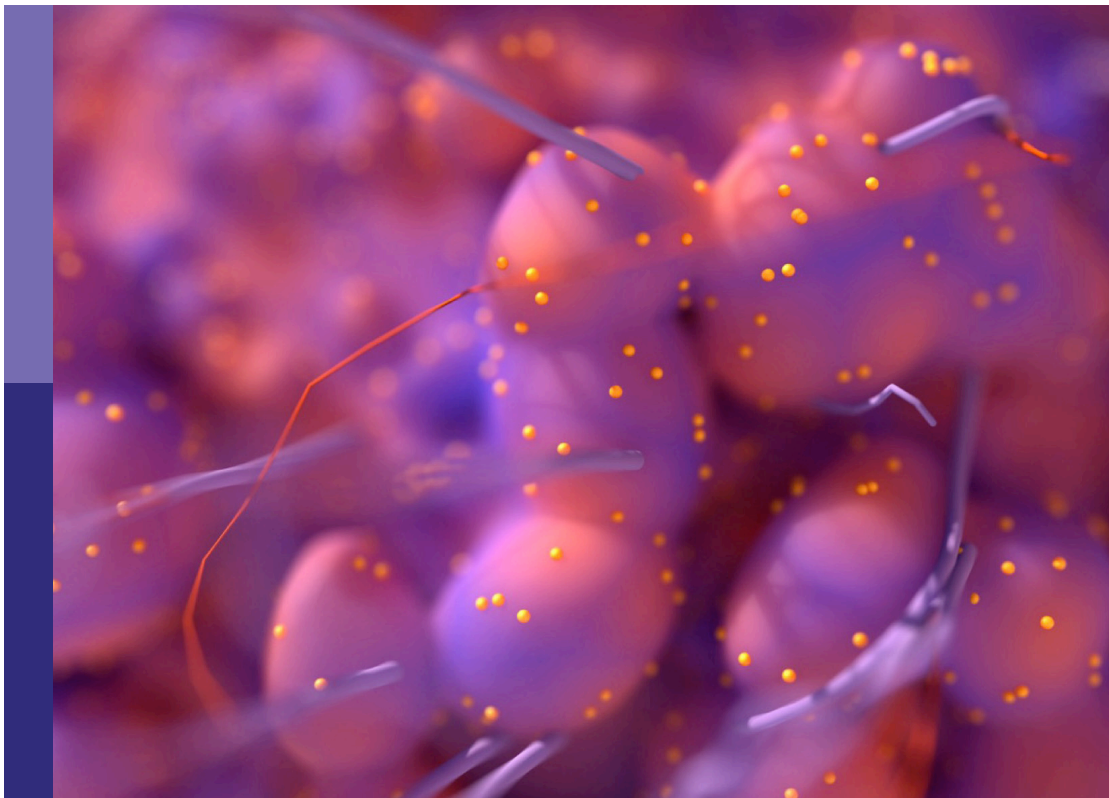
frontiers | Research Topics

Deciphering the role of signature genes in cancer prognosis and therapy resistance

Edited by
Dalila Luciola Zanette, Manoj K. Pandey and Jawed A. Siddiqui

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-6271-0
DOI 10.3389/978-2-8325-6271-0

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

Deciphering the role of signature genes in cancer prognosis and therapy resistance

Topic editors

Dalila Luciola Zanette — Oswaldo Cruz Foundation (Fiocruz), Brazil

Manoj K. Pandey — Cooper Medical School of Rowan University, United States

Jawed A. Siddiqui — University of Mississippi Medical Center, United States

Citation

Zanette, D. L., Pandey, M. K., Siddiqui, J. A., eds. (2025). *Deciphering the role of signature genes in cancer prognosis and therapy resistance*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-6271-0

Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 08 April 2025
DOI 10.3389/fonc.2025.1591256

OPEN ACCESS

EDITED AND REVIEWED BY
Tao Liu,
University of New South Wales, Australia

*CORRESPONDENCE
Jawed A. Siddiqui
jsiddiqui@umc.edu

RECEIVED 10 March 2025
ACCEPTED 27 March 2025
PUBLISHED 08 April 2025

CITATION
Zanette DL, Pandey MK and Siddiqui JA (2025) Editorial: Deciphering the role of signature genes in cancer prognosis and therapy resistance. *Front. Oncol.* 15:1591256. doi: 10.3389/fonc.2025.1591256

Editorial: Deciphering the role of signature genes in cancer prognosis and therapy resistance

Dalila L. Zanette[1], Manoj K. Pandey[2] and Jawed A. Siddiqui[3,4]*

[1]Carlos Chagas Institute, Oswaldo Cruz Foundation (Fiocruz/Paraná), Rio de Janeiro, Brazil, [2]Department of Biomedical Sciences, Cooper Medical School of Rowan University, Camden, NJ, United States, [3]Department of Cell and Molecular Biology, University of Mississippi Medical Center, Jackson, MS, United States, [4]Cancer Center Research Institute, University of Mississippi Medical Center, Jackson, MS, United States

KEYWORDS

cancer, therapy resistance, tumor microenvironment, immune checkpoint inhibitor (ICI), metastasis

Editorial on the Research Topic

Deciphering the role of signature genes in cancer prognosis and therapy resistance

Despite decades of research on cancer and its effects on patient care, it continues to pose a significant health challenge. The characterization of the multi-faceted disease is significantly influenced by tumor angiogenesis and therapy response, which are driven by genetic and epigenetic alterations. Preventive medicine, therapy design, and personalized care are interconnected concepts that relate to the variability in molecular determinants of cancer or gene signatures, which affect prognosis and treatment strategies. Gene signatures facilitate the recognition and functional characterization of entities through high-throughput genome and transcriptome analysis, RNA sequencing methodologies, single-cell omics investigations, and microarray profiling. The altered expression profile of various gene signatures indicates dysregulated cellular processes. The irregular expression of the gene-signature profile is directly linked to enhanced cell proliferation, immune system evasion, disruption of apoptotic pathways, epithelial to mesenchymal transition, alterations in the tumor microenvironment, and therapy resistance.

This Research Topic, titled *"Deciphering the Role of Signature Genes in Cancer Prognosis and Therapy Resistance,"* focuses on recent advancements in understanding the genetic signatures that influence cancer progression, as well as their contributions to therapy resistance and metastatic potential. This Research Topic comprises nine articles: seven original research articles (Liu et al., Mao et al., Li et al., Wang et al., Benevides et al., Terrones et al., and Sonnemann et al.), one review article (Wu et al.), and one mini review (Coelho et al.).

The application of immune checkpoint inhibitors (ICIs) has gained widespread acceptance for the treatment of patients with advanced Hepatocellular carcinoma (HCC). Liu et al. conducted a study involving 54 patients with hepatocellular carcinoma (HCC), utilizing SVM-RFE modeling on six immune-related genes (IRGs): CMTM7, HDAC1, HRAS, PSMD1, RAET1E, and TXLNA. They developed a novel approach to identify overall survival and the impact of immunotherapy in HCC [1]. A study by Mao

et al. on HCC identified six survival-related genes (BMI1, CCR3, CDC25C, CFL1, LDHA, RAC1) associated with the CCL18 signaling pathway. Additionally, enhanced cell proliferation, migration, and stemness were reported in response to the overexpression of these six genes [2].

Gastric cancer (GC) is recognized as a prevalent malignancy, and conventional treatment methods are insufficient for achieving favorable patient prognoses. Li et al. conducted a study utilizing the TCGA database to develop a novel prognostic system for gastric cancer patients. The authors emphasized a notable prognostic distinction between high and low-risk GC groups, indicating that risk scoring has a more substantial impact on prognosis than tumor stage identification in GC. This study is the inaugural effort in developing a model that emphasizes the significance of platelet-related genes in gastric cancer progression, metastasis, and resistance to therapy [3]. A study by Wang et al. developed a predictive model for glycosylation-related genes to elucidate the broader implications of immunotherapy for gastric cancer. GLT8D2 has been identified as a significant prognostic marker with a robust correlation to Tumor Infiltrating Lymphocytes (TILs), encompassing CD8+ T cells, CD4+ T cells, Treg cells, B cells, neutrophils, dendritic cells (DCs), natural killer (NK) cells, and monocytes, especially macrophages in GC [4].

The expression of long noncoding RNAs (lncRNAs), particularly HOTAIR, is well-documented in solid tumors. A study by Benevides et al. investigated the expression of HOTAIR and PTGS2 in 87 patients with Chronic Myeloid Leukemia (CML). Samples of CML exhibit significant downregulation in the expression of these two genes. Additionally, they emphasized the inverse correlation between BCR: ABL1 expression and HOTAIR and PTGS2 in CML patients undergoing imatinib treatment, highlighting the possible regulatory interactions between these factors in the context of the CML therapy [5].

ROS1+ non-small cell lung cancer (NSCLC) represents a molecular subgroup comprising approximately 2% of newly diagnosed lung cancers each year. Terrones et al. conducted an analysis of the transcriptomic characteristics of ROS1+ NSCLC samples utilizing data from The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases [6]. The authors observed an upregulation in pathways related to nucleotide synthesis and cell adhesion. The downregulation of NOTCH1 correlates with the reduced expression of PD-L1 in ROS1+ NSCLC. This study highlights the significance of nucleotide synthesis and cell adhesion in elucidating the pathophysiology of ROS1+ NSCLC [6].

A study by Sonnemann et al. investigated the transcriptional activator Vestigial-like 1 (VGLL1), elucidating its cellular function and downstream targets in placental, breast, and pancreatic cancer cells [7]. The authors conducted ChIP-seq analysis to identify eight transcription factors with VGLL1-binding motifs. Additionally, increased expression of VGLL1 was associated with enhanced cell invasion and proliferation, highlighting its potential as a key player in cancer progression [7].

The complexity of prostate cancer (PCa) is significantly influenced by the genetic signature of the individual patient. Wu et al. summarize that the progression of prostate cancer (PCa) and its eventual advancement to therapy resistance and lethality is driven by the genetic interactions among the androgen receptor (AR), retinoblastoma (Rb), PTEN, WNT, p53, and MYC. This review elaborates on the functional and therapeutic potential of these key genes, which hold promise for future breakthroughs and the development of novel drugs to tailor treatment options for PCa [8].

Coelho et al. have described urinary mRNA-based biomarkers as potential tools for studying aggressiveness in non-muscle invasive bladder cancer (NMIBC), a type characterized by high proliferation and recurrence [9]. The presence of both shared (IGF2, ANAXA10, CRH) and exclusive (ABC1 and UPK1B) mRNA-based biomarkers for NMIBC is proposed to enhance prognosis and inform targeted therapy design for this type of bladder cancer [9].

This Research Topic will be an essential resource for researchers investigating the intricacies of hallmark genes in cancer prognosis and therapeutic resistance. We seek to offer an extensive understanding of how these genetic fingerprints affect disease development, treatment response, and resistance mechanisms by integrating advanced studies and expert insights. We anticipate that the insights presented in this Research Topic will augment existing knowledge and stimulate innovative approaches for advancing cancer diagnostics, prognostication, and therapeutic interventions.

Author contributions

DZ: Writing – review & editing. MP: Writing – review & editing. JS: Writing – review & editing, Writing – original draft.

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 06 March 2024
DOI 10.3389/fonc.2024.1371990

OPEN ACCESS

EDITED BY
Dalila Luciola Zanette,
Oswaldo Cruz Foundation (Fiocruz), Brazil

REVIEWED BY
Yin Celeste Cheuk,
The University of Hong Kong, Hong Kong SAR, China
Yu-gang Huang,
Hubei University of Medicine, China

*CORRESPONDENCE
Yi Pan
panyi@cpu.edu.cn
Liang Jin
ljstemcell@cpu.edu.cn

†These authors have contributed equally to this work

RECEIVED 17 January 2024
ACCEPTED 19 February 2024
PUBLISHED 06 March 2024

CITATION
Mao J, Tao Y, Wang K, Sun H, Zhang M, Jin L and Pan Y (2024) Identification of hub genes within the CCL18 signaling pathway in hepatocellular carcinoma through bioinformatics analysis.
Front. Oncol. 14:1371990.
doi: 10.3389/fonc.2024.1371990

Identification of hub genes within the CCL18 signaling pathway in hepatocellular carcinoma through bioinformatics analysis

Jinlei Mao†, Yuhang Tao†, Keke Wang, Hanru Sun, Manqi Zhang, Liang Jin* and Yi Pan*

State Key Laboratory of Natural Medicines, Jiangsu Key Laboratory of Druggability of Biopharmaceuticals, School of Life Science and Technology, China Pharmaceutical University, Nanjing, Jiangsu, China

Introduction: Hepatocellular carcinoma (HCC) is an aggressive malignancy, and CCL18, a marker of M2 macrophage activation, is often associated with tumor immune suppression. However, the role of CCL18 and its signaling pathway in HCC is still limited. Our study focuses on investigating the prognostic impact of CCL18 and its signaling pathway in HCC patients and biological functions *in vitro*.

Methods: HCC-related RNA-seq data were obtained from TCGA, ICGC, and GEO. The 6 hub genes with the highest correlation to prognosis were identified using univariate Cox and LASSO regression analysis. Multivariate Cox regression analysis was performed to assess their independent prognostic potential and a nomogram was constructed. *In vitro* experiments, including CCK8, EdU, RT-qPCR, western blot, and transwell assays, were conducted to investigate the biological effects of exogenous CCL18 and 6 hub genes. A core network of highly expressed proteins in the high-risk group of tumors was constructed. Immune cell infiltration was evaluated using the ESTIMATE and CIBERSORT packages. Finally, potential treatments were explored using the OncoPredict package and CAMP database.

Results: We identified 6 survival-related genes (BMI1, CCR3, CDC25C, CFL1, LDHA, RAC1) within the CCL18 signaling pathway in HCC patients. A nomogram was constructed using the TCGA_LIHC cohort to predict patient survival probability. Exogenous CCL18, as well as overexpression of BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1, can promote proliferation, migration, invasion, stemness, and increased expression of PD-L1 protein in LM3 and MHCC-97H cell lines. In the high-risk group of patients from the TCGA_LIHC cohort, immune suppression was observed, with a strong correlation to 21 immune-related genes and suppressive immune cells.

Conclusion: Exogenous CCL18 promotes LM3 and MHCC-97H cells proliferation, migration, invasion, stemness, and immune evasion. The high expression of BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 can serve as a biomarkers for immune evasion in HCC.

KEYWORDS
hepatocellular carcinoma, CCL18, tumor microenvironment, prognostic, diagnosis

Introduction

Liver cancer is the sixth most common primary malignant tumor and the fourth leading cause of cancer-related death worldwide, with a five-year survival rate of 21%. Hepatocellular carcinoma (HCC) accounts for more than 90% of liver cancer cases (1, 2). Although hepatitis B virus (HBV), hepatitis C virus (HCV), and alcohol remain important risk factors, the prevalence of obesity and diabetes has made non-alcoholic fatty liver disease (NAFLD) or non-alcoholic steatohepatitis (NASH) a dominant risk factor for HCC (3).

HCC presents intricate molecular characteristics and various pathological subtypes in a more natural manner, and the recommended treatment strategy for patients with advanced HCC continues to be systemic therapy, utilizing first-line agents like Sorafenib and Lenvatinib (3, 4). In recent times, there has been a growing focus on immune checkpoint inhibitors (ICIs) for the treatment of HCC. The combination of Atezolizumab (anti-programmed death-ligand 1) and Bevacizumab (anti-vascular endothelial growth factor) has emerged as a new standard for patients with advanced HCC (5), offering a therapy that modulates the HCC microenvironment. However, it is important to note that this treatment is only effective in a minority of HCC patients (6). In this regard, further research is needed to better understand the tumor microenvironment in HCC. This will allow for the identification of biomarkers that can be used to develop personalized treatment strategies.

Tumor microenvironment (TME) is a complex ecosystem that encompasses diverse immune cells, including dendritic cells (DC), monocytes, macrophages, B cells, and T cells (7). The TME of HCC accelerates tumor cell proliferation, invasion, and metastasis by forming an immunosuppressive environment (8). Tumor-associated macrophages (TAMs) have a high proportion in HCC TMEs and contribute to angiogenesis, cancer cell progression, and treatment resistance (9, 10). Macrophages can be classified into two main types: M1 and M2. M1 macrophages are involved in the immune response against cancer cells and express the CD86 marker. On the other hand, M2 macrophages have immunosuppressive functions and express the CD163 and CD206 markers (11). Of note, Guo et al. demonstrated that there is a subgroup of M2 macrophages ($CD68^+$ $CD206^+$) with high expression of chemokine ligand 18 (CCL18) in the HCC microenvironment and may be involved in the HCC process (12).

CCL18 is a chemokine secreted by TAMs and serves as a biomarker for M2 macrophages. It has been shown to promote tumor cell proliferation and facilitate immune evasion, aiding in the progression of tumor growth (13, 14). Lin et al. reported that CCL18 can promote HCC cell migration and invasion (15). However, research on the immunosuppressive effects of CCL18 in HCC is relatively limited. In this study, we systematically investigated the impact of the CCL18 signaling pathway on the prognosis of HCC patients. Consequently, six hub genes associated with prognosis were determined through bioinformatics analysis. These genes are denoted as BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1. Subsequently, we validated the biological functions of exogenous CCL18 and these six genes in LM3 and MHCC-97H cells through experimental assays. Then, we conducted immune cell infiltration of high-risk group and verified the influence of exogenous CCL18 and the expression of hub genes on PD-L1 protein. Eventually, potential treatments were explored using computational tools. The aim of this research is to gain a better understanding of the mechanisms underlying the development of immunosuppressive malignant HCC associated with CCL18. Identifying relevant tumor biomarkers may serve as a reference for diagnostic and immunotherapy of HCC.

Abbreviations: HCC, hepatocellular carcinoma; CCL18, chemokine ligand 18; PD-L1, programmed death-ligand 1; HBV, hepatitis B virus; HCV, hepatitis C virus; NAFLD, non-alcoholic fatty liver disease, NASH, non-alcoholic steatohepatitis; ICIs, immune checkpoint inhibitors; TME, tumor microenvironment; DC dendritic cells; TAMs, tumor-associated macrophages; TCGA, The Cancer Genome Atlas; ICGA, International Cancer Genome Consortium; GEO, Gene Expression Omnibus; ROC, receiver operating characteristic; OS, overall survival; RFS, recurrence-free survival; PFS, progression-free survival; DSS, disease-specific survival; DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction; GAPDH, Glyceraldehyde-3-phosphate dehydrogenase; HR, hazard ration; AUC, area under the curve; EMT, Epithelial-Mesenchymal Transition; MIR-4, macrophage inflammatory; PARC, pulmonary and activation-regulated chemokine; DC-CK1, dendritic cell chemokine 1; AMAC-1, alternative macrophage activation-associated CC chemokine 1; LDH, lactate dehydrogenase isoenzymes.

Materials and methods

Database selection and data acquisition

In this study, we acquired gene expression matrix (RNA-seq) and clinical information of HCC patients from The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/; TCGA_LIHC; 362 patient samples), International Cancer Genome Consortium (ICGC) (https://dcc.icgc.org/; ICGC_JP; 230 patient samples), and Gene Expression Omnibus (GEO) (GSE14520; 225 patient samples) (Table 1). The tumor samples for RNA-seq included 374 cases for TCGA_LIHC, 243 cases for ICGC_JP, and 225 cases for GSE14520. Accordingly, the adjacent normal samples of RNA-seq included 50 cases for TCGA_LIHC, 202 cases for ICGC_JP, and 220 cases for GSE14520. TCGA_LIHC dataset was used as the internal training cohort while ICGC_JP and GSE14520 datasets were used as the

TABLE 1 Clinical characteristics of the HCC patients in this research.

TCGA_LIHC			ICGC_JP			GSE14520		
Number of tumor patients 362			Number of tumor patients 230			Number of tumor patients 225		
Age	>=60	198(54.25%)	**Age**	>=60	185(80.43%)	**Age**	>=60	43(19.11%)
	<60	164(45.75%)		<60	45(19.57%)		<60	178(79.11%)
Gender	Male	245(67.12%)	**Gender**	Male	170(73.91%)		unknow	4
	Female	117(32.88%)		Female	60(26.09%)	**Gender**	Male	191(84.89%)
T_stage	T1	178(49.17%)	**Stage**	stageI	35(15.22%)		Female	30(13.33%)
	T2	90(24.86%)		stageII	105(45.65%)		unknow	4
	T3	78(21.55%)		stageIII	71(30.87%)	**TNM_stage**	stageI	93(41.33%)
	T4	13		stageIV	19(8.26%)		stageII	77(34.22%)
	TX	1					stageIII	48(21.33%)
	Unknow	2					unknow	6
N_stage	N0	247(68.23%)				**BCLC_stage**	Stage_A	148(65.78%)
	N1	4					Stage_B	22(9.78%)
	NX	110(30.39%)					Stage_C	29(12.89%)
	Unknow	1					unknow	26
M_stage	M0	260(71.82%)				**CLIP_stage**	Stage_0	97(43.11%)
	M1	3					Stage_1	74(32.89%)
	MX	99(27.35%)					Stage_2	35(15.56%)
Stage	stageI	168(83.46%)					Stage_3	9
	stageII	83(22.92%)					Stage_4	3
	stageIII	83(22.92%)					Stage_5	1
	stageIV	4					unknow	6
	unknow	24						

external testing cohorts. GSE14520 dataset was downloaded through "GEOquery" R package, and clinical data was acquired from the website. All TCGA_LIHC and ICGC_JP data, gene-expressed profile and clinical details were manually downloaded from the website.

Meanwhile, 99 genes that are implicated in the CCL18 signaling pathway were obtained from WikiPathways (https://www.wikipathways.org/) (Table S1). Immune-related genes were obtained from the ImmPort database (https://www.immport.org/shared/home).

Identification of hub genes

We utilized the "survival" and "survminer" R packages to perform univariate Cox regression analysis in the TCGA_LIHC training cohort with the aim of identifying genes that have a substantial impact on survival. Consequently, we observed significant variations in the expression of 39 genes. Subsequently, the "glmnet" R package was employed to conduct LASSO regression, enabling the selection of the most crucial variables from this gene set. As a result, a subset of 6 genes, specifically BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1, were identified as prominent hub genes.

Construction of prediction model

The multivariate Cox regression analysis was employed using the 6 hub genes. The risk score for individual patients in TCGA_LIHC cohort was calculated using the following formula.

$$Risk\ \ score = \sum_{i=1}^{n} coefficient(i)\ \times\ gene(i)$$

In this formula, coefficients were acquired from multivariate Cox regression, where gene(i) represents mRNA expression. The patients in both the training and test cohorts were classified into two groups, namely the high- and low-risk groups, based on the median value of the risk score. Additionally, we examined whether the risk score independently served as a prognostic factor. The clinical characteristics within TCGA_LIHC cohort including age, gender, M, N, T stage, stage and risk score were analyzed through univariate and multivariate Cox regression analysis. Analogous analyses were performed using the ICGC_JP and GSE14520 cohorts.

A predictive nomogram was developed using the 6 hub genes to estimate the survival probability for individual HCC patients. The predictive capability of the nomogram was evaluated in both the training and test cohorts through the utilization of receiver operating characteristic (ROC) curves and calibration curves. The "rms", "timeROC", and "pROC" R packages were utilized for these assessments.

Survival and pathway correlation analysis

To evaluate the influence of the six hub genes on the overall survival rate of TCGA_LIHC cohort, survival curves were generated using the Kaplan-Meier method from the "survival" R package, and the median value was taken as the best cut-off. The Kaplan-Meier Plotter (http://kmplot.com/analysis/) platform was employed to examine the association between the expression levels of prognostic genes and various clinical endpoints, including overall survival (OS), recurrence-free survival (RFS), progression-free survival (PFS), and disease-specific survival (DSS). Additionally, the genes encompassed within pathways were collected and subjected to analysis using the

"GSVA" R package, with the parameter method = 'ssgsea' being specifically chosen. Subsequently, the correlation between the prognostic genes and the pathways was assessed using Spearman correlation analysis (16).

Functional enrichment analysis and establishment of a PPI network

The "limma" R package was utilized to identify differentially expressed genes (DEGs) in the TCGA_LIHC cohort. This included comparisons between HCC tumor and normal samples, as well as between high- and low-risk groups (Log2 fold change > 1, p value< 0.05). A total of 308 genes were identified as the intersection between the genes exhibiting high expression in tumors and the genes within the high-risk score group. The findings were visually represented through the utilization of volcano and Venn diagrams using "ggvenn", "tidyverse", and "ggrepel" R packages, effectively illustrating the intersection of 308 genes exhibiting high expression in tumors and the genes within the high-risk score group.

The functional enrichment analysis of DEGs was systematically performed using the "clusterProfiler" and "org.Hs.eg.db" R packages, which facilitated Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses for a comprehensive understanding of the biological functions involved. The protein-protein interaction (PPI) network was meticulously examined utilizing the STRING database (https://cn.string-db.org/), and a significant molecular cluster was identified through the application of MODE, a Cytoscape plugin (Cytoscape software version 3.9.1).

Estimation of immune cell infiltration

The estimation of stromal, immune, and ESTIMATE scores was performed using the "ESTIMATE" R package, which provides a computational approach to calculate and quantify the stromal and immune components within the tumor microenvironment. Meanwhile, the "CIBERSORT" R package was utilized to estimate the impact of the risk score on the proportions of 22 immune cell subtypes in the TCGA_LIHC training cohort.

Drug sensitivity analysis

The "OncoPredict" R package was utilized to predict drug sensitivity based on gene expression profiles. This approach enabled the calculation of drug sensitivity values for each sample, with lower values indicating higher efficacy of the drug. Comprehensive drug sensitivity analysis was conducted on all samples to determine the drug sensitivity values for commonly used tumor drugs.

CAMP (https://clue.io/) is an extensive database and analysis platform that offers valuable resources and tools to delve into and comprehend gene expression profiles and drug perturbations. In order to identify potential small-molecule drugs for the treatment of high-risk patients, we utilized the CAMP online database. Specifically, we selected downregulated genes from the pool of significant genes, as well as upregulated genes from the top 150 genes. By inputting this gene set into the CAMP database, we conducted a meticulous screening to identify promising small-molecule drugs capable of modulating the dysregulated gene expression patterns associated with high-risk patients.

Cell culture and transfection

L02, Huh7, HepG2, LM3 and MHCC-97H cells were purchased from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). Cells were incubated in DMEM medium with 10% fetal bovine serum (FBS) and maintained in penicillin (100 IU/mL) and streptomycin (100 mg/mL) in 5% CO2 at 37°C. The plasmids encoding the BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 genes were constructed by cloning the sequence of the coding region using the appropriate primers (Table S2) and inserting the fragment into the pcDNA3.1 (+) plasmid. The cells were transfected with the plasmids using Lipofectamine 2000 Transfection Reagent (Invitrogen) and then the medium was changed 6h after transfection.

RNA extraction and quantitative real-time PCR

Total RNA was isolated using the TRIzol (Invitrogen) RNA extraction method, following the manufacturers' instructions. qRT-PCR measurements were performed as described previously (17) with the appropriate primers listed in Table S3. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was regarded as the internal reference, and the $2^{-\Delta\Delta Ct}$ method was applied to express the ratio of the target gene expression in the experimental group compared to the control group.

Protein extraction and western blotting

Protein extraction and western blotting analysis were performed using previously standard procedures (17). The following antibodies were used for western blotting: anti-E-cadherin antibody(#20874-1-AP Proteintech, China), anti-N-cadherin antibody(#22018-1-AP Proteintech, China), anti-ZEB1 antibody(#66279-1-Ig Proteintech, China), anti-Vimentin antibody(10366-1-AP Proteintech, China), anti-SOX2 antibody (#11064-1-AP Proteintech, China), anti-GAPDH antibody (#10494-1-AP Proteintech, China), and anti-PD-L1 antibody (#66248-1-Ig Proteintech, China).

Proliferation assay

Cell proliferation was detected using the Cell Counting Kit-8 (CCK-8, #K1018, APExBIO, USA) and the EdU cell proliferation assay kit (#C0071S, Beyotime Biotech, China), following the manufacturer's instructions. The colony-formation assay was also performed to assess cell proliferation. 5000 cells were plated per well

in triplicate in 6-well plates. The culture medium was changed every 3 days. Once visible clones were observed, each well was washed with PBS three times, fixed with methanol for 30 minutes at room temperature, and then stained with 0.05% crystal violet for 30 minutes. After washing, the colonies were counted and imaged.

Migration and invasion assay

According to the published method (18), transwell migration (without Matrigel) and Matrigel (Matrigel, Corning, China) invasion assays were performed to evaluate cell migration and invasion abilities, respectively. Additionally, cell migration was measured using wound healing assays as previously described (18).

Sphere formation assay

A sphere formation assay was performed to assess the stemness properties of LM3 and MHCC-97H cells. 1×10^5 cells were seeded into the 6-well ultra-low attachment plates (Corning, China) in sphere formation medium (#CCM0012, Minneapolis, USA). The cells were incubated in a CO_2 incubator for two weeks, and the number of spheres was counted under a stereomicroscope (Olympus).

Statistical analysis

Statistical analysis was performed using R Studio (R version 4.2.3) and GraphPad Prism 9.0.2. For genes with multiple probes, the maximum expression was selected. Cox regression analysis was conducted using the "survival" package to assess the association between variables and prognosis, including hazard ratios and 95% confidence intervals (CI). Lasso analysis was employed as a variable selection method to refine the scope of variables, with the lambda value chosen for optimal regularization. All data were presented as the mean and standard error of the mean (mean ± SD, n = 3). GraphPad Prism was used to create bar graphs. A two-tailed Student's t-test was used to compare the means between two groups, and an ANOVA test was used to assess significant differences among various experimental groups. The p-values in multiple comparisons were adjusted to control the false discovery rate (FDR) using the Benjamini-Hochberg method. The OS was evaluated using a Kaplan-Meier (K-M) curve, with statistical significance assessed using a log-rank test. The correlation between two variables conforming to a normal distribution was calculated employing the Pearson method. $P< 0.05$ was considered as statistically significant.

Results

Identification of prognostic-related genes

HCC raw datasets were obtained from the TCGA, ICGC, and GEO databases. Prior to analysis, these datasets were normalized using the log2(TPM + 1) transformation. 99 genes within CCL18 signaling pathway were downloaded from WikiPathways (Table S1). The expression patterns of these genes in the TCGA_LIHC cohort (n = 424) were illustrated using a heatmap (Figure S1A). Next, 39 genes exhibiting significant differences in hazard ration (HR) were identified from the TCGA_LIHC cohort by univariate Cox regression analysis (Figure 1A). Six hub genes, namely BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1, were identified using Lasso regression analysis (Figures 1B, C). Then, we compared the expression of these six hub genes in TCGA_LIHC, ICGC_JP (n = 445), and GSE14520 (n = 445) cohorts between normal and tumor samples (Figures 1D–I). The high expression of CCR3 and LDHA in tumor tissue is not prominent. This study shows that BMI1, CCR3, CDC25C, CFL1, and RAC1 consistently have higher expression levels, while LDHA only shows increased expression in ICGC_JP, not in TCGA_LIHC and GSE14520 cohorts. Additionally, the OS analysis revealed that patients with high expression of 6 hub genes exhibited a shorter survival time (Figures 1J–O) in the TCGA_LIHC cohort. Similarly, the Kaplan-Meier Plotter analyzed the correlation between the expression of six hub genes and survival, including OS, RFS, PFS, DSS (Figures S1B–E), and the results basically indicated that hub gene expression is associated with a poor prognosis in patients with HCC. On the other hand, the RT-qPCR assay was applied to detect the expression of hub genes in L02 normal hepatocytes and Huh7, HepG2, LM3, and MHCC-97H hepatoma cells (Figure 1P). Six hub genes were significantly up-regulated in LM3 and MHCC-97H cell lines, while only BMI1 and CDC25C were up-regulated in hepG2 cells, and CFL1 and LDHA did not show up-regulation in Huh7 cells. Additionally, the expression of hub genes was significantly upregulated upon stimulation with CCL18 in most HCC cells (Figure 1Q).

Construction and verification of prognostic model

We calculated the risk score of individual HCC patients in all cohorts using the following formula: Risk score = [BMI1 expression × (0.1971976)] + [CCR3 expression × (0.2586062)] + [CDC25C expression × (0.181479)] + [CFL1 expression × (0.2796217)] + [LDHA expression × (0.3388622)] + [RAC1 expression × (0.1223471)]. Applying the median score as the best cut-off value, patients were divided into two groups: the high- and low-risk group. The K-M survival curve analysis revealed that the risk score served as a robust prognostic indicator for HCC patients. Notably, patients with higher risk scores exhibited significantly worse prognosis compared to those in the low-risk group (TCGA_LIHC $p<0.001$; ICGC_JP $p<0.01$; GSE14520 $p<0.01$) (Figures 2A–C).

To assess the potential of the risk score as an independent prognostic factor, we conducted univariate and multivariate Cox regression analyses to examine its association with other clinical characteristics, such as age, gender, and stage. We found that age and gender were not identified as independent predictors of prognosis in HCC patients, and the risk score served as an

10.3389/fonc.2024.1371990

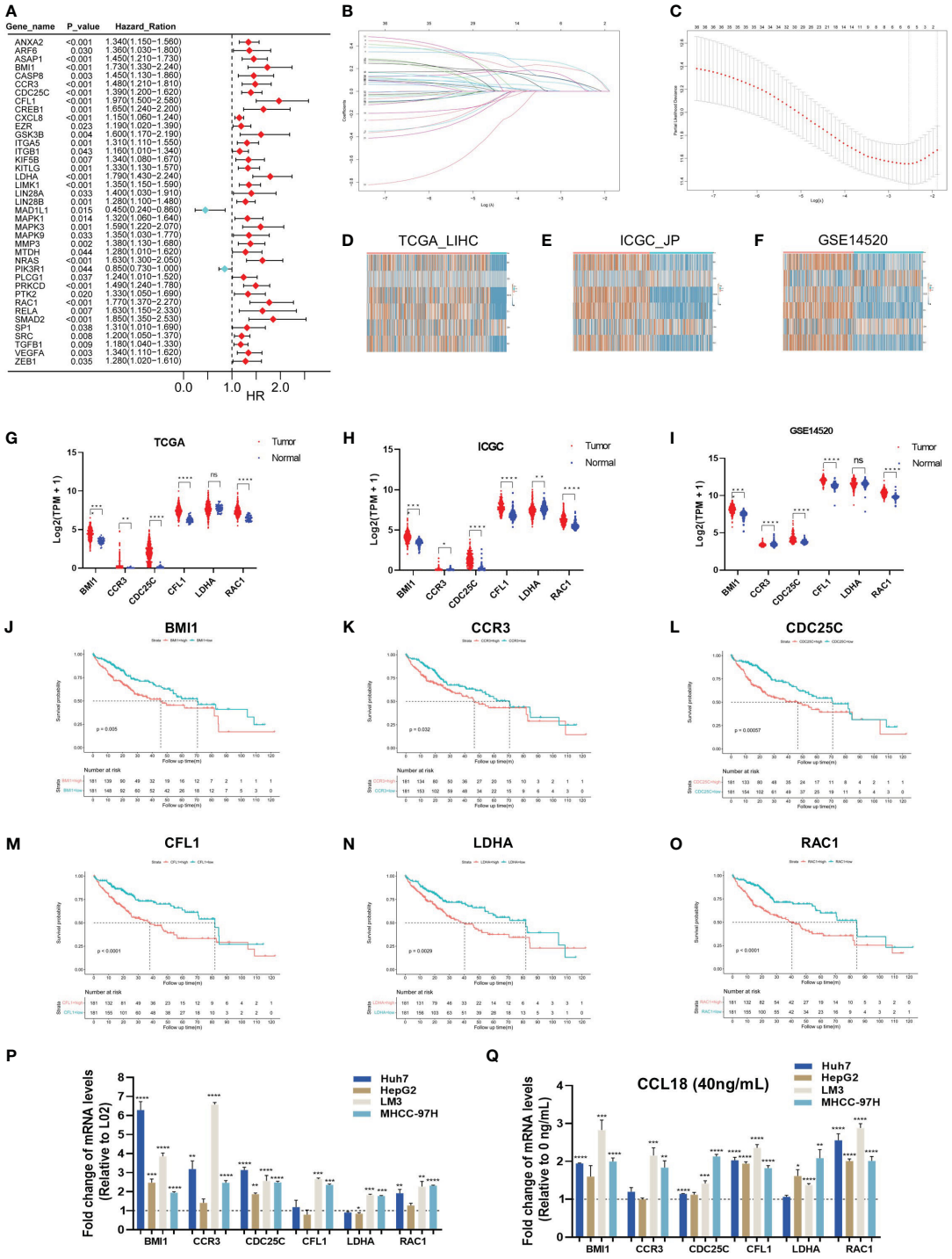


FIGURE 1
Six hub genes associated with prognosis in CCL18 signaling pathway in HCC. **(A)** Univariate Cox regression analysis for CCL18 signaling pathway-related genes in TCGA_LIHC training cohort. **(B, C)** The hub genes (n=6) were determined by the minimum lambda value of the LASSO regression analysis. **(D–I)** The heatmaps **(D–F)** and box plots **(G–I)** showed the transcription expression of six hub genes in TCGA_LIHC, ICGC_JP, and GSE14520 cohorts, consistently. **(J–O)** Survival analysis showed that all hub genes were associated with shorter survival. **(P)** RT-qPCR analysis of hub genes expression levels in L02, Huh7, HepG2, LM3 and MHCC-97H cells. **(Q)** RT-qPCR analysis of hub genes levels in Huh7, HepG2, LM3 and MHCC-97H cells after 48h of no stimulation or stimulation with 40 ng/mL CCL18. All experiments were performed with three experimental replicates, each measured with qPCR once. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

10.3389/fonc.2024.1371990

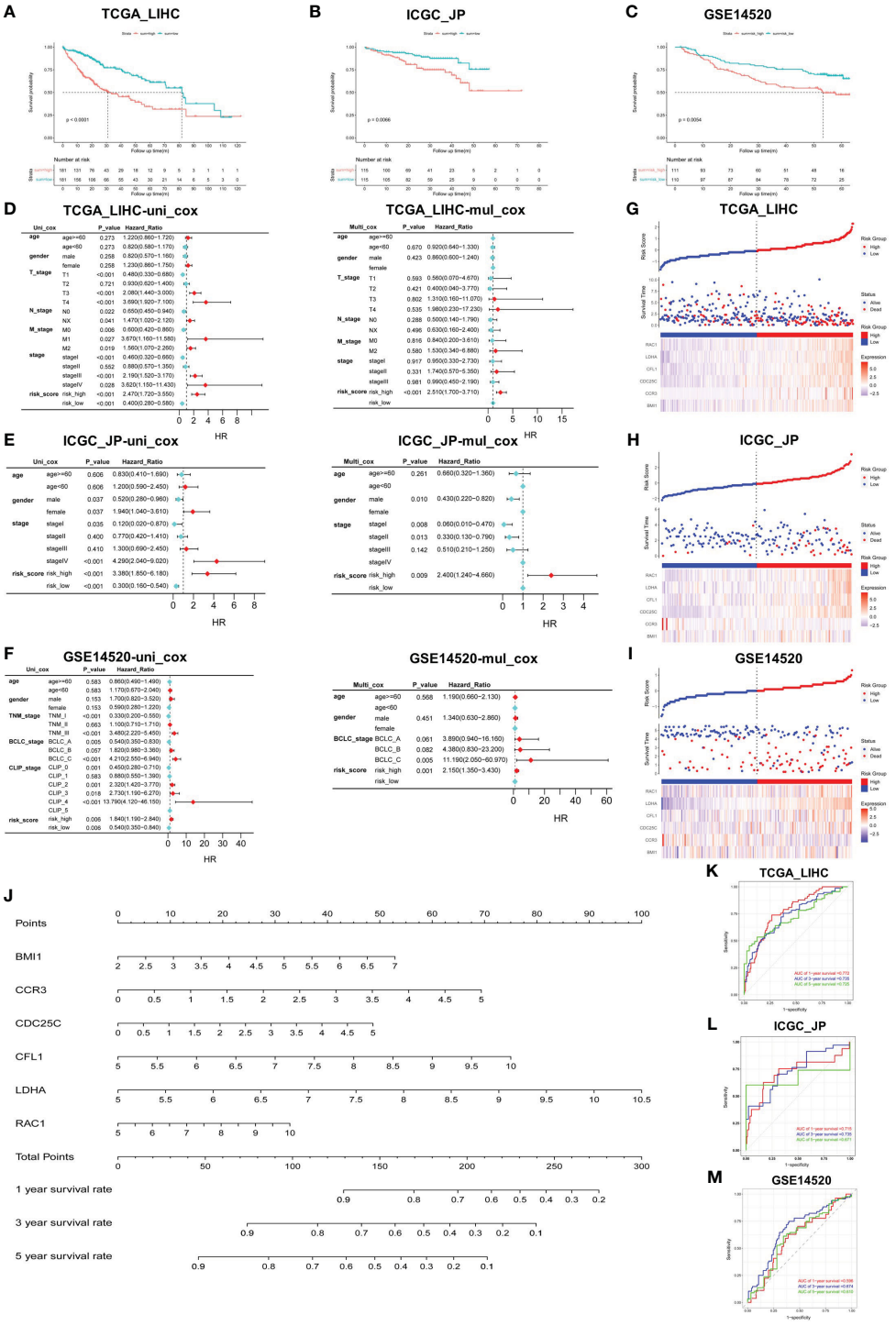


FIGURE 2
Construction and verification of prediction model. **(A–C)** Validation of risk score in TCGA_LIHC, ICGC_JP, and GSE14520 cohorts for OS. **(D–F)** Univariate and multivariate Cox regression analysis between risk score and other clinical characteristics in all cohorts. **(G–I)** Triplet graph showed the relationship between risk score, survival status, and gene expression. **(J)** The prognostic nomogram was built based on the 6 hub genes using TCGA_LIHC cohort. **(K–M)** The ROC curve for the prognostic performance of the nomogram in each cohort, including TCGA_LIHC **(K)**, ICGC_JP **(L)**, and GSE14520 **(M)**.

10.3389/fonc.2024.1371990

independent risk factor for OS in all cohorts (p<0.01) (Figures 2D–F). Risk score triptychs show the corresponding risk score, gene expression, and survival status (Figures 2G-I).

Additionally, a predictive nomogram was constructed using the TCGA_LIHC cohort, incorporating the expression levels of the six hub genes. This nomogram provided a quantitative assessment of the 1-, 3-, and 5-year survival rates for each HCC patient (Figure 2J), thereby offering potential clinical utility. The ROC curves showed the excellent predictive performance of the hub genes (Figures 2K–M), with area under the curve (AUC) values of 0.772 at 1-year, 0.735 at 3-year, and 0.725 at 5-year in the training cohort (Figure 2K). Moreover, the calibration curves were drawn to evaluate the consistency between the predictive survival possibility and the actual probability in TCGA_LIHC, ICGC_JP and GSE14520 cohorts (Figures S2A–H). These results highlight the remarkable precision and accuracy of the constructed nomogram.

Exogenous CCL18 enhanced HCC cells' proliferation, migration, invasion and stem cell-like phenotype

CCL18 exerts distinct effects in various cancer types, demonstrating its ability to enhance the proliferation of specific malignancies, such as ovarian cancer and osteosarcoma. Furthermore, CCL18 acts as an inducer, promoting tumor metastasis[13]. However, the effects of CCL18 on HCC cells have not been deeply studied. Hence, our study aimed to investigate the effects of exogenous CCL18 on HCC cells.

To investigate the biological functions of CCL18 in LM3 and MHCC-97H cells, we stimulated the cells with or without 40ng/mL CCL18 for 48h observed its impact on cell proliferation, migration, invasion and stemness properties. CCK-8 and EdU assays revealed that CCL18 promoted the proliferation abilities of LM3 and MHCC-97H cells (Figures 3A, B). Transwell migration and invasion assays showed that CCL18 significantly promoted the migration and

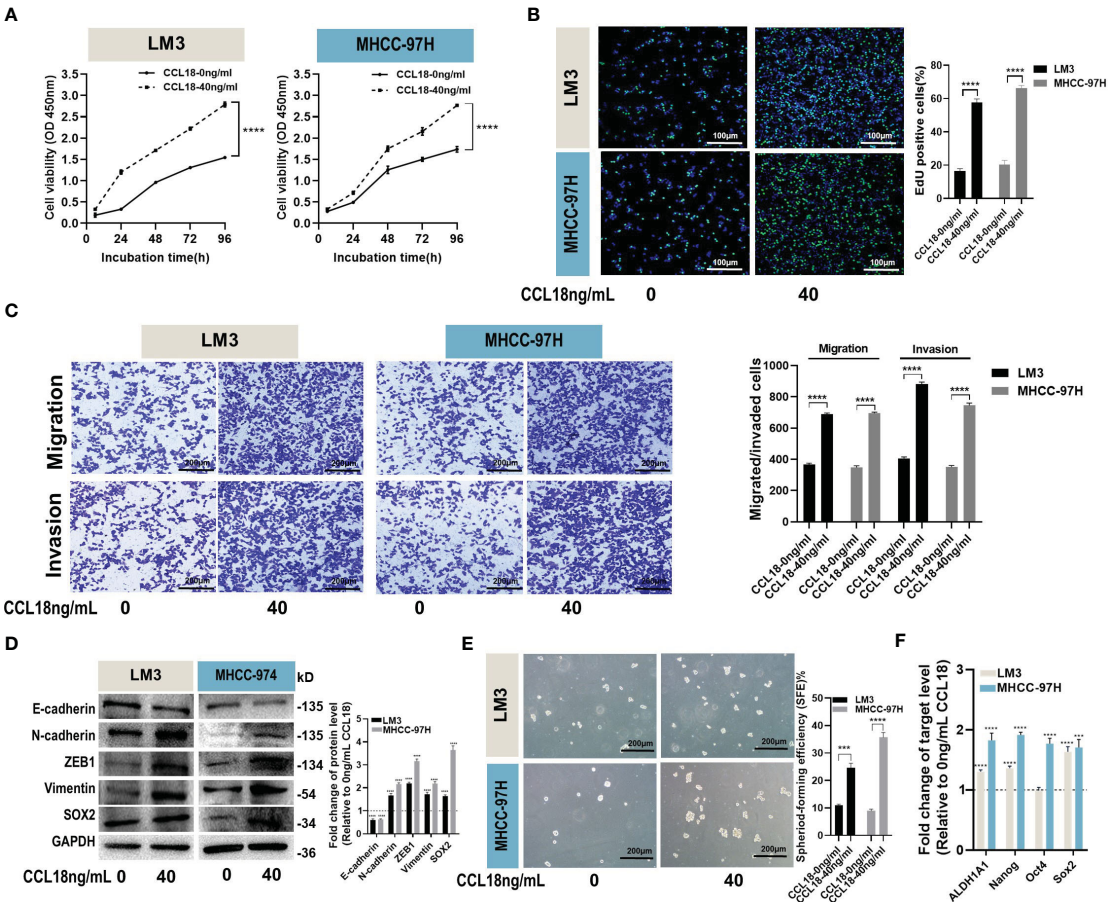


FIGURE 3
CCL18 promotes proliferation, migration, invasion and stemness properties of HCC cells *in vitro*. **(A, B)** Proliferation of LM3 and MHCC-97H cells after 48h of no stimulation or stimulation with 40 ng/mL CCL18 was examined by CCK-8 assays **(A)** and EdU assays ((**B**), Scale bar:100 μm). **(C)** Migration and invasion of LM3 and MHCC-97H after 48h of no stimulation or stimulation with 40 ng/mL CCL18 were detected by Transwell assays. Scale bar:200 μm. **(D)** The protein levels of EMT and stemness related gene in LM3 and MHCC-97H cells after 48h of no stimulation or stimulation with 40 ng/mL CCL18. **(E)** Sphere-formation abilities of LM3 and MHCC-97H cells were assessed after 48h of no stimulation or stimulation with 40 ng/mL CCL18. Scale bar:200 μm. **(F)** The expression of stemness-related genes in LM3 and MHCC-97H cells after 48h of no stimulation or stimulation with 40 ng/mL CCL18. All data are shown as the mean ± SD. **P*<0.05, ***P*<0.01, ****P*<0.001, and *****P*<0.0001 by two-tailed Student's *t*-test.

invasion abilities of LM3 and MHCC-97H cells (Figure 3C). Alterations in the expression of Epithelial-Mesenchymal Transition (EMT)-associated proteins were detected by western blotting, indicating that CCL18 facilitated the EMT process (Figure 3D). Cell stemness was determined by the sphere formation assay (Figure 3E) and the detection of stemness gene expressions (Figure 3F). The results suggest that CCL18 promotes stemness of LM3 and MHCC-97H cells. Taken together, these data collectively indicate that CCL18 may play a promoting role in HCC cell proliferation, migration, invasion, and stemness *in vitro*.

PPI construction of high-risk tumor samples and functional enrichment analysis

To investigate the genomic composition of the high-risk group in tumors, we reanalyzed the TCGA_LIHC cohort. Applying the thresholds of p< 0.05 and log2FoldChange > 1, we identified a total of 2722 DEGs in HCC tumor samples (Figure 4A) and 475 DEGs in the high-risk group (Figure 4B). By intersecting these two gene sets, we identified a total of 308 genes (Figure 4C).

Next, to better understand the function and specific mechanism of these genes, we utilized the "clusterProfiler" R package to conduct GO and KEGG enrichment analysis. The analysis of biological processes indicated a significant enrichment of genes specifically expressed in the high-risk group of tumors in processes such as "mitotic nuclear division" and "sister chromatid segregation" (Figure 4D). Furthermore, the analysis of cellular components highlighted the enrichment of these genes in cellular locations such as the "chromosomal region" and "spindle" (Figure 4E). Moreover, the molecular function analysis demonstrated a significant enrichment of these genes in functions such as "DNA replication origin binding" and "catalytic activity, acting on DNA" (Figure 4F). Additionally, the KEGG pathway analysis indicated a significant enrichment of these genes in pathways such as "cell cycle" and "DNA replication"

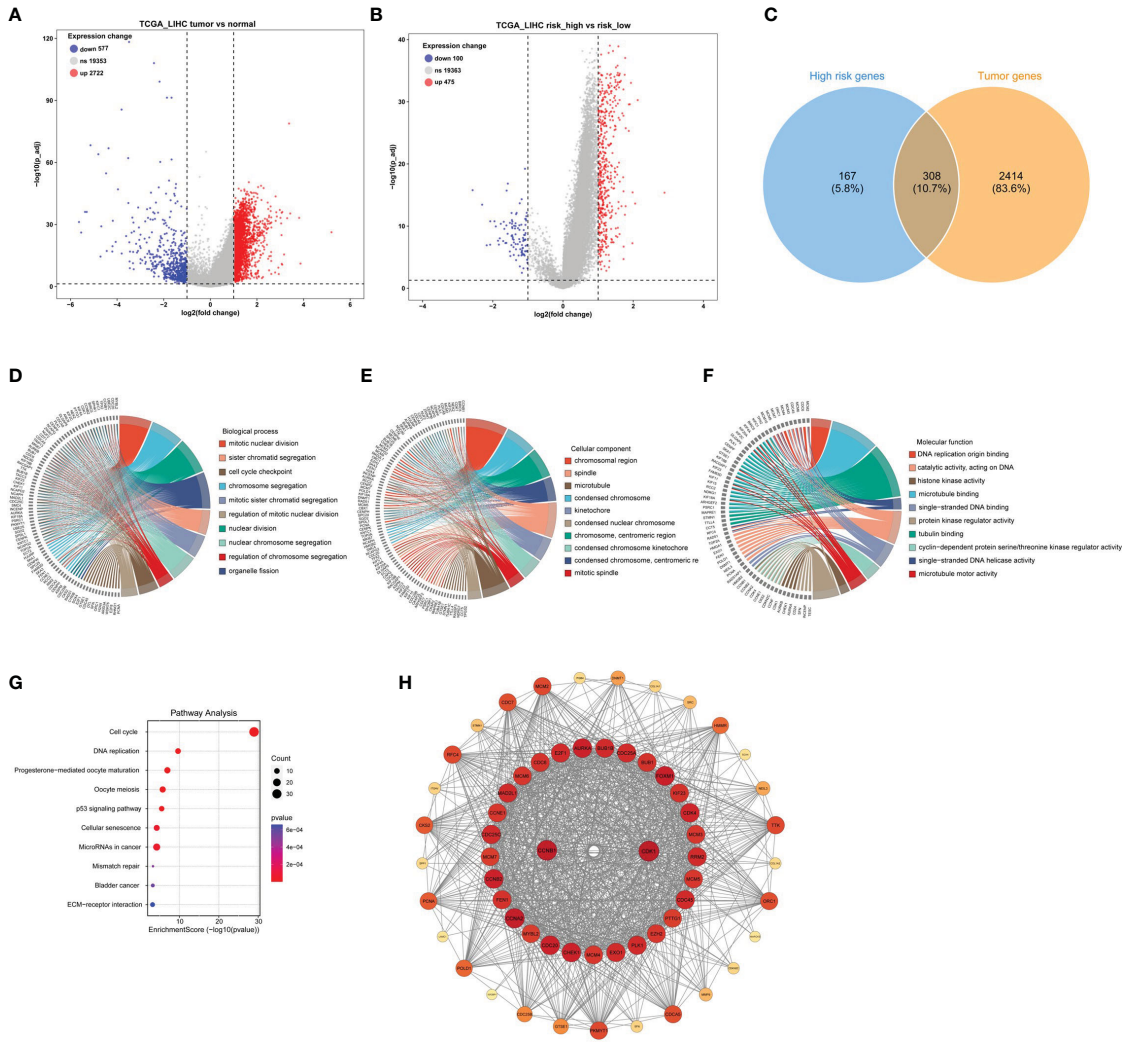


FIGURE 4
Identification of DEGs in tumors with high-risk samples and PPI functional enrichment analysis. **(A)** The volcano plot showing the DEGs in the high- and low-risk scores in TCGA_LIHC cohort. **(B)** The volcano plot revealed the DEGs in tumor and normal samples in TCGA_LIHC cohort. **(C)** The Venn plot displayed the intersection of high-risk score samples and tumor samples. **(D–G)** The chord diagrams showed the results of GO enrichment analysis, including biological process **(D)**, molecular function **(E)**, and cellular component **(F)**, and KEGG result **(G)**. **(H)** The key module analyzed by MODE included 104 genes out of the 308 genes shown by the Cytoscape software.

(Figure 4G). Furthermore, the interaction network of the 308 proteins was analyzed using the STRING database. Subsequently, the MCODE plugin identified a highly significant cluster consisting of 104 proteins (Figure 4H). This observation suggests that these 104 proteins play a crucial role as the main regulatory agents within the high-risk group of HCC.

Evaluation of immune cell infiltration

The direct chemotactic effect of CCL18 on Treg cells and its role in modulating the immunosuppressive tumor microenvironment are well-established (19). Using the ESTIMATE R package, we performed calculations of the relevant indicators for both the high- and low-risk groups. Our results revealed no significant differences in the Stromal Score and ESTIMATE Score values between the high- and low-risk groups. However, a notable disparity was observed in the Immune Score values, indicating a significantly higher level of immune cell infiltration in the tumor samples from the high-risk group (Figure 5A). This finding provides evidence that the high-risk group of tumor samples exhibits a greater extent of immune cell infiltration compared to the low-risk group.

Additionally, the CIBERSORT R package was employed to investigate the composition of immune cells and explore the correlation between hub genes and immune cells. The box plots illustrate the estimated proportions of 22 immune cell types in both the high- and low-risk groups (Figure 5B). The findings suggest that immune suppression may indeed occur in the high-risk group of HCC. This conclusion is supported by the higher relative abundance of Memory B cells, resting dendritic cells, eosinophils, M0 macrophages, neutrophils, T cells CD4 memory resting, T cells follicular helper, and Tregs observed in this group. Conversely, the low-risk group exhibits a higher relative abundance of M2 macrophages, resting mast cells, activated/resting NK cells, resting T cells CD4 memory, and T cells gamma delta. The scatter plot shows the proportions of the 22 immune cells at the individual patient level (Figure 5C). Additionally, the Pearson correlation coefficient was calculated to assess the correlation between hub genes and immune cell types (Figure 5D). The Venn diagram illustrates the high expression of 21 immune-related genes in the high-risk group of tumor samples (Figure 5E). We computed the correlation between these genes and immune cell types, and the results were similar to previous findings (Figure 5F).

Eventually, up-regulation in PD-L1 protein expression was observed in HCC cells after stimulation of CCL18 (Figure 5G) or transfection with CCR3, CDC25C, CFL1, LDHA or RAC1 plasmid (Figure 5H). Together, these results imply that CCL18 signaling pathway is associated with immune cell infiltration and immune escape in the HCC microenvironment.

Functional analysis of hub genes on proliferation, migration, invasion and stemness of HCC cells

We performed an analysis to examine the correlation between six hub genes and various signaling pathways. By applying a significance threshold of $p< 0.05$ and $|cor| > 0.3$, we successfully identified specific hub genes associated with different pathways. Within the DNA replication pathway, we found BMI1, CDC25C, and CFL1 to be the hub genes of interest. Similarly, the G2M pathway revealed the presence of BMI1, CDC25C, CFL1, and RAC1 as significant hub genes. Moving to the PI3K pathway, our analysis highlighted BMI1, CCR3, CFL1, and LDHA as the hub genes involved. Lastly, within the EMT pathway, the hub genes CCR3, CFL1, and RAC1 were found to play crucial roles (Figure 6A).

In order to elucidate the biological functions of the hub genes, various analyses including proliferation, migration, invasion, and stemness were conducted on HCC cells. Overexpression plasmids of BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 were constructed and subsequently transfected into LM3 and MHCC-97H cells. The transfection efficiency was confirmed through qPCR analysis (Figure 6B). The CCK-8 and EdU assays demonstrated that all the aforementioned genes played a role in promoting HCC cell proliferation (Figures 6C, D). Transwell migration and invasion assays suggested that overexpression of CCR3, LDHA, and RAC1 enhanced the migration and invasion abilities of both LM3 and MHCC-97H cells (Figures 6E, F). The results of the sphere formation assay revealed that all of the above genes supported the maintenance of stemness properties in HCC cells (Figure 6G), and the detection of stemness marker genes suggested that BMI1, CCR3, CDC25C, and RAC1 maintained the stemness of LM3 and MHCC-97H cells by upregulating stemness transcription factors (Figure 6H). Western blot experiments showed that overexpression of BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 could induce EMT (Figure 6I).

Together, these results imply that different hub genes are involved in different processes that contribute to HCC progression.

Identification of candidate agents in high-risk score patients

To provide better clinical recommendations, we calculated the sensitivity score of common drugs in TCGA_LIHC patients using the "OncoPredict" R package. We then calculated the Pearson correlation between the drug sensitivity score and risk score. Based on a significance level of $p<0.05$ and a correlation coefficient of $cor>0.3$ or $cor<-0.6$, we selected 14 anti-tumor drugs for horizontal lollipop mapping (Figures 7A–C). Meanwhile, we utilized the CAMP database to further predict potential small molecule compounds for the treatment of high-risk group patients. Based on the lowest scores, we identified the top 10 ranking compounds as PD-198306, fenretinide, MK-2206, wortmannin, vemurafenib, WYE-125132, BMS-754807, selumetinib, BGT-226 and GSK-269962 (Figure 7D). However, it is important to note that further experimental validation and in-depth studies are required to confirm the therapeutic potential and efficacy of these compounds.

Discussion

HCC is a highly malignant cancer that requires immediate investigation of new therapeutic strategies. Traditional first-line

10.3389/fonc.2024.1371990

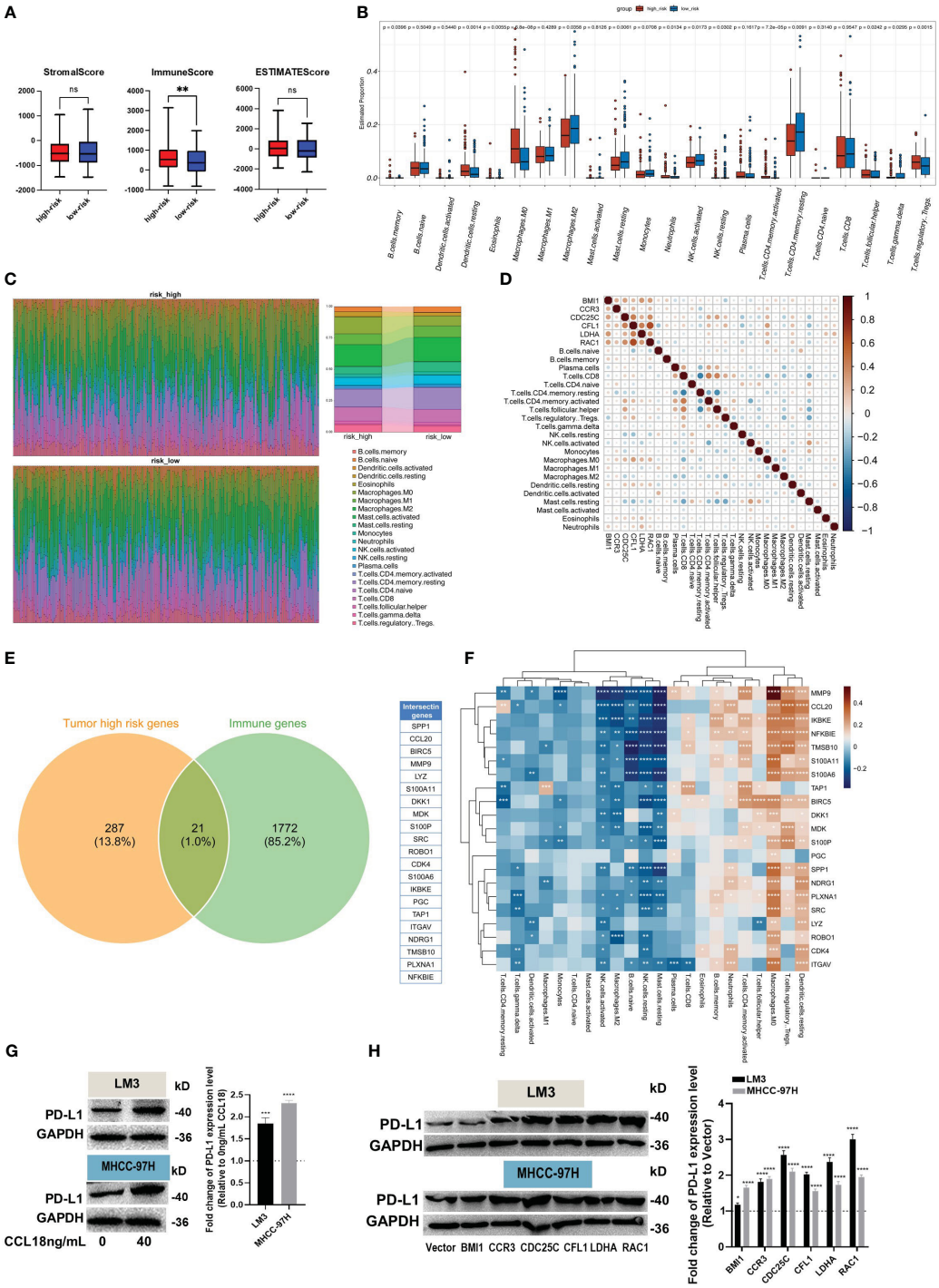


FIGURE 5
Evaluation of immune infiltration and immune escape. **(A)** The Estimate R package was used to calculate stromal, immune, and ESTIMATE scores. **(B, C)** The box plot **(B)** showed the different levels of 22 immune cell subtypes, and the relative proportion of immune cells is shown by the accumulation diagram **(C)**, both in the CIBERSORT package. **(D)** The correlation between hub genes and immune cells. **(E)** The Venn plot displays the 21 immune-related genes in high-risk samples. **(F)** The correlation heat map shows the correlation of 21 immune-related genes with 22 immune cells. **(G, H)** The protein levels of PD-L1 in LM3 and MHCC-97H cells were measured after stimulation without or with 40 ng/mL CCL18 for 48h **(G)**, or transfection with BMI1, CCR3, CDC25C, CFL1, LDHA and RAC1 overexpression plasmids for 48h **(H)**. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

10.3389/fonc.2024.1371990

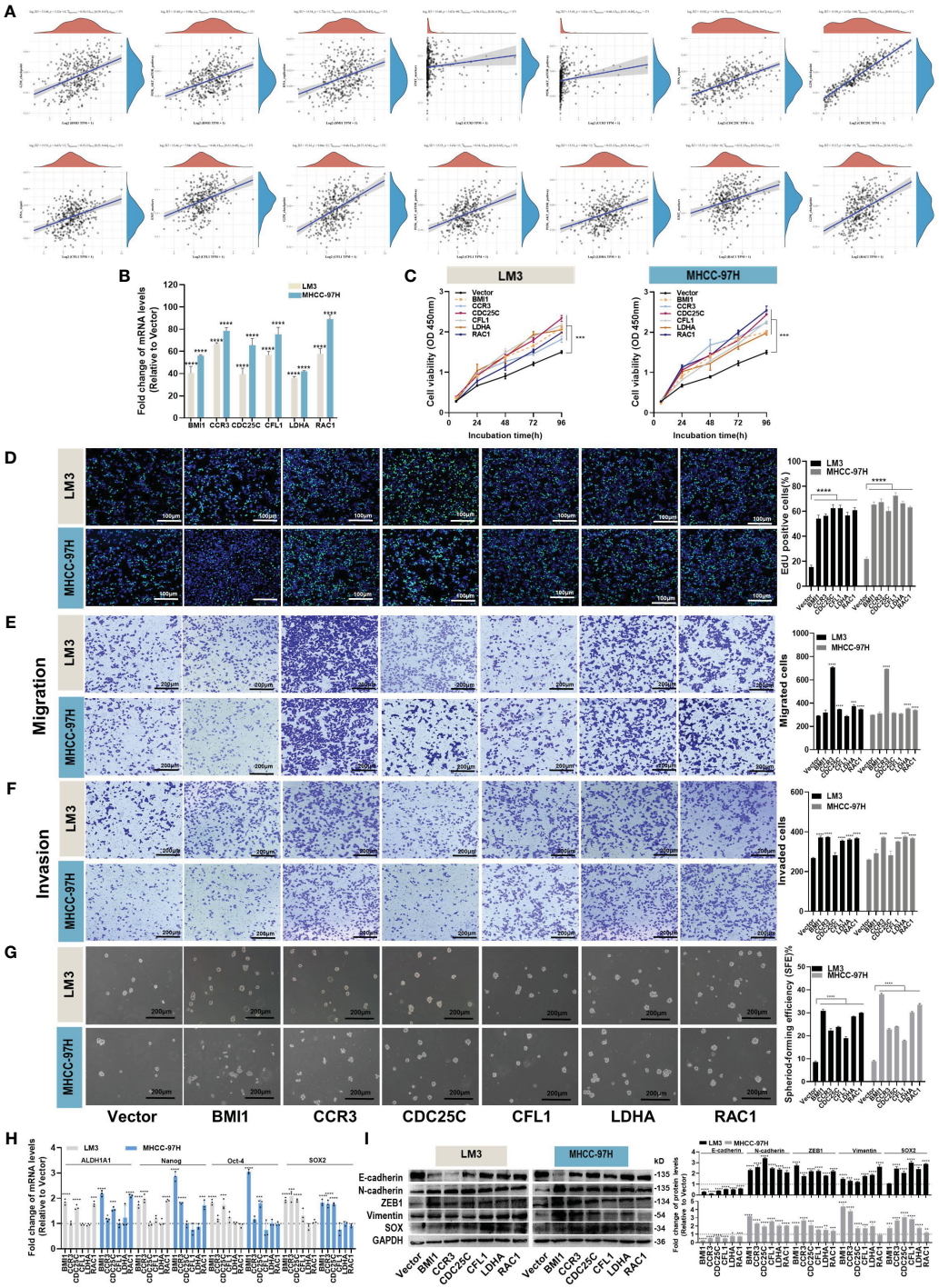
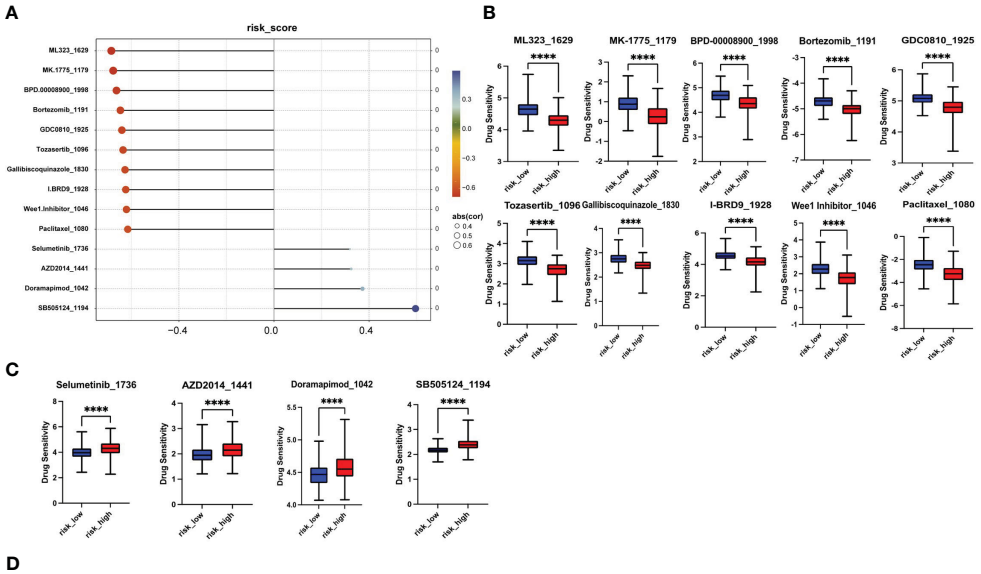


FIGURE 6
Effects of hub genes on proliferation, migration, invasion and stemness properties in HCC cells. **(A)** The correlations between six hub genes and pathway score were analyzed using Spearman. The abscissa represents the distribution of gene expression, and the ordinate represents the distribution of pathway score. The density curve on the right represents the trend in the distribution of pathway immune score, while the upper density curve represents the trend in the distribution of gene expression. The value on the top represents the correlation p value, correlation coefficient and correlation calculation method. **(B)** The overexpression efficiency of BMI1, CCR3, CDC25C, CFL1, LDHA and RAC1 overexpression plasmid was assessed by RT-qPCR. **(C, D)** Proliferation of LM3 and MHCC-97H cells after transfection with BMI1, CCR3, CDC25C, CFL1, LDHA and RAC1 overexpression plasmid for 48h was examined by CCK-8 assays **(C)** and EdU assays (**(D)**, Scale bar:100 μm). **(E, F)** Migration **(E)** and invasion **(F)** of LM3 and MHCC-97H cells were detected after transfection with BMI1, CCR3, CDC25C, CFL1, LDHA and RAC1 overexpression plasmid for 48h were detected by Transwell assays. Scale bar:200 μm. **(G)** Sphere-formation abilities of LM3 and MHCC-97H cells were observed after transfection with BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 overexpression plasmids for 48 hours. Scale bar: 200 μm. **(H)** The expression of stemness-related genes in LM3 and MHCC-97H cells was analyzed after transfection with BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 overexpression plasmids for 48h. **(I)** The protein levels of EMT-related genes in LM3 and MHCC-97H cells were measured after transfection with BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 overexpression plasmids for 48h. All data are shown as the mean ± SD. *P<0.05, **P<0.01, ***P<0.001, and ****P<0.0001 by two-tailed Student's t-test.

10.3389/fonc.2024.1371990



D

Pert_iname	MOA	Target_name	FDR_q_nlog10	Norm_cs
PD-198306	MAP kinase inhibitor|MEK inhibitor	MAPK1|MAPK3|MAP2K1|MAP2K2	15.6536	-2.1246
fenretinide	Apoptosis stimulant|Retinoid receptor agonist	RARA|CYP2C8|CYP3A5|RARG	15.6536	-2.0524
MK-2206	AKT inhibitor	AKT1|AKT2|AKT3	15.6536	-1.9997
wortmannin	PI3K inhibitor	PLK1|PIK3CA|PIK3CG|ATM|ATR|MTOR|PLK3|PI4KA|PI4KB|PIK3CD|PIK3R1|PRKDC	15.6536	-1.9108
vemurafenib	RAF inhibitor	BRAF|CYP2C19|CYP3A4|CYP3A5|RAF1	15.6536	-1.8593
WYE-125132	MTOR inhibitor	MTOR|PIK3CA	15.6536	-1.8442
BMS-754807	IGF-1 inhibitor	IGF1R|AKT1	15.6536	-1.8418
selumetinib	MEK inhibitor	MAP2K1|MAP2K2	15.6536	-1.8276
BGT-226	PI3K inhibitor	MTOR|PIK3CG|PIK3CA|PIK3CB	15.6536	-1.8009
GSK-269962	Rho associated kinase inhibitor	ROCK1|ROCK2	15.6536	-1.793

FIGURE 7
Screening of small molecule compounds with potential therapeutic effects in high-risk group. **(A)** The lollipop chart shows the results of OncoPredict R package according to p<0.05, cor>0.3 or cor<-0.6. **(B)** Potential drugs for the treatment of high-risk HCC patients. **(C)** Potentially inappropriate drugs for the treatment of high-risk HCC patients. **(D)** The top 10 compounds with the highest negative scores according to CAMP. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

treatment drugs, such as Sorafenib and Lenvatinib, have only shown slight improvements in OS, with an extension of approximately 2.8 and 4.4 months (3). Over the past five years, significant advancements have been made in the field of immunotherapy, specifically in ICIs. However, in advanced HCC patients, monotherapy with ICIs has only demonstrated objective response rates of 15-20%, without any significant improvement in OS. Furthermore, specific biomarkers for this subgroup of patients have yet to be identified (20). The tumor microenvironment of HCC is characterized by a significant presence of non-tumor stromal cells, including cancer-associated fibroblasts, endothelial cells, TAMs, B cells, and T cells. These cells play crucial roles in the progression of cancer (7). Among these cells, our particular focus lies on the role of macrophages.

CCL18, also known as macrophage inflammatory protein 4 (MIP-4), pulmonary and activation-regulated chemokine (PARC), dendritic cell chemokine 1 (DC-CK1), and alternative macrophage activation-associated CC chemokine 1 (AMAC-1), belongs to the family of CC chemokines and acts as a chemoattractant. CCL18 is located on chromosome 17 in the human genome and shares the highest amino acid identity (65%) with CCL3. It encodes a protein consisting of 89 amino acids, with the mature active form comprising 69 amino acids without a terminal alanine at the C-terminus (13, 21). In recent years, more and more studies have revealed the existence of multiple subtypes within M1 and M2 macrophages, including further subdivisions of M2 macrophages into M2a, M2b, M2c, and M2d subgroups (14). However, the specific M2 macrophage subtype responsible for secreting CCL18 remains to be conclusively determined. The progression of cancer necessitates evasion of immune surveillance, and CCL18, which is secreted by M2 macrophages, serves as a hallmark of macrophage activation. It acts as a chemotactic factor that promotes immune suppression and immune escape, thereby facilitating tumor development (13). Studies have demonstrated the crucial role of CCL18 in the progression of fibrotic immune diseases (22) and tumors. CCL18 has been shown to promote immunosuppressive states and progression in esophageal squamous cell carcinoma (23), multiple myeloma (24), osteosarcoma (25), ovarian cancer (26), and renal cell carcinoma (27). In a word, the CCL18 signaling pathway has demonstrated its prognostic significance in patients with tumors.

Several gene prognostic models related to signaling pathways have been reported in HCC, including STING pathway genes (28), hypoxia-related and immune-associated genes (29), and chromatin

organization-related genes (30). However, the biological functions and prognostic impact of CCL18 signaling pathway genes in HCC remain largely unknown. In this study, we acquired CCL18 signaling pathway-associated genes from the Wikipathway website and obtained RNA-seq sequencing data of HCC from publicly available databases such as TCGA, ICGC, and GEO. By employing a diverse range of well-established bioinformatics methodologies, we successfully identified six key genes (BMI1, CCR3, CDC25C, CFL1, LDHA, RAC1) that have a substantial impact on prognosis. Subsequently, we validated the independent prognostic value of the expression of these genes in predicting the prognosis of HCC patients through multivariate Cox regression analysis. Furthermore, we utilized these genes to construct a nomogram that enables the prediction of patients' OS rates at 1-, 3-, and 5- years. In summary, these findings provide robust evidence supporting the prognostic evaluation and personalized treatment of HCC, thereby contributing to the enhancement of patients' survival rates and treatment efficacy. Additionally, we analyzed the core protein network of highly expressed proteins in the high-risk group and compared the differences in immune cell infiltration between the high- and low-risk groups. The results revealed a higher proportion of immune-inhibitory cells in the high-risk group, suggesting that the overexpression of these hub genes indeed induces immune suppression in tumors. In the experimental section, we investigated the effects of exogenous CCL18 on the biological functions of LM3 and MHCC-97H cells. We observed that exogenous CCL18 promoted cell proliferation, migration, invasion, and stemness. Additionally, by overexpressing hub genes in the cell lines, we identified that BMI1, CCR3, CDC25C, CFL1, LDHA, and RAC1 participated in promoting different functions of HCC cells. Furthermore, we discovered that both exogenous CCL18 and overexpression of CCR3, CDC25C, CFL1, LDHA, and RAC1 could induce the expression of PDL1 in the HCC cell lines, which is consistent with the occurrence of immune suppression in HCC. However, the molecular processes and regulatory mechanisms underlying these findings require further investigation in future studies.

BMI1, a polycomb-group protein, is involved in the regulation of embryonic development and DNA damage repair. It is also an oncogene, with dysregulated expression frequently associated with various cancers. H. Wang et al. discovered that exogenous CCL18 can promote the up-regulation of OCT4 and BMI1 mRNA and protein expression (31). Moreover, up-regulation of BMI1 expression in HCC has been linked to its role in blocking the INK4a/ARF locus, NF-κB signaling pathway, and TGFβ2/SMAD signaling axis, while simultaneously activating the Wnt/β-catenin signaling axis, thereby promoting the development of HCC (32). Inhibition of BMI1 has been shown to enhance immune checkpoint blockade in CCA cells (33). Additionally, BMI1 has been implicated in promoting breast cancer (34) and endometrial cancer (35).

The role of CCR3 in tumor cells is relatively limited, as it is primarily highly expressed in inflammatory cells such as mast cells, eosinophils, basophils, and Th2 cells. It plays a significant role in inflammatory responses (36). CCL18 acts as a neutral CCR3 antagonist. CCR3 exerts its functions through various ligands. In breast cancer, it promotes cancer progression through the CCL5-CCR3 axis (37). Compared to primary prostate tumors, CCR3 exhibits high expression in bone and visceral metastases, potentially exerting its effects via the CCR3/CCL7 axis (38). In renal cell carcinoma, CCR3 facilitates tumor proliferation and metastasis through the CCL11/CCR3 axis (39).

CDC25C is one of the three isoforms of the CDC25 phosphatase family, and it plays a crucial role in regulating the G2/M transition and mediating DNA damage repair during cell division. Extensive research has demonstrated that abnormal expression of CDC25C is associated with the progression of various types of cancer (40). CFL1, a 166-amino acid phosphoprotein, is one of the five components representing actin-binding proteins. It regulates the polymerization and depolymerization of F-actin and G-actin (41). Similarly, CFL1 also contributes to the proliferation, invasion, and metastasis of malignant tumors (42–44). The lactate dehydrogenase isoenzymes (LDH) are tetramers composed of LDHA and LDHB, and their aberrant expression is often associated with cellular metabolism and tumor progression. Chen et al. discovered that the expression of LDHA was upregulated and LDHB was downregulated in prostate cancer cells by exogenous CCL18 at both mRNA and protein levels (45). In addition, LDHA can also serve as a biomarker for various malignant tumors (46–48). RAC1, a small GTP-binding protein, belongs to the Rac subfamily of the Rho GTPase family, and it is involved in various biological functions, including regulating cell migration, signal transduction, and promoting cell polarization. Lihong Shi et al. discovered that elevated levels of CCL18 promote lymph node metastasis and distant metastasis in NSCLC patients. They demonstrated that CCL18 activates RAC1 to regulate cellular migration and invasion, ultimately leading to cytoskeletal remodeling *in vitro* (49). RAC1 has been extensively discussed for its role in promoting proliferation, participating in angiogenesis, facilitating tumor migration and invasion, as well as its involvement in stemness in tumor cells (50).

Immunotherapy remains a promising trend for HCC patients in the future. In our study, we found that exogenous CCL18, as well as CCR3, CDC25C, CFL1, LDHA or RAC1 plasmid could promote the production of PD-L1 protein in LM3 and MHCC-97H cells. This suggests that immune suppression may occur as a result. Furthermore, the high-risk group derived from the six hub genes also exhibits a positive correlation with immune-suppressive cells such as Treg cells.

Our study provides the first systematic elucidation of the six hub genes in the CCL18 signaling pathway that impact the prognosis of HCC patients. Additionally, we constructed a protein-protein interaction network of key proteins in the high-risk group and analyzed the immune cell infiltration in the high-risk group. These findings contribute to our understanding of immune evasion genes in HCC.

Our study has several limitations. Firstly, in terms of variable selection, it would be advantageous to explore a range of machine learning methods, such as random forest and support vector machines, to enhance the accuracy of our analysis. Secondly, among the identified hub genes, CDC25C, CFL1, and LDHA are

primarily associated with abnormal physiological activities, such as cytoskeletal dynamics, cellular respiration, and cell replication and proliferation, rather than directly interacting with the CCL18 chemokine. Lastly, in our cellular experiments, we have primarily focused on functional experiments related to gene overexpression. To further deepen our understanding, future investigations should include knockdown and rescue experiments.

Conclusion

In summary, our study conducted a comprehensive analysis of the prognostic impact of six hub genes within the CCL18 signaling pathway in HCC patients. Our investigation demonstrated that exogenous CCL18 enhances key oncogenic processes in HCC cell lines LM3 and MHCC-97H, including proliferation, migration, invasion, and up-regulation of the immune-suppressive marker PD-L1 protein. We also investigated the functions of six key genes, revealing their potential involvement in liver cancer development. We further identified 21 immune-related genes that exhibit strong correlations with immune suppressive cells. Collectively, these findings significantly contribute to our understanding of immune evasion within the tumor microenvironment and the underlying oncogenic processes in HCC patients.

Data availability statement

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

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used. The manuscript does not involve any animal experimentation.

Author contributions

JLM: Writing – review & editing. YHT: Writing – review & editing. KKW: Writing – review & editing. HRS: Writing – review & editing. MQZ: Writing – review & editing. LJ: Writing – review & editing. YP: Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (No. 82073227, 82373925 and No. 82070801).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1371990/full#supplementary-material

References

1. Brown ZJ, Tsilimigras DI, Ruff SM, Mohseni A, Kamel IR, Cloyd JM, et al. Management of hepatocellular carcinoma: A review. *JAMA Surg*. (2023) 158:410–20. doi: 10.1001/jamasurg.2022.7989

2. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin*. (2023) 73:17–48. doi: 10.3322/caac.21763

3. Vogel A, Meyer T, Sapisochin G, Salem R and Saborowski A. Hepatocellular carcinoma. *Lancet*. (2022) 400:1345–62. doi: 10.1016/S0140-6736(22)01200-4

4. Vij M, Calderaro J. Pathologic and molecular features of hepatocellular carcinoma: An update. *World J Hepatol*. (2021) 13:393–410. doi: 10.4254/wjh.v13.i4.393

5. Zhu AX, Abbas AR, de Galarreta MR, Guan Y, Lu S, Koeppen H, et al. Molecular correlates of clinical response and resistance to atezolizumab in combination with bevacizumab in advanced hepatocellular carcinoma. *Nat Med*. (2022) 28:1599–611. doi: 10.1038/s41591-022-01868-2

6. Ruf B, Heinrich B and Greten TF. Immunobiology and immunotherapy of HCC: spotlight on innate and innate-like immune cells. *Cell Mol Immunol*. (2021) 18:112–27. doi: 10.1038/s41423-020-00572-w

7. Donne R, Lujambio A. The liver cancer immune microenvironment: Therapeutic implications for hepatocellular carcinoma. *Hepatology*. (2023) 77:1773–96. doi: 10.1002/hep.32740

8. Chen C, Wang Z, Ding Y and Qin Y. Tumor microenvironment-mediated immune evasion in hepatocellular carcinoma. *Front Immunol*. (2023) 14:1133308. doi: 10.3389/fimmu.2023.1133308

9. Cheng K, Cai N, Zhu J, Yang X, Liang H and Zhang W. Tumor-associated macrophages in liver cancer: From mechanisms to therapy. *Cancer Commun (Lond)*. (2022) 42:1112–40. doi: 10.1002/cac2.12345

10. Huang Y, Ge W, Zhou J, Gao B, Qian X and Wang W. The role of tumor associated macrophages in hepatocellular carcinoma. *J Cancer*. (2021) 12:1284–94. doi: 10.7150/jca.51346

11. Elliott LA, Doherty GA, Sheahan K and Ryan EJ. Human tumor-infiltrating myeloid cells: phenotypic and functional diversity. *Front Immunol*. (2017) 8:86. doi: 10.3389/fimmu.2017.00086

12. Song G, Shi Y, Zhang M, Goswami S, Afridi S, Meng L, et al. Global immune characterization of HBV/HCV-related hepatocellular carcinoma identifies macrophage

and T-cell subsets associated with disease progression. *Cell Discovery*. (2020) 6:90. doi: 10.1038/s41421-020-00214-5

13. Korbecki J, Olbromski M and Dziegiel P. CCL18 in the progression of cancer. *Int J Mol Sci*. (2020) 21:7955. doi: 10.3390/ijms21217955

14. Qin R, Ren W, Ya G, Wang B, He J, Ren S, et al. Role of chemokines in the crosstalk between tumor and tumor-associated macrophages. *Clin Exp Med*. (2023) 23:1359–73. doi: 10.1007/s10238-022-00888-z

15. Lin Z, Li W, Zhang H, Wu W, Peng Y, Zeng Y, et al. CCL18/PITPNM3 enhances migration, invasion, and EMT through the NF-kappaB signaling pathway in hepatocellular carcinoma. *Tumor Biol*. (2016) 37:3461–8. doi: 10.1007/s13277-015-4172-x

16. Wei J, Huang K, Chen Z, Hu M, Bai Y, Lin S, et al. Characterization of glycolysis-associated molecules in the tumor microenvironment revealed by pan-cancer tissues and lung cancer single cell data. *Cancers (Basel)*. (2020) 12:1788. doi: 10.3390/cancers12071788

17. Li X, Yang J, Ni R, Chen J, Zhou Y, Song H, et al. Hypoxia-induced lncRNA RBM5-AS1 promotes tumorigenesis via activating Wnt/beta-catenin signaling in breast cancer. *Cell Death Dis*. (2022) 13:95. doi: 10.1038/s41419-022-04536-y

18. Tang T, Zhu Q, Li X, Zhu G, Deng S, Wang Y, et al. Protease Nexin I is a feedback regulator of EGF/PKC/MAPK/EGR1 signaling in breast cancer cells metastasis and stemness. *Cell Death Dis*. (2019) 10:649. doi: 10.1038/s41419-019-1882-9

19. Cardoso AP, Pinto ML, Castro F, Costa AM, Marques-Magalhaes A, Canha-Borges A, et al. The immunosuppressive and pro-tumor functions of CCL18 at the tumor microenvironment. *Cytokine Growth Factor Rev*. (2021) 60:107–19. doi: 10.1016/j.cytogfr.2021.03.005

20. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers*. (2021) 7:6. doi: 10.1038/s41572-020-00240-3

21. Guan P, Burghes AH, Cunningham A, Lira P, Brissette WH, Neote K, et al. Genomic organization and biological characterization of the novel human CC chemokine DC-CK-1/PARC/MIP-4/SCYA18. *Genomics*. (1999) 56:296–302. doi: 10.1006/geno.1998.5635

22. Zeng W, Xiong L, Wu W, Li S, Liu J, Yang L, et al. CCL18 signaling from tumor-associated macrophages activates fibroblasts to adopt a chemoresistance-inducing phenotype. *Oncogene*. (2023) 42:224–37. doi: 10.1038/s41388-022-02540-2

23. Sui X, Chen C, Zhou X, Wen X, Shi C, Chen G, et al. Integrative analysis of bulk and single-cell gene expression profiles to identify tumor-associated macrophage-derived CCL18 as a therapeutic target of esophageal squamous cell carcinoma. *J Exp Clin Cancer Res*. (2023) 42:51. doi: 10.1186/s13046-023-02612-5

24. Qiao B, Chen L, Cheng Q, Wang G, Li Q, Zhang B, et al. CCL18 promotes migration and invasion of multiple myeloma cells and is associated with poor prognosis. *Carcinogenesis*. (2023) 44:38–45. doi: 10.1093/carcin/bgac097

25. Li C, Xiang F, Gong Y, Fu Y, Chen G, Wang Z, et al. Tumor-derived microparticles promoted M2-like macrophages polarization to stimulate osteosarcoma progression. *Int J Biochem Cell Biol*. (2023) 166:106494. doi: 10.1016/j.biocel.2023.106494

26. Liu S, Tao Z, Lou J, Li R, Fu X, Xu J, et al. CD4(+)CCR8(+) Tregs in ovarian cancer: a potential effector Tregs for immune regulation. *J Transl Med*. (2023) 21:803. doi: 10.1186/s12967-023-04686-3

27. Nguyen TN, Nguyen-Tran HH, Chen CY, Hsu T. IL6 and CCL18 mediate cross-talk between VHL-deficient kidney cells and macrophages during development of renal cell carcinoma. *Cancer Res*. (2022) 82:2716–33. doi: 10.1158/0008-5472.CAN-21-3749

28. Pu Z, Liu J, Liu Z, Peng F, Zhu Y, Wang X, et al. STING pathway contributes to the prognosis of hepatocellular carcinoma and identification of prognostic gene signatures correlated to tumor microenvironment. *Cancer Cell Int*. (2022) 22:314. doi: 10.1186/s12935-022-02734-4

29. Hu B, Yang XB, Sang XT. Development and verification of the hypoxia-related and immune-associated prognosis signature for hepatocellular carcinoma. *J Hepatocell Carcinoma*. (2020) 7:315–30. doi: 10.2147/JHC.S272109

30. Chen J, Chen X, Li T, Wang L and Lin G. Identification of chromatin organization-related gene signature for hepatocellular carcinoma prognosis and predicting immunotherapy response. *Int Immunopharmacol*. (2022) 109:108866. doi: 10.1016/j.intimp.2022.108866

31. Wang H, Liang X, Li M, Tao X, Tai S, Fan Z, et al. Chemokine (CC motif) ligand 18 upregulates Slug expression to promote stem-cell like features by activating the mammalian target of rapamycin pathway in oral squamous cell carcinoma. *Cancer Sci*. (2017) 108:1584–93. doi: 10.1111/cas.13289

32. Wang R, Fan H, Sun M, Lv Z and Yi W. Roles of BMI1 in the initiation, progression, and treatment of hepatocellular carcinoma. *Technol Cancer Res Treat*. (2022) 21:15330338211070689. doi: 10.1177/15330338211070689

33. Liu Z, Hu C, Zheng L, Liu J, Li K, Li X, et al. BMI1 promotes cholangiocarcinoma progression and correlates with antitumor immunity in an exosome-dependent manner. *Cell Mol Life Sci*. (2022) 79:469. doi: 10.1007/s00018-022-04500-1

34. Srinivasan M, Bharali DJ, Sudha T, Khedr M, Guest I, Sell S, et al. Downregulation of Bmi1 in breast cancer stem cells suppresses tumor growth and proliferation. *Oncotarget*. (2017) 8:38731–42. doi: 10.18632/oncotarget.16317

35. Zhao Y, Yang W, Zheng K, Chen J and Jin X. The role of BMI1 in endometrial cancer and other cancers. *Gene*. (2023) 856:147129. doi: 10.1016/j.gene.2022.147129

36. Tang S, Shu X. [Effect of CCR3 gene on related inflammatory cells in respiratory allergic diseases]. *Lin Chuang Er Bi Yan Hou Tou Jing Wai Ke Za Zhi*. (2021) 35:80–4. doi: 10.13201/j.issn.2096-7993.2021.01.021

37. Yamaguchi M, Takagi K, Narita K, Miki Y, Onodera Y, Miyashita M, et al. Stromal CCL5 promotes breast cancer progression by interacting with CCR3 in tumor cells. *Int J Mol Sci*. (2021) 22:1918. doi: 10.3390/ijms22041918

38. Guerard A, Laurent V, Fromont G, Esteve D, Gilhodes J, Bonnelye E, et al. The chemokine receptor CCR3 is potentially involved in the homing of prostate cancer cells to bone: Implication of bone-marrow adipocytes. *Int J Mol Sci*. (2021) 22:1994. doi: 10.3390/ijms22041994

39. Johrer K, Zelle-Rieser C, Perathoner A, Moser P, Hager M, Ramoner R, et al. Up-regulation of functional chemokine receptor CCR3 in human renal cell carcinoma. *Clin Cancer Res*. (2005) 11:2459–65. doi: 10.1158/1078-0432.CCR-04-0405

40. Liu K, Zheng M, Lu R, Du J, Zhao Q, Li Z, et al. The role of CDC25C in cell cycle regulation and clinical cancer therapy: a systematic review. *Cancer Cell Int*. (2020) 20:213. doi: 10.1186/s12935-020-01304-w

41. Halder SK, Rafi MO, Shahriar EB, Albogami S, El-Shehawi AM, Daullah S, et al. Identification of the most damaging nsSNPs in the human CFL1 gene and their functional and structural impacts on cofilin-1 protein. *Gene*. (2022) 819:146206. doi: 10.1016/j.gene.2022.146206

42. Yao B, Li Y, Chen T, Niu Y, Wang Y, Yang Y, et al. Hypoxia-induced cofilin 1 promotes hepatocellular carcinoma progression by regulating the PLD1/AKT pathway. *Clin Transl Med*. (2021) 11:e366. doi: 10.1002/ctm2.366

43. Li ZF, Yao YD, Zhao YY, Liu Y, Liu ZH, Hu P, et al. Effects of PAK4/LIMK1/Cofilin-1 signaling pathway on proliferation, invasion, and migration of human osteosarcoma cells. *J Clin Lab Anal*. (2020) 34:e23362. doi: 10.1002/jcla.23362

44. Kang X, Zhao C, Liu Y and Wang G. The phosphorylation level of Cofilin-1 is related to the pathological subtypes of gastric cancer. *Med (Baltimore)*. (2022) 101:e31309. doi: 10.1097/MD.0000000000031309

45. Chen G, Cai ZD, Lin ZY, Wang C, Liang YX, Han ZD, et al. ARNT-dependent CCR8 reprogrammed LDH isoform expression correlates with poor clinical outcomes of prostate cancer. *Mol Carcinog*. (2020) 59:897–907. doi: 10.1002/mc.23201

46. Bolanos-Suarez V, Alfaro A, Espinosa AM, Medina-Martinez I, Juarez E, Villegas-Sepulveda N, et al. The mRNA and protein levels of the glycolytic enzymes lactate dehydrogenase A (LDHA) and phosphofructokinase platelet (PFKP) are good predictors of survival time, recurrence, and risk of death in cervical cancer patients. *Cancer Med*. (2023) 12:15632–49. doi: 10.1002/cam4.6123

47. Comandatore A, Franczak M, Smolenski RT, Morelli L, Peters GJ, Giovannetti E. Lactate Dehydrogenase and its clinical significance in pancreatic and thoracic cancers. *Semin Cancer Biol*. (2022) 86:93–100. doi: 10.1016/j.semcancer.2022.09.001

48. Wu Y, Kou Q, Sun L and Hu X. Effects of anoxic prognostic model on immune microenvironment in pancreatic cancer. *Sci Rep*. (2023) 13:9104. doi: 10.1038/s41598-023-36413-9

49. Shi L, Zhang B, Sun X, Zhang X, Lv S, Li H, et al. CC chemokine ligand 18 (CCL18) promotes migration and invasion of lung cancer cells by binding to Nir1 through Nir1-ELMO1/DOC180 signaling pathway. *Mol Carcinog*. (2016) 55:2051–62. doi: 10.1002/mc.22450

50. Liang J, Oyang L, Rao S, Han Y, Luo X, Yi P, et al. Rac1, A potential target for tumor therapy. *Front Oncol*. (2021) 11:674426. doi: 10.3389/fonc.2021.674426

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 22 May 2024
DOI 10.3389/fimmu.2024.1370367

OPEN ACCESS

EDITED BY
Manoj K. Pandey,
Cooper Medical School of Rowan University, United States

REVIEWED BY
Rajesh Mani,
University of Kentucky, United States
Jiaxiong Tan,
Tianjin Medical University Cancer Institute and Hospital, China

*CORRESPONDENCE
Junchang Zhang
zhangjunchang1992@163.com
Yong Li
liyong@gdph.org.cn

†These authors have contributed equally to this work

RECEIVED 14 January 2024
ACCEPTED 03 May 2024
PUBLISHED 22 May 2024

CITATION
Wang H, Zheng J, Ma Q, Zhang J and Li Y (2024) GLT8D2 is a prognostic biomarker and regulator of immune cell infiltration in gastric cancer. *Front. Immunol.* 15:1370367. doi: 10.3389/fimmu.2024.1370367

GLT8D2 is a prognostic biomarker and regulator of immune cell infiltration in gastric cancer

Han Wang[1,2†], Jiabin Zheng[2†], Qingyang Ma[3], Junchang Zhang[3*] and Yong Li[2*]

[1]Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [2]Department of Gastrointestinal Surgery, Department of General Surgery, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, China, [3]Department of Gastrointestinal Surgery, First Affiliated Hospital of Jinan University, Guangzhou, Guangdong, China

Because of the considerable tumor heterogeneity in gastric cancer (GC), only a limited group of patients experiences positive outcomes from immunotherapy. Herein, we aim to develop predictive models related to glycosylation genes to provide a more comprehensive understanding of immunotherapy for GC. RNA sequencing (RNA-seq) data and corresponding clinical outcomes were obtained from GEO and TCGA databases, and glycosylation-related genes were obtained from GlycoGene DataBase. We identified 48 differentially expressed glycosylation-related genes and established a prognostic model (seven prognosis genes including *GLT8D2, GALNT6, ST3GAL6, GALNT15, GBGT1, FUT2, GXYLT2*) based on these glycosylation-related genes using the results from Cox regression analysis. We found that these glycosylation-related genes revealed a robust correlation with the abundance of Tumor Infiltrating Lymphocytes (TILs), especially the GLT8D2 which is associated with many TILs. Finally, we employed immunohistochemistry and Multiplex Immunohistochemical to discover that GLT8D2 serves as a valuable prognostic biomarker in GC and is closely associated with macrophage-related markers. Collectively, we established a prognostic model based on glycosylation-related genes to provide a more comprehensive understanding of prediction for GC prognosis, and identified that GLT8D2 is closely correlated with adverse prognosis and may underscore its role in regulating immune cell infiltration in GC patients.

KEYWORDS

gastric cancer, glycosylation, prognosis, biomarker, GLT8D2

1 Introduction

According to the 2020 Global Cancer Report, GC is one of the most prevalent malignancies worldwide, ranking fourth in mortality and fifth in morbidity (1). At present, a significant number of GC patients are diagnosed at an advanced tumor stage, resulting in a poor prognosis (2). Immunotherapy has emerged as a prominent therapeutic approach for advanced GC patients and has demonstrated remarkable efficacy (3). However, the efficacy of immunotherapy is limited due to the substantial tumor heterogeneity in GC, as only a small subset of patients benefited from immunotherapy, which is potentially linked to the immune microenvironment of tumors. Therefore, identifying useful biomarkers for immune checkpoint inhibitors and developing novel immunotherapeutic strategies are urgently needed.

The development of GC is a multifaceted process influenced by various factors, including environmental stimuli, epigenetic mechanisms, and protein modifications. Glycosylation represents a prevalent form of protein modification closely intertwined with numerous tumorigenesis processes. In GC, the glycosylation landscape is dramatically altered, often as a result of dysregulation of glycosyltransferases, glycosidases, and other related enzymes. Research indicates that overexpression of GnT-V induces mislocalization of E-cadherin within GC cells, consequently compromising its functionality (4, 5). In contrast, GnT-III can counteract the activity of GnT-V by regulating the glycosylation modification of E-cadherin (4). GALNT10 exhibited a positive correlation with the histological type and degree of differentiation in GC (6). GALNT2 mediates O-glycosylation of EGFR, resulting in reduced EGFR phosphorylation and inhibition of the EGFR-Akt signaling pathway, thereby impeding the onset and progression of GC (7). This suggests that glycosylation plays an important role in the occurrence and development of GC.

Acknowledging the pivotal role of glycosylation in GC pathogenesis, significant endeavors have been undertaken to delineate glycosylated gene profiles and assess their efficacy as diagnostic and therapeutic biomarkers. Currently, certain glycoprotein and glycan-associated biomarkers (referred to as carbohydrate biomarkers) are employed in human cancer screening, diagnosis, and treatment, including CA19–9, carcinoembryonic antigen (CEA), CA125, AFP, and HER2 (8). Glycosylation patterns also hold significant potential in guiding personalized treatment approaches. Expression of certain glycosylation genes correlates with response to chemotherapy, targeted agents, and immunotherapy. For instance, elevated expression of the sialyltransferase ST6GalNAc1 is linked to resistance against trastuzumab in HER2-positive GC (9). Hence, elucidating the precise pathological regulatory mechanisms underlying glycosylation modifications in GC may pave the way for novel avenues in the comprehensive treatment of GC.

Glycosylation plays a role in numerous cancer-related biological processes, yet the involvement of tumor glycosylation in immune evasion is often overlooked (10–12). Aberrant tumor glycosylation can alter the way of immune system perceives tumors, thus driving immune suppression within the tumor microenvironment (13–15). Previous studies have indicated that glycosylated histones of tumor cells can interact with lectin receptors expressed by immune cells, such as Sialic Acid-Binding Immunoglobulin-like Lectins (SIGLECs) and Macrophage Galactose-Specific Lectin (MGL), to mediate immune evasion. O-glycosylation of MUC1, CD43, and CD45, as well as the glycolipids GM2 and GD2, which carry terminal N-acetylglucosamine, can interact with MGL on macrophages, leading to increased IL-10 production and the induction of effector T-cell apoptosis, driving immune suppression processes (16). On the other hand, N-glycans stabilize PD-L1 by reducing proteasomal degradation, thereby enhancing its immune inhibitory activity (17, 18). However, the function and mechanism of glycosylation in the immune evasion of GC remain unclear. Therefore, gaining insights into the interplay between glycosylation and immune cell infiltration could offer a more comprehensive perspective on the effectiveness of cancer immunotherapy.

Investigations into prognostic signatures linked to glycosylation in cancer have yielded promising results in various malignancies, including hepatocellular carcinoma, clear cell renal cell carcinoma, and pancreatic cancer (19–21). However, similar investigations in the context of GC are scarce. Previous studies have explored cancer-related prognostic signatures associated with glycosylation in GC, these investigations have primarily focused on a restricted set of pertinent genes, possibly neglecting other critical components within the immune microenvironment. Moreover, these studies have largely remained confined to bioinformatics analysis without employing pertinent experimental validation methods, resulting in a gap in our comprehension of specific glycosylation-related genes influencing the prognosis and immune status of GC. Therefore, it is imperative to systematically analyze the relationship between glycosylation and GC, and to further explore potential novel prognostic biomarkers and therapeutic targets.

In this study, we conducted a systematic profiling of expression data specific to STAD and correlated clinical outcomes sourced from both The Cancer Genome Atlas (TCGA) and GEO databases. Additionally, we identified glycosylation-related genes utilizing data extracted from the GlycoGene DataBase. Then, we evaluated the differentially expressed glycosylation-related genes between GC tissues and adjacent normal tissues, screened for signatures associated with survival, and established a prognostic model based on glycosylation-related genes to predict the prognosis of GC patients. Furthermore, we explored the prognostic value of glycosyltransferase 8 domain-containing 2 (GLT8D2) and its potential predictive role in immunotherapy efficacy via the Tumor Immune Estimation Resource (TIMER) and immunohistochemistry. This study revealed the association between glycosylation and the immune microenvironment in GC and the possible connection and mechanism by which GLT8D2 may regulate TILs. High expression of GLT8D2 promotes the proliferation and migration of GC cells, and was also shown to be associated with a worse prognosis in GC patients.

2 Methods

2.1 Data source

RNA sequencing (RNA-seq) data for stomach adenocarcinoma (STAD), referred to as TCGA-STAD, were obtained from the

TCGA database (https://portal.gdc.cancer.gov/). Additional data, including counts and fragments per kilobase of transcript per million mapped reads (FPKMs), as well as clinical information corresponding to the respective patients, were also obtained. The RNA expression data, which included the GSE19826, GSE26899, GSE54129, GSE84433 and GSE84437 datasets and contained normal and tumor tissues, were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo). To ensure data standardization, all the information was subjected to quantile normalization and transformed into a log2 scale. When multiple probes were used to detect a single gene symbol, the mean expression levels were calculated for analysis. Therefore, a total of 170 glycosylation-related genes obtained from the GlycoGene DataBase (GGDB; https://acgg.asia/ggdb2/) were selected as candidate genes. This study adhered to the publication guidelines stipulated by the GEO and TCGA databases.

2.2 Differentially expressed glycosylation-related genes

The identification of differentially expressed genes (DEGs) between tumor and adjacent normal tissues was conducted using the GEO datasets GSE19826, GSE26899 and GSE54129. This analysis was performed within the RStudio environment (version 1.2.5001) using the "limma" package, applying the following cutoff criteria for adjustment: p value < 0.05 and |log2FC| ≥ 1. Subsequently, the "heatmap" package was used to visualize the magnitude of differences across the three datasets. A Venn diagram was subsequently drawn from the selected glycosylation-related genes to determine the intersection between the candidate genes and the DEGs.

Afterwards, functional analysis was performed using the Metascape Online platform (https://Metascape.org/gp/index.html#/main/step1) (22). The differentially expressed glycosylation-related genes were input into Metascape for comprehensive functional analysis, including the construction of a protein–protein interaction (PPI) network. We applied the MCODE algorithm to identify densely connected regions within the network. A significance threshold of $p < 0.05$ was used for this analysis. Furthermore, functional enrichment analysis was also conducted to assess the biological functions of the differentially expressed glycosylation-related genes using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. The criteria for GO term enrichment and KEGG signaling pathway enrichment were set at FDR < 0.05. The 10 most significant GO terms and KEGG signaling pathways were subsequently visualized using the R package "ggplot2".

2.3 Construction and validation of the glycosylation-related gene prognostic model

The present study utilized the TCGA-STAD and GEO datasets (GSE84433 and GSE84437, respectively) to develop a prognostic signature based on glycosylation-related genes. The TCGA-STAD cohort served as the training cohort, while the GEO datasets GSE84433 and GSE84437 were used as the validation cohort. Univariate Cox analysis of overall survival (OS) was initially conducted to identify glycosylation-related genes associated with OS, considering a p value <0.05 to indicate statistical significance. Subsequently, the optimal model relying on prognosis-related glycosylation-related genes was identified using the Least Absolute Shrinkage and Selection Operator (LASSO) penalized Cox proportional hazards regression method through the R package "glmnet". The signature was then established using these independent prognostic genes in accordance with their respective coefficients. Patients were divided into two groups according to the median risk score: low-risk and high-risk. Survival comparisons between the low-risk and high-risk groups were conducted using Kaplan–Meier (K–M) survival curves generated with the R package "survival".

2.4 Clinical relevance investigation and prognostic nomogram construction

To furnish a quantitative predictive tool for assessing survival risk in GC patients, a nomogram was developed using differentially expressed glycosylation-related genes and clinical parameters. Additionally, calibration curves were generated to compare the predictive outcomes with actual survival data, thereby evaluating the predictive accuracy of the nomograms. The construction of the nomogram and the calibration curves was accomplished using the R package "rms".

2.5 Tumor immune estimation resource database

The TIMER2.0 (https://timer.cistrome.org/) is a web-based interactive platform designed for comprehensive immune infiltration analysis across various malignancies. Six advanced algorithms were used to provide a more robust assessment of TILs levels using data from the TCGA and other tumor-related datasets. In this study, we investigated the associations between GLT8D2 expression and the expression of gene markers specific to TILs, namely, CD8+ and CD4+ T cells, B cells, monocytes, natural killer (NK) cells, dendritic cells (DCs), tumor-associated macrophages (TAMs), M1 macrophages, M2 macrophages and neutrophils, using correlation modules. To visualize the expression patterns between pairs of custom genes in GC and determine the statistical significance of the correlations, Spearman's correlation coefficients were computed, and expression dispersion maps were generated. The gene expression levels are represented as log2 RSEM values.

2.6 TISIDB

TISIDB (http://cis.hku.hk/TISIDB/index.php) is an online platform that integrates diverse data sources to explore the intricate interplay between tumors and the immune system. This database

proves invaluable for shedding light on the interactions between tumors and immune cells, predicting responses to immunotherapy, and identifying novel targets for immunotherapeutic interventions. It is a valuable resource for advancing research and therapies in the field of cancer immunology. In this study, we harnessed the ability of TISIDB to investigate the correlation between GLT8D2 and a comprehensive set of immune components, such as 28 TILs, in the context of GC.

2.7 Immunohistochemistry and multiplex immunohistochemical

This study entailed the analysis of 150 paraffin-embedded GC specimens and 30 normal specimens procured from the Shanghai Outdo Biotech Company between January 2010 and December 2015. The inclusion criteria stipulated that all samples were acquired from patients with histologically confirmed gastric adenocarcinoma and validated by expert gastrointestinal pathologists. Patient records comprised comprehensive data encompassing age, sex, tumor location, TNM stage, histological grade, Lauren's classification, treatment history, and detailed follow-up information for survival analysis. The exclusion criteria encompassed patients who had undergone chemotherapy or radiotherapy before surgery and those with synchronous or metachronous malignancies. Multiplex Immunohistochemistry (PANOVUE kit, #10234100050) was employed to assess the expression levels of GLT8D2 and CD68, aiming to establish a correlation between GLT8D2 expression and CD68 expression. Anti-GLT8D2 (1:1000 dilution; Bioss, bs-8302R) and anti-CD68 (1:2000 dilution; Cell Signaling Technology, #97778) antibodies were utilized. Immunohistochemistry was carried out according to the DAB kit of Fujian Maxim Company (DAB-0031), Anti-GLT8D2 (1:100 dilution; Bioss, bs-8302R) antibody was used. Staining intensities were classified into four categories: negative (-), weak (+), moderate (++), and strong (+++).

2.8 Cell transfection and lentiviral infection

Gastric cancer cell line AGS, purchased from the Cell Bank of Chinese Academy of Sciences (Shanghai), was used in this study and treated with DMEM/F12 (Gibco, CAT# C11330500BT, Beijing, China) medium combined with 10% fetal bovine serum FBS (Gibco, CAT# 10099141C, Beijing, China). China). GeneChem(Shanghai, China) provided the GLT8D2-knockdown lentiviral vector. GLT8D2 was cloned into GV341 vector (GeneChem, Shanghai, China) to construct GLT8D2 lentiviral expression vector. Lentivirus transduction was generated and purified according to the manufacturer's instructions. Puromycin (2 μg/ml) was added to screen the transgenic cells.

2.9 Cell viability, colony formation, and wound healing assays

A quantity ranging from 1000 to 1500 cells were evenly distributed across the wells of 96-well culture plates. Subsequently, the assessment of cell viability was carried out using a Cell Counting Kit-8 (CCK-8) (Beyotime, CAT# C0048M) following a 2-hour incubation period at 37°C. This evaluation was conducted at multiple time points, specifically 0, 24, 48, and 72 hours post-seeding, in strict adherence to the guidelines provided by the manufacturer. In colony formation assays, 500 cells were seeded per well in six-well plates for experiments, and the cells were cultured for two weeks. Subsequently, the colonies were fixed with 4% paraformaldehyde for 15 minutes and stained with crystal violet (Beyotime Biotechnology, CAT# C0121) for 15 minutes. In the wound healing assays, we used cell culture dishes to create a defined wound and observed the migration capability of the cells during the healing process. At specific time intervals (0 h, 12 h, and 24 h), we documented and measured the extent of wound closure to assess the cell migration and healing ability.

2.10 Cell migration assays

Migration assays were carried out using transwell plates with 8-μm pores. In the migration assay, cells were placed in the top compartment with 0.2 ml of serum-free medium, while 0.8 ml of culture media supplemented with 10% fetal bovine serum was added to the bottom chamber. After the cells were incubated for 24 hours, they were fixed with 4% paraformaldehyde for 15 minutes and subsequently stained with crystal violet for 15 minutes. Unmigrated cells were then removed from the top layer using cotton swabs. Migrating cells were observed and imaged using a 10× microscope (Olympus CKX53).

2.11 Statistical analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS, version 26.0) and GraphPad Prism (version 8.0). K–M plots were generated to construct survival curves. In these KM plots, as well as in the analysis conducted using the TIMER2.0 and TISIDB tools, hazard ratios (HRs) and p values were computed using the log-rank test. Spearman's correlation coefficient was utilized to evaluate the correlation between GLT8D2 expression and immune infiltration. Univariate and multivariate Cox regression analyses were executed with the R package "survival", providing HRs along with their corresponding 95% confidence intervals (CIs). Additionally, the differences among various clinical factors were evaluated using independent t tests, with statistical significance denoted by a p value < 0.05.

3 Results

3.1 Differentially expressed glycosylation-related gene signatures in GC

The GEO datasets used in this study are provided in Supplementary Table S1. After conducting the differential gene analysis, a total of 984 dysregulated genes were identified from the

GEO dataset GSE19826, with 338 genes exhibiting upregulation and 646 genes exhibiting downregulation (Figure 1A). Additionally, from the GEO dataset GSE26899, 527 dysregulated genes were found, consisting of 174 upregulated genes and 353 downregulated genes (Figure 1B). Finally, the GEO dataset GSE54129 yielded 2,583 dysregulated genes, of which 1,134 genes were upregulated and 1449 genes were downregulated (Figure 1C). Furthermore, the dysregulated genes in the three aforementioned datasets were visualized in a more intuitive manner using volcano plots (Figures 1D–F). The glycosylation genes obtained from the GlycoGene DataBase were experimentally validated and are listed in Supplementary Table S2. To obtain the "differentially expressed glycosylation-related genes", the differentially expressed genes (DEGs) from the GEO datasets were compared with the glycosylation-related gene set using a Venn diagram, which revealed 48 intersecting glycosylation-related genes among the four datasets (Figure 1G).

To investigate the mechanisms underlying glycosylation signatures in GC, a comprehensive functional analysis was conducted using Metascape Online. Our findings indicated that the dysregulated glycosylation genes are primarily associated with various biological processes, such as the response to glycoprotein biosynthetic process, O-glycan processing, carbohydrate metabolic process, and metabolism of carbohydrates, as revealed by Gene Ontology (GO) analysis (Figure 2A). Moreover, KEGG pathway analysis also demonstrated that these dysregulated glycosylation genes were significantly enriched in pathways related to glycoprotein biosynthetic process, O-glycan and N-linked glycosylation, and cellular polysaccharide metabolic process (Figure 2B). These results prompted us to explore the correlation between the glycosylation gene set and the progression of GC. Furthermore, through the utilization of the protein–protein interaction (PPI) network and the MCODE plugin in Metascape Online, we identified significant modules within these glycosyltransferase genes (Figure 2C). Module 1 included *FUT2, FUT3, FUT4, FUT9, GCNT2, B4GALT1, B4GALT4, B3GNT3, and ST3GAL6*. Module 2 includes *GCNT1, GALNT7, GALNT12, B3GNT6, and ST6GALNAC1*. By utilizing the TCGA database, we found that the most enriched terms in terms of biological process (BP), cellular component (CC), and molecular function (MF) were "transferase activity", "Golgi stack", and "glycoprotein biosynthetic process", respectively (Figure 2D). Moreover, functional enrichment analysis revealed that the signaling pathway most relevant to the glycosyltransferase genes was O-glycan biosynthesis (Figure 2E).

3.2 The establishment and verification of a glycosylation-related prognostic model

Initially, the genes significantly associated with prognosis were detected through the application of univariate Cox regression analysis. As depicted in Figures 3A–I, nine glycosylation-related genes were identified as prognostic genes: *GLT8D2, CHSY3,*

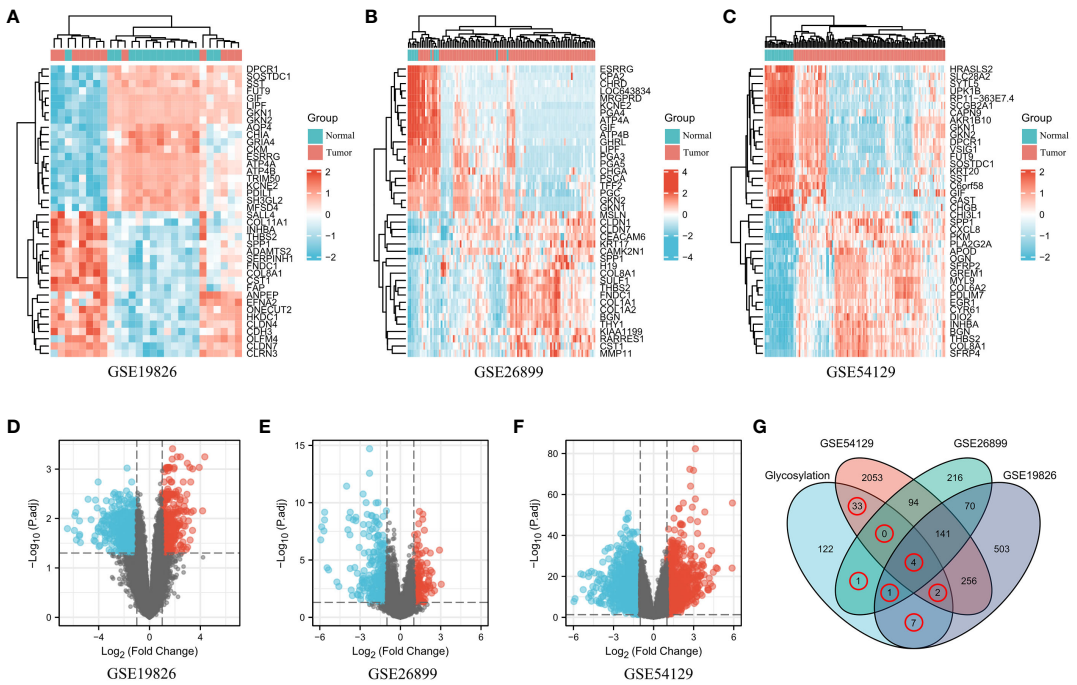


FIGURE 1
Differentially expressed glycosylation-related gene signatures in GC. **(A–C)**. The expression patterns of glycosylation-related genes in both normal and tumor samples were examined across the GEO datasets: GSE19826 **(A)**, GSE26899 **(B)**, and GSE54129 **(C)**. Genes were categorized based on their expression levels, with high expression represented by the color red and low expression represented by the color blue. **(D–F)**. Volcano plots showing the dysregulated glycosylation-related genes in the three aforementioned GEO datasets. **(G)** Venn diagram showing the dysregulated glycosylation-related genes common to the four datasets.

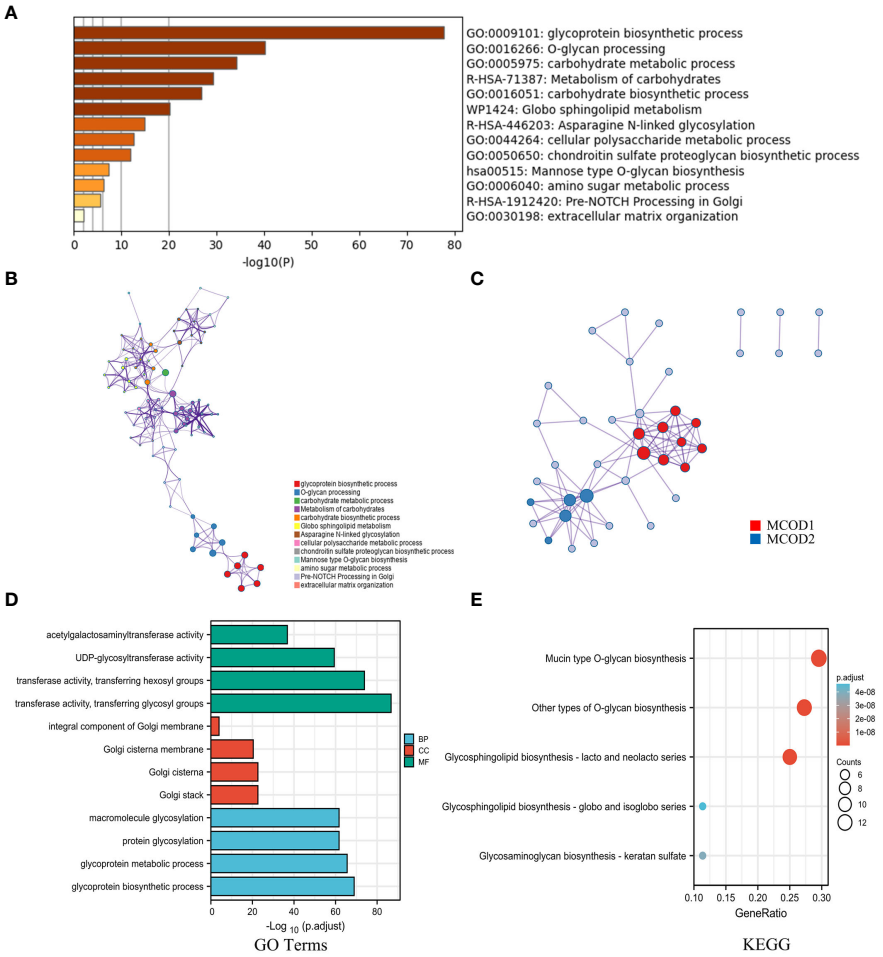


FIGURE 2
The mechanisms underlying glycosylation signatures in GC. **(A)** Bar plot showing the distribution and relationships of the different functions according to the GO and KEGG analyses based on Metascape Online. **(B)** Network showing the distribution and relationships of the different functions according to the GO and KEGG analyses based on Metascape Online. **(C)** PPI network and MCODE showing the hub genes among the glycosylation-related genes. **(D)** GO enrichment analysis; BP, biological process; CC, cellular component; MF, molecular function. **(E)** KEGG pathway annotation.

GALNT6, ST3GAL6, GALNT15, GBGT1, FUT2, B4GALNT3, and GXYLT2. Figure 3J displays a forest plot presenting the outcomes of the univariate Cox regression analysis. Therefore, a prognostic model was established based on the Cox regression coefficient as follows: risk score= [expression level of FUT2 × (-0.11)] + [expression level of ST3GAL6 × 0.26] + [expression level of GALNT6 × (-0.03)] + [expression level of GALNT15 × 0.01] + [expression level of GLT8D2 × 0.14] + [expression level of GXYLT2 × 0.01] + [expression level of GBGT1 × (-0.05)].

3.3 The predictive model construction for GC patients

In the training cohort, LASSO regression was adopted to analyze the data according to the univariate analysis procedure described above (Figure 4A). After conducting calculations that involved combining the coefficients from the LASSO analysis with the levels of gene expression, we identified a set of seven prognostic genes: *GLT8D2, GALNT6, ST3GAL6, GALNT15, GBGT1, FUT2, and GXYLT2* (Figure 4B). Employing these seven genes, we computed an individualized risk score for each patient, and the threshold for distinguishing between the high-risk and low-risk categories was established at the median value (Figure 4C). This study revealed that OS was significantly worse in high-risk patients than in low-risk patients in the TCGA training set (Figure 4D, P<0.05). Similar results were obtained in the validation sets GSE84433 and GSE84437 (Figure 4D, P<0.05).

3.4 Construction of the nomogram

A predictive glycosylation-related prognostic nomogram was established via multivariate analysis. Therefore, using seven prognostic genes and certain clinicopathological factors, we developed a prognostic nomogram that serves as a valuable

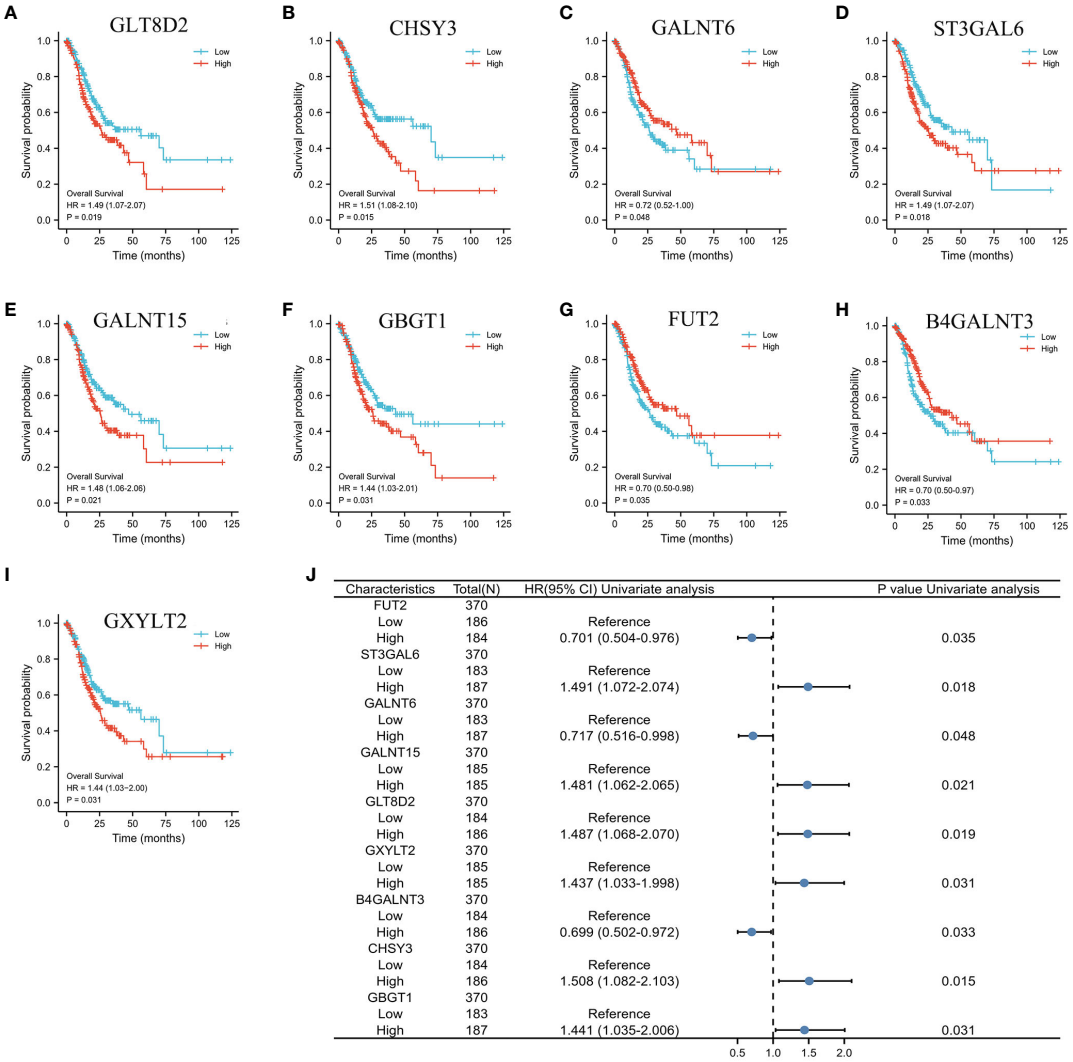


FIGURE 3
K–M plots and forest plot of glycosylation-related prognostic factors. **(A–I)** Kaplan–Meier plots showing the glycosylation-related genes with prognostic value. **(J)** Forest plot showing the results of the univariate Cox regression analyses.

quantitative tool for predicting the survival prospects of individual patients (Figure 5A). Furthermore, the predictive accuracy for overall survival was evaluated via calibration curves. Importantly, the calibration curves of this prognostic nomogram demonstrated excellent agreement between the predicted and actual survival rates at the 1-, 3-, and 5-year milestones across the entire TCGA cohort (Figure 5B).

3.5 Correlation between immune infiltration and GLT8D2 expression in GC

Immune cell infiltration is crucial in tumor progression. Therefore, to further explore the association of glycosylation-related genes with immunity, correlation analysis was conducted between seven glycosylation-related genes and immune functions via the TISIDB platform. As shown in Figure 6A, these glycosylation-related genes were strongly correlated with the abundance of TILs, especially GLT8D2, which is associated with many TILs.

As such, we utilized the TIMER platform to evaluate the association between GLT8D2 expression and immune cell infiltration in STAD. GLT8D2 expression was adversely correlated with the purity of the STAD cells (rho = -0.17, $p < 0.00088$). Our findings revealed a robust correlation between GLT8D2 and TILs. Specifically, a high level of GLT8D2 expression was positively associated with the degree of infiltration by various immune cell populations, including macrophages (rho = 0.744), $CD8^+$ T cells (rho = 0.403), $CD4^+$ T cells (rho = 0.26), B cells (rho = 0.276), monocytes (rho = 0.464), neutrophils (rho = 0.345), T-cell regulatory cells (rho = 0.276), NK cells (rho = 0.219), and myeloid dendritic cells (rho = 0.54) (Figure 6B). Importantly, all p values were markedly less than 0.001. The TCGA database was also used to assess the difference in immune cells between patients with

10.3389/fimmu.2024.1370367

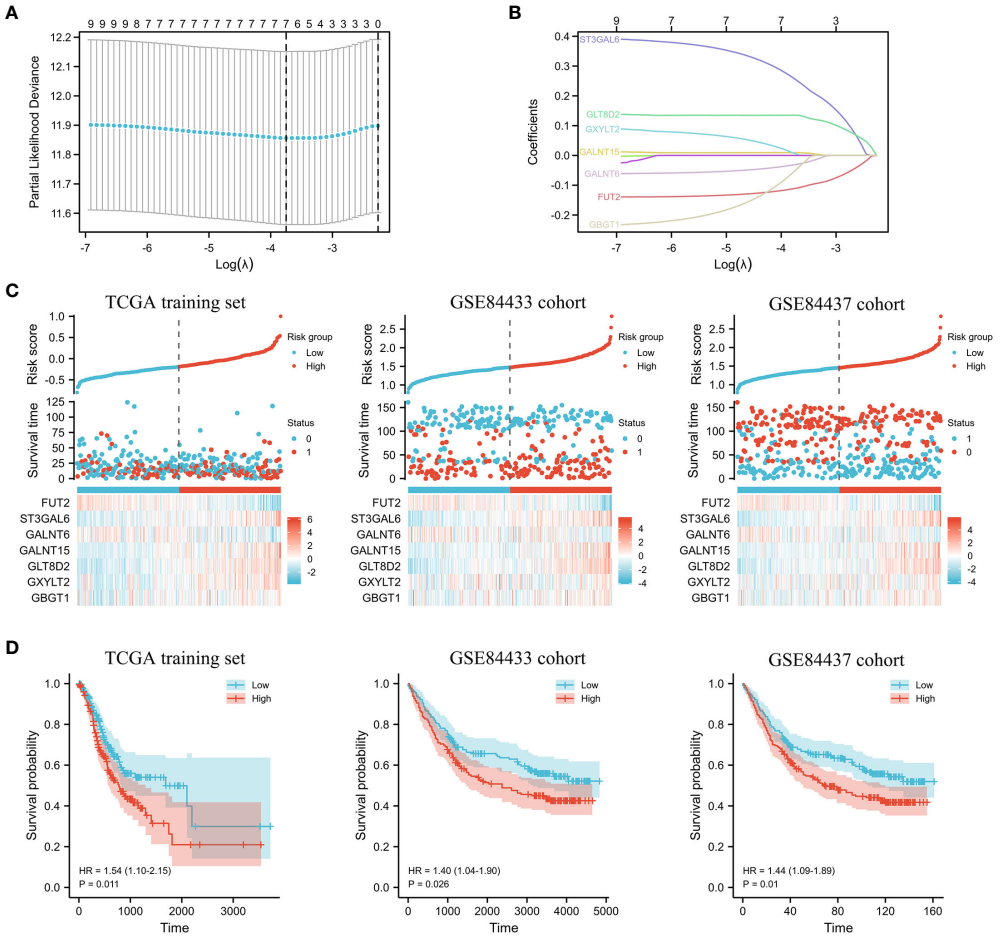


FIGURE 4
The predictive model constructed for GC patients. **(A)** Partial likelihood deviance of DEGs. **(B)** LASSO regression and coefficient values of DEGs. **(C)** Risk score distribution, survival status, and expression of 7 DEGs for GC patients in the low- and high-risk groups in the TCGA training set and the GSE84433 and GSE84437 cohorts. **(D)** KM survival analyses of 7 DEGs for GC patients in the low- and high-risk groups in the TCGA training set and the GSE84433 and GSE84437 cohorts.

high- or low-grade GLT8D2 tumors. Similar results were obtained (Figure 6C). These findings collectively underscore the pivotal role of GLT8D2 in orchestrating immune infiltration within the context of GC.

3.6 GLT8D2 expression is correlated with macrophage-related marker expression and poor prognosis in GC patients

Evidently, GLT8D2 exhibited a significant correlation with a majority of the marker sets associated with tumor-associated macrophages (TAMs), M1-type macrophages, and M2-type macrophages in STAD. Specifically, this study revealed strong correlations between GLT8D2 and TAM markers, including CD68, chemokine ligand (CCL)-2 and Interleukin 10 (IL10), in STAD. Additionally, GLT8D2 displayed robust correlations with M1 phenotype markers, such as Interferon Regulatory Factor 5 (IRF5) and Prostaglandin-Endoperoxide Synthase 2 (PTGS2), as well as with M2 phenotype markers, including CD163, V-Set and Immunoglobulin Domain Containing 4 (VSIG4), and Membrane Spanning 4-Domains A4A (MS4A4A) (Figures 7A–C). All p values were markedly less than 0.001. Moreover, we employed a multiplex immunohistochemical approach to assess the correlation between GLT8D2 and CD68 expression. Our findings demonstrated that elevated GLT8D2 expression was associated with increased CD68 infiltration (Figure 7D). Concurrently, we investigated the relationship between GLT8D2 expression and clinicopathological characteristics in GC patients via immunohistochemistry (IHC) (Table 1). By scoring the staining intensity, we classified the expression levels of GLT8D2 into four groups: negative (–), weak (+), moderate (++) and strong (+++) staining (Figure 8A). The results of this study indicated that high GLT8D2 expression was correlated with poorer OS and disease-free survival (DFS) in GC patients (Figures 8B, C). Furthermore, univariate and multivariate Cox proportional hazards regression analyses of OS and DFS in GC patients revealed that GLT8D2 was an independent prognostic risk factor (Tables 2, 3). Consequently, our findings support the assertion that GLT8D2 is a valuable prognostic biomarker in GC and is closely associated with immune infiltration.

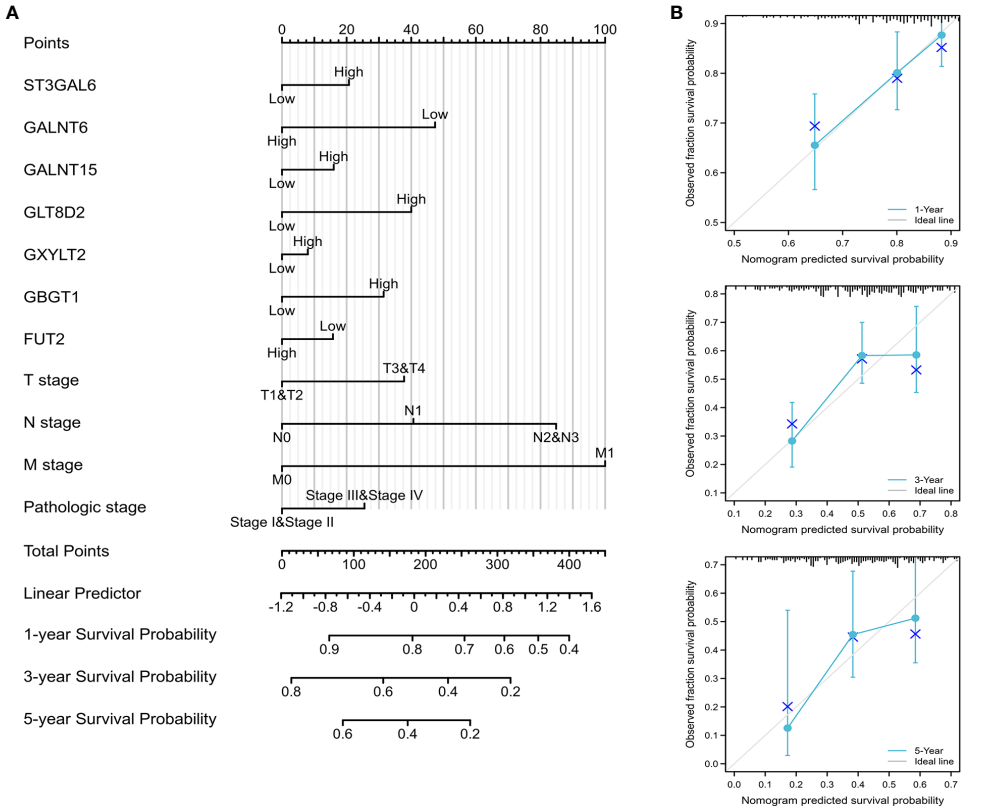


FIGURE 5
Construction of the prognostic nomogram **(A)** Nomogram designed to predict 1-, 3-, and 5-year OS in the complete TCGA cohort. **(B)** Calibration curves of the nomogram assessing the agreement between projected and observed 1-, 3-, and 5-year survival rates across the entire TCGA cohort. A dashed line at 45° indicates a flawless prediction, with the actual performance of our nomogram depicted by the blue lines.

3.7 GLT8D2 knockdown blocks the proliferation and metastasis of GC cells *in vitro*

To gain a deeper understanding of the impact of GLT8D2 in GC, we explored phenotypic alterations in GC cells following GLT8D2 knockdown. The effectiveness of GLT8D2 knockdown was verified by western blotting (Figure 9A). Colony formation assays and cell viability demonstrated a reduction in the clonogenic capacity of GC cells following GLT8D2 knockdown (Figures 9B, C). Furthermore, the wound healing assay revealed wider wounds after the same 24-hour interval in the GLT8D2 deficiency groups than in the shCtrl group (Figure 9D). In addition, GLT8D2 knockdown significantly diminished the migratory ability of GC cells, as evidenced by cell migration assays (Figure 9E). Collectively, these findings indicate a pivotal role for GLT8D2 in the proliferation and migration of GC cells.

4 Discussion

Increasing evidence indicates that glycosylation plays a pivotal role in tumorigenesis and the efficacy of cancer treatments. In the present study, we focused on glycosylation-related genes and investigated their impact on the prognosis of GC patients. Our objective was to elucidate glycosylation-related prognostic models and their relationship with the GC immune microenvironment, aiming to further identify potential biomarkers for prognosis assessment and targeted therapy. A comprehensive bioinformatics study was subsequently performed to systematically analyze glycosylation-related genes associated with poor prognosis in GC patients, and a glycosylation-based prognostic model was established by using the GEO, TCGA, and GlycoGene databases. Additionally, we showed that high expression of GLT8D2 was associated with poor prognosis in GC patients and revealed novel insights into the key role of GLT8D2, which may serve as a prognostic biomarker associated with immune infiltration in GC.

Glycosylation plays a role in numerous cancer-related biological processes, including inflammation, immune surveillance, cell–cell adhesion (4, 5), cell-matrix interactions (23), intercellular and intracellular signal transduction (24–27), and cellular metabolism (28, 29). Tumor classification studies based on glycosylation-related gene expression profiles have begun to emerge (30–32). Subtypes of colorectal cancer patients with poor prognoses have been identified using glycan gene markers, among which the loss of GALNT6 gene expression has been associated with cancer cell invasion and chemoresistance and has been highlighted as a prognostic biomarker (33). Research has indicated that the overexpression of

10.3389/fimmu.2024.1370367

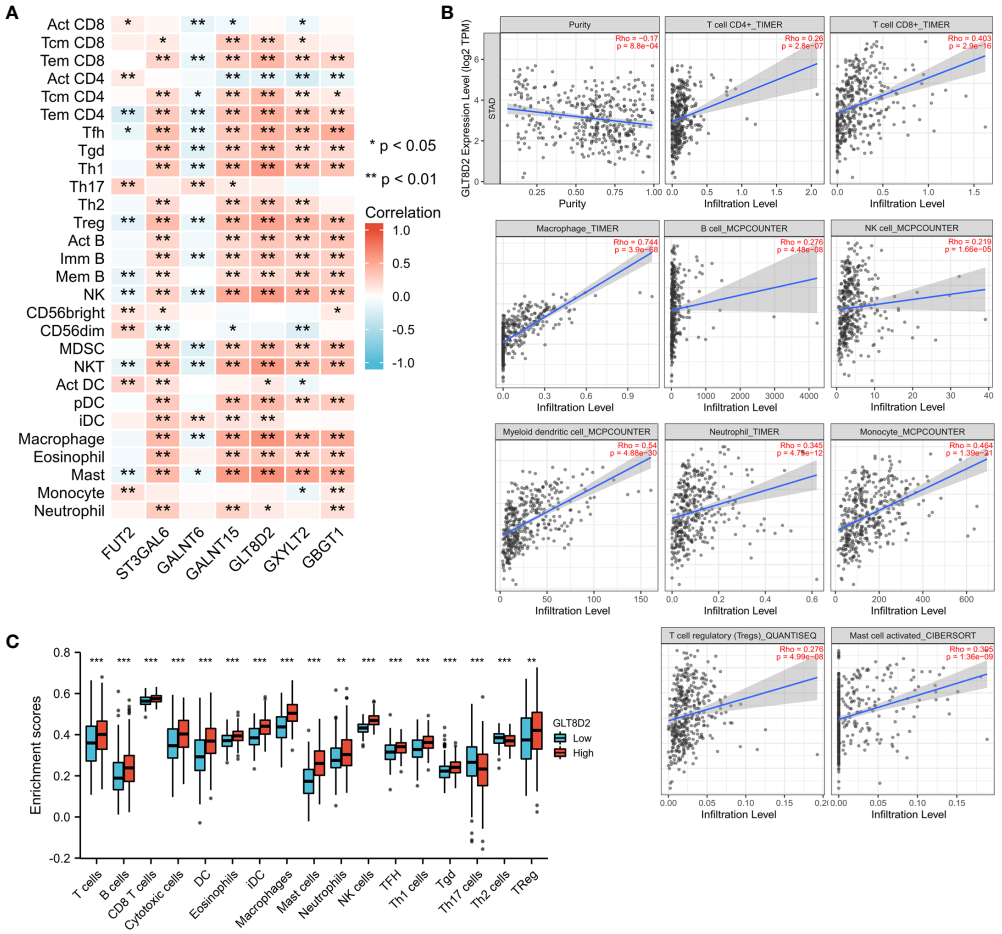


FIGURE 6
Correlations between immune infiltration and glycosylation-related genes in GC. **(A)** Correlation matrix of seven glycosylation-related genes and the abundance of TILs. The red dots represent a positive correlation, and the blue dots represent a negative correlation. **(B)** Association between GLT8D2 expression and immune cell infiltration in STAD according to TIMER data. **(C)** Differences in immune cells between patients with high or low GLT8D2 expression in tumors in the TCGA database. *, **, and *** represent P<0.05, P<0.01, and P<0.001 respectively.

GnT-V results in the mislocalization of E-cadherin within GC cells, leading to functional impairment. The primary mechanism involves the addition of N-glycan chains with β-1,6-GlcNAc branches mediated by GnT-V to E-cadherin, promoting incorrect assembly and ineffective adhesive connections, thereby affecting cell–cell adhesion and subsequently contributing to tumor metastasis (4, 5). Therefore, exploring the biological significance of glycosylation in GC is advantageous for deciphering the pathological regulatory mechanisms involved in cancer biology, which may help in identifying novel biomarkers for prognosis and targeted therapy.

In this study, we analyzed DEGs from the GEO datasets GSE19826, GSE26899, and GSE75241 and intersected them with a glycosylation-related gene set obtained from the GlycoGene DataBase, resulting in the identification of 48 glycosylation-related genes. Subsequently, functional analysis of these 48 glycosylation-related genes revealed that these genes were associated with various biological processes, including the response to glycoprotein biosynthetic processes and O-glycan processing. We also employed univariate Cox and multivariateCox regression analyses to identify 9 out of 48 adverseprognosis-associated glycosylation-related genes and establish a glycosylation-based prognostic model. Among these proteins, GALNT6 has been reported to promote the occurrence of breast cancer through abnormal glycosylation of the mucin protein MUC1 (34). In addition, previous research has revealed that the hypermethylation of ST3GAL6 is strongly correlated with Epstein–Barr virus-associated gastric carcinomas (35). In addition, GXYLT2 has also been reported to be a potential diagnostic and prognostic biomarker for GC by bioinformatics analysis (36). This study discovered that GLT8D2 is highly expressed in GC and is closely associated with poor prognosis by bioinformatics analysis and clinical samples. Cellular functional studies also suggested that GLT8D2 affects the proliferation and migration of GC cells. As a glycosyltransferase, GLT8D2 may modify the substrate protein by glycosylation, thereby affecting its stability, localization, interaction and activity, and then regulate the occurrence and development of tumors (24–27). Moreover, glycosylation is essential for the function of adhesion molecules such as integrins and cadherins on the cell membrane (27). Abnormal expression of GLT8D2 may affect the glycan modification of these molecules, reduce the adhesion of cells to the extracellular matrix, enhance cell migration and invasion, and

10.3389/fimmu.2024.1370367

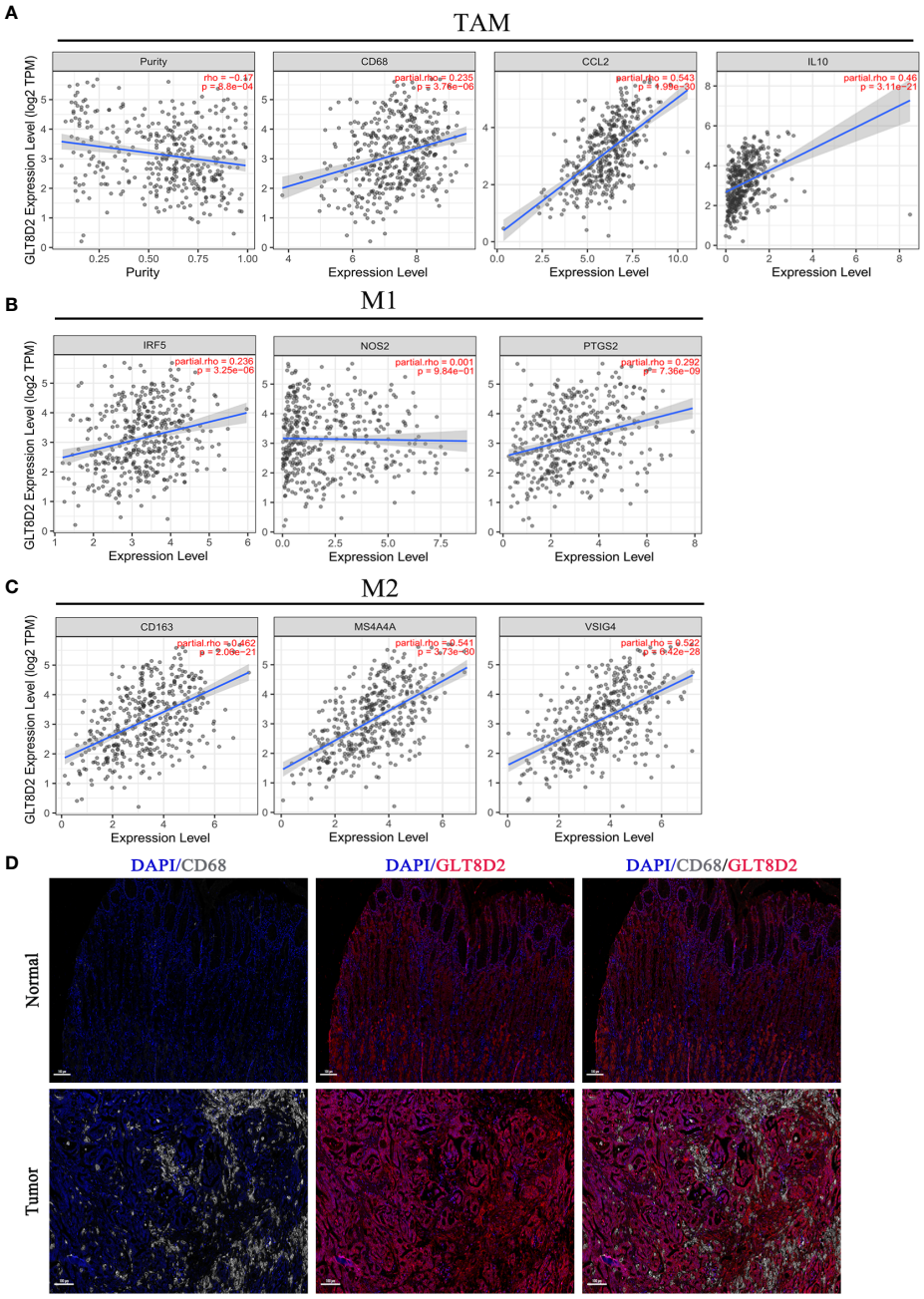


FIGURE 7
GLT8D2 expression is correlated with macrophage-related marker expression and poor prognosis in GC patients. **(A–C)** Associations between GLT8D2 expression and macrophage-related markers, including tumor-associated macrophages (TAMs), M1-type macrophages, and M2-type macrophages, in STAD according to TIMER analysis. **(D)** Multiplex immunohistochemical staining showing the correlation between GLT8D2 (red) and CD68 (white) expression.

promote tumor spread and growth. These studies suggest that glycosylation-related genes especially GLT8D2 may play a crucial role in the development and progression of GC. Therefore, we believe that our study may contribute to providing new insights into GC treatment.

In the cancer microenvironment, TILs have been demonstrated to play a crucial role in the initiation and progression of cancer (37–39). They may show the characteristics of promoting or inhibiting tumor growth in different types of cancer and different stages of the same type of cancer (40, 41). For CD8+ CTLs (cytotoxic T lymphocytes), they act as tumor suppressors by triggering a cytolytic reaction by recognizing tumor-specific antigens presented by the major histocompatibility complex (MHC) (42). In addition, regulatory T cells (Tregs) and myeloid-derived suppressor cells (MDSCs) can create an immunosuppressive tumor microenvironment by secreting inhibitory cytokines (such as IL-10 and TGF-β), depleting trophic factors, and directly inhibiting effector T cell function, thereby promoting tumor progression (43).

10.3389/fimmu.2024.1370367

TABLE 1 Associations of GLT8D2 expression with clinical parameters in GC patients.

Characteristic	GLT8D2		
	Low (%)	High (%)	*P*
Age (years)			0.077
<60	37(50.7)	36(49.3)	
≥60	28(36.4)	49(63.6)	
Gender			0.848
Male	40(44.0)	51(56.0)	
Female	25(42.4)	34(57.6)	
Tumor size			0.174
≤5 cm	41(47.7)	45(52.3)	
>5 cm	23(36.5)	40(63.5)	
Borrmann type			0.981
I-II	11(42.3)	15(57.7)	
III-IV	37(42.0)	61(58.0)	
Differentiation			**0.004**
Well+ moderate	22(64.7)	12(35.3)	
poor	34(35.8)	61(64.2)	
pTNM stage			**0.030**
I-II	32(54.2)	27(45.8)	
III-IV	33(36.3)	58(63.7)	
Depth of invasion			**0.023**
T1/2	20(60.6)	13(39.4)	
T3/4	45(38.5)	72(61.5)	
Lymph node metastasis			0.13
N0	26(52.0)	24(48.0)	
N+	39(30.0)	61(70.0)	
Distant metastasis			0.272
M0	61(45.2)	74(54.8)	
M1	4(26.7)	11(73.3)	
CEA level (μg/L)			0.155
≤5	57(46.0)	67(54.0)	
>5	8(30.8)	18(69.2)	
LVI			**0.027**
Yes	10(27.0)	27(73.0)	
No	48(48.0)	52(52.0)	
PNI			**0.046**
Yes	3(18.8)	13(81.3)	
No	53(44.9)	65(55.1)	

Bold values indicate P < 0.05.

Consequently, we observed a robust correlation between glycosylation-related genes and immune-infiltrating cells. To ensure the depth and practical applicability of our research, it is imperative to focus on the comprehensive exploration of the most promising or scientifically significant genes in subsequent studies. Subsequently, our analysis revealed that the GLT8D2 gene, a novel glycosyltransferase situated on chromosomal region 12q23.3, exhibited noteworthy correlations and biological significance across multiple dimensions. Previous research revealed that GLT8D2 is involved in the pathogenesis of nonalcoholic fatty liver disease (NAFLD) by negatively regulating microsomal triglyceride transfer protein (MTP) in HepG2 cells (44), and the GLT8D2/FGFR/PI3K/AKT signaling axis was found to be a significant contributor to platinum-based chemotherapy resistance in ovarian cancer (45). However, the biological functions and association with immune infiltration of GLT8D2 in GC remain unclear.

Thus, we systematically investigated the association between GLT8D2 expression and the degree of immune infiltration in GC. Our study revealed strong correlations between GLT8D2 expression and TILs, including $CD8^+$ T cells, $CD4^+$ T cells, Treg cells, B cells, neutrophils, dendritic cells (DCs), natural killer (NK) cells, and monocytes, particularly macrophages. Macrophages are a distinct type of immune cell classified into M1 and M2 subtypes and play critical roles in angiogenesis (46), invasion (47), and antitumor immunity (48). Additionally, our research demonstrated the association between GLT8D2 expression and macrophage markers in GC via TIMER. Clearly, GLT8D2 expression was strongly correlated with TAM markers, including CD68, CCL-2 and IL10. Moreover, by employing an immunohistochemical approach, this study demonstrated that elevated GLT8D2 expression was associated with increased CD68 infiltration and led to poor prognosis in GC patients. This suggested that GLT8D2-regulated TILs mainly play a role in promoting tumor progression, and the mechanism may be related to the immune escape caused by galactose-specific lectin (MGL) of macrophages, which can leading to increased IL-10 production and the induction of effector T-cell apoptosis, driving immune suppression processes (16). These findings discovered that GLT8D2 may potentially regulate TAM polarization, which could enhance the effectiveness of immunotherapy by targeting GLT8D2. High expression of GLT8D2 was also shown to be associated with a worse prognosis in GC patients. Taken together, these findings indicate that GLT8D2 plays a significant role in recruiting and modulating TILs in GC. Further investigations into the molecular mechanisms and functions of GLT8D2 in regulating macrophages are warranted and will provide additional insights.

5 Conclusions

Our study established a prognostic model based on glycosylation-related genes, which could contribute to assisting in clinical decision-making by predicting patient outcomes and recognizing responsiveness to particular therapies. Furthermore,

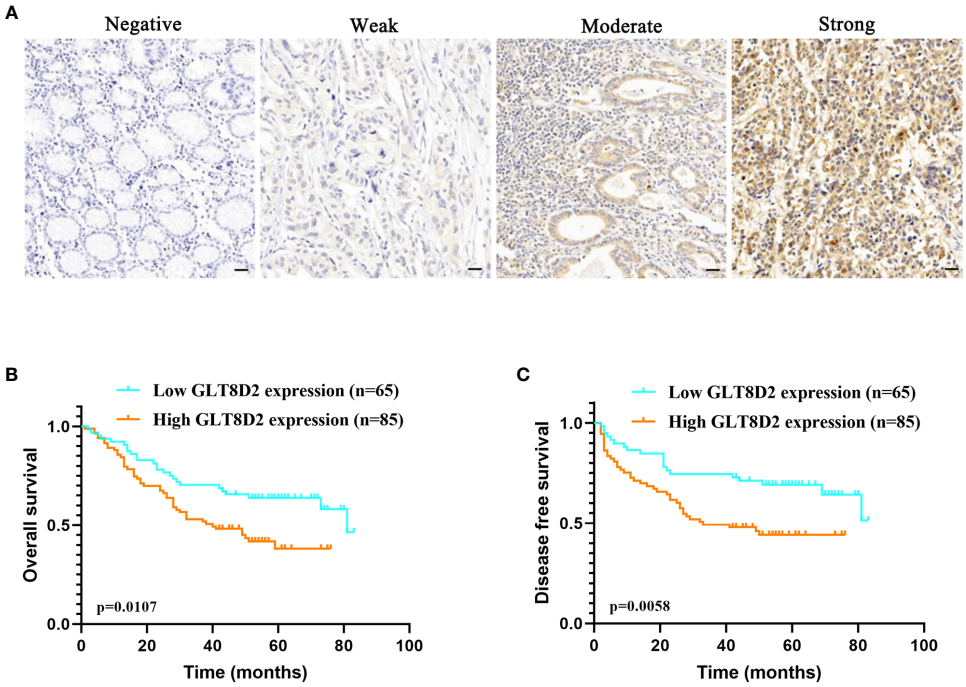


FIGURE 8
GLT8D2 serves as a valuable prognostic biomarker in GC. **(A)** Immunohistochemistry image showing GLT8D2 expression based on the immunohistochemical score. Patients were classified into four groups: negative (−), weak (+), moderate (++) and strong (+++) staining. **(B–C)** Relationships between GLT8D2 expression and overall survival (OS) **(B)** or disease-free survival (DFS) **(C)** outcomes in GC patients.

TABLE 2 Univariate and multivariate analyses for OS in GC patients.

Variable	Univariate			Multivariate		
	HR	95%CI	*P*	HR	95%CI	*P*
Age (years)						
≥60 vs.<60	0.633	0.346–1.155	0.136			
Gender						
Male vs. Female	1.921	1.079–1.3.419	**0.026**	1.946	1.127–3.361	**0.017**
Tumor size						
>5 cm vs. ≤5 cm	1.774	0.958–3.284	0.068	1.866	1.060–3.285	**0.031**
Borrmann type						
III-IV vs. I-II	3.327	1.107–9.998	**0.032**	3.999	1.344–11.89	**0.013**
Differentiation						
poor vs. Well+ moderate	0.5922	0.27–1.297	0.190			
Depth of invasion						
T3–4 vs. T1–2	4.368	0.568–33.594	0.157			
Lymph node metastasis						
N+ vs. N0	4.050	1.632–10.048	**0.003**	4.815	2.05–11.306	**0.001**

(Continued)

TABLE 2 Continued

Variable	Univariate			Multivariate		
	HR	95%CI	*P*	HR	95%CI	*P*
CEA level (μg/L)						
>5 vs. ≤5	1.052	0.498–2.223	0.895			
LVI						
Present vs. none	1.291	0.680–2.453	0.435			
PNI						
Present vs. none	1.047	0.372–2.944	0.931			
GLT8D2						
High vs. Low	2.165	1.149–4.078	**0.017**	2.078	1.155–3.738	**0.015**

Bold values indicate P < 0.05.

TABLE 3 Univariate and multivariate analyses for DFS in GC patients.

Variable	Univariate			Multivariate		
	HR	95%CI	*P*	HR	95%CI	*P*
Age (years)						
≥60 vs.<60	0.672	0.367–1.231	0.672			
Gender						
Male vs. Female	1.675	0.950–2.952	0.074	1.752	1.015–3.023	**0.044**
Tumor size						
>5 cm vs. ≤5 cm	1.832	0.981–3.421	0.058	2.008	1.133–3.557	**0.017**
Borrmann type						
III-IV vs. I-II	3.332	1.114–9.962	**0.031**	3.976	1.339–11.81	**0.013**
Differentiation						
Poor vs. Well+ moderate	0.622	0.287–1.347	0.229			
Depth of invasion						
T3–4 vs. T1–2	4.532	0.591–34.738	0.146			
Lymph node metastasis						
N+ vs. N0	3.981	1.608–9.855	**0.003**	5.062	2.157–11.88	**0.001**
CEA level (μg/L)						
>5 vs. ≤5	0.961	0.459–2.012	0.915			
LVI						
Present vs. none	1.708	0.906–3.222	0.098			
PNI						
Present vs. none	0.885	0.317–2.475	0.836			
GLT8D2						
High vs. Low	2.142	1.149–3.993	**0.017**	2.091	1.167–3.748	**0.013**

Bold values indicate P < 0.05.

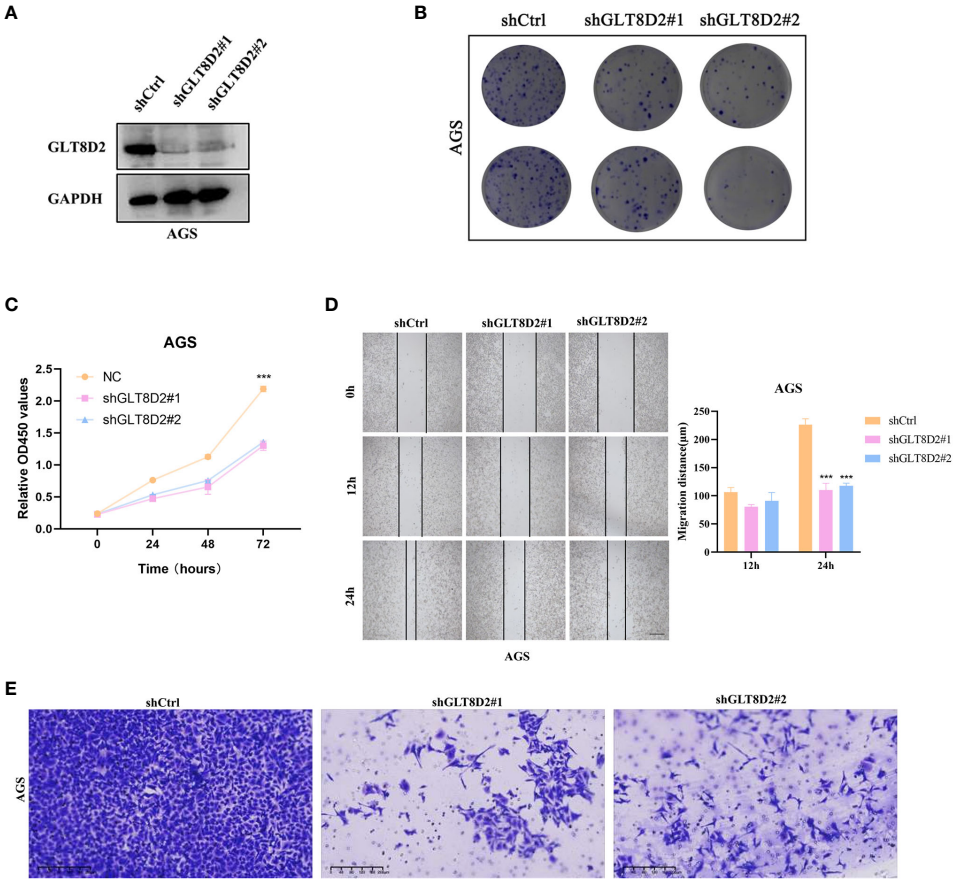


FIGURE 9
GLT8D2 knockdown blocks the proliferation and metastasis of GC cells *in vitro*. **(A)** Evaluation of the efficiency of shRNA via Western blotting. **(B)** Colony formation assays showing the clonogenic capacity of GC cells following GLT8D2 knockdown. **(C)** Cells growth ability after GLT8D2 knockdown were determined by CCK8 assay. **(D)** The wound healing assay showing the migration ability of GC cells following GLT8D2 knockdown. **(E)** Transwell assay showing the migratory capacity of GC cells following GLT8D2 knockdown.

increased expression of GLT8D2 is closely correlated with adverse prognosis and may underscore its role in regulating immune cell infiltration in GC patients, shedding new light on its potential key role as a prognostic biomarker related to immune infiltration in 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

This study was approved by the ethics committee of Guangdong Provincial People's Hospital. The informed consent form signed by all patients was obtained from Shanghai Outdo Biotech Company. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

HW: Writing – original draft, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation. JiZ: Writing – original draft, Data curation, Investigation, Methodology. QM: Data curation, Writing – original draft. JuZ: Funding acquisition, Resources, Supervision, Writing – review & editing. YL: Funding acquisition, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the National Natural Science Foundation of China (32370836), the National Key Clinical Specialty Construction Project (2021-2024, No. 2022YW030009), the Fundamental Research Funds for the Central Universities (21623303), the Guangdong Basic and Applied Basic Research Foundation (2023A1515110954), and Funding by Science and Technology Projects in Guangzhou (2024A04J4101). All these study sponsors had no role in the study design or in the collection, analysis or interpretation of the data.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1370367/full#supplementary-material

References

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

2. Zong L, Abe M, Seto Y, Ji J. The challenge of screening for early gastric cancer in China. *Lancet*. (2016) 388:2606. doi: 10.1016/S0140-6736(16)32226-7

3. Joshi SS, Badgwell BD. Current treatment and recent progress in gastric cancer. *CA Cancer J Clin*. (2021) 71:264–79. doi: 10.3322/caac.21657

4. Pinho SS, Figueiredo J, Cabral J, Carvalho S, Dourado J, Magalhães A, et al. E-cadherin and adherens-junctions stability in gastric carcinoma: functional implications of glycosyltransferases involving N-glycan branching biosynthesis, N-acetylglucosaminyltransferases III and V. *Biochim Biophys Acta*. (2013) 1830:2690–700. doi: 10.1016/j.bbagen.2012.10.021

5. Pinho SS, Reis CA, Paredes J, Magalhães AM, Ferreira AC, Figueiredo J, et al. The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Hum Mol Genet*. (2009) 18:2599–608. doi: 10.1093/hmg/ddp194

6. Gao Y, Liu Z, Feng J, Sun Q, Zhang B, Zheng W, et al. Expression pattern of polypeptide N-acetylgalactosaminyltransferase-10 in gastric carcinoma. *Oncol letters*. (2013) 5:113–6. doi: 10.3892/ol.2012.980

7. Hu WT, Yeh CC, Liu SY, Huang MC, Lai IR. The O-glycosylating enzyme GALNT2 suppresses the Malignancy of gastric adenocarcinoma by reducing EGFR activities. *Am J Cancer Res*. (2018) 8(9):1739–51.

8. Kümler I, Tuxen MK, Nielsen DL. A systematic review of dual targeting in HER2-positive breast cancer. *Cancer Treat Rev*. (2014) 40:259–70. doi: 10.1016/j.ctrv.2013.09.002

9. Blangé D, Stroes CI, Derks S, Bijlsma MF, van Laarhoven HWM. Resistance mechanisms to HER2-targeted therapy in gastroesophageal adenocarcinoma: A systematic review. *Cancer Treat Rev*. (2022) 108:102418. doi: 10.1016/j.ctrv.2022.102418

10. Fang R, Xu F, Shi H, Wu Y, Cao C, Li H, et al. LAMTOR5 raises abnormal initiation of O-glycosylation in breast cancer metastasis via modulating GALNT1 activity. *Oncogene*. (2020) 39:2290–304. doi: 10.1038/s41388-019-1146-2

11. Nguyen AT, Chia J, Ros M, Hui KM, Saltel F, Bard F. Organelle specific O-glycosylation drives MMP14 activation, tumor growth, and metastasis. *Cancer Cell*. (2017) 32:639–53.e6. doi: 10.1016/j.ccell.2017.10.001

12. Li Y, Zhang R, Hei H. Advances in post-translational modifications of proteins and cancer immunotherapy. *Front Immunol*. (2023) 14:1229397. doi: 10.3389/fimmu.2023.1229397

13. Schjoldager KT, Narimatsu Y, Joshi HJ, Clausen H. Global view of human protein glycosylation pathways and functions. *Nat Rev Mol Cell Biol*. (2020) 21:729–49. doi: 10.1038/s41580-020-00294-x

14. Varki A. Since there are PAMPs and DAMPs, there must be SAMPs? Glycan "self-associated molecular patterns" dampen innate immunity, but pathogens can mimic them. *Glycobiology*. (2011) 21:1121–4. doi: 10.1093/glycob/cwr087

15. Srivastava AK, Guadagnin G, Cappello P, Novelli F. Post-translational modifications in tumor-associated antigens as a platform for novel immuno-oncology therapies. *Cancers (Basel)*. (2022) 15(1):138. doi: 10.3390/cancers15010138

16. van Vliet SJ, Gringhuis SI, Geijtenbeek TB, van Kooyk Y. Regulation of effector T cells by antigen-presenting cells via interaction of the C-type lectin MGL with CD45. *Nat Immunol*. (2006) 7:1200–8. doi: 10.1038/ni1390

17. Li CW, Lim SO, Xia W, Lee HH, Chan LC, Kuo CW, et al. Glycosylation and stabilization of programmed death ligand-1 suppresses T-cell activity. *Nat Commun*. (2016) 7:12632. doi: 10.1038/ncomms12632

18. Feng C, Zhang L, Chang X, Qin D, Zhang T. Regulation of post-translational modification of PD-L1 and advances in tumor immunotherapy. *Front Immunol*. (2023) 14:1230135. doi: 10.3389/fimmu.2023.1230135

19. Shen C, Zheng B, Chen Z, Zhang W, Chen X, Xu S, et al. Identification of prognostic models for glycosylation-related subtypes and tumor microenvironment infiltration characteristics in clear cell renal cell cancer. *Heliyon*. (2024) 10:e27710. doi: 10.1016/j.heliyon.2024.e27710

20. Hu H, He B, He M, Tao H, Li B. A glycosylation-related signature predicts survival in pancreatic cancer. *Aging*. (2023) 15:13710–37. doi: 10.18632/aging.v15i23

21. Dai T, Li J, Liang RB, Yu H, Lu X, Wang G. Identification and experimental validation of the prognostic significance and immunological correlation of glycosylation-related signature and ST6GALNAC4 in hepatocellular carcinoma. *J hepatocellular carcinoma*. (2023) 10:531–51. doi: 10.2147/JHC.S400472

22. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. (2019) 10:1523. doi: 10.1038/s41467-019-09234-6

23. Zhao Y, Sato Y, Isaji T, Fukuda T, Matsumoto A, Miyoshi E, et al. Branched N-glycans regulate the biological functions of integrins and cadherins. *FEBS J*. (2008) 275:1939–48. doi: 10.1111/j.1742-4658.2008.06346.x

24. Takeuchi H, Haltiwanger RS. Significance of glycosylation in Notch signaling. *Biochem Biophys Res Commun*. (2014) 453:235–42. doi: 10.1016/j.bbrc.2014.05.115

25. Gomes C, Osório H, Pinto MT, Campos D, Oliveira MJ, Reis CA. Expression of ST3GAL4 leads to SLe(x) expression and induces c-Met activation and an invasive phenotype in gastric carcinoma cells. *PloS One*. (2013) 8:e66737. doi: 10.1371/journal.pone.0066737

26. de-Freitas-Junior JC, Carvalho S, Dias AM, Oliveira P, Cabral J, Seruca R, et al. Insulin/IGF-I signaling pathways enhances tumor cell invasion through bisecting GlcNAc N-glycans modulation. an interplay with E-cadherin. *PloS One*. (2013) 8:e81579. doi: 10.1371/journal.pone.0081579

27. Boscher C, Dennis JW, Nabi IR. Glycosylation, galectins and cellular signaling. *Curr Opin Cell Biol*. (2011) 23:383–92. doi: 10.1016/j.ceb.2011.05.001

28. Bassagañas S, Carvalho S, Dias AM, Pérez-Garay M, Ortiz MR, Figueras J, et al. Pancreatic cancer cell glycosylation regulates cell adhesion and invasion through the modulation of α2β1 integrin and E-cadherin function. *PloS One*. (2014) 9(5):e98595. doi: 10.1371/journal.pone.0098595

29. Dennis JW, Nabi IR, Demetriou M. Metabolism, cell surface organization, and disease. *Cell*. (2009) 139:1229–41. doi: 10.1016/j.cell.2009.12.008

30. Wagatsuma T, Nagai-Okatani C, Matsuda A, Masugi Y, Imaoka M, Yamazaki K, et al. Discovery of pancreatic ductal adenocarcinoma-related aberrant glycosylations: A multilateral approach of lectin microarray-based tissue glycomic profiling with public transcriptomic datasets. *Front Oncol*. (2020) 10:338. doi: 10.3389/fonc.2020.00338

31. Potapenko IO, Lüders T, Russnes HG, Helland Å, Sørlie T, Kristensen VN, et al. Glycan-related gene expression signatures in breast cancer subtypes; relation to survival. *Mol Oncol*. (2015) 9:861–76. doi: 10.1016/j.molonc.2014.12.013

32. Milde-Langosch K, Karn T, Schmidt M, zu Eulenburg C, Oliveira-Ferrer L, Wirtz RM, et al. Prognostic relevance of glycosylation-associated genes in breast cancer. *Breast Cancer Res Treat*. (2014) 145:295–305. doi: 10.1007/s10549-014-2949-z

33. Noda M, Okayama H, Tachibana K, Sakamoto W, Saito K, Thar Min AK, et al. Glycosyltransferase gene expression identifies a poor prognostic colorectal cancer subtype associated with mismatch repair deficiency and incomplete glycan synthesis. *Clin Cancer Res*. (2018) 24:4468–81. doi: 10.1158/1078-0432.CCR-17-3533

34. Mao Y, Zhang Y, Fan S, Chen L, Tang L, Chen X, et al. GALNT6 promotes tumorigenicity and metastasis of breast cancer cell via β-catenin/MUC1-C signaling pathway. *Int J Biol Sci*. (2019) 15:169–82. doi: 10.7150/ijbs.29048

35. Kawamura YI, Toyota M, Kawashima R, Hagiwara T, Suzuki H, Imai K, et al. DNA hypermethylation contributes to incomplete synthesis of carbohydrate determinants in gastrointestinal cancer. *Gastroenterology*. (2008) 135:142–51.e3. doi: 10.1053/j.gastro.2008.03.031

36. Zhao Y, Hu S, Zhang J, Cai Z, Wang S, Liu M, et al. Glucoside xylosyltransferase 2 as a diagnostic and prognostic marker in gastric cancer via comprehensive analysis. *Bioengineered*. (2021) 12:5641–54. doi: 10.1080/21655979.2021.1967067

37. Klauschen F, Müller KR, Binder A, Bockmayr M, Hägele M, Seegerer P, et al. Scoring of tumor-infiltrating lymphocytes: From visual estimation to machine learning. *Semin Cancer Biol*. (2018) 52:151–7. doi: 10.1016/j.semcancer.2018.07.001

38. Curigliano G. Gyneco-oncological genomics and emerging biomarkers for cancer treatment with immune-checkpoint inhibitors. *Semin Cancer Biol*. (2018) 52:253–8. doi: 10.1016/j.semcancer.2018.05.004

39. Liang C, Zhao Y, Chen C, Huang S, Deng T, Zeng X, et al. Higher TOX genes expression is associated with poor overall survival for patients with acute myeloid leukemia. *Front Oncol*. (2021) 11:740642. doi: 10.3389/fonc.2021.740642

40. Zhao Y, Liao P, Huang S, Deng T, Tan J, Huang Y, et al. Increased TOX expression associates with exhausted T cells in patients with multiple myeloma. *Exp Hematol Oncol*. (2022) 11:12. doi: 10.1186/s40164-022-00267-0

41. Huang S, Zhao Y, Liao P, Wang J, Li Z, Tan J, et al. Different expression patterns of VISTA concurrent with PD-1, Tim-3, and TIGIT on T cell subsets in peripheral blood and bone marrow from patients with multiple myeloma. *Front Oncol*. (2022) 12:1014904. doi: 10.3389/fonc.2022.1014904

42. Rosenberg SA, Restifo NP. Adoptive cell transfer as personalized immunotherapy for human cancer. *Sci (New York NY)*. (2015) 348:62–8. doi: 10.1126/science.aaa4967

43. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell*. (2017) 168:707–23. doi: 10.1016/j.cell.2017.01.017

44. Zhan Y, Zhao F, Xie P, Zhong L, Li D, Gai Q, et al. Mechanism of the effect of glycosyltransferase GLT8D2 on fatty liver. *Lipids Health Dis*. (2015) 14:43. doi: 10.1186/s12944-015-0040-3

45. Huang S, Liang S, Chen G, Chen J, You K, Ye H, et al. Overexpression of glycosyltransferase 8 domain containing 2 confers ovarian cancer to CDDP resistance by activating FGFR/PI3K signalling axis. *Oncogenesis*. (2021) 10:55. doi: 10.1038/s41389-021-00343-w

46. Sammarco G, Gadaleta CD, Zuccalà V, Albayrak E, Patruno R, Milella P, et al. Tumor-associated macrophages and mast cells positive to tryptase are correlated with angiogenesis in surgically-treated gastric cancer patients. *Int J Mol Sci*. (2018) 19 (4):1176. doi: 10.3390/ijms19041176

47. Zhang D, Qiu X, Li J, Zheng S, Li L, Zhao H. TGF-β secreted by tumor-associated macrophages promotes proliferation and invasion of colorectal cancer via miR-34a-VEGF axis. *Cell Cycle*. (2018) 17:2766–78. doi: 10.1080/15384101.2018.1556064

48. Wynn TA, Chawla A, Pollard JW. Macrophage biology in development, homeostasis and disease. *Nature*. (2013) 496:445–55. doi: 10.1038/nature12034

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 07 June 2024
DOI 10.3389/fonc.2024.1403052

OPEN ACCESS

EDITED BY
Dalila Luciola Zanette,
Oswaldo Cruz Foundation (Fiocruz), Brazil

REVIEWED BY
Kevin Roarty,
Baylor College of Medicine, United States
Mateus Aoki,
Oswaldo Cruz Foundation (Fiocruz), Brazil

*CORRESPONDENCE
Gregory Lizée,
glizee@mdanderson.org

RECEIVED 18 March 2024
ACCEPTED 10 May 2024
PUBLISHED 07 June 2024

CITATION
Sonnemann HM, Pazdrak B, Nassif B, Sun Y, Elzohary L, Talukder AH, Katailiha AS, Bhat K and Lizée G (2024) Placental co-transcriptional activator Vestigial-like 1 (VGLL1) drives tumorigenesis via increasing transcription of proliferation and invasion genes.
Front. Oncol. 14:1403052.
doi: 10.3389/fonc.2024.1403052

Placental co-transcriptional activator Vestigial-like 1 (VGLL1) drives tumorigenesis via increasing transcription of proliferation and invasion genes

Heather M. Sonnemann[1,2], Barbara Pazdrak[2], Barbara Nassif[1,2], Yimo Sun[1,2], Lama Elzohary[2], Amjad H. Talukder[2], Arjun S. Katailiha[2], Krishna Bhat[3] and Gregory Lizée[2,4]*

[1]University of Texas MD Anderson Cancer Center, UTHealth Graduate School of Biomedical Sciences, Houston, TX, United States, [2]Department of Melanoma Medical Oncology, UT MD Anderson Cancer Center, Houston, TX, United States, [3]Department of Translational Molecular Pathology, UT MD Anderson Cancer Center, Houston, TX, United States, [4]Department of Immunology, UT MD Anderson Cancer Center, Houston, TX, United States

Introduction: Vestigial-like 1 (VGLL1) is a co-transcriptional activator that binds to TEA domain-containing transcription factors (TEADs). Its expression is upregulated in a variety of aggressive cancer types, including pancreatic and basal-like breast cancer, and increased transcription of VGLL1 is strongly correlated with poor prognosis and decreased overall patient survival. In normal tissues, VGLL1 is most highly expressed within placental trophoblast cells, which share the common attributes of rapid cellular proliferation and invasion with tumor cells. The impact of VGLL1 in cancer has not been fully elucidated and no VGLL1-targeted therapy currently exists.

Methods: The aim of this study was to evaluate the cellular function and downstream genomic targets of VGLL1 in placental, pancreatic, and breast cancer cells. Functional assays were employed to assess the role of VGLL1 in cellular invasion and proliferation, and ChIP-seq and RNAseq assays were performed to identify VGLL1 target genes and potential impact using pathway analysis.

Results: ChIP-seq analysis identified eight transcription factors with a VGLL1-binding motif that were common between all three cell types, including TEAD1-4, AP-1, and GATA6, and revealed ~3,000 shared genes with which VGLL1 interacts. Furthermore, increased VGLL1 expression led to an enhancement of cell invasion and proliferation, which was supported by RNAseq analysis showing transcriptional changes in several genes known to be involved in these processes.

Discussion: This work expands our mechanistic understanding of VGLL1 function in tumor cells and provides a strong rationale for developing VGLL1-targeted therapies for treating cancer patients.

KEYWORDS

pancreatic cancer, breast cancer, migration and invasion, mechanism of tumor progression, mechanism of transcription, cell proliferation

Introduction

Vestigial-like 1 (VGLL1) is a mammalian co-transcriptional activator that has been shown to play an important role in placental development and is also highly expressed in various cancer types. It was first discovered as a co-transcriptional activator in drosophila called Vestigial (vg), where it functions as a key master regulator of wing development (1, 2). Substitution of drosophila vg with the human ortholog VGLL1 (previously designated tondu) rescued wing development, underscoring the high level of conservation of this gene and its encoded protein (2). Studies over the past decade have begun illuminating the role of mammalian VGLL1 and have focused largely on two broad areas of developmental biology: normal placental development and tumorigenesis. VGLL1 is expressed at a low to negligible level in most normal tissues but demonstrates by far the highest expression in placental trophoblasts (3–5). These cells are highly proliferative and invasive, attributes that are critical for embryonic implantation (6, 7). Although little is known about the precise role VGLL1 plays in placental development, aberrantly elevated expression of this gene has also been observed in several aggressive cancer types. Pancreatic, triple negative breast, gastric, and HPV-related cancers such as ovarian cancer show the highest levels of VGLL1 mRNA expression, and increased transcription of VGLL1 is strongly correlated with poor prognosis and lower overall patient survival (3, 4, 8–11).

Crucial clues about the function of VGLL1 came from molecular studies designed to understand the structure-function relationship between TEA domain-containing transcription factors (TEADs) and different co-transcriptional activators. These studies revealed that the binding of VGLL1 to TEADs showed high structural similarity to that of co-transcriptional activators in the Hippo pathway, Yes-associated protein 1 (YAP1), and WW Domain containing transcription regulator 1 (WWTR1/TAZ) (2, 12). Since the hippo pathway has been strongly implicated in the tumorigenesis of many cancer types, this study shed important light on the potential functional role of VGLL1 in cancer (13). Furthermore, YAP1/TAZ both have well-characterized roles in the development of many normal tissues, including cytotrophoblast differentiation in the placenta (7). The intriguing relationship between tumorigenesis and placental development has been widely speculated for several decades (14, 15). However, recent evidence is emerging that VGLL1 may provide a crucial mechanistic link between these two processes.

Cancer cells and placental trophoblasts share at least three cardinal attributes: rapid cellular proliferation, tissue invasion, and the induction of immune suppression (6, 16, 17). Therefore, in the current study we investigated the effect of VGLL1 on cell proliferation and invasion by downregulation or upregulation of its expression levels in cancer cells originating from pancreas, breast, and placenta using siRNA knockdown or retrovirus overexpression systems. In addition, we performed RNA sequencing analysis of these tumor cells to evaluate transcriptional changes associated with regulating VGLL1 protein expression. Finally, ChIP-seq analysis was performed to identify and compare the chromosomal loci with which VGLL1 interacts in the three different cancer cell types. These studies confirm a key role for VGLL1 in driving cell proliferation and invasion by promoting the expression of specific genes involved in these processes and provide novel mechanistic insights into how VGLL1 may contribute to tumor aggressiveness and poor clinical outcomes in cancer patients.

Methods

Cell culture

Human pancreatic cell lines PANC1 (cat#CRL-1469, RRID: CVCL_0480), PANC10.05 (RRID: CVCL_1639, cat#CRL-2547), Capan1 (RRID: CVCL_0237, cat#HTB-79), Capan2 (RRID: CVCL_0026, cat#HTB-80), SU8686 (RRID: CVCL_3881, cat#CRL-1837) and human breast cancer cell lines BT20 (RRID: CVCL_0178, cat#HTB-19), MDA-MB-468 (RRID: CVCL_0419, cat#HTB-25), MDA-MB-175-VII (RRID: CVCL_1400, cat#HTB-132), and human choriocarcinoma cells, Bewo (RRID: CVCL_0044, cat#CCL-98), were obtained from ATCC and tested negative for mycoplasma contamination. PANC1 and PANC10.05 cells were grown in DMEM (Gibco, cat#11965-092) or RPMI (Gibco, cat#11875-093) media containing 10% heat inactivated BenchMark FBS (Gemini, cat#100-106). Bewo and BT20 cells were cultured in Ham's F-12K (Gibco, cat#21-127-022) or EMEM (ATCC, cat#30-2003) media, respectively, also supplemented with 10% FBS. Retroviral-producing cells, Phoenix Eco (RRID: CVCL_H717, cat#CRL-3214) and PG-13 (RRID: CVCL_4273, cat#CRL-10686) cells were obtained from ATCC and tested negative for mycoplasma contamination. The cells were cultured

in DMEM media supplemented with 10% FBS. Cell lines were authenticated by the Cytogenetic and Cell Authentication Core at MD Anderson Cancer Center. MycoAlert Mycoplasma Detection Kit (Lonza, cat#LT07-318) was used to test cells for mycoplasm throughout the study.

Cloning of human VGLL1-Myc cDNA into retrovirus MG-neo vector

Human VGLL1 with MYC tag containing restriction enzyme sites Not1 and Sal1 (hVGLL1-MYC, sequence below) was synthesized by GeneArt and inserted in pENTR221 vector. NEB-10 beta competent cells (New England Biolabs, cat#C3019H) were used to expand the vectors following manufacturer protocol. To clone hVGLL1-MYC cDNA into retrovirus MG-neo vector, hVGLL1-MYC pENTR221 and MG-neo plasmids were digested using Not1-HF (New England Biolabs, cat#R3189S) and Sal1-HF (New England Biolabs, cat#R3138S) for 1 hour at room temperature. Gel purified hVGLL1-MYC was ligated to the MG-neo vector destination plasmid using T4 DNA ligase (New England Biolabs, cat#101228-180) for 1 hr at room temperature. Next, bacteria transformation was performed for hVGLL1-MYC MG-neo plasmid expansion, and the plasmid was purified using Wizard Plus SV Minipreps DNA purification systems (Promega, cat#A1470). To verify the hVGLL1-MYC insert sequence, the plasmid was digested using restriction enzymes Not1 and Sal1 and ran on a 1% agarose gel (Lonza, cat#50002) followed by purification by QIAquick Gel Extraction Kit (Qiagen, cat# 28704). The insert sequence was confirmed using Advanced Technology Genomics Core at MD Anderson Cancer Center using Forward primer: AATTCGCCAGCACAGTGGAGATC and Reverse primer: CGGCAATATGGTGGAAAATAACCGG. The sequencing results were viewed on Chromas 2.6.6 (RRID: SCR_000598) software and sequences were aligned using nucleotide BLAST (BLASTn suite, RRID: SCR_001598) (18). Next, hVGLL1-MYC MG-neo plasmid DNA was expanded and purified using a high-speed plasmid midi kit (Qiagen, cat#12643). The plasmid concentration was determined using a nanodrop.

cDNA hVGLL1-MYC Sequence with restriction enzyme sites (bold):

T C **G C G G C C G C** C A T G G A A G A A A T G A A G A A
G A C T G C C A T C C G G C T G C C C A A A G G C A A A
C A G A A G C C T A T A A A G A C G G A A T G G A A T T C
C C G G T G T G T C C T T T T C A C C T A C T T C C A A G G
G G A C A T C A G C A G C G T A G T G G A T G A A C A C T T
C T C C A G A G C T C T G A G C A A T A T C A A G A G C C C C
C A G G A A T T G A C C C C C T C G A G T C A G A G T G A A G G
T G T G A T G C T G A A A A A C G A T G A T A G C A T G T C T C C A A
A T C A G T G G C G T T A C T C G T C T C C A T G G A C A A A G C C A C A
A C C A G A A G T A C C T G T C A C A A A C C G T G C C G C C A
A C T G C A A C T T G C A T G T G C C T G G T C C C A T G G C T G T G A
A T C A G T T C T C A C C G T C C C T G G C T A G G A G G G C C T C T G T T
C G G C C T G G G G A G C T G T G G C A T T T C T C C T C C C T G G C G G
G C A C C A G C T C C T T A G A G C C T G G C T A C T C T C A T C C C T T C
C C C G C T C G G C A C C T G G T T C C A G A G C C C C A G C C T G A T G G
G A A A C G T G A G C C T C T C C T A A G T C T C C T C C A G C A A G A C A G A
T G C C T A G C C C G T C C T C A G G A A T C T G C C G C C A
G G G A G A A T G G C A A C C C T G G C C A G A T A G C T G
G A A G C A C A G G G T T G C T C T T C A A C C T G C C T C C C
G G C T C A G T T C A C T A T A A G A A A C T A T A T G T A T C T C
G T G G A T C T G C C A G T A C C A G C C T T C C A A A T G A A A C
T C T T T C A G A G T T A G A G A C A C C T G G G A A A T A C T C A
C T T A C A C C A C C A A A C C A C T G G G G C C A C C C A C A T C
G A T A C C T G C A G C A T C T T **G A G G G G A T C C** G G A G G
A G G C G G A T C T **G A G C A G A A A C T C A T C A G**
T G A A G A G G A C C T G T G A **G T C G A C** G

Not1 (GCGGCCGC)
BamH1 (GGATCC)
MYC (GAGCAGAAACTCATCAGTGAAGAGGACCTG)
Sal1 (GTCGAC)

Production of hVGLL1-MYC MG-neo retrovirus

For retrovirus production, Phoenix Eco cells plated in a 60 mm dish were incubated overnight to reach about 60-80% confluency. To transduce the cells, 5ug of hVGLL1-MYC MG-neo plasmid DNA and 30ul TransIT-293 (Mirus bio, Qiagen, cat#MIR2704) transfection reagent were each diluted in 300uL of Opti-MEM (Gibco, cat#51985-034) in separated tubes. After 10 minutes incubation at room temperature, the diluted DNA was added into a tube containing the diluted transfection reagent in a drop-wise fashion and incubated for another 10 minutes at room temperature. Next, media from Phoenix Eco cells was replaced with 1mL Opti-MEM and 600uL of the DNA mixture was added to the cells in a drop-wise fashion and cells were cultured overnight in the incubator. The next day, media were changed to DMEM with 10% FBS (without antibiotics) and the cells were incubated overnight. The next day, the supernatant from hVGLL1-MYC MG-neo transduced Phoenix Eco cells was collected, filter sterilized and mixed with 10ug/mL polybrene (Sigma, cat#TR-1003-G). To generate stock of viral-producing cell we transduced PG-13 cells with the supernatant prepared from transduced Phoenix Eco cells mixed with polybrene. Human VGLL1 gene in these PG-13 hVGLL1-MYC MG-neo cells was detectable after 48 hr.

Transduction of PANC1 cells with hVGLL1-MYC MG-neo retrovirus plasmid

To generate PANC1 cells overexpressing hVGLL1-MYC for the experiments, overnight cultured Panc1 cells with 80% confluency were incubated with the filtered supernatant from PG-13 hVGLL1-MYC MG-neo cells mixed with polybrene. PANC1 cells expressing hVGLL1-MYC were selected by cell culture in the presence of G418 sulfate antibiotic at a concentration of 0.5mg/ml for 4 days.

siRNA knockdown

PANC10.05, BT20, and Bewo cells were plated in 6-well plates and cultured overnight. The cells at 60-80% confluency were used for transfection with siVGLL1 (ThermoFisher Scientific, cat#4392420 s28152) or siScramble control (ThermoFisher Scientific, cat#4390846). Briefly, 9uL lipofectamine RNAiMAX transfection reagent (Invitrogen, cat#13778030) and 30pmol of siVGLL1 (ThermoFisher Scientific, cat#4392420 s28152) or siScramble (ThermoFisher Scientific, cat#4390846) were diluted in 150uL Opti-mem media (Gibco, cat#51985-034). Next, diluted siRNA was added to diluted lipofectamine RNAiMAX at 1:1 ratio and incubated for 5 minutes at room temperature. 250uL of combined siRNA and lipofectamine were added in a drop-wise fashion into the cells. The transfected cells incubated with the siRNA solution for 48 hrs were used for the indicated experiments.

Western blot

The indicated cancer cells were plated on 6-well plates at a density to reach about 90% confluency after 48h cultured. Next, the cells were washed twice with cold PBS (Corning, cat#21-040-CV) and lysed with 350uL of lysis buffer (50mM HEPES (Corning, cat#25-060-C1), 150mM NaCl (Invitrogen, cat#AM9760G), 1mM EDTA (Sigma, cat#03690-100ml), 1% triton X100 (Millipore, cat#T9284-100ml), 1mM PMSF (Sigma, cat#P-7626), and 1X Halt protease and phosphatase inhibitor (ThermoScientific, cat#78-138) for 45 minutes on ice. Protein concentrations were quantified using BCA assay (ThermoFisher) with BSA standard set (Bio-Rad). 10ug of protein from each sample was loaded into well of 10% Tris-Glycine gel (Invitrogen) and ran at 125V. The resolved proteins were transferred to a PVDF membrane (BioRad, cat#1704156) for 7 minutes at 25V using Bio-Rad Trans-Blot Turbo Transfer System (cat#1704150). Next, the membrane was blocked with EveryBlot Blocking Buffer (Bio-Rad) for 30 min at room temperature and probed with primary antibodies against VGLL1 (ProteinTech, cat#10124-2AP, RRID: AB_2218174), MYC (Cell Signaling, cat#2276), or GAPDH (Cell Signaling, cat#5174, RRID: AB_10622025) overnight at 4°C followed by incubation with anti-mouse (Cell Signaling, cat#7076, RRID: AB_330924) or anti-rabbit (Cell Signaling, cat#7074, RRID: AB_2099233) HRP-conjugated antibodies for 45 min at room temperature. Immunoblots were developed using Clarity Western ECL Substrate (Biorad, cat#1705060) and imaged by ChemiDoc MP (Biorad, cat#12003154).

VGLL1 immunoprecipitation

For immunoprecipitation (IP) of VGLL1 protein from cancer cells we used Pierce™ Classic Magnetic IP/Co-IP Kit (ThermoFisher Scientific, cat#88804). VGLL1 was IP from 200ug of beads precleared cell lysate using 5ug VGLL1 Ab (ProteinTech, cat#10124-2AP, RRID: AB_2218174) per sample. VGLL1-IP proteins were washed and eluted from the beads with Laemmli sample buffer. The resolved protein was detected by western blot with anti-VGLL1 Ab. We also included IP control samples without VGLL1 Ab to determine potential non-specific protein binding to the beads. Sample inputs contained 10ug of protein from the indicated cell lysate.

Real-time quantitative PCR

Real time quantitative PCR (RT-qPCR) was performed to determine hVGLL1 mRNA expression in PANC1, PANC10.05, BT20, Bewo cells. RNA was isolated from the cells cultured on 6-well plates (80-90% confluency) using RNAeasy Plus mini kit (Qiagen, cat#74134) according to the manufacturer's protocol. The sample RNA was qualified using a NanoDrop Onec (ThermoFisher). High-capacity RNA-to-cDNA kit (Applied Biosystems, cat#4387406) was used to generate cDNA from 0.5ug RNA following the manufacturer's instructions. cDNA was quantified using a NanoDrop. RT-qPCR amplifications were done with PowerUP SYBR Green Master Mix (Applied Biosystems, cat#A25742) using 500nM primers (see Table 1) and 10ng cDNA per reaction in a total volume of 20ul in 96-well plate. The samples were run in triplicate on a QuantStudio3 System (Applied Biosystems). VGLL1 mRNA expression was normalized to the mRNA expression level of GAPDH in the same sample. The results were then calculated as a relative VGLL1 mRNA expression compared to empty vector or siScramble controls that were expressed as 1 value.

Proliferation assay

The effect of VGLL1 expression on cancer cell growth was determined by daily counting of cells during seven days of culture. The cells were plated in triplicate at a density of 30,000 cells per well of 6-well plate. To count the cells, they were washed twice with PBS, trypsinized using 50uL of 0.25% Trypsin (Gibco, cat#325200-056), and collected with 1mL media. After centrifugation, the cell pellets were resuspended in 50uL of media and 50uL of 0.2% trypan blue (Lonza BioWhittaker, cat#17-942E). 20uL of the cell suspension was transferred into the cellometer slide (Nexcelom, cat#SD100) and analyzed for the number of live cells by a Nexelom cell counter.

TABLE 1 Table of primer sequences used in qRT-PCR.

Primer Name	Primer Sequence
Human GAPDH Forward	ACA ACT TTG GTA TCG TGG AAG G
Human GAPDH Reverse	GCC ATC ACG CCA CAG TTT C
Human VGLL1 Forward	CCA AAG GCA AAC AGA AGC CTA
Human VGLL1 Reverse	CAT CCAC ACC TTC ACT CTG ACT C
Human B-actin Forward	GCG AGA AGA TGA CC AGA TC
Human B-actin Reverse	CCA GTG GTA CGG CCA GAG G

Invasion assay

CHEMICON cell invasion assay (Millipore Sigma cat#ECM550) was used to determine the effect of VGLL1 expression on cancer cell invasion. Briefly, 500,000 cells were plated in 300uL serum-free media on top of a transwell coated with a basement membrane matrix and the bottom chamber was filled with 500ul media supplemented with 10% FBS. After 24h incubation, migrated cells on the bottom of the transwell were stained with the cell staining solution provided in the kit. Images were acquired using an Axiovert 200 microscope and quantified using ImageJ software (RRID: SCR_003070). Samples were plated in triplicate and four representative images were taken from each well.

Chromatin immunoprecipitation

For VGLL1 chromatin immunoprecipitation sequencing (ChIP-Seq) analysis, PANC10.05, BT20 and Bewo cells were sent to Active Motif (Carlsbad, CA). To prepare chromatin, the cells were fixed with 1% formaldehyde for 15 min and quenched with 0.125 M glycine. Chromatin was isolated by adding lysis buffer and disrupted with a Dounce homogenizer. Lysates were sonicated and the DNA sheared to an average length of 300-500 bp with Active Motif's EpiShear probe sonicator (cat#53051). Genomic DNA (Input) was prepared by treating aliquots of chromatin with RNase, proteinase K, and heat for de-crosslinking, followed by SPRI beads clean up (Beckman Coulter) and quantitation by Clariostar (BMG Labtech). For ChIP, 30 ug chromatin was precleared with protein G agarose beads (Invitrogen) and genomic DNA regions of interest were isolated using 5ug of VGLL1 Ab (Protein Tech, cat#10124-2-AP). IP complexes were washed, eluted from the beads with SDS buffer, and subjected to RNase and proteinase K treatment. Crosslinks were reversed by incubation overnight at 65°C. DNA from ChIP was purified by phenol-chloroform extraction and ethanol precipitation.

Quantitative PCR (QPCR) reactions were carried out only for PANC10.05 in triplicate on specific genomic regions using SYBR Green Supermix (Bio-Rad). The resulting signals were normalized for primer efficiency by carrying out QPCR for each primer pair using Input DNA.

ChIP sequencing (RRID: SCR_001237)

Illumina sequencing libraries were prepared from the ChIP and Input DNAs by the standard consecutive enzymatic steps of end-polishing, dA-addition, and adaptor ligation. Steps were performed on an automated system (Apollo 342, Wafergen Biosystems/Takara). After a final PCR amplification step, the resulting DNA libraries were quantified and sequenced on Illumina's NovaSeq 6000 (75 nt reads, single end). Reads were aligned to the human genome (hg38) using the BWA algorithm (default settings, RRID: SCR_010910) (19). Duplicate reads were removed, and only uniquely mapped reads (mapping quality >= 25) were used for further analysis. Alignments were extended in silico at their 3'-ends to a length of 200 bp, which is the average genomic fragment length in the size-selected library and assigned to 32-nt bins along the genome. The resulting histograms (genomic "signal maps") were stored in bigWig files. Peak locations were determined using the MACS algorithm (v2.1.0) (20) with a cutoff of p-value = 1e-7. Peaks that were on the ENCODE blacklist of known false ChIP-Seq peaks were removed. Signal maps and peak locations were used as input data to Active Motifs proprietary analysis program, which creates Excel tables containing detailed information on sample comparison, peak metrics, peak locations and gene annotations. Other key software used: bcl2fastq2 (v2.20) (processing of Illumina base-call data and demultiplexing), Samtools (v0.1.19, RRID: SCR_002105) (processing of BAM files), BEDtools (v2.25.0, RRID: SCR_006646) (processing of BED files), wigToBigWig (v4) (generation of bigWIG files).

ChIP-Atlas was used to compare our ChIP-seq data to publicly available ChIP-seq datasets using ChIP-Atlas online tool (https://chip-atlas.org/enrichment_analysis, RRID: SCR_015511) (21, 22).

RNA sequencing

RNA sequencing analysis from total RNA isolated from different cancer cells was performed by Active Motif (Carlsbad, CA). For each sample, 500ng of total RNA was then used in Illumina's TruSeq Stranded mRNA Library kit (Cat# 20020594). Libraries were sequenced on Illumina NovaSeq 6000 as paired-end 150-nt reads. Sequence reads were analyzed with the STAR alignment–DESeq2 software pipeline (RRID: SCR_004463). GO analysis was used to analyze several gene sets using the GO online tool (http://geneontology.org/).

Statistical analyses

Graphpad Prism 10 (RRID: SCR_000306) was used to perform unpaired t test analysis for cell invasion assays and two-way ANOVA analysis for proliferation assays. A $p<0.05$ was considered statistically significant.

Results

VGLL1 expression promotes tumor cell proliferation and invasion

This study aimed to identify common and unique VGLL1-dependent functions in tumor cells derived from 3 different cancer types that may contribute to disease progression. Based on the elevated VGLL1 transcript expression reported in particular tumor types, we screened pancreatic cancer (PDAC), basal-like breast cancer (BLBC), and choriocarcinoma (a placenta-derived tumor) cells for VGLL1 protein expression. Western blot analysis demonstrated relatively high VGLL1 protein levels in PANC10.05 (PDAC), BT20 (BLBC), MDA-MB-468 (BLBC) and Bewo

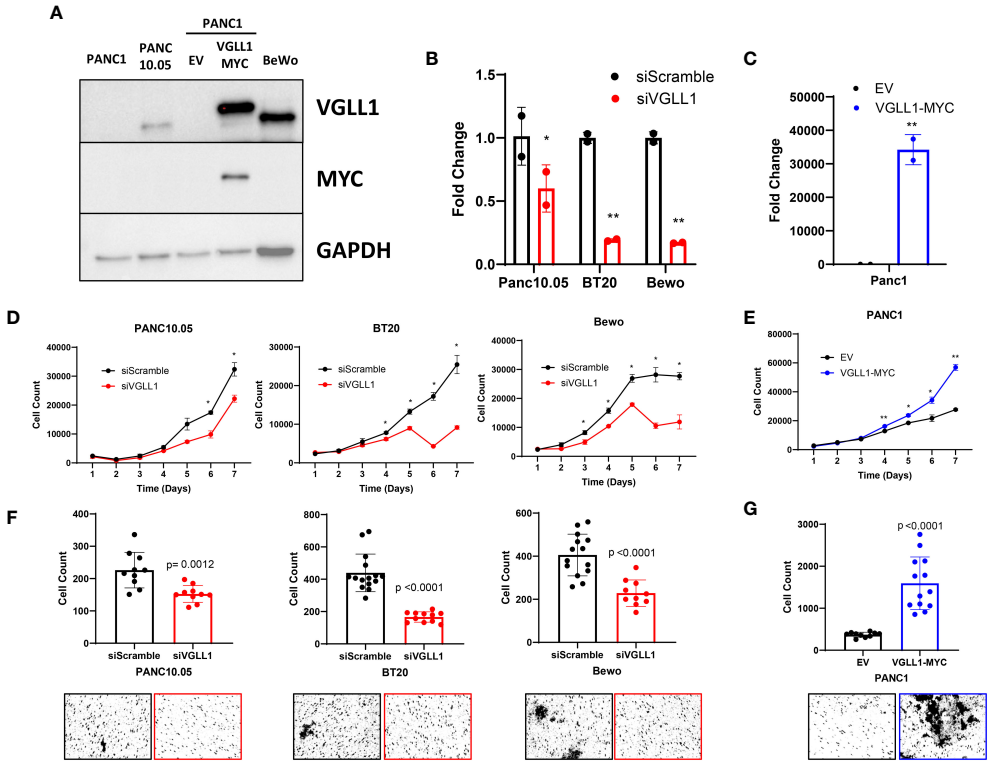


FIGURE 1
VGLL1 expression drives tumor cell proliferation and invasion. **(A)** Western blot analysis to assess VGLL1 protein expression in the cell lines PANC1, PANC10.05, Bewo, and PANC1 cells transduced to express either hVGLL1-MYC or empty vector (EV). **(B)** RT-qPCR analysis of VGLL1 mRNA expression in siScramble- (black) or siVGLL1- (red) treated cells after 48h. **(C)** RT-qPCR analysis of VGLL1 mRNA expression after retroviral transduction of PANC1 cells with hVGLL1-MYC MG-neo plasmid (blue) or empty MG-neo vector (black). **(D, E)** Cell proliferation analysis of siScramble- (black) or siVGLL1-treated PANC10.05, BT20, and Bewo cells (red), and PANC1 cells transduced with empty MG-neo vector (black) or hVGLL1-MYC MG-neo plasmid (blue). **(F, G)** Cell invasion assay results comparing siScramble (black) or siVGLL1-treated PANC10.05, BT20, and Bewo cells (red), and PANC1 cells transduced with empty MG-neo vector (black) or VGLL1-MYC MG-neo plasmid (blue). The indicated cells were plated onto the top of a transwell chamber and after 24 hr incubation, migrated invading cells were stained. Images were taken and quantified using ImageJ, with representative images displayed. *p<0.05, **p<0.01.

(choriocarcinoma) cells (Supplementary Figure 1). Based on these results we selected PANC10.05, BT20 and Bewo cells for our study, which represent cancer cells originating from pancreas, breast, and placenta tumors, respectively. By contrast, VGLL1 expression was undetectable in PANC1 PDAC cells (Figure 1; Supplementary Figure 1). Thus, PANC1 cells represent the ~60% of PDAC patients whose tumors do not express VGLL1. Since recent studies have suggested that VGLL1 is critical for the transition from classical to a more aggressive basal-like form of PDAC (11), we utilized PANC1 cells to assess phenotypic changes induced by ectopic VGLL1 overexpression.

Uncontrolled tumor cell proliferation and invasion are amongst the most important hallmarks of cancer progression. To determine how VGLL1 may impact these cellular processes and drive associated transcriptional changes, we utilized two *in vitro* models: siRNA knockdown and retrovirus overexpression systems. For downregulating VGLL1 expression, PANC10.05, BT20, and Bewo cells (expressing high levels of endogenous VGLL1) were transfected with the VGLL1-targeting siRNA or siScramble control. For overexpressing VGLL1, PANC1 cells were transduced with the hVGLL1-MYC MG-neo retroviral vector. RT-qPCR analysis of RNA isolated from these cells showed about 50% (PANC10.05) to 75% (BT20 and Bewo) reduction of VGLL1 transcript levels in cells transfected with siVGLL1 compared to siScramble control (Figure 1B). On the other hand, PANC1 cells transduced with the hVGLL1-MYC retroviral vector demonstrated more than a 3-log increase in VGLL1 mRNA levels compared to empty vector-transduced cells (Figure 1C). Furthermore, immunoblotting with anti-VGLL1 or anti-MYC Abs confirmed high VGLL1-MYC protein expression in transduced PANC1 cells (Figure 1A).

To evaluate the impact of VGLL1 expression on cancer cell proliferation, we monitored cell growth of these genetically altered and control cells for 7 days (Figures 1D, E). Tumor cells were plated at a density of 30,000 cells per well in a 6-well plate, and the number of live cells was quantitated daily. A 1.5-fold reduction in proliferation rate was observed by day 5 in BT20 and Bewo cells treated with siVGLL1 compared with control siScramble-treated cells (Figure 1D). A similar 1.7-fold reduction in proliferation of siVGLL1-treated PANC10.05 cells was observed by day 6 of culture (Figure 1D). Of note, in addition to inhibition of cell growth, a significant increase in cell death was observed in BT20 and Bewo cells by day 6 of culture leading to about 3-fold reduction in the number of live cells compared to siScramble-treated cells. This

observation was consistent with our initial attempts at CRISPR-mediated complete VGLL1 knockdown, which was lethal to the cells (data not shown). Therefore, to maintain the integrity of tumor cell lines, all further experiments were performed within 48 hrs of siRNA knockdown. By contrast, ectopic overexpression of VGLL1 in PANC1 cells clearly increased the rate of tumor cell proliferation, with a greater than 2-fold increase in cell number by day 7 of culture compared to empty vector-transduced cells (Figure 1E). These results demonstrated that while augmentation of VGLL1 expression enhanced cell proliferation rate, downregulation of VGLL1 expression was associated with inhibition of cancer cell growth and induction of cell death.

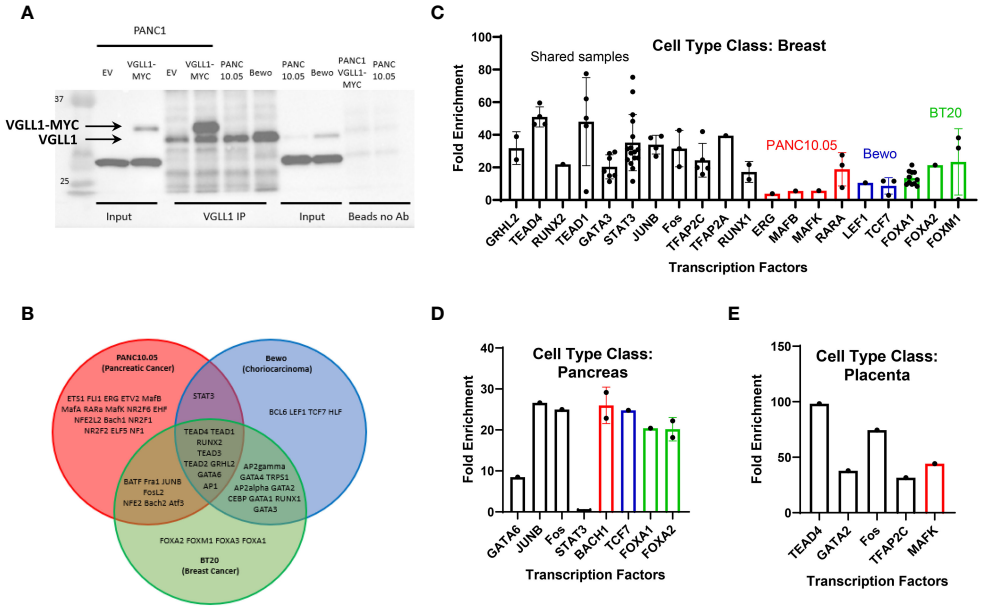
Tumor metastasis is the primary cause of mortality in most cancer patients. Since cellular invasion is critical for the metastatic process (23), we next assessed the impact of modulating VGLL1 expression on this process. The invasion capacity of cells was measured by their ability to break down a basement membrane matrix and subsequently migrate through the pores on the bottom side of a transwell. Quantitative analysis of the cell invasion assay revealed that siVGLL1-treated PANC10.05, BT20, and Bewo cells all demonstrated a 1.5 to 2.6-fold decrease in invasion compared to control siScramble-treated cells (Figure 1F). By contrast, VGLL1-transduced PANC1 cells showed a 4.3-fold increase in migrated cells compared to empty vector-transduced control cells (Figure 1G). These results support the notion that VGLL1 overexpression in tumors may play a substantial role in promoting tissue invasion, consistent with its association with tumor aggressiveness and poor patient prognosis.

VGLL1 interacts with distinct transcription factors to regulate transcription in different cancer cells

To identify the transcription factors (TFs) that can interact with VGLL1 in different cancer types, the chromosomal loci bound by this transcriptional co-activator were determined by chromatin immunoprecipitation sequencing (ChIP-seq) analysis performed on PANC10.05, BT20, and Bewo tumor cells. First, we validated the utility of anti-VGLL1 Ab by evaluating its specificity by Western blot analysis. As shown in Figure 2A, VGLL1 immunoprecipitation from lysates of both PANC10.05 and Bewo cells showed a single predominant band representing the native form of VGLL1 protein at approximately 29 kDa, consistent with its predicted molecular weight. As expected, MYC-tagged VGLL1 protein derived from transduced PANC1 cells migrated slower at about 32 kDa. A nonspecific band was observed at approximately 26 kDa in all input cell lysates; however, no such band was detected in the VGLL1 immunoprecipitated samples (Figure 2A). Further validation of the VGLL1 Ab was performed by rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) on VGLL1-MYC transduced PANC1 cells. The mass spectrometry results demonstrated a 68% coverage of the VGLL1 protein, including 28 unique peptides (data not shown). Collectively, these results confirmed that the VGLL1-specific antibody met the quality standards required for ChIP-seq analysis.

HOMER motif analysis performed on the ChIP-seq data derived from PANC10.05, BT20 and Bewo cells revealed multiple potential binding partners for VGLL1. The Venn diagram in Figure 2B depicts all the TFs that showed highly significant p-values of less than 1E-10. Some TFs were specific for individual cell lines, others were shared between two cell lines, and a total of eight TFs were shared amongst all three tumor types, including TEAD1, TEAD2, TEAD3, TEAD4, GRHL2, RUNX2, AP1, and GATA6. To validate our findings, we used ChIP-Atlas Enrichment analysis to compare our data to publicly available ChIP-seq data (21, 22). This analysis compares prior TF immunoprecipitation results derived from different tissue types or cell lines to our findings, with higher fold enrichment scores indicating a higher degree of similarity to our VGLL1 ChIP-seq data (Figures 2C–E). The breast-derived samples showed a high degree of similarity for multiple TFs, as shown in Figure 2C. Each dot represented a single database entry for the indicated specific TF, with results distinguished by TFs found to bind VGLL1 in multiple tumor cell lines (black), or in individual cell lines: PANC10.05 (red), BT20 (green) and Bewo (blue). High ChIP-Atlas enrichment scores were also observed for multiple TFs associated with pancreas and placenta, though the database entries for these tissue types were significantly less (Figures 2D, E). Overall, a significant degree of overlap was observed in the TF profiles from the three tumor cell lines found in our HOMER motif analysis and those identified in tissue-specific ChIP-seq database. By contrast, low enrichment scores were seen in the breast tissue-derived ChIP-seq database samples for TFs found only in the PANC10.05 cell line (for example, MAFB and ERG, Figure 2C).

Comparing the eight TFs identified to interact with VGLL1 in all three tumor cell lines, we assessed the similarity in target loci found in the different cancer cells with publicly available ChIP-seq data for the same TFs (Figure 2F). This comparison showed a very high degree of overlap in the percentage of target loci where VGLL1 bound compared to those targets previously identified for TEAD1, TEAD2, TEAD3, and TEAD4 (p < 10e-370 for all 3 tumor cell lines). For example, the VGLL1 ChIP-seq analysis of PANC10.05 cells demonstrated a 46.88% overlap with a TEAD4 ChIP-seq analysis from placental trophoblast cells and low background binding outside these targets. Similar results were observed for BT20 and Bewo cells (46.70% and 53.46% target overlap with TEAD4, respectively). The high degree of overlap observed for the 4 TEADs in all 3 tumor cell lines suggest that VGLL1 interacts with these TEADs globally and not in a tissue-specific manner (Figure 2F). Interestingly, AP-1 targets identified from ChIP-seq overlapped significantly with VGLL1 target loci in PANC10.05 and BT20 cells (48.58% and 32.43%, respectively), but not Bewo cells. By contrast, GATA6 targets showed strong overlap with VGLL1 target loci identified in Bewo cells (40.81%), but less overlap in PANC10.05 cells and BT20 cells (12.91% and 19.89%, respectively). The TFs GATA3 and AP2gamma (TFAP2C) also demonstrated significant overlap with VGLL1 target loci in BT20 and Bewo cells, but not in PANC10.05 cells (Figure 2F). These results confirm previous studies reporting that VGLL1 can cooperate with TEAD4, GATA3 and TFAP2C to initiate transcription (7, 12, 24), and expand the list of candidate TFs that can potentially interact with the VGLL1 co-activator in either a global or tissue-specific fashion.



F

		PANC10.05			BT20			Bewo		
Motif	Name	pvalue	%of Targets	% of Background	pvalue	%of Targets	% of Background	pvalue	%of Targets	% of Background
	TEAD4(TEA)/Tropoblast-Tead4-ChIP-Seq(GSE37350)/Homer	1e-437	46.88%	8.37%	1e-405	46.70%	8.61%	1e-434	53.46%	8.82%
	TEAD3(TEA)/HepG2-TEAD3-ChIP-Seq(Encode)/Homer	1e-466	54.93%	11.55%	1e-430	55.41%	12.27%	1e-413	59.52%	12.47%
	TEAD2(TEA)/Py2T-Tead2-ChIP-Seq(GSE55709)/Homer	1e-370	35.26%	5.04%	1e-337	34.81%	5.20%	1e-383	41.33%	5.24%
	TEAD1(TEAD)/HepG2-TEAD1-ChIP-Seq(Encode)/Homer	1e-470	51.68%	9.81%	1e-420	51.46%	10.50%	1e-419	56.39%	10.66%
	RUNX2(Runt)/PCa-RUNX2-ChIP-Seq(GSE33889)/Homer	1e-466	17.50%	7.52%	1.00E-57	18.16%	6.92%	1.00E-33	5.15%	0.90%
	AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	1e-538	48.58%	6.98%	1.00E-248	32.43%	6.72%	1.00E-02	8.08%	6.37%
	Gata6(Zf)/HUG1N-GATA6-ChIP-Seq(GSE51936)/Homer	1.00E-17	12.91%	7.30%	1.00E-50	19.89%	8.61%	1.00E-235	40.81%	9.31%
	GRHL2(CP2)/HBE-GRHL2-ChIP-Seq(GSE46194)/Homer	1.00E-34	8.62%	2.84%	1.00E-51	10.43%	2.88%	1.00E-21	8.34%	3.13%

FIGURE 2
VGLL1 interacts with common and distinct transcription factors to regulate transcription in tumor cells. **(A)** Western blot analysis showing relative VGLL1 protein expression either in cell lysates or immunoprecipitated from PANC10.05 and Bewo cells expressing endogenous VGLL1, or PANC1 cells transduced with empty vector (EV) or hVGLL1-MYC plasmid (VGLL1-MYC). Samples without VGLL1 Ab (Beads no Ab) confirmed the specificity of the anti-VGLL1 Ab for IP applications. **(B)** VGLL1 ChIP-seq analysis was performed on native PANC10.05, BT20, and Bewo tumor cell lines. Homor Motif analysis of the VGLL1 chromatin binding regions revealed several transcription factors (TFs) likely to interact with VGLL1. The Venn diagram shows the common VGLL1 TFs between PANC10.05 (red), BT20 (green) and Bewo (blue) tumor cells. **(C–E)** ChIP-Atlas was used to compare our results with previously published ChIP-seq data. Each dot represents a different sample (cell line, tissue, etc.) used as a source to pull down individual TFs. Higher enrichment scores indicate higher similarity to the VGLL1 ChIP-seq results. Some samples demonstrated overlap in more than one tumor cell type (black) and others overlapped only with individual tumor cell lines PANC10.05 (red), BT20 (green), Bewo (blue). **(F)** Table of TFs that were identified in all three tumor cell types. The table contains the TF name and analyzed sample tissue, chromatin motif sequence recognized, percentage of overlapping targets compared to background and p-value. The same color scheme as above indicates the different tumor cell lines analyzed.

VGLL1 regulates transcription of common genes in cancer cells derived from different tumor types

We next analyzed the ChIP-seq data from PANC10.05, BT20 and Bewo cells to explore the exact location of the VGLL1-TF complex binding to specific genes. Figure 3A shows a representative image of the ChIP-seq peaks showing the specificity of immunoprecipitated VGLL1 compared with input control cell lysates from the 3 tumor cell lines. Notably, the number of chromatin binding regions identified corresponded to the level of VGLL1 expression (Figure 3B); BT20 and Bewo cells, demonstrating the highest VGLL1 expression, showed approximately 28,000 and 35,0000 total merge peaks, respectively compared to ~9,000 total merged peaks for PANC10.05 cells, which express lower levels of VGLL1 (Figure 3B). Importantly, VGLL1 bound at 2786 specific gene loci that were common to all three cell lines (Figure 3B). Supplementary Table 1 contains a comprehensive list of genes shown to bind VGLL1, as

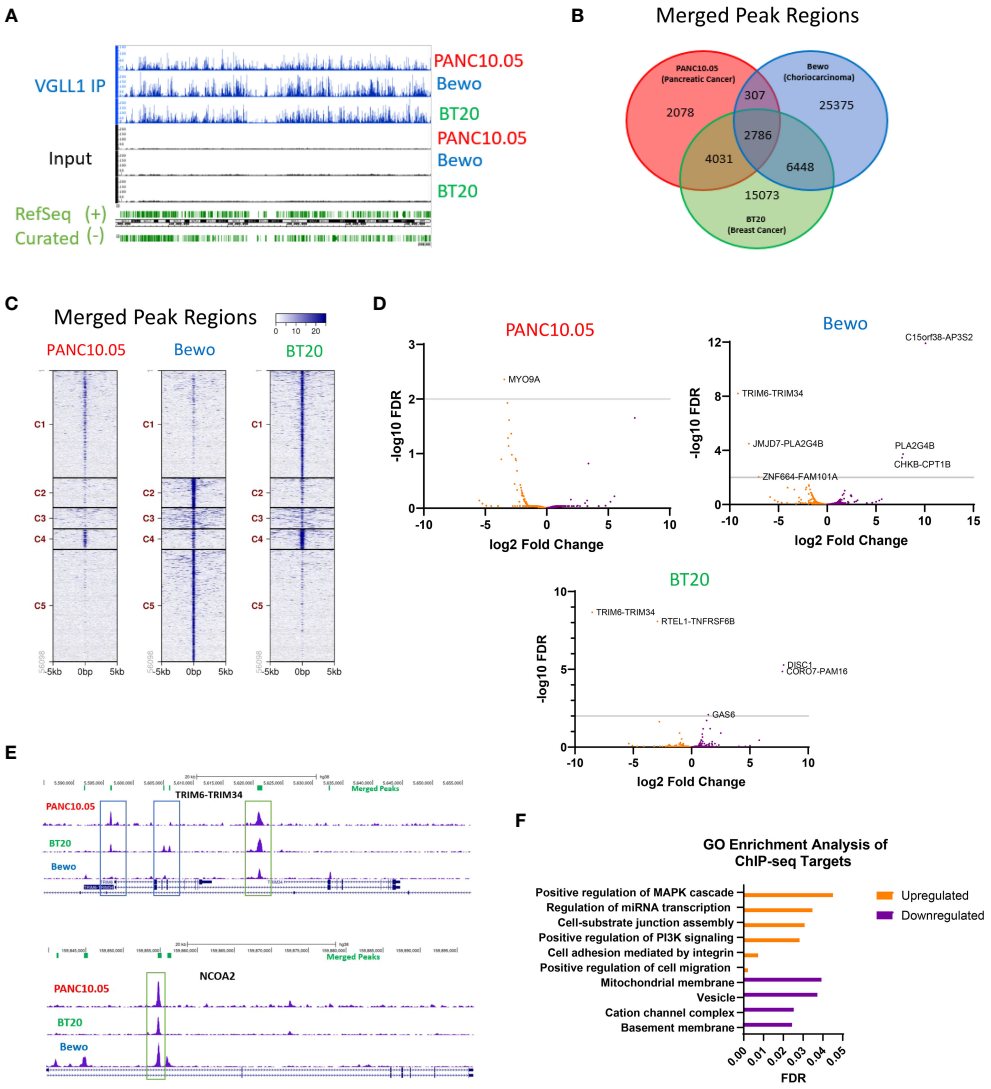


FIGURE 3
VGLL1 regulates transcription of common and unique genes in different tumor cell types. **(A)** A representative image of the ChIP-seq chromatin peaks detected for all tumor line samples comparing immunoprecipitated VGLL1 (VGLL1 IP) to input cell lysate. **(B)** Venn diagram shows the number of merged peaks detected and overlap between the 3 cell lines PANC10.05 (red), BT20 (green) and Bewo (blue). **(C)** Comparison of merged peak regions showing common and unique VGLL1 binding clusters for each tumor cell line. **(D)** RNA-seq analysis was performed on the indicated tumor cell lines treated with either siVGLL1 or siScramble. Changes in transcript expression for genes associated with VGLL1-binding regions identified in the merged peaks analysis are shown as Volcano plots for each tumor cell line, illustrating genes upregulated (orange) or downregulated (purple) by VGLL1 expression. **(E)** Representative images of the ChIP-seq chromatin peaks for each tumor cell line are shown for TRIM6-TRIM34, a read-through transcript upregulated by VGLL1 expression (top), NCOA2, a gene whose expression was upregulated by VGLL1 expression (bottom). Green boxes highlight examples of merged peaks present in all 3 tumor cell types and blue boxes indicate peaks not present in all cell lines. **(F)** GO-enrichment analysis was performed on VGLL1 ChIP-seq target genes modulated in response to VGLL1 knockdown. Upregulated pathways are shown in orange and downregulated pathways are in purple. .

determined by ChIP-seq analysis (Supplementary Table 1). Unexpectedly, most of the merged peak regions were found within the intron regions of genes rather than annotated promoter regions (Supplementary Figures 2A-C). Co-transcriptional activators such as YAP1 typically show a higher proportion of interactions at promoter regions than introns (25); however, VGLL1 interacted mainly with introns and distal intergenic regions in all 3 cell types studied (Supplementary Figure 2A). Peak-centered histograms of VGLL1 ChIP-seq peaks revealed significant overlap between all cell lines with no enrichment observed in the input samples (Supplementary Figure 2C). These results also demonstrated a significant increase in merged peak intron regions over the number of promoter region interactions (Supplementary Figure 2C). A ChIP-seq peaks heatmap of gene regions using coordinates centered around VGLL1 +/-5Kb were organized into five clusters and sorted by descending average value within each cluster for each of the 3 tumor cell lines (Figure 3C; Supplementary Figure 2B). Again, the promoter region heatmap demonstrated relatively few VGLL1 interactions (Supplementary Figure 2B). However, the merged region heatmap revealed tissue-specific clustering; as shown in Figure 3C, Cluster 4 corresponded to

the overlapping binding regions found within all 3 cell lines. By contrast, Cluster 1 corresponded to BT20-specific binding regions and Cluster 5 was enriched in Bewo-specific binding regions.

To identify transcriptional regulation events potentially impacted by VGLL1, we compared the VGLL1 ChIP-seq findings to RNA-seq data obtained from PANC10.05, BT20, and Bewo cells treated with siVGLL1 versus siScramble. Volcano plots showing differential gene transcript expression of VGLL1-bound genes identified by ChIP-seq demonstrated that many of the genes were regulated in a similar manner. For example, the TRIM6-TRIM34 fusion transcript was upregulated by VGLL1 in both BT20 and Bewo cells. (Figure 3D). Since VGLL1 expression levels impacted the number of chromatin binding locations identified, separate plots included VGLL1-binding regions identified in two of the three tumor cell lines (Supplementary Figure 2D). For each of the 3 cell lines, genes downregulated or upregulated by siVGLL1-knockdown are indicated (Supplementary Figure 2D). Supplementary Table 2 contains the complete list of genes impacted by modulation of VGLL1 expression, and a brief description of protein function. Figure 3E shows representative VGLL1 ChIP-seq signal peak tracks for two genes, TRIM6-TRIM34 and NCOA2, that were both shown to bind VGLL1 in all three tumor cell lines. Interestingly, while TRIM6-TRIM34 transcript levels were upregulated in BT20 and Bewo cells, NCOA2 expression was upregulated in PANC10.05 and Bewo cell lines (Figure 3E). Furthermore, some of the observed intragenic VGLL1-binding peaks were present in all 3 cell lines, while others were detected in only two of the three lines (Figure 3E).

To identify pathways directly regulated by VGLL1 expression, gene ontology (GO) analysis was performed on genes that were identified by ChIP-seq analysis to be bound by VGLL1 and that were either upregulated or downregulated at the transcript level following VGLL1 knockdown in at least two of the three tumor cell lines. This analysis revealed a potential role for VGLL1 in upregulating genes involved in cell migration, MAPK and PI3K signaling pathways (Figure 3F). On the other hand, genes involved in basement membrane and cation channel complex function were downregulated by VGLL1 expression. These findings were consistent with our functional results demonstrating that VGLL1 expression enhanced tumor cell invasion capacity (Figures 1F, G).

VGLL1 expression regulates transcription of genes involved in cell proliferation and invasion

To determine the overall impact of VGLL1 expression on the transcriptomic profiles of the different cancer cells, global RNA-seq data from PANC10.05, BT20, and Bewo cells treated with siScramble versus siVGLL1 was analyzed without prior filtering on the genes shown to bind VGLL1 by ChIP-seq. The results for all differentially expressed genes (DEGs) for each cell line are shown in Figure 4A. This analysis revealed several gene clusters that were either upregulated or downregulated by VGLL1 in a global or tumor cell-specific fashion. The top 20 overall DEGs for each tumor cell line are depicted in Figure 4B. In addition, volcano plots for each cell line in Figure 4C illustrate the upregulation or downregulation of specific gene transcription in response to VGLL1 expression.

To better understand these VGLL1-dependent genes, we sorted the list of DEGs by protein function, highlighting the genes that were also identified by ChIP-seq analysis to interact with VGLL1 (Supplementary Table 2). While 34 of the 128 genes (26.6%) were either uncharacterized or were predicted not to form a protein, 39 of the genes (30.5%) were known to be involved with transcription or other nuclear processes. Interestingly, 28 of the genes (21.9%) had known roles in cell proliferation and invasion, while 13 genes (10.2%) were involved with developmental processes. Moreover, gene ontology (GO) analysis performed on DEGS within each tumor cell line revealed a number of pathways upregulated in multiple cell lines, including angiogenesis, chromatin remodeling, positive regulation of transcription, in addition to cell proliferation and invasion. (Figure 4D). Furthermore, several of the genes identified as being potentially regulated by VGLL1 were not included in the GO analyses, most notably TRIM6-TRIM34, NCOA2, and ASLX2. These genes were not only highly upregulated by VGLL1 at the mRNA level but were also identified in the ChIP-seq analysis, indicating that VGLL1-TF complexes can interact with them directly. Interestingly, several studies have shown these genes to have roles in cancer cell proliferation, invasion, angiogenesis, and EMT transition (26–38). Taken together, our findings provide additional mechanistic evidence that VGLL1 plays an important role in modulating the expression of specific genes involved in driving tumor progression.

Discussion

Transcriptomic analyses of normal human tissues have shown that the co-transcriptional activator VGLL1 demonstrates uniquely elevated expression within the placenta (4, 5, 39). VGLL1 is also overexpressed in multiple cancer types, demonstrating predominant expression in basal-like breast and pancreatic cancers with transcription increased by 8- to 60-fold compared to normal breast and pancreatic tissue (3, 4, 40). Emerging evidence has suggested that tumor cells may co-opt attributes of normal placental VGLL1 function to promote invasion, proliferation, and tumor progression (3, 8, 9). In this study, we demonstrated that modulation of VGLL1 expression has a marked impact on cell invasion and proliferation in multiple cancer cell types, and that it can interact with several known TFs to drive the expression of multiple genes involved in these pathways to promote tumor growth and metastasis (41).

Under normal physiological circumstances, VGLL1 acts as a master regulator of early placenta development. Yang et al. have demonstrated that VGLL1 plays an important role in placenta formation at the trophectoderm stage (7). Specifically, VGLL1 was shown to be critical for appropriate syncytiotrophoblast (SCT) and extravillous trophoblast (EVT) development since knockdown of VGLL1 during SCT and EVT differentiation led to a decrease in cell proliferation, failure to differentiate and widespread cell death (7). This was reminiscent of the reduced proliferation and cell death we

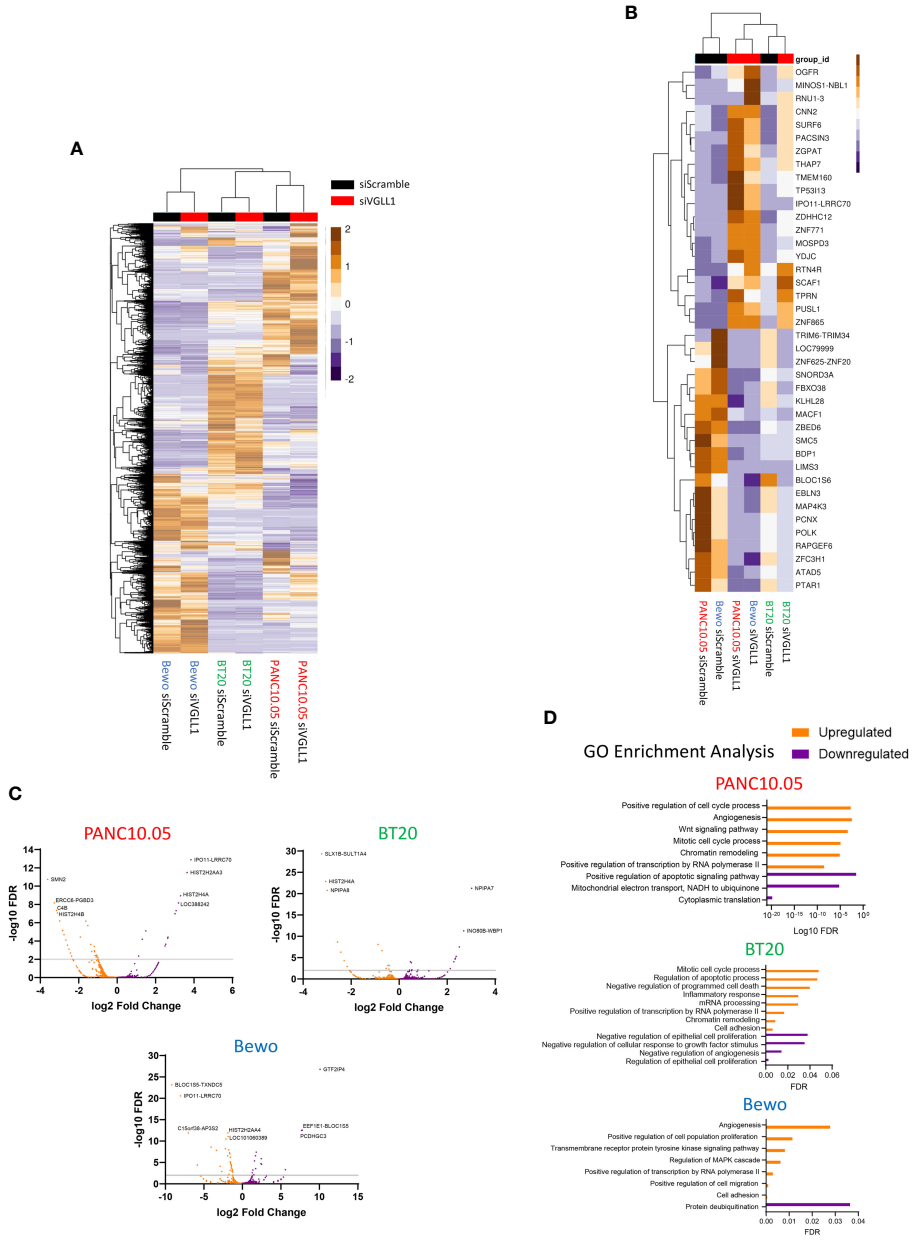


FIGURE 4
VGLL1 expression regulates transcription of genes involved in cellular proliferation and invasion. **(A)** Global heatmap showing all differentially expressed genes (DEGs) identified from RNAseq analysis of PANC10.05, BT20, and Bewo cells following treatment with either siScramble or siVGLL1. **(B)** Heatmap showing the top 20 upregulated or downregulated genes in response to VGLL1 knockdown that were common to all three tumor cell lines analyzed. **(C)** Volcano plots of DEGs identified in each tumor cell line. **(D)** Results of GO-pathway analysis using the DEGs identified for each tumor cell line. Upregulated genes and pathways are shown in orange and downregulated genes and pathways are indicated in purple.

observed in tumor cells following VGLL1 knockdown. Disruption of VGLL1 expression in the placenta can also lead to pathological conditions associated with preeclampsia and choriocarcinoma (42, 43). Choriocarcinoma is a trophoblastic malignancy that shows uncontrolled invasion and highly proliferative features unlike normal trophoblast cells, which demonstrate a controlled invasive phenotype during embryo implantation (44). In this study, we showed that VGLL1 is highly expressed in a choriocarcinoma cell line and that VGLL1 expression drove enhanced cell proliferation and invasion in these tumor cells. Further work will be required to characterize the precise mechanisms by which VGLL1 acts to promote these processes in both normal placental development and in placenta-derived tumors.

In addition to choriocarcinoma, VGLL1 has been shown to be highly overexpressed in a variety of other cancer types. Echoing its role in the placenta, VGLL1 expression has previously been linked to increased cell proliferation in basal-like breast cancer, estrogen receptor-positive breast cancer, gastric cancer, pancreatic cancer, prostate cancer and HPV-related cervical cancer (8–12, 24). In addition, VGLL1 overexpression has been shown to increase cell

invasion in gastric cancer cells (9). However, the molecular mechanisms leading to these phenotypic changes have not been well-defined. Although its interactions with TEAD4 strongly suggest a primary function for VGLL1 as a co-transcriptional activator, its downstream target genes and potential interactions with other transcription factors are still largely unknown. Using a combination of cellular assays, ChIP-seq, and RNA-seq analyses, our study confirmed and extended prior findings by comprehensively assessing the downstream impact of VGLL1 expression in pancreatic, basal-like breast and placenta-derived cancer cells. We showed in all 3 tumor cell lines that downregulation of VGLL1 expression inhibited cell proliferation and invasion capacity. In addition, this study revealed that reduced VGLL1 expression led to a decrease in the transcription of several genes associated with cell growth, invasion, proliferation, and angiogenesis. Interestingly, many of the genes regulated by VGLL1 have not yet been mapped to a specific pathway but may play a substantial role in driving cancer progression.

TRIM6-TRIM34, a read-through transcript highly upregulated in the context of VGLL1 expression, encodes a fusion protein with no currently known function. However, TRIM6 and TRIM34 individually encode ubiquitin ligases that may play roles in cancer and placenta development. TRIM6 has been reported to increase cell proliferation, invasion, metastasis, and angiogenesis in cancer (26–30). In addition, it has been linked to maintaining pluripotent embryonic stem cells in mice (45). TRIM34 has been shown to facilitate cell fusion of epithelial cells, also known as multinucleated goblet cells (46). Interestingly, multinucleated trophoblast cells are required for embryo implantation and to develop a fused extravillous trophoblastic cell shell (47). The placenta also contains syncytiotrophoblast cells made up of fused cytotrophoblasts that are responsible for the transport of resources between the mother and fetus (48). Although it is currently unknown if TRIM34 plays a role in placental cell fusion, this gene does show moderate expression within normal human placenta (5). The ASXL2 gene, also shown to be upregulated by VGLL1, is reported to regulate EMT transition during trophoblast differentiation and has been linked to the promotion of tumorigenesis and cell proliferation in a variety of cancer types (31–34). This evidence collectively supports the notion that tumor cells may indeed co-opt VGLL1 function to promote invasion, proliferation, and tumor progression; however, further investigations into the roles and contributions of individual genes are required before firm conclusions can be made in this regard.

In addition to identifying individual candidate genes under the transcriptional control of VGLL1, ChIP-seq analysis identified chromatin targets throughout the genome capable of specifically interacting with VGLL1. Motif analysis of these interacting chromatin regions led to the identification of both known and novel TFs that may cooperate with VGLL1 to initiate transcription. For example, GATA3 and AP2 gamma (TFAP2C), two TFs that were recently reported to cooperate with TEAD4 and VGLL1 to differentiate pluripotent stem cells into trophoblast stem cells (7), were also identified from our ChIP-seq analysis of breast and placental tumor cells. Furthermore, previous studies have shown that VGLL1 can directly interact with TEAD1 and TEAD4 in cancer cells (7, 10, 12, 24). Our results validated these findings and expanded the list of TFs that VGLL1 can potentially interact with to include TEAD2, TEAD3, RUNX2, GATA6, AP-1 and GRHL2. These TFs have been linked to the hippo pathway and epithelial tissue development, PI3K/AKT pathway activation, and the promotion of tumorigenesis (49–51). Notably, we found that VGLL1 bound mostly at intron regions. Enhancer regions are enriched within intron regions, especially in tissue-specific gene regulation (52–55). Enhancers are clusters of transcription factor binding sites that activate the expression of target genes (52, 53, 56). This activation can impact RNA splicing, stabilize lncRNAs, and increase the rate of transcription (53, 54, 56–58). Further studies will be required to delineate how VGLL1 cooperates with these TFs at intron regions to drive normal development in the placenta and disease progression in cancer cells.

Clinical studies have shown that higher tumor VGLL1 expression correlates with lower overall survival in pancreatic cancer, gastric cancer, basal-like breast cancer, and breast cancers that become resistant to SERD therapy (4, 8, 9, 11, 24). Furthermore, our study demonstrated the dependency of tumor cells on VGLL1 expression by showing that downregulation of VGLL1 diminished cell invasiveness, proliferation, and survival. Together, these observations provide a compelling rationale for developing a VGLL1-targeted therapy for those cancer patients whose tumors demonstrate high VGLL1 expression. Attempts have been made to develop therapeutics against other co-transcriptional activators associated with cancer development, such as WWTR1 and YAP1 (59, 60). However, the high expression of these genes in most normal human tissues can induce toxicity, limiting the utility of such approaches. By contrast, VGLL1 demonstrates very low to negligible expression in most normal tissues apart from placenta, suggesting that targeting VGLL1 in tumors may prove feasible without causing dose-limiting side effects. Although many questions remain regarding the precise roles VGLL1 plays in promoting gene transcription, emerging evidence supports that effective targeting of this co-transcriptional activator may greatly benefit many patients suffering from different types of cancer.

Data availability statement

The data that support the findings of this study are openly available in the Gene Expression Omnibus (GEO) repository at https://www.ncbi.nlm.nih.gov/geo/, reference numbers GSE267817 (ChIP-seq) and GSE267818 (RNA-seq).

Author contributions

HS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. BP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Validation, Visualization, Writing – review & editing. BN: Data curation, Investigation, Methodology,

Validation, Visualization, Writing – review & editing. YS: Data curation, Investigation, Methodology, Writing – review & editing. LE: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – review & editing. AT: Investigation, Methodology, Writing – review & editing. AK: Project administration, Writing – review & editing. KB: Conceptualization, Methodology, Supervision, Writing – review & editing. GL: Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The Advanced Technology Genomics Core was supported in part by The University of Texas MD Anderson Cancer Center and P30CA016672.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1403052/full#supplementary-material

References

1. Halder G, Polaczyk P, Kraus M, Hudson A, Kim J, Laughon A, et al. The Vestigial and Scalloped proteins act together to directly regulate wing-specific gene expression in Drosophila. *Genes Dev*. (1998) 12:3900–9. doi: 10.1101/gad.12.24.3900

2. Vaudin P, Delanoue R, Davidson I, Silber J, Zider A. TONDU (TDU), a novel human protein related to the product of vestigial (vg) gene of Drosophila melanogaster interacts with vertebrate TEF factors and substitutes for Vg function in wing formation. *Development*. (1999) 126:4807–16. doi: 10.1242/dev.126.21.4807

3. Sonnemann HM, Pazdrak B, Antunes DA, Roszik J, Lizee G. Vestigial-like 1 (VGLL1): An ancient co-transcriptional activator linking wing, placenta, and tumor development. *Biochim Biophys Acta Rev Cancer*. (2023) 1878:188892. doi: 10.1016/j.bbcan.2023.188892

4. Bradley SD, Talukder AH, Lai I, Davis R, Alvarez H, Tiriac H, et al. Vestigial-like 1 is a shared targetable cancer-placenta antigen expressed by pancreatic and basal-like breast cancers. *Nat Commun*. (2020) 11:5332. doi: 10.1038/s41467-020-19141-w

5. Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Proteomics. *Tissue-based map Hum proteome Science*. (2015) 347:1260419. doi: 10.1126/science.1260419

6. Tarrade A, Lai Kuen R, Malassiné A, Tricottet V, Blain P, Vidaud M, et al. Characterization of human villous and extravillous trophoblasts isolated from first trimester placenta. *Lab investigation; J Tech Methods pathology*. (2001) 81:1199–211. doi: 10.1038/labinvest.3780334

7. Yang Y, Jia W, Luo Z, Li Y, Liu H, Fu L, et al. VGLL1 cooperates with TEAD4 to control human trophectoderm lineage specification. *Nat Commun*. (2024) 15:583. doi: 10.1038/s41467-024-44780-8

8. Castilla MA, Lopez-Garcia MA, Atienza MR, Rosa-Rosa JM, Diaz-Martin J, Pecero ML, et al. VGLL1 expression is associated with a triple-negative basal-like phenotype in breast cancer. *Endocr Relat Cancer*. (2014) 21:587–99. doi: 10.1530/ERC-13-0485

9. Kim B-K, Cheong J-H, Im J-Y, Ban HS, Kim S-K, Kang M-J, et al. PI3K/AKT/β-catenin signaling regulates vestigial-like 1 which predicts poor prognosis and enhances Malignant phenotype in gastric cancer. *Cancers*. (2019) 11:1923. doi: 10.3390/cancers11121923

10. Mori S, Takeuchi T, Ishii Y, Kukimoto I. The transcriptional cofactor VGLL1 drives transcription of human papillomavirus early genes *via* TEAD1. *J Virol*. (2020) 94:e01945-19. doi: 10.1128/JVI.01945-19

11. Kim S, Leem G, Choi J, Koh Y, Lee S, Nam SH, et al. Integrative analysis of spatial and single-cell transcriptome data from human pancreatic cancer reveals an intermediate cancer cell population associated with poor prognosis. *Genome Med*. (2024) 16:20. doi: 10.1186/s13073-024-01287-7

12. Pobbati AV, Chan SW, Lee I, Song H, Hong W. Structural and functional similarity between the Vgll1-TEAD and the YAP-TEAD complexes. *Structure*. (2012) 20:1135–40. doi: 10.1016/j.str.2012.04.004

13. Zanconato F, Cordenonsi M, Piccolo S. YAP/TAZ at the roots of cancer. *Cancer Cell*. (2016) 29:783–803. doi: 10.1016/j.ccell.2016.05.005

14. Nin DS, Deng LW. Biology of cancer-testis antigens and their therapeutic implications in cancer. *Cells*. (2023) 12:926. doi: 10.3390/cells12060926

15. Gurchot C. The trophoblast theory of cancer (John Beard, 1857-1924) revisited. *Oncology*. (1975) 31:310–33. doi: 10.1159/000225037

16. RA W DH. Hallmarks of cancer: the next generation. *Cell*. (2011) 144:646–74. doi: 10.1016/j.cell.2011.02.013

17. Svensson-Arvelund J, Mehta R, Lindau R, Mirrasekhian E, Rodriguez-Martinez H, Berg G, et al. The human fetal placenta promotes tolerance against the semiallogeneic fetus by inducing regulatory T cells and homeostatic M2 macrophages. *J Immunol (Baltimore Md 1950)*. (2015) 194:1534–44. doi: 10.4049/jimmunol.1401536

18. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res*. (2022) 50:D20–D6. doi: 10.1093/nar/gkab1112

19. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. (2009) 25:1754–60. doi: 10.1093/bioinformatics/btp324

20. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of chIP-seq (MACS). *Genome Biol*. (2008) 9:R137. doi: 10.1186/gb-2008-9-9-r137

21. Zou Z, Ohta T, Miura F, Oki S. ChIP-Atlas 2021 update: a data-mining suite for exploring epigenomic landscapes by fully integrating ChIP-seq, ATAC-seq and Bisulfite-seq data. *Nucleic Acids Res*. (2022) 50:W175–W82. doi: 10.1093/nar/gkac199

22. Oki S, Ohta T, Shioi G, Hatanaka H, Ogasawara O, Okuda Y, et al. ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep*. (2018) 19:e46255. doi: 10.15252/embr.201846255

23. Valster A, Tran NL, Nakada M, Berens ME, Chan AY, Symons M. Cell migration and invasion assays. *Methods*. (2005) 37:208–15. doi: 10.1016/j.ymeth.2005.08.001

24. Gemma C, Singh AK, Belfiore A, Lai C-F, Periyasamy M, Abuelmaaty S, et al. VGLL1-directed TEAD activation drives endocrine therapy resistance in estrogen receptor positive breast cancer. *bioRxiv*. [Preprint] (2020). doi: 10.1101/2020.11.29.402842

25. LeBlanc L, Lee BK, Yu AC, Kim M, Kambhampati AV, Dupont SM, et al. Yap1 safeguards mouse embryonic stem cells from excessive apoptosis during differentiation. *Elife*. (2018) 7:e40167. doi: 10.7554/eLife.40167

26. Wei C, Wu J, Liu W, Lu J, Li H, Hai B. Tripartite motif-containing protein 6 facilitates growth and migration of breast cancer through degradation of STUB1. *Eur J Histochem*. (2021) 65:3214. doi: 10.4081/ejh.2021.3214

27. Liu X, Zhao J, Dong P, Du X, Lu W, Feng Y, et al. TRIM6 silencing for inhibiting growth and angiogenesis of gliomas by regulating VEGFA. *J Chem Neuroanat*. (2023) 132:102291. doi: 10.1016/j.jchemneu.2023.102291

28. Guo J, Feng S, Liu H, Chen Z, Ding C, Jin Y, et al. TRIM6: an upregulated biomarker with prognostic significance and immune correlations in gliomas. *Biomolecules*. (2023) 13:1298. doi: 10.20944/preprints202308.0544.v1

29. Zhao H, Huang J, Chen M, Li B, Chen X, Zhou M. Tripartite motif protein 6 promotes colorectal cancer cell migration and metastasis *via* SOCS2-STAT3 signaling. *Front Oncol*. (2021) 11:695525. doi: 10.3389/fonc.2021.695525

30. Zheng S, Zhou C, Wang Y, Li H, Sun Y, Shen Z. Correction: TRIM6 promotes colorectal cancer cells proliferation and response to thiostrepton by TIS21/FoxM1. *J Exp Clin Cancer Res*. (2022) 41:221. doi: 10.1186/s13046-022-02434-x

31. Perez-Garcia V, Lea G, Lopez-Jimenez P, Okkenhaug H, Burton GJ, Moffett A, et al. BAP1/ASXL complex modulation regulates epithelial-mesenchymal transition during trophoblast differentiation and invasion. *Elife*. (2021) 10:e63254. doi: 10.7554/eLife.63254

32. Wang G, Yang L, Gao J, Mu H, Song Y, Jiang X, et al. Identification of candidate biomarker ASXL2 and its predictive value in pancreatic carcinoma. *Front Oncol*. (2021) 11:736694. doi: 10.3389/fonc.2021.736694

33. Park UH, Kang MR, Kim EJ, Kwon YS, Hur W, Yoon SK, et al. ASXL2 promotes proliferation of breast cancer cells by linking ERalpha to histone methylation. *Oncogene*. (2016) 35:3742–52. doi: 10.1038/onc.2015.443

34. Cui R, Yang L, Wang Y, Zhong M, Yu M, Chen B. Elevated expression of ASXL2 is associated with poor prognosis in colorectal cancer by enhancing tumorigenesis and inducing cell proliferation. *Cancer Manag Res*. (2020) 12:10221–8. doi: 10.2147/CMAR.S266083

35. Lin Z, Yang F, Lu D, Sun W, Zhu G, Lan B. Knockdown of NCOA2 inhibits the growth and progression of gastric cancer by affecting the wnt signaling pathway-related protein expression. *Technol Cancer Res Treat*. (2020) 19:1533033820928072. doi: 10.1177/1533033820928072

36. Watson S, LaVigne CA, Xu L, Surdez D, Cyrta J, Calderon D, et al. VGLL2-NCOA2 leverages developmental programs for pediatric sarcomagenesis. *Cell Rep*. (2023) 42:112013. doi: 10.1016/j.celrep.2023.112013

37. Tanaka M, Homme M, Teramura Y, Kumegawa K, Yamazaki Y, Yamashita K, et al. HEY1-NCOA2 expression modulates chondrogenic differentiation and induces mesenchymal chondrosarcoma in mice. *JCI Insight*. (2023) 8:e160279. doi: 10.1172/jci.insight.160279

38. Xing D, Meyer CF, Gross JM, Argani P, Hung CF, Wu TC, et al. Uterine MEIS1::NCOA2 fusion sarcoma with lung metastasis: A case report and review of the literature. *Int J Gynecol Pathol*. (2024) 43:47–55. doi: 10.1097/PGP.0000000000000951

39. Consortium GT. The genotype-tissue expression (GTEx) project. *Nat Genet*. (2013) 45:580–5. doi: 10.1038/ng.2653

40. Tomczak K, Czerwinska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. (2015) 19:A68–77. doi: 10.5114/wo.2014.47136

41. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discovery*. (2022) 12:31–46. doi: 10.1158/2159-8290.CD-21-1059

42. Turco MY, Moffett A. Development of the human placenta. *Development*. (2019) 146:dev163428. doi: 10.1242/dev.163428

43. Enquobahrie DA, Meller M, Rice K, Psaty BM, Siscovick DS, Williams MA. Differential placental gene expression in preeclampsia. *Am J Obstet Gynecol*. (2008) 199:566 e1–11. doi: 10.1016/j.ajog.2008.04.020

44. Baergen RN. Gestational choriocarcinoma. *Gen Diagn Pathol*. (1997) 143:127–41.

45. Sato T, Okumura F, Ariga T, Hatakeyama S. TRIM6 interacts with Myc and maintains the pluripotency of mouse embryonic stem cells. *J Cell Sci*. (2012) 125:1544–55. doi: 10.1242/jcs.095273

46. Sun D, An X, Ji B. TRIM34 facilitates the formation of multinucleated giant cells by enhancing cell fusion and phagocytosis in epithelial cells. *Exp Cell Res*. (2019) 384:111594. doi: 10.1016/j.yexcr.2019.111594

47. Stanek J, Biesiada J. Sensitivity and specificity of finding of multinucleate trophoblastic giant cells in decidua in placentas from high-risk pregnancies. *Hum Pathol*. (2012) 43:261–8. doi: 10.1016/j.humpath.2011.03.012

48. Goldman-Wohl D, Yagel S. United we stand not dividing: the syncytiotrophoblast and cell senescence. *Placenta*. (2014) 35:341–4. doi: 10.1016/j.placenta.2014.03.012

49. Wang Z, Coban B, Wu H, Chouaref J, Daxinger L, Paulsen MT, et al. GRHL2-controlled gene expression networks in luminal breast cancer. *Cell Commun Signal*. (2023) 21:15. doi: 10.1186/s12964-022-01029-5

50. Cohen-Solal KA, Boregowda RK, Lasfar A. RUNX2 and the PI3K/AKT axis reciprocal activation as a driving force for tumor progression. *Mol Cancer*. (2015) 14:137. doi: 10.1186/s12943-015-0404-3

51. Deng X, Jiang P, Chen J, Li J, Li D, He Y, et al. GATA6 promotes epithelial-mesenchymal transition and metastasis through MUC1/beta-catenin pathway in cholangiocarcinoma. *Cell Death Dis*. (2020) 11:860. doi: 10.1038/s41419-020-03070-z

52. Ko JY, Oh S, Yoo KH. Functional enhancers as master regulators of tissue-specific gene regulation and cancer development. *Mol Cells*. (2017) 40:169–77. doi: 10.14348/molcells.2017.0033

53. Rose AB. Introns as gene regulators: A brick on the accelerator. *Front Genet*. (2018) 9:672. doi: 10.3389/fgene.2018.00672

54. Shaul O. How introns enhance gene expression. *Int J Biochem Cell Biol*. (2017) 91:145–55. doi: 10.1016/j.biocel.2017.06.016

55. Borsari B, Villegas-Miron P, Perez-Lluch S, Turpin I, Laayouni H, Segarra-Casas A, et al. Enhancers with tissue-specific activity are enriched in intronic regions. *Genome Res*. (2021) 31:1325–36. doi: 10.1101/gr.270371.120

56. Panigrahi A, O'Malley BW. Mechanisms of enhancer action: the known and the unknown. *Genome Biol*. (2021) 22:108. doi: 10.1186/s13059-021-02322-1

57. Wang Y, Ma M, Xiao X, Wang Z. Intronic splicing enhancers, cognate splicing factors and context-dependent regulation rules. *Nat Struct Mol Biol*. (2012) 19:1044–52. doi: 10.1038/nsmb.2377

58. Ntini E, Marsico A. Functional impacts of non-coding RNA processing on enhancer activity and target gene expression. *J Mol Cell Biol*. (2019) 11:868–79. doi: 10.1093/jmcb/mjz047

59. Kakiuchi-Kiyota S, Schutten MM, Zhong Y, Crawford JJ, Dey A. Safety considerations in the development of hippo pathway inhibitors in cancers. *Front Cell Dev Biol*. (2019) 7. doi: 10.3389/fcell.2019.00156

60. Shibata M, Ham K, Hoque MO. A time for YAP1: Tumorigenesis, immunosuppression and targeted therapy. *Int J Cancer*. (2018) 143(9):2133–44. doi: 10.1002/ijc.31561

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 12 June 2024
DOI 10.3389/fimmu.2024.1371829

OPEN ACCESS

EDITED BY
Manoj K. Pandey,
Cooper Medical School of Rowan University, United States

REVIEWED BY
Yanlong Shi,
Nanjing Medical University, China
Apurva Patel,
Gujarat Cancer & Research Institute, India

*CORRESPONDENCE
Xundi Xu
xuxundi@csu.edu.cn
Chunmeng Wang
cmwang1975@163.com

RECEIVED 17 January 2024
ACCEPTED 31 May 2024
PUBLISHED 12 June 2024

CITATION
Liu Z, Yang L, Liu C, Wang Z, Xu W, Lu J, Wang C and Xu X (2024) Identification and validation of immune-related gene signature models for predicting prognosis and immunotherapy response in hepatocellular carcinoma.
Front. Immunol. 15:1371829.
doi: 10.3389/fimmu.2024.1371829

Identification and validation of immune-related gene signature models for predicting prognosis and immunotherapy response in hepatocellular carcinoma

Zhiqiang Liu[1], Lingge Yang[2,3], Chun Liu[1], Zicheng Wang[1], Wendi Xu[1], Jueliang Lu[1], Chunmeng Wang[2,3]* and Xundi Xu[1,4]*

[1]Department of General Surgery, The Second Xiangya Hospital of Central South University, Changsha, China, [2]Department of Musculoskeletal Oncology, Fudan University Shanghai Cancer Center, Shanghai, China, [3]Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China, [4]Department of General Surgery, South China Hospital of Shenzhen University, Shenzhen, China

Background: This study seeks to enhance the accuracy and efficiency of clinical diagnosis and therapeutic decision-making in hepatocellular carcinoma (HCC), as well as to optimize the assessment of immunotherapy response.

Methods: A training set comprising 305 HCC cases was obtained from The Cancer Genome Atlas (TCGA) database. Initially, a screening process was undertaken to identify prognostically significant immune-related genes (IRGs), followed by the application of logistic regression and least absolute shrinkage and selection operator (LASSO) regression methods for gene modeling. Subsequently, the final model was constructed using support vector machines-recursive feature elimination (SVM-RFE). Following model evaluation, quantitative polymerase chain reaction (qPCR) was employed to examine the gene expression profiles in tissue samples obtained from our cohort of 54 patients with HCC and an independent cohort of 231 patients, and the prognostic relevance of the model was substantiated. Thereafter, the association of the model with the immune responses was examined, and its predictive value regarding the efficacy of immunotherapy was corroborated through studies involving three cohorts undergoing immunotherapy. Finally, the study uncovered the potential mechanism by which the model contributed to prognosticating HCC outcomes and assessing immunotherapy effectiveness.

Results: SVM-RFE modeling was applied to develop an OS prognostic model based on six IRGs (CMTM7, HDAC1, HRAS, PSMD1, RAET1E, and TXLNA). The performance of the model was assessed by AUC values on the ROC curves, resulting in values of 0.83, 0.73, and 0.75 for the predictions at 1, 3, and 5 years, respectively. A marked difference in OS outcomes was noted when comparing the high-risk group (HRG) with the low-risk group (LRG), as demonstrated in both the initial training set (P <0.0001) and the subsequent validation cohort (P <0.0001). Additionally, the SVMRS in the HRG demonstrated a notable positive correlation with key immune checkpoint genes (CTLA-4, PD-1, and PD-L1). The results obtained from the examination of three cohorts undergoing

immunotherapy affirmed the potential capability of this model in predicting immunotherapy effectiveness.

Conclusions: The HCC predictive model developed in this study, comprising six genes, demonstrates a robust capability to predict the OS of patients with HCC and immunotherapy effectiveness in tumor management.

KEYWORDS

hepatocellular carcinoma, machine learning, predictive model, immunotherapy efficacy, immune checkpoint inhibitors

Introduction

Hepatocellular carcinoma (HCC), a predominant subtype of primary liver cancer, constitutes approximately 75–85% of all such cases and ranks as one of the most prevalent malignancies affecting the digestive system on a global scale (1). Data from CLOBOCAN 2020 reveal that the global annual incidence of new liver cancer cases per year has increased to 905,677 (1). Correspondingly, it is ranked the sixth most common form of malignancy. With a death toll of 830,1.8 million, it is also ranked third in terms of mortality (1). In addition, a persistent increase is projected in the number of new liver cancer cases, potentially establishing it as the foremost cause of cancer-related deaths in many countries (2). However, patients with early-stage HCC can improve their chances of survival by undergoing radical treatment. Nonetheless, it is important to note that even with this aggressive approach, the 5-year recurrence rate remains as high as 70% (3). Hence, facilitating the identification of patients at risk of unfavorable clinical outcomes is essential for the prompt detection of recurrence and metastasis in HCC. Such progress is key for timely treatment and mitigation of the disease burden on patients with HCC.

At present, immune checkpoint inhibitors (ICIs) are widely used to treat patients with advanced HCC. In this context, two extensive clinical trials have demonstrated that combining ICIs and VEGFR-targeted drugs is superior to sorafenib in treating patients with advanced HCC (4, 5). Among these studies, the IMbrave150 trial highlighted that administering a combination therapy of atezolizumab (PD-L1 inhibitor) and bevacizumab outperformed sorafenib in augmenting overall survival (OS) and progression-free survival (PFS) outcomes in individuals with advanced, unresectable HCC when utilized as the first-line treatment (4). The ORIENT-32 trial, which assessed the effectiveness of combining Sintilimab (PD-1 inhibitor) with IBI305 (bevacizumab biosimilar) in treating unresectable HBV-associated HCC in a Chinese patient cohort, revealed notable improvements in both OS and PFS compared to the sole administration of sorafenib as a first-line treatment (5). Thus, the pivotal challenge lies in pinpointing biomarkers capable of accurately predicting the response of patients with HCC to ICIs. Such identification is essential to protect patients from unbeneficial therapies and to diminish both healthcare and financial burdens (6).

Our study initially screened immune-related genes (IRGs) associated with patient prognosis, followed by the application of various machine learning techniques for modeling. Once the optimal model was selected, its clinical application value was validated using our cohort and multiple data sets from various sources. Thus, the primary objective was to establish a model that can be utilized for prognostic evaluation to precisely diagnose HCC, while also predicting adverse clinical outcomes and offering timely intervention. Furthermore, the model demonstrated a notable potential in predicting the efficacy of immunotherapy, with an objective to categorize the immunotherapeutic responses in patients with HCC and to aid in the optimization of clinical pharmacotherapy.

Materials and methods

Data acquisition and processing

The TCGAbiolinks package (version: 2.28.4) (7) in R (version: 4.2.2) software was utilized to retrieve liver hepatocellular carcinoma (LIHC) patient data, including clinical information and gene expression spectrum matrix, from The Cancer Genome Atlas Program (TCGA) database (https://portal.gdc.cancer.gov/) on November 15, 2022. The training set consisted of 305 samples with HCC, PFS, and OS data for a period of at least 30 days. These samples had complete clinical stage information, prognostic follow-up information, and expression profile matrix. The transcript abundance data measured in transcripts per kilobase (TPM) and the gene symbol table associated with the ENSID were acquired separately. When there were multiple matches between the gene symbol and ENSID, the median value of expression data was selected. Additionally, genes with no expression (TPM = 0) in more than half of patients with LIHC were excluded.

In addition, the DNA methylation- and RNA-based stemness score (RNAss), was obtained from the UCSC Xena database

(https://xenabrowser.net/datapages) (8) on December 11, 2022. This data was used for the subsequent analysis of tumor stemness correlation in the training sets.

The validation datasets of this study consisted of 54 patients who were diagnosed and treated for stage I-IV HCC between January 1, 2015, and December 31, 2016. These patients were identified based on strict admission criteria using pathology. The primary inclusion criteria were as follows: 1) initial diagnosis made at our hospital; 2) absence of any other malignancies; 3) availability of complete follow-up data. OS was delineated as the duration from the initial identification of HCC in a patient to the point of either their death or the conclusion of the follow-up period. The follow-up concluded on December 31, 2021, incorporating outpatient consultations and telephone-based follow-ups. The present research received approval from the Clinical Research Ethics Committee of the Second Xiangya Hospital, Central South University (Approval No. LYF2022070). Prior to hospitalization, all patients had provided their explicit consent by signing informed consent forms. Another independent ICGC–LIRI–JP HCC cohort data were extracted from 231 patients with HCC and their corresponding clinical information in the International Cancer Genome Consortium (ICGC) database (https://dcc.icgc.org/).

Additional three validation datasets containing gene expression and clinical information were obtained from TIGER (http://tiger.canceromics.org) (9) on March 20, 2023. These datasets, which are independent of the main dataset, include three cohorts: Melanoma-phs000452 (10), non-small cell lung cancer (NSCLC)-GSE135222 (11), and renal cell carcinoma (RCC)-Braun_2020 (12). Specifically, the Melanoma-phs000452 cohort involved 153 patients receiving an anti-PD-1 drug, the NSCLC-GSE135222 cohort had 27 patients undergoing treatment with an anti-PD-1 drug, and the RCC-Braun_2020 cohort consisted of 311 patients who received treatment with a combination of anti-PD-1 and Everolimus drugs. The IRGs were singled out from the ImmPort database (accessible at https://www.immport.org/shared/genelists) (13) on December 20, 2022. A total of 1793 IRG were identified after eliminating duplicate genes.

Univariate survival analysis

To pinpoint the genes implicated in the OS and DFS among patients in the training set, a univariate survival analysis was carried out using the Survival package (version 3.3–1) in the R 4.2.2 software environment. Next, the two gene sets mentioned above were compared with IRG. After identifying the overlapping genes, the expression matrix of these genes was extracted from the verification set for further modeling. The survival curve was generated utilizing the ggsurvplot function from the survminer package (version: 0.4.9). The optimal threshold for gene classification was established using the surv_cutpoint function, which facilitated classifying genes into high- and low-expression groups based on this threshold. Subsequently, the hazard ratio (HR), along with its 95% confidence interval (CI) and P-values for the genes incorporated in the model, were graphically depicted using the forestploter package (version 1.0.0).

Logistic and least absolute shrinkage and selection operator regression analysis

In medical research, logistic regression stands out as a crucial statistical method for analyzing the intricate relationship between diseases and their pathogenic factors, providing valuable insights into the underlying mechanisms of disease development (14). LASSO regression, on the other hand, offers the advantage of flexibility in handling both continuous and discrete dependent variables with minimal data requirements, making it widely applicable, while also facilitating variable selection and reducing model complexity (15).

Following the processing of the gene expression profile and patient survival data, Logistic and LASSO regression models were built using the glmnet package (version: 4.1–6) (16, 17). Among them, logistic regression was employed to model the survival status of patients, using it as the dependent variable. The regression analysis involved the use of glm and step functions, with the direction set as "both". Finally, variables with a significance level of P <0.05 were selected for the construction of the logistic regression model. In the LASSO regression analysis, the family was set to "binomial", alpha was set to "1", nfolds was set to "10", and lambda.1se was selected based on the coef function the acquisition of the regression coefficients of each gene. Subsequently, the LASSO regression model was then constructed based on this selection. The risk scores of the two aforementioned models were obtained using the predict function and designated as the logistic regression risk score (LRRS) and the LASSO-associated risk score (LARS), respectively.

Support vector machine-recursive feature elimination analysis and modeling

SVM-RFE, as an embedded method widely utilized in pattern recognition and machine learning, demonstrates its practical value by effectively employing structural risk minimization to enhance learning performance, utilizing a sequential backward selection algorithm to iteratively refine feature sets, and ultimately enabling the construction of prognostic models and analysis of immunotumor microenvironment correlations through targeted gene screening (18, 19). The significant variables identified in the logistic regression analysis were combined with the variables used in constructing the LASSO regression model. The resulting Venn diagram was generated using the Vennerable package (version 3.0). After integrating the above-mentioned variables with the patient survival information, the caret package (version: 6.0–94) (20) was used for the Recursive Feature Elimination (RFE); herein, the function was set to "caretFuncs", the method was set to "cv", and the number was set to "10". After filtering out the optimal factors for modeling, SVM modeling was conducted using the e1071 package (version 1.7–14), the type was specified as "C-classification", and the kernel was set to "radial". The decision values of this model were utilized as risk scores and designated as support vector machine risk score (SVMRS).

Model evaluation

Cutoff values for risk scores across the three models were established through the surv_cutpoint function in the survminer package. Based on these values, patients were categorized into high-risk group (HRG) and low-risk group (LRG). The risk scores and groupings from the three models were then combined with the patient data for further evaluation of the models.

To assess the differentiation of the aforementioned models, we utilize the cindex function from the pec package (version: 2023.04.12) (21) for both evaluation and visualization purposes. Subsequently, decision curve analysis (DCA) was conducted utilizing the rmda package (version 1.6) to ascertain the clinical net benefit derived from the three models. The predictive performance of each mode was examined by computing the area under the curve (AUC) values at three different time intervals: 1 year, 3 years, and 5 years. For these calculations, the timeROC package (version: 0.4) (22) was utilized, and the results were visually represented through the receiver operating characteristic (ROC) curves. In addition, the confusion matrices were examined and graphed using the yardstick package (version 1.2.0). To assess the model's ability to accurately recall patients who experienced a fatal clinical outcome, we utilized modEvA (version: 3.9.3) (23) to generate precision-recall curves (PRC) and calculate their AUC values. Finally, to evaluate the level of calibration of the models, the calibration curves and nomogram diagrams were drawn using the calibrate functions and nomogram functions of the rms package (version: 6.5.0). The evaluation and comparison of all the aforementioned differentiations and calibrations were consolidated to establish the ultimate prognostic model suitable for the study cohort. The risk factor correlation diagrams, ROC curves, and survival curves were generated based on the risk score, patient survival, survival time, and gene expression of each model.

Correlation analysis between SVMRS or IRGs and clinical pathological features of TCGA-LIHC

From the clinical information of TCGA-LIHC patients, a representative set of features, including stage, comorbidities, and Eastern cooperative oncology group (ECOG) performance scores, were selected. We then analyzed whether these clinical pathological features exhibited significant differences and correlations between patient groups based on SVMRS values and the TPM expression levels of the six IRGs utilized in the construction of models.

Real-time quantitative polymerase chain reaction and validation of the prognostic value of models

For RT-qPCR analysis, total RNA from 54 liver tissue samples (from patients in our cohort) embedded in paraffin was isolated utilizing the BIOG RNA FFPE Tissue Kit in accordance with the guidelines specified by Baidai (Changzhou, China). The synthesis of cDNA was accomplished utilizing the Evo M-MLV RT Mix kit complemented with gDNA Clean (Accurate Biotechnology, Hunan, China). To ascertain the SVMRS, the detection of the expression of the genes to be tested was conducted through qPCR utilizing the SYBR® Green Pro Taq HS qPCR KIT (Accurate Biotechnology, Hunan, China). The gene expression levels were standardized using the 18S rRNA as a reference. The primers and their corresponding sequences are documented in Supplementary Table S1. We utilized the SVMRS (Support Vector Machine Regression Score) of each patient, along with their prognosis and survival information, as well as gene relative expressions in our cohort. Subsequently, the risk factor correlation diagrams, ROC curves, and survival curves were generated to validate the prognostic significance of the model. The aforementioned analytical approaches were also utilized in the independent ICGC–LIRI–JP HCC cohort to validate the prognostic predictive value of the model.

Tumor stemness and immune cell infiltration analysis

The data from the TCGA database exhibited a positive association between the stemness score of HCC and unfavorable clinical outcomes in patients. This finding implies a notable correlation of the tumor stemness score with the OS and PFS in the context of TCGA-LIHC (24). Consequently, we examined the disparities in six tumor stemness scores between the HRG and LRG. This comparison was done as per the tumor stemness scores derived from 305 samples and the corresponding risk groups of patients in the training set. Moreover, a prominent association was identified between the stemness scores and the tumor immune microenvironment (TIME) (24). Consequently, the distribution of 22 different types of immune cells in the training set was analyzed using the CIBERSORT package (version: 0.1.0) (25). Subsequently, we examined the variations in immune cell types between groups based on the HRG and LRG of patients.

Analysis and validation of immunotherapy efficacy prediction

The genes examined in this study were IRG, which may possess specific prognostic significance for immunotherapy effectiveness. To substantiate this hypothesis, we initially examined the variations in expression levels of four frequently utilized immunotherapy drug targets: CTLA-4, PDCD1 (PD-1), CD274 (PD-L1), and PDCD1LG2 (PD-L2), between HRG and LRG. Subsequently, the expression correlations between SVMRS and the aforementioned four genes were examined based on the classification into HRG and LRG. This analysis aimed to make an initial assessment of the potential immune prediction value.

Subsequently, we employed three distinct validation cohorts that had undergone immunotherapy. These cohorts were then classified into HRG and LRG using SVM-RFE modeling following

the same methodology. The study then focused on examining the differences in SVMRS between the HRG and LRG, with an emphasis on evaluating the immunotherapy responses in the validation cohorts. The predictive performance of the model for the immunotherapy responses was further verified by conducting survival analyses in conjunction with the prognostic information of patients.

Functional enrichment and pathway analysis

To delve into the mechanistic underpinnings of differentiating the HRG from the LRG, we initially analyzed the differentially expressed genes (DEG) using the limma package (version: 3.56.2) (26). This investigation was carried out with a fold change threshold of "2" and a false discovery rate (FDR) of "0.05". The list of DEGs was used to conduct Gene Ontology (GO) enrichment analysis through the application of the clusterProfiler package (version: 4.8.3) (27, 28). The data were graphically depicted using the GOplot package (version 1.0.2) (29). Subsequently, the gene list and fold change value were utilized to conduct gene-set enrichment analysis based on the Kyoto Encyclopedia of Genes and Genomes (GSEA-KEGG). A threshold of P <0.05 was set to ascertain the statistical significance of the results. Visualization of the GSEA results was achieved through the dotplotGsea function in the GseaVis package (version: 0.0.9) and the gseaNb function from the same package. Additionally, the cnetplot function from the enrichplot package (version: 1.20.3) was used for visualization.

The identified pathways of interest were retrieved from the PathCards database (https://pathcards.genecards.org/) (30). The expression matrix of these genes was extracted and used for expression correlation analysis with SVMRS. Each gene was analyzed individually. The ComplexHeatmap software (version 2.16.0) (31, 32) was used for visualization.

Statistical analysis

The data in this study underwent statistical analysis using GraphPad Prism software (version 9.0.0, San Diego, California, USA) for both statistical analysis and image rendering. The software package of the method utilized default parameters for the parameters that were not specified. Additionally, the ggplot2 package (version: 3.3.5) (33) was employed for data visualization, which was not explicitly mentioned. Spearman correlation analysis was used for correlation analysis. The scatterplots were generated using Sangerbox (http://www.sangerbox.com/tool) (34). Additionally, the study utilized the Mann-Whitney rank sum test for the analysis of continuous variables of skewed distribution between two groups. In contrast, when data conformed to a normal distribution with consistent variance, the Student's t-test was utilized for executing a comparative analysis between the two groups. For the comparative analysis of multiple sets of data that satisfy the assumptions of homogeneity of variance and normal distribution, Ordinary one-way ANOVA should be employed, followed by Holm-Šídák's multiple comparisons test for pairwise comparisons within the groups. However, if these assumptions are not met, the Kruskal-Wallis test should be utilized for the comparison among multiple groups, accompanied by Dunn's multiple comparisons test for within-group comparisons. For discrete variables, the Chi-square test was used for comparison between groups. A significance level of P <0.05 was used to determine statistical significance (*P <0.05, **P <0.01, ***P <0.001, ****P <0.0001).

Results

The SVM-RFE model was developed using 6 prognosis related genes

The analytical flow chart for this study is shown in Figure 1. As observed, the univariate survival analysis revealed that there were 6608 genes influencing PFS and 9772 genes influencing OS in the training set. Furthermore, 81 genes were obtained after the intersection with IRG, which were used for subsequent modeling (Supplementary Figure S1). Additionally, logistic regression analysis identified eight genes, while LASSO regression analysis screened seven genes. Among these, two genes were found to be common to both methods. Therefore, a total of 13 genes were selected for SVM-RFE modeling (Figure 2A). Following the RFE analysis, a total of six genes were selected to be included in the construction of the final model (Figure 2B). The HR, 95% CI, and P-values for these genes in the univariate analysis are shown in Figure 2C, demonstrating that all six genes were identified as risk factors. Moreover, these six genes were identified as prognostic markers and were found to have an impact on the OS of patients in TCGA-LIHC. The specific genes are as follows: CMTM7 (P <0.0001, HR = 1.05, Figure 2D), HDAC1 (P <0.0001, HR = 1.01, Figure 2E), HRAS (P <0.0001, HR = 1.01, Figure 2F), PSMD1 (P <0.0001, HR = 1.07, Figure 2G), PAET1E (P = 0.00017, HR = 6.97, Figure 2H), TXLNA (P <0.0001, HR = 1.03, Figure 2I).

SVM-RFE model was found to be the best model in this study

As per the aforementioned outcomes, the SVM-RFE model exhibits the advantage of having a limited number of constituent genes. Thus, to further verify the optimal nature of this model in our study, we conducted additional evaluations focusing on differentiation and clinical applicability. Upon analyzing the fluctuation of the C-statistic in relation to the OS time, we determined that the SVM-RFE model exhibited the highest level of effectiveness (Figure 3A). Similarly, the DCA outcomes indicated that all three models were capable of enhancing the net benefit, with the SVM-RFE model exhibiting the greatest increase in net benefit (Figure 3B). Additionally, the time-dependent ROC curve analysis

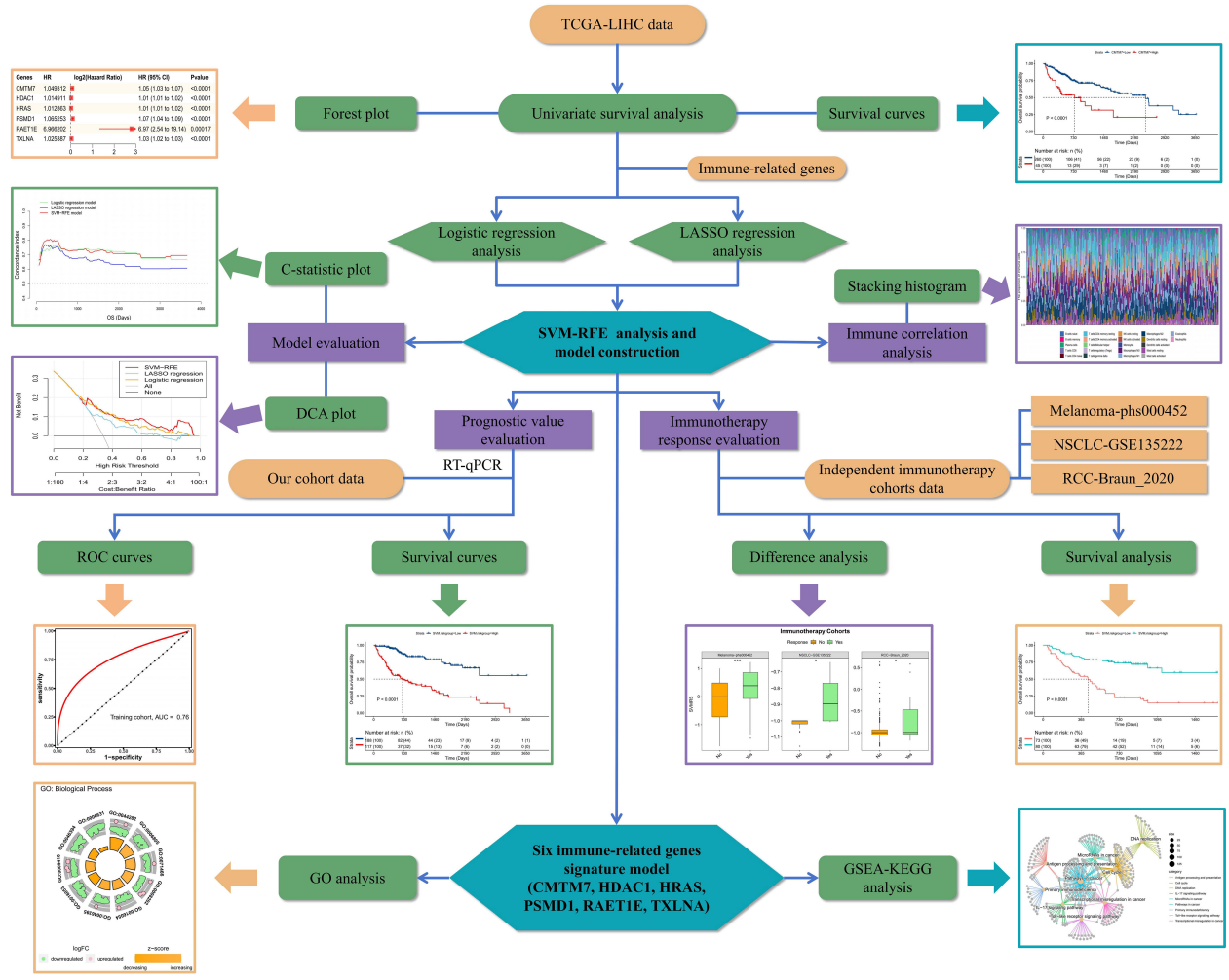


FIGURE 1
Flow chart of this study.

revealed that the SVM-RFE model achieved AUC values of 0.83, 0.73, and 0.75 for the 1-, 3-, and 5-year OS predictions, respectively (Figure 3C). In addition, the SVM-RFE model was found to outperform both the logistic regression model (Supplementary Figure S2A) and the LASSO regression model (Supplementary Figure S2B). Furthermore, the accuracy of the SVM-RFE model (75.08%) was higher (Figure 3D) than that of the logistic regression model (70.16%, Supplementary Figure S2C), as well as the LASSO regression model (69.51%, Supplementary Figure S2D). Meanwhile, we constructed the Precision-Recall Curve (PRC) to evaluate the efficacy of these three models in accurately identifying dead samples. As observed, the use of the logistic, LASSO, and SVM-RFE models improved the probability of detecting dead cells from an initial 33.77% (103/105) to 63.2% (Supplementary Figure S2E), 52.3% (Supplementary Figure S2F), and 68.9% (Figure 3E), respectively, with the SVM-RFE model having the highest precision-recall rate among the three models. Thus, the SVM-RFE model proved to be the most effective model in this study. In addition, the risk score of the model also indicated a high goodness of fit (Figures 3F, G).

The model has potential value in prognostic prediction

By calculating the Survival-associated Variable Model Risk Score (SVMRS) for each patient and integrating the survival status and gene expression values, risk factor association diagrams were generated to assess the prognostic prediction of the risk score for the 305 patients. Figure 4A demonstrates the arrangement of patients based on their risk scores, ranging from low to high. The optimal cut-off value for SVMRS (-0.9214) was employed to classify patients into HRG and LRG. The mortality rate in the HRG was significantly higher than that in the LRG, and all the genes with elevated expression levels were exclusively found in the HRG, indicating that these genes were all associated with increased risk. Additionally, the ROC curve demonstrated an AUC value of 0.76 for this model's ability to predict patient mortality in the training set (Figure 4B). Furthermore, a statistically noteworthy contrast in the OS rate was found between the HRG and LRG (Figure 4C). This result indicates that individuals with an elevated SVMRS risk score were more prone to unfavorable outcomes.

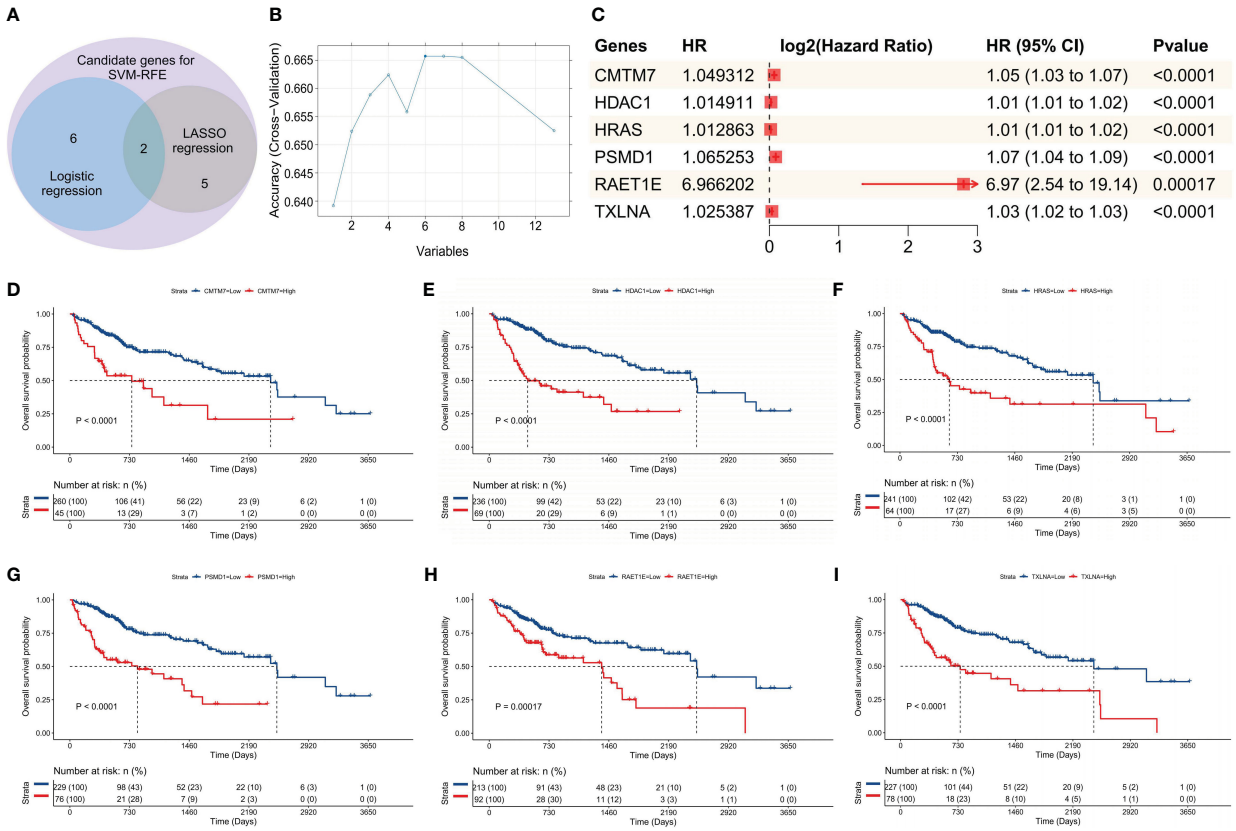


FIGURE 2
Construction of the SVM-RFE model. **(A)** Venn diagram of the genes included in the three models during the initial screening; **(B)** Line chart for the change in accuracy with the increase in variables during the analysis; **(C)** Univariate survival analysis forest plot, based on the gene symbol, HR, 95% CI, and P value; D–I. Survival curves plotted as per the optimal cut-off value for each gene group; the genes were: **(D)** CMTM7, **(E)** HDAC1, **(F)** HRAS, **(G)** PSMD1, **(H)** RAET1E, and **(I)** TXLNA, respectively.

Similarly, the prognostic performance of this model was verified using our validation cohort. As indicated by the risk factor diagram (Figure 4D), all six genes were confirmed to be risk factors. In addition, the ROC curve of our validation cohort indicated that the model accurately predicted patient mortality, as evidenced by an AUC value of 0.77 (Figure 4E). The results of survival analysis also demonstrated a significantly poorer clinical prognosis in the HRG compared to the LRG (Figure 4F). In the independent ICGC–LIRI–JP HCC cohort, we also observed consistent results with those mentioned previously. Specifically, the risk factor diagram (Supplementary Figure S3A), ROC curves (Supplementary Figures S3B, C), survival curve (Supplementary Figure S3D), and confusion matrix (Supplementary Figure S3E) all indicated that the model effectively stratified patients into risk groups and accurately predicted their OS. Thus, based on the analysis and validation conducted, it can be concluded that this model holds promise in predicting OS in patients with HCC.

SVMRS and the six IRGs were correlated with selected clinicopathologic features of HCC patients

Further clinical correlation analysis of SVMRS and the expression profiles of six IRGs utilized in model construction revealed that HRAS (Supplementary Figure S4A), PSMD1 (Supplementary Figure S4B), and SVMRS (Supplementary Figure S4C) were associated with T stage, with their values generally increasing alongside T stage progression. Additionally, PSMD1 (Supplementary Figure S4D) was found to be related to N stage, exhibiting significantly higher expression in the N1 group. HRAS (Supplementary Figure S4E), PMSD1 (Supplementary Figure S4F), TXLNA (Supplementary Figure S4G), and SVMRS (Supplementary Figure S4H) were associated with stage, where higher values corresponded to later stages. These findings further validate the prognostic predictive value of our constructed model.

Moreover, from the perspective of patients' comorbidities, RAET1E (Supplementary Figure S5A) and SVMRS (Supplementary Figure S5B) were associated with comorbidities. Specifically, TXLNA was linked to hepatitis B (Supplementary Figure S5C), SVMRS to hepatitis C (Supplementary Figure S5D), and RAET1E to non-alcoholic fatty liver disease (Supplementary Figure S5E). Regarding ECOG performance scores, the expression levels of HDAC1 (Supplementary Figure S5F), PSMD1 (Supplementary Figure S5G), and CMTM7 (Supplementary Figure S5H) correlated with them, where higher scores corresponded to increased gene expression. Similarly, SVMRS demonstrated a correlation with ECOG scores, as shown in Supplementary Figure S5I. This suggested that higher SVMRS values are associated with comorbidities and increased

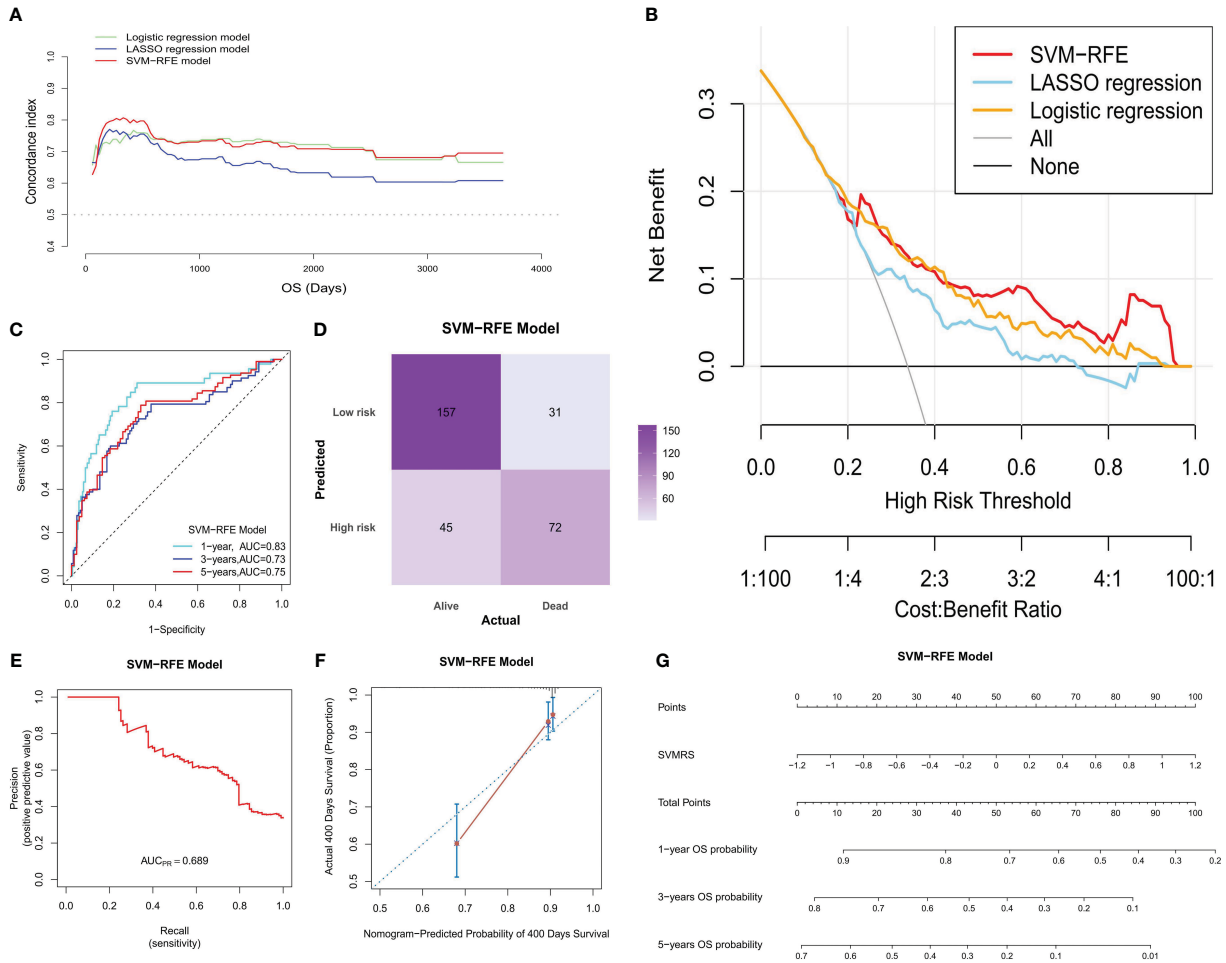


FIGURE 3
Model evaluation. In the training set: **(A)** The C-statistic (vertical coordinate) was plotted with the patient's OS (horizontal coordinate) changes; **(B)** DCA of the three models; the horizontal coordinate represents the risk threshold, and the vertical coordinate represents the net benefit. **(C)** The ROC curves were drawn based on the risk score of the SVM-RFE model and the 1-, 3- and 5-year OS time recorded. **(D)** The confusion matrix was plotted according to the classification of HRG and LRG of patients by the model, combined with the actual death status of patients; **(E)** The PRC was drawn according to the accuracy and recall rate of the model; **(F)** The calibration curve of the SVM-RFE model was plotted at 400 days (the time point comprised the best calibration degree). The horizontal coordinate denotes the predicted survival situation, and the vertical coordinate denotes the actual survival situation. Every 100 people were divided into groups and resampled 1000 times. **(G)** Nomogram developed according to the risk score, the total points, and its corresponding 1-, 3-, and 5-year OS probability.

ECOG scores, indicating a poorer prognosis for patients with HCC. The clinical correlation analysis of risk grouping exhibited high consistency with SVMRS, except for hepatitis C, where no significant statistical differences were observed. However, the differences in T stage, overall Stage, comorbidities, and ECOG scores aligned with SVMRS (Supplementary Table S2), suggesting the feasibility of the risk grouping approach employed in this study.

Differences in RNAss and immune cell infiltration between HRG and LRG

The investigation into tumor stemness unveiled a noteworthy difference in the RNAss and the epigenetically regulated RNAss (EREG.EXPss) between the HRG and LRG. Accordingly, the values of RNAss (P = 0.0035) and EREG.EXPss (P = 0.0217) were prominently augmented in the HRG compared to the LRG (Figure 5A). Given the correlation between this value and the TIME, we proceeded to perform a detailed analysis of the differences in the proportion of 22 different types of immune cell subtypes between the two groups. The distribution of immune cells in 305 samples in the training set was depicted in Figure 5B, while the comparison of proportions between the HRG and LRG was illustrated in Figure 5C. The analysis comparing HRG and LRG revealed that the LRG exhibited a higher proportion of T cell CD4 memory resting (P = 0.0113), monocytes (P = 0.0003), and mast cells resting (P = 0.0055). Moreover, the HRG exhibited a higher proportion of T cells CD4 memory activated (P = 0.0415), macrophages M0 (P <0.0001), and neutrophils (P = 0.0264) (Figure 5D).

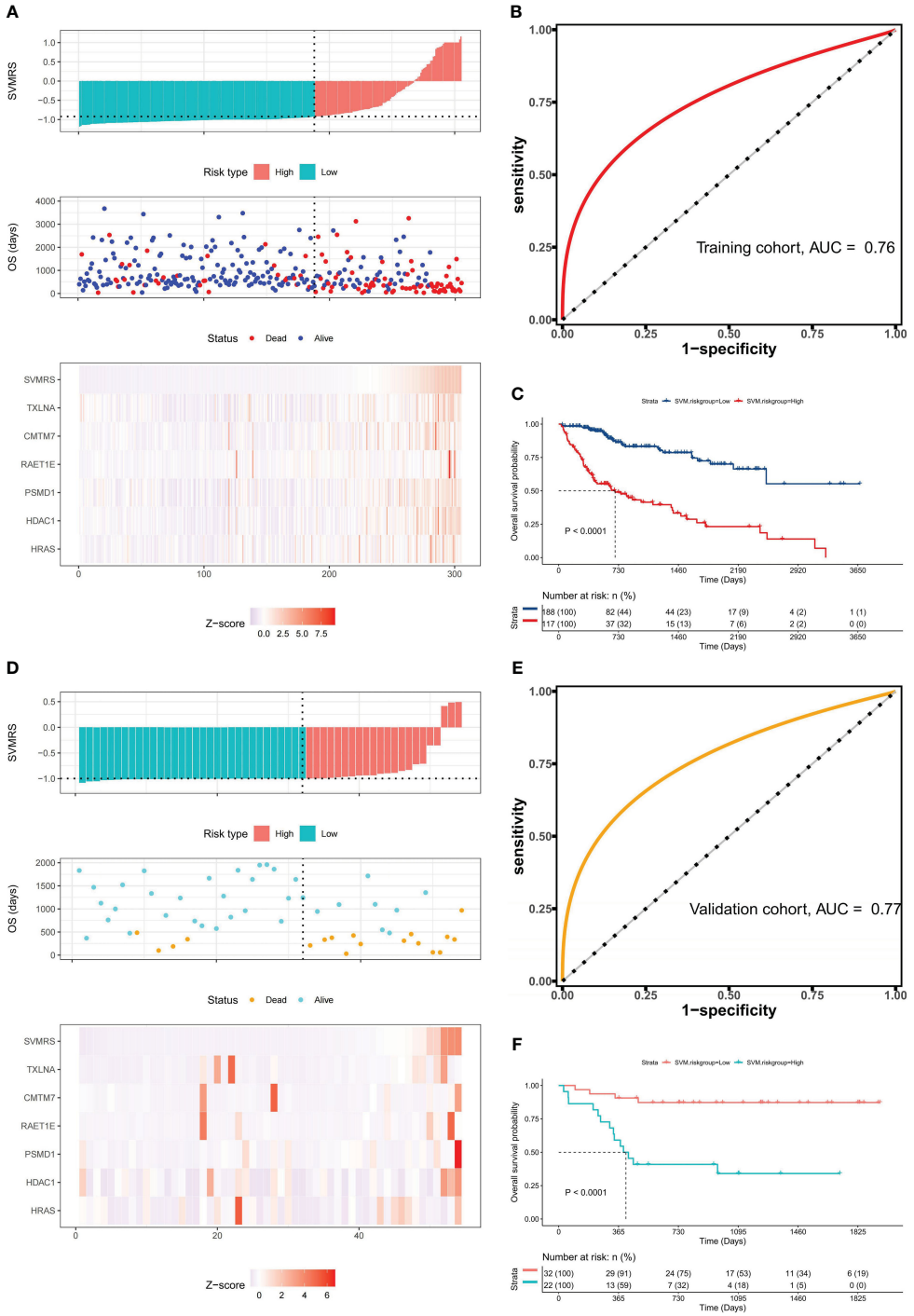


FIGURE 4
Prognostic value of the model. **(A)** Risk factor association diagram of the model in the training set includes a histogram of the high and low distribution of the patient's risk score, the scatter plot of the patient's survival situation distribution, and the heat map of the change of gene expression value with the associated risk scores. The horizontal coordinate represents the number of patients ranked by risk score from low to high; the ordinates represent risk score, OS, and model-related genes. **(B)** ROC curve drawn as per the risk score calculated by the model and the survival state of the patient in the training set; **(C)** Survival curve was drawn as per the optimal value of the risk score (SVMRS = -0.9214) in the training set. **(D)** Risk factor correlation diagram of the model in our cohort; **(E)** ROC curve based on the model's risk score and the survival status of patients in our cohort. **(F)** Survival curve drawn by the HRG and LRG as per the optimal value of the risk score (SVMRS = -0.9981) in our cohort.

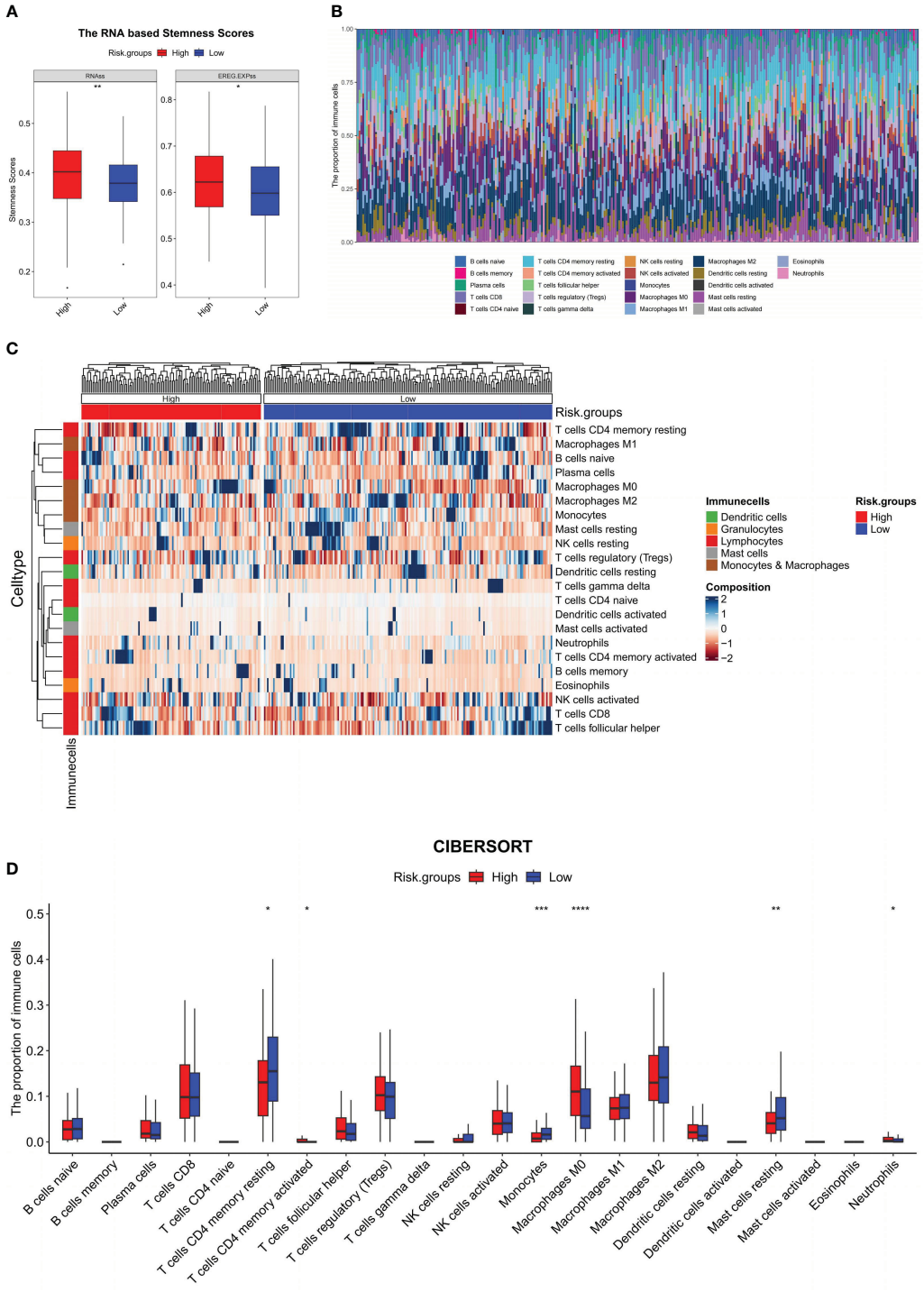


FIGURE 5
Immune microenvironment analysis of the model. In the training set: **(A)** RNA-based stemness scores were analyzed between HRG and LRG and represented using a boxplot; **(B)** Stacking histogram representing the proportion of 22 types of immune cells; **(C)** Heat map representing the proportion of 22 kinds of immune cells between the HRG and LRG; **(D)** Boxplot for the difference of 22 types of immune cells between the HRG and LRG; *P <0.05, **P <0.01, ***P <0.001, ****P <0.0001.

The model has potential value in predicting immunotherapy response

The outcomes of the analytical investigations revealed significant differences in immune checkpoint gene (ICG) expression between HRG and LRG. Specifically, the expression levels of four widely utilized immunotherapy drug targets, *viz.*, CTLA-4 (P <0.0001), PD-1 (P <0.0001), PD-L1 (P = 0.0124), and PD-L2 (P = 0.0182), were notably higher in the HRG compared to the LRG (Figure 6A). In addition, the correlation analysis indicates

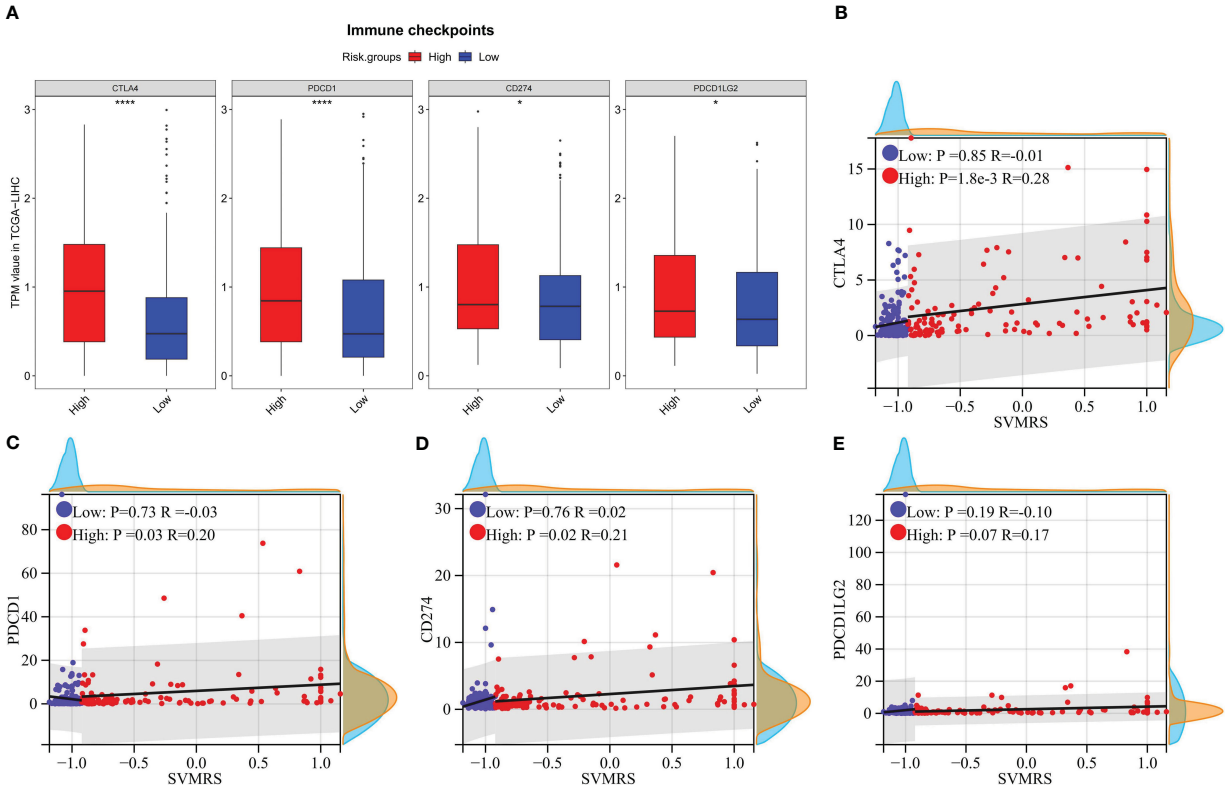


FIGURE 6
Correlation analysis between risk score and the expression of ICGs. In the training set: **(A)** The differences of 4 commonly used ICGs between HRG and LRG were analyzed and shown using a boxplot; B–E. Scatter plots for the correlation analysis between SVMRS and **(B)** CTLA-4, **(C)** PD-1, **(D)** PD-L1, and **(E)** PD-L2 in the HRG and LRG. *P <0.05, ****P <0.0001.

that CTLA-4 (P = 0.0018, R = 0.28, Figure 6B), PD-1 (P = 0.03, R = 0.20, Figure 6C), and PD-L1 (P = 0.03, R = 0.21, Figure 6D) exhibited a significant positive correlation with SVMRS in the HRG. However, no correlation was observed in the LRG. Thus, while the statistical significance of the correlation between PD-L2 and SVMRS in HRG was not established, a noticeable trend could be observed (P = 0.07, R = 0.17, Figure 6E). Hence, we hypothesized that the HRG may exhibit greater susceptibility to immunotherapy.

To verify the aforementioned hypothesis, three independent cohorts from different platforms were utilized as immunotherapy validation datasets for further analysis. All patients in the three cohorts received treatment with anti-PD-1 medications, and both the effectiveness of the drugs and the prognosis of the patients were recorded. The analysis of differences revealed that the SVMRS in the group of individuals who responded was significantly greater than that in the group of individuals who did not respond in the Melanoma-phs000452 cohort (P = 0.0004). Similar results were observed in the NSCLC-GSE135222 cohort (P = 0.0112) and the RCC-Braun_2020 cohort (P = 0.0236) (Figure 7A). Furthermore, the survival analysis demonstrated statistically significant disparities between the HRG and LRG in all three cohorts: Melanoma phs000452 (P <0.0001, Figure 7B), NSCLC-GSE135222 (P = 0.0001, Figure 7C), and RCC-Braun_2020 (P <0.0001, Figure 7D). Previous analyses have revealed that the HRG, which had a worse prognosis, experienced significantly longer survival after undergoing immunotherapy. This survival advantage was notably superior to that of the LRG, indicating that the HRG could derive substantial benefits from immunotherapy. Thus, it was verified that the model possesses the capability to predict immunotherapy effectiveness.

Signaling pathways related to tumorigenesis and immune progression were activated in the HRG

The differential expression analysis of DEGs between HRG and LRG identified a total of 341 genes. Among these DEGs, 89 exhibited upregulated expression and 252 exhibited downregulated expression in the HRG (as shown in Figure 8A; Supplementary Table S3). Further, the GO enrichment analysis of these 341 DEGs pointed to their potential roles in the regulation of the top 10 biological processes, cell components, and molecular functions (Figures 8B–D). The details, as well as the corresponding GO enrichment results, are shown in Supplementary Table S4. As observed, these genes are found to be primarily associated with tumor metabolism. Further analysis of the GSEA-KEGG pathway revealed the top 10 pathways that were either suppressed or activated. These pathways are presented in Figure 8E and are ranked based on the normalized enrichment score (NES). Additionally, it is evident that the inhibited pathways exhibited an increase in metabolic activity, whereas the stimulated pathways were associated with the development of tumors and immune processes.

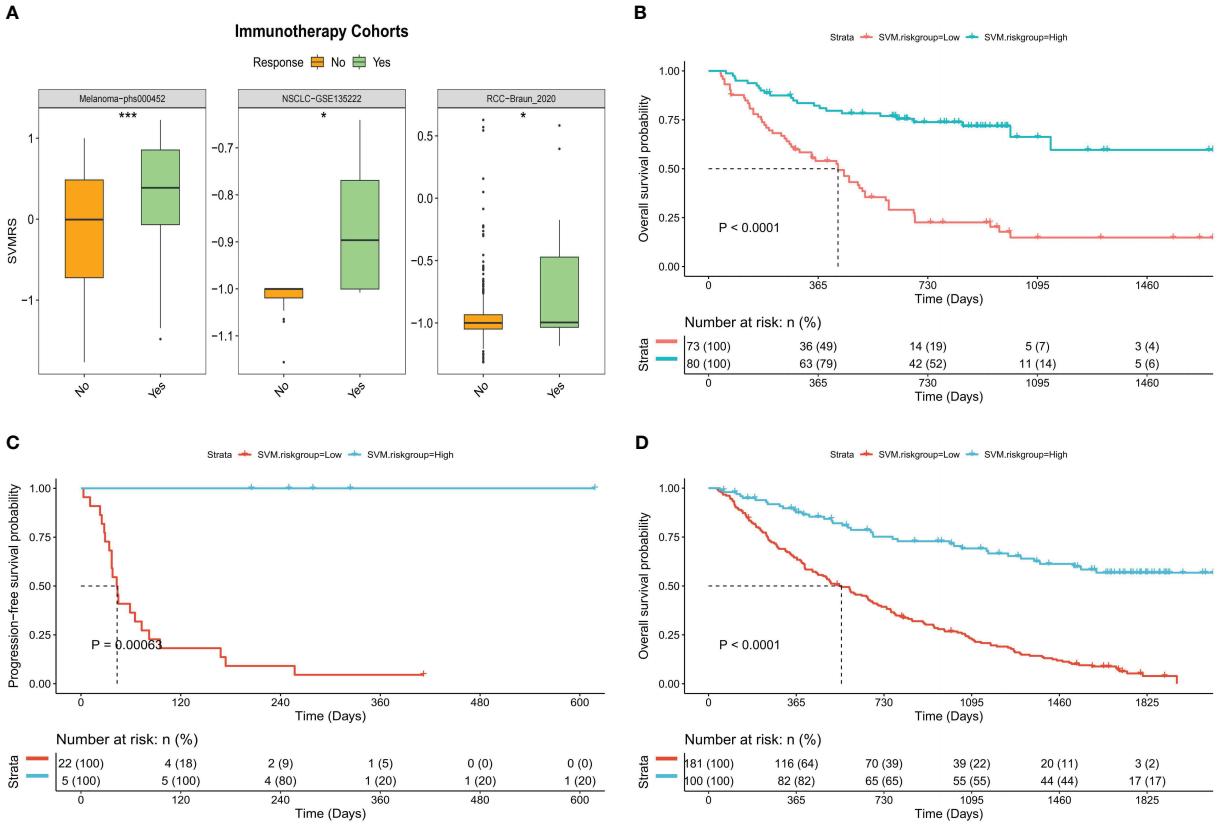


FIGURE 7
Predictive value of the model for immunotherapy response. **(A)** Boxplot of the difference of SVMRS between the response and no response group using the three immunotherapy validation datasets. **(B–D)**. Survival curve as per the patient's risk groups and the survival status of patients in three immunotherapy validation cohorts, **(B)** Melanoma-phs000452, **(C)** NSCLC-GSE135222, and **(D)** RCC-Braun_2020; *P <0.05, ***P <0.001.

Nine pathways that were activated in the HRG were selected from all the relevant pathways (as shown in Supplementary Table S5). The pathways associated with tumorigenesis comprised the cell cycle, DNA replication, and the Toll-like receptor (TLR) signaling pathway (Figure 8F). In addition, the tumor progression pathways identified were microRNAs in cancer, transcriptional dysregulation in cancer, and pathways in cancer (Figure 8G). Moreover, the immune-related pathways identified were antigen processing and presentation, primary immunodeficiency, and IL-17 signaling pathway (Figure 8H). Subsequently, visualization of the interconnection network among the aforementioned 9 exemplary pathways was performed, and a strong correlation between all 9 pathways was observed (Figure 9A). Finally, the study focused on examining the expression profiles of key genes involved in the TLR signaling pathway, which are associated with tumorigenesis and immune processes. Notably, a correlation analysis was conducted to examine the relationship between the SVMRS and the expression of the six genes constituting the model, as depicted in Figure 9B. All the key genes in the TLR signaling pathway exhibited statistically significant correlations with SVMRS or the six genes utilized in the modeling. This indicates that these key genes in the model are likely to have an immunoregulatory and cancer-promoting function by participating in the regulation of this pathway. This could also be one of the intrinsic mechanisms contributing to the unfavorable prognosis and heightened vulnerability to immunotherapy in the HRG of patients.

Discussion

In recent years, numerous studies have focused on using sequencing data to identify markers that can impact the prognosis of patients with HCC and develop corresponding models, aiming to enhance the accuracy of patient prognosis prediction and provide guidance for clinical practice (35–37). For instance, Wang et al. (35) developed four gene signature models associated with disulfidptosis to predict OS outcomes in the context of HCC. Herein, the calculated AUC values of ROC for OS at 1, 3, and 5 years in the training set were 0.766, 0.736, and 0.699, respectively, demonstrating noteworthy potential in predicting the effectiveness of anti-tumor therapy. Furthermore, Chen et al. (36) constructed a model to predict the OS of patients with HCC using five genes related to cuproptosis. In the training set, the calculated AUC values of ROC for the model were recorded at 0.775, 0.685, and 0.670 for 1, 3, and 5 years, respectively. In addition, Shi et al. (37) have successfully developed a ten epithelial-mesenchymal transition (EMT)-related genes signature prognostic model for HCC, validating its accuracy in stratifying patients into high and low-risk groups using datasets

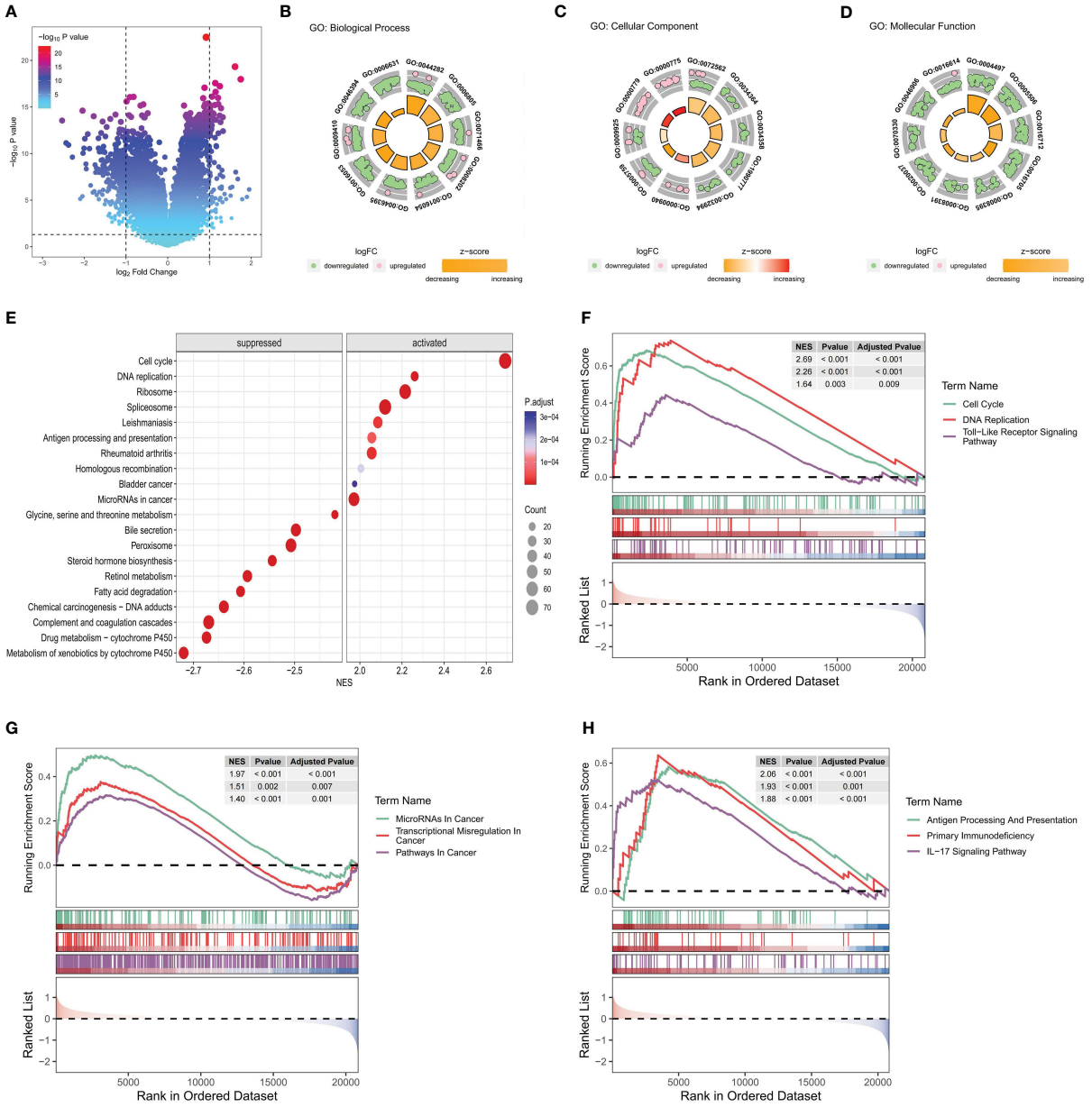


FIGURE 8
Functional enrichment analysis of key differential genes in the model. In the training set: **(A)** The volcano plot is based on risk groups for differential analysis, where the horizontal ordinate denotes the log2 Fold Change and the longitudinal coordinates denote the -log10 P value. Use |log2fold change|=1 to draw the vertical dotted line and P = 0.05 to draw the horizontal dotted line. **(B–D)** Ranked by P value, chord diagrams of the top 10 results of **(B)** biological process, **(C)** cellular component, **(D)** molecular function plotted from GO enrichment analysis of DEG. **(E)** Ordered by NES, the top 10 suppressed or activated pathways were shown according to the GSEA-KEGG pathway analysis results. **(F–H)**. Pathways associated with **(F)** tumorigenesis, **(G)** tumor progression, and **(H)** immune progression that were activated in the HRG, as selected from the GSEA-KEGG pathway analysis results.

from TCGA and International Cancer Genome Consortium (ICGC), and its risk score tightly correlates with tumor stage, grade, and immune cell infiltration, exhibiting significant prognostic value with ROC AUC values of 0.767, 0.694, and 0.680 at 1-, 2-, and 3- year OS in the training set, respectively. In our study, we developed a prognostic model based on machine learning consisting of six IRGs for predicting the survival of patients with HCC. Accordingly, the calculated AUC values of ROC for 1-, 3-, and 5-year OS were 0.83, 0.73, and 0.75, respectively, demonstrating the significance of predicting OS in patients with HCC.

All of the six key genes used to build the model were IRG, of which CMTM7, belonging to the Chemokine-like factor (CKLF)-like MARVEL transmembrane domain-containing proteins (CMTM) family, plays a crucial function in the immune system and is abundantly expressed in immune tissues (38). CMTM7 functions as a tumor suppressor in various types of cancer within the field of cancer research. For example, knockdown of CMTM7 was observed to impair the process of autophagy and accelerate the development of tumors in lung cancer (39). Moreover, CMTM7 was also found to serve as a potential biomarker for identifying

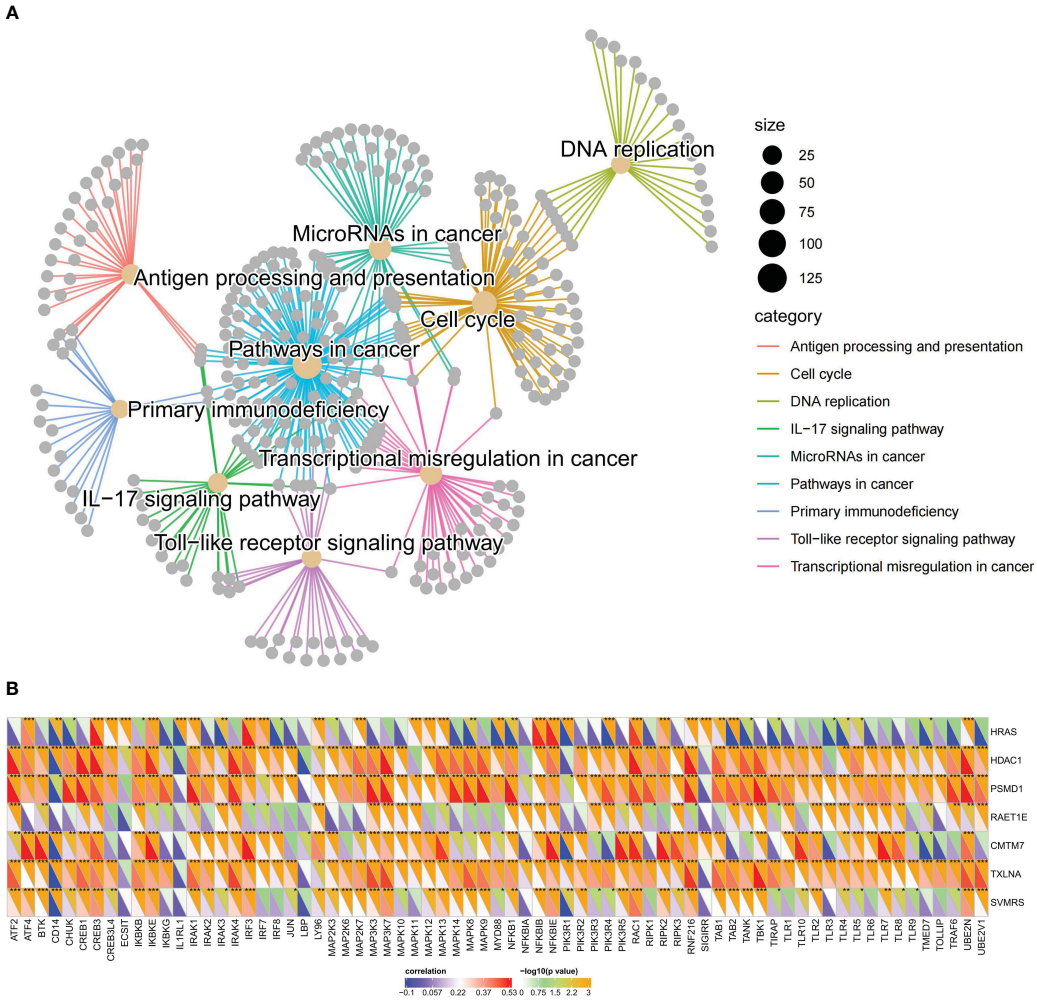


FIGURE 9
Pathways regulated by the key genes of the model. **(A)** The interaction network constructed by the nine representative pathways; each gray dot represents a gene; different pathways were represented using different colors; **(B)** SVMRS and the six genes used in modeling correlated well with the expression of key genes involved in the TLR signaling pathway. The lower left half triangle in each column represents the correlation coefficient. Blue represents a negative correlation and red represents a positive correlation; the darker the color, the stronger the correlation. The upper right half triangle represents the results of Spearman correlation analysis; *P <0.05, **P <0.01, ***P <0.001, ****P <0.0001.

immunological traits and predicting immunotherapy effectiveness in breast cancer (40). On the other hand, HDAC1 is critically involved in regulating gene expression by modulating the acetylation of both histone and non-histone proteins (41). Correspondingly, its overexpression has been frequently associated with the progression, metastatic potential, and prognostic outcomes of multiple cancer types, including colon, gastric, prostate, and breast cancers. In addition, HDAC1 is linked to unfavorable prognosis and resistance to chemotherapy in cases of pancreatic cancer (42). Moreover, HDAC inhibitors exhibit potent anticancer effects in hematological malignancies and hold promise as potential therapeutic agents for treating colorectal cancer (43) and triple-negative breast cancer (44). HRAS also comprises a prevalent oncogene, which is positioned upstream of the RAS/MAPK signaling pathway and plays a pivotal role in transmitting signals from the extracellular environment to the nucleus, leading to cell growth, division, proliferation, and differentiation (45). HRAS mutations have the ability to trigger YAP1-AXL signaling, leading to metastasis in head and neck cancer (46). Moreover, HRAS overexpression in gastroenteropancreatic neuroendocrine tumors is strongly associated with a notable response to lenvatinib (47). Additionally, PSMD1 is classified as an innate immune gene, and its up-regulation is strongly associated with the progression of different types of cancers. Correspondingly, it has been used as a prognostic marker for conditions like oropharyngeal cancer (48), chronic myeloid leukemia (49), and HCC (50), among others. Furthermore, in HCC, PSMD1 is found to be significantly correlated with changes in the TIME as well as immune cells (50). In addition, RAET1E, belonging to the RAET1 gene family, is classified as a major histocompatibility complex class I–related molecule (51). Earlier studies have shown that elevated levels of RAET1E expression may be linked to poor prognosis in both cervical cancer (52) and ovarian cancer (53). Conversely, TXLNA, also referred to as IL-14, is identified as a high-

molecular-weight B cell growth factor, and its ectopic expression has often been linked to dismal prognostic outcomes in glioma (54). Earlier research indicates an association between TXLNA expression and the proliferative activity and low differentiation of HCC cells (55). This suggests a poor prognosis, thus leading to its use as a marker for assessing the malignancy of HCC (55). To summarize, these six IRGs that constitute our model are all related to the formation and progression of tumors to a certain degree. Thus, additional studies are necessary to explore these connections in more detail.

Throughout the progression of cancer, tumor cells undergo a gradual loss of their original differentiation phenotype and acquire certain stem-like characteristics. This transformation enables tumor cells to have stronger abilities for proliferation and migration, thus facilitating the progression of cancer (24, 56, 57). In this context, Malta et al. (24) discovered a correlation between the RNAss and the prognosis of TCGA-LIHC. Accordingly, a higher stemness score indicated a worse prognosis in terms of OS and PFS in patients with TCGA-LIHC. Our study yielded similar findings, indicating that as the patient prognosis worsened, both the SVMRS and the RNAss increased. Furthermore, they also highlighted a notable association between the RNAss and TIME (24). Therefore, we conducted an examination of the disparity in the ratio of immune cell infiltration between the HRG and LRG. Our findings indicate notable distinctions between the groups in five distinct immune cell types, namely CD4 memory T cells, monocytes, macrophages M0, mast cells, and neutrophils. Among them, CD4 memory T cells have been reported to have the ability to recognize and attack tumor cells, thereby aiding in the regulation of tumor growth and metastatic potential (58). The findings of our study show that the HRG had a greater percentage of activated T cell CD4 memory. This suggests that the HRG may be more susceptible to immunotherapy, which was confirmed during subsequent analysis. Monocytes exhibit dual roles in tumor immunity. On the one hand, monocytes have the ability to influence the TIME through different mechanisms, induce immune tolerance and angiogenesis, and increase the proliferation of tumor cells; on the other hand, monocytes can also produce antitumor effectors and activate antigen-presenting cells (59, 60). Moreover, monocytes also have the ability to differentiate into macrophages; these M0 macrophages, in their initial state, are also referred to as naive macrophages (61). Exosomes released by lung tumor cells have been documented to expedite the macrophage transformation of the M0 phenotype into the M2 phenotype, thereby promoting carcinogenic activities (62). Earlier studies have shown that patients with HCC with a high level of infiltration of macrophage M0 cells tend to have a negative prognosis (63). Furthermore, genes associated with macrophage M0 cells may offer insights into potential clinical treatment approaches for patients with HCC (63). The results of our study also revealed that the HRG, characterized by a poor prognosis, exhibited elevated levels of macrophage M0 infiltration. In addition, mast cells can facilitate the onset and progression of HCC by increasing the population of immunosuppressive cells, resulting in a poor prognosis (64). In our study, we found that the HRG had a lower proportion of mast cells in a resting state. Additionally, although not statistically significant, a higher proportion of mast cells in the HRG were in an activated state, indicating that patients in the HRG may have a poorer outcome. Additionally, the involvement of tumor-infiltrating neutrophils has been identified as a key factor in the malignant phenotypes of HCC. On the one hand, tumor-infiltrating neutrophils express a protein called PD1 ligand PDL1, which hinders the function of CD4+ and CD8+T cells by binding to PD1, and this interaction promotes the evasion of the immune system by the tumor (65–67). On the other hand, neutrophils release substances called CCL2 and CCL17, which attract immunosuppressive macrophages and Treg cells (65–67). Furthermore, the presence of both peritumoral and intratumoral neutrophils in patients with HCC has been linked to a negative prognosis. This observation implies that neutrophils may offer promising avenues for targeted therapeutic strategies (67). In our study, the HRG, which had a worse prognosis, demonstrated a greater proportion of neutrophils, aligning with the previous findings. Thus, the findings of our study are corroborated by numerous prior studies, indicating the rationality of the HRG and LRG employed in our study.

At present, the use of immunotherapy for HCC is in its initial stage, and there is a lack of definitive biomarkers to predict its effectiveness. However, the implementation of immunotherapy has shown promising advancements in the treatment outlook for advanced HCC (68–70). Hence, identifying biomarkers capable of precisely predicting immunotherapy effectiveness is expected to emerge as a prominent avenue in the treatment of HCC. Indicators, such as immune cell infiltration, PD-1/PD-L1, and tumor mutational burden/microsatellite instability in the TIME, are considered to hold considerable potential in predicting therapeutic efficacy (68–70). Our study revealed a notable increase in the PD-1/PD-L1 expression in the HRG and highlighted a positive correlation with the SVMRS. The analysis of immune cell infiltration results also indicates that the HRG may exhibit greater responsiveness to immunotherapy. Thus, we confirmed this conjecture through the analysis of patients in the three cohorts undergoing immunotherapy. Consequently, the model employed in this study was found to exhibit promising potential in predicting immunotherapy effectiveness. Further functional enrichment analysis revealed that the HRG exhibited activation of pathways associated with tumorigenesis and immune processes. This activation may contribute to the improved efficacy of immunotherapy, suggesting an internal mechanism. Among them, TLR is a pattern recognition receptor found in many different cells, which plays a crucial role in the innate immune response. Correspondingly, TLRs on tumor cells can enhance the stemness, proliferation, and metastasis of tumor cells, and resist cytotoxic lymphocyte attack (71). In HCC, the signal transduction pathway of TLRs is frequently associated with the progression (72). TLR3 and TLR4, among these receptors, hold potential as candidate prognostic indicators for treating HCC (72). Moreover, the TLR4 signaling pathway activation has been noted to foster the growth, mobility, and invasive capabilities of HCC cells, hinder programmed cell death, and accelerate resistance to tumor drugs (73). This suggests that targeting the TLR4 pathway could be a promising approach for immunotherapy in HCC. Moreover, TLR serves as a crucial link connecting the innate and acquired immune

systems, playing a significant role in the body's immune response (74). In the context of immunotherapy, the TLR signaling pathway participates in the regulation of PD-1/PD-L1 and PD-L2 expression (75). Furthermore, TLR9 agonists have also undergone extensive research for their potential use in tumor treatment, either as standalone therapies or in combination with other agents (76). In this context, the potential clinical application prospects of targeting TLR alone or in combination with other drugs have been demonstrated (77, 78). In our study, both the SVMRS and the 6 key genes of the model were significantly correlated with the majority of the key genes in the TLR signaling pathway. Additionally, pathway analysis revealed that the pathway was activated in the HRG. Thus, the activation of the TLR signaling pathway may contribute to a negative prognosis and enhance immunotherapy effectiveness in individuals at high risk. Consequently, targeting this pathway may serve as a promising therapeutic approach for this specific patient population.

Our study presents a novel approach for predicting OS and immunotherapy effectiveness for HCC using six IRG. However, there are still certain constraints that need to be acknowledged. First, the study relies on information obtained from a publicly available dataset. While we did utilize our own cohort to validate the findings, additional experimental evidence is required to definitively confirm the proposed hypothesis. However, our study revealed, via functional enrichment analysis, that the six IRGs play a role in regulating various pathways associated with tumor formation and progression. This finding is likely to provide valuable insights for future research and facilitate further investigation into the underlying molecular mechanisms. Secondly, the presence of diverse detection platforms and training methods in the datasets leads to variations in sequencing backgrounds and normalization techniques. Consequently, it becomes challenging to determine a universally applicable cut-off value for SVMRS across all datasets. Hence, it is necessary to initially acquire the threshold value of SVMRS through small sample detection prior to its application, and subsequently refine the threshold value through extensive clinical prospective studies. In order to determine immunotherapy effectiveness, it is necessary to conduct extensive prospective clinical trials to validate the use of high SVMRS as a predictive factor. Thirdly, it should be noted that the cohort of 54 patients under study did not undergo immunotherapy, thereby rendering the prediction analysis of immunotherapy response unfeasible. Meanwhile, the three cohorts used to validate immunotherapy effectiveness were all composed of patients with cancers other than HCC. Thus, further validation is required to determine if the predictive value of our model for immunotherapy efficacy in HCC cohorts is consistent.

Conclusions

To conclude, the current study successfully resulted in the development of a prediction model for HCC using bioinformatics analysis and machine learning. This model, based on a 6-IRG signature, has the potential to accurately predict immunotherapy response. The risk score and risk groups of our model exhibited substantial variations in tumor stemness, tumor immune cell infiltration levels, ICG expression, and immunotherapy effectiveness. The key genes in our model likely participate in the regulation of various pathways associated with tumorigenesis and immune processes. Thus, our study introduces a novel approach for predicting the prognosis of HCC and evaluating immunotherapy effectiveness, providing promising prospects for clinical application.

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 humans were approved by Clinical Research Ethics Committee of the Second Xiangya Hospital, Central South University (Approval No.: LYF2022070). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ZL: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft. LY: Conceptualization, Data curation, Formal analysis, Validation, Visualization, Writing – original draft. CL: Data curation, Formal analysis, Software, Validation, Visualization, Writing – review & editing. ZW: Data curation, Formal analysis, Software, Validation, Writing – review & editing. WX: Data curation, Methodology, Validation, Visualization, Writing – review & editing. JL: Methodology, Software, Visualization, Writing – review & editing. CW: Conceptualization, Supervision, Writing – review & editing. XX: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the platform funding of Hunan Provincial Key Laboratory of Hepatobiliary Disease Research. (No. 2017TP1007).

Acknowledgments

We would like to thank the patients who shared their experiences with our oncologists, and we are grateful to Bullet Edits Limited for copyediting the manuscript.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1371829/full#supplementary-material

References

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

2. Rumgay H, Arnold M, Ferlay J, Lesi O, Cabasag CJ, Vignat J, et al. Global burden of primary liver cancer in 2020 and predictions to 2040. *J Hepatol*. (2022) 77:1598–606. doi: 10.1016/j.jhep.2022.08.021

3. Llovet JM, Zucman-Rossi J, Pikarsky E, Sangro B, Schwartz M, Sherman M, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers*. (2016) 2:16018. doi: 10.1038/nrdp.2016.18

4. Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N Engl J Med*. (2020) 382:1894–905. doi: 10.1056/NEJMoa1915745

5. Ren Z, Xu J, Bai Y, Xu A, Cang S, Du C, et al. Sintilimab plus a bevacizumab biosimilar (IBI305) versus sorafenib in unresectable hepatocellular carcinoma (ORIENT-32): a randomized, open-label, phase 2–3 study. *Lancet Oncol*. (2021) 22:977–90. doi: 10.1016/S1470–2045(21)00252–7

6. Greten TF, Villanueva A, Korangy F, Ruf B, Yarchoan M, Ma L, et al. Biomarkers for immunotherapy of hepatocellular carcinoma. *Nat Rev Clin Oncol*. (2023) 20:780–98. doi: 10.1038/s41571–023-00816–4

7. Mounir M, Lucchetta M, Silva TC, Olsen C, Bontempi G, Chen X, et al. New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. *PloS Comput Biol*. (2019) 15:e1006701. doi: 10.1371/journal.pcbi.1006701

8. Goldman MJ, Craft B, Hastie M, Repecka K, McDade F, Kamath A, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol*. (2020) 38:675–8. doi: 10.1038/s41587–020-0546–8

9. Chen Z, Luo Z, Zhang D, Li H, Liu X, Zhu K, et al. TIGER: A web portal of tumor immunotherapy gene expression resource. *Genomics Proteomics Bioinf*. (2023) 21:337–48. doi: 10.1016/j.gpb.2022.08.004

10. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science*. (2015) 350:207–11. doi: 10.1126/science.aad0095

11. Jung H, Kim HS, Kim JY, Sun JM, Ahn JS, Ahn MJ, et al. DNA methylation loss promotes immune evasion of tumors with high mutation and copy number load. *Nat Commun*. (2019) 10:4278. doi: 10.1038/s41467–019-12159–9

12. Braun DA, Hou Y, Bakouny Z, Ficial M, Sant' AM, Forman J, et al. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat Med*. (2020) 26:909–18. doi: 10.1038/s41591-020-0839-y

13. Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci Data*. (2018) 5:180015. doi: 10.1038/sdata.2018.15

14. Stoltzfus JC. Logistic regression: a brief primer. *Acad Emerg Med*. (2011) 18:1099–104. doi: 10.1111/j.1553-2712.2011.01185.x

15. Alhamzawi R, Ali H. The Bayesian adaptive lasso regression. *Math Biosci*. (2018) 303:75–82. doi: 10.1016/j.mbs.2018.06.004

16. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. (2010) 33:1–22. doi: 10.18637/jss.v033.i01

17. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization paths for cox's proportional hazards model via coordinate descent. *J Stat Softw*. (2011) 39:1–13. doi: 10.18637/jss.v039.i05

18. Sanz H, Valim C, Vegas E, Oller JM, Reverter F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinf*. (2018) 19:432. doi: 10.1186/s12859–018-2451–4

19. Chen D, Liu J, Zang L, Xiao T, Zhang X, Li Z, et al. Integrated machine learning and bioinformatic analyses constructed a novel stemness-related classifier to predict prognosis and immunotherapy responses for hepatocellular carcinoma patients. *Int J Biol Sci*. (2022) 18:360–73. doi: 10.7150/ijbs.66913

20. Kuhn M. Building predictive models in R using the caret package. *J Stat SOFTW*. (2008) 28:1–26. doi: 10.18637/jss.v028.i05

21. Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. *J Stat Softw*. (2012) 50:1–23. doi: 10.18637/jss.v050.i11

22. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. (2013) 32:5381–97. doi: 10.1002/sim.5958

23. Barbosa AM, Real R, Munoz AR, Brown JA. New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Divers Distrib*. (2013) 19:1333–8. doi: 10.1111/ddi.12100

24. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell*. (2018) 173:338–54. doi: 10.1016/j.cell.2018.03.034

25. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol*. (2019) 37:773–82. doi: 10.1038/s41587–019-0114–2

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

27. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. (2021) 2:100141. doi: 10.1016/j.xinn.2021.100141

28. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118

29. Walter W, Sanchez-Cabo F, Ricote M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics*. (2015) 31:2912–4. doi: 10.1093/bioinformatics/btv300

30. Belinky F, Nativ N, Stelzer G, Zimmerman S, Iny ST, Safran M, et al. PathCards: multi-source consolidation of human biological pathways. *Database (Oxford)*. (2015) 2015:bav006. doi: 10.1093/database/bav006

31. Gu Z. Complex heatmap visualization. *iMeta*. (2022) 1:e43. doi: 10.1002/imt2.43

32. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. (2016) 32:2847–9. doi: 10.1093/bioinformatics/btw313

33. Wickham H. *ggplot2: elegant graphics for data analysis*. new york: Springer-verlag (2016). doi: 10.1007/978-3-319-24277-4

34. Shen W, Song Z, Zhong X, Huang M, Shen D, Gao P, et al. Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta*. (2022) 1:e36. doi: 10.1002/imt2.36

35. Wang T, Guo K, Zhang D, Wang H, Yin J, Cui H, et al. Disulfidptosis classification of hepatocellular carcinoma reveals correlation with clinical prognosis and immune profile. *Int Immunopharmacol*. (2023) 120:110368. doi: 10.1016/j.intimp.2023.110368

36. Chen Y, Tang L, Huang W, Abisola FH, Zhang Y, Zhang G, et al. Identification of a prognostic cuproptosis-related signature in hepatocellular carcinoma. *Biol Direct*. (2023) 18:4. doi: 10.1186/s13062-023-00358-w

37. Shi Y, Wang J, Huang G, Zhu J, Jian H, Xia G, et al. A novel epithelial–mesenchymal transition gene signature for the immune status and prognosis of

hepatocellular carcinoma. *Hepatol Int*. (2022) 16:906–17. doi: 10.1007/s12072–022-10354–3

38. Liu Z, Liu Y, Li T, Wang P, Mo X, Lv P, et al. CMTM7 plays key roles in TLR-induced plasma cell differentiation and p38 activation in murine B-1 B cells. *Eur J Immunol*. (2020) 50:809–21. doi: 10.1002/eji.201948363

39. Liu B, Lu Y, Zhang T, Yu X, Wang Q, Chi Y, et al. CMTM7 as a novel molecule of ATG14L-Beclin1-VPS34 complex enhances autophagy by Rab5 to regulate tumorigenicity. *Cell Commun Signal*. (2021) 19:77. doi: 10.1186/s12964–021-00720–3

40. Jiang X, Qian Z, Chen Y, Zhou T, Zhao C, Yin Y. CMTM7 recognizes an immune-hot tumor microenvironment and predicts therapeutic response of immunotherapy in breast cancer well. *Front Genet*. (2022) 13:1051269. doi: 10.3389/fgene.2022.1051269

41. Li Y, Seto E. HDACs and HDAC inhibitors in cancer development and therapy. *Cold Spring Harb Perspect Med*. (2016) 6:a026831. doi: 10.1101/cshperspect.a026831

42. Wright CA, Gordon ER, Cooper SJ. Genomic analysis reveals HDAC1 regulates clinically relevant transcriptional programs in Pancreatic cancer. *BMC Cancer*. (2023) 23:1137. doi: 10.1186/s12885–023-11645–0

43. Duan N, Hu X, Qiu H, Zhou R, Li Y, Lu W, et al. Targeting the E2F1/Rb/HDAC1 axis with the small molecule HR488B effectively inhibits colorectal cancer growth. *Cell Death Dis*. (2023) 14:801. doi: 10.1038/s41419–023-06205–0

44. Pang Y, Shi R, Chan L, Lu Y, Zhu D, Liu T, et al. The combination of the HDAC1 inhibitor SAHA and doxorubicin has synergic efficacy in triple negative breast cancer. *vivo. Pharmacol Res*. (2023) 196:106926. doi: 10.1016/j.phrs.2023.106926

45. Simanshu DK, Nissley DV, McCormick F. RAS proteins and their regulators in human disease. *Cell*. (2017) 170:17–33. doi: 10.1016/j.cell.2017.06.009

46. Jagadeeshan S, Prasad M, Badarni M, Ben-Lulu T, Liju VB, Mathukkada S, et al. Mutated HRAS activates YAP1-AXL signaling to drive metastasis of head and neck cancer. *Cancer Res*. (2023) 83:1031–47. doi: 10.1158/0008–5472.CAN-22–2586

47. Liverani C, Spadazzi C, Ibrahim T, Pieri F, Foca F, Calabrese C, et al. HRAS overexpression predicts response to Lenvatinib treatment in gastroenteropancreatic neuroendocrine tumors. *Front Endocrinol (Lausanne)*. (2022) 13:1045038. doi: 10.3389/fendo.2022.1045038

48. Park HC, Kim H, Kim JY, Lee HY, Lee J, Cha W, et al. PSMD1 as a prognostic marker and potential target in oropharyngeal cancer. *BMC Cancer*. (2023) 23:1242. doi: 10.1186/s12885–023-11689–2

49. Bencomo-Alvarez AE, Rubio AJ, Olivas IM, Gonzalez MA, Ellwood R, Fiol CR, et al. Proteasome 26S subunit, non-ATPases 1 (PSMD1) and 3 (PSMD3), play an oncogenic role in chronic myeloid leukemia by stabilizing nuclear factor-kappa B. *ONCOGENE*. (2021) 40:2697–710. doi: 10.1038/s41388–021-01732–6

50. Chen X, Liu G, Wu B. Analysis and experimental validation of the innate immune gene PSMD1 in liver hepatocellular carcinoma and pan-cancer. *Heliyon*. (2023) 9:e21164. doi: 10.1016/j.heliyon.2023.e21164

51. Rodriguez JM, Wolfrum S, Robblee M, Chen KY, Gilbert ZN, Choi JH, et al. Altered expression of Raet1e, a major histocompatibility complex class 1-like molecule, underlies the atherosclerosis modifier locus Ath11 10b. *Circ Res*. (2013) 113:1054–64. doi: 10.1161/CIRCRESAHA.113.302052

52. Cho H, Chung JY, Kim S, Braunschweig T, Kang TH, Kim J, et al. MICA/B and ULBP1 NKG2D ligands are independent predictors of good prognosis in cervical cancer. *BMC Cancer*. (2014) 14:957. doi: 10.1186/1471–2407-14–957

53. McGilvray RW, Eagle RA, Rolland P, Jafferji I, Trowsdale J, Durrant LG. ULBP2 and RAET1E NKG2D ligands are independent predictors of poor prognosis in ovarian cancer patients. *Int J Cancer*. (2010) 127:1412–20. doi: 10.1002/ijc.25156

54. Hu B, Chen D, Li Y, Yu S, Kuang L, Ma X, et al. Expression of TXLNA in brain gliomas and its clinical significance: a bioinformatics analysis. *Chin Neurosurg J*. (2023) 9:27. doi: 10.1186/s41016–023-00341–4

55. Ohtomo N, Tomiya T, Tanoue Y, Inoue Y, Nishikawa T, Ikeda H, et al. Expression of alpha-taxilin in hepatocellular carcinoma correlates with growth activity and Malignant potential of the tumor. *Int J Oncol*. (2010) 37:1417–23. doi: 10.3892/ijo_00000793

56. Shibue T, Weinberg RA. EMT. CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat Rev Clin Oncol*. (2017) 14:611–29. doi: 10.1038/nrclinonc.2017.44

57. Ge Y, Gomez NC, Adam RC, Nikolova M, Yang H, Verma A, et al. Stem cell lineage infidelity drives wound repair and cancer. *Cell*. (2017) 169:636–50. doi: 10.1016/j.cell.2017.03.042

58. Kunzli M, Masopust D. CD4(+) T cell memory. *Nat Immunol*. (2023) 24:903–14. doi: 10.1038/s41590–023-01510–4

59. Olingy CE, Dinh HQ, Hedrick CC. Monocyte heterogeneity and functions in cancer. *J Leukoc Biol*. (2019) 106:309–22. doi: 10.1002/JLB.4RI0818-311R

60. Ugel S, Cane S, De Sanctis F, Bronte V. Monocytes in the tumor microenvironment. *Annu Rev Pathol*. (2021) 16:93–122. doi: 10.1146/annurev-pathmechdis-012418–013058

61. Chaintreuil P, Kerreneur E, Bourgoin M, Savy C, Favreau C, Robert G, et al. The generation, activation, and polarization of monocyte-derived macrophages in human Malignancies. *Front Immunol*. (2023) 14:1178337. doi: 10.3389/fimmu.2023.1178337

62. Pritchard A, Tousif S, Wang Y, Hough K, Khan S, Strenkowski J, et al. Lung tumor cell-derived exosomes promote M2 macrophage polarization. *Cells-Basel*. (2020) 9:1303. doi: 10.3390/cells9051303

63. Xu X, Wang J. Prognostic prediction and multidimensional dissections of a macrophages M0-related gene signature in liver cancer. *Front Endocrinol (Lausanne)*. (2023) 14:1153562. doi: 10.3389/fendo.2023.1153562

64. Huang S, Wu H, Luo F, Zhang B, Li T, Yang Z, et al. Exploring the role of mast cells in the progression of liver disease. *Front Physiol*. (2022) 13:964887. doi: 10.3389/fphys.2022.964887

65. Arvanitakis K, Mitroulis I, Germanidis G. Tumor-associated neutrophils in hepatocellular carcinoma pathogenesis, prognosis, and therapy. *Cancers (Basel)*. (2021) 13:2899. doi: 10.3390/cancers13122899

66. Chen H, Zhou XH, Li JR, Zheng TH, Yao FB, Gao B, et al. Neutrophils: Driving inflammation during the development of hepatocellular carcinoma. *Cancer Lett*. (2021) 522:22–31. doi: 10.1016/j.canlet.2021.09.011

67. Geh D, Leslie J, Rumney R, Reeves HL, Bird TG, Mann DA. Neutrophils as potential therapeutic targets in hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol*. (2022) 19:257–73. doi: 10.1038/s41575–021-00568–5

68. Pallozzi M, Di Tommaso N, Maccauro V, Santopaolo F, Gasbarrini A, Ponziani FR, et al. Non-invasive biomarkers for immunotherapy in patients with hepatocellular carcinoma: current knowledge and future perspectives. *Cancers (Basel)*. (2022) 14:4631. doi: 10.3390/cancers14194631

69. He Y, Lu M, Che J, Chu Q, Zhang P, Chen Y. Biomarkers and future perspectives for hepatocellular carcinoma immunotherapy. *Front Oncol*. (2021) 11:716844. doi: 10.3389/fonc.2021.716844

70. Ren D, Hua Y, Yu B, Ye X, He Z, Li C, et al. Predictive biomarkers and mechanisms underlying resistance to PD1/PD-L1 blockade cancer immunotherapy. *Mol Cancer*. (2020) 19:19. doi: 10.1186/s12943–020-1144–6

71. Cao LL, Kagan JC. Targeting innate immune pathways for cancer immunotherapy. *IMMUNITY*. (2023) 56:2206–17. doi: 10.1016/j.immuni.2023.07.018

72. Huang C, Zhou Y, Cheng J, Guo X, Shou D, Quan Y, et al. Pattern recognition receptors in the development of nonalcoholic fatty liver disease and progression to hepatocellular carcinoma: An emerging therapeutic strategy. *Front Endocrinol (Lausanne)*. (2023) 14:1145392. doi: 10.3389/fendo.2023.1145392

73. Papadakos SP, Arvanitakis K, Stergiou IE, Lekakis V, Davakis S, Christodoulou MI, et al. The role of TLR4 in the immunotherapy of hepatocellular carcinoma: can we teach an old dog new tricks? *Cancers (Basel)*. (2023) 15:2795. doi: 10.3390/cancers15102795

74. Akira S, Takeda K, Kaisho T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat Immunol*. (2001) 2:675–80. doi: 10.1038/90609

75. Curran CS, Gupta S, Sanz I, Sharon E. PD-1 immunobiology in systemic lupus erythematosus. *J Autoimmun*. (2019) 97:1–9. doi: 10.1016/j.jaut.2018.10.025

76. Dongye Z, Li J, Wu Y. Toll-like receptor 9 agonists and combination therapies: strategies to modulate the tumor immune microenvironment for systemic anti-tumor immunity. *Br J Cancer*. (2022) 127:1584–94. doi: 10.1038/s41416–022-01876–6

77. Kaur A, Baldwin J, Brar D, Salunke DB, Petrovsky N. Toll-like receptor (TLR) agonists as a driving force behind next-generation vaccine adjuvants and cancer therapeutics. *Curr Opin Chem Biol*. (2022) 70:102172. doi: 10.1016/j.cbpa.2022.102172

78. Tran TH, Tran T, Truong DH, Nguyen HT, Pham TT, Yong CS, et al. Toll-like receptor-targeted particles: A paradigm to manipulate the tumor microenvironment for cancer immunotherapy. *Acta Biomater*. (2019) 94:82–96. doi: 10.1016/j.actbio.2019.05.043

frontiers | Frontiers in Oncology

TYPE Mini Review
PUBLISHED 09 August 2024
DOI 10.3389/fonc.2024.1441883

OPEN ACCESS

EDITED BY
Sergio Luis Felisbino,
São Paulo State University, Brazil

REVIEWED BY
Sanja Stifter-Vretenar,
Skejby Sygehus, Denmark

*CORRESPONDENCE
Dalila Lucíola Zanette
dalila.zanette@fiocruz.br

RECEIVED 31 May 2024
ACCEPTED 16 July 2024
PUBLISHED 09 August 2024

CITATION
Coelho KBCA, Wosniaki DK, Marin AM, Fabris L, Borges dos Reis R, Aoki MN and Zanette DL (2024) Urinary mRNA-based biomarkers for non-muscle-invasive bladder cancer: a mini-review. *Front. Oncol.* 14:1441883. doi: 10.3389/fonc.2024.1441883

Urinary mRNA-based biomarkers for non-muscle-invasive bladder cancer: a mini-review

Karoline Brito Caetano Andrade Coelho[1], Denise Kusma Wosniaki[2], Anelis Maria Marin[2], Laura Fabris[3], Rodolfo Borges dos Reis[1], Mateus Nóbrega Aoki[2] and Dalila Lucíola Zanette[1,2]*

[1]Uro-Oncology Laboratory, Surgery and Anatomy Department, Ribeirao Preto Medical School, University of Sao Paulo, Ribeirão Preto, São Paulo, Brazil, [2]Laboratory for Applied Science and Technology in Health, Carlos Chagas Institute, Oswaldo Cruz Foundation (Fiocruz), Curitiba, Paraná, Brazil, [3]Department of Applied Science and Technology, Politecnico di Torino, Torino, Italy

Bladder cancer (BC) is the second most common type of cancer of the urinary system. Approximately 75% of the cases are non-muscle invasive bladder cancer (NMIBC), which has a high recurrence and progression rate. Current diagnosis and surveillance methods present challenges, including risks to the patients. For this reason, urinary biomarkers have been proposed as alternatives to the methods. The goal of this mini-review is to describe urinary mRNA-based biomarkers available in current literature for NMIBC tumors, using the PubMed database. The search included the following keywords: "biomarkers" AND "bladder cancer" AND "urine" and "RNA" and "non-muscle". The search yielded 11 original researchers utilizing mRNA-based urinary biomarkers. Although there is a wide variety of biomarkers described, the cohorts of the studies were not exclusively NMIBC, which is the subtype of BC that would mostly benefit from the introduction of a good follow-up biomarker, highlighting the need for randomized interventional trials for NMIBC.

KEYWORDS

non-muscle invasive bladder cancer (NMIBC), urinary biomarkers, mRNA-based, muscle-invasive bladder cancer (MIBC), surveillance, aggressiveness

Introduction

Bladder cancer (BC) is the second most common type of cancer of the urinary system and the thirteenth most common cause of cancer death worldwide (https://gco.iarc.fr/). Although there are several risk factors for BC, 82% of all cases are due to modifiable risk factors (lifestyle and occupational exposure). Tobacco is the most recognized risk factor for BC (1).

Over 90% of BC cases are classified as Urothelial Cell Carcinoma (UCC) or Transitional Cell Carcinoma (TCC), which originates from the urothelium of the bladder. UCC can be subdivided into non-muscle-invasive (NMIBC), muscle-invasive (MIBC), or metastatic.

Approximately 75% of all cases present the subtype NMIBC, while 25% have MIBC or metastatic disease. Tumors that remain confined to the epithelium (urothelium) are defined as NMIBC (Stages Ta, T1, and Tis), and the tumors that invade the muscle layer of the bladder, (Stage T2), perivesical fat (Stage T3), or the adjacent organs (Stage T4) are defined as MIBC (2–4).

Cystoscopy is the gold standard for the diagnosis and surveillance of BC (5, 6) Patient with non-invasive disease (NMIBC) present a higher risk of recurrence and progression to muscle-invasive disease and their follow-up requires a greater frequency of cystoscopies. Nevertheless, cystoscopy is an invasive exam for the patient and expensive for public health systems. In addition, it may generate infection, pain, and, in some cases, hematuria (4, 7). Urinary cytology can be useful as a noninvasive, inexpensive, and highly specific tool to complement cystoscopy. Cytology has a moderate sensitivity to detect high-grade lesions, but its sensitivity is low, around 20 to 50% for low-grade papillary tumors (8). As a result, most patients with a cytologic diagnosis of a low-grade urothelial neoplasm prove not to have a tumor. The false-positive rates of urine cytology range from 1.3% to 15%, and false positives occur in patients with bladder stones, human polyomavirus infections, and prior chemotherapy (9).

For this reason, non-invasive, ancillary tools that allow a longer interval between cystoscopies are needed to reduce risks to the patients and costs to the healthcare systems, especially in the case of NMIBC tumors. Based on this rationale, urinary biomarkers for detection and surveillance have been proposed for decades as alternatives to cystoscopy. The urine is a perfect candidate as a biological sample, not only because it is obtained in a non-invasive way, but also because of its continuous contact with the bladder tumoral tissue, which enables it to provide useful transcriptomic, epigenetic, and genomic insights that may be related to BC (10) In this mini-review, we will focus on describing of urinary mRNA-based biomarkers to identify potential biomarkers to identify aggressiveness in NMIBC.

Urinary biomarkers

Studies have proposed urinary biomarkers for the detection and surveillance of bladder cancer including urine cytology, protein-based, cell-based, genomic, and transcriptomic approaches. Cell and protein-based biomarkers have been approved by the Food and Drug Administration (FDA), such as Cytology, uCyt+, and UroVysion (exfoliated cells in urine sample) and NMP22 enzyme and Bladder Tumor Antigen – BTA – (protein in urine sample) (11). Although they show increased sensitivity for low-grade tumors, their specificity still doesn't surpass cytology. Currently, the recommendation in the European Association of Urology (EAU) guidelines continues to be cytology, in association with cystoscopy.

Despite substantial efforts, there is still a need for randomized interventional trials that are multicentric and prospective. The currently available literature is still discrepant, with small cohorts, usually performed in only one center, with analytics divergences, with the result that the biomarkers that have been identified to date do not present superior accuracy to the gold standard. Because of these limitations, these biomarkers have not been incorporated into current clinical practice (12).

Numerous reports have proposed proteins, DNA or RNA biomarkers for diagnosis of BC in urine. Until the year 2000, protein biomarkers were dominant in the literature, but more recently the proportion of DNA and RNA studies has increased (13). Protein-based biomarkers are susceptible to conditions in which the presence of protein is increased in the urine, such as inflammation, hematuria, and kidney stones. DNA-based biomarkers assess genetic alterations (point mutations, copy number alterations, and epigenetic changes including DNA methylation). The stability of the DNA molecules is an advantage over messenger RNA since the collection and transportation of samples would be simpler for DNA, but mRNA has great potential, so it has been recently described in an increasing number of studies. Compared with protein biomarkers, RNA biomarkers can be detected with greater sensitivity and specificity, and in general primers and probes are cheaper than antibodies, which are used to detect proteins. There are several possible RNA biomarkers, both coding and non-coding RNAs, including microRNAs, long non-coding RNAs, and circRNAs, which have been studied as potential biomarkers in bladder cancer in the past few years (11).

Studies using extracellular vesicles still present many challenges in clinical practice due to the lack of standardization in the methodologies for isolation and analysis and the lack of multicentric validations, despite all the efforts from the scientific community to standardize methods and results in this area (14).

mRNAs have advantages over protein and DNA biomarkers that compensate its instability that requires special conditions of collection and transportation since the methods to detect mRNA biomarkers have lower costs when compared to protein biomarkers and provide dynamic insights into cellular states and regulatory processes compared to DNA biomarkers (15). This review describes the utility and accuracy of messenger RNAs as biomarkers to monitor NMIBC by evaluating differentially expressed transcripts present in cell-free urine or urine cells (Supplementary Figure 1).

Potential mRNA-based urinary biomarkers for NMIBC

5-mRNA (*ABL1, ANXA10, UPK1B, CRH*, and *IGF2*)

Pichler and colleagues (2018) (16) analyzed the 5-mRNA (*ABL1, ANXA10, UPK1B, CRH*, and *IGF2*) model proposed by Wallace and colleagues (2018) (17), now named Xpert BC Monitor and showed that it presents sensitivity superior to cytology, even in NMIBC low-grade and pTa disease, while overall specificity is similar. Xpert BC Monitor successfully discriminated between tumor stages, grades, size, and number of tumors, and previous intravesical instillations didn't increase the rate of false positivity. In addition, combining this test with barbotage cytology (bladder washing) did not enhance diagnostic accuracy compared with the

test alone (AUC=0.85 vs. AUC=0.87). In contrast, a prospective study with 230 patients with NMIBC tumors showed that overall sensitivity for Xpert BC Monitor was higher than for cytology and when combined, Xpert BC Monitor and cytology, it was superior to cytology alone. However, the overall specificity for cytology is better (18).

Another report showed that Xpert had an overall high diagnostic capability to detect residual tumors in repeat biopsy after initial complete Transurethral Resection of Bladder Tumor (TURBT) of T1BC (Stage T1 of BC) in NMIBC patients, with a sensitivity of 86% and negative predictive value (NPV) of 89%. The results of the Xpert test were independently associated with early tumor recurrence, suggesting that Xpert can detect genetic abnormalities before macroscopic existence by checking cystoscopy (19). The approach could help reduce invasiveness in follow-up of these patients due to the partially reduced need for cystoscopy and, consequently, could improve adherence. In concordance, other studies have described that Xpert could also be a promising tool in follow-up of recurrent NMIBC patients and could function as a predictive tool to determine the presence of residual tumors after primary TURBT. The Xpert Monitor presented higher sensitivity and an improved NPV when compared with UroVysion and cytology in patients under follow-up for BC. The specificity was minimally improved compared with UroVysion and was lower compared with cytology. Xpert was more sensitive for both high-grade and low-grade BC. The high NPV for high-grade disease is particularly important for NMIBC monitoring, as high NPV gives high confidence that the test is truly negative, allowing to reliably exclude recurrent disease. This reliability would allow waiving one cystoscopy if the Xpert result is negative, as the currently advised follow-up schedule for low-risk NMIBC consists of cystoscopy at three and twelve months after TURBT. Additionally, Xpert showed robust reproducibility and good specificity in non-BC patients (18).

Briefly, the biomarkers found in the Xpert BC Monitor test are mRNAs translated into proteins that are related to cell pathways such as cell division, adhesion, differentiation, and response to stress (*ABL1*), cell growth and signal transduction (*ANXA10*), epigenetic dysregulation in BC (*UPK1B*), neuroendocrine stress response, immunity, and inflammation (*CRH*), and proliferation and survival (*IGF2*) (16).

ABL proto-oncogene 1, non-receptor tyrosine kinase (*ABL1*) encodes a protein tyrosine kinase involved in a variety of cellular processes. The BCR region of ABL1 presents retrotransposon repeats that have been associated with bladder cancer (20). Annexin A10 (*ANXA10*) encodes a member of the annexin family of calcium-dependent phospholipid-binding proteins. This protein was found to play a role in the regulation of cellular growth and signal transduction pathways in BC. *UPK1B* encodes a uroplakin. Four different uroplakin proteins are known at present. These proteins heterodimerize and form urothelial plaques on the surface of urothelial cells. Uroplakins are significantly downregulated during urothelial transformation and tumorigenesis. In BC, *UPK1B* gene transcription is regulated epigenetically via CpG methylation. The corticotropin-releasing hormone (*CRH*) system was initially identified as a hypothalamus-directed mediator of neuroendocrine stress response, while recent studies suggest a link between *CRH* and the development of solid cancers. Preclinical studies showed the proinflammatory and procarcinogen nature of CRH family peptides and their receptors, and the fact that they modulate immunity, inflammation, and tumor cell growth. The last gene in this panel, Insulin-like growth factor 2 (*IGF2*) is a mitogenic peptide hormone overexpressed in aggressive tumors and during embryonic development. The binding of *IGF2* to its receptor, IGF1R, initiates breast and lung tumorigenesis and promotes the progression of endometrial and gastric cancers. Overexpression of *IGF2* is at least partly caused by loss of imprinting in prostate, and colon cancers, but its deregulation may also be attributable to an abnormal expression of transcription factors. Thus IGF2/IGF1R signaling enhances tumor progression in several cancers (21), but its contribution to BC progression is still unclear despite its good performance as one of the biomarkers of the XPert BC Monitor test.

According to these authors, the limitation of the use of mRNA-based techniques is the difficulty of obtaining enough high-quality RNA from voided urine. In the Xpert test, the *ABL1* mRNA functions as a sample adequacy control to verify that the sample contains human cells and human RNA. Moreover, there are some discrepancies between the studies utilizing Xpert: 1) variability in sensibility (85.9%, 84%, and 46%), 2) variability in specificity (72.3%, 91%, and 77%), and 3) lack of validation (Table 1). Although promising results, the test accuracy was discrepant between the studies, and research on long-term follow-up is needed.

Potential mRNA-based urinary biomarkers for NMIBC and MIBC

Cxbladder Monitor Test (*CDC2, HOXA13, MDK, CXCR2*, and *IGFBP5*)

O'Sullivan and colleagues (2012) (26) developed 2 classifiers for risk stratification of urothelial cancer from their mRNA assay data (29). The classifier Cxbladder-D included the fifth marker, neutrophil marker (*CXCR2*), to reduce the risk of false-positive results in the inflamed urothelium. The second classifier, Cxbladder-S, was able to stratify tumors into low-risk - low-grade stage Ta, with a sensitivity of 91% and a specificity of 90%, respectively. In addition, the same group showed that the quantitative measurement of these five gene expression markers presented high sensitivity and negative predictive value to rule out recurrent urothelial carcinoma during surveillance (23).

Furthermore, the CxBladder Monitor showed superior performance compared with currently available, FDA-approved urine tests used as adjuncts to cystoscopy. Subgroup analyses demonstrated superior sensitivity and NPV for Cxbladder Monitor regardless of patient age and sex, or recurrent tumor size, stage, or grade by comparison with NPM22 Elisa, NPM22 Bladder Check and cytology. CxBladder Monitor had a superior sensitivity compared to NMP22 enzyme-linked immunosorbent assay, NMP22 BladderChek, and UroVysion fluorescence and

TABLE 1 Original researchers utilizing mRNA-based biomarkers.

References	Year	Only NMIBC	NMIBC sample	Markers	Urine use (Isolation RNA)	Method	Recurrence (%)	SN[a] (%)	SP[b] (%)	NPV[c] (%)	PPV[d] (%)	AUC[e]	Validation	Aim
PMID: 33785220 (19)	2021	Yes	NMIBC, n=254	*ABL1, ANXA10, UPK1B, CRH*, and *IGF2*	Xpert Urine Transport Reagent Kit	RT-PCR	24	85.9	72.3	88.9	66.4	0.78	No	Recurrence
PMID: 28941000 (16)	2018	Yes	NMIBC, n=140	*ABL1, ANXA10, UPK1B, CRH*, and *IGF2*	Xpert Urine Transport Reagent Kit	RT-PCR	NA	84.0	91.0	93.0	NA	0.87	No	Surveillance
PMID: 30553612 (18)	2019	No	–	*ABL1, ANXA10, UPK1B, CRH*, and *IGF2*	Xpert Urine Transport Reagent Kit	RT-PCR	18	74	80	93	27.8	NA	No	Recurrence
PMID: 30355587 (22)	2019	Yes	NMIBC, n=230	*ABL1, ANXA10, UPK1B, CRH*, and *IGF2*	Xpert Urine Transport Reagent Kit	RT-PCR	22	46.2	77.0	83	36.9	0.65	No	Diagnosis
PMID: 29061538 (17)	2018	No	NMIBC, n=49	*ABL1, ANXA10, UPK1B, CRH*, and *IGF2*	Xpert Urine Transport Reagent Kit	RT-qPCR	NA	73.0	NA	NA	NA	0.87	Yes	Diagnosis
PMID: 27986532 (23)	2017	No	–	*CDC2, HOXA13, MDK, CXCR2, and IGFBP5*	Not reported		NA	92.0	97.0	0.96	NA	0.66	NA	Surveillance
PMID: 28366272 (24)	2017	No	NMIBC, n=957	*CDC2, HOXA13, MDK, CXCR2, and IGFBP5*	The voided mid-stream urine was stabilized according to the manufacturer's instructions for each comparator test.		NA	91.0	NA	0.96	NA	NA	No	Diagnosis
PMID: 22818138 (25)	2012	No	NMIBC, n=55	*CDC2, HOXA13, MDK, CXCR2, and IGFBP5*	Voided urine was mixed with an equal volume of Cxbladder storage buffer		NA	91.0	90.0	NA	NA	0.87	No	Risk stratification
PMID: 33766467 (26)	2021	No	NMIBC, n=59	*ROBO1, WNT5A, CDC42BPB, ABL1, CRH, IGF2, ANXA10, and UPK1B*	Not reported	RT-PCR	NA	92.5	73.5	97.4	47.1	0.923	No	Risk stratification
PMID: 30771285 (27)	2019	No	NMIBC, n=127	*ANXA10, IGF2,KIFC3, KRT20,LCN2, MAGEA3, RPS21, and SLC1A6*	Cell pellet (TRIzol reagent, Invitro)	RNA-seq and nCounter	NA	94.0	NA	98.0	NA	0.823	Yes	Diagnosis
PMID: 24852426 (28)	2014	No	NMIBC, n=50	XIAP	Pellet (RNA purification kit, Norgen Biotek)	RT-PCR	44	82.91	78.38	NA	NA	0.85	NA	Diagnosis

[a]Sensibility.
[b]Specificity.
[c]Negative predictive value.
[d]Positive predictive value.
[e]Area under curve.
NA, Non applicable.

10.3389/fonc.2024.1441883

urine cytology, in patients Ta, Tis, and ≥T1 undergoing monitoring for recurrence. The clinical utility of Cxbladder Monitor was demonstrated as a confirmatory negative test that may be used as an adjunctive to cystoscopy, improving the monitoring for recurrent UC, or as a direct rule-out test for patients identified as being at low risk for recurrent disease (24). Thus, CxBladder may be useful as an adjunct tool to cystoscopy to risk stratify and monitor recurrence in patients with urothelial cancer. Moreover, a study from Li and collaborators showed that the use of CxMonitor (CxM) as a home urine test allowed patients to skip their scheduled surveillance cystoscopy in the presence of a CxM-negative test. The authors report that 66 CxM-negative patients skipped cystoscopy, and none had findings on follow-up cystoscopy that required biopsies (30).

The biomarkers of the CxBladder Monitor are distinct from those in the XPert BC test. In summary, the biomarkers found in the Cxbladder Monitor Test are mRNAs that are translated into proteins related to cell pathways such as cell cycle (*CDC2*), gene expression regulation, morphogenesis, and differentiation (*HOXA13*), migration, growth, and angiogenesis (*MDK*), and cellular response, regulation of smooth muscle cell migration and proliferation (*IGFBP5*) (https://www.ncbi.nlm.nih.gov/home/genes/). *CDC2* (CDK1) CDK1 phosphorylates TFCP2L1, a pluripotency-associated transcription factor, and the CDK1-TFCP2L1 pathway is activated in BC cells, stimulating their proliferation, self-renewal, and invasion. In patients with BC, high co-expression of *TFCP2L1* and *CDK1* was associated with unfavorable clinical characteristics including tumor grade and distant metastasis (31). *HOXA13* gene is higher in low-grade tumors compared to high-grade BC tumor samples, which suggests its potential as a diagnostic marker in NMIBC. The expression level of *HOXA13* has also been reported to be higher in NMIBC urine samples than in normal controls. *HOXA13* gene expression has been tested as a diagnostic marker for NMIBC along with *PLK1* and *FGFR3* by Valizadeh and colleagues, who describe *HOXA13* with a greater sensitivity compared to *PLK1* and *FGFR3* (32). Midkine (*MDK*) is a heparin-binding growth factor that is overexpressed in bladder tumor tissue and urine from BC patients when compared to healthy individuals. It has been shown that microscopic hematuria and infection were not obstacles to detecting BC by *MDK* mRNA test (PMID: (33). IGFBP5 prolongs insulin-like growth factors (IGFs) half-life and restricts their function, affecting the IGF signaling pathway, which plays a role in cellular growth, differentiation, and apoptosis. IGFBP5 overexpression strongly correlates with several adverse prognostic factors in BC (34).

3-marker urinary panel (*ROBO1, CRH*, and *IGF2*)

A 3-marker urinary mRNA panel was proposed by Shkolyar and colleagues (2021) (26) to identify intermediate and high-risk BC patients undergoing surveillance. The *ROBO1*, *CRH*, and *IGF2* gene expression levels were associated with increased risk with a sensitivity of 92.5% and specificity of 73.5%. This panel consists of two genes that were already included in Xpert BC (*CRH* and *IGF2*), while *ROBO1* was included for the first time in a panel of biomarkers of BC. Robo1 protein (ROBO1) is overexpressed in human bladder cancer tissues and paracarcinoma tissues (35). Despite this being a promising panel, the study included a small cohort in a single center, therefore further validation is needed.

x8-gene expression classifier (*ANXA10, LCN2, KRT20, SLC1A6, RPS21, IGF2, MAGEA3*, and *KIFC3*)

This panel was developed in serial steps, from the discovery to the validation phases, performed by the same group and in a multicentric international cohort. Logistic regression analysis was used to generate an 8-gene expression classifier (*ANXA10, IGF2, KIFC3, KRT20, LCN2, MAGEA3, RPS21*, and *SLC1A6*) that showed an area under the curve (AUC) of 0.893 for detecting BC. The 8-gene classifier was also tested in an independent multicentric, international cohort composed of patients in follow-up for BC. The 8-gene classifier performed equally in all BC risk groups, with high and comparable overall sensibility in low-grade and high-grade tumors. The authors reported that their 8-gene expression classifier outperforms the current gold standard (cystoscopy) as well as the previously developed gene expression tests in terms of sensitivity and negative predictive value. It was reported that by using the classifier, around 17% of BC patients under follow-up in their validation cohort could safely skip cystoscopy, while the remaining patients should undergo cystoscopy. The 8-gene classifier is described as safe to guarantee the detection of potential life-threatening tumors in cases of high-risk NMIBC, due to its high sensitivity and NPV (27). *MAGEA3* is a cancer-testis antigen that has been reported to be overexpressed in 15% of the patients with BC by immunohistochemistry. Kaplan-Meier analysis revealed significantly worse 5-year progression-free survival associated with a strong expression of MAGEA3 (36) *LCN2/MMP-9* pathway has been associated with an aggressive phenotype of bladder cancer and the elevated NPV of this protein complex makes them candidate markers of exclusion test for bladder cancer (37). Up-regulated expression of *KIFC3* has been described in many types of cancer and is associated with Epithelial-Mesenchymal-Transition and other important events in tumor development and progression (38) *KRT20* (cytokeratin 20) gene was selected as a surrogate marker (along with 3 uroplakin genes) for luminal MIBC subtype by Olkhov-Mitsel and colleagues in the analysis performed with tumoral tissue (39). To our knowledge, *RPS21* was not previously reported in BC, except from the 8-gene classifier described by Montalbo and colleagues.

X-linked inhibitor of apoptosis (*XIAP*)

The X-linked inhibitor of apoptosis (*XIAP*) is an IAP protein family member that acts as an inhibitor of the caspase/apoptosis pathway. Urinary *XIAP* gene expression was investigated as a biomarker in BC by Srivastava and colleagues (2014) (28). These authors demonstrated a better sensitivity for *XIAP* gene expression

in primary NMIBC cases when compared with voided urine cytology. However, the same study showed that *XIAP* had lower sensitivity than the cytology for recurrent cases. *XIAP* was more sensitive than cytology for the diagnosis of BC patients with early stage. Similarly, a better sensitivity of *XIAP* was detected for both higher- and lower-grade TCC cases as compared to cytology. Expression significantly correlated with tumor grade in this study, but the authors discuss the lack of previous evidence (28).

Discussion

The mRNA urinary panels of BC markers described here have exclusive and shared markers. Among the shared markers, we highlight the presence of *IGF2* in the Xpert, in the 3-marker urinary panel (Panel_3), and the x8-gene expression classifier (Panel_8), The *CRH* gene is part of the Xpert and 3-marker urinary panel. Similarly, the *ANXA10* gene is part of two panels, the Xpert and x8-gene expression classifier. The other genes are mutually exclusive between panels (Figure 1). Although the 3-marker urinary panel and x8-gene expression classifier had mixed cohorts (MIBC and NMIBC tumors), they contain markers that were present in the previously mentioned Xpert test, although two of the genes found in the Xpert Test are exclusive (*ABL1* and *UPK1B*). While *CRH* is shared with the 3-marker urinary panel, *ANXA10* is shared with the 8-gene expression classifier, and *IGF2* is shared with the two panels (Figure 1).

The high rate of recurrence in NMIBC requires follow-up with cystoscopies every three months for two years, and subsequently in more spaced intervals. Moreover, cystoscopy is an invasive procedure, associated with some risks, while urine cytology is inexpensive but has a limited performance for low-grade tumors. These are the main reasons that justify the search for more efficient and less invasive biomarkers for the surveillance of NMIBC. The ideal biological sample would be urine, given its non-invasive nature. Despite several efforts and studies seeking an ideal panel of urinary biomarkers, there is no consensus.

As previously discussed, only three studies included cohorts exclusively of NMIBC, which makes a more comprehensive review difficult. However, we described biomarkers available in the current literature for both types of BC, reinforcing the need for validation and new studies for patients with NMIBC tumors. Urine is the biological sample that is the most reliable and non-invasive source of biomarkers in bladder cancer, as well as in other urological malignancies since the tumor mass is in close and direct contact with urine. This makes urine a liquid biopsy sample but also a source of exfoliated cancer cells. NMIBC has a high risk of recurrence and progression to muscle-invasive disease, requiring follow-up with repeated cystoscopies, which are invasive and expensive. This is the main reason for the extensive research to find new biomarkers and improve those that are already described. Good biomarkers to evaluate the diagnosis and progression of bladder cancer would facilitate follow-up and increase the quality

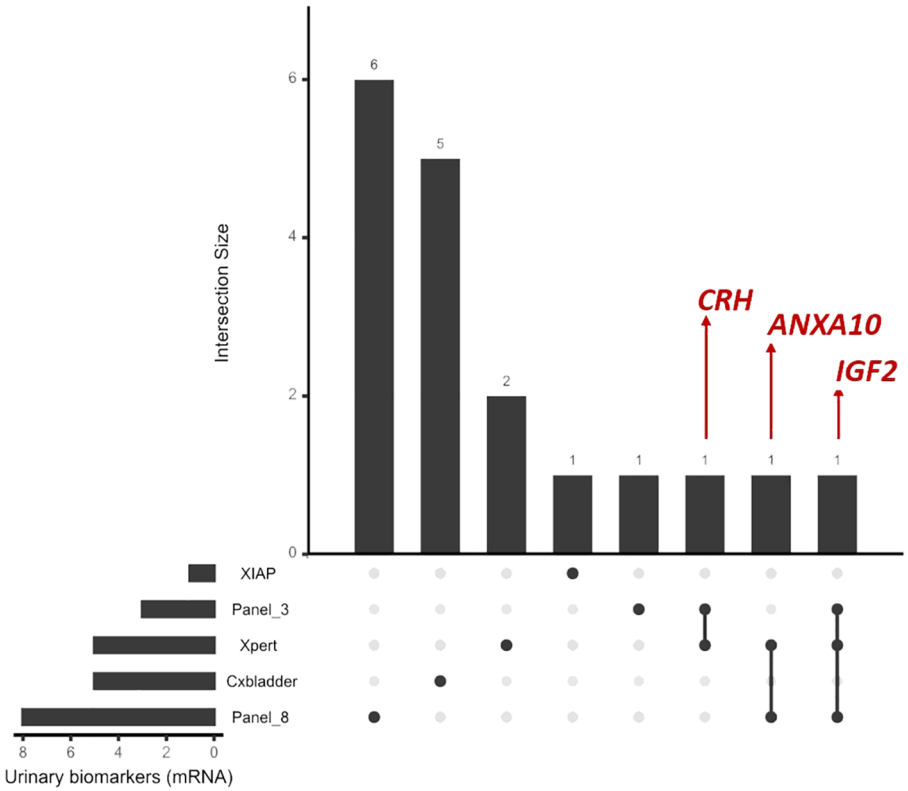


FIGURE 1
Visualization of the intersection between the markers (mRNA) of the tests described using the UpSetR package in R.

of life of BC patients. There are a lot of unexplored possibilities to be studied for the discovery and validation in this field, so this mini-review describes the existing panels of mRNAs that act as biomarkers in bladder cancer, with a special focus on NMIBC.

Author contributions

KC: Conceptualization, Validation, Writing – review & editing, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft. DW: Writing – review & editing. AM: Writing – review & editing. MA: Writing – review & editing. LF: Writing – review & editing. RB: Writing – review & editing. DZ: Conceptualization, Data curation, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. KC was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1441883/full#supplementary-material

References

1. Al-Zalabani AH, Stewart KFJ, Wesselius A, Schols AMWJ, Zeegers MP. Modifiable risk factors for the prevention of bladder cancer: a systematic review of meta-analyses. *Eur J Epidemiol*. (2016) 31:811–51. doi: 10.1007/s10654-016-0138-6

2. Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. *Lancet*. (2009) 374:239–49. doi: 10.1016/S0140-6736(09)60491-8

3. Burger M, Catto JWF, Dalbagni G, Grossman HB, Herr H, Karakiewicz P, et al. Epidemiology and risk factors of urothelial bladder cancer. *Eur Urol*. (2013) 63:234–41. doi: 10.1016/j.eururo.2012.07.033

4. Sanli O, Dobruch J, Knowles MA, Burger M, Alemozaffar M, Nielsen ME, et al. Bladder cancer. *Nat Rev Dis Primers*. (2017) 3:17022. doi: 10.1038/nrdp.2017.22

5. Leal J, Luengo-Fernandez R, Sullivan R, Witjes JA. Economic burden of bladder cancer across the European union. *Eur Urol*. (2016) 69:438–47. doi: 10.1016/j.eururo.2015.10.024

6. Babjuk M, Burger M, Capoun O, Cohen D, Compérat EM, Dominguez Escrig JL, et al. European association of urology guidelines on non–muscle-invasive bladder cancer (Ta, T1, and carcinoma in situ). *Eur Urol*. (2022) 81:75–94. doi: 10.1016/j.eururo.2021.08.010

7. Koguchi D, Matsumoto K, Shiba I, Harano T, Okuda S, Mori K, et al. Diagnostic potential of circulating tumor cells, urinary microRNA, and urinary cell-free DNA for bladder cancer: A review. *Int J Mol Sci*. (2022) 23:9148. doi: 10.3390/ijms23169148

8. Sullivan PS, Chan JB, Levin MR, Rao J. Urine cytology and adjunct markers for detection and surveillance of bladder cancer. *Am J Transl Res*. (2010) 2:412–40.

9. Sun T, Hutchinson L, Tomaszewicz K, Caporelli ML, Meng X, McCauley K, et al. Diagnostic value of a comprehensive, urothelial carcinoma-specific next-generation sequencing panel in urine cytology and bladder tumor specimens. *Cancer Cytopathol*. (2021) 129:537–47. doi: 10.1002/cncy.22410

10. Lokeshwar SD, Lopez M, Sarcan S, Aguilar K, Morera DS, Shaheen DM, et al. Molecular oncology of bladder cancer from inception to modern perspective. *Cancers (Basel)*. (2022) 14:2578. doi: 10.3390/cancers14112578

11. Maas M, Todenhöfer T, Black PC. Urine biomarkers in bladder cancer — current status and future perspectives. *Nat Rev Urol*. (2023) 20:597–614. doi: 10.1038/s41585-023-00773-8

12. Matuszczak M, Kiljańczyk A, Salagierski M. A liquid biopsy in bladder cancer—The current landscape in urinary biomarkers. *Int J Mol Sci*. (2022) 23:8597. doi: 10.3390/ijms23158597

13. Humayun-Zakaria N, Ward DG, Arnold R, Bryan RT. Trends in urine biomarker discovery for urothelial bladder cancer: DNA, RNA, or protein? *Transl Androl Urol*. (2021) 10:2787–808. doi: 10.21037/tau-20-1327

14. Dudzik D, Macioszek S, Struck-Lewicka W, Kordalewska M, Buszewska-Forajta M, Waszczuk-Jankowska M, et al. Perspectives and challenges in extracellular vesicles untargeted metabolomics analysis. *TrAC Trends Analytical Chem*. (2021) 143:116382. doi: 10.1016/j.trac.2021.116382

15. Xi X, Li T, Huang Y, Sun J, Zhu Y, Yang Y, et al. RNA biomarkers: frontier of precision medicine for cancer. *Noncoding RNA*. (2017) 3. doi: 10.3390/ncrna3010009

16. Pichler R, Fritz J, Tulchiner G, Klinglmair G, Soleiman A, Horninger W, et al. Increased accuracy of a novel mRNA -based urine test for bladder cancer surveillance. *BJU Int*. (2018) 121:29–37. doi: 10.1111/bju.14019

17. Wallace E, Higuchi R, Satya M, McCann L, Sin MLY, Bridge JA, et al. Development of a 90-minute integrated noninvasive urinary assay for bladder cancer detection. *J Urol*. (2018) 199:655–62. doi: 10.1016/j.juro.2017.09.141

18. van Valenberg FJP, Hiar AM, Wallace E, Bridge JA, Mayne DJ, Beqaj S, et al. Prospective validation of an mRNA-based urine test for surveillance of patients with bladder cancer. *Eur Urol*. (2019) 75:853–60. doi: 10.1016/j.eururo.2018.11.055

19. Elsawy AA, Awadalla A, Abdullateef M, Ahmed AE, Abol-Enein H. Can repeat biopsy be skipped after initial complete resection of T1 bladder cancer? The role of a novel urinary mRNA biomarker. *Urologic Oncology: Semin Original Investigations*. (2021) 39:437.e11–437.e19. doi: 10.1016/j.urolonc.2021.02.009

20. Kim MH, Yang GE, Jeong MS, Mun JY, Lee SY, Nam JK, et al. VNTR polymorphism in the breakpoint region of ABL1 and susceptibility to bladder cancer. *BMC Med Genomics*. (2021) 14:121. doi: 10.1186/s12920-021-00968-1

21. Chiu YF, Wu CC, Kuo MH, Miao CC, Zheng MY, Chen PY, et al. Critical role of SOX2-IGF2 signaling in aggressiveness of bladder cancer. *Sci Rep*. (2020) 10:8261. doi: 10.1038/s41598-020-65006-z

22. D´Elia C, Pycha A, Folchini DM, Mian C, Hanspeter E, Schwienbacher C, et al. Diagnostic predictive value of Xpert Bladder Cancer Monitor in the follow-up of patients affected by non-muscle invasive bladder cancer. *J Clin Pathol*. (2019) 72:140–4. doi: 10.1136/jclinpath-2018-205393

23. Kavalieris L, O'Sullivan P, Frampton C, Guilford P, Darling D, Jacobson E, et al. Performance characteristics of a multigene urine biomarker test for monitoring for recurrent urothelial carcinoma in a multicenter study. *J Urol*. (2017) 197:1419–26. doi: 10.1016/j.juro.2016.12.010

24. Lotan Y, O'Sullivan P, Raman JD, Shariat SF, Kavalieris L, Frampton C, et al. Clinical comparison of noninvasive urine tests for ruling out recurrent urothelial carcinoma. *Urol Oncol*. (2017) 35:531.e15–531.e22. doi: 10.1016/j.urolonc.2017.03.008

25. O'Sullivan P, Sharples K, Dalphin M, Davidson P, Gilling P, Cambridge L, et al. A multigene urine test for the detection and stratification of bladder cancer in patients presenting with hematuria. *J Urol*. (2012) 188:741–7. doi: 10.1016/j.juro.2012.05.003

26. Shkolyar E, Zhao Q, Mach KE, Teslovich NC, Lee TJ, Cox S, et al. Bladder cancer risk stratification using a urinary mRNA biomarker panel – A path towards cystoscopy triaging. *Urologic Oncology: Semin Original Investigations*. (2021) 39:497.e9–497.e15. doi: 10.1016/j.urolonc.2021.02.011

27. Montalbo R, Lozano JJ, Izquierdo L, Ingelmo-Torres M, BaÑos C, Palou J, et al. Ability of a urine gene expression classifier to reduce the number of follow-up cystoscopies in bladder cancer patients. *Transl Res*. (2019) 208:73–84. doi: 10.1016/j.trsl.2019.02.003

28. Srivastava AK, Singh PK, Singh D, Dalela D, Rath SK, Goel MM, et al. Evaluation of urinary XIAP as a diagnostic biomarker of carcinoma of urinary bladder. *Tumour Biol*. (2014) 35:8243–8. doi: 10.1007/s13277-014-2026-6

29. Holyoake A, O'Sullivan P, Pollock R, Best T, Watanabe J, Kajita Y, et al. Development of a multiplex RNA urine test for the detection and stratification of transitional cell carcinoma of the bladder. *Clin Cancer Res*. (2008) 14:742–9. doi: 10.1158/1078-0432.CCR-07-1672

30. Li KD, Chu CE, Patel M, Meng MV, Morgan TM, Porten SP. Cxbladder Monitor testing to reduce cystoscopy frequency in patients with bladder cancer. *Urol Oncol*. (2023) 41:326.e1–8. doi: 10.1016/j.urolonc.2023.01.009

31. Heo J, Noh BJ, Lee S, Lee HY, Kim Y, Lim J, et al. Phosphorylation of TFCP2L1 by CDK1 is required for stem cell pluripotency and bladder carcinogenesis. *EMBO Mol Med*. (2020) 12:e10880. doi: 10.15252/emmm.201910880

32. Valizadeh S, Taghiyar S, Vahidi S, Abazari O, Akhavan Tafti M, Zavar Reza J. Application of PLK1 and HOXA13 gene expression levels in urine in the diagnosis of non-muscle invasive bladder cancer. *Biochem Genet*. (2024). doi: 10.1007/s10528-024-10735-3

33. Lin H, Zhou Q, Wu W, Ma Y. Midkine is a potential urinary biomarker for non-invasive detection of bladder cancer with microscopic hematuria. *Onco Targets Ther*. (2019) 12:11765–75. doi: 10.2147/OTT

34. Liang PI, Wang YH, Wu TF, Wu WR, Liao AC, Shen KH, et al. IGFBP-5 overexpression as a poor prognostic factor in patients with urothelial carcinomas of upper urinary tracts and urinary bladder. *J Clin Pathol*. (2013) 66:573–82. doi: 10.1136/jclinpath-2012-201278

35. Li Y, Cheng H, Xu W, Tian X, Li X, Zhu C. Expression of Robo protein in bladder cancer tissues and its effect on the growth of cancer cells by blocking Robo protein. *Int J Clin Exp Pathol*. (2015) 8:9932–40.

36. Lausenmeyer EM, Braun K, Breyer J, Gierth M, Denzinger S, Burger M, et al. Strong expression of cancertestis antigens CTAG1B and MAGEA3 is correlated with unfavourable histopathological features and MAGEA3 is associated with worse progression-free survival in urothelial bladder cancer. *Urol Int*. (2019) 102:77–82. doi: 10.1159/000493577

37. Candido S, Di Maso M, Serraino D, McCubrey JA, Bortolus R, Zanin M, et al. Diagnostic value of neutrophil gelatinase-associated lipocalin/matrix metalloproteinase-9 pathway in transitional cell carcinoma of the bladder. *Tumour Biol*. (2016) 37:9855–63. doi: 10.1007/s13277-016-4872-x

38. Lu S, Liu Y, Tian S, He Y, Dong W. KIFC3 regulates progression of hepatocellular carcinoma via EMT and the AKT/mTOR pathway. *Exp Cell Res*. (2023) 426:113564. doi: 10.1016/j.yexcr.2023.113564

39. Olkhov-Mitsel E, Yu Y, Lajkosz K, Liu SK, Vesprini D, Sherman CG, et al. Development of a clinically applicable nanoString-based gene expression classifier for muscle-invasive bladder cancer molecular stratification. *Cancers (Basel)*. (2022) 14. doi: 10.3390/cancers14194911

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 10 October 2024
DOI 10.3389/fonc.2024.1443346

OPEN ACCESS

EDITED BY
Jixin Dong,
University of Nebraska Medical Center, United States

REVIEWED BY
Apurva Patel,
Gujarat Cancer & Research Institute, India
Rui Liu,
Xi'an Jiaotong University, China
Yanlong Shi,
Nanjing Medical University, China

*CORRESPONDENCE
Mateus Nóbrega Aoki
mateus.aoki@fiocruz.br
Dalila Luciola Zanette
dalila.zanette@fiocruz.br

RECEIVED 03 June 2024
ACCEPTED 16 September 2024
PUBLISHED 10 October 2024

CITATION
Kubaski Benevides AP, Marin AM, Wosniaki DK, Oliveira RN, Koerich GM, Kusma BN, Munhoz EC, Zanette DL and Aoki MN (2024) Expression of *HOTAIR* and *PTGS2* as potential biomarkers in chronic myeloid leukemia patients in Brazil. *Front. Oncol.* 14:1443346. doi: 10.3389/fonc.2024.1443346

Expression of *HOTAIR* and *PTGS2* as potential biomarkers in chronic myeloid leukemia patients in Brazil

Ana Paula Kubaski Benevides[1], Anelis Maria Marin[1], Denise K. Wosniaki[1], Rafaela Noga Oliveira[1], Gabriela Marino Koerich[1], Bianca Nichele Kusma[1], Eduardo Cilião Munhoz[2], Dalila Luciola Zanette[1*] and Mateus Nóbrega Aoki[1*]

[1]Laboratory for Applied Science and Technology in Health, Carlos Chagas Institute, Oswaldo Cruz Foundation (Fiocruz), Curitiba, Brazil, [2]Hematology and Oncology Clinic, Erasto Gaertner Hospital, Curitiba, Brazil

Chronic myeloid leukemia (CML) is a clonal myeloproliferative neoplasm in which all the patients has the translocation (*9;22*) that generates de BCR::ABL1 tyrosine kinase. Despite this disease possessing a good biomarker (*BCR::ABL1* transcripts level) for diagnosis and prognosis, many studies has been performed to investigate other molecules, such as the long noncoding RNAs (lncRNAs) and mRNAs, as potential biomarkers with the aim of predicting a change in *BCR::ABL1* levels and as an associated biomarker. A RNAseq was performed comparing 6 CML patients with high *BCR::ABL1* expression with 6 healthy control individuals, comprising the investigation cohort to investigate these molecules. To validate the results obtained by RNAseq, samples of 87 CML patients and 42 healthy controls were used in the validation cohort by RT-qPCR assays. The results showed lower expression of *HOTAIR* and *PTGS2* in CML patients. The *HOTAIR* expression is inversely associated with *BCR::ABL1* expression in imatinib-treated CML patients, and to *PTGS2* showing that CML patients with high *BCR::ABL1* expression showed reduced *PTGS2* expression.

KEYWORDS

CML, biomarkers, lncRNAs, *HOTAIR*, *PTGS2*, RNAseq

Introduction

Chronic myeloid leukemia (CML) is a clonal myeloproliferative neoplasm of the hematopoietic system, characterized by the presence of the Philadelphia chromosome (Ph), a fusion chromosome t (*9,22*) (q34;q11) resulting from the reciprocal translocation between chromosomes 9 and 22. This translocation generates a chimeric gene between the *BCR*

(breakpoint cluster region protein) and *ABL1* (Abelson murine leukemia viral oncogene homolog 1) genes. This genetic translocation produces a constitutively active tyrosine-kinase, leading to the general imbalance found in CML (1). Notably, CML patients are treated with tyrosine-kinase inhibitors (TKIs), especially with the first-generation TKI imatinib, which shows great efficiency and safety (2). However, about one-third of CML imatinib-treated patients switch to second or third-generation TKIs (dasatinib, nilotinib, ponatinib) due to resistance or toxicity (3).

Long noncoding RNAs (lncRNAs) are noncoding transcripts longer than 200 nucleotides that play crucial regulatory roles in gene expression, translation, genome organization, and cell structure, both in physiological and pathological contexts (4–7). These molecules typically exhibit restricted expression patterns and are often highly cell-specific, except for MALAT1 and NEAT1 (8–10). While many lncRNAs are found in the nucleus, a significant fraction is located in the cytoplasm (11). Nuclear lncRNAs are more abundant but less stable than their cytoplasmic counterparts, and their nuclear instability reflects a fine-tuning regulation of transcriptional programs. In the cytoplasm, lncRNAs mostly sequester miRNAs to regulate their activity and levels, which in turn signaling pathways by regulating miRNA target proteins and mRNAs (12).

Physiologically, lncRNAs display dynamic expression during the differentiation of various cell types, such as muscle, immune, and neural cells, underscoring their role in cellular differentiation (10, 13). Due to their important roles, lncRNAs have gained prominence in scientific and translational research as biomarkers for diagnosis, prognosis, and treatment response or resistance, particularly in the oncology field (14–17). The lncRNA *HOTAIR* has been described in solid tumors, such as breast cancer (18, 19), hepatocellular carcinoma (20, 21), glioma (22, 23) and colon cancer (24, 25), where it plays an oncogenic role related to a worse prognosis and reduced chance of complete remission. A functional mechanism of *HOTAIR* may be represented by its participation and interaction with epigenetic regulators such as the PRC2 complex and the Lysine Demethylase 1 (LSD1) in chromatin remodeling and transcription (26). Furthermore, it was functional associated as regulator of Wnt/β-catenin signaling pathway (27), suppressing *TGF-β1* and *ZEB1* (28) and sponging microRNAs such miR-331-3p (29) and miR-126 (30). One of the first HOTAIR expression role in oncohematology was observed as regulating cell cycle progression during myeloid maturation in human promyelocytic leukemia cells (31) and modulating *c-KIT* through expression sponging miR-193a in acute myeloid leukemia (32). More recently, studies also showed this transcript associated with molecular pathways and process in leukemias (33–35).

The prostaglandin-endoperoxide synthase 2 (PTGS2), also known as cyclooxygenase-2 (COX-2), is an enzyme encoded by the *PTGS2* gene. It plays a role in the conversion of arachidonic acid into prostaglandin H_2, which is further transformed into five primary prostaglandins (PGD2, PGE2, PGF2α, PGI2, and TXA2) by cell-specific synthases (36). This enzyme is associated with inflammatory diseases, carcinogenesis, angiogenesis, metastasis (37, 38) and apoptosis resistance (39, 40). Its expression pattern is also linked to carcinogenesis (41). *PTGS2* research in oncohematological diseases advanced to a knowledge of its involvement and therapeutical approach (42–44), highlighting it importance as an active player in bone marrow and blood cellular context and homeostasis.

This study aimed to evaluate lncRNAs as potential biomarkers in CML by comparing the transcriptomes of CML patients with high levels of *BCR::ABL1* transcripts to those of healthy donors. The selected transcripts were then validated in a larger cohort of patients, demonstrating the differential expression of several transcripts that may be correlated with CML prognosis. The hypothesis was that mRNA and lncRNA expression in white blood cells from CML patients with *BCR::ABL1* positive/negative expression could represent an *BCR::ABL1* associated biomarker, linking for molecular pathways and preliminary data for further studies.

Methods

Study cohort — This study was conducted after receiving approval from the Ethics Committee of the Hospital Erasto Gaertner (CAAE 08809419.0.0000.0098) and Hospital do Trabalhador (CAAE 77979417.8.3001.5225). Clinical samples from patients with a confirmed diagnosis of CML and volunteers without a current or previous history of any kind of cancer were recruited from Hospital Erasto Gaertner and Hospital do Trabalhador, in Curitiba, Brazil, from May 2020 to December 2023, following Brazilian guidelines and regulations. The study was described in detail to all participants, who read, discussed, and signed an informed consent form before sample collection. For each patient, 4 mL of peripheral blood was collected in EDTA tubes and processed within 24 hours. CML patients had blood collected several times during treatment and clinical follow-up throughout the project period. The blood was centrifuged to obtain buffy coats and plasma. Personal and clinical data such as age, gender, date of diagnosis, and treatment were also accessed from patients' clinical records. For RNA sequencing (RNAseq), six samples from CML patients with *BCR::ABL1* expression higher than 10% and treated with imatinib were selected. The control group comprised six healthy volunteers paired for gender and age. For the validation cohort, a total of 87 CML clinical samples from patients under treatment with imatinib were selected and subdivided according to the *BCR::ABL1* expression levels as follows: 30 CML samples with *BCR::ABL1* higher than 2% (*BCR::ABL1* high); 28 CML samples with *BCR::ABL1* lower than 2% (*BCR::ABL1* low), and 29 CML samples with undetectable *BCR::ABL1* expression. The control group included 42 individuals without any current or previous history of any history of cancer, paired for gender and age.

RNA extraction and cDNA synthesis: Total RNA was extracted from the buffy coat using QIAmp RNA Blood Mini Kit (QIAGEN, Hilden, Germany). The extracted RNA was quantified using a NanoDrop™ One spectrophotometer (ThermoFisher, Waltham, MA, USA). cDNA was synthesized from 1μg of total RNA using random primers and SuperScript III (ThermoFisher, Waltham, MA, USA), following the manufacturer's instructions. The resulting cDNA was then diluted (1:2) and used for validation with RT-qPCR reactions.

BCR::ABL1 quantification: *BCR::ABL1* expression analysis was performed following an in-house one-step duplex qPCR methodology, as previously described by Marin et al. (2023) (45).

Library construction and sequencing (ribosome RNA depletion): The quantity and purity of the total RNA were assessed using a Bioanalyzer 2100 with an RNA 6000 Nano LabChip Kit (Agilent, CA, USA), and only RNAs with RIN number > 7.0 were used. Approximately 5ug of total RNA was used to deplete ribosomal RNA according to the Ribo-Zero Gold rRNA Removal Kit (Illumina, Cat.MRZG12324, San Diego, USA). After removing ribosomal RNAs, the remaining RNAs were fragmented into short fragments using divalent cations at high temperatures. The cleaved RNA fragments were then reverse transcribed using SuperScriptTM II Reverse Transcriptase (Invitrogen, USA). The cDNA was subsequently used to synthesize U-labeled second-stranded DNAs with *E. coli* DNA polymerase I (NEB, USA), RNase H (NEB, USA), and dUTP Solution (Thermo Fisher, USA). An A-base was added to the blunt ends of each strand to prepare them for ligation to the indexed T-base adapters. Single- or dual-index adapters are ligated to the fragments, and size selection was performed with AMPureXP beads. The ligated products were amplified by PCR, resulting in a final cDNA library with an average insert size of 300 ± 50bp. Finally, 2×150bp paired-end sequencing was performed on an Illumina NovaseqTM 6000, following the vendor's recommended protocol.

Sequence and filtering of clean reads: A total of million 2 x 150 bp paired-end reads were generated and subsequently filtered using Cutadapt (version: cutadapt-1.9, https://cutadapt.readthedocs.io/en/stable/) with using quality controls parameters. The sequence quality was verified using FastQC, including the Q20, Q30, and GC content of the clean data.

Mapping with reference genome: The reads from all samples were aligned to the human reference genome using HISAT2 (version: hisat2-2.0.4, https://daehwankimlab.github.io/hisat2/) package. HISAT2 allows multiple alignments per read and permits a maximum of two mismatches when mapping the reads to the reference. It also builds a database of potential splice junctions and confirms these by comparing the previously unmapped reads against the database of putative junctions.

Quantification of gene abundance: The mapped reads of each sample were assembled using StringTie (version: stringtie-1.3.4d, http://ccb.jhu.edu/software/stringtie/) with default parameters. Then, all transcriptomes from all samples were merged to reconstruct a comprehensive transcriptome using gffcompare software (version: gffcompare-0.9.8, http://ccb.jhu.edu/software/stringtie/gffcompare.shtml). Once the final transcriptome was generated, StringTie and Ballgown (http://www.bioconductor.org/packages/release/bioc/html/ballgown.html) were used to estimate the expression levels of all transcripts. The expression abundance for mRNAs was quantified by calculating the FPKM (fragment per kilobase of transcript per million mapped reads) value.

Differentially expressed genes (DEGs) analysis: The differential expression analysis was performed using DESeq2 (R package) between two different groups, and by edgeR was used for analysis between two samples. Genes with the parameter of false discovery rate (FDR) below 0.05 and absolute fold change ≥ 2 were considered differentially expressed.

Principal component analysis (PCA): PCA was performed using the princomp function in R (http://www.r-project.org/).

GO enrichment analysis: Gene Ontology (GO) is an international standardized gene functional classification system with three ontologies: molecular function, cellular component, and biological process. All DEGs were mapped to GO terms in the Gene Ontology database (http://www.geneontology.org/). The number of genes associated with each term was calculated, and significantly enriched GO terms in DEGs, compared to the genome background, were identified using a hypergeometric test.

Pathway enrichment analysis (KEGG): Genes usually interact with each other to perform specific biological functions. Pathway-based analysis helps to further understand the biological functions of genes. KEGG is a major public pathway-related database used for this purpose.

Gene set enrichment analysis (GSEA): We performed gene set enrichment analysis using the GSEA (v4.1.0) software and MSigDB database to identify whether a set of genes in specific GO terms and KEGG pathways with significant differences in two groups. Briefly, we input the gene expression matrix and rank genes by the Signal2Noise normalization method. Enrichment scores and p-values were calculated using default parameters. Parameters meeting the conditions of |NES|>1, NOM p-val<0.05, and FDR q-val<0.25 were considered significantly different between the two groups.

RT-qPCR: cDNAs were used as templates in RT-qPCR using TaqMan™ specific assays (ThermoFisher Scientific, Inc., MA, USA) and TaqPath™ Pro Amp™ Master Mix (ThermoFisher, Waltham, MA, USA). B-Actin and GAPDH were selected as housekeeping genes using TaqMan assays with the same amplification protocol. All reactions were carried out in a QuantStudio™ 5 Flex (Applied Biosystems, Foster City, CA, USA), with the following thermal cycling conditions: 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute.

Housekeeping selection: the initial analysis excluded all outlier samples using the Outlier calculator from GraphPad Prism (alpha value = 0.05), based on the Ct values of each housekeeping gene from all 129 clinical samples. Subsequently, the best housekeeping gene was selected using BestKeeper software (Pfaffl et al., 2004), using the Ct values for *GAPDH* and *Beta-Actin* from samples not eliminated in the previous step.

Statistical analysis: For gender data chi-square test was used, while for age was used t-student. For qRT-PCR data the relative expression was calculated using the 2-ΔΔCt method, with one healthy control (HC) sample defined as the calibrator, and t-student test was used. All statistical analyses were used as significant p value > 0,05.

Results

Study cohort description

The cohort description is described in Tables 1, 2, where the first one shows the characterization of CML patients subjected to RNASeq, indicating age, gender, BCR::ABL1 percentage, and clinical data for blood cells count and creatinine.

10.3389/fonc.2024.1443346

The Table 2 shows the characterization of all CML and healthy controls study cohort. There were no significant differences in age or gender between the groups. The mean *BCR::ABL1* levels for the *BCR::ABL1* high and *BCR::ABL1* low groups were 82.84% and 0.145%, respectively. The white blood cells and blast count varies especially in high *BCR::ABL1* high expression patients, where from thirty patients eight showed high white blood cells count (ranging from 39,300 to 471,180/mm^3). For *BCR::ABL1* low and absent expression, white blood cells count were in reference range. Just 4 patients of *BCR::ABL1* high expression group had blast detection with percentage ranging from 3% to 5%.

RNA sequencing

To clarify, HC samples were labeled as Ao1, Ao3, Ao4, Ao5, Ao6, and Ao9, while CML samples were labeled as Ao10, Ao13, Ao14, Ao15, Ao16 and Ao17. The principal component analysis (PCA) showed that only one CML sample (Ao17) clustered with the HC samples (Figure 1A). As expected, the Pearson correlation between samples indicated a higher correlation with the CML group and HC group separately, and a lower correlation when comparing CML to HC (Figure 1B).

The RNAseq analysis identified 133 genes with significantly different expression between CML and the HC group. Of these, 67 genes were upregulated, and 66 genes were downregulated in CML compared to healthy subjects (Figures 2A, B). When the analysis was extended to transcripts, 1,621 transcripts showed differential expression between CML and HC groups, from with 953 upregulated and 668 downregulated in CML compared to HC (Figures 2C, D). In a more restricted analysis, 583 lncRNAs were differentially expressed, with 224 upregulated and 359 downregulated in CML compared to healthy subjects (Figures 2E, F).

GO enrichment and KEGG pathways analyses for mRNA are shown in Figure 3, highlighting an enrichment of protein binding in the Molecular Function component (3A) and the *JAK-STAT* signaling pathway (3B).

For GSEA data, it was analyzed compared GO and KEGG showed 76 and 37 statistical different functions/pathways between CML and healthy controls, respectively, exemplifying "cytoplasmatic translation" in GO and "Oxidative Phosphorilation" in KEGG, with size of 87 and 117, respectively, demonstrated in Figure 4.

Validation by RT-qPCR

A critical analysis of differentially expressed mRNA and lncRNA in CML samples was performed with the following criteria: 1) only genes with fold changes higher than 1.2 or lower than 0.5 were considered; 2) Sample positivity percentage (SPP) was used to indicate the number of samples with detectable levels of a given gene/lncRNA within the two groups, CML or HC. A 100% SPP value indicates a gene/lncRNA expressed in all samples from that group, while 0.00% refers to undetectable levels of mRNA/lncRNA in any of the samples in the group. Table 3 describes this

TABLE 1 Demographic and clinical data for 6 CML patients subjected to RNASeq, indicating BCR::ABL1 expression, blood cell count and creatinine quantification.

ID	Age ± SD	Gender	BCR::ABL1%	Total leukocytes*	Lymphocytes*	Neutrophils*	Basophils*	Platelets*	Creatinine**
CML Ao10	50	M	18,72	7.110	1.920	4.479	0	264.000	1,1
CML Ao13	49	F	20,52	5.230	2.249	2.563	52	174.000	0,6
CML Ao14	38	M	34,13	4.900	2.352	2.205	0	41.000	1,1
CML Ao15	63	M	12,62	2.820	1.523	1.128	0	27.000	0,8
CML Ao16	60	M	13,05	4.470	626	3.572	0	110.000	1,2
CML Ao17	69	F	37,35	4.190	712	2.682	84	387.000	1,27

M, Male; F, Female; *Cells/mm^3; **mg/dl.

TABLE 2 Study cohort characterization, describing sample number, mean age, and gender for CML and control groups.

Group	Samples	Mean age ± SD (years)	Gender		BCR::ABL1 expression range	White blood cells per mm3 (mean±StDev)
			Male	Female		
BCR::ABL1 high	30	54.03 ± 15.53	63.33%	36.67%	82,84%-34,1%	51,726±115,326
BCR::ABL1 low	28	50.68 ± 12.76	64.29%	35.71%	1%-0,008%	6,927±2,518
undetectable *BCR::ABL1*	29	62.36 ± 11.82	44.83%	55.17%	ND	6,454±1,801
Healthy control	42	54.65 ± 12.28	40.48%	59.52%	ND	NA

ND, Not detected; NA, Not available.

critical analysis of the mRNAs and lncRNAs selected from RNAseq for the validation cohort. Additionally, three lncRNAs (*MALAT, HOTAIR*, and *Ddx3y*) that did not show a statistical difference in RNAseq results but are correlated with oncohematological diseases (though not yet described in CML) according to scientific literature were included.

From the 129 clinical samples, one sample from the *BCR::ABL1* low group and four samples from the HC group were removed due to outlier criteria and/or housekeeping Ct values higher than 32, indicating poor nucleic acid quality. Following this, housekeeping gene selection with BestKeeper indicated that *GAPDH* was the most stable option; hence, it was used for normalization. For MALAT it was observed a significant higher expression in *BCR::ABL1* lower expression in comparison to healthy controls (Figure 5A), while for *Adgre5*, *Arghef1* and Ssr2 showed no different expression (Figures 5B–D). However as depicted in Figure 5E, the lncRNA validation showed that *HOTAIR* expression was significantly lower in the CML *BCR::ABL1* high and *BCR::ABL1* low groups compared to the *BCR::ABL1* absent group and HC groups.

Only samples from *BCR::ABL1* high and HC groups were used for mRNA validation. Figure 6 shows the significantly lower expression of *PTGS2* (6A), *Rhob* (6B), and *Cd83* (6C) in CML patients compared to HC subjects.

Following this, a critical analysis of *PTGS2* and *HOTAIR* expression in *BCR::ABL1* high-expression patients was performed by subdividing the samples according to their fold change, considering a fold change higher than 1.5 and lower than 0.1. Regarding no finding about *HOTAIR* expression, an interesting data was found in *PTGS2*, where Table 4 lists the 13 patients included in this subgroup; five presented a higher fold change while eight showed a lower fold change. Interestingly, all patients with PTGS2 fold change ≥1.5 experienced successful imatinib treatment, with no need to switch to another class of TKI. Conversely, six (75%) of the eight patients with a *PTGS2* fold change ≤0.1 had unsatisfactory results with imatinib and therefore had to switch their treatment to the second-line TKI, dasatinib. The follow-up for these cohorts ranged from 1,066 to 532 days, a sufficient period to demonstrate whether the patient responded to imatinib.

Discussion

CML was the first disease to have a targeted therapy, with the development of imatinib, a TKI that specifically targets the BCR::ABL1 protein in leukemia cells (1). However, about 10-15% of CML patients develop resistance to imatinib, which led to the development of other generations of TKIs, such as dasatinib and Nilotinib, including clonal evolution, mutations in the *BCR::ABL1* kinase domain, and activation of *BCR::ABL1* independent pathways (46–48).

We describe here a comparative transcriptome analysis of CML patients and HC subjects, highlighting the lower expression of *HOTAIR* and *PTGS2* in CML patients, which may be related to

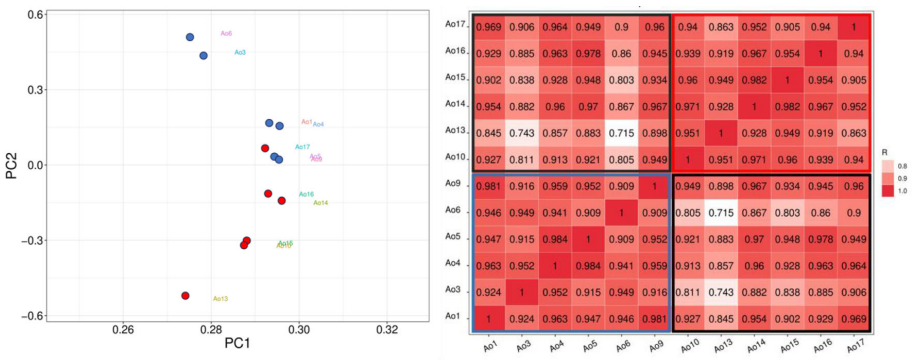


FIGURE 1
Correlation analyses between CML and HC groups. (Left) PCA analysis demonstrates a separation between CML samples (red dots) and healthy subjects (blue dots), except for one CML sample (Ao17), which clustered with the HC samples. (Right) Person correlation between samples, showing CML sample correlations in upper right (red box), and HC sample correlations in lower left (blue box), and CML *vs* HC sample correlations in the box delimited by black line.

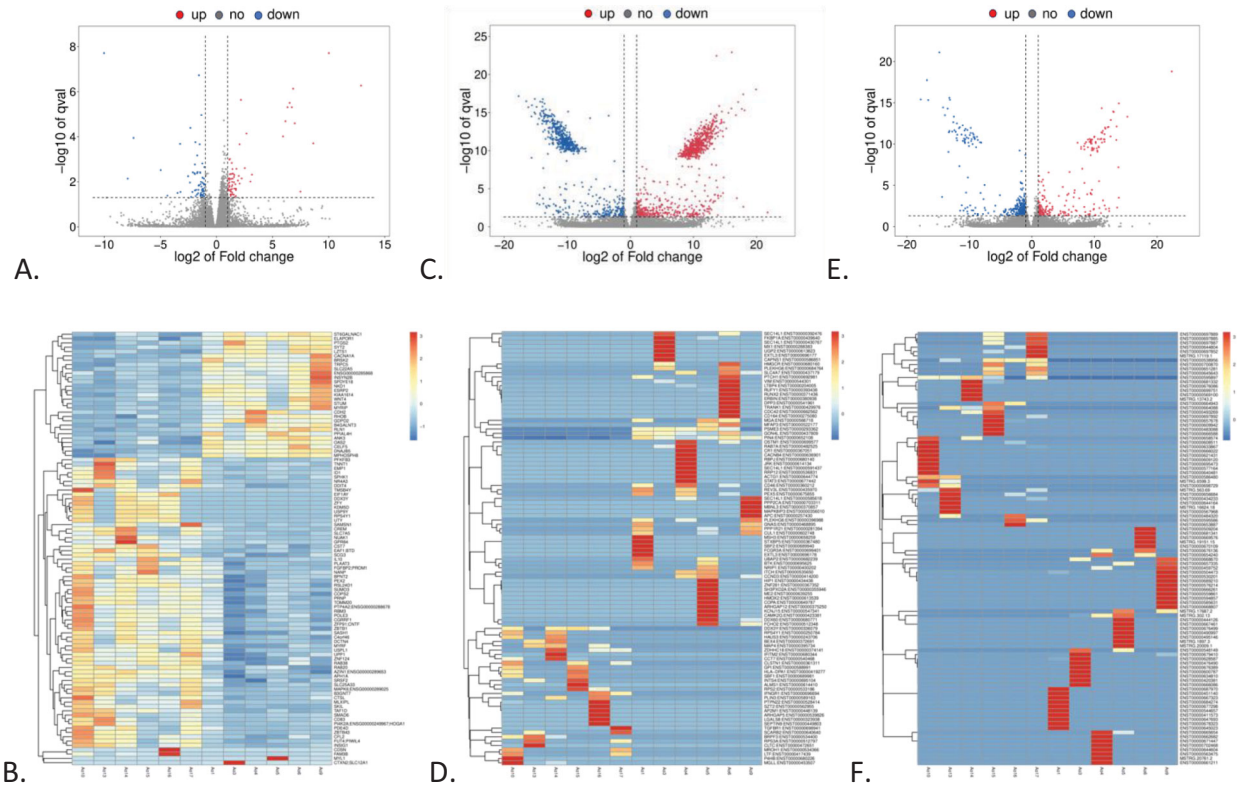


FIGURE 2
RNAseq general data from CML *versus* healthy control groups relative to genes, transcripts, and lncRNA. **(A)** Volcano plot showing differentially expressed genes in CML *vs* HCs (HC). **(B)** Heatmap demonstrating individual expression of the most differentially expressed genes. **(C)** Volcano plot showing differentially expressed transcripts in CML *vs* HC. **(D)** Heatmap demonstrating individual expression of the most differentially expressed transcripts in CML *vs* HC. **(E)** Volcano plot showing differentially expressed lncRNAs in CML *vs* HC. **(F)** Heatmap demonstrating individual expression of the most differentially expressed lncRNAs in CML versus HC.

disease status and imatinib treatment. Despite the pivotal role of the Ph chromosome in CML, some studies have explored the transcriptome of CML patients in search of additional lncRNA and mRNA biomarkers. Giustacchini et al. (2017) conducted a single-cell transcriptomic analysis of CML patients and found an expression signature for normal hematopoietic stem cells (HSCs), BCR::ABL1^{+} and BCR::ABL1^{-}, including transcripts previously implicated in CML pathogenesis. When the authors analyzed the top 245 differentially expressed genes, BCR::ABL1^{+} cells clustered separately from BCR::ABL1^{-} cells, indicating a specific, consistent transcriptomic profile across different CML patients (49). BCR-ABL^{+} stem cells (SCs) at diagnosis did not cluster according to response category, with most BCR::ABL^{-} SCs from poor-responder patients contained within the poor-responder cluster of genes. They also identified a subpopulation of highly quiescent BCR::ABL^{+} SCs, already present at diagnosis and markedly selected during treatment

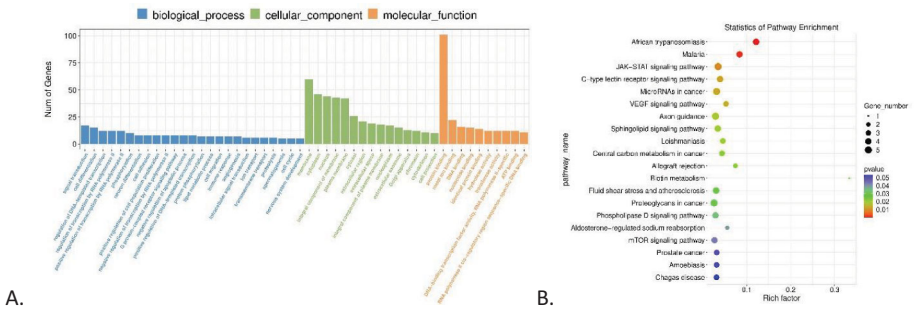


FIGURE 3
GO and KEGG enrichment analysis. **(A)** GO enrichment analysis of differentially expressed genes in CML, showing the number of genes correlated with biological process, cellular component, and molecular function, highlighting signal transduction, membrane, and protein binding, respectively. **(B)** KEGG pathway enrichment analysis of differentially expressed genes in CML, highlighting JAK-STAT signaling pathway and microRNAs in cancer.

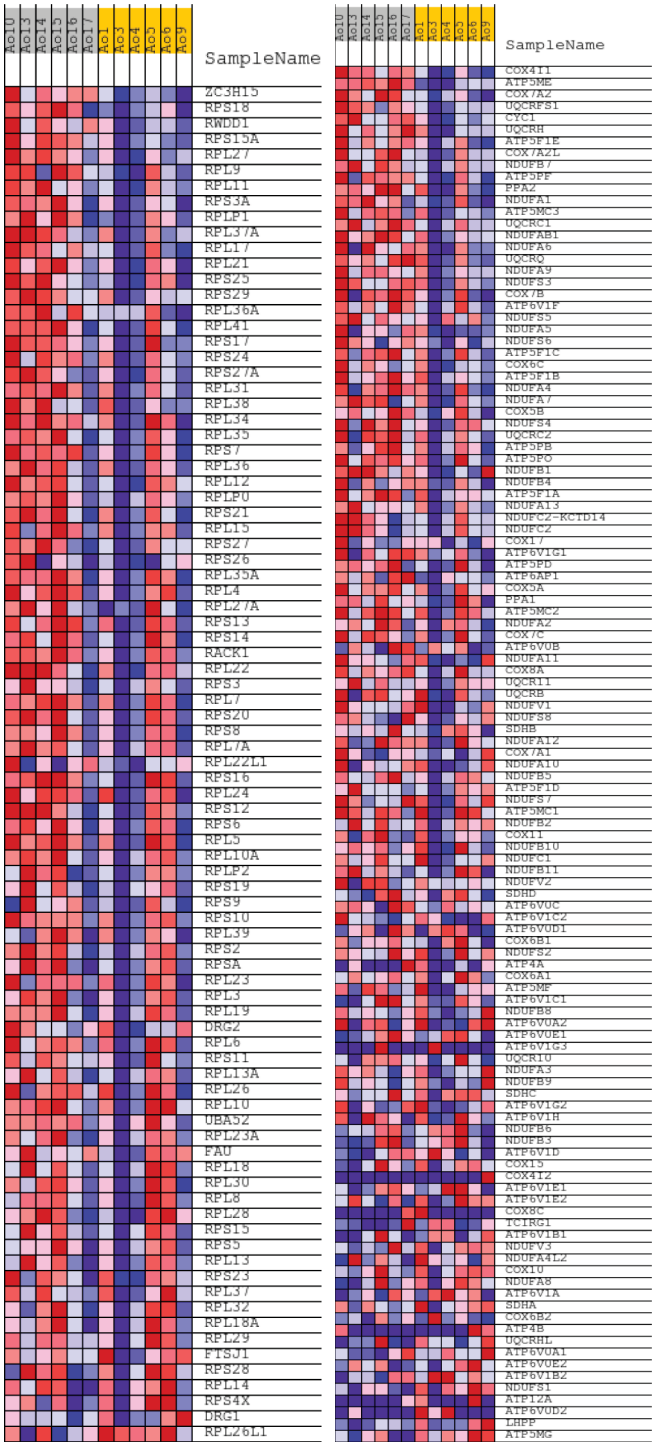


FIGURE 4
Gene Set Enrichment Analysis (GSEA) for RNASeq of CML and healthy controls exemplifying GO and KEGG pathways for cytoplasmatic translation (left) and Oxidative Phosphorylation (right), respectively.

in patients who failed to achieve a deep molecular response (49). Youn et al. (2021) identified 398 differentially expressed genes (258 upregulated and 140 downregulated) in pediatric CML CD34+ cells compared to adult ones. Rho pathways were most significantly downregulated in pediatric CML patients, highlighting VAV2 and ARHGAP27 (50).

PTGS2 codes for an enzyme that belongs to a family responsible for catalyzing the synthesis of prostaglandins from arachidonic acid, which is a stress-inducible protein typically expressed at low levels under normal physiological conditions (51). Its biological function and activation are directly induced by pro-inflammatory cytokines and growth factors correlated with activated intracellular

TABLE 3 mRNAs and lncRNAs selected from RNA sequencing for the validation cohort by RT-qPCR.

Transcript	Name	CML		Healthy controls		Fold-change	Regulation in RNA Seq
		Sample Positivity % (SPP)	FPKM Mean (SD)	Sample Positivity % (SPP)	FPKM Mean (SD)		
mRNA	*PTGS2*	100.00%	24.26 (14.90)	100.00%	104.92 (54.42)	-2.11	Down
	RHOB	100.00%	29.71 (10.48)	100.00%	70.10 (17.48)	-1.24	Down
	CD83	100.00%	19.18 (13.41)	100.00%	5.57 (1.55)	1.78	Up
	RPS4Y1	100.00%	14.37 (10.90)	83.33%	0.13 (0.10)	6.83	Up
lncRNA	*ARHGEF1*	83.33%	2.25 (1.39)	0.00%	ND	NA	Up
	SSR2	33.33%	3.92 (6.43)	0.00%	ND	NA	Up
	ADGRE5	100.00%	5.65 (0.85)	100.00%	16.13 (9.79)	0.35	Down

SPP, Sample Positivity Percentage.
FPKM, Fragments Per Kilobase Million.
Selected transcripts demonstrating the percentage of samples expressing the transcript, FPKM mean and standard deviation (SD), fold change and regulation in CML patients in correlation with HC. ND, not detected; NA, not available.

inflammation-related pathways (52). Several reports indicate that *PTGS2* expression levels are directly related to the carcinogenesis process, particularly due to the inflammatory and, consequently, immunological context. In innate immunity, natural killer cells are inhibited from exerting cytotoxic effects, migrating and secreting interferon γ by PGE2, the main enzymatic product of PTGS2 activity (53, 54). Göbel et al. (2014) showed that *PTGS2* overexpression can initiate and promote carcinogenesis by inhibiting the proliferation of B and T lymphocytes, particularly natural killer T cells, thereby acting as an immunosuppressor (55). PGE2 directly inhibits the proliferation and effector functions of CD4+ and CD8+ T cells and promotes their differentiation in regulatory T cells (Tregs), contributing to the immune response associated with melanoma (56). Several reports have linked the immune system, cancer, and *PTGS2*, demonstrating that its expression induces Tregs, and these cells support cancer-mediated immune suppression (57). PGE2 is directly involved in the complex tumor inflammatory microenvironment, inhibiting inflammatory

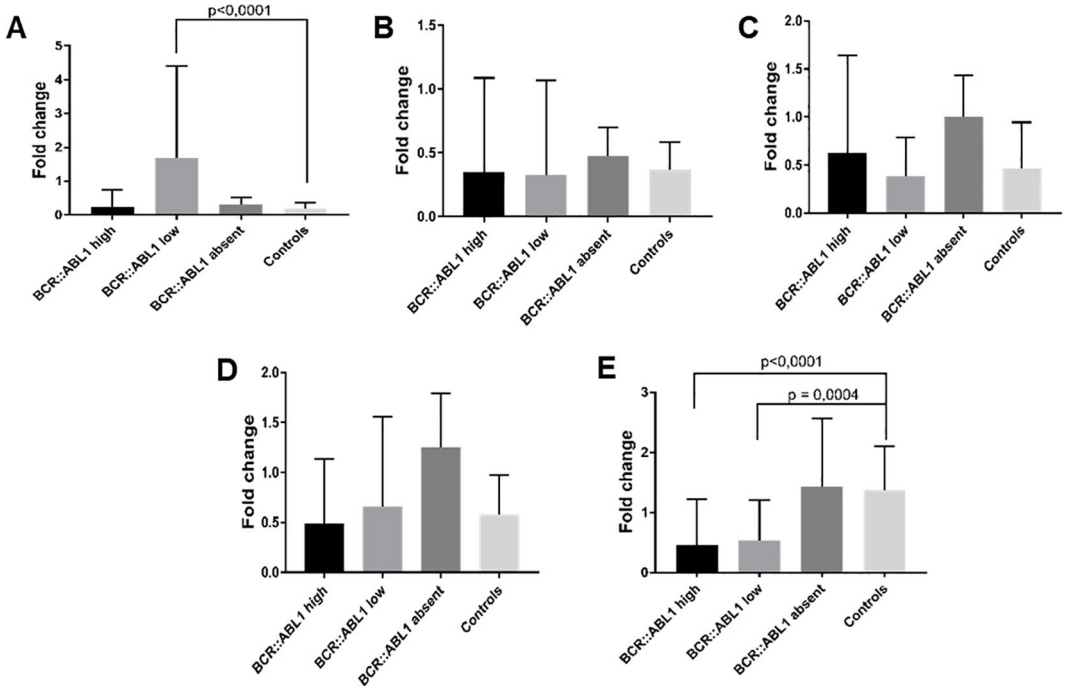


FIGURE 5
lncRNA expression validation by RT-qPCR. **(A)** *MALAT*, **(B)** *Adgre5*, **(C)** *Arghef1*, **(D)** Ssr2 and **(E)** *HOTAIR* in *BCR::ABL1* high, low, and absent expression and healthy controls. The results indicate that *MALAT* was significantly overexpressed in *BCR::ABL1* low expression samples compared to healthy controls, while *HOTAIR* was down-expressed in both *BCR::ABL1* high and low expression samples compared to controls.

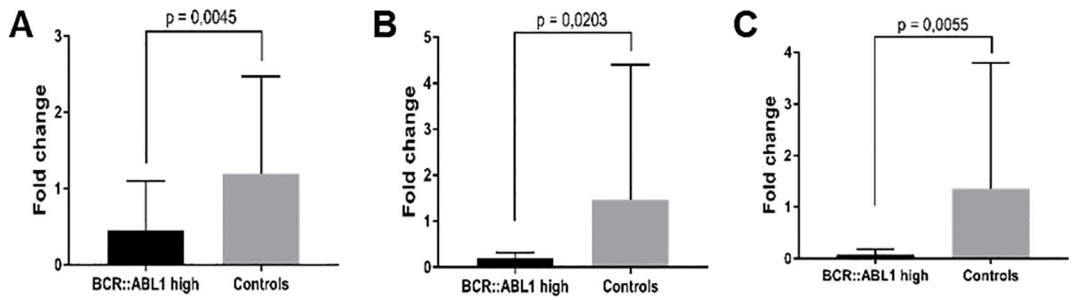


FIGURE 6
mRNA expression validation by RT-qPCR. **(A)** *PTGS2*, **(B)** *Rhob*, and **(C)** *Cd83* in *BCR::ABL1* high expression patients and HC subjects.

chemokines CCL3 and CCL4, preventing the accumulation of activated immune cells. It also reduces IL-10 secretion and shifts the microenvironment in favor of Th1 immune response (58).

There are only a few reports of *PTGS2* in the oncohematological context, with inconclusive results regarding its role in leukemias. A recent study on Mexican adults with acute lymphoblastic leukemia showed that *PTGS2* is highly downregulated compared to controls using transcriptomic microarray analysis (59). In a recent lymphoma report, Qi et al. (2023) demonstrated a consistent correlation between *PTGS2* upregulation and chromosome 17p deletions in human B-cell lymphomas, linking arachidonate metabolism as a new susceptibility factor for lymphoma (60). Regarding CML, the role of *PTGS2* remains unclear primarily due to a lack of clinical studies. However, Giles et al. (2002) showed higher PTGS2 protein levels in the bone marrow of 149 chronic-phase CML patients compared to normal controls, and higher PTGS2 levels were significantly associated with shorter survival. Most PTGS2 studies in CML are *in vitro*, focusing on pharmacologically suppressing *PTGS2* activity (60). Celecoxib, a specific *PTGS2* inhibitor, has been shown to reduce the growth of both K562 and primary CML cells and induce apoptosis in a dose-dependent manner. This effect is accompanied by the downregulation of cyclin D1, cyclin E, and p-Rb expression, the upregulation of *P16(INK4a)* and *P27KIP* expression, and a G1-S phase arrest of the cell cycle (61). Another study using PTGS2-inhibitors showed similar results of reducing growth and inducing apoptosis in a K562 cell line, including indomethacin (62), DuP-697 (63) and nabumetone (64). Together, these results indicate that the *in vitro* pharmacological suppression of *PTGS2* may positively affect CML treatment. However, our study presents a controversial result, showing that CML patients with high *BCR::ABL1* expression showed reduced *PTGS2* expression.

TABLE 4 *PTGS2* fold change in CML patients.

PTGS2 (relative expression)	TKI		Follow-up (days)
1.9643	I		1,066
2.3022	I		1,008
1.6189	I		874
4.2752	I	TFR	889
2.6440	I		890
0.0742	I	D	763
0.0730	I	D	749
0.0219	I	D	637
0.0218	I		742
0.0789	I	D	532
0.0352	I	D	667
0.0938	I		640
0.0745	I	D	532

The patient's treatment during the follow-up, indicating if the TKI was changed from imatinib to dasatinib. I, imatinib; D, dasatinib; TFR, treatment-free remission.

The lncRNAs are non-protein coding RNAs enrolled in several biological processes, including the regulation of gene expression under both normal physiological and disease conditions (65). Recent results have shown that the intracellular location of lncRNAs is crucial; those located in the nucleus can regulate chromatin and transcription, while the ones in the cytoplasm are involved in mRNA stability and translation regulation (66). LncRNAs have been implicated in the initiation and progression of leukemia (67), *NEAT1* (nuclear paraspeckle assembly transcript 1) is correlated with the poor progression in CML (68) and with multidrug resistance in pediatric acute lymphocytic leukemia (69). Other lncRNAs, such as *MALAT* (70), *GAS5* (71), *ANRIL* (72), *TUG* (73) and *PANDAR* (74) have been shown to be downregulated in lymphoid and myeloid leukemias, underscoring their significance in hematological malignancies.

We found that lncRNA *HOTAIR* expression is inversely associated with *BCR::ABL1* expression in imatinib-treated CML patients, as those with absent *BCR::ABL1* showed *HOTAIR* expression levels similar to control subjects. Its lower expression may represent a biomarker associated with *BCR::ABL1* expression, even in early disease or in response to TKI treatment. In oncohematology, *HOTAIR* high expression correlates with the poor prognosis in acute myeloid leukemia, suggesting a relationship with exacerbated proliferation of SCs, a higher number of blasts, and lower disease-free and overall survival (75) Information on *HOTAIR* expression in CML is limited; still, Wang et al. (2017) showed that *HOTAIR* was significantly upregulated in multidrug resistance protein 1 positive patients and in the K562-

imatinib-resistant cells. Knockdown of *HOTAIR* in the K562-imatinib-resistance cells resulted in higher sensitivity to imatinib treatment and attenuation of the PI3K/Akt pathway, suggesting that *HOTAIR* plays a crucial role in acquired resistance to imatinib (76). Two reports showed *HOTAIR* expression in bone marrow of Chinese CML patients, divided into chronic phase, accelerated phase and blast crisis. Contrary to our finding, both reports showed a significant higher expression of HOTAIR in CML patients in comparison to healthy controls, however both also demonstrated that accelerated phase and blast crisis samples showed HOTAIR expression significantly higher than chronic phase (77, 78). It was also demonstrated that *HOTAIR* had higher expression in K562 and KCL-22 cells compared to bone marrow mononuclear cells, accompanied by lower expression of PTEN, and functional results indicated that *HOTAIR* downregulation reduces the proliferation, colony formation, and cell cycle progression while increasing the apoptosis rate of CML cells (78).

So here we described the transcriptomic profile of *BCR::ABL1* Brazilian CML patients and demonstrated differential expression of mRNA and lncRNA compared to healthy controls subjects. Selected transcripts were validated by RT-qPCR in a separate cohort, confirming the downregulation of *HOTAIR* and *PTGS2* in *BCR::ABL1* high-expression Brazilian CML patients. Our results suggest a relationship between imatinib response and downregulation of *PTGS2* and *HOTAIR*, as CML patients with higher expression of these transcripts successfully responded to imatinib. Despite the patient stratification according to *BCR::ABL1* expression levels, a weakness of this study is was the use of white blood cells, not analyze specific cell subsets, such as CD34+ cells, and instead examined a heterogeneous population of peripheral blood leukocytes. This results are limited and described to this scenario, however comprises data that *HOTAIR* and *PTGS2* expression is not consolidated in CML studies, variating according study design and intended goals. However, as a strength considering the results were validated in a larger cohort with extensive clinical data and long follow-up, supporting the value of our findings. It was observed that HOTAIR and PTGS2 may represent an important role in CML progression, as biomarker of TKI efficiency and resistance, however more studies are required, especially for basic cellular and molecular contexts. Additionally, our group is following and generating data on specific cell type and follow-up recruitments in CML patient, aiming describing more specific and detailed findings.

Data availability statement

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

Ethics statement

The studies involving humans were approved by Ethics Committee of the Hospital Erasto Gaertner and Hospital do Trabalhador. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

AK: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. AM: Conceptualization, Formal analysis, Methodology, Supervision, Validation, Visualization, Writing – review & editing. DW: Investigation, Methodology, Validation, Writing – review & editing. RO: Formal analysis, Investigation, Methodology, Writing – review & editing. GK: Investigation, Methodology, Validation, Writing – review & editing. BK: Formal analysis, Investigation, Validation, Writing – review & editing. EM: Formal analysis, Investigation, Writing – review & editing. DZ: Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Writing – review & editing, Writing – original draft. MA: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The project was supported by Carlos Chagas Institute, Fiocruz PR, through the PEP grant.

Acknowledgments

The authors would like to thank Carlos Chagas Institute and Erasto Gaertner Hospital's support.

Conflict of interest

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

Publisher's note

References

1. Apperley JF. Chronic myeloid leukaemia. *Lancet*. (2015) 385:1447–59. doi: 10.1016/S0140-6736(13)62120-0

2. An X, Tiwari AK, Sun Y, Ding PR, Ashby CR, Chen ZS. BCR-ABL tyrosine kinase inhibitors in the treatment of Philadelphia chromosome positive chronic myeloid leukemia: A review. *Leuk Res*. (2010) 34:1255–68. doi: 10.1016/j.leukres.2010.04.016

3. Gibbons DL, Pricl S, Kantarjian H, Cortes J, Quintás-Cardama A. The rise and fall of gatekeeper mutations? The *BCR-ABL1* T315I paradigm. *Cancer*. (2012) 118:293–9. doi: 10.1002/cncr.v118.2

4. Halley P, Kadakkuzha BM, Faghihi MA, Magistri M, Zeier Z, Khorkova O, et al. Regulation of the apolipoprotein gene cluster by a long noncoding RNA. *Cell Rep*. (2014) 6:222–30. doi: 10.1016/j.celrep.2013.12.015

5. Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*. (2011) 477:295–300. doi: 10.1038/nature10398

6. Bolha L, Ravnik-Glavač M, Glavač D. Long noncoding RNAs as biomarkers in cancer. *Dis Markers*. (2017) 2017:1–14. doi: 10.1155/2017/7243968

7. Hacisuleyman E, Shukla CJ, Weiner CL, Rinn JL. Function and evolution of local repeats in the Firre locus. *Nat Commun*. (2016) 7:11021. doi: 10.1038/ncomms11021

8. Gloss BS, Dinger ME. The specificity of long noncoding RNA expression. *Biochim Biophys Acta (BBA) - Gene Regul Mechanisms*. (2016) 1859:16–22. doi: 10.1016/j.bbagrm.2015.08.005

9. Eißmann M, Gutschner T, Hämmerle M, Günther S, Caudron-Herger M, Groß M, et al. Loss of the abundant nuclear non-coding RNA *MALAT1* is compatible with life and development. *RNA Biol*. (2012) 9:1076–87. doi: 10.4161/rna.21089

10. Flynn RA, Chang HY. Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell*. (2014) 14:752–61. doi: 10.1016/j.stem.2014.05.014

11. Mas-Ponte D, Carlevaro-Fita J, Palumbo E, Hermoso Pulido T, Guigo R, Johnson R. LncATLAS database for subcellular localization of long noncoding RNAs. *RNA*. (2017) 23:1080–7. doi: 10.1261/rna.060814.117

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

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

14. Sideris N, Dama P, Bayraktar S, Stiff T, Castellano L. LncRNAs in breast cancer: a link to future approaches. *Cancer Gene Ther*. (2022) 29:1866–77. doi: 10.1038/s41417-022-00487-w

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

16. Ahmad M, Weiswald LB, Poulain L, Denoyelle C, Meryet-Figuiere M. Involvement of lncRNAs in cancer cells migration, invasion and metastasis: cytoskeleton and ECM crosstalk. *J Exp Clin Cancer Res*. (2023) 42:173. doi: 10.1186/s13046-023-02741-x

17. Bhan A, Soleimani M, Mandal SS. Long noncoding RNA and cancer: A new paradigm. *Cancer Res*. (2017) 77:3965–81. doi: 10.1158/0008-5472.CAN-16-2634

18. Schorderet P, Duboule D. Structural and functional differences in the long non-coding RNA hotair in mouse and human. *PloS Genet*. (2011) 7:e1002071. doi: 10.1371/journal.pgen.1002071

19. Raju GSR, Pavitra E, Bandaru SS, Varaprasad GL, Nagaraju GP, Malla RR, et al. HOTAIR: a potential metastatic, drug-resistant and prognostic regulator of breast cancer. *Mol Cancer*. (2023) 22:65. doi: 10.1186/s12943-023-01765-3

20. Lumkul L, Jantaree P, Jaisamak K, Wongkummool W, Lapisatepun W, Orrapin S, et al. Combinatorial gene expression profiling of serum HULC, HOTAIR, and UCA1 lncRNAs to differentiate hepatocellular carcinoma from liver diseases: A systematic review and meta-analysis. *Int J Mol Sci*. (2024) 25. doi: 10.3390/ijms25021258

21. Yang Z, Zhou L, Wu LM, Lai MC, Xie HY, Zhang F, et al. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann Surg Oncol*. (2011) 18:1243–50. doi: 10.1245/s10434-011-1581-y

22. Ahmad F, Sudesh R, Ahmed AT, Haque S. Roles of HOTAIR long non-coding RNA in gliomas and other CNS disorders. *Cell Mol Neurobiol*. (2024) 44:23. doi: 10.1007/s10571-024-01455-8

23. Zhang JX, Han L, Bao ZS, Wang YY, Chen LY, Yan W, et al. HOTAIR, a cell cycle–associated long noncoding RNA and a strong predictor of survival, is preferentially expressed in classical and mesenchymal glioma. *Neuro Oncol*. (2013) 15:1595–603. doi: 10.1093/neuonc/not131

24. Svoboda M, Slyskova J, Schneiderova M, Makovicky P, Bielik L, Levy M, et al. HOTAIR long non-coding RNA is a negative prognostic factor not only in primary tumors, but also in the blood of colorectal cancer patients. *Carcinogenesis*. (2014) 35:1510–5. doi: 10.1093/carcin/bgu055

25. Tufail M. HOTAIR in colorectal cancer: structure, function, and therapeutic potential. *Med Oncol*. (2023) 40:259. doi: 10.1007/s12032-023-02131-5

26. Bhan A, Mandal SS. LncRNA HOTAIR: A master regulator of chromatin dynamics and cancer. *Biochim Biophys Acta (BBA) - Rev Cancer*. (2015) 1856:151–64. doi: 10.1016/j.bbcan.2015.07.001

27. Hakami MA, Hazazi A, Abdulaziz O, Almasoudi HH, Alhazmi AYM, Alkhalil SS, et al. HOTAIR: A key regulator of the Wnt/β-catenin signaling cascade in cancer progression and treatment. *Pathol Res Pract*. (2024) 253:154957. doi: 10.1016/j.prp.2023.154957

28. Zhou Y, Zhang Y, Shao Y, Yue X, Chu Y, Yang C, et al. LncRNA HOTAIR down-expression inhibits the invasion and tumorigenicity of epithelial ovarian cancer cells by suppressing TGF-β1 and ZEB1. *Discover Oncol*. (2023) 14:228. doi: 10.1007/s12672-023-00846-5

29. Buranjiang G, Abuduwanke A, Li X, Abulizi G. LncRNA HOTAIR enhances RCC2 to accelerate cervical cancer progression by sponging miR-331-3p. *Clin Transl Oncol*. (2023) 25:1650–60. doi: 10.1007/s12094-022-03059-4

30. Sun Y, Hu ZQ. LncRNA HOTAIR aggravates myocardial ischemia-reperfusion injury by sponging microRNA-126 to upregulate SRSF1. *Eur Rev Med Pharmacol Sci*. (2020) 24:9046–54. doi: 10.26355/eurrev_202009_22850

31. Zhang X, Weissman SM, Newburger PE. Long intergenic non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid maturation in NB4 human promyelocytic leukemia cells. *RNA Biol*. (2014) 11:777–87. doi: 10.4161/rna.28828

32. Xing Cy, Hu X, Xie Fy, Yu Zj, Li Hy, Bin-Zhou, et al. Long non-coding RNA HOTAIR modulates c-KIT expression through sponging miR-193a in acute myeloid leukemia. *FEBS Lett*. (2015) 589:1981–7. doi: 10.1016/j.febslet.2015.04.061

33. Shagerdi Esmaeli N, Asadi S, Bashash D, Salari S, Hamidpour M. Involvement Value of FLT-3, c-Myc, STAT3, p27, and HOTAIR Gene Expression in Acute Myeloid Leukemia Patients: A Molecular Perspective to a Novel Leukemogenesis Mechanism. *Int J Hematol Oncol Stem Cell Res*. (2023) 17:145–55. doi: 10.18502/ijhoscr.v17i3.13304

34. Liu JM, Li M, Luo W, Sun HB. Curcumin attenuates Adriamycin-resistance of acute myeloid leukemia by inhibiting the lncRNA HOTAIR/miR-20a-5p/WT1 axis. *Lab Invest*. (2021) 101:1308–17. doi: 10.1038/s41374-021-00640-3

35. Zhou W, Xu S, Chen X, Wang C. HOTAIR suppresses PTEN via DNMT3b and confers drug resistance in acute myeloid leukemia. *Hematology*. (2021) 26:170–8. doi: 10.1080/16078454.2021.1880733

36. Liu H, Huang J, Peng J, Wu X, Zhang Y, Zhu W, et al. Upregulation of the inwardly rectifying potassium channel Kir2.1 (KCNJ2) modulates multidrug resistance of small-cell lung cancer under the regulation of miR-7 and the Ras/MAPK pathway. *Mol Cancer*. (2015) 14:59. doi: 10.1186/s12943-015-0298-0

37. Tsujii M, Kawano S, Tsuji S, Sawaoka H, Hori M, DuBois RN. Cyclooxygenase regulates angiogenesis induced by colon cancer cells. *Cell*. (1998) 93:705–16. doi: 10.1016/S0092-8674(00)81433-6

38. Tsujii M, DuBois RN. Alterations in cellular adhesion and apoptosis in epithelial cells overexpressing prostaglandin endoperoxide synthase 2. *Cell*. (1995) 83:493–501. doi: 10.1016/0092-8674(95)90127-2

39. Nzeako UC, Guicciardi ME, Yoon JH, Bronk SF, Gores GJ. COX-2 inhibits Fas-mediated apoptosis in cholangiocarcinoma cells. *Hepatology*. (2002) 35:552–9. doi: 10.1053/jhep.2002.31774

40. Sharma S, Stolina M, Yang SC, Baratelli F, Lin JF, Atianzar K, et al. Tumor cyclooxygenase 2-dependent suppression of dendritic cell function. *Clin Cancer Res*. (2003) 9:961–8.

41. Hashemi Goradel N, Najafi M, Salehi E, Farhood B, Mortezaee K. Cyclooxygenase-2 in cancer: A review. *J Cell Physiol*. (2019) 234:5683–99. doi: 10.1002/jcp.v234.5

42. Kelkka T, Tyster M, Lundgren S, Feng X, Kerr C, Hosokawa K, et al. Anti-COX-2 autoantibody is a novel biomarker of immune aplastic anemia. *Leukemia*. (2022) 36:2317–27. doi: 10.1038/s41375-022-01654-6

43. Subkorn P, Norkaew C, Deesrisak K, Tanyong D. Punicalagin, a pomegranate compound, induces apoptosis and autophagy in acute leukemia. *PeerJ*. (2021) 9:e12303. doi: 10.7717/peerj.12303

44. Salimi A, Aghvami M, Azami Movahed M, Zarei MH, Eshghi P, Zarghi A, et al. Evaluation of Cytotoxic Potentials of Novel Cyclooxygenase-2 Inhibitor against ALL Lymphocytes and Normal Lymphocytes and Its Anticancer Effect through Mitochondrial Pathway. *Cancer Invest*. (2020) 38:463–75. doi: 10.1080/07357907.2020.1808898

45. Marin AM, Wosniaki DK, Sanchuki HBS, Munhoz EC, Nardin JM, Soares GS, et al. Molecular BCR::ABL1 quantification and ABL1 mutation detection as essential tools for the clinical management of chronic myeloid leukemia patients: results from a Brazilian single-center study. *Int J Mol Sci*. (2023) 24:10118. doi: 10.3390/ijms241210118

46. Hochhaus A, Baccarani M, Silver RT, Schiffer C, Apperley JF, Cervantes F, et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. *Leukemia*. (2020) 34:966–84. doi: 10.1038/s41375-020-0776-2

47. Redaelli S, Mologni L, Rostagno R, Piazza R, Magistroni V, Ceccon M, et al. Three novel patient-derived BCR/ABL mutants show different sensitivity to second and

third generation tyrosine kinase inhibitors. *Am J Hematol*. (2012) 87. doi: 10.1002/ajh.v87.11

48. Breccia M, Alimena G. Second-generation tyrosine kinase inhibitors (TKI) as salvage therapy for resistant or intolerant patients to prior TKIs. *Mediterr J Hematol Infect Dis*. (2014) 6:e2014003. doi: 10.4084/mjhid.2014.003

49. Giustacchini A, Thongjuea S, Barkas N, Woll PS, Povinelli BJ, Booth CAG, et al. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat Med*. (2017) 23:692–702. doi: 10.1038/nm.4336

50. Youn M, Smith SM, Lee AG, Chae HD, Spiteri E, Erdmann J, et al. Comparison of the transcriptomic signatures in pediatric and adult CML. *Cancers (Basel)*. (2021) 13:6263. doi: 10.3390/cancers13246263

51. Martín-Vázquez E, Cobo-Vuilleumier N, López-Noriega L, Lorenzo PI, Gauthier BR. The PTGS2/COX2-PGE $_2$ signaling cascade in inflammation: Pro or anti? A case study with type 1 diabetes mellitus. *Int J Biol Sci*. (2023) 19:4157–65. doi: 10.7150/ijbs.86492

52. Klein T, Shephard P, Kleinert H, Kömhoff M. Regulation of cyclooxygenase-2 expression by cyclic AMP. *Biochim Biophys Acta (BBA) - Mol Cell Res*. (2007) 1773:1605–18. doi: 10.1016/j.bbamcr.2007.09.001

53. Gualde N, Harizi H. Prostanoids and their receptors that modulate dendritic cell-mediated immunity. *Immunol Cell Biol*. (2004) 82:353–60. doi: 10.1111/j.0818-9641.2004.01251.x

54. Martinet L, Jean C, Dietrich G, Fournié JJ, Poupot R. PGE2 inhibits natural killer and γδ T cell cytotoxicity triggered by NKR and TCR through a cAMP-mediated PKA type I-dependent signaling. *Biochem Pharmacol*. (2010) 80:838–45. doi: 10.1016/j.bcp.2010.05.002

55. Göbel C, Breitenbuecher F, Kalkavan H, Hähnel PS, Kasper S, Hoffarth S, et al. Functional expression cloning identifies COX-2 as a suppressor of antigen-specific cancer immunity. *Cell Death Dis*. (2014) 5:e1568–8. doi: 10.1038/cddis.2014.531

56. Mougiakakos D, Johansson CC, Trocme E, All-Ericsson C, Economou MA, Larsson O, et al. Intratumoral forkhead box P3-positive regulatory T cells predict poor survival in cyclooxygenase-2–positive uveal melanoma. *Cancer*. (2010) 116:2224–33. doi: 10.1002/cncr.v116:9

57. Mahic M, Yaqub S, Johansson CC, Taske'n K, Aandahl EM. FOXP3+CD4+CD25+ Adaptive regulatory T cells express cyclooxygenase-2 and suppress effector T cells by a prostaglandin E2-dependent mechanism. *J Immunol*. (2006) 177:246–54. doi: 10.4049/jimmunol.177.1.246

58. Chen EP, Markosyan N, Connolly E, Lawson JA, Li X, Grant GR, et al. Myeloid Cell COX-2 deletion reduces mammary tumor growth through enhanced cytotoxic T-lymphocyte function. *Carcinogenesis*. (2014) 35:1788–97. doi: 10.1093/carcin/bgu053

59. Cruz-Miranda GM, Olarte-Carrillo I, Bárcenas-López DA, Martínez-Tovar A, Ramírez-Bello J, Ramos-Peñafiel CO, et al. Transcriptome analysis in mexican adults with acute lymphoblastic leukemia. *Int J Mol Sci*. (2024) 25:1750. doi: 10.3390/ijms25031750

60. Qi L, Pan X, Chen X, Liu P, Chen M, Zhang Q, et al. COX-2/PGE2 upregulation contributes to the chromosome 17p-deleted lymphoma. *Oncogenesis*. (2023) 12:5. doi: 10.1038/s41389-023-00451-9

61. Zhang G, Liu D, Dai C, Li R. Antitumor effects of celecoxib on K562 leukemia cells are mediated by cell-cycle arrest, caspase-3 activation, and downregulation of Cox-2 expression and are synergistic with hydroxyurea or imatinib. *Am J Hematol*. (2006) 81:242–55. doi: 10.1002/ajh.20542

62. Zhang Gs, Tu Cq, Zhang Gy, Zhou Gb, Zheng W. Indomethacin induces apoptosis and inhibits proliferation in chronic myeloid leukemia cells. *Leuk Res*. (2000) 24:385–92. doi: 10.1016/S0145-2126(99)00198-8

63. Peng HL, Zhang GS, Liu JH, Gong FJ, Li RJ. Dup-697, a specific COX-2 inhibitor, suppresses growth and induces apoptosis on K562 leukemia cells by cell-cycle arrest and caspase-8 activation. *Ann Hematol*. (2008) 87:121–9. doi: 10.1007/s00277-007-0385-4

64. Vural F, Özcan MA, Özsan GH, Ateş H, Demirkan F, Pişkin >Ö, et al. Cyclo-oxygenase 2 inhibitor, nabumetone, inhibits proliferation in chronic myeloid leukemia cell lines. *Leuk Lymphoma*. (2005) 46:753–6. doi: 10.1080/10428190400027860

65. Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics*. (2013) 193:651–69. doi: 10.1534/genetics.112.146704

66. Quinn JJ, Zhang QC, Georgiev P, Ilik IA, Akhtar A, Chang HY. Rapid evolutionary turnover underlies conserved lncRNA–genome interactions. *Genes Dev*. (2016) 30:191–207. doi: 10.1101/gad.272187.115

67. Gasic V, Karan-Djurasevic T, Pavlovic D, Zukic B, Pavlovic S, Tosic N. Diagnostic and therapeutic implications of long non-coding RNAs in leukemia. *Life*. (2022) 12:1770. doi: 10.3390/life12111770

68. Rostami M, Kharajo RS, Parsa-kondelaji M, Ayatollahi H, Sheikhi M, Keramati MR. Altered expression of NEAT1 variants and P53, PTEN, and BCL-2 genes in patients with acute myeloid leukemia. *Leuk Res*. (2022) 115:106807. doi: 10.1016/j.leukres.2022.106807

69. Pouyanrad S, Rahgozar S, Ghodousi ES. Dysregulation of miR-335-3p, targeted by NEAT1 and MALAT1 long non-coding RNAs, is associated with poor prognosis in childhood acute lymphoblastic leukemia. *Gene*. (2019) 692:35–43. doi: 10.1016/j.gene.2019.01.003

70. Huang JL, Liu W, Tian LH, Chai TT, Liu Y, Zhang F, et al. Upregulation of long non-coding RNA MALAT-1 confers poor prognosis and influences cell proliferation and apoptosis in acute monocytic leukemia. *Oncol Rep*. (2017) 38:1353–62. doi: 10.3892/or.2017.5802

71. Gasic V, Stankovic B, Zukic B, Janic D, Dokmanovic L, Krstovski N, et al. Expression pattern of long non-coding RNA growth arrest-specific 5 in the remission induction therapy in childhood acute lymphoblastic leukemia. *J Med Biochem*. (2019) 38:292–8. doi: 10.2478/jomb-2018-0038

72. Tang X, Ren H, Guo M, Qian J, Yang Y, Gu C. Review on circular RNAs and new insights into their roles in cancer. *Comput Struct Biotechnol J*. (2021) 19:910–28. doi: 10.1016/j.csbj.2021.01.018

73. Qin J, Bao H, Li H. Correlation of long non-coding RNA taurine-upregulated gene 1 with disease conditions and prognosis, as well as its effect on cell activities in acute myeloid leukemia. *Cancer Biomarkers*. (2018) 23:569–77. doi: 10.3233/CBM-181834

74. Yang L, Zhou JD, Zhang TJ, Ma JC, Xiao GF, Chen Q, et al. Overexpression of lncRNA em PANDAR em predicts adverse prognosis in acute myeloid leukemia. *Cancer Manag Res*. (2018) 10:4999–5007. doi: 10.2147/CMAR.S180150

75. Hao S, Shao Z. HOTAIR is upregulated in acute myeloid leukemia and that indicates a poor prognosis. *Int J Clin Exp Pathol*. (2015) 8:7223–8.

76. Wang X, Liu S, Cao L, Zhang T, Yue D, Wang L, et al. miR-29a-3p suppresses cell proliferation and migration by downregulating IGF1R in hepatocellular carcinoma. *Oncotarget*. (2017) 8:86592–603. doi: 10.18632/oncotarget.21246

77. Li Z, Luo J. Epigenetic regulation of HOTAIR in advanced chronic myeloid leukemia. *Cancer Manag Res*. (2018) 10:5349–62. doi: 10.2147/CMAR.S166859

78. Song H, Chen L, Liu W, Xu X, Zhou Y, Zhu J, et al. Depleting long noncoding RNA HOTAIR attenuates chronic myelocytic leukemia progression by binding to DNA methyltransferase 1 and inhibiting PTEN gene promoter methylation. *Cell Death Dis*. (2021) 12:440. doi: 10.1038/s41419-021-03637-4

frontiers | Frontiers in Immunology

TYPE Original Research
PUBLISHED 13 November 2024
DOI 10.3389/fimmu.2024.1477427

OPEN ACCESS

EDITED BY
Mariana Aris,
FUCA, Argentina

REVIEWED BY
Huaichao Luo,
Sichuan Cancer Hospital, China
Alexander Zaslavsky,
University of Michigan, United States

*CORRESPONDENCE
Ying Liu
zlyyliuying1664@zzu.edu.cn

RECEIVED 07 August 2024
ACCEPTED 25 October 2024
PUBLISHED 13 November 2024

CITATION
Li Q, Zhang C, Ren Y, Qiao L, Xu S, Li K and Liu Y (2024) A novel platelets-related gene signature for predicting prognosis, immune features and drug sensitivity in gastric cancer. *Front. Immunol.* 15:1477427. doi: 10.3389/fimmu.2024.1477427

A novel platelets-related gene signature for predicting prognosis, immune features and drug sensitivity in gastric cancer

Qun Li, Cheng Zhang, Yulin Ren, Lei Qiao, Shuning Xu, Ke Li and Ying Liu*

Department of Medical Oncology, The Affiliated Cancer Hospital of Zhengzhou University and Henan Cancer Hospital, Zhengzhou, China

Background: Platelets can dynamically regulate tumor development and progression. Nevertheless, research on the predictive value and specific roles of platelets in gastric cancer (GC) is limited. This research aims to establish a predictive platelets-related gene signature in GC with prognostic and therapeutic implications.

Methods: We downloaded the transcriptome data and clinical materials of GC patients (n=378) from The Cancer Genome Atlas (TCGA) database. Prognostic platelets-related genes screened by univariate Cox regression were included in Least Absolute Shrinkage and Selection Operator (LASSO) analysis to construct a risk model. Kaplan-Meier curves and receiver operating characteristic curves (ROCs) were performed in the TCGA cohort and three independent validation cohorts. A nomogram integrating the risk score and clinicopathological features was constructed. Functional enrichment and tumor microenvironment (TME) analyses were performed. Drug sensitivity prediction was conducted through The Cancer Therapeutics Response Portal (CTRP) database. Finally, the expression of ten signature genes was validated by quantitative real-time PCR (qRT-PCR).

Results: A ten-gene (*SERPINE1*, *ANXA5*, *DGKQ*, *PTPN6*, *F5*, *DGKB*, *PCDH7*, *GNG11*, *APOA1*, and *TF*) predictive risk model was finally constructed. Patients were categorized as high- or low-risk using median risk score as the threshold. The area under the ROC curve (AUC) values for the 1-, 2-, and 3-year overall survival (OS) in the training cohort were 0.670, 0.695, and 0.707, respectively. Survival analysis showed a better OS in low-risk patients in the training and validation cohorts. The AUCs of the nomogram for predicting 1-, 2-, and 3-year OS were 0.708, 0.763, and 0.742, respectively. TME analyses revealed a higher M2 macrophage infiltration and an immunosuppressive TME in the high-risk group. Furthermore, High-risk patients tended to be more sensitive to thalidomide, MK-0752, and BRD-K17060750.

Conclusion: The novel platelets-related genes signature we identified could be used for prognosis and treatment prediction in GC.

KEYWORDS

gastric cancer, platelets, prediction model, prognosis, gene signature, tumor microenvironment

1 Introduction

Gastric cancer (GC) is one of the most prevalent cancers of the digestive tract. In 2020, there were 1.09 million new cases globally, and 0.77 million deaths due to GC, making it the 5th most common type of cancer and the 4th leading cause of cancer-related death in the world (1). GC is mostly diagnosed at advanced stages owing to its occult onset and atypical early symptoms, which is associated with a dismal overall prognosis with a 5-year survival rate of 19–31% in European and American countries and 28% in China (2). The prognosis of GC patients is not reliably predicted by conventional prognostic approaches including tumor staging systems and histopathological diagnosis, partly due to molecular heterogeneity within similar tumor stages and classifications. Alternative methods for predicting the prognosis of GC patients and directing clinical management of GC treatment are still needed to be investigated.

Beyond their well-established role in pathological thrombosis and hemostasis, platelets (PLT) are increasingly being recognized for their important roles in inflammation, tissue repair, tumor growth, and tumor metastasis (3, 4). Preclinical research has shown that PLT and tumor cells interact in both direct and indirect ways, facilitating tumor cell growth and metastasis (5, 6), immune evasion (7), and chemoresistance (8). Growing clinical data has demonstrated a strong correlation between increased PLT count and poor prognosis in cancer patients (9). Cancers with hematogenous metastases, such as breast cancer, lung cancer, hepatocellular cancer, and GC, are reported to have greater prevalence of thrombocytosis (10–13), which implies that elevated PLT count could be employed to monitor the progress of certain cancers.

Tumor-educated PLT have been shown to have a role in maintaining the primary tumor microenvironment (TME). When PLT get into contact directly with cancer cells, they could be activated and form microaggregates around tumor cells, preventing the cells from being recognized by the immune system (14). Besides, PLT may additionally produce a variety of immune-modulating molecules in a contact-independent way (15, 16), helping to maintain the microenvironments of both primary and metastatic tumors. PLT, along with other non-tumor cells and extracellular matrix, collectively contribute to the immunosuppressive TME that promotes tumor cell proliferation, aids in tumor evasion of immune surveillance, and inhibits anti-tumor immune responses (3).

With the extensive use of RNA-sequencing technology, protein profiling and functional tests, comprehensive analysis of tumor-educated PLT has progressed substantially, making PLT a potential target for cancer treatment and a promising liquid biopsy marker for treatment response monitoring and tumor progression. Nevertheless, research on the specific roles played by PLT in GC is limited. In recent years, the identification of survival-associated genes using array-based databases has been utilized for guiding individualized treatment plans for GC patients (17–19). Therefore, we collected platelet-related genes (PRGs) and developed a reliable PLT-related prognostic risk signature in GC via bioinformatics analysis. The immunological status and biological function of GC patients at high and low risk were then examined. Overall, our study indicate that the PLT-related prognostic risk signature is a reliable gene signature for the prediction of GC patients' prognosis and may strengthen the recognition of GC pathogenesis and the exploration of novel therapeutic targets for GC patients.

2 Materials and methods

2.1 Data acquisition and processing

The transcriptome data and clinical materials of GC (n=378) were downloaded from TCGA (https://portal.gdc.cancer.gov/). The clinical features are detailed in Supplementary Table S1. Moreover, we selected three independent validation datasets (GSE15459, GSE62254, GSE84437) from the GEO database (http://www.ncbi.nlm.nih.gov/geo/) and obtained their normalized microarray gene expression data and clinical data. We obtained a list of 300 PRGs from previous literature (Supplementary Data Sheet 1) (20).

2.2 Identification of prognostic PRGs

After collection and preprocessing the data of GC, the Univariate Cox regression analysis was performed on the PRGs collected to identify PRGs with prognostic value ($P<0.05$). Genetic mutations of the prognostic PRGs were analyzed on the cBioportal online tool (https://www.cbioportal.org/) using the Stomach Adenocarcinoma (TCGA, Firehouse Legacy) dataset. Moreover, a PPI network diagram of the prognostic PRGs was constructed with the STRING database (http://string-db.org/) and graphed with the Cytoscape software (21) (version 3.7.2). Differential expression analysis for the prognostic PRGs between tumor and normal tissues was performed using the "limma" package in R (version 4.2.3) (22).

2.3 Construction and verification of PLT signature

The least absolute shrinkage and selection operator (LASSO) regression analysis of the candidate prognostic PRGs was performed to construct a prognostic gene signature. Then, we calculated each patient's risk score. The calculation formula is as follows:

$$Risk\ \ score = \sum_{i=1}^{n} Coef\ \ mRNA(i) \times Expression\ \ mRNA(i)$$

Based on the median value of the risk score, patients in the TCGA training group were divided into high- and low-risk groups. To further verify the predictive ability of the model, three independent validation datasets (GSE15459, GSE62254, and GSE84437) were included in our study. Risk scores were

calculated separately for each sample in the training cohort and the GEO validation cohorts based on the same risk formula. Based on the median risk score, we could divide the patients into two subgroups of high risk and low risk to explore the prognostic differences between the two groups. The Kaplan-Meier curves and receiver operating characteristic curves (ROCs) were constructed for the training cohort and validation cohorts.

2.4 Independent prognostic analysis and nomogram construction

To determine if the PLT signature may serve as a standalone predictive factor in patients with GC, we preformed multivariate Cox regression analysis. A nomogram for predicting overall survival (OS) at 1, 2, and 3 years in clinical patients was constructed using the "rms" R package based on the patient's age, histologic grade, gender, stage and risk scores.

2.5 Functional enrichment analysis

We utilized the "limma" R package (22) to identify differentially expressed genes (DEGs) between the high-risk group and the low-risk group with the criteria of fold change (FC) > 2 and false discovery rate (FDR) < 0.05. To further investigate the function of the DEGs, the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were analyzed using hypergeometric distribution testing by the "ClusterProfiler" R package (23). "circlize" R package (24) visualizes the GO and KEGG results. Finally, Gene Set Enrichment Analysis (GSEA) with the Kolmogorov–Smirnov (KS) test was performed to find enriched KEGG pathways, the ridge plot was used to present the details of GSEA via the "ggstatsplot" R package.

2.6 Risk model's association with TME

The Immuno-Oncology Biological Research (IOBR) R package (25) (version 0.99.9) was used to analyze the immune features and immune cell infiltration in high- and low-risk groups. Based on the 186 TME-associated signatures in the R packet IOBR, we calculated the sample enrichment score. We assessed the expression of common immune checkpoint genes between high-risk and low-risk groups. The relationship between risk score and immune checkpoint genes was analyzed with Pearson correlation test. The CIBERSORT algorithm in the IOBR package was used for calculating the relative abundance of 22 kinds of immune cells in TCGA-GC cohort, and the ESTIMATE algorithm for calculating each sample's matrix and immune scores.

2.7 Drug sensitivity analysis

The Cancer Therapeutics Response Portal (CTRP) database contains data on the sensitivity of different tumor cells to different chemotherapeutic drugs. We employed the R package "oncoPredict" (26) to calculate the sensitivity of individual GC patient to different chemotherapeutic drugs based on the gene expression data (log2(TPM + 1)). Then, the difference in the area under the dose–response curve (AUC) values between high-risk and low-risk groups was evaluated.

2.8 Validation of expression patterns of signature genes via the human protein atlas

The protein expression of ten signature genes in GC and normal tissues was determined using immunohistochemistry (IHC) from the Human Protein Atlas (HPA) (https://www.proteinatlas.org/), which is a valuable database providing extensive transcriptome and proteomic data for specific human tissues and cells.

2.9 Quantitative real-time PCR validation of signature genes

The normal human gastric epithelial cell line GES-1, and four human gastric cancer cell lines, MKN45, N87, HGC27, and KATO-3, were authenticated by STR profiling. All cell lines were cultured in RPMI-1640 containing 1% penicillin/streptomycin and 10% fetal bovine serum. Cells were grown in 5% CO2 at 37°C. Total RNA was extracted using TRIzol (TransGen Biotech, China). Complementary DNA (cDNA) was synthesized using the GoScript™ Reverse Transcription Mix and Oligo(dT) kit (Promega, United States). Real-time PCR was performed using SYBR Green PCR Master Mix (FastStart Universal SYBR Green Master, Roche). Relative gene expression levels were normalized to the levels of *GAPDH* using the ΔCt method. Each experiment was operated in technical triplicate. The amplification primer sequences of each gene are detailed in Supplementary Data Sheet 6.

2.10 Statistical analysis

R software version 4.2.3 was used to conduct the statistical analysis, and p-values and FDR q-values below 0.05 were regarded as statistically significant.

3 Results

3.1 Identification of prognosis–related PRGs in GC patients

The primary design of this study was depicted in the flow chart (Figure 1). A total of 30 PRGs were significantly associated with prognosis of GC patients based on the Univariate Cox regression analysis (Supplementary Data Sheet 2). Genetic mutations of the 30 prognostic PRGs were analyzed through cBioPortal online tool for GC patients. Genes with a mutation rate no less than 5% are shown

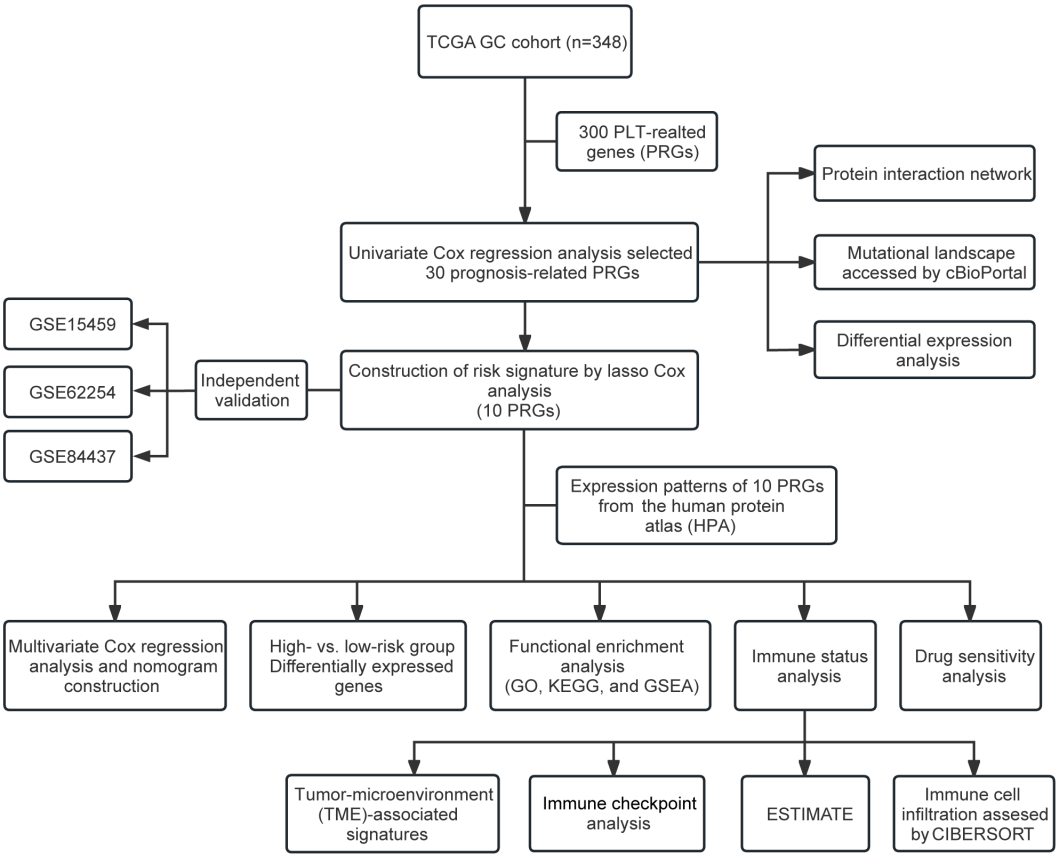


FIGURE 1
Flow chart of the study.

in Figure 2A. *COL1A2* had the highest mutation rate (13%) among 478 patients, followed by *F5* (7%), *GNG11* (7%), *FN1* (7%), and *DGKG* (7%). We constructed a protein interaction network for the 30 genes based on the STRING database. As shown in Figure 2B, the *FN1*, *ALB*, *SERPINE1*, *COL1A1*, *COL1A2*, and *ITGB1* genes were at the core of the protein network interaction. The differential expression analysis of the candidate prognostic PRGs revealed 23 differentially expressed genes, including 18 upregulated and 5 down-regulated genes (Figure 2C).

3.2 PLT signature establishment

The forest plot of the 30 prognostic PRGs obtained by univariate Cox regression analysis was shown in Figure 3A. Then we constructed a predictive prognostic model consisting of 10 PRGs by LASSO regression analysis (Figures 3B–D). They were *SERPINE1*, *ANXA5*, *DGKQ*, *PTPN6*, *F5*, *DGKB*, *PCDH7*, *GNG11*, *APOA1*, and *TF*. The coefficient and HR value of multivariate Cox regression analysis is shown in the form of a forest map (Figure 3B). A linear prediction model was developed based on the weighted regression coefficients of 10 prognostic PRGs, calculated as *risk score* = (−01357× *SERPINE1* exp) + (0.12273 × *ANXA5* exp) + (−0.0927) × *DGKQ* exp + (−0.0722 × *PTPN6* exp) + (0.05354 × *F5* exp) + (0.04579 × *DGKB* exp) + (0.04079 × *PCDH7* exp) + (0.03418 × *GNG11* exp) + (0.03202×*APOA1* exp) + (0.02374 × *TF* exp). Of these, *SERPINE1*, *ANXA5*, *F5*, *DGKB*, *PCDH7*, *GNG11*, *APOA1*, and *TF* showed significant positive correlations with risk scores, while *DGKQ* and *PTPN6* showed significant negative correlation with risk scores.

3.3 Validation of the PLT signature

After establishing the predictive prognostic model based on 10 prognostic PRGs for GC, we computed the risk score for each GC patient based on the LASSO coefficients and expression value for each PRG (Supplementary Data Sheet 3). We contrasted the distribution of risk score, the survival status and the heatmap of GC patients in the TCGA cohort (Figure 4A). The risk curves and scatter plots implied that mortality was positively related to the risk score in the TCGA cohort. The heatmap unveiled that higher *DGKQ* and *PTPN6* expression were detected in the low-risk group, while the other eight genes (*F5*, *APOA1*, *TF*, *ANXA5*, *SERPINE1*, *PCDH7*, *DGKB*, and *GNG11*) were highly expressed

10.3389/fimmu.2024.1477427

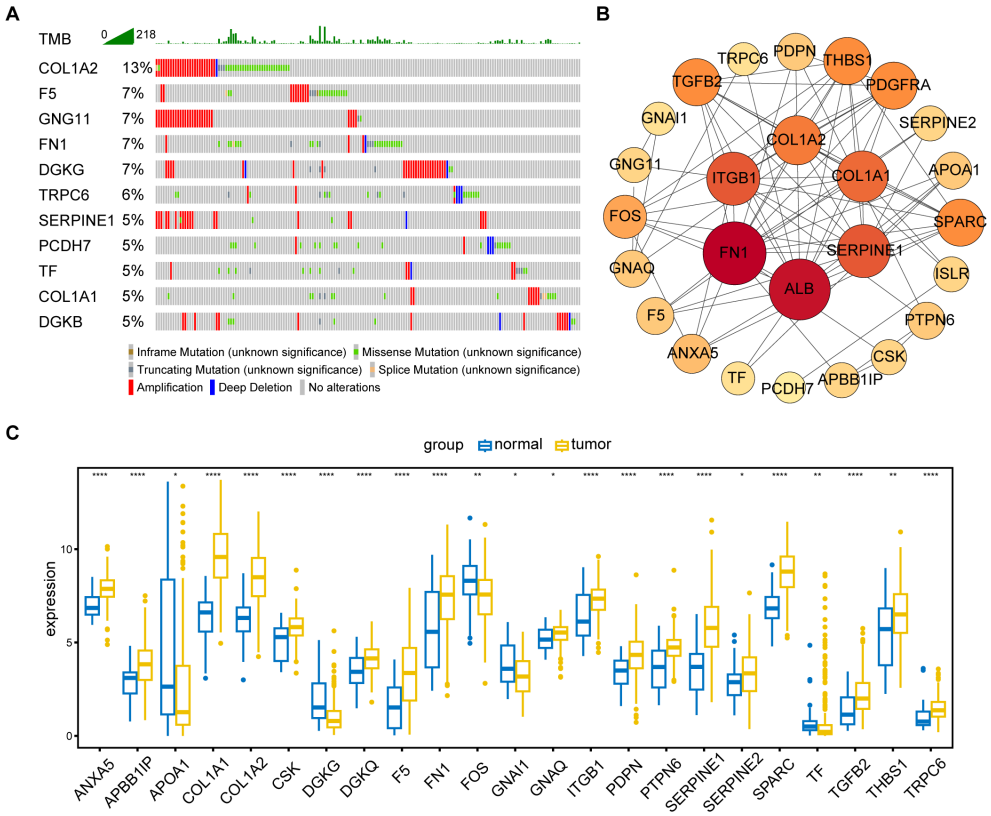


FIGURE 2
Identification and analysis of prognostic PRGs. **(A)** Gene mutation analysis of 30 prognostic PRGs in patients with GC by cBioPortal analysis (only genes altered in ≥ 5% of 478 samples are displayed). **(B)** Protein–protein interaction (PPI) network analysis of 30 prognosis-related PRGs using STRING database. **(C)** Differential expression of 30 prognostic PRGs between tumor and normal tissues (Only genes with P < 0.05 are displayed). *P < 0.05, **P < 0.01, ****P < 0.0001.

in the high-risk group. Kaplan-Meier analysis was used to analyze the survival and prognosis of GC patients in TCGA. As shown in the Figure 4B, patients in the low-risk group had a better prognosis, while patients in the high-risk group had a worse prognosis (P<0.001). The AUCs of 1-year, 2-year, and 3-year survival ROC curves predicted by the PLT signature were 0.670, 0.695, and 0.707, respectively, suggesting the efficiency of PLT signature in predicting prognosis for GC to a certain extent (Figure 4C).

To further demonstrate the stability and reliable generalization of our model, the GSE15459, GSE62254, and GSE84437 cohorts were used as the external validation cohorts. The Kaplan-Meier curves showed a significant difference in prognosis between the high-risk and low-risk patients in these three cohorts, respectively, with a more significant survival advantage for patients in the low-risk group (P = 0.001, P = 0.003, P < 0.001, respectively) (Figures 4D–F). The ROC curve was used as a tool to predict the survival time of patients at 1-, 2-, and 3- years. The AUCs at 1-, 2-, and 3- years for the GSE15459 cohort were 0.670, 0.633, and 0.662, respectively (Figure 4G). The AUCs for the GSE62254 cohort were 0.667, 0.606, and 0.608, respectively (Figure 4H). The AUCs for the GSE84437 cohort were 0.599, 0.608, and 0.611, respectively (Figure 4I). This indicates that the model has an excellent predictive effect.

3.4 Creation of nomograms based on PLT signatures combined with clinical characteristics

To validate the reliability and clinical value of the biological signature constructed based on PRGs as a predictor of prognosis, we conducted multivariate Cox regression analysis including common clinical characteristics (Supplementary Data Sheet 4). It is shown that in the multifactorial cox analysis, tumor stage (P<0.001) and risk score (P<0.001) were all independent prognostic factors significantly associated with patient prognosis (Figure 5A). Based on the above analysis, in order to be able to predict patients' prognosis quantitatively and to inform clinical decision-making, we integrated the risk score and clinical indicators to construct Nomogram plots as a means of predicting the probability of prognostic survival at 1, 2, and 3 years (Figure 5B). We then used time-dependent ROC curve analysis to compare the predictive accuracy between the nomogram, risk score, and common clinicopathological features (Figure 5C). The results showed that risk score had a much greater AUC value than the rest of the individual clinicopathological features, and the nomogram model suggested higher prognostic accuracy at 1-, 2-, and 3-year OS with a larger AUC than risk score. The time-dependent AUCs of the

10.3389/fimmu.2024.1477427

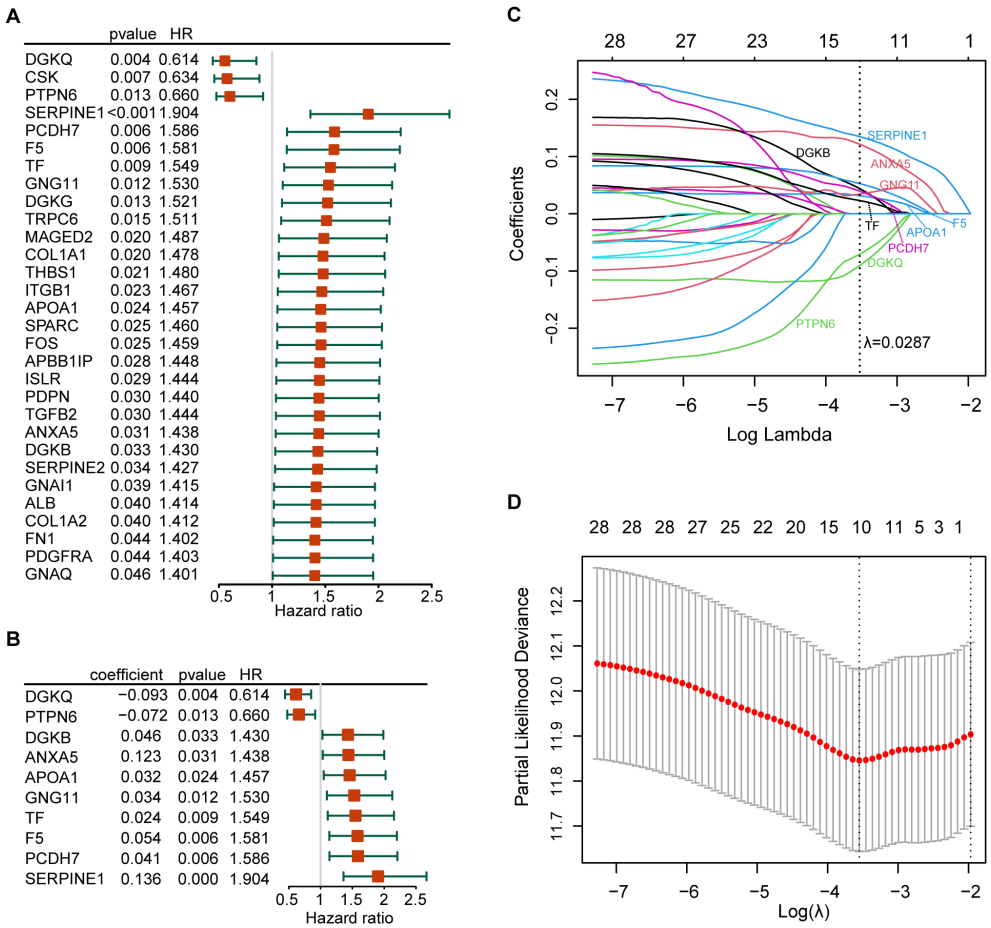


FIGURE 3
Establishment of the risk model. **(A)** The forest plot of the 30 prognosis-related PRGs obtained by univariate Cox regression analysis; **(B)** Construction of prognostic signatures based on lasso Cox analysis; **(C)** LASSO coefficient profiles of 30 prognosis-related PRGs, genes are represented by different colors; **(D)** LASSO regression with tenfold cross-validation, and selection of the optimal parameter (lambda) in the LASSO model.

nomogram for predicting 1-, 2-, and 3-year OS were 0.708, 0.763, and 0.742, respectively. Combined with these results, this suggests that our PLT signature is more practical and influential for clinical decision making and is more suitable as a clinical decision tool for predicting the prognosis of patients with GC in the clinical setting.

3.5 Identification of DEGs between high-risk and low-risk groups and function enrichment analysis

We performed DEGs analysis between high-risk and low-risk groups on the TCGA cohort, and the results showed that 2,442 DEGs were differentially expressed between the high-risk group and the low-risk group based on the criteria of $P < 0.05$. Among that, 2,249 genes were up-regulated, and 193 genes were down-regulated. The volcano plot of DEGs were displayed in Figure 6A. All of the upregulated and downregulated genes were demonstrated in Supplementary Data Sheet 5. The results of GO analysis can be divided into three categories: biological process, cellular component, and molecular function. Where in biological processes, such as axonogenesis, extracellular matrix organization, extracellular structure organization; Cellular components, such as collagen–containing extracellular matrix and synaptic membrane; And molecular functions, such as extracellular matrix structural constituent, G protein–coupled peptide receptor activity, peptide receptor activity, and glycosaminoglycan binding were significantly enriched (Figure 6B). KEGG pathways were enriched in Neuroactive ligand–receptor interaction, cyclic adenosine monophosphate (cAMP) signaling pathway, Calcium signaling pathway, Cell adhesion molecules, and extracellular matrix (ECM)–receptor interaction (Figure 6C). Then, the GSEA method was applied to identify the significantly enriched KEGG pathways in the high-risk samples. The ridgeplot showed that several pathways, such as calcium signaling pathway, cAMP signaling pathway, ECM-receptor interaction, focal adhesion, and neuroactive ligand-receptor interaction, were significantly enriched in the high-risk group (Figure 6D). DNA replication, base excision repair, homologous recombination, nucleotide excision repair were the pathways that were substantially enriched in the low-risk group. GSEA plot of important pathways enriched in the high-risk group was shown in Figure 6E.

10.3389/fimmu.2024.1477427

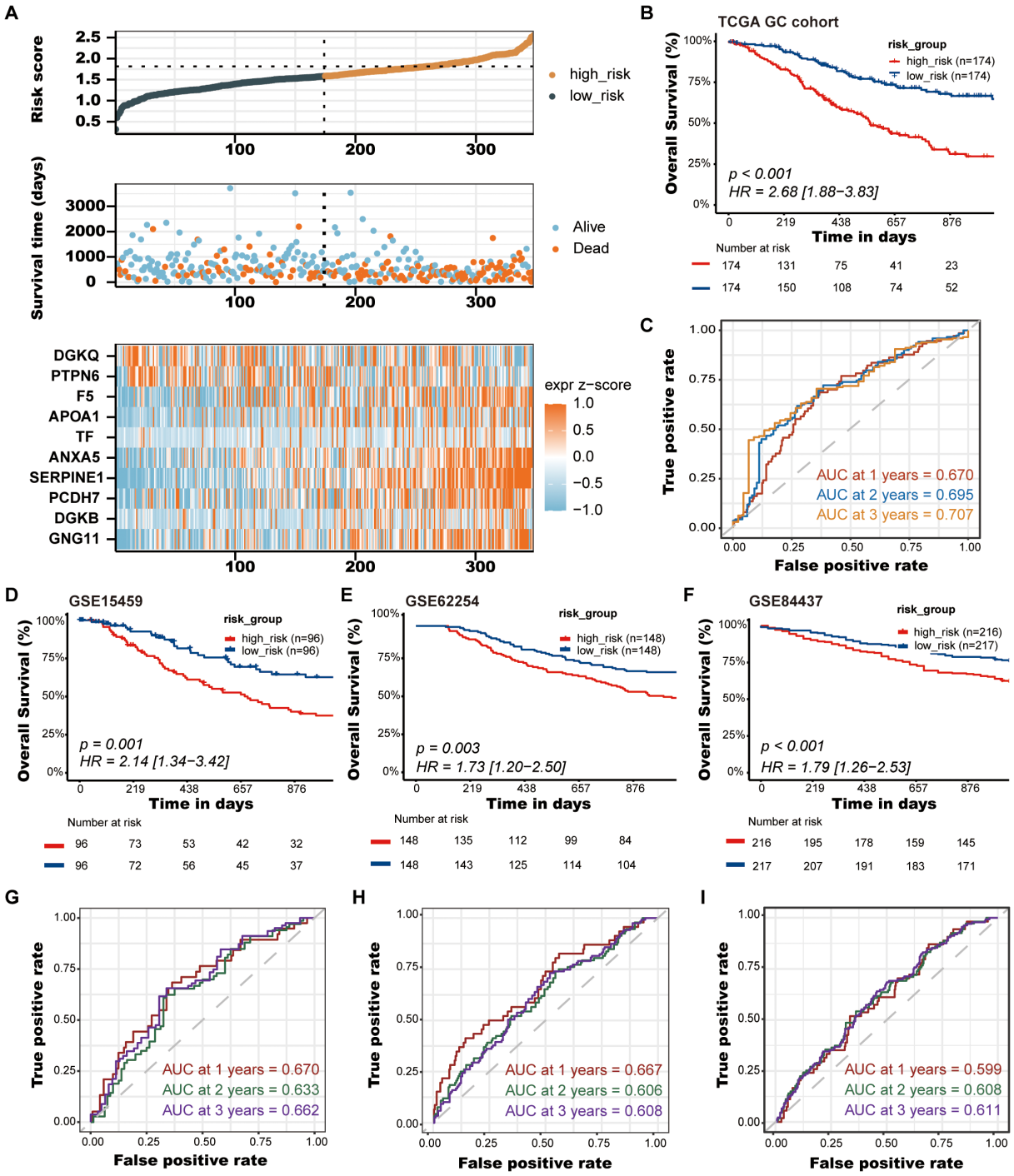


FIGURE 4
Validation of the PLT signature. **(A)** The distribution of risk score, the survival status and the heatmap of GC patients in the TCGA cohort; **(B)** Kaplan-Meier survival curves of OS between low-risk and high-risk groups in the TCGA cohort; **(C)** Time-dependent ROC curves of 1-, 2-, and 3-years of GC patients in TCGA cohort; **(D-F)** Kaplan-Meier survival curves of OS between low-risk and high-risk groups in the GSE15459, GSE62254, and GSE84437 cohorts, respectively; **(G-I)** Time-dependent ROC curves of 1-, 2-, and 3-years of GC patients in the GSE15459, GSE62254, and GSE84437 cohorts, respectively.

3.6 Immune signatures between high-risk and low-risk groups

To further elucidate differences in the immune microenvironment of patients between high-risk and low-risk groups, we compared the enrichment scores of TME cells-related signatures between two groups. The results showed that T cell-related signatures [T cell accumulation, T cell exhaustion, T cell regulatory (27)] and tumor-associated macrophage-related signatures [Macrophages Bindea et al. (28), TAM_Peng_et_al (27)] had significantly higher enrichment

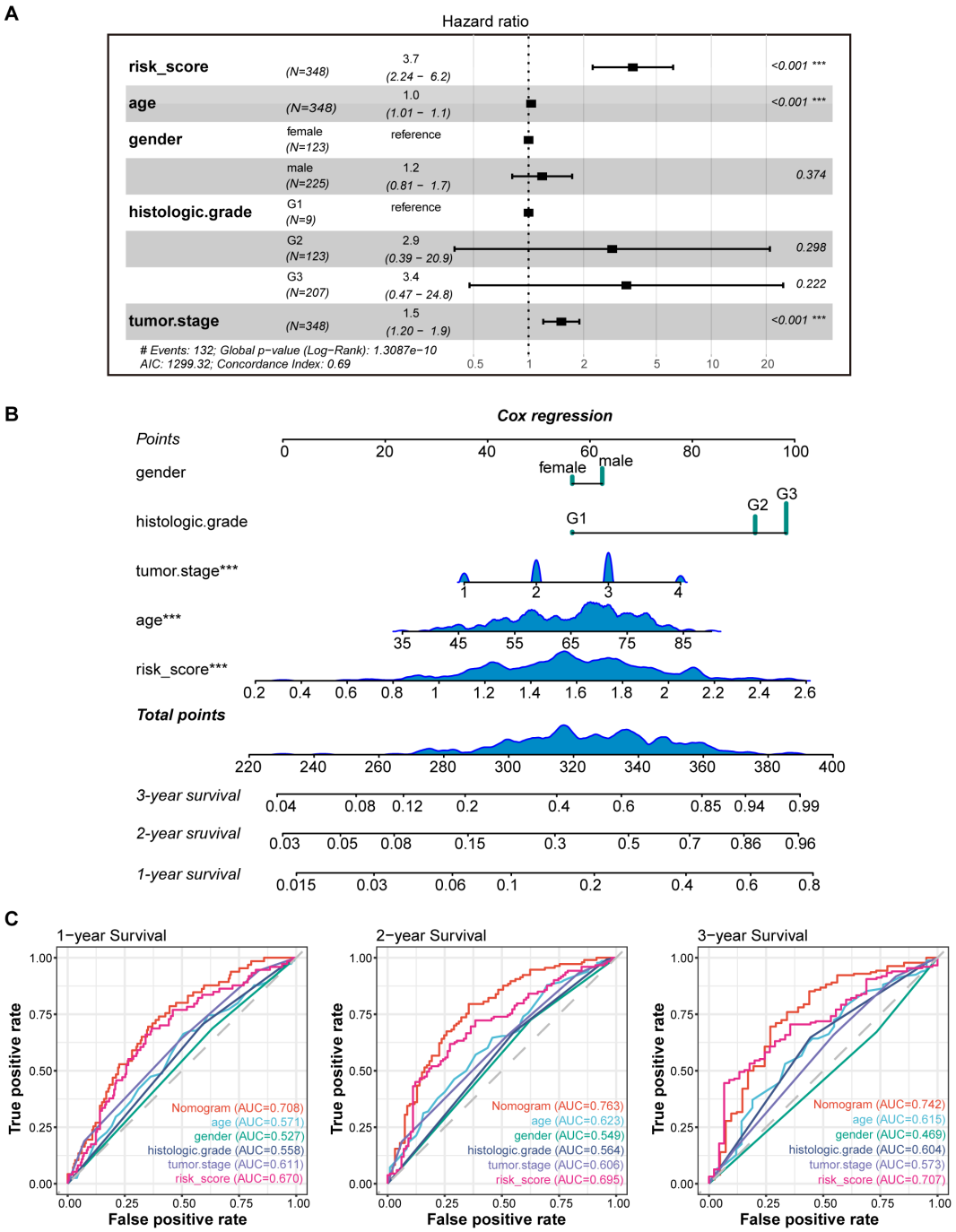


FIGURE 5
Construction and validation of the nomogram model. **(A)** Multivariate Cox analyses indicated that risk score was an independent prognostic factor significantly associated with OS in TCGA cohort; **(B)** Nomogram for predicting 1-, 2-, and 3-year OS; **(C)** Time-dependent ROC curve analyses of the nomogram, risk score, age, gender, histologic grade and tumor stage in TCGA cohort. *** P < 0.001.

scores in the high-risk group compared to the low-risk group (Supplementary Figure S1). Additional examination of TME signatures employing the IOBR package unveiled an immunosuppressive, exclusive, and exhausted TME in the high-risk group (Supplementary Figures S2A–C). Furthermore, patients in the low-risk group exhibited higher scores in DNA damage response (DDR), mismatch repair, and homologous recombination (Supplementary Figure S2D), suggesting that they may be more sensitive to immunotherapy. High-risk patients demonstrated more pronounced epithelial-mesenchymal transition (EMT) signatures (Supplementary Figure S2E). Taken together, these findings suggest an immunosuppressive TME in the high-risk group. The extent of immune cell infiltration in patients in the TCGA cohort was then assessed. The results of ESTIMATE suggested that stromal score, and ESTIMATE score were higher in the high-risk group (Figure 7A). We then estimated the proportion of 22 types of immune cells in each

10.3389/fimmu.2024.1477427

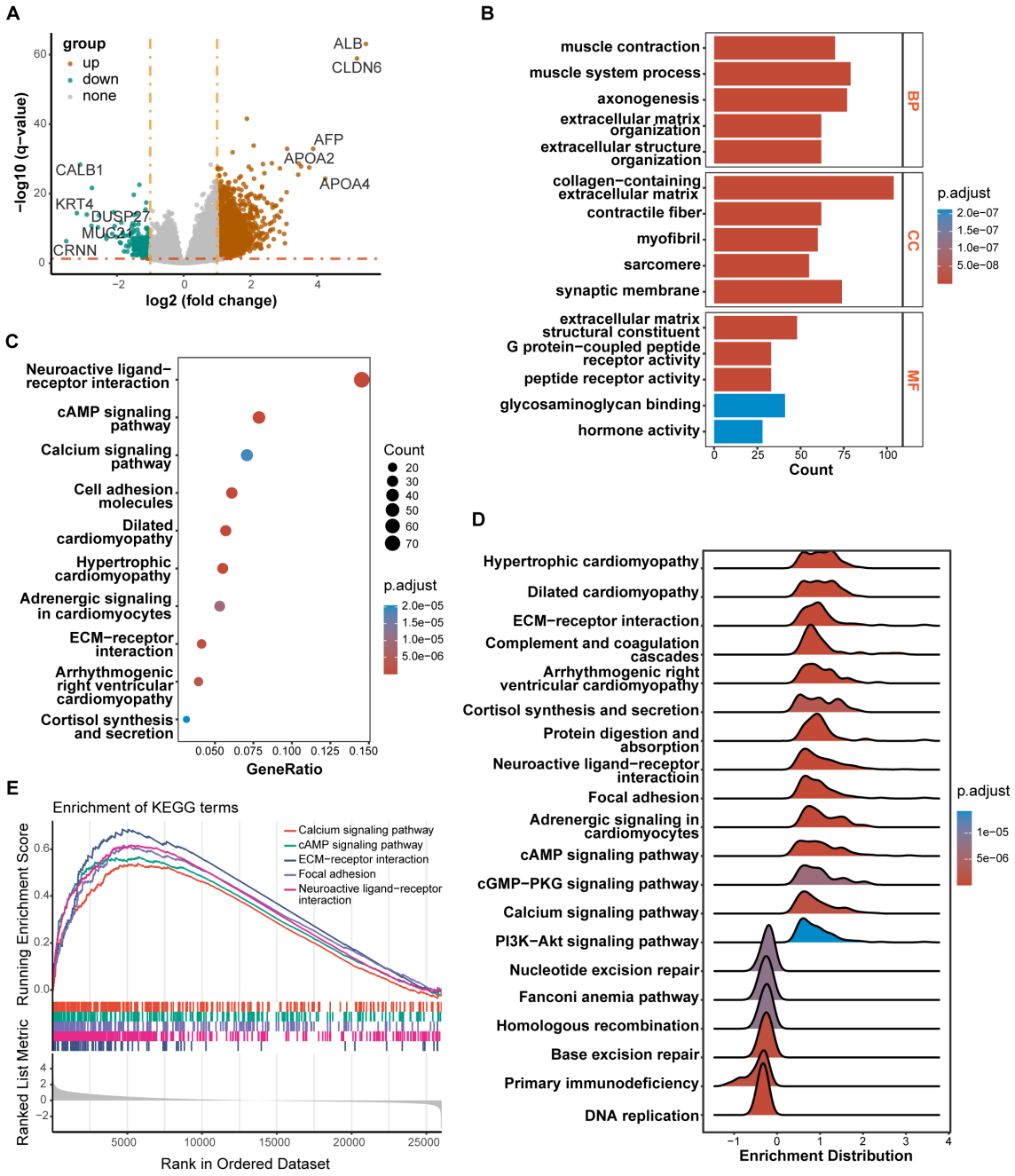


FIGURE 6
Results of differentially expressed genes (DEGs) and function enrichment. **(A)** Volcano plot of DEGs between high-risk and low-risk groups; **(B)** GO enrichment analysis of DEGs between two groups; **(C)** KEGG enrichment analysis of DEGs between two groups; **(D)** Ridgeplot of KEGG by GSEA. **(E)** GSEA plot of important pathways in comparison between two groups.

sample by CIBERSORT algorithm. The difference in the proportion of each type of immune cell between two risk groups was shown Figure 7B. The results revealed that compared with the low-risk group, memory B cells ($P < 0.001$), follicular helper T cells ($P < 0.001$) exhibited lower infiltrating levels in the high-risk group. However, samples in the high-risk group had a significant increase in the fraction of naïve B cells ($P < 0.01$), monocytes ($P < 0.001$) and macrophages M2 ($P < 0.01$). We also explored the relationship between risk score and common immune checkpoint genes, including programmed cell death 1 (PDCD1), PDCD1 ligand 1 (PDCD1L1/CD274), cytotoxic T-lymphocyte-associated antigen 4 (CTLA4), PDCD1 ligand 2 (PDCD1LG2), hepatitis A virus cellular receptor 2/T-cell immunoglobulin mucin receptor 3 (HAVCR2/TIM3), lymphocyte activating 3 (LAG3), and T cell immunoreceptor with immunoglobulin and ITIM domain (TIGIT). As displayed in Supplementary Figure S3A, the levels of HAVCR2 and

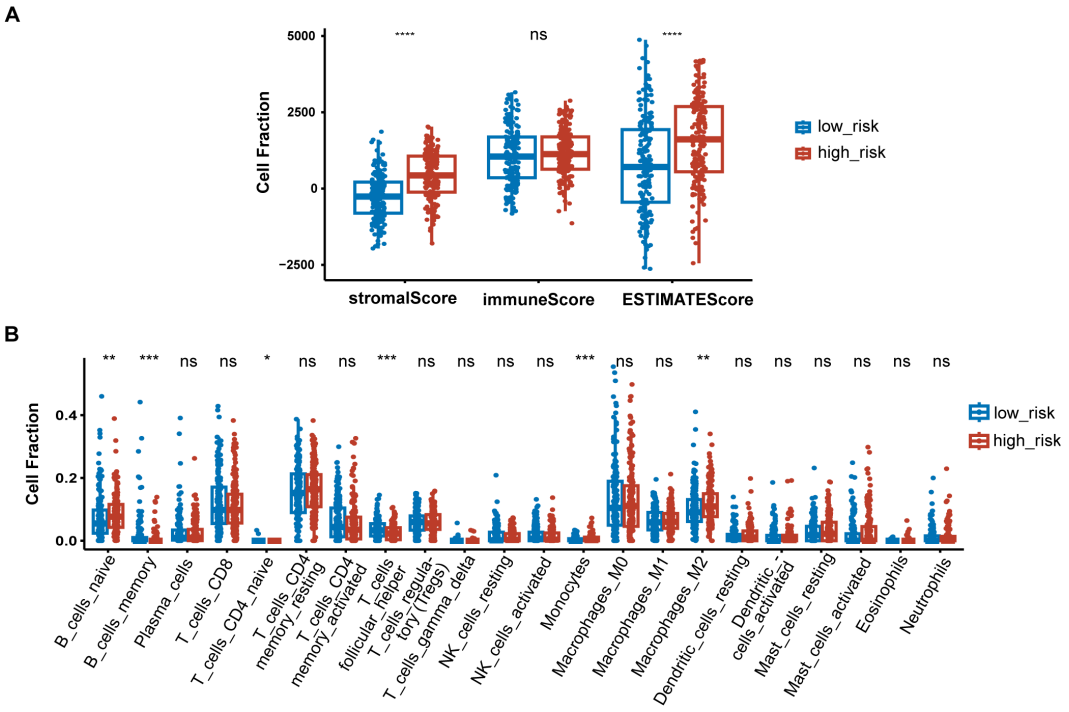


FIGURE 7
Immune signatures between high-risk and low-risk groups. **(A)** Difference of ESTIMATE immune infiltration between different risk groups in the TCGA cohort; **(B)** The proportion of immune cell components assessed by CIBERSORT in the TCGA cohort. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, ns, not significant.

PDCD1LG2 were increased in the high-risk group in GC. However, no significant correlation was observed between the immune checkpoint molecules and the risk score (Supplementary Figure S3B).

3.7 Relationship between risk scores and response to chemotherapy

In order to find more effective drugs for patients in the high-risk group, we further studied the sensitivity of tumor cells to chemotherapeutic drugs between different risk groups based on CTRP database. The AUC value represents the degree of drug sensitivity. An increasing AUC value represents a lower drug sensitivity. We found that patients in the high-risk group tended to be less sensitive to oxaliplatin, doxorubicin, and mitomycin, but more sensitive to thalidomide, MK-0752, and BRD-K17060750 (Figure 8).

3.8 Validation of the expression levels of signature genes in clinical samples

IHC results of the protein expression of the signature genes from HPA database were displayed in Figure 9. SERPINE1 was expressed at a low level in stomach normal tissues and was not detected or expressed at a low level in tumor tissues. ANXA5 was expressed at a low to medium level in both tumor and stomach normal tissues. DGKQ was not detected or expressed at a low level in GC tumor tissues but had medium to high expression levels in stomach normal tissues. PCDH7 was expressed at a high level in stomach normal tissues and had diverse expression levels in GC tumor tissues, ranging from low expression and medium to high expression. GNG11 was not detected in stomach normal tissues but had various expression levels in GC tumor tissues, which was from not detected and low to medium expression. APOA1 was not detected in both tumor and stomach normal tissues. TF was not detected or expressed at a low level in normal tissues and had diverse expression levels in GC tumor tissues, which was from not detected and low expression to medium expression. Data for F5 and DGKB were lacking and therefore not presented.

3.9 Validation of ten signature PRGs by qRT-PCR

The qRT-PCR analysis was performed to further verify ten signature genes in normal and tumor cells (Figure 10). The results showed that the expression of PTPN6, F5, DGKB, PCDH7, and TF was elevated in GC cell lines, whereas the expression of SERPINE, ANXA5, DGKQ, and GNG11 was decreased. APOA1 expression was not detected in the GES-1 and four human gastric cancer cell lines.

4 Discussion

GC is one of the most common primary malignant tumors in the digestive tract with high rates of incidence and mortality.

10.3389/fimmu.2024.1477427

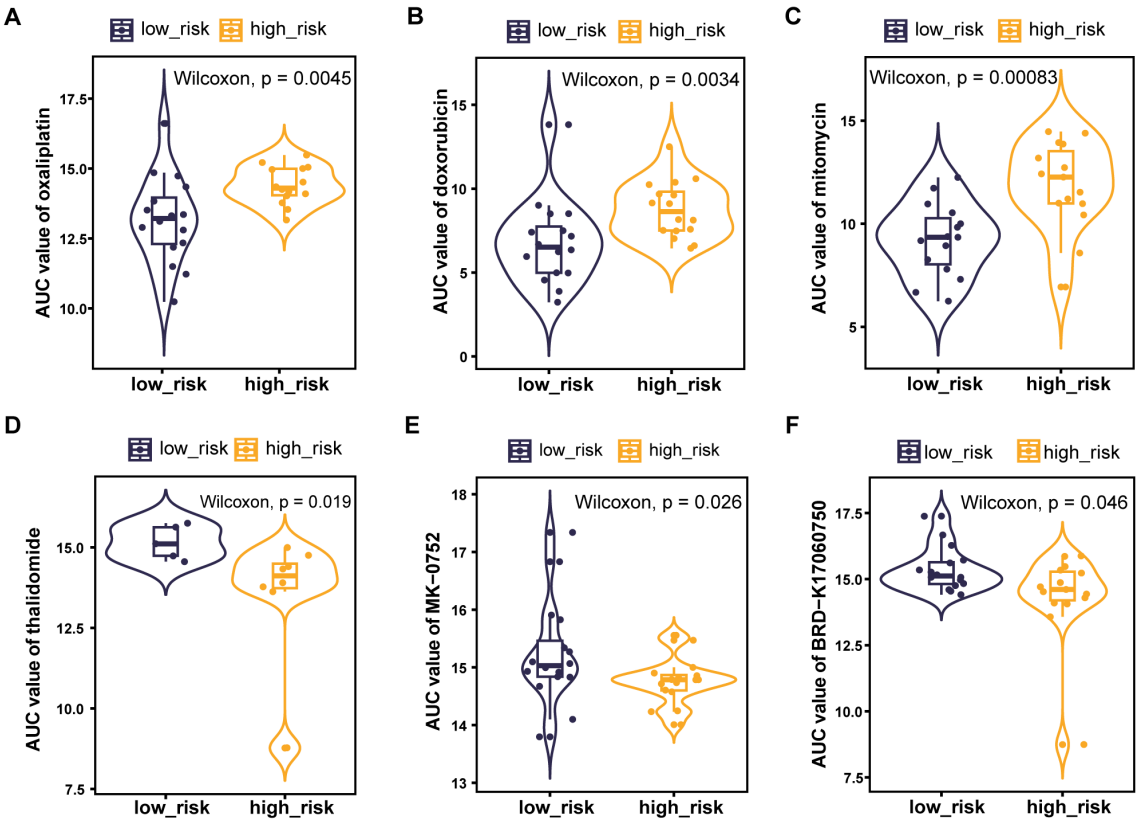


FIGURE 8
PLT signature predicts the sensitivity of chemotherapy. **(A)** Oxaliplatin; **(B)** Doxorubicin; **(C)** Mitomycin; **(D)** Thalidomide; **(E)** MK-0752; **(F)** BRD-K17060750.

Despite recent advances in immunotherapy and molecular targeted therapies, the prognosis of advanced GC patients is still miserably poor. The TNM staging system used by the American Joint Committee on Cancer (AJCC) is a major factor influencing prognosis and treatment decisions of GC (29). The Asian Cancer Research Group (ACRG) newly proposed a molecular classification system where GC is divided into four subtypes: microsatellite stable (MSS)/TP53 activation, MSS/TP53 loss, microsatellite instability (MSI), and MSS/EMT. The result of survival analysis demonstrated that the MSS/EMT group had the worst prognosis due to its easy metastasis and the MSI group had a better prognosis (30). Nonetheless, the existing prognostic stratification systems are not sufficient to accurately predict the prognosis in GC patients. Hence, it is still urgently necessary to explore novel and effectual molecular prognostic biomarkers for GC.

In recent years, PLT in cancers have gotten wide attention due to its roles in regulating tumor proliferation, metastasis and TME through several mechanisms (5, 7). PLT can secrete growth factors like epidermal growth factor (EGF) and vascular endothelial growth factor (VEGF), promoting tumor cell proliferation and angiogenesis (4). PLT also release transforming growth factor-β (TGFβ) and serotonin, creating an immunosuppressive microenvironment by suppressing T cell activity and promoting the transition of M1 macrophages to the M2 phenotype (31). Additionally, PLT facilitate EMT, increasing the invasiveness of tumor cells (5). Moreover, PLT can form microaggregates around circulating tumor cells, protecting them from immune detection and enhancing their ability to metastasize (32). It has been reported that PLT reduction was associated with improved OS and progression-free survival (PFS) rates in patients with stage IV GC (33). However, the effect of PLT-related mRNAs in GC and the mechanism of how PLT alterations affect the tumor biological processes of GC remains unclear to date.

In this study, we integrated PLT-related gene expression profiles from the TCGA-GC dataset and screened 10 genes to construct a new prognostic model for GC patients using LASSO regression analysis. The PLT signature we constructed was shown to be an independent prognostic factor for GC, and a substantial prognostic difference was discovered between the high and low risk groups. Furthermore, the HR of the risk score in multivariate Cox analyses was 3.7 (2.24-6.2), while the HR of tumor stage was only 1.5 (1.20–1.9). Risk score seems more pronounced than tumor stage in prognosis prediction of GC. A nomogram integrated with age, gender, histologic grade, tumor grade and risk score also showed a good prediction of GC patients' survival in 1-, 2-, and 3- years. It helps improve clinicians' decision-making and optimize the personalized treatment plans of GC patients. ROC curves demonstrated the PLT signature's superiority to the other

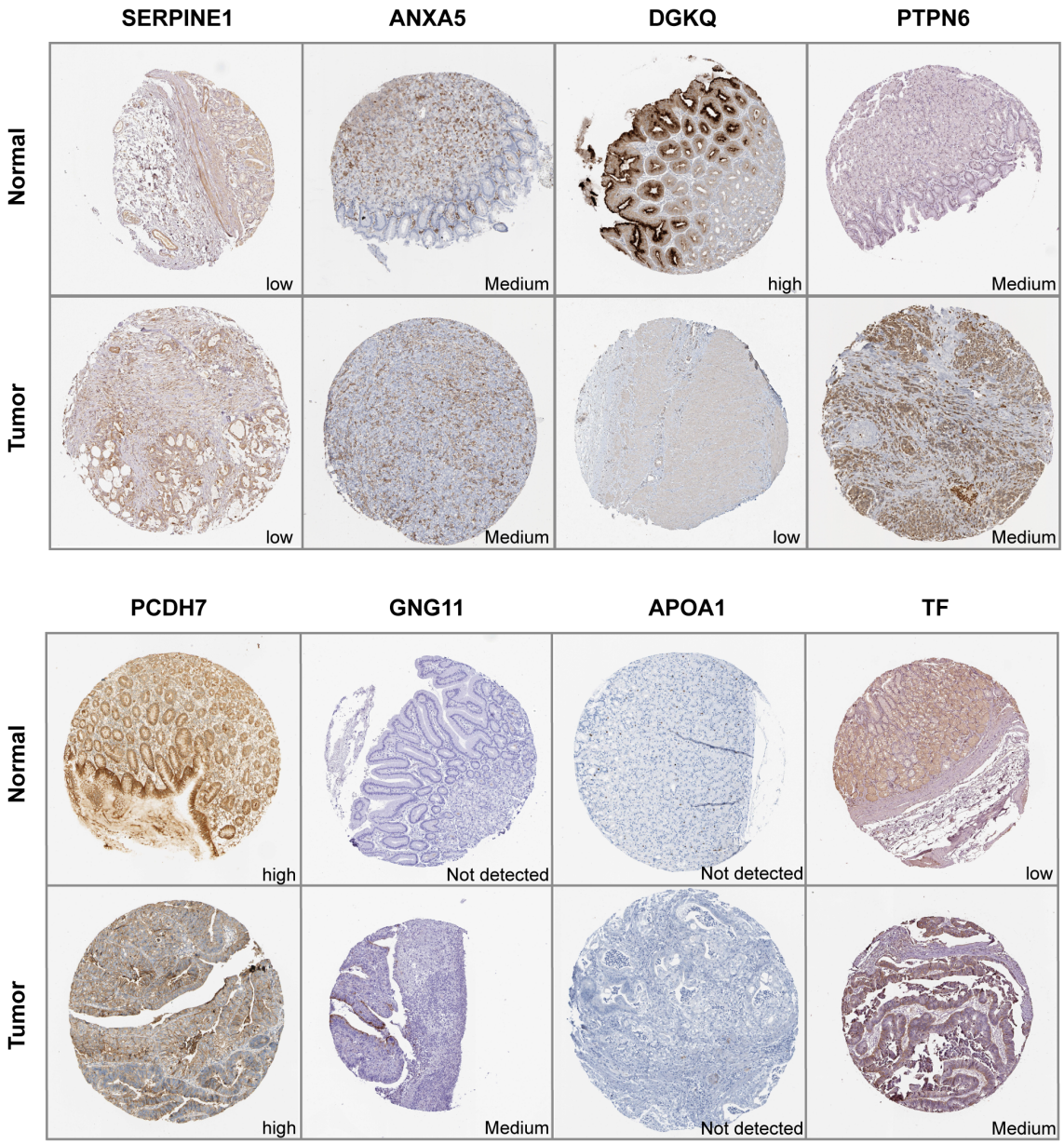


FIGURE 9
Immunohistochemistry of signature genes in GC and normal samples from the HPA database.

traditional clinical indicators such as age, gender, histologic grade, and tumor grade.

After a comprehensive review of the literature, we reviewed the roles of the signature genes in platelet function and their relationship with cancer, highlighting that most genes included in the PLT risk model are closely associated with cancer to varying extents. *Serine protease inhibitor clade E member 1* (*SERPINE1*) plays key roles in regulating the fibrinolytic system (34). It has been detected in various cancers and implicated in tumor progression and angiogenesis in multiple cancer types (35–37). It was reported that *SERPINE1* contributes to tumor proliferation, invasion and migration by regulating EMT in GC (38). *Annexin A5* (*ANXA5*) was identified as an anticoagulant protein and soon reported as a potential apoptosis biomarker due to its binding to phosphatidylserine (39). *ANXA5* contributes to an immunostimulatory profile in the TME and serves as a link between the innate and adaptive immune systems (39). *ANXA5* may potentially affect the prognosis of GC patients as well as the immune therapy response due to its influence on the angiogenesis phenotype (40). *Diacylglycerol Kinase Beta* (*DGKB*) and *Diacylglycerol Kinase Theta* (*DGKQ*) encode different isotypes of Diacylglycerol kinases, which regulates the intracellular concentration of the second messenger diacylglycerol. Inhibition of diacylglycerol kinase augmented platelet secretion and aggregation (41). Diacylglycerol pathways influence the tumor ecosystem by mediating the intricate and dynamic interactions between cancer cells and the tumor immune

10.3389/fimmu.2024.1477427

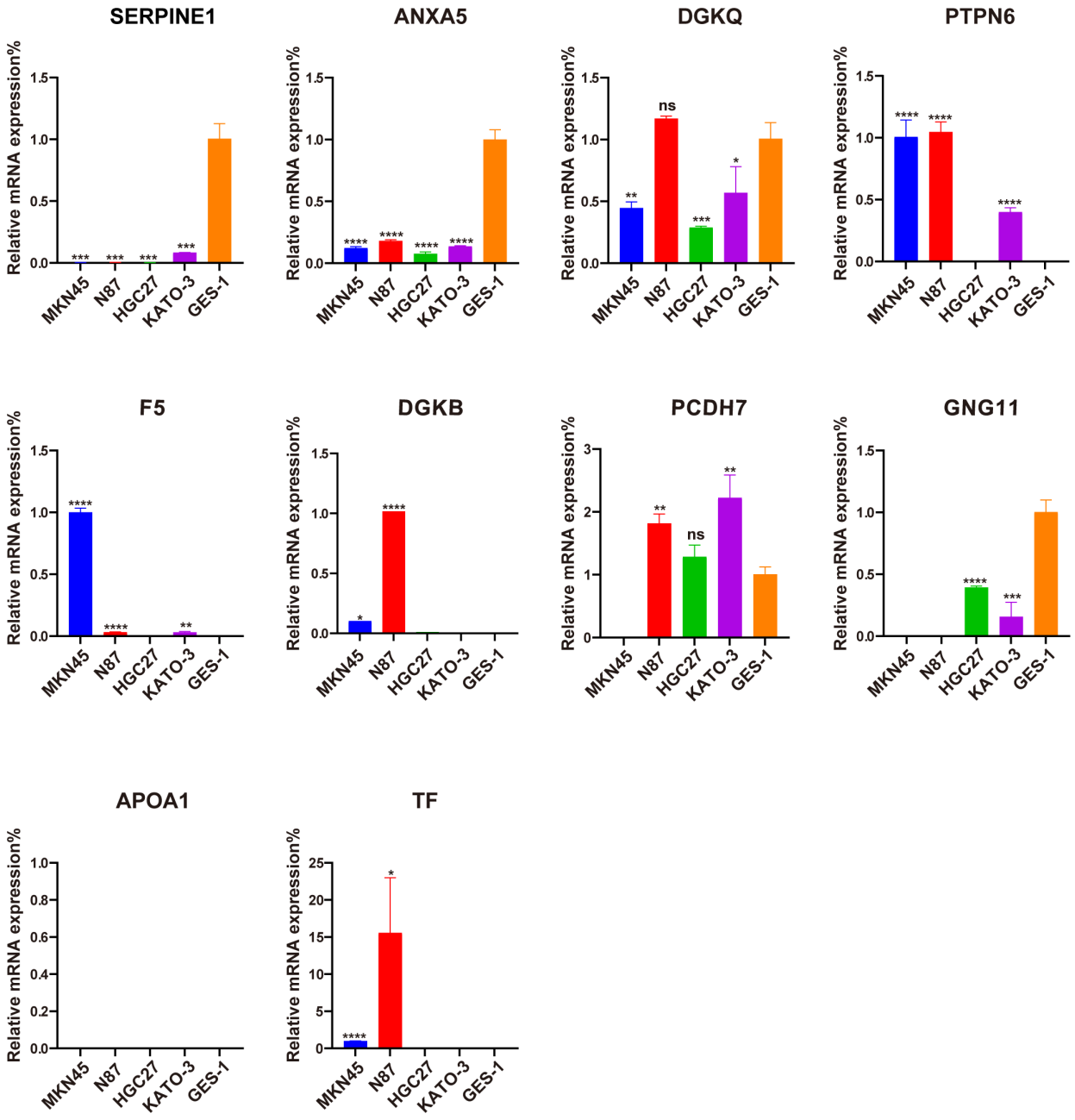


FIGURE 10
Evaluation of the expression of ten PLT-related signature genes in normal and GC cells. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, ns, not significant.

environment (42). The protein encoded by *Protein Tyrosine Phosphatase Non-Receptor Type 6* (*PTPN6*) is a member of the protein tyrosine phosphatase family (43). *PTPN6* has been shown to inhibit platelet apoptosis and necroptosis during sepsis (44), and its elevated expression is linked to poor prognosis and increased immune infiltration in cancer (45). *Coagulation factor V* (*F5*) plays an essential role in coagulation as both a procoagulant cofactor and an anticoagulant cofactor (46). High *F5* expression was associated with aggressive tumors, but also with improved survival in breast cancer (47). *Protocadherin 7* (*PCDH7*) belongs to the cadherin superfamily and plays a role in the pathways of platelet activation, signaling, and aggregation (48). Zhou et al. found that *PCDH7* could suppress cell migration and invasion through E-cadherin inhibition in GC cell lines (49). *G Protein Subunit Gamma 11* (*GNG11*) is a member of guanine nucleotide-binding protein gamma family, which regulated G-protein coupled receptors-dependent platelet function (50). Jiang et al. discovered that high expression of *GNG11* was associated with poor prognosis of ovarian cancer patients (51). *Apolipoprotein A1* (*APOA1*) encodes

apolipoprotein A-I, which was shown to inhibit platelet activation and reduce both clot strength and stability *in vivo* (52). A preclinical study showed that reduced plasma APOA1 level is associated with gastric tumor growth in mouse cancer xenograft model (53). *Transferrin* (*TF*) encodes iron binding transport proteins (54). Although iron metabolism has been reported closely related to cancer progression (55), the role of *TF* in cancer has yet to be investigated.

We found that patients in the high-risk group had significantly higher activities of calcium signaling pathway, cAMP signaling pathway and ECM-receptor interaction. Ca2+ signaling is closely implicated in platelet function. Release of Ca2+ from the dense tubular system into the cytosol initiated by activated PLCγ2 can amplify platelet activation (56). It is becoming evident that dysregulated Ca2+ homeostasis may serve an important role in carcinogenesis or tumorigenesis (57). cAMP is recognized for its significant role in regulating platelets, and platelet activators are known to disrupt the cAMP signaling pathway at various levels (58). Cancer cells, including glioblastoma, ovarian cancer, colorectal cancer and breast cancer, utilize the cAMP/PKA signaling pathway to facilitate invasion, migration, adhesion, clonal development, and other malignant traits (59). Deregulation of ECM remodeling, characterized by excess matrix deposition and increased stiffness, is associated with bone marrow pathologies that can lead to defects in platelet production and function (60). ECM-receptor interaction pathway plays an essential role in tumor shedding, adhesion, and mobility (61). It has been demonstrated that in GC, ECM -receptor interaction pathway takes involvement in the process of tumor invasion and metastasis (62). Taken together, the enrichment of these pathways demonstrates to some extent the mechanism of poorer prognosis in patients with higher risk scores.

As tumor-educated PLT play significant roles in modulating the immune environment, we further explored the immune features of high-risk patients. Immune-related gene signature suggest an immunosuppressive TME in the high-risk group. Given that immune checkpoint inhibitors are less effective in an immunosuppressive microenvironment (63), we speculated that patients in the high-risk group would benefit less from immunotherapy. We further investigated the relationship between the risk scores and immune checkpoint molecules, which have been considered potential biomarkers of response to ICIs. Although the expression of HAVCR2 and PDCD1LG2 was elevated in the high-risk group, the correlations between the immune genes and the risk score were not significant. Additional biomarkers, such as tumor mutation burden (TMB) and human leukocyte antigen (HLA), merit additional investigation to gain a more comprehensive understanding of the potential relationship between the risk score and efficacy of immunotherapy.

The investigation of immune cell infiltration in different risk groups of GC patients can help clinicians to gain a better knowledge of the overall immune landscape of patients. Our findings demonstrated that the high-risk group had higher M2 macrophage infiltration and that tumor-associated macrophage-related signatures were enriched in this group. Li et al. found that GC-derived mesenchymal stromal cells can induce the polarization of macrophages into the M2 subtype, which promotes the migration and invasion of gastric cancer cells via advancing the process of EMT (64). It has been shown that M2 phenotype polarization of macrophage may contribute to acquired trastuzumab resistance in HER2-positive GC (65). Thus, we hypothesized that M2 macrophage polarization may contribute to the poor prognosis in high-risk patients.

Drug resistance is a major cause of death in cancer patients (66). We investigated the potential correlation between drug sensitivity and PLT risk scores using CTRP database. The results suggest that patients in the high-risk group tended to be less sensitive to classical antitumor drugs including oxaliplatin, doxorubicin, and mitomycin. The PLT signature may be used as a predictor of tumor response to chemotherapy. On the other hand, we identified three unconventional antitumor compounds including thalidomide, MK-0752, and BRD-K17060750, with potential advantages for patients with high PLT risk scores. Thalidomide combined with capecitabine has been shown in a prior study to be a safe and mildly effective treatment for elderly patients with advanced GC (67). MK-0752 is a potent inhibitor of gamma secretase, an enzyme required for Notch pathway activation. This agent has been investigated in phase 1 clinical trials in solid tumors (68, 69). The efficacy of these drugs for GC patients in the high-risk group is expected for further investigation.

There are some strengths of the present study. Firstly, our signature is based on TCGA data and GEO databases with relatively large sample sizes. Secondly, our risk model is more cost-effective and has good clinical practicability because it's based on a specific gene set. Thirdly, our risk model and nomogram have great clinical implications for the prognostic evaluation and selection of treatment options for GC patients.

Certain limitations of our study should be addressed. Firstly, more data from prospective clinical GC cohorts need to be collected to externally validate the utility of the model in the next step. Secondly, the public databases provide only a limited amount of information on clinical features and may not include other clinical factors, such as treatment history, and molecular types that can influence prognosis. Lastly, we validated the expression of the risk model genes using qPCR, but further mechanistic studies *in vivo* and vitro need to be conducted to better comprehend the mechanisms by which PLT-related genes affect TME and immunotherapy sensitivity.

In conclusion, our study firstly constructed a reliable PLT-related risk model for predicting survival in GC patients. The independence and predictive performance of this model was further validated using external validation data. This study deepens our understanding of platelet-related genes in GC and provides new potential prognostic and therapeutic biomarkers for individualized treatment.

Data availability statement

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

Author contributions

QL: Conceptualization, Formal analysis, Funding acquisition, Writing – original draft, Writing – review & editing. CZ: Software, Visualization, Writing – original draft, Writing – review & editing. YR: Visualization, Writing – original draft, Writing – review & editing. LQ: Visualization, Writing – original draft, Writing – review & editing. SX: Visualization, Writing – original draft, Writing – review & editing. KL: Visualization, Writing – original draft, Writing – review & editing. YL: Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by PhD Start-up Fund of Henan Cancer Hospital.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2024.1477427/full#supplementary-material

References

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

2. Kole C, Charalampakis N, Tsakatikas S, Kouris N-I, Papaxoinis G, Karamouzis MV, et al. Immunotherapy for gastric cancer: a 2021 update. *Immunotherapy*. (2022) 14:41–64. doi: 10.2217/imt-2021-0103

3. Gaertner F, Massberg S. Patrolling the vascular borders: platelets in immunity to infection and cancer. *Nat Rev Immunol*. (2019) 19:747–60. doi: 10.1038/s41577-019-0202-z

4. Li S, Lu Z, Wu S, Chu T, Li B, Qi F, et al. The dynamic role of platelets in cancer progression and their therapeutic implications. *Nat Rev Cancer*. (2023) 24:72–87. doi: 10.1038/s41568-023-00639-6

5. Labelle M, Begum S, Hynes RO. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell*. (2011) 20:576–90. doi: 10.1016/j.ccr.2011.09.009

6. Cho MS, Bottsford-Miller J, Vasquez HG, Stone R, Zand B, Kroll MH, et al. Platelets increase the proliferation of ovarian cancer cells. *Blood*. (2012) 120:4869–72. doi: 10.1182/blood-2012-06-438598

7. Placke T, Örgel M, Schaller M, Jung G, Rammensee H-G, Kopp H-G, et al. Platelet-derived MHC class I confers a pseudonormal phenotype to cancer cells that subverts the antitumor reactivity of natural killer immune cells. *Cancer Res*. (2012) 72:440–8. doi: 10.1158/0008-5472.CAN-11-1872

8. Radziwon-Balicka A, Medina C, O'Driscoll L, Treumann A, Bazou D, Inkielewicz-Stepniak I, et al. Platelets increase survival of adenocarcinoma cells challenged with anticancer drugs: mechanisms and implications for chemoresistance. *Br J Pharmacol*. (2012) 167:787–804. doi: 10.1111/j.1476-5381.2012.01991.x

9. Lin RJ, Afshar-Kharghan V, Schafer AI. Paraneoplastic thrombocytosis: the secrets of tumor self-promotion. *Blood*. (2014) 124:184–7. doi: 10.1182/blood-2014-03-562538

10. Abdulrahman GO, Das N, Lutchman Singh K. The predictive role of thrombocytosis in benign, borderline and Malignant ovarian tumors. *Platelets*. (2020) 31:795–800. doi: 10.1080/09537104.2019.1686755

11. Li X, Li M, Hu Z, Zhou L, Zheng M, Jiao D, et al. Tumor-infiltrating platelets promote the growth of lung adenocarcinoma. *Transl Oncol*. (2024) 39:101813. doi: 10.1016/j.tranon.2023.101813

12. Pang Q, Gong X, Pan H, Wang Y, Hu X, Liu H, et al. Platelet count as a predictor of vascular invasion and extrahepatic metastasis in hepatocellular carcinoma: A systematic review and meta-analysis. *Heliyon*. (2024) 10:e28173. doi: 10.1016/j.heliyon.2024.e28173

13. Hwang SG, Kim KM, Cheong JH, Kim HI, An JY, Hyung WJ, et al. Impact of pretreatment thrombocytosis on blood-borne metastasis and prognosis of gastric cancer. *Eur J Surg Oncol*. (2012) 38:562–7. doi: 10.1016/j.ejso.2012.04.009

14. Jiang X, Wong KHK, Khankhel AH, Zeinali M, Reategui E, Phillips MJ, et al. Microfluidic isolation of platelet-covered circulating tumor cells. *Lab Chip*. (2017) 17:3498–503. doi: 10.1039/c7lc00654c

15. Kuznetsov HS, Marsh T, Markens BA, Castaño Z, Greene-Colozzi A, Hay SA, et al. Identification of luminal breast cancers that establish a tumor-supportive macroenvironment defined by proangiogenic platelets and bone marrow-derived cells. *Cancer Discovery*. (2012) 2:1150–65. doi: 10.1158/2159-8290.CD-12-0216

16. Hu Q, Hisamatsu T, Haemmerle M, Cho MS, Pradeep S, Rupaimoole R, et al. Role of platelet-derived tgfβ1 in the progression of ovarian cancer. *Clin Cancer Res*. (2017) 23:5611–21. doi: 10.1158/1078-0432.CCR-16-3272

17. Wang L, Li Z, Li Z, Ren Y, Qian L, Yu Y, et al. Identification of A novel gene signature combining ferroptosis- and immunity-related genes for prognostic prediction, immunotherapy and potential therapeutic targets in gastric cancer. *J Cancer*. (2023) 14:3457–76. doi: 10.7150/jca.87223

18. Zhang M, Cao C, Li X, Gu Q, Xu Y, Zhu Z, et al. Five EMT-related genes signature predicts overall survival and immune environment in microsatellite instability-high gastric cancer. *Cancer Med*. (2023) 12:2075–88. doi: 10.1002/cam4.4975

19. Xiang T, Wei Z, Ye C, Liu G. Prognostic impact and immunotherapeutic implications of NETosis-related gene signature in gastric cancer patients. *J Cell Mol Med*. (2023) 28:e18087. doi: 10.1111/jcmm.18087

20. Li X, Zhao K, Lu Y, Wang J, Yao W. Genetic analysis of platelet-related genes in hepatocellular carcinoma reveals a novel prognostic signature and determines PRKCD as the potential molecular bridge. *Biol Proced Online*. (2022) 24:22. doi: 10.1186/s12575-022-00185-9

21. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. (2003) 13:2498–504. doi: 10.1101/gr.1239303

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

23. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118

24. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. *Bioinformatics*. (2014) 30:2811–2. doi: 10.1093/bioinformatics/btu393

25. Zeng D, Ye Z, Shen R, Yu G, Wu J, Xiong Y, et al. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front Immunol*. (2021) 12:687975. doi: 10.3389/fimmu.2021.687975

26. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting *in vivo* or cancer patient drug response and biomarkers from cell line screening data. *Brief Bioinform*. (2021) 22:bbab260. doi: 10.1093/bib/bbab260

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

28. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. (2013) 39:782–95. doi: 10.1016/j.immuni.2013.10.003

29. He X, Wu W, Lin Z, Ding Y, Si J, Sun L-M. Validation of the American Joint Committee on Cancer (AJCC) 8th edition stage system for gastric cancer patients: a population-based analysis. *Gastric Cancer*. (2018) 21:391–400. doi: 10.1007/s10120-017-0770-1

30. Cristescu R, Lee J, Nebozhyn M, Kim K-M, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med*. (2015) 21:449–56. doi: 10.1038/nm.3850

31. Lj G, Felding-Habermann B. Contribution of platelets to tumour metastasis. *Nat Rev Cancer*. (2011) 11:123–34. doi: 10.1038/nrc3004

32. Ren J, He J, Zhang H, Xia Y, Hu Z, Loughran P, et al. Platelet TLR4-ERK5 axis facilitates NET-mediated capturing of circulating tumor cells and distant metastasis after surgical stress. *Cancer Res*. (2021) 81:2373–85. doi: 10.1158/0008-5472.CAN-20-3222

33. Konopka K, Frączek P, Lubaś M, Micek A, Kwinta Ł, Streb J, et al. Reduction of cancer-induced thrombocytosis as a biomarker of improved outcomes in advanced gastric cancer. *J Clin Med*. (2022) 11:1213. doi: 10.3390/jcm11051213

34. Huttinger ZM, Haynes LM, Yee A, Kretz CA, Holding ML, Siemieniak DR, et al. Deep mutational scanning of the plasminogen activator inhibitor-1 functional landscape. *Sci Rep*. (2021) 11:18827. doi: 10.1038/s41598-021-97871-7

35. Humphries BA, Buschhaus JM, Chen Y-C, Haley HR, Qyli T, Chiang B, et al. Plasminogen activator inhibitor 1 (PAI1) promotes actin cytoskeleton reorganization and glycolytic metabolism in triple-negative breast cancer. *Mol Cancer Res*. (2019) 17:1142–54. doi: 10.1158/1541-7786.MCR-18-0836

36. Sakamoto H, Koma Y-I, Higashino N, Kodama T, Tanigawa K, Shimizu M, et al. PAI-1 derived from cancer-associated fibroblasts in esophageal squamous cell carcinoma promotes the invasion of cancer cells and the migration of macrophages. *Lab Invest*. (2021) 101:353–68. doi: 10.1038/s41374-020-00512-2

37. Wang B, Gu B, Zhang T, Li X, Wang N, Ma C, et al. Good or bad: Paradox of plasminogen activator inhibitor 1 (PAI-1) in digestive system tumors. *Cancer Lett*. (2023) 559:216117. doi: 10.1016/j.canlet.2023.216117

38. Yang J-D, Ma L, Zhu Z. SERPINE1 as a cancer-promoting gene in gastric adenocarcinoma: facilitates tumour cell proliferation, migration, and invasion by regulating EMT. *J Chemother*. (2019) 31:408–18. doi: 10.1080/1120009X.2019.1687996

39. Woodward A, Faria GNF, Harrison RG. Annexin A5 as a targeting agent for cancer treatment. *Cancer Lett*. (2022) 547:215857. doi: 10.1016/j.canlet.2022.215857

40. Hong Z, Wen P, Wang K, Wei X, Xie W, Rao S, et al. The macrophage-associated prognostic gene ANXA5 promotes immunotherapy resistance in gastric cancer through angiogenesis. *BMC Cancer*. (2024) 24:141. doi: 10.1186/s12885-024-11878-7

41. Shulga YV, Topham MK, Epand RM. Regulation and functions of diacylglycerol kinases. *Chem Rev*. (2011) 111:6186–208. doi: 10.1021/cr1004106

42. Cooke M, Kazanietz Mg. Overarching roles of diacylglycerol signaling in cancer development and antitumor immunity. *Sci Signaling*. (2022) 15:eabo0264. doi: 10.1126/scisignal.abo0264

43. Han Y, Zhang J, Pang Y, Wang Y, Zhang X, Zhang H. The role of Src homology region 2 domain-containing phosphatase-1 hypermethylation in the classification of patients with myelodysplastic syndromes and its association with signal transducer and activator of transcription 3 phosphorylation in skm-1 cells. *J Int Med Res*. (2021) 49:300060521999550. doi: 10.1177/0300060521999550

44. Jiang J, Li W, Zhou L, Liu D, Wang Y, An J, et al. Platelet ITGA2B inhibits caspase-8 and Rip3/Mlkl-dependent platelet death though PTPN6 during sepsis. *iScience*. (2023) 26:107414. doi: 10.1016/j.isci.2023.107414

45. Cui P, Lian J, Liu Y, Zhang D, Lin Y, Lu L, et al. Pan-cancer analysis of the prognostic and immunological roles of SHP-1/ptpn6. *Sci Rep*. (2024) 14:23083. doi: 10.1038/s41598-024-74037-9

46. Nicolaes GAF, Dahlbäck B. Factor V and thrombotic disease: description of a janus-faced protein. *Arterioscler Thromb Vasc Biol*. (2002) 22:530–8. doi: 10.1161/01.atv.0000012665.51263.b7

47. Tinholt M, Garred Ø, Borgen E, Beraki E, Schlichting E, Kristensen V, et al. Subtype-specific clinical and prognostic relevance of tumor-expressed F5 and regulatory F5 variants in breast cancer: the CoCaV study. *J Thromb Haemost*. (2018) 16:1347–56. doi: 10.1111/jth.14151

48. Yoshida K, Yoshitomo-Nakagawa K, Seki N, Sasaki M, Sugano S. Cloning, expression analysis, and chromosomal localization of BH-protocadherin (PCDH7), a novel member of the cadherin superfamily. *Genomics*. (1998) 49:458–61. doi: 10.1006/geno.1998.5271

49. Chen H-F, Ma R-R, He J-Y, Zhang H, Liu X-L, Guo X-Y, et al. Protocadherin 7 inhibits cell migration and invasion through E-cadherin in gastric cancer. *Tumour Biol*. (2017) 39:1010428317697551. doi: 10.1177/1010428317697551

50. Alarabi AB, Karim ZA, Hinojos V, Lozano PA, Hernandez KR, Ramirez JEM, et al. The G-protein βγ subunits regulate platelet function. *Life Sci*. (2020) 262:118481. doi: 10.1016/j.lfs.2020.118481

51. Jiang M-M, Zhao F, Lou T-T. Assessment of significant pathway signaling and prognostic value of GNG11 in ovarian serous cystadenocarcinoma. *Int J Gen Med*. (2021) 14:2329–41. doi: 10.2147/IJGM.S314911

52. Jones WL, Ramos CR, Banerjee A, Moore EE, Hansen KC, Coleman JR, et al. elevated in trauma patients, inhibits platelet activation and decreases clot strength. *Platelets*. (2022) 33:1119–31. doi: 10.1080/09537104.2022.2078488

53. Chong P-K, Lee H, Zhou J, Liu S-C, Loh MCS, So JBY, et al. Reduced plasma APOA1 level is associated with gastric tumor growth in MKN45 mouse xenograft model. *J Proteomics*. (2010) 73:1632–40. doi: 10.1016/j.jprot.2010.04.005

54. Brandsma ME, Jevnikar AM, Ma S. Recombinant human transferrin: beyond iron binding and transport. *Biotechnol Adv*. (2011) 29:230–8. doi: 10.1016/j.bioteChadv.2010.11.007

55. Morales M, Xue X. Targeting iron metabolism in cancer therapy. *Theranostics*. (2021) 11:8412–29. doi: 10.7150/thno.59092

56. Zhou Y, Zhang D, Tan P, Xian B, Jiang H, Wu Q, et al. Mechanism of platelet activation and potential therapeutic effects of natural drugs. *Phytomedicine*. (2023) 108:154463. doi: 10.1016/j.phymed.2022.154463

57. Cui C, Merritt R, Fu L, Pan Z. Targeting calcium signaling in cancer therapy. *Acta Pharm Sin B*. (2017) 7:3–17. doi: 10.1016/j.apsb.2016.11.001

58. Smolenski A. Novel roles of cAMP/cGMP-dependent signaling in platelets. *J Thromb Haemost*. (2012) 10:167–76. doi: 10.1111/j.1538-7836.2011.04576.x

59. Fajardo AM, Piazza GA, Tinsley HN. The role of cyclic nucleotide signaling pathways in cancer: targets for prevention and treatment. *Cancers (Basel)*. (2014) 6:436–58. doi: 10.3390/cancers6010436

60. Leiva O, Leon C, Kah Ng S, Mangin P, Gachet C, Ravid K. The role of extracellular matrix stiffness in megakaryocyte and platelet development and function. *Am J Hematol*. (2018) 93:430–41. doi: 10.1002/ajh.25008

61. Bao Y, Wang L, Shi L, Yun F, Liu X, Chen Y, et al. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell Mol Biol Lett*. (2019) 24:38. doi: 10.1186/s11658-019-0162-0

62. Yan P, He Y, Xie K, Kong S, Zhao W. In silico analyses for potential key genes associated with gastric cancer. *PeerJ*. (2018) 6:e6092. doi: 10.7717/peerj.6092

63. Nakamura K, Smyth MJ. Myeloid immunosuppression and immune checkpoints in the tumor microenvironment. *Cell Mol Immunol*. (2020) 17:1–12. doi: 10.1038/s41423-019-0306-1

64. Li W, Zhang X, Wu F, Zhou Y, Bao Z, Li H, et al. Gastric cancer-derived mesenchymal stromal cells trigger M2 macrophage polarization that promotes metastasis and EMT in gastric cancer. *Cell Death Dis*. (2019) 10:918. doi: 10.1038/s41419-019-2131-y

65. Hu X, Ma Z, Xu B, Li S, Yao Z, Liang B, et al. Glutamine metabolic microenvironment drives M2 macrophage polarization to mediate trastuzumab resistance in HER2-positive gastric cancer. *Cancer Commun (Lond)*. (2023) 43:909–37. doi: 10.1002/cac2.12459

66. Vasan N, Baselga J, Hyman DM. A view on drug resistance in cancer. *Nature*. (2019) 575:299–309. doi: 10.1038/s41586-019-1730-1

67. Li Y, Chu Y, Song R, Xu F. Thalidomide combined with chemotherapy in treating elderly patients with advanced gastric cancer. *Aging Clin Exp Res*. (2018) 30:499–505. doi: 10.1007/s40520-017-0790-z

68. Krop I, Demuth T, Guthrie T, Wen PY, Mason WP, Chinnaiyan P, et al. Phase I pharmacologic and pharmacodynamic study of the gamma secretase (Notch) inhibitor MK-0752 in adult patients with advanced solid tumors. *J Clin Oncol*. (2012) 30:2307–13. doi: 10.1200/JCO.2011.39.1540

69. Piha-Paul SA, Munster PN, Hollebecque A, Argilés G, Dajani O, Cheng JD, et al. Results of a phase 1 trial combining ridaforolimus and MK-0752 in patients with advanced solid tumours. *Eur J Cancer*. (2015) 51:1865–73. doi: 10.1016/j.ejca.2015.06.115

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 16 December 2024
DOI 10.3389/fonc.2024.1408697

OPEN ACCESS

EDITED BY
Dalila Luciola Zanette,
Oswaldo Cruz Foundation (Fiocruz), Brazil

REVIEWED BY
Haigang Wu,
Henan University, China
Kulbhushan Thakur,
University of Delhi, India

*CORRESPONDENCE
Geert Vandeweyer
geert.vandeweyer@uza.be

RECEIVED 28 March 2024
ACCEPTED 29 November 2024
PUBLISHED 16 December 2024

CITATION
Terrones M, Op de Beeck K, Van Camp G, Vandeweyer G and Mateiu L (2024) Transcriptomic analysis of *ROS1*+ non-small cell lung cancer reveals an upregulation of nucleotide synthesis and cell adhesion pathways.
Front. Oncol. 14:1408697.
doi: 10.3389/fonc.2024.1408697

Transcriptomic analysis of *ROS1*+ non-small cell lung cancer reveals an upregulation of nucleotide synthesis and cell adhesion pathways

Marc Terrones[1,2], Ken Op de Beeck[1,2], Guy Van Camp[1,2], Geert Vandeweyer[1*] and Ligia Mateiu[1]

[1]Center of Medical Genetics, University of Antwerp and Antwerp University Hospital, Edegem, Belgium, [2]Center for Oncological Research, University of Antwerp and Antwerp University Hospital, Wilrijk, Belgium

Introduction: The transcriptomic characteristics of *ROS1*+ non-small cell lung cancer (NSCLC) represent a crucial aspect of its tumor biology. These features provide valuable insights into key dysregulated pathways, potentially leading to the discovery of novel targetable alterations or biomarkers.

Methods: From The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases, all available *ROS1*+ (n = 10), *ALK*+ (n = 5) and *RET*+ (n = 5) NSCLC tumor and *ROS1*+ cell line (n = 7) RNA-sequencing files were collected. In addition, 10 healthy lung RNA-seq samples were included. Differential gene expression with DESeq2 (R package) and gene co-expression (WGCNA, R package) analyses were performed. Functional annotation was performed through Gene Set Enrichment Analysis (GSEA) using Webgestalt and RNAseqChef, Over-Representation Analysis (ORA) through Enrichr. iRegulon was used to identify enriched transcription factors that regulate a gene co-expression module.

Results: *ROS1*+ NSCLC samples were significantly enriched for the nucleotide synthesis and cell adhesion KEGG pathways compared to *ALK*+ and *RET*+ samples. Moreover, *NOTCH1* was significantly downregulated in *ROS1*+ NSCLC and PD-L1 was weakly expressed. When comparing *ROS1*+ tumor versus cell line transcriptomes, an upregulation of *MYC* and *MET* was found in cell lines together with a significantly decreased expression of *HER3*, *HER4* and *BRAF*. Within *ROS1*-tumors, *GJB2* was overexpressed in the *CD74*- and *CLTC*-*ROS1*+ subgroups. The differential expression of *IL20RB* and *GJB2* in cell lines was confirmed through RT-qPCR. Finally, the gene co-expression analysis unveils a gene cluster involving cell cycle-related genes which significantly correlates with the disease stage of patients. In addition, we propose *TFDP1* and *ISL1* as key *ROS1*-specific transcription factors.

Conclusion: This study highlights cell adhesion and nucleotide synthesis as crucial signatures in *ROS1*+ NSCLC. The upregulation of *GJB2* may serve as a prognostic biomarker, along with *IL20RB*, a known mediator of bone metastases. Furthermore, *TDFP1* and *ISL1* were identified as relevant transcription factors that could potentially regulate the biological processes in *ROS1*-rearranged NSCLC.

KEYWORDS

ROS1+ NSCLC, RNA-sequencing, gene co-expression, nucleotide synthesis, cell adhesion, prognostic biomarker

1 Introduction

ROS1+ non-small cell lung cancer (NSCLC) is a molecular subgroup of malignancies which account for approximately 2% of newly diagnosed lung cancer cases every year (1, 2). Chromosomal rearrangements that involve the 6q22 locus harboring *ROS1* result in the formation of oncogenic fusion proteins. Thus, *ROS1*+ NSCLC belongs to the category of oncogene-addicted tumors like *ALK*+, *RET*+ and *NTRK*+ NSCLC, among others (3, 4). The overexpression of a constitutively activated tyrosine kinase-containing fusion protein promotes cell growth, proliferation and migration through the stimulation of the MAPK, PI3K/mTOR and JAK/STAT signaling pathways (1). Oncogene-addicted lung adenocarcinomas are predominantly diagnosed in young, non-smoker patients. They present a low tumor mutational burden (TMB) and show a poor response towards immune checkpoint inhibitors (5). Concerning the treatment of *ROS1*+ NSCLC, tyrosine kinase inhibitors (TKIs) are the most effective targeted therapies in first line. However, resistance is observed in the majority of patients after a certain period of treatment, caused by intrinsic (e.g. kinase point mutations that prevent inhibitor binding) or extrinsic resistance mechanisms, such as histological transformation (e.g. to small cell lung cancer) or the activation of bypass signaling cascades such as EGFR or c-MET (6–9). Although last-generation TKIs such as repotrectinib and the investigational NVL-520 have shown potent activity against the aggressive ROS1 kinase domain point mutations like G2032R (10, 11), heavily pre-treated patients presenting either extrinsic resistance or brain, bone or liver metastases represent a clinical challenge. Unless other targetable alterations are present driving resistance, patients in this setting rely on chemotherapy-based regimes, with a limited benefit (12). In addition, a remarkable heterogeneity in disease outcomes and metastatic patterns is typically observed among patients, highlighting the need for a deeper understanding of the *ROS1*+ NSCLC tumor biology (13, 14).

Intriguingly, the biological processes defining *ROS1*+ NSCLC beyond the tyrosine kinase-mediated signaling remain largely underexplored. This pathway canonically induces the neoplastic transformation of alveolar type II cells. However, the dysregulation of molecular processes not directly orchestrated by the ROS1 signaling requires further elucidation. It is known that *ROS1* rearrangements, due to their pro-tumorigenic role, are mainly mutually exclusive with alterations affecting homologous kinases such as ALK, RET and NTRK. Despite converging in the same cellular signaling axes, each oncogenic fusion type results in different disease trajectories. For instance, *ALK*+ and *RET*+ NSCLC patients are at higher risk of developing brain metastases compared to *ROS1*+ NSCLC during treatment (15). Moreover, the presence of additional mutations that might explain the tumor evolution, such as tumor-suppressor inactivating variants or copy number gains of oncogenes, has not been fully characterized within the *ROS1*+ NSCLC setting. Although keystone findings in the initiation and evolution of lung adenocarcinoma have been recently published, little attention is paid to *ROS1*-driven tumors (16). This can be explained by the low prevalence of *ROS1* rearrangements across NSCLC patients; which restricts the access to the substantial amount of tumor specimens required to conduct solid studies.

The culprits behind the heterogeneous disease outcomes, poor response to immune checkpoint inhibitors and metastases development of *ROS1*+ NSCLC have not been highlighted yet. Nevertheless, developmental signaling pathways such as the Notch and the Sonic Hedgehog (Shh) pathways are widely known to be aberrantly activated in NSCLC, and co-exist with driver oncogene mutations (17–21). Importantly, Notch signaling orchestrates the epithelial-to-mesenchymal transition (EMT) of NSCLC cells, a critical process at the initial phase of metastasis and drug resistance (22). On top of that, non-canonical oncogenic signaling pathways have lately drawn the attention to the lung cancer field, as exemplified by the SPP1 (osteopontin/phosphoprotein 1) pathway (23). A study by Liu et al. unveiled through single cell RNA-sequencing the pro-tumorigenic role of *GJB2* in LUAD (24). This gene encodes the gap junction beta-2 protein belonging to the connexin family, which was described to modulate intercellular communication and extracellular matrix remodeling. Hence, assessing the expression levels of genes involved in the Notch and Shh pathways together with *GJB2* will shed light on the *ROS1*+ NSCLC pathomechanisms.

Some specific insights into *ROS1*+ tumor evolution were provided in a recent study by Neel et al., demonstrating that the subcellular localization determined by the *ROS1* fusion partner gene modulates

the downstream signaling pathway that will be activated (25). This phenomenon will likely be reflected at a gene expression level. Besides contrasting these findings using publicly available data, in this study we hypothesized that *ROS1*+ NSCLC is defined by specific transcriptomic signatures compared to other oncogene-driven tumors like *ALK*+ or *RET*+ lung adenocarcinomas. In parallel to the deeper exploration of the biology of *ROS1*+ NSCLC, the second goal of this exploratory analysis consists in identifying candidate genes whose dysregulated expression in a *ROS1*+-specific manner might be considered as novel therapeutic targets or biomarkers.

2 Methods

2.1 Obtaining gene expression data

Gene expression RNA-seq data files were obtained from publicly available databases Gene Expression Omnibus (GEO) (fastq) and The Cancer Genome Atlas (TCGA) (raw gene counts) after selecting for *ROS1*-fusion containing lung cancer RNA-seq samples. The raw read counts for GEO samples were acquired adhering to the specifications outlined in the published GDC mRNA quantification analysis pipeline, which was also utilized for TCGA samples (GRCh38.d1.vd1_gencode.v36 GDC reference genome, STAR 2.7.5c).

2.2 Patient-derived cell lines

HCC-78 cells were obtained from the German Collection of Microorganisms and Cell Cultures GmbH (DSMZ, Germany). CUTO-28, CUTO-37, CUTO-38 and CUTO-27 were kindly provided by Prof. Dr. Robert C. Doebele (Division of Medical Oncology, University of Colorado School of Medicine, Anschutz Medical Campus). Cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and L-glutamine 1% v/v; at 37°C, 5% CO_2 in a humidified incubator.

2.3 RT-qPCR

To validate the chosen significant differentially expressed genes (DEGs), 2 µg of total RNA extracted with QIAgen RNeasy RNA isolation kit was employed to synthesize the complementary DNA (cDNA) using the SuperScriptTM III First-Strand Synthesis kit (Thermo Fisher, cat. # 18080051). The following primers were designed to assess the expression levels of the genes (5'➔3'): *GJB2* (FWD: TGGTGGACCTACACAAGCA, REV: TGGAGAAGCCGT CGTACAT), *IL20RB* (FWD: CTGAAGGTCCTGAGTGTGATG, REV: GAGGTCTGTGAGCCCAATG), *GAPDH* (FWD: TGCACCACCAACTGCTTAGC, REV:

GGCATGGACTGTGGTCATGAG), *HPRT1* (FWD: TGACACTGGCAAAACAATGCA, REV:

GGTCCTTTTCACCAGCAAGCT) and *YWHAZ* (FWD: CGAAGCTGAAGCAGGAGAAG, REV: TTTGTGGG ACAGCATGGATG). The quantitative PCR was performed using the SYBR Green Master Mix 2x (Eurogentec, cat. # RT-SN2X-03+) in a Bio-Rad CFX96 real-time PCR system. Data was analyzed with the Bio-Rad CFX Maestro software (Bio-Rad v2.3) and qbasePLUS (Biogazelle). Gene expression levels are reported as calibrated and normalized relative quantities (CNRQ) ± standard error (SE) obtained from the gene normalization considering *GAPDH*, *HPRT1* and *YWHAZ* as reference target genes.

2.4 Statistical methods

2.4.1 Differential gene expression

The read counts associated with protein coding genes from all samples were consolidated into a matrix, serving as input for the DESeq2, R package (26). Gene counts related to Y chromosome, mitochondrial, and ribosomal RNA were excluded from subsequent analysis. In the differential gene expression analysis, the batch effect correction was applied using the removeBatchEffect function from the limma R package (27). Multiple contrasts were defined to assess differential expression across various groups within the study (Supplementary File 1). In each comparison, genes with a p-value less than 0.05, following multiple testing correction using the Benjamini-Hochberg (BH) method, were considered to be statistically significant and thus identified as differentially expressed. The cut-offs for the BH method include $\log_2$(fold change) = $|2|$ and false-discovery rate (FDR) < 0.05. The clustering of the cell line samples was performed through the K-means method. K=7 was chosen given that 7 different cell lines were analyzed and the three biological replicates of each sample clustered together, as verified upon the DESeq2 analysis. Correlation between *ROS1* expression and oncogenes of interest across *ROS1*+ tumor samples was calculated using the Spearman rank method in RNAseqChef (28). Regarding the RT-qPCR experiments, technical triplicates were used per each reaction and every RT-qPCR was performed twice. Differences in *GJB2* and *IL20RB* expression between cell lines were assessed using 1-way ANOVA, a Bonferroni correction for multiple comparisons and α=0.05 in GraphPad Prism v9.

2.4.2 Gene co-expression

The investigation of dysregulated genes involved the utilization of Weighted Gene Co-expression Network Analysis (WGCNA, R package) to identify co-expressed modules (i.e. clusters of genes exhibiting similar expression patterns across samples) and hub genes (i.e. genes playing a central and highly connected role within a co-expression module) (29). It is based on the assumption that highly correlated genes within a module (cluster) are involved in common biological processes. For this analysis, we selected the 70% most variable genes from the gene expression data, after filtering, normalization and batch correction. Then, a signed correlation matrix was created by calculating the biweight midcorrelation (a robust alternative to the Pearson correlation) across all gene pairs. This adjacency matrix was obtained using a soft threshold power of 8 to establish a scale-free topology. The

topological overlap measure (TOM) was then computed for all genes, considering both direct pairwise correlations and shared correlations with other genes. Unsupervised clustering of genes in the hierarchical cluster tree, based on a dissimilarity threshold of 1-TOM, resulted in the formation of gene modules. In this study, the minimum module size was defined as 100 genes, and the module-merging cut height was set at 0.3.

2.4.3 Functional annotation

The functional characterization of the differentially expressed genes was carried out through different methods. RNAseqChef was used to perform a gene set enrichment analysis (GSEA) and over-representation analysis (ORA) together with Webgestalt and Enrichr (30, 31).

2.4.4 Kaplan-Meier survival curves

Survival data obtained from the lung adenocarcinoma (LUAD) cohort available at TCGA database was plotted with GEPIA2 (32). All the subtypes were included in the cohort: proximal inflammatory (PI), proximal proliferative (PP) and terminal respiratory unit (TRU).

3 Results

3.1 Patient characteristics

With regard to the *ROS1*+ patient characteristics shown in Table 1, among samples gathered in this study (n=10), the median age at diagnosis was 63 years-old and both sexes are represented in a balanced manner, with 50% males and 50% females respectively. Diverse ethnicities are also present in the study, with 50% of Caucasian (5/10), 33% of Asian (3/10) and 10% of African-American (1/10) descent. Concerning disease stage, half of the sequenced samples were categorized as a IIB (5/10), followed by stage IIIA and IB in equal proportions each (2/10) and only one sample (1%) was collected at stage IV (1/10). Finally, previous treatment received by the patients was not reported for all samples. Within the available annotations, chemotherapy alone or in combination with radiotherapy or immunotherapy was the most common approach (3/10). Only one patient received radiation, immuno- and chemotherapy previously.

3.2 Nucleotide synthesis and cell adhesion pathways are enriched signatures in *ROS1* + NSCLC

To begin with, we sought to confirm the overexpression of the rearranged tyrosine kinase in each of the three tumor types under investigation (Figure 1A). While oncogene overexpression was evident across all subtypes, statistical significance was achieved in the following comparisons: *ALK*+ vs *ROS1*+ tumors ($p < 0.001$), *RET*+ vs *ROS1*+ ($p < 0.001$), *ROS1*+ vs *RET*+ tumors ($p = 0.01$), and *ROS1*+ vs normal lung tissue ($p < 0.001$).

Next, in an effort to profile a *ROS1*+ specific expression signature, transcriptomes of *ROS1*+ tumor samples and *ROS1*+ patient-derived cell lines were compared to normal lung tissue or *ALK*/*RET*+ NSCLC specimens. The number of overlapping significant differentially expressed genes in both contrasts are shown in Figure 1B. 118 genes were differentially expressed in *ROS1*+ compared to both normal tissue and non-*ROS1*+ NSCLC. 805 genes were found to be specifically and significantly dysregulated in *ROS1*+ samples compared to normal lung tissue. In addition, 169 specifically and significantly differentially expressed genes (DEGs) were obtained from the comparison between *ROS1* and *ALK*/*RET*-driven tumors. We then proceeded with an over-representation analysis (ORA) for the 118 *ROS1*+ specific DEGs to determine which biological functions were enriched in *ROS1*-rearranged tumors and cell lines based on the overlapping 118 gene group. The top 10 enriched categories and the log2(fold change) of the genes associated with each biological function are illustrated in Figure 1C. Firstly, metabolic processes related to nucleotide synthesis, such as the response to purine-containing compound, amine or pyridine-containing metabolic process and the response to organophosphorus, were the most enriched. The *ROS1*-specific overexpression of *NMNAT2*, an enzyme involved in nicotinamide adenine dinucleotide (NAD) synthesis, and *ISL1*, a crucial transcription factor regulating glycolysis and tumorigenesis, highlight the metabolic dependencies of *ROS1*+ tumor specimens and cell lines. Moreover, several significantly overexpressed genes, including *IL20RB*, were associated with the positive regulation of cytokine production. Considering the role of *IL20RB* in promoting bone metastasis in lung adenocarcinoma (33), we selected this gene for further validation. RT-qPCR was performed to confirm the expression of *IL20RB* in cell lines, as shown in Figure 1D. Furthermore, the analysis revealed an overrepresentation of biological processes related to cell adhesion, including integrin activation (characterized by *RASIP1* overexpression), homotypic cell-cell interaction, and receptor-mediated endocytosis.

To obtain a more comprehensive understanding of the gene expression landscape, we explored the primary differences between *ROS1*+ NSCLC samples and cell lines compared to normal lung tissue. This approach aimed to highlight the crucial pathways that may contribute to the malignant transformation process in this specific tumor subtype. For this purpose, we carried out an over-representation analysis (ORA). As shown in Figure 1E, the MAPK signaling pathway (Gene ratio = 1.78, $p = 0.4$) and the focal adhesion KEGG pathways (GR = 2.31, $p = 0.15$) contained the highest amount of the significant DEGs. In addition, the choline metabolism in cancer (GR = 3.2, $p = 0.15$), the neurotrophin signaling pathway (GR = 2.56, $p = 0.24$) and the *Vibrio cholerae* infection were also over-represented categories (GR = 3.97, $p = 0.18$). To complement the initial findings, we expanded the GSEA-based analysis to identify significantly upregulated or downregulated pathways in *ROS1*+ lung adenocarcinomas and cell lines in comparison to *ALK*/*RET*+ tumor samples (Figure 1F). Interestingly, the analysis revealed an upregulation of gene sets associated with the cell cycle and homologous recombination, accompanied by a downregulation of gene sets related to tyrosine and beta-alanine metabolism, lysosome function, and the chemokine signaling pathway (all p-values < $2e^{-16}$).

10.3389/fonc.2024.1408697

TABLE 1 Characteristics of the samples retrieved from GEO and TCGA.

Sample	GEO Number	Sequenced material	Oncogenic fusion	Age	Sex	Ethnicity	Disease stage	Previous treatment
S39	GSM993681	Tumor + normal	*CD74-ROS1*	58	M	Asian	IIB	n/a
S9	GSM993651	Tumor + normal	*CCDC6-ROS1*	69	M	Asian	IIB	n/a
S48	GSM993690	Tumor + normal	*SLC34A2-ROS1*	48	F	Asian	IB	n/a
TCGA-64-1680-01	n/a	Tumor	*CD74-ROS1*	63	M	Caucasian	IV	RT
TCGA-86-8278-01	n/a	Tumor	*CD74-ROS1*	63	F	Caucasian	IIB	RT, CT and TKI
TCGA-44-2665-01	n/a	Tumor	*CLTC-ROS1*	55	F	Caucasian	IIB	CT and IT
TCGA-55-6986-01	n/a	Tumor	*EZR-ROS1*	74	F	Caucasian	IB	None
TCGA-NJ-A7XG-01	n/a	Tumor	*EZR-ROS1*	49	M	African American	IIIA	CT
TCGA-05-4426-01	n/a	Tumor	*SLC34A2-ROS1*	71	M	n/a	IB	n/a
TCGA-62-A46Y-01	n/a	Tumor	*SLC34A2-ROS1*	70	F	Caucasian	IIIA	RT and CT
S26	GSM993668	Tumor + normal	*EML4-ALK*	70	F	Asian	IB	n/a
TCGA-50-8460-01	n/a	Tumor	*EML4-ALK*	74	M	Caucasian	IA	RT
TCGA-67-6215-01	n/a	Tumor	*EML4-ALK*	52	F	Caucasian	IB	CT
TCGA-78-7163-01	n/a	Tumor	*EML4-ALK*	60	M	Caucasian	n/a	n/a
TCGA-86-A4P8-01	n/a	Tumor	*EML4-ALK*	59	F	Caucasian	n/a	n/a
S2	GSM993645	Tumor + normal	*KIF5B-RET*	62	M	Asian	IIIA	n/a
S6	GSM993649	Normal	*KIF5B-RET*	58	M	Asian	IA	n/a
S42	GSM993684	Normal	*KIF5B-RET*	62	F	Asian	IIIB	n/a
TCGA-75-6203-01	n/a	Tumor	*CCDC6-RET*	n/a	F	n/a	IIIA	n/a
TCGA-55-6543-01	n/a	Tumor	*TRIM33-RET*	n/a	n/a	n/a	IA	n/a
TCGA-55-8616-11A	n/a	Normal	–	58	F	Caucasian	n/a	n/a
TCGA-50-5931-11A	n/a	Normal	–	75	F	Caucasian	n/a	n/a
TCGA-55-1592-11A	n/a	Normal	–	n/a	M	Caucasian	n/a	n/a
TCGA-55-6968-11A	n/a	Normal	–	61	M	Caucasian	n/a	n/a
TCGA-38-4632-11A	n/a	Normal	–	42	M	African American	n/a	n/a
CUTO23	GSM7675355-57	Cell line	*CD74-ROS1*	n/a	n/a	n/a	n/a	n/a
CUTO27	GSM7675361-63	Cell line	*CD74-ROS1*	n/a	n/a	n/a	n/a	n/a

(Continued)

TABLE 1 Continued

Sample	GEO Number	Sequenced material	Oncogenic fusion	Age	Sex	Ethnicity	Disease stage	Previous treatment
CUTO28	GSM7675367-69	Cell line	*TPM3-ROS1*	n/a	n/a	n/a	n/a	n/a
CUTO33	GSM7675373-75	Cell line	*CD74-ROS1*	n/a	n/a	n/a	n/a	n/a
CUTO37	GSM7675379-81	Cell line	*CD74-ROS1*	n/a	n/a	n/a	n/a	n/a
CUTO38	GSM7675386-88	Cell line	*CD74-ROS1*	n/a	n/a	n/a	n/a	n/a
HCC78	GSM7675391-93	Cell line	*SLC34A2-ROS1*	65	M	n/a	n/a	n/a

M, male; F, female; RT, radiotherapy; CT, chemotherapy; TKI, tyrosine kinase inhibitor; n/a, not available.

3.3 Inflammation-related pathways differentiate *ROS1*+ from *ALK*+ and *RET*+ NSCLC

Whole tumor transcriptome analysis enables accurate profiling of both the tumor cells and the stromal compartment, which comprise a heterogeneous population of cells that work together to enhance the fitness of the tumor cells. Therefore, directly comparing *ROS1*+ lung adenocarcinomas with *ALK*+ or *RET*+ tumors, while excluding cell lines, provides further insights into the tumor microenvironment. Figure 2A shows the heatmap resulting from clustering *ROS1*+ and *ALK*+ tumor transcriptomes based on the significant DEGs between the two groups. Samples clustered together according to their driver mutation, indicating an oncogene-dependent modulation of the gene expression profile. Subsequent GSEA revealed a significant upregulation of the interleukin-17 signaling pathway, the ribosome signature and the systemic lupus erythematosus KEGG pathways in *ROS1*+ tumors compared to *ALK*+ specimens (Figure 2B). In addition, downregulation of the cGMP-PKG signaling pathway, the aldosterone synthesis and secretion signature together with primary immunodeficiency and vascular smooth muscle contraction gene sets was observed in *ROS1*+ vs *ALK*+ tumors. Thus, both up and downregulated gene sets implicate differences concerning inflammatory and protein synthesis pathways. Similarly, *ROS1*+ and *RET*+ specimens were collated, revealing a mixed clustering pattern (Figure 2C). In this case, three out of the four *RET*+ samples grouped together whilst one shared higher similarity with *ROS1*+ samples. GSEA unveiled significant upregulation of immune response, hematopoietic cell lineage, cell adhesion and cholesterol metabolism pathways in *ROS1*-rearranged tumors. Interestingly, no significantly upregulated gene sets were observed in *RET*+ specimens (Figure 2D).

We explored the impact of the Notch signaling pathway within *ROS1*+ NSCLC, as it is known to exert a pro-tumorigenic role in NSCLC (Figure 2E). Although no genes regulated through the Notch signaling pathway were significantly dysregulated, we found that *NOTCH1* expression decreased significantly in *ROS1*+ tumor samples compared to *RET*+ specimens (p=0.04). With regard to the other members of the *NOTCH* gene family, *NOTCH2* and *NOTCH3* reflected similar expression trends across tumor types. Interestingly, *NOTCH4* exhibited a modest expression in *ROS1*+ tumors compared to *ALK*+ and *RET*+ specimens, although the difference did not reach statistical significance. Considering the role of the Notch pathway in promoting epithelial-to-mesenchymal transition (EMT), we conducted a comprehensive analysis to explore the association between *NOTCH1* expression and the metastatic characteristics of lung adenocarcinoma (LUAD) patients across multiple studies. As illustrated in Supplementary Figure 3A, the group exhibiting low *NOTCH1* expression levels comprised a significantly higher proportion of individuals with TNM stage M0 (p=0.0021), suggesting an absence of metastatic activity in the tumor. To extend our understanding of the patterns of metastatic progression, we retrieved the clinical details of *ROS1*+, *ALK*+ and *RET*+ NSCLC patients from the "Metastatic NSCLC study" by Jee et al. (34). Supplementary Figure 3B shows that *ROS1*+ NSCLC patients presented lower rates of extrathoracic metastases, being the central nervous system (CNS) and liver the most common sites. Contrarily, *ALK*+ and *RET*+ NSCLC patients were diagnosed with bone and soft tissue metastatic lesions as well as CNS and liver. Thus, higher *NOTCH1* expression in *RET*+ specimens might be related to the EMT-promoting effect and enhanced metastatic disease in *RET*+ NSCLC patients.

In addition, changes in expression of canonical tumor-suppressor genes such as *BRCA1*, *BRCA2* and *TP53* as well as oncogenes like *BRAF*, *EGFR*, *KRAS*, *MET* and *MYC* were assessed (Figure 2F). No significant differences were observed between the various tumor types, suggesting that the expression of these genes is not dependent on the specific rearrangement. Furthermore, we investigated the expression levels of *CD274*, a gene that encodes the programmed death-ligand 1 (PD-L1), a cell surface protein expressed by neoplastic cells that interacts with its receptor, programmed cell death protein 1 (PD-1). The latter is present in activated T, natural killer (NK) and B lymphocytes, macrophages, dendritic cells (DCs) and monocytes. The interaction between PD-1 and PD-L1 leads to the suppression of cellular immunity against tumor cells. Across the three tumor types, a relatively low level of normalized *CD274* counts was observed, with no significant differences detected. This finding supports the classification of *ROS1*+, *ALK*+, and *RET*+ NSCLC as "cold tumors," which is consistent with the modest benefit of immunotherapy observed in these patient subsets (35, 36).

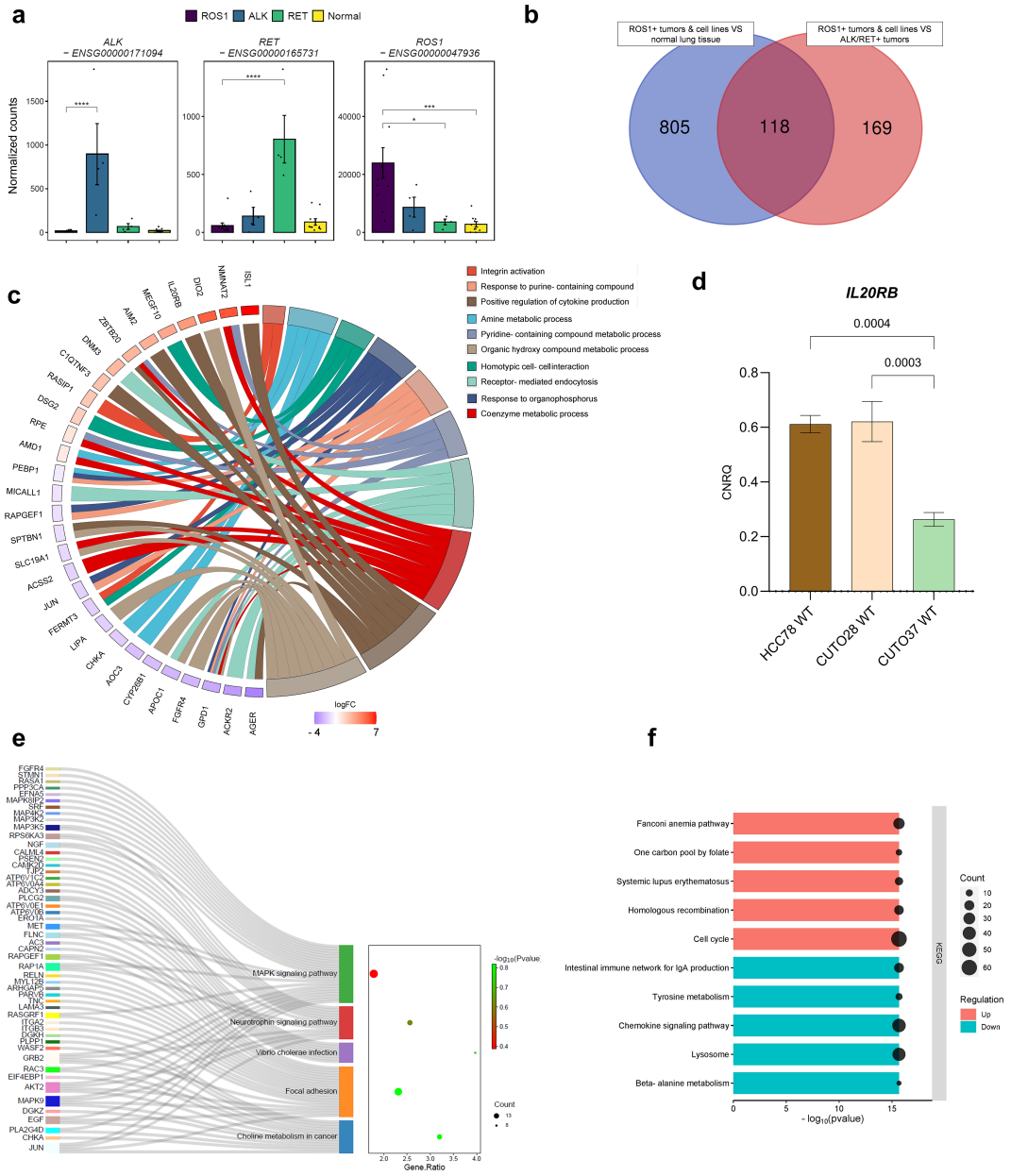


FIGURE 1
ROS1+ NSCLC signature. **(A)** Gene expression levels of the oncogenic kinases in *ALK*+, *RET*+, *ROS1*+ tumors and normal lung tissue. **(B)** Venn diagram reflecting the significant differentially expressed genes (DEGs) between two comparisons: (1) *ROS1*+ tumor specimens and cell lines versus normal lung tissue and (2) *ROS1*+ tumor specimens and cell lines versus *ALK*+ and *RET*+ tumor samples. **(C)** Circos plot reflecting the top 10 enriched biological processes in *ROS1*+ NSCLC. **(D)** *IL20RB* mRNA levels in HCC-78, CUTO-28 and CUTO-37 cell lines expressed as calibrated and normalized relative mRNA quantity (CNRQ) ± SEM. A 1-way ANOVA test was performed considering a Bonferroni correction and a chosen $\alpha = 0.05$. **(E)** Over-representation analysis (ORA) depicting the key enriched pathways in *ROS1*+ tumor samples and cell lines versus normal lung tissue. **(F)** GSEA summarizing the dysregulated KEGG pathways in *ROS1*+ NSCLC compared to normal lung tissue and *ALK*+/*RET*+ tumor specimens.

3.4 Differences between *ROS1*+ specimens and tumor-derived cell lines comprise cell cycle, DNA repair and inflammation pathways

Given that both *ROS1*+ NSCLC tumor specimens and *ROS1*+ patient-derived cell lines were used in this study, evaluating the key transcriptomic differences between tumor and cell line samples is particularly important. As shown in Figure 3A, CUTO and HCC-78 cell lines cluster separately from *ROS1*+ NSCLC tumor samples. This remarkable difference is reflected by the 6,803 significantly DEGs. Interestingly, the HCC-78 cell line gene expression pattern clustered between the one of CUTO-23 and CUTO-38 cell lines, indicating that HCC-78 cells share common transcriptomic traits with CUTO cell lines. Among the significant DEGs shown in in Figure 3B, tumor specimens overexpressed genes like *MUC5B*, encoding for the glycoprotein mucin, *AEBP1 (AE binding protein 1)*, *SPARCL1 (SPARC-like protein 1)*, *RBX1 (ring box-1) and ELN (elastin)*. In

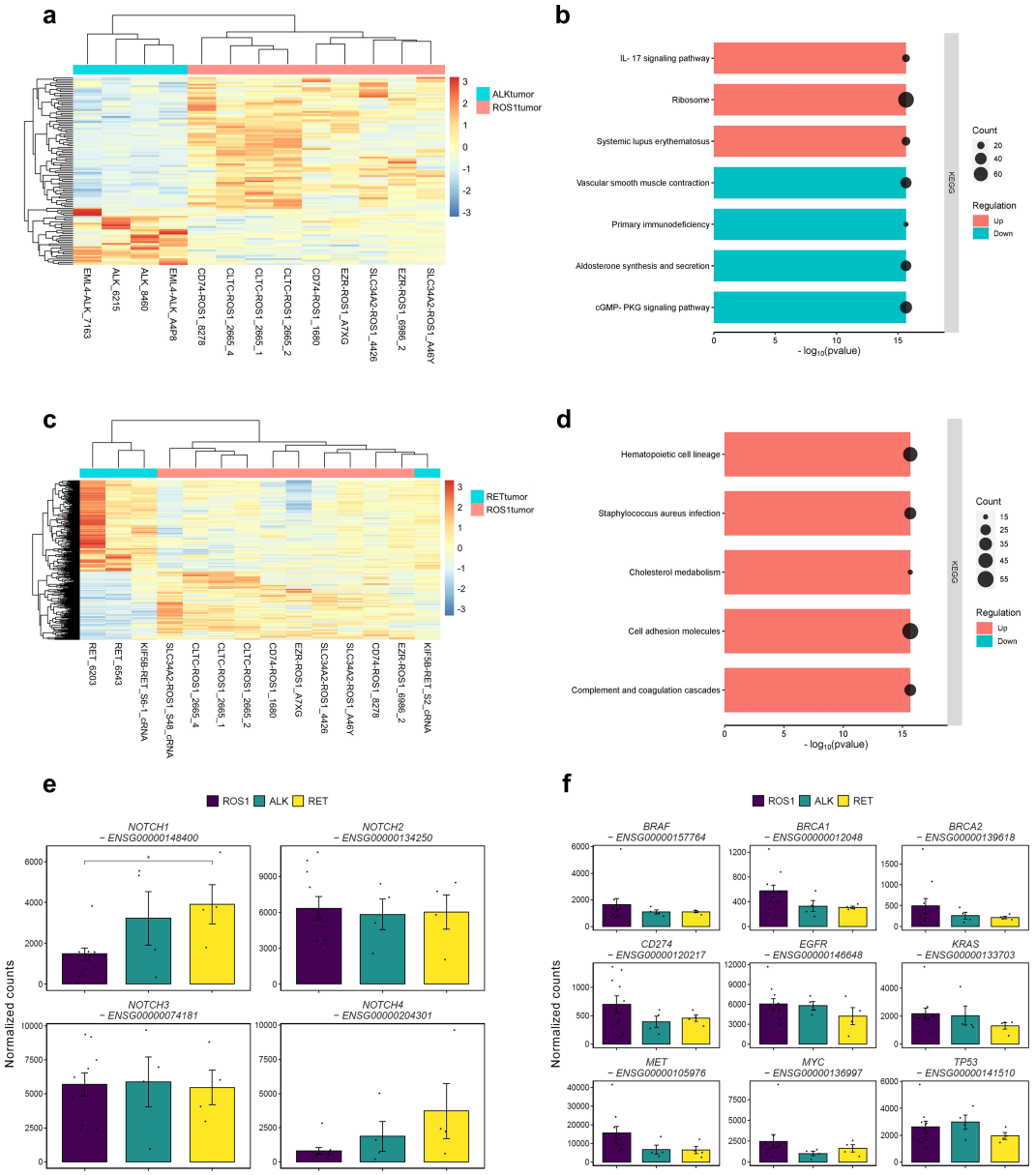


FIGURE 2
Differences between *ROS1+* and *ALK+/RET+* tumors. **(A)** Clustergram reflecting the differences between *ROS1+* and *ALK+* tumor specimens. **(B)** GSEA resulting from the significantly DEGs between *ROS1+* and *ALK+* tumors. **(C)** Clustergram reflecting the differences between *ROS1+* and *RET+* tumor specimens. **(D)** GSEA resulting from the significantly DEGs between *ROS1+* and *RET+* tumors. **(E)** Gene expression levels of the *NOTCH* family and **(F)** canonical tumor-suppressor genes and oncogenes.

contrast, genes such as *ALDOA (aldolase A), U2AF1 (U2 small nucleolar RNA auxiliary factor 1), MFSD14A (major facilitator superfamily domain containing 14A) and FKBP2 (FKBP prolyl isomerase 2)* were upregulated in patient-derived cell lines.

In order to showcase the representativity of the patient-derived cell lines used in the study, special attention was paid to genes involved in extrinsic resistance mechanisms towards TKIs described in *ROS1+* NSCLC patients (Figure 3C). The statistical tests were performed relatively to *CD74-ROS1* tumor specimens given that it constitutes the most frequent rearrangement subtype. *EGFR* was significantly overexpressed in CUTO-38 cells ($p < 0.001$), CUTO-37 ($p = 0.012$), CUTO-33 ($p = 0.04$) and CUTO-23 ($p = 0.003$). *ERBB2*, whose product is HER2, was significantly downregulated in *CLTC-ROS1* tumors ($p < 0.001$), CUTO-37, CUTO-28, CUTO-27 and CUTO-33 ($p < 0.001$). Contrarily, it was upregulated in CUTO38 cells ($p = 0.001$). A similar trend was observed in *ERBB3* expression, except for a significant downregulation in HCC-78 cells ($p = 0.006$) instead of an overexpression in CUTO-38 cells. *GJB2* was clearly expressed in *CD74-* and *CLTC-ROS1* tumors, being significantly higher in the latter case ($p < 0.001$). In contrast, this gene was practically not expressed in the remaining tumor subtypes and cell lines. In parallel, *MET* was upregulated in CUTO38 and CUTO-37 ($p < 0.001$) together with CUTO-33 ($p = 0.024$) and CUTO28 ($p = 0.011$). *MYC* was overexpressed in HCC-78 cells ($p = 0.021$),

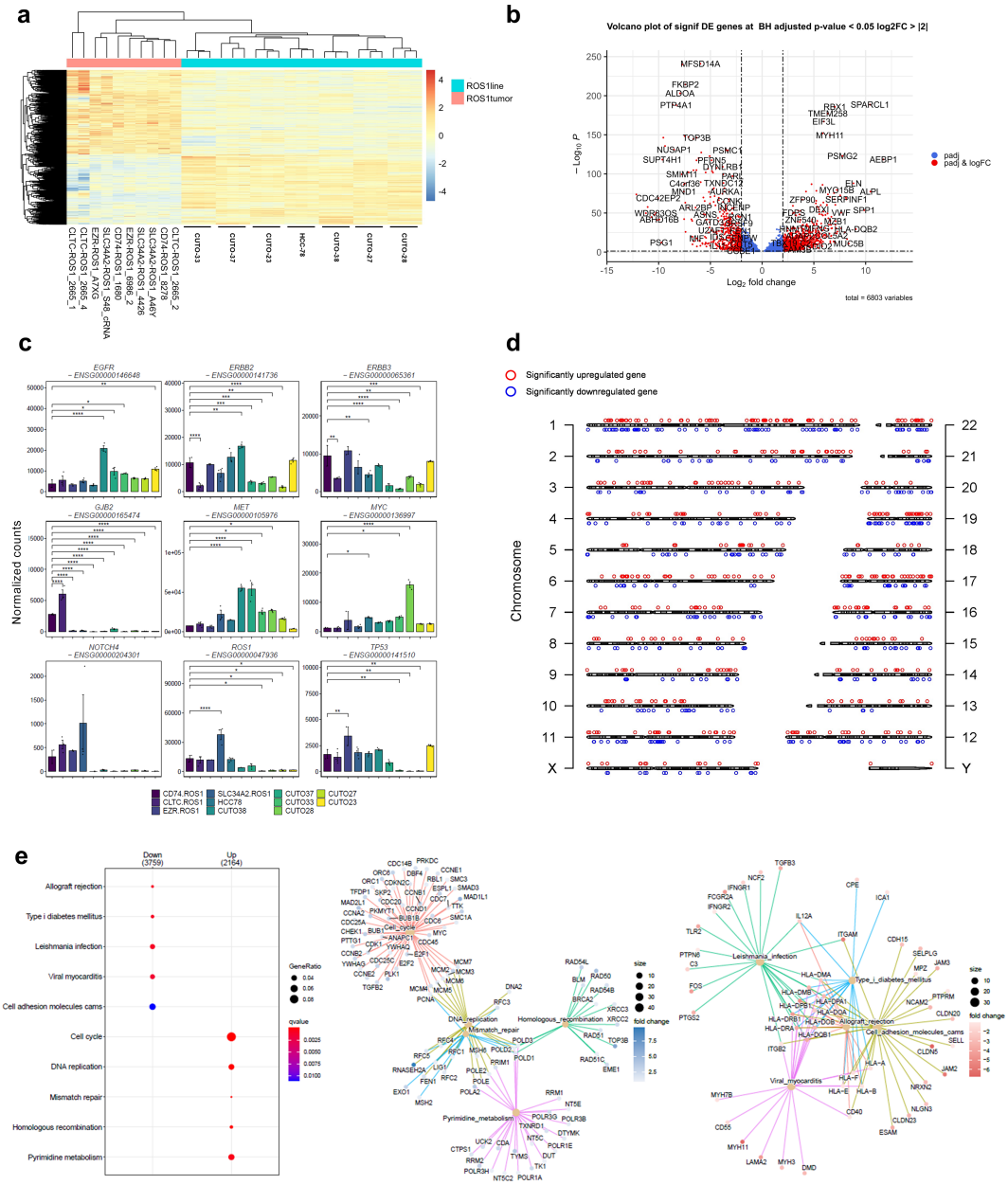


FIGURE 3
Transcriptomic differences between *ROS1*+ tumors and patient-derived cell lines. **(A)** Clustergram highlighting the differences between *ROS1*+ tumor specimens and patient-derived cell lines. **(B)** Volcano plot containing more than 6,000 DEGs. **(C)** Differential expression of relevant oncogenes and tumor-suppressor genes across *ROS1*+ NSCLC tumor subtypes and patient-derived cell lines. **(D)** Distribution across chromosomes of the top 500 significant up- and downregulated genes between tumor specimens and cell lines. **(E)** GSEA elucidating the KEGG pathways up- or down-regulated in patient-derived cell lines versus tumor specimens.

CUTO33 ($p = 0.017$) and CUTO28 ($p < 0.001$). No differences were detected concerning *NOTCH4* expression. However, the endogenous *ROS1* levels varied across samples. *SLC34A2-ROS1* tumors reflected an upregulation ($p < 0.001$) whilst CUTO-33, CUTO-28, CUTO-27 and CUTO-23 showed lower *ROS1* levels ($p = 0.01$). Finally, *TP53* was upregulated in *EZR-ROS1* tumors ($p<0.001$) and downregulated in CUTO33 ($p = 0.007$), CUTO-28 ($p = 0.004$) and CUTO-27 cells ($p = 0.005$).

To investigate whether specific genomic regions contained a notable concentration of differentially expressed genes (DEGs), the top 500 significant DEGs were mapped to their corresponding locations on the genome (Figure 3D). This visualization aimed to provide a graphical representation of the DEG distribution. The analysis revealed that the distribution of these genes was relatively uniform across the genome, with the exception of chromosome 7, which exhibited a higher density of DEGs compared to other chromosomes. GSEA that included the 6,803 DEGs and using the "chromosomalLocation" function revealed a significant enrichment of overexpressed genes in cell lines located in the cytogenetic band chr7p22.3 (Supplementary Figure 1). Interestingly, copy number

gains involving this genomic region known to harbor oncogenes such as *UNCX, FAM20C, MAD1L1* and *PDGFA* have been reported during the immortalization process of patient-derived small cell lung cancer lines (37).

GSEA identified pathways whose activation differed between tumor samples and cell lines (Figure 3E). As expected, immune response-related gene sets like allograft rejection, *Leishmania* infection and viral myocarditis together with cell adhesion molecules (CAMs) were significantly overexpressed in tumor specimens. Contrarily, gene sets corresponding to the cell cycle, DNA replication, DNA repair and RNA transport were enriched in CUTO and HCC-78 cell lines. Therefore, these results highlight the major biological differences between tumor samples and cell lines and can be summarized in cell cycle, DNA repair and inflammatory signatures.

3.5 Differences between *ROS1*+ NSCLC cell lines encompass EMT- and Myc-related hallmarks

The subsequent exploratory analysis focused exclusively on *ROS1*+ NSCLC patient-derived cell lines. As illustrated in Figure 4A, these cell lines exhibit a remarkable phenotypic heterogeneity. Their morphology and colony formation do not adhere to a single pattern, leading us to hypothesize that such differences might be reflected at the transcriptomic level. Despite the limited sample size, the differential gene expression analysis unveiled intriguing features, as depicted in the principal component analysis (PCA) shown in Figure 4B. The majority of cell lines clustered together, with a clear separation along the first principal component, which accounted for 72.44% of the variance. CUTO-37 and CUTO-38 were the two cell lines that clustered more distinctly from the rest. This observation is further supported by the dendrogram presented in Figure 4C.

Next, we identified similar expression patterns per each cell line through a k-means sample clustering. Assuming that each cell line constitutes a separate group and given that the biological replicates clustered together in each line, a K=7 was chosen. The resulting clustergram is shown in Figure 4D, and the following clusters are: 1: HCC-78 (450 genes), 2: CUTO-23 (314 genes), 3: CUTO-38 (764 genes), 4: CUTO-28 (1094 genes), 5: CUTO-37 (1127 genes), 6: CUTO-27 (164 genes) and 7: CUTO-33 (724 genes). The subsequent GSEA performed in order to identify the significantly enriched hallmarks in each cluster is shown in Figure 4E. HCC-78 transcriptome was enriched in KRas-mediated signaling and interferon responses, CUTO-38 cells showed an enrichment in cholesterol and reactive oxygen species (ROS) together with androgen response and bile acid metabolism. CUTO-37 cells were enriched in the EMT, hypoxia and TNFα signaling via NF-ƙB, among other hallmarks. CUTO-33 cells were EMT- and UV response-enriched hallmarks. CUTO-28 cells showed an enrichment in estrogen responses, Myc targets, mTORC1 signaling and E2F targets hallmarks. Finally, CUTO-27 depicted an upregulation in TNFα signaling via NF-ƙB hallmark. No significant hallmarks were identified in CUTO-23 cells. The networks of the GSEA hallmarks identified in the two most distant cell lines according to the dendrogram, CUTO-28 and CUTO-37 are represented in Figure 4F (CUTO-28) and Figure 4G (CUTO-37) respectively.

3.6 The *ROS1* fusion partner genes modulate modestly the tumor transcriptome

The molecular subtypes defined by the ROS1 fusion partner genes have been shown to influence downstream signaling due to the different subcellular localizations of the resulting fusion proteins (25). Therefore, we hypothesized that the differential activation of signaling pathways might be reflected at the gene expression level. Three out of the four *ROS1*-rearranged tumor subtypes were included in the analysis, as there were at least two independent samples per fusion type (*CD74-ROS1*, *SLC34A2-ROS1*, and *EZR-ROS1*). The PCA plot containing PC1 and PC2 indicates that *EZR-ROS1* tumors cluster remarkably further from *CD74-* and *SLC34A2-ROS1* tumor specimens, as shown in Figure 5A. The complementary dendrogram showcases the similarity between *CD74-* and *SLC34A2-ROS1* tumors (Figure 5B).

Next, a DEG analysis was conducted to unveil the distinctive traits of each *ROS1* rearrangement type, done through pairwise comparisons. The cut-off conditions were $\log_2$(fold change |2|) > and FDR < 0.05. The obtained volcano plots and heatmaps per each rearrangement are shown in Supplementary Figure 2B, for the *CD74-ROS1* signature, 2c for *SLC34A2-ROS1* signature and 2d for *EZR-ROS1* signature respectively. The functional annotation of the DEGs in each subtype revealed a significant upregulation of *CD74-ROS1* tumors in EMT hallmark and a downregulation of TGF-β signaling (Figure 5C). In addition, *EZR-ROS1* tumor specimens were characterized by an enrichment in β-catenin signaling and the downregulation of a G2M checkpoint hallmark (Figure 5D). It is worth noting that no significant enrichment was observed in the *SLC34A2-ROS1* subtype.

Finally, the correlation between *ROS1* expression and selected oncogenes of interest was determined for the *CD74-ROS1*, *SLC34A2-ROS1*, and *EZR-ROS1* subtypes (Supplementary Figure 2). Notably, *MET* (Spearman correlation coefficient = 0.964), *BRAF* (0.963), *MYC* (0.714), and *EGFR* (0.39) exhibited positive correlations with ROS1 expression. Conversely, the expression of transcription factors known to drive epithelial-mesenchymal transition (EMT), such as *SNAI1* (-0.678) and *TWIST1* (-0.785), negatively correlated with *ROS1* expression. However, these correlations did not reach statistical significance, which might be attributed to the limited sample size in each group.

3.7 The connexin-encoding *GJB2* is expressed in a *CD74-* and *CLTC-ROS1*-dependent manner

Based on our hypothesis, we proceeded to evaluate *GJB2* expression as a potential prognostic marker for ROS1-rearranged

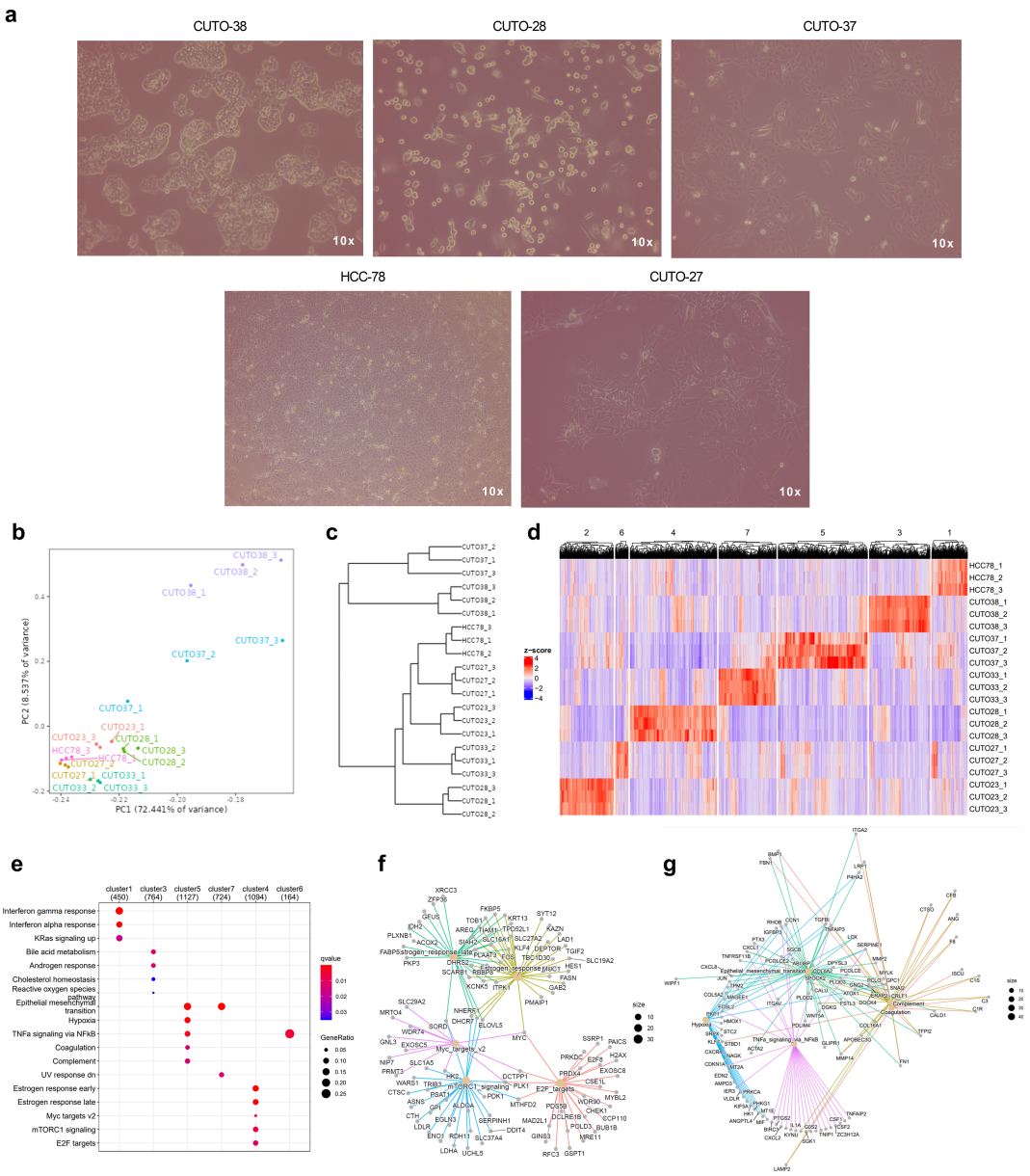


FIGURE 4
Transcriptomic traits of *ROS1*+ patient-derived NSCLC cell lines. **(A)** Phenotypes of some of the cell lines characterized in the study. Phase contrast microscopy images taken at 10X magnification. **(B)** Principal component plot. **(C)** Dendrogram and **(D)** clustergram reflecting the differences across cell lines. **(E)** GSEA indicating the significantly enriched hallmarks in each cell line. **(F)** Enriched hallmarks in CUTO-28 and **(G)** CUTO-37 cell lines.

NSCLC, considering its upregulation compared to *ALK*+ and *RET*+ tumors, as well as normal adjacent lung tissue. Statistical significance was achieved only when compared to normal lung tissue ($p = 0.042$) (Figure 6A). This observation might be attributed to the fact that *GJB2* overexpression occurs only in specific *ROS1* fusion subtypes. Figure 6B illustrates the normalized *GJB2* counts across the investigated *ROS1* rearrangement types, revealing that *GJB2* expression was present only in *CD74-ROS1* and *CLTC-ROS1* samples, with a significant upregulation in the latter subtype ($p = 0.016$). These findings were validated through RT-qPCR in our cell line models, where transcript levels are expressed as calibrated and normalized relative mRNA quantities (CNRQ) ± SEM (Figure 6C). *GJB2* expression was not detected in HCC-78 cells harboring the *SLC34A2-ROS1* fusion, in contrast to CUTO-28 (*TPM3-ROS1*) and CUTO-37 (*CD74-ROS1*) cell lines, which showed significant *GJB2* expression ($p < 0.0001$ in both comparisons). Consequently, the higher levels of *GJB2* transcript in CUTO-28 and CUTO-37 lines suggest that *GJB2* is actively expressed by tumor cells and that its expression varies depending on the specific *ROS1* fusion type.

Moreover, we evaluated the prognostic significance of high *GJB2* expression in the lung adenocarcinoma (LUAD) cohort from The Cancer Genome Atlas (TCGA) (Figure 6D). LUAD patients with *GJB2* overexpression exhibited a poorer prognosis compared to those in the low-*GJB2* expression group (Hazard ratio for high *GJB2* expression = 1.8; p-value for high *GJB2* expression < 0.001). Next, we validated this finding in an independent LUAD

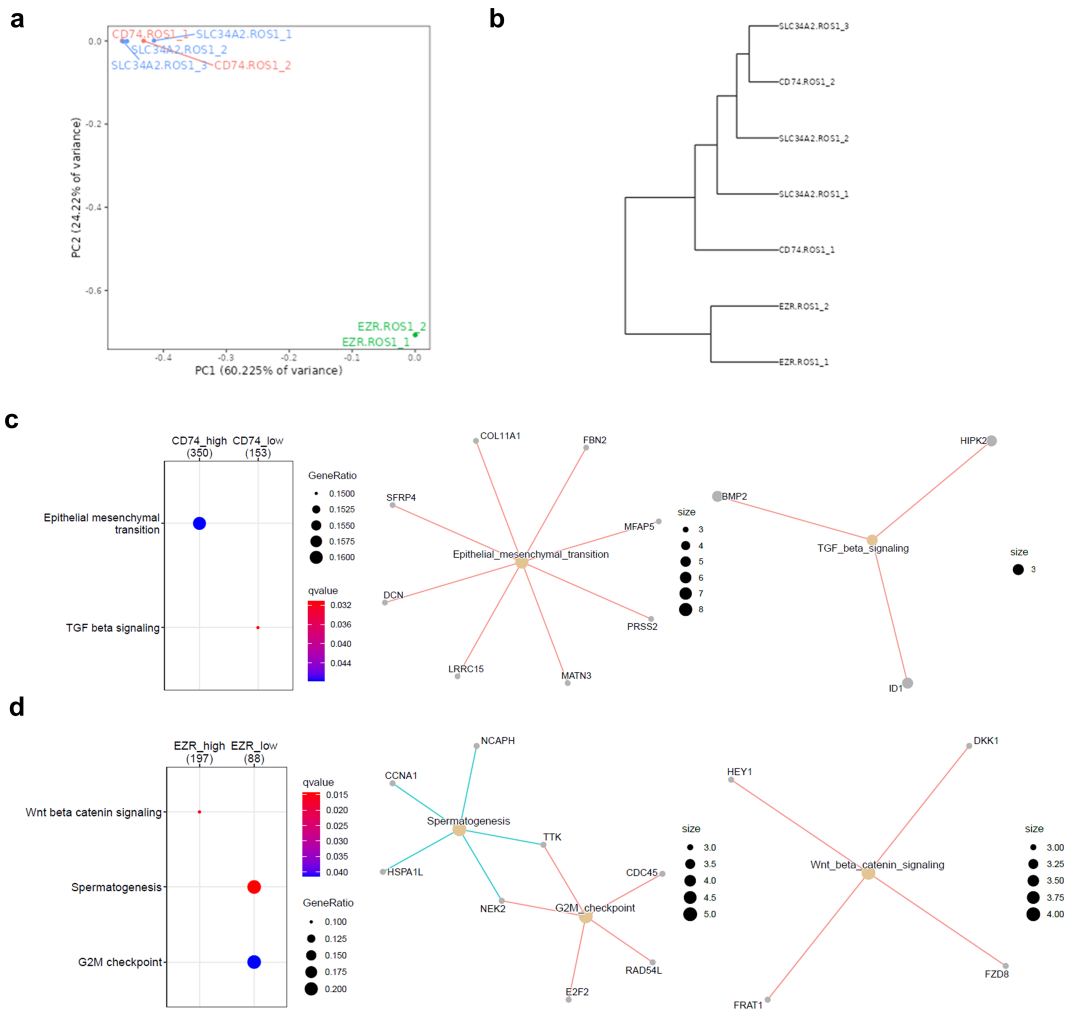


FIGURE 5
Impact of the *ROS1* fusion partner in the tumor transcriptome. **(A)** Principal component analysis (PCA) of *CD74-*, *SLC34A2-* and *EZR-ROS1* tumor samples. **(B)** Clustergram of the cell lines based on their transcriptomic traits **(C)** GSEA depicting the enriched hallmarks in *CD74-ROS1* and **(D)** *EZR-ROS1* NSCLC samples.

patient cohort that combined several studies. As shown in Supplementary Figure 4, the results were concordant with the TCGA cohort. Taking these observations into account, *GJB2* expression might explain the variability among *ROS1*+ NSCLC patient disease outcomes. As reported by Liu et al., the potential mechanism behind the pro-tumorigenic role of high *GJB2* expression might be the upregulation of the SPP1 signaling pathway (24). Furthermore, we mapped the reported *GJB2* variants within the same TCGA LUAD cohort onto the gene structure, as illustrated in Figure 6E, to identify potential gain-of-function mutations. Among the variants, four were missense mutations: L10M and G130C, located in the connexin domain; and R165Q and V178M, present in the cysteine-rich connexin domain of GJB2. Interestingly, only the V178M variant has been clinically described in the literature, reported as a pathogenic mutation identified in patients with autosomal-recessive hearing loss (38). This suggests that V178M is unlikely to be a gain-of-function mutation in the context of hearing loss. However, the impact of the GJB2 V178M variant on tumor cell biology remains to be elucidated.

3.8 Gene co-expression analysis proposes TFDP1 as master transcription factor in *ROS1*+ NSCLC

Unsupervised methods, such as gene co-expression analysis, enable the identification of gene clusters with positively or negatively correlated expression patterns, independent of sample category or experimental condition. In this study, we identified 18 gene clusters, each characterized by a correlation coefficient and a p-value associated with a specific trait or phenotype. Supplementary Figure 5 illustrates the resulting matrix, with columns representing the traits of the tumor samples included in the analysis. The strongest correlation was observed for the green-yellow cluster (coefficient = -0.48, p = 0.007) with respect to disease stage,

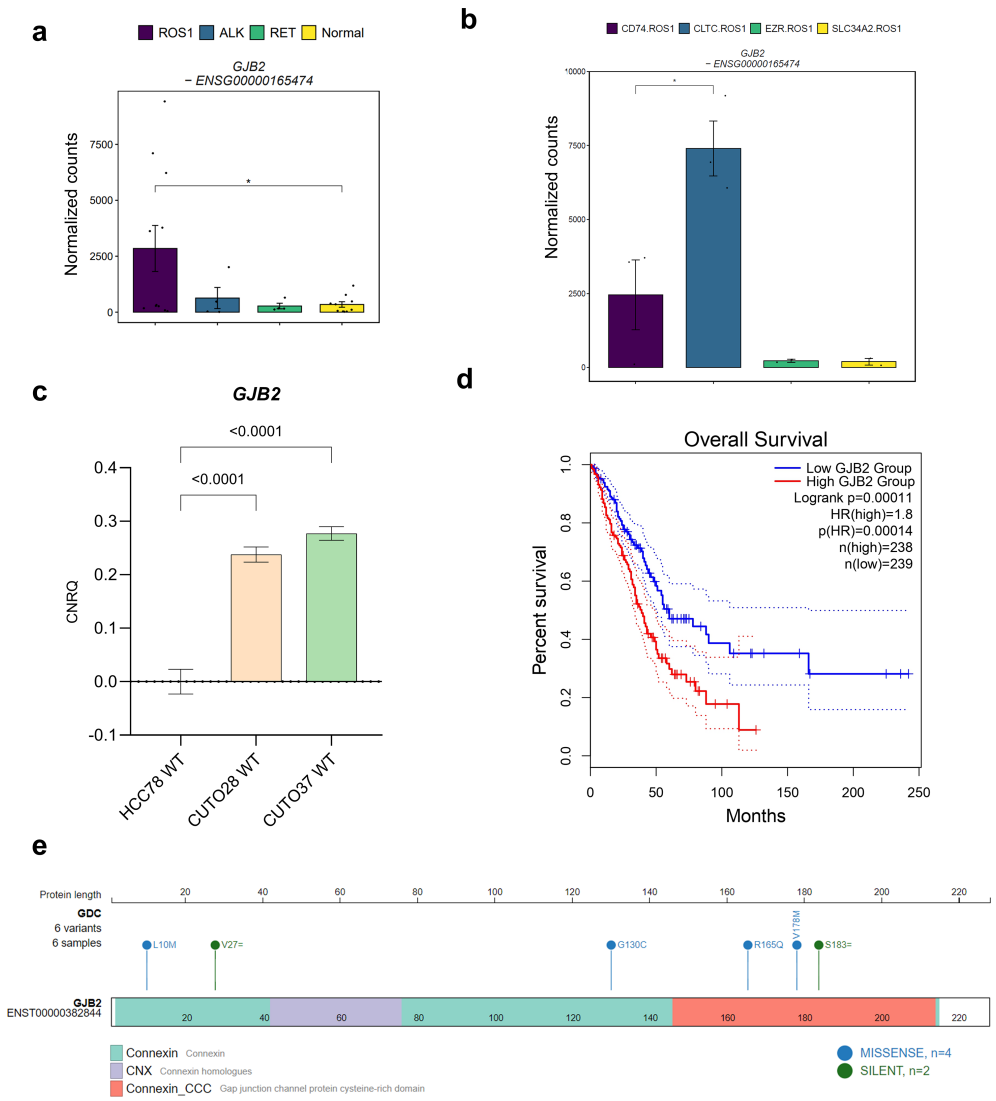


FIGURE 6
The role of *GJB2* expression in *ROS1*+ NSCLC. **(A)** *GJB2* expression levels across oncogene-driven NSCLC and normal lung tissue. **(B)** *GJB2* expression across different *ROS1* rearrangements in tumor specimens. **(C)** *GJB2* mRNA levels in HCC-78, CUTO-28 and CUTO-37 cell lines expressed as calibrated and normalized relative mRNA quantity (CNRQ) ± SEM. A 1-way ANOVA test was performed considering a Bonferroni correction and a chosen $\alpha = 0.05$. **(D)** Kaplan-Meier survival plot of The Cancer Genome Atlas (TCGA) lung adenocarcinoma (LUAD) cohort divided in *GJB2*-high or low expression levels. **(E)** *GJB2* single-nucleotide variants identified in LUAD patients.

suggesting that the expression patterns of genes within this cluster collectively explain a portion of the variability attributed to the disease stage of the samples. The observed eigengene expression patterns of the green-yellow module across samples (Figure 7A) highlight the variability in the direction and magnitude of gene expression among different sample types. These discrepancies can be attributed to the moderate correlation coefficient of the module, indicating that while the genes within the cluster share a common expression pattern, other factors may also influence their expression. The functional annotation of this gene cluster through over-representation analysis (ORA) revealed the involvement of genes in critical cellular processes such as cell cycle, DNA repair and replication, and the P53 signaling pathway (Figure 7B). These findings suggest that the disease stage of tumor samples can be partially explained by the coordinated expression of genes regulating these essential pathways.

The following module which showed moderate correlation between the genes is the red one (0.44, p=0.01). The trait that defines this gene cluster is the comparison between *ROS1*+ tumor specimens and normal or healthy lung samples. When carrying out ORA, the steroid biosynthesis, amino sugar and nucleotide sugar metabolism together with the small cell lung cancer KEGG pathways were identified (Figure 7C). Finally, with the aim to unveil master regulatory transcription factors that orchestrate the *ROS1*+ NSCLC transcriptome, an iRegulon-based analysis upon the red module was performed followed by an ORA annotation (Figure 7D). The top-ranked transcription factor (TF) for the red gene module was TFDP1. Considering only the target genes of the

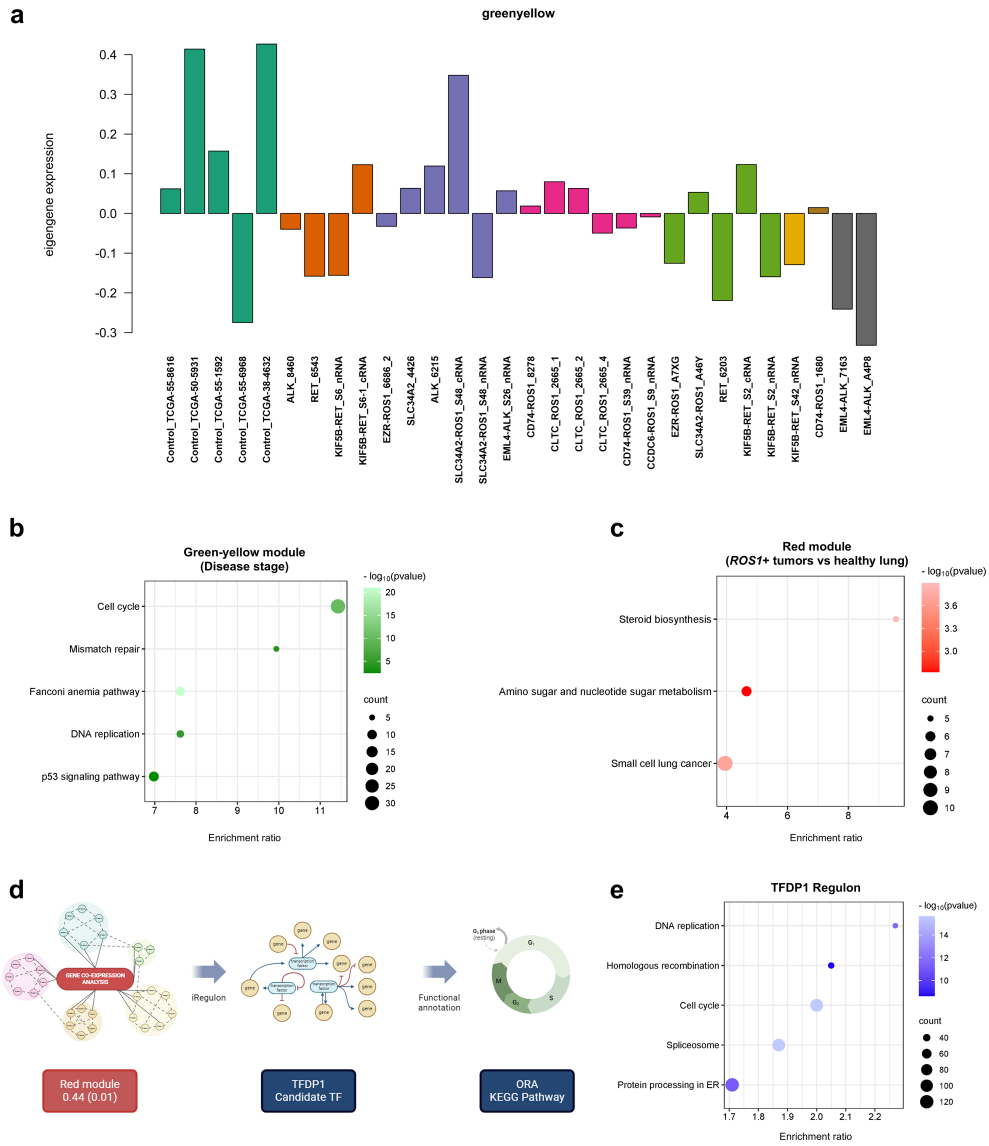


FIGURE 7
Gene co-expression analysis. **(A)** Eigengene expression patterns of the green-yellow module across samples. **(B)** ORA comprising the green-yellow gene module and **(C)** red module. **(D)** Workflow performed to identify key transcription factors (TFs) that regulate the expression of genes in a given co-expression module. **(E)** ORA performed with the TFDP1 regulon.

aforementioned TF, ORA was performed (Figure 7E) and it unveiled that DNA replication, homologous recombination, cell cycle, spliceosome and protein processing in ER were the significantly over-represented KEGG pathways in the TFDP1 regulon.

4 Discussion

The biology of oncogene-addicted lung neoplasms has been heterogeneously approached within the last decades. Partially due to differences in the prevalence of each molecular subtype, restricting access to patient-derived samples. In consequence, study of *ROS1*-, *ALK*- and *RET*- rearranged tumors remains modest beyond the kinase domain of their oncogenic fusion proteins. To unveil distinctive traits between oncogene-addicted lung adenocarcinomas, particularly between rare subtypes like *ROS1*+ NSCLC, aspects such as tumor transcriptome profiling might be enlightening. Although existing transcriptome studies lack the necessary depth and specificity required for a comprehensive characterization of *ROS1*-driven NSCLC, we hypothesized that reanalysis based on combined data can yield a more representative characterization of this tumor type (39–41).

Our findings point towards a relevant contribution of the nucleotide synthesis and cell adhesion pathways to the proliferation of *ROS1*+ lung adenocarcinomas. The highest overexpressed gene in *ROS1*+ tumor specimens and cell lines, *ISL1* provides new insights about this molecular subtype. This gene encodes a LIM homeodomain transcription factor which regulates insulin expression during homeostasis. Moreover, its pathological role in tumorigenesis has

been reported through the regulation of cyclin D1 and c-Myc genes in neuroblastoma, gastric cancer and non-Hodgkin lymphoma (42–44). *ISL1* has been recently reported to mediate lung morphogenesis, particularly, the adequate branching of the respiratory epithelium through the Sonic hedgehog (Shh) signaling pathway (45). In addition, *ISL1* is known to regulate processes like epithelial cell differentiation, adhesion and migration in NSCLC (46, 47). Importantly, Shh mediates processes like EMT, therefore a dysregulation of *ISL1* expression might enhance tumorigenesis via the acquisition of a mesenchymal phenotype of neoplastic cells. Li et al. provided evidence supporting this hypothesis by demonstrating that triptolide-mediated inhibition of the Sonic hedgehog (Shh) pathway could reverse chemoresistance in NSCLC cells (48). Regarding the upregulated cell adhesion signature in *ROS1*+ NSCLC, the gene encoding Ras interacting protein 1 (*RASIP1*) was found to be significantly overexpressed. RASIP1 plays a crucial role in the integrin activation process, which is essential for cell migration (49). Therefore, pharmacological inhibition of RASIP1 or its associated pathways might represent a potential strategy to prevent or reduce the development of metastasis in *ROS1*+ NSCLC.

Additionally, the significantly increased levels of *NOTCH1* in *RET*+ tumors might partially explain the higher rates of extrathoracic metastases in *RET*+ compared to *ROS1*+ NSCLC patients. The potential underlying mechanism might be the epithelial-to-mesenchymal transition (EMT)-promoting role of the Notch pathway that enhances cell migration (19, 22, 50). As depicted by two studies in NSCLC, Notch-1 and Notch-3 respectively promote EGFR-TKI resistance and maintain a stem-like status. Remarkably, their pharmacological inhibition decreased cell growth and migration, autophagy and increased the cell apoptotic activity (18, 20, 21). Hence, the Notch pathway constitutes a crucial player regarding the regulation of EMT in NSCLC, becoming a drug target for metastatic disease.

It is known that higher *GJB2* expression levels in a subset of cancer-associated fibroblasts (CAFs) induce stromal tumor fibrosis, enhancing chemotherapy resistance in solid tumors (51). These observations were complemented by the significant positive correlation between *GJB2* expression and the tumor infiltration by CAFs (52). Specifically in LUAD, the promoter of *GJB2* was found to be hypomethylated, possibly leading to gene overexpression. Moreover, the functional annotation of DEGs in *GJB2*-overexpressing LUAD revealed an enrichment of the PI3K/AKT and ECM-receptor interaction KEGG pathways (53). However, the functional characterization of *GJB2* missense variants identified in patients is currently lacking, limiting our ability to draw definitive conclusions regarding their impact on gene function and disease pathogenesis. In parallel, the upregulation of *IL20RB* in *ROS1*+ NSCLC tumors and cell lines might constitute one of the mechanisms that mediate bone metastasis. According to a recent study performed in The Netherlands, around 1/3 of stage IV *ROS1*+ NSCLC patients present bone metastases at diagnosis (54). IL20RB has been described to activate the JAK/STAT pathway upon stimulation with interleukin 19 (IL-19). In consequence, phosphorylated STAT3 translocates to the nucleus of tumor cells promoting the secretion of granulocyte-macrophage colony stimulating factor (GM-CSF), a cytokine that induces IL-19 synthesis in osteoclasts (33). Importantly, IL20RB neutralization with antibodies reduced the metastatic potential of tumor cells *in vivo*; therefore this becomes a relevant strategy to be evaluated in clinical trials in order to manage bone metastases in advanced stages.

The identification of the enriched systemic lupus erythematosus (SLE) signature in *ROS1*+ NSCLC might be *a priori* a puzzling finding. SLE is an autoimmune disease characterized by a dysregulation of cytokines, T cells, B cells and macrophages (55). However, this signature has been previously identified in lung cancer with a higher incidence in the adenocarcinoma histological subtype and non-smoker women (56). Although the causal explanation of this findings is not fully understood, the manifestations of the SLE signature are consistent with the higher risk of thromboembolic events reported in *ROS1*+ NSCLC patients (57–59). Mechanistically, SLE-induced coagulopathy is caused by autoantibodies which target endothelial cells. The resulting damage triggers the coagulation cascade (60). Therefore, in the context of *ROS1*+ NSCLC, the expression of a SLE gene signature might be mediated by cell adhesion molecules; another important gene set identified in this study (61, 62). The potential aberrant expression of cell adhesion molecules like integrins and gap junction proteins could initiate the coagulation cascade, explaining the increased susceptibility of thromboembolic events observed in patients.

Importantly, the main limitation of this study is the low sample size of *ROS1*-, *ALK*- and *RET*-rearranged tumors, which is explained by the low prevalence of each tumor subtype. The limited number of tumor specimens and cell lines included in our analysis may affect the generalizability of our findings, necessitating further validation using additional patient-derived samples and experimental models. Moreover, the patients involved in the study received various treatment regimens, which likely impacted gene expression profiles, introducing another potential limitation of our study. However, the identified signatures emerged from analyzing tumor samples and patient-derived cell lines independently collected and sequenced, increasing the robustness of the presented evidence. By employing RT-qPCR, we validated the expression of *IL20RB* in our *ROS1*+ cell lines as well as the fusion-specific *GJB2* expression pattern. Consequently, we confirm the validity of such findings in independent samples and using an orthogonal method to NGS like RT-qPCR. In addition, complementing the dataset with patient-derived cell lines provides two main advantages. First, the confirmation of the observations in tumor samples, a biologically comprehensive sample type. Second, a detailed assessment of the cell line transcriptomes unveiled converging gene expression patterns with tumor samples, which can be summarized in the enrichment of cell adhesion molecules and nucleotide synthesis signatures. The shared traits offer a new insight about the most representative features of cell lines' tumor physiology, shown by a dependency on *ISL1* expression which links its activity with glucose metabolism and the acquisition of a mesenchymal phenotype.

As seen in our study, differences were noticed between tumor samples and cell lines, reflected by the more than 6000 significantly DEGs. Among them, key oncogenes like *MYC* and *MET* were found to be upregulated in cell lines; possibly as a result of the spontaneous

immortalization of cell lines. Moreover, the observed cluster of overexpressed genes on chromosome 7, in line with previously reported copy number gains of 7p22.3 in cell lines, must be considered when extrapolating results obtained *in vitro*. Because this region is rich in oncogenes, differences should be closely regarded when designing experiments involving pathways in which these genes participate (16, 37). It is also worth noting that HCC-78 cells clustered among CUTO lines, indicating that their transcriptomes are very similar despite being cell lines established in different time points. Nevertheless, the phenotypical and transcriptomic variability among *ROS1*+ patient-derived cell lines supports the heterogeneous *ROS1*+ NSCLC patient outcomes in the clinical setting, reinforcing them as valuable experimental models.

The mild impact of different *ROS1* fusion partner genes contrasts with the work by Neel et al. They demonstrated SLC34A2-ROS1 and SDC4-ROS1 fusions strongly activate the MAPK pathway, which was not seen in CD74-ROS1, due to their differential subcellular localization. The discordance with our results might be explained by three reasons. First, the included data did not include determination of phosphorylated Erk 1/2 protein through immunoblotting, which would have been the most specific method to determine the activation levels of the MAPK pathway. Second, the overlap between the target genes of the MAPK pathway and other ROS1 downstream signaling pathways, such as the JAK/STAT or mTOR/AKT, might mask the differential MAPK pathway activation levels from a transcriptome perspective. Third, Neel et al. also mentioned that a shorter CD74-ROS1 isoform can localize in endosomes and plasma membrane instead of the ER, and as such activate the MAPK pathway as well (25).

Nonetheless, the *GJB2* overexpression found in *CLTC-ROS1* and *CD74-ROS1* subtypes, might have implications in disease progression. First because *GJB2*, expressed by tumor cells and cancer-associated fibroblasts (CAFs), contributes to ECM remodeling and activates the SPP1/PI3K/AKT signaling pathway in lung adenocarcinoma (24, 63). Second, due to the pro-metastatic role of *KRT16*, which was found to be upregulated in the same two patient subsets. (Supplementary Figure 6). This cytokeratin upregulates the synthesis of vimentin in lung cancer cells (64). Consequently, this hypothesis is concordant with the increased likelihood of *CD74-ROS1*+ NSCLC patients to develop brain metastases (65). On top of that, Wang et al. identified that *CD74-ROS1*+ bone metastatic NSCLC cells secreted CCL5 through STAT3 activation to recruit macrophages. In this interaction, the tumor-promoting M2 macrophages stimulate tumor cells and induce EMT via TGF-β pathway stimulation (66). These results align with the upregulated EMT signature that we report in *CD74-ROS1* specimens. Similarly, the downregulation of the TGF-β pathway that we found in tumor specimens can be explained by the lack of interaction between tumor cells and the bone niche in the samples that we analyzed.

Applying gene co-expression analysis in the tumor samples lead to the detection of the cell cycle and amino sugar and nucleotide sugar metabolism pathway over representation within the *ROS1*+ lung adenocarcinoma subtype. This approach allowed an unsupervised characterization of samples aimed to identify functionally related genes (67). The confirmation of these observations in additional samples, experimental models and independent cohorts might open the door to find new actionable targets in a *ROS1*-specific manner. In parallel, this approach combined with the iRegulon tool pointed towards TFDP1, a transcription factor known to interact with E2F, another essential regulator whose targets were significantly enriched in CUTO-28 cells. Moreover, *TFDP1* amplification has been described in lung cancer and esophageal squamous cell carcinomas (68). It is important to note that an activation of the TFDP1/E2F1 axis in lung cancer results in the attenuation of the p53 pathway, mediated by COMMD9 (69). Thus, this transcription factor represents a novel research object holding therapeutic potential.

Focused efforts concerning *ROS1*+ NSCLC patients are currently oriented towards refining the treatment scheme upon improved patient stratification, overcoming drug resistance and understanding the disease risk factors (70). Our study contributes to two of these aspects. First, treatment-wise, combination strategies of TKIs and monoclonal antibodies targeting IL20RB might be beneficial to treat bone metastases. Second, further functional studies are required to confirm the role of the nucleotide synthesis pathway in *ROS1*+ NSCLC, its targeting might enhance the sensitivity towards checkpoint inhibitors as reported by Wu et al. (71) Besides, the inhibition of nucleotide synthesis through mTORC1/IMPDH targeting is known to induce replication stress, ultimately resulting in apoptosis (72). Thus, exploring combinations of TKIs and IMPDH inhibitors could be useful to address heavily pre-treated cases. Third, *GJB2* expression could be employed as a prognostic biomarker. Collectively, such observations have the potential to enhance the tailoring of therapies to different patient subsets and help to predict disease outcomes.

5 Conclusion

In the present study we perform an in-depth characterization of *ROS1*+ NSCLC using two complementary approaches: differential gene expression and gene co-expression analysis. Our results point towards *IL20RB*, the nucleotide synthesis and cell adhesion pathways as specific signatures compared to *ALK*+ and *RET*+ tumors. Importantly, they constitute targetable alterations which could be co-inhibited together with ROS1. Moreover, we report differences in oncogene expression such as *MYC*, *MET* and *BRAF* between *ROS1*+ tumor samples and cell lines, which should be taken into account when interpreting *in vitro* experiments. Finally, we propose ISL1 and TFDP1 as candidate transcription factors that complement the oncogenic dependencies of *ROS1*+ NSCLC through cyclin D1, c-Myc and the Sonic hedgehog (Shh) pathway. Furthermore, the identification of the enriched systemic lupus erythematosus (SLE) signature might be related to the higher risk of thromboembolic events in *ROS1*+ NSCLC patients. In addition, *GJB2* was found overexpressed in *CD74*- and *CLTC-ROS1*+ tumor specimens and cell lines, which positively correlates with patients

presenting a poor prognosis. Despite the limited sample size, the robustness of our evidences is supported by the independent validation of *IL20RB* and *GJB2* expression using RT-qPCR. Collectively, the present study broadens our understanding of the molecular alterations in *ROS1+* NSCLC, paving the path towards novel therapeutic strategies.

Data availability statement

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

Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used. Ethical approval was not required for the studies on animals in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

MT: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. KO: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. GV: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. GV: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing. LM: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This project was founded by FWO (Research Foundation Flanders) -"Kom op tegen kanker". Project number: G094820N.

Acknowledgments

We would like to thank Prof. Dr. Robert C. Doebele and his team for kindly providing us the CUTO cell lines, as part of a research collaboration agreement with the University of Colorado.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1408697/full#supplementary-material

References

1. Drilon A, Jenkins C, Iyer S, Schoenfeld A, Keddy C, Davare MA. ROS1-dependent cancers - biology, diagnostics and therapeutics. *Nat Rev Clin Oncol*. (2021) 18:35–55. doi: 10.1038/s41571-020-0408-9

2. Parikh DA, Walia G, Freeman-Daily J, Hennink M, Tomalia T, Buonanno L, et al. Characteristics of patients with ROS1+ Cancers: results from the first patient-designed, global, pan-cancer ROS1 data repository. *JCO Oncol Pract*. (2020) 16:e183–9. doi: 10.1200/JOP.19.00135

3. Russo A, Lopes AR, McCusker MG, Garrigues SG, Ricciardi GR, Arensmeyer KE, et al. New targets in lung cancer (excluding EGFR, ALK, ROS1). *Curr Oncol Rep*. (2020) 22:48. doi: 10.1007/s11912-020-00909-8

4. Remon J, Pignataro D, Novello S, Passiglia F. Current treatment and future challenges in ROS1- and ALK-rearranged advanced non-small cell lung cancer. *Cancer Treat Rev*. (2021) 95:102178. doi: 10.1016/j.ctrv.2021.102178

5. Choudhury NJ, Schneider JL, Patil T, Zhu VW, Goldman DA, Yang SR, et al. Response to immune checkpoint inhibition as monotherapy or in combination with chemotherapy in metastatic ROS1-rearranged lung cancers. *JTO Clin Res Rep*. (2021) 2. doi: 10.1016/j.jtocrr.2021.100187

6. Roys A, Chang X, Liu Y, Xu X, Wu Y, Zuo D. Resistance mechanisms and potent-targeted therapies of ROS1-positive lung cancer. *Cancer Chemother Pharmacol*. (2019). doi: 10.1007/s00280-019-03902-6

7. Dziadziuszko R, Le AT, Wrona A, Jassem J, Camidge DR, Varella-Garcia M, et al. An activating KIT mutation induces crizotinib resistance in ROS1-positive lung cancer. *J Thorac Oncol*. (2016) 11:1273–81. doi: 10.1016/j.jtho.2016.04.001

8. Ku BM, Bae YH, Lee KY, Sun J-M, Lee S-H, Ahn JS, et al. Entrectinib resistance mechanisms in ROS1-rearranged non-small cell lung cancer. *Invest New Drugs*. (2020) 38:360–8. doi: 10.1007/s10637-019-00795-3

9. Lin JJ, Langenbucher A, Gupta P, Yoda S, Fetter IJ, Rooney M, et al. Small cell transformation of ROS1 fusion-positive lung cancer resistant to ROS1 inhibition. *NPJ Precis Oncol*. (2020) 4:21. doi: 10.1038/s41698-020-0127-9

10. Drilon A, Ou S-HI, Cho BC, Kim D-W, Lee J, Lin JJ, et al. Repotrectinib (TPX-0005) is a next-generation ROS1/TRK/ALK inhibitor that potently inhibits ROS1/TRK/ALK solvent- front mutations. *Cancer Discovery*. (2018) 8:1227–36. doi: 10.1158/2159-8290.CD-18-0484

11. Drilon A, Horan JC, Tangpeerachaikul A, Besse B, Ou SHI, Gadgeel SM, et al. NVL-520 is a selective, TRK-sparing, and brain-penetrant inhibitor of ROS1 fusions and secondary resistance mutations. *Cancer Discovery*. (2023) 13:598–615. doi: 10.1158/2159-8290.CD-22-0968

12. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, et al. NCCN guidelines insights: non-small cell lung cancer, version 2.2021. *J Natl Compr Cancer Network*. (2021) 19:254–66. doi: 10.6004/jnccn.2021.0013

13. Park S, Ahn B-C, Lim SW, Sun J-M, Kim HR, Hong MH, et al. Characteristics and outcome of ROS1-positive non-small cell lung cancer patients in routine clinical practice. *J Thorac Oncol*. (2018) 13:1373–82. doi: 10.1016/j.jtho.2018.05.026

14. Chang X, Liu Z, Man S, Roys A, Li Z, Zuo D, et al. Metastasis manners and the underlying mechanisms of ALK and ROS1 rearrangement lung cancer and current possible therapeutic strategies. *RSC Adv*. (2019) 9:17921–32. doi: 10.1039/C9RA02258A

15. Drilon A, Lin JJ, Filleron T, Ni A, Milia J, Bergagnini I, et al. Frequency of brain metastases and multikinase inhibitor outcomes in patients with RET–rearranged lung cancers. *J Thorac Oncol*. (2018) 13:1595–601. doi: 10.1016/j.jtho.2018.07.004

16. Al Bakir M, Huebner A, Martínez-Ruiz C, Grigoriadis K, Watkins TBK, Pich O, et al. The evolution of non-small cell lung cancer metastases in TRACERx. *Nature*. (2023) 616:534–42. doi: 10.1038/s41586-023-05729-x

17. Du J, Chen W, Yang L, Dai J, Guo J, Wu Y, et al. Disruption of SHH signaling cascade by SBE attenuates lung cancer progression and sensitizes DDP treatment. *Sci Rep*. (2017) 7. doi: 10.1038/s41598-017-02063-x

18. Xie M, He CS, Wei SH, Zhang L. Notch-1 contributes to epidermal growth factor receptor tyrosine kinase inhibitor acquired resistance in non-small cell lung cancer. *Vitro vivo. Eur J Cancer*. (2013) 49:3559–72. doi: 10.1016/j.ejca.2013.07.007

19. Galluzzo P, Bocchetta M. Notch signaling in lung cancer. *Expert Rev Anticancer Ther*. (2011) 11:533–40. doi: 10.1586/era.10.158

20. Ma Y, Li M, Si J, Xiong Y, Lu F, Zhang J, et al. Blockade of Notch3 inhibits the stem-like property and is associated with ALDH1A1 and CD44 via autophagy in non-small lung cancer. *Int J Oncol*. (2016) 48:2349–58. doi: 10.3892/ijo.2016.3464

21. Mur EB, Bernardo S, Papon L, Mancini M, Fabbrizio E, Goussard M, et al. Notch inhibition overcomes resistance to tyrosine kinase inhibitors in EGFR-driven lung adenocarcinoma. *J Clin Invest*. (2020) 130:612–24. doi: 10.1172/JCI126896

22. Yuan X, Wu H, Han N, Xu H, Chu Q, Yu S, et al. Notch signaling and EMT in non-small cell lung cancer: Biological significance and therapeutic application. *J Hematol Oncol*. (2014) 7. doi: 10.1186/s13045-014-0087-z

23. Ji X, Liu Y, Mei F, Li X, Zhang M, Yao B, et al. SPP1 overexpression is associated with poor outcomes in ALK fusion lung cancer patients without receiving targeted therapy. *Sci Rep*. (2021) 11:14031. doi: 10.1038/s41598-021-93484-2

24. Liu Z, Xiao Z, Wang X, Zhang L, Zhang Z. Ion channel gene GJB2 influences the intercellular communication by Up-regulating the SPP1 signaling pathway identified by the single-cell RNA sequencing in lung adenocarcinoma. *Front Oncol*. (2023) 13:1146976. doi: 10.3389/fonc.2023.1146976

25. Neel DS, Allegakoen DV, Olivas V, Mayekar MK, Hemmati G, Chatterjee N, et al. Differential subcellular localization regulates oncogenic signaling by ROS1 kinase fusion proteins. *Cancer Res*. (2019) 79:546–56. doi: 10.1158/0008-5472.CAN-18-1492

26. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. (2014) 15. doi: 10.1186/s13059-014-0550-8

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

28. Etoh K, Nakao M. A web-based integrative transcriptome analysis, RNAseqChef, uncovers the cell/tissue type-dependent action of sulforaphane. *J Biol Chem*. (2023) 299. doi: 10.1016/j.jbc.2023.104810

29. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinf*. (2008) 9. doi: 10.1186/1471-2105-9-559

30. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. (2019) 47:W199–205. doi: 10.1093/nar/gkz401

31. Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. (2016) 44:W90–7. doi: 10.1093/nar/gkw377

32. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. (2017) 45: W98–W102. doi: 10.1093/nar/gkx247

33. He Y, Luo W, Liu Y, Wang Y, Ma C, Wu Q, et al. IL-20RB mediates tumoral response to osteoclastic niches and promotes bone metastasis of lung cancer. *J Clin Invest*. (2022) 132. doi: 10.1172/JCI157917

34. Jee J, Lebow ES, Yeh R, Das JP, Namakydoust A, Paik PK, et al. Overall survival with circulating tumor DNA-guided therapy in advanced non-small-cell lung cancer. *Nat Med*. (2022) 28:2353–63. doi: 10.1038/s41591-022-02047-z

35. Negrao MV, Skoulidis F, Montesion M, Schulze K, Bara I, Shen V, et al. Oncogene-specific differences in tumor mutational burden, PD-L1 expression, and outcomes from immunotherapy in non-small cell lung cancer. *J Immunother Cancer*. (2021) 9. doi: 10.1136/jitc-2021-002891

36. Lee J, Park CK, Yoon H-K, Sa YJ, Woo IS, Kim HR, et al. PD-L1 expression in ROS1-rearranged non-small cell lung cancer: A study using simultaneous genotypic screening of EGFR, ALK, and ROS1. *Thorac Cancer*. (2019) 10:103–10. doi: 10.1111/1759-7714.12917

37. Coe BP, Lee EHL, Chi B, Girard L, Minna JD, Gazdar AF, et al. Gain of a region on 7p22.3, containing MADILI, is the most frequent event in small-cell lung cancer cell lines. *Genes Chromosomes Cancer*. (2006) 45:11–9. doi: 10.1002/gcc.20260

38. Posukh OL. Genetic etiology of hearing loss in Russia. *Hum Genet*. (2022) 141:649–63. doi: 10.1007/s00439-021-02327-7

39. Seo JS, Ju YS, Lee WC, Shin JY, Lee JK, Bleazard T, et al. The transcriptional landscape and mutational profile of lung adenocarcinoma. *Genome Res*. (2012) 22:2109–19. doi: 10.1101/gr.145144.112

40. Martínez-Ruiz C, Black JRM, Puttick C, Hill MS, Demeulemeester J, Larose Cadieux E, et al. Genomic–transcriptomic evolution in lung cancer and metastasis. *Nature*. (2023) 616:543–52. doi: 10.1038/s41586-023-05706-4

41. Dhanasekaran SM, Balbin OA, Chen G, Nadal E, Kalyana-Sundaram S, Pan J, et al. Transcriptome meta-analysis of lung cancer reveals recurrent aberrations in NRG1 and Hippo pathway genes. *Nat Commun*. (2014) 5. doi: 10.1038/ncomms6893

42. Liu Z, Hu W, Qin Y, Sun L, Jing L, Lu M, et al. Isl1 promotes gene transcription through physical interaction with Set1/Mll complexes. *Eur J Cell Biol*. (2023) 102. doi: 10.1016/j.ejcb.2023.151295

43. Li M, Sun C, Bu X, Que Y, Zhang L, Zhang Y, et al. ISL1 promoted tumorigenesis and EMT via Aurora kinase A-induced activation of PI3K/AKT signaling pathway in neuroblastoma. *Cell Death Dis*. (2021) 12. doi: 10.1038/s41419-021-03894-3

44. Zhang Q, Yang Z, Jia Z, Liu C, Guo C, Lu H, et al. ISL-1 is overexpressed in non-Hodgkin lymphoma and promotes lymphoma cell proliferation by forming a p-STAT3/p-c-Jun/ISL-1 complex. *Mol Cancer*. (2014) 13:181. doi: 10.1186/1476-4598-13-181

45. Huang R, Zhang C, Zheng Y, Zhang W, Huang H, Qiu M, et al. ISL1 regulates lung branching morphogenesis via Shh signaling pathway. *J Biol Chem*. (2023) 299. doi: 10.1016/j.jbc.2023.105034

46. Li H, Yin C, Zhang B, Sun Y, Shi L, Liu N, et al. PTTG1 promotes migration and invasion of human non-small cell lung cancer cells and is modulated by miR-186. *Carcinogenesis*. (2013) 34:2145–55. doi: 10.1093/carcin/bgt158

47. Wooten DJ, Groves SM, Tyson DR, Liu Q, Lim JS, et al. Systems-level network modeling of Small Cell Lung Cancer subtypes identifies master regulators and destabilizers. 15(10):e1007343 doi: 10.1371/journal.pcbi.1007343

48. Li LB, Yang LX, Liu L, Liu FR, Li AH, Zhu YL, et al. Targeted inhibition of the HNF1A/SHH axis by triptolide overcomes paclitaxel resistance in non-small cell lung cancer. *Acta Pharmacol Sin*. (2024) 45:1060–76. doi: 10.1038/s41401-023-01219-y

49. Chen Y, Zhang L, Liu L, Sun S, Zhao X, Wang Y, et al. Rasip1 is a RUNX1 target gene and promotes migration of NSCLC cells. *Cancer Manag Res*. (2018) 10:4537–52. doi: 10.2147/CMAR.S168438

50. Gainor JF, Tseng D, Yoda S, Dagogo-Jack I, Friboulet L, Lin JJ, et al. Patterns of metastatic spread and mechanisms of resistance to crizotinib in ROS1-positive non-small-cell lung cancer. *JCO Precis Oncol*. (2017) 2017. doi: 10.1200/PO.17.00063

51. Cho SJ, Oh JH, Baek J, Shin Y, Kim W, Ko J, et al. Intercellular cross-talk through lineage-specific gap junction of cancer-associated fibroblasts related to stromal fibrosis and prognosis. *Sci Rep*. (2023) 13. doi: 10.1038/s41598-023-40957-1

52. Jia Y, Guo B, Zhang W, Wang F, Zhang Y, Zhang Q, et al. Pan-cancer analysis of the prognostic and immunological role of GJB2: a potential target for survival and immunotherapy. *Front Oncol*. (2023) 13:1110207. doi: 10.3389/fonc.2023.1110207

53. Lu A, Shi Y, Liu Y, Lin J, Zhang H, Guo Y, et al. Integrative analyses identified ion channel genes GJB2 and SCNN1B as prognostic biomarkers and therapeutic targets for lung adenocarcinoma. *Lung Cancer*. (2021) 158:29–39. doi: 10.1016/j.lungcan.2021.06.001

54. ten Berge DMHJ, Damhuis RAM, Aerts JGJV, Dingemans AMC. Real-world treatment patterns and survival of patients with ROS1 rearranged stage IV non-squamous NSCLC in the Netherlands. *Lung Cancer*. (2023) 181. doi: 10.1016/j.lungcan.2023.107253

55. Kaul A, Gordon C, Crow MK, Touma Z, Urowitz MB, Van Vollenhoven R, et al. Systemic lupus erythematosus. *Nat Rev Dis Primers*. (2016) 2. doi: 10.1038/nrdp.2016.39

56. Rosenberger A, Sohns M, Friedrichs S, Hung RJ, Fehringer G, McLaughlin J, et al. Gene-set meta-analysis of lung cancer identifies pathway related to systemic lupus erythematosus. *PloS One*. (2017) 12. doi: 10.1371/journal.pone.0173339

57. Muñoz-Unceta N, Zugazagoitia J, Manzano A, Jiménez-Aguilar E, Olmedo ME, Cacho JD, et al. High risk of thrombosis in patients with advanced lung cancer harboring rearrangements in ROS1. *Eur J Cancer*. (2020) 141:193–8. doi: 10.1016/j.ejca.2020.10.002

58. Beninato T, Lo Russo G, Garassino MC, De Braud F, Platania M. Recurrent thrombosis followed by Lazarus response in ROS1 rearranged NSCLC treated with crizotinib: a case report. *Tumori*. (2020) 106(6):300891620905665. doi: 10.1177/0300891620905665

59. Nichetti F, Russo G, Prelaj A, Provenzano L, de Braud F, Cabiddu M, et al. ALK/ROS1 rearrangements: A real hallmark for thromboembolic events in cancer patients? *Thromb Res*. (2020) 194:176–7. doi: 10.1016/j.thromres.2020.06.041

60. Moschetti L, Piantoni S, Vizzardi E, Sciatti E, Riccardi M, Franceschini F, et al. Endothelial dysfunction in systemic lupus erythematosus and systemic sclerosis: A common trigger for different microvascular diseases. *Front Med (Lausanne)*. (2022) 9:849086. doi: 10.3389/fmed.2022.849086

61. Nieswandt B, Varga-Szabo D, Elvers M. Integrins in platelet activation. *J Thromb Haemostasis*. (2009) 7:206–9. doi: 10.1111/j.1538-7836.2009.03370.x

62. Vaiyapuri S, Jones CI, Sasikumar P, Moraes LA, Munger SJ, Wright JR, et al. Gap junctions and connexin hemichannels underpin hemostasis and thrombosis. *Circulation*. (2012) 125:2479–91. doi: 10.1161/CIRCULATIONAHA.112.101246

63. Tang Y, Zhang YJ, Wu ZH. High GJB2 mRNA expression and its prognostic significance in lung adenocarcinoma: A study based on the TCGA database. *Med (United States)*. (2020) 99. doi: 10.1097/MD.0000000000019054

64. Wang W, Zhu L, Zhou J, Liu X, Xiao M, Chen N, et al. Targeting the KRT16-vimentin axis for metastasis in lung cancer. *Pharmacol Res*. (2023) 193. doi: 10.1016/j.phrs.2023.106818

65. Li Z, Shen L, Ding D, Huang J, Zhang J, Chen Z, et al. Efficacy of crizotinib among different types of ROS1 fusion partners in patients with ROS1-rearranged non–small cell lung cancer. *J Thorac Oncol*. (2018) 13:987–95. doi: 10.1016/j.jtho.2018.04.016

66. Wang Z, Lei Z, Wang Y, Wang S, Wang J-P, Jin E, et al. Bone-metastatic lung adenocarcinoma cells bearing CD74-ROS1 fusion interact with macrophages to promote their dissemination. *Oncogene*. (2024) 43:2215–27. doi: 10.1038/s41388-024-03072-7

67. van Dam S, Võsa U, van der Graaf A, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene-disease predictions. *Brief Bioinform*. (2018) 19:575–92. doi: 10.1093/bib/bbw139

68. Zhan W, Wang W, Han T, Xie C, Zhang T, Gan M, et al. COMMD9 promotes TFDP1/E2F1 transcriptional activity via interaction with TFDP1 in non-small cell lung cancer. *Cell Signal*. (2017) 30:59–66. doi: 10.1016/j.cellsig.2016.11.016

69. Tang M, Burgess JT, Fisher M, Boucher D, Bolderson E, Gandhi NS, et al. Targeting the COMMD4–H2B protein complex in lung cancer. *Br J Cancer*. (2023) 129:2014–24. doi: 10.1038/s41416-023-02476-8

70. Boulanger MC, Schneider JL, Lin JJ. Advances and future directions in *ROS1* fusion-positive lung cancer. *Oncologist*. (2024). doi: 10.1093/oncolo/oyae205

71. Wu Hl, Gong Y, Ji P, Xie Yf, Jiang YZ, Liu G. Targeting nucleotide metabolism: a promising approach to enhance cancer immunotherapy. *J Hematol Oncol*. (2022) 15. doi: 10.1186/s13045-022-01263-x

72. Valvezan AJ, Turner M, Belaid A, Lam HC, Miller SK, McNamara MC, et al. mTORC1 couples nucleotide synthesis to nucleotide demand resulting in a targetable metabolic vulnerability. *Cancer Cell*. (2017) 32:624–38.e5. doi: 10.1016/j.ccell.2017.09.013

frontiers | Frontiers in Oncology

TYPE Review
PUBLISHED 23 January 2025
DOI 10.3389/fonc.2025.1467540

OPEN ACCESS

EDITED BY
Dalila Luciola Zanette,
Oswaldo Cruz Foundation (Fiocruz), Brazil

REVIEWED BY
Antimo Migliaccio,
University of Campania Luigi Vanvitelli, Italy
Guru P. Sonpavde,
Advent Health Orlando, United States

*CORRESPONDENCE
Fan Wu
1142085668@qq.com
Miaomiao Hao
2937144713@qq.com

RECEIVED 20 July 2024
ACCEPTED 06 January 2025
PUBLISHED 23 January 2025

CITATION
Wu F, Zhang H and Hao M (2025) Interactions between key genes and pathways in prostate cancer progression and therapy resistance. *Front. Oncol.* 15:1467540. doi: 10.3389/fonc.2025.1467540

Interactions between key genes and pathways in prostate cancer progression and therapy resistance

Fan Wu[1*], Hengsen Zhang[2] and Miaomiao Hao[1*]

[1]Department of Pathology, Shanghai Ninth People's Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China, [2]Department of Neurosurgery, Affiliated Hospital of Jiangnan University, Wuxi, China

Prostate cancer is one of the most prevalent malignant tumors in men, particularly in regions with a high Human Development Index. While the long-term survival rate for localized prostate cancer is relatively high, the mortality rate remains significantly elevated once the disease progresses to advanced stages, even with various intensive treatment modalities. The primary obstacle to curing advanced prostate cancer is the absence of comprehensive treatment strategies that effectively target the highly heterogeneous tumors at both genetic and molecular levels. Prostate cancer development is a complex, multigenic, and multistep process that involves numerous gene mutations, alteration in gene expression, and changes in signaling pathways. Key genetic and pathway alterations include the amplification and/or mutation of the androgen receptor, the loss of Rb, PTEN, and p53, the activation of the WNT signaling pathway, and the amplification of the MYC oncogene. This review summarizes the mechanisms by which these genes influence the progression of prostate cancer and highlights the interactions between multiple genes and their relationship with prostate cancer. Additionally, we reviewed the current state of treatments targeting these genes and signaling pathways, providing a comprehensive overview of therapeutic approaches in the context of prostate cancer.

KEYWORDS

signaling pathway, therapeutic targets, CRPC, prostate cancer, gene mutation

1 Introduction

Prostate cancer is currently the most common malignancy among men in the United States, with an incidence of 29% (1). In 2024, it is the most common cause of male cancer death after lung and bronchial cancer (1). Globally, prostate cancer mortality is slightly lower than that of lung cancer in the male population (2). Therefore, prostate cancer ranks high in both incidence and mortality rates. Prostate cancer is influenced by various risk

factors, including age, family history, obesity, and unhealthy dietary habits (3). Age is a primary risk factor for prostate cancer. The incidence is rare in men under 50 years old (1 in 350), but it increases sharply to 1 in 52 by age 59, and by age 65, the rate rises to more than 1 in 2. Men with a family history of the disease have more than double the risk of developing prostate cancer compared to those without such a history (4–6). Additionally, race plays a role in prostate cancer risk. Research from 2010 found, compared to White male patients, Black male patients exhibit a more rapid progression of prostate cancer and may develop invasive prostate cancer at an earlier stage (7, 8).

Early-stage prostate cancer often lacks noticeable symptoms, making it difficult to detect and delaying timely and effective treatment. Currently, the screening and diagnosis of prostate cancer mainly include serum Prostate-specific Antigen (PSA), Magnetic Resonance Imaging fusion ultrasound-guided prostate biopsy (MRI-TRUS), and digital rectal examination. Despite the availability of these methods, PSA remains the most widely used screening tool for early diagnosis of prostate cancer worldwide. Although PSA is highly sensitive for early detection, it lacks specificity of the properties of prostate tissue. This means it cannot differentiate between high-risk and low-risk tumors and may also be elevated in cases of enlarged prostate, aging, prostatitis, certain urological diseases, and specific drug treatments. Consequently, PSA screening may lead to overtreatment of prostate cancer (9).

In recent years, alongside PSA, other tumor markers such as p53, MDM2 and Ki67 have been used to monitor the progression and treatment of prostate cancer. Additionally, the application of next generation sequencing (NGS) technology in cancer diagnosis and treatment has deepened researchers' understanding of prostate cancer and its molecular biology. Drug therapies targeting prostate cancer-related genes are also under investigation and some of them have been used in clinical treatment, but none of the therapeutic effects are very satisfactory, and the treatment of advanced prostate cancer is still an urgent problem to be solved. This article primarily reviews the treatment, drug resistance, and prognosis of genes related to prostate cancer. The interactions between related genes are further summarized and it is suggested that combination therapy targeting such multiple genes may be more effective in the treatment of advanced prostate cancer.

2 Androgen receptor

2.1 Role of AR in prostate cancer

Androgen receptor (AR), a nuclear transcription factor in the steroid hormone receptor family, is central to prostate cancer pathogenesis. When testosterone or 5-alpha-dihydrotestosterone (DHT) binds to AR, the receptor dimerizes and translocations to the nucleus, where it binds to the androgen response element (ARE) (10). This interplay participates in the transcriptional activity of genes that prevent apoptosis and induce cell proliferation. AR supports proper development in normal prostate, whereas elevated AR expression drives disease progression in prostate cancer (11).

2.2 Mechanisms of resistance to ADT

Androgen deprivation therapy (ADT) is a treatment designed to reduce or block the production of androgens (male hormones, such as testosterone) that fuel the growth of prostate cancer. ADT is initially effective in treating prostate cancer (12). As the disease progresses, most patients eventually develop castrate-resistant prostate cancer (CRPC) and metastases after ADT (Figure 1). There are two main mechanisms behind this resistance. First, although early-stage prostate cancer is primarily driven by androgen-dependent cancer cells, the disease is heterogeneous, not only composed of androgen-dependent cells. Castration resistance occurs due to the growth of androgen-independent cells, which arises from genetic alterations in the AR (13). Second, apart from the androgens produced by the adrenal glands and testis that stimulate AR, intra-tumoral secretion of enzymes involved in testosterone synthesis, such as cytochrome P450 17-alpha hydroxysteroid dehydrogenase (CYP17), also supports tumor survival and growth (14). Moreover, a new mechanism about the resistance has been found in recent years. AR splice variants are more common in CRPC, and they are characterized primarily by the loss of ligand domains, which retain the ability to bind to DNA in the absence of androgens (15). There are many variants of AR spliceosome, among which AR-V7 is one of the most studied variants. AR-V7 can complete nuclear transfer in the absence of androgen binding and recruit cofactors to complete transcriptional activation of downstream genes, followed by aberrant activation of the AR signaling pathway (16). Interestingly, AR-V7 also predicted treatment response to AR-targeting drugs, and AR-V7-positive patients who received enzalutamide and abiraterone had shorter progression-free survival and shorter overall survival than AR-V7-negative patients (17). This also provides strong evidence for AR-V7 as a biomarker for prostate cancer.

2.3 Emerging therapies and challenges

To target CRPC, new drugs that inhibit androgen-producing enzymes or block AR have been developed in recent years, such as second-generation nonsteroidal AR antagonists (enzalutamide, apalutamide, and darolutamide) and the androgen biosynthesis inhibitor abiraterone (18). In a phase 3 trial of enzalutamide, which randomized 1,125 male patients with metastatic castration sensitive prostate cancer (mCSPC) into groups of ADT in combination with either enzalutamide (N = 563) or a standard nonsteroidal antiandrogen agent (bicalutamide, flutamide, or nilutamide; N = 562) until progression or unacceptable toxicity. the enzalutamide arm had fewer deaths than the standard-care group (102 *vs* 143; HR 0.67; 95% CI 0.52-0.86; $P = 0.002$) and 3-year overall survival (OS) estimated at 80% (based on 94 events) *vs* 72% (based on 130 events), respectively (19). In another clinical trial, 297 patients with high-risk metastatic hormone-sensitive prostate cancer (mHSPC) treated with abiraterone, 127 with enzalutamide, and 142 with apalutamide were compared. There were no differences in time to CRPC ($p = 0.13$), OS ($p = 0.7$), and cancer-specific survival (CSS) ($p = 0.5$) among the three ARPIs, but abiraterone was significantly better in 99% PSA decline achievement compared to apalutamide (72% vs. 57%, $p = 0.003$) (20). However, over time, most patients still develop resistance to these treatments

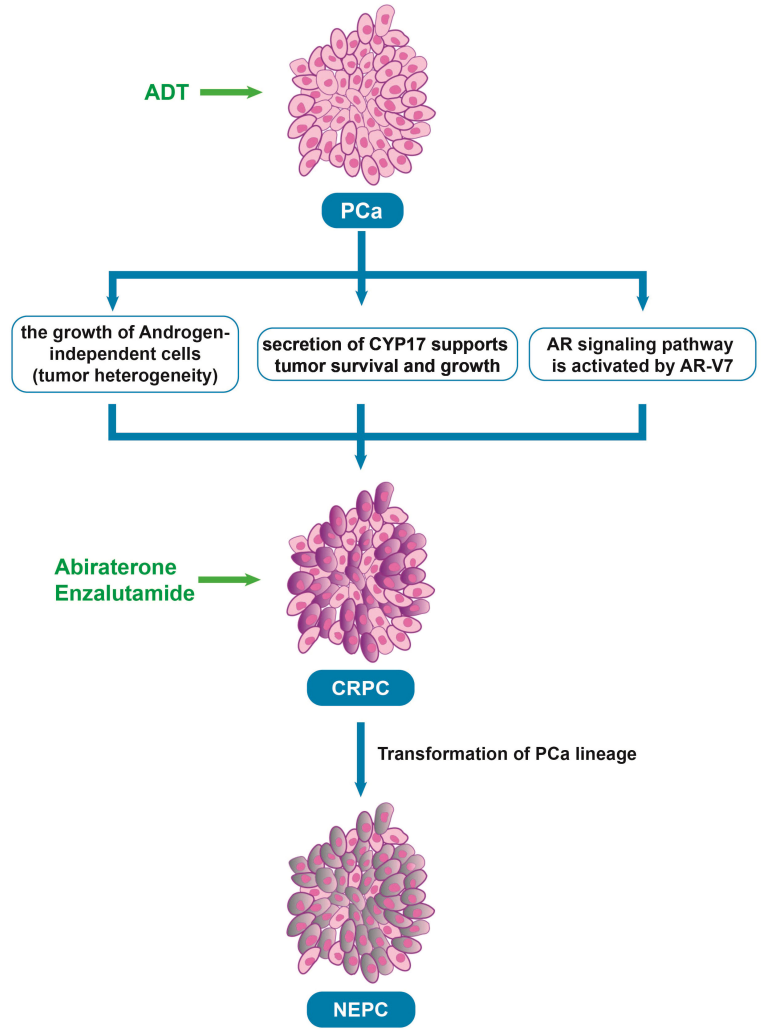


FIGURE 1
The process underlying the development of CRPC after ADT. PCa, Prostate cancer; ADT, Androgen deprivation therapy; CRPC, Castrate resistant prostate cancer; AR, Androgen receptor; NEPC, neuroendocrine prostate cancer, Abiraterone and Enzalutamide: The androgen biosynthesis inhibitor and the novel AR inhibitor.

(Table 1) (21). Some studies have found that after treatment with anti-androgen drugs, prostate cancer cells undergo a lineage shift, which refers to the conversion of cells from luminal and basal cells to neuroendocrine-type cells caused by adaptation to the environment (22–24). Thus, the prostate cancer cells can evade drug-targeted therapy, causing treatment-resistant neuroendocrine prostate cancer.

TABLE 1 Genes associated with prostate cancer progression.

Gene	Function	Interaction between genes
AR	Regulation of AR signaling pathway	Rb, p53, MYC, WNT
Rb	Regulation of cell cycle	AR, p53, PTEN
PTEN	Regulation of PI3K/AKT signaling pathway	Rb, p53, MYC
WNT	Regulation of WNT/β-catenin signaling pathway	AR
p53	Regulation of cell cycle	Rb, PTEN, AR
MYC	Regulation of gene expression and key cellular processes	AR, PTEN

3 Retinoblastoma

3.1 Mechanisms of cell cycle regulation by retinoblastoma

Retinoblastoma is a malignant tumor, and Retinoblastoma (Rb) is a tumor suppressor gene identified in this tumor (25, 26). The Rb gene

is located on chromosome 13q14.2 and was the first human tumor suppressor gene to be cloned (25). The Rb protein family includes Rb, p107 and p130, collectively referred to as "pocket proteins", which are involved in cell cycle regulation (27). The cell cycle is the series of events in which cellular components are doubled, and then accurately segregated into daughter cells. In eukaryotes, the cell cycle consists of four phases, S-phase, in which DNA replication occurs, M-phase, in which mitosis occurs, and two interphases, G1 and G2, between S-phase and M-phase, which are the times when the cell acquires mass, integrates growth signals, organizes the replication of the genome, and prepares the chromosomes for segregation (28). In its low phosphorylation state, Rb can inhibit the transcriptional activity of E2F by binding to its downstream transcription factors (E2F), thereby suppressing the expression of genes involved in the cell cycle and arresting the cell cycle in the G1 phase (29). However, in late G1, Rb transitions from a low phosphorylation state to a high phosphorylated, inactive state, releasing E2F and allowing cells to enter the S-phase, thereby promoting cell proliferation (29). The cyclin-cyclin dependent kinase (CDK) complex promotes cell cycle progression by phosphorylating members of the Rb family during G1. Cyclin D expression leads to CDK4 (and CDK6)- dependent phosphorylation of Rb, reducing its binding to E2Fs and promoting early cell cycle gene expression (30). CDK inhibitors (such as p16 and p21) can prevent CDK from phosphorylating Rb by inhibiting the activity of CDK4 and CDK6, thereby promoting Rb function (31).

3.2 Role of retinoblastoma in prostate cancer

The inactivation of Rb is closely related to all stages of prostate cancer formation (32). Rb-mediated loss of cell cycle control only leads to the occurrence of prostatic proliferative diseases and is not sufficient to cause malignant tumors (33). It has been shown that Rb deletion can promote angiogenesis, metastasis and neuroendocrine differentiation (NED), a process by which epithelial tumor cells acquire features of neuroendocrine cells, resulting in a more aggressive phenotype in human prostate cancer cells (34). In addition, Rb can promote epithelial-mesenchymal transition (EMT) and tumor cell invasion by regulating downstream target genes (35). Recently, Jin, X., et al. reported that the Rb-NF-κB axis can be used to overcome cancer immune escape induced by conventional or targeted therapies (36). Thus, while the absence of Rb does not cause the occurrence of prostate cancer, it can lead to the proliferation of prostate cells and plays an essential role in the metastasis, EMT and NED of prostate cancer.

In addition to promoting the development of prostate cancer through the aforementioned mechanisms, Rb loss also participates in the AR signaling pathway. Androgens are known prostatic epithelial cell growth factors (37) and play an important role in prostate cancer development. Androgens can activate Rb by regulating CDK4/cyclin D1 and CDK2 complexes, thereby initiating the cell cycle (38). After androgen castration treatment, the level of cyclin D protein is reduced, maintaining low Rb phosphorylation, causing cell cycle arrest, and inhibiting tumor development (39). Sharma, A., et al. have found that CRPC that develops after castration-resistant treatment shows decreased Rb expression and increased AR expression (40). Subsequently, Gupta, S. et al. also have found that AR overexpression in CRPC was associated with Rb inactivation (41). we believe that there are several mechanisms for this phenomenon: 1) deletion of Rb activates E2F, which acts downstream of it to increase AR expression (42, 43); 2) Rb loss increases AR recruitment to homologous promoters, resulting in increased AR target gene expression (44); 3) AR induces signals that promote CDK activity and promotes phosphorylation of Rb to inactivate it (Figure 2) (45).

3.3 Emerging therapies and challenges

Given the above mechanism of cell cycle regulation by Retinoblastoma in prostate cancer, inhibition of Rb phosphorylation can be used as a therapeutic strategy for prostate cancer. By binding CDK inhibitors to CDK4 and CDK6, Rb phosphorylation is inhibited to prevent the G1-S phase transition and induce cell cycle arrest. At present, there is evidence that highly selective small molecule inhibitors of CDK4 and CDK6, Palbociclib, Ribociclib and Abemaciclib, are effective in the treatment of breast cancer (46), but the therapeutic effect of prostate cancer is not clear. In breast cancer, data from the latest MONARCH-3 study showed that at a median follow-up time of 8.1 years, treatment with Abemaciclib in combination with an nonsteroidal aromatase inhibitor (NSAI) numerically prolonged Overall Survival (OS) compared to NSAI therapy alone in patients with HR+, HER2- advanced breast cancer, however, unfortunately, the difference did not reach statistical significance (P=0.0664) (47). Ribociclib is the only CDK4/6 inhibitor that has achieved positive OS results in all three phase III studies, with stable and consistent OS benefit, whether targeting premenopausal or postmenopausal populations, as a first- or second-line treatment, or in combination with an aromatase inhibitor (AI) or fulvestrant. This is based on several unique mechanisms of action. Firstly, Ribociclib can induce tumors cell senescence to achieve a long-term response (48); secondly, Ribociclib significantly affects peripheral innate and adaptive immune responses, and achieves long-term efficacy through immune activation (49). These are all characteristics that Abemaciclib does not possess. Likely due to the above reasons, the most recent phase 3 study of Abemaciclib with abiraterone in patients with metastatic CRPC (mCRPC) did not show a significant increase in radiographic progression-free survival (rPFS) for the addition of Abemaciclib to abiraterone, the medians rPFS were 21.96 months for the Abemaciclib plus abiraterone group vs 20.28 months for the placebo (PBO) plus abiraterone group (50).

4 PTEN

4.1 Mechanisms of PTEN in cell proliferation and apoptosis

Phosphatase and tensin homolog gene (PTEN) is a tumor suppressor gene with phosphatase activity, which is located in

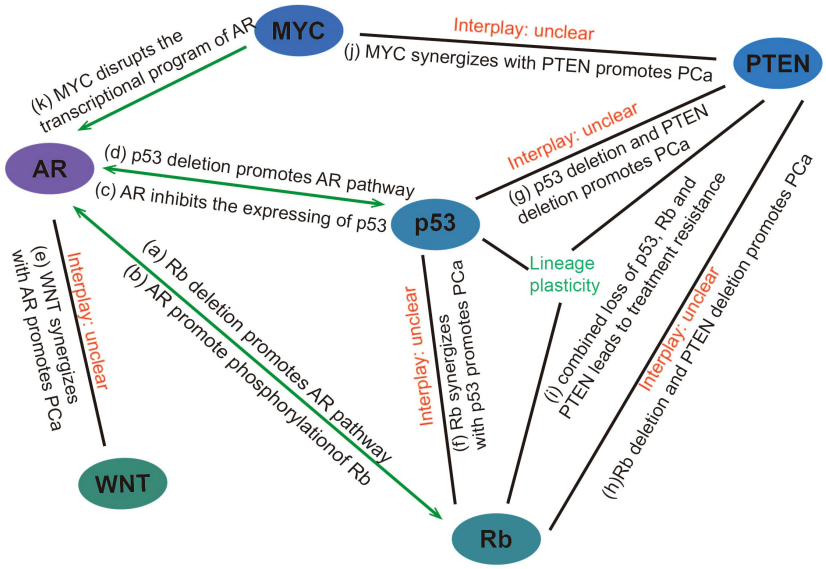


FIGURE 2
Interaction between genes/pathways in prostate cancer **(A)** Rb deletion promotes the expression of genes downstream of the AR pathway through transcriptional activation of E2F and facilitating promoter recruitment of the AR. **(B)** AR induces signals that promote CDK activity and promotes phosphorylation of Rb to inactivate it. **(C)** AR can promote prostate cancer progression by reducing p53 expression through G3BP3, which promotes the nuclear translocation of P53. **(D)** p53 overexpression inhibits androgen-induced transactivation of NKX3.1 by repressing the promoter of the AR gene and blocking AR-DNA binding activity. Conversely, p53 deletion promotes prostate cancer progression by facilitating the AR signaling pathway. **(E)** WNT can synergize with AR to promote the development of aggressive prostate cancer. **(F)** p53 and Rb deletion can mediate lineage plasticity, thereby enabling prostate cancer to evade targeted therapies and progress to CRPC. **(G)** The deletion of PTEN and p53 can lead to changes in the lineage of prostate cancer, resulting in the formation of CRPC. **(H, I)** Rb loss facilitates lineage plasticity and metastasis of prostate adenocarcinoma initiated by PTEN mutation. The additional loss of p53 causes resistance to antiandrogen therapy. **(J)** MYC overexpression or targeted PTEN loss can each produce early prostate adenocarcinomas but are not sufficient to induce genetic instability or metastases with high penetrance. However, MYC activation and PTEN deletion induced genomic instability and aggressive prostate cancer. **(K)** MYC promotes the development of mCRPC by disrupting the transcriptional program of AR. PCa, Prostate cancer.

chromosome 10q23 and spans 200kb in full length (51). PTEN is involved in tumor progress by inhibiting the phosphatidylinositol 3-kinase (PI3K)/serine-threonine kinase (AKT)/mammalian target of the rapamycin (mTOR) pathway and its reduction or loss of expression caused by methylation, mutation or deletion is closely related to the occurrence and development of various tumors (52). The PI3K/AKT/mTOR pathway is crucial for cell signal transduction. PTEN enables dephosphorylate phosphatidylinositol (3,4,5)-trisphosphate (PIP3) of PI3K/AKT/mTOR pathway to generate Phosphatidylinositol (4,5)-bisphosphate (PIP2). When PTEN is lost, PIP2 is phosphorylated to PIP3 by PI3K, activating a series of kinases in the signaling pathway, including AKT. AKT affects cell apoptosis through serine phosphorylation of Bcl-2-associated death promoter (BAD) and Caspase-9, and influences cell proliferation, differentiation, and survival through regulation of transcription, translation, and cell cycle (53).

4.2 Role of PTEN in prostate cancer

The loss of PTEN promotes overactivation of the PI3K/AKT/mTOR signaling pathway leading to cell transformation and tumorigenesis (54). In a study on the PTEN deleted mouse model of prostate cancer, a blockade of mTOR inhibited prostate tumorigenesis in epithelial cells (55). Numerous studies have found that the occurrence of prostate cancer and its hormone-independent transformation course are related to the loss of PTEN gene expression (56, 57). Studies have shown that PTEN is absent in 15% to 20% of primary prostate cancers, and the frequency of PTEN deletion is higher in CRPC and mCRPC tissues, reaching 40% to 60% (58). PTEN deletion is positively correlated with Gleason score, pathological grade, clinical stage and metastasis of prostate cancer (59). The above evidence suggests that PTEN loss is closely associated with prostate cancer progression and tumorigenesis.

4.3 Emerging therapies and challenges

Several inhibitors (rapamycin analogs) targeting the PI3K/AKT/mTOR pathway have been investigated to counteract the mechanism by which PTEN deletion promotes prostate cancer progression (60), but their antitumor effects have been disappointing. Although rapamycin inhibited PI3K/AKT/mTOR pathway, long-term treatment caused resistance and was not suitable for monotherapy. Wang Y. et al. have found that combination of rapamycin and bicalutamide (anti-androgenic drug) improved anti-prostate cancer effect due to the suppression of mTOR stimulated AR transcriptional activity (61). A clinical trial for mCRPC demonstrated that the combination of the PI3K inhibitor samotolisib with enzalutamide, which causes an

improved PFS in mCRPC patients progressing on abiraterone, median Prostate Cancer Clinical Trials Working Group criteria (PCWG2)-PFS and rPFS was significantly longer in the samotolisib/enzalutamide versus placebo/enzalutamide arm (3.8 vs. 2.8 months; P = 0.003 and 10.2 vs. 5.5 months; P = 0.03), respectively (62). Another phase 3 study combining the AKT inhibitor ipatasertib with abiraterone acetate, which has shown a significant positive impact on PFS in mCRPC patients with PTEN loss, in the 521 (47%) patients who had tumors with PTEN loss (261 in the placebo-abiraterone group and 260 in the ipatasertib-abiraterone group), median rPFS was 16.5 months (95% CI 13.9-17.0) in the placebo-abiraterone group and 18.5 months (16.3-22.1) in the ipatasertib-abiraterone group (hazard ratio [HR] 0.77 [95% CI 0.61-0.98]; p=0.034) (63). Combined AKT and androgen-receptor signaling pathway inhibition is a potential treatment for men with PTEN-loss mCRPC, a population with a poor prognosis. This result suggests that combination therapy targeting multiple genes or pathways may become a major direction for future prostate cancer treatment. Moreover, several natural bioactive compounds including afrocyclamin A, apigenin, arctigenin, curcumin, cryptotanshinone, oridonin, salidroside, and vitexin were reported to target the PI3K/AKT/mTOR pathway, however, some compounds are currently under examination in clinical trials (64).

5 WNT

5.1 Mechanisms of WNT in cell proliferation and apoptosis

WNT codes a family of proteins involved in the cell signaling process. The WNT signaling pathway is a highly conserved signaling pathway with multiple downstream channels stimulated by the binding of WNT ligand proteins to membrane protein receptors. This pathway plays a crucial in embryonic development, cell proliferation, cell migration and apoptosis. Abnormalities in the WNT signaling pathway are closely associated with the development and progression of various diseases, including cancer (65).

The WNT/β-catenin pathway is the canonical pathway of WNT signaling. Extracellular WNT signaling molecules prevent the phosphorylation of β-catenin, allowing it to accumulate in the cytoplasm. When the concentration of β-catenin in the cytoplasm reaches a certain level, it translocated to the nucleus and combines with the intracellular transcription factor T-cell factor/lymphoid enhancer factor (TCF/LEF) to form a complex. This complex activates the proto-oncogenes Cyclin D1 and c-MYC, leading to tumor cell proliferation, differentiation and maturation (66).

5.2 Role of WNT in prostate cancer

Bisson, I. and D.M. Prowse have shown that the WNT/β-catenin signaling pathway is highly active in tumor stem cells and may play a role in the self-renewal of prostate cancer stem cells (67). Wang, B.E., et al. have found that targeting prostate cancer stem cells with WNT/β-catenin signaling inhibitors has been shown to enhance the therapeutic effect of prostate cancer treatments (68). Similar to other genes, WNT signaling is strongly associated with advanced prostate cancer, and Wang, Y., et al. have found that WNT signaling promotes bone metastasis of prostate cancer (69). In addition, β-catenin can interact with other pathways (AR) to coordinate proliferation during tumor growth (70). The above findings suggest that the WNT/β-catenin signaling pathway plays an important role in prostate cancer, especially advanced prostate cancer. This feature may provide a key therapeutic target for the treatment of advanced prostate cancer.

5.3 Emerging therapies and challenges

Currently, there is still no effective drug therapy targeting the WNT/β-catenin signaling pathway. However, there are a number of drugs in clinical trials. A new β-catenin mimic small molecule inhibitor, CWP232291, is currently in clinical trials. CWP232291 induce endoplasmic reticulum stress and cell apoptosis, ultimately leading to β-catenin degradation (71). In addition, Cirmtuzumab and Foxy-5 are in Phase 1 trials. Cirmtuzumab is a monoclonal antibody that targets the receptor called ROR1 of the non-canonical Wnt pathway and is suspected to contribute to prostate cancer growth and progression (72). Foxy-5 mimic the effects of Wnt-5a to impair migration of epithelial cancer cells and thereby acting anti-metastatic (73). Given the correlative role of the wnt pathway with AR and MYC, combination therapy with an AR inhibitor or a MYC inhibitor may be useful in the treatment of advanced prostate cancer in the future.

6 p53

6.1 Mechanisms of p53 in cell cycle and DNA repair

The p53 gene is an important tumor suppressor gene in human cancer, first identified in extracts of transformed cells (74). It plays a vital role in regulating cell cycle and DNA repair. p53 regulate both the G1-S phase (75) and the G2-M arrest (76), thus providing a checkpoint function and repair of genes in the cell cycle. In terms of apoptosis, p53 can induce apoptosis by directly activating its downstream apoptotic genes, such as Bax, Puma and Noxa, etc. in cells that fail to repair DNA damage (77).

6.2 Role of p53 in prostate cancer

Deletion of p53 or loss of function due to p53 mutations is detectable in many cancers (78). There are various types of p53 mutations in prostate cancer, including deep deletion, Fusion, shallow deletion, missense mutation, truncating mutation, splice mutation, in-frame mutation and amplification (79). Cotter et al. found that in localized prostate cancer the mutation types of p53 were mainly deep deletion and mutation, while in advanced prostate

cancer the mutation types of p53 were deep deletion, mutation and amplification (80). The incidence of p53 mutations is not the same at different stages of prostate cancer, ranging from 31.4% in CRPC to 66.7% in neuroendocrine prostate cancer (81–83). Wang, Y., et al. found that p53 deletion promotes invasion and metastasis in advanced prostate cancer, via enhancing the FAK-Src signaling pathway (84). Actually, p53 mutations occur not only in the advanced stage of prostate cancer but also in its early stage (85). The frequency of these mutations gradually increases as the cancer progresses, reaching the highest level in CRPC (77, 86). These findings suggest that p53 plays a key role in multiple stages of prostate cancer development. In addition, Fonseca, G.N., et al. have shown that the expression of mutant p53 is positively correlated with tumor staging (87). More p53 mutations are found in metastatic prostate cancers than in early-stage prostate cancers, making p53 a potential independent predictor of recurrence of low- and intermediate-grade prostate cancers (88).

In 2006, a study specifically knocked out the Rb and p53 genes in mouse prostate epithelium, and found that after knocking out the Rb gene or p53 gene alone, mice could only develop prostate intraepithelial neoplasia (PIN), but could not develop prostate cancer (89). Only after the simultaneous knockout of Rb and p53 genes, the mice can develop prostate cancer and become highly metastatic (89). It suggests that the loss of Rb and p53 may play a synergistic role in the development and progression of prostate cancer. In a recent study, it was found that in prostate cancer with p53 and Rb deletion, overexpression of the transcription factor SOX2 can mediate lineage plasticity, thereby enabling prostate cancer to evade targeted therapies and lead to CRPC (90). In addition, the deletion of PTEN and p53 can also lead to changes in the lineage of prostate cancer, resulting in the formation of CRPC (91, 92). Ku, S.Y., et al. have found that Rb loss facilitates lineage plasticity and metastasis of prostate adenocarcinoma initiated by PTEN mutation, additional loss of p53 causes resistance to antiandrogen therapy (93). These results indicate that the lineage change of prostate cancer is involved in the deletion of multiple genes, and the specific mechanism of the lineage change of prostate cancer remains to be further studied. This also makes the treatment of advanced prostate cancer more difficult and complex.

Androgen castration is a common treatment for prostate cancer, but most cancers eventually develop androgen independence. Relevant studies have proved that the loss of p53 is associated with CRPC. Inhibition of p53 expression can reduce AR-mediated signal transduction, while overexpression of wild-type p53 can reduce androgen function (94). This is because p53 overexpression inhibits androgen-induced transactivation of NKX3.1 by repressing the promoter of the AR gene and blocking AR-DNA binding activity (95). Therefore, the basic physiological level of wild-type p53 is necessary for AR signal and has a protective effect on it, but the balance between p53 and AR is eliminated as cancer progresses (94), and deletion of p53 leads to androgen-induced transactivation of NKX3.1, which promotes prostate cancer progression. AR also promotes the inactivation of p53. A Study in 2017 showed that AR can induce the translocation of p53 from the nucleus to the cytoplasm via the downstream target gene G3BP2, thereby inhibiting the function of p53 (96).

6.3 Emerging therapies and challenges

p53 inactivation may limit the effectiveness of radiation therapy in localized prostate cancer because the effectiveness of treatment relies on p53-mediated cell senescence and apoptosis. Consequently, the p53 pathway can be used as a specific target to enhance the radiosensitivity of prostate cancer cells. For example, using potent radiosensitizers for prostate cancer cells that retain the functional allele of p53 can improve the efficacy of radiation therapy (97). For p53-deficient CRPC, flubendazole is a well-known anti-malarial drug and a potential anti-tumor drug that has been shown to induce cell cycle arrest in the G2/M phase, promote cell death *in vitro* by inducing p53 expression, and inhibit the growth of CRPC tumors in xenograft models (98). But these drugs have had limited clinical trials and their safety has not been proven, there are still many challenges in the treatment of advanced prostate cancer. The findings that p53 interacts with Rb, PTEN and AR in advanced prostate cancer, and synergizes with Rb in the development of prostate cancer, have important implications for the treatment of advanced prostate cancer, and that exploring gene interactions and combining therapies may be of immense help in addressing drug resistance in advanced prostate cancer.

7 MYC

7.1 Mechanisms of MYC in cell proliferation and apoptosis

The MYC family of proto-oncogenes consists of three homologs: c-MYC (MYC), n-MYC (MYCN), and l-MYC (MYCL), located on chromosomes 8, 2, and 1, respectively. Although MYC family genes encode proteins with similar structural architecture and function, different timing of expression and tissue specificity is exhibited during development (99–101). These genes are involved in regulating integral gene expression and key cellular processes including proliferation, differentiation, cell cycle, metabolism and apoptosis.

7.2 Role of MYC in prostate cancer

c-MYC (MYC) is a major promoter of prostate cancer tumorigenesis and progression (102, 103). Under normal conditions, its expression and function are strictly controlled, but overexpression of MYC is frequently observed in prostate cancer (104). Amplification of MYC has been reported to be associated with aggressiveness and poor prognosis in prostate cancer (103). Studies have shown that MYC overexpression in normal luminal cells of the mouse prostate is sufficient to cause PIN and prostate cancer (105, 106). This indicates that dysregulated MYC protein expression is a key oncogenic event driving prostate carcinogenesis. Furthermore, overexpression of MYCN mediates the transformation of CRPC to neuroendocrine prostate cancer (107).

The interplay of MYC with other signaling pathways also exerts a significant role in the development of prostate cancer. Overexpression

of MYC leads to the pausing of RNA polymerase II at the promoter-proximal regions of AR-dependent genes, disrupting the AR transcriptional program promote the initiation and progression of prostate tumors (102). Arriaga et al. have recently reported a MYC and RAS co-activation signature associated with metastatic progression and failure to anti-androgen treatments (108). Gretchen et al. found that MYC activation and PTEN deletion in mouse prostate luminal cells induced genomic instability and aggressive prostate cancer in the absence of induced telomere dysfunction or p53 loss of function (109). These studies indicate that MYC can cooperate with other pathways to promote the development of prostate cancer.

7.3 Emerging therapies and challenges

Given its key role in prostate cancer, MYC is considered a potential therapeutic target. MYC inhibitors that disrupt MYC and Max dimerization sensitize enzalutamide-resistant prostate cancer cells to growth inhibition by enzalutamide (110). Bromodomain extra-terminal enhancer inhibitors can affect MYC transcription by targeting upstream MYC pathways and have shown preclinical efficacy in MYC-driven CRPC models (111, 112). Kirchner et al. reported that inhibition of PIM, a family of serine-threonine kinase, with the pan-PIM kinases inhibitor AZD-1208 was effective in limiting MYC-driven lesion progression (113). Additionally, a study found that dual inhibitors targeting MYCN and Aurora A kinase (AURKA) could be potential therapies for neuroendocrine prostate cancer (114). Despite these advances, there are still no clinically approved drugs targeting MYC for the treatment of prostate cancer.

8 Discussion

In recent years, the incidence of prostate cancer has been steadily increasing. The continuous proliferation and metastasis of prostate cancer cells are critical clinical features and the main causes of mortality in advanced prostate cancer. These processes are regulated by a series of genetic alterations (Table 1). It is challenging to elucidate the mechanisms underlying prostate cancer through a single gene mutation or deletion.

The development of prostate cancer involves complex interactions among multiple genes and pathways (Figure 2). The molecular mechanisms involving the interaction among multiple genes and pathways remain to be further explored. Further investigation into the synergistic effects of Rb and p53, MYC and PTEN, and WNT and AR in prostate cancer, as well as the identification of common downstream target genes among these interacting genes or pathways, could lead to the discovery of novel targeted therapies. Such research may offer new avenues for treating CRPC and addressing the lineage plasticity of prostate cancer. Currently, resistance to prostate cancer treatment remains a significant challenge. There are many ongoing clinical trials targeting different genes and pathways for the treatment of different stages of prostate cancer, but they still have different limitations, which further suggests that it is critical to explore the interactions of multiple genes and pathways (Table 2).

The co-deletion of Rb, PTEN and p53 has been shown to confer resistance to antiandrogen therapy. By exploring the molecular mechanisms associated with this co-deletion, we may uncover more effective and sensitive tumor markers and therapeutic targets, thereby improving treatment strategies for advanced prostate cancer. There is still no effective solution to the problem of chemotherapy drug resistance in advanced prostate cancer, but in breast cancer it has been found that drug resistance can be solved through multigene interactions. In HR+/HER2-advanced breast cancer, the medical community has been exploring new therapeutic options for patients who develop resistance after CDK4/6 inhibitors combined with endocrine therapy. Some researchers have found that PI3K pathway inhibitors can alleviate resistance to chemotherapy drug, CDK4/6 inhibitors, in advanced-stage patients. In patients with HR/HER2-advanced breast cancer after progression on the CDK4/6 inhibitor, the patients who applied endocrine therapy in combination with the mTOR inhibitor had a median PFS benefit of 5.1 months (115). This evidence suggests that exploring the mechanisms of multigene interactions could help address chemotherapy resistance in advanced tumors.

In addition to the genes discussed in above, there are a number of genes associated with prostate progression. For example, breast cancer susceptibility gene 1 (BRCA1) and breast cancer susceptibility gene 2 (BRCA2) have been shown to be closely associated with prostate cancer aggressiveness and patient prognosis (116). Both are oncogenes, which can regulate the cell cycle through synergistic effects with other repair mechanisms in the organism and other oncogenes, ensuring the proliferation and apoptosis of normal cells (117). The correlation between BRCA mutation and prostate cancer is still in the research stage, and it is controversial whether BRCA mutation carriers are the high-risk group for prostate cancer, and at present, there is no evidence to show the most suitable method for the treatment of BRCA mutation-associated prostate cancer. Studies have shown that BRCA mutation carriers in the mCRPC population have better treatment outcomes compared to non-carriers, and that patients with either BRCA1 or BRCA2 mutations benefit from treatment with abiraterone or enzalutamide (118). Therefore, exploring the interrelationships of BRCA1 or BRCA2 with other genes and pathways may offer further assistance in the treatment of BRCA mutation-associated prostate cancer.

Src/Ras/extracellular signal-regulated kinase (Erk) pathway also associated with prostate cancer progression. Src is a non-receptor protein tyrosine kinase (119). Src could activate multiple downstream signaling pathways, including the PI3K/AKT pathway and the Ras/Erk pathway, which are important for cell proliferation and DNA synthesis (120, 121). In prostate cancer cells, androgens trigger the binding of AR to Src, this interaction activates Src/Ras/Erk pathway and affects G1 to S cell cycle progression (122). Migliaccio et al. identified an amino acid peptide that inhibits the AR/Src interaction, which inhibits the binding of AR to Src and the activation of the Src/Ras/Erk pathway (123). However, the peptide had no such inhibitory effect in AR-negative prostate cancer cell lines, suggesting that the peptide can only inhibit the androgen receptor-dependent Src pathway in prostate cancer. In addition, Src/Ras/Erk plays an important role in breast cancer,

TABLE 2 The clinical trials that are ongoing to treat prostate cancer at different stages.

Gene	PCa (Phase)	CPRC (Phase)	mCRPC (Phase)	Limitations
AR	Apalutamide (2) NCT01790126 Goserelin (2) NCT00298155 ARN-509 (2) NCT01790126 SHR3680 (3) NCT03520478	Enzalutamide (3) NCT00974311 Apalutamide (4) NCT04108208	Enzalutamide (4) NCT02116582 ARV-110 (1/2) NCT03888612 Apalutamide (1) NCT03523442 JNJ-56021927 (1) NCT02162836 TRC253 (1/2) NCT02987829 ARV-110 (1) NCT05177042	Inevitability of castration resistance
Rb (CDK4/6)	Abemaciclib (1/2) NCT05617885	–	Palbociclib (2) NCT02905318 Ribociclib (1/2) NCT02494921 Abemaciclib (2/3) NCT03706365 TQB3616 (2) NCT05156450	Limited clinical trials Questionable safety profile
PTEN (PI3K/AKT)	AZD2014 (1) NCT02064608	LY3023414 (2) NCT02407054 AZD8186 (1) NCT01884285	Perifosine (2) NCT00060437 GSK2636771(1) NCT02215096 Afuresertib (1/2) NCT04060394 Ipatasertib (3) NCT03072238	Limited clinical trials Biomarkers needed for patient selection
WNT	FOXY-5 (1) NCT02020291	–	Cirmtuzumab (1) NCT05156905	Bone-related toxicity Limited clinical trials
p53	PC14586 (1/2) NCT04585750	–	APR-246 (1) NCT00900614 Arsenic trioxide (2) NCT00004149	Limited clinical trials Questionable safety profile
MYC	–	–	ZEN-3694 (2) NCT04471974	Limited clinical trials

which has led to several studies of Src inhibitors (124). In an ongoing phase 2 trial in prostate cancer, the effect of combining an Src inhibitor with an AR inhibitor versus an AR inhibitor alone on the development of EMT in prostate cancer was compared, but no definitive results have been published (125).

In recent years, the development of immune checkpoint inhibitors (ICIs) has transformed the treatment landscape for various genitourinary malignancies. ICIs are innovative tumor therapeutic agents that restore the body's anti-tumor immunity by blocking the tumor immune escape mechanism. However, the efficacy of ICIs in prostate cancer remains limited, especially in cases of CRPC, which is challenging to control with traditional therapies. Prostate cancer is often considered an "immune-cold" tumor, characterized by a tumor microenvironment with low immune activity, low tumor mutational burden, interferon signaling dysregulation, and a complex microenvironment, making it less responsive to monotherapy with immunotherapy (126, 127). Recent studies have reported interactions between genetic mutations and immune checkpoints in prostate cancer, indicating that the loss of PTEN and p53 induces the expression of B7-H3, an immune checkpoint molecule, and that elevated B7-H3 contributes to tumor growth and immune suppression of T cells and NK cells in PTEN/p53-deficient tumors (128). Additionally, anti-angiogenesis therapy not only prunes blood vessels essential for cancer growth and metastasis but also reprograms the tumor immune microenvironment (129). Consequently, combination therapy with ICIs and anti-angiogenesis agents can effectively induce tumor regression in some cancer patients. Nevertheless, achieving durable remission remains challenging for advanced prostate cancer patients. Further research has revealed a connection between gene mutations and anti-angiogenic therapy. In prostate cancer, restoring PTEN activity by inhibiting the PI3K-Akt pathway can re-sensitize cancer cells to anti-angiogenic therapy (130). AR can upregulate epidermal growth factor receptor expression in prostate cancer cells (131). These findings suggest that further exploration into the relationship between genomic mutations, immune checkpoints, and anti-angiogenesis may offer innovative approaches to prostate cancer treatment.

The treatment for patients with metastatic prostate cancer includes radiopharmaceuticals in addition to the drugs listed above. There are many types of radiopharmaceuticals used in mCRPC patients. Strontium-89 (89Sr) has been shown to be very effective in the treatment of patients with chemotherapy-refractory bone metastases (132). Samarium-153 (153Sm) lexidronam (EDTMP) has also been shown to provide significant pain relief in patients with bone metastases (133). The most recent radiopharmaceutical available is lutetium-177 (177Lu). The newest radiopharmaceutical currently on the market is lutetium-177 (177Lu)-PSMA-617, which was approved on 23 March 2022 by the US Food and Drug Administration. Patients are eligible for this treatment if they have mCRPC, have been previously treated with Androgen pathway inhibitors (ARPI) and type chemotherapy, and have positive prostate-specific membrane antigen (PSMA) imaging, indicating PSMA expression in metastatic lesions (134). More research into PSMA-targeted therapies is currently underway. Over the next decade, radiopharmaceuticals may play a central role in the treatment of patients with advanced prostate cancer.

9 Conclusion

Rb, PTEN, WNT, p53, MYC and AR and their interactions play important roles in regulating prostate cancer development. Investigating the mechanisms of interaction between various pathways and genes can help to identify new common target genes and provide more effective therapeutic strategies to address drug resistance in CRPC. In addition, treatment of these genes and pathways in combination with immune checkpoints, anti-

angiogenesis and radiopharmaceuticals may offer innovative approaches to prostate cancer treatment. Such insights could inform the selection of therapeutic strategies, thereby establishing a robust foundation for the treatment of prostate cancer.

Author contributions

FW: Writing – original draft, Writing – review & editing. HZ: Writing – review & editing. MH: Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

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

Publisher's note

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. (2024) 74:12–49. doi: 10.3322/caac.21820

2. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2024) 74:229–63. doi: 10.3322/caac.21834

3. Bergengren O, Pekala KR, Matsoukas K, Fainberg J, Mungovan SF, Bratt O, et al. 2022 Update on prostate cancer epidemiology and risk factors-A systematic review. *Eur Urol*. (2023) 84:191–206. doi: 10.1016/j.eururo.2023.04.021

4. Steinberg GD, Carter BS, Beaty TH, Childs B, Walsh PC. Family history and the risk of prostate cancer. *Prostate*. (1990) 17:337–47. doi: 10.1002/pros.2990170409

5. Oczkowski M, Dziendzikowska K, Pasternak-Winiarska A, Włodarek D, Gromadzka-Ostrowska J. Dietary factors and prostate cancer development, progression, and reduction. *Nutrients*. (2021) 13:496. doi: 10.3390/nu13020496

6. Rivera-Izquierdo M, Pérez de Rojas J, Martínez-Ruiz V, Pérez-Gómez B, Sánchez M-J, Khan KS, et al. Obesity as a risk factor for prostate cancer mortality: A systematic review and dose-response meta-analysis of 280,199 patients. *Cancers*. (2021) 13:4169. doi: 10.3390/cancers13164169

7. Dovey ZS, Nair SS, Chakravarty D, Tewari AK. Racial disparity in prostate cancer in the African American population with actionable ideas and novel immunotherapies. *Cancer Rep (Hoboken NJ)*. (2021) 4:e1340. doi: 10.1002/cnr2.1340

8. Powell IJ, Bock CH, Ruterbusch JJ, Sakr W. Evidence Supports a Faster Growth Rate and/or Earlier Transformation to Clinically Significant Prostate Cancer in Black than in White American Men, and Influences Racial Progression and Mortality Disparity. *J Urol*. (2010) 183:1792–6. doi: 10.1016/j.juro.2010.01.015

9. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. Active monitoring, radical prostatectomy and radical radiotherapy in PSA-detected clinically localised prostate cancer: the protecT three-arm RCT. *Health Technol Assess*. (2020) 24:1–176. doi: 10.3310/hta24370

10. Senapati D, Kumari S, Heemers HV. Androgen receptor co-regulation in prostate cancer. *Asian J Urol*. (2020) 7:219–32. doi: 10.1016/j.ajur.2019.09.005

11. Schweizer MT, Yu EY. AR-signaling in human Malignancies: prostate cancer and beyond. *Cancers (Basel)*. (2017) 9:7. doi: 10.3390/cancers9010007

12. Nanda JS, Koganti P, Perri G, Ellis L. Phenotypic plasticity - alternate transcriptional programs driving treatment resistant prostate cancer. *Crit Rev Oncog*. (2022) 27:45–60. doi: 10.1615/CritRevOncog.2022043096

13. Pienta KJ, Bradley D. Mechanisms underlying the development of androgen-independent prostate cancer. *Clin Cancer Res*. (2006) 12:1665–71. doi: 10.1158/1078-0432.CCR-06-0067

14. Montgomery RB, Mostaghel EA, Vessella R, Hess DL, Kalhorn TF, Higano CS, et al. Maintenance of intratumoral androgens in metastatic prostate cancer: A mechanism for castration-resistant tumor growth. *Cancer Res*. (2008) 68:4447–54. doi: 10.1158/0008-5472.CAN-08-0249

15. Cato L, Shomali M. AR structural variants and prostate cancer. *Adv Exp Med Biol*. (2022) 1390:195–211. doi: 10.1007/978-3-031-11836-4_11

16. Fletcher C. AR-v7 liquid biopsy for treatment stratification in prostate cancer: how close are we? *Curr Opin Urol*. (2017) 27:500–9. doi: 10.1097/MOU.0000000000000416

17. Antonarakis ES, Lu C, Luber B, Wang H, Chen Y, Nakazawa M, et al. Androgen receptor splice variant 7 and efficacy of taxane chemotherapy in patients with metastatic castration-resistant prostate cancer. *JAMA Oncol*. (2015) 1:582–91. doi: 10.1001/jamaoncol.2015.1341

18. Desai K, McManus JM, Sharifi N. Hormonal therapy for prostate cancer. *Endocr Rev*. (2021) 42:354–73. doi: 10.1210/endrev/bnab002

19. Davis ID, Martin AJ, Stockler MR, Begbie S, Chi KN, Chowdhury S, et al. Enzalutamide with standard first-line therapy in metastatic prostate cancer. *N Engl J Med*. (2019) 381:121–31. doi: 10.1056/NEJMoa1903835

20. Yanagisawa T, Fukuokaya W, Hatakeyama S, Narita S, Muramoto K, Katsumi K, et al. Comparison of abiraterone, enzalutamide, and apalutamide for metastatic hormone-sensitive prostate cancer: A multicenter study. *Prostate*. (2024) 85(2):165–74. doi: 10.1002/pros.24813

21. Lu C, Terbuch A, Dolling D, Yu J, Wang H, Chen Y, et al. Treatment with abiraterone and enzalutamide does not overcome poor outcome from metastatic castration-resistant prostate cancer in men with the germline homozygous HSD3B1 c.1245C genotype. *Ann Oncol*. (2020) 31:1178–85. doi: 10.1016/j.annonc.2020.04.473

22. Scher HI, Fizazi K, Saad F, Taplin M-E, Sternberg CN, Miller K, et al. Increased survival with enzalutamide in prostate cancer after chemotherapy. *N Engl J Med*. (2012) 367:1187–97. doi: 10.1056/NEJMoa1207506

23. de Bono JS, Logothetis CJ, Molina A, Fizazi K, North S, Chu L, et al. Abiraterone and increased survival in metastatic prostate cancer. *N Engl J Med*. (2011) 364:1995–2005. doi: 10.1056/NEJMoa1014618

24. Nijhout HF. Development and evolution of adaptive polyphenisms. *Evol Dev*. (2003) 5:9–18. doi: 10.1046/j.1525-142x.2003.03003.x

25. Lee WH, Bookstein R, Hong F, Young LJ, Shew JY, Lee EY. Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science*. (1987) 235:1394–9. doi: 10.1126/science.3823889

26. Friend SH, Bernards R, Rogelj S, Weinberg RA, Rapaport JM, Albert DM, et al. Human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature*. (1986) 323:643–6. doi: 10.1038/323643a0

27. Henley SA, Dick FA. The retinoblastoma family of proteins and their regulatory functions in the mammalian cell division cycle. *Cell Div*. (2012) 7:10. doi: 10.1186/1747-1028-7-10

28. Barnum KJ, O'Connell MJ. Cell cycle regulation by checkpoints. *Methods Mol Biol*. (2014) 1170:29–40. doi: 10.1007/978-1-4939-0888-2_2

29. Weinberg RA. The retinoblastoma protein and cell cycle control. *Cell*. (1995) 81:323–30. doi: 10.1016/0092-8674(95)90385-2

30. Narasimha AM, Kaulich M, Shapiro GS, Choi YJ, Sicinski P, Dowdy SF. Cyclin D activates the rb tumor suppressor by mono-phosphorylation. *Elife*. (2014) 3:e02872. doi: 10.7554/eLife.02872

31. Serrano M, Hannon GJ, Beach DA. New regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature*. (1993) 366:704–7. doi: 10.1038/366704a0

32. Qiu X, Pascal LE, Song Q, Zang Y, Ai J, O'Malley KJ, et al. Physical and functional interactions between ELL2 and RB in the suppression of prostate cancer cell proliferation, migration, and invasion. *Neoplasia*. (2017) 19:207–15. doi: 10.1016/j.neo.2017.01.001

33. Maddison LA, Sutherland BW, Barrios RJ, Greenberg NM. Conditional deletion of Rb causes early stage prostate cancer. *Cancer Res*. (2004) 64:6018–25. doi: 10.1158/0008-5472.CAN-03-2509

34. Labrecque MP, Takhar MK, Nason R, Santacruz S, Tam KJ, Massah S, et al. The retinoblastoma protein regulates hypoxia-inducible genetic programs, tumor cell invasiveness and neuroendocrine differentiation in prostate cancer cells. *Oncotarget*. (2016) 7:24284–302. doi: 10.18632/oncotarget.8301

35. Thangavel C, Boopathi E, Liu Y, Haber A, Ertel A, Bhardwaj A, et al. RB loss promotes prostate cancer metastasis. *Cancer Res*. (2017) 77:982–95. doi: 10.1158/0008-5472.CAN-16-1589

36. Jin X, Ding D, Yan Y, Li H, Wang B, Ma L, et al. Phosphorylated RB promotes cancer immunity by inhibiting NF-κB activation and PD-L1 expression. *Mol Cell*. (2019) 73:22–35.e6. doi: 10.1016/j.molcel.2018.10.034

37. Cunha GR, Donjacour AA, Cooke PS, Mee S, Bigsby RM, Higgins SJ, et al. The endocrinology and developmental biology of the prostate. *Endocr Rev*. (1987) 8:338–62. doi: 10.1210/edrv-8-3-338

38. Fribourg AF, Knudsen KE, Strobeck MW, Lindhorst CM, Knudsen ES. Differential requirements for ras and the retinoblastoma tumor suppressor protein in the androgen dependence of prostatic adenocarcinoma cells. *Cell Growth Differ*. (2000) 11:361–72.

39. Knudsen KE, Arden KC, Cavenee WK. Multiple G1 regulatory elements control the androgen-dependent proliferation of prostatic carcinoma cells. *J Biol Chem*. (1998) 273:20213–22. doi: 10.1074/jbc.273.32.20213

40. Sharma A, Yeow W-S, Ertel A, Coleman I, Clegg N, Thangavel C, et al. The retinoblastoma tumor suppressor controls androgen signaling and human prostate cancer progression. *J Clin Invest*. (2010) 120:4478–92. doi: 10.1172/JCI44239

41. Gupta S, Vanderbilt C, Abida W, Fine SW, Tickoo SK, Al-Ahmadie HA, et al. Immunohistochemistry-based assessment of androgen receptor status and the AR-null phenotype in metastatic castrate resistant prostate cancer. *Prostate Cancer Prostatic Dis*. (2020) 23:507–16. doi: 10.1038/s41391-020-0214-6

42. Han W, Liu M, Han D, Li M, Toure AA, Wang Z, et al. RB1 loss in castration-resistant prostate cancer confers vulnerability to LSD1 inhibition. *Oncogene*. (2022) 41:852–64. doi: 10.1038/s41388-021-02135-3

43. Mandigo AC, Shafi AA, McCann JJ, Yuan W, Laufer TS, Bogdan D, et al. Novel oncogenic transcription factor cooperation in RB-deficient cancer. *Cancer Res*. (2022) 82:221–34. doi: 10.1158/0008-5472.CAN-21-1159

44. Macleod KF. The RB tumor suppressor: A gatekeeper to hormone independence in prostate cancer? *J Clin Invest*. (2010) 120:4179–82. doi: 10.1172/JCI45406

45. Balk SP, Knudsen KE. AR, the cell cycle, and prostate cancer. *Nucl Recept Signal*. (2008) 6:e001. doi: 10.1621/nrs.06001

46. Finn RS, Rugo HS, Gelmon KA, Cristofanilli M, Colleoni M, Loi S, et al. Long-term pooled safety analysis of palbociclib in combination with endocrine therapy for hormone receptor-positive/human epidermal growth factor receptor 2-negative advanced breast cancer: updated analysis with up to 5 Years of follow-up. *Oncologist*. (2021) 26:e749–55. doi: 10.1002/onco.13684

47. Goetz MP, Toi M, Huober J, Sohn J, Trédan O, Park IH, et al. Abemaciclib plus a nonsteroidal aromatase inhibitor as initial therapy for HR+, HER2- advanced breast cancer: final overall survival results of MONARCH 3. *Ann Oncol*. (2024) 35:718–27. doi: 10.1016/j.annonc.2024.04.013

48. Mayayo-Peralta I, Faggion B, Hoekman L, Morris B, Lieftink C, Goldsbrough I, et al. Ribociclib induces broad chemotherapy resistance and EGFR dependency in ESR1 wildtype and mutant breast cancer. *Cancers*. (2021) 13:6314. doi: 10.3390/cancers13246314

49. Peuker CA, Yaghobramzi S, Grunert C, Keilholz L, Gjerga E, Hennig S, et al. Treatment with ribociclib shows favourable immunomodulatory effects in patients with hormone receptor-positive breast cancer-findings from the RIBECCA trial. *Eur J Cancer (Oxford England: 1990)*. (2022) 162:45–55. doi: 10.1016/j.ejca.2021.11.025

50. Smith MR, Piulats JM, Todenhöfer T, Lee J-L, Arranz Arija JA, Mazilu L, et al. CYCLONE 2: A phase 3 study of abemaciclib with abiraterone in patients with metastatic castration-resistant prostate cancer. *JCO*. (2024) 42:5001–1. doi: 10.1200/JCO.2024.42.16_suppl.5001

51. Georgescu MM, Kirsch KH, Akagi T, Shishido T, Hanafusa H. The tumor-suppressor activity of PTEN is regulated by its carboxyl-terminal region. *Proc Natl Acad Sci U.S.A*. (1999) 96:10182–7. doi: 10.1073/pnas.96.18.10182

52. Fruman DA, Chiu H, Hopkins BD, Bagrodia S, Cantley LC, Abraham RT. The PI3K pathway in human disease. *Cell*. (2017) 170:605–35. doi: 10.1016/j.cell.2017.07.029

53. Maehama T, Dixon JE. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem*. (1998) 273:13375–8. doi: 10.1074/jbc.273.22.13375

54. Choudhury AD. PTEN-PI3K pathway alterations in advanced prostate cancer and clinical implications. *Prostate*. (2022) 82 Suppl 1:S60–72. doi: 10.1002/pros.24372

55. Majumder PK, Febbo PG, Bikoff R, Berger R, Xue Q, McMahon LM, et al. mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. *Nat Med*. (2004) 10:594–601. doi: 10.1038/nm1052

56. Abou-Ouf H, Assem H, Ghosh S, Karnes RJ, Stoletov K, Palanisamy N, et al. High serine-arginine protein kinase 1 expression with PTEN loss defines aggressive phenotype of prostate cancer associated with lethal outcome and decreased overall survival. *Eur Urol Open Sci*. (2021) 23:1–8. doi: 10.1016/j.euros.2020.11.005

57. Alzoubi A, Al Bashir S, Smairat A, Alrawashdeh A, Haddad H, Kheirallah K. PTEN loss is not a determinant of time to castration-resistance following androgen-deprivation therapy in prostate cancer: A study from Jordan. *J Med Life*. (2023) 16:593–8. doi: 10.25122/jml-2023-0034

58. Jamaspishvili T, Berman DM, Ross AE, Scher HI, De Marzo AM, Squire JA, et al. Clinical implications of PTEN loss in prostate cancer. *Nat Rev Urol*. (2018) 15:222–34. doi: 10.1038/nrurol.2018.9

59. Lotan TL, Wei W, Morais CL, Hawley ST, Fazli L, Hurtado-Coll A, et al. PTEN loss as determined by clinical-grade immunohistochemistry assay is associated with worse recurrence-free survival in prostate cancer. *Eur Urol Focus*. (2016) 2:180–8. doi: 10.1016/j.euf.2015.07.005

60. Roudsari NM, Lashgari N-A, Momtaz S, Abaft S, Jamali F, Safaiepour P, et al. Inhibitors of the PI3K/Akt/mTOR pathway in prostate cancer chemoprevention and intervention. *Pharmaceutics*. (2021) 13:1195. doi: 10.3390/pharmaceutics13081195

61. Wang Y, Mikhailova M, Bose S, Pan C-X, deVere White RW, Ghosh PM. Regulation of androgen receptor transcriptional activity by rapamycin in prostate cancer cell proliferation and survival. *Oncogene*. (2008) 27:7106–17. doi: 10.1038/onc.2008.318

62. Sweeney CJ, Percent IJ, Babu S, Cultrera JL, Mehlhaff BA, Goodman OB, et al. Phase Ib/II study of enzalutamide with samotolisib (LY3023414) or placebo in patients with metastatic castration-resistant prostate cancer. *Clin Cancer Res*. (2022) 28:2237–47. doi: 10.1158/1078-0432.CCR-21-2326

63. Sweeney C, Bracarda S, Sternberg CN, Chi KN, Olmos D, Sandhu S, et al. Ipatasertib plus abiraterone and prednisolone in metastatic castration-resistant prostate cancer (IPATential150): A multicentre, randomised, double-blind, phase 3 trial. *Lancet*. (2021) 398:131–42. doi: 10.1016/S0140-6736(21)00580-8

64. Nitulescu GM, Margina D, Juzenas P, Peng Q, Olaru OT, Saloustros E, et al. Akt inhibitors in cancer treatment: the long journey from drug discovery to clinical use (Review). *Int J Oncol*. (2016) 48:869–85. doi: 10.3892/ijo.2015.3306

65. Nusse R, Clevers H. Wnt/β-catenin signaling, disease, and emerging therapeutic modalities. *Cell*. (2017) 169:985–99. doi: 10.1016/j.cell.2017.05.016

66. Clevers H, Nusse R. Wnt/β-catenin signaling and disease. *Cell*. (2012) 149:1192–205. doi: 10.1016/j.cell.2012.05.012

67. Bisson I, Prowse DM. WNT signaling regulates self-renewal and differentiation of prostate cancer cells with stem cell characteristics. *Cell Res*. (2009) 19:683–97. doi: 10.1038/cr.2009.43

68. Wang B, Wang X, Long JE, Eastham-Anderson J, Firestein R, Junttila MR. Castration-resistant Lgr5(+) cells are long-lived stem cells required for prostatic regeneration. *Stem Cell Rep*. (2015) 4:768–79. doi: 10.1016/j.stemcr.2015.04.003

69. Wang Y, Singhal U, Qiao Y, Kasputis T, Chung J-S, Zhao H, et al. Wnt signaling drives prostate cancer bone metastatic tropism and invasion. *Transl Oncol*. (2020) 13:100747. doi: 10.1016/j.tranon.2020.100747

70. Lee SH, Luong R, Johnson DT, Cunha GR, Rivina L, Gonzalgo ML, et al. Androgen signaling is a confounding factor for β-catenin-mediated prostate tumorigenesis. *Oncogene*. (2016) 35:702–14. doi: 10.1038/onc.2015.117

71. Pak S, Park S, Kim Y, Park J-H, Park C-H, Lee K-J, et al. The small molecule WNT/β-catenin inhibitor CWP232291 blocks the growth of castration-resistant prostate cancer by activating the endoplasmic reticulum stress pathway. *J Exp Clin Cancer Res*. (2019) 38:342. doi: 10.1186/s13046-019-1342-5

72. Mckay R. *A Phase 1b Trial Investigating Docetaxel Combined with Cirmtuzumab in Patients with Metastatic Castration Resistant Prostate Cancer*. clinicaltrials.gov (2024).

73. WntResearch AB. *Phase I Dose Escalating Study to Evaluate the Safety, Tolerability, Anti-Tumour Activity and Pharmacokinetic and Pharmacodynamic Profiles of Foxy-5 in Patients With Metastatic Breast, Colon or Prostate Cancer*. clinicaltrials.gov (2016).

74. Linzer DI, Levine AJ. Characterization of a 54K Dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell*. (1979) 17:43–52. doi: 10.1016/0092-8674(79)90293-9

75. Harper JW, Adami GR, Wei N, Keyomarsi K, Elledge SJ. The P21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell*. (1993) 75:805–16. doi: 10.1016/0092-8674(93)90499-g

76. Martín-Caballero J, Flores JM, García-Palencia P, Serrano M. Tumor susceptibility of P21(Waf1/Cip1)-deficient mice. *Cancer Res*. (2001) 61:6234–8.

77. Heidenberg HB, Bauer JJ, McLeod DG, Moul JW, Srivastava S. The role of the P53 tumor suppressor gene in prostate cancer: A possible biomarker? *Urology*. (1996) 48:971–9. doi: 10.1016/s0090-4295(96)00365-2

78. Donehower LA, Soussi T, Korkut A, Liu Y, Schultz A, Cardenas M, et al. Integrated analysis of TP53 gene and pathway alterations in the cancer genome atlas. *Cell Rep*. (2019) 28:1370–1384.e5. doi: 10.1016/j.celrep.2019.07.001

79. Ding D, Blee AM, Zhang J, Pan Y, Becker NA, Maher LJ, et al. Gain-of-function mutant P53 together with ERG proto-oncogene drive prostate cancer by beta-catenin activation and pyrimidine synthesis. *Nat Commun*. (2023) 14:4671. doi: 10.1038/s41467-023-40352-4

80. Cotter K, Rubin MA. The evolving landscape of prostate cancer somatic mutations. *Prostate*. (2022) 82:S13-S24. doi: 10.1002/pros.24353

81. Pa M, Kh V. Mutant P53 in cancer: new functions and therapeutic opportunities. *Cancer Cell*. (2014) 25:304–17. doi: 10.1016/j.ccr.2014.01.021

82. Ecke TH, Schlechte HH, Schiemenz K, Sachs MD, Lenk SV, Rudolph BD, et al. TP53 gene mutations in prostate cancer progression. *Anticancer Res*. (2010) 30:1579–86.

83. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat Med*. (2016) 22:298–305. doi: 10.1038/nm.4045

84. Wang Y, Zhang YX, Kong CZ, Zhang Z, Zhu YY. Loss of P53 facilitates invasion and metastasis of prostate cancer cells. *Mol Cell Biochem*. (2013) 384:121–7. doi: 10.1007/s11010-013-1789-1

85. Liu W, Xie CC, Thomas CY, Kim S-T, Lindberg J, Egevad L, et al. Genetic markers associated with early cancer-specific mortality following prostatectomy. *Cancer*. (2013) 119:2405–12. doi: 10.1002/cncr.27954

86. Navone NM, Troncoso P, Pisters LL, Goodrow TL, Palmer JL, Nichols WW, et al. P53 protein accumulation and gene mutation in the progression of human prostate carcinoma. *J Natl Cancer Inst*. (1993) 85:1657–69. doi: 10.1093/jnci/85.20.1657

87. Fonseca GN, Srougi M, Leite KRM, Nesrallah LJ, Ortiz V. The role of HER2/Neu, BCL2, P53 genes and proliferating cell nuclear protein as molecular prognostic parameters in localized prostate carcinoma. *Sao Paulo Med J*. (2004) 122:124–7. doi: 10.1590/s1516-31802004000300009

88. Schlomm T, Iwers L, Kirstein P, Jessen B, Köllermann J, Minner S, et al. Clinical significance of P53 alterations in surgically treated prostate cancers. *Mod Pathol*. (2008) 21:1371–8. doi: 10.1038/modpathol.2008.104

89. Zhou Z, Flesken-Nikitin A, Corney DC, Wang W, Goodrich DW, Roy-Burman P, et al. Synergy of P53 and Rb deficiency in a conditional mouse model for metastatic prostate cancer. *Cancer Res*. (2006) 66:7889–98. doi: 10.1158/0008-5472.CAN-06-0486

90. Mu P, Zhang Z, Benelli M, Karthaus WR, Hoover E, Chen C-C, et al. SOX2 promotes lineage plasticity and antiandrogen resistance in TP53- and RB1-deficient prostate cancer. *Science*. (2017) 355:84–8. doi: 10.1126/science.aah4307

91. Blee AM, He Y, Yang Y, Ye Z, Yan Y, Pan Y, et al. TMPRSS2-ERG controls luminal epithelial lineage and antiandrogen sensitivity in PTEN and TP53-mutated prostate cancer. *Clin Cancer Res*. (2018) 24:4551–65. doi: 10.1158/1078-0432.CCR-18-0653

92. Martin P, Liu Y-N, Pierce R, Abou-Kheir W, Casey O, Seng V, et al. Prostate epithelial Pten/TP53 loss leads to transformation of multipotential progenitors and epithelial to mesenchymal transition. *Am J Pathol*. (2011) 179:422–35. doi: 10.1016/j.ajpath.2011.03.035

93. Ku SY, Rosario S, Wang Y, Mu P, Seshadri M, Goodrich ZW, et al. Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science*. (2017) 355:78–83. doi: 10.1126/science.aah4199

94. Cronauer MV, Schulz WA, Burchardt T, Ackermann R, Burchardt M. Inhibition of P53 function diminishes androgen receptor-mediated signaling in prostate cancer cell lines. *Oncogene*. (2004) 23:3541–9. doi: 10.1038/sj.onc.1207346

95. Jiang A, Yu C, Zhang P, Chen W, Liu W, Hu X, et al. P53 overexpression represses androgen-mediated induction of NKX3.1 in a prostate cancer cell line. *Exp Mol Med*. (2006) 38:625–33. doi: 10.1038/emm.2006.74

96. Ashikari D, Takayama K, Tanaka T, Suzuki Y, Obinata D, Fujimura T, et al. Androgen induces G3BP2 and SUMO-mediated P53 nuclear export in prostate cancer. *Oncogene*. (2017) 36:6272–81. doi: 10.1038/onc.2017.225

97. Lehmann BD, McCubrey JA, Jefferson HS, Paine MS, Chappell WH, Terrian DMA. Dominant role for P53-dependent cellular senescence in radiosensitization of human prostate cancer cells. *Cell Cycle*. (2007) 6:595–605. doi: 10.4161/cc.6.5.3901

98. Zhou X, Zou L, Chen W, Yang T, Luo J, Wu K, et al. Flubendazole, FDA-approved anthelmintic, elicits valid antitumor effects by targeting P53 and promoting ferroptosis in castration-resistant prostate cancer. *Pharmacol Res*. (2021) 164:105305. doi: 10.1016/j.phrs.2020.105305

99. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human C-Myc Onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci U.S.A*. (1982) 79:7824–7. doi: 10.1073/pnas.79.24.7824

100. Zelinski T, Verville G, White L, Hamerton JL, McAlpine PJ, Lewis M. Confirmation of the assignment of MYCL to chromosome 1 in humans and its position relative to RH, UMPK, and PGM1. *Genomics*. (1988) 2:154–6. doi: 10.1016/0888-7543(88)90097-3

101. Ruiz-Pérez MV, Henley AB, Arsenian-Henriksson M. The MYCN protein in health and disease. *Genes (Basel)*. (2017) 8:113. doi: 10.3390/genes8040113

102. Qiu X, Boufaied N, Hallal T, Feit A, de Polo A, Luoma AM, et al. MYC drives aggressive prostate cancer by disrupting transcriptional pause release at androgen receptor targets. *Nat Commun*. (2022) 13:2559. doi: 10.1038/s41467-022-30257-z

103. Zhang E, Chen Z, Liu W, Lin L, Wu L, Guan J, et al. NCAPG2 promotes prostate cancer Malignancy and stemness via STAT3/c-MYC signaling. *J Transl Med*. (2024) 22:12. doi: 10.1186/s12967-023-04834-9

104. Clegg NJ, Couto SS, Wongvipat J, Hieronymus H, Carver BS, Taylor BS, et al. MYC cooperates with AKT in prostate tumorigenesis and alters sensitivity to mTOR inhibitors. *PloS One*. (2011) 6:e17449. doi: 10.1371/journal.pone.0017449

105. Ellwood-Yen K, Graeber TG, Wongvipat J, Iruela-Arispe ML, Zhang J, Matusik R, et al. Myc-driven murine prostate cancer shares molecular features with human prostate tumors. *Cancer Cell*. (2003) 4:223–38. doi: 10.1016/s1535-6108(03)00197-1

106. Iwata T, Schultz D, Hicks J, Hubbard GK, Mutton LN, Lotan TL, et al. MYC overexpression induces prostatic intraepithelial neoplasia and loss of Nkx3.1 in mouse luminal epithelial cells. *PloS One*. (2010) 5:e9427. doi: 10.1371/journal.pone.0009427

107. Dardenne E, Beltran H, Benelli M, Gayvert K, Berger A, Puca L, et al. N-Myc induces an EZH2-mediated transcriptional program driving neuroendocrine prostate cancer. *Cancer Cell*. (2016) 30:563–77. doi: 10.1016/j.ccell.2016.09.005

108. Arriaga JM, Panja S, Alshalalfa M, Zhao J, Zou M, Giacobbe A, et al. A MYC and RAS co-activation signature in localized prostate cancer drives bone metastasis and castration resistance. *Nat Cancer*. (2020) 1:1082–96. doi: 10.1038/s43018-020-00125-0

109. Hubbard GK, Mutton LN, Khalili M, McMullin RP, Hicks JL, Bianchi-Frias D, et al. Combined MYC activation and Pten loss are sufficient to create genomic instability and lethal metastatic prostate cancer. *Cancer Res*. (2016) 76:283–92. doi: 10.1158/0008-5472.CAN-14-3280

110. Bai S, Cao S, Jin L, Kobelski M, Schouest B, Wang X, et al. A positive role of C-Myc in regulating androgen receptor and its splice variants in prostate cancer. *Oncogene*. (2019) 38:4977–89. doi: 10.1038/s41388-019-0768-8

111. Wyce A, Degenhardt Y, Bai Y, Le B, Korenchuk S, Crouthame M-C, et al. Inhibition of BET bromodomain proteins as a therapeutic approach in prostate cancer. *Oncotarget*. (2013) 4:2419–29. doi: 10.18632/oncotarget.1572

112. Asangani IA, Dommeti VL, Wang X, Malik R, Cieslik M, Yang R, et al. Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature*. (2014) 510:278–82. doi: 10.1038/nature13229

113. Kirschner AN, Wang J, van der Meer R, Anderson PD, Franco-Coronel OE, Kushner MH, et al. PIM kinase inhibitor AZD1208 for treatment of MYC-driven prostate cancer. *J Natl Cancer Inst*. (2015) 107:dju407. doi: 10.1093/jnci/dju407

114. Ton A-T, Singh K, Morin H, Ban F, Leblanc E, Lee J, et al. Dual-inhibitors of N-myc and AURKA as potential therapy for neuroendocrine prostate cancer. *Int J Mol Sci*. (2020) 21:8277. doi: 10.3390/ijms21218277

115. Mo H, Renna CE, Moore HCF, Abraham J, Kruse ML, Montero AJ, et al. Real-world outcomes of everolimus and exemestane for the treatment of metastatic hormone receptor-positive breast cancer in patients previously treated with CDK4/6 inhibitors. *Clin Breast Cancer*. (2022) 22:143–8. doi: 10.1016/j.clbc.2021.10.002

116. Castro E, Goh C, Olmos D, Saunders E, Leongamornlert D, Tymrakiewicz M, et al. Germline BRCA mutations are associated with higher risk of nodal involvement, distant metastasis, and poor survival outcomes in prostate cancer. *J Clin Oncology: Off J Am Soc Clin Oncol*. (2013) 31:1748–57. doi: 10.1200/JCO.2012.43.1882

117. Varol U, Kucukzeybek Y, Alacacioglu A, Somali I, Altun Z, Aktas S, et al. BRCA genes: BRCA 1 and BRCA 2. *J BUON: Off J Balkan Union Oncol*. (2018) 23:862–6.

118. Antonarakis ES, Lu C, Luber B, Liang C, Wang H, Chen Y, et al. Germline DNA-repair gene mutations and outcomes in men with metastatic castration-resistant prostate cancer receiving first-line abiraterone and enzalutamide. *Eur Urol*. (2018) 74:218–25. doi: 10.1016/j.eururo.2018.01.035

119. Roskoski R. Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors. *Pharmacol Res*. (2015) 94:9–25. doi: 10.1016/j.phrs.2015.01.003

120. Machida K, Matsuda S, Yamaki K, Senga T, Thant AA, Kurata H, et al. V-Src suppresses SHPS-1 expression via the Ras-MAP kinase pathway to promote the oncogenic growth of cells. *Oncogene*. (2000) 19:1710–8. doi: 10.1038/sj.onc.1203497

121. Liu H, Xu J, Zhou L, Yun X, Chen L, Wang S, et al. Hepatitis B virus large surface antigen promotes liver carcinogenesis by activating the Src/PI3K/Akt pathway. *Cancer Res*. (2011) 71:7547–57. doi: 10.1158/0008-5472.CAN-11-2260

122. Migliaccio A, Castoria G, Di Domenico M, de Falco A, Bilancio A, Lombardi M, et al. Steroid-induced androgen receptor-oestradiol receptor beta-Src complex triggers prostate cancer cell proliferation. *EMBO J*. (2000) 19:5406–17. doi: 10.1093/emboj/19.20.5406

123. Migliaccio A, Varricchio L, De Falco A, Castoria G, Arra C, Yamaguchi H, et al. Inhibition of the SH3 domain-mediated binding of Src to the androgen receptor and its effect on tumor growth. *Oncogene*. (2007) 26:6619–29. doi: 10.1038/sj.onc.1210487

124. Migliaccio A, Di Domenico M, Castoria G, Nanayakkara M, Lombardi M, de Falco A, et al. Steroid receptor regulation of epidermal growth factor signaling through Src in breast and prostate cancer cells: steroid antagonist action. *Cancer Res*. (2005) 65:10585–93. doi: 10.1158/0008-5472.CAN-05-0912

125. Jonsson Comprehensive Cancer Center. *An Open-Label, Neoadjuvant Phase 2 Study Comparing the Effects of AR Inhibition With and Without SRC or MEK Inhibition on the Development of EMT in Prostate Cancer*. clinicaltrials.gov (2024).

126. Venkatachalam S, McFarland TR, Agarwal N, Swami U. Immune checkpoint inhibitors in prostate cancer. *Cancers*. (2021) 13:2187. doi: 10.3390/cancers13092187

127. Claps M, Mennitto A, Guadalupi V, Sepe P, Stellato M, Zattarin E, et al. Immune-checkpoint inhibitors and metastatic prostate cancer therapy: learning by making mistakes. *Cancer Treat Rev*. (2020) 88:102057. doi: 10.1016/j.ctrv.2020.102057

128. Shi W, Wang Y, Zhao Y, Kim JJ, Li H , Meng C, et al. Immune checkpoint B7-H3 is a therapeutic vulnerability in prostate cancer harboring PTEN and TP53 deficiencies. *Sci Transl Med*. (2023) 15:eadf6724. doi: 10.1126/scitranslmed.adf6724

129. Yi M, Jiao D, Qin S, Chu Q, Wu K, Li A. Synergistic effect of immune checkpoint blockade and anti-angiogenesis in cancer treatment. *Mol Cancer*. (2019) 18:60. doi: 10.1186/s12943-019-0974-6

130. Makhov PB, Golovine K, Kutikov A, Teper E, Canter DJ, Simhan J, et al. Modulation of Akt/mTOR signaling overcomes sunitinib resistance in renal and prostate cancer cells. *Mol Cancer Ther*. (2012) 11:1510–7. doi: 10.1158/1535-7163.MCT-11-0907

131. Pignon J-C, Koopmansch B, Nolens G, Delacroix L, Waltregny D, Winkler R. Androgen receptor controls EGFR and ERBB2 gene expression at different levels in prostate cancer cell lines. *Cancer Res*. (2009) 69:2941–9. doi: 10.1158/0008-5472.CAN-08-3760

132. Gunawardana DH, Lichtenstein M, Better N, Rosenthal M. Results of strontium-89 therapy in patients with prostate cancer resistant to chemotherapy. *Clin Nucl Med*. (2004) 29:81–5. doi: 10.1097/01.rlu.0000109721.58471.44

133. Serafini AN, Houston SJ, Resche I, Quick DP, Grund FM, Ell PJ, et al. Palliation of pain associated with metastatic bone cancer using samarium-153 lexidronam: A double-blind placebo-controlled clinical trial. *J Clin Oncol*. (1998) 16:1574–81. doi: 10.1200/JCO.1998.16.4.1574

134. FDA approves pluvicto/locametz for metastatic castration-resistant prostate cancer. *J Nucl Med: Off Publication Soc Nucl Med*. (2022) 63:13N.

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact