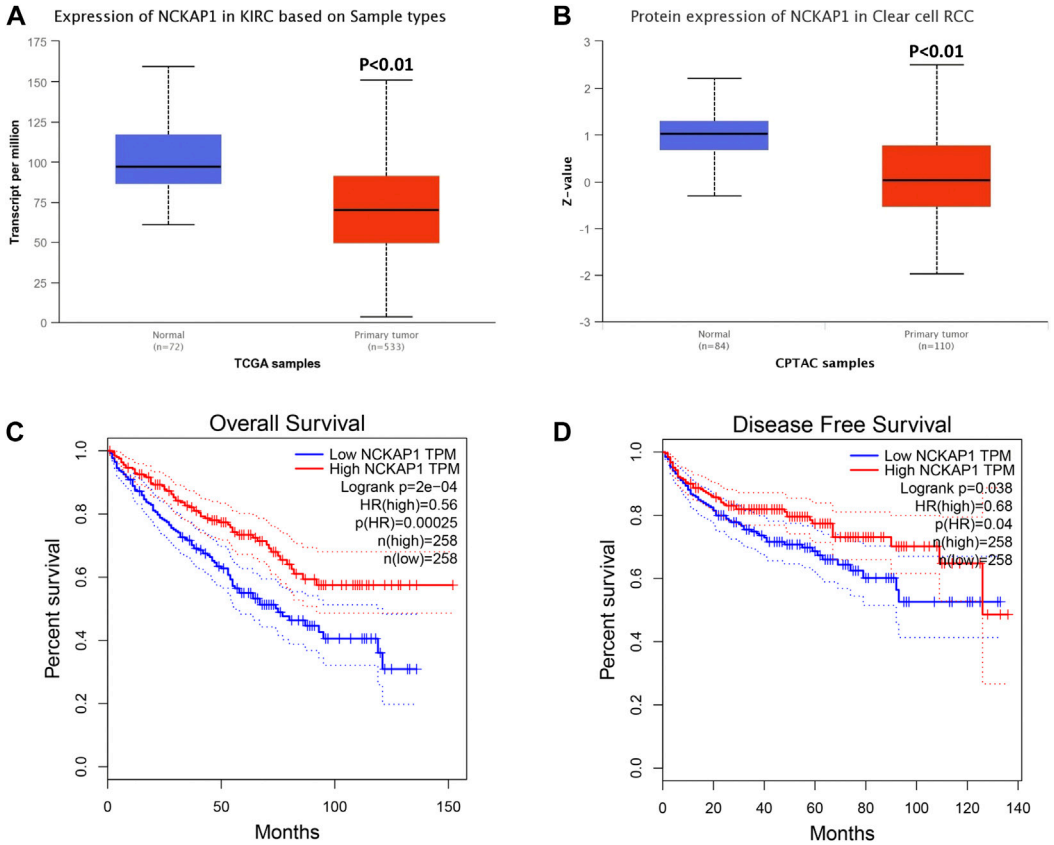


FIGURE 1 | Analysis of NCKAP1 expression and survival curve in ccRCC on the basis of the UALCAN database and GEPIA survival analysis platform. **(A)** mRNA expression and **(B)** protein expression of NCKAP1 in ccRCC and normal tissues. The overall survival **(C)** and disease-free survival **(D)** curve of NCKAP1 in ccRCC. *$p <$ 0.05 is statistically significant.

Membranes were blocked with skimmed milk followed by NCKAP1 primary antibody (1:1,000; Proteintech, China) overnight incubation at 4°C. Subsequent secondary antibody (1:5,000; Abcam) was done at room temperature following membrane washing. Blots were developed using a chemiluminescence detection kit. GAPDH (1:1,000; Proteintech, China) was used as the loading control.

Cell Viability Assay

A Cell Counting Kit-8 (Beyotime Biotechnology, Jiangsu, China) was employed to quantify the cell viability in accordance with the prescribed approach. Briefly, 96 well plates were used to seed 2×10^3 cells followed by 5 days of culture under normal conditions. At the end of each period (1, 2, 3, 4 or 5 days), incubation of the plates was done at 37°C for 2 hours following the addition of 10 µL of CCK-8 to each well. Absorbance values measured at 540 nm to quantify the cell viability. Triplicate assays were conducted and repeated three times.

Colony Formation Assay

6 well plates were employed to seed 1×10^3 cells followed by incubation at 37°C. After 2 weeks of culture, 4% paraformaldehyde was utilized to fix colonies and subsequent crystal violet (0.5%) staining at room temperature. Enumeration of the colonies was done utilizing digital images of the well obtained from each of the three replicate wells.

Wound Healing Assay

This was done to evaluate the ability of the cells to migrate. Briefly, 24 well plates were utilized to seed 5×10^5 cells and allowed to adhere overnight. Cells grew to confluence and then artificial wounds were introduced using to mark a line down the center of the cell layer. Cells were then cultured in serum-free medium. The wounded areas were imaged right away (0 h) and at 24 h after the wound was introduced using an inverted microscope (Olympus Corp). Triplicate experiments were done.

Transwell Assay

In order to quantify cell invasion this assay involved the coating of Transwellinserts (8 µm pore size, Corning, NY, United States) with matrigel (BD Biosciences, NJ, United States) and 5×10^4 cells were added into the upper compartment. RPMI-1640 containing 20% FBS was then added to the lower chamber of the transwell and the cells followed by a 24-h incubation. Post-migration of cells to the lower chamber from the upper chamber, the membranes were then stained and the migrated cells were

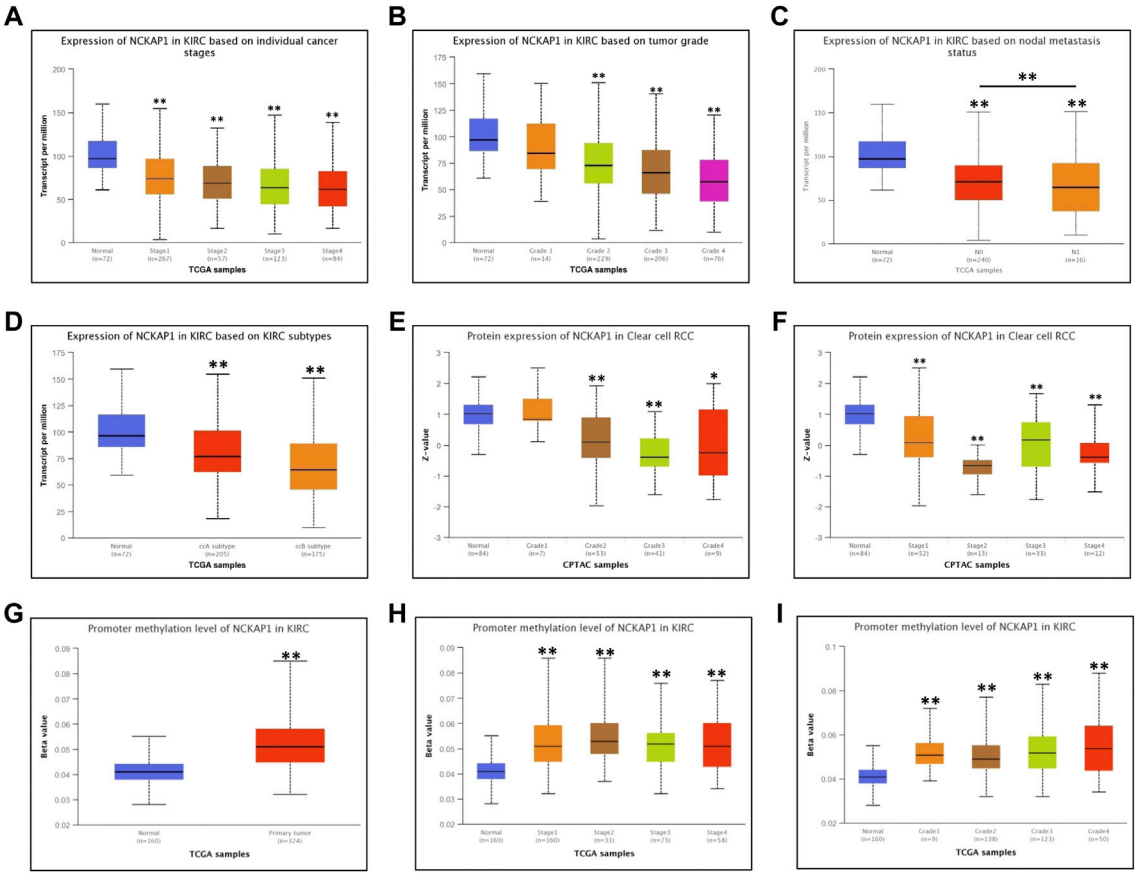


FIGURE 2 | Correlation between NCKAP1 expression and clinicopathologic characteristics in ccRCC tissues. mRNA expression of NCKAP1 in ccRCC sub-groups based on individual cancer stages **(A)**, tumor grade **(B)**, node metastasis **(C)** and ccRCC subtypes **(D)**; protein expression of NCKAP1 in ccRCC sub-groups based on individual cancer stages **(E)** and tumor grade **(F)**. Promoter methylation level of NCKAP1 in normal tissues and ccRCC **(G)**; Promoter methylation level of NCKAP1 in ccRCC of individual tumor stage and tumor grade **(H,I)**. *$p < 0.05$ and **$p < 0.001$ are statistically significant.

enumerated. The number of cells was scored from five randomly selected fields of view on the lower membrane. The assay was performed in triplicate.

Animal Experiments

Male BALB/c athymic nude mice (4–6 weeks old) were purchased from Hunan SJA Laboratory Animal Co., Ltd. 3.0×10^6 cells transfected with ACHN-NCKAP1 or ACHN-vector were injected subcutaneously into mice to set up the ccRCC xenograft model. When tumors were palpable, their sizes were measured every 3 days for 14 days. 2 weeks post-monitored, the sacrifice of both groups of animals was done and tumors were isolated for growth and IHC analyses. Tumor dimensions were gauged employing calipers and the volumes were calculated utilizing the expression V = (shorter diameter^2 × longer diameter)/2. These experiments were approved by the Ethics Committee of Shantou University Medical College.

Statistical Analysis

The outcomes of the survival curve, GEPIA databases are represented by the HR and p or the COX P-values of a log-rank test. Assessment of the correlation of gene expression was done using the LinkedOmics, GSCALite and TIMER databases and compared with Pearson Correlation analysis. The Pearson χ^2 test was utilized for quantifying the correlation between the expression of NCKAP1 and the patient clinic-pathological parameters. Other data were statistically evaluated using a Student's t-test. Significance was at P-values of < 0.05.

RESULTS

NCKAP1 Expression is Significantly Decreased in ccRCC and Correlated With Patient Outcomes

Using the UALCAN database, we evaluated NCKAP1 mRNA and protein expression in ccRCC. We found that NCKAP1 mRNA expression was decreased in ccRCC tissues compared to normal tissues, consistent with NCKAP1 protein expression data (**Figures 1A,B**). To investigate the

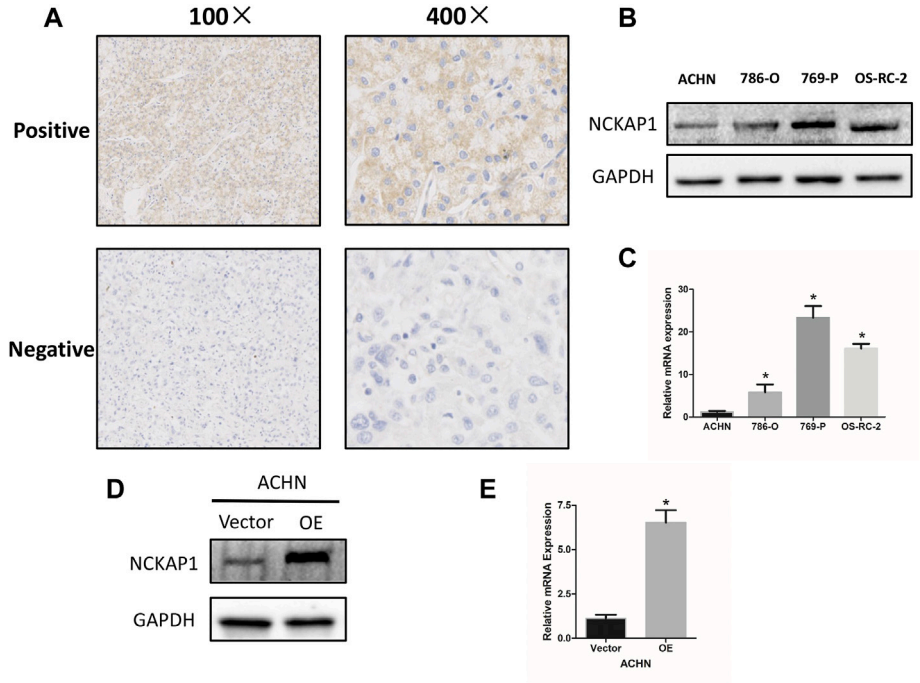


FIGURE 3 | NCKAP1 expression in ccRCC tissues and cell lines. **(A)** Positive (upper) and negative (lower) expressions of NCKAP1in ccRCC. (×100 left, ×400 right). **(B)**Western blotting and **(C)** Quantitative real-time PCR (qPCR) results show that ACHN cells exhibited low expression compared to that of 786-O, 769-P and OS-RC-2 cells. GAPDH was used as a loading control. Overexpression of NCKAP1 (OE) in a transfected ACHN cell line verified by western blotting **(D)** and qPCR **(E)** compared to that of ACHN cells transfected with the control vector (Vector). GAPDH was used as a loading control.

TABLE 1 | Relationship between NCKAP1 expression and clinicopathological factors in ccRCC.

Characteristics	Number	Negative	Positive	P Value
n	75	53	22	
Age				0.144
≦55	28	17	11	
〉55	47	36	11	
Gender				0.525
Male	47	32	15	
Female	28	21	7	
Position				0.776
Left	36	26	10	
Right	39	27	12	
TNM Stage				0.039
I + II	52	33	19	
III + IV	23	20	3	
Tumor Size				0.315
≦7 cm	41	27	14	
〉7 cm	34	26	8	
Grade				0.027
I + II	55	35	20	
III + IV	20	18	2	

prognostic value of NCKAP1, we used the GEPIA database to determine NCKAP1 OS and DFS between mRNA expression and ccRCC were analyzed. **Figures 1C,D** show that low expression of NCKAP1 mRNA may indicate worse OS and DFS in ccRCC.

Correlation of NCKAP1 Expression With Clinical Features of ccRCC

Using the UALCAN database, we examined the correlation between NCKAP1 expression and various clinicopathological features of ccRCC and found that NCKAP1 mRNA expression was significantly associated with tumor grade, TNM stage, and lymph node metastasis (**Figures 2A–C**). These data were similar to the results observed for NCKAP1 protein expression (**Figures 2E,F**); NCKAP1 mRNA expression was lower in advanced cancers compared to early-stage cancers (ccA subtype vs. ccB subtype; $p < 0.001$). Furthermore, methylation levels of the NCKAP1 promoter were increased in ccRCC compared to normal tissue (**Figure 2G**), confirming that this was the opposite of NCKAP1 expression. ccRCC stages, tumor grades 1, 2, 3, and 4 also showed elevated NCKAP1 promoter methylation levels (**Figures 2H,I**).

To confirm the association between NCAKP1 expression and the clinical features in ccRCC, IHC staining of tissue microarrays was performed. Division of the 75 specimens on the TMA was done into a negative NCKAP1 expression group ($n = 53$) and a positive NCKAP1 expression group ($n = 22$) (**Figure 3A**). The correlations between NCKAP1 expression and clinical features are summarized in **Table 1**. A significant association was observed between low NCKAP1 expression and TNM stage, tumor size and pathological grade, whereas NCKAP1 expression displayed no apparent associations with age, gender and tumor position.

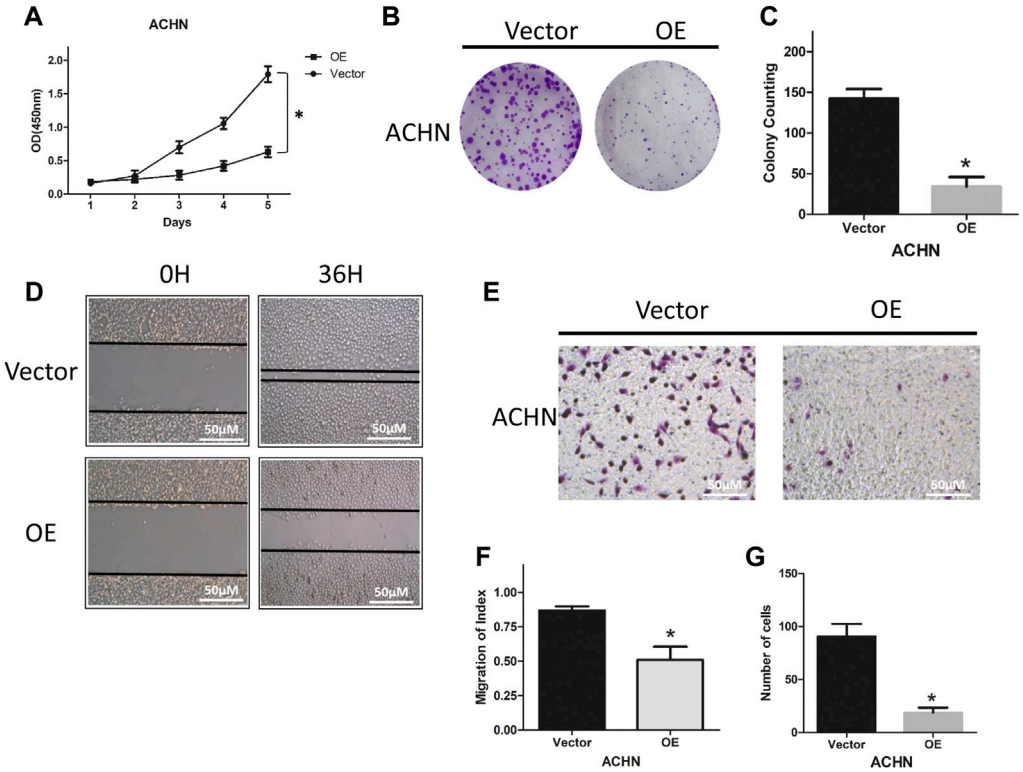


FIGURE 4 | NCKAP1 inhibited cell growth, migration and invasion *in vitro*. CCK-8 assay results showed the cell viability in **(A)** ACHN-OE cells compared with ACHN-Vector cells. Cell colony formation assay showed a statistically significant decrease of **(B,C)** ACHN-OE cells compared with ACHN-Vector cells. Wound-healing assay results showed a statistically significant decrease of migration **(D,F)** ACHN-OE cells compared with ACHN-Vector cells. Scale bar = 50 μm. Transwell invasion assay results showed a statistically significant decrease of invaded **(E,G)** ACHN-OE cells compared with ACHN-Vector cells. Scale bar = 50 μm. The results are mean ± SD values from three independent experiments, $^*p < 0.05$.

NCKAP1 Expression in ccRCC Cell Lines and Overexpression of NCKAP1 in ACHN Cells

A significant association was found between low expression of NCKAP1 and clinicopathological features, suggesting that NCKAP1 may be very much involved in ccRCC tumorigenesis. We examined the expression of NCKAP1 in the ccRCC cell lines described above and found that ACHN cells had significantly lower levels of NCKAP1 expression at the mRNA and protein levels relative to other cell lines (**Figures 3B,C**). Next, for further functional verification, an overexpression plasmid (pEZ-Lv201-NCKAP1) or control vector (pEZLv201) was transfected into ACHN cells. After transfection, qRT-PCR and Western blotting were employed to confirm the expression levels of NCKAP1 mRNA and protein, respectively (**Figures 3D,E**).

Overexpression of NCKAP1 Targets the Ability of ACHN Cells to Proliferate, Migrate, and Invade

To explore the functional role of NCKAP1, we characterized the ability of ACHN cells overexpressing NCKAP1 to proliferate, migrate, and invade. As shown in **Figures 4A–C**, overexpression of NCKAP1 in ACHN cells suppressed the cell growth rate and reduced the number and size of colonies formed. The number and size of colonies formed were reduced compared to the vector control group. Overexpression of NCKAP1 also reduced wound closure (**Figures 4D,F**) and invasiveness (**Figures 4E,G**) relative to vector controls. These data indicate a potentially significant effect of NCKAP1 on the ability of ACHN cells to proliferate, migrate, and invade.

Overexpression NCKAP1 inhibits RCC Progression *in vivo*

We investigated the impact of NCKAP1 overexpression on the tumor growth properties of ACHN cells by establishing a xenograft model established in nude mice. We found that overexpression of NCKAP1 inhibited the formation of tumors compared to vector cells (**Figure 5A**). The growth rate and weights in the ACHN-OE mice were distinctly lower against the control group (**Figures 5B,C**, $p < 0.05$) indicating that NCKAP1 can suppress the growth of RCC tumor xenografts. Morever, in contrast with the NCKAP1-Vector, the NCKAP1-OE group significantly decreased expression of Ki-67 and E-cad in tumor tissues (**Figure 5B**). Taken together, these results demonstrated that NCKAP1 inhibited the growth and metastasis of ACHN tumor.

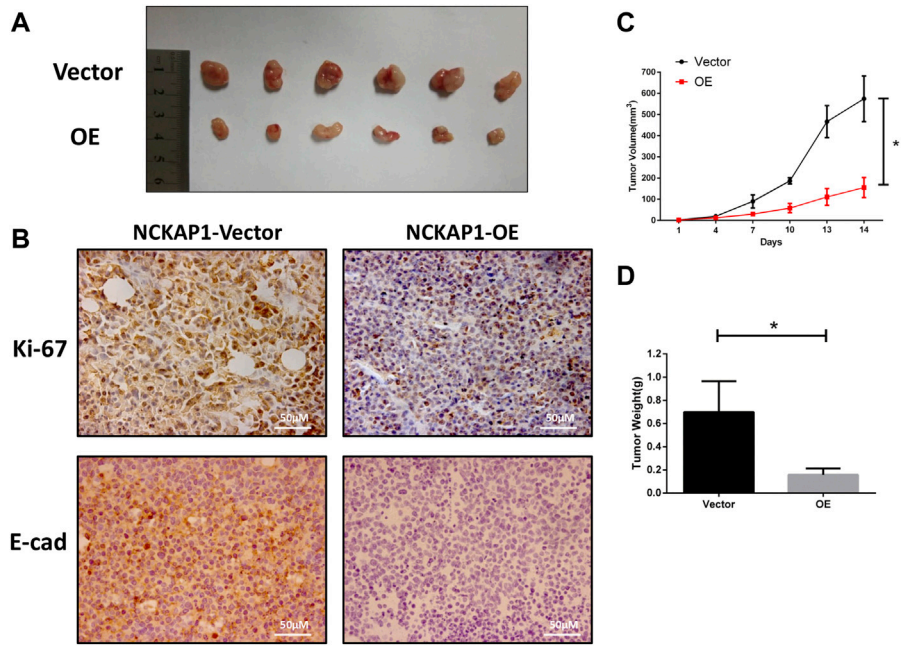


FIGURE 5 | NCKAP1 suppressed RCC progression *in vivo*. **(A)** Representative Data showed that NCKAP1-OE significantly inhabited tomor growth in nude mice xenograft. **(B)** Immunohistochemistry (IHC) staining showed that the expression of Ki-67 and E-CAD differed in tumor tissues. (Scale bar = 50 μm) **(C)** Tumor volume and **(D)** tumor weight was decreased significantly in NCKAP1-OE cells' mice model, $^*p < 0.05$.

Enrichment Analysis of NCKAP1-Related Co-expressed Genes in ccRCC

To explore potential molecular mechanisms of NCKAP1 in ccRCC, we investigated NCKAP1-related coexpressed genes and determined their enrichment function in LinkedOmics using data from 533 TCGA ccRCC patients. Volcano plots showing these up- and down-regulated genes are shown in **Figure 6A**; the top 50 genes positively and negatively associated with NCKAP1 are shown as heat maps in **Figures 6B,C**. As shown in **Supplementary Figure S1**, STAM2 (Cor = 0.7816, p = 6.515e-111), ATF2 (Cor = 0.7585, p = 8.011e-101) and LANCL1 (Cor = 0.7568, p = 3.944e-100) showed strong positive correlation with NCKAP1; IRF3 (Cor = −0.7433, p = 8.929e-95), OGFR (Cor = −0.7132, p = 5.77e-84) and C9orf142 (Cor = −0.7126, p = 8.602e-84) showed strong negative correlation with NCKAP1

GO Enrichment Analysis Was Performed to Explore the Cellular Components, Biological Processes, and Molecular Functions of NCKAP1

The results (**Figures 6D–F**) revealed that genes significantly correlated with NCKAP1 are primarily involved in the liquid immune response, ribosomal NADGH dehydrogenase, mitochondrial respiratory chain complex assembly, and ATPase activity. KEGG pathway analysis (**Figure 6G**) revealed that NCKAP1 is involved in the enriched in ribosomal signaling, oxidative phosphorylation, TGF-β signaling pathway, cytoplasmic DNA-sensing pathway, and IgA synthesis immune network in the intestine. In addition, NCKAP1 and the top three significant positive or negative genes, including STAM2, ATF2, LANCL1, IRF3, OGFR, and C9orf142, were selected as hub genes for pathway analysis by the GSCALite platform. We explored the role of hub genes in well-known cancer-related pathways including TSC/mTOR, RTK, RAS/MAPK, PI3K/AKT, hormone ER, hormone AR, EMT, DNA damage response, cell cycle, and apoptosis pathway. We found that NCAKP1 is involved in the activation of PI3K/AKT, RTK, Hormone ER, RAS/MAPK pathway, and Hormone AR, DNA Damage Response, Cell Cycle pathway, whereas NCAKP1 is involved in the inhibition of EMT and TSC/mTOR pathways. Furthermore, by using GEPIA2 database, NCKAP1 expression was found to be significantly positively related to the expression of EMT signaling genes VIM (p = 3.9e-07), E-cad (p = 3.2e-05), and N-cad (p = 0) in the TCGA-KIRC cohort (**Supplementary Figure S4**). These results are shown in **Supplementary Figure S2, S4**.

Regulatory Factor NCKAP1 Network in ccRCC

The NCKAP1 network of kinase targets in ccRCC was examined. As shown in **Supplementary Table S1**, the top five kinases are mainly Ataxia Telangiectasia Mutated (ATM), Ribosomal Protein S6 Kinase B1 (RPS6KB1), cyclin-dependent kinase 1 (CDK1), cyclin-dependent kinase 5 (CDK5) and Large Tumor Suppressor Kinase 1 (LATS1). We also explored potential miRNA targets of NCKAP1 in ccRCC (**Supplementary Table S1**). The top five miRNA targets were identified as (ATGTTAA) MIR-302C, (CTTGTAT) MIR-381, (ATAGGAA) MIR-202, (GTATTAT) MIR-369–3P and (ATATGCA) MIR-448.

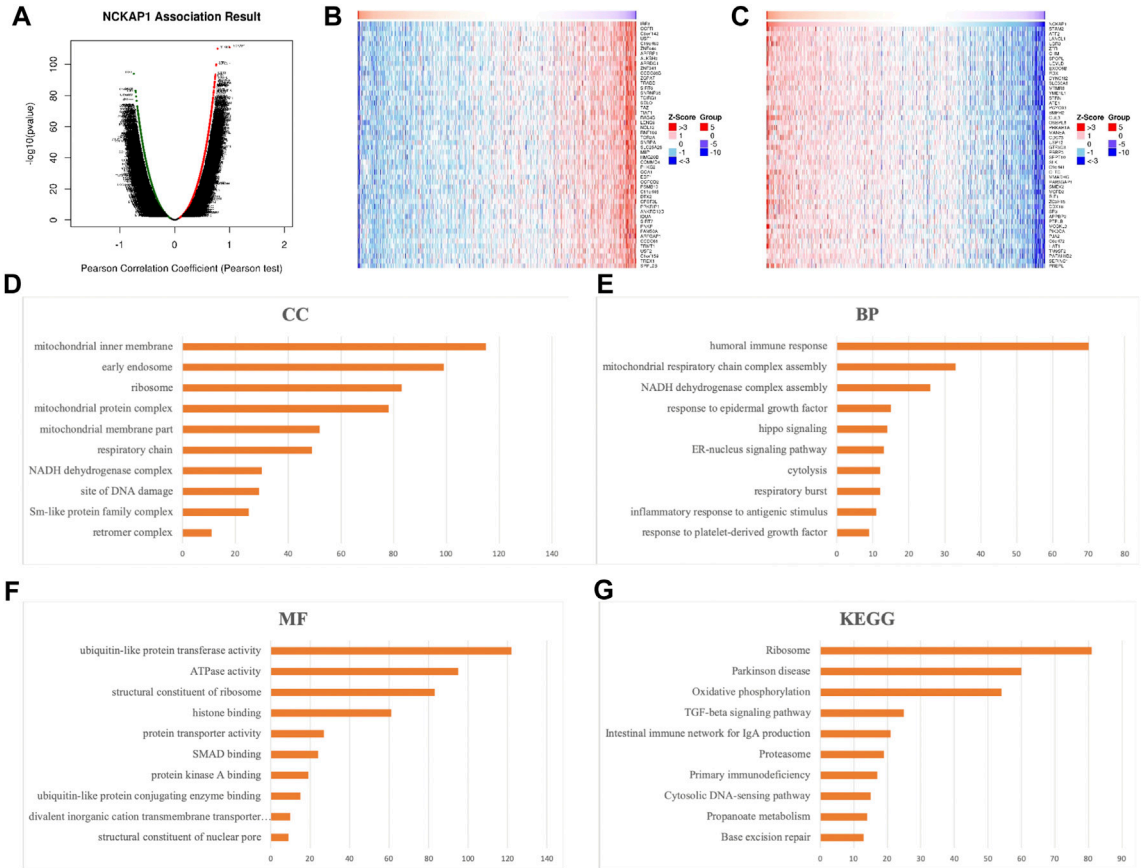


FIGURE 6 | Significant genes correlated with NCKAP1 and enrichment analysis in RCC (LinkedOmics). **(A)** The positive and negative genes correlated with NCKAP1 in ccRCC (Pearson test). **(B)** The top 50 positive genes and **(C)** the top negative genes correlated with NCKAP1 in ccRCC. Red indicated positively correlated genes, while green indicated negatively correlated genes. **(D)** Cellular components. **(E)** Biological processes. **(F)** Molecular functions. **(G)** KEGG pathway analysis.

Characterization of Immune Cells and NCKAP1 Expression in CcRCC

Tumor immunity is extremely involved in tumorigenesis and prognosis in ccRCC Using the TIMER web portal, we correlated NCKAP1 expression with the intensity of immune infiltrating cells NCKAP1 expression was found in B cells (Cor = 0. 216, p = 3.16e-06), $CD8^+$ T cells (Cor = 0. 175, p = 2.31e-04), macrophages (Cor = 0.308, p = 2.61e-11), neutrophils (Cor = 0.283, p = 6.88e-10) and dendritic cells (Cor = 0.234, p = 4.27e-07), showing a clear positive correlation (**Figure 7A**). Furthermore, somatic copy number changes of NCKAP1 can indeed inhibit the infiltration of immune cells such as $CD8^+$ T cells, B cells, neutrophils, dendritic cells, macrophages and $CD4^+$ T cells in ccRCC (**Figure 7B**). After adjusting by tumor purity, we found an important biomarker of immune cell correction by NCKAP1. As shown in **Table 2**, markers of Monocyte (CD86, CSF1R), TAM (IL10), M1 macrophage (NOS2, PTGS2), M2 macrophage (CD163, VSIG4, MS4A4A) and Dendritic cell (NRP1) which showed significant correlations with NCKAP1 were obtained. Similar results were obtained for Th1 (STAT1), Th2 (STAT6), Tfh (BCL6), Th17 (STAT3), and Treg (CCR8, STAT5B). In summary, we expected a specific correlation between NCKAP1 expression and the intensity of immune infiltration.

Immune Checkpoint, TMB, MSI, and NCKAP1 Expression in ccRCC

Immune checkpoints play an essential role in targeted immunotherapy and are considered an important method of tumor therapy. In this study, we analyzed the expression of NCKAP1 and SIGLEC15, HAVCR2, CTLA4, TIGIT, PDCD1LG2, CD274, LAG3, and PDCD1 immune checkpoint-related genes. The results showed that immune checkpoint markers (HAVCR2, CTLA4, TIGIT, CD274, LAG3, PDCD1) were significantly correlated with NCKAP1 expression; TMB and MSI were considered prognostic markers to predict response to immunotherapy in CCRCC (Pang et al., 2016). However, our results showed that TMB and MSI were not significantly correlated with NCKAP1 (**Figure 8B**).

DISCUSSION

ccRCC is one of the universal urinary tract malignancies (Chen et al., 2016; Pang et al., 2016). ccRCC is challenging to diagnose

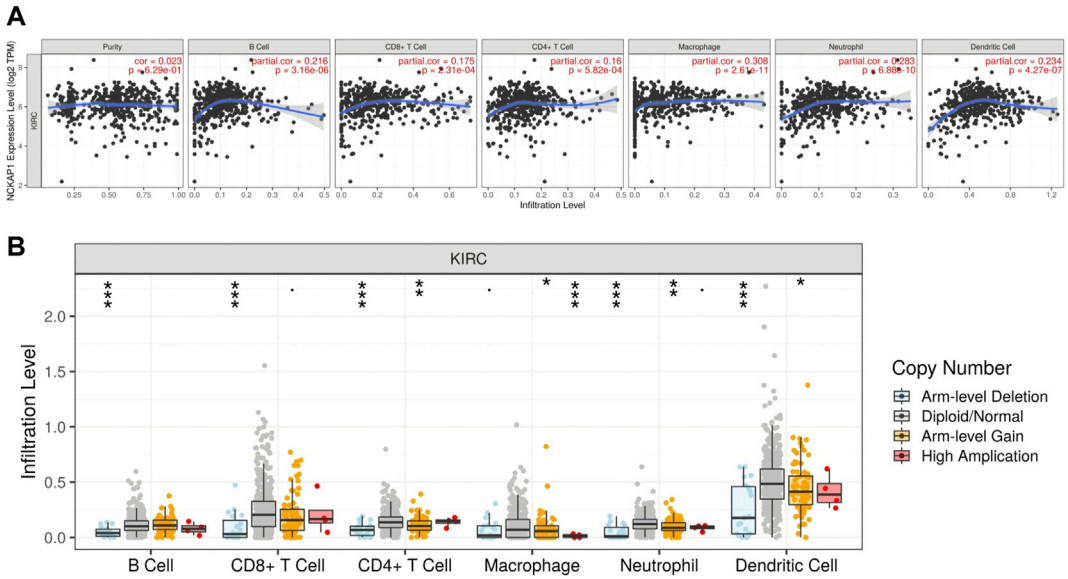


FIGURE 7 | Correlation between NCKAP1and immune cell infiltration in ccRCC. **(A)** The correlation between NCKAP1 and the immune infiltration level in ccRCC (TIMER). **(B)** The correlation between copy number alteration of NCKAP1 and immune cell infiltration in ccRCC.

early, and patients often present with advanced metastatic cancer (Rydzanicz et al., 2013). Therefore, identifying new biomarkers in ccRCC that can be used as a therapeutic system for early diagnosis and new treatment design is much needed.

Previous studies have shown that NCKAP1 is associated with multiple cancer types; high NCKAP1 levels are clearly associated with the clinical features of human non-small cell lung cancer (NSCLC) (Zhu et al., 2021). NCKAP1 also interacts with HSP90 (heat shock protein 90) and is highly implicated in NSCLC cell invasion and metastasis (Xiong et al., 2019a). John et al. used a mouse xenograft breast cancer metastasis model to show that NCKAP1 plays a critical role in invasion and metastasis by regulating WASF3 stability and function (Cowell et al., 2017), and Karthic et al. found that targeted deletion of NCKAP1 inhibited melanoma progression using a BRAF/PTEN-deficient mouse model (Swaminathan et al., 2021). Zhu et al. also analyzed gene expression profiles in the GEO database and showed that NCKAP1 is an autophagy-related gene and is significantly associated with event-free survival in several melanoma patients (Zhu et al., 2019). Xiao et al. found that the miR-34c-3p target NCKAP1 promotes the progression of hepatocellular carcinoma and is associated with poor prognosis (Xiao et al., 2017). These results indicate that NCKAP1 may function as an oncogene in a variety of cancer types. On the other hand, NCKAP1 has been shown to have tumor suppressor activity that regulates the HCC cell cycle through the regulation of Rb1/p53, but not the WASF pathway (Zhong et al., 2019). These differences may be due to selective activation of target genes of NCKAP1-related pathways in specific tissues.

The association between NCKAP1 expression and clinical features in ccRCC patients remains largely unknown. We first detected NCKAP1 mRNA and protein expression levels using the TCGA and CPTAC datasets and found that NCKAP1 expression was downregulated in tumors compared to normal tissues. When compared between NCKAP1 expression, tumor grade, TNM stage, and lymph node metastasis, a negative correlation was observed, suggesting that NCKAP1 may have an antitumor effect in ccRCC. Other studies have shown that transcriptional repression of tumor suppressor genes occurs with tumor progression due to hypermethylation of promoter regions (Ruiz de la Cruz et al., 2021). Therefore, we speculated that low expression of NCAKP1 might be associated with high methylation levels in ccRCC, which is consistent with our previous results. Our IHC results showed that low expression of NCKAP1 correlated with tumor size, stage, and grade, confirming our previous findings. GEPIA survival analysis showed that low NCKAP1 expression was significantly associated with poor prognosis in ccRCC patients. Functional assays also showed that NCKAP1 affects cancer cell proliferation, migration, and invasion, as well as inhibits tumor growth *in vivo*.

Next, we examined the expression of genes significantly associated with NCKAP1 and its function in ccRCC. The results showed that genes associated with NCKAP1 function primarily in humoral immune responses, NADH dehydrogenase and mitochondrial respiratory chain complex assembly, oxidative phosphorylation, cytoplasmic DNA sensing and TGF-β signaling pathways. It has been shown that among these biological processes, immune response and metabolic alterations played essential functions in tumorigenesis (Smith et al., 2018; Monette et al., 2019; Chakraborty et al., 2021). The TGF-β signaling pathway is frequently downregulated in tumor cells and may increase or hinder tumor growth (Bao et al., 2021). Prior studies have shown that TGF-β can modify tumor activity by inhibiting host tumor immune surveillance and directly regulating oncogenic metabolism in EMT, cellular invasion and metastasis (Huber-Ruano et al., 2017; Lee et al., 2017;

TABLE 2 | Correlation analysis between NCAKP1 and gene biomarkers of immune cells in ccRCC (TIMER).

Immune Cells	Biomarkers	None		Purity	
		Cor	*p*-Value	Cor	*p*-Value
CD8+T cell	CD8A	−0.096	2.62e-02	0.009	8.50e-1
	CD8B	−0.195	*	−0.067	1.52e-01
T cell (general)	CD3D	−0.234	**	−0.104	2.60e-02
	CD3E	−0.252	***	−0.126	7.22e-03
	CD2	−0.159	*	−0.002	9.72e-01
B cell	CD19	−0.159	*	−0.056	2.32e-01
	CD79A	−0.235	**	−0.134	3.98e-03
Monocyte	CD86	0.013	7.82e-01	0.196	*
	CD115(CSF1R)	−0.004	9.27e-01	0.156	*
TAM	CCL2	−0.063	1.74e-01	0.074	1.12e-01
	CD68	0.02	6.57e-01	0.125	7.64e-03
	IL10	0.07	1.32e-01	0.218	**
M1 macrophage	INOS(NOS2)	0.004	9.37e-01	0.01	8.29e-01
	IRF5	−0.104	2.47e-02	0.033	4.81e-01
	COX2(PTGS2)	0.201	**	0.251	***
M2 macrophage	CD163	0.147	1.44e-03	0.299	***
	VSIG4	0.036	4.34e-01	0.162	*
	MS4A4A	0.057	2.14e-01	0.213	**
Neutrophils	CD66b(CEACAM8)	0.113	1.45e-02	0.112	1.61e-02
	CD11b(ITGAM)	0.003	9.41e-01	0.128	6.20e-03
	CCR7	−0.177	*	−0.029	5.38e-01
NK	KIR2DL1	−0.109	1.79e-02	−0.037	4.25e-01
	KIR2DL3	−0.128	5.24e-03	−0.027	5.64e-01
	KIR2DL4	−0.17	*	−0.067	1.54e-01
	KIR3DL1	−0.149	1.21e-03	−0.065	1.68e-01
	KIR3DL2	−0.199	**	−0.086	6.48e-02
	KIR3DL3	−0.108	1.93e-02	−0.077	1.02e-01
	KIR2DS4	−0.161	4.4e-04	−0.098	3.60e-02
Dendritic cell	HLA−DPB1	−0.174	*	−0.029	5.41e-01
	HLA−DQB1	−0.167	2.8e-04	−0.031	5.02e-01
	HLA−DRA	−0.105	2.33e-02	0.061	1.31e-01
	HLA−DPA1	−0.14	2.29e-03	−0.003	9.41e-01
	BDCA−1(CD1C)	−0.075	1.02e-01	0.062	1.88e-01
	BDCA−4(NRP1)	0.313	***	0.4	***
	CD11c(ITGAX)	−0.084	6.91e-02	0.04	3.90e-01
Th1	T−bet (TBX21)	−0.16	4.86e-04	−0.016	7.27e-01
	STAT4	0.018	6.92e-01	0.176	*
	STAT1	0.267	***	0.399	***
	IFN−γ(IFNG)	−0.047	3.04e-01	0.104	2.59e-02
	TNF−α(TNF)	−0.098	3.31e-02	0.046	3.27e-01
Th2	GATA3	−0.137	2.84e-03	0.035	4.50e-01
	STAT6	0.249	***	0.25	***
	STAT5A	0.076	9.76e-02	0.086	6.69e-02
	IL13	−0.056	2.25e-01	−0.005	9.13e-01
Tfh	BCL6	0.168	*	0.239	**
	IL21	0.03	5.21e-01	0.125	7.25e-03
Th17	STAT3	0.389	***	0.438	***
	IL17A	−0.033	4.8e-01	−0.008	8.65e-01
Treg	FOXP3	−0.211	**	−0.08	8.68e-02
	CCR8	0.092	4.53e-02	0.264	***
	STAT5B	0.298	***	0.31	***
	TGFβ(TGFB1)	−0.034	4.57e-01	0.059	2.09e-01
T cell exhaustion	PD−1(PDCD1)	−0.228	**	−0.11	1.84e-02
	CTLA4	−0.073	1.14e-01	0.042	3.67e-01
	LAG3	−0.203	*	−0.085	7.03e-02
	TIM3(HAVCR2)	−0.029	5.34e-01	0.15	1.30e-03
	GZMB	−0.214	**	−0.086	6.50e-02

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

Ungefroren, 2019). Epithelialmesenchymal transition (EMT), which play vital roles intumor cell migration and invasion, is essential step in the process of metastasis. To our best knowedge, WAVE complex has been shown to be a promoter of cell invasion in various cancer cell types. NCKAP1, as a part of WAVE complex, it is therefore required for WAVE function and its regulation of

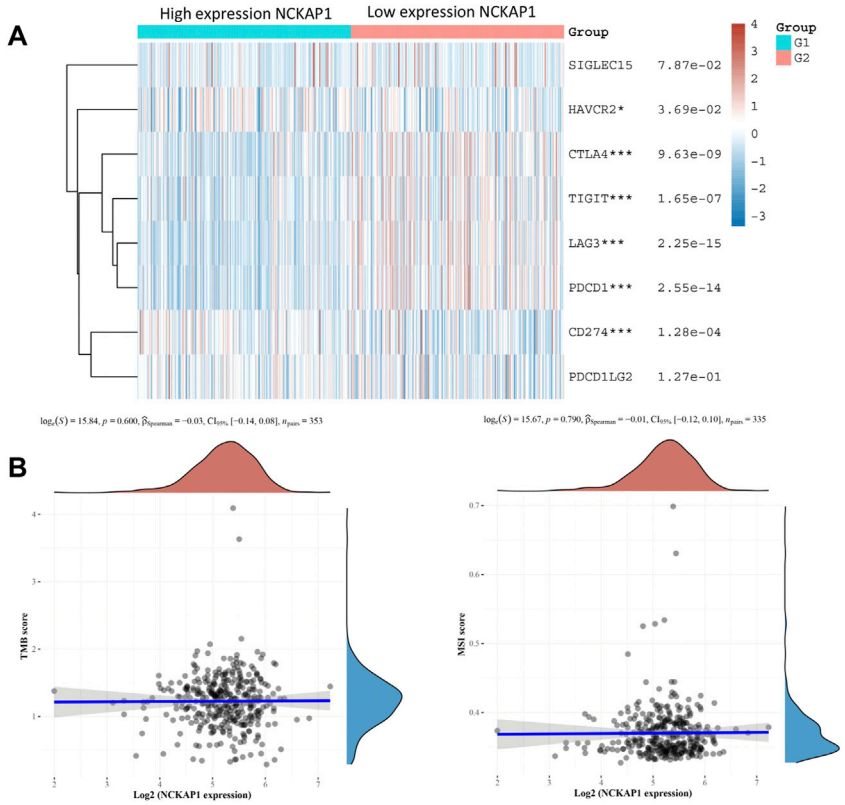


FIGURE 8 | Association of the NCKAP1 with immune checkpoints, TMB, and MSI in ccRCC. **(A)** Correlation of NCKAP1 expression with immune checkpoint genes. **(B)** Correlation of NCKAP1 expression with tumor mutational burden (TMB). Correlation of NCKAP1 expression with microsatellite instability (MSI).

invasion. Targeting NCKAP1 is thus reported to lead to the suppression of metastasis (Teng et al., 2016). Taken together, NCKAP1 and its related genes are primarily involved in tumor-related functions and playing important roles in EMT-related signaling pathways, suggesting that NCKAP1 may mediate ccRCC tumorigenesis and development.

Tumor-infiltrating immune cells are a crucial element in the tumor microenvironment and have been found to influence proliferation, invasion, and metastasis in various cancer types (Bremnes et al., 2016; Xiong et al., 2019b; Yang et al., 2020). Tumor-infiltrating immune cells and immune checkpoints are thought to play an essential role in immunotherapy, making them hotspot studies in ccRCC treatment (Xu et al., 2021; Wu et al., 2022). Our results revealed a clear correlation between NCKAP1 levels and immune cell counts and biomarker levels. Furthermore, the association with various immune checkpoints (HAVCR2, CTLA4, TIGIT, CD274, LAG3, and PDCD1) strongly suggested that NCKAP1 is a co-regulator of immune checkpoints in ccRCC. These results suggest that NCKAP1 may play an essential role in regulating tumor immune invasion and immunotherapy, which in turn may affect the prognosis of ccRCC.

This study has several limitations, and future studies should validate these results in a more significant number of cases and thoroughly investigate the detailed mechanisms by which NCKAP1 is involved in ccRCC.

CONCLUSION

In this study, using multiple portal databases, we found that low NCKAP1 expression levels were negatively correlated with clinical features and prognosis of ccRCC. Furthermore, our results indicate that NCAKP1 may play an important role in oncogenesis *in vitro* and *in vivo*. In addition, NCKAP1 and its related genes function primarily in metabolic-related signaling pathways and immune cell infiltration, predictably suggesting that NCKAP1 is a prognostic biomarker for ccRCC.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of The Second Affiliated Hospital of Shantou University Medical College. Written informed consent for participation was not required for this

study in accordance with the national legislation and the institutional requirements. The animal study was reviewed and approved by These experiments were approved by the Ethics Committee of Shantou University Medical College.

AUTHOR CONTRIBUTIONS

JC and JG designed the study, implemented the experiments and data analyses. XZ and ST planned the project and authored the core manuscript text. JC, WZ, XX collected the data. ST, XZ conceived the research study and wrote the manuscript. All authors reviewed the manuscript.

FUNDING

This project was supported by grants from the National Natural Science Foundation of China (82002068); Natural Science Foundation of Guangdong Province (2021A1515010949); Shantou Science and Technology project (200623175260683; 200624095260243).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.764957/full#supplementary-material

REFERENCES

Bao, J. M., Dang, Q., Lin, C. J., Lo, U. G., Feldkoren, B., Dang, A., et al. (2021). SPARC Is a Key Mediator of TGF-β-Induced Renal Cancer Metastasis. *J. Cell Physiol.* 236 (3), 1926–1938. doi:10.1002/jcp.29975

Bremnes, R. M., Busund, L.-T., Kilvær, T. L., Andersen, S., Richardsen, E., Paulsen, E. E., et al. (2016). The Role of Tumor-Infiltrating Lymphocytes in Development, Progression, and Prognosis of Non-small Cell Lung Cancer. *J. Thorac. Oncol.* 11 (6), 789–800. doi:10.1016/j.jtho.2016.01.015

Capitanio, U., Bensalah, K., Bex, A., Boorjian, S. A., Bray, F., Coleman, J., et al. (2019). Epidemiology of Renal Cell Carcinoma. *Eur. Urol.* 75 (1), 74–84. doi:10.1016/j.eururo.2018.08.036

Chakraborty, S., Balan, M., Sabarwal, A., Choueiri, T. K., and Pal, S. (2021). Metabolic Reprogramming in Renal Cancer: Events of a Metabolic Disease. *Biochimica Biophysica Acta (BBA) - Rev. Cancer* 1876 (1), 188559. doi:10.1016/j.bbcan.2021.188559

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V. S. K., et al. (2017). UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

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 (2), 115–132. doi:10.3322/caac.21338

Cowell, J. K., Teng, Y., Bendzunas, N. G., Ara, R., Arbab, A. S., and Kennedy, E. J. (2017). Suppression of Breast Cancer Metastasis Using Stapled Peptides Targeting the WASF Regulatory Complex. *Cancer Growth Metastasis* 10, 117906441771319. doi:10.1177/1179064417713197

Huber-Ruano, I., Raventós, C., Cuartas, I., Sánchez-Jaro, C., Arias, A., Parra, J. L., et al. (2017). An Antisense Oligonucleotide Targeting TGF-B2 Inhibits Lung Metastasis and Induces CD86 Expression in Tumor-Associated Macrophages. *Ann. Oncol.* 28 (9), 2278–2285. doi:10.1093/annonc/mdx314

Innocenti, M., Zucconi, A., Disanza, A., Frittoli, E., Areces, L. B., Steffen, A., et al. (2004). Abi1 Is Essential for the Formation and Activation of a WAVE2 Signalling Complex. *Nat. Cell Biol.* 6 (4), 319–327. doi:10.1038/ncb1105

Kotecha, R. R., Motzer, R. J., and Voss, M. H. (2019). Towards Individualized Therapy for Metastatic Renal Cell Carcinoma. *Nat. Rev. Clin. Oncol.* 16 (10), 621–633. doi:10.1038/s41571-019-0209-1

Lee, S. Y., Jeong, E. K., Ju, M. K., Jeon, H. M., Kim, M. Y., Kim, C. H., et al. (2017). Induction of Metastasis, Cancer Stem Cell Phenotype, and Oncogenic Metabolism in Cancer Cells by Ionizing Radiation. *Mol. Cancer* 16 (1), 10. doi:10.1186/s12943-016-0577-4

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.CAN-17-0307

Monette, A., Morou, A., Al-Banna, N. A., Rousseau, L., Lattouf, J.-B., Rahmati, S., et al. (2019). Failed Immune Responses across Multiple Pathologies Share Pan-Tumor and Circulating Lymphocytic Targets. *J. Clin. investigation* 129 (6), 2463–2479. doi:10.1172/JCI125301

Pang, C., Guan, Y., Li, H., Chen, W., and Zhu, G. (2016). Urologic Cancer in China. *Jpn. J. Clin. Oncol.* 46 (6), 497–501. doi:10.1093/jjco/hyw034

Patel, C., Ahmed, A., and Ellsworth, P. (2012). Renal Cell Carcinoma: a Reappraisal. *Urol. Nurs.* 32 (4), 182–190. doi:10.7257/1053-816x.2012.32.4.182

Perazella, M. A., Dreicer, R., and Rosner, M. H. (2018). Renal Cell Carcinoma for the Nephrologist. *Kidney Int.* 94 (3), 471–483. doi:10.1016/j.kint.2018.01.023

Rai, A., Greening, D. W., Xu, R., Suwakulsiri, W., and Simpson, R. J. (2020). Exosomes Derived from the Human Primary Colorectal Cancer Cell Line SW480 Orchestrate Fibroblast-Led Cancer Invasion. *Proteomics* 20 (14), 2000016. doi:10.1002/pmic.202000016

Ruiz de la Cruz, M., de la Cruz Montoya, A. H., Rojas Jiménez, E. A., Martínez Gregorio, H., Díaz Velásquez, C. E., Paredes de la Vega, J., et al. (2021). Cis-Acting Factors Causing Secondary Epimutations: Impact on the Risk for Cancer and Other Diseases. *Cancers* 13 (19), 4807. doi:10.3390/cancers13194807

Rydzanicz, M., Wrzesiński, T., Bluyssen, H. A. R., and Wesoły, J. (2013). Genomics and Epigenomics of Clear Cell Renal Cell Carcinoma: Recent Developments and Potential Applications. *Cancer Lett.* 341 (2), 111–126. doi:10.1016/j.canlet.2013.08.006

Smith, C. C., Beckermann, K. E., Bortone, D. S., De Cubas, A. A., Bixby, L. M., Lee, S. J., et al. (2018). Endogenous Retroviral Signatures Predict Immunotherapy Response in Clear Cell Renal Cell Carcinoma. *J. Clin. investigation* 128 (11), 4804–4820. doi:10.1172/JCI121476

Suzuki, T., Nishiyama, K., Yamamoto, A., Inazawa, J., Iwaki, T., Yamada, T., et al. (2000). Molecular Cloning of a Novel Apoptosis-Related Gene, Human Nap1 (NCKAP1), and its Possible Relation to Alzheimer Disease. *Genomics* 63 (2), 246–254. doi:10.1006/geno.1999.6053

Swaminathan, K., Campbell, A., Papalazarou, V., Jaber-Hijazi, F., Nixon, C., McGhee, E., et al. (2021). The RAC1 Target NCKAP1 Plays a Crucial Role in the Progression of Braf;Pten-Driven Melanoma in Mice. *J. Investigative Dermatology* 141 (3), 628–637. doi:10.1016/j.jid.2020.06.029

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a Web Server for Cancer and Normal Gene Expression Profiling and Interactive Analyses. *Nucleic Acids Res.* 45, W98–W102. doi:10.1093/nar/gkx247

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

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2019). The Immune Landscape of Cancer. *Immunity* 51 (2), 411–412. doi:10.1016/j.immuni.2019.08.004

Ungefroren, H. (2019). Blockade of TGF-β Signaling: a Potential Target for Cancer Immunotherapy? *Expert Opin. Ther. Targets* 23, 679–693. doi:10.1080/14728222.2019.1636034

Vasaikar, S. V., Straub, P., Wang, J., and Zhang, B. (2018). LinkedOmics: Analyzing Multi-Omics Data within and across 32 Cancer Types. *Nucleic Acids Res.* 46, D956–D963. doi:10.1093/nar/gkx1090

Whitelaw, J. A., Swaminathan, K., Kage, F., and Machesky, L. M. (2020). The WAVE Regulatory Complex Is Required to Balance Protrusion and Adhesion in Migration. *Cells* 9 (7), 1635. doi:10.3390/cells9071635

Wu, Z., Chen, Q., Qu, L., Li, M., Wang, L., Mir, M. C., et al. (2022). Adverse Events of Immune Checkpoint Inhibitors Therapy for Urologic Cancer Patients in

Clinical Trials: A Collaborative Systematic Review and Meta-Analysis. *Eur. Urol.* 81 (4), 414–425. doi:10.1016/j.eururo.2022.01.028

Xiao, C.-Z., Wei, W., Guo, Z.-X., Zhang, M.-Y., Zhang, Y.-F., Wang, J.-H., et al. (2017). MicroRNA-34c-3p Promotes Cell Proliferation and Invasion in Hepatocellular Carcinoma by Regulation of NCKAP1 Expression. *J. Cancer Res. Clin. Oncol.* 143 (2), 263–273. doi:10.1007/s00432-016-2280-7

Xiong, Y., He, L., Shay, C., Lang, L., Loveless, J., Yu, J., et al. (2019). Nck-associated Protein 1 Associates with HSP90 to Drive Metastasis in Human Non-small-cell Lung Cancer. *J. Exp. Clin. Cancer Res.* 38 (1), 122. doi:10.1186/s13046-019-1124-0

Xiong, Y., Liu, L., Xia, Y., Qi, Y., Chen, Y., Chen, L., et al. (2019). Tumor Infiltrating Mast Cells Determine Oncogenic HIF-2α-Conferred Immune Evasion in Clear Cell Renal Cell Carcinoma. *Cancer Immunol. Immunother.* 68 (5), 731–741. doi:10.1007/s00262-019-02314-y

Xu, H., Zheng, X., Zhang, S., Yi, X., Zhang, T., Wei, Q., et al. (2021). Tumor Antigens and Immune Subtypes Guided mRNA Vaccine Development for Kidney Renal Clear Cell Carcinoma. *Mol. Cancer* 20 (1), 159. doi:10.1186/s12943-021-01465-w

Yang, S., Liu, T., Nan, H., Wang, Y., Chen, H., Zhang, X., et al. (2020). Comprehensive Analysis of Prognostic Immune-Related Genes in the Tumor Microenvironment of Cutaneous Melanoma. *J. Cell. Physiology* 235 (2), 1025–1035. doi:10.1002/jcp.29018

Zhang, F., Xie, S., Zhang, Z., Zhao, H., Zhao, Z., Sun, H., et al. (2020). A Novel Risk Model Based on Autophagy Pathway Related Genes for Survival Prediction in Lung Adenocarcinoma. *Med. Sci. Monit.* 26, e924710. doi:10.12659/MSM.924710

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

Zhu, F. X., Wang, X. T., Zeng, H. Q., Yin, Z. H., and Ye, Z. Z. (2019). A Predicted Risk Score Based on the Expression of 16 Autophagy-related Genes for Multiple Myeloma Survival. *Oncol. Lett.* 18 (5), 5310–5324. doi:10.3892/ol.2019.10881

Zhu, B., V, M., Finch-Edmondson, M., Lee, Y., Wan, Y., Sudol, M., et al. (2021). miR-582-5p Is a Tumor Suppressor microRNA Targeting the Hippo-YAP/TAZ Signaling Pathway in Non-small Cell Lung Cancer. *Cancers* 13 (4), 756. doi:10.3390/cancers13040756

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.

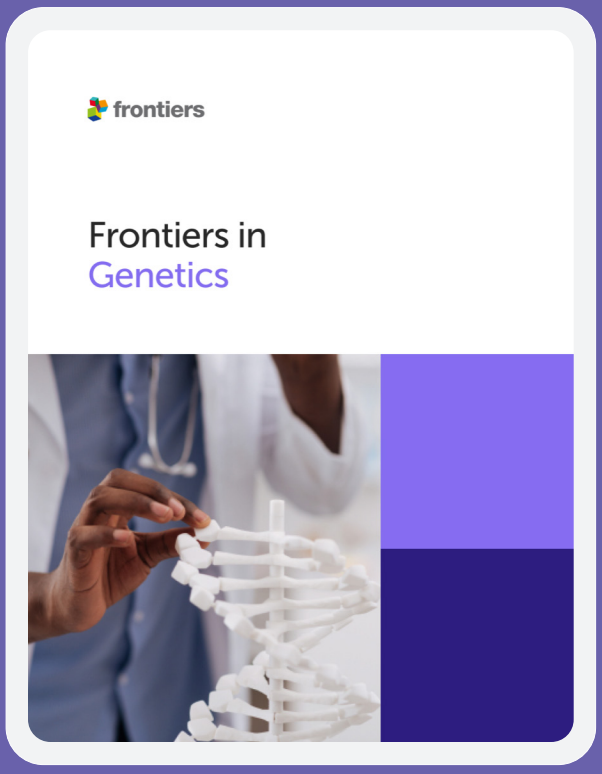
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact