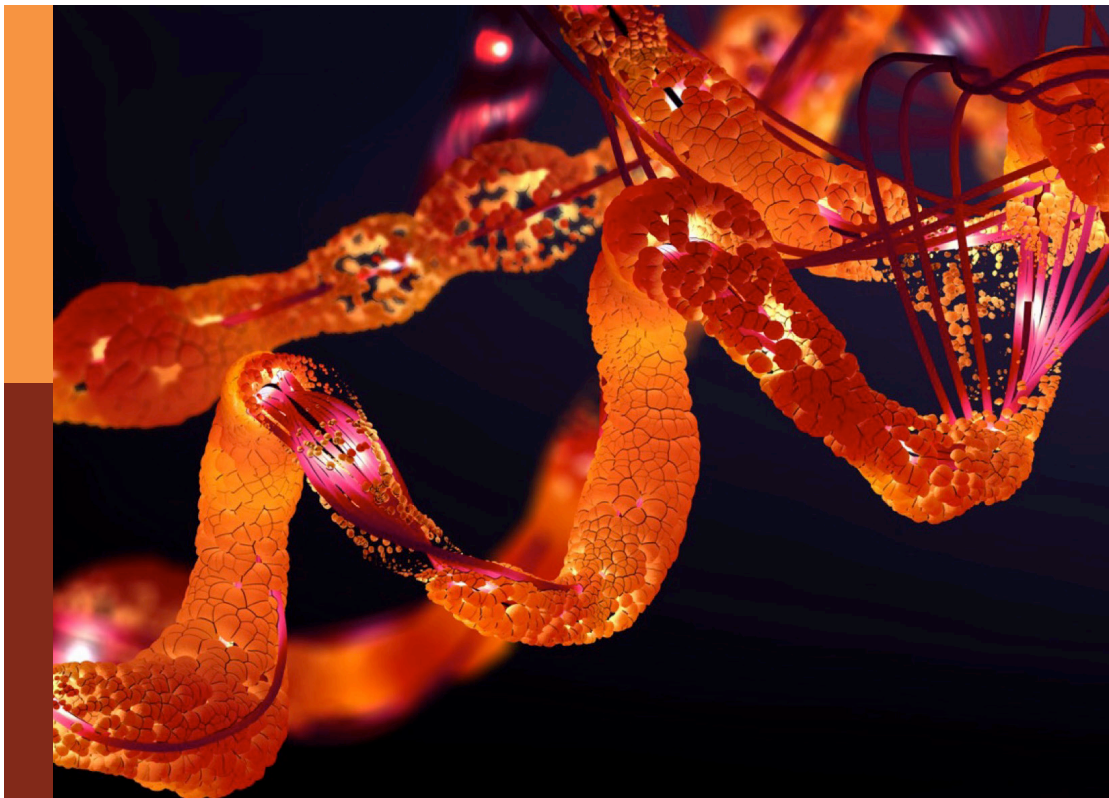
frontiers | Research Topics

# Cancer diagnostics in solid tumors - from pathology to precision oncology

**Edited by**
Pedro Borralho, Fernando Schmitt, Liang Wang and Umberto Malapelle

**Published in**
Frontiers in Molecular Biosciences



ISSN 1664-8714
ISBN 978-2-83251-837-3
DOI 10.3389/978-2-83251-837-3

## About Frontiers

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

## Frontiers journal series

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

## Dedication to quality

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

## What are Frontiers Research Topics?

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

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

# Cancer diagnostics in solid tumors - from pathology to precision oncology

**Topic editors**

Pedro Borralho — Squad Lead—Novartis Oncology, Portugal

Fernando Schmitt — University of Porto, Portugal

Liang Wang — Department of Tumor Biology, Moffitt Cancer Center & Research Institute, United States

Umberto Malapelle — University of Naples Federico II, Italy

**Citation**

Borralho, P., Schmitt, F., Wang, L., Malapelle, U., eds. (2023). *Cancer diagnostics in solid tumors - from pathology to precision oncology*.

Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-837-3

# Table of contents

frontiers | Frontiers in Molecular Biosciences

TYPE Editorial
PUBLISHED 21 February 2023
DOI 10.3389/fmolb.2023.1150641



OPEN ACCESS

EDITED BY
William C. Cho,
QEH, Hong Kong SAR, China

REVIEWED BY
Yingcheng Wu,
Fudan University, China

*CORRESPONDENCE
Fernando Schmitt,
fschmitt@ipatimup.pt

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 24 January 2023
ACCEPTED 14 February 2023
PUBLISHED 21 February 2023

CITATION
Malapelle U, Borralho P, Wang L and Schmitt F (2023), Editorial: Cancer diagnostics in solid tumors-from pathology to precision oncology. *Front. Mol. Biosci.* 10:1150641. doi: 10.3389/fmolb.2023.1150641



# Editorial: Cancer diagnostics in solid tumors-from pathology to precision oncology

Umberto Malapelle[1], Pedro Borralho[2], Liang Wang[3] and Fernando Schmitt[4]*

[1]Department of Public Health, University of Naples Federico II, Naples, Italy, [2]Novartis Farma - Produtos Farmacêuticos, Porto Salvo, Portugal, [3]Department of Tumor Biology, Moffitt Cancer Center & Research Institute Tampa, Tampa, FL, United States, [4]Department of Pathology, Faculty of Medicine of University of Porto, Porto, Portugal

KEYWORDS

**precision oncology, molecular pathology, tumors, molecular biology, pathologist**

Editorial on the Research Topic
Cancer diagnostics in solid tumors-from pathology to precision oncology

In the novel era of target treatments for advanced stage cancer patients, personalized medicine is a rapidly evolving field. This phenomenon led to a paradigm shift in disease diagnosis and treatment, from the traditional patient stratification based on phenotypic biomarkers to ever-increasing customization and tailoring of diagnostic and therapeutic interventions. In this scenario, molecular testing on biological specimens, including tissue and liquid biopsies, plays a key role in the assessment of clinically relevant biomarkers status with diagnostic, prognostic, and predictive values to tailor and track, therapeutic interventions, and outcomes along the full patient treatment continuum. (La Thangue and Kerr, 2011). In this setting, modern pathologists, who are long-standing gatekeepers of cancer diagnostics, should evolve with the development of a novel figure of molecular pathologist. (Pisapia et al., 2022). Currently, this figure plays a pivotal role in multidisciplinary teams and molecular tumor boards, bridging anatomic pathology with cellular and molecular biology, being increasingly involved in multidisciplinary treatment decisions and response monitoring, far beyond their traditional role in cancer diagnosis. (Pisapia et al., 2022).

Overall, in this Research Topic of Frontiers in Molecular Biosciences, we attempt to address the state of art and the future perspectives in the field of precision medicine and integrative pathology.

Regarding lung cancer, the adoption of novel molecular approaches enables the possibility to add more pathological, biomolecular, and clinical information to rare and non-codified clinical entities, with an increase in the knowledge and the expertise to manage this specific histotypes, as reported by Parente et al. Another interesting field of investigation is represented by micro-RNA (miRNA). As reported by Tiang et al., miR-125b-5p may be adopted as a prognostic marker and a therapeutic target for lung adenocarcinoma. (Tang et al.). Molecular biology may be used not only in advanced stage of disease but also to predict the prognosis of early-stage lung adenocarcinoma and guide the selection of adjuvant therapy for early-stage lung adenocarcinoma patients, as demonstrated by Zhao et al. (Zhao Y et al., 2022).

As far as central nervous system neoplasms are concerned, Zheng et al. highlighted that a 21 gene-pair signature based on relative expression ordering may be useful to identify high-risk low grade gliomas patients in order to guide timely intervention. (Zheng et al.). In addition, further

advances have been made in the field of glioblastoma management with the identification of TSPAN4 expression as a prognostic and immune target and RCN1 as a potential targeted in these patients. (Lu et al.), (Zheng et al.).

Considering breast cancer, novel therapeutic strategies have been developed. In particular, triple negative breast cancer patients may benefit from immune-checkpoint inhibitors after PD-L1 evaluation, (Ribeiro et al.), whereas *PIK3CA* mutated advanced ER+/HER2-breast cancer patients can be treated with specific PIK3CA inhibitors (Peixoto et al.), with a significant improvement in clinical outcomes.

Interesting results have also been obtained in gastric cancer. In particular, Hong *et al* proposed an integrated analysis of prognostic mRNA signature in early- and progressive-stage gastric adenocarcinomas (Hong et al.).

Regarding hepatocellular carcinoma, Zhao et al. reported that ALG3 was overexpressed and can be considered a potential indicator of survival in these patients. (Zhao et al., 2022). Remarkably, as it has been reported by Roa-Colomo et al. the identification and characterization of circulating epithelial cells by ASGR1 and/or miR-122-5p expression may be used as a risk-stratification and independent prognostic tool in liver cirrhosis and early stage hepatocellular carcinoma patients. (Roa-Colomo et al.).

Considering pancreatic cancer, it has been demonstrated that a comprehensive analysis of Lysyl Oxidase family members, in particular LOX and LOXL2, might have a prognostic and predictive value in pancreatic adenocarcinoma. (Jiang et al.).

A novel prognostic stratification model based on neutrophil extracellular traps signature has been proposed by Chen et al. for Head and Neck squamous cell carcinoma. (Chen et al.).

As far as genital-urinary tract is concerned, important advances have been made. In particular, Xia et al. proposed a novel gene signature associated with "E2F target" to predict the prognosis in prostate cancer (Xia et al.). Considering kidney, a significant prognostic role was associated with mTORC1 signaling pathway-related genes and pyroptosis-related genes and tumor microenvironment infiltration (Zhang et al.; Zhanghuang et al.).

Beyond a specific cancer type, Wu et al. highlighted the oncogenic role of SNRPB expression. (Wu et al.).

Overall, this Research Topic has highlighted the role of molecular testing and of the novel figure of molecular pathologist in the management of cancer patients.

Considering the implementation of next-generation approaches, based on spatial transcriptomics and digital pathology, and in particular the possibility to simultaneously combine these data with the molecular status, the evolving field of modern Anatomic Pathology will be definitely revolutionized during the next years. In this setting, pathologists should take care not only of morphological aspects but also of the integrated data coming from different cancer types, also by using Artificial Intelligence based tools. (Bonizzi et al.; Souza da Silva et al.). Ongoing research is warranted to improve the clinical outcome of cancer patients.

## Author contributions

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

## Conflict of interest

Author PB is employed by Novartis Farma—Produtos Farmacêuticos, S.A., Porto Salvo, Portugal. UM has received personal fees (as consultant and/or speaker bureau) from Boehringer Ingelheim, Roche, MSD, Amgen, Thermo Fisher Scientifics, Eli Lilly, Diaceutics, GSK, Merck and AstraZeneca, Janssen, Diatech, Novartis, and Hedera., for work performed outside of the current study.

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

## Publisher's note



## References

La Thangue, N. B., and Kerr, D. J. (2011). Predictive biomarkers: A paradigm shift towards personalized cancer medicine. *Nat. Rev. Clin. Oncol.* 8, 587–596. doi:10.1038/nrclinonc.2011.121

Pisapia, P., L'Imperio, V., Galuppini, F., Sajjadi, E., Russo, A., Cerbelli, B., et al. (2022). The evolving landscape of anatomic pathology. *Crit. Rev. Oncol. Hematol.* 178, 103776. doi:10.1016/j.critrevonc.2022.103776

Tang, L., Yuan, Y., Zhai, H., Wang, J., Zhang, D., Liang, H., et al. (2022). MicroRNA-125b-5p correlates with prognosis and lung adenocarcinoma progression. *Front. Mol. Biosci.* 8, 788690. doi:10.3389/fmolb.2021.788690

Zhao, Y., Qing, B., Xu, C., Zhao, J., Liao, Y., Cui, P., et al. (2022). DNA damage response gene-based subtypes associated with clinical outcomes in early-stage lung adenocarcinoma. *Front. Mol. Biosci.* 9, 901829. doi:10.3389/fmolb.2022.901829

Zheng, Y., Lang, Y., Qi, B., Wang, Y., Gao, W., and Li, T. (2023). TSPAN4 is a prognostic and immune target in Glioblastoma multiforme. *Front. Mol. Biosci.* 9, 1030057. doi:10.3389/fmolb.2022.1030057

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 13 October 2021
doi: 10.3389/fmolb.2021.736947

# Integrative Analyses and Verification of the Expression and Prognostic Significance for RCN1 in Glioblastoma Multiforme

*Weicheng Lu[1†], Hong Chen[2†], Bo Liang[3†], Chaopeng Ou[1], Mingwei Zhang[4]*, Qiuyuan Yue[5]* and Jingdun Xie[1]**

[1]State Key Laboratory of Oncology in Southern China, Department of Anesthesiology, Sun Yat-sen University Cancer Center, Collaborative Innovation for Cancer Medicine, Guangzhou, China, [2]Department of Gastrointestinal Surgery, Fujian Provincial Hospital, Fuzhou, China, [3]Nanjing University of Chinese Medicine, Nanjing, China, [4]Department of Radiation Oncology, The First Affiliated Hospital of Fujian Medical University, Fuzhou, China, [5]Department of Radiology, Fujian Cancer Hospital and Fujian Medical University Cancer Hospital, Fuzhou, China

OPEN ACCESS

**Edited by:**
Liang Wang,
Moffitt Cancer Center and Research Institute, United States

**Reviewed by:**
Anupam Nath Jha,
Tezpur University, India
Xiaoyi Huang,
Harbin Medical University Cancer Hospital, China

***Correspondence:**
Jingdun Xie
xiejd6@mail.sysu.edu.cn
Qiuyuan Yue
circlessoo@sina.cn
and Mingwei Zhang
zhangmingwei28@sina.cn

†These authors have contributed equally to this work

**Specialty section:**
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

**Received:** 06 July 2021
**Accepted:** 28 September 2021
**Published:** 13 October 2021

**Citation:**
Lu W, Chen H, Liang B, Ou C, Zhang M, Yue Q and Xie J (2021) Integrative Analyses and Verification of the Expression and Prognostic Significance for RCN1 in Glioblastoma Multiforme. Front. Mol. Biosci. 8:736947. doi: 10.3389/fmolb.2021.736947

Glioblastoma multiform is a lethal primary brain tumor derived from astrocytic, with a poor prognosis in adults. Reticulocalbin-1 (RCN1) is a calcium-binding protein, dysregulation of which contributes to tumorigenesis and progression in various cancers. The present study aimed to identify the impact of RCN1 on the outcomes of patients with Glioblastoma multiforme (GBM). The study applied two public databases to require RNA sequencing data of Glioblastoma multiform samples with clinical data for the construction of a training set and a validation set, respectively. We used bioinformatic analyses to determine that RCN1 could be an independent factor for the overall survival of Glioblastoma multiform patients. In the training set, the study constructed a predictive prognostic model based on the combination of RCN1 with various clinical parameters for overall survival at 0.5-, 1.0-, and 1.5-years, as well as developed a nomogram, which was further validated by validation set. Pathways analyses indicated that RCN1 was involved in KEAS and MYC pathways and apoptosis. *In vitro* experiments indicated that RCN1 promoted cell invasion of Glioblastoma multiform cells. These results illustrated the prognostic role of RCN1 for overall survival in Glioblastoma multiform patients, indicated the promotion of RCN1 in cell invasion, and suggested the probability of RCN1 as a potential targeted molecule for treatment in Glioblastoma multiform.

**Keywords: glioblastoma multiform, invasion, prognosis, nomogram, RCN1**

## INTRODUCTION

Glioblastoma multiforme (GBM) represents the most prevalent brain cancer in adults and has a dismal prognosis and poor quality of life (Omuro and DeAngelis, 2013). The current treatment strategies for GBM are maximum surgical resection followed by a combination of chemotherapy and radiotherapy (Gilbert et al., 2013; Ostrom et al., 2015). Even with the advancement in therapeutic options over recent decades, recent studies have demonstrated that the median survival of GBM patients is 16.6 months, which decreases after 2 years, with a survival rate of only 34% (Gilbert et al., 2013). Several studies have illustrated that some omics markers within tumors could impact patients' survival, like the status of Isocitrate dehydrogenase 1/2 (IDH1/2) mutation, glioma-CpG island

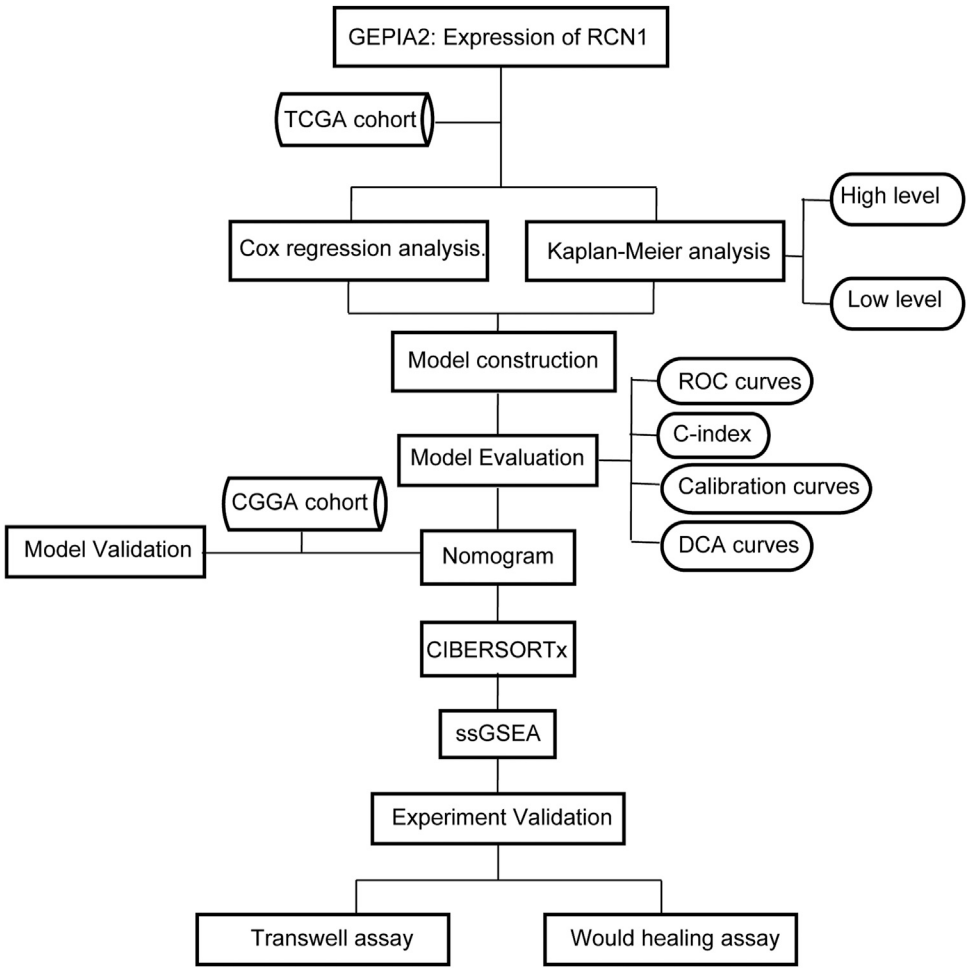



**FIGURE 1 |** Flow chart of this study.

methylator phenotype (G-CIMP), methylation of O-6-methylguanine-DNA methyltransferase (MGMT), and codeletion for chromosome 1p and 19q (1p/19q codeletion) (Hartmann et al., 2010; Wick et al., 2013; Hainfellner et al., 2014; Louis et al., 2016). In addition, further improvements have been made with subspecialized care, improved resection methods precisely targeted radiotherapy, and early systemic salvage therapies (Jayamanne et al., 2018). However, patients with GBM still have a poor prognosis due to the GBM's aggressive behavior, rapid progression, and frequent recurrence (Soffietti et al., 2014). Thus, it is imperative to search for a novel biomarker with good prediction for the prognostic signature of GBM via various methods, to explore the molecular mechanisms for precisely targeted treatments in GBM.

Reticulocalbin-1 (RCN1), a calcium-binding protein, contains six conserved regions and is located in the endoplasmic reticulum (Ozawa and Muramatsu, 1993), which regulates calcium-dependent activities combined with reticulocalbin 2 (RCN2) (Nakakido et al., 2016). Dysregulation of RCN1 protein has been reported in multifarious diseases, including cancer, cardiovascular, and neuromuscular diseases (Grzeskowiak et al., 2003; Liu et al., 1997; Zhang et al., 2006). It has been found that RCN1 is involved in breast cancer (Nakakido et al., 2016), colorectal cancer (Nimmrich et al., 2000), liver cancer (Lu et al., 2015), kidney cancer (Giribaldi et al., 2013), and non-small cell lung cancer (Chen et al., 2019). In addition, it is reported that down-regulation of RCN1 facilitates apoptosis and necroptosis in prostate cancer cells (Liu et al., 2018). Overall, the above findings have revealed that dysregulation of RCN1 could contribute to tumorigenesis and progression. However, the relationship between RCN1 and prognosis, nor its biological functions in GBM, have been completely investigated.

In our study, GBM samples in the Cancer Genome Atlas (TCGA) database were enrolled in the training set, and cases in the Chinese Glioma Genome Atlas (CGGA) were used for the external validation set, to assess the RCN1-related prognostic signature in GBM. Gene Expression Profiling and Interactive Analyses (GEPIA2) (Tang et al., 2017) (http://gepia2.cancer-pku.cn/#index) were used to profile the tissue-wise expression of RCN1 in GBM. In addition, the next version of Cell type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORTx, (Newman et al., 2019) https://

cibersortx.stanford.edu/) was used to illustrate the abundances of the infiltration immune cells correlating with expression of RCN1 in GBM, and the single sample gene search enrichment analyses (ssGSEA) (Barbie et al., 2009) was utilized to determine potential pathways of RCN1 involved (**Figure 1**).

# MATERIALS AND METHODS

## Data Acquisition

The RNA sequencing data (level 3) of GBM with corresponding clinical data were downloaded from the TCGA database to assess the prognostic impact of RCN1-signature on patients with GBM, and were enrolled into the training set. Then, GBM samples from the Chinese Glioma Genome Atlas (CGGA) dataset were employed as an external validation set for further verification.

## Differential Analysis of Reticulocalbin-1 Expression in Glioblastoma Multiforme

Gene Expression Profiling and Interactive Analyses (GEPIA) (Tang et al., 2017) is an online resource for gene expression analysis, and GEPIA2 (http://gepia2.cancer-pku.cn/#index) (Tang et al., 2019) was an updated version of GEPIA, containing 163 GBM cases of TCGA and 207 normal brain samples of the Genome Tissues Expression database (GETx) (GTEx Consortium, 2015). We performed a differential analysis of RCN1 expression through GEPIA2 and used a boxplot to show the results with log2 of transcript count per million [$\log_2$ (TPM + 1)] presenting the expression level of RCN1.

## Survival Analysis

As described previously (Ogłuszka et al., 2019), we used *survMisc* package in R to divide GBM samples into high and low RCN1 groups based on the optimum cutoff. For the exploration of the association between the expression of RCN1 and survival, Kaplan-Meier (KM) analysis with log-rank test was conducted by using *survival* package.

## Evaluation of Reticulocalbin-1-Based Prognostic Value

To explore the effect of the variation of RCN1 expression on OS in GBM patients, we estimated the relative risk and log2-based transformation for survival status, followed by fitting smooth line using a locally estimated scatterplot smoothing (LOESS) method (Zheng et al., 2020). We then compared the predictive effect of RCN1 with different prognostic factors by receiver operating characteristic (ROC) curve (Zhou et al., 2019). We regarded the RCN1 expression as a single continuous covariate for further regression analysis based on the result of LOESS analysis. We adjusted for clinical features, which changed the regression coefficients of RCN1 by more than 10%, or of $p < 0.1$ through the univariate analysis (Kernan et al., 2000). Next, adjust I model was determined after adjusting confounders, while all clinical factors were enrolled as adjust II model. Then, univariate and multivariate analyses using Cox proportional hazard regression model were performed. We also applied LOESS method to visually assess the relationship between RCN1 expression and OS in adjust I and adjust II models, respectively. Interaction test and stratified analysis (Soria et al., 2015) were also carried out to assess the differential prognostic value of RCN1 in accordance with a different model. A two-tailed $p < 0.05$ was considered to be statistically significant.

## Construction and Comparison of Three Prognostic Models

After determining the prognostic features, five clinical prognostic factors in GBM (age, gender, IDH status, chemotherapy, as well as radiotherapy) were enrolled to construct the prognostic model as model 1. Model 2 (the expression of RCN1) and model 3 (model 1 + model 2) were compared with model 1 to estimate the robustness of different prognostic features to GBM prognosis, respectively. We used discrimination, calibration, and model improvement capability to evaluate the different models. The discrimination of these models was estimated by ROC curve, concordance index (C-index) (Harrell et al., 1996), as well as decision curve analysis (DCA) (Vickers and Elkin, 2006; Vickers et al., 2008; Rousson and Zumbrunn, 2011; Kerr et al., 2016). Three corresponding models were constructed with *coxph* function in *survival* package, and then the risk value of patients in each model was calculated with *predict* function, and thus the ROC curves were constructed by *timeROC* (Heagerty et al., 2000; Blanche et al., 2013) with those risk values. Three models were firstly built with *cph* function, and then their C-index at different times with the prediction error curve was calculated by *cindex* function in *pec* package. The *stdca* package was applied to visualization for DCA. The calibration curves of different models were completed via *caoplot* function in *pec* package. Notably, we used the bootstrap method with 1,000 resamples both for analyses of discrimination and calibration. Moreover, the improvement capability of the model was assessed through the net reclassification improvement (NRI) and the integrated discrimination improvement (IDI) by using *IDI.INF.OUT* function in *survIDINRI* package (Pencina et al., 2008). After the best model was determined, *regplot* package was employed to construct the diagram of the nomogram.

## External Validation of the Prognostic Signature

The prognostic capability of the RCN1-based signature was externally validated in the validation set. The gene expression information with corresponding clinical factors of GBM was obtained from the CGGA database. Similarly, we divided GBM patients into high- and low-RCN1 groups by the optimum cut-off value, and then performed the KM survival curve. Afterward, using *predict* function, the time-independent ROC curves, C-index, and calibration curves were employed to assess the accuracy of prognostic model RCN1 signature-based.

## Infiltrative Immune Cell Analysis

CIBERSORTx (Newman et al., 2019), the next-generation version of CIBERSORT, a machine learning tool, provides an estimation of the abundances of member cell types in a condition with a

mixed cell population, by using gene expression data (Newman et al., 2019). We applied CIBERSORTx to analyze the infiltration level of 22 immune cells in high- and low-RCN1 groups, using samples from TCGA and CGGA datasets, respectively. After enabling batch correction, performing the "Bulk mode", and selecting the quantile normalization algorithm, the results were represented with an absolute score for the proportion of 22 immune cell subsets of GBM samples. Consecutively, the samples with $p < 0.05$ were retained after repeating the crossover operation 500 times (Ali et al., 2016). Wilcoxon rank-sum test was applied to identify the differences between the two groups.

## ssGSEA

The ssGSEA was used to identify the differentially enriched hallmarks for a single sample (Barbie et al., 2009). To identify key pathways related to RCN1, we chose to focus on 50 hallmark gene sets, which were designed to highlight gene sets contained in the Molecular Signatures Database (MSigDB) (Subramanian et al., 2005). Gene symbol profiles for *homo sapiens* were downloaded from MSigDB database (Liberzon et al., 2015). Then we estimated the degree of each hallmark's ssGSEA profile in two groups, using the *gsva* package, both in the training and validation cohort. Next, by *limma* package, differential analysis was performed; and $|t| > 1$ or adjusted $p < 0.05$ were considered as statistically significant.

**TABLE 1 |** The clinical characteristics of patients and expression of RCN1 in GBM.

| Characteristics | TCGA (*N* = 153) | CGGA (*N* = 208) |
|---|---|---|
| Age | 60 (52~69) | 53 (43~60) |
| Male | 54 (35.29%) | 82 (39.42%) |
| RCN1 | 59.60 ± 12.29 | 6.12 (5.14~6.62) |
| IDH status | | |
| Mutant | 10 (6.54%) | 31 (14.90%) |
| Wild type | 143 (93.46%) | 177 (85.10%) |
| Radiotherapy | | |
| Yes | 130 (84.97%) | 184 (88.46%) |
| No | 23 (15.03%) | 24 (11.54%) |
| Chemotherapy | | |
| Yes | 112 (73.20%) | 168 (80.77%) |
| No | 41 (26.80%) | 40 (19.23%) |

*Values are expressed as median (interquartile range), number of cases (%), or mean ± standard deviation. TCGA: The Cancer Genome Atlas, CGGA: Chinese Glioma Genome Atlas, RCN1: Reticulocalbin-1, IDH: isocitrate dehydrogenase.*

## Validation of the Effect of Reticulocalbin-1 on Glioblastoma Multiforme Cells

### Cell Culture

Human GBM cell lines U87 and A172 were purchased from American Type Culture Collection (ATCC, Manassas, Virginia, United States), which were authenticated with a short tandem repeat. Cells were set on a humidified incubator with 5% $CO_2$ at 37°C, as well as cultured in Dulbecco's Modified Eagle's Medium (GIBCO, Billings, MT, United States) added with 10% fetal bovine serum (GIBCO).

### Small Interfering RNA Transfection

The small interfering RNA of RCN1 (si-RCN1) sequences and the corresponding negative control were designed and purchased from RiboBio (Guangzhou, China). For transient silencing, A172 and U87 GBM cell lines were transfected with negative control or si-RCN1 by Lipofectamine™ 3,000 Reagent (Invitrogen, United States) according to the manufacturer's instruction. After 48 h, cells were harvested for quantitative real-time-polymerase chain reaction (qRT-PCR) analysis. Target sequences for transient silencing were as follows: si-RCN1-1: GAAGCTAACTAAAGA GGAA; si-RCN1-2: CCAGGCATCTGGTATATGA; negative control siRNA was obtained from RiboBio (Guangzhou, China).

### qRT-PCR

By using ReverTra Ace® qPCR RT Master Mix with gDNA Remover (TOYOBO, Shanghai, China), total RNA was extracted and then was used to synthesize the first complementary DNA (cDNA) strand according to the manufacturer's protocol. The qRT-PCR reaction was carried out to estimate the RNA levels and *ACTIN* was used as the internal reference. The primers used for qRT-PCR were as follows: RCN1 forward 5′-AAGGAGAGGCTAGGGAAGATT-3′ and reverse 5′-ATCCAGGTTTTCAGCTCCTCA-3'; ACTIN forward 5′- CACCATTGGCAATGAGCGGTTC-3′ and reverse 5′- AGG TCTTTGCGGATGTCCACGT -3'. The relative normalized expression of the target genes was compared with that of ACTIN, and the mRNA expression of each gene was calculated with the $2^{-\Delta\Delta Ct}$ method (Liang et al., 2021).

### Cell Invasion Assays

Cell invasion was measured through wound healing and transwell migration assays following the manufacturers' instructions. In brief, cells were plated in 6-well plates and cultured at 37°C with 5% $CO_2$. After 24 h, at which cells were reached on 80% confluence, we used a sterile 10 μL disposable serological pipette to make a straight-line scratch, and then cells were harvested after 48 h. Images of the scratch width were taken using an inverted microscope (Olympus IX73 Inverted Microscope, Olympus, Beijing China) at 0 and 48 h after the scratch, and then calculated by ImageJ software (version 1.52, National Institutes of Health, United States).

As for the transwell migration assay, it was performed using the Boyden chamber with a gelatin-coated polycarbonate filter with an 8-μm pore size (Neuro Probe, Gaithersburg, MD, United States). Cells were added to the upper chamber in 24-well plates at a density of $5.0 \times 10^4$ cells per well, and the lower chamber was filled with 800 μL 10% FBS for 24 h. Cells in the transwell chamber were fixed with 4% paraformaldehyde for 15 min, stained with 0.1% crystal violet for 30 min, and then observed by a BDS500 Inverted Biological Microscope (Chongqing Optec Instrument Co., Ltd., China).

### Statistical Analysis

All the data were presented as mean ± standard deviation and the statistical analyses were performed by IBM SPSS Statistics 25.0 software (International Business Machines Corporation, United States). The student's t-test was performed to evaluate the significant difference between the two groups.

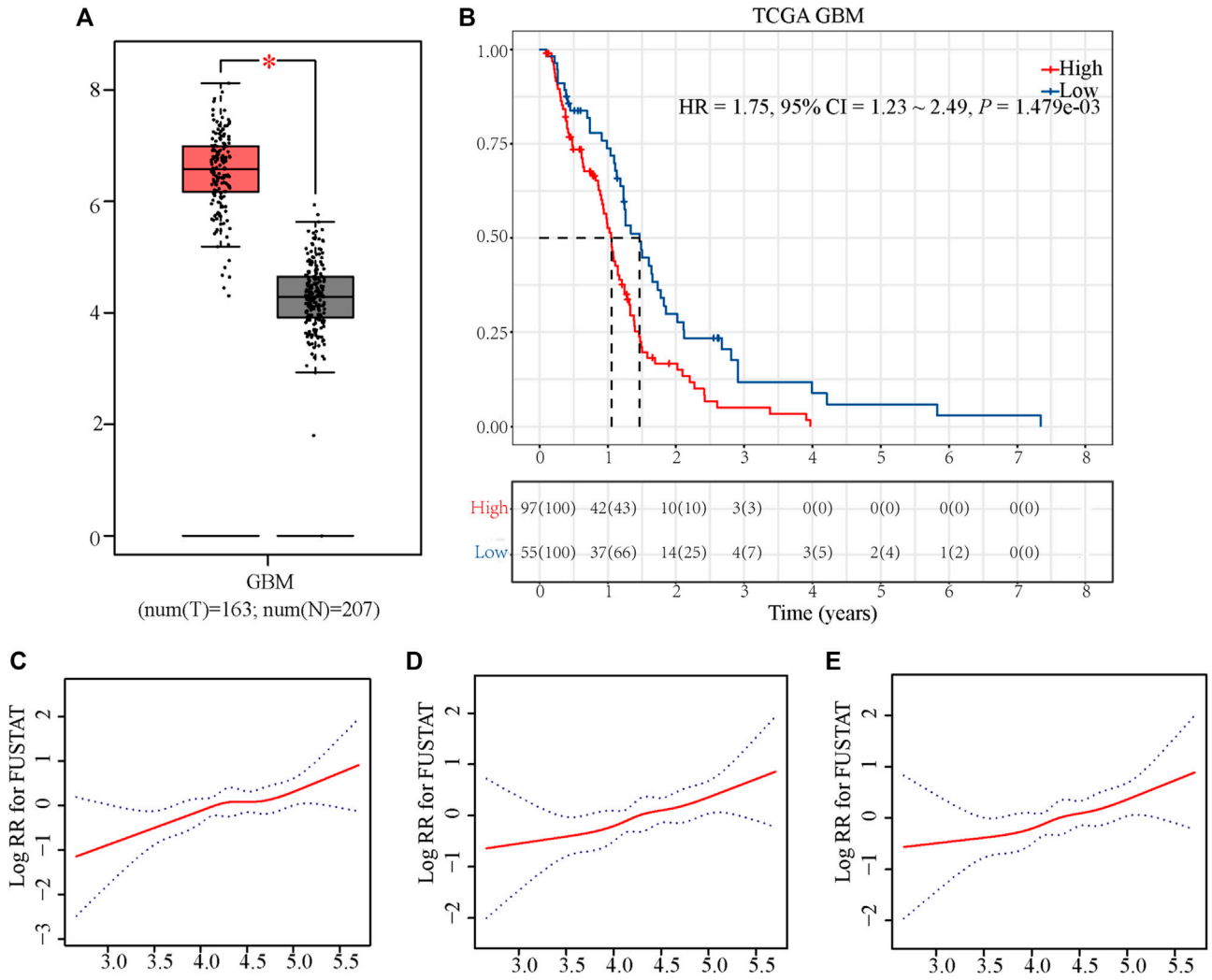



**FIGURE 2 |** Exploration of the association between RCN1 and OS of patients with GBM in TCGA. **(A)** Boxplot. **(B)** Kaplan-Meier curve with log-rank test in TCGA. **(C)** Association between the variation of RCN1 expression and OS in non-adjusted. **(D)** Association between the variation of RCN1 expression and OS in adjusted I model. **(E)** Association between the variation of RCN1 expression and OS in adjusted II model.

# RESULTS

## Patients Characteristics

A total of 361 GBM samples (153 patients from TCGA as the training cohort and 208 patients from CGGA as the validation cohort) were obtained in our study, as shown in **Table 1**.

## Reticulocalbin-1 Was Elevated in Glioblastoma Multiforme and May Act as an Oncogene

As shown in **Figure 2A**, we found that the expression of RCN1 was higher in the GBM samples ($n$ = 163) than in the normal brain tissues ($n$ = 207), by using the GEPIA2 tool. In the TCGA database, as the near-linear correlation between the variation of RCN1 expression and OS revealed through LOESS (**Figure 2C**), RCN1 expression was considered as a single continuous variation for further analysis. A total of 153 samples were clustered into the high- ($n$ = 97) or low-RCN1 group ($n$ = 56) by the optimal cut-off value (4.144). Patients with higher RCN1 expression had a worse OS than those with a low one in GBM ($p$ = 0.001) (**Figure 2B**).

## Reticulocalbin-1 is an Independent Prognostic Signature for Glioblastoma Multiforme

Considering the interference of the confounding factors, identifying and then adjusting for potential confounding factors was conducted. We firstly found that RCN1 may be an independent prognostic signature when compared with other signatures (**Supplementary Figure S1**). In the training cohort, we then identified variants (age and radiotherapy) to be adjusted and

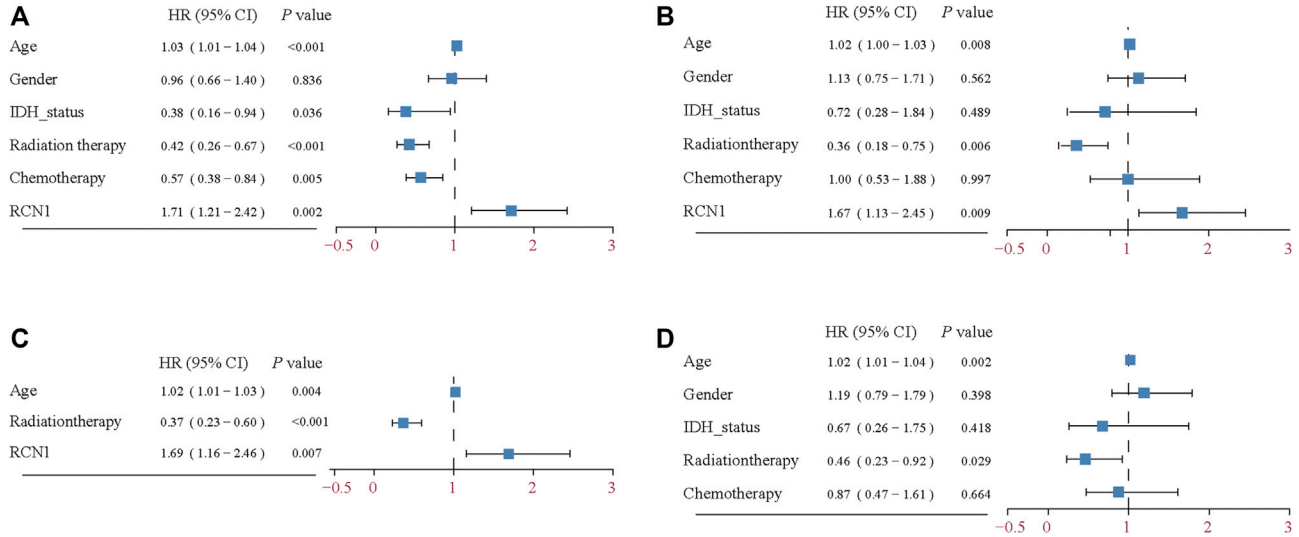



**FIGURE 3 |** Forest plot of the Cox regression analysis. **(A)** Univariable Cox regression analysis for the training cohort. **(B)** Multivariable Cox regression analysis for the training cohort. **(C)** Multivariable Cox regression analysis for adjusted I model in the training cohort. **(D)** Multivariable Cox regression analysis for clinical features in training cohort.

**TABLE 2 |** Cox proportional hazards regression models of various prognostic parameters in patients with GBM in TCGA.

| Characteristic | Non-adjusted model | | Adjusted I model | | Adjusted II model | |
|---|---|---|---|---|---|---|
| | HR (95% CI) | *p* value | HR (95% CI) | *p* value | HR (95% CI) | *p* value |
| Age | 1.03 (1.01–1.04) | <0.001 | 1.02 (1.01–1.03) | 0.004 | 1.02 (1.0–1.03) | 0.008 |
| Gender | 0.96 (0.66–1.4) | 0.836 | — | — | 1.13 (0.75–1.71) | 0.562 |
| IDH status | 0.38 (0.16–0.94) | 0.036 | — | — | 0.72 (0.28–1.84) | 0.489 |
| Radiotherapy | 0.42 (0.26–0.67) | <0.001 | 0.37 (0.23–0.6) | <0.001 | 0.36 (0.18–0.75) | 0.006 |
| Chemotherapy | 0.57 (0.38–0.84) | 0.005 | — | — | 1.0 (0.53–1.88) | 0.997 |
| RCN1 | 1.71 (1.21–2.42) | 0.002 | 1.69 (1.16–2.46) | 0.007 | 1.67 (1.13–2.45) | 0.009 |

*Adjusted I model: Factors including age, radiotherapy for adjustment, Adjusted II model: Factors including age, gender, IDH, radiotherapy and chemotherapy for adjustment. HR: hazard ratio, CI: confidence interval.*

then enrolled these clinical features into the adjusted I model. The adjustments for age, gender, IDH status, chemotherapy, and radiotherapy were included in the adjusted II model. Both non-adjusted and two adjusted models were analyzed by using the Cox regression analysis to further investigate whether RCN1 could estimate OS independently (**Figure 3**). As shown in **Table 2**, in the non-adjusted model, prognosis was correlated with age (HR = 1.03, 95% CI 1.01~1.04, $p < 0.001$), radiotherapy (HR = 0.42, 95% CI: 0.26~0.67, $p < 0.001$), IDH status (HR = 0.38, 95% CI: 0.16~0.94, $p = 0.036$), chemotherapy (HR = 0.57, 95% CI 0.38~0.84, $p = 0.005$), and RCN1 expression (HR = 1.71, 95% CI 1.21~2.42, $p = 0.002$) in the training cohort. In adjust I model, after adjusting for confounding factors (age and radiotherapy), RCN1 was still associated with OS (HR = 1.69, 95% CI 1.16~2.46, $p = 0.007$) (**Table 2**). Furthermore, after adjusting for five predominant clinical features (age, gender, IDH status, chemotherapy, and radiotherapy), RCN1 independently predicted prognosis in the training cohort (HR = 1.67, 95% CI 1.13~2.45, $p = 0.009$) (**Table 2**). In the same way, we also found the near-linear correlation between the variation of RCN1 expression and OS both in adjust I model (**Figure 2D**) and adjust II model (**Figure 2E**), thus, we enrolled RCN1 as a single continuous variation for further analysis. In addition, subgroup analysis showed that there were no statistical differences neither in the non-adjusted model nor adjusted I model, except for chemotherapy ($P_{interaction} = 0.0019$ in the non-adjusted model, $P_{interaction} = 0.0118$ in the adjusted I model, respectively) (**Table 3**), revealing that RCN1 might be an independent prognostic factor for OS in patients with GBM.

## Construction and Evaluation of Three Prognostic Models

The clinical features and RCN1 were enrolled to construct the prognostic models for GBM. We firstly built three prognostic models (model 1: five clinical variants, model 2: the expression of RCN1, and model 3: model 1 + model 2) and then evaluated them. Model 3 had a higher area of under curve (AUC), better C-index, and lower prediction error compared with model 2 and model 1 (**Figures 4A–C**). DCA showed that the net benefit of

**TABLE 3 |** Subgroup analysis of the associations between OS and RCN1 of patients with GBM in TCGA.

| Covariates | Total (*N* = 153) | RCN1 | | RCN1[a] | |
|---|---|---|---|---|---|
| | | HR (95% CI) | *P* for interaction | HR (95%CI) | *P* for interaction |
| Age | — | — | 0.2553 | — | — |
| <60 y | 72 | 2.02 (1.24–3.29)[c] | — | — | — |
| >=60 y | 81 | 1.36 (0.84–2.21) | — | — | — |
| Total | 153 | 1.66 (1.16–2.36)[c] | — | — | — |
| Gender | — | — | 0.8929 | — | 0.7652 |
| Male | 54 | 1.67 (0.98–2.83) | — | 1.83 (0.96–3.51) | — |
| Female | 99 | 1.75 (1.11–2.75)[b] | — | 1.62 (1.00–2.63) | — |
| Total | 153 | 1.71 (1.21–2.42)[c] | — | 1.69 (1.15–2.50)[c] | — |
| Radiotherapy | — | — | 0.6713 | — | — |
| No | 23 | 2.70 (0.50–14.62) | — | — | — |
| Yes | 130 | 1.86 (1.28–2.70)[c] | — | — | — |
| Total | 153 | 1.89 (1.31–2.73)[d] | — | — | — |
| Chemotherapy | — | — | 0.0019 | — | 0.0118 |
| No | 41 | 0.62 (0.29–1.33) | — | 0.81 (0.38–1.73) | — |
| Yes | 112 | 2.33 (1.53–3.53)[d] | — | 2.55 (1.63–3.98)[d] | — |
| Total | 153 | 1.66 (1.17–2.34)[c] | — | 1.91 (1.30–2.80)[c] | — |

[a]*Covariates were adjusted as in Adjusted I model (**Table 2**). HR (95% CI) were derived from Cox proportional hazards regression models.*
[b]P < *0.05.*
[c]< *0.01.*
[d]P < *0.001. HR: hazard ratio, CI: confidence interval.*

model 3 in 0.5 and 1 year is better than the other two models, but there was no significant difference in 1.5 years (**Figures 4D–F**). It was found that the calibration of model 3 was better than that of model 1 and model 2 in 0.5 and 1 year, while the calibration of the three models was poor in 1.5 years (model 2 was better than others) (**Figures 4G–I**). As for model improvement capability, when model 1 was considered as the reference, the NRI and IDI of model 3 were both positive, in which the NRI of 1.5 years was increased by 16.7% ($p$ = 0.064) and the IDI of 1.5 years was increased by 3.7% ($p$ = 0.034); on the contrary, the NRI and IDI of model 2 were both negative, where the IDI of 0.5 year was decreased by 11.4% ($p$ = 0.0016) meanwhile the NRI was decreased by 27.4% ($p$ = 0.040) (**Table 4**). From the above results, it can be determined that model 3 had good discrimination and calibration in the prediction of OS. Therefore, we developed a nomogram in accordance with model 3 to assess OS at 0.5-, 1.0-, and 1.5-years in the TCGA dataset, in which each signature was assigned points according to its risk contribution to OS (**Figure 5**).

## External Validation for Nomogram in Chinese Glioma Genome Atlas

To assess whether an RCN1 signature-based model can similarly play a prognostic value in different populations, a total of 208 GBM samples from the CGGA database as an external validation cohort were used to assess its prediction performance. According to the optimum cutoff (5.521), high- ($n$ = 147) and low- ($n$ = 61) RCN1 groups were determined. Consistent with the findings in the training cohort, the Kaplan-Meier curve revealed patients with high RCN1 represented a worse OS than those with low expression ($p$ = 0.0047) (**Figure 6A**). Moreover, the time-dependent ROC curves were performed and the AUCs of the 0.5-, 1.0-, and 1.5-year survival for the constructed nomogram in the training cohort were 0.737, 0.673, and 0.694, respectively, in the validation cohort (**Figure 6B**). The nomogram shared a C-index of more than 0.6 at different times (**Figure 6C**) and a relatively low prediction error (**Figure 6D**). Finally, the calibration curves for this nomogram in the validation cohort at 0.5 year was poor, while those at 1 year and 1.5 years were good (**Figures 6E–G**).

## Differential Abundances of Infiltrative Immune Cells

By using the CIBERSORTx algorithm, the relative proportions of 22 immune cells between two groups in GBM were obtained. In the TCGA dataset, there were three types of infiltrative immune cells with significant difference at different groups, whereas in CGGA there was fourteen, in which T cells CD4 memory resting, eosinophils, and macrophages M0 were differentially expressed in two data at the same time (**Figure 7**). With more detail as shown by bar plots in **Figure 7A**, in the training cohort, the infiltration level of eosinophils was significantly higher in the low-risk group, whereas the infiltration level of macrophages M0 and T cells CD4 memory resting was significantly higher in the high-risk group. In the validation cohort, the infiltration levels of B cells naive, dendritic cells resting, mast cells activated, macrophages M1, eosinophils, T cells CD4 memory resting, neutrophils, NK cells activated, and monocytes were significantly higher in the low-risk group, whereas the B cells memory, macrophages (M0, M2), T cells regulatory (Tregs), and plasma cells were significantly higher in the high-risk group (**Figure 7B**).

## Pathway Enrichment Analysis

The ssGSEA was used to identify signaling pathways RCN1-involved in GBM, and then demonstrate significant differences in the enrichment of MSigDB hallmark gene set in the TCGA and

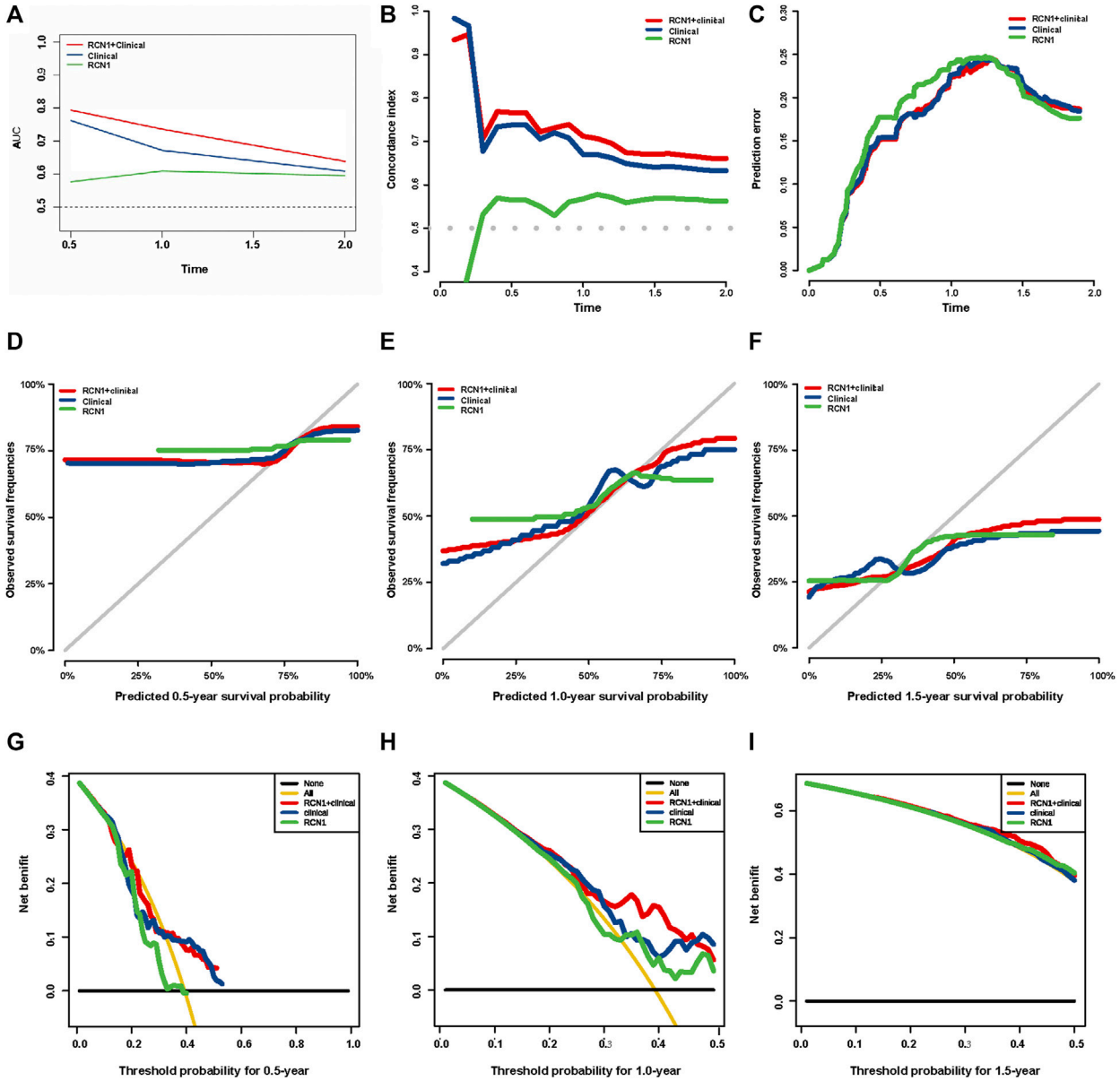



**FIGURE 4 |** Evaluation of three prognostic models in TCGA. **(A)** ROC curves of the three models at different times. **(B)** C-index curves of the three models at different times. **(C)** Prediction error probability curves of the three models at different times. **(D–F)** Calibration curves of the three models at different times. **(G–I)** DCA curves of the three models at different times.

**TABLE 4 |** The model improvement capability in three prognostic models.

| Year | Model 3 vs Model 1 | | | | Model 2 vs Model 1 | | | |
|---|---|---|---|---|---|---|---|---|
| | NRI | | IDI | | NRI | | IDI | |
| | Value | *p* value | Value | *p* value | Value | *p* value | Value | *p* value |
| 0.5 | 0.129 | 0.284 | 0.018 | 0.176 | −0.274 | 0.040 | −0.114 | 0.016 |
| 1.0 | 0.122 | 0.156 | 0.024 | 0.094 | −0.082 | 0.384 | −0.088 | 0.056 |
| 1.5 | 0.164 | 0.064 | 0.037 | 0.034 | −0.035 | 0.707 | −0.031 | 0.322 |

*Model 1 represents the prognostic model with age, gender, IDH status, chemotherapy, and radiotherapy; Model 2 only comprises the expression of RCN1; Model 3 integration of model 1 and 2, including all factors (age, gender, IDH status, chemotherapy, radiotherapy, and the expression of RCN.*

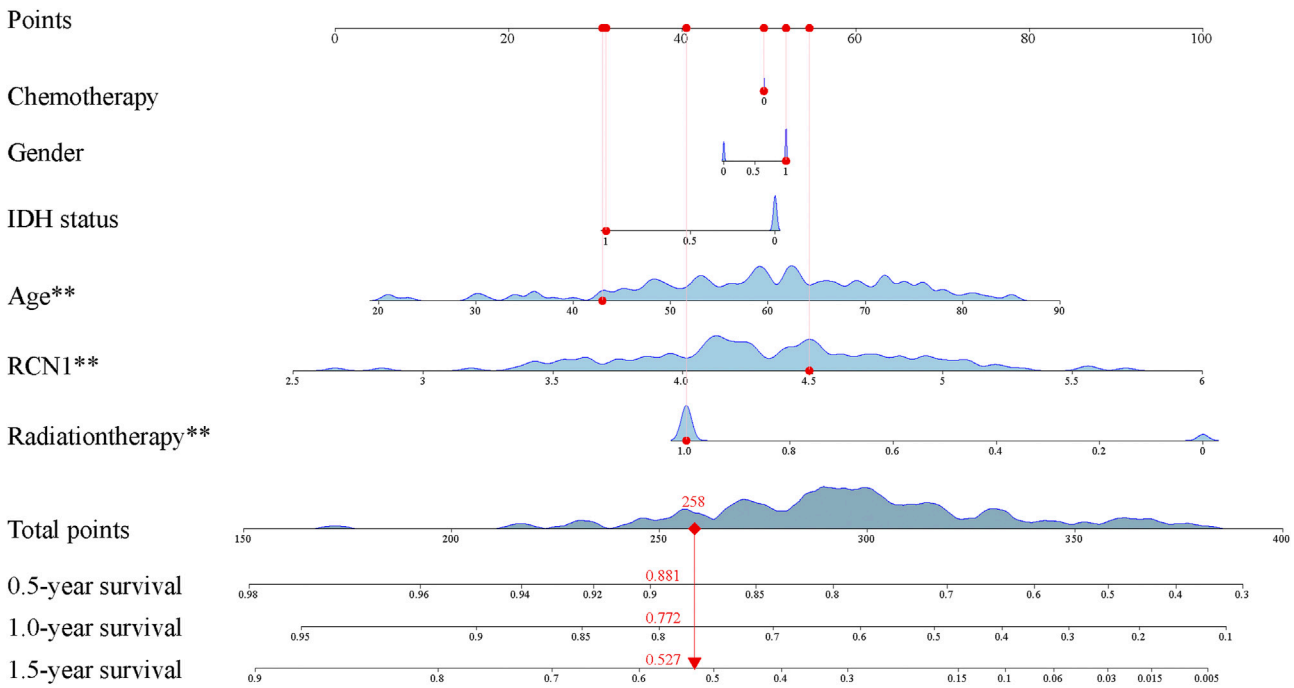



**FIGURE 5 |** Construction of prognostic nomogram.

CGGA databases, respectively. The results indicated that there was no significant pathway screened in the TCGA database (**Figure 8A**), whereas the KRAS-signaling-DN pathway was significantly involved in the low-RCN1 group and some pathways, including reactive oxygen species, MYC targets V2 and V1, apoptosis, and DNA repair pathways, in the high-RCN1 group in the CGGA database (**Figure 8B**).

## Knockdown of Reticulocalbin-1Inhibited Cell Invasion in Glioblastoma Multiforme

Finally, in order to elucidate the effect of RCN1on GBM cell invasion, we conducted a series of morphological and molecular biological experiments. The si-RCN1-1 and si-RCN1-2 could effectively reduce endogenous RCN1 mRNA expression by mRNA levels in both U87 and A172 cells (**Figures 9A,B**). Later, we confirmed that si-RCN1 could decrease cell invasion (**Figures 9C–F**).

# DISCUSSION

GBM is the most aggressive brain tumor. The prognosis of patients with GBM remains poor, although treatment strategies, including maximum surgical resection, radiation, and chemotherapy, have been conducted. The role and mechanism of biomarkers in GBM tumorigenesis are very important for the development of treatment of patients with GBM (Sasmita et al., 2018). Recently, ample molecular markers have been identified, which could provide new insight regarding GBM formation and progression. The discoveries and roles of molecular biomarkers, GBM-specific microRNAs, and GBM stem cells were delineated (Sasmita et al., 2018; Hassn Mesrati et al., 2020). There are some genetic mutation features, epigenetic modification, and some molecular alterations *in situ*. As highlighted by Alireza Mansouri and his colleagues, the methylation of *MGMT* promoter has been identified to provide better outcome prediction when GBM patients receive temozolomide chemotherapy (Mansouri et al., 2019). Meanwhile, the amplification of *CCND2* suggests better clinical outcomes in *IDH*-mutant patients while the elevated total copy number variation, co-amplification of *CDK4/MDM2*, amplification of *PDGFRA*, and *CDKN2A* show worse association (Zheng et al., 2013; Sasmita et al., 2018; Mirchia et al., 2019; Louis et al., 2021). Large-scale studies further demonstrated that *IDH*-mutant GBMs tend to have a higher concentration of 2-hydroxyglutarate (an oncometabolite produced as a result of the *IDH1/2* mutation), which may promote infiltration via upregulation of HIF-1α and VEGF(K. Mirchia and Richardson, 2020; Sasmita et al., 2018). For IDH-wide type GBM patients, the mutation of *BRAF* may benefit them while the co-alteration of *EGFR/PTEN/CDKN2A* and mutation of *PIK3CA* and *H3K27M* correlates with worse clinical outcomes (Mirchia and Richardson, 2020; Umehara et al., 2019). Circulating biomarkers have also been developed in GBM because of their non-invasive potential (Jelski and Mroczko, 2021; Kefayat et al., 2021), but the sensitivity and specificity are still problems to be solved (Müller Bark et al., 2020; Raza et al., 2020; Jones et al., 2021). Additionally, there are

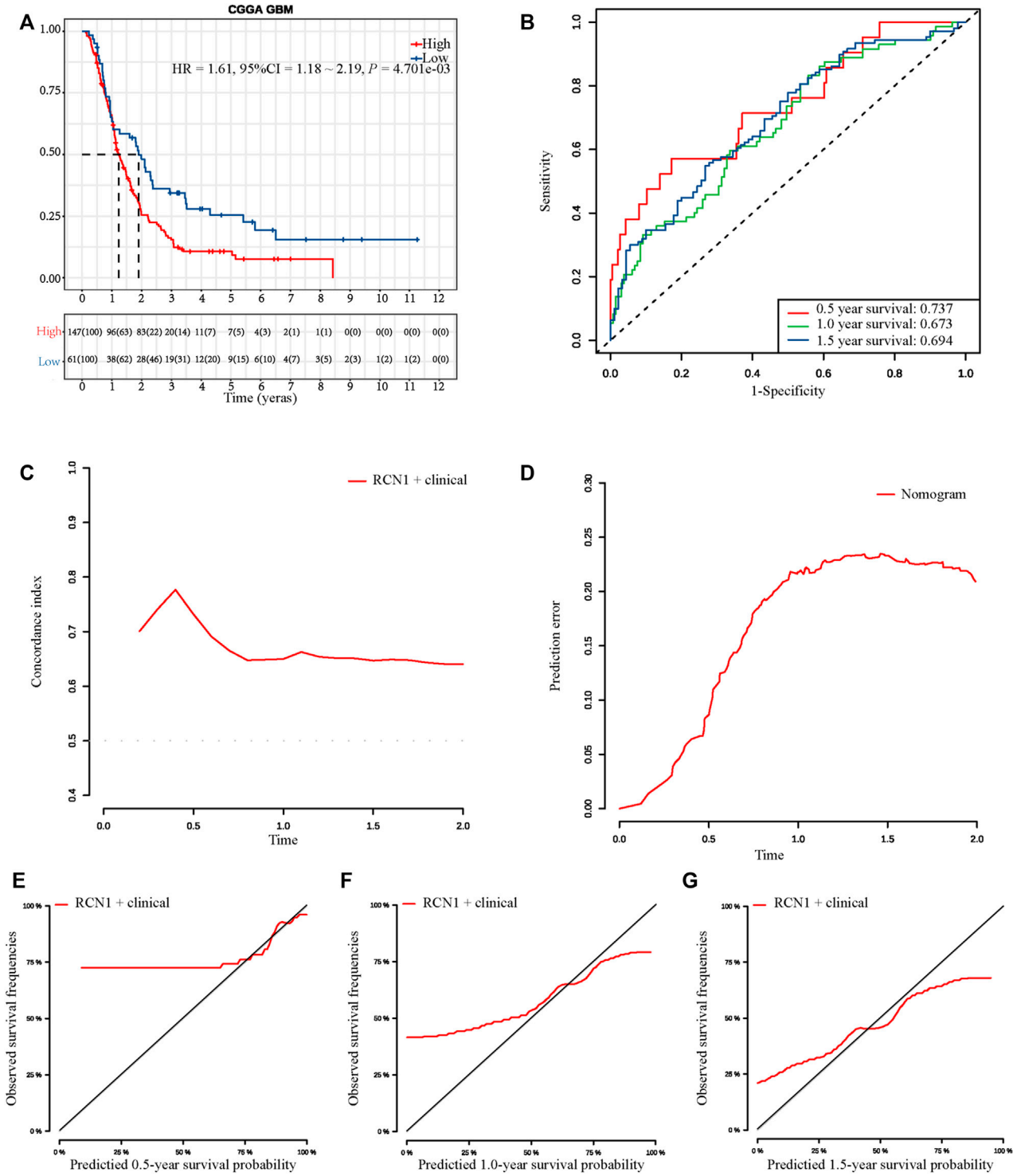



**FIGURE 6 |** External validation of the prognostic signature in CGGA. **(A)** Kaplan-Meier curve with log-rank test in CGGA. **(B)** ROC curves of the three models at different times. **(C)** C-index curves of the three models at different times. **(D)** Prediction error probability curves of the three models at different times. **(E–G)** Calibration curves of the three models at different times.

still no clinically validated circulating biomarkers for GBM patients because of the limitation in blood-brain-barrier, low concentration, and their short half-life (Müller Bark et al., 2020). It could not even be proven the release of potential biomarkers were exclusive from tumor cells or not. According to a systematic review, about 133 distinct biomarkers were identified from 1853

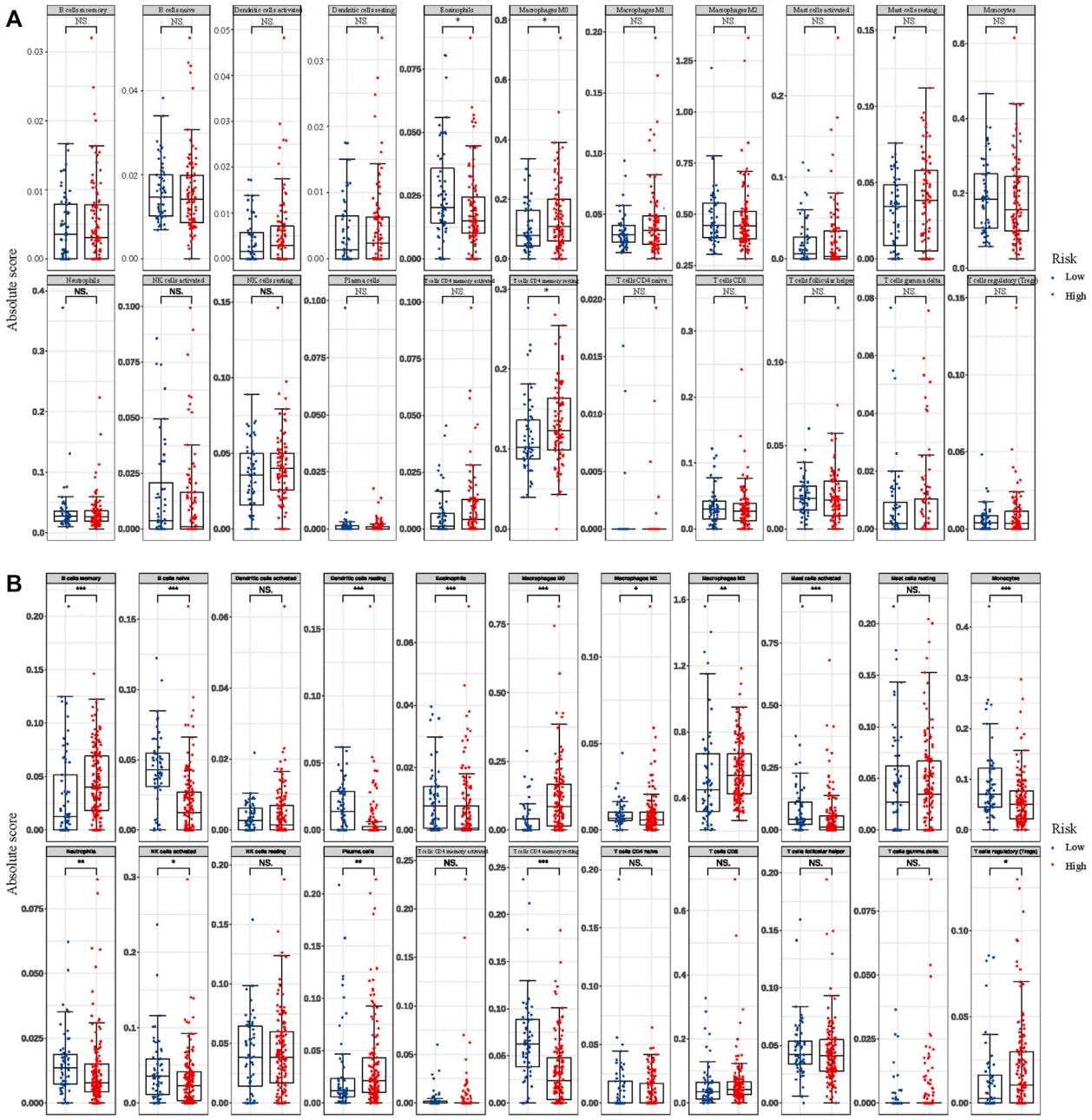



**FIGURE 7 |** Differential abundances of infiltrative immune cells between high- and low-RCN1 groups. **(A)** TCGA. **(B)** CGGA.

patients and evaluated based on level of evidence (IA~IVD) using an adapted framework from the National Comprehensive Cancer Network guideline (Raza et al., 2020). Nevertheless, few can reach level IA and further refinements are needed. Therefore, it is urgent to identify new biomarkers that can illustrate the in-depth molecular mechanism, enhance the diagnosis, respond to treatment, and provide prognostic prediction of GBM. Here, we identified the potential of RCN in GBM. RCN1, as a calcium-binding protein located in the lumen of the endoplasmic reticulum, containing six conserved regions with similarity to a high-affinity calcium-binding motif, the EF-hand. Recently, several studies have revealed that RCN1 acts as an oncogene involved in tumor progression. In our study, we found that the expression of RCN1 is higher in GBM samples than normal brain tissues by utilizing the GEPIA2 online tool, which is consistent with previous findings that high RCN1 expression is present in various malignancies (Amatschek et al., 2004). Recently, RCN1 has been demonstrated to be a prognostic marker in various

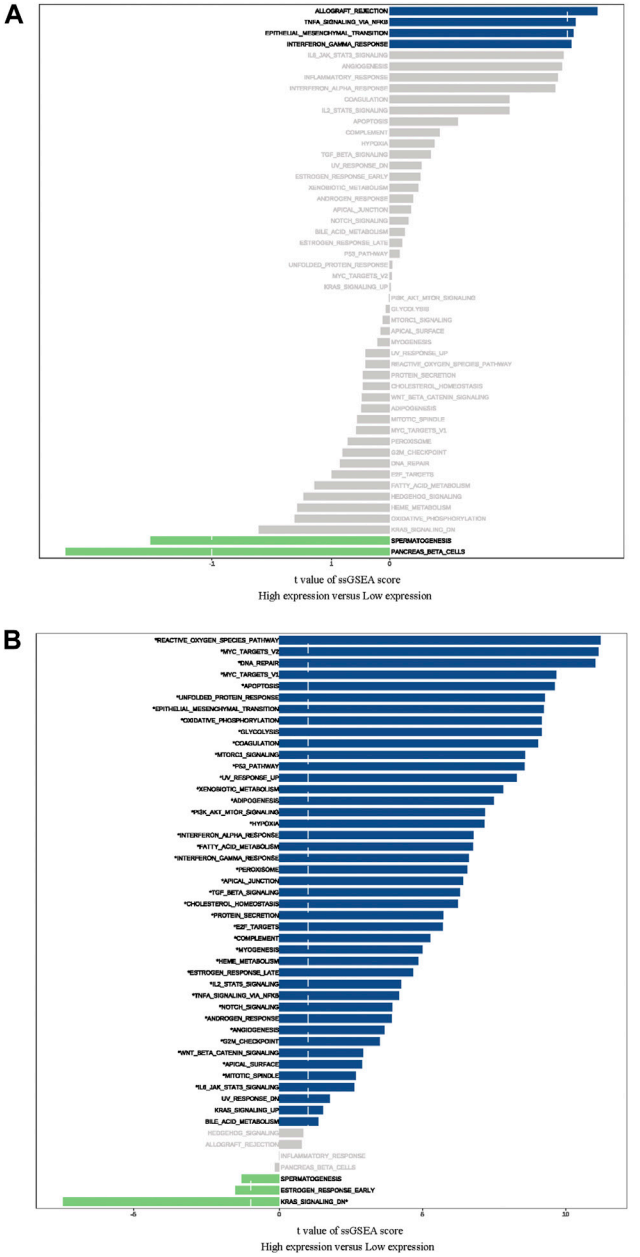



**FIGURE 8 |** Difference of Hallmark gene set in high- and low-RCN1 groups. **(A)** TCGA. **(B)** CGGA.

cancers, such as non-small cell lung cancer (Chen et al., 2019) and renal cell carcinoma (Giribaldi et al., 2013). However, the function and the underlying mechanism of RCN1 involved in GBM remains a vague notion. Based on the observation in previous studies (Chen et al., 2019; Giribaldi et al., 2013; Liu et al., 2018), we hypothesized that RCN1 could be a promising prognostic marker or as a response predictor for targeted therapy in GBM. Interestingly, we confirmed our hypothesis through TCGA and validated them in CGGA and experiments. At the same time, the mechanisms behind RCN1 were also explored and the identified pathways were consistent with previous studies (Fukasawa et al., 2021; Pucci et al., 2021; Sighel et al., 2021; Xue et al., 2021).

The convenient access to the public database allows us for the application of large-scale gene expression profiling and database mining for potential correlation between genes and overall survival of a variety of malignancies including GBM (Aldape et al., 2015; Zhou et al., 2021). In the present study, RCN1 was identified as an independent factor for the prognosis of GBM patients in TCGA, according to the results of univariate and multivariate analysis, with or without adjustment for confounders on the survival difference as described in a previous study (Erturk and Tas, 2017). By identifying the confounding factors, we further revealed the independent prognostic role of RCN1 in GBM patients. Since several clinical characteristics have been identified to influence the outcomes of GBM patients, interaction may exist (Brankovic et al., 2019) and it is plausible to perform subgroup analysis to estimate the interaction between RCN1 and selective clinical characteristics (Gönen, 2003). Here we innovatively identified confounding factors to increase the reliability of the results. Discrimination and calibration are the most commonly used indicators in evaluating prediction models. However, a systematic review found that 63% of the studies reported the discrimination information of prediction models, but only 36% of the studies reported the calibration information (Wessler et al., 2015); we reported both discrimination and calibration.

Recently, endoplasmic reticulum (ER) stress has been reported to have a considerable impact on cell growth, proliferation, metastasis, invasion, angiogenesis, and chemoradiotherapy resistance in various cancers. Accumulated studies have demonstrated that ER stress governs multiple pro-tumoural attributes in the cancer cell mainly via reprogramming the function of immune cells. The tumor microenvironment can be shaped because of the functional impact of ER stress responses in endothelial cells, cancer-associated fibroblasts, and other stromal cells during cancer progression (Chen and Cubillos-Ruiz, 2021). ER stress and downstream autophagy in the regulation of cell fate may function in temozolomide treatment and they have the potential to be therapeutic targets in GBM (He et al., 2019). Besides, it was revealed that the ER stress-related genes-based risk model can serve as a prognostic factor to predict the outcome for patients and be correlated with immune and inflammation responses in glioma (Zhang et al., 2021). Among these genes, RCN1, as an ER-resident calcium-binding protein, is verified as one of ER stress-related genes, of which the depletion causes the ER stress-induced cells' apoptosis in various cancers (Huang et al., 2020; Liu et al., 2018; Xu et al., 2017). In the early study, it was demonstrated that RCN1 was identified as a genuine phagocytosis ligand to stimulate microglial phagocytosis of apoptotic neurons that were subsequently targeted by phagosomes (Ding et al., 2015). It was also observed that RCN1 interacts with SEC63p to activate protein translocation and quality control pathways in the ER (Honoré, 2009). As mentioned above, RCN1 proteins

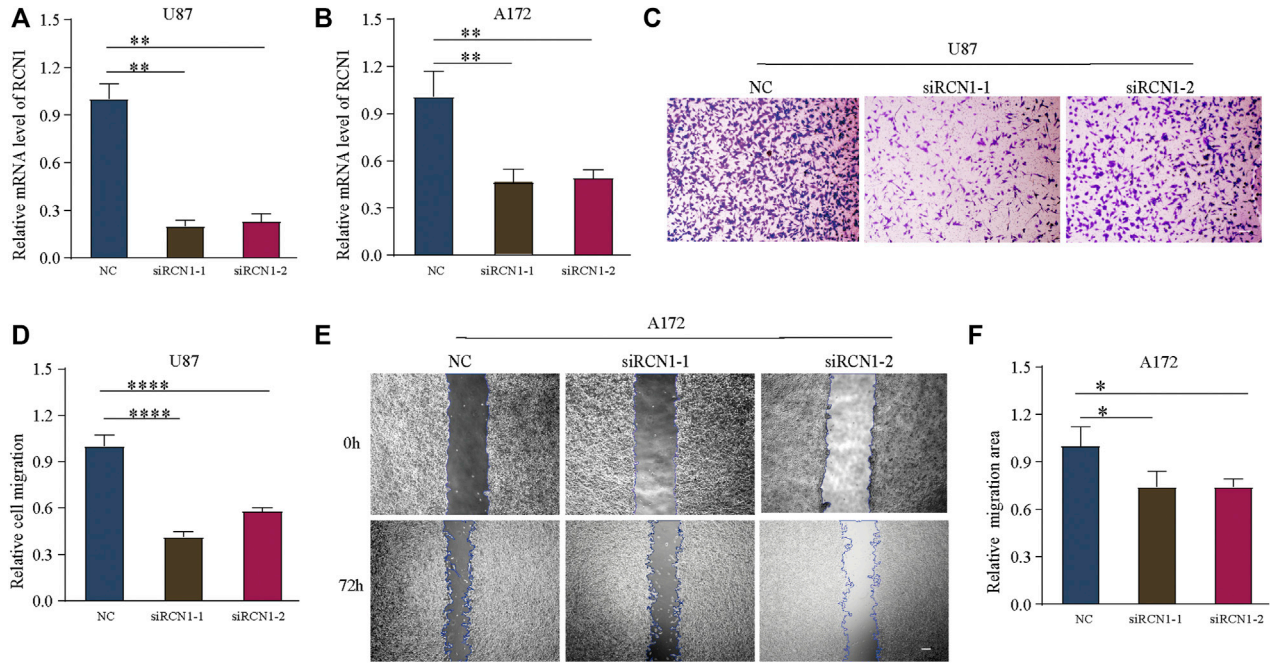



**FIGURE 9 |** Down-expression of RCN1 inhibits cell migration. **(A,B)** The mRNA of RCN1 knockdown on U87 and A172 cells after transfection. **(C,D)** Transwell assay. **(E,F)** Would healing assay.

can not only have intrinsic roles in tumor development but also serve as a regulator involved in immune-related activities. Considering the powerful role of tumor microenvironment cells contributing significantly to prognosis, we further investigated the immune cell infiltration between the high- and low-risk groups in the TCGA and CGGA datasets, respectively. It was revealed that the infiltration levels of T cells CD4 memory resting, eosinophils, and macrophages M0 were differentially changed in both TCGA and CGGA cohorts. Furthermore, we found that macrophages M0 was significantly higher in the high-risk group both in the TCGA and CGGA datasets, representing a certain relationship between them.

Furthermore, our results also suggested that high RCN1 may have a certain biological function in GBM. Therefore, to further investigate the role RCN1 played in GBM, we performed *in vitro* experiments and verified that the downregulation of RCN1 inhibited the cell invasion in GBM cell lines. The results of functional experiments are consistent with the above mentioned in our study, and further confirmed the critical role of RCN1 in GBM.

Several limitations should be noted in our study. Firstly, although we selected clinical samples from two different databases for mutual validation, verification in a larger sample is needed in the future. In addition, the variables involved in our prognostic model are easy to obtain, which is undoubtedly very convenient for clinical application, but it also needs to be further confirmed in clinical practice. Thirdly, although we verified the results of rigorous data mining, more animal or clinical exploration is still urgently needed. In addition, as for the absence of 1p19q characterization in TCGA, the status of 1p19q codeletion was not included for analyses. Finally, considering a lack of verification from clinical samples, the effectiveness of the RCN1-signature in GBM patients needs to be further verified by well-designed investigations in clinic. And how RCN1 regulates ER stress biology still needs to be further illustrated by experiments.

Overall, we found that high RCN1 has a poor OS for GBM patients, and confirmed RCN1 as an independent prognostic factor and developed a prognostic predictive model based on RCN1, which performed well in the prediction of OS for GBM patients.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

WL performed the validation experiment and wrote the manuscript. HC collected the data and analysed in bioinformatics ways. BL supported the bioinformatics analyses and wrote the manuscript. CO analysed the statistics results. JX, QY, and MZ conceptualize the study and revised the original manuscript. JX also designed the validation experiment *in vitro*.

## FUNDING

This study was supported by the GuangDong Natural Science Foundation for Distinguished Young Scholars (2019B151502010), National Natural Science Foundation of China (81870878, 82003386), GuangDong Natural Science Foundation (2018A030313092), and Sun Yat-sen University Training Project Foundation for Young Teachers (19ykpy183).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.736947/full#supplementary-material

**Supplementary Figure S1** | RCN1 is an independent prognostic signature compared with others. Different potential prognostic signatures were evaluated based on the data from TCGA and CCGA using the ROC curve.

## REFERENCES

Aldape, K., Zadeh, G., Mansouri, S., Reifenberger, G., and von Deimling, A. (2015). Glioblastoma: Pathology, Molecular Mechanisms and Markers. *Acta Neuropathol.* 129 (6), 829–848. doi:10.1007/s00401-015-1432-1

Ali, H. R., Chlon, L., Pharoah, P. D. P., Markowetz, F., and Caldas, C. (2016). Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *Plos Med.* 13 (12), e1002194. doi:10.1371/journal.pmed.1002194

Amatschek, S., Koenig, U., Auer, H., Steinlein, P., Pacher, M., Gruenfelder, A., et al. (2004). Tissue-Wide Expression Profiling Using cDNA Subtraction and Microarrays to Identify Tumor-Specific Genes. *Cancer Res.* 64 (3), 844–856. doi:10.1158/0008-5472.can-03-2361

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA Interference Reveals that Oncogenic KRAS-Driven Cancers Require TBK1. *Nature* 462 (7269), 108–112. doi:10.1038/nature08460

Blanche, P., Dartigues, J.-F., and Jacqmin-Gadda, H. (2013). Estimating and Comparing Time-Dependent Areas under Receiver Operating Characteristic Curves for Censored Event Times with Competing Risks. *Statist. Med.* 32 (30), 5381–5397. doi:10.1002/sim.5958

Brankovic, M., Kardys, I., Steyerberg, E. W., Lemeshow, S., Markovic, M., Rizopoulos, D., et al. (2019). Understanding of Interaction (Subgroup) Analysis in Clinical Trials. *Eur. J. Clin. Invest.* 49 (8), e13145. doi:10.1111/eci.13145

Chen, X., and Cubillos-Ruiz, J. R. (2021). Endoplasmic Reticulum Stress Signals in the Tumour and its Microenvironment. *Nat. Rev. Cancer* 21 (2), 71–88. doi:10.1038/s41568-020-00312-2

Chen, X., Shao, W., Huang, H., Feng, X., Yao, S., and Ke, H. (2019). Overexpression of RCN1 Correlates with Poor Prognosis and Progression in Non-Small Cell Lung Cancer. *Hum. Pathol.* 83, 140–148. doi:10.1016/j.humpath.2018.08.014

Ding, Y., Caberoy, N. B., Guo, F., LeBlanc, M. E., Zhang, C., Wang, W., et al. (2015). Reticulocalbin-1 Facilitates Microglial Phagocytosis. *PLoS One* 10 (5), e0126993. doi:10.1371/journal.pone.0126993

Erturk, K., and Tas, F. (2017). Effect of Biology on the Outcome of Female Melanoma Patients. *Mol. Clin. Onc* 7 (6), 1093–1100. doi:10.3892/mco.2017.1446

Fukasawa, K., Kadota, T., Horie, T., Tokumura, K., Terada, R., Kitaguchi, Y., et al. (2021). CDK8 Maintains Stemness and Tumorigenicity of Glioma Stem Cells by Regulating the C-MYC Pathway. *Oncogene* 40 (15), 2803–2815. doi:10.1038/s41388-021-01745-1

Gilbert, M. R., Wang, M., Aldape, K. D., Stupp, R., Hegi, M. E., Jaeckle, K. A., et al. (2013). Dose-Dense Temozolomide for Newly Diagnosed Glioblastoma: A Randomized Phase III Clinical Trial. *J. Clin. Oncol.* 31 (32), 4085–4091. doi:10.1200/jco.2013.49.6968

Giribaldi, G., Barbero, G., Mandili, G., Daniele, L., Khadjavi, A., Notarpietro, A., et al. (2013). Proteomic Identification of Reticulocalbin 1 as Potential Tumor Marker in Renal Cell Carcinoma. *J. Proteomics* 91, 385–392. doi:10.1016/j.jprot.2013.07.018

Gönen, M. (2003). Planning for Subgroup Analysis: A Case Study of Treatment-Marker Interaction in Metastatic Colorectal Cancer. *Controlled Clin. Trials* 24 (4), 355–363. doi:10.1016/s0197-2456(03)00006-0

Grzeskowiak, R., Witt, H., Drungowski, M., Thermann, R., Hennig, S., Perrot, A., et al. (2003). Expression Profiling of Human Idiopathic Dilated Cardiomyopathy. *Cardiovasc. Res.* 59 (2), 400–411. doi:10.1016/s0008-6363(03)00426-7

GTEx Consortium (2015). Human Genomics. The Genotype-Tissue Expression (GTEx) Pilot Analysis: Multitissue Gene Regulation in Humans. *Science* 348 (6235), 648–660. doi:10.1126/science.1262110

Hainfellner, J., Louis, D. N., Perry, A., and Wesseling, P. (2014). Letter in Response to David N. Louiset Al, International Society of Neuropathology-Haarlem Consensus Guidelines for Nervous System Tumor Classification and Grading, Brain Pathology, Doi: 10.1111/bpa.12171. *Brain Pathol.* 24 (6), 671–672. doi:10.1111/bpa.12187

Harrell, F. E., Jr., Lee, K. L., and Mark, D. B. (1996). Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors. *Stat. Med.* 15 (4), 361–387. doi:10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4

Hartmann, C., Hentschel, B., Wick, W., Capper, D., Felsberg, J., Simon, M., et al. (2010). Patients with IDH1 Wild Type Anaplastic Astrocytomas Exhibit Worse Prognosis Than IDH1-Mutated Glioblastomas, and IDH1 Mutation Status Accounts for the Unfavorable Prognostic Effect of Higher Age: Implications for Classification of Gliomas. *Acta Neuropathol.* 120 (6), 707–718. doi:10.1007/s00401-010-0781-z

Hassn Mesrati, M., Behrooz, A. B., Y. Abuhamad, A., and Syahir, A. (2020). Understanding Glioblastoma Biomarkers: Knocking a Mountain with a Hammer. *Cells* 9 (5), 1236. doi:10.3390/cells9051236

He, Y., Su, J., Lan, B., Gao, Y., and Zhao, J. (2019). Targeting Off-Target Effects: Endoplasmic Reticulum Stress and Autophagy as Effective Strategies to Enhance Temozolomide Treatment. *OncoTargets Ther.* 12, 1857–1865. doi:10.2147/OTT.S194770

Heagerty, P. J., Lumley, T., and Pepe, M. S. (2000). Time-Dependent ROC Curves for Censored Survival Data and a Diagnostic Marker. *Biometrics* 56 (2), 337–344. doi:10.1111/j.0006-341x.2000.00337.x

Honoré, B. (2009). The Rapidly Expanding CREC Protein Family: Members, Localization, Function, and Role in Disease. *BioEssays* 31 (3), 262–277. doi:10.1002/bies.200800186

Huang, Z.-H., Qiao, J., Feng, Y.-Y., Qiu, M.-T., Cheng, T., Wang, J., et al. (2020). Reticulocalbin-1 Knockdown Increases the Sensitivity of Cells to Adriamycin in Nasopharyngeal Carcinoma and Promotes Endoplasmic Reticulum Stress-Induced Cell Apoptosis. *Cell Cycle* 19 (13), 1576–1589. doi:10.1080/15384101.2020.1733750

Jayamanne, D., Wheeler, H., Cook, R., Teo, C., Brazier, D., Schembri, G., et al. (2018). Survival Improvements with Adjuvant Therapy in Patients with Glioblastoma. *ANZ J. Surg.* 88 (3), 196–201. doi:10.1111/ans.14153

Jelski, W., and Mroczko, B. (2021). Molecular and Circulating Biomarkers of Brain Tumors. *Int. J. Mol. Sci.* 22 (13), 7039. doi:10.3390/ijms22137039

Jones, J., Nguyen, H., Drummond, K., and Morokoff, A. (2021). Circulating Biomarkers for Glioma: A Review. *Neurosurgery* 88 (3), E221–E230. doi:10.1093/neuros/nyaa540

Kefayat, A., Amouheidari, A., Ghahremani, F., and Alirezaei, Z. (2021). Diagnostic and Prognostic Value of Stem Cell Factor Plasma Level in Glioblastoma Multiforme Patients. *Cancer Med.* 10 (15), 5154–5162. doi:10.1002/cam4.4073

Kernan, W. N., Viscoli, C. M., Brass, L. M., Broderick, J. P., Brott, T., Feldmann, E., et al. (2000). Phenylpropanolamine and the Risk of Hemorrhagic Stroke. *N. Engl. J. Med.* 343 (25), 1826–1832. doi:10.1056/nejm200012213432501

Kerr, K. F., Brown, M. D., Zhu, K., and Janes, H. (2016). Assessing the Clinical Impact of Risk Prediction Models with Decision Curves: Guidance for Correct Interpretation and Appropriate Use. *J. Clin. Oncol.* 34 (21), 2534–2540. doi:10.1200/JCO.2015.65.5654

Liang, B., Liang, Y., Li, R., Zhang, H., and Gu, N. (2021). Integrating Systematic Pharmacology-Based Strategy and Experimental Validation to Explore the Synergistic Pharmacological Mechanisms of Guanxin V in Treating

Ventricular Remodeling. *Bioorg. Chem.* 115, 105187. doi:10.1016/j.bioorg.2021.105187

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. *Cel Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004

Liu, X., Zhang, N., Wang, D., Zhu, D., Yuan, Q., Zhang, X., et al. (2018). Downregulation of Reticulocalbin-1 Differentially Facilitates Apoptosis and Necroptosis in Human Prostate Cancer Cells. *Cancer Sci.* 109 (4), 1147–1157. doi:10.1111/cas.13541

Liu, Z., Brattain, M. G., and Appert, H. (1997). Differential Display of Reticulocalbin in the Highly Invasive Cell Line, MDA-MB-435, Versus the Poorly Invasive Cell Line, MCF-7. *Biochem. Biophysical Res. Commun.* 231 (2), 283–289. doi:10.1006/bbrc.1997.6083

Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., et al. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a Summary. *Acta Neuropathol.* 131 (6), 803–820. doi:10.1007/s00401-016-1545-1

Louis, D. N., Perry, A., Wesseling, P., Brat, D. J., Cree, I. A., Figarella-Branger, D., et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: A Summary. *Neuro Oncol.* 23 (8), 1231–1251. doi:10.1093/neuonc/noab106

Lu, J.-J., Lu, D.-Z., Chen, Y.-F., Dong, Y.-T., Zhang, J.-R., Li, T., et al. (2015). Proteomic Analysis of Hepatocellular Carcinoma HepG2 Cells Treated with Platycodin D. *Chin. J. Nat. Medicines* 13 (9), 673–679. doi:10.1016/s1875-5364(15)30065-0

Mansouri, A., Hachem, L. D., Mansouri, S., Nassiri, F., Laperriere, N. J., Xia, D., et al. (2019). MGMT Promoter Methylation Status Testing to Guide Therapy for Glioblastoma: Refining the Approach Based on Emerging Evidence and Current Challenges. *Neuro Oncol.* 21 (2), 167–178. doi:10.1093/neuonc/noy132

Mirchia, K., and Richardson, T. E. (2020). Beyond IDH-Mutation: Emerging Molecular Diagnostic and Prognostic Features in Adult Diffuse Gliomas. *Cancers* 12 (7), 1817. doi:10.3390/cancers12071817

Mirchia, K., Sathe, A. A., Walker, J. M., Fudym, Y., Galbraith, K., Viapiano, M. S., et al. (2019). Total Copy Number Variation as a Prognostic Factor in Adult Astrocytoma Subtypes. *Acta Neuropathol. Commun.* 7 (1), 92. doi:10.1186/s40478-019-0746-y

Müller Bark, J., Kulasinghe, A., Chua, B., Day, B. W., and Punyadeera, C. (2020). Circulating Biomarkers in Patients with Glioblastoma. *Br. J. Cancer* 122 (3), 295–305. doi:10.1038/s41416-019-0603-6

Nakakido, M., Tamura, K., Chung, S., Ueda, K., Fujii, R., Kiyotani, K., et al. (2016). Phosphatidylinositol Glycan Anchor Biosynthesis, Class X Containing Complex Promotes Cancer Cell Proliferation through Suppression of EHD2 and ZIC1, Putative Tumor Suppressors. *Int. J. Oncol.* 49 (3), 868–876. doi:10.3892/ijo.2016.3607

Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., et al. (2019). Determining Cell Type Abundance and Expression from Bulk Tissues with Digital Cytometry. *Nat. Biotechnol.* 37 (7), 773–782. doi:10.1038/s41587-019-0114-2

Nimmrich, I., Erdmann, S., Melchers, U., Finke, U., Hentsch, S., Moyer, M. P., et al. (2000). Seven Genes that Are Differentially Transcribed in Colorectal Tumor Cell Lines. *Cancer Lett.* 160 (1), 37–43. doi:10.1016/s0304-3835(00)00553-x

Ogłuszka, M., Orzechowska, M., Jędroszka, D., Witas, P., and Bednarek, A. K. (2019). Evaluate Cutpoints: Adaptable Continuous Data Distribution System for Determining Survival in Kaplan-Meier Estimator. *Comput. Methods Programs Biomed.* 177, 133–139. doi:10.1016/j.cmpb.2019.05.023

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

Ostrom, Q. T., Gittleman, H., Stetson, L., Virk, S. M., and Barnholtz-Sloan, J. S. (2015). Epidemiology of Gliomas. *Cancer Treat. Res.* 163, 1–14. doi:10.1007/978-3-319-12048-5_1

Ozawa, M., and Muramatsu, T. (1993). Reticulocalbin, a Novel Endoplasmic Reticulum Resident Ca(2+)-Binding Protein with Multiple EF-Hand Motifs and a Carboxyl-Terminal HDEL Sequence. *J. Biol. Chem.* 268 (1), 699–705. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/8416973. doi:10.1016/s0021-9258(18)54208-3

Pencina, M. J., D' Agostino, R. B., D' Agostino., Vasan, R. S., Jr., and Vasan, R. S. (2008). Evaluating the Added Predictive Ability of a New Marker: From Area under the ROC Curve to Reclassification and beyond. *Statist. Med.* 27 (2), 157–172. discussion 207-112. doi:10.1002/sim.2929

Pucci, C., Marino, A., Şen, Ö., De Pasquale, D., Bartolucci, M., Iturrioz-Rodríguez, N., et al. (2021). Ultrasound-Responsive Nutlin-Loaded Nanoparticles for Combined Chemotherapy and Piezoelectric Treatment of Glioblastoma Cells. *Acta Biomater.* S1742-7061 (21), 00242–00247. doi:10.1016/j.actbio.2021.04.005

Raza, I. J., Tingate, C. A., Gkolia, P., Romero, L., Tee, J. W., and Hunn, M. K. (2020). Blood Biomarkers of Glioma in Response Assessment Including Pseudoprogression and Other Treatment Effects: A Systematic Review. *Front. Oncol.* 10, 1191. doi:10.3389/fonc.2020.01191

Rousson, V., and Zumbrunn, T. (2011). Decision Curve Analysis Revisited: Overall Net Benefit, Relationships to ROC Curve Analysis, and Application to Case-Control Studies. *BMC Med. Inform. Decis. Mak* 11, 45. doi:10.1186/1472-6947-11-45

Sasmita, A. O., Wong, Y. P., and Ling, A. P. K. (2018). Biomarkers and Therapeutic Advances in Glioblastoma Multiforme. *Asia-pac J. Clin. Oncol.* 14 (1), 40–51. doi:10.1111/ajco.12756

Sighel, D., Notarangelo, M., Aibara, S., Re, A., Ricci, G., Guida, M., et al. (2021). Inhibition of Mitochondrial Translation Suppresses Glioblastoma Stem Cell Growth. *Cel Rep.* 35 (4), 109024. doi:10.1016/j.celrep.2021.109024

Soffietti, R., Bertero, L., Pinessi, L., and Rudà, R. (2014). Pharmacologic Therapies for Malignant Glioma: A Guide for Clinicians. *CNS Drugs* 28 (12), 1127–1137. doi:10.1007/s40263-014-0215-x

Soria, J.-C., Felip, E., Cobo, M., Lu, S., Syrigos, K., Lee, K. H., et al. (2015). Afatinib versus Erlotinib as Second-Line Treatment of Patients with Advanced Squamous Cell Carcinoma of the Lung (LUX-Lung 8): An Open-Label Randomised Controlled Phase 3 Trial. *Lancet Oncol.* 16 (8), 897–907. doi:10.1016/s1470-2045(15)00006-6

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

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: An Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Res.* 47 (W1), W556–w560. doi:10.1093/nar/gkz430

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: A Web Server for Cancer and Normal Gene Expression Profiling and Interactive Analyses. *Nucleic Acids Res.* 45 (W1), W98–w102. doi:10.1093/nar/gkx247

Umehara, T., Arita, H., Yoshioka, E., Shofuda, T., Kanematsu, D., Kinoshita, M., et al. (2019). Distribution Differences in Prognostic Copy Number Alteration Profiles in IDH-Wild-Type Glioblastoma Cause Survival Discrepancies across Cohorts. *Acta Neuropathol. Commun.* 7 (1), 99. doi:10.1186/s40478-019-0749-8

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

Vickers, A. J., and Elkin, E. B. (2006). Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med. Decis. Making* 26 (6), 565–574. doi:10.1177/0272989x06295361

Wessler, B. S., Lai Yh, L., Kramer, W., Cangelosi, M., Raman, G., Lutz, J. S., et al. (2015). Clinical Prediction Models for Cardiovascular Disease. *Circ. Cardiovasc. Qual. Outcomes* 8 (4), 368–375. doi:10.1161/circoutcomes.115.001693

Wick, W., Meisner, C., Hentschel, B., Platten, M., Schilling, A., Wiestler, B., et al. (2013). Prognostic or Predictive Value of MGMT Promoter Methylation in Gliomas Depends on IDH1 Mutation. *Neurology* 81 (17), 1515–1522. doi:10.1212/WNL.0b013e3182a95680

Xu, S., Xu, Y., Chen, L., Fang, Q., Song, S., Chen, J., et al. (2017). RCN1 Suppresses ER Stress-Induced Apoptosis via Calcium Homeostasis and PERK-CHOP Signaling. *Oncogenesis* 6 (3), e304. doi:10.1038/oncsis.2017.6

Xue, Y., Fu, Y., Zhao, F., Gui, G., Li, Y., Rivero-Hinojosa, S., et al. (2021). Frondoside A Inhibits an MYC-Driven Medulloblastoma Model Derived from Human-Induced Pluripotent Stem Cells. *Mol. Cancer Ther.* 20, 1199–1209. doi:10.1158/1535-7163.MCT-20-0603

Zhang, Q., Guan, G., Cheng, P., Cheng, W., Yang, L., and Wu, A. (2021). Characterization of an Endoplasmic Reticulum Stress-related Signature to

Evaluate Immune Features and Predict Prognosis in Glioma. *J. Cel Mol Med* 25, 3870–3884. doi:10.1111/jcmm.16321

Zhang, Y., Ye, J., Chen, D., Zhao, X., Xiao, X., Tai, S., et al. (2006). Differential Expression Profiling between the Relative normal and Dystrophic Muscle Tissues from the Same LGMD Patient. *J. Transl Med.* 4, 53. doi:10.1186/1479-5876-4-53

Zheng, S., Fan, J., Yu, F., Feng, B., Lou, B., Zou, Q., et al. (2020). Viral Load Dynamics and Disease Severity in Patients Infected with SARS-CoV-2 in Zhejiang Province, China, January-March 2020: Retrospective Cohort Study. *BMJ* 369, m1443. doi:10.1136/bmj.m1443

Zheng, S., Fu, J., Vegesna, R., Mao, Y., Heathcock, L. E., Torres-Garcia, W., et al. (2013). A Survey of Intragenic Breakpoints in Glioblastoma Identifies a Distinct Subset Associated with Poor Survival. *Genes Develop.* 27 (13), 1462–1472. doi:10.1101/gad.213686.113

Zhou, J.-G., Liang, B., Jin, S.-H., Liao, H.-L., Du, G.-B., Cheng, L., et al. (2019). Development and Validation of an RNA-Seq-Based Prognostic Signature in Neuroblastoma. *Front. Oncol.* 9, 1361. doi:10.3389/fonc.2019.01361

Zhou, J.-G., Liang, B., Liu, J.-G., Jin, S.-H., He, S.-S., Frey, B., et al. (2021). Identification of 15 lncRNAs Signature for Predicting Survival Benefit of Advanced Melanoma Patients Treated with Anti-PD-1 Monotherapy. *Cells* 10 (5), 977. doi:10.3390/cells10050977

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

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

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 03 December 2021
doi: 10.3389/fmolb.2021.784876

# Mixed Pulmonary Adenocarcinoma and Atypical Carcinoid: A Report of Two Cases of a Non-codified Entity With Biological Profile

*Paola Parente[1], Antonio Rossi[2†], Angelo Sparaneo[3], Federico Pio Fabrizio[3], Antonella Centonza[2], Marco Taurchini[4], Tommaso Mazza[5], Maurizio Cassano[1], Giuseppe Miscio[1], Flavia Centra[3], Gian Maria Ferretti[4], Concetta Martina Di Micco[2], Paolo Graziano[1] and Lucia Anna Muscarella[3]**

[1]Unit of Pathology, Fondazione IRCCS Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy, [2]Unit of Oncology, Fondazione IRCCS Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy, [3]Laboratory of Oncology, Fondazione IRCCS Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy, [4]Surgical Thoracic Unit, Fondazione IRCCS Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy, [5]Bioinformatics Unit, Fondazione IRCCS Ospedale Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy

OPEN ACCESS

**Edited by:**
Umberto Malapelle,
University of Naples Federico II, Italy

**Reviewed by:**
Silvia Uccella,
University of Insubria, Italy
Linda Pattini,
Politecnico di Milano, Italy

***Correspondence:**
Lucia Anna Muscarella
l.muscarella@operapadrepio.it

**†Present address:**
Antonio Rossi,
Oncology Center of Excellence,
Therapeutic Science & Strategy Unit,
IQVIA, Milan, Italy

**Specialty section:**
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

**Received:** 28 September 2021
**Accepted:** 03 November 2021
**Published:** 03 December 2021

**Citation:**
Parente P, Rossi A, Sparaneo A, Fabrizio FP, Centonza A, Taurchini M, Mazza T, Cassano M, Miscio G, Centra F, Ferretti GM, Di Micco CM, Graziano P and Muscarella LA (2021) Mixed Pulmonary Adenocarcinoma and Atypical Carcinoid: A Report of Two Cases of a Non-codified Entity With Biological Profile. Front. Mol. Biosci. 8:784876. doi: 10.3389/fmolb.2021.784876

Pulmonary carcinoids combined with a non-neuroendocrine component have rarely been described, and this histological subtype is not included as a specific entity in the current World Health Organization classification of pulmonary neoplasms. Here, we described the molecular and histological features of two rare cases of mixed lung neoplasms, composed of atypical carcinoid and adenocarcinoma. The targeted next-generation sequencing analysis covering single nucleotide variations, copy number variations, and transcript fusions in a total of 161 cancer genes of the two different tumor components shows a similar molecular profile of shared and private gene mutations. These findings suggest their monoclonal origin from a transformed stem/progenitor tumor cell, which acquires a divergent differentiation during its development and progression and accumulates novel, specific mutations.

**Keywords: lung, atypical carcinoid, adenocarcinoma, mixed neoplasm, monoclonality, next-generation sequencing**

## INTRODUCTION

Combined or collision pulmonary tumors are primary lung neoplasms with two or more histologically distinct phenotypes (Olofson and Tafe, 2018). The World Health Organization (WHO) classification of the Tumors of the lung recognizes combined malignancies among neuroendocrine carcinomas (e.g., small cell lung cancer (SCLC) with large cell neuroendocrine lung cancer (LCNEC)] and combined SCLC or LCNEC with non-small cell lung cancer (NSCLC) histotype (WHO, 2021). Nevertheless, rare cases of typical and atypical carcinoids associated with squamous cell carcinoma or adenocarcinoma have been reported and extensively profiled at the molecular level (La Rosa et al., 2018). These combinations are not included as a specific entity in the WHO classification (WHO, 2021), and the pathogenesis of mixed histology is not well elucidated due to the rarity of presentation. The first proposed theory suggested a polyclonal origin with two independent precursors harboring different neoplasms (collision tumors). The second and more supported hypothesis indicates a monoclonal origin with a divergent differentiation of the two phenotypes starting from the same precursor clone (combined tumors) (La Rosa et al., 2018; Olofson and Tafe, 2018).

To date, only two cases of pulmonary adenocarcinoma (PA) mixed with atypical carcinoid (AC) with a detailed biological profile have been reported, and a monoclonal origin of the two components from a common transformed stem/progenitor tumor cell, which acquired divergent differentiation during neoplastic development, was suggested (i.e., epithelial carcinomatous, high grade, and epithelial neuroendocrine, low grade) (La Rosa et al., 2018; Olofson and Tafe, 2018). In both cases, the two components shared mutations in some genes, and additional mutations unique for each component were described in one case (La Rosa et al., 2018). However, the rarity of these tumors affects their deep and complete biological knowledge and the possibility to set an adequate medical therapy in metastatic/advanced patients.

Here, we present two novel additional cases of mixed PA and AC that were also profiled at the genetic level using next-generation sequencing (NGS) with a large gene panel including single nucleotide variations (SNVs), copy number variations (CNVs), and gene fusions.

# MATERIAL AND METHODS

## Patients Selection

Both patients presented as Case 1 and Case 2 underwent curative surgery at Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Casa Sollievo della Sofferenza Hospital (San Giovanni Rotondo, FG, Italy). The clinical–pathological information and the biological material used in this study were collected following the Declaration of Helsinki, after the Local Ethics Committee Approval and with the informed consent of patients for genetic analysis of the lesions.

## Pathological Evaluation

Surgical specimens were examined and processed according to the current College of American Pathologists (CAP) guidelines. The specimens were fixed in 10% buffered formalin for 24–48 h, sampled and embedded in paraffin. Three-micrometer-thick tissue sections were cut and stained with hematoxylin and eosin (H&E). Immunohistochemical staining of sections representative of both neoplastic histotypes of primitive tumors and metastasis was carried out with the following antibodies: chromogranin A, synaptophysin, transcriptional thyroid factor-1 (TTF1, clone 8G7G3/1; Dako, Glostrup, Denmark), pan-cytokeratins (AE1-3 clone, Dako), cytokeratin 7, anaplastic lymphoma kinase (ALK, D5F3 clone, Ventana Medical Systems, Inc.), c-ROS oncogene 1 (ROS1, D4D6 clone, Cell Signaling), and tumor proportion score (TPS) of programmed death-ligand 1 (PDL1, clone 22C3, Dako) were also recorded based on recent guidelines.

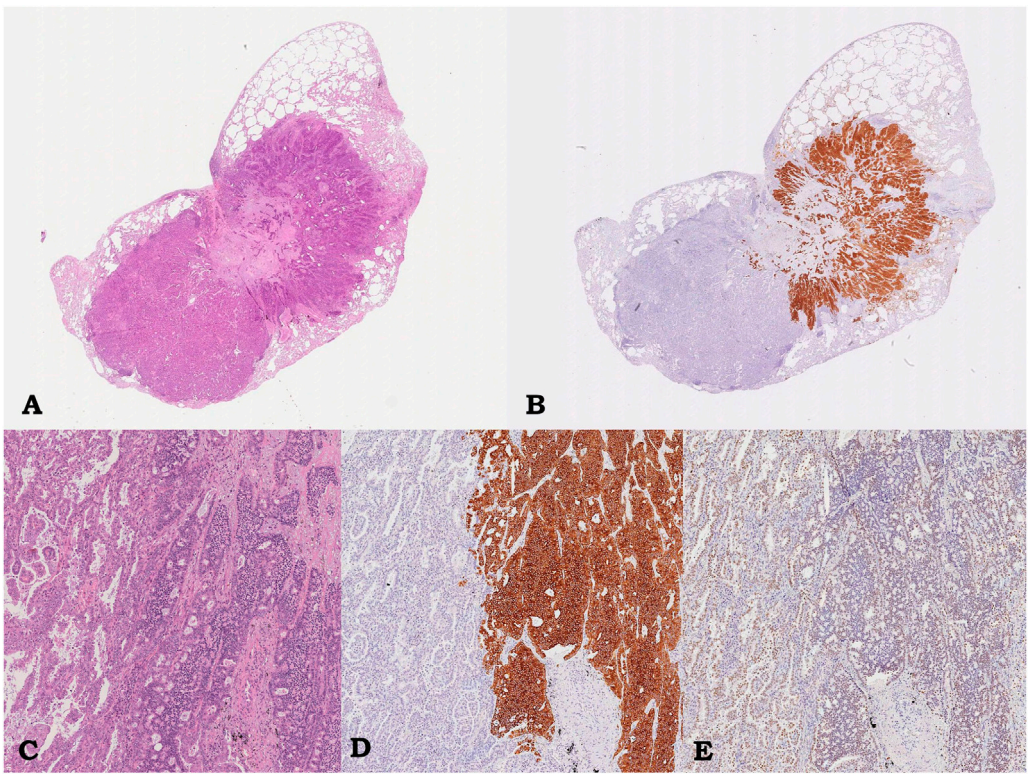



**FIGURE 1 |** Case 1, pulmonary lesion. Intraparenchymal neoplasm consisting of glandular component corresponding to lepidic-papillary pulmonary adenocarcinoma, on the left, and of adhered trabecular proliferation of monomorphic epithelioid cells corresponding to atypical carcinoid, on the right; hematoxylin/eosin, 0.5x **(A)** and 5.5x **(C)**. Immunostaining showing diffuse and strong immunoreactivity for Synaptophysin only in the carcinoid component **(**on the right, 5.5x; **B, D)** and immunoreactivity for TTF-1 in both component **(**5.5x, **D)**.

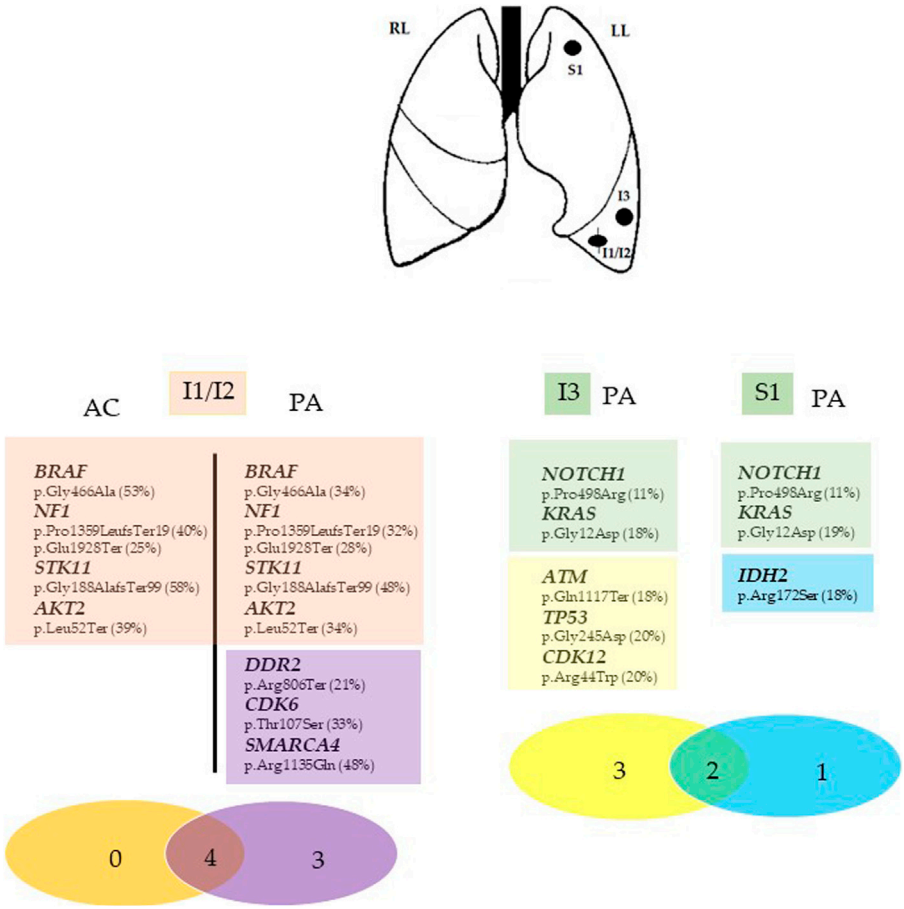



**FIGURE 2 |** Summary of the genomic findings of Case 1. Genes with mutations identified in each tumor entity of the inferior and superior left lung (I1/I2, I3, and S1) are listed as symbol and aminoacidic changes (% mutant allele fraction). The Venn diagram was used to show the shared mutations (red circle in I1/I2 nodule and green circle in I3 and S1 nodules) and exclusive mutations for each tumor fraction. The different numbers represent the somatic mutations in the corresponding samples, whereas the number in the overlapped regions red (I1/I2) and green (I3 and S1) are ubiquitous somatic mutations shared by the two concurrent tumor components (I1/I2) and by the two isolated adenocarcinoma nodules (I3 and S1) in the same patient. I1/I2, mixedcarcinoid/adenocarcinoma tumor; I3, inferior adenocarcinoma tumor, S1, superior adenocarcinoma tumor. Asterisks (*) indicate the different mutation in the same gene. LL, Left lung; RL, right lung. AC, atypical carcinoid; PA, pulmonary adenocarcinoma.

## Molecular Biology

Unstained formalin fixed-paraffin embedded (FFPE) cancer tissue sections were microdissected to enrich for at least 60% neoplastic cellularity of each histological component of tumor. DNA and RNA were extracted using the QIamp DNA Micro kit (Qiagen) and RecoverAll™ Total Nucleic Acid Isolation kit (Ambion), respectively, and quantified with the Qubit fluorimeter (Thermo Fisher Scientific, Waltham, MA, United States). Relevant SNVs, CNVs, and gene fusions from a total of 161 cancer driver genes were searched using targeted NGS with Oncomine™ Comprehensive Assay v3M panel (Thermo Fisher, Life technologies Inc. division) and are listed in **SupplementaryTable S1**. Amplifications were performed starting from 30 ng of DNA/RNA, and libraries were run on GeneStudio S5. Raw signal data were analyzed using Torrent Suite version 5.10.1, and the annotation of variants was performed by the Ion Reporter Server System v5.16 (https://ionreporter.thermofisher.com). A threshold of at least 20 mutated reads and an allelic frequency of 10% was used to perform mutation calling. Additionally, alignments were visually confirmed by the genome Browse software v3.0. Non-synonymous, insertions/deletions (indels) and mononucleotide variants (MNVs) located in splicesite_5′, splicesite_3′ region that produce missense and nonsense mutations, non-frameshift/frameshift insertion were considered.

The *ROS1* gene status was also assessed by fluorescence *in situ* hybridization (FISH) analysis using ZytoLight® SPEC ROS1 Dual Color Break Apart Probe (ZytoVision GmbH, Bremerhaven, Germany).

# RESULTS

## Case Report 1

On July 2008, a 59-year-old man, ex-smoker (45 packs/year), underwent upper right lung lobectomy and regional lymph

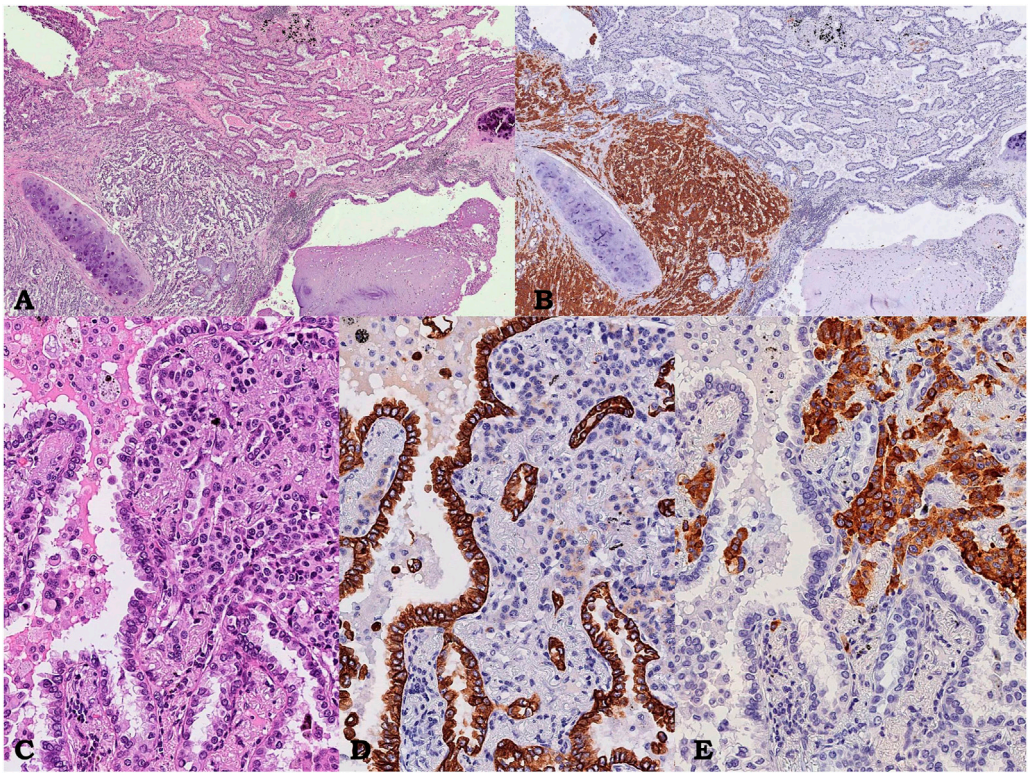



**FIGURE 3 |** Case 2, pulmonary neoplasm. Intraparenchymal neoplasm showing a glandular component corresponding to an acinar pulmonary adenocarcinoma on the right intermingled with an organoid proliferation of polygonal shaped uniform tumor cells on the left, near bronchus wall, corresponding to atypical carcinoid; hematoxylin/eosin, 3.5x **(A)**. Immunostaining showing diffuse and strong immunoreactivity for Synaptofisin only in the carcinoid component **(**on the left, 3.5x; **B)**, At higher power field, composite neoplasm **(**hematoxylin/eosin, **C**, 15x**)** showing diffuse and strong immunoreactivity for CK7 only in the adenocarcinomatous **(D** 15x**)** and chromogranin only in the neuroendocrine component **(E**, 15x**)**, respectively.

adenectomy with the diagnosis of stage I (pT2, pN0) lung adenocarcinoma, solid pattern. On July 2020, the chest CT scan revealed the presence of an upper left lung lobe and two lower left lung lobe lesions. On August 2020, the patient underwent a wedge resection of the upper and the lower lung lobes lesions with N1 and N2 nodal sampling.

Gross examination of the surgical specimens of the first atypical pulmonary resection of left inferior lobe revealed an Intraparenchymal, peripheral, solid, yellow-white lesion measuring 1.4 cm in greatest dimension. Histological assessment identified a well-circumscribed lesion composed of two different morphological components, tightly adhered but not intermingled each other. The first component represented about 60% of the whole neoplasm and was characterized by a solid and trabecular proliferation of polygonal-shaped uniform tumor cells, with nuclei with finely granular chromatin and inconspicuocus nucleoli, consistent with carcinoid. Four mitosis/2 $mm^2$ were identified, without tumoral necrosis. The second component, which represented about 40% of whole neoplasm, showed a main lepidic, non-mucinous pattern with secondary papillary architecture, corresponding to a lepidic-papillary pattern PA. Immunoreactivity for chromogranin A, synaptophysin, TTF-1, and pan-cytokeratins AE1–3 was documented in carcinoid component, while adenocarcinomatous component was positive only for TTF-1 and cytokeratins (**Figure 1**). CK7 immunoreactivity was selectively documented in the adenocarcinomatous component. A final diagnosis of "combined pulmonary adenocarcinoma with atypical carcinoid" was made. No immunoreactivity for ALK and ROS1 was documented in both components. TPS for PD-L1 was <1% in both components. Both two other lesions on atypical pulmonary resections of the upper left lobe and inferior left lobe showed intraparenchymal neoplasm of 1.8 and 1.5 cm, respectively, corresponding both to PA with a main solid pattern and lepidic, non-mucinous second pattern. A positivity ROS1 score of 2+ with gene rearrangement on FISH was documented in the upper pulmonary lobe neoplasia. No immunoreactivity for ALK and a PD-L1 TPS of 1%–49% were observed. No immunoreactivity for ALK and ROS1 with PD-L1 TPS of >50% were documented in the lower lobe neoplasia. No pleural invasion was observed. No lymph nodes metastases were found. The final UICC 2017 stage pT4 (m-3) (PL0) pN0 was assigned.

Targeted NGS analysis on Ion Torrent NGS platform using the Ocav3 panel was performed on DNA and RNA extracted from each microdissected histological components of the primary mixed neoplasm (I1/I2) and in the other two adenocarcinomas located in the inferior (I3) and superior (S1) left pulmonary lobes. Summary of genetic results are

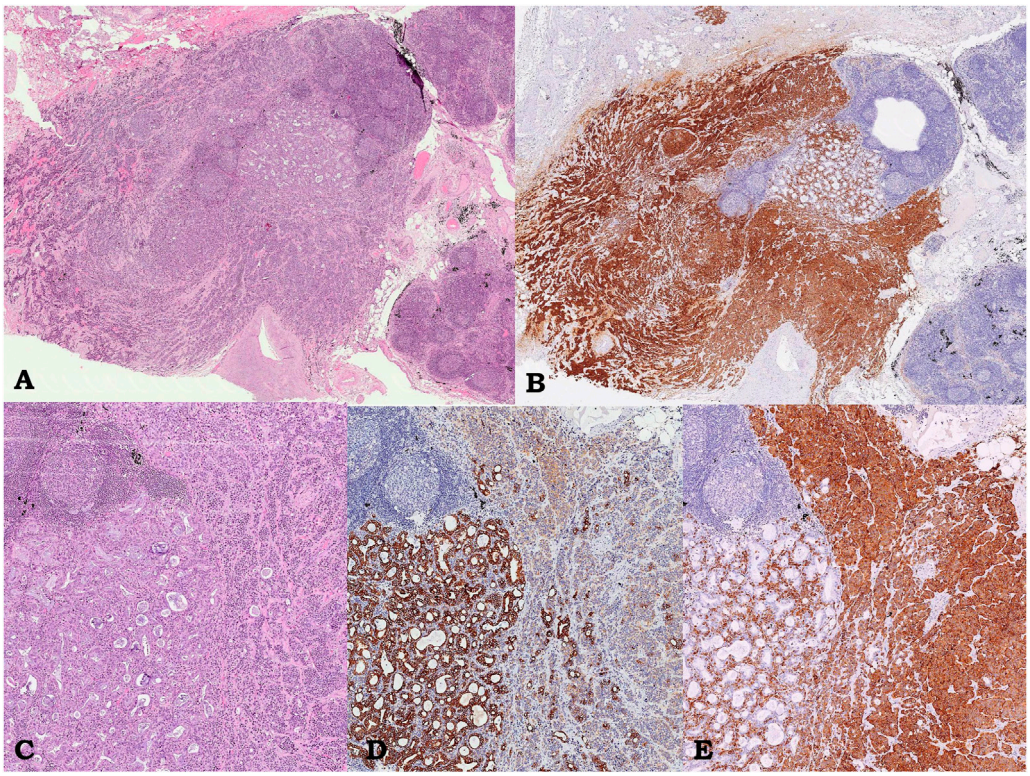



**FIGURE 4 |** Case 2, lymph node metastasis. Lymph node showing adenocarcinoma component intermingled with carcinoid; hematoxylin/eosin, 3.5× **(A)**. Immunostaining showing diffuse and strong immunoreactivity for Synaptofisin only in the carcinoid component **(**on the left, 3.5×; **B)**. At higher power field, composite neoplasm **(**hematoxylin/eosin, **C**, 4×**)** showing diffuse and strong immunoreactivity for CK7 only in the adenocarcinomatous **(D**, 4×**)** and chromogranin only in the neuroendocrine component **(E**, 4×**)**, respectively.

shown in **Figure 2.** Among a total of 161 investigated genes included in the NGS panel, both common and private somatic genes mutations were identified in each entity (**Figure 2**). The adenocarcinoma I1 and carcinoid I2 components of the mixed neoplasm shared the same mutations in *BRAF* (p.Gly466Ala), *NF1* (p.Pro1359LeufsTer19 and p.Glu1928Ter), *STK11* (p.Gly188AlafsTer99), and *AKT2* (p.Leu52Ter) genes. Four additional somatic mutations were detected in the *DDR2* (p.Arg806Ter), *CDK6* (p.Thr107Ser), and *SMARCA4* gene (p.Arg1135Gln) were identified in the adenocarcinoma component, whereas no specific mutations were identified in the carcinoid component. A different genetic profile emerged from the molecular analysis of the two additional adenocarcinoma nodules I3 and S1, lacking those somatic mutations identified in the mixed nodules, but shared two somatic missense mutations in the *KRAS* (p.Gly12Asp) and *NOTCH1* (p.Pro498Arg). In addition, the I3 adenocarcinoma showed a specific, somatic mutations in *ATM* (p.Gln1117Ter), *TP53* (p.Gly245Asp), and *CDK12* (p.Arg44Trp) genes, whereas the S1 adenocarcinoma showed a specific, somatic, missense mutations in *IDH2* gene (p.Arg172Ser).

Since November 2020, considering the ROS1 positivity, the patient started crizotinib, which is still ongoing, reporting a stable disease.

## Case Report 2

On October 2016, a 66-year-old woman, never smoker, underwent upper left lung lobectomy and regional lymph adenectomy, for a pulmonary mass discovered on radiological examination for shoulder pain present for a long time.

The gross examination of the surgical specimens showed an intraparenchymal, subpleural, whitish mass, measuring 8.5 cm in greatest dimension. Histologically, a well-circumscribed neoplasia composed of two different morphological components, which were separated in some areas and intermingled in others, was documented. The first component represented about 70% of the tumor burden and consisted of an organoid proliferation of polygonal, shaped uniform tumor cells with nuclei with finely granular chromatin and inconspicuous nucleoli, consistent with carcinoid. Four mitosis/2 $mm^2$ and diffuse, punctate necrosis were identified. The second component, representing about 30% of the tumor burden, showed a glandular architecture corresponding to acinar pattern of PA. Neoplastic infiltration of parietal pleura was documented (PL3). Immunoreactivity for chromogranin A, synaptophysin, TTF-1, and pan-cytokeratins (AE1-3 clone) was observed in carcinoid component, while glandular component was positive only for TTF-1 and cytokeratins; cytokeratin 7 was selectively expressed in the adenocarcinomatous component (**Figure 3**). Metastatis constituted by both neuroendocrine and non-neuroendocrine

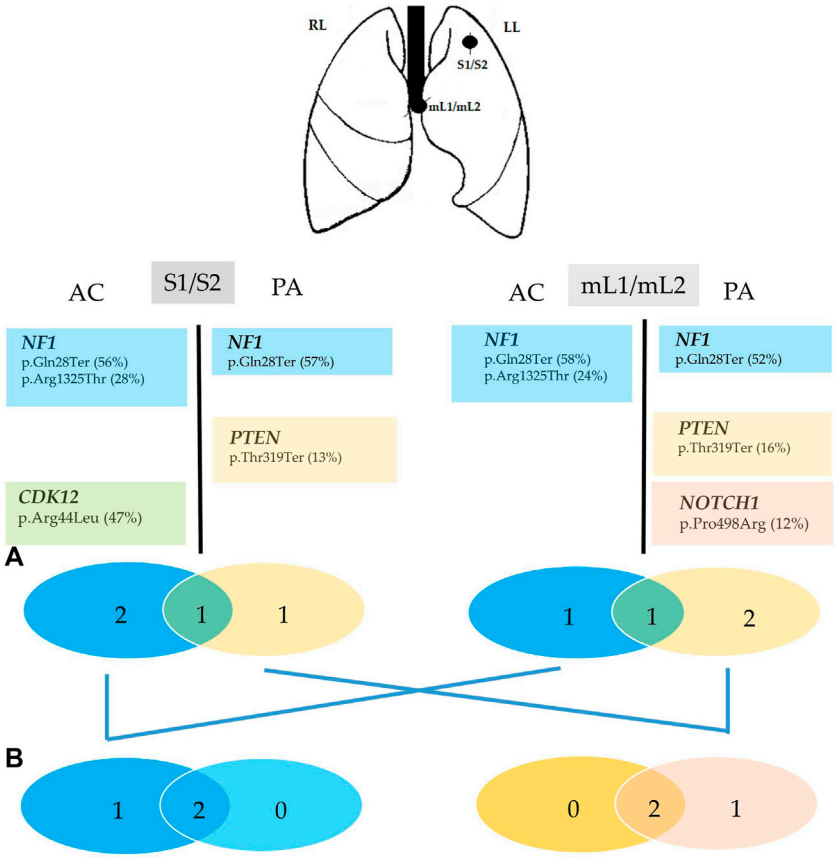



**FIGURE 5 |** Summary of the genomic findings of Case 2. Genes with mutations identified in each tumor entity of the superior left lung (S1/S2) and hilar lymphnode metastasis are listed as symbol and aminoacidic changes (% mutant allele fraction). The Venn diagrams **(A)** were used to show the relationship of somatic mutations between the S1 (blue circle) and S2 (yellow circle), nL1 (blue circle) and mL2 (yellow circle) in the patient. The number represents the somatic mutations identified in the corresponding samples, and the numbers in the overlapped green regions are the ubiquitous somatic mutations between the two concurrent tumor components (S1/S2, mL1/mL2) in the same patient. The Venn diagrams **(B)** were used to show the relationship of somatic mutations between the carcinoid component of S1 and between the adenocarcinoma component of S2 and mL2. The numbers represent the somatic mutations identified in the corresponding samples, and the numbers in the overlapped regions are ubiquitous somatic mutations shared by the components with the same histology in primary and metastatic sites.S1/S2, mixed carcinoid/adenocarcinoma tumor of the superior lung; mL1/mL2, mixed carcinoid/adenocarcinoma hilar lymphnode metatsasis. LL, left lung; RL, right lung. AC, atypicalcarcinoid; PA, pulmonary adenocarcinoma.

component was found in four out of six hilar lymph nodes, showing the same immunoreactivity in different neoplastic areas such as primitive lesion (**Figure 4**). Immunoreactivity for ALK and ROS1 resulted negative in both component, in primitive tumor and in lymph nodal metastasis. TPS for PD-L1 was <1% in both component, in primitive tumor and in lymph nodal metastasis. A final diagnosis of "combined pulmonary adenocarcinoma with atypical carcinoid" was made with stage pT3 (PL3) pN2 sec UICC 2017. Each neoplastic component in both primitive and metastatic lesions was individually microdissected from unstained FFPE slides for DNA and RNA extraction. Targeted NGS analysis on Ion Torrent NGS platform was performed using the Ocav3 panel.

Among the 161 investigated genes, both common and private genes mutations were identified in each investigated entity of primary and metastatic sites (**Figure 5**). No shared point mutations, CNVs, or gene fusions were identified in both adenocarcinoma and carcinoid components from any sites (S1/S2 and mL1/mL2). By contrast, one somatic mutation in the *PTEN* gene (p.Thr319Ter) was detected specifically in the adenocarcinoma components of both primary and metastatic sites (S2 and mL2), whereas the carcinoid component of both sites shared one somatic mutation in the NF1 gene (p.Arg1325Thr). Primary carcinoid S1 showed one private somatic mutation in the *CDK12* gene (p.Arg44Leu), whereas the adenocarcinoma component of metastatic site showed a private mutation in the *NOTCH1* gene (p.Pro498Arg).

On December 2016, she progressed due to the onset of osteoblastic bone lesions detected by TAC with a negative octreoscan performed on January 2017. In February 2017, cisplatin plus pemetrexed regimen started for six cycles, reporting a partial response after four cycles. In November 2017, due to a massive bone metastasis with bone marrow infiltration, the patient died.

**TABLE 1 |** Clinical and pathological features of reported combined pulmonary tumors.

| References | NE component | NSCLC histotype | Pathological stadiation | Smoking habit (p/y) | Prognosis | Molecular profile |
|---|---|---|---|---|---|---|
| Okazaki et al. Okazaki et al. (2015) | AC | SqCC | pT3pN0cM0 | Former smoker (75 p/y) | PD 9m death 21m | No |
| Owens et al. Owens and Fraire (2011) | TC | SqCC | T2N2M0 | Former smoker (15 p/y) | nr | No |
| Sen et al. Sen and Borczuk (1998) | TC | PA | pT1pN1cM0 | Current smoker (45 p/y) | nr | No |
| Abbi et al. Abbi et al. (2014) | TC | PA | T1aN0m0 (PA) pT2N0M0 (TC) | Former smoker (50 p/y) | nr | No |
| Olofson et al. Olofson and Tafe (2018) | AC | PA | pT1apNxcM0 | Former smoker (43 p/y) | Alive* | Yes |
| La Rosa et al. La Rosa et al. (2018) | AC | PA | pT2apN2cM0 | Former smoker (nr p/y) | PD 18m Alive 26m | Yes |
| Parente et al. (present work, Case 1) | AC | PA | pT4(m-3)pN0M0 | Former smoker (45 p/y) | SD 14m Alive 14m | yes |
| Parente et al. (present work, Case 2) | AC | PA | pT3pN2M0 | Never smoker | PD 2m Death 13m | Yes |

*AC, atypical carcinoid; TC, typical carcinoid; PA, pulmonary adenocarcinoma; NE, neuroendocrine; SqCC, squamous cell carcinoma; NSCLC, non-small cell lung cancer; p/y, pack/year; SD, stable disease; PD, progression of disease; m, months; nr, not reported*
**no follow-up time reported.*

## DISCUSSION

The WHO classification of thoracic tumors defines combined pulmonary neoplasms as an admixture of SCLC and LCNEC or as a coexistence of NSCLC with a carcinomatous neuroendocrine component (WHO, 2021). Rare cases of pulmonary carcinoids combined with a non-carcinoid component have been described, but these tumor types are not yet included in the current WHO classification as a specific entity (La Rosa et al., 2018; Olofson and Tafe, 2018). The coexistence of neuroendocrine and non-neuroendocrine neoplasms is a well-known phenomenon, and available molecular data published point towards a common origin of both components, thus supporting the hypothesis of a monoclonal neoplastic proliferation modality (La Rosa et al., 2018). Molecular studies on colorectal neuroendocrine/non-neuroendocrine carcinomas clearly suggest that the neoplastic proliferations arise from a single (monoclonal) non-neuroendocrine precursor, following the acquisition of one non-neuroendocrine profile that progress towards a more malignant neuroendocrine phenotype, through epigenetic, transcriptional, or translational events (Jesinghaus et al., 2017). Due to their exceptional rarity, very few molecular data supporting this hypothesis are available for mixed lung neoplasms with epithelial carcinomatous component combined with a well-differentiated neuroendocrine component (La Rosa et al., 2018; Olofson and Tafe, 2018).

Specifically, only two cases of combined NSCLC, squamous histotype, and carcinoid (Okazaki et al., 2015; Owens and Fraire, 2011) and two cases of combined NSCLC, adenocarcinoma histotype with typical carcinoid were reported to date (Sen and Borczuk, 1998; Abbi et al., 2014). Unfortunately, none of these cases was investigated at the molecular level, and their genetic profiles are unknown. More recently, two additional cases of combined lung atypical carcinoid and adenocarcinoma were described with a detailed histological and molecular characterization, the last consisting of a common mutational profile of both components, based on NGS panels (Olofson and Tafe, 2018; La Rosa et al., 2018). Specifically, in the first case a *BRAF p.Val600Glu* mutation in both adenocarcinoma and carcinoid components was documented (Olofson and Tafe, 2018). In the second case, the same missense mutations in *KRAS* (p.Gly13Asp), *PAPPA2* (p.Arg901Leu), and *NF1* (p.Val2106Phe) genes were described in both neoplastic components. Moreover, an additional somatic mutation was also detected in neuroendocrine components of both primary and metastatic sites, while a missense mutation in *SMARCA4* gene (p.Pro171Leu) was identified only in the adenocarcinoma component of primary site (**Table 1**) (La Rosa et al., 2018).

Our results expand the present knowledge about the hypothesis of a clonal origin of these mixed lung tumors and the ability of each histological component to enhance the metastatic process. In Case 1, the molecular profiling revealed both common and private somatic mutations in multiple genes in both neuroendocrine and non-neuroendocrine components, totally different from those identified in the two synchronous, non-mixed lung adenocarcinoma (**Table 2**). To note, we found common mutations in both carcinoid and adenocarcinoma components of Case 1 in *BRAF* and *STK11* genes that are rarely mutated in primary lung carcinoid (Armengol et al., 2015; Clinical Lung Cancer Genome Project (CLCGP), 2013). In addition, a mutation in the chromatin-remodeling *SMARCA4* gene was identified in the adenocarcinoma component, frequently associated with pure carcinoid histology (Simbolo et al., 2017).

Molecular and histological characterization of Case 2 gave additional information in supporting the common origin of both neoplastic components. To note, as in the case no.1, we

**TABLE 2 |** Molecular profile of published combined pulmonary adenocarcinoma and atypical carcinoids of lung primary site of neoplasms.

| Reference | Mutated genes | | Type of mutation (aa change) | | |
|---|---|---|---|---|---|
| | NE (AC) component | NSCLC (PA) component | NE (AC) component | NSCLC (PA) component | |
| Olofson et al. Olofson and Tafe (2018) | ***BRAF*** | ***BRAF*** | **p.Val600Glu** | **p.Val600Glu** | |
| La Rosa et al. La Rosa et al. (2018) | ***KRAS*** ***PAPPA2*** ***NF1*** | ***KRAS*** ***PAPPA2*** ***NF1*** *SMARCA4* | **p.Gly13Asp p.Arg901Leu p.Val2106Phe** | **p.Gly13Asp p.Arg901Leu p.Val2106Phe** p.Pro171Leu | |
| *Present work* | ***BRAF*** ***NF1*** ***NF1*** ***STK11*** ***AKT2*** *DDR2* *CDK6* *SMARCA4* | ***BRAF***, ***NF1***, ***NF1***, ***STK11***, ***AKT2*** | **p.Gly466Ala, p.Pro1359LeufsTer19, p.Glu1928Ter, p.Gly188AlafsTer99, p.Leu52Ter**, p.Arg806Ter, p.Thr107Ser, p.Arg1135Gln | **p.Gly466Ala, p.Pro1359LeufsTer19, p.Glu1928Ter, p.Gly188AlafsTer99, p.Leu52Ter** | Case 1 |
| *Present work* | ***NF1*** *NF1* *CDK12* | ***NF1***, *PTEN*, *NOTCH1* | **p.Gln28Ter**, p.Arg1325Thr, p.Arg44Leu | **p.Gln28Ter**, p.Thr319Ter, p.Pro498Arg | Case 2 |

*AC, atypical carcinoid; PA, pulmonary adenocarcinoma; m, months; NE, neuroendocrine; NSCLC, non-small cell lung cancer; aa change, aminoacidic change.*
*Shared gene mutations are highlighted in bold.*

identified in case no.2 non-sense pathogenic mutations in *NF1* gene shared by the two neoplastic components both of primary and of metastatic sites. The *NRF1* mutations, mainly associated with the NSCLC, were reported by La Rosa as shared by both component adenocarcinoma/carcinoid, but their role in the pathogenesis of mixed component remains unknown (Philpott et al., 2017). Finally, the presence of *NOTCH1* private mutation in the adenocarcinoma component only in the primary site, in addition to a common mutation in *PTEN* gene shared by the adenocarcinoma component in primary and nodal metastasis, should support their/its role in inducing metastasis. In this context, transcriptional analysis of such tumors could be also useful to better assess the effect of such mutations and distinguish genetic determinants from passenger mutations, and to assess the existence of intermediate gene expression programs.

Our data add new interesting insight into the understanding of the biology of this unclassified entity, since, if integrated with the previously described, they favor a monoclonal origin of both components, with clinical and therapeutic implications. At this point, some clinical questions arise, such as the following: is there a role for adjuvant chemotherapy? What is the most appropriate regimen for the treatment of advanced disease? Even if the carcinoid component of a mixed tumor should be less aggressive than the NSCLC component, it should be relatively resistant to chemotherapy and radiotherapy, and there is no proven optimal therapy for metastatic unresectable carcinoids tumors (Caplin et al., 2015; Baudin et al., 2021).

Adding more pathological, biomolecular, and clinical information to these rare clinical entities will increase the knowledge and the expertise to manage this specific histotype. In this context, a multicentric study collecting a case series of similar cases is needed to have a better comprehension of molecular mechanisms underlying the pathogenesis of this uncommon tumor entity.

Finally, we propose to classify this rare not-codified entity as "mixed" neoplasm, as yet codified in the 5th edition of Digestive System Tumors in the Mixed Neuroendocrine–Non-neuroendocrine Neoplasm (MiNENs) (WHO, 2019).

## 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 below: BioProject PRJNA768323.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Commitee of IRCCS Casa Sollievo della Sofferenza. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

Conceptualization, supervision, and writing—original draft preparation: PP, AR, and LM. Investigations, methodology, and validation: AS, FF, FC, MC, and GP. Data curation: PP, AR, AS, FF, CM, AC, TM, and LM. Recourses: PG and MT. Writing—review and editing: PP, AR, AS, FPF, CM,

MT, GF, AC, PG, TM, and LM. Funding acquisition: LM. All authors have read and agreed to the published version of the article.

## FUNDING

This research was funded by the Italian Ministry of Health, Ricerca Corrente 2021, by the "5 × 1000" voluntary contributions to Fondazione IRCCS Casa Sollievo della Sofferenza.

## REFERENCES

Abbi, K. K. S., Hameed, M. K., Jiang, Y., De Las Casas, L. E., and Schwann, T. A. (2014). Pulmonary Collision Tumor Consisting of Adenocarcinoma and Typical Carcinoid-A Case Report and Review of Literature. *Am. J. Ther.* 21, e234–e238. doi:10.1097/MJT.0b013e318293b0b0

Armengol, G., Sarhadi, V. K., Rönty, M., Tikkanen, M., Knuuttila, A., and Knuutila, S. (2015). Driver Gene Mutations of Non-Small-Cell Lung Cancer Are Rare in Primary Carcinoids of the Lung: NGS Study by Ion Torrent. *Lung* 193, 303–308. doi:10.1007/s00408-015-9690-1

Baudin, E., Caplin, M., Garcia-Carbonero, R., Fazio, N., Ferolla, P., Filosso, P. L., et al. (2021). Lung and Thymic Carcinoids: ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up☆. *Ann. Oncol.* 32, 439–451. doi:10.1016/j.annonc.2021.01.003

Caplin, M. E., Baudin, E., Ferolla, P., Filosso, P., Garcia-Yuste, M., Lim, E., et al. (2015). Pulmonary Neuroendocrine (Carcinoid) Tumors: European Neuroendocrine Tumor Society Expert Consensus and Recommendations for Best Practice for Typical and Atypical Pulmonary Carcinoids. *Ann. Oncol.* 26, 1604–1620. doi:10.1093/annouc/mdv04110.1093/annonc/mdv041

Clinical Lung Cancer Genome Project (CLCGP) (2013). A Genomics-Based Classification of Human Lung Tumors. *Sci. Transl Med.* 5, 209ra153. doi:10.1126/scitranslmed.3006802

Jesinghaus, M., Konukiewitz, B., Keller, G., Kloor, M., Steiger, K., Reiche, M., et al. (2017). Colorectal Mixed Adenoneuroendocrine Carcinomas and Neuroendocrine Carcinomas Are Genetically Closely Related to Colorectal Adenocarcinomas. *Mod. Pathol.* 30, 610–619. doi:10.1038/modpathol.2016.220

La Rosa, S., Simbolo, M., Franzi, F., Uccella, S., Imperatori, A., Nardecchia, E., et al. (2018). Combined Adenocarcinoma-Atypical Carcinoid of the Lung. Targeted Next-Generation Sequencing (NGS) Suggests a Monoclonal Origin of the Two Components. *Diagn. Histopathology* 24, 120–123. doi:10.1016/j.mpdhp.2018.02.002

Okazaki, M., Sano, Y., Soga, Y., Katayama, H., Sakao, N., Shikatani, Y., et al. (2015). Combined Atypical Carcinoid Tumour and Squamous Cell Carcinoma of the Lung. *Intern. Med.* 54, 1385–1388. doi:10.2169/internalmedicine.54.3846

Olofson, A. M., and Tafe, L. J. (2018). A Case of a Primary Lung Cancer Comprised of Adenocarcinoma and Atypical Carcinoid Tumor with Both Components Harboring BRAF p.V600E Mutation. *Exp. Mol. Pathol.* 104, 26–28. doi:10.1016/j.yexmp.2017.12.004

Owens, C. L., and Fraire, A. E. (2011). Combined Carcinoid Tumor and Squamous Cell Carcinoma of Lung: Expanding the Spectrum of Combined Tumors of Lung. *Int. J. Surg. Pathol.* 19, 273–275. doi:10.1177/1066896910397883

Philpott, C., Tovell, H., Frayling, I. M., Cooper, D. N., and Upadhyaya, M. (2017). The NF1 Somatic Mutational Landscape in Sporadic Human Cancers. *Hum. Genomics.* 11, 13. doi:10.1186/s40246-017-0109-3

Sen, F., and Borczuk, A. C. (1998). Combined Carcinoid Tumor of the Lung: A Combination of Carcinoid and Adenocarcinoma. *Lung Cancer* 21, 53–58. doi:10.1016/s0169-5002(98)00042-7

Simbolo, M., Mafficini, A., Sikora, K. O., Fassan, M., Barbi, S., Corbo, V., et al. (2017). Lung Neuroendocrine Tumours: Deep Sequencing of the Four World Health Organization Histotypes Reveals Chromatin-Remodelling Genes as Major Players and a Prognostic Role forTERT,RB1,MEN1andKMT2D. *J. Pathol.* 241, 488–500. doi:10.1002/path.4853

WHO (2019). *WHO Classification of Digestive System Tumours*. Fifth edition. Lyon: IARC Press.

WHO (2021). *WHO Classification of Thoracic Tumours*. Fifth edition. Lyon: IARC Press.

## ACKNOWLEDGMENTS

The authors are grateful to the patients and their families for participating to this study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2021.784876/full#supplementary-material

**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 handling editor declared a past co-authorship with the authors.

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

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 03 February 2022
doi: 10.3389/fmolb.2021.788690

# MicroRNA-125b-5p Correlates With Prognosis and Lung Adenocarcinoma Progression

*Lin Tang[1,2†], Yixiao Yuan[1,2†], Haoqing Zhai[3,4†], Juan Wang[1,2†], Dahang Zhang[1,2], Huasu Liang[2], Yulin Shi[2], Lincan Duan[1,2*] and Xiulin Jiang[3,4*]*

[1]The Department of Thoracic Surgery, The Third Affiliated Hospital of Kunming Medical University, Kunming, China, [2]Department of Graduate School of Kunming Medical University, Kunming, China, [3]Key Laboratory of Animal Models and Human Disease Mechanisms of Chinese Academy of Sciences and Yunnan Province, Kunming Institute of Zoology, Kunming, China, [4]Kunming College of Life Science, University of Chinese Academy of Sciences, Beijing, China

A growing number of studies have focused on investigating microRNAs as crucial regulators in the progression of multiple cancer types. Nevertheless, the biological effects and immunological role of miR-125b-5p in non-small cell lung cancer (lung adenocarcinoma, LUAD) have not been determined. The present study aimed to examine the function of miR-125b-5p on cell proliferation and the outcomes of LUAD patients. We utilized diverse public databases in the analysis of the expression, prognosis, diagnostic value, and immune role of miR-125b-5p in non-small cell lung cancer. The growth curve, colony formation, flow cytometry, and Transwell and invasion assays were utilized to determine the function of miR-125b-5p in LUAD progression. In this study, we found that miR-125b-5p was decreased in LUAD and correlated with poor prognosis. Pathway analyses revealed that miR-125b-5p was mainly involved in cell proliferation and immune regulation. Moreover, *in vitro* experiments indicated that the overexpression of miR-125b-5p significantly inhibited cell proliferation, migration, and invasion and induced cell apoptosis of LUAD. Finally, we discovered that miR-125b-5p correlated with immune cell infiltration. In summary, these results demonstrated that miR-125b-5p serves as a prognostic marker and a therapeutic target for LUAD.

**Keywords: miRNA-125b-5p, non-small cell lung cancer, diagnostics, biomarkers, cell proliferation, cell migration, cell apoptosis**

OPEN ACCESS

***Edited by:***
*Umberto Malapelle,*
*University of Naples Federico II, Italy*

***Reviewed by:***
*Carmelina Antonella Iannuzzi,*
*G. Pascale National Cancer Institute Foundation (IRCCS), Italy*
*Qiang Guo,*
*Hubei University of Medicine, China*
*Chen Li,*
*Free University of Berlin, Germany*

***Correspondence:**
*Lincan Duan*
*duanmumuhuosan@163.com*
*Xiulin Jiang*
*jiangxiulin@mail.kiz.ac.cn*

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

***Specialty section:***
*This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences*

***Received:*** *03 October 2021*
***Accepted:*** *23 December 2021*
***Published:*** *03 February 2022*

***Citation:***
*Tang L, Yuan Y, Zhai H, Wang J, Zhang D, Liang H, Shi Y, Duan L and Jiang X (2022) MicroRNA-125b-5p Correlates With Prognosis and Lung Adenocarcinoma Progression. Front. Mol. Biosci. 8:788690. doi: 10.3389/fmolb.2021.788690*

## INTRODUCTION

Lung cancer remains the most common deadly disease, with an estimated 2.09 million new cases and 1.76 million deaths each year (Rivera and Wakelee, 2016). Non-small cell lung cancers (NSCLCs), the most common lung cancers, are known to have diverse pathological features. As the most common histologic subtype of lung cancer, lung adenocarcinoma (LUAD) usually results in a lower 5-year survival rate in lung cancer (Schwartz and Cote, 2016; Liang et al., 2021). Although there are many diverse treatments, including chemoradiotherapy, targeted therapy, and immunotherapy, the prognosis of LUAD remains poor because it lacks effective diagnostic markers. Therefore, it is urgent to further uncover specific prognostic prediction methods for the diagnosis and treatment of lung cancer.

MicroRNAs (miRNAs) are non-coding RNAs that often inhibit gene expression at the posttranscriptional level. Mounting evidence has demonstrated that miRNAs play an important

role in modulating lung cancer cell proliferation, invasion, migration, and apoptosis (Iqbal et al., 2019). For instance, Qu et al. found that miR-133b was downregulated in renal cell carcinoma cell lines. Overexpression of miR-133b significantly inhibits cell proliferation, migration, and invasion by targeting MMP9 in renal cell carcinoma (Wu et al., 2014). Furthermore, Wu et al. showed that miR-125b was decreased in bladder cancer specimens. An elevated level of miR-125b inhibits cell growth and migration and induces cell cycle arrest in the G1 phase, with a further study showing miR-125b targeting SphK1 and inhibiting bladder cancer progression (Zhao et al., 2015). Previous studies have shown that miR-125b plays a tumor suppressor role in breast cancer (Wang et al., 2020), cervical cancer (Sun et al., 2019), colorectal cancer (Yagi et al., 2019), and prostate cancer (Zedan et al., 2019). However, the underlying mechanisms of how miR-125b-5p regulates LUAD progression and metastasis remain elusive, which enhances the significance of the findings of the present study.

In this study, we found that miR-125b-5p was notably downregulated in LUAD tissues and cells and inhibited the proliferative and metastatic phenotypes in LUAD cell lines. Significantly, we demonstrated that miR-125b-5p is correlated with immune cell infiltration. In summary, these results showed that miR-125b-5p is a potential molecular marker for poor prognosis in LUAD and provides potential diagnostic and immunotherapeutic biomarkers for LUAD in the future.

# METHODS

## Expression, Prognosis, and Clinical Information Analysis

We mainly used The Cancer Genome Atlas (TCGA) and the Kaplan–Meier plotter (http://kmplot.com/analysis/) (Nagy et al., 2021) to analyze the expression, prognosis, and clinical information of miR-125b-5p in cancers.

## Gene Set Enrichment and Immune Infiltration Analysis

We utilized the clusterProfiler package for the analysis of the potential signaling pathway and molecular function of the miR-125b-5p target gene in LUAD (Subramanian et al., 2005; Yu et al., 2012). We used the GSVA package in R to examine the LUAD immune infiltration of 24 tumor-infiltrating immune cells in tumor samples through single-sample gene set enrichment analysis (ssGSEA) (Bindea et al., 2013; Hänzelmann et al., 2013). The correlation between miR-125b-5p and the infiltration levels of immune cells was analyzed with Spearman's correlation, and the immune cells with different expression groups of miR-125b-5p were analyzed using the Wilcoxon rank-sum test.

## Cell Culture, Constructs, and Transfection

The BEAS-2B cell line was purchased from the Cell Bank of Kunming Institute of Zoology and cultured in BEGM media (CC-3170; Lonza, Basel, Switzerland). Lung cancer cell lines, including A549, H1299, H358, HCC827, SPC-A1, H1650, and H1975, were purchased from Cobioer (Nanjing, China), with STR document, and were cultured in RPMI-1640 medium (Corning, Corning, NY, United States) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. The normal control (NC), miR-125b-5p mimics, and miR-125b-5p inhibitors were purchased from RiboBio (Guangzhou, China). Cells were transfected with the indicated miRNA mimics or NC using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, United States) and then collected for various experiments.

## Cell Proliferation and Colony Formation Assays

Cell proliferation and colony formation assays were performed as previously documented (Xiong et al., 2021). Briefly, for the cell proliferation assay, the indicated cells were plated into 12-well plates at a density of $1.5 \times 10^4$, and the cell numbers were subsequently counted each day using the automatic cell analyzer Countstar (IC 1000; Shanghai Ruiyu Biotech Co., Shanghai, China). For the colony formation assay, the indicated cells were seeded in a 6-well plate (cat. 703001; NEST, Jiangsu, China) with 600 cells per well supplemented with 2 ml cell culture medium. The cell culture medium was changed every 3 days for 2–3 weeks, and then the indicated cells were fixed with 4% paraformaldehyde (PFA) and stained with 0.5% crystal violet. For the tumor sphere formation assay, the indicated cells were plated in an ultralow-attachment 6-well plate (cat. 3471; Corning), cultured in serum-free DMEM/F12 supplemented with B27 (cat. 2309544; Gibco, Waltham, MA, USA), 20 ng/ml epidermal growth factor (EGF) and 20 ng/ml basic fibroblast growth factor (bFGF), and 4 μg/ml heparin. From 10 to 14 days after culture, the spheres were photographed and counted using a Nikon inverted microscope (Ti-S).

## Cell Migration Assay

Cell migration assays were performed as previously documented (Xiong et al., 2021). Briefly, to produce a wound, the monolayer cells on the 6-well plate were scraped in a straight line with pipette tips. The plate was then washed with PBS to remove detached cells. Photographs of the scratch were taken at indicated time points using the Nikon inverted microscope (Ti-S). Gap width was calculated using GraphPad Prism software. For the Transwell assay, $2.5 \times 10^4$ cells in 100 μl serum-free medium were plated in a 24-well plate chamber insert (cat. 3422; Corning Life Sciences) with a medium containing 10% FBS at the bottom of the insert. The cells were incubated for 24 h and then fixed with 4% PFA for 20 min. After washing, the cells were stained with 0.5% crystal violet.

## Flow Cytometry Assay

The Annexin V FITC Apoptosis Detection Kit I (556547; BD, Shanghai, China) was used to evaluate cellular apoptosis according to the manufacturer's instructions. For cell flow cytometry experiments, the indicated cells were digested and

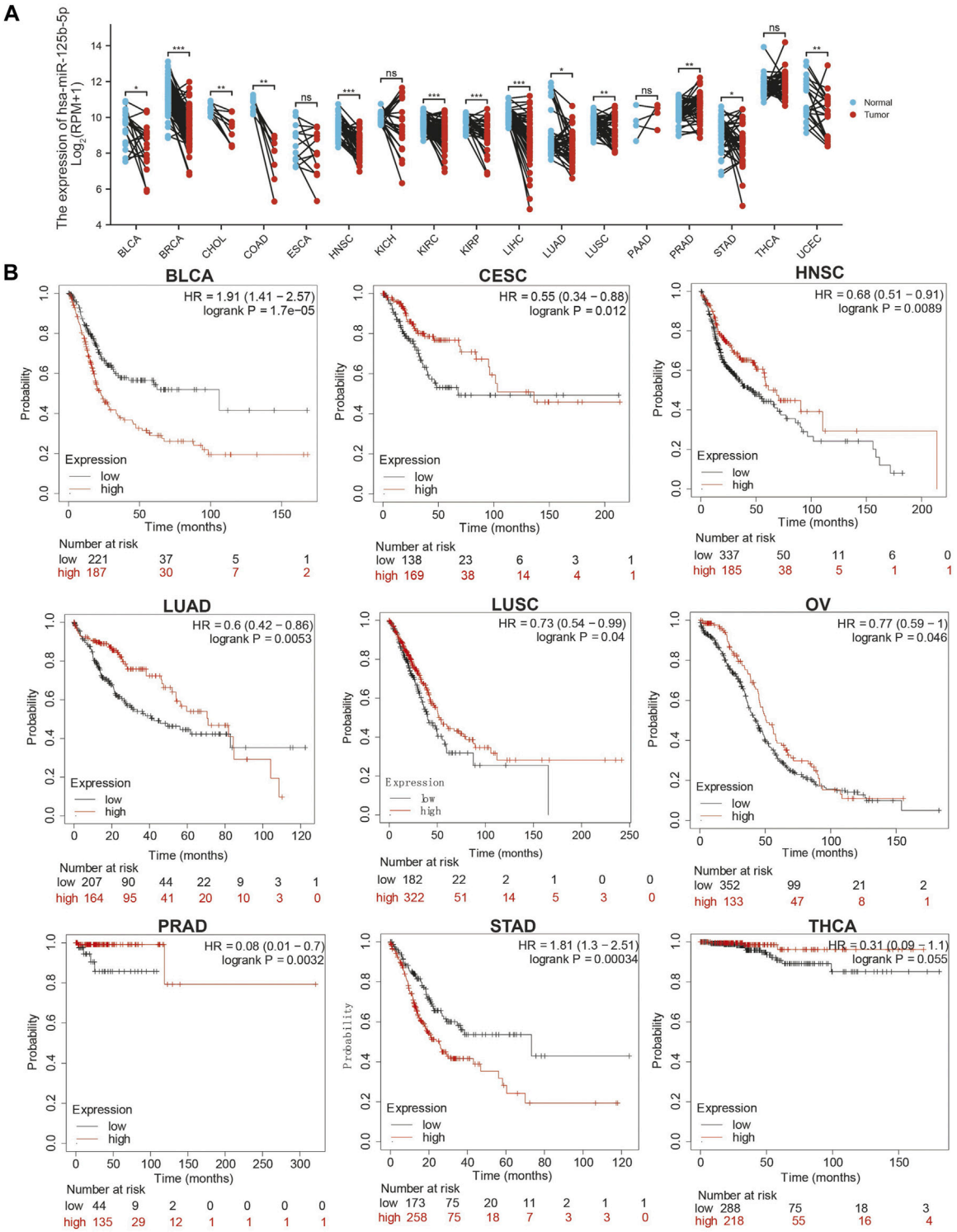



**FIGURE 1 |** Expression analysis for miR-125b-5p in human cancers. **(A)** Expression of miR-125b-5p in pan-cancer. **(B)** Prognosis of miR-125b-5p in pan-cancer examined using the Kaplan–Meier plotter database.

washed with PBS twice and then fixed in 75% alcohol overnight at −20°C. The fixed cells were washed three times and then stained with propidium iodide (PI) buffer at room temperature for 30 min in the dark. The cell cycle was then analyzed using the FACSAria SORP machine (BD, San Diego, CA, United States).

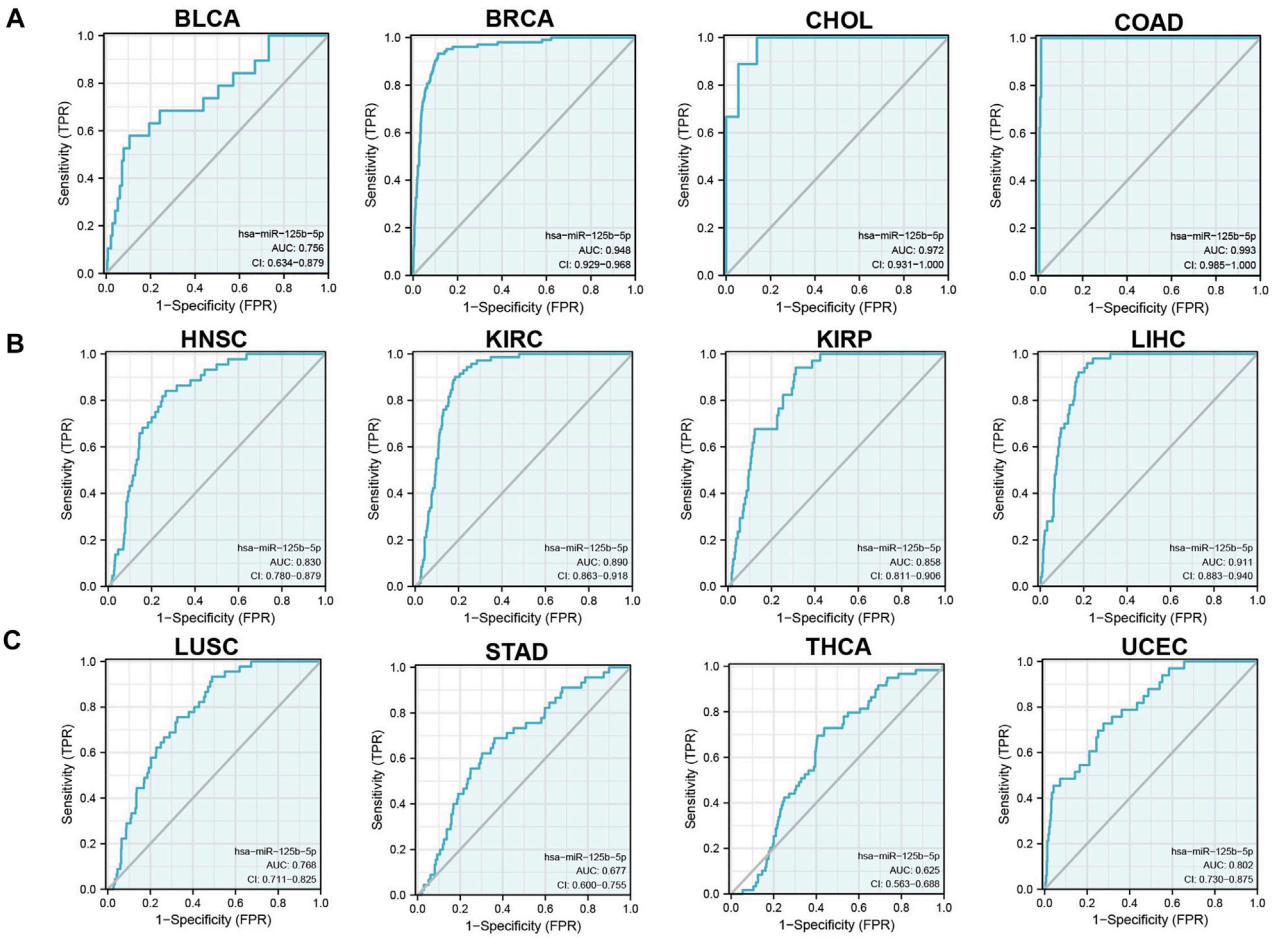



**FIGURE 2 |** Analysis of the receiver operating characteristic (ROC) curve for miR-125b-5p in human cancers. **(A–C)** ROC curve analysis and area under the curve (AUC) values for miR-125b-5p in diverse human cancers.

## Real-Time RT-PCR Assay

Quantitative real-time PCR (qRT-PCR) assay was performed as documented (Jiang et al., 2018). The primer sequences were as follows: miR-125b-5p primer: TCCCTGAGACCCTAACTTGTG A; U6-F: GGTCGGGCAGGAAAGAGGGC; and U6-R: GCT AATCTTCTCTGTATCGTTCC. The expression quantification was obtained with the $2^{-\Delta\Delta Ct}$ method.

## Western Blot Assay

The Western blot and immunohistochemistry staining assay was performed as documented. Briefly, cell lysates were collected, Western blot was performed, and the primary antibody was incubated overnight, followed by secondary antibody incubation. Finally, develop using instruments. Detailed information of the antibodies utilized in the study is as follows:

| Antibody name | Catalog no. | Dilution | Supplier | Species |
|---|---|---|---|---|
| β-actin | 60008-1-1g | 1:5,000 | Proteintech | Mouse |
| PARP | 9542S | 1:1,000 | CST | Rabbit |
| Bcl-2 | 15071S | 1:500 | CST | Mouse |
| Bax | ab77566 | 1:1,000 | Abcam | Mouse |

## Statistical Analysis

Significance of the data between two experimental groups was determined by Student's t-test, and multiple group comparisons were analyzed using one-way ANOVA. Values of $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$ were considered significant.

# RESULTS

## MiR-125b-5p is Decreased in Pan-Cancer

We initially examined the expression of miR-125b-5p in multiple cancer types using the TCGA database. The results showed that the expression of miR-125b-5p was significantly lower in breast invasive carcinoma, bladder urothelial carcinoma, cholangiocarcinoma, colon adenocarcinoma, head and neck squamous cell carcinoma, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma, liver hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, stomach adenocarcinoma, and uterine corpus endometrial carcinoma. On the contrary, miR-125b-5p showed a

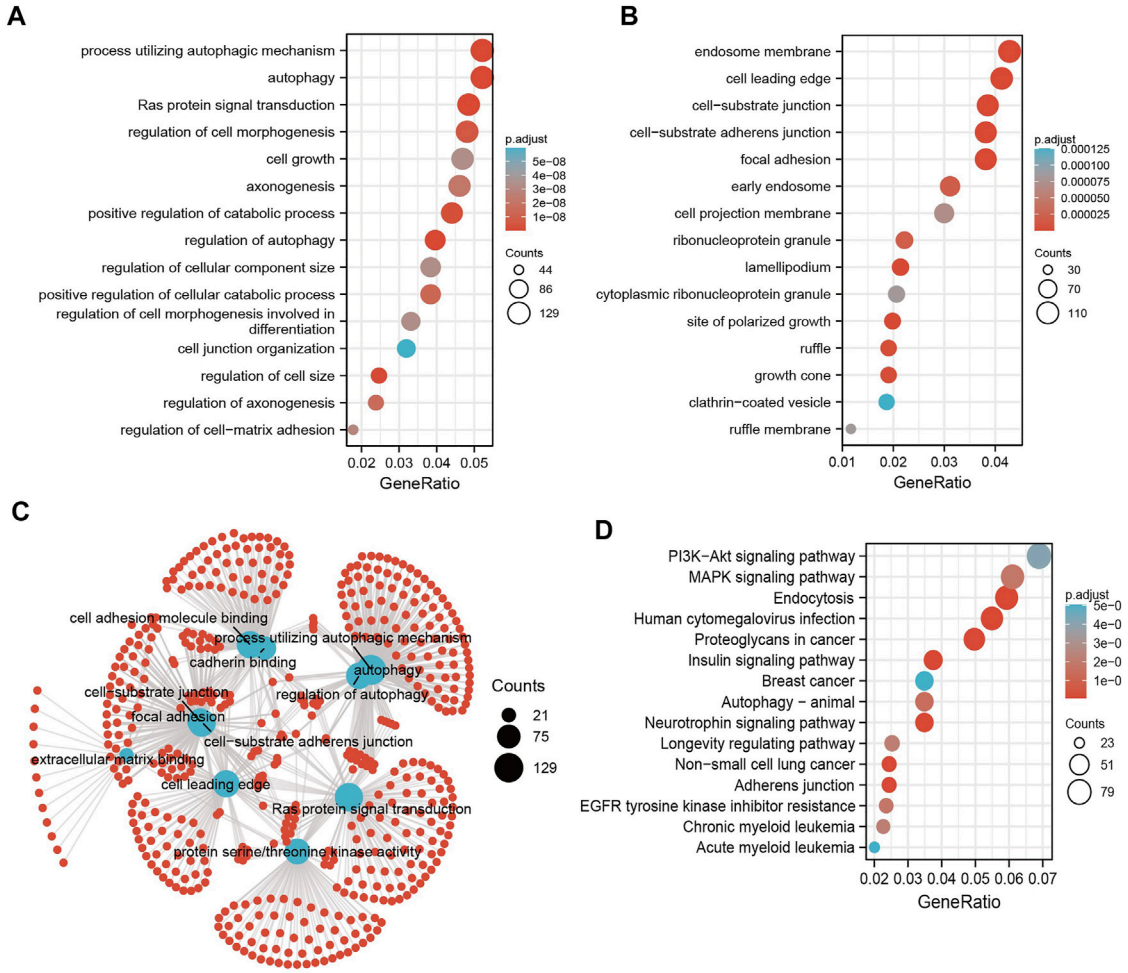



**FIGURE 3 |** Analysis of the functions of the downstream target genes of miR-125b-5p in human cancers. **(A–C)** Gene Ontology (GO) terms of the downstream target genes of miR-125b-5p in pan-cancer analysis. **(D)** Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathways of the downstream target genes of miRNA-125b-5p in pan-cancer analysis by starBase.

significantly higher expression in prostate adenocarcinoma (**Figure 1A**). These data confirmed that miR-125b-5p mainly decrease in multiple cancer types. Furthermore, we analyzed the prognostic value of miR-125b-5p in pan-cancer. The results demonstrated that a low expression of miR-125b-5p was correlated with poor prognosis in cervical squamous cell carcinoma and endocervical adenocarcinoma, head and neck squamous cell carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, ovarian serous cystadenocarcinoma, prostate adenocarcinoma, and thyroid carcinoma. A low expression of miR-125b-5p was also correlated with better prognosis in bladder urothelial carcinoma and stomach adenocarcinoma (**Figure 1B**). Given that miR-125b-5p significantly affected the prognosis of multiple cancer types, therefore, we next examined its diagnostic value in human cancer patients. Receiver operating characteristic (ROC) curve analysis demonstrated higher sensitivity and specificity of miR-125b-5p for detection in pan-cancer (**Figure 2**). These data support the potential of miR-125b-5p as a biomarker in multiple cancer types.

## Analysis of the Function of the Target Genes of miR-125b-5p

MicroRNAs mainly regulate the expression of messenger RNAs (mRNAs) and play an important role in diverse cancer progression. Therefore, we utilized starBase to predict the target genes of miR-125b-5p and used these genes to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. The top 16 significant terms in the enrichment analysis of biological process (BP), molecular function (MF), and cellular component (CC) were presented. Notably, in terms of BP, the target genes of miR-125b-5p were enriched in the regulation of autophagy, Ras protein signal transduction, autophagy, process utilizing autophagic mechanism, regulation of cell size, positive regulation of catabolic process, regulation of cell morphogenesis,

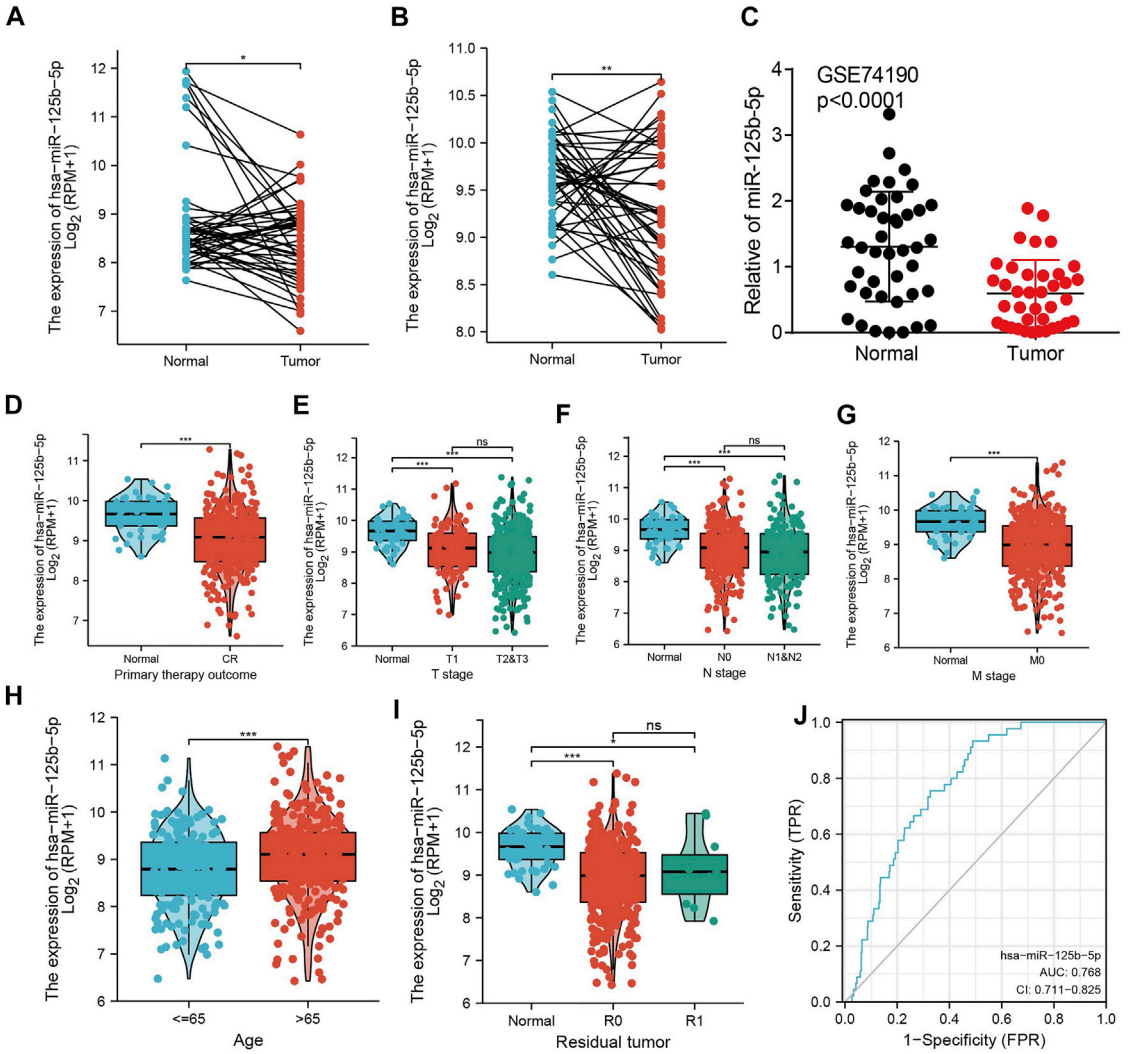



**FIGURE 4 |** miR-125b-5p was decreased in lung adenocarcinoma (LUAD). **(A–C)** Examination of the expression of miR-125b-5p in LUAD in The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) datasets. **(D–I)** Correlation between miR-125b-5p expression and clinicopathological characteristics of LUAD, including primary therapy outcome, TNM stage, age, and residual tumor. **(J)** Receiver operating characteristic (ROC) curve analysis and area under the curve (AUC) values for miR-125b-5p in LUAD. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

positive regulation of cellular catabolic process, regulation of axonogenesis, axonogenesis, regulation of cell–matrix adhesion, regulation of cell morphogenesis involved in differentiation, cell growth, regulation of cellular component size, cell junction organization, and endomembrane system organization (**Figure 3A**).

In terms of CC, the target genes of miR-125b-5p were enriched in the cell leading edge, focal adhesion, cell–substrate junction, cell–substrate adherens junction, endosome membrane, lamellipodium, site of polarized growth, growth cone, ruffle, early endosome, ribonucleoprotein granule, cell projection membrane, ruffle membrane, cytoplasmic ribonucleoprotein granule, and clathrin-coated vesicles (**Figure 3B**). In terms of MF, the target genes of miR-125b-5p were enriched in cell adhesion molecule binding, extracellular matrix binding, protein serine/threonine kinase activity, cadherin binding, mitogen-activated protein kinase binding, Ras GTPase binding, miRNA binding, regulatory RNA binding, modification-dependent protein binding, small GTPase binding, and p53 binding (**Figure 3C**). Furthermore, the top 15 KEGG pathways for the target genes of miR-125b-5p were enriched in human cytomegalovirus infection, endocytosis, neurotrophin signaling pathway, adherens junction, insulin signaling pathway, non-small cell lung cancer, proteoglycans in cancer, autophagy—animal, MAPK signaling pathway, EGFR tyrosine kinase inhibitor resistance, longevity regulating pathway, chronic myeloid leukemia, PI3K/Akt signaling pathway, and breast cancer and acute myeloid leukemia (**Figure 3D**).

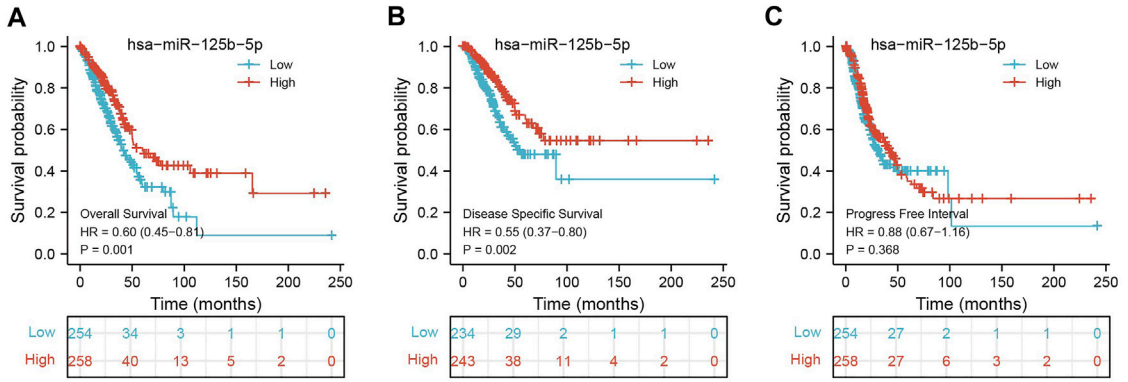



**FIGURE 5 |** Prognostic value of miR-125b-5p in lung adenocarcinoma (LUAD). Prognostic value of miR-125b-5p in LUAD, including overall survival **(A)**, disease-specific survival **(B)**, and progression-free interval **(C)**.

## Analysis of the Correlation Between miR-125b-5p Expression and Clinicopathological Characteristics in LUAD

To examine the expression of miR-125b-5p in LUAD, we analyzed the datasets from TCGA and Genetic Testing Ontology (GTO) and found that miR-125b-5p was decreased in LUAD tissues compared to that in normal tissues (**Figures 4A–C**). Furthermore, a lower expression correlated with poor clinicopathological features, including the primary therapy outcome, TNM stage, age, and residual tumor (**Figures 4D–I**). The ROC curve analysis of miR-125b-5p showed an area under the curve (AUC) value of 0.768 for LUAD patients (**Figure 4J**). Given that the expression of miR-125b-5p correlated with poor clinicopathological features in LUAD, therefore, we next explored the prognostic value of miR-125b-5p in non-small cell lung cancer. The results demonstrated that a lower expression of miR-125b-5p was significantly associated with poor overall survival and disease-specific survival (**Figures 5A–C**).

## Overexpression of miR-125b-5p Inhibits LUAD Progression

To elucidate the biological functions of miR-125b-5p in LUAD cells, we transfected miR-125b-5p mimics and inhibitors in A549 and H1975 cells. The efficiency of miR-125b-5p knockdown and overexpression in A549 and H1975 cells was verified by the qRT-PCR assay (**Figure 6B**). Growth curve and colony formation assays were carried out to evaluate cell proliferation, with the results showing that the overexpression of miR-125b-5p notably inhibited the proliferation of A549 and H1975 cells, while miR-125b-5p knockdown inhibited increased cell growth (**Figures 6C,D**). Additionally, we found that the overexpression of miR-125b-5p promoted cell apoptosis, with miR-125b-5p inhibiting the reduction of cell apoptosis, which was further confirmed by Western blot examining the expression of marker genes critical for cellular apoptosis, such as Bax, Bcl-2, and PARP (**Figure 6F**). Finally, the Transwell and wound healing assay revealed that the overexpression of miR-125b-5p reduced the migration and invasion of LUAD cells, while miR-125b-5p knockdown had the opposite effect (**Figures 7A,B**). Taken together, these data suggest that miR-125b-5p exerts the role of a suppressor gene in LUAD cells.

## Correlation of miR-125b-5p and Immune Cells in LUAD

We used ssGSEA to examine the correlation between miR-125b-5p and immune cell infiltration in LUAD. The results showed that miR-125b-5p was positively correlated with the infiltration levels of mast cells, natural killer (NK) cells, dendritic cells (DCs), plasmacytoid DCs (pDCs), interstitial DCs (iDCs), macrophages, T follicular helper (TFH) cells, eosinophils, B cells, type 1 T helper (Th1) cells, T cells, CD8 T cells, cytotoxic cells, neutrophils, and central memory (TCM) and effector memory (TEM) T cells (**Figures 8A–E**). However, Ted and Th2 cells were all negatively correlated with miR-125b-5p expression (**Figures 8A–E**). These results demonstrate that miR-125b-5p plays a significant role in the immune response of LUAD.

# DISCUSSION

Lung cancer has become the leading cause of cancer-related deaths worldwide, with a rising trend of incidence and mortality. LUAD accounts for 85% of all lung cancers (Reck et al., 2014). It has been shown that miRNAs play a major role in regulating the occurrence and progression of cancer (Lee and Cheah, 2019). For instance, Qu et al. found that miR-133b had a low expression in renal cell carcinoma. The overexpression of miR-133b significantly inhibits cell proliferation, migration, and invasion of renal cell carcinoma. Research on its mechanism showed that miR-133b exerts this role *via* modulating the expression of MMP9 and inhibiting the progression of renal

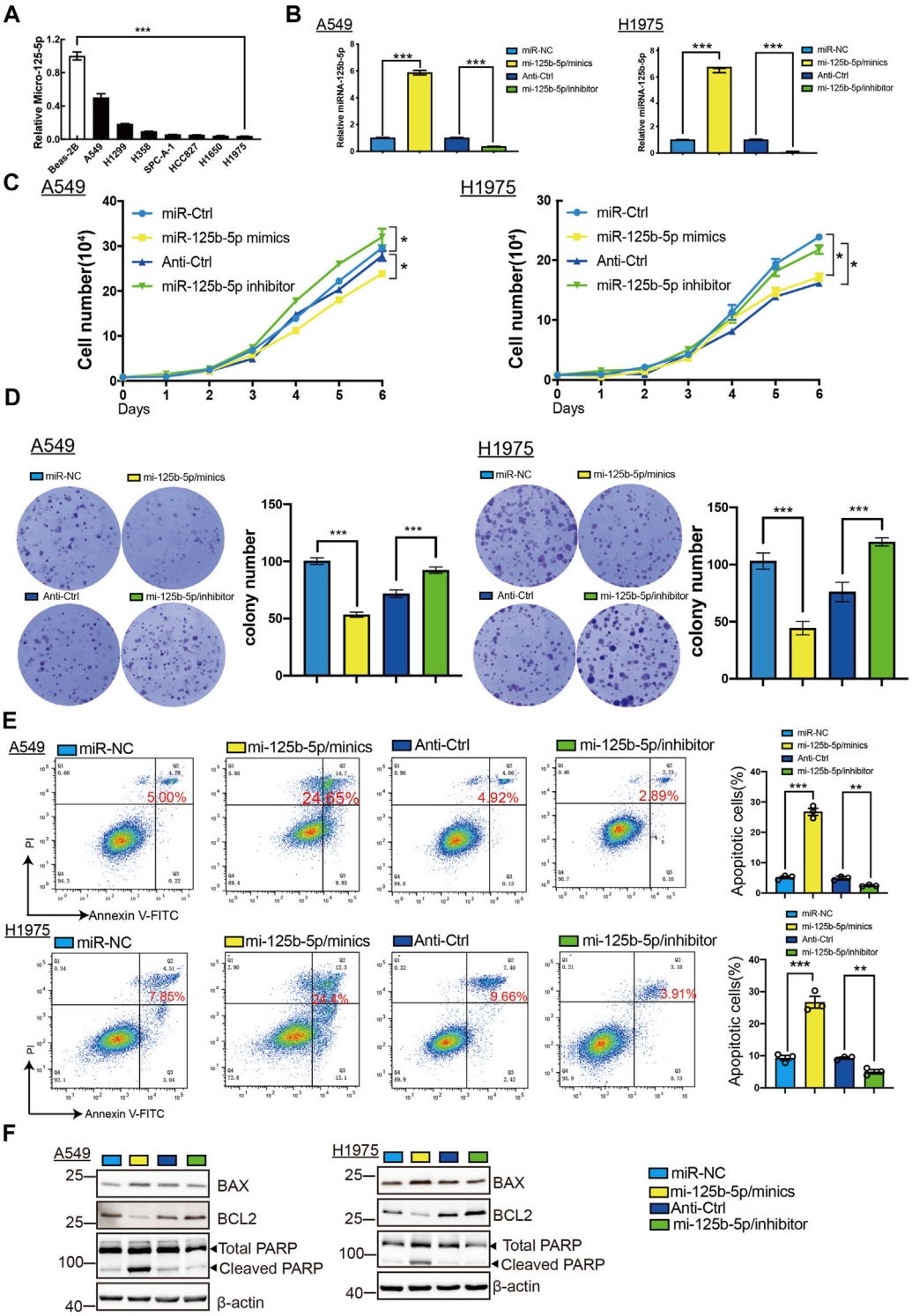



**FIGURE 6 |** Overexpression of miR-125b-5p inhibited lung adenocarcinoma (LUAD) cell proliferation. **(A)** Expression of miR-125b-5p in LUAD cell lines examined by the quantitative real-time PCR (qRT-PCR) assay. **(B)** Expression of miR-125b-5p in LUAD cell lines after its overexpression examined using the qRT-PCR assay. **(C,D)** Overexpression of miR-125b-5p on cell growth ability examined using the growth curve and colony formation assays. Scale bar, 50 μm. **(E,F)** Overexpression of miR-125b-5p on cell apoptosis examined by flow cytometry and Western blot assay. Quantification data are also indicated. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

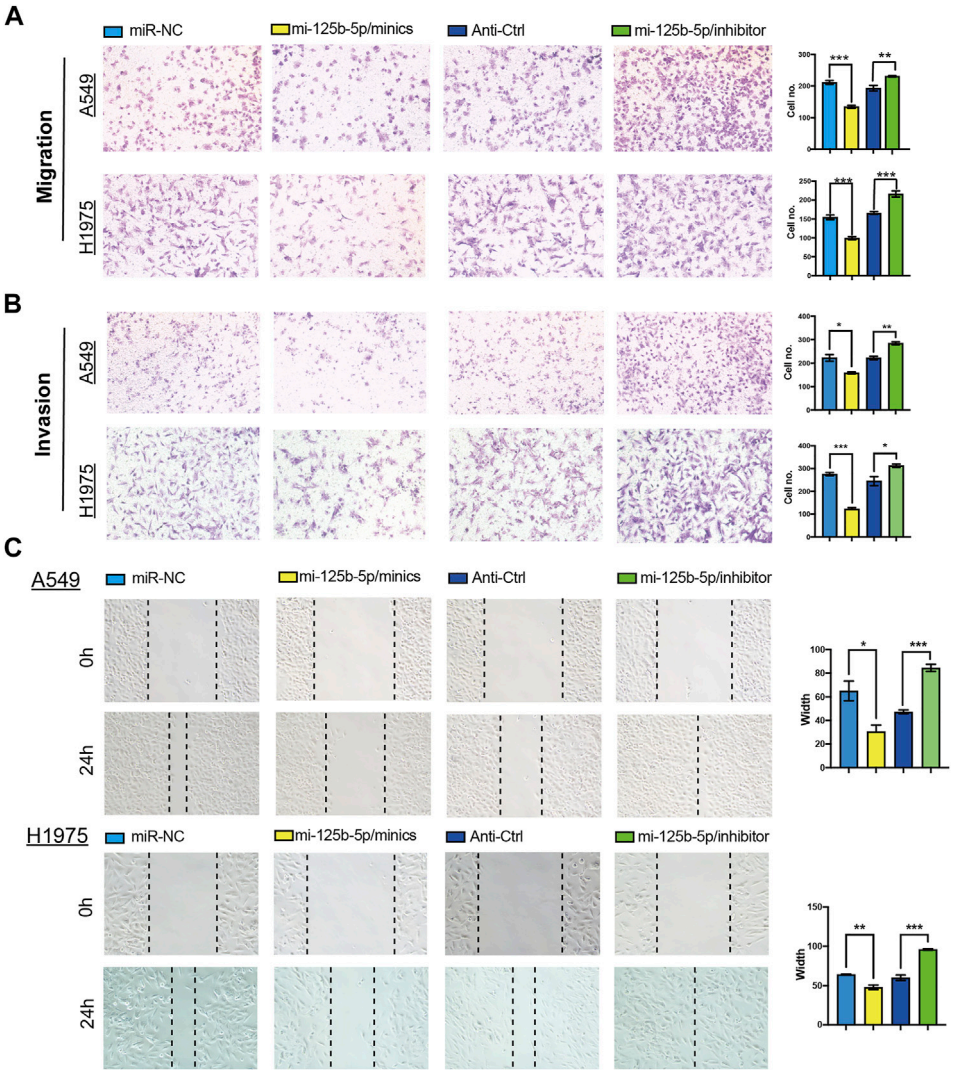



**FIGURE 7 |** Overexpression of miR-125b-5p inhibited lung adenocarcinoma (LUAD) migration and invasion. **(A–C)** Overexpression of miR-125b-5p on cell migration and invasion abilities examined by the Transwell and wound healing assays. Quantification data are also indicated. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

cell carcinoma (Wu et al., 2014). Furthermore, Zhang et al. showed that miR-125b targeted MMP11 and breast cancer progression (Wang et al., 2020). A previous study also showed that miR-125b inhibited bladder cancer cell proliferation and migration by targeting SphK1 (Zhao et al., 2015). However, the clinical significance and functional role of miR-125b-5p in LUAD remain unclear.

In the current study, we demonstrated that the expression levels of miR-125b-5p were significantly downregulated in LUAD cells compared with normal bronchial epithelial cells. The low expression of miR-125b-5p correlated with poor overall survival, disease-specific survival, and poor clinicopathological features, including the primary therapy outcome, TNM stage, age, and residual tumor. The ROC curve analysis of miR-125b-5p showed an AUC value of 0.768 for LUAD patients. These results indicate that miR-125b-5p has great potential as a detection index for the diagnosis of LUAD with high sensitivity and specificity.

For the function of miR-125b-5p in LUAD, we performed KEGG enrichment analysis and found that the target genes of miR-125b-5p were enriched in human cytomegalovirus infection, endocytosis, neurotrophin signaling pathway, adherens junction, insulin signaling pathway, non-small cell lung cancer, proteoglycans in cancer, autophagy animal, MAPK signaling pathway, EGFR tyrosine kinase inhibitor resistance, longevity regulating pathway, chronic myeloid leukemia, PI3K/Akt signaling pathway, and breast cancer and acute myeloid leukemia. These results suggest that miR-125b-5p may play an important role in the progression of LUAD.

Previous studies reported that miR-125b played a suppressor role in diverse cancer progression. For example, miR-125b

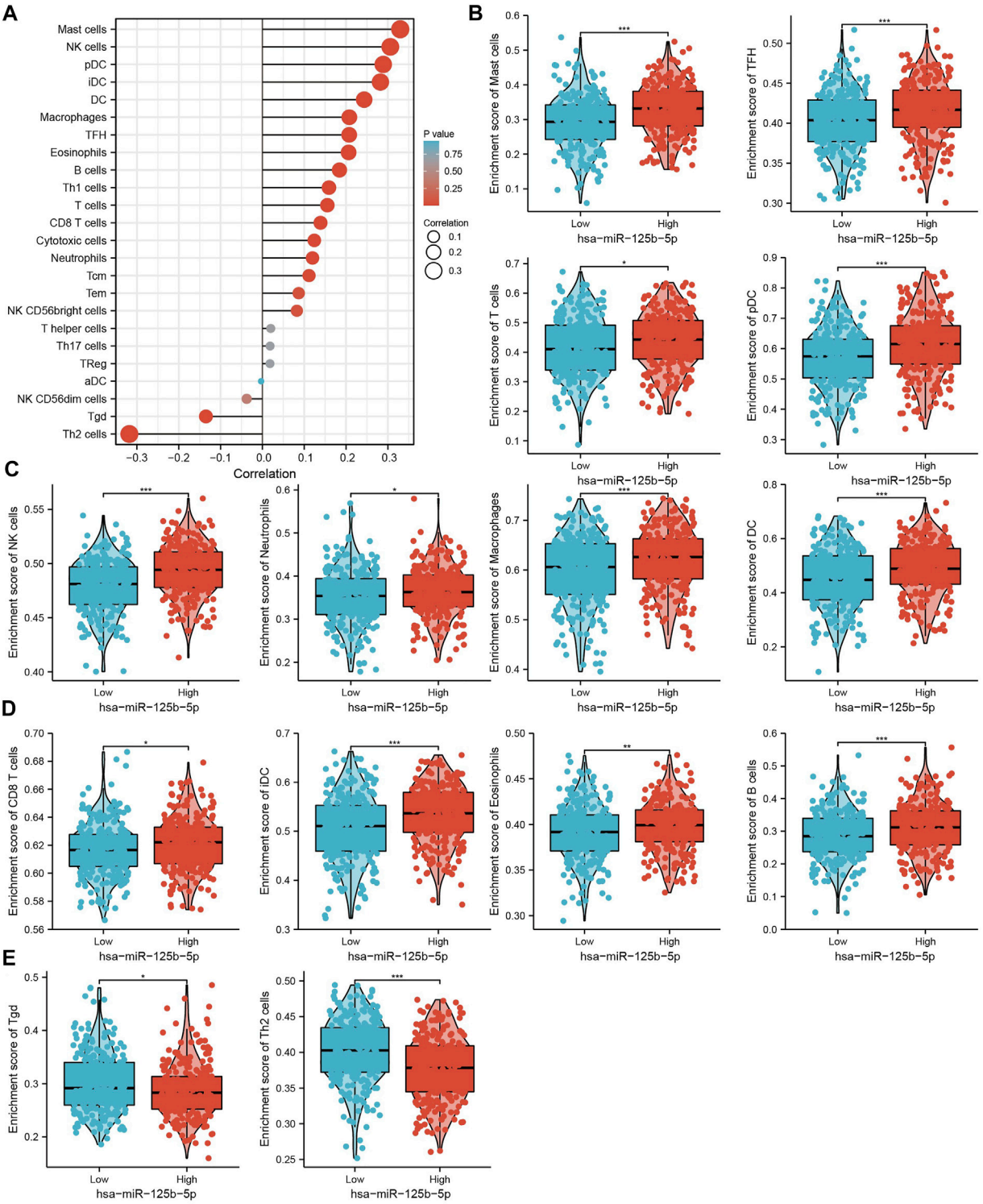



**FIGURE 8 |** Analysis of the correlation between the expression of miR-125b-5p and immune infiltration. **(A)** Correlation between the relative abundance of 24 immune cells and the expression level of miR-125b-5p. **(B–E)** Diverse proportions of immune cell subtypes in tumor samples in high and low miR-125b-5p expression groups. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

regulated HMGA1 and inhibited cervical cancer progression. In this study, we found that miR-125b-5p knockdown promoted the proliferation, migration, and invasion of LUAD cells. Additionally, we found that knockdown of miR-125b-5p inhibited cell apoptosis, but its overexpression showed the opposite effect and promoted cell apoptosis, which was further confirmed by Western blot examining the expression of marker genes critical for cellular apoptosis, such as Bax, Bcl-2, and PARP. These results indicate that miR-125b-5p may inhibit the proliferation, migration, and invasion of LUAD cells. Above all, these results show that miR-125b-5p plays a suppressor role in LUAD cancer progression.

We further studied the immune role of miR-125b-5p in the tumor microenvironment. We analyzed the correlation between the expression of miR-125b-5p and diverse immune cell infiltration in LUAD. The results showed that miR-125b-5p was positively correlated with the infiltration levels of mast cells, NK cells, pDCs, iDCs, DCs, macrophages, TFH cells, eosinophils, B cells, Th1 cells, T cells, CD8 T cells, cytotoxic cells, neutrophils, TCM, and TEM. These results at least partially confirm that miR-125b-5p plays a crucial role in immune response regulation.

## CONCLUSION

In summary, we demonstrated that miR-125b-5p plays a significant role in the progression of LUAD. This study has thus provided a new perspective in the research on miRNAs in LUAD, as well as a scientific basis for novel developments in the diagnosis and treatment of LUAD.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

LT, YY, HZ, and JW designed this work, performed related assays, and analyzed the data. DZ, HL, and YS contributed to the study materials. LD and XJ supervised and wrote the manuscript. All authors have read and approved the final version of the manuscript.

## FUNDING

This work was supported by the National Nature Science Foundation of China (82160508), Yunnan Applied Basic Research Projects (YNWRMY-2019-067), Yunnan Province Specialized Training Grant for High-Level Healthcare Professionals (D-201614), and Yunnan Province Applied Basic Research Foundation (2019FE001) to LD.

## ACKNOWLEDGMENTS

The authors would like to thank support from the Department of Thoracic Surgery and Department of Pathology at The Third Affiliated Hospital of Kunming Medical University.

## REFERENCES

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

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics* 14, 7.

Iqbal, M. A., Arora, S., Prakasam, G., Calin, G. A., and Syed, M. A. (2019). MicroRNA in Lung Cancer: Role, Mechanisms, Pathways and Therapeutic Relevance. *Mol. Aspects Med.* 70, 3–20. doi:10.1016/j.mam.2018.07.003

Jiang, L.-P., Fan, S.-Q., Xiong, Q.-X., Zhou, Y.-C., Yang, Z.-Z., Li, G.-F., et al. (2018). GRK5 Functions as an Oncogenic Factor in Non-small-cell Lung Cancer. *Cell Death Dis* 9 (3), 295. doi:10.1038/s41419-018-0299-1

Lee, S. S., and Cheah, Y. K. (2019). The Interplay between MicroRNAs and Cellular Components of Tumour Microenvironment (TME) on Non-small-cell Lung Cancer (NSCLC) Progression. *J. Immunol. Res.* 2019, 3046379. doi:10.1155/2019/3046379

Liang, R., Li, X., Li, W., Zhu, X., and Li, C. (2021). DNA Methylation in Lung Cancer Patients: Opening a "window of Life" under Precision Medicine. *Biomed. Pharmacother.* 144, 112202. doi:10.1016/j.biopha.2021.112202

Nagy, Á., Munkácsy, G., and Győrffy, B. (2021). Pancancer Survival Analysis of Cancer Hallmark Genes. *Sci. Rep.* 11 (1), 6047. doi:10.1038/s41598-021-84787-5

Reck, M., Popat, S., Reinmuth, N., De Ruysscher, D., Kerr, K. M., and Peters, S. (2014). Metastatic Non-small-cell Lung Cancer (NSCLC): ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-Up. *Ann. Oncol.* 25 (Suppl. 3), iii27–iii39. doi:10.1093/annonc/mdu199

Rivera, G. A., and Wakelee, H. (2016). Lung Cancer in Never Smokers. *Adv. Exp. Med. Biol.* 893, 43–57. doi:10.1007/978-3-319-24223-1_3

Schwartz, A. G., and Cote, M. L. (2016). Epidemiology of Lung Cancer. *Adv. Exp. Med. Biol.* 893, 21–41. doi:10.1007/978-3-319-24223-1_2

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

Sun, B., Zhang, Y., Zhou, L., Yin, L., Li, F., Li, C., et al. (2019). The Proliferation of Cervical Cancer Is Promoted by miRNA-125b through the Regulation of the HMGA1. *Ott* Vol. 12, 2767–2776. doi:10.2147/ott.s197740

Wang, Y., Wei, Y., Fan, X., Zhang, P., Wang, P., Cheng, S., et al. (2020). MicroRNA-125b as a Tumor Suppressor by targetingMMP11in Breast Cancer. *Thorac. Cancer* 11 (6), 1613–1620. doi:10.1111/1759-7714.13441

Wu, D., Pan, H., Zhou, Y., Zhou, J., Fan, Y., and Qu, P. (2014). microRNA-133b Downregulation and Inhibition of Cell Proliferation, Migration and Invasion by Targeting Matrix Metallopeptidase-9 in Renal Cell Carcinoma. *Mol. Med. Rep.* 9 (6), 2491–2498. doi:10.3892/mmr.2014.2116

Xiong, Q., Jiang, L., Liu, K., Jiang, X., Liu, B., Shi, Y., et al. (2021). miR-133b Targets NCAPH to Promote β-catenin Degradation and Reduce Cancer Stem Cell Maintenance in Non-small Cell Lung Cancer. *Sig Transduct Target. Ther.* 6 (1), 252. doi:10.1038/s41392-021-00555-x

Yagi, T., Iinuma, H., Hayama, T., Matsuda, K., Nozawa, K., Tsukamoto, M., et al. (2019). Plasma Exosomal microRNA-125b as a Monitoring Biomarker of Resistance to mFOLFOX6-Based Chemotherapy in Advanced and Recurrent Colorectal Cancer Patients. *Mol. Clin. Oncol.* 11 (4), 416–424. doi:10.3892/mco.2019.1911

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

Zedan, A. H., Hansen, T. F., Assenholt, J., Madsen, J. S., and Osther, P. J. S. (2019). Circulating miRNAs in Localized/locally Advanced Prostate Cancer Patients after Radical Prostatectomy and Radiotherapy. *Prostate* 79 (4), 425–432. doi:10.1002/pros.23748

Zhao, X., He, W., Li, J., Huang, S., Wan, X., Luo, H., et al. (2015). MiRNA-125b Inhibits Proliferation and Migration by Targeting SphK1 in Bladder Cancer. *Am. J. Transl Res.* 7 (11), 2346–2354.

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

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

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 18 February 2022
doi: 10.3389/fmolb.2022.833771

# A Neutrophil Extracellular Traps Signature Predicts the Clinical Outcomes and Immunotherapy Response in Head and Neck Squamous Cell Carcinoma

*Naifei Chen*[1†], *Dongsheng He*[2†] *and Jiuwei Cui*[1*]

[1]*Cancer Center, The First Hospital of Jilin University, Changchun, China,* [2]*Department of Medical Oncology, The First Hospital of Putian, Teaching Hospital, Fujian Medical University, Putian, China*

OPEN ACCESS

**Edited by:**
*Liang Wang,*
*Moffitt Cancer Center and Research Institute, United States*

**Reviewed by:**
*Srinivasa Reddy Bonam,*
*Institut National de la Santé et de la Recherche Médicale (INSERM), France*
*Udayan Bhattacharya,*
*NewYork-Presbyterian, United States*

***Correspondence:**
*Jiuwei Cui*
*cuijw@jlu.edu.cn*

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

**Specialty section:**
*This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences*

**Received:** *12 December 2021*
**Accepted:** *04 February 2022*
**Published:** *18 February 2022*

**Citation:**
*Chen N, He D and Cui J (2022) A Neutrophil Extracellular Traps Signature Predicts the Clinical Outcomes and Immunotherapy Response in Head and Neck Squamous Cell Carcinoma. Front. Mol. Biosci. 9:833771. doi: 10.3389/fmolb.2022.833771*

**Background:** Neutrophil extracellular traps (NETs) play an important role in the occurrence, metastasis and immune escape of cancers. This study aimed to investigate NET-related genes, their clinical prognostic value and their correlation with immunotherapy and anticancer drugs in patients with head and neck squamous cell carcinoma (HNSCC).

**Methods:** Differentially expressed NET-related genes in HNSCC were identified based on multiple public databases. To improve the clinical practicability and avoid overfitting, univariable, least absolute shrinkage and selection operator (LASSO) and multivariable Cox algorithms were used to construct a prognostic risk model. A nomogram was further used to explore the clinical value of the model. Internal and external validation were conducted to test the model. Furthermore, the immune microenvironment, immunophenoscore (IPS) and sensitivity to anticancer drugs in HNSCC patients with different prognostic risks were explored.

**Results:** Six NET-related genes were screened to construct the risk model. In the training cohort, Kaplan–Meier (K-M) analysis showed that the overall survival (OS) of low-risk HNSCC patients was significantly better than that of high-risk HNSCC patients ($p < 0.001$). The nomogram also showed a promising prognostic value with a better C-index (0.726 vs 0.640) and area under the curve (AUC) (0.743 vs 0.706 at 3 years, 0.743 vs 0.645 at 5 years) than those in previous studies. Calibration plots and decision curve analysis (DCA) also showed the satisfactory predictive capacity of the nomogram. Internal and external validation further strengthened the credibility of the clinical prognostic model. The level of tumor mutational burden (TMB) in the high-risk group was significantly higher than that in the low-risk group ($p = 0.017$), and the TMB was positively correlated with the risk score ($R = 0.11$; $p = 0.019$). Moreover, the difference in immune infiltration was significant in HNSCC patients with different risks ($p < 0.05$). Furthermore, the IPS analysis indicated that anti-PD-1 ($p < 0.001$), anti-CTLA4 ($p < 0.001$) or combining immunotherapies ($p < 0.001$) were more beneficial for low-risk HNSCC patients. The response to anticancer drugs was also closely correlated with the expression of NET-related genes ($p < 0.001$).

**Conclusion:** This study identified a novel prognostic model that might be beneficial to develop personalized treatment for HNSCC patients.

**Keywords: HNSCC, nets, prognosis, immune microenvironment, immunophenoscore, anticancer drugs**

## INTRODUCTION

Head and neck cancer, the sixth most common malignancies worldwide, leads to unacceptable mortality, with more than 450,000 deaths reported in 2020 (International Agency for Research on Cancer and World Health Organization, 2021). As the most common primary component of head and neck cancer, the main risk factors for head and neck squamous cell carcinoma (HNSCC) are smoking, chronic exposure to alcohol, different forms of chewing tobacco (such as betel palm) and HPV infection (Gillison et al., 2015). Currently, the overall survival (OS) of HNSCC patients is not satisfactory even in the context of surgery, radiotherapy, chemotherapy, targeted therapy and immunotherapy. Increasing amounts of evidence highlights that distant failure is 20–30% in locoregionally advanced HNSCC patients, while the percentage of locoregional failure is approximately 40–50%. HNSCC has become a serious global public health problem (Samra et al., 2018; Muzaffar et al., 2021). Therefore, exploring novel therapeutic targets and developing a novel prognostic model to improve personalized treatment is urgent.

Neutrophils are the most abundant component of circulating immune cells and they play an irreplaceable role in the response against pathogens. The immune cells can release neutrophil extracellular traps (NETs) under *in vitro* stimulation or pathological conditions (Kolaczkowska and Kubes, 2013; Li et al., 2020; Németh et al., 2020). NETs are extracellular web-like structures consisting of mitochondrial and nuclear DNA fibers decorated with histones and granular antimicrobial enzymes, which have been recently considered a host defense mechanism to entrap, constrain and kill invasive bacteria and other pathogens (Pruchniak and Demkow, 2019). The process of classical NETs formation is termed "neutrophil extracellular traposis (NETosis)", which has been defined as a unique form of regulated cell death distinguished from other programmed cell deaths, such as necroptosis, autophagy and apoptosis (Manda-Handzlik et al., 2020). NETs play a crucial role in the pathogenesis of various diseases, such as COVID-19, cystic fibrosis, small vessel vasculitis and cancers (Demkow, 2021).

Although NETs might theoretically exert antitumorigenic effects by trapping and killing cancer cells, increasing amounts of evidence indicates that NETs might exert pro-tumorigenic effects. NETs can be triggered by cancer cells and cancer-associated fibroblasts, which directly or indirectly facilitate the proliferation and metastasis of cancer cells by hijacking the antimicrobial immune system (Park et al., 2016). It has been suggested that NETs are related to the recurrence of abdominal cancers and a poor prognosis of patients with colorectal cancer and Ewing sarcoma (Berger-Achituv et al., 2013; Richardson et al., 2017; Kanamaru et al., 2018). Moreover, NETs have also been reported to be enriched in the liver metastases of patients with colon and breast cancers, and serum NETs might predict the occurrence of liver metastases in patients with early-stage breast cancer (Yang et al., 2020).

Epithelial mesenchymal transition (EMT) arms cancer cells with motility and invasiveness, and NETs can induce this transition. The web-like structure might increase the adhesive ability of cancer cells and awaken dormant cancer cells (Ireland and Oliver, 2020; Teijeira et al., 2020; Yang and Liu, 2021). Furthermore, NETs enhance the immune escape of cancer cells. It has been suggested that the NETs might to reduce the curative effect of immune checkpoint inhibitors (ICIs) and chimeric antigen receptor T-cell (CAR-T) immunotherapy (Teijeira et al., 2020; Volkov et al., 2021). In summary, NETs have become a new field of investigation in oncology. Nevertheless, NET-related genes, their prognostic value and their relationship with immunotherapy in HNSCC remain largely unknown.

In the present study, we first constructed a prognostic risk model related to NETs in HNSCC based on multiple public databases, and internal and external validations were used to assess the accuracy of the risk model. Furthermore, the clinical value, immune microenvironment and drug sensitivity based on the model were investigated. To the best of our knowledge, there are no previous studies exploring NET-related genes, their prognostic value or their relationship with the immune microenvironment in HNSCC. This study found that NET-related genes might be potential prognostic markers and therapeutic targets in HNSCC patients and could be used to further improve the efficacy of treatment in patients with HNSCC through personalized treatment.

## MATERIALS AND METHODS

### Data Acquisition

The mRNA sequencing data (FPKM) of HNSCC patients (44 head and neck normal samples and 502 HNSCC samples) in The Cancer Genome Atlas (TCGA) database and mRNA sequencing data (FPKM) of 55 head and neck normal samples (salivary gland) in the Genotype-Tissue Expression (GTEx) database were obtained from UCSC Xena (https://gtexportal.org/home/). The clinical data of the HNSCC patients were downloaded from the TCGA database. A total of 170 NET-related genes were obtained from previous studies and are shown in **Supplementary Table S1** (Dwyer et al., 2014; Papayannopoulos, 2018).

### Data Processing

The sequencing data in TCGA and GETx were transformed with log2 (FPKM+1) and normalized with the limma R package, and 99 normal head and neck samples and 502 HNSCC samples were integrated (Ritchie et al., 2015). The 170 NET-related genes were

matched with the mRNA sequencing data, and the limma R package was used to screen out the differentially expressed NETs genes (DEGs) with log2 (fold change) > 1 and adjusted $p < 0.05$ (Wu et al., 2020).

## Construction and Assessment of the Prognostic Risk Score Model

The expression of NET-related genes in HNSCC patients were integrated with the corresponding survival data. The patients with HNSCC in the TCGA database were randomly separated into training and testing groups at a ratio of 8:2. In the training group ($n$ = 400), the NET-related genes were included in a univariate Cox regression analysis ($p < 0.1$) to identify candidate genes (Jiao et al., 2021). Then, differentially expressed candidate genes -related NETs were identified by intersecting the DEGs and candidate genes. The least absolute shrinkage and selection operator (LASSO) Cox regression algorithm was applied to avoid overfitting. Furthermore, multivariable Cox regression analysis was used to compute the coefficient of the prognostic risk score model. The HNSCC patients in the training group were divided into high- and low-risk cohorts based on the median risk score. The Kaplan-Meier method, receiver operating characteristic (ROC) curve, distribution of risk score and survival status were used to evaluate the efficiency of the prognostic risk model. Moreover, gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) analysis of the screened genes was performed by employing the "clusterProfiler" R package (Liu et al., 2021).

## Internal and External Validation of the Multigene Prognostic Model

To further evaluate the value of the prognostic risk model of HNSCC patients, a testing cohort ($n$ = 99) and a whole cohort ($n$ = 499) in the TCGA database were used to conduct internal validation. The mRNA sequencing and clinical data of HNSCC patients ($n$ = 108) from the E-MTAB-8588 dataset in the ArrayExpress database were used to conduct an external validation, and the "sva" R package was used to diminish the batch effect of different datasets (Leek et al., 2012). Internal and external validation was conducted based on the medium risk score in the training cohort.

## Construction and Assessment of the Nomogram for Patients With Head and Neck Squamous Cell Carcinoma

To further weight the possibility of the risk score based on genes related to NETs being an independent prognostic parameter, the risk score of HNSCC patients was integrated with the corresponding clinical parameters (age, sex, grade, stage, margin status, chemotherapy, and radiotherapy) in the training cohort. The parameters were included in a univariate Cox regression algorithm to screen out the characteristics correlated with overall survival (OS) in HNSCC patients with a $p$ value less than 0.05. A multivariable Cox regression algorithm was further used to identify independent prognostic parameters. Then, a nomogram was constructed based on the parameters. To weigh the capability of the clinical prognostic model to forecast an individual's OS, the concordance index (C-index), ROC curve, calibration plot and decision curve analysis (DCA) were calculated. The nomogram was further validated for testing the overall and external cohorts for internal and external validation.

## Relationship Between Immune Infiltration and Prognostic Risk Score in Patients With Head and Neck Squamous Cell Carcinoma

The single nucleotide variant (SNV) data in the TCGA database were downloaded to calculate the tumor mutational burden (TMB) for each HNSCC patient. Spearman's algorithm was used to analyze the correlation between the risk score and TMB. Furthermore, the "Cell Type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT)" deconvolution algorithm with 1,000 permutations was applied to quantify 22 types of tumor-infiltrating lymphocytes (TILs) in the microenvironment of low- and high-risk patients with HNSCC with a $p$ value less than 0.05 (Becht et al., 2016). Moreover, the data of the immunophenoscore (IPS) of the HNSCC patients were downloaded from The Cancer Immunome Atlas (TCIA, https://tcia.at/), and the immunotherapy response for anti-PD-1 and anti-CTLA4 in low- and high-risk patients with HNSCC were further investigated (Charoentong et al., 2017).

## Exploration of Drug Sensitivity Based on the Prognostic Model

To explore the anticancer drugs targeted to the NET-related genes that constructed the prognostic model, the sensitivity data of anticancer drugs approved by the United States Food and Drug Administration were downloaded from the CellMiner database (https://discover.nci.nih.gov/cellminer/) (Foy et al., 2017). Pearson analysis was used to explore the correlation between anticancer drug sensitivity and NET-related genes to construct the prognostic risk model (Reinhold et al., 2019).

# RESULTS

## Identification of Differentially Expressed Genes Related to Neutrophil Extracellular Traps in Head and Neck Squamous Cell Carcinoma

The gene expression of NETs in HNSCC was filtered by matching mRNA sequencing data of HNSCC and 170 NETs. Then, a total of 31 differentially expressed genes (19 upregulated and 12 downregulated) that were related to NETs were screened out in HNSCC by using the limma R package (**Figure 1A**).

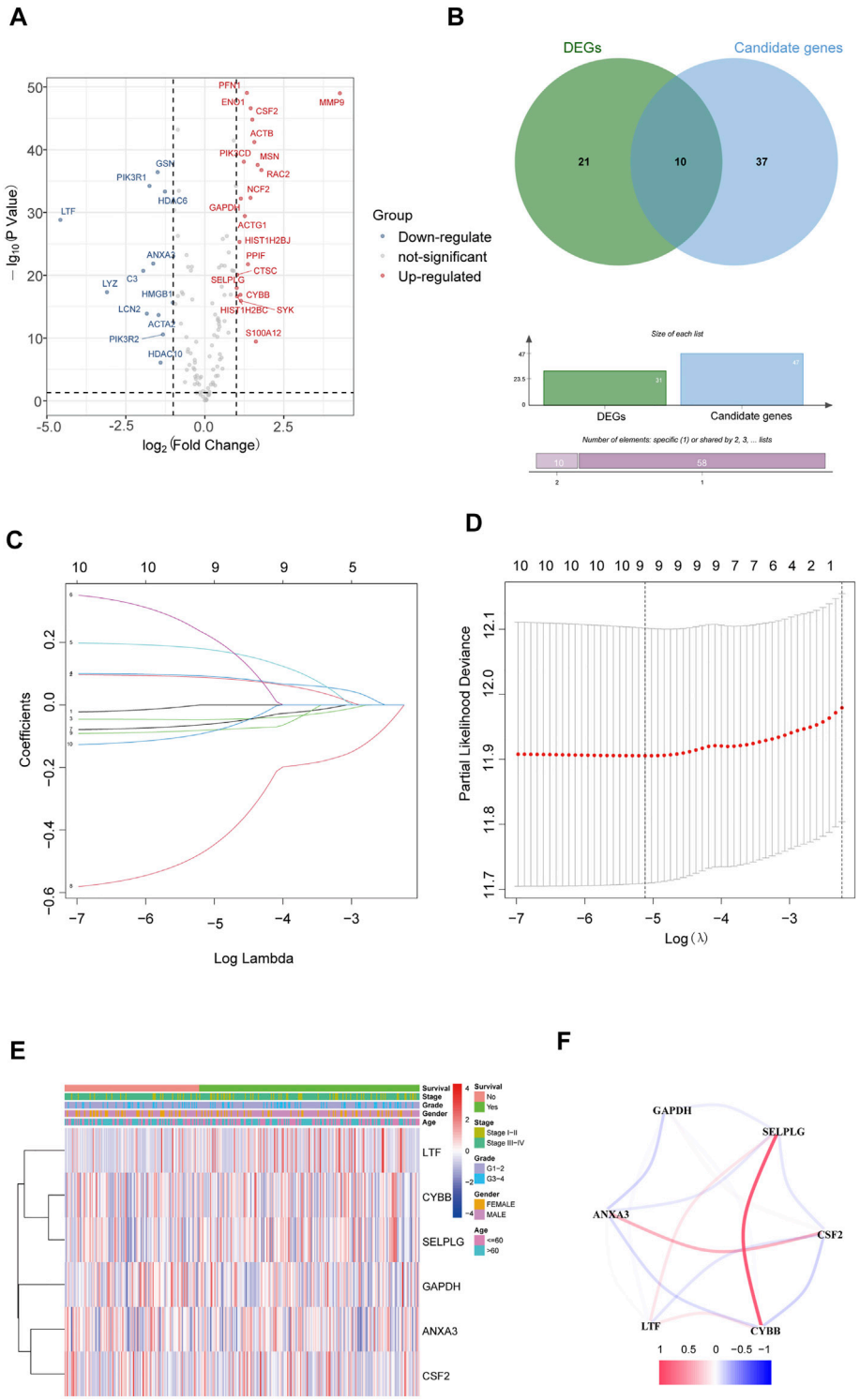



**FIGURE 1 |** Identification of NET-related signatures in HNSC**C. (A)** Differentially expressed genes related NETs in HNSCC. **(B)** Differentially expressed candidate genes related NETs in HNSCC. **(C)** LASSO coefficient profiles of 9 genes related NETs. **(D)** Cross-validation for tuning parameter selection in the proportional hazards model. **(E)** The heatmap of the 6 genes related NETs. **(F)** The correlation of the 6 genes related NETs.

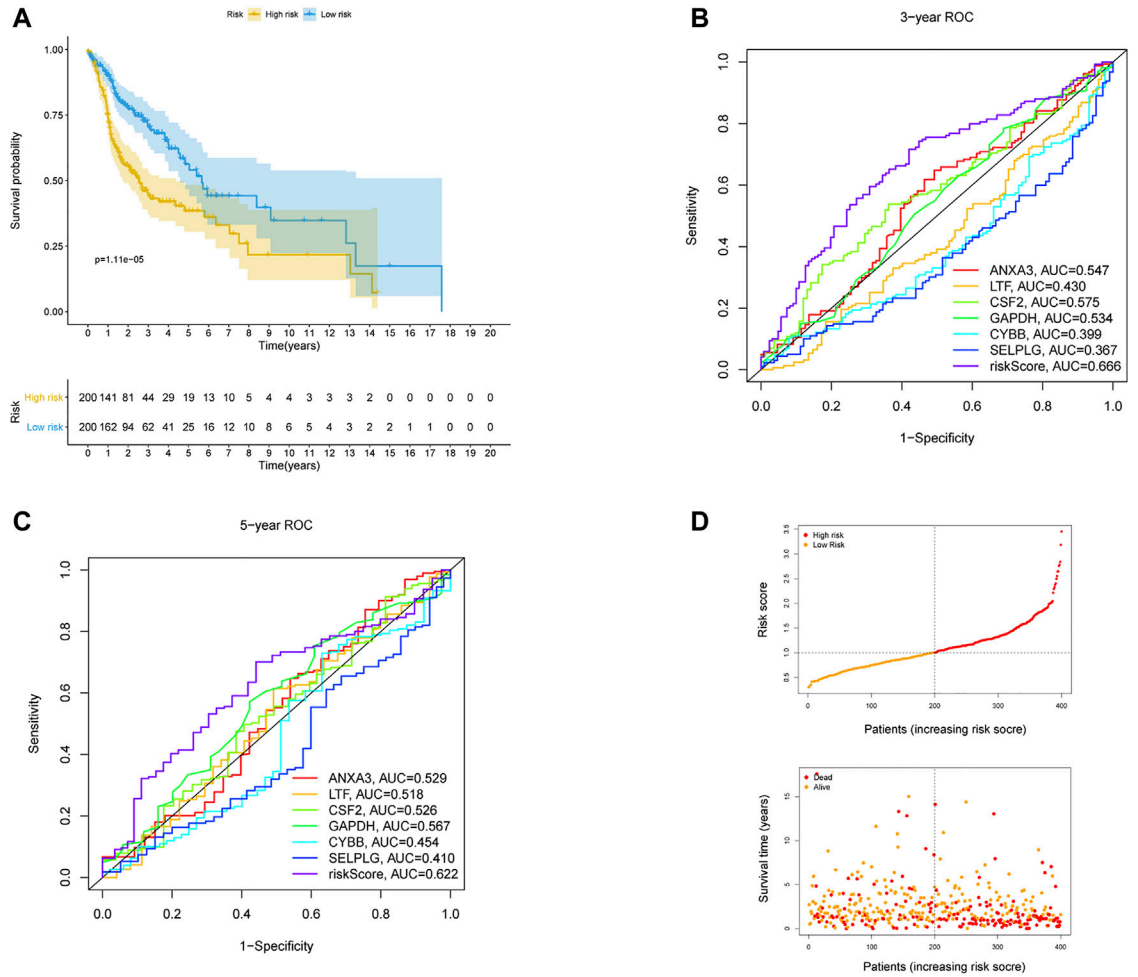



**FIGURE 2 |** Validation of the risk prognostic model in training cohort. **(A)** K-M survival analysis of the model. **(B,C)** ROC curves analysis of the model at 3 and 5 years. **(D)** Distribution of risk score and survival status based on the prognostic model.

## Construction and Assessment of a Prognostic Model Based on Gene-Related Neutrophil Extracellular Traps in Head and Neck Squamous Cell Carcinoma

The 170 NET-related genes was matched with the expression and clinical data of HNSCC patients. To construct a convincing prognostic risk model, HNSCC patients were randomly separated into training (n = 400) and testing cohorts (n = 99) at a ratio of 8:2. In the training cohort, a univariate Cox regression algorithm was used to investigate the candidate genes related to NETs and 10 differentially expressed candidate genes were filtered (**Figure 1B**). The 10 genes were included in the LASSO Cox regression algorithm to avoid overfitting (**Figure 1C**), and cross validation was conducted, which filtered out 9 prognostic genes (**Figure 1D**). To further improve the clinical practicability, 6 gene signatures were purified by multivariable Cox regression analysis and used to build a prognostic risk model (**Figure 1E**). The risk score was calculated as follows: Risk score= (0.116 × ANXA3 expression level) + (-0.062 × LTF expression level) + (0.120 × CSF2 expression level) + (0.184 × GAPDH expression level) + (CYBB × 0.306 expression level) + (-0.604 × SELPLG expression level). The correlations of the 6 genes are shown in **Figure 1F**. Moreover, the credibility of the model was assessed. In the K-M analysis, the OS between patients with high and low risk was significantly different ($p < 0.001$) (**Figure 2A**). The area under the curve (AUC) values for the ROC curves at 3 and 5 years were 0.666 and 0.622, respectively (**Figures 2B,C**). Furthermore, the gene function and potential pathways of the 6 genes are displayed in **Supplementary Figure S1**. The most highly enriched function of the 6 screened genes were neutrophil degranulation and neutrophil activation involved in immune response. KEGG pathway analysis suggested that the 6 genes might participate in NETs formation or HIF−1 signaling pathway. The HIF-1 has been reported to regulate neutrophil from anti-tumor N1 to pro-tumor N2, the latter might promote the NETs formation (Guglietta et al., 2016; Triner and Shah, 2019).

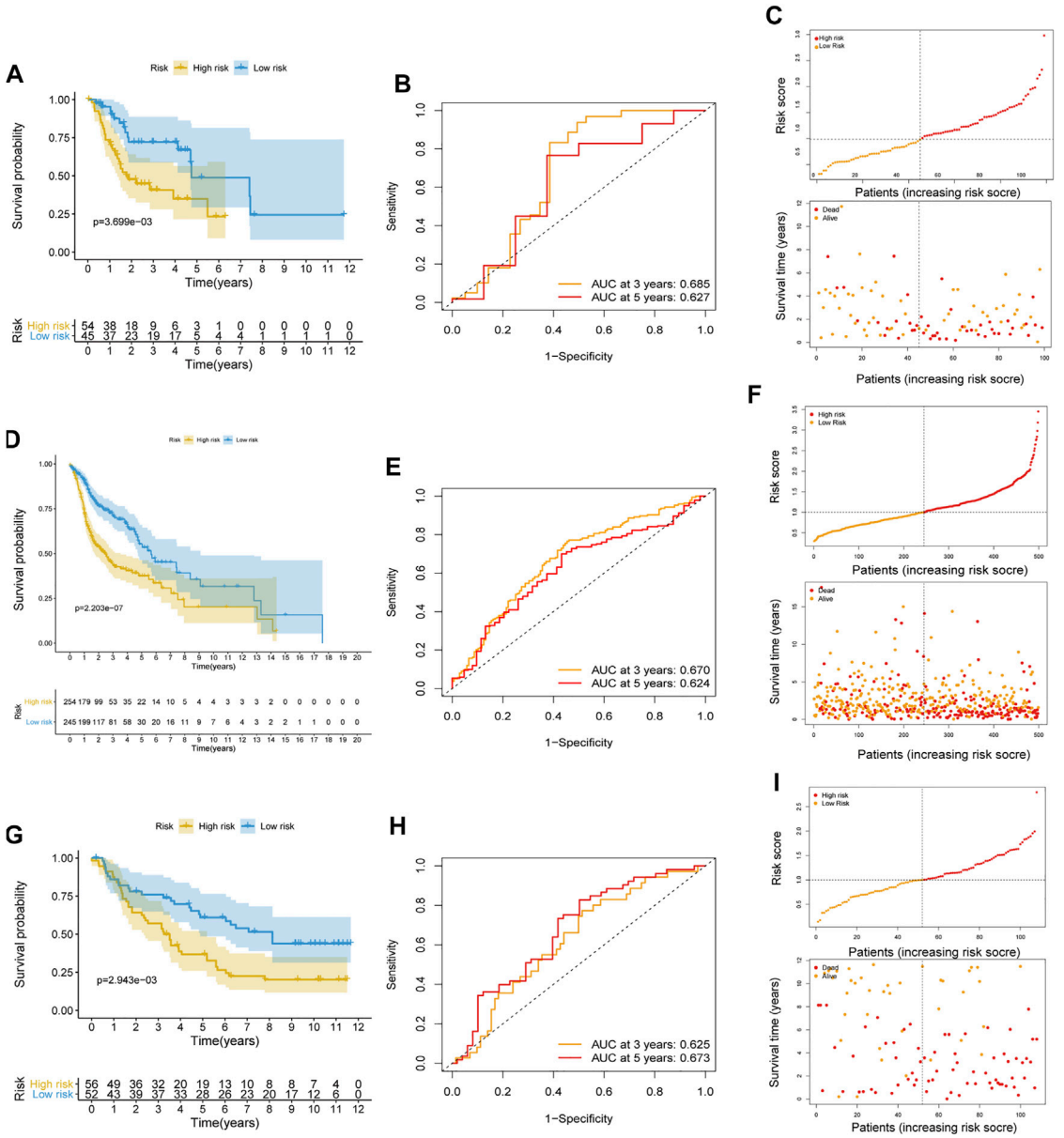



**FIGURE 3 |** Internal and external validation of the risk prognostic model. **(A–C)** K-M analysis, ROC curves, distribution of risk score and survival status in testing cohort. **(D–F)** K-M analysis, ROC curves, distribution of risk score and survival status in entire cohort. **(G–I)** K-M analysis, ROC curves, distribution of risk score and survival status in external validated cohort.

## Internal and External Validation of the Prognostic Risk Model

To further weight the predictive capability of the model in the training cohort, the testing cohort and entire cohort were used for internal validation, and the E-MTAB-8588 dataset was used for external validation. The HNSCC patients were separated into high- (n = 54) and low- (n = 45) risk groups in the testing cohort based on the medium risk score in the training cohort. The difference in OS between the two groups was significant ($p < 0.01$) (**Figure 3A**), and the AUC values at 3 and 5 years were 0.685 and 0.627, respectively (**Figures 3B,C**). In the entire cohort, the difference in OS between the high-risk group (n = 254) and the low-risk group (n = 245) was significant ($p < 0.001$) (**Figure 3D**). In addition, the AUCs at 3 and 5 years were 0.670 and 0.624, respectively (**Figures 3E,F**). Similar to the internal validation, external validation also showed acceptable K-M analysis ($p < 0.01$) (**Figure 3G**) and AUCs (0.625 at 3 years; 0.673 at 5 years) that increased year by year (**Figures 3H,I**).

## Construction and Assessment of a Clinical Prognostic Model

To assess the possibility of the risk score being an independent prognostic factor, the risk score of HNSCC patients in the training

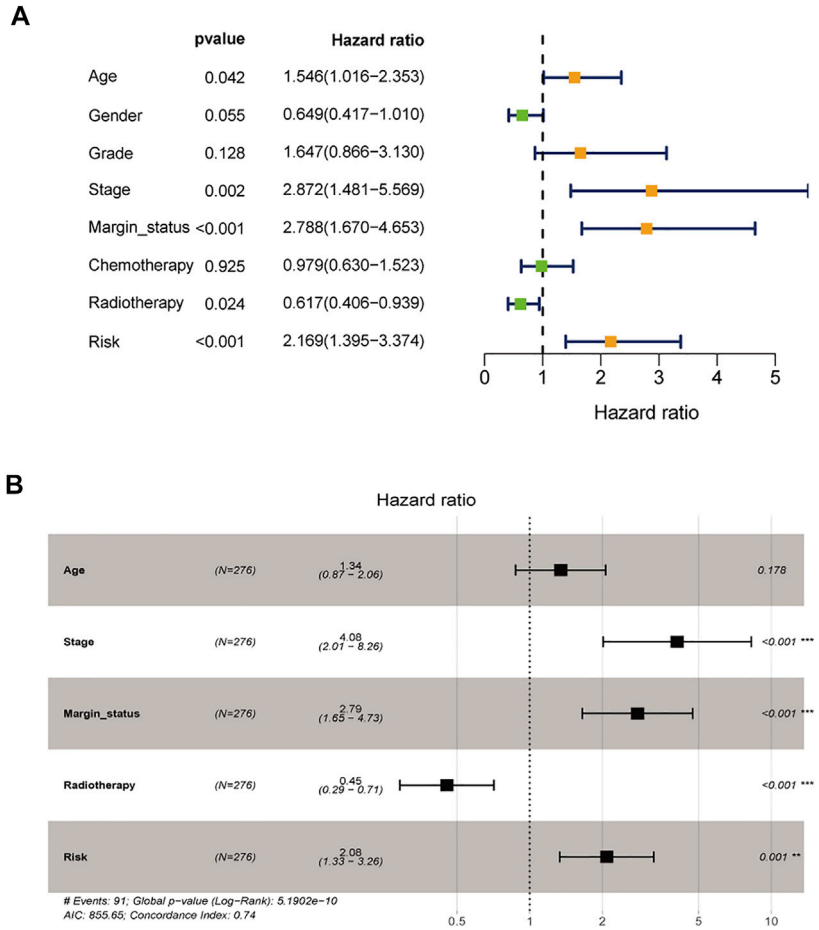



**FIGURE 4 |** Identification of independent prognostic parameters in HNSCC. **(A)** The univariate Cox regression analysis of clinical parameters. **(B)** The multivariate Cox regression analysis of clinical parameters.

cohort was matched with the clinical parameters (age, sex, grade, stage, margin status, radiotherapy and risk score), and 276 individuals were included in the univariate Cox regression analysis. Age, stage, margin status, radiotherapy and risk score were correlated with the prognosis of the HNSCC patients (**Figure 4A**). A multivariable Cox algorithm was used to further screen out four independent prognostic parameters of the HNSCC patients (stage, margin status, radiotherapy and risk score) with a $p$ value less than 0.001 (**Figure 4B**). Based on the four parameters, a nomogram with an optimistic C-index (0.726) was built to predict an individual's prognosis at 3 and 5 years (**Figure 5A**). ROC curve analysis showed satisfactory AUC values at 3 and 5 years (0.743 and 0.743, respectively) (**Figures 5B,C**). The actual curve was also close to the ideal curve in the calibration plot at 3 and 5 years (**Figures 5D,E**). Moreover, DCA also further confirmed the creditability of the prognostic accuracy of the nomogram (**Figures 5F,G**).

## Internal and External Validation of the Clinical Prognostic Model

To further explore the credibility of the nomogram in the training cohort, the risk score in the testing, entire cohort and E-MTAB-8588 dataset were matched with the corresponding clinical parameters for internal and external validation. In the testing cohort, the C-index was 0.667, and the AUC values at 3 and 5 years were 0.628 and 0.660, respectively, increasing year by year (**Figure 6A**). In the entire cohort, the C-index was 0.686, and the AUC values at 3 and 5 years were 0.715 and 0.719, respectively, increasing year by year (**Figure 6B**). In the externally validated cohort, the C-index was 0.590 when the margin status data were missing, and the AUC values at 3 and 5 years were 0.602 and 0.658, respectively, increasing year by year (**Figure 6C**). The internal and external validations together suggested the reliability of the nomogram in the training cohort. Moreover, the risk score is significantly correlated with T stage, perineural invasion, extracapsular spread and different cancer statue and subtype ($p < 0.05$) (**Supplementary Figure S2**).

## Relationship Between Immune Infiltration and Prognostic Risk Score in Patients With Head and Neck Squamous Cell Carcinoma

The TMB has been indicated to be a market for the curative effect of immune checkpoint inhibitors. The TMB of the high- and low-

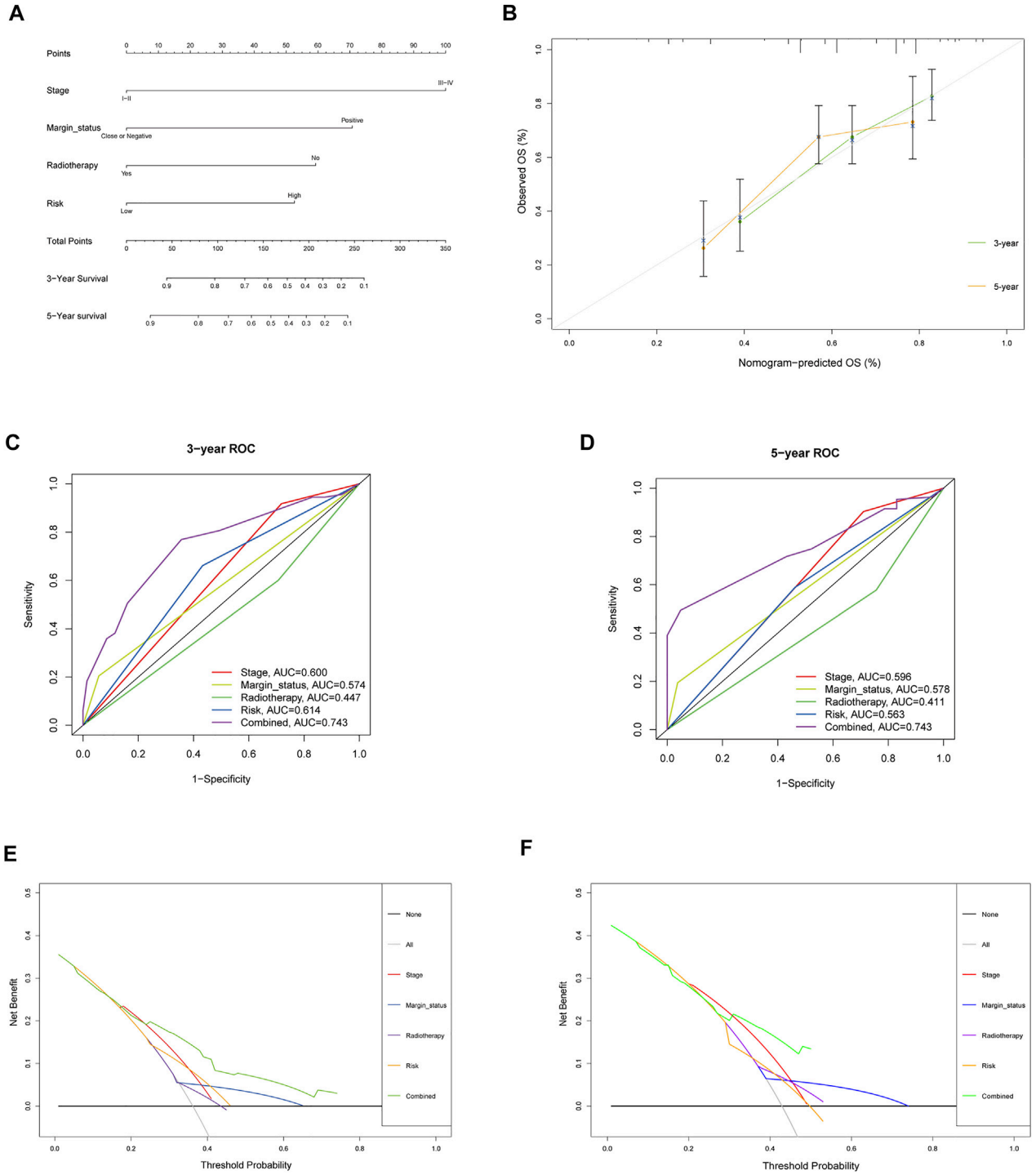



**FIGURE 5 |** Nomogram to predict 3- and 5- year OS and its validation in training cohort. **(A)** Nomogram to predict 3- and 5- year OS of HNSCC patients. **(B)** Calibration plot analysis to assess the accuracy of nomogram to predict 3- and 5- year OS in HNSCC patients. **(C-D)** ROC curves to assess the accuracy of nomogram to predict 3- and 5- year OS in HNSCC patients. **(E,F)** DCA to assess the accuracy of nomogram to predict 3- and 5- year OS in HNSCC patients.

risk groups of HNSC patients is shown in **Figures 7A,B**. The TMB difference was significant between the two groups ($p$ = 0.017) (**Figure 7C**), and the risk score was positively correlated with TMB (R = 0.11; $p$ = 0.019) (**Figure 7D**). Moreover, the OS of the high-risk and high-TMB group was lower than that of the low-risk and low-TMB group, with a $p$ value less than 0.001 (**Figure 7E**). However, a TMB with a high level might fail to accurately predict the immune checkpoint blockade response

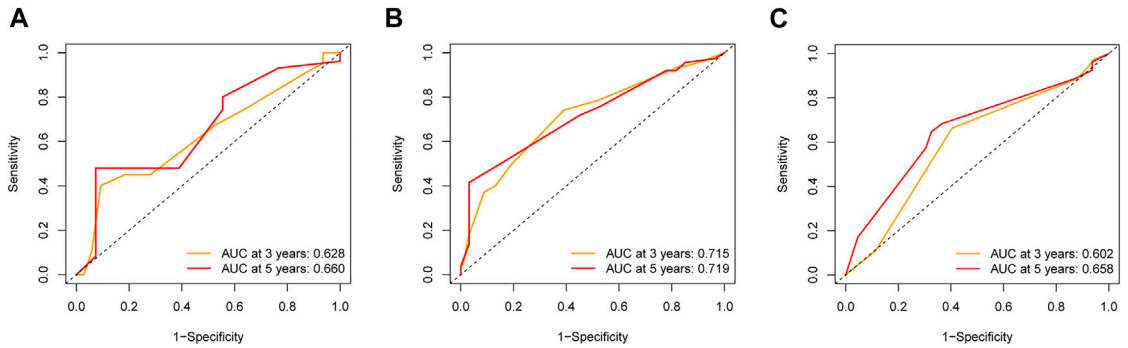



**FIGURE 6 |** The validation of nomogram in internal and external cohorts. **(A)** ROC curves to assess the accuracy of nomogram to predict 3- and 5- year OS in testing cohort. **(B)** ROC curves to assess the accuracy of nomogram to predict 3- and 5- year OS in entire cohort. **(C)** ROC curves to assess the accuracy of nomogram to predict 3- and 5- year OS in external cohort.

across all cancer types, and the immune microenvironment in HNSCC patients was further explored based on the risk predictive model related to NETs (**Figure 8**).

In low-risk patients with HNSCC, higher expression of markers of memory B cells ($p < 0.01$), plasma cells ($p < 0.001$), CD8 T cells ($p < 0.001$), CD4 memory activated T cells ($p < 0.001$), regulatory T cells (Tregs) ($p < 0.001$), and resting mast cells ($p < 0.001$) was observed (**Figures 11A–E**), while higher expression of markers of CD4 memory resting T cells ($p < 0.001$), resting NK cells ($p < 0.01$), M0 macrophages ($p < 0.001$), resting dendritic cells ($p < 0.05$), activated dendritic cells ($p < 0.05$), activated mast cells ($p < 0.001$), eosinophils ($p < 0.001$), and neutrophils ($p < 0.05$) were observed than in the high-risk group (**Figure 9**).

Moreover, differences in immune checkpoint genes in low- and high-risk HNSCC patients were investigated. The expression levels of CTLA4, PD-1, TNFRSF14, BTLA, VSIR, LAIR1, HAVCR2, LGALS9, TIMD4, CD244, CD48, TIGIT, LAG3, IDO1, IDO2, NOS2, CXCL12 and CCL2 were higher in the low-risk group than in the high-risk group ($p < 0.001$). In contrast, the high-risk group had higher CXCL8, VEGFA, and ARG2 expression than the low-risk group ($p < 0.001$) (**Figure 10**). The corresponding correlations between the immune checkpoint genes and risk scores are shown in **Figure 11**.

## Immunophenoscore Analysis of Head and Neck Squamous Cell Carcinoma in the High and Low Risk Groups

To further strengthen the credibility of the immune checkpoint genes based on the risk predictive model, the IPS were analyzed in different risk groups. The IPS might be useful to predict the curative effect of immunotherapy. As shown in **Figures 12A–C**, the three subgroups shown the IPS is significantly higher in low risk group than that in high risk group, which indicated that the low-risk group might receive more benefits from anti-PD-1 therapy, anti-CTLA4 therapy or combined immunotherapy with anti-PD-1 and anti-CTLA4 than the high-risk HNSCC patients.

## Exploration of Drug Sensitivity Based on the Prognostic Model

This model can be used not only to predict the survival of HNSCC patients but also to differentiate the difference in the immune microenvironment in HNSCC patients with different levels of risk. Therefore, it is necessary to investigate anticancer drugs targeting the model. The 16 representative correlation analyses are shown in **Figure 13**. LTF, SELPLG and CYBB were most strongly correlated with anticancer drugs. In the representative correlation analyses, LTF and CTBB were both positively correlated with isotretine, and SELPLG was positively correlated with carmustine.

# DISCUSSION

Convincing evidence has shown that NETs participate in the initiation and development of cancers. NETs are formed when neutrophils develop NETosis, and the web-like structure has been found to play a procancer role in cancer invasion, evasion and metastasis (Demkow, 2021). The OS of HNSCC patients is unsatisfactory even in the context of comprehensive treatment, and the 5-years survival rate of HNSCC patients is further reduced by recurrence, at a 50% rate for the locally advanced stage (Muzaffar et al., 2021). To the best of our knowledge, NETs, an emergent hotspot in oncology, have not been comprehensively investigated in previous studies related to HNSCC.

In the study, a risk prognostic model related to NETs in HNSCC was constructed, the optimistic clinical value of the model was shown by multiple validations. Moreover, the tumor immune microenvironment, immune checkpoint genes, IPS and drug sensitivity in HNSCC based on the prognostic model were further analyzed, which all indicated the potential value of the model related to NETs in immunotherapy of HNSCC patients.

In the present study, we first systematically explored the correlation between NET-related genes and HNSCC. To improve its credibility, the sequencing data in the TCGA and GTEx databases were integrated, which maximally diminished

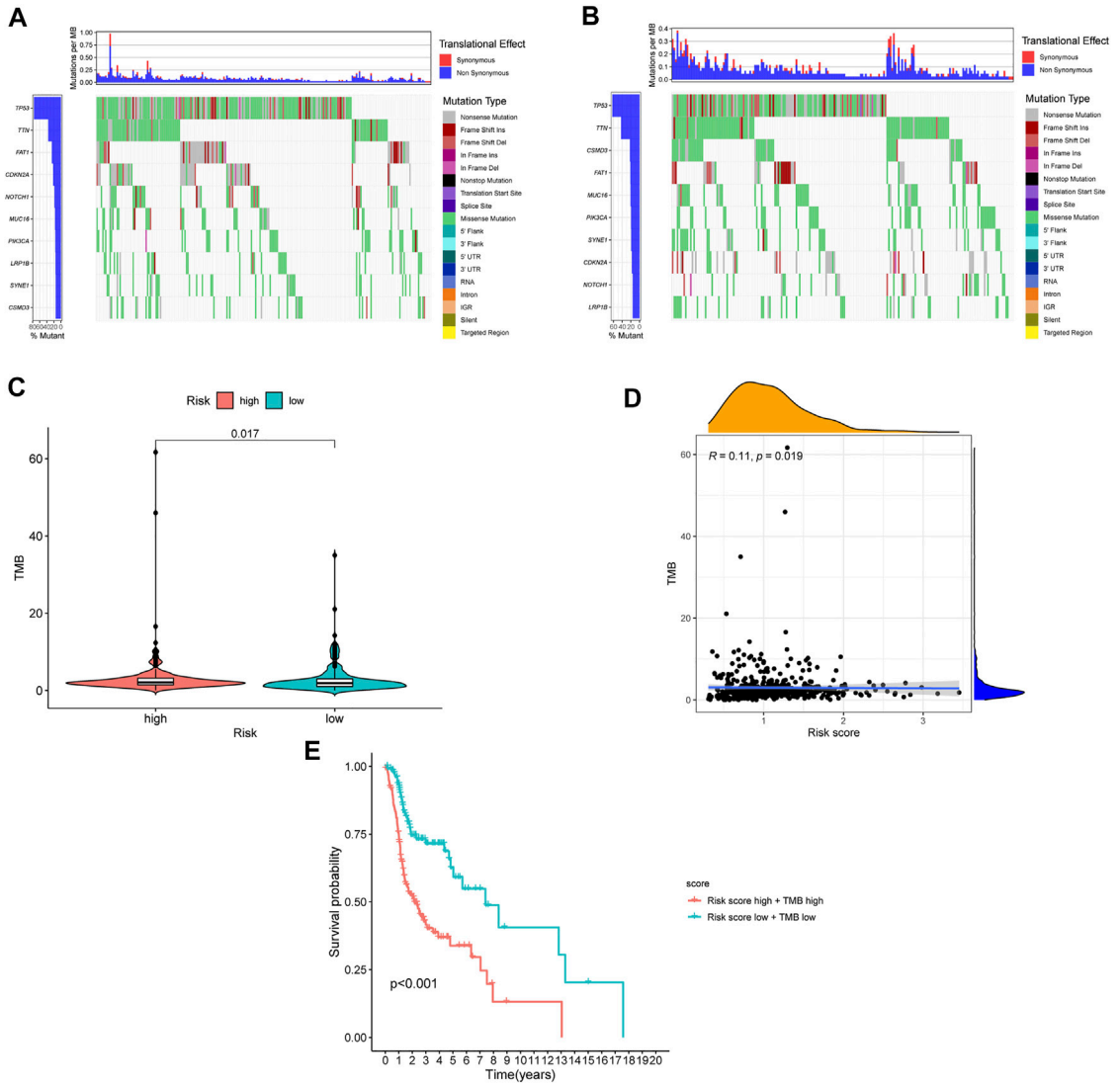



**FIGURE 7 |** The differences of TMB in high and low risk HNSCC patients. **(A)** The TMB in high risk HNSCC patients. **(B)** The TMB in low risk HNSCC patients. **(C)** The difference of TMB was significant in HNSCC patients with different risk. **(D)** The TMB was positively correlated with the risk score in patients with HNSCC. **(E)** The OS of the HNSCC patients of high risk and high TMB were lower than those in HNSCC patients with low risk and low TMB.

the imbalance between cancer samples and noncancer samples. The 6 NET-related genes were screened out by univariate, LASSO, and multivariable logistic regression algorithms to construct the prognostic risk model, which avoided overfitting while strengthening the clinical practicability of the model. The majority of the 6 signature genes in the risk model have been found to be involved in multiple cancers. However, the role of some of these genes in HNSCC is still unclear.

Annexin A3 (ANXA3), an important member of the Annexin multigene family, plays a pivotal role in signaling pathways and malignant biological behaviors of cancer cells, such as proliferation and apoptosis. The abnormal expression of ANXA3 is correlated with the development, occurrence, metastasis and drug resistance of cancers (Park et al., 2005; Harashima et al., 2008; Bandorowicz-Pikuła et al., 2012). The gene expression was exclusively restricted to neutrophils, it's reported that the ANXA3 participate in the formation of NETs in several diseases such as rheumatoid arthritis and systemic lupus erythematosus (Chapman et al., 2019; Toufiq et al., 2020). Lactotransferrin (LTF) is an important member of the transferrin gene family, and it encodes a glycoprotein that is a major iron-binding protein and is widely expressed in saliva, trachea, milk, nasal secretions and neutrophil particles (Bournazou et al., 2010). Some studies have shown that LTF has an antitumor function and can inhibit the proliferation and metastasis of cancer cells, and the expression of this gene is downregulated in multiple types of cancers (Lee et al., 2003; Ni et al., 2020). The LTF has been indicated to suppress the formation and release of NETs, which might be correlated to the anticancer role of the gene (Okubo et al., 2016). Colony-

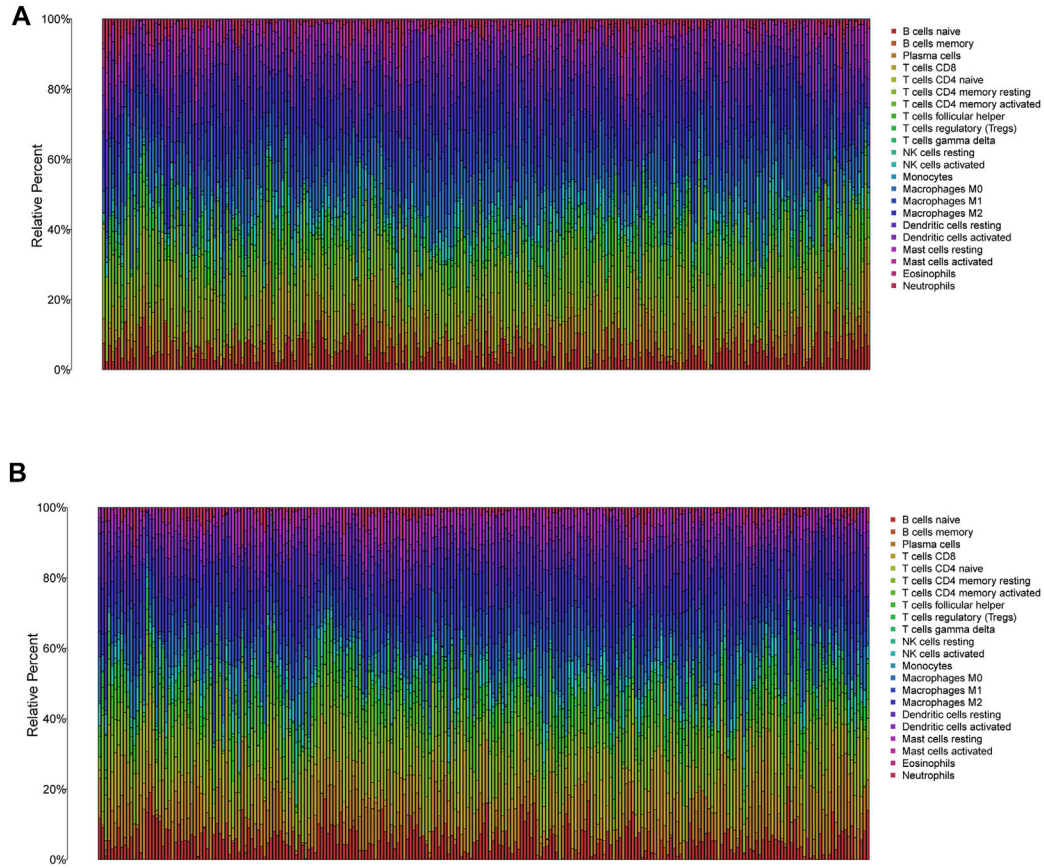



**FIGURE 8 |** The immune infiltration of 22 immune cell types in high and low risk patients with HNSCC. **(A)** The immune infiltration of 22 immune cell types in high risk patients with HNSCC. **(B)** The immune infiltration of 22 immune cell types in low risk patients with HNSCC.

stimulating factor 2 (CSF2) is a granulocyte macrophage-colony stimulating factor that can stimulate the production of monocytes and granulocytes. This gene is associated with neutrophil counts and has been indicated to be an oncogene in several cancers, such as colon cancer and urothelial carcinoma (Lee et al., 2016; Xu et al., 2019). The colony-stimulating factors increase the number of neutrophils and induce their activation, NETs can be released by the systemic release of colony-stimulating factors in cancers, which might be one of carcinogenic effects of CSF2 (Demers and Wagner, 2013). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a key glycolytic pathway-related enzyme that catalyzes redox reactions. However, convincing evidence suggests that GAPDH may play a nonenzymatic role, which is correlated with DNA repair, autophagy and apoptosis (Hara and Snyder, 2006; Colell et al., 2007; Azam et al., 2008). It has been reported that GAPDH is overexpressed in cancer cells, promoting their proliferation and metastasis (Liu et al., 2017). The GAPDH has been indicated to participate in the formation of NETs, the production of NETs increase when the expression level of GAPDH is up-regulated (Dwyer et al., 2014; Agraz-Cibrian et al., 2019). Cytochrome b-245 beta chain (CYBB, also known as NOX2) is a compound enzyme complex that is only expressed in myeloid cells such as macrophages and neutrophilic granulocytes (Martner et al., 2019). The generation of NETs, an early effect of NOX2 activation in neutrophils, is linked to NOX2 activation (Singel and Segal, 2016). Selectin P ligand (SELPLG, also known as PSGL-1) is mainly expressed in inflammatory and immune cells and participates in the recruitment of inflammatory and immune cells to the site of inflammation by tethering and rolling (Laszik et al., 1996). SELPLG act a positive effect to enhances neutrophil recruitment and NET formation, its deficiency might affect neutrophil function and immune cell differentiation and therefore act on tumor growth (Coffelt et al., 2015; Yago et al., 2018).

The HNSCC patients have been divided high- and low risk-groups based on the 6 genes related to NETs, which suggested that the difference of NETs in different risk groups. The OS of HNSCC patients with low NETs was significantly better than that of HNSCC with high NETs, K-M and AUCs at 3 and 5 years showed acceptable values in the training cohort. Furthermore, the internal and external validation improve the credibility of the model. Overall, the practicability of the model in the training cohort was validated by using multiple methods.

To further investigate the clinical prognostic value of the model related to NETs, the risk score of HNSCC was

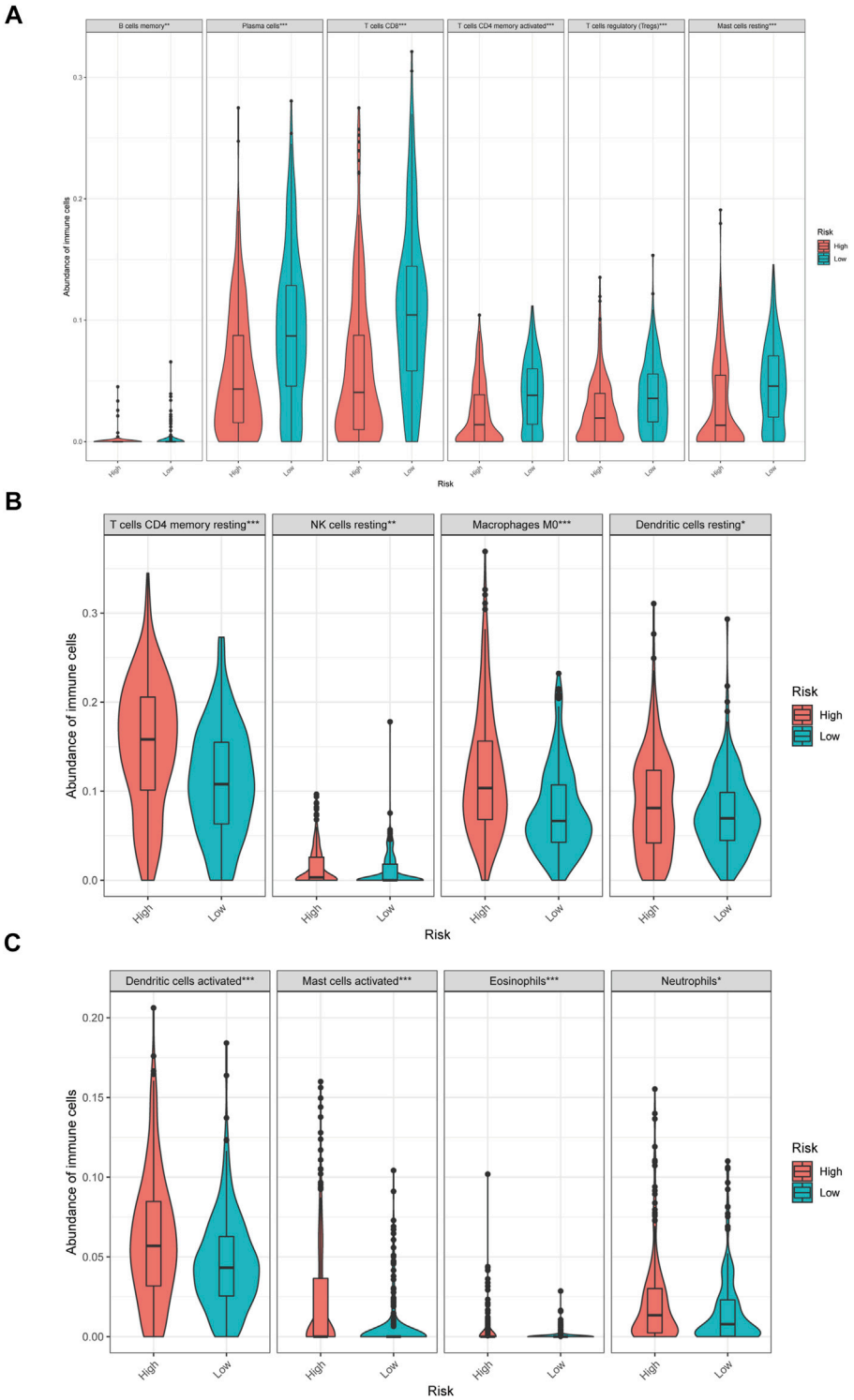



**FIGURE 9 |** The different immune infiltration in the high and low risk HNSCC patients. **(A)** The expression of six types of immune cells is higher in low risk group compared with high risk group. **(B,C)** The expression of eight types of immune cells is higher in high risk group compared with low risk patients.

integrated with clinical parameters. Univariate and multivariable Cox regression algorithms showed that the risk score could be an independent prognostic parameter of HNSCC patients, and the risk score has a synergistic effect with other clinical parameters to improve the value of the nomogram to calculate an individual prognosis. The

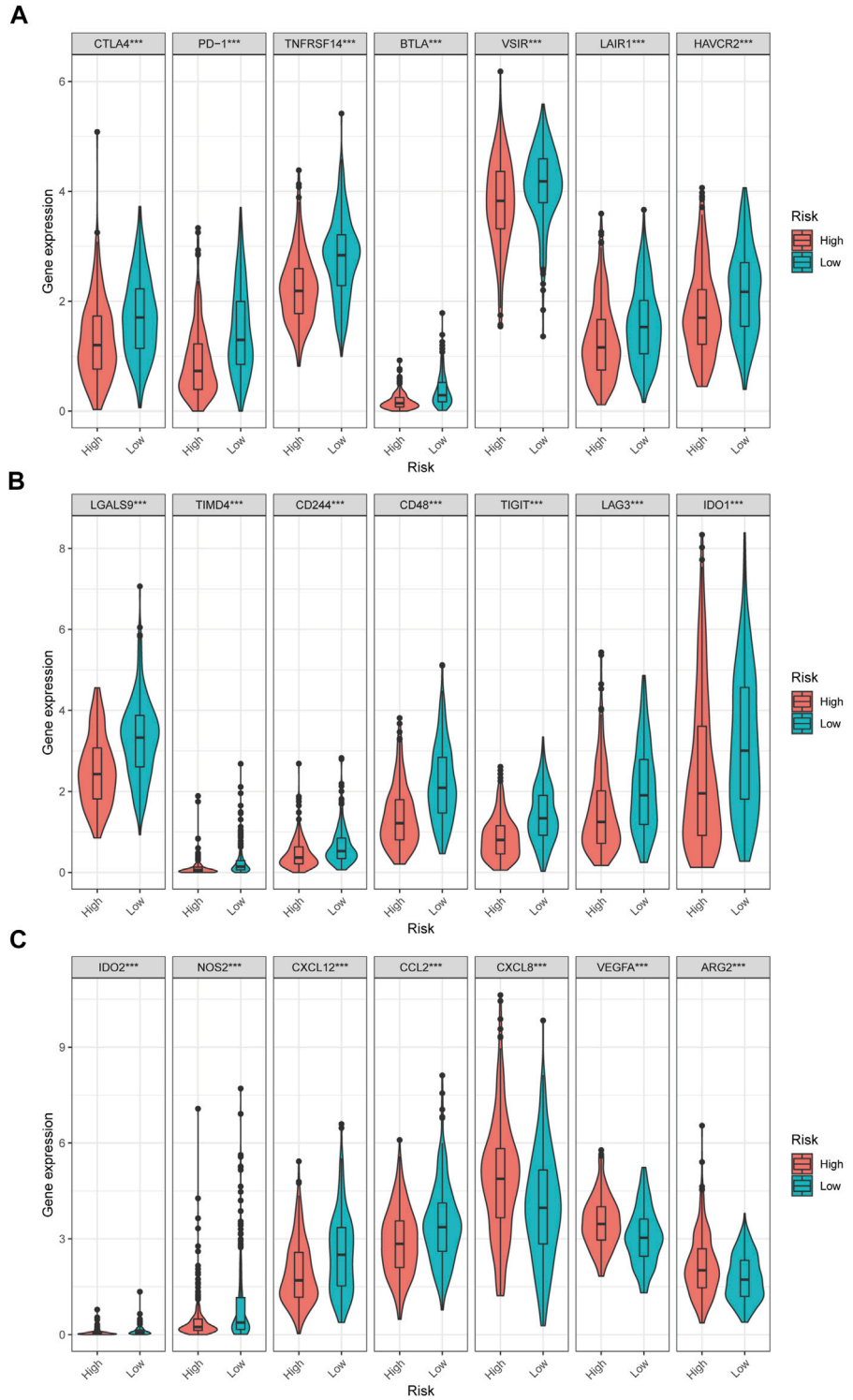



**FIGURE 10 |** The difference of immune checkpoint genes in the high and low risk HNSCC patients.

nomogram used to predict individual prognosis in this study was more effective than that in previous studies with a superior C-index (0.726 vs 0.640) and AUC for 3 (0.743 vs 0.706) and 5 years (0.743 vs 0.645) (He et al., 2021). In the internal and external validation cohorts, the C-index and AUC for 3 and 5 years also showed satisfactory values, which

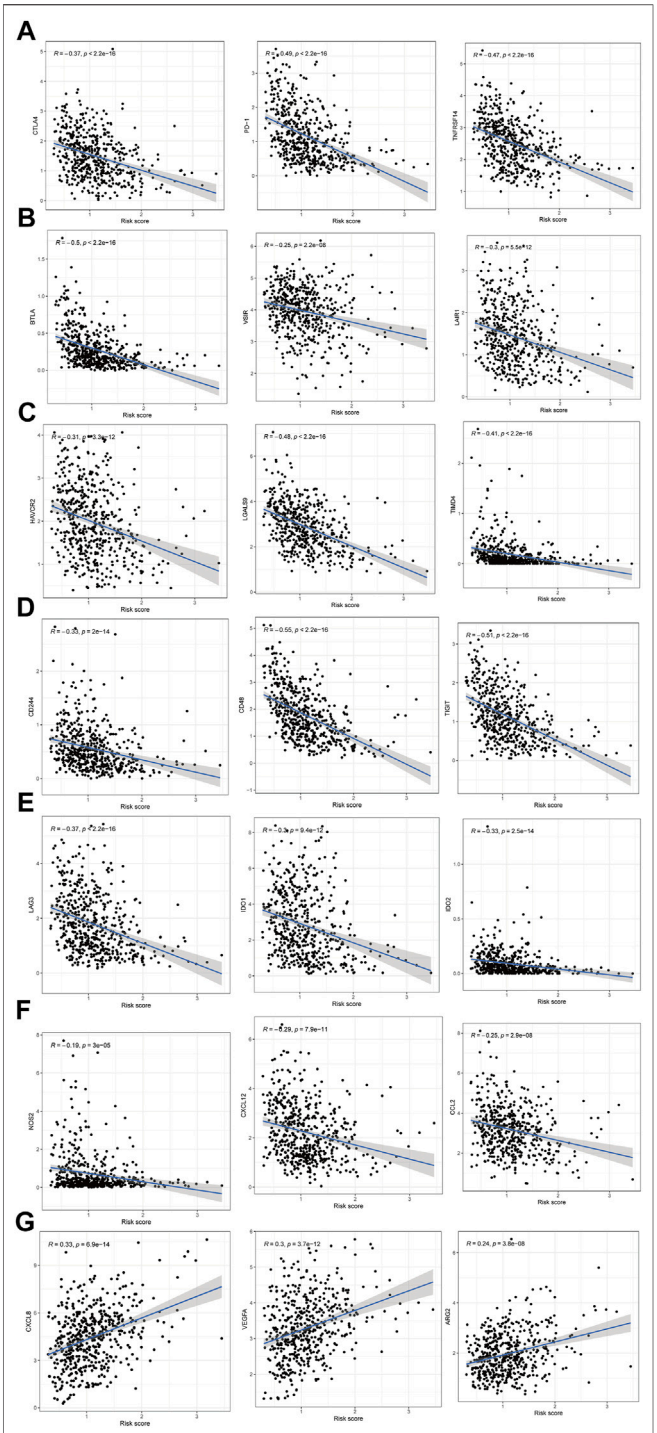



**FIGURE 11 |** The correlations between the immune checkpoint genes and risk score in HNSCC patients. **(A–F)** The immune checkpoint genes was negatively correlated with the risk score. **(G)** The immune checkpoint genes was positively correlated with the risk score.

further strengthened the credibility of the clinical prognostic model. The results further indicated that NETs might acts a synergistic effect with other clinical parameters to predict the clinical outcomes of HNSCC patients.

Moreover, NETs have been suggested to play an important role in immune evasion, wrapping and protecting cancer cells from the anticancer effect of neighboring immune cells such as NK cells and $CD8^+$ T cells in tumor immune microenvironment (TME) (Ireland and Oliver, 2020; Teijeira et al., 2020). NETs acts a "physical barrier" to cover the cancer cells and reduce the curative effect of ICIs and CAR-T. It has been reported that the responsiveness of tumor to PD-1 plus CTLA-4 dual checkpoint blockade can be improved by inhibiting NETs. Moreover, the efficiency of CAR-T might be increased by reducing the NETs in TME (Teijeira et al., 2020; Volkov et al., 2021). Furthermore, the NETs has been suggested to be a biomarker to prognose the response of anti-PD-1 therapy in melanoma based on the significant effect of NETs in immunotherapy (Jensen et al., 2021). Therefore, further exploration of the correlation between the model based on NETs and the immune microenvironment of HNSCC patients is necessary.

TMB has been reported to be a marker for predicting the curative effect of immunotherapy in cancer patients (Chan et al., 2019). In this study, the TMB of HNSCC patients was positively correlated with the risk score, which indicated that the prediction result of TMB and NETs is inconsistent. It has been reported that a high TMB fails to predict the immune checkpoint blockade response across all cancers, which is consistent with our result (McGrail et al., 2021). Therefore, the immune microenvironment in high- and low-risk HNSCC patients needs to be further investigated.

In the present study, the differences in immune infiltration and immune checkpoint genes in high- and low-risk HNSCC patients were illustrated, we found the differences of immune infiltration and immune checkpoint genes were significant. Regulatory T cells (Tregs) are a specialized subpopulation of T cells that act to suppress immune response in some cancers such as ovarian carcinoma, prostate cancer, and non-small cell lung cancer (Curiel et al., 2004; Tao et al., 2012; Flammiger et al., 2013). However, the role of Tregs is associated with a favorable outcome in colorectal carcinoma, Triple-negative breast cancer, and HNSCC (Salama et al., 2009; Bron et al., 2013; Song et al., 2015). In the TME of HNSCC patients, the infiltration abundance of Tregs in high risk group is lower than that in low risk group, the favorable role of Tregs in HNSCC might be correlated with the anti-inflammatory capacity of Tregs which restric inflammation-related carcinogenesis (Shang et al., 2015). The difference of TME in HNSCC patients might be beneficial for improving personalized treatment and immunotherapy effects by using different immune checkpoint inhibitors. Furthermore, the difference of IPS in high- and low-risk HNSCC patients further strengthen the possibility.

Certainly, the present study had several potential limitations. The in-depth molecular mechanisms of the NET-related genes used to construct the prognostic model

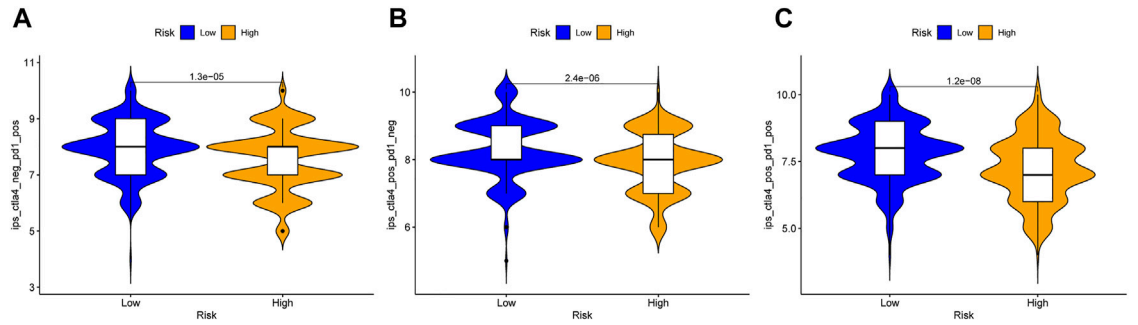



**FIGURE 12 |** Immunophenoscore analysis in the high and low risk HNSCC patients. **(A)** Low-risk group might receive more benefits from anti-PD-1 therapy than the high-risk HNSCC patients. **(B)** Low-risk group might receive more benefits from anti-CTLA4 therapy than the high-risk HNSCC patients. **(C)** Low-risk group might receive more benefits from combined immunotherapy than the high-risk HNSCC patients.

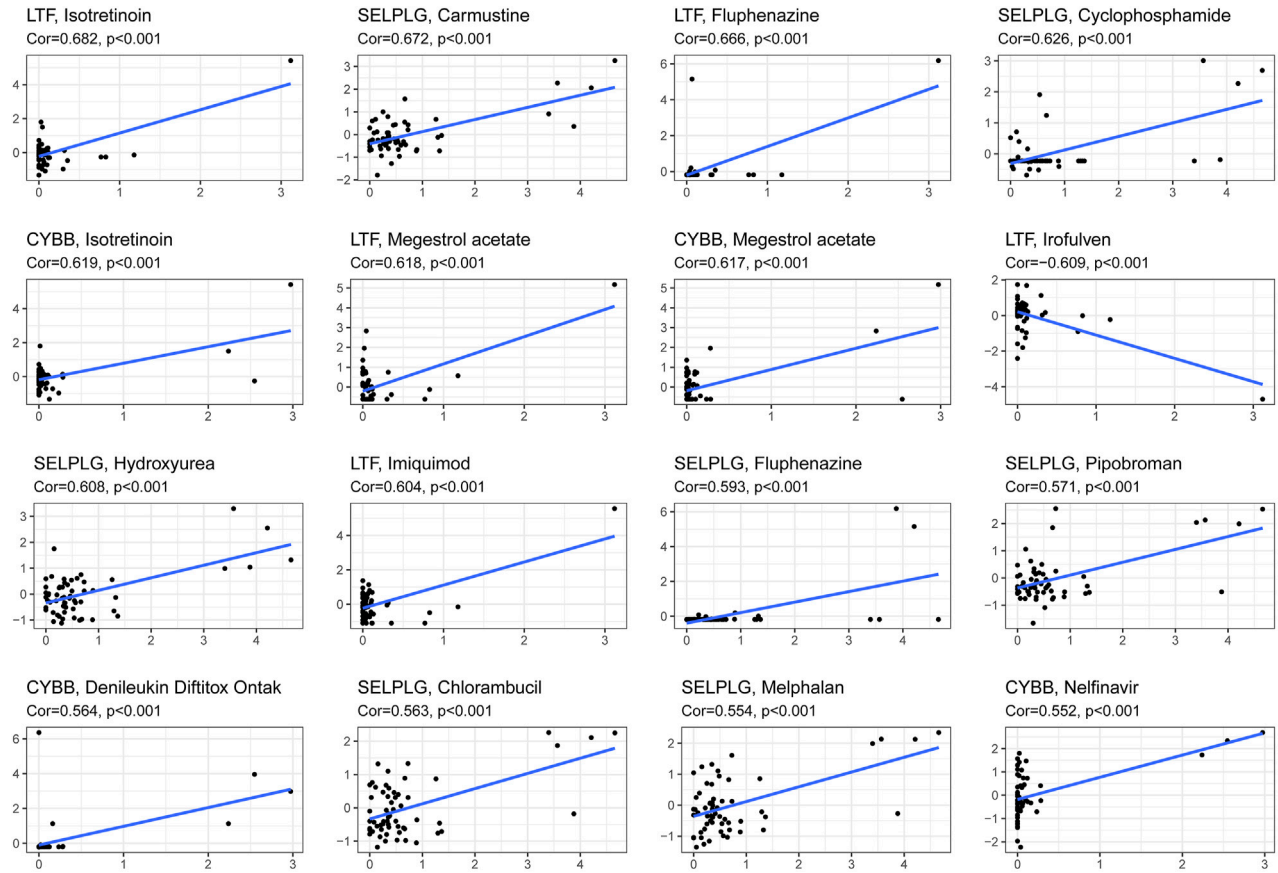



**FIGURE 13 |** The correlation between genes related to NETs to construct the prognostic model and drug sensitivity.

must be further verified in experimental studies. Moreover, this study was based only on research data from public databases, which might contribute to selection bias. Thus, a multicenter and large-scale study is necessary to further validate the clinical utility of our model.

## CONCLUSION

In summary, for the first time, this study identified a novel prognostic model of HNSCC patients based on 6 NET-related genes. Furthermore, the value of the model is promising for

predicting the individual prognosis with other clinical parameters, immunotherapy, and drug sensitivity, which suggests that this novel model related to NETs might be beneficial in improving individualized treatment, thereby improving the curative effect for HNSCC patients.

## DATA AVAILABILITY STATEMENT

The datasets used and analyzed for this study were obtained from TCGA (https://portal.gdc.cancer.gov/), UCSC Xena (https://xena.ucsc.edu/), ArrayExpress (https://www.ebi.ac.uk/arrayexpress/), TCIA (https://tcia.at/) and CellMiner database (https://discover.nci.nih.gov/cellminer/).

## AUTHOR CONTRIBUTIONS

JC: conceptualization. NC and DH: methodology. NC and DH: formal analysis. NC and DH: investigation. NC and DH: writing–original draft preparation. JC: writing–review and editing. NC and DH: visualization. JC: supervision. NC and DH: project administration. NC, DH and JC: funding acquisition. All authors support publishing.

## FUNDING

This work was supported by grants from the National Natural Science Foundation of China—Youth Science Fund (32000617), Scientific Research Project of Putian University (No. 2021049), Jilin Provincial Science and Technology Department (20190303146SF), Jilin Provincial Department of Finance Project (JLSWSRCZX 2020-0023) and Jilin Province Biotherapeutic Science and Technology Innovation Center Project (20200602032ZP).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.833771/full#supplementary-material

**Supplementary Figure S1 |** Representative results of GO and KEGG analyses. **(A)** The GO analysis of the 6 screened genes. **(B)** The potential signaling pathway of the screened 6 genes.

**Supplementary Figure S2 |** The correlation between risk score and different clinical characteristics in HNSCC patients.

**Supplementary Table S1 |** The NET-related genes.

## REFERENCES

Agraz-Cibrian, J. M., Giraldo, D. M., and Urcuqui-Inchima, S. (2019). 1,25-Dihydroxyvitamin D3 Induces Formation of Neutrophil Extracellular Trap-like Structures and Modulates the Transcription of Genes Whose Products Are Neutrophil Extracellular Trap-Associated Proteins: A Pilot Study. *Steroids* 141, 14–22. doi:10.1016/j.steroids.2018.11.001

Azam, S., Jouvet, N., Jilani, A., Vongsamphanh, R., Yang, X., Yang, S., et al. (2008). Human Glyceraldehyde-3-Phosphate Dehydrogenase Plays a Direct Role in Reactivating Oxidized Forms of the DNA Repair Enzyme APE1. *J. Biol. Chem.* 283, 30632–30641. doi:10.1074/jbc.m801401200

Bandorowicz-Pikuła, J., Woś, M., and Pikuła, S. (2012). Participation of Annexins in Signal Transduction, Regulation of Plasma Membrane Structure and Membrane Repair Mechanisms. *Postepy Biochem.* 58, 135–148.

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. *Genome Biol.* 17, 218. doi:10.1186/s13059-016-1070-5

Berger-Achituv, S., Brinkmann, V., Abed, U. A., Kühn, L. I., Ben-Ezra, J., Elhasid, R., et al. (2013). A Proposed Role for Neutrophil Extracellular Traps in Cancer Immunoediting. *Front. Immunol.* 4, 48. doi:10.3389/fimmu.2013.00048

Bournazou, I., Mackenzie, K. J., Duffin, R., Rossi, A. G., and Gregory, C. D. (2010). Inhibition of Eosinophil Migration by Lactoferrin. *Immunol. Cel Biol.* 88, 220–223. doi:10.1038/icb.2009.86

Bron, L., Jandus, C., Andrejevic-Blant, S., Speiser, D. E., Monnier, P., Romero, P., et al. (2013). Prognostic Value of Arginase-II Expression and Regulatory T-Cell Infiltration in Head and Neck Squamous Cell Carcinoma. *Int. J. Cancer* 132, E85–E93. doi:10.1002/ijc.27728

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

Chapman, E. A., Lyon, M., Simpson, D., Mason, D., Beynon, R. J., Moots, R. J., et al. (2019). Caught in a Trap? Proteomic Analysis of Neutrophil Extracellular Traps in Rheumatoid Arthritis and Systemic Lupus Erythematosus. *Front. Immunol.* 10, 423. doi:10.3389/fimmu.2019.00423

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

Coffelt, S. B., Kersten, K., Doornebal, C. W., Weiden, J., Vrijland, K., Hau, C.-S., et al. (2015). IL-17-producing γδ T Cells and Neutrophils Conspire to Promote Breast Cancer Metastasis. *Nature* 522, 345–348. doi:10.1038/nature14282

Colell, A., Ricci, J.-E., Tait, S., Milasta, S., Maurer, U., Bouchier-Hayes, L., et al. (2007). GAPDH and Autophagy Preserve Survival after Apoptotic Cytochrome C Release in the Absence of Caspase Activation. *Cell.* 129, 983–997. doi:10.1016/j.cell.2007.03.045

Curiel, T. J., Coukos, G., Zou, L., Alvarez, X., Cheng, P., Mottram, P., et al. (2004). Specific Recruitment of Regulatory T Cells in Ovarian Carcinoma Fosters Immune Privilege and Predicts Reduced Survival. *Nat. Med.* 10, 942–949. doi:10.1038/nm1093

Demers, M., and Wagner, D. D. (2013). Neutrophil Extracellular Traps. *Oncoimmunology* 2, e22946. doi:10.4161/onci.22946

Demkow, U. (2021). Neutrophil Extracellular Traps (NETs) in Cancer Invasion, Evasion and Metastasis. *Cancers* 13, 4495. doi:10.3390/cancers13174495

Dwyer, M., Shan, Q., D'Ortona, S., Maurer, R., Mitchell, R., Olesen, H., et al. (2014). Cystic Fibrosis Sputum DNA Has NETosis Characteristics and Neutrophil Extracellular Trap Release Is Regulated by Macrophage Migration-Inhibitory Factor. *J. Innate Immun.* 6, 765–779. doi:10.1159/000363242

Flammiger, A., Weisbach, L., Huland, H., Tennstedt, P., Simon, R., Minner, S., et al. (2013). High Tissue Density of FOXP3+ T Cells Is Associated with Clinical Outcome in Prostate Cancer. *Eur. J. Cancer* 49, 1273–1279. doi:10.1016/j.ejca.2012.11.035

Foy, J.-P., Bazire, L., Ortiz-Cuaran, S., Deneuve, S., Kielbassa, J., Thomas, E., et al. (2017). A 13-gene Expression-Based Radioresistance Score Highlights the Heterogeneity in the Response to Radiation Therapy across HPV-Negative HNSCC Molecular Subtypes. *BMC Med.* 15, 165. doi:10.1186/s12916-017-0929-y

Gillison, M. L., Chaturvedi, A. K., Anderson, W. F., and Fakhry, C. (2015). Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma. *J. Clin. Oncol.* 33, 3235–3242. doi:10.1200/jco.2015.61.6995

Guglietta, S., Chiavelli, A., Zagato, E., Krieg, C., Gandini, S., Ravenda, P. S., et al. (2016). Coagulation Induced by C3aR-dependent NETosis Drives

Protumorigenic Neutrophils during Small Intestinal Tumorigenesis. *Nat. Commun.* 7, 11037. doi:10.1038/ncomms11037

Hara, M. R., and Snyder, S. H. (2006). Nitric Oxide-GAPDH-Siah: a Novel Cell Death cascade. *Cell Mol Neurobiol.* 26, 525–536. doi:10.1007/s10571-006-9011-6

Harashima, M., Harada, K., Ito, Y., Hyuga, M., Seki, T., Ariga, T., et al. (2008). Annexin A3 Expression Increases in Hepatocytes and Is Regulated by Hepatocyte Growth Factor in Rat Liver Regeneration. *J. Biochem.* 143, 537–545. doi:10.1093/jb/mvm248

He, F., Chen, Z., Deng, W., Zhan, T., Huang, X., Zheng, Y., et al. (2021). Development and Validation of a Novel Ferroptosis-Related Gene Signature for Predicting Prognosis and Immune Microenvironment in Head and Neck Squamous Cell Carcinoma. *Int. Immunopharmacol.* 98, 107789. doi:10.1016/j.intimp.2021.107789

International Agency for Research on Cancer; World Health Organization (2021). Cancer Today. Available at: https://gco.iarc.fr/(accessed March 10, 2021).

Ireland, A. S., and Oliver, T. G. (2020). Neutrophils Create an impeNETrable Shield between Tumor and Cytotoxic Immune Cells. *Immunity* 52, 729–731. doi:10.1016/j.immuni.2020.04.009

Jensen, C., Kverneland, A., Donia, M., Mortensen, J. H., Karsdal, M. A., Svane, I. M., et al. (2021). Evaluation of CPa9-HNE-A Peripheral Biomarker Based on Neutrophil Extracellular Traps (NETs)-For Predicting Outcome in Patients with Metastatic Melanoma Treated with Anti-PD-1 Therapy. *Cancer Res.* 2021, 386. doi:10.1158/1538-7445

Jiao, X., Wei, X., Li, S., Liu, C., Chen, H., Gong, J., et al. (2021). A Genomic Mutation Signature Predicts the Clinical Outcomes of Immunotherapy and Characterizes Immunophenotypes in Gastrointestinal Cancer. *Npj Precis. Onc.* 5, 36. doi:10.1038/s41698-021-00172-5

Kanamaru, R., Ohzawa, H., Miyato, H., Matsumoto, S., Haruta, H., Kurashina, K., et al. (2018). Low Density Neutrophils (LDN) in Postoperative Abdominal Cavity Assist the Peritoneal Recurrence through the Production of Neutrophil Extracellular Traps (NETs). *Sci. Rep.* 8, 632. doi:10.1038/s41598-017-19091-2

Kolaczkowska, E., and Kubes, P. (2013). Neutrophil Recruitment and Function in Health and Inflammation. *Nat. Rev. Immunol.* 13, 159–175. doi:10.1038/nri3399

Laszik, Z., Jansen, P., Cummings, R., Tedder, T., McEver, R., and Moore, K. (1996). P-selectin Glycoprotein Ligand-1 Is Broadly Expressed in Cells of Myeloid, Lymphoid, and Dendritic Lineage and in Some Nonhematopoietic Cells. *Blood* 88, 3010–3021. doi:10.1182/blood.V88.8.3010.bloodjournal8883010

Lee, J.-Y., Eom, E.-M., Kim, D.-S., Ha-Lee, Y. M., and Lee, D.-H. (2003). Analysis of Gene Expression Profiles of Gastric normal and Cancer Tissues by SAGE. *Genomics* 82, 78–85. doi:10.1016/s0888-7543(03)00098-3

Lee, Y.-Y., Wu, W.-J., Huang, C.-N., Li, C.-C., Li, W.-M., Yeh, B.-W., et al. (2016). CSF2 Overexpression Is Associated with STAT5 Phosphorylation and Poor Prognosis in Patients with Urothelial Carcinoma. *J. Cancer* 7, 711–721. doi:10.7150/jca.14281

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments. *Bioinformatics* 28, 882–883. doi:10.1093/bioinformatics/bts034

Li, T., Zhang, Z., Li, X., Dong, G., Zhang, M., Xu, Z., et al. (2020). Neutrophil Extracellular Traps: Signaling Properties and Disease Relevance. *Mediators Inflamm.* 2020, 1–14. doi:10.1155/2020/9254087

Liu, K., Tang, Z., Huang, A., Chen, P., Liu, P., Yang, J., et al. (2017). Glyceraldehyde-3-phosphate Dehydrogenase Promotes Cancer Growth and Metastasis through Upregulation of SNAIL Expression. *Int. J. Oncol.* 50, 252–262. doi:10.3892/ijo.2016.3774

Liu, M., Li, F., Liu, B., Jian, Y., Zhang, D., Zhou, H., et al. (2021). Profiles of Immune Cell Infiltration and Immune-Related Genes in the Tumor Microenvironment of Esophageal Squamous Cell Carcinoma. *BMC Med. Genomics* 14, 75. doi:10.1186/s12920-021-00928-9

Manda-Handzlik, A., Bystrzycka, W., Cieloch, A., Glodkowska-Mrowka, E., Jankowska-Steifer, E., Heropolitanska-Pliszka, E., et al. (2020). Nitric Oxide and Peroxynitrite Trigger and Enhance Release of Neutrophil Extracellular Traps. *Cell. Mol. Life Sci.* 77, 3059–3075. doi:10.1007/s00018-019-03331-x

Martner, A., Aydin, E., and Hellstrand, K. (2019). NOX2 in Autoimmunity, Tumor Growth and Metastasis. *J. Pathol.* 247, 151–154. doi:10.1002/path.5175

McGrail, D. J., Pilié, P. G., Rashid, N. U., Voorwerk, L., Slagter, M., Kok, M., et al. (2021). High Tumor Mutation burden Fails to Predict Immune Checkpoint Blockade Response across All Cancer Types. *Ann. Oncol.* 32, 661–672. doi:10.1016/j.annonc.2021.02.006

Muzaffar, J., Bari, S., Kirtane, K., and Chung, C. H. (2021). Recent Advances and Future Directions in Clinical Management of Head and Neck Squamous Cell Carcinoma. *Cancers* 13, 338. doi:10.3390/cancers13020338

Németh, T., Sperandio, M., and Mócsai, A. (2020). Neutrophils as Emerging Therapeutic Targets. *Nat. Rev. Drug Discov.* 19, 253–275. doi:10.1038/s41573-019-0054-z

Ni, L., Yuan, C., Zhang, C., Xiang, Y., Wu, J., Wang, X., et al. (2020). Co-expression Network Analysis Identified LTF in Association with Metastasis Risk and Prognosis in clear Cell Renal Cell Carcinoma. *Onco Targets Ther.* 13, 6975–6986. doi:10.2147/OTT.S251000

Okubo, K., Kamiya, M., Urano, Y., Nishi, H., Herter, J. M., Mayadas, T., et al. (2016). Lactoferrin Suppresses Neutrophil Extracellular Traps Release in Inflammation. *EBioMedicine* 10, 204–215. doi:10.1016/j.ebiom.2016.07.012

Papayannopoulos, V. (2018). Neutrophil Extracellular Traps in Immunity and Disease. *Nat. Rev. Immunol.* 18, 134–147. doi:10.1038/nri.2017.105

Park, J. E., Lee, D. H., Lee, J. A., Park, S. G., Kim, N.-S., Park, B. C., et al. (2005). Annexin A3 Is a Potential Angiogenic Mediator. *Biochem. Biophysical Res. Commun.* 337, 1283–1287. doi:10.1016/j.bbrc.2005.10.004

Park, J., Wysocki, R. W., Amoozgar, Z., Maiorino, L., Fein, M. R., Jorns, J., et al. (2016). Cancer Cells Induce Metastasis-Supporting Neutrophil Extracellular DNA Traps. *Sci. Transl. Med.* 8, 361ra138. doi:10.1126/scitranslmed.aag1711

Pruchniak, M. P., and Demkow, U. (2019). Potent NETosis Inducers Do Not Show Synergistic Effects *In Vitro*. *Cent. Eur. J. Immunol.* 44, 51–58. doi:10.5114/ceji.2019.84017

Reinhold, W. C., Varma, S., Sunshine, M., Elloumi, F., Ofori-Atta, K., Lee, S., et al. (2019). RNA Sequencing of the NCI-60: Integration into CellMiner and CellMiner CDB. *Cancer Res.* 79, 3514–3524. doi:10.1158/0008-5472.CAN-18-2047

Richardson, J. J. R., Hendrickse, C., Gao-Smith, F., and Thickett, D. R. (2017). Neutrophil Extracellular Trap Production in Patients with Colorectal Cancer *In Vitro*. *Int. J. Inflamm.* 2017, 1–11. doi:10.1155/2017/4915062

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007

Salama, P., Phillips, M., Grieu, F., Morris, M., Zeps, N., Joseph, D., et al. (2009). Tumor-infiltrating FOXP3+ T Regulatory Cells Show strong Prognostic Significance in Colorectal Cancer. *J. Clin. Oncol.* 27, 186–192. doi:10.1200/jco.2008.18.7229

Samra, B., Tam, E., Baseri, B., and Shapira, I. (2018). Checkpoint Inhibitors in Head and Neck Cancer: Current Knowledge and Perspectives. *J. Investig. Med.* 66, 1023–1030. doi:10.1136/jim-2018-000743

Shang, B., Liu, Y., Jiang, S.-j., and Liu, Y. (2015). Prognostic Value of Tumor-Infiltrating FoxP3+ Regulatory T Cells in Cancers: a Systematic Review and Meta-Analysis. *Sci. Rep.* 5, 15179. doi:10.1038/srep15179

Singel, K. L., and Segal, B. H. (2016). NOX2-dependent Regulation of Inflammation. *Clin. Sci. (Lond)* 130, 479–490. doi:10.1042/CS20150660

Song, Q.-K., Ren, J., Zhou, X.-N., Wang, X.-L., Song, G.-H., Di, L.-J., et al. (2015). The Prognostic Value of Peripheral CD4+CD25+ T Lymphocytes Among Early Stage and Triple Negative Breast Cancer Patients Receiving Dendritic Cells-Cytokine Induced Killer Cells Infusion. *Oncotarget* 6, 41350–41359. doi:10.18632/oncotarget.5534

Tao, H., Mimura, Y., Aoe, K., Kobayashi, S., Yamamoto, H., Matsuda, E., et al. (2012). Prognostic Potential of FOXP3 Expression in Non-small Cell Lung Cancer Cells Combined with Tumor-Infiltrating Regulatory T Cells. *Lung Cancer* 75, 95–101. doi:10.1016/j.lungcan.2011.06.002

Teijeira, Á., Garasa, S., Gato, M., Alfaro, C., Migueliz, I., Cirella, A., et al. (2020). CXCR1 and CXCR2 Chemokine Receptor Agonists Produced by Tumors Induce Neutrophil Extracellular Traps that Interfere with Immune Cytotoxicity. *Immunity* 52, 856–871.e8. doi:10.1016/j.immuni.2020.03.001

Toufiq, M., Roelands, J., Alfaki, M., Syed Ahamed Kabeer, B., Saadaoui, M., Lakshmanan, A. P., et al. (2020). Annexin A3 in Sepsis: Novel Perspectives from an Exploration of Public Transcriptome Data. *Immunology* 161, 291–302. doi:10.1111/imm.13239

Triner, D., and Shah, Y. M. (2019). Hypoxic Regulation of Neutrophils in Cancer. *Int. J. Mol. Sci.* 20, 4189. doi:10.3390/ijms20174189

Volkov, D. V., Tetz, G. V., Rubtsov, Y. P., Stepanov, A. V., and Gabibov, A. G. (2021). Neutrophil Extracellular Traps (NETs): Opportunities for Targeted Therapy. *Acta naturae* 13, 15–23. doi:10.32607/actanaturae.11503

Wu, D., Zhang, P., Li, F., Shen, Y., Chen, H., Feng, Y., et al. (2020). CD138-Multiple Myeloma Cells Express High Level of CHK1 Which Correlated to Overall Survival in MM Patient. *Aging* 12, 23067–23081. doi:10.18632/aging.104066

Xu, Z., Zhang, Y., Xu, M., Zheng, X., Lin, M., Pan, J., et al. (2019). Demethylation and Overexpression of CSF2 Are Involved in Immune Response, Chemotherapy Resistance, and Poor Prognosis in Colorectal Cancer. *Onco Targets Ther.* 12, 11255–11269. doi:10.2147/OTT.S216829

Yago, T., Liu, Z., Ahamed, J., and McEver, R. P. (2018). Cooperative PSGL-1 and CXCR2 Signaling in Neutrophils Promotes Deep Vein Thrombosis in Mice. *Blood* 132, 1426–1437. doi:10.1182/blood-2018-05-850859

Yang, D., and Liu, J. (2021). Neutrophil Extracellular Traps: a New Player in Cancer Metastasis and Therapeutic Target. *J. Exp. Clin. Cancer Res.* 40, 233. doi:10.1186/s13046-021-02013-6

Yang, L., Liu, Q., Zhang, X., Liu, X., Zhou, B., Chen, J., et al. (2020). DNA of Neutrophil Extracellular Traps Promotes Cancer Metastasis via CCDC25. *Nature* 583, 133–138. doi:10.1038/s41586-020-2394-6

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

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

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 23 March 2022
doi: 10.3389/fmolb.2022.871602

# Comprehensive Analysis of Pyroptosis-Related Genes and Tumor Microenvironment Infiltration Characterization in Papillary Renal Cell Carcinoma

*Chiyu Zhang*[†], *Ruizhen Huang*[†] *and Xiaoqing Xi*[*]

*Department of Urology Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, China*

OPEN ACCESS

***Edited by:***
*Umberto Malapelle,*
*University of Naples Federico II, Italy*

***Reviewed by:***
*Hao Chen,*
*Guangdong Academy of Medical Sciences, China*
*Jian Lu,*
*Peking University Third Hospital, China*

****Correspondence:***
*Xiaoqing Xi*
*xixiaoqing500@sina.com*

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

***Specialty section:***
*This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences*

***Received:*** *08 February 2022*
***Accepted:*** *07 March 2022*
***Published:*** *23 March 2022*

***Citation:***
*Zhang C, Huang R and Xi X (2022) Comprehensive Analysis of Pyroptosis-Related Genes and Tumor Microenvironment Infiltration Characterization in Papillary Renal Cell Carcinoma. Front. Mol. Biosci. 9:871602. doi: 10.3389/fmolb.2022.871602*

**Background:** Immunotherapy has emerged as an important technique for treating a variety of cancers. The dynamic interplay between tumor cells and invading lymphocytes in the tumor microenvironment is responsible for the good response to immunotherapy (TME). Pyroptosis, or inflammation-induced cell death, is closely linked to a number of cancers. However, in papillary renal cell carcinoma (KIRP), the association between pyroptosis and clinical prognosis, immune cell infiltration, and immunotherapy impact remains unknown.

**Methods:** We carefully investigated the link between pyroptosis and tumor growth, prognosis, and immune cell infiltration by evaluating 52 pyroptosis-related genes. The PRG score was utilized to measure a single tumor patient's pyroptosis pattern. After that, we looked at how well these values predicted prognoses and therapy responses in KIRP.

**Results:** We discovered that PRG differences between subgroups were linked to clinical and pathological aspects, prognosis, and TME in two separate genetic subtypes. After that, a PRG score for estimating overall survival (OS) was developed, and its predictive potential in KIRP patients was confirmed. As a result, we developed a very precise nomogram to improve the PRG score's clinical usefulness. A low PRG score, which is determined by mutation load and immune activation, suggests a good chance of survival. Furthermore, the PRG score was linked to chemotherapeutic drug sensitivity in a substantial way.

**Conclusions:** The possible functions of PRGs in the TME, clinical and pathological characteristics, and prognosis were established in our thorough investigation of PRGs in KIRP. These results might help us better understand PRGs in KIRP and offer a new avenue for prognostic evaluation and the development of more effective immunotherapy treatments.

**Keywords: pyroptosis, papillary renal cell carcinoma, gene mutation, prognostic model, drug sensitivity**

## BACKGROUND

Renal cell carcinoma (RCC) is a broad term that refers to a variety of malignancies. Each kind has different histologic characteristics, a distinct genetic profile, and a varied clinical history and response to treatment (Moch et al., 2016). Papillary renal cell carcinoma (KIRP) is the second most frequent RCC subtype, accounting for 10–15% of patients (Akhtar et al., 2019). Papillary renal cell carcinoma is divided into two categories based on their histological characteristics (Mendhiratta et al., 2021). When compared to non-papillary RCC subtypes, KIRP often present as homogenous, solid masses with relative hypovascularity (Vikram et al., 2009; Tordjman et al., 2020). Targeted therapy for KIRP has previously failed due to a lack of understanding of the molecular underpinnings of these cancers (Williamson, 2021). It is hoped that, as we get a better understanding of the etiology of the pathogenesis of KIRP, effective targeted therapeutics will emerge in due course.

Pyroptosis, defined as a programmed cell death caused by the Gasdermin family, is accompanied by inflammatory and immunological responses (Xia et al., 2019; Meng et al., 2021; Wu et al., 2021). Pyroptosis and cancer have a complicated interaction; pyroptosis can prevent tumor incidence and progression while simultaneously acting as a factor promoting inflammation to generate a milieu conducive to tumorigenesis (Wu et al., 2021; Xiang et al., 2021). The fundamental cause of pyroptosis is the stimulation of inflammasome, which occurs through the caspase inflammasome pathways (Zheng and Li, 2020; Song et al., 2021). The importance of pyroptosis in the TME is becoming clearer, while the molecular basis of pyroptosis in the KIRP immune microenvironment is yet unknown.

Employing computational techniques, this work analyzed the mRNA transcription of pyroptosis-related genes (PRGs) in detail and created a complete picture of the tumoral immune landscape. To begin with, patients with KIRP were divided into two distinct subgroups based on their PRG transcription levels. On the basis of differentially expressed genes (DEGs), patients were then divided into three gene subgroups. Additionally, we developed a grading scale to estimate overall survival (OS) and to define the immunological landscape of KIRP, which effectively predicted clinical outcomes and immunotherapy responsiveness.

## MATERIALS AND METHODS

### KIRP Data Source and Preprocessing

The TCGA platform was used to obtain the transcriptome RNA sequences and clinical data for 471 KIRP samples. The Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) platform was used to collect the GEO cohort (GSE2748) KIRP samples. Incomplete clinical information was deleted from samples, and FPKM values in TCGA-KIRP were transformed to Transcripts Per Kilobase Million (TPM) values and utilized for copy number variation (CNV) analysis (Zhao et al., 2021). The transcriptome RNA sequences from the TCGA-KIRP and GSE2748 datasets were combined after correction. R software version 4.1.1 was used to process the raw data.

### Unsupervised Clustering Analysis of PRGs

First and foremost, PRGs were identified from published literature (Man and Kanneganti, 2015; Wang and Yin, 2017; Karki and Kanneganti, 2019). The transcription of 52 PRGs was utilized to identify distinct pyroptosis types and to identify patients for further research using an unsupervised clustering approach. To do the aforementioned study, we utilized the R package "ConsensuClusterPlus" and 1,000 repeats to ensure clustering stability (Wilkerson and Hayes, 2010). Our best guess for how many clusters there should be was arrived at by using the consensus clustering method. Gene set variation analysis (GSVA) was performed on the MSigDB hallmark gene set in order to investigate the variations in PRGs that occur throughout cellular mechanisms. Furthermore, the Kaplan–Meier curves generated by the R packages "survival" and "survminer" were used to investigate the differences in overall survival across subtypes.

### Relationship Between Molecular Subgroups and TME in KIRP

We utilized the ESTIMATE method to assess each patient's immunological and stromal scores. KIRP samples' fractions of 22 different kinds of human immune cells were also determined using the CIBERSORT technique (Chen et al., 2018). To evaluate the degree of immune cell infiltration in the TME of KIRP, a single-sample gene set enrichment analysis (ssGSEA) approach was utilized in conjunction with the other techniques (Subramanian et al., 2005).

### Identification of DEGs in Various PRG Clusters

We utilized the R package "limma" to compare DEGs between PRG clusters (Ritchie et al., 2015). The Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses were carried out using the "clusterProfiler" program (Gene Ontology Consortium, 2015; Kanehisa et al., 2017).

### Generation of PRG Score

To objectively measure pyroptosis in individual KIRP patients, a score system was developed. The following is the procedure for developing the scoring system: The DEGs found in various pyroptosis clusters were first standardized across all samples, and then the overlap genes were retrieved. Then, for each gene, we ran a univariate Cox regression analysis. For the following phase of the investigation, a list of genes with a high prognosis was created. The following formula was used to determine pyroptosis scores using principal component analysis (PCA):

$$\text{PRG Score} = \Sigma\ (\text{Expi} * \text{Coefi})$$

### The Creation and Testing of a Nomogram Grading System

The clinical characteristics and risk score from the independent prognostic study were used to generate a nomogram using the

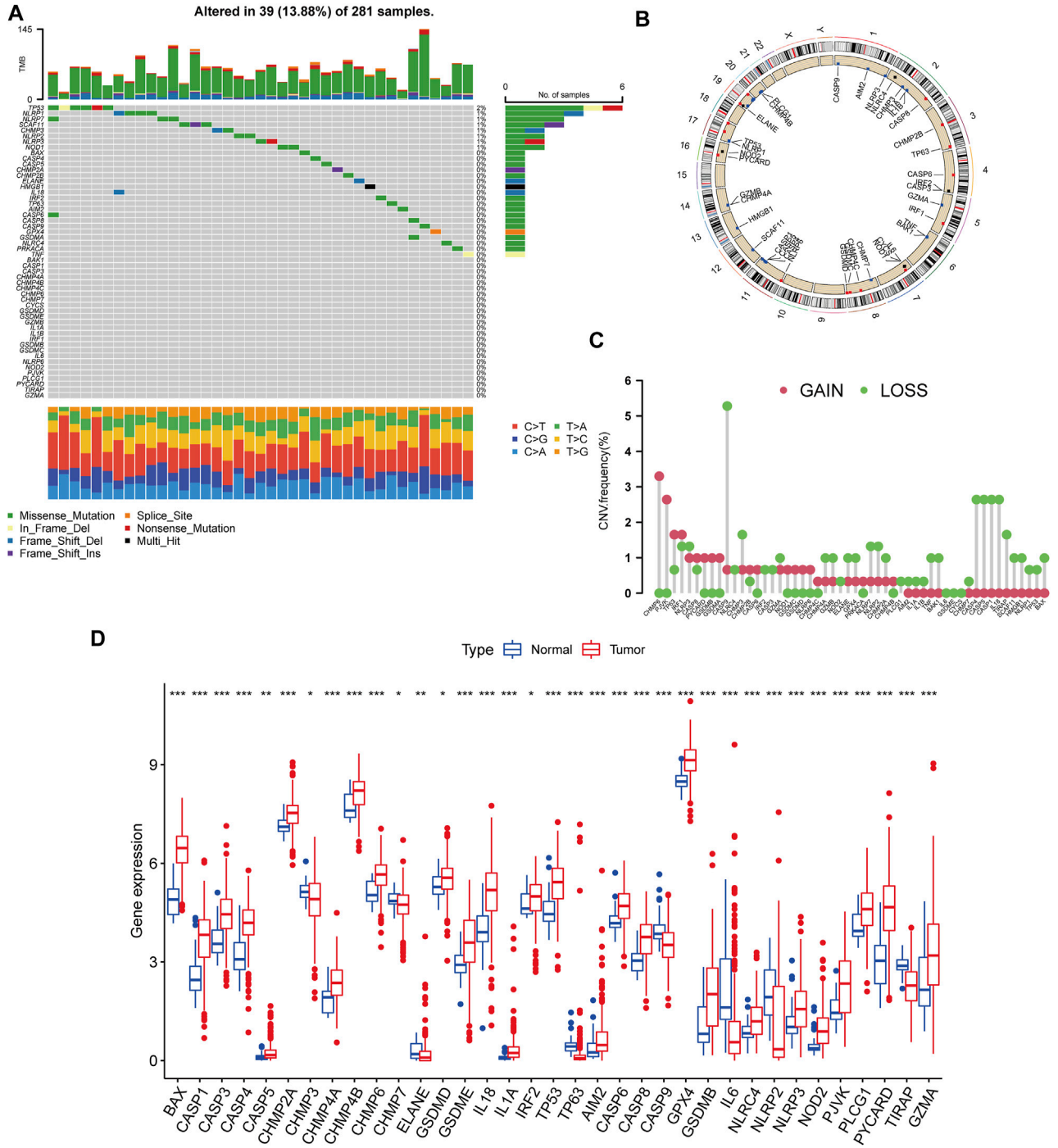



**FIGURE 1 |** Genetic and transcriptional variation of PRGs in KIRP. **(A)** The number of genetic mutations found in KIRP patients. The PRGs' mutation frequency was indicated by the number on the right. Each column represents one of the 39 TCGA samples in which the PRGs had at least one genetic change. **(B)** On the chromosomes, the position of the PRGs CNV changes. **(C)** The PRGs' CNV mutation frequency. The frequency of modification was displayed in the column. The frequency of amplification is shown by a red dot, whereas the frequency of deletion is indicated by a green dot. **(D)** PRGs differential expression study in normal and KIRP tissues. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

"rms" program (Park, 2018). In the nomogram scoring approach, a score was given to each variable, and the total score was derived by adding the scores from all variables in each sample. With the help of nomogram calibration plots, it was possible to see the relationship between surviving events and their virtually observed consequences.

## Analyzing Drug Susceptibility and Mutations

The mutation annotation file from the TCGA was developed using the "maftools" R package to determine the genetic alterations of KIRP patients in various categories (Mayakonda et al., 2018). For each patient with KIRP in the two groups, we estimated the tumor mutation burden (TMB) value. Using the "pRRophetic" software, we estimated the semi-inhibitory concentration (IC50) values of medications frequently used to treat KIRP in patients to see if there were any variations in their therapeutic effects (Geeleher et al., 2014).

# RESULTS

## Genetic and Transcriptional Variation of PRGs in KIRP

We began by looking at their CNV and mutation frequency in TCGA-KIRP. PRG mutations were found in 39 of 281 TCGA-KIRP samples, accounting for 13.88 percent of the total. The top four PRGs with the greatest mutation frequencies were TP53 (2%), NLRP1 (1%), NLRP7 (1%), and SCAF11 (1%) (**Figure 1A**). Further investigating the predictive impact of PRGs, we discovered that most PRGs, such as TP53, GZMB, AIM2, CHMP7, and GSDMC, were substantially linked with the overall prognosis of KIRP patients. **Figure 1B** depicts the 39 PRG positions on the chromosomal CNV alterations. CNV change was widespread in the 52 PRGs that exhibited mostly copy number amplification, such as CHMP6 and PJVK, whereas CASP9, CASP4, CASP5, CASP1, and IL18 showed considerable copy number decrease (**Figure 1C**). The levels of PRGs in KIRP patients and renal tissues were examined to see whether the genetic variants mentioned above had an influence on PRG transcription in KIRP patients (**Figure 1D**). According to our results, variations in CNV may be major factors driving changes in PRG regulator expression. CNV-amplified PRG transcription was considerably greater in KIRP compared to normal renal tissue (e. g., CHMP6 and PJVK) and vice versa (e. g., CHMP3, CHMP7, ELANE, CASP9, NLRP2, and TIRAP). This study found considerable PRG inherited heterogeneity and variance in KIRP patients, indicating that PRG transcription imbalances were important in the onset and development of KIRP.

## Construction of Pyroptosis Subtypes in KIRP

The patients from two suitable KIRP cohorts (TCGA-KIRP and GSE2748) were included in our study for additional analysis in order to thoroughly recognize the PRG expression levels implicated in carcinogenesis. We created a pyroptosis network map to demonstrate the interactions of PRGs in KIRP patients and their relationship to prognosis in these patients (**Figure 2A**). We employed a consensus clustering technique to identify patients according to the transcript level of the 52 PRGs to learn more about their expression features in KIRP. Our findings indicated that k = 2 was the best choice for categorizing the whole patients into subgroup A (n = 177) and subgroup B (n = 146) (**Figures 2B,C**). According to the results of the PCA analysis, the PRG expression patterns of the two subgroups were considerably different (**Figure 2D**). In addition, comparing the clinicopathological characteristics of various KIRP subtypes revealed substantial disparities in PRG transcription and clinical features (**Figure 2E**). Cluster A had the highest level of expression for PRGs.

Following that, we used GSVA enrichment to investigate the biologic effects of several PRG groups. There was significant enrichment of immune cell activation pathways in PRG cluster A, including T cell and B cell receptor signaling pathways (**Figure 3A**). Then, we employed ssGSEA to compare immune cell infiltration in the tumor microenvironment amongst distinct PRG clusters. Surprisingly, we found a considerable variation in immune cell concentration amongst PRG clusters, with cluster A was considerably enriched in immune cells (**Figure 3B**).

## PRG Signature Generation

We utilized the "limma" R package to explore 2461 PRG phenotype-associated DEGs, which were designated as PRG signature genes, to learn more about the probable biologic activities of distinct PRG clusters. Surprisingly, it was shown that DEGs are greatly overrepresented in immune-related pathways. DEGs were found to be concentrated in biological processes of BP related to T cell activation and leukocyte cell-cell adhesion. DEGs were considerably enriched in cytokine receptor binding, cytokine binding, and immune receptor activity throughout MF processes (**Figure 4A**). Additionally, DEGs were considerably enriched in immune-related pathways in KEGG pathway enrichment analyses: cytokine-cytokine receptor interaction, Human T-cell leukemia virus infection, and chemokine signaling pathway (**Figure 4B**).

Following that, we used univariate Cox analysis to screen for prognostically related DEGs, yielding 659 genes related to survival time. We employed unsupervised cluster analysis to separate the KIRP patients into three gene groups based on the genes that were chosen for inclusion (**Figures 5A,B**). The survival research showed that patients in subgroup C had the highest overall survival rate, while those in subgroup A had the lowest OS (**Figure 5C**). In gene cluster A, all PRGs were found to be overexpressed (**Figure 5D**). This revealed that greater PRG expression might be related to a worse prognosis in KIRP patients. Furthermore, the heatmap revealed that these prognostically associated DEGs were common in subgroup A, which confirmed the findings of the previous study (**Figure 5E**).

## Generation of PRG the Signatures Scoring System

We established the PRG score, a scoring method that evaluates the pyroptosis types of each KIRP patient, to further examine the activities of pyroptosis in KIRP. In order to begin, we separate the patients into two groups: those who will be trained and those who

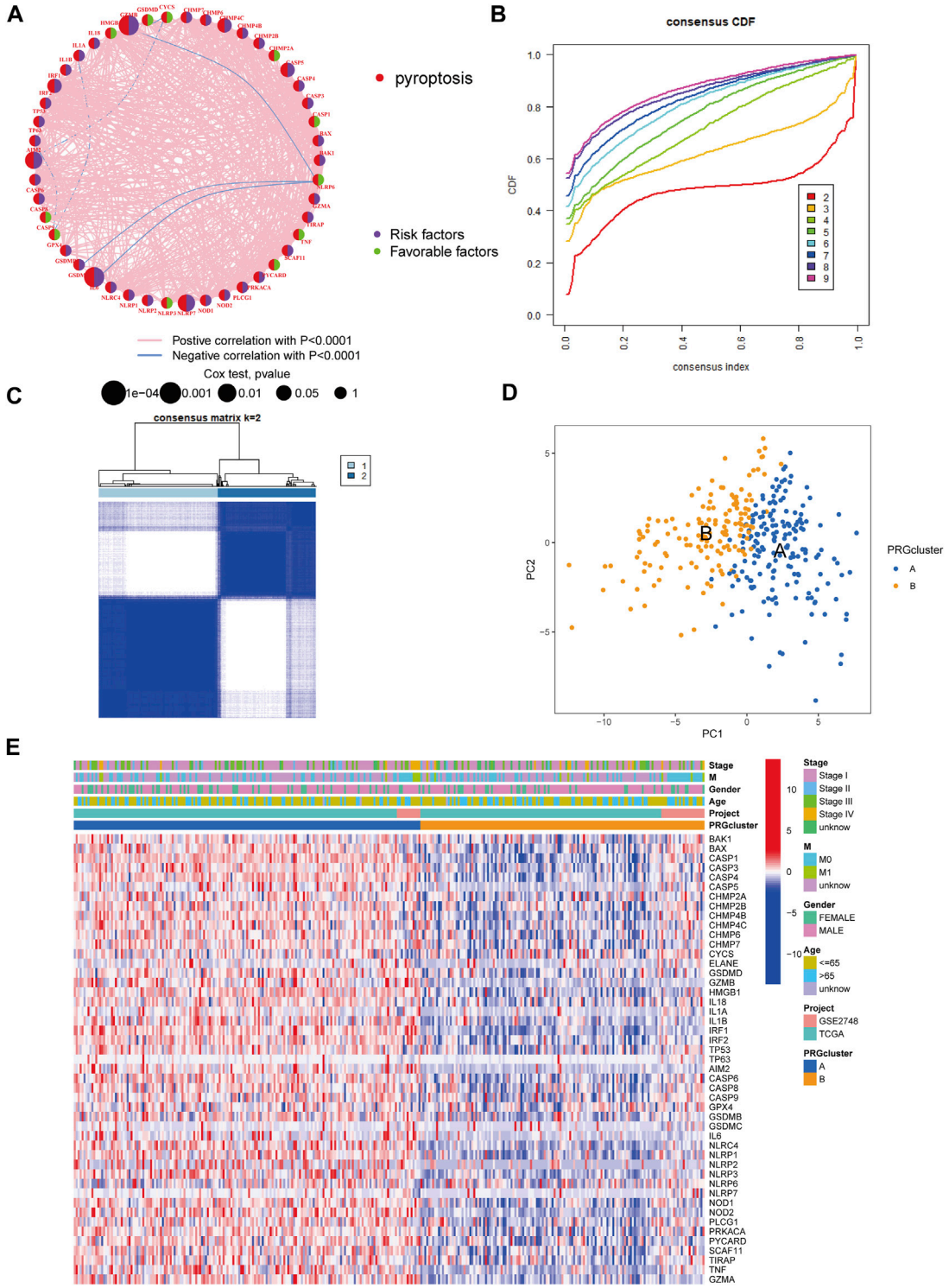



**FIGURE 2 |** Unsupervised cluster analysis divided KIRP patients into subclusters. **(A)** Interactions of PRGs in KIRP. **(B)** The cumulative distribution function (CDF) for consensus clustering. **(C)** Heatmap of the consensus matrix identifying two clusters and their correlation area. **(D)** Principal component analysis (PCA) between the two subtypes. **(E)** PRG heatmap in KIRP between two distinct subtypes.

will be tested. LASSO and multivariate Cox analyses were performed on prognostic DEGs associated with pyroptosis subtypes to further select the best prognostic features (**Figures 6A,B**). In order to depict variations in the characteristics of specific patients, an alluvial diagram was employed (**Figure 6C**). Between subgroups, we identified a statistically

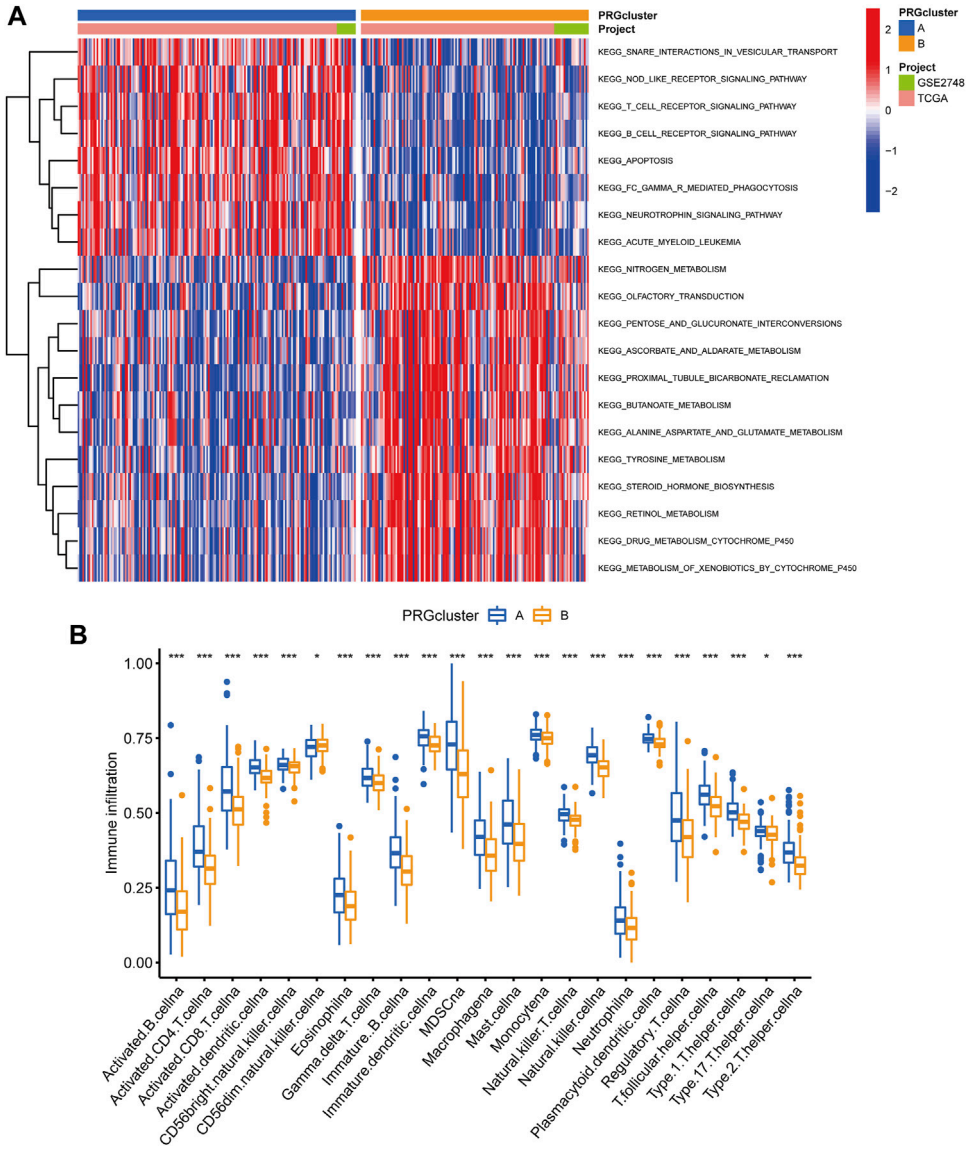



**FIGURE 3 |** Analysis of GSVA enrichment and TME immune cell infiltration. **(A)** The functional pathways of two PRG subgroups were identified using GSVA enrichment. **(B)** Immune cell infiltration levels differ between PRG subgroups.

significant variance in PRG score. Additionally, both gene cluster A (**Figure 6D**) and PRG cluster A had a high PRG score (**Figure 6E**). On the basis of the median PRG score, patients were split into high and low-risk categories. When the PRG score is increased, the survival duration reduces, and the recurrence rate rises, as shown by the risk distribution chart for the PRG score (**Figures 6F,G**). Furthermore, we discovered substantial disparities in the expression of PRGs between risk categories (**Figure 6H**).

## Survival Prediction and Nomogram

In all sets ($p < 0.001$, **Figure 7A**), training sets ($p = 0.008$, **Figure 7B**), and testing sets ($p < 0.001$, **Figure 7C**), those with low scores had considerably better OS than subjects with elevated scores, according to the K–M curves. In all sets, AUC values of 0.948, 0.809, and 0.820 were used to reflect the 1-, 3-, and 5-year survival rates of the PRG score (**Figure 7D**). Similarly, the AUC in the training set was 0.967, 0.707, and 0.713 (**Figure 7E**), and in the test set, they were 0.936, 0.916, and 0.959 (**Figure 7F**). According to this, the PRG score may accurately predict the clinical outcome with KIRP patients. Because the PRG score's clinical utility in evaluating OS in patients of KIRP is unsatisfactory, a nomogram integrating the PRG score and clinical characteristics was developed to predict survival rates (**Figure 7H**). The calibration chart revealed that the PRG

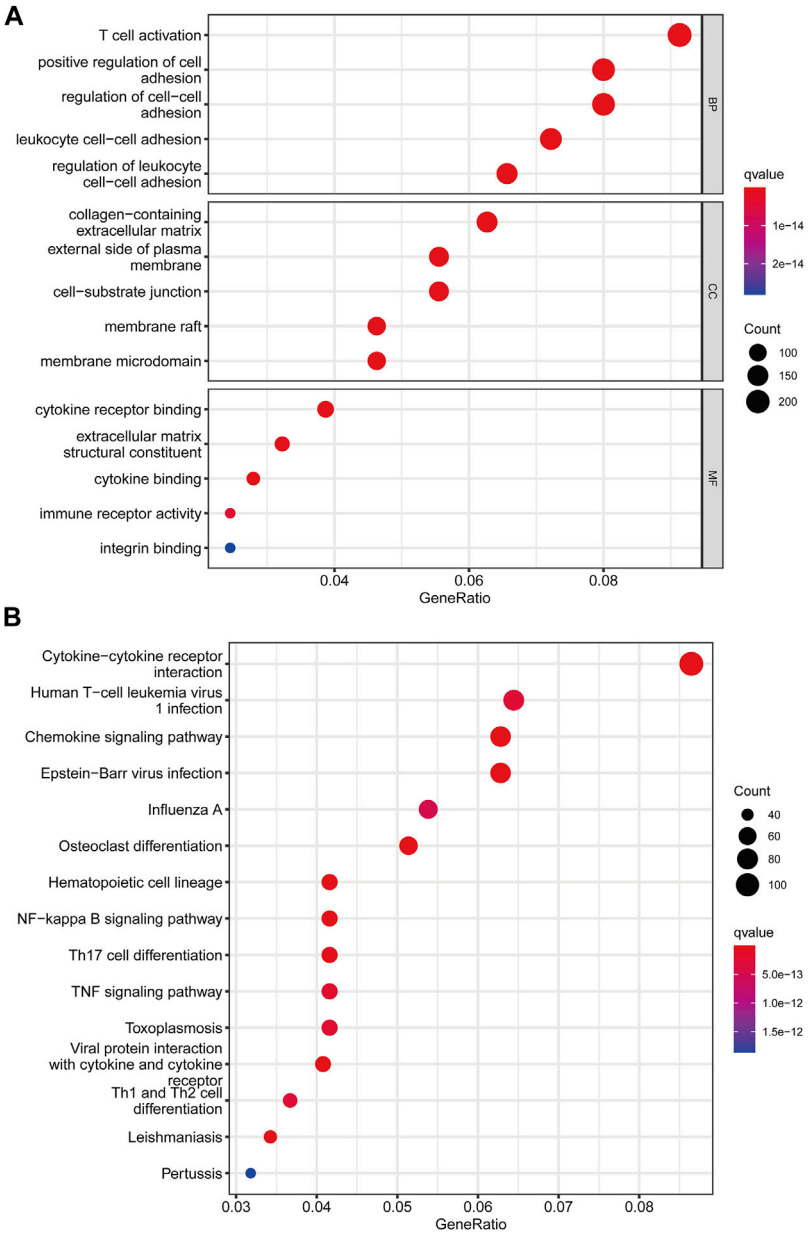



**FIGURE 4 |** Functional enrichment analysis among two pyroptosis subtypes. **(A)** Enrichment of GO functions. **(B)** The KEGG pathway.

score performed well, with a good match between the projected and actual survival rates (**Figure 7G**).

## Assessment of TME Between the Two Groups

The CIBERSORT technique was utilized to determine the association between the PRG score and immune cell abundance. The PRG score was favorably connected with B cells and dendritic cells resting, macrophages M1, and other cells, as indicated in the scatter diagrams, and negatively correlated with macrophages M0, macrophages M2 (**Figure 8A–L**). Additionally, we examined the relationship between the genes and the number of immune cells in the body. We discovered that the 13 genes were highly linked with the majority of immune cells, especially GBP1 and macrophages M1, KCNJ5 and macrophages M2 (**Figure 8M**).

## Analyses of Gene Mutation and Drug Sensitivity

Then, in the TCGA-KIRP cohort, we looked at how the distribution of somatic mutations differed across two PRG

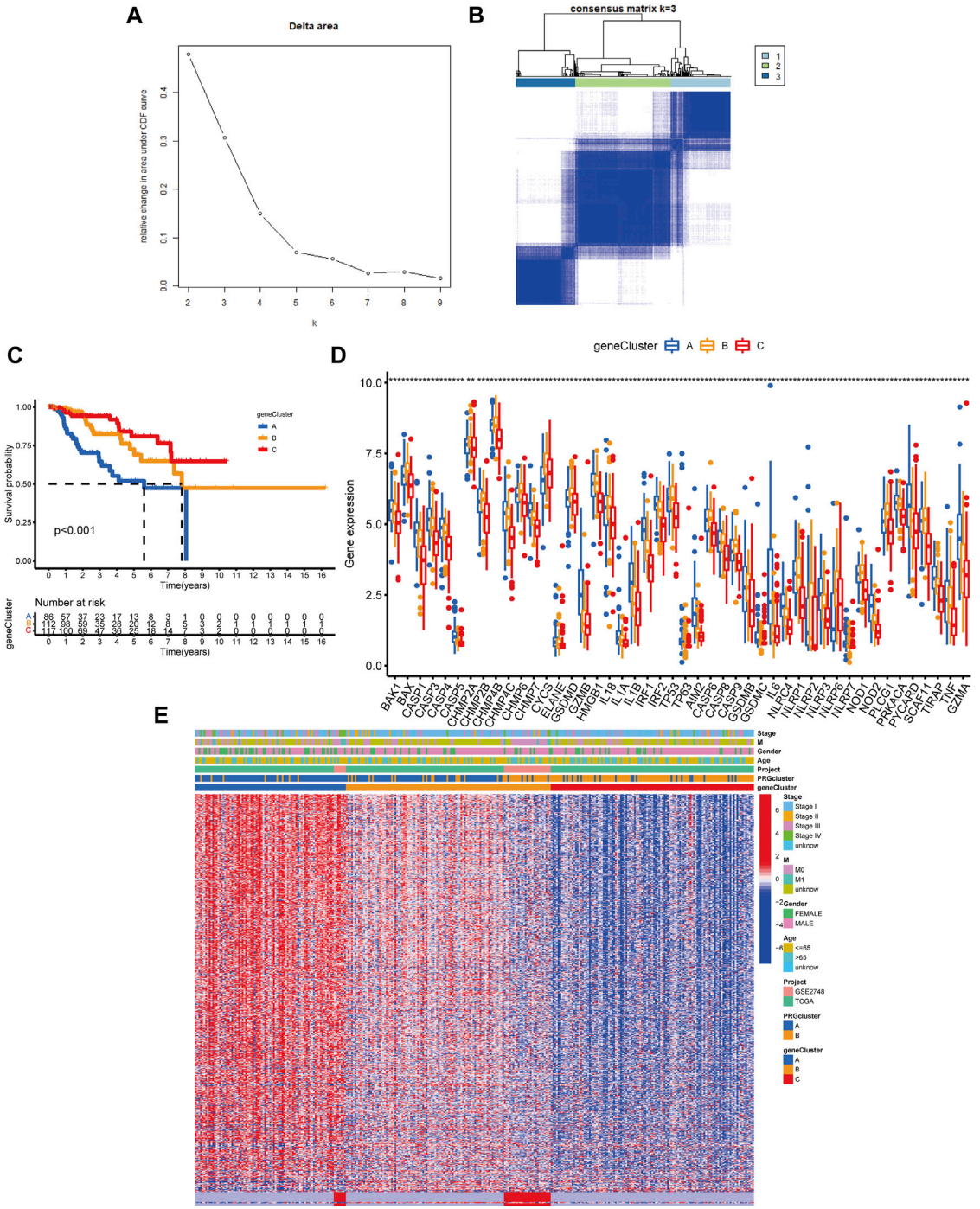



**FIGURE 5 |** Classification of genetic subgroups. **(A)** Changes in the area under the CDF curve. **(B)** The K = 3 consensus clustering matrix. **(C)** Survival curves for subgroups A, B, and C. **(D)** PRG expression between gene clusters. **(E)** Heatmap of genes based on unsupervised clustering techniques.

score groups. As it turned out, the mutation rates between the two groups varied significantly (64.18 percent and 59.44 percent). TTN, MUC16, MET, KMT2C, KIAA1109, SETD2, USH2A, MUC4, KMT2D, and WDFY3 were the remarkable mutant genes in the various risk groups (**Figures 9A,B**). TTN, MUC16, and MET mutations were more common in patients with lower PRG scores than in those with higher PRG values. KMT2C and KIAA1109, on the other hand, have the exact opposite mutation levels. KMT2C and KIAA1109 exhibited a greater frequency of missense mutations and a lower rate of frame-shift mutations in patients with high PRG score. We next chose

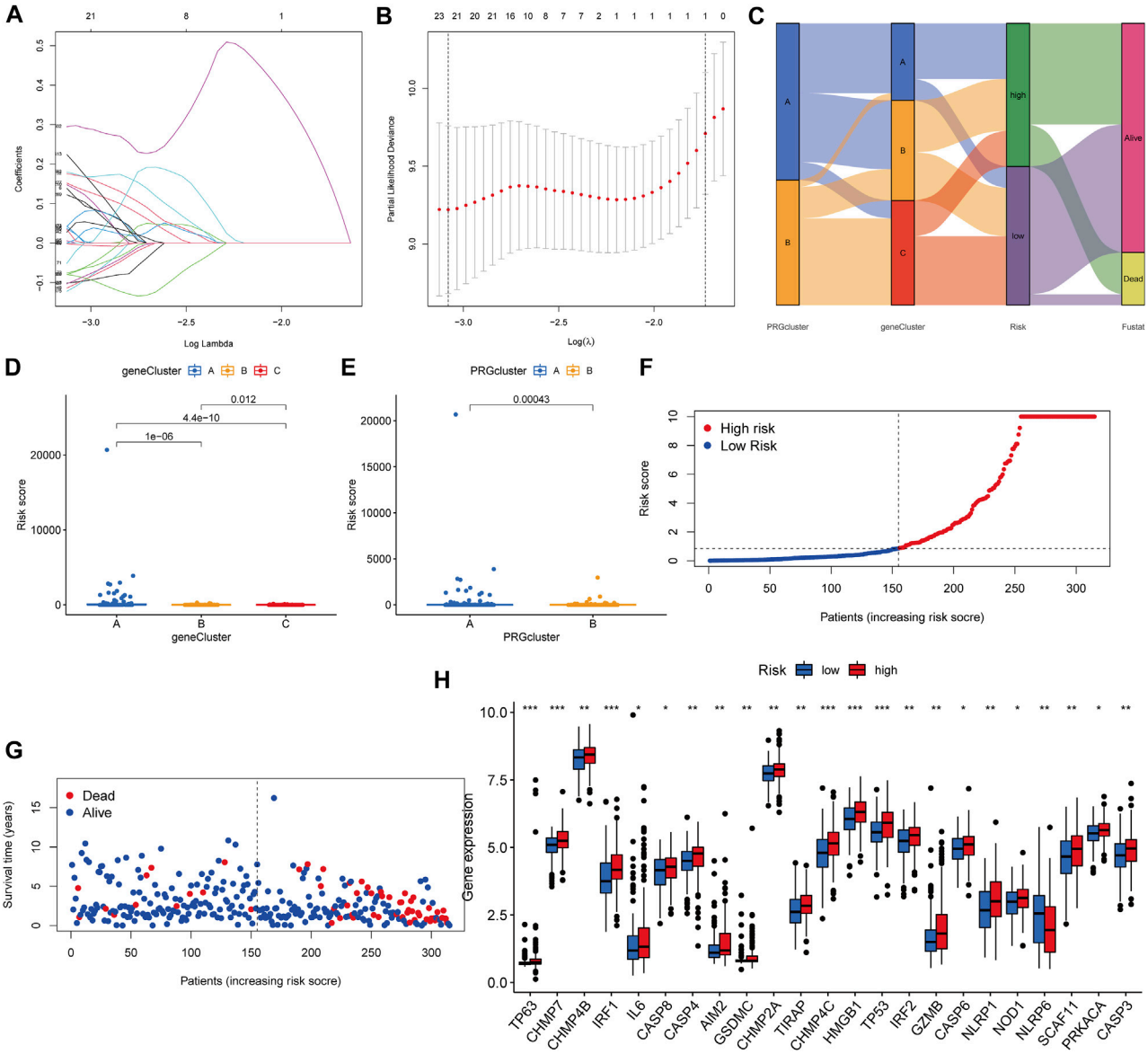



**FIGURE 6 |** Generation of the PRG signature scoring system. **(A–B)** LASSO regression analysis and partial likelihood deviance on prognostic genes. **(C)** An alluvial plot illustrating the link between the pyroptosis cluster, the gene cluster, and the pyroptosis score. **(D)** The PRG score in different genecluster subgroups. **(E)** The distribution of PRG scores in various PRGcluster subgroups. **(F–G)** The ranked dot plot depicts the distribution of PRG scores, whereas the scatter plot depicts the survival state. **(H)** Variations in PRG expression across risk categories.

chemotherapeutic medications that are presently used to treat KIRP to assess the sensitivities of individuals in distinct risk categories to these treatments. We discovered individuals with high PRG scores had lower IC50 values with A-443654 (specific inhibitor of Akt), ATRA (all-trans retinoic acid), AZD-2281 (olaparib), AZD8055 (mTOR inhibitor), and BI-D1870 (RSK inhibitor). However, the IC50 values of chemotherapeutics like AZD6244 (MEK inhibitor selumetinib) and BMS-708163 (gamma secretase inhibitor) were much lower in individuals with a lower PRG score. The findings suggested that PRGs were linked to medication sensitivity (**Figure 9C–N**). While most of the ideas about how to diagnose and treat cancer have come from research on clear cell RCC, the unique characteristics of KIRP may have significance for illnesses.

## DISCUSSION

KIRP is a heterogeneous subtype of RCC that differs from clear cell RCC in terms of clinicopathologic and molecular characteristics (Chen et al., 2019). When compared to clear cell RCC, KIRP is less likely to develop distant metastases upon diagnosis (Murugan et al., 2021). Because the majority of

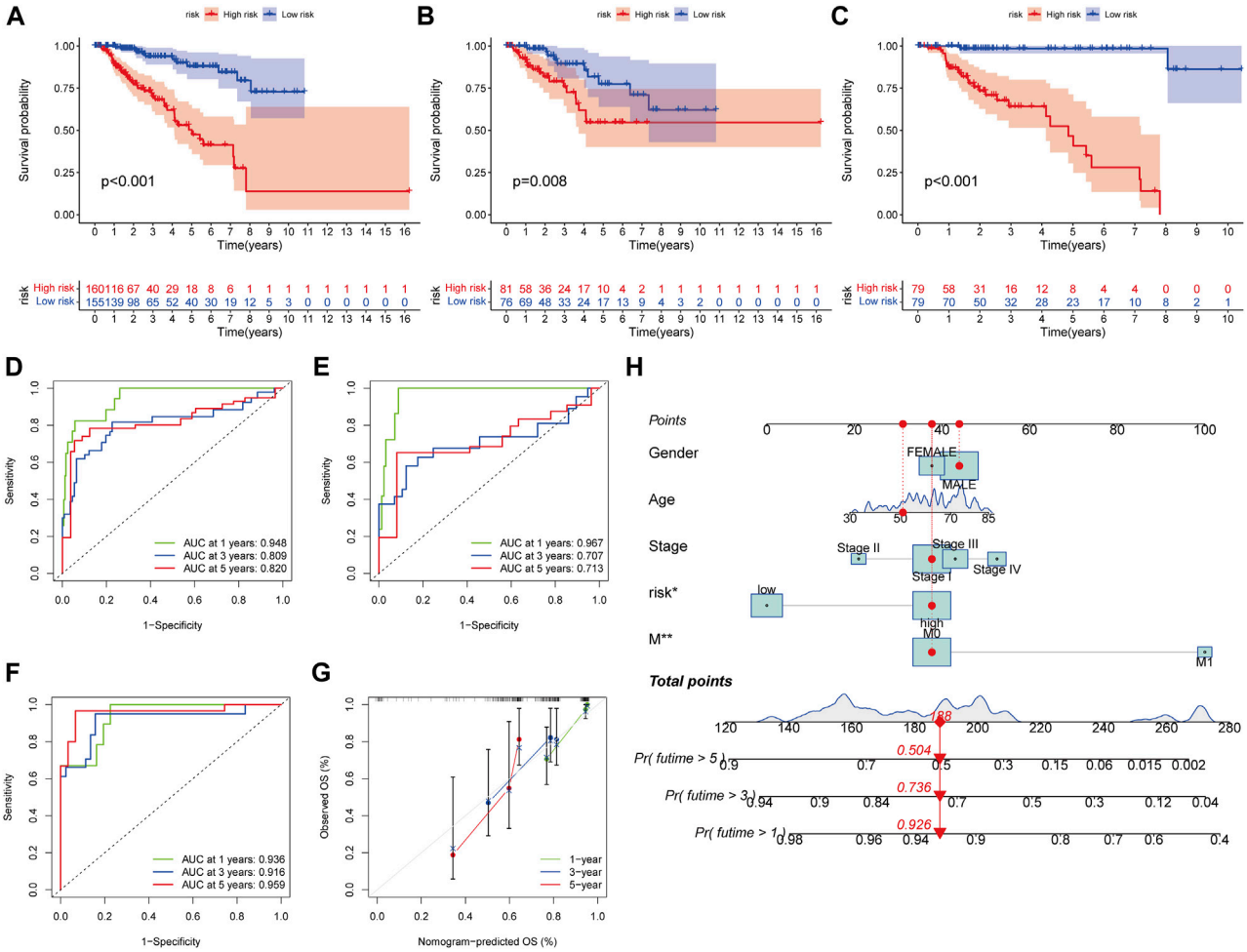



**FIGURE 7 |** Survival prediction and nomogram. **(A–C)** Kaplan-Meier analysis of the OS between the groups in the overall, training, testing, sets. **(D–F)** ROC curves to estimate the sensitivity and specificity of 1-, 3-, and 5-year survival based on the PRG score in the overall, training, testing sets. **(G)** Calibration plots for nomograms to predict OS at 1-, 3- and 5-year. **(H)** Nomogram for predicting the 1-, 3-, and 5-year OS of KIRP patients.

current knowledge about how to diagnose and treat renal tumors is derived from studies on clear cell RCC, the unique characteristics of KIRP may be critical to the illness (Mendhiratta et al., 2021; Yamamoto et al., 2021). For advanced KIRP, there is currently no agreed-upon treatment regimen, which poses a huge challenge for clinicians (Courthod et al., 2015). Molecularly targeted therapy and immunotherapy are also being tested in this subtype, and the results so far are favorable (Tachibana et al., 2021). KIRP has a better prognosis in the early stages than clear cell malignancies, but the situation changes drastically after the cancer has spread to distant sites (Zhao and Eyzaguirre, 2019). Pyroptosis has been linked to the emergence and progression of cancers in a growing body of data (Liu et al., 2021; Zhang et al., 2021; Zhou et al., 2021). Pyroptosis has been found to slow the development of tumors in multiple cancers (Sharma et al., 2019; Chen et al., 2020; Fenini et al., 2020). Pyroptosis can activate the immune system, inhibit tumor cell proliferation by altering the TME, and even kill malignant cells (Fu and Song, 2021; Sun et al., 2021; Wei et al., 2021). We don't know how it affects the KIRP microenvironment or immune function just yet.

Using the TCGA–KIRP and GSE2748 cohorts, we first investigated the genetic variants and expression of PRGs. However, even while PRGs had a low mutation rate, the majority of KIRP patients' genes were dysregulated and were linked with a poor prognosis. Second, based on 52 PRGs, we found two unique molecular subtypes. The TME's properties changed greatly across the two subtypes. The KIRP subtypes were also distinguished by a high level of immunological activation, which included the T cell and the B cell receptor signaling pathway. Differences in mRNA transcriptome across pyroptosis subtypes were also shown to be strongly connected to PRGs and immune-associated molecular mechanisms. Third, we employed unsupervised cluster analysis to separate the KIRP patients into three gene groups based on the genes that were chosen for inclusion. Therefore, our results imply that PRGs may be useful in predicting the clinical prognosis and therapeutic

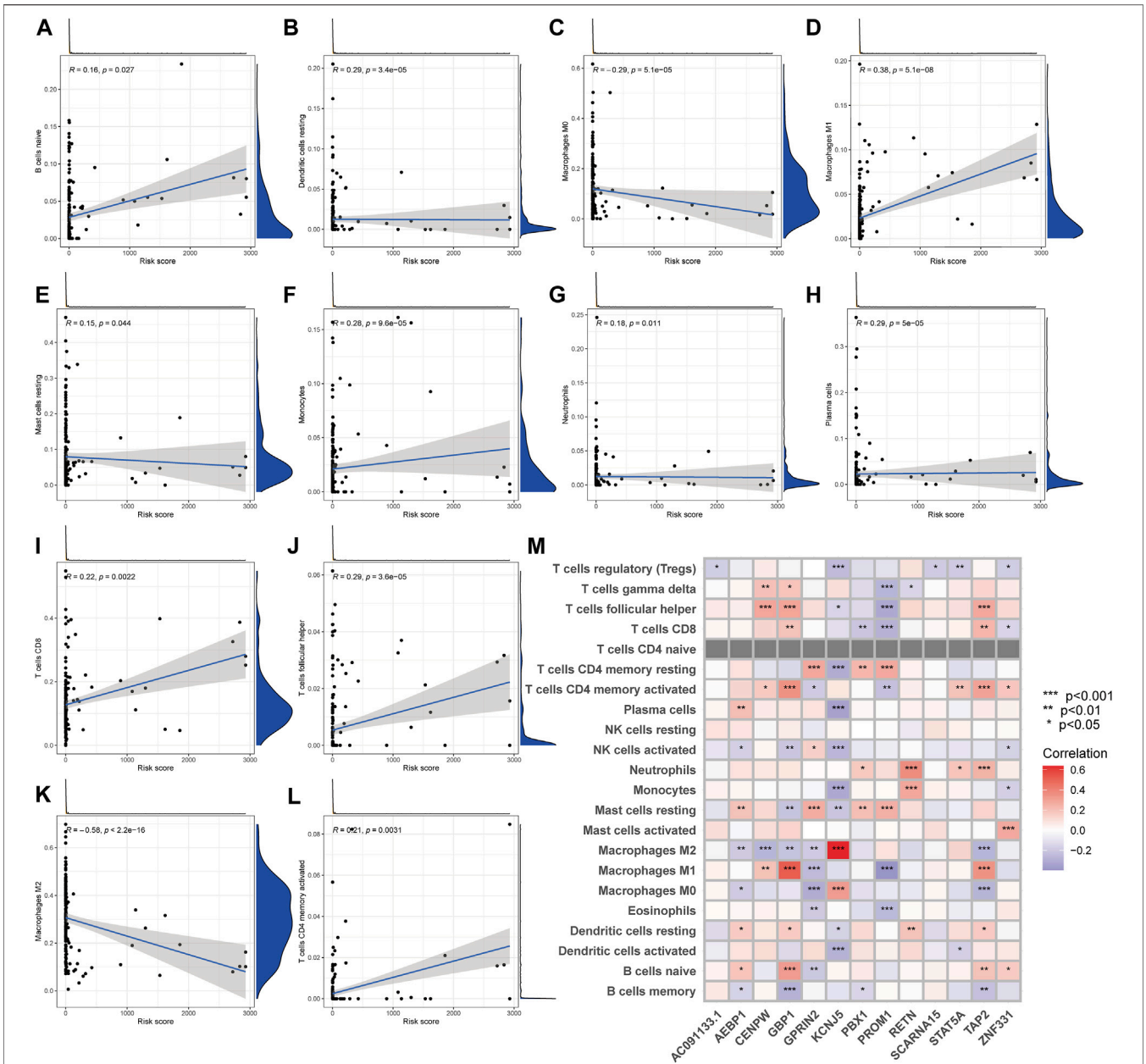



**FIGURE 8 |** TME comparison between the two subgroups. **(A–L)** The relationship between PRG score and immune cell type. **(M)** The relationship between the abundance of 22 immune cells and the abundance of 13 genes.

responsiveness of KIRP patients. Fourth, we developed a reliable and accurate predicted score and established the predictive power. Transcript levels of the score-containing genes were also evaluated in KIRP patients. On the basis of the median PRG score, patients were split into high and low-risk categories. When the PRG score is increased, the survival duration reduces and the recurrence rate rises. Finally, by combining the score and clinical features, we created a quantitative nomogram that increased performance and made the PRG score easier to use. The prognostic model may be used to stratify the prognosis of patients, which will aid in better identifying the pathogenesis of KIRP and will bring innovation for targeted therapeutics. KIRP patients may benefit from the development of a prognostic model that can be used to stratify their prognosis. It will also assist in the better recognition of the pathogenesis of KIRP and the development of novel targeted therapeutics.

Despite recent breakthroughs in immunotherapy, KIRP patients' outcomes remain heterogeneous, underlining the importance of TME in carcinogenesis and therapy. The cells of the TME participate in a variety of immune activations, including the pro-survival inflammatory response organized by malignancies (Seager et al., 2017). Additionally, evidence

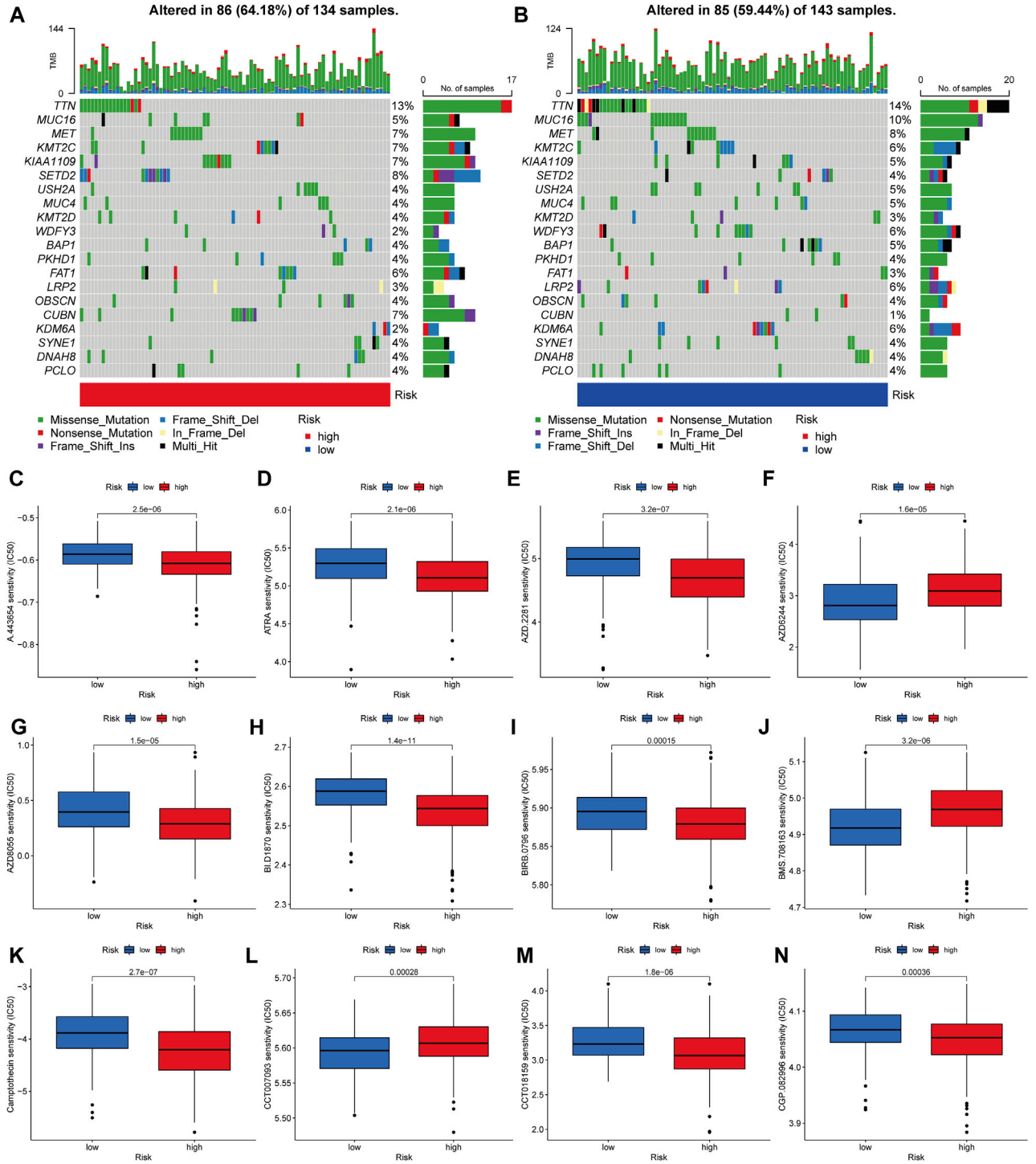



**FIGURE 9 |** Analyses of gene mutation and drug sensitivity. **(A–B)** The waterfall plot of gene mutation characteristics generated with different scores. **(C–N)** Correlations between PRG score and drug sensitivity.

demonstrates that the TME has a major influence on tumor growth, progression, and treatment resistance (Hinshaw and Shevde, 2019). In the current research, the subtype B defined by immunological inhibition was shown to be linked with a greater PRG score, while the subtype A marked by immune activation was found to be associated with a decreased score. The features of the tumor microenvironment, as well as the abundance of various tumor-infiltrating immune cells, were

found to vary considerably across the two molecular subtypes and distinct PRG scores. According to the findings of this research, PRGs play a significant role in the development of KIRP. It has been shown that subtype A and a lower PRG score are associated with greater infiltration of different activated T lymphocytes, suggesting that they play a beneficial role in the formation of KIRP. The presence of regulatory cells, which limit the immune response to tumors, has been linked with a bad prognosis (Tanaka and Sakaguchi, 2017). This is consistent with our findings that patients with subgroup B and elevated scores had more Tregs in the TME than those in the low-risk category.

Advanced KIRP treatment options have progressed significantly during the 1980s and 1990s, from cytokine and cytotoxic chemotherapy to genetic targeted treatments and immune checkpoint inhibitors at present (Ronnen et al., 2006). Despite biological differences, the early use of VEGF and mTOR inhibitors in KIRP was primarily supported by data from medical trials examining their effectiveness in clear cell RCC patients. Over the last several years, over a dozen studies have been conducted to specifically target KIRP, either on its own or in comprehensive experiments incorporating different renal tumors. Encouragingly, a series of cohort studies have revealed that these treatments are efficacious in treating KIRP (Motzer et al., 2002; Gore et al., 2009). In an attempt to elucidate the molecular basis of KIRP in the population, a comprehensive genomic analysis was performed on patients with advanced KIRP (D'Avella et al., 2020). The most frequently found mutations in individuals with type 1 illness were MET (33%), CKDN2A/B (18%), and EGFR (18%). CDKN2A/B (18%) and MET (18%) mutations were the most prevalent type 2 mutations. This work emphasizes the role of MET mutations in KIRP and another possible CDKN2A mutation for further investigation. CDK4/6 inhibitors are now licensed for the treatment of metastatic breast cancer and may be used in KIRP in the future. Whereas the best order of systemic medications is uncertain, limited therapeutic evidence shows that first-line VEGF therapy is associated with superior results when compared to mTOR inhibitors. Additionally, the ESPN study found no difference in survival time between everolimus and sunitinib as first-line therapy (Tannir et al., 2016). Due to the poor effectiveness of currently available therapies for hereditary papillary renal cell carcinoma, it is critical to understand how MET inhibitors affect it. Cabozantinib (Choueiri et al., 2016; Prisciandaro et al., 2019) and crizotinib (Schöffski et al., 2017), both of which are c-MET tyrosine kinase inhibitors, are two further targeted treatments with potential effectiveness. The objective response rate was 25.4 percent in a sample of 118 KIRP with an 11-month median follow-up, and the median duration of response was not attained, indicating potential anticancer activity in a tumor that had previously been deemed resistant to immunotherapy (McDermott et al., 2021). Combining immunotherapy with specific medicines like MET inhibitors is now being tested in clinical studies. Only p53 deletion, which has been substantially linked with poor survival, has demonstrated an association with clinical outcomes at the molecular level (Leroy et al., 2002; Perret et al., 2008; Ricketts et al., 2018). Mutations in the TP53 gene, CDKN2A gene, PBRM1 gene, and the hypermethylation genomic cluster are related to survival in KIRP (Pitra et al., 2019).

## CONCLUSION

Our thorough examination of PRGs revealed a complex regulatory system via which they influence TME, clinical and pathological characteristics, and prognosis. Apart from this, we also looked at how PRGs may help with immunotherapy as well as other forms of targeted therapy. The findings underscore the essential therapeutic implications of PRGs and provide novel strategies for targeting immunotherapy therapies for KIRP patients.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

CZ wrote the paper. RH edited the paper. XX made the images out. All authors contributed to the article and approved the submitted version.

## REFERENCES


Akhtar, M., Al-Bozom, I. A., and Al Hussain, T. (2019). Papillary Renal Cell Carcinoma (PRCC): An Update. *Adv. Anat. Pathol.* 26, 124–132. doi:10.1097/pap.0000000000000220

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol. Biol. (Clifton, N.J.)* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Chen, Q., Cheng, L., and Li, Q. (2019). The Molecular Characterization and Therapeutic Strategies of Papillary Renal Cell Carcinoma. *Expert Rev. anticancer Ther.* 19, 169–175. doi:10.1080/14737140.2019.1548939

Chen, Z., He, M., Chen, J., Li, C., and Zhang, Q. (2020). Long Non-coding RNA SNHG7 I-nhibits NLRP3-dependent P-yroptosis by T-argeting the miR-34a/SIRT1 axis in L-iver C-ancer. *Oncol. Lett.* 20, 893–901. doi:10.3892/ol.2020.11635

Choueiri, T. K., Escudier, B., Powles, T., Tannir, N. M., Mainwaring, P. N., Rini, B. I., et al. (2016). Cabozantinib versus Everolimus in Advanced Renal Cell Carcinoma (METEOR): Final Results from a Randomised, Open-Label, Phase 3 Trial. *Lancet Oncol.* 17, 917–927. doi:10.1016/s1470-2045(16)30107-3

Courthod, G., Tucci, M., Di Maio, M., and Scagliotti, G. V. (2015). Papillary Renal Cell Carcinoma: A Review of the Current Therapeutic Landscape. *Crit. Rev. oncology/hematology* 96, 100–112. doi:10.1016/j.critrevonc.2015.05.008

D'Avella, C., Abbosh, P., Pal, S. K., and Geynisman, D. M. (2020). Mutations in Renal Cell Carcinoma. *Urol. Oncol. Semin. Original Invest.* 38, 763–773. doi:10.1016/j.urolonc.2018.10.027

Fenini, G., Karakaya, T., Hennig, P., Di Filippo, M., and Beer, H.-D. (2020). The NLRP1 Inflammasome in Human Skin and beyond. *Ijms* 21, 4788. doi:10.3390/ijms21134788

Fu, X.-W., and Song, C.-Q. (2021). Identification and Validation of Pyroptosis-Related Gene Signature to Predict Prognosis and Reveal Immune Infiltration in Hepatocellular Carcinoma. *Front. Cel Dev. Biol.* 9, 748039. doi:10.3389/fcell.2021.748039

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

Gene Ontology Consortium (2015). Gene Ontology Consortium: Going Forward. *Nucleic Acids Res.* 43, D1049–D1056. doi:10.1093/nar/gku1179

Gore, M. E., Szczylik, C., Porta, C., Bracarda, S., Bjarnason, G. A., Oudard, S., et al. (2009). Safety and Efficacy of Sunitinib for Metastatic Renal-Cell Carcinoma: an Expanded-Access Trial. *The Lancet OncologyOncology* 10, 757–763. doi:10.1016/s1470-2045(09)70162-7

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

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: New Perspectives on Genomes, Pathways, Diseases and Drugs. *Nucleic Acids Res.* 45, D353–d361. doi:10.1093/nar/gkw1092

Karki, R., and Kanneganti, T.-D. (2019). Diverging Inflammasome Signals in Tumorigenesis and Potential Targeting. *Nat. Rev. Cancer* 19, 197–214. doi:10.1038/s41568-019-0123-y

Leroy, X., Zini, L., Leteurtre, E., Zerimech, F., Porchet, N., Aubert, J.-P., et al. (2002). Morphologic Subtyping of Papillary Renal Cell Carcinoma: Correlation with Prognosis and Differential Expression of MUC1 between the Two Subtypes. *Mod. Pathol.* 15, 1126–1130. doi:10.1097/01.Mp.0000036346.88874.25

Liu, D., Yang, X., and Wu, X. (2021). Tumor Immune Microenvironment Characterization Identifies Prognosis and Immunotherapy-Related Gene Signatures in Melanoma. *Front. Immunol.* 12, 663495. doi:10.3389/fimmu.2021.663495

Man, S. M., and Kanneganti, T.-D. (2015). Regulation of Inflammasome Activation. *Immunol. Rev.* 265, 6–21. doi:10.1111/imr.12296

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

McDermott, D. F., Lee, J.-L., Bjarnason, G. A., Larkin, J. M. G., Gafanov, R. A., Kochenderfer, M. D., et al. (2021). Open-Label, Single-Arm Phase II Study of Pembrolizumab Monotherapy as First-Line Therapy in Patients with Advanced Clear Cell Renal Cell Carcinoma. *Jco* 39, 1020–1028. doi:10.1200/jco.20.02363

Mendhiratta, N., Muraki, P., Sisk, A. E., Jr., and Shuch, B. (2021). Papillary Renal Cell Carcinoma: Review. *Urol. Oncol. Semin. Original Invest.* 39, 327–337. doi:10.1016/j.urolonc.2021.04.013

Meng, J., Huang, X., Qiu, Y., Zheng, X., Huang, J., Wen, Z., et al. (2021). Pyroptosis-related Gene Mediated Modification Patterns and Immune Cell Infiltration Landscapes in Cutaneous Melanoma to Aid Immunotherapy. *Aging* 13, 24379–24401. doi:10.18632/aging.203687

Moch, H., Cubilla, A. L., Humphrey, P. A., Reuter, V. E., and Ulbright, T. M. (2016). The 2016 WHO Classification of Tumours of the Urinary System and Male Genital Organs-Part A: Renal, Penile, and Testicular Tumours. *Eur. Urol.* 70, 93–105. doi:10.1016/j.eururo.2016.02.029

Motzer, R. J., Bacik, J., Mariani, T., Russo, P., Mazumdar, M., and Reuter, V. (2002). Treatment Outcome and Survival Associated with Metastatic Renal Cell Carcinoma of non-clear-cell Histology. *Jco* 20, 2376–2381. doi:10.1200/jco.2002.11.123

Murugan, P., Jia, L., Dinatale, R. G., Assel, M., Benfante, N., Al-Ahmadie, H. A., et al. (2021). Papillary Renal Cell Carcinoma: a Single Institutional Study of 199 Cases Addressing Classification, Clinicopathologic and Molecular Features, and Treatment Outcome. *Mod. Pathol.* 23. doi:10.1038/s41379-021-00990-9

Park, S. Y. (2018). Nomogram: An Analogue Tool to Deliver Digital Knowledge. *J. Thorac. Cardiovasc. Surg.* 155, 1793. doi:10.1016/j.jtcvs.2017.12.107

Perret, A. G., Clemencon, A., Li, G., Tostain, J., and Peoc'h, M. (2008). Differential Expression of Prognostic Markers in Histological Subtypes of Papillary Renal Cell Carcinoma. *BJU Int.* 102, 183–187. doi:10.1111/j.1464-410X.2008.07605.x

Pitra, T., Pivovarcikova, K., Alaghehbandan, R., and Hes, O. (2019). Chromosomal Numerical Aberration Pattern in Papillary Renal Cell Carcinoma: Review Article. *Ann. Diagn. Pathol.* 40, 189–199. doi:10.1016/j.anndiagpath.2017.11.004

Prisciandaro, M., Ratta, R., Massari, F., Fornarini, G., Caponnetto, S., Iacovelli, R., et al. (2019). Safety and Efficacy of Cabozantinib for Metastatic Nonclear Renal Cell Carcinoma. *Am. J. Clin. Oncol.* 42, 42–45. doi:10.1097/coc.0000000000000478

Ricketts, C. J., De Cubas, A. A., Fan, H., Smith, C. C., Lang, M., Reznik, E., et al. (2018). The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma. *Cell Rep* 23, 3698–4326.e315. doi:10.1016/j.celrep.2018.03.07510.1016/j.celrep.2018.06.032

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007

Ronnen, E. A., Kondagunta, G. V., Ishill, N., Spodek, L., Russo, P., Reuter, V., et al. (2006). Treatment Outcome for Metastatic Papillary Renal Cell Carcinoma Patients. *Cancer* 107, 2617–2621. doi:10.1002/cncr.22340

Schöffski, P., Wozniak, A., Escudier, B., Rutkowski, P., Anthoney, A., Bauer, S., et al. (2017). Crizotinib Achieves Long-Lasting Disease Control in Advanced Papillary Renal-Cell Carcinoma Type 1 Patients with MET Mutations or Amplification. EORTC 90101 CREATE Trial. *Eur. J. Cancer* 87, 147–163. doi:10.1016/j.ejca.2017.10.014

Seager, R. J., Hajal, C., Spill, F., Kamm, R. D., and Zaman, M. H. (2017). Dynamic Interplay between Tumour, Stroma and Immune System Can Drive or Prevent Tumour Progression. *Converg. Sci. Phys. Oncol.* 3, 034002. doi:10.1088/2057-1739/aa7e86

Sharma, B. R., Karki, R., and Kanneganti, T. D. (2019). Role of AIM2 Inflammasome in Inflammatory Diseases, Cancer and Infection. *Eur. J. Immunol.* 49, 1998–2011. doi:10.1002/eji.201848070

Song, W., Ren, J., Xiang, R., Kong, C., and Fu, T. (2021). Identification of Pyroptosis-Related Subtypes, the Development of a Prognosis Model, and Characterization of Tumor Microenvironment Infiltration in Colorectal Cancer. *Oncoimmunology* 10, 1987636. doi:10.1080/2162402x.2021.1987636

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

Sun, Z., Jing, C., Guo, X., Zhang, M., Kong, F., Wang, Z., et al. (2021). Comprehensive Analysis of the Immune Infiltrates of Pyroptosis in Kidney Renal Clear Cell Carcinoma. *Front. Oncol.* 11, 716854. doi:10.3389/fonc.2021.716854

Tachibana, H., Kondo, T., Ishihara, H., Fukuda, H., Yoshida, K., Takagi, T., et al. (2021). Modest Efficacy of Nivolumab Plus Ipilimumab in Patients with Papillary Renal Cell Carcinoma. *Jpn. J. Clin. Oncol.* 51, 646–653. doi:10.1093/jjco/hyaa229(2021)

Tanaka, A., and Sakaguchi, S. (2017). Regulatory T Cells in Cancer Immunotherapy. *Cell Res* 27, 109–118. doi:10.1038/cr.2016.151

Tannir, N. M., Jonasch, E., Albiges, L., Altinmakas, E., Ng, C. S., Matin, S. F., et al. (2016). Everolimus versus Sunitinib Prospective Evaluation in Metastatic Non-Clear Cell Renal Cell Carcinoma (ESPN): A Randomized Multicenter Phase 2 Trial. *Eur. Urol.* 69, 866–874. doi:10.1016/j.eururo.2015.10.049

Tordjman, M., Dbjay, J., Chamouni, A., Morini, A., Timsit, M. O., Mejean, A., et al. (2020). Clear Cell Papillary Renal Cell Carcinoma: A Recent Entity with Distinct Imaging Patterns. *Am. J. Roentgenology* 214, 579–587. doi:10.2214/ajr.19.21681

Vikram, R., Ng, C. S., Tamboli, P., Tannir, N. M., Jonasch, E., Matin, S. F., et al. (2009). Papillary Renal Cell Carcinoma: Radiologic-Pathologic Correlation and Spectrum of Disease. *RadioGraphics* 29, 741–754. doi:10.1148/rg.293085190

Wang, B., and Yin, Q. (2017). AIM2 Inflammasome Activation and Regulation: A Structural Perspective. *J. Struct. Biol.* 200, 279–282. doi:10.1016/j.jsb.2017.08.001

Wei, R., Li, S., Yu, G., Guan, X., Liu, H., Quan, J., et al. (2021). Deciphering the Pyroptosis-Related Prognostic Signature and Immune Cell Infiltration Characteristics of Colon Cancer. *Front. Genet.* 12, 755384. doi:10.3389/fgene.2021.755384

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: a Class Discovery Tool with Confidence Assessments and Item Tracking. *Bioinformatics (Oxford, England)* 26, 1572–1573. doi:10.1093/bioinformatics/btq170

Williamson, S. R. (2021). Clear Cell Papillary Renal Cell Carcinoma: an Update after 15 Years. *Pathology* 53, 109–119. doi:10.1016/j.pathol.2020.10.002

Wu, J., Zhu, Y., Luo, M., and Li, L. (2021). Comprehensive Analysis of Pyroptosis-Related Genes and Tumor Microenvironment Infiltration Characterization in Breast Cancer. *Front. Immunol.* 12, 748221. doi:10.3389/fimmu.2021.748221

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

Xiang, R., Ge, Y., Song, W., Ren, J., Kong, C., and Fu, T. (2021). Pyroptosis Patterns Characterized by Distinct Tumor Microenvironment Infiltration Landscapes in Gastric Cancer. *Genes* 12, 1535. doi:10.3390/genes12101535

Yamamoto, T., Gulanbar, A., Hayashi, K., Kohno, A., Komai, Y., Yonese, J., et al. (2021). Is Hypervascular Papillary Renal Cell Carcinoma Present? *Abdom. Radiol. (N Y)* 46, 1687–1693. doi:10.1007/s00261-020-02809-8

Zhang, Y., He, R., Lei, X., Mao, L., Jiang, P., Ni, C., et al. (2021). A Novel Pyroptosis-Related Signature for Predicting Prognosis and Indicating Immune Microenvironment Features in Osteosarcoma. *Front. Genet.* 12, 780780. doi:10.3389/fgene.2021.780780

Zhao, J., and Eyzaguirre, E. (2019) Clear Cell Papillary Renal Cell Carcinoma. 143, 1154–1158. doi:doi:10.5858/arpa.2018-0121-RS

Zhao, Y., Li, M-C., Konaté, M. M., Chen, L., Das, B., Karlovich, C., et al. (2021). TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-Seq Data from the NCI Patient-Derived Models Repository. *J. translational Med.* 19, 269. doi:10.1186/s12967-021-02936-w

Zheng, Z., and Li, G. (2020). Mechanisms and Therapeutic Regulation of Pyroptosis in Inflammatory Diseases and Cancer. *Ijms* 21, 1456. doi:10.3390/ijms21041456

Zhou, S., Sun, Y., Chen, T., Wang, J., He, J., Lyu, J., et al. (2021). The Landscape of the Tumor Microenvironment in Skin Cutaneous Melanoma Reveals a Prognostic and Immunotherapeutically Relevant Gene Signature. *Front. Cel Dev. Biol.* 9, 739594. doi:10.3389/fcell.2021.739594

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

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

frontiers | Frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 01 April 2022
doi: 10.3389/fmolb.2022.778857

# Comprehensive Analysis on Prognosis and Immune Infiltration of Lysyl Oxidase Family Members in Pancreatic Adenocarcinoma With Experimental Verification

*Chao Jiang*[1,2], *Meng Wang*[1], *Weikai Yao*[3], *Guoyue Lv*[1], *Xueyan Liu*[4]* and *Guangyi Wang*[1]*

[1]Department of Hepatobiliary Pancreatic Surgery I, The First Hospital of Jilin University, Changchun, China, [2]Multi-Organ Transplant Program, University Health Network, Toronto, ON, Canada, [3]Department of Pathology, The First Hospital of Jilin University, Changchun, China, [4]Cardiovascular Department, China-Japan Union Hospital of Jilin University, Changchun, China

OPEN ACCESS

**Edited by:**
Pedro Borralho,
Novartis Oncology, Portugal

**Reviewed by:**
Emanuel Della Torre,
Ospedale San Raffaele (IRCCS), Italy
Marta Bento Afonso,
Research Institute for Medicines (iMed.ULisboa), Portugal

***Correspondence:**
Xueyan Liu
liuxueyan@jlu.edu.cn
Guangyi Wang
guangyi@jlu.edu.cn

**Specialty section:**
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

**Received:** 17 September 2021
**Accepted:** 07 March 2022
**Published:** 01 April 2022

**Citation:**
Jiang C, Wang M, Yao W, Lv G, Liu X and Wang G (2022) Comprehensive Analysis on Prognosis and Immune Infiltration of Lysyl Oxidase Family Members in Pancreatic Adenocarcinoma With Experimental Verification. Front. Mol. Biosci. 9:778857. doi: 10.3389/fmolb.2022.778857

**Background:** Pancreatic adenocarcinoma (PDAC) is the most aggressive among all solid malignancies with delayed disease detection and limited effective treatment. However, due to the intricate heterogeneity and exclusive tumor microenvironment (TME), the development of effective therapy has been facing enormous challenges. The lysyl oxidases (LOXs) underpin the shaping of the TME to promote cancer growth, metastasis and modulate response to treatment.

**Materials and Methods:** The mRNA expression, prognostic, and clinicopathological data for LOXs in PDAC from multiple open-access databases were summarized and analyzed. The protein expression was verified by immunohistochemistry (IHC). Co-expressed genes of LOXs were predicted and elaborated by LinkedOmics. Functional enrichment analysis of LOXs co-expressed genes was performed using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). TIMER and TISIDB were applied to analyze the relationship between LOXs expression and immune infiltration.

**Results:** The mRNA expression levels of LOX, LOXL1 and LOXL2 were significantly higher in PDAC, the expression levels of LOXL3 and LOXL4 were contrary in different databases. High mRNA levels of LOX and LOXL2 were associated with advanced PDAC stage, while elevated LOX and LOXL3 expression correlated with high tumor grade. The IHC staining showed higher expression levels of LOX, LOXL1 and LOXL2, lower expression level of LOXL3 in PDAC tissues, while the protein expression of LOXL4 made no difference. Functional enrichment analysis showed a close relationship with extracellular matrix (ECM) regulation, except that LOXL3 and its ligands were highly associated with immune-related functions. Further analysis suggested that LOX and LOXL3 strongly correlated with tumor-infiltrating lymphocytes (TILs), various immune signatures, and immune checkpoints. Finally, survival analysis revealed high LOX and LOXL2 expression predicted worse overall survival (OS), progression-free interval (PFI), and disease-specific survival (DSS).

**Conclusion:** These findings indicated that the LOX family, especially LOX and LOXL2, might have a prospective value in PDAC oncogenesis, and they may become prognostic biomarkers, revealing a promising field in targeted therapy.

**Keywords: pancreatic adenocarcinoma, lysyl oxidase, LOX family members, bioinformatics analysis, tumor microenvironment, immune infiltration, prognostic**

## INTRODUCTION

Pancreatic adenocarcinoma (PDAC), a highly fatal disease with a poor prognosis, ranks as the seventh leading cause of cancer mortality globally (Sung et al., 2021), is projected to become the second leading cancer-related death in America and Europe in 2030 (Rahib et al., 2014; Quante et al., 2016). However, over the past decades, albeit considerable advances have been made in the diagnosis, surgery, chemotherapy, radiotherapy and systemic therapy, however, due to non-specific symptoms, lacking reliable markers, low surgical resection rate and therapy resistance, only months of overall survival (OS) achieved (Conroy et al., 2011; Winter et al., 2012), with the 5-years survival remaining a grim 10% (Siegel et al., 2021). Therefore, further investigation of pathological features and the pathogenesis of PDAC is imperative for exceptional therapeutic strategies.

Such poor outcomes of PDAC have fueled ongoing efforts to exploit the tumor microenvironment (TME). The TME is a highly complex ecosystem composed of cancer cells, stromal cells, immune cells, and the extracellular matrix (ECM) interaction. Growing evidence has illustrated the pivotal roles of TME components in immune suppression and tumor progression. It is worth noting that the surrounding ECM has long been implicated in accelerating PDAC progression by directly promoting cellular transformation and metastasis. Furthermore, the ECM and its reorganization are of paramount importance in the evolution of both tumor and immune microenvironment and their interaction (Setargew et al., 2021). These alterations may greatly affect the biological behavior of tumor cells and stromal cells, and ultimately influence the prognosis of PDAC patients.

Lysyl oxidase (LOX) family includes five copper-dependent amino enzymes: LOX and LOX-like (LOXL) 1-4 (Molnar et al., 2003; Barker et al., 2012), of which the canonical function is catalyzing collagen and elastin cross-linking (Harris, 1976; Kagan and Li, 2003; Zhao et al., 2018). Lysyl Oxidases (LOXs) contribute to ECM stiffness and homeostasis, and their dysregulation is involved in several diseases, including tissue fibrosis (Al-U'datt et al., 2019; Li et al., 2020) and cancer (Wang et al., 2016; Tan et al., 2020). In reality, the activity of LOXs integrates a complex network within the TME linking the ECM and the immunological components. Intriguingly, LOXs play different roles in various tumor progression, exhibiting either enhancing (Wilgus et al., 2011; Kasashima et al., 2016)or suppressing (Johnston and Lopez, 2018; Zhao et al., 2020) function in different types of human malignancies.

It has been demonstrated that overexpressed LOXs are critically involved in the PDAC progression. Le Calve et al. (Le Calve et al., 2016) found higher mRNA expression of LOX, LOXL1 and LOXL2 in PDAC. Ma et al. (Ma et al., 2019) found that LOX expression might be an independent prognostic factor of progression-free survival (PFI) in PDAC patients. Higher LOXL2 expression elevated the invasiveness of PDAC cells and correlated with lower survival of PDAC patients (Kim et al., 2019). However, the potential roles of LOXs in the PDAC immune response have not been elucidated. Herein, various large-scale open-accessed databases were applied to explore the expression pattern, cancer-related functional aspects, prognostic value, and relationship with the immune infiltrating of LOXs in PDAC.

## MATERIALS AND METHODS

### Oncomine

Oncomine gene expression array dataset (https://www.oncomine.org/resource/login.html) (Rhodes et al., 2004; Rhodes et al., 2007) was applied to analyze the transcripts levels of LOXs in different cancers. The data were compared by the *t*-test with a cut-off *p*-value of 0.001 and a fold change of 2.

### GEPIA

Gene Expression Profiling Interactive Analysis (GEPIA, http://gepia2.cancer-pku.cn/#index) (Tang et al., 2017) was used to analyze the relative transcriptional expression of LOXs between PDAC tissues and normal tissues. $p < 0.01$ was considered statistically significant.

### Human Protein Atlas

The Human Protein Atlas (HPA, https://www.proteinatlas.org/) (Uhlen et al., 2015; Uhlen et al., 2017) was introduced to explore the direct contrast of protein expression of LOXL1-4 between normal tissues and PDAC tissues based on immunohistochemistry (IHC). The data of LOX were not available in HPA.

### UALCAN

The UALCAN (http://ualcan.path.uab.edu/index.html) (Chandrashekar et al., 2017) was used to explore the protein expression of LOXs in PADC. The protein assembly data relative abundance was obtained from the Pancreatic Ductal Adenocarcinoma (PDAC) Discovery Study of Clinical Proteomics Tumor Analysis Consortium (CPTAC, https://proteomics.cancer.gov/programs/cptac). $p < 0.01$ was considered statistically significant.

### Immunohistochemistry

As the feedback regulation loop of LOXs, the regulation of transcripts level and protein level are not necessarily the same. To further explore the LOXs protein level in PDAC patients, LOXs expression was

evaluated through IHC in PDAC tissue and paired adjacent tissue($n$ = 6). The study was approved by the First Hospital of Jilin University Ethics Committee (2019-180). All patients provided written informed consent for the subsequent use of their resected tissues in this study. All clinical investigations were conducted following the principles of the Declaration of Helsinki. The PDAC tissues and paired adjacent tissues were prepared into 4 μm paraffin sections and incubated overnight at 4 °C with rabbit polyclonal antibodies against LOXs (1: 150, Novus Biologicals, USA. LOX: NB100-2527; LOXL1: H00004016-D01P; LOXL2: NBP1-32954; LOXL3: NBP2-75964; LOXL4: NBP2-32692). After washing, the sections were conjugated with ImmPRESS anti-Rabbit kit for 30 min (LSBio, USA. LS-J1066). The negative controls demonstrated the specific signals of the primary antibody. Subsequently, all fields were observed under light microscopy. The average integrated option density (AOD) of LOXs was determined using ImageJ software (National Institutes of Health, USA).

## Metascape

Metascape (https://metascape.org/gp/index.html#/main) (Zhou et al., 2019) was applied to explore the enrichment of the co-expressed genes with LOXs categorized by process and pathway. The most statistically significant 20 terms were shown in the diagram. Terms with $p$ < 0.01, enrichment factor>1.5, and minimum count three were considered significant. The protein-protein interaction (PPI) was conducted using the Molecular Complex Detection (MCODE) algorithm.

## LinkedOmics Database

The LinkedOmics database (http://www.linkedomics.org/login.php) (Vasaikar et al., 2018), a helpful third-party online tool containing The Cancer Genome Atlas (TCGA) data(Tomczak et al., 2015), was used to identify the co-expressed genes of LOXs. The correlation was expressed by the Pearson coefficient.

## TIMER

Tumor Immune Estimation Resource (TIMER) database (http://timer.cistrome.org/) (Li et al., 2017; Li et al., 2020) was applied to assess the association of LOXs expression with tumor-infiltrating lymphocytes (TILs) via the Gene module and immune biomarker sets (Siemers et al., 2017) (including immune checkpoint gene sets) via Correlation module. The correlation was adjusted by tumor purity and expressed by the Spearman coefficient (Chen et al., 2020).

## TISIDB

TISIDB (http://cis.hku.hk/TISIDB/) (Ru et al., 2019) was introduced to validate the correlation of LOXs expression with the abundance of TILs, and the relationship of LOXs expression with clinical characteristics in PDAC patients. The correlation was expressed by the Spearman coefficient.

## Survival Analysis

The mRNA expression of LOXs in PDAC was obtained from c-BioPortal (www.cbioportal.org) (Cerami et al., 2012; Gao et al., 2013), and the clinical data of PDAC was extracted from integrated clinical data resources(Liu et al., 2018) to perform survival analysis. Patients were divided into two groups based on the median value of mRNA expression.

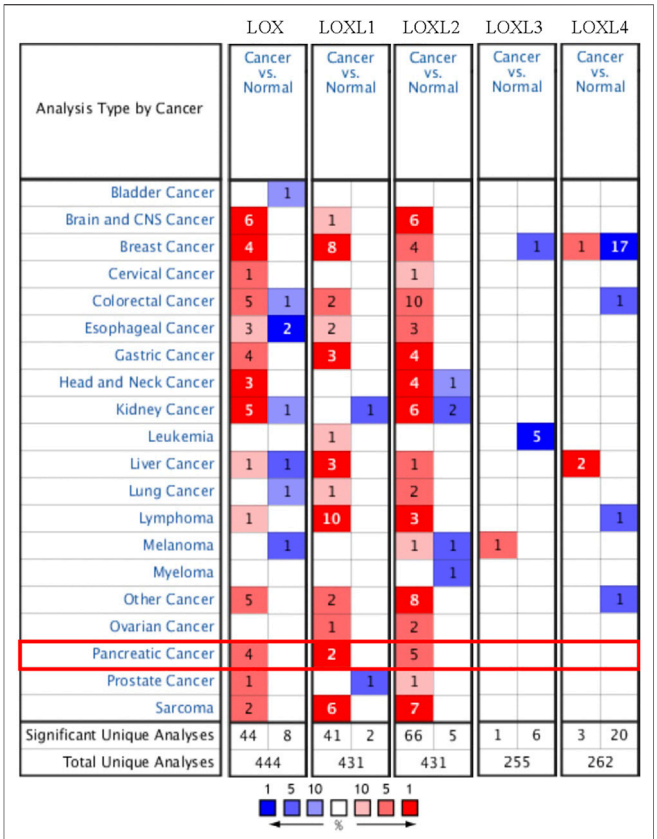



**FIGURE 1 |** The transcriptional expression of distinct LOX family members in 20 types of cancers (Oncomine database). The numbers in each table represented the studies with statistically significant tumor tissue mRNA overexpression (red) or down-expression (blue), and LOX, LOXL1, and LOXL2 expression were upregulated in PDAC tissues.

## Data Visualization and Analysis

The Venn diagram was drawn via a friendly online tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) to show the overlapping co-expressed genes of LOXs; the heatmap was conducted by TBtools (Chen et al., 2020); the KEGG Mapper (Shannon et al., 2003) was applied to conduct the cancer network using the overlapping genes retrieved from UniProt (https://www.uniprot.org/) (UniProt, 2021). The AOD of LOXs was expressed as the mean ± standard deviation. Data were analyzed with a two-tailed Student's t-test for a parametric test. $p$ < 0.05 was considered to indicate a statistically significant difference.

# RESULTS

## High mRNA Expression of LOXs in PDAC Tissues

The mRNA expression of LOX, LOXL1, and LOXL2 was significantly higher in PDAC tissues than healthy tissues based

**TABLE 1 |** The significant changes of LOXs expression in transcriptional level between PAAD and normal pancreas tissues (Oncomine).

| | Type of PAAD *vs.* pancreas | Fold change | *p*-value | *t*-test | PMID |
|---|---|---|---|---|---|
| LOX | Pancreatic adenocarcinoma | 4.433 | 2.78E-04 | 5.864 | 12750293 |
| | Pancreatic ductal adenocarcinoma | 8.835 | 9.96E-13 | 9.142 | 19260470 |
| | Pancreatic carcinoma | 2.933 | 6.00E-4 | 4.353 | 15867264 |
| LOXL1 | Pancreatic ductal adenocarcinoma | 5.245 | 3.97E-16 | 11.346 | 19260470 |
| | Pancreatic adenocarcinoma | 2.084 | 9.01E-04 | 3.937 | 12651607 |
| LOXL2 | Pancreatic adenocarcinoma | 5.235 | 4.56E-06 | 7.695 | 12750293 |
| | Pancreatic carcinoma | 4.989 | 2.63E-07 | 6.39 | 19732725 |
| | Pancreatic adenocarcinoma | 3.334 | 5.61E-04 | 4.255 | 12651607 |
| | Pancreatic ductal adenocarcinoma | 3.172 | 5.96E-11 | 7.589 | 19260470 |
| | Pancreatic carcinoma | 3.014 | 5.13E-4 | 4.277 | 15867264 |

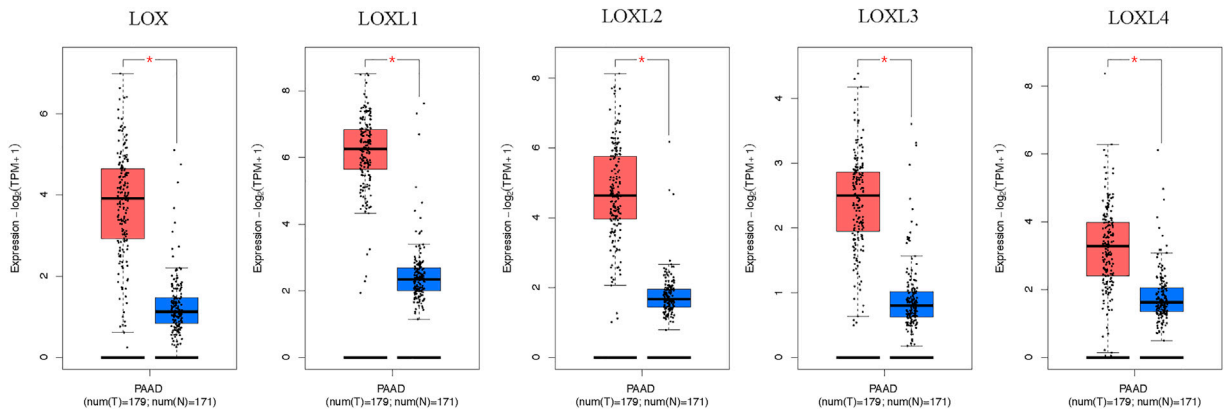



**FIGURE 2 |** Boxplot showing the expression levels of LOXs in PDAC. The mRNA expression of LOXs was significantly elevated in PDAC than normal pancreas tissues (GEPIA database). The number of normal samples is 171 (blue box), and PDAC samples are 179 (red box), red star means $p < 0.01$.

on the Oncomine analysis (**Figure 1** and **Table 1**). In Logsdon PDAC dataset(Logsdon et al., 2003), LOX and LOXL2 transcripts were upregulated in PDAC tissues by fold changes of 4.433 and 5.235 ($p$ = 2.78E-4, 4.56E-6), respectively. Iacobuzio-Donahue (Iacobuzio-Donahue et al., 2003) found a 2.084-fold increase of LOXL1 and a 3.334-fold increase of LOXL2 in PDAC tissues. Moreover, other studies were added to support the upregulation of LOX, LOXL1, and LOXL2 in PDAC (Badea et al., 2008; Pei et al., 2009).

Furthermore, the mRNA expression patterns of LOXs were further measured by GEPIA based on the TCGA database. The GEPIA analyses showed all LOXs were higher in PDAC than in normal tissues (**Figure 2**).

## Protein Expression of LOXs in PDAC Tissues

To further explore the protein level in PDAC, IHC was performed in PDAC and adjacent pancreas. The IHC results demonstrated elevated LOX, LOXL1 and LOXL2 protein expression in the PDAC tissues (LOX: $p < 0.01$; LOXL1: $p < 0.05$; LOXL2: $p < 0.05$ **Figures 3A–C**). However, LOXL3 showed a remarkable decrease in PDAC tissues ($p < 0.05$, **Figure 3D**), and LOXL4 made no difference (**Figure 3E**).

We also analyzed the protein expression patterns of LOXs in PDAC using HPA and CPTAC datasets. Consistent with our IHC results, HPA results indicated increased LOXL1 and decreased LOXL3 expression in PDAC tissues (**Supplementary Figure S1** and **Supplementary Table S1**), while no significant difference was identified in LOXL2 and LOXL4 compared with exocrine glandular cells in normal tissue. However, the expression of LOLX3 was contrary in different databases. The protein expression level of LOXL3 was identified higher in PDAC tissues than normal tissues verified by CPTAC data (**Figure 4**).

## Relationship Between LOXs Expression and the Clinical Characteristics of PDAC Patients

We further investigated the association between the mRNA expression of LOXs and the clinical features of PDAC. As indicated in **Figure 5**, the higher PDAC stage was related to the higher mRNA levels of LOX and LOXL2. Meanwhile, tumor grade was significantly correlated with the expression of LOX and LOXL3.

## LOXs Correlated Genes and Functional Enrichment Analysis in PDAC

To gain insight into the potential role of LOXs in the PDAC process, we conducted interactome, gene ontology (GO), and

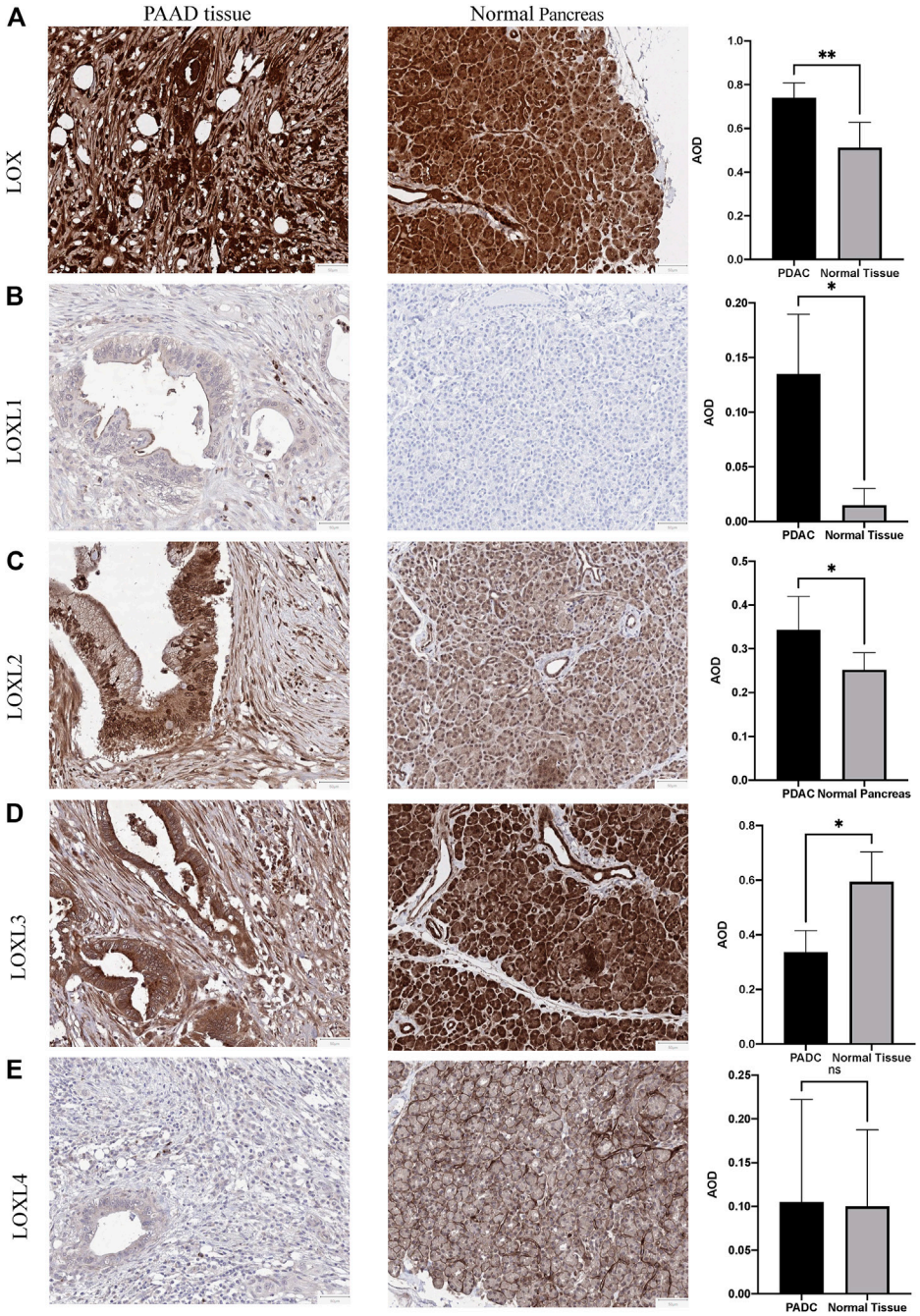



**FIGURE 3 |** Representative immunohistochemistry of LOXs in PDAC tissues and normal pancreas tissues. **(A)** IHC staining of LOX and quantification showed higher expression in PDAC than normal pancreas tissues. **(B)** IHC staining and quantification showed higher expression of LOXL1 in PDAC. **(C)** IHC staining and quantification showed higher expression of LOXL2 in PDAC. **(D)** IHC staining and quantification showed lower expression of LOXL3 in PDAC. **(E)** LOXL4 expression levels were similar in PDAC and normal tissues. *$p < 0.01$; NS, no significant difference. $n = 6$, repeated 5 times.

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of the top 200 correlated genes of each LOX family member extracted from LinkedOmics. The results revealed that LOX, LOXL1, 2, 4 were mainly involved in ECM regulation, such as extracellular matrix structural constituent (GO: 0005201), collagen-containing extracellular matrix (GO: 0062023), extracellular matrix (GO: 0031012), external encapsulating structure (GO: 0030312), extracellular matrix organization (GO: 0030198), and extracellular structure organization (GO: 0043062), while LOXL3 showed a close relationship with immune-related functions, such as myeloid leukocyte activation (GO: 0002274), leukocyte activation involved in

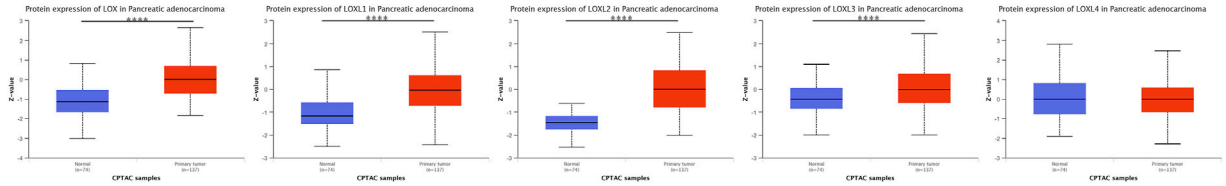



**FIGURE 4 |** Protein expression levels of LOXs in PDAC tissues and normal tissues (CPTAC database). Protein expression levels of LOX, LOXL1, 2, 3 were significantly up-regulated. ****$p < 0.0001$.

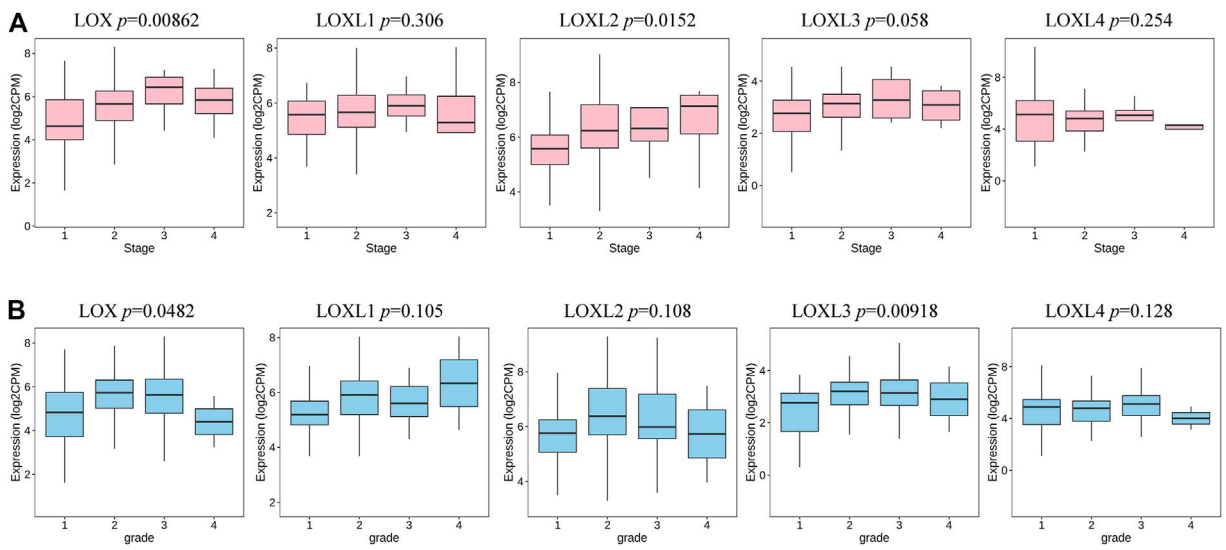



**FIGURE 5 |** Transcription levels of LOXs in different stages and grades of patients with PDAC (TISIDB). **(A)** The expression of LOX and LOXL2 was highly correlated to the stage of the PDAC. **(B)** The expression of LOX and LOXL3 was significantly associated with tumor grade in PDAC.

immune response (GO: 0002366), myeloid leukocyte mediated immunity (GO: 0002444), myeloid cell activation involved in immune response (GO: 002275), regulation of leukocyte activation (GO: 0002694), leukocyte differentiation (GO: 0002521), and lymphocyte activation (GO: 0046649) (**Figure 6** and **Supplementary Table S2**).

Besides the specific role of each LOX member plays in cancer, the involvement of several members of LOXs has also been illustrated. Thus, the 168 overlapping genes of each LOX member's top 200 correlated genes were used to investigate the combined function of the LOX family in PDAC (**Figure 7**). As indicated in **Figure 8A**, we observed that functions of those genes were enriched in ECM remodeling, such as extracellular matrix organization (GO: 0030198), extracellular matrix structural constituent conferring tensile strength (GO: 0030020), blood vessel development (GO: 0001568), and collagen metabolic process (GO: 0032963). The networks of enrichment terms of LOXs according to cluster-ID were displayed in **Figure 8B**. Furthermore, the PPI network for LOXs and overlapping co-expressed genes was conducted, wherein the significant densely connected network constituents were assigned in a unique color (**Figure 8C**).

Similar functional enrichment results were achieved using DAVID (https://david.ncifcrf.gov/) (Huang da et al., 2009a; Huang da et al., 2009b) (**Supplementary Figure S2**). Considering the strong correlation between LOXs-related genes and pathways in cancer (hsa05200) as shown in the Metascape analysis, the KEGG pathway map was generated (**Figure 9**). The cancer network revealed that those correlated genes were essentially abundant in pathways in cancer (hsa05200), PI3K-AKT signaling pathway (has04151), focal adhesion (has04510), and ECM-receptor interaction (hsa04512), the details illustrated in **Supplementary Figure S3**.

## Association of LOXs With Immune Infiltration in PDAC

Considering the critical role of immune cells in the TME contributing to the progression of PDAC and a close relationship of LOXL3 with immune response demonstrated in the current study, we further investigated the correlation of LOXs transcriptional level with immune infiltration in PDAC using the TIMER and TISIDB database. The TIMER results showed that the LOXs expression was negatively related to tumor purity. The

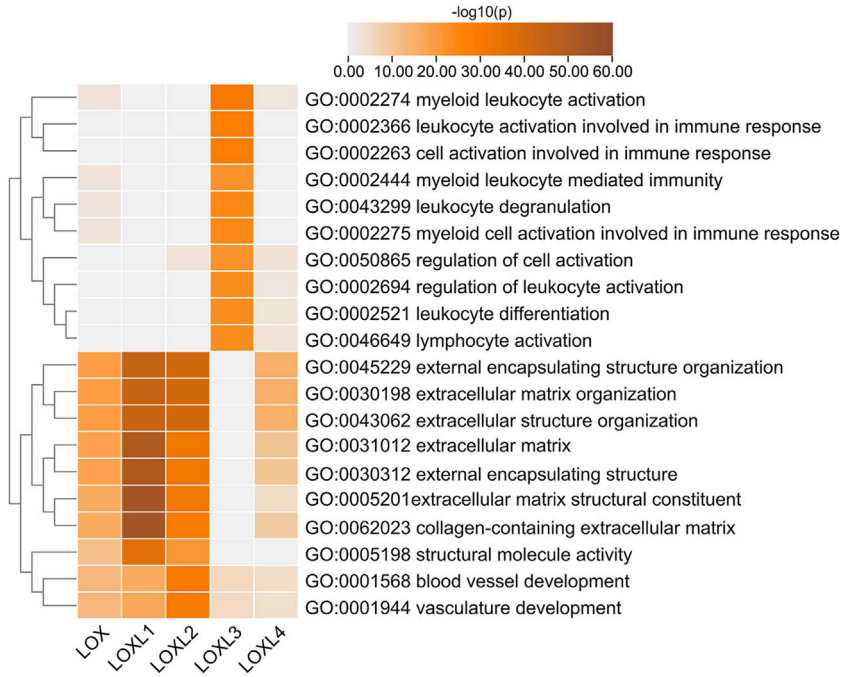



**FIGURE 6 |** Top 20 statistically significant pathways associated with each LOXs identified by Metascape. Heatmap of the molecular function, biological processes, or pathways enriched with distinct LOX co-expressed genes. The bar color shade was decided by the *p-value*, the deeper the shade the less the *p*-value.

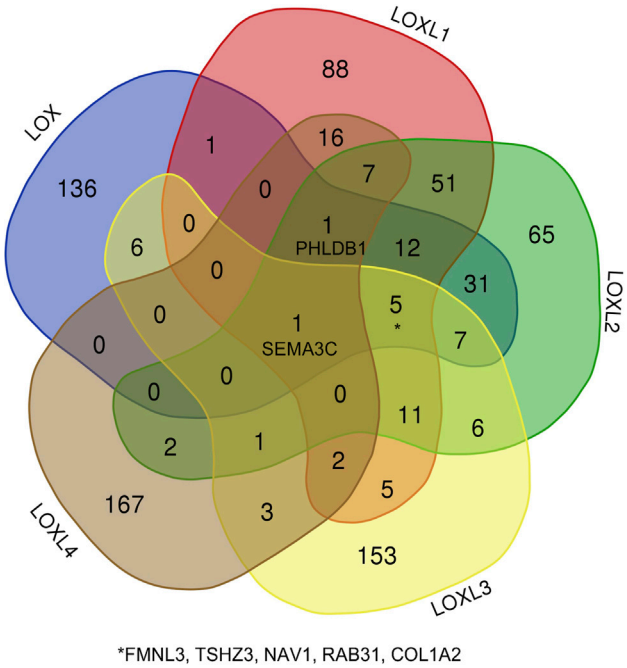



**FIGURE 7 |** The Venn diagram represents the correlated genes of LOXs based on the Pearson correlation coefficient.

mRNA expression of LOX, LOXL1, 3, 4 was remarkably correlated with $CD8^+$ T cells. The mRNA expression levels of LOX, LOXL1-3 had a significant correlation with infiltrating levels of macrophages and NK cells. In addition, the LOX family was highly associated with neutrophil and dendritic cell infiltration (**Figures 10A–E**). It was worth noting that LOX and LOXL3 remarkedly correlated with those TILs.

Then, we explored the correlation between the abundances of TILs and LOXs expression via TISIDB. In line with the TIMER results, the TISIDB analysis demonstrated that high LOX and LOXL3 expression was strongly correlated with the Tcm $CD8^+$ T cell, Tem $CD8^+$ T cell, Tfh cell, Th1 cell, Treg cell, NK cell, MDSC cell, act DC cell, pDC cell, macrophage, and mast cell infiltration. Meanwhile, the expression of LOXL1, 2, 4 showed a relatively weak correlation with them (**Figure 10F** and **Supplementary Table S3**). These results suggested that LOXs, particularly LOX and LOXL3, might affect PDAC by influencing immune infiltration.

## Correlation Between mRNA Expression of LOXs and Immune Biomarkers in PDAC

To better understand LOXs' crosstalk with immune-related genes, we analyzed the association between the mRNA expression of LOXs and various immune biomarkers using TIMER. The results indicated that the expression of LOX and LOXL3 was highly related to biomarkers of $CD8^+$ T cell, T cell, B cell, monocyte, TAM, M1 macrophage, M2 macrophage, neutrophil, DC cell, Th1, Th2, and Th17 in PDAC (**Figure 11**).

We further explored the immune checkpoint expression level of PDAC cells. The results showed that LOX and LOXL3 were strongly correlated to PD-1, PD-L1, CLTA4, GZMB, TIM-3, and LAG3, while other members of LOX showed no relationship

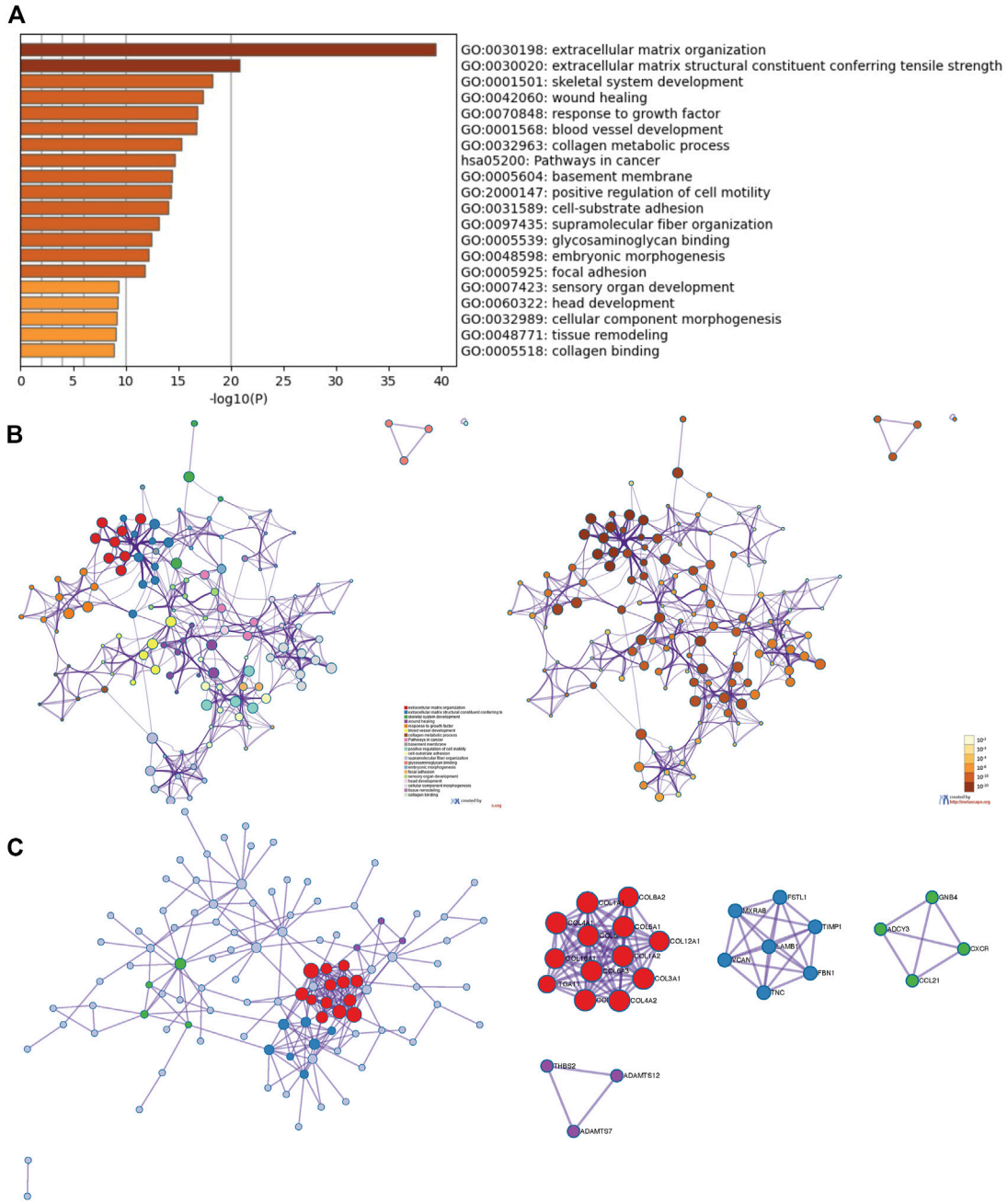



**FIGURE 8 |** The enrichment analysis and visualized protein-protein interaction (PPI) enrichment analysis of co-expressed genes of LOXs (Metascape). **(A)** Heatmap of enriched terms regarding Gene Ontology analysis colored by *p*-values. **(B)** An interactive network of the top 20 enrichment terms. **(C)** PPI network and densely connected network components are identified by the Molecular Complex Detection (MCODE).

(**Table 2**). In brief, these results affirmed the role of LOX and LOXL3, especially LOXL3 in the immune response in PDAC.

## Prognostic Value of mRNA Expression of LOXs in PDAC Patients

As displayed in **Figures 12A,B, C** the mRNA levels of LOX and LOXL2 were significantly correlated with poor OS ($p$ = 0.048 and 6e-04), progression-free interval (PFI) ($p$ = 0.0029 and 0.018), and disease-specific survival (DSS) ($p$ = 0.014 and 0.0031) in PDAC. However, the mRNA levels of LOXL1, 3, 4 showed no relationship with the overcomes in PDAC patients. The mRNA levels of LOX and LOXL2 highly correlated with the outcomes of PDAC patients, indicating LOX and LOXL2 may be served as biomarkers for predicting the prognosis.

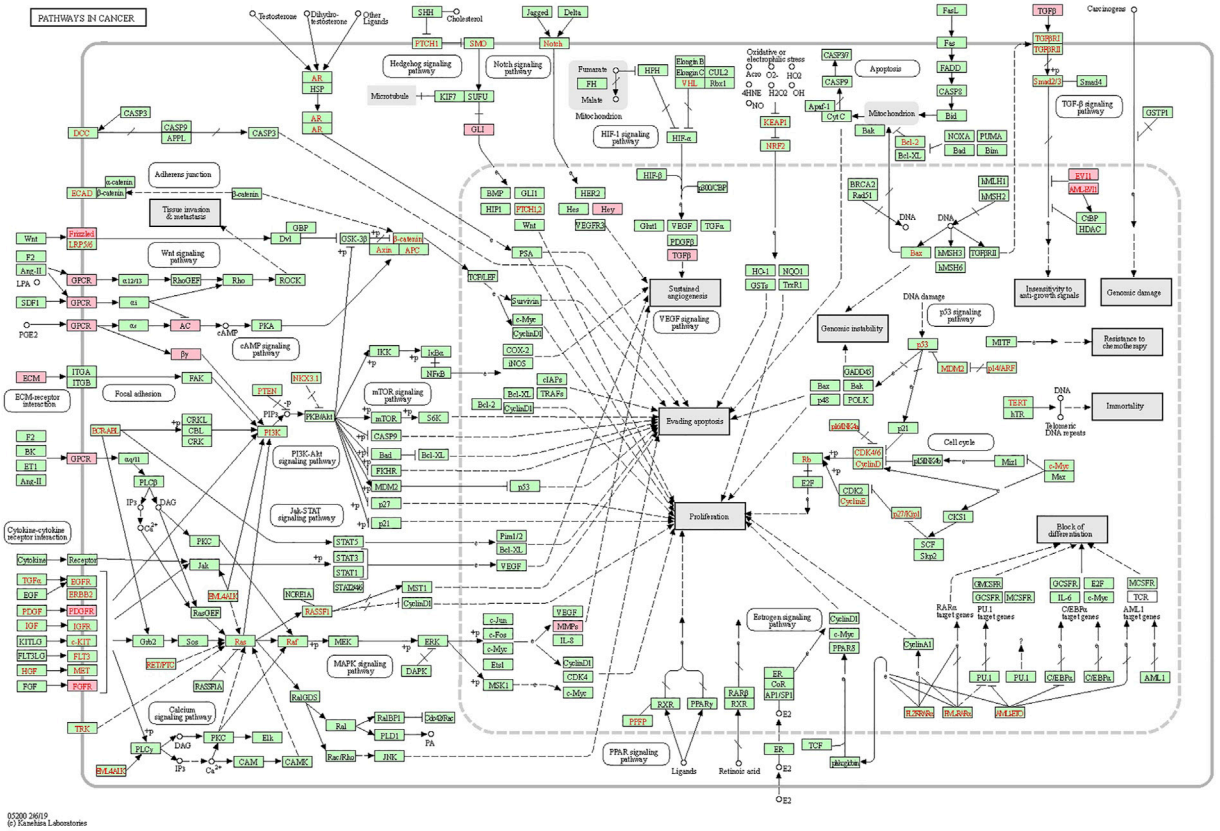



**FIGURE 9 |** Positional relationship of LOXs co-expressed genes in the cancer network. The overlapping genes were mainly enriched in cancer-related pathways, such as pathways in cancer (hsa05200), PI3K-AKT signaling pathway (has04151), focal adhesion (has04510), ECM-receptor interaction (hsa04512). The relevant positions of the corresponding genes in the pathway are marked with pink circles.

## DISCUSSION

As a kind of secreted protein family, LOXs play a pivotal role in PDAC pathogenesis by regulating the TME in two aspects, ECM remodeling and immune-related responses. It has been well known that the TME could significantly affect the behavior of tumor cells and ultimately influence the outcomes of patients (Setargew et al., 2021). The TME of PDAC is characterized by dense desmoplasia and extensive immunosuppression (Neesse et al., 2015; Ren et al., 2018). LOXs are major drivers in the biogenesis of the desmoplastic matrix in cancer. LOXs prevent collagen degradation and improve tissue stiffness through catalyzing the cross-linking process of collagen and elastin, and finally contribute to dense desmoplasia in PDAC. Dense desmoplasia plays a critical role in metastasis and limiting chemotherapy efficacy (Incio et al., 2016; Ren et al., 2018; Schizas et al., 2020). With the increasing interest in anti-stromal therapies, the LOX family has emerged as a potential target. Although the role of the LOX family in tumorigenesis and prognosis of several cancers has been partially confirmed, further bioinformatics analysis of PADC has yet to be performed. In this study, we used multitalented public databases to reveal the dysregulated expression of the LOXs and their relationship with tumor stage and prognosis. As for extensive immunosuppression in PDAC, we found that LOXs, particularly LOX and LOXL3 are closely related to TILs, various immune signatures, and immune checkpoints, however, the function in the immune microenvironment remains unclear which calls for further studies.

Elevated LOX expression has been noted in PDAC tissues compared to adjacent tissues and correlated with poor OS and DSS (Miller et al., 2015; Ma et al., 2019). A high expression of LOX was detected in metastatic pancreatic cancer in the mice model. Conversely, LOX inhibition increased immune cell recruitment, vascularization, and decreased fibrillar collagen (Miller et al., 2015). In line with this finding, Le Calve et al. (2016) found that LOX protein was upregulated in PDAC tissues. Compared to adjacent pancreas tissues, we found higher mRNA expression of LOX in PDAC tissues, which was correlated with tumor stage and grade of PDAC. The IHC showed a higher LOX protein level in PDAC tissues than normal pancreatic tissues. The functional enrichment analysis showed that LOX and its ligands were mainly involved in the ECM regulatory. The transcriptional level of LOX is strongly related to the infiltration of $CD8^+$ T cell, macrophage, NK cell, neutrophil, dendritic cell infiltration, as well as various immune biomarkers and immune checkpoints.

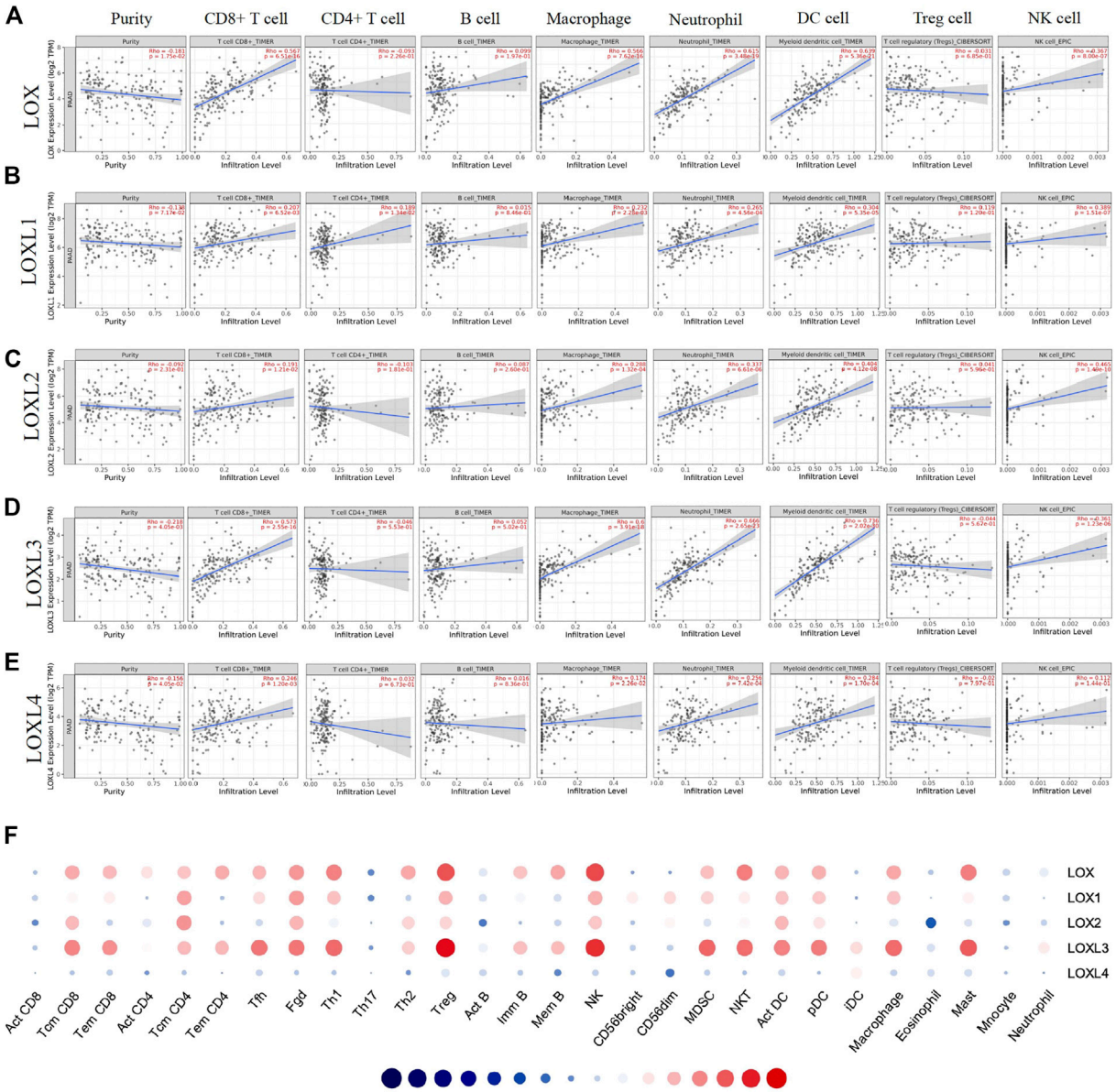



**FIGURE 10 |** Correlation analysis of LOXs expression and immune infiltration in TME of PDAC. **(A–E)** (TIMER). The correlation between the abundance of immune cells and the expression of **(A)** LOX, **(B)** LOXL1, **(C)** LOXL2, **(D)** LOXL3, **(E)** LOXL4 in PDAC. **(F)** Correlation between the relative abundances of 28 tumor-infiltrating lymphocytes (TILs) and LOXs expression in PDAC (TISIDB).

Furthermore, a high LOX expression was significantly associated with poor OS, PFI, and DSS in PDAC patients. These results together indicate the oncogenic role and prognostic value of LOX in PDAC, as well as the role in the immune response.

LOXL1 transcript was also detected upregulated in PDAC patients and associated with chemotherapy resistance (Le Calve et al., 2016). In our study, the mRNA and protein expression of LOXL1 in PDAC was higher than that in normal tissues. However, the mRNA expression of LOXL1 showed no relationship with tumor stage, grade, and outcomes of PDAC patients. We also revealed the role of LOXL1 in ECM shaping and a close relationship with infiltrating levels of $CD8^+$ T cell, macrophage and NK cell, neutrophil, and dendritic cell. However, more studies should be carried out to identify the potential function of LOXL1 in PDAC progression.

Higher LOXL2 expression correlated with clinicopathological features of advanced disease and predicted a worse OS of PDAC patients (Park et al., 2016; Kim et al., 2019). Previous studies have revealed that LOXL2 induced epithelial-mesenchymal transition (EMT) through the PI3K-AKT signaling pathway (Wu and Zhu, 2015) to enhance the invasive and migratory capacity of PDAC cells (Park et al., 2016). In the current study, the expression of LOXL2 was higher in human pancreatic cancer tissues than in normal tissues, which highly correlated to the tumor stage. The

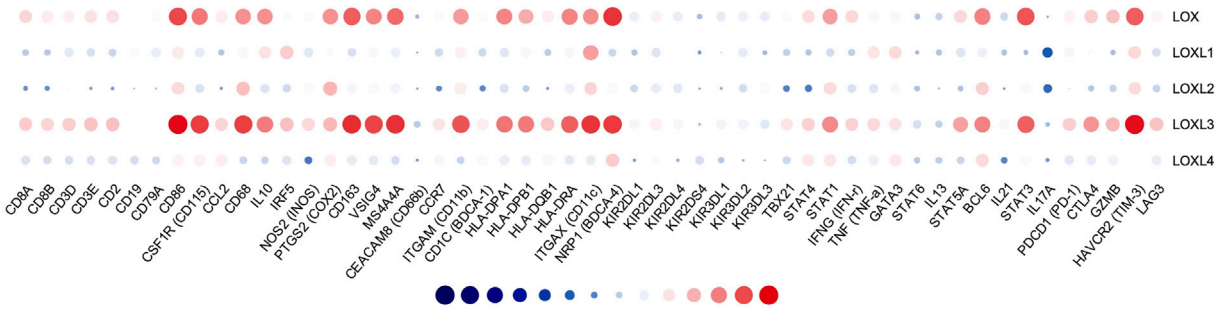



**FIGURE 11 |** Correlation analysis of LOXs and immune signatures in PDAC (TIMER). The correlation was adjusted by purity.

functional enrichment analysis showed LOXL2 also had a close relationship with ECM regulation. In addition, LOXL2 expression showed a significant relationship with macrophage and NK cell, neutrophil, and dendritic cell infiltration. And Minici's study (Minici et al., 2020) has demonstrated LOXL2 is expressed by infiltrating plasmablasts in PDAC. It is worth noting the high expressed LOXL2 predicted an inferior OS, PFI, and DSS in PDAC patients. Consequently, LOXL2 might be a representative biomarker for PDAC. However, phase II clinical trials failed to demonstrate the utility of LOXL2 inhibition in PDAC (Benson et al., 2017), which may be due to cancer-associated fibroblasts (CAFs). The discovery of the tumor-restraining functions of CAFs provides a potential explanation for the unsuccessful clinical trials of therapeutic agents targeting CAFs or stromal components (Van Cutsem et al., 2020). These observations suggest that future therapeutic strategies should avoid generic targeting of tumor-restraining CAF subpopulations in favor of precise reprogramming and normalization of tumor-promoting CAF subsets (Chen et al., 2021).

In this study, overexpressed mRNA of LOXL3 was identified in PDAC tissues and correlated with tumor grade, but showed no relationship with the outcomes of PDAC patients according to the GEPIA database and CPTAC database. Intriguingly, the LOXL3 protein level was found to be downregulated according to our IHC results and HPA database as well, which was controversial. As is known, the expression of LOXL3 was detected mainly in the nucleus correlated with tumor invasion, lymph node metastasis, and poorer prognosis of patients (Kasashima et al., 2018) through interacting with SNAIL and contributing to proliferation and metastasis by inducing epithelial-mesenchymal transition in pancreatic ductal adenocarcinoma cells (Eiseler et al., 2012). Additionally, TGF-induced LOXL3 upregulation in gastric cancer cells suggested that LOXL3 was downstream from the TGF-signaling pathway (Kasashima et al., 2018). However, studies on LOXL3 were mainly based on evidence from cell lines, limited evidence from *in vivo* studies was found. Meanwhile, LOXL3 expression in some tumors was contrary as in some cases LOXL3 was downregulated. Li Ma (Ma et al., 2017) found that LOXL3 was a dual-specificity enzyme involved in STAT3 deacetylation and deacetylimination modulation which reduced the activity of signal transducer and activator of transcription 3 (STAT3). STAT3 activation helps to promote tumor progression (Sun et al., 2021) and is associated with PDAC through PDAC patients' biopsy study(Laklai et al., 2016). De-activation of STAT3 by LOXL3 would be a protective role in PDAC pathogenesis. Furthermore, unlike other LOX family members, the functional analysis of LOXL3 and its co-expressed genes showed enrichment in the immune response. LOXL3 expression showed a high correlation with $CD8^+$ T cell, macrophage, NK cell, neutrophil, and dendritic cell infiltrating levels. In addition, LOXL3 was highly related to biomarkers of $CD8^+$ T cell, T cell, B cell, monocyte, TAM, M1macrophage, M2 macrophage, neutrophil, DC cell, Th1, Th2, and Th17. In addition, the LOXL3 expression was positively associated with immune checkpoints. LOXL3 was also found to induce downregulation of Th17 and Treg activation as well (Ma et al., 2017) which play a role in cancer pathogenesis (Marques et al., 2021). Thus, we speculate LOXL3 plays a crucial role in the modulation and recruitment of immune cells and affects the expression of immune signatures in PDAC, and more studies should be carried out to validate.

Accumulating evidence suggest LOXL4 an oncoprotein in several tumors, such as hepatocellular carcinoma (Li et al., 2019), lung cancer (Zhang et al., 2019) and breast cancer (Choi et al., 2017). However, the expression and prognostic role of LOXL4 in PDAC have yet to be investigated. This report demonstrated a higher mRNA expression of LOXL4 in PDAC tissues with no correlation with tumor stage, grade, or outcomes in PDAC patients. The protein expression of LOLX4 showed no difference in PDAC and normal pancreas. The functional analysis of LOXL4 identified similar results to the canonical function of LOXs. In addition, the expression of LOXL4 showed a weak correlation with immune biomarkers. These results together suggest that LOXL4 function in pancreatic tumorigenesis needs to be further determined.

The involvement of one or more LOX members was verified in different cancers. Thus, the LOX family members are likely to present distinct functions specific for each member and partially overlapping activity. We further analyzed the synergy function of LOXs with overlapping co-expressed genes by GO enrichment analysis and KEGG pathway enrichment. GO analysis demonstrated the role of LOXs in ECM regulation. The cancer network showed that changes in LOXs expression mainly affect

**TABLE 2 |** Correlation analysis of LOXs and immune biomarkers in PAAD (TIMER).

| Description | Gene marker | LOX | | LOXL1 | | LOXL2 | | LOXL3 | | LOXL4 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | rho | p | rho | p | rho | p | rho | p | rho | p |
| CD8+T cell | CD8A | 0.325 | ** | 0.100 | 0.300 | −0.062 | 0.533 | 0.346 | *** | 0.162 | 0.074 |
| | CD8B | 0.281 | * | 0.095 | 0.326 | −0.067 | 0.495 | 0.327 | *** | 0.148 | 0.107 |
| T cell | CD3D | 0.271 | * | 0.181 | 0.044 | −0.005 | 0.972 | 0.340 | *** | 0.153 | 0.094 |
| | CD3E | 0.299 | ** | 0.156 | 0.085 | −0.032 | 0.781 | 0.364 | *** | 0.160 | 0.079 |
| | CD2 | 0.303 | ** | 0.129 | 0.167 | −0.037 | 0.732 | 0.365 | *** | 0.168 | 0.064 |
| B cell | CD19 | 0.234 | 0.005 | 0.224 | 0.010 | −0.012 | 0.917 | 0.236 | * | 0.152 | 0.098 |
| | CD79A | 0.250 | * | 0.193 | 0.030 | −0.014 | 0.905 | 0.235 | * | 0.137 | 0.135 |
| Monocyte | CD86 | 0.603 | *** | 0.283 | * | 0.314 | ** | 0.687 | *** | 0.263 | * |
| | CSF1R (CD115) | 0.512 | *** | 0.187 | 0.036 | 0.160 | 0.067 | 0.621 | *** | 0.267 | * |
| TAM | CCL2 | 0.279 | * | 0.086 | 0.379 | −0.031 | 0.789 | 0.326 | *** | 0.276 | * |
| | CD68 | 0.461 | *** | 0.255 | * | 0.368 | *** | 0.619 | *** | 0.129 | 0.161 |
| | IL10 | 0.463 | *** | 0.298 | ** | 0.167 | 0.055 | 0.505 | *** | 0.124 | 0.178 |
| M1 Macrophage | IRF5 | 0.233 | 0.005 | 0.340 | *** | 0.140 | 0.113 | 0.367 | *** | 0.171 | 0.059 |
| | NOS2 (INOS) | 0.248 | * | 0.179 | 0.047 | 0.225 | 0.008 | 0.319 | *** | −0.130 | 0.159 |
| | PTGS2 (COX2) | 0.460 | *** | 0.244 | * | 0.388 | *** | 0.384 | *** | 0.275 | * |
| M2 Macrophage | CD163 | 0.565 | *** | 0.157 | 0.085 | 0.212 | 0.013 | 0.648 | *** | 0.189 | 0.033 |
| | VSIG4 | 0.486 | *** | 0.221 | 0.011 | 0.252 | 0.003 | 0.611 | *** | 0.182 | 0.042 |
| | MS4A4A | 0.537 | *** | 0.201 | 0.023 | 0.208 | 0.015 | 0.640 | *** | 0.150 | 0.103 |
| Neutrophils | CEACAM8 (CD66b) | 0.048 | 0.625 | 0.016 | 0.894 | 0.021 | 0.855 | 0.106 | 0.217 | −0.029 | 0.790 |
| | CCR7 | 0.249 | * | 0.143 | 0.120 | −0.079 | 0.408 | 0.294 | ** | 0.216 | 0.014 |
| | ITGAM (CD11b) | 0.444 | *** | 0.268 | * | 0.278 | * | 0.583 | *** | 0.127 | 0.239 |
| Dendritic cell | CD1C (BDCA-1) | 0.245 | 0.006 | 0.135 | 0.180 | −0.085 | 0.463 | 0.275 | * | 0.251 | 0.006 |
| | HLA-DPA1 | 0.486 | *** | 0.171 | 0.075 | 0.163 | 0.099 | 0.522 | *** | 0.192 | 0.049 |
| | HLA-DPB1 | 0.411 | *** | 0.256 | * | 0.099 | 0.375 | 0.509 | *** | 0.189 | 0.052 |
| | HLA-DQB1 | 0.278 | * | 0.106 | 0.310 | −0.026 | 0.842 | 0.352 | *** | 0.258 | * |
| | HLA-DRA | 0.490 | *** | 0.215 | 0.019 | 0.195 | 0.040 | 0.559 | *** | 0.181 | 0.067 |
| | ITGAX (CD11c) | 0.466 | *** | 0.435 | *** | 0.323 | ** | 0.635 | *** | 0.185 | 0.059 |
| | NRP1 (BDCA-4) | 0.639 | *** | 0.131 | 0.100 | 0.251 | * | 0.625 | *** | 0.345 | *** |
| Natural killer cell | KIR2DL1 | 0.196 | 0.036 | 0.085 | 0.424 | 0.136 | 0.189 | 0.215 | 0.011 | −0.032 | 0.805 |
| | KIR2DL3 | 0.265 | * | 0.169 | 0.080 | 0.190 | 0.047 | 0.272 | * | 0.231 | 0.013 |
| | KIR2DL4 | 0.192 | 0.041 | −0.001 | 0.993 | 0.175 | 0.071 | 0.217 | 0.010 | 0.040 | 0.753 |
| | KIR2DS4 | 0.069 | 0.588 | −0.040 | 0.744 | 0.024 | 0.858 | 0.024 | 0.822 | 0.152 | 0.143 |
| | KIR3DL1 | 0.184 | 0.053 | −0.006 | 0.978 | 0.027 | 0.837 | 0.209 | 0.013 | 0.089 | 0.438 |
| | KIR3DL2 | 0.201 | 0.031 | 0.185 | 0.052 | 0.149 | 0.138 | 0.207 | 0.015 | 0.021 | 0.875 |
| | KIR3DL3 | 0.189 | 0.045 | 0.057 | 0.628 | 0.250 | 0.006 | 0.176 | 0.042 | 0.013 | 0.928 |
| Th1 | TBX21 | 0.229 | 0.008 | 0.117 | 0.243 | −0.104 | 0.284 | 0.293 | ** | 0.136 | 0.172 |
| | STAT4 | 0.311 | ** | 0.083 | 0.420 | −0.118 | 0.219 | 0.332 | *** | 0.287 | * |
| | STAT1 | 0.438 | *** | 0.146 | 0.131 | 0.311 | ** | 0.476 | *** | 0.292 | * |
| | IFNG (IFN-γ) | 0.318 | ** | 0.115 | 0.250 | 0.161 | 0.080 | 0.332 | *** | 0.148 | 0.134 |
| | TNF (TNF-α) | 0.255 | * | 0.292 | * | 0.170 | 0.064 | 0.314 | ** | 0.201 | 0.028 |
| Th2 | GATA3 | 0.255 | * | 0.318 | ** | 0.218 | 0.014 | 0.297 | ** | 0.287 | * |
| | STAT6 | 0.077 | 0.450 | 0.043 | 0.696 | 0.078 | 0.432 | 0.158 | 0.063 | 0.096 | 0.360 |
| | IL13 | 0.126 | 0.185 | 0.082 | 0.426 | 0.019 | 0.873 | 0.118 | 0.174 | 0.108 | 0.301 |
| | STAT5A | 0.319 | ** | 0.171 | 0.068 | 0.078 | 0.432 | 0.429 | *** | 0.127 | 0.209 |
| Tfh | BCL6 | 0.494 | *** | 0.200 | 0.030 | 0.336 | ** | 0.494 | *** | 0.313 | ** |
| | IL21 | 0.137 | 0.146 | 0.082 | 0.424 | −0.023 | 0.844 | 0.220 | 0.008 | −0.092 | 0.389 |
| Th17 | STAT3 | 0.581 | *** | 0.058 | 0.595 | 0.252 | * | 0.559 | *** | 0.225 | 0.012 |
| | IL17A | 0.022 | 0.855 | −0.210 | 0.021 | −0.163 | 0.076 | 0.050 | 0.601 | −0.046 | 0.689 |
| T cell exhaustion | PDCD1 (PD-1) | 0.263 | * | 0.215 | 0.018 | 0.006 | 0.956 | 0.339 | *** | 0.177 | 0.062 |
| | CTLA4 | 0.339 | *** | 0.241 | 0.007 | 0.086 | 0.385 | 0.456 | *** | 0.187 | 0.046 |
| | GZMB | 0.365 | *** | 0.062 | 0.562 | 0.122 | 0.202 | 0.382 | *** | 0.206 | 0.024 |
| | HAVCR2 (TIM-3) | 0.566 | *** | 0.317 | ** | 0.313 | ** | 0.682 | *** | 0.240 | 0.007 |
| | LAG3 | 0.262 | * | 0.156 | 0.101 | 0.123 | 0.197 | 0.361 | *** | 0.193 | 0.038 |
| | CD274 (PD-L1) | 0.562 | *** | 0.099 | 0.196 | 0.284 | ** | 0.551 | *** | 0.291 | ** |

*p < *0.01*, **p < *0.001*, ***p < *0.0001*.

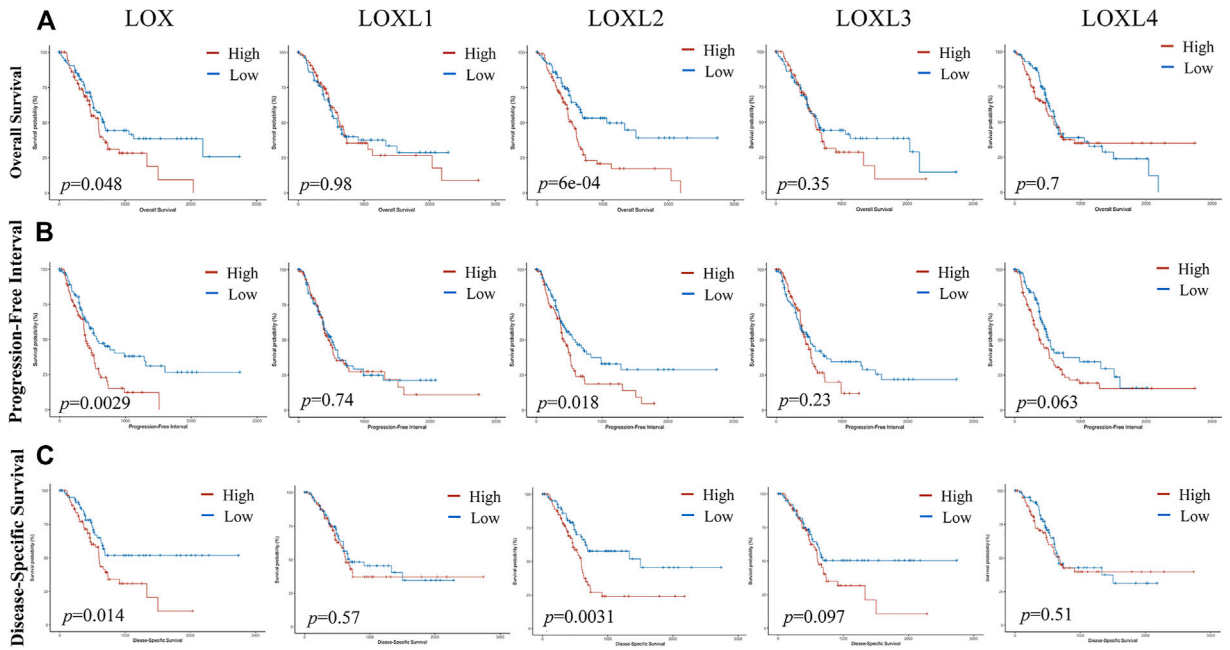



**FIGURE 12 |** The prognosis value of LOXs expression in PDAC. **(A)** Upregulated LOX and LOXL2 were highly correlated with poor overall survival (OS) in PDAC patients. **(B)** Upregulated LOX, LOXL2, and LOXL4 were significantly associated with inadequate progression-free interval (PFI). **(C)** LOX and LOXL2 expression was negatively related to disease-specific survival (DSS).

pathways in cancer (hsa05200), PI3K-AKT signaling pathway (has04151), focal adhesion (has04510), ECM-receptor interaction (hsa04512). These pathways are significantly associated with PDAC progression. Considering the major ECM component in PDAC, the predominant role of LOXs in the ECM remodeling, and involvement in the cancer-related pathways, we speculate that differentially expressed LOXs in PDAC are potential targets for drug therapy.

Current therapeutic strategies aim to deconstruct the surrounding desmoplastic stroma and target immunosuppressive pathways that have largely failed in PDAC. Even immunotherapeutic strategies, such as immune-checkpoint blockade, gained remarkable success in many malignancies that have not yet translated to PDAC (Johnson et al., 2017; Stromnes et al., 2017; Morrison et al., 2018). In addition, ablating the stromal barriers that restrict drug delivery has also demonstrated disappointing and contradictory responses. The key contributor is the multi-faceted functions of the complex TME components (Chronopoulos et al., 2016; Stromnes et al., 2017). Thus, combination strategies that target multiple features of the TME simultaneously might succeed. LOXs inhibitors reduce the fibrotic ECM and desmoplasia and overcome chemoresistance in triple-negative breast cancer by increasing the diffusion of chemotherapeutics into the tumor (Saatci et al., 2020). As demonstrated LOXs played a crucial role in modulating immune cells and affecting the expression of immune signatures in PDAC. Considering the close relationship of LOXs and the immune checkpoints, LOXs inhibitors might also play a role in promoting the anti-tumor effect of immune checkpoint inhibitors. A combination of LOXs inhibitor treatment may improve patient outcomes by priming the TME. However, further preclinical *in vivo* investigations and clinical trials need to be carried out to confirm our speculation.

In conclusion, our research identifies upregulated LOX, LOXL1 and LOXL2 mRNA and protein expression in PDAC tissue. LOX and LOXL2 could serve as novel biomarkers with prognostic significance in PDAC patients. The function of LOXs and their ligands is mainly involved in ECM regulation and pathways in cancer. Intriguing, LOXL3 and its co-expressed genes hold remarked relationship with immune response. Furthermore, a close association between the expression of LOX and LOXL3 with tumor immune infiltration was verified, uncovering the potential molecular mechanism underlying the carcinogenesis in PDAC. Our findings inspire new insight into prognostic biomarkers and new immunotherapeutic targets for PDAC. However, more experimental studies are needed to reveal the link between LOXs and PDAC, thereby promoting the clinical application of LOXs as prognostic indicators and combine therapy targets in PDAC.

## DATA AVAILABILITY STATEMENT

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

## ETHICS STATEMENT

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

## AUTHOR CONTRIBUTIONS

GW was responsible for the study concept and design; CJ and XL were involved in data collection, data screening, statistical analysis, and manuscript writing; WY was involved in the pathology. MW and GL modified and took charge of supervising the manuscript. All authors have read and agreed to the published version of the manuscript.

## FUNDING

This work was supported by the China Scholarship Council (201906175060) and the Jilin Provincial Department of Finance (JLSWSRCZX2021-016). However, the funders had no role in study design, data collection, analysis, decision to publish, or manuscript preparation.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.778857/full#supplementary-material

## REFERENCES

Al-U'datt, D. A., Allen, B. G., and Nattel, S. (2019). Role of the Lysyl Oxidase Enzyme Family in Cardiac Function and Disease. *Cardiovasc. Res.* 13, 1820–1837. doi:10.1093/cvr/cvz176

Badea, L., Herlea, V., Dima, S. O., Dumitrascu, T., and Popescu, I. (2008). Combined Gene Expression Analysis of Whole-Tissue and Microdissected Pancreatic Ductal Adenocarcinoma Identifies Genes Specifically Overexpressed in Tumor Epithelia. *Hepatogastroenterology* 55, 2016–2027.

Barker, H. E., Cox, T. R., and Erler, J. T. (2012). The Rationale for Targeting the LOX Family in Cancer. *Nat. Rev. Cancer* 12, 540–552. doi:10.1038/nrc3319

Benson, A. B., 3rd, Wainberg, Z. A., Hecht, J. R., Vyushkov, D., Dong, H., Bendell, J., et al. (2017). A Phase II Randomized, Double-Blind, Placebo-Controlled Study of Simtuzumab or Placebo in Combination with Gemcitabine for the First-Line Treatment of Pancreatic Adenocarcinoma. *The oncologist* 22, 241–e15. doi:10.1634/theoncologist.2017-0024

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

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

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* 13, 1194–1202. doi:10.1016/j.molp.2020.06.009

Chen, Y., McAndrews, K. M., and Kalluri, R. (2021). Clinical and Therapeutic Relevance of Cancer-Associated Fibroblasts. *Nat. Rev. Clin. Oncol.* 18, 792–804. doi:10.1038/s41571-021-00546-5

Choi, S. K., Kim, H. S., Jin, T., and Moon, W. K. (2017). LOXL4 Knockdown Enhances Tumor Growth and Lung Metastasis through Collagen-dependent Extracellular Matrix Changes in Triple-Negative Breast Cancer. *Oncotarget* 8, 11977–11989. doi:10.18632/oncotarget.14450

Chronopoulos, A., Robinson, B., Sarper, M., Cortes, E., Auernheimer, V., Lachowski, D., et al. (2016). ATRA Mechanically Reprograms Pancreatic Stellate Cells to Suppress Matrix Remodelling and Inhibit Cancer Cell Invasion. *Nat. Commun.* 7, 12630. doi:10.1038/ncomms12630

Conroy, T., Desseigne, F., Ychou, M., Bouché, O., Guimbaud, R., Bécouarn, Y., et al. (2011). FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. *N. Engl. J. Med.* 364, 1817–1825. doi:10.1056/NEJMoa1011923

Eiseler, T., Köhler, C., Nimmagadda, S. C., Jamali, A., Funk, N., Joodi, G., et al. (2012). Protein Kinase D1 Mediates anchorage-dependent and -independent Growth of Tumor Cells via the Zinc finger Transcription Factor Snail1. *J. Biol. Chem.* 287, 32367–32380. doi:10.1074/jbc.M112.370999

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci. Signal.* 6, pl1. doi:10.1126/scisignal.2004088

Harris, E. D. (1976). Copper-induced Activation of Aortic Lysyl Oxidase *In Vivo*. *Proc. Natl. Acad. Sci.* 73, 371–374. doi:10.1073/pnas.73.2.371

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists. *Nucleic Acids Res.* 37, 1–13. doi:10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources. *Nat. Protoc.* 4, 44–57. doi:10.1038/nprot.2008.211

Iacobuzio-Donahue, C. A., Maitra, A., Olsen, M., Lowe, A. W., van Heek, N. T., Rosty, C., et al. (2003). Exploration of Global Gene Expression Patterns in Pancreatic Adenocarcinoma Using cDNA Microarrays. *Am. J. Pathol.* 162, 1151–1162. doi:10.1016/S0002-9440(10)63911-9

Incio, J., Liu, H., Suboj, P., Chin, S. M., Chen, I. X., Pinter, M., et al. (2016). Obesity-Induced Inflammation and Desmoplasia Promote Pancreatic Cancer Progression and Resistance to Chemotherapy. *Cancer Discov.* 6, 852–869. doi:10.1158/2159-8290.CD-15-1177

Johnson, B. A., 3rd, Yarchoan, M., Lee, V., Laheru, D. A., and Jaffee, E. M. (2017). Strategies for Increasing Pancreatic Tumor Immunogenicity. *Clin. Cancer Res.* 23, 1656–1669. doi:10.1158/1078-0432.CCR-16-2318

Johnston, K. A., and Lopez, K. M. (2018). Lysyl Oxidase in Cancer Inhibition and Metastasis. *Cancer Lett.* 417, 174–181. doi:10.1016/j.canlet.2018.01.006

Kagan, H. M., and Li, W. (2003). Lysyl Oxidase: Properties, Specificity, and Biological Roles inside and outside of the Cell. *J. Cel. Biochem.* 88, 660–672. doi:10.1002/jcb.10413

Kasashima, H., Yashiro, M., Kinoshita, H., Fukuoka, T., Morisaki, T., Masuda, G., et al. (2016). Lysyl Oxidase Is Associated with the Epithelial-Mesenchymal Transition of Gastric Cancer Cells in Hypoxia. *Gastric Cancer* 19, 431–442. doi:10.1007/s10120-015-0510-3

Kasashima, H., Yashiro, M., Okuno, T., Miki, Y., Kitayama, K., Masuda, G., et al. (2018). Significance of the Lysyl Oxidase Members Lysyl Oxidase like 1, 3, and 4 in Gastric Cancer. *Digestion* 98, 238–248. doi:10.1159/000489558

Kim, I.-k., Lee, Y. S., Kim, H. S., Dong, S. M., Park, J. S., and Yoon, D. S. (2019). Specific Protein 1(SP1) Regulates the Epithelial-Mesenchymal Transition via Lysyl Oxidase-like 2(LOXL2) in Pancreatic Ductal Adenocarcinoma. *Sci. Rep.* 9, 5933. doi:10.1038/s41598-019-42501-6

Laklai, H., Miroshnikova, Y. A., Pickup, M. W., Collisson, E. A., Kim, G. E., Barrett, A. S., et al. (2016). Genotype Tunes Pancreatic Ductal Adenocarcinoma Tissue Tension to Induce Matricellular Fibrosis and Tumor Progression. *Nat. Med.* 22, 497–505. doi:10.1038/nm.4082

Le Calvé, B., Griveau, A., Vindrieux, D., Maréchal, R., Wiel, C., Svrcek, M., et al. (2016). Lysyl Oxidase Family Activity Promotes Resistance of Pancreatic Ductal Adenocarcinoma to Chemotherapy by Limiting the Intratumoral Anticancer Drug Distribution. *Oncotarget* 7, 32100–32112. doi:10.18632/oncotarget.8527

Li, R., Wang, Y., Zhang, X., Feng, M., Ma, J., Li, J., et al. (2019). Exosome-mediated Secretion of LOXL4 Promotes Hepatocellular Carcinoma Cell Invasion and Metastasis. *Mol. Cancer* 18, 18. doi:10.1186/s12943-019-0948-8

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

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for Analysis of Tumor-Infiltrating Immune Cells. *Nucleic Acids Res.* 48, W509–W514. doi:10.1093/nar/gkaa407

Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* 173, 400–e11. doi:10.1016/j.cell.2018.02.052

Logsdon, C. D., Simeone, D. M., Binkley, C., Arumugam, T., Greenson, J. K., Giordano, T. J., et al. (2003). Molecular Profiling of Pancreatic Adenocarcinoma and Chronic Pancreatitis Identifies Multiple Genes Differentially Regulated in Pancreatic Cancer. *Cancer Res.* 63, 2649–2657.

Ma, L., Huang, C., Wang, X.-J., Xin, D. E., Wang, L.-s., Zou, Q. C., et al. (2017). Lysyl Oxidase 3 Is a Dual-Specificity Enzyme Involved in STAT3 Deacetylation and Deacetylimination Modulation. *Mol. Cel.* 65, 296–309. doi:10.1016/j.molcel.2016.12.002

Ma, W., Li, T., Wu, S., Li, J., Wang, X., and Li, H. (2019). LOX and ACSL5 as Potential Relapse Markers for Pancreatic Cancer Patients. *Cancer Biol. Ther.* 20, 787–798. doi:10.1080/15384047.2018.1564565

Marques, H. S., de Brito, B. B., da Silva, F. A. F., Santos, M. L. C., de Souza, J. C. B., Correia, T. M. L., et al. (2021). Relationship between Th17 Immune Response and Cancer. *Wjco* 12, 845–867. doi:10.5306/wjco.v12.i10.845

Miller, B. W., Morton, J. P., Pinese, M., Saturno, G., Jamieson, N. B., McGhee, E., et al. (2015). Targeting the LOX/Hypoxia axis Reverses many of the Features that Make Pancreatic Cancer Deadly: Inhibition of LOX Abrogates Metastasis and Enhances Drug Efficacy. *EMBO Mol. Med.* 7, 1063–1076. doi:10.15252/emmm.201404827

Minici, C., Rigamonti, E., Lanzillotta, M., Monno, A., Rovati, L., Maehara, T., et al. (2020). B Lymphocytes Contribute to Stromal Reaction in Pancreatic Ductal Adenocarcinoma. *Oncoimmunology* 9, 1794359. doi:10.1080/2162402X.2020.1794359

Molnar, J., Fong, K. S. K., He, Q. P., Hayashi, K., Kim, Y., Fong, S. F. T., et al. (2003). Structural and Functional Diversity of Lysyl Oxidase and the LOX-like Proteins. *Biochim. Biophys. Acta (Bba) - Proteins Proteomics* 1647, 220–224. doi:10.1016/s1570-9639(03)00053-0

Morrison, A. H., Byrne, K. T., and Vonderheide, R. H. (2018). Immunotherapy and Prevention of Pancreatic Cancer. *Trends Cancer* 4, 418–428. doi:10.1016/j.trecan.2018.04.001

Neesse, A., Algül, H., Tuveson, D. A., and Gress, T. M. (2015). Stromal Biology and Therapy in Pancreatic Cancer: a Changing Paradigm. *Gut* 64, 1476–1484. doi:10.1136/gutjnl-2015-309304

Park, J. S., Lee, J.-h., Lee, Y. S., Kim, J. K., Dong, S. M., and Yoon, D. S. (2016). Emerging Role of LOXL2 in the Promotion of Pancreas Cancer Metastasis. *Oncotarget* 7, 42539–42552. doi:10.18632/oncotarget.9918

Pei, H., Li, L., Fridley, B. L., Jenkins, G. D., Kalari, K. R., Lingle, W., et al. (2009). FKBP51 Affects Cancer Cell Response to Chemotherapy by Negatively Regulating Akt. *Cancer cell* 16, 259–266. doi:10.1016/j.ccr.2009.07.016

Quante, A. S., Ming, C., Rottmann, M., Engel, J., Boeck, S., Heinemann, V., et al. (2016). Projections of Cancer Incidence and Cancer-related Deaths in Germany by 2020 and 2030. *Cancer Med.* 5, 2649–2656. doi:10.1002/cam4.767

Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., and Matrisian, L. M. (2014). Projecting Cancer Incidence and Deaths to 2030: the Unexpected burden of Thyroid, Liver, and Pancreas Cancers in the United States. *Cancer Res.* 74, 2913–2921. doi:10.1158/0008-5472.CAN-14-0155

Ren, B., Cui, M., Yang, G., Wang, H., Feng, M., You, L., et al. (2018). Tumor Microenvironment Participates in Metastasis of Pancreatic Cancer. *Mol. Cancer* 17, 108. doi:10.1186/s12943-018-0858-1

Rhodes, D. R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J., Briggs, B. B., et al. (2007). Oncomine 3.0: Genes, Pathways, and Networks in a Collection of 18,000 Cancer Gene Expression Profiles. *Neoplasia* 9, 166–180. doi:10.1593/neo.07112

Rhodes, D. R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., et al. (2004). ONCOMINE: a Cancer Microarray Database and Integrated Data-Mining Platform. *Neoplasia* 6, 1–6. doi:10.1016/s1476-5586(04)80047-2

Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., et al. (2019). TISIDB: an Integrated Repository portal for Tumor-Immune System Interactions. *Bioinformatics* 35, 4200–4202. doi:10.1093/bioinformatics/btz210

Saatci, O., Kaymak, A., Raza, U., Ersan, P. G., Akbulut, O., Banister, C. E., et al. (2020). Targeting Lysyl Oxidase (LOX) Overcomes Chemotherapy Resistance in Triple Negative Breast Cancer. *Nat. Commun.* 11, 2416. doi:10.1038/s41467-020-16199-4

Schizas, D., Charalampakis, N., Kole, C., Economopoulou, P., Koustas, E., Gkotsis, E., et al. (2020). Immunotherapy for Pancreatic Cancer: A 2020 Update. *Cancer Treat. Rev.* 86, 102016. doi:10.1016/j.ctrv.2020.102016

Setargew, Y. F. I., Wyllie, K., Grant, R. D., Chitty, J. L., and Cox, T. R. (2021). Targeting Lysyl Oxidase Family Meditated Matrix Cross-Linking as an Anti-stromal Therapy in Solid Tumours. *Cancers* 13, 491. doi:10.3390/cancers13030491

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

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2021. *CA A. Cancer J. Clin.* 71, 7–33. doi:10.3322/caac.21654

Siemers, N. O., Holloway, J. L., Chang, H., Chasalow, S. D., Ross-MacDonald, P. B., Voliva, C. F., et al. (2017). Genome-wide Association Analysis Identifies Genetic Correlates of Immune Infiltrates in Solid Tumors. *PloS one* 12, e0179726. doi:10.1371/journal.pone.0179726

Stromnes, I. M., Hulbert, A., Pierce, R. H., Greenberg, P. D., and Hingorani, S. R. (2017). T-cell Localization, Activation, and Clonal Expansion in Human Pancreatic Ductal Adenocarcinoma. *Cancer Immunol. Res.* 5, 978–991. doi:10.1158/2326-6066.CIR-16-0322

Sun, L., Zhang, X., Song, Q., Liu, L., Forbes, E., Tian, W., et al. (2021). IGFBP2 Promotes Tumor Progression by Inducing Alternative Polarization of Macrophages in Pancreatic Ductal Adenocarcinoma through the STAT3 Pathway. *Cancer Lett.* 500, 132–146. doi:10.1016/j.canlet.2020.12.008

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Tan, H. Y., Wang, N., Zhang, C., Chan, Y. T., Yuen, M. F., and Feng, Y. (2021). Lysyl Oxidase-Like 4 Fosters an Immunosuppressive Microenvironment during Hepatocarcinogenesis. *Hepatology* 73, 2326–2341. doi:10.1002/hep.31600

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a Web Server for Cancer and normal Gene Expression Profiling and Interactive Analyses. *Nucleic Acids Res.* 45, W98–W102. doi:10.1093/nar/gkx247

Tomczak, K., Czerwińska, P., and Wiznerowicz, M. (2015). Review the Cancer Genome Atlas (TCGA): an Immeasurable Source of Knowledge. *wo* 1A, 68–77. doi:10.5114/wo.2014.47136

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Tissue-based Map of the Human Proteome. *Science* 347, 1260419. doi:10.1126/science.1260419

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

UniProt, C. (2021). UniProt: the Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* 49, D480–D9. doi:10.1093/nar/gkaa1100

Van Cutsem, E., Tempero, M. A., Sigal, D., Oh, D.-Y., Fazio, N., Macarulla, T., et al. (2020). Randomized Phase III Trial of Pegvorhyaluronidase Alfa with Nab-Paclitaxel Plus Gemcitabine for Patients with Hyaluronan-High Metastatic Pancreatic Adenocarcinoma. *J. Clin. Oncol.* 38, 3185–3194. doi:10.1200/JCO.20.00590

Vasaikar, S. V., Straub, P., Wang, J., and Zhang, B. (2018). LinkedOmics: Analyzing Multi-Omics Data within and across 32 Cancer Types. *Nucleic Acids Res.* 46, D956–D963. doi:10.1093/nar/gkx1090

Wang, T.-H., Hsia, S.-M., and Shieh, T.-M. (2016). Lysyl Oxidase and the Tumor Microenvironment. *Int. J. Mol. Sci.* 18, 62. doi:10.3390/ijms18010062

Wilgus, M.-L., Borczuk, A. C., Stoopler, M., Ginsburg, M., Gorenstein, L., Sonett, J. R., et al. (2011). Lysyl Oxidase: a Lung Adenocarcinoma Biomarker of Invasion and Survival. *Cancer* 117, 2186–2191. doi:10.1002/cncr.25768

Winter, J. M., Brennan, M. F., Tang, L. H., D'Angelica, M. I., Dematteo, R. P., Fong, Y., et al. (2012). Survival after Resection of Pancreatic Adenocarcinoma: Results from a Single Institution over Three Decades. *Ann. Surg. Oncol.* 19, 169–175. doi:10.1245/s10434-011-1900-3

Wu, L., and Zhu, Y. (2015). The Function and Mechanisms of Action of LOXL2 in Cancer (Review). *Int. J. Mol. Med.* 36, 1200–1204. doi:10.3892/ijmm.2015.2337

Zhang, Y., Jiang, W. L., Yang, J. Y., Huang, J., Kang, G., Hu, H. B., et al. (2019). Downregulation of Lysyl Oxidase-like 4 LOXL4 by miR-135a-5p Promotes Lung Cancer Progression *In Vitro* and *In Vivo*. *J. Cel Physiol* 234, 18679–18687. doi:10.1002/jcp.28508

Zhao, R., Liu, W., Wang, M., Zhang, Y., Pan, L., Feng, F., et al. (2020). Lysyl Oxidase Inhibits TNF-α Induced Rat Nucleus Pulposus Cell Apoptosis via Regulating Fas/FasL Pathway and the P53 Pathways. *Life Sci.* 260, 118483. doi:10.1016/j.lfs.2020.118483

Zhao, W., Yang, A., Chen, W., Wang, P., Liu, T., Cong, M., et al. (2018). Inhibition of Lysyl Oxidase-like 1 (LOXL1) Expression Arrests Liver Fibrosis Progression in Cirrhosis by Reducing Elastin Crosslinking. *Biochim. Biophys. Acta (Bba) - Mol. Basis Dis.* 1864, 1129–1137. doi:10.1016/j.bbadis.2018.01.019

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat. Commun.* 10, 1523. doi:10.1038/s41467-019-09234-6

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

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

frontiers | Frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 13 April 2022
doi: 10.3389/fmolb.2022.838654

# A Novel Gene Signature Associated With "E2F Target" Pathway for Predicting the Prognosis of Prostate Cancer

*Haoran Xia*[1,2†], *Miaomiao Wang*[1,2†], *Xiaonan Su*[3], *Zhengtong Lv*[1,2], *Qiuxia Yan*[1,4], *Xiaoxiao Guo*[1,2*] *and Ming Liu*[1,2*]

[1]*Department of Urology, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China,* [2]*Graduate School of Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China,* [3]*Department of Urology, Zoucheng People's Hospital, Zoucheng, China,* [4]*Peking University Fifth School of Clinical Medicine, Beijing, China*

OPEN ACCESS

***Edited by:***
*Liang Wang,*
*Moffitt Cancer Center and Research Institute, United States*

***Reviewed by:***
*Raffaella Pippa,*
*Sanford Burnham Prebys Medical Discovery Institute, United States*
*Pankaj Pathak,*
*National Institutes of Health (NIH), United States*

****Correspondence:***
*Xiaoxiao Guo*
*ensitern@163.com*
*Ming Liu*
*liumingbjh@126.com*

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

***Specialty section:***
*This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences*

***Received:*** *18 December 2021*
***Accepted:*** *15 March 2022*
***Published:*** *13 April 2022*

***Citation:***
*Xia H, Wang M, Su X, Lv Z, Yan Q, Guo X and Liu M (2022) A Novel Gene Signature Associated With "E2F Target" Pathway for Predicting the Prognosis of Prostate Cancer. Front. Mol. Biosci. 9:838654. doi: 10.3389/fmolb.2022.838654*

**Background:** The effect of the adenoviral early region 2 binding factors (E2Fs) target pathway on prostate cancer is not clear. It is necessary to establish an E2F target-related gene signature to predict prognosis and facilitate clinical decision-making.

**Methods:** An E2F target-related gene signature was established by univariate and LASSO Cox regression analyses, and its predictive ability was verified in multiple cohorts. Moreover, the enrichment pathway, immune microenvironment, and drug sensitivity of the activated E2F target pathway were also explored.

**Results:** The E2F target-related gene signature consisted of *MXD3*, *PLK1*, *EPHA10*, and *KIF4A*. The patients with high-risk scores showed poor prognosis, therapeutic resistance, and immunosuppression, along with abnormal growth characteristics of cells. Tinib drugs showed high sensitivity to the expression of *MXD3* and *EPHA10* genes.

**Conclusion:** Our research established an E2F target-related signature for predicting the prognosis of prostate cancer. This study provides insights into formulating individualized detection and treatment as well as provides a theoretical basis for future research.

**Keywords: "E2F target" pathway, prostate cancer, gene signature, prognosis, immune infiltration, therapeutic resistance**

## INTRODUCTION

Prostate cancer is the second most common male malignant neoplasia and the fifth leading cause of cancer death in men worldwide. With an increase in the aging population, 2.3 million new cases of prostate cancer and 740,000 deaths are expected to occur globally in the next 20 years. Despite a decline in global incidence rates, the incidence in China shows an annual increase of 2.6% (Culp et al., 2020). Early diagnosis plays a key role in the prognosis of prostate cancer. To further improve patient outcomes, new molecular markers need to be identified to allow a more reliable diagnosis and prognosis.

The adenoviral early region 2 binding factors (E2Fs) of the transcription factor family are critical regulators of cell cycle progression (Sun et al., 2007). In response to mutation or phosphorylation, RB1 is inactivated, causing E2Fs to detach from the E2F–RB1 complex to bind

with certain promoters of the E2F target genes (Van den Heuvel and Dyson, 2008; Hallstrom and Nevins, 2009). The high expression of the E2F target gene plays a pivotal role in tumorigenesis and is related to poor prognosis in many tumors, including neuroblastoma (Molenaar et al., 2012), breast cancer (Oshi et al., 2021), high-grade serous ovarian cancer (Dahl et al., 2019), and prostate cancer (Wang et al., 2021).

In this study, we aimed to find a novel prognosis gene signature to guide further clinical decision-making for patients with prostate cancer. Briefly, the effect of the E2F target pathway on the poor prognosis of prostate cancer was determined by single-sample gene set enrichment analysis (ssGSEA), and then the prognostic gene set related to the E2F target pathway was established by weighted gene co-expression network analysis (WGCNA) and differentially expressed gene (DEG) analysis. Using a cohort from The Cancer Genome Atlas (TCGA) database, a gene signature was obtained by univariate and least absolute shrinkage and selection operator (LASSO) Cox regression analyses, and the risk value of the E2F target pathway was calculated. Then, we verified the enrichment of the E2F target pathway and the worse prognosis in the high-risk group in two separate cohorts from the International Cancer Genome Consortium (ICGC) and Gene Expression Omnibus (GEO) databases. In addition, the study explored functional differences among different risk groups. Our developed gene signature could facilitate early screening, predict prognosis, and provide patients with more individualized treatments.

# METHODS

## Data Preparation and Processing

In total, 1,145 patients with prostate adenocarcinoma (PRAD) from TCGA, ICGC, and GEO databases were enrolled in this study. Of these, the data of 540 patients with PRAD from TCGA database were downloaded as the training cohort (https://portal.gdc.cancer.gov/). Two test cohorts were used, namely, test cohort I consisting of 357 patients with PRAD from the ICGC database (https://dcc.icgc.org/projects/PRAD-CA) and test cohort II consisting of 248 patients with PRAD from the GSE116918 database (radical radiotherapy with ADT) (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE116918). All downloaded data included fragments per kilobase of sequences per million mapped reads (FPKM)-normalized RNA sequencing (RNA-seq) data, clinical characteristics annotation, and follow-up information such as biochemical recurrence (BCR), metastasis (Met), and overall survival (OS). Before further analysis, all RNA-seq data included were log2-transformed and normalized by the R package "sva."

## Construction of E2F Target Signature

First, based on the HALLMARKS gene set from the Molecular Signatures Database (MSigDB) (http://www.gsea-msigdb.org/gsea/msigdb/search.jsp), ssGSEA was calculated using the R package "GSVA" in the training set (Lee et al., 2008). The R package "WGCNA" was used to perform WGCNA by using TCGA database mRNA matrix (Langfelder and Horvath, 2008). An adjacency matrix was constructed to describe the correlation strength between the nodes and then transformed to a topological overlap matrix (TOM). Next, hierarchical clustering was performed to identify the modules by setting the minimum number of genes in each module to 60. After merging similar modules, the module genes with the highest correlation to E2F target scores in ssGSEA results were identified. These genes were then combined with the E2F target gene set in HALLMARKS to obtain a new gene set, denoted as gene set A.

Simultaneously, Kaplan–Meier survival and log-rank tests (R package "survival") were performed using gene set A and BCR information to screen genes associated with prognosis—which were labeled as gene set B. DEGs between TCGA-PRAD tumors and normal tissues were obtained by using R package "limma" (false discovery rate [FDR] < 0.05, log|fold change [FC]|> 1) (Ritchie et al., 2015). E2F target-related candidate prognostic genes were obtained by intersecting DEGs with gene set B. Furthermore, these genes were enrolled for inclusion in LASSO Cox regression (R package "glmnet") (Tibshirani, 1997). Finally, the remaining genes after 10,000 contractions were selected, and their coefficients were recorded to obtain an E2F target-related gene signature.

## Statistical and Bioinformatics Analyses

All analyses and graphs were executed using R v4.0.3 (http://www.r-project.org) and GraphPad v8.0.3 (https://www.graphpad.com/). The Pearson correlation test was used to calculate the established inter-gene correlations in the gene signature. After the E2F target-related risk score was calculated based on the signature, the training set (TCGA cohort) and the two test sets (cohorts from ICGC and GEO) were divided into high- and low-risk groups according to the median risk score in the training set. Differences between the groups were calculated using *t*-tests and chi-square tests. Scatter maps and heat maps were used to visualize the survival distribution and gene expression patterns of patients in different groups (R packages "ggplot2" and "scales"). Furthermore, the dimension of high-latitude data was reduced by using *t*-distributed stochastic neighbor embedding (tSNE) and principal components analysis (PCA) of the R package "Rtsne" to test whether the gene signature could divide patients into different groups (Reich et al., 2008; van der Maaten and Hinton, 2008). The prognostic ability of the gene signature was reflected by the Kaplan–Meier survival analysis, receiver operating characteristic (ROC) curve, and time-dependent ROC (T-ROC) curve (R packages "survival" and "timeROC"). The multivariate Cox stepwise regression models of the R package "survival" were used to test whether the gene signature was an independent predictor of prognosis.

The function between groups was analyzed by GSEA v4.1.0. For this algorithmic analysis, it was verified whether the E2F target pathway was activated in different groups. In addition, GSEA based on Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases was used to study functional differences according to the DEGs in different groups (R package "GOplot" and "limma") (Subramanian et al., 2005).

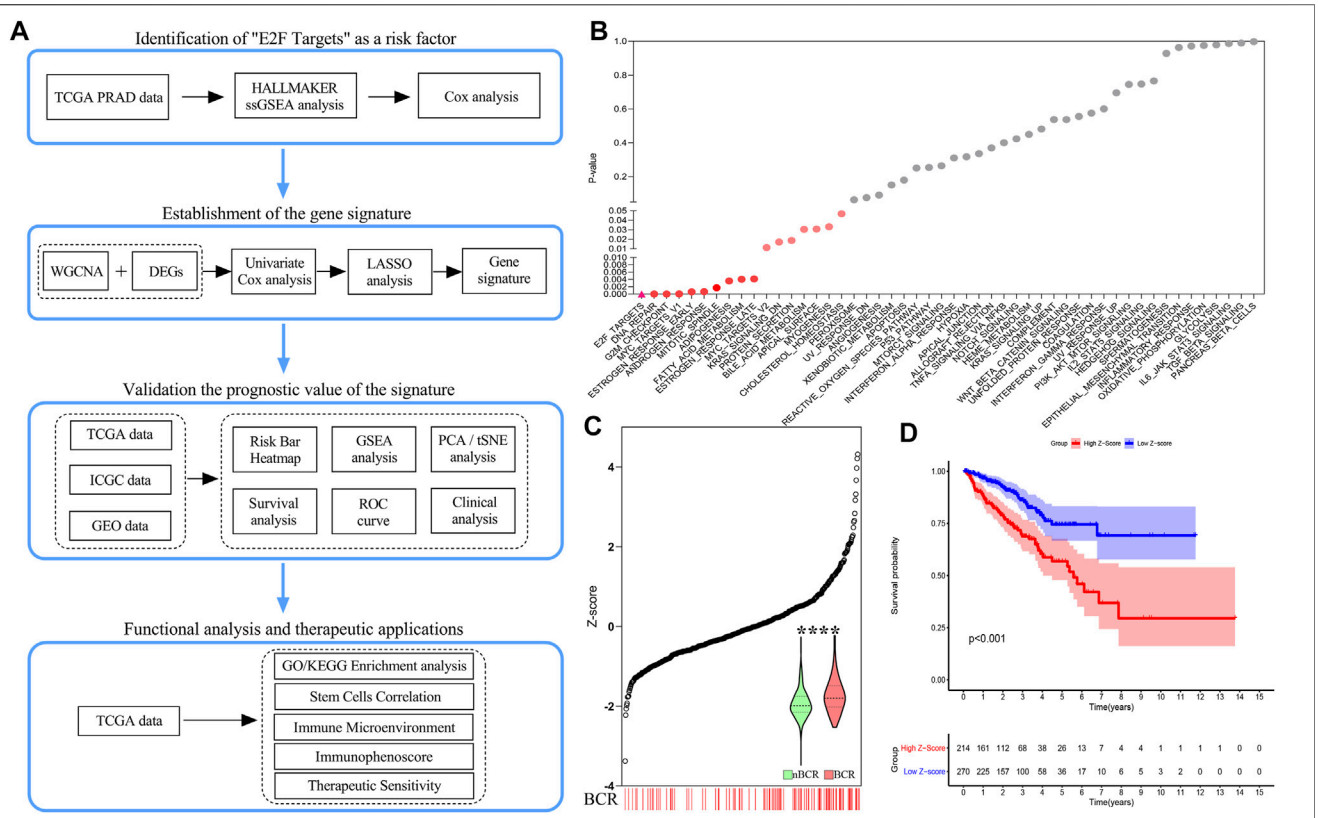



**FIGURE 1 |** Schematic diagram and identification of risk factors for the "E2F target" pathway. **(A)** Research design and process diagram. **(B)** Univariate Cox regression analysis revealed a significant correlation between the "E2F target" pathway and poor prognosis. **(C)** The number of BCR patients increased significantly with increasing z-score. **(D)** The prognosis of the high z-score group was worse than that of the low z-score group by Kaplan–Meier analysis.

The immune cells were divided into different types based on previous studies. Meanwhile, the differences of immune infiltration were explored by ssGSEA (immune gene set from MSigDB) and the CIBERSORT algorithm (https://cibersortx.stanford.edu/), as well as the immunophenoscore (IPS) and immunophenotyping (Newman et al., 2015; Charoentong et al., 2017; Thorsson et al., 2018). Finally, gene correlation analysis based on the CellMiner database (https://discover.nci.nih.gov/cellminer/home.do) explored the sensitivity between the genes and drugs to screen for new medicines and therapeutic targets (Pommier et al., 2012).

# RESULTS

## E2F Target Pathway Identified as a Prognostic Risk Factor

The overall process is shown by a schematic diagram in **Figure 1A**. In the first step, ssGSEA scores of 50 pathways in each sample were included in univariate Cox regression analysis. For the BCR of the training cohort, the E2F target pathway had the greatest statistical significance ($p < 0.0001$, **Figure 1B**). The E2F target pathway score was z-score-transformed and divided into high- and low-z-score groups according to their median value. The incidence of BCR in the high-z-score group was significantly higher than that of the low-group ($p < 0.0001$, **Figure 1C**). Moreover, the Kaplan–Meier survival curves revealed a worse prognosis in the high-z-score group ($p < 0.001$, **Figure 1D**). The details of the training cohort are presented in **Supplementary Table S1**.

## Candidate Gene Screening and Gene Signature Construction

A total of 35 gene modules were identified by WGCNA, among which the expression of the black module was most related to the increase in the E2F target z-score, with a total of 806 genes (**Figures 2A,B** and **Supplementary Table S2**). Then, 1,006 genes related to the E2F target pathway were obtained by combining the 806 genes with those of the E2F target pathway from MSigDB (**Supplementary Table S2**). After univariate Cox regression analysis, 359 E2F-related prognostic genes were screened out (**Figure 2C** and **Supplementary Table S3**). Concurrently, 107 DEGs were identified by differential analysis between cancer and normal prostate tissue (**Figure 2D** and **Supplementary Table S3**), and a total of 44 candidate "E2F target"-related prognostic genes were obtained by intersecting these DEGs with the previously screened genes (**Figure 2E** and **Supplementary Table S3**). Subsequently, the 44 candidate genes were further screened by a LASSO Cox regression model, optimized when the

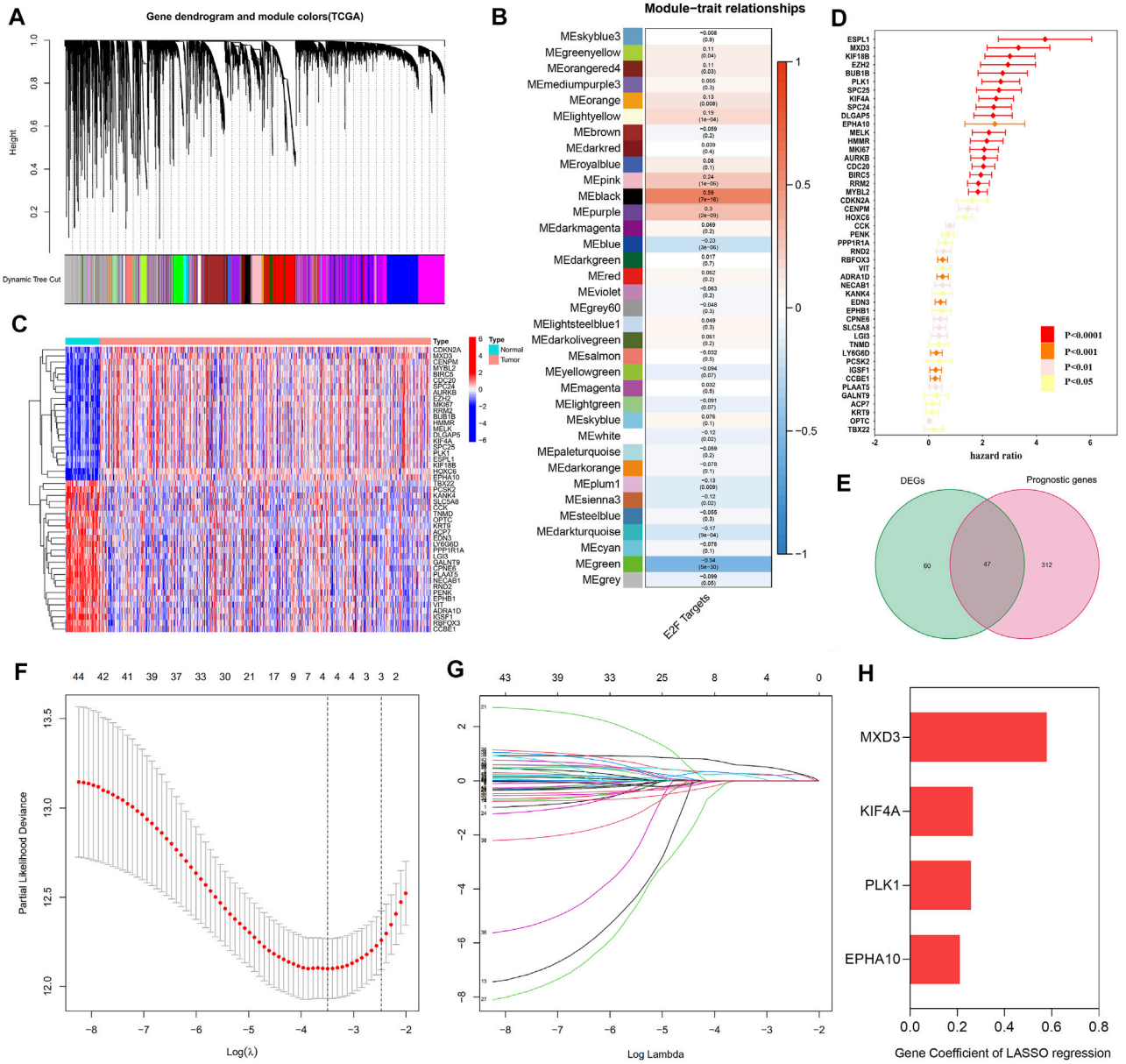



**FIGURE 2 |** Establishment of the "E2F target" risk score. **(A–B)** WGCNA showed that the gene in MEblack had the highest correlation with the activation of the "E2F target" pathway. **(C)** The heat map showed the top part of DEGs between the prostate tumor and normal prostate tissue. **(D)** Prognostic genes with statistical significance were screened by univariate Cox regression analysis. **(E)** The Venn diagram showed 47 prognostic differential genes obtained after the intersection of DEGs and prognostic genes. **(F–G)** LASSO Cox regression was used to establish the signature; the best log(λ) value was -3.5, and 4 indicators remained. **(H)** The LASSO coefficients of the 4 genes in the signature.

minimum λ value was 0.03035, with four genes remaining, including MAX dimerization protein 3 (MXD3), polo-like kinase 1 (PLK1), EPH receptor A10 (EPHA10), and kinesin family member 4A (KIF4A) (**Figures 2F,G**). **Figure 2H** and **Supplementary Table S3** show the coefficients of these four genes, and the final risk score was calculated as follows:

Risk score = MXD3* 0.579083868008915 + PLK1* 0.258020225608399 + EPHA10* 0.211234188276561 + KIF4A * 0.266044684281992.

## High E2F Target-Related Risk Score Associated With Poor Prognosis in Training Cohort

As the risk score increased, the gene expression of the signature rose, the number of patients with BCR in the training cohort increased significantly, and the BCR-free survival time decreased (**Figure 3A**). The genetic correlation analysis did not show excessive high correlations (**Figure 3B** and **Supplementary**

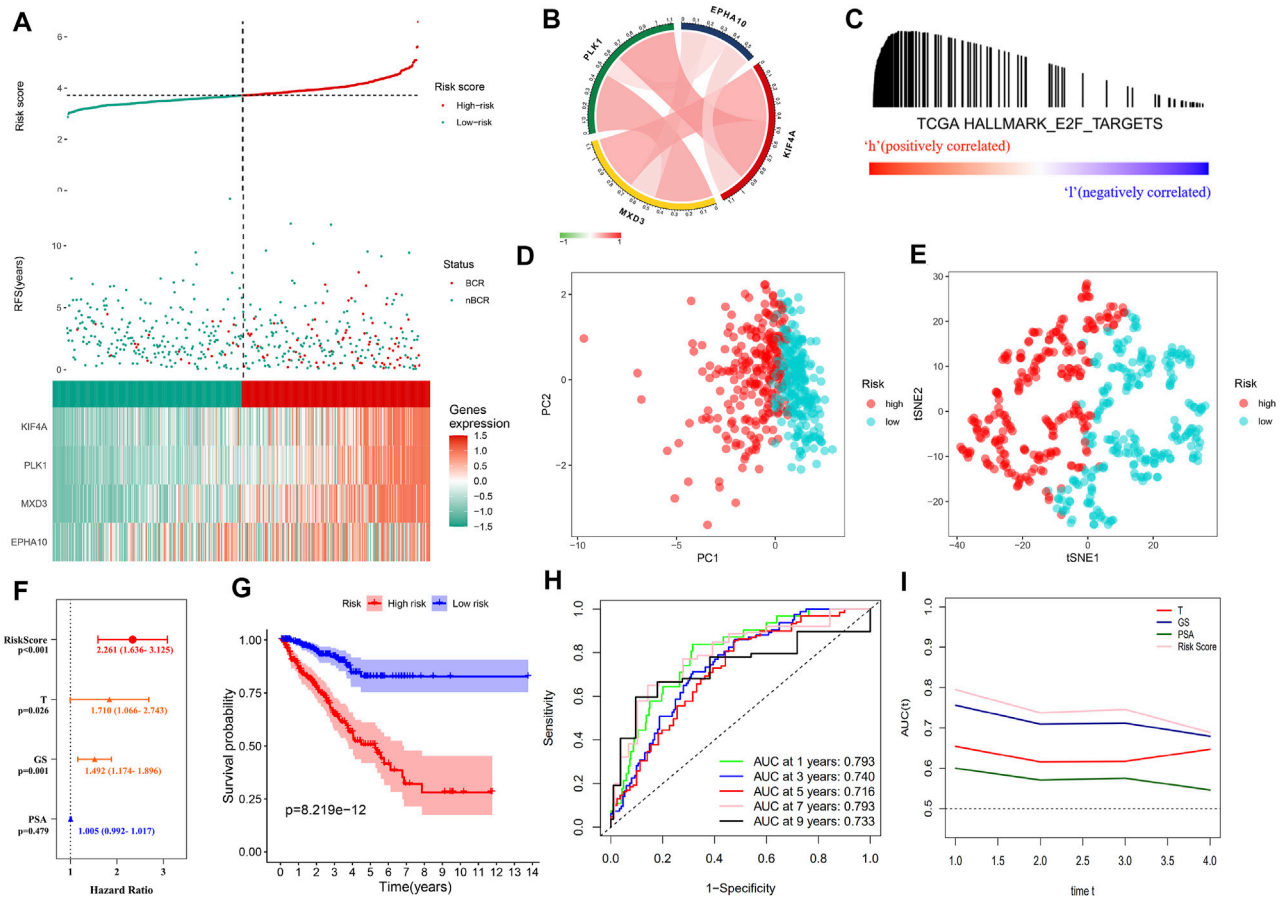



**FIGURE 3 |** Distribution of the "E2F target" risk score in TCGA cohort. **(A)** Patient risk score–survival distribution map and gene expression heat map of "E2F target"-related genes. **(B)** Gene correlation heat map showing low correlation among genes. **(C)** GSEA proved that in the high-risk group, the "E2F Target" pathway was activated. **(D–E)** PCA and tSNE analysis indicated that the model could be used to distinguish between different risk groups. **(F)** The "E2F target" risk score was an independent prognostic factor for poor prognosis. **(G)** Kaplan–Meier survival analysis confirmed that patients with a high-risk had a worse prognosis. **(H)** The 10-year AUC determined by ROC analysis of the gene signature was relatively high, suggesting that the predictive ability of the signature was good. **(I)** Compared with the T stage, Gleason score, and PSA, the average AUC of the risk score was the highest, indicating that the predictive ability of the risk score was the best.

**Table S4**). According to the median risk score of 3.713494217, the patients were divided into high- and low-risk groups. GSEA confirmed the enrichment of the E2F target pathway genes in the high-risk group (**Figure 3C**). PCA and tSNE analysis showed that patients in the high- and low-risk groups could be completely distinguished from each other (**Figures 3D,E**). Multivariate Cox regression analysis showed that the risk score was an independent predictor of prognosis and the strongest predictor of BCR along with clinical features (HR: 2.261, 95% CI: 1.636–3.125, $p < 0.001$, **Figure 3F**). The Kaplan–Meier survival curve showed that the BCR-free survival time in the high-risk group was significantly lower than that in the low-risk group ($p < 0.0001$, **Figure 3G**). **Figure 3H** shows that the area under the ROC curve (AUC) of the risk score in the training group within 9 years was all higher than 0.7, which indicates that the average predictive ability of the risk score was strong. The T-ROC curve revealed that the AUC of the risk score was always higher than that of other clinical features and tended to be stable over time (**Figure 3I**).

## Verification of the Prognostic Effect of E2F Target-Related Risk Score in the Test Cohort

The risk score of each sample was calculated in test cohorts I and II, which were divided into high- and low-risk groups according to the median risk score of the training cohort (i.e., 3.713494217). The four-gene signature showed the same expression pattern as the training cohort, and the distribution of deaths in test cohort I and metastatic patients in test cohort II increased with the risk score (**Figures 4A–C**). In addition, in the high-risk groups of the two cohorts, the E2F target pathway was shown to be activated (**Figure 4D**). In addition, the composition of the gene signature was verified to distinguish between patients at different risks and to predict poor prognosis within 10 years in the two cohorts (**Figures 4E,F**). The detailed risk information and gene expression of the training and test cohorts are represented in **Supplementary Table S5**, and the relevant information generated in the process of GSEA is shown in **Supplementary Table S6**.

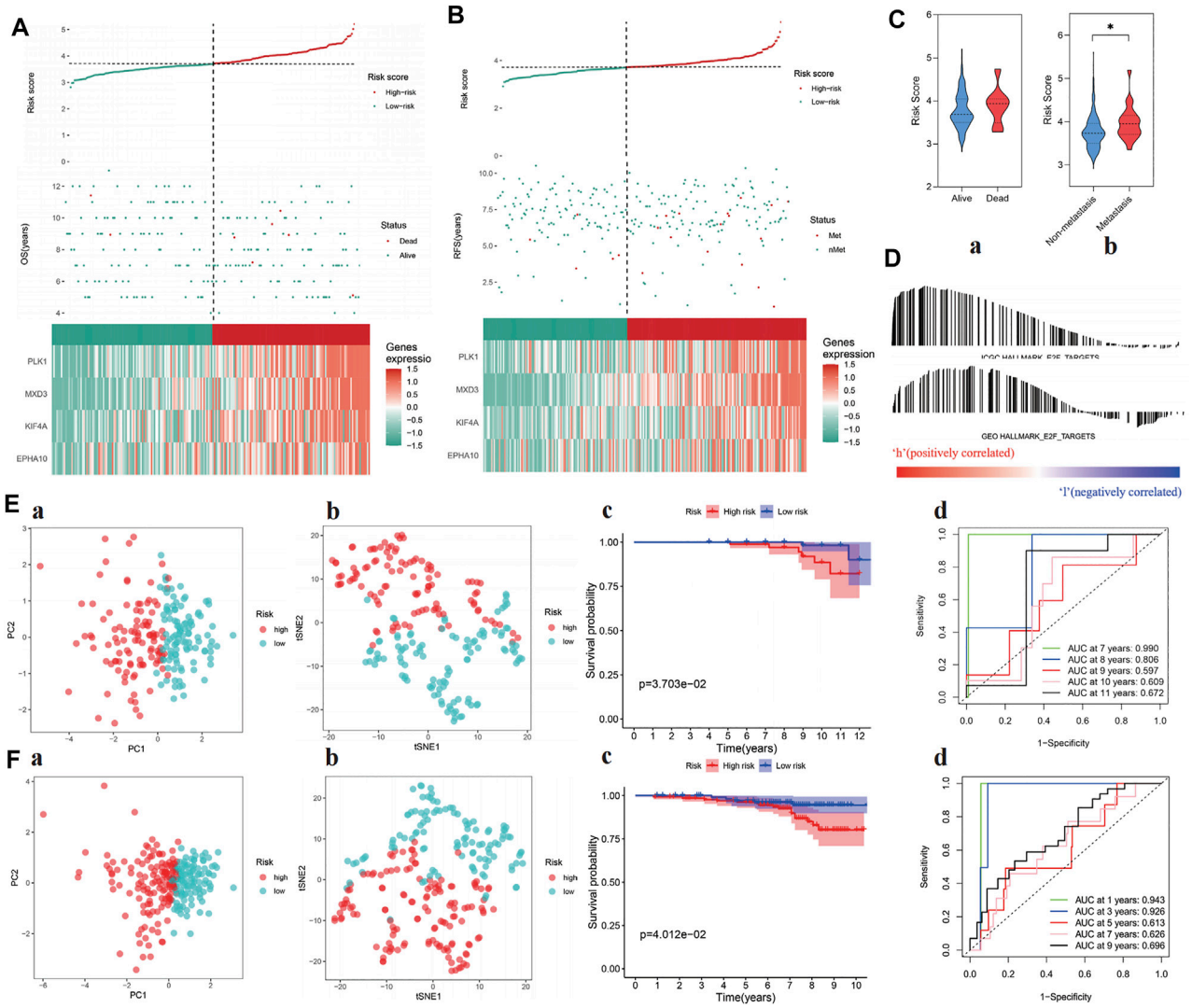



**FIGURE 4 |** Verification of prognostic ability of "E2F target" risk score in ICGC and GEO cohort. **(A–B)** Patients' risk score–survival distribution map and gene expression heat map of "E2F target"-related genes in ICGC (left) and GEO (right) cohorts. **(C)** The "E2F target" risk score of patients with poor prognosis increased in both ICGC (left) and GEO (right) cohorts. **(D)** GSEA proved that at the high-risk group, the "E2F target" pathway was activated in both cohorts. **(E–F)** The ability of the "E2F target" risk score to distinguish high-risk patients was verified by PCA and tSNE in ICGC (Ea, Eb) and GEO (Fa, Fb) cohorts, and the prognosis of the high-risk group was significantly worse than the low-risk group (Ec, Ed, Fc, and Fd).

## Correlation Between Gene Signature and Clinical Characteristics

Pearson correlation analysis was used to study the distribution of patients among different groups with various clinical feature stratifications. As shown in **Figure 5A**, in the training cohort, the high-risk group was closely associated with higher PSA, Gleason score (GS), and T stages. Similarly, in test cohorts I and II, the PSA, GS, and T stages tended to increase in the high-risk group (**Figures 5B,C**). Unfortunately, this trend was not statistically significant. In this regard, we gave a cautious explanation. In test cohort I, as OS was the primary endpoint, a shortage of other endpoints arose, which thus affected the statistically significant results. In contrast, test cohort II received radiotherapy plus androgen deprivation therapy, which, from the perspective of the E2F target pathway, may have been more beneficial to patients, thus affecting OS and the statistically significant results.

## Functional Differences Among Different Risk Groups

The DEGs in the high- and low-risk groups were calculated using the "limma" algorithm. **Figure 6A** shows all the up- and downregulated DEGs screened by using $p < 0.05$ as a threshold. The functional enrichment analysis of DEGs was performed, and the results of KEGG analysis showed that DNA replication, base excision repair, and mismatch repair

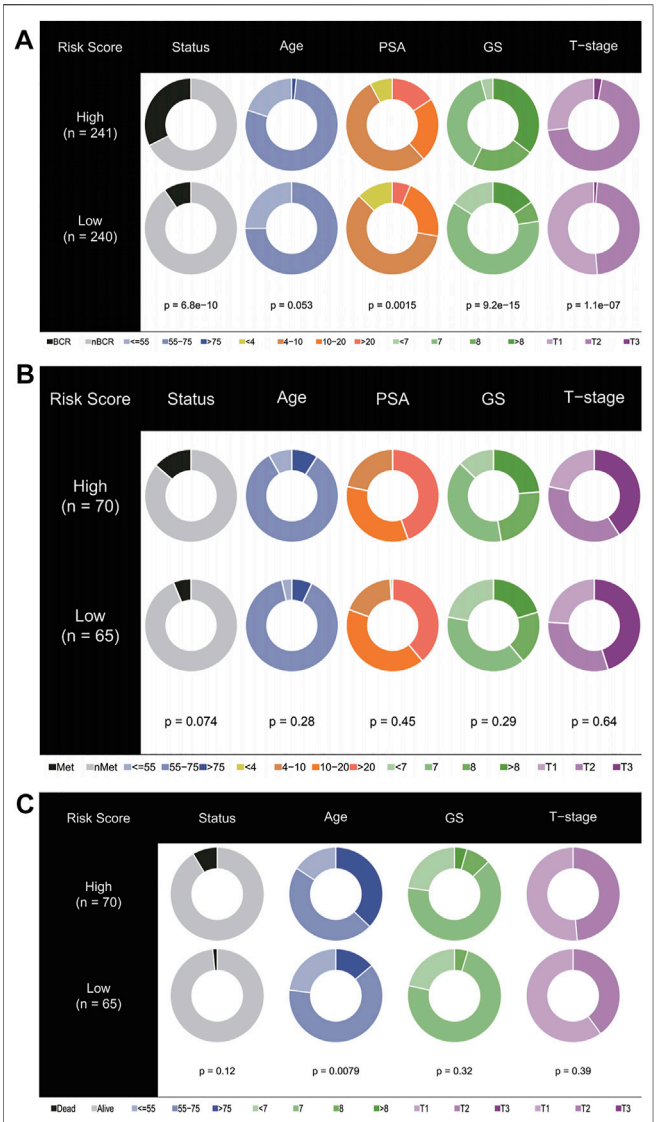



**FIGURE 5 |** Stratified proportion of clinical characteristics of patients with prostate cancer in high- and low-risk groups. In the TCGA **(A)**, ICGC **(B)**, and GEO **(C)** cohorts, patients with higher PSA, GS, and T stage accounted for an increase in the high-risk group.

were active in the high-risk group. Surprisingly, histidine metabolism was suppressed in this group (**Figure 6B**). As shown in **Figures 6C,D** and **Supplementary Table S7**, the GO analysis showed that in the high-risk group, the activation of the E2F target pathway participated in the binding of many factors and proteins and may regulate cell growth, transcription, apoptosis, and other metabolic activities by interfering with binding processes. The study also found the regulation of the Wnt and p53 signaling pathways in the high-risk group, suggesting that these pathways may be abnormally activated. In addition, the study found that the activation of the E2F target pathway was also related to remodeling of the extracellular matrix, which may enhance the invasive ability of tumor cells, promote cancer metastasis, and lead to poor prognosis. Moreover, the correlation between the risk scores and the characteristics of tumor stem cells was studied. A significant positive correlation was found between the risk score and stem cell characteristics (**Figures 6E,F**). All relevant information to support these results is presented in **Supplementary Table S7**.

## Immune Microenvironment and Therapeutic Sensitivity

**Figure 7A** and **Supplementary Table S8** show the immune infiltration of high- and low-risk groups using different algorithms. From the ssGSEA and CIBERSORT thermograms, the immune infiltration of patients in the high-risk group was more complex with both immune activation and immunosuppression and varying degrees. However, it is worth noting that both the algorithms showed significant infiltration of T regulatory cells in the high-risk group (**Figure 7B**). Concomitantly, negative immune regulatory genes, such as *EZH2*, *HAVCR1/2*, and *DNMT1* were also observed to be significantly expressed in the high-risk groups. When focusing on the immune subtype, it was found that the C3 type with good prognosis was concentrated in patients with low-risk scores, while the higher the risk score, the more the C1, C2, and C4 types were linked to poor prognosis (**Figures 7C,D**). Continuously, the IPS algorithm was used to calculate the immunophenotype of the samples with the top 10 and the last 10 risk scores, which once again verified the high expression of immune suppressor cells (SCs) and a decrease in the IPS in the high-risk group (**Figure 7E** and **Supplementary Figure S1**).

GSEA and sensitivity analysis of commonly used drugs reflected the resistance of high expression of different genes to treatment (**Figures 8A,B** and **Supplementary Table S9**). *MXD3* had certain sensitivity to commonly used drugs, although not that strong. In addition, sensitive drugs were screened using the CellMiner database (**Supplementary Table S9**), and the results showed that -tinib drugs were more sensitive to *MXD3* and *EPHA10*, suggesting that they may become potential new therapeutic targets.

# DISCUSSION

The progression of prostate cancer is interrelated with multiple genes. Certain tumor molecular markers can predict the prognosis of patients for more precise personalized treatment plans in patients with cancer. Retinoblastoma tumor suppressor protein (RB) is a pivotal regulator of the cell cycle and functionally inactive in most cancers, including prostate cancer (Classon and Harlow, 2002; Jarrard et al., 2002). A previous study showed that RB plays a tumor suppressor function in prostate cancer (Bookstein et al., 1990) and plays a key role in the cell cycle regulation by regulating the adenoviral early region 2 binding factors (E2Fs) of transcription factor family (Fischer and Müller, 2017; Wu and Wu, 2021). The E2F family of transcription factors (E2Fs) play an important role in cell cycle regulation; E2F1, E2F3, and E2F7 were reportedly involved in the G1-S transition procession of the cell cycle (Xie et al., 2021). E2F transcription

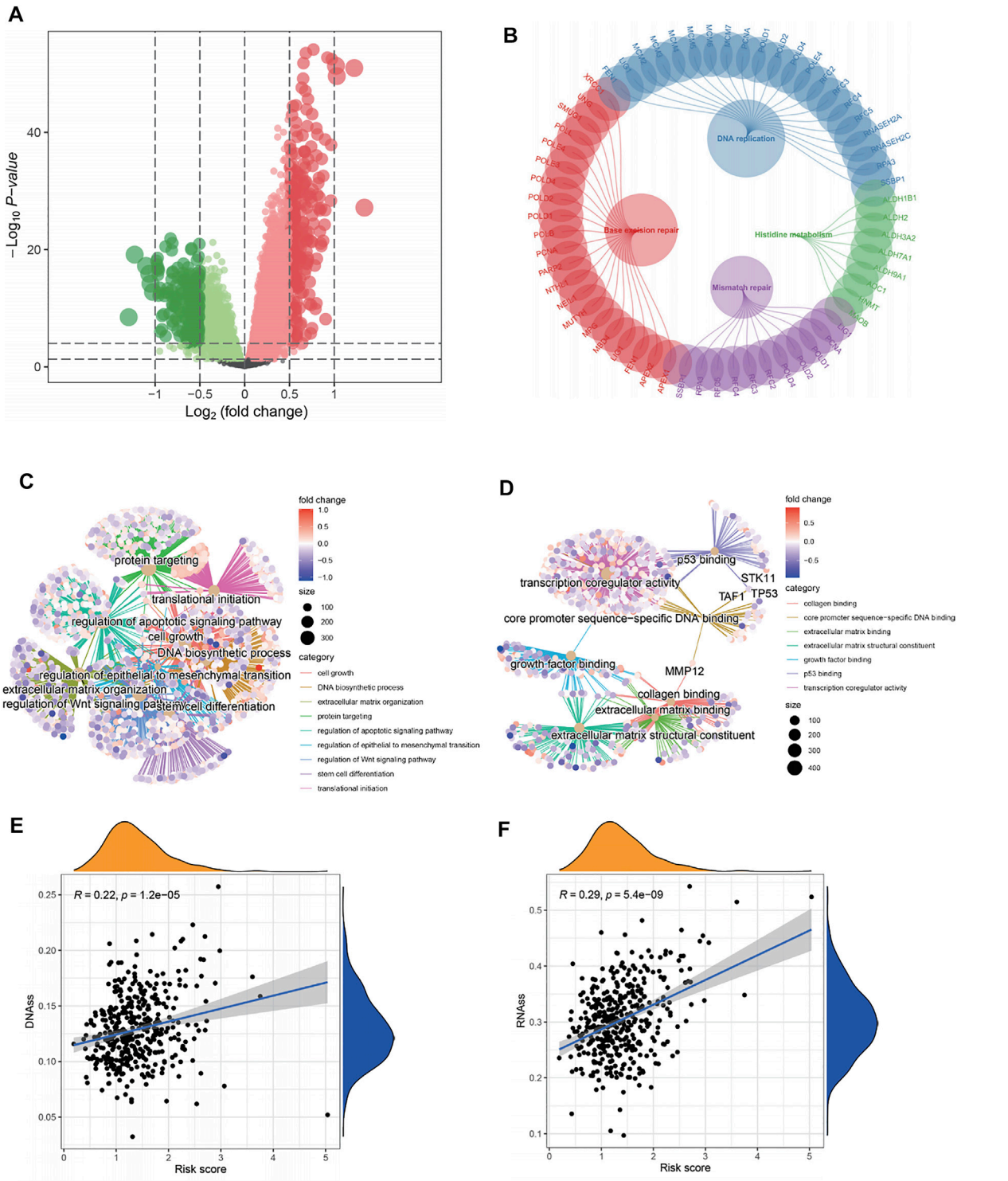



**FIGURE 6 |** Functional analysis and correlation analysis of stem cell characteristics and risk score. **(A)** Volcano diagram of DEGs between high- and low-risk groups. **(B–D)** KEGG and GO enrichment analysis showed that DNA replication, mismatch repair, cell growth, extracellular matrix remodeling, and binding pathways were active in patients with a high-risk, as well as the p53 pathway, wnt pathway, and stem cell differentiation. On the contrary, histidine metabolism was inhibited. **(E–F)** The "E2F target" risk score was positively correlated with DNAs and RNAs.

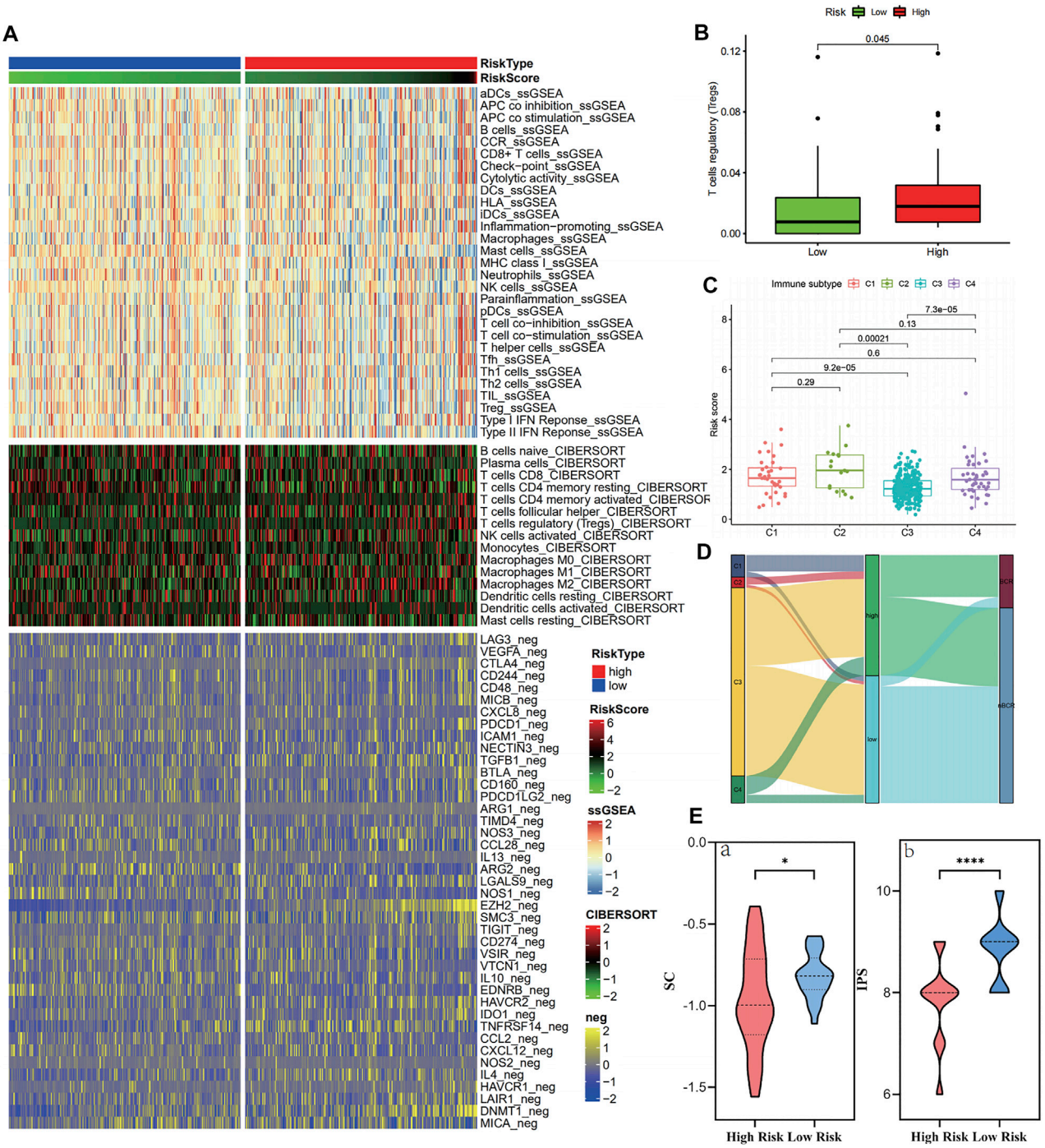



**FIGURE 7 |** Immune cell infiltration in the high- and low-risk groups in the training cohort. **(A–B)** Tregs and negative immunoregulatory genes were highly expressed in the high-risk score group. **(C–D)** Immunotyping results showed that C1, C2, and C4 types were increased in the high-risk group. **(E)** The IPS algorithm showed a significant increase in SC (a) and CP (b) in the high-risk group.

factors not only regulate the expression of target genes but also ensure that target genes are mainly transcribed in a cell cycle-dependent manner (Fang et al., 2020). Therefore, E2F transcription factor abnormalities play a role in tumorigenesis. Mutations in RB are inactivated in prostate cancer, leading to RB–E2F complex dissociation, followed by free E2F binding to the promoter of certain E2F target genes, in turn controlling the progression of tumorigenesis. Various E2F target genes are

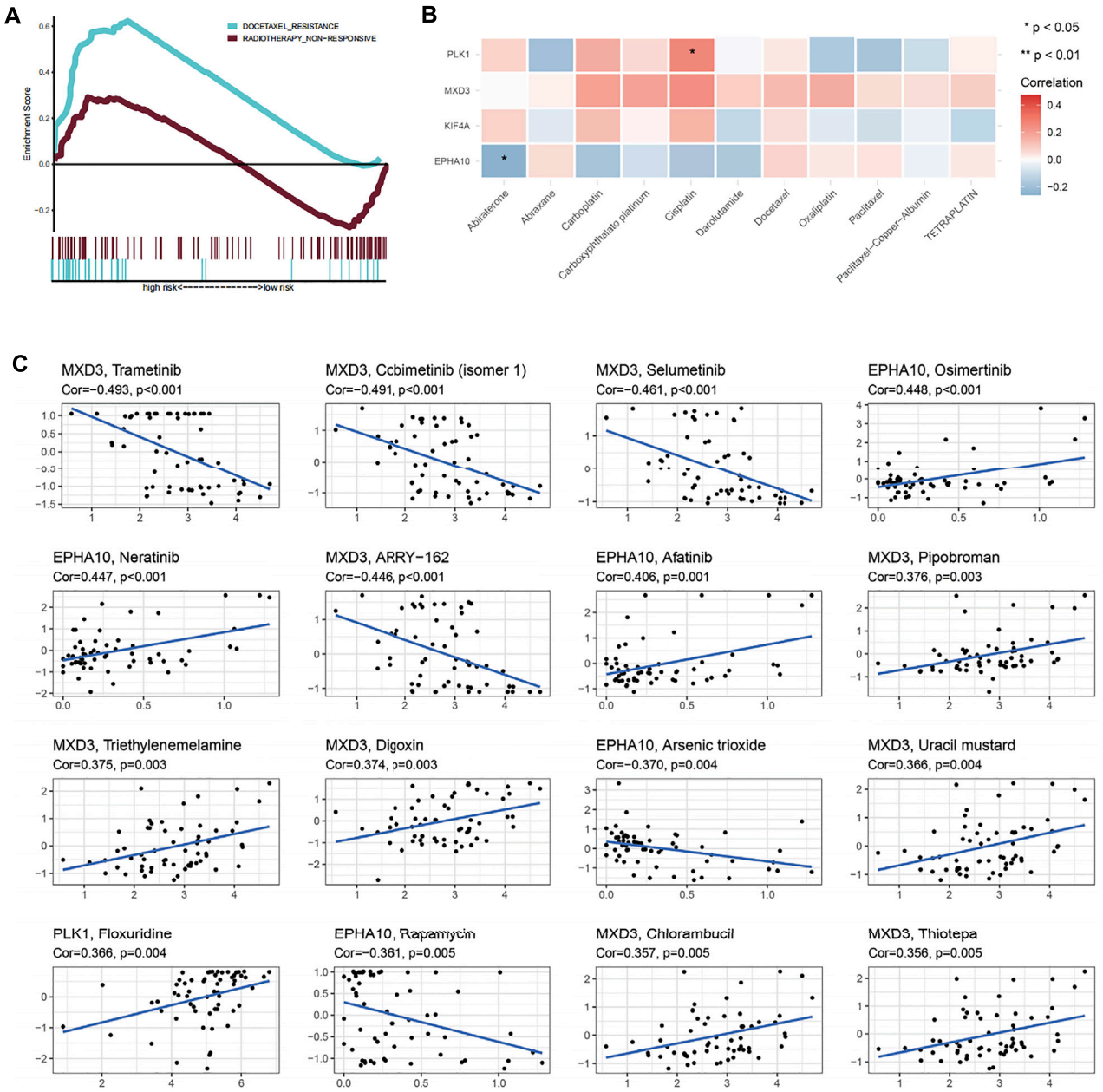



**FIGURE 8 |** Therapeutic response and drug sensitivity analysis. **(A)** Patients in the high-risk group showed significant docetaxel resistance. **(B)** Sensitivity analysis of genes in "E2F target"-related signature and common drugs of prostate cancer. **(C)** Drug sensitivity analysis showed that the top 16 drugs in the CellMiner database, which were sensitive to genes in "E2F target"-related signature.

related to one another and are expected to form a gene signature to predict patient prognosis (Kent and Leone, 2019). Therefore, E2F target genes play an important role in cancer development.

In this study, bioinformatics analysis was used to find four E2F target-related genes as a novel prognosis gene signature to guide further clinical decision-making. The four gene signatures (*MDX3*, *PLK1*, *EPHA10*, and *KIF4A*) showed a strong correlation with prostate cancer prognosis in cases selected from TCGA database. These genes may predict the prognosis of prostate cancer more accurately than existing signatures. We also verified the enrichment of the E2F target pathway and the worse prognosis in high-risk groups from two separate cohorts of the ICGC and GEO databases. In addition, exploring the functional differences among different risk groups showed that in high-risk groups, the binding process of a variety of proteins and molecules was abnormally activated, and the properties of stem cells increased. Simultaneously, T regulatory cells and immunosuppressive genes were highly expressed. These may lead to abnormal proliferation, apoptosis, extracellular matrix remodeling, and immune escape of tumor cells. It is worth noting

that severe inhibition of histidine metabolism was observed in the high-risk group. Some studies have found that accelerating histidine metabolism may improve the therapeutic effect of anticancer drugs (Kanarek et al., 2018). Therefore, the activation of the E2F target pathway may cause resistance to the treatment by inhibiting the histidine metabolism. Accordingly, our drug analysis demonstrated the low sensitivity of drugs and hinted at the potential of *MXD3* and *EPHA10* as therapeutic targets.

Recent studies show that PLK1 and KIF4A as biomarkers have a high prognosis value in patients with prostate cancer (Gao et al., 2018; Cao et al., 2020; Liu et al., 2020; Wu et al., 2020; Das et al., 2021). PLK1 has been proven to be a potent and promising target for prostate cancer treatment (Mao et al., 2018; Shin et al., 2019). MXD3 is a transcription target of E2F1 (Yun et al., 2007) and belongs to the MYC/MAX/MAD network, which can compete with MYC to regulate the cell cycle and proliferation (Ayer et al., 1993; Grandori et al., 2000). MXD3 has been shown to predict poor prognosis in clear cell renal cell carcinoma (Zhang et al., 2021) and hepatocellular carcinoma (Xu et al., 2019), as well as having been indicated as a new molecular targeted site to treat neuroblastoma (Yoshida et al., 2020). *PLK1* (Iliaki et al., 2021), *EPHK10*, and *KIF4A* also play important roles as prognostic indicators or as targeted therapy sites for the progression of multiple tumors, such as pancreatic cancer (Zhu et al., 2021), esophageal squamous cell carcinoma (Li et al., 2021), bladder cancer (Zheng et al., 2021), and lung cancer (Kirienko et al., 2021). Furthermore, EPHA10 has already been approved as a potential therapeutic target of prostate cancer (Nagano et al., 2014), and high *EPHA10* expression correlated with lymph node metastasis of breast cancer (Nagano et al., 2013)—the type of malignant behavior that usually predicts a poor prognosis.

Our study is not without limitations. First, although the study used advanced bioanalysis algorithms for analysis, the study was based on the analysis of mRNA levels, and no experiments were performed to validate the effect of gene expression on phenotype. Second, due to the complexity of prostate cancer treatment and defects in the use of public databases, different patients in the same cohort may have received different treatments, which may have a certain impact on gene expression and cause bias in the results. Moreover, because prostate cancer is inert cancer, the study chose different prognostic endpoints in different cohorts, which may affect the comparison of risk values. To address these problems, it is necessary to establish prospective cohorts and perform in-depth experiments to elucidate potential pathways and mechanisms.

Nevertheless, this research introduces many innovations and commendable points. This study established the first E2F target-related gene signature to predict the prognosis of prostate cancer. It used cohorts from TCGA and ICGC databases with the same treatments, minimizing cohort heterogeneity to ensure the reliability of the analysis results. In addition, a cohort of radiotherapy plus androgen deprivation therapy from the GEO database was used to further verify the ability of the gene signature to predict the prognosis of patients at different stages of treatment. Consequently, further functional analysis and immune microenvironment analysis provided a preliminary explanation for the tumor-promoting effect of the E2F target pathway and revealed potential therapeutic target genes and sensitive drugs.

## CONCLUSION

This research established the first gene signature related to the E2F target pathway to predict the prognosis of prostate cancer. Furthermore, possible explanations for how the activation of the E2F target pathway results in the occurrence and development of prostate cancer have been provided. The guidance offered by this study can be useful for the individualized detection and treatment of patients with prostate cancer and lay a theoretical basis for further research into the therapeutic potential of the E2F target pathway.

## 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 review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

HX conducted data search, download, collation, statistics, and analysis, as well as manuscript writing and revision. MW conducted statistics, analysis, and manuscript writing. XS, ZL, and QY helped to search the data. ML gave financial support to the research. XG and ML reviewed the manuscript.

## FUNDING

This research was supported by the Beijing Hospital Clinical Research 121 Project (BJ-2018-090) and Discipline Construction Project of Peking Union Medical College (201920202101).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.838654/full#supplementary-material

**Supplementary Figure S1 |** Pie chart displayed the z-score of MHC, EC, CP, and SC of IPS algorithm in the top 10 and the last 10 "E2F target" risk score of patients.

## REFERENCES


Ayer, D. E., Kretzner, L., and Eisenman, R. N. (1993). Mad: a Heterodimeric Partner for Max that Antagonizes Myc Transcriptional Activity. *Cell* 72 (2), 211–222. doi:10.1016/0092-8674(93)90661-9

Bookstein, R., Rio, P., Madreperla, S. A., Hong, F., Allred, C., Grizzle, W. E., et al. (1990). Promoter Deletion and Loss of Retinoblastoma Gene Expression in Human Prostate Carcinoma. *Proc. Natl. Acad. Sci. U.S.A.* 87 (19), 7762–7766. doi:10.1073/pnas.87.19.7762

Cao, Q., Song, Z., Ruan, H., Wang, C., Yang, X., Bao, L., et al. (2020). Targeting the KIF4A/AR Axis to Reverse Endocrine Therapy Resistance in Castration-Resistant Prostate Cancer. *Clin. Cancer Res.* 26 (6), 1516–1528. doi:10.1158/1078-0432.ccr-19-0396

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

Classon, M., and Harlow, E. (2002). The Retinoblastoma Tumour Suppressor in Development and Cancer. *Nat. Rev. Cancer* 2 (12), 910–917. doi:10.1038/nrc950

Culp, M. B., Soerjomataram, I., Efstathiou, J. A., Bray, F., and Jemal, A. (2020). Recent Global Patterns in Prostate Cancer Incidence and Mortality Rates. *Eur. Urol.* 77 (1), 38–52. doi:10.1016/j.eururo.2019.08.005

Dahl, E. S., Buj, R., Leon, K. E., Newell, J. M., Imamura, Y., Bitler, B. G., et al. (2019). Targeting IDH1 as a Prosenescent Therapy in High-Grade Serous Ovarian Cancer. *Mol. Cancer Res.* 17 (8), 1710–1720. doi:10.1158/1541-7786.mcr-18-1233

Das, R., Sjöström, M., Shrestha, R., Yogodzinski, C., Egusa, E. A., Chesner, L. N., et al. (2021). An Integrated Functional and Clinical Genomics Approach Reveals Genes Driving Aggressive Metastatic Prostate Cancer. *Nat. Commun.* 12 (1), 4601. doi:10.1038/s41467-021-24919-7

Fang, Z., Lin, M., Li, C., Liu, H., and Gong, C. (2020). A Comprehensive Review of the Roles of E2F1 in colon Cancer. *Am. J. Cancer Res.* 10 (3), 757–768.

Fischer, M., and Müller, G. A. (2017). Cell Cycle Transcription Control: DREAM/MuvB and RB-E2f Complexes. *Crit. Rev. Biochem. Mol. Biol.* 52 (6), 638–662. doi:10.1080/10409238.2017.1360836

Gao, H., Chen, X., Cai, Q., Shang, Z., and Niu, Y. (2018). Increased KIF4A Expression Is a Potential Prognostic Factor in Prostate Cancer. *Oncol. Lett.* 15 (5), 7941–7947. doi:10.3892/ol.2018.8322

Grandori, C., Cowley, S. M., James, L. P., and Eisenman, R. N. (2000). The Myc/Max/Mad Network and the Transcriptional Control of Cell Behavior. *Annu. Rev. Cel Dev. Biol.* 16, 653–699. doi:10.1146/annurev.cellbio.16.1.653

Hallstrom, T. C., and Nevins, J. R. (2009). Balancing the Decision of Cell Proliferation and Cell Fate. *Cell Cycle* 8 (4), 532–535. doi:10.4161/cc.8.4.7609

Iliaki, S., Beyaert, R., and Afonina, I. S. (2021). Polo-like Kinase 1 (PLK1) Signaling in Cancer and beyond. *Biochem. Pharmacol.* 193, 114747. doi:10.1016/j.bcp.2021.114747

Jarrard, D. F., Modder, J., Fadden, P., Fu, V., Sebree, L., Heisey, D., et al. (2002). Alterations in the p16/pRb Cell Cycle Checkpoint Occur Commonly in Primary and Metastatic Human Prostate Cancer. *Cancer Lett.* 185 (2), 191–199. doi:10.1016/s0304-3835(02)00282-3

Kanarek, N., Keys, H. R., Cantor, J. R., Lewis, C. A., Chan, S. H., Kunchok, T., et al. (2018). Histidine Catabolism Is a Major Determinant of Methotrexate Sensitivity. *Nature* 559, 632–636. doi:10.1038/s41586-018-0316-7

Kent, L. N., and Leone, G. (2019). The Broken Cycle: E2F Dysfunction in Cancer. *Nat. Rev. Cancer* 19 (6), 326–338. doi:10.1038/s41568-019-0143-7

Kirienko, M., Sollini, M., Corbetta, M., Voulaz, E., Gozzi, N., Interlenghi, M., et al. (2021). Radiomics and Gene Expression Profile to Characterise the Disease and Predict Outcome in Patients with Lung Cancer. *Eur. J. Nucl. Med. Mol. Imaging* 48, 3643. doi:10.1007/s00259-021-05371-7

Langfelder, P., and Horvath, S. (2008). WGCNA: an R Package for Weighted Correlation Network Analysis. *BMC bioinformatics* 9, 559. doi:10.1186/1471-2105-9-559

Lee, E., Chuang, H.-Y., Kim, J.-W., Ideker, T., and Lee, D. (2008). Inferring Pathway Activity toward Precise Disease Classification. *Plos Comput. Biol.* 4, e1000217. doi:10.1371/journal.pcbi.1000217

Li, Y., Zhu, X., Yang, M., Wang, Y., Li, J., Fang, J., et al. (2021). YAP/TEAD4-induced KIF4A Contributes to the Progression and Worse Prognosis of Esophageal Squamous Cell Carcinoma. *Mol. carcinogenesis* 60 (7), 440–454. doi:10.1002/mc.23303

Liu, Z., Zhong, J., Cai, C., Lu, J., Wu, W., and Zeng, G. (2020). Immune-related Biomarker Risk Score Predicts Prognosis in Prostate Cancer. *Aging (Albany NY)* 12 (22), 22776–22793. doi:10.18632/aging.103921

Mao, F., Li, J., Luo, Q., Wang, R., Kong, Y., Carlock, C., et al. (2018). Plk1 Inhibition Enhances the Efficacy of BET Epigenetic Reader Blockade in Castration-Resistant Prostate Cancer. *Mol. Cancer Ther.* 17 (7), 1554–1565. doi:10.1158/1535-7163.mct-17-0945

Molenaar, J. J., Koster, J., Ebus, M. E., van Sluis, P., Westerhout, E. M., de Preter, K., et al. (2012). Copy Number Defects of G1-Cell Cycle Genes in Neuroblastoma Are Frequent and Correlate with High Expression of E2F Target Genes and a Poor Prognosis. *Genes Chromosom. Cancer* 51 (1), 10–19. doi:10.1002/gcc.20926

Nagano, K., Kanasaki, S. i., Yamashita, T., Maeda, Y., Inoue, M., Higashisaka, K., et al. (2013). Expression of Eph Receptor A10 Is Correlated with Lymph Node Metastasis and Stage Progression in Breast Cancer Patients. *Cancer Med.* 2 (6), 972–977. doi:10.1002/cam4.156

Nagano, K., Yamashita, T., Inoue, M., Higashisaka, K., Yoshioka, Y., Abe, Y., et al. (2014). Eph Receptor A10 Has a Potential as a Target for a Prostate Cancer Therapy. *Biochem. biophysical Res. Commun.* 450 (1), 545–549. doi:10.1016/j.bbrc.2014.06.007

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

Oshi, M., Angarita, F. A., Tokumaru, Y., Yan, L., Matsuyama, R., Endo, I., et al. (2021). A Novel Three-Gene Score as a Predictive Biomarker for Pathologically Complete Response after Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. *Cancers* 13 (10), 2401. doi:10.3390/cancers13102401

Pommier, Y., Reinhold, W. C., Sunshine, M., Varma, S. H., Kohn, K. W., and Doroshow, J. H. (2012). 268 CellMiner: a Web-Based Suite of Genomic and Pharmacologic Tools to Explore Transcript and Drug Patterns in the NCI-60 Cell Line Set. *Eur. J. Cancer* 48, 82. doi:10.1016/s0959-8049(12)72066-x

Reich, D., Price, A. L., and Patterson, N. (2008). Principal Component Analysis of Genetic Data. *Nat. Genet.* 40, 491–492. doi:10.1038/ng0508-491

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007

Shin, S. B., Woo, S. U., and Yim, H. (2019). Cotargeting Plk1 and Androgen Receptor Enhances the Therapeutic Sensitivity of Paclitaxel-Resistant Prostate Cancer. *Ther. Adv. Med. Oncol.* 11, 1758835919846375. doi:10.1177/1758835919846375

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

Sun, A., Bagella, L., Tutton, S., Romano, G., and Giordano, A. (2007). From G0 to S Phase: a View of the Roles Played by the Retinoblastoma (Rb) Family Members in the Rb-E2f Pathway. *J. Cel. Biochem.* 102 (6), 1400–1404. doi:10.1002/jcb.21609

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The Immune Landscape of Cancer. *Immunity* 48, 812–e14. doi:10.1016/j.immuni.2018.03.023

Tibshirani, R. (1997). The Lasso Method for Variable Selection in the Cox Model. *Statist. Med.* 16, 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3

Van den Heuvel, S., and Dyson, N. J. (2008). Conserved Functions of the pRB and E2F Families. *Nat. Rev. Mol. Cel Biol* 9 (9), 713–724. doi:10.1038/nrm2469

van der Maaten, L. J. P., and Hinton, G. E. (2008). Visualizing High-Dimensional Data Using T-SNE. *J. Machine Learn. Res.* 9 (Nov), 2579–2605.

Wang, D., Tang, W., Zhang, P., Liu, Z., Lyu, F., Xiao, Y., et al. (2021). Comprehensive Analysis of the Functional and Prognostic Value of E2F Transcription Factors in Human Prostate Cancer through Data Mining and Experimental Validation. *Translational Cancer Res.* 10, 5095. doi:10.21037/tcr-21-1532

Wu, T., and Wu, L. (2021). The Role and Clinical Implications of the Retinoblastoma (RB)-E2F Pathway in Gastric Cancer. *Front. Oncol.* 11, 655630. doi:10.3389/fonc.2021.655630

Wu, X., Lv, D., Eftekhar, M., Khan, A., Cai, C., Zhao, Z., et al. (2020). A New Risk Stratification System of Prostate Cancer to Identify High-Risk Biochemical Recurrence Patients. *Transl Androl. Urol.* 9 (6), 2572–2586. doi:10.21037/tau-20-1019

Xie, D., Pei, Q., Li, J., Wan, X., and Ye, T. (2021). Emerging Role of E2F Family in Cancer Stem Cells. *Front. Oncol.* 11, 723137. doi:10.3389/fonc.2021.723137

Xu, B., Lv, W., Li, X., Zhang, L., and Lin, J. (2019). Prognostic Genes of Hepatocellular Carcinoma Based on Gene Coexpression Network Analysis. *J. Cell. Biochem.* doi:10.1002/jcb.28441

Yoshida, S., Duong, C., Oestergaard, M., Fazio, M., Chen, C., Peralta, R., et al. (2020). MXD3 Antisense Oligonucleotide with Superparamagnetic Iron Oxide Nanoparticles: A New Targeted Approach for Neuroblastoma. *Nanomedicine: Nanotechnology, Biol. Med.* 24, 102127. doi:10.1016/j.nano.2019.102127

Yun, J.-S., Rust, J. M., Ishimaru, T., and Di´az, E. (2007). A Novel Role of the Mad Family Member Mad3 in Cerebellar Granule Neuron Precursor Proliferation. *Mol. Cel Biol* 27 (23), 8178–8189. doi:10.1128/mcb.00656-06

Zhang, F., Liu, L., Wu, P., Li, S., and Wei, D. (2021). Overexpression of MAX Dimerization Protein 3 (MXD3) Predicts Poor Prognosis in clear Cell Renal Cell Carcinoma. *Transl Androl. Urol.* 10 (2), 785–796. doi:10.21037/tau-20-1187

Zheng, P., Wu, K., Gao, Z., Li, H., Li, W., Wang, X., et al. (2021). KIF4A Promotes the Development of Bladder Cancer by Transcriptionally Activating the Expression of CDCA3. *Int. J. Mol. Med.* 47 (6). doi:10.3892/ijmm.2021.4932

Zhu, H., Li, Q., Zhao, Y., Peng, H., Guo, L., Zhu, J., et al. (2021). Vaccinia-related Kinase 2 Drives Pancreatic Cancer Progression by Protecting Plk1 from Chfr-Mediated Degradation. *Oncogene* 40 (28), 4663–4674. doi:10.1038/s41388-021-01893-4

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

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

## GLOSSARY

**AUC** area under the curve
**BCR** biochemical recurrence
**DEGs** differentially expressed genes
**E2F** adenoviral early region 2 binding factor
**EPHA10** EPH receptor A10
**FC** fold change
**FDR** false discovery rate
**FPKM** fragments per kilobase million
**GEO** Gene Expression Omnibus
**GO** Gene Ontology
**GS** Gleason score
**ICGC** International Cancer Genome Consortium
**IPS** immunophenoscore
**KEGG** Kyoto Encyclopedia of Genes and Genomes
**KIF4A** kinesin family member 4A
**LASSO** least absolute shrinkage and selection operator
**Met** metastasis
**MSigDB** Molecular Signatures Database
**MXD3** MAX dimerization protein 3
**OS** overall survival
**PCA** principal components analysis
**PLK1** polo-like kinase 1
**PRAD** prostate adenocarcinoma
**RB** retinoblastoma tumor suppressor protein
**RNA-seq** RNA-sequencing
**ROC** receiver operating characteristic
**SCs** suppressor cells
**ssGSEA** single-sample gene set enrichment analysis
**TCGA** The Cancer Genome Atlas
**TOM** topological overlap matrix
**T-ROC** time-dependent ROC
**tSNE** t-distributed stochastic neighbor embedding
**WGCNA** weighted gene co-expression network analysis.

frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 14 April 2022
doi: 10.3389/fmolb.2022.806727

# A Qualitative Signature to Identify *TERT* Promoter Mutant High-Risk Tumors in Low-Grade Gliomas

*Weicheng Zheng*[1,2†], *Ruolan Zhang*[2†], *Ziru Huang*[2†], *Jianpeng Li*[2], *Haonan Wu*[2], *Yuwei Zhou*[2], *Jinwei Zhu*[2] *and Xianlong Wang*[2*]

[1]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China, [2]Department of Bioinformatics, School of Medical Technology and Engineering, Key Laboratory of Medical Bioinformatics, Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fuzhou, China

**Background:** Telomerase reverse transcriptase promoter (*TERT*-p) mutation has been frequently found, but associated with contrary prognosis, in both low-grade gliomas and glioblastomas. For the low-grade gliomas (Grades II-III), *TERT*-p mutant patients have a better prognosis than the wildtype patients, whereas for the GBMs (Grade IV), *TERT*-p mutation is related to a poor prognosis. We hypothesize that there exist high-risk patients in LGGs who share GBM-like molecular features, including *TERT*-p mutation, and need more intensive treatment than other LGGs. A molecular signature is needed to identify these high-risk patients for an accurate and timely treatment.

**Methods:** Using the within-sample relative expression orderings of gene pairs, we identified the gene pairs with significantly stable REOs, respectively, in both the *TERT*-p mutant LGGs and GBMs but with opposite directions in the two groups. These reversely stable gene pairs were used as the molecular signature to stratify the LGGs into high-risk and low-risk groups.

**Results:** A signature consisting of 21 gene pairs was developed, which can classify LGGs into two groups with significantly different overall survival. The high-risk group has a similar genetic mutation profile and a similar survival profile as GBMs, and these high-risk tumors may progress to a more malignant state.

**Conclusion:** The 21 gene-pair signature based on REOs is capable of identifying high-risk patients in LGGs and guiding the clinical choice for appropriate and timely intervention.

**Keywords: glioma, TERT promoter, biomarker, transcriptome, relative expression orderings**

OPEN ACCESS

***Edited by:***
*Umberto Malapelle,*
*University of Naples Federico II, Italy*

***Reviewed by:***
*Pankaj Pathak,*
*National Institutes of Health (NIH), United States*
*Pasquale Pisapia,*
*University of Naples Federico II, Italy*

***Correspondence:**
*Xianlong Wang*
*wang.xianlong@139.com*

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

***Specialty section:***
*This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences*

***Received:*** *01 November 2021*
***Accepted:*** *07 March 2022*
***Published:*** *14 April 2022*

***Citation:***
*Zheng W, Zhang R, Huang Z, Li J, Wu H, Zhou Y, Zhu J and Wang X (2022) A Qualitative Signature to Identify TERT Promoter Mutant High-Risk Tumors in Low-Grade Gliomas. Front. Mol. Biosci. 9:806727. doi: 10.3389/fmolb.2022.806727*

## INTRODUCTION

Primary tumors account for more than half of central nervous system tumors, of which gliomas are the most frequent, accounting for 45–50% of all primary malignant brain tumors (Williams et al., 2018; Ostrom et al., 2019). Based on histopathological criteria, gliomas are classified into low-grade gliomas (LGGs) and glioblastoma (GBMs, Grade IV). The LGGs consist of oligodendroglioma (OD, Grades II-III) and astrocytoma (A, Grades II-III) (Louis et al., 2016). A frustrating reality is that not much progress has been made in prediction and treatment of gliomas for a long time (Filbin and Suva, 2016). Gliomas have a characteristic diffusely infiltrative pattern of growth in brain, and it is impossible to achieve complete resection (Cancer Genome Atlas Research et al., 2015). Prediction of

the clinical outcome is often inaccurate because current standards are often subjective and dependent on pathologists' experience, and it is almost impossible to distinguish the mixed histological appearance of glioma tissues (Cancer Genome Atlas Research et al., 2015). High interobserver variability results in high risk of misclassification, which has documented serious clinical consequences and may affect the choice of treatment options and prognosis of patients (Foote et al., 2015).

In recent years, various molecular markers have been discovered for gliomas, including isocitrate dehydrogenase 1 and 2 (*IDH 1/2*) mutations (Lu et al., 2012), codeletion of chromosome arms 1p and 19q (1p19q co-del) (Cairncross et al., 2013; van den Bent et al., 2013; Buckner et al., 2017) and telomerase reverse transcriptase promoter (*TERT*-p) hotspot mutations (Chan et al., 2015; Foote et al., 2015; Heidenreich et al., 2015; Masui et al., 2018). The 2016 World Health Organization (WHO) Classification of Tumors of the Central Nervous System (CNS) integrated *IDH* and 1p/19q co-del into clinical diagnosis (Louis et al., 2016). These molecular diagnostic markers are challenging our prior assumptions concerning the definitions of gliomas and becoming one of the most important factors for glioma prognosis (Buckner et al., 2017).

TERT is an enzyme that maintains the length of telomeres (O'Sullivan and Karlseder, 2010; Heidenreich and Kumar, 2017). Approximately 83% of primary GBMs (Killela et al., 2013) and 79.3% of OD (Killela et al., 2014) harbor a mutation of C to T in a mutually exclusive manner in the promoter of *TERT* at either -124 or -146 bp upstream from the transcription start site (Heidenreich et al., 2015). The mutations result in an additional ETS (E26 transformation-specific family transcription factor) binding site recognized by GABPA (GA-binding protein A), which facilitates reactivation of telomerases (Bollam et al., 2018). Previous research found that mutations of *TERT*-p lead to an increase in transcriptional activity by 2–4 times (Heidenreich et al., 2015).

The *TERT*-p mutations mainly occur in tumors from tissues with low rates of self-renewal (Heidenreich and Kumar, 2017). While being first discovered in melanoma, the *TERT*-p mutations have been found defining subsets of patients with adverse disease outcomes in hepatocellular carcinoma, urothelial carcinoma and other tumors (Heidenreich and Kumar, 2017; Heidenreich and Kumar, 2017). The prognostic impact of *TERT*-p mutation in diffuse gliomas appears to be ambivalent (Arita et al., 2016). In LGGs, *TERT*-p mutation is a hallmark for better prognosis, whereas in GBM it tends to be associated with poor prognosis (Reitman et al., 2013; Simon et al., 2015). Therefore, the consequence of *TERT*-p mutation and its interaction with other molecular markers in glioma pathogenesis remain to be fully understood (Bollam et al., 2018). High-throughput genomic sequencing is often used for the detection of *TERT*-p mutations, but the sensitivity is affected by the proportion of tumor cells and sequencing depth.

Treatment of glioma patients is critical to the prognosis. Gliomas of different grades or specific molecular states should receive corresponding treatments. For example, LGGs especially OD are known to be sensitive to procarbazine, lomustine and vincristine (PCV)-based chemotherapy. The response rate of LGG to PCV as an initial therapy ranges from 52 to 100% (Buckner et al., 2017). GBMs are generally treated with temozolomide (TMZ) after resection (Hegi et al., 2005). While *TERT*-p mutations mainly occur in OD and GBM, accurate classification to distinguish them is critical for choosing the right treatment. Misclassified patients are at risk of either over-treatment or under-treatment, which is detrimental to clinical decision-making and patient survival. Under-treatment is one of the most serious problems in cancer treatment and directly related to the survival. There is already ample evidence that some gliomas are misclassified as LGGs. They have malignant molecular alterations, but not yet exhibit histopathological characteristics (Bruner et al., 1997; Olar and Sulman, 2015). These gliomas usually have a strong tendency to develop into more malignant states, but indistinguishable due to their mixed histological appearance. Conventional histopathological diagnosis may suffer from morphological ambiguity and interobserver discordance (Arita et al., 2016). It is necessary to develop a robust and objective molecular marker to identify them.

In this study, we aim to develop a robust molecular marker by comparing the transcriptional profiles of LGGs and GBMs, in order to identify patients with high-grade characteristics (GBM-like patients) from LGGs. For these patients, a more aggressive treatment is recommended.

# MATERIALS AND METHODS

## Data Sources and Data Preprocessing

The gene expression profiles used in this study were downloaded from various databases. Datasets GSE61374, GSE16011, GSE43388 and GSE68848 were downloaded from the Gene Expression Omnibus (Barrett et al., 2007) (GEO, http://www.ncbi.nlm.nih.gov/geo/). Datasets TCGA-LGG and TCGA-GBM were downloaded from The Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov/) (Hoadley et al., 2018). Dataset E-MTAB-3892 was downloaded from ArrayExpress (http://www.ebi.ac.uk/arrayexpress/). Dataset CGGA was downloaded from the Chinese Glioma Genome Atlas (CGGA, http://www.cgga.org.cn/) (Bao et al., 2014; Zhao et al., 2017; Hu et al., 2018). The GBM-cohort include 301 sample with *TERT* mutant status (Arita et al., 2016). Clinical information for each dataset is shown in **Table 1** and **Supplementary Table S1**. The TCGA dataset was used as the training set, which include 130 grades II and III samples and 23 grade IV samples. The *TERT*-p mutation status of grade IV samples is unknown. In order to further improve the transferability of the classification signature on different profiling platforms, two GEO datasets, GSE16011 ($n = 159$) and GSE61374 ($n = 133$, low-grade samples only) were also used for training.

## Selection of Differentially Expressed Genes

Differentially expressed genes (DEGs) were selected between the *TERT*-p mutant low-grade gliomas (LGGs) and *TERT*-p mutant GBM (high-grade gliomas) samples using the following procedure. Genes in each sample were ranked by their expression levels while excluding those with a median value of 0 in a training set. The median rank of each gene was then calculated in the two groups of the training set and top 200 genes were selected with the largest difference in the median rank

**TABLE 1 |** Clinical characteristics of the training datasets.

| | TCGA | | GSE16011-61374 | |
|---|---|---|---|---|
| | ***TERT*-p mut (*n* = 149)** | ***TERT*-p wt (*n* = 158)** | ***TERT*-p mut (*n* = 204)** | ***TERT*-p wt (*n* = 88)** |
| Grade and Histological Type — no.(%) | | | | |
| Grade II | | | | |
| Oligodendroglioma | 43 (72.9) | 26 (31.7) | 3 (21.4) | 0 (0.0) |
| Oligoastrocyma | 13 (22) | 29 (35.4) | 6 (42.9) | 7 (15.6) |
| Astrocytoma | 3 (5.1) | 27 (32.9) | 5 (35.7) | 38 (84.4) |
| Grade III | | | | |
| Oligodendroglioma | 38 (53.5) | 9 (12.5) | 2 (6.5) | 1 (2.3) |
| Oligoastrocyma | 11 (15.5) | 20 (27.8) | 18 (58.0) | 12 (27.9) |
| Astrocytoma | 22 (31) | 43 (59.7) | 11 (35.5) | 30 (69.8) |
| Grade IV | | | | |
| Glioblastoma | 19 (-) | 4 (-) | 159 (-) | 0 (-) |
| Age at diagnosis — years | | | | |
| Mean | 50.14 ± 12.6 | 37.3 ± 12.0 | 53.0 ± 13.6 | 39.9 ± 12.4 |
| Range | 20–76 | 14–70 | 15–81 | 21–80 |
| Male sex — no./total (%) | 79/149 (53.0) | 93/158 (58.9) | 134/204 (65.7) | 56/88 (63.6) |
| *IDH* status — no. (%) | | | | |
| *IDH1/2*-mut | 93 (62.4) | 129 (98.0) | 68 (33.3) | 77 (87.5) |
| *IDH*-wt | 55 (36.9) | 19 (12.0) | 105 (51.5) | 1 (12.5) |
| 1p19q status — no. (%) | | | | |
| Co-del | 86 (57.7) | 2 (1.3) | - | - |
| No co-del | 63 (42.3) | 156 (98.7) | - | - |
| Median survival — months | 13.4 | 17.3 | 12.7 | 49.0 |

between the LGG group and the GBM group (100 genes up-regulated in the LGG group and 100 genes up-regulated in the GBM group). The common DEGs obtained from the two training sets (TCGA dataset and GEO dataset) were selected as the candidate gene set to develop the gene-pair signature.

## Selection of Significantly Stable Gene Pairs and Reversely Stable Gene Pairs

The relationship between a pair of genes is denotated as $G_1 > G_2$ if the expression level of $G_1$ is greater than that of $G_2$ in a sample. In a cohort of $n$ samples, the probability to observe this relative expression ordering (REO) (Li et al., 2016) in $m$ samples can be represented using the binomial distribution model,

$$p = 1 - \sum_{i=0}^{m-1} \binom{n}{i} P_0^i (1 - P_0)^{n-i}$$

where $P_0$ ($P_0$ = 50%) is the probability to observe the REO in a sample by chance. A gene pair with $p < 0.01$ was considered with a significantly stable REO.

A reversely stable gene pair refers to a gene pair which is significantly stable in both the GBM cohort and the LGG cohort but with the reversal REO directions, e.g., $G_1 > G_2$ in one cohort but $G_1 < G_2$ in the other.

## Consistency Assessment of Reversely Stable Gene Pairs

Candidate gene pairs with reversely stable REOs were identified in each train set and the overlapping pairs were considered as consistent reversal gene pairs. The binomial distribution model was used to evaluate the significance of consistency between the two groups of gene pairs.

## Classification Rules

Reclassification of a glioma sample into high-risk and low-risk groups is based on how many gene pairs in the signature showing the REO patterns of the GBM cohort using the majority voting rule. If more than 10 gene pairs among 21 pairs in a sample showing the REO patterns of GBM, the sample is classified as high-risk and also called as the GBM-like sample. Otherwise, it is reclassified as a low-risk sample.

## Survival Analysis

The Kaplan-Meier method was used to estimate the survival curves. The significance of the difference between two survival curves was tested using the log-rank method (Jones and Crowley, 1989). The Cox proportional-hazards model was used to evaluate the hazard ratios of the signature and other factors. All analyses were performed using Python 3.4.1 or R 3.5.1.

# RESULTS

## Ambivalent Prognostic Impact of *TERT*-P Mutation

In order to investigate the potential of *TERT* promoter mutation status as a prognostic factor, we compared the overall survival (OS) of *TERT*-p mutant and wildtype gliomas using E-MTAB-3892 (Kamoun et al., 2016) dataset and a GBM cohort from a previous study (Arita et al., 2016) (**Figure 1**). In LGGs, we observed a survival advantage of *TERT*-p mutant patients ($p$ = 6.4E-03, **Figure 1A**). On the contrary, in the GBM cohort, *TERT*-p

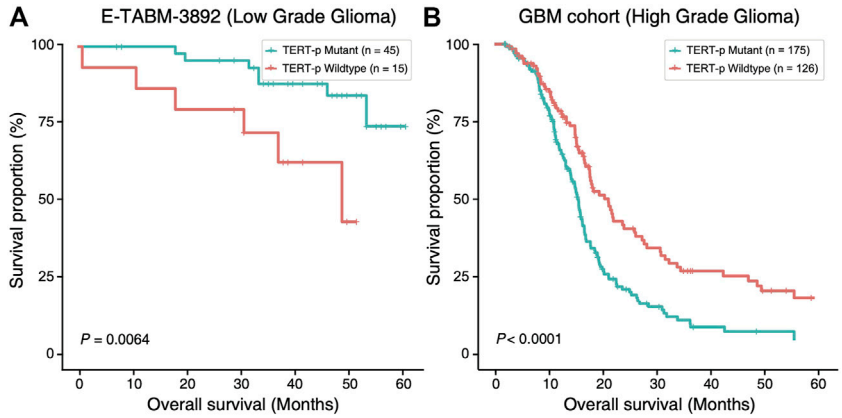



**FIGURE 1 |** Ambivalent prognostic impact of *TERT*-p mutation. In LGGs (E-TABM-3892 dataset), **(A)** *TERT*-p mutant samples have a significantly longer OS than the wildtype samples whereas in GBMs (GBM cohort dataset), **(B)** *TERT*-p mutant samples have a significantly shorter OS.

mutant patients showed a significantly shorter OS than the wildtype patients ($p$ < 1E-04, **Figure 1B**). The median OS of low-grade samples with *TERT*-p mutations in the TCGA training dataset is shorter than 20 months, which leads us to suspect that there exist GBM-like patients in the low-grade *TERT*-p mutant samples.

## Development of the Classification Signature Consisting of 21 Gene-Pairs

Top 200 DEGs with the largest differences in the median ranks between the LGGs and the GBMs were selected in the two training datasets, respectively. There exist 48 genes shared by the two training datasets with the same differential expression pattern. They were used to develop the classification signature based on REOs. From the GSE61374-GSE16011 dataset and the TCGA dataset, 559 and 485 reversal gene pairs were identified (adjusted $p$ = 0.01), respectively, and with an intersection of 483 gene pairs. The maximum matching algorithm in graph theory is applied to remove the redundancy in the 483 pairs. Finally, 21 pairs of disjoint gene pairs were obtained, which are denoted as 21-GPS. For each gene pair in 21-GPS, the REO patterns (Gene A > Gene B) in **Supplementary Table S2** are associated with a worse survival. If more than 10 gene pairs in a sample have these patterns, this sample was classified into the high-risk (GBM-like) group; otherwise, it was classified into the low-risk group. A scheme is given in **Figure 2** to show the classification process.

## Significant Survival Difference Found Between the Two Groups Classified by 21-GPS

The 21-GPS signature was first applied to the *TERT*-p mutant LGGs in both the TCGA training set and an independent validation dataset of CGGA (**Figure 3**). A group of high-risk samples was identified, respectively, in each dataset (32 in TCGA and 19 in CGGA), which even has a worse OS than the *TERT*-p wildtype samples. Furthermore, the OS is similar as that of the GBM samples. These samples should receive a more aggressive and early treatment.

## Survival Prediction Performance of 21-GPS for Samples Without *TERT*-P Status

In four datasets without *TERT*-p mutation status (**Supplementary Table S1**), 81, 53, 22 and 50 high-risk samples were identified, respectively, in CGGA, GSE68848, GSE43388 and GSE16011, using 21-GPS. All the high-risk samples have a significantly worse OS than the low-risk samples in each dataset (**Figure 4**). The median OS is shorter than 20 months in GSE68848, GSE43388 and GSE16011.

In the CGGA dataset, approximately 45% of LGGs were reclassified as high-risk samples. Among the high-risk samples, the proportion of *IDH* wildtype samples is significantly higher than that in the low-risk group (39.5 vs. 17.0%) and the proportion of grade III patients is also significantly higher (60.5 vs. 4.2%). Therefore, the high-risk group is enriched with grade III and *IDH* wildtype patients who have a greater survival risk.

## Higher Prognostic Value of 21-GPS Than *IDH* Mutation Status

Mutation statuses of *IDH* and *TERT* promoter have been used to classify gliomas in previous studies. Therefore, we sought to investigate the prognostic value of 21-GPS in comparison with the classification results based on the *IDH* and *TERT*-p mutation statuses using survival analysis. In the CGGA dataset of LGGs, a similar classification performance was observed between 21-GPS and *IDH* mutation status in the *TERT*-p mutant group ($p$ < 0.0001 vs. $p$ < 0.0001, log-rank test, **Figures 5A,B**). However, in the *TERT*-p wildtype samples, the 21-GPS classifier showed a much better classification performance than *IDH* mutation status ($p$ < 0.0001 vs. $p$ = 0.023, log-rank test, **Figures 5C,D**). In both the *TERT*-p mutant and wildtype groups, the 21-GPS classifier achieved higher hazard ratios than the *IDH* mutation status. Furthermore, we also carried out receiver operating characteristic (ROC) curve analysis to evaluate the classification performance of 21-GPS and *IDH* mutation status, using 3-year survival as the threshold to distinguish the reference high risk and low-risk groups (**Supplementary Figure S1**). The area under the ROC

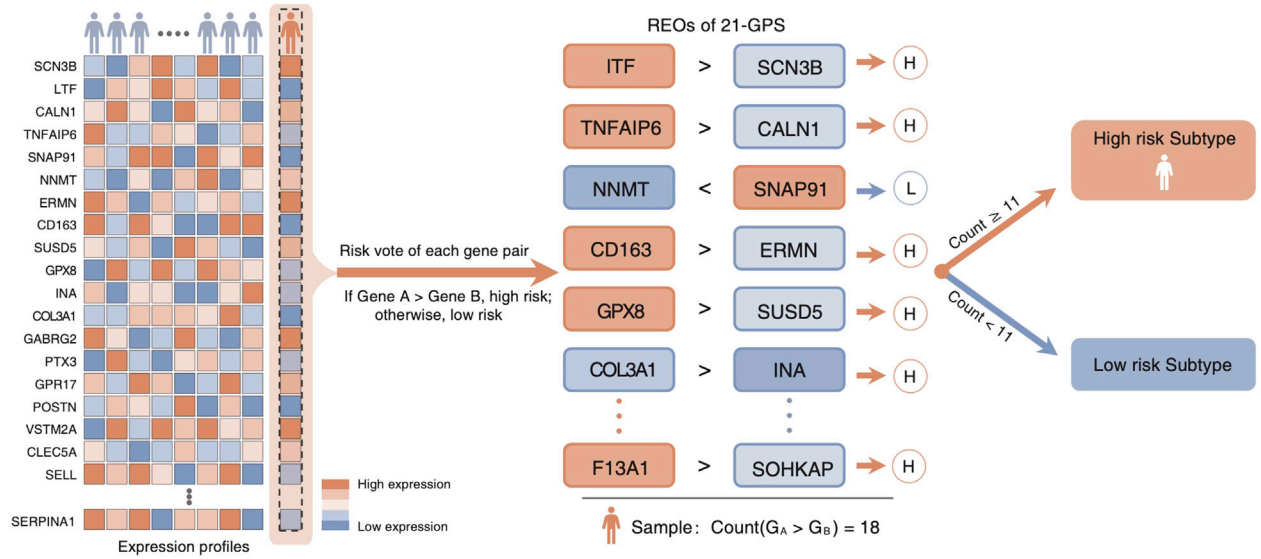



**FIGURE 2 |** Classification scheme of the 21-GPS signature. For a given sample, we evaluate the REO of 21 gene pairs of the signature. If more than or equal to 11 pairs have a REO of gene A > gene B, which is the high-risk pattern, the sample is identified as a high-risk sample.

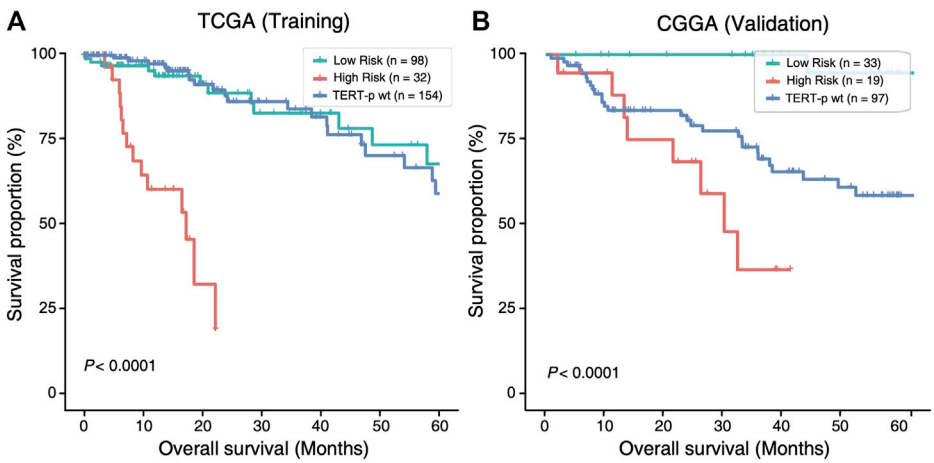



**FIGURE 3 |** The high-risk samples identified by 21-GPS have significantly worse prognosis than the low-risk samples in the *TERT*-p mutant samples of the training TCGA dataset **(A)** and the validation CGGA dataset **(B)**. The survival curves of *TERT*-p wildtype samples are also shown for comparison.

curve (AUC) is also higher for 21-GPS than for *IDH* mutation in both the *TERT*-p mutant group (0.955 vs. 0.886) and the *TERT*-p wildtype group (0.862 vs. 0.660). Thus, 21-GPS has a higher prognostic value than *IDH* mutation in stratifying LGGs.

## 21-GPS as an Independent Prognostic Factor Revealed by Cox Regression Analysis

Univariate and multivariate Cox regression models were used to evaluate the prognostic value of 21-GPS along common clinicopathological factors such as age, gender and *IDH* mutation status (**Supplementary Table S3**). In the TCGA and CGGA datasets, the Cox regression analysis were performed separately in the *TERT*-p mutant samples and wildtype samples. As the multivariate Cox regression result shows, the 21-GPS classification result is an independent prognostic factor in all five subsets, including both the training set and independent validation set (TCGA-LGG: HR = 11.33, $p < 0.0001$; CGGA: HR = 28.11, $p < 0.0001$; GSE16011-LGG: HR = 1.96, $p = 0.0017$; GSE68848-LGG: HR = 2.32, $p < 0.0001$; GSE43388-LGG: HR = 3.39, $p = 0.0009$). In the multivariate Cox regression result, the 21-GPS classification still shows an excellent prediction performance as an independent prognostic factor. The significance of 21-GPS in predicting OS is stronger than that of *IDH* mutation. In addition, the classification performance of 21-GPS is stable not only in the datasets obtained on

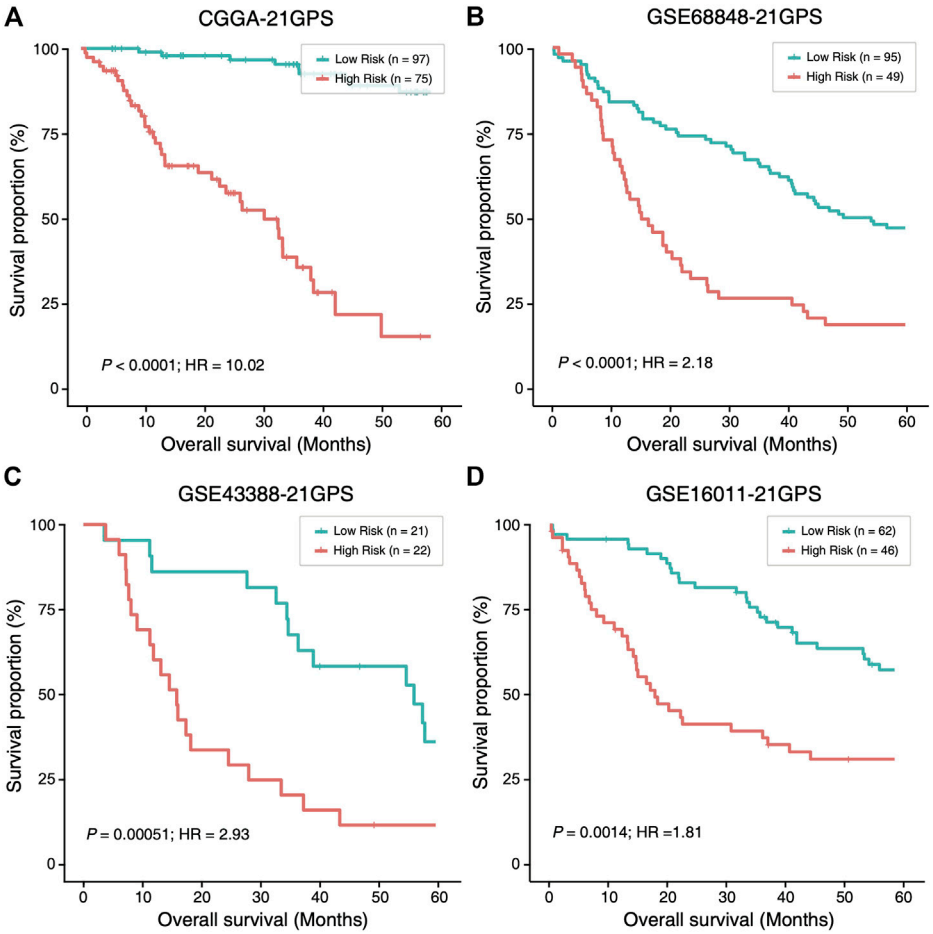



**FIGURE 4 |** The high-risk samples identified by 21-GPS also have significantly worse prognosis than the low-risk samples in four independent datasets without the *TERT*-p mutation information, CGGA **(A)**, GSE68848 **(B)**, GSE43388 **(C)** and GSE16011 **(D)**. Hazard ratio (HR) associated with IDH mutation is comparing the wildtype ones with respect to the mutant ones in this figure and **Figure 5**.

RNA-sequencing platform and but also in the datasets obtained on microarray platforms.

## GBM-like Mutation Landscape Found in High-Risk Samples Identified by 21-GPS

In order to further reveal the molecular characteristics of the high-risk group identified by 21-GPS, somatic mutational landscape is compared between the low-risk and high-risk samples in the TCGA dataset in which the somatic mutation information is available (**Figure 6**). In **Figure 6**, we selected the top 10 genes with mutation frequency greater than 3%, e.g., *IDH*, *TERT*-p and O-6-methylguanine-DNA methyltransferase promoter (*MGMT*-p) and some tumor-specific genes, e.g., tumor protein p53 (*TP53*) and phosphatase and tensin homolog (*PTEN*). The samples were grouped either by the original high- and low-grade labels (left two panels in **Figure 6B**) or by the combined *IDH* and *TERT*-p mutation statuses (right two panels in **Figure 6B**). For the latter classification method, only the *IDH* wildtype and *TERT*-p mutant subsets are displayed, which are expected to have the worst and best prognosis, respectively, among four possible subsets.

Among the low-risk samples, the proportion of *IDH1* and *IDH2* mutations, in a mutually exclusive state, is very high (91.6%). The ATRX chromatin remodeler gene (*ATRX*) and *TP53* show a tendency of co-mutation, similar as in low-grade astrocytoma. The proportions of malignant prognostic factors, e.g., mutations of *PTEN* and the epidermal growth factor receptor gene (*EGFR*), are very low. In short, the mutational landscape of the low-risk samples shows the classic molecular characteristics of LGGs.

In the high-risk group, we found that most of samples are *IDH* wildtype, and most of *IDH* mutant samples are accompanied with *ATRX* and *TP53* mutations or *TP53* mutations alone. The *TERT*-p mutation frequency is more than 80% in the high-risk group. *ATRX* mutations and *TERT*-p mutations show a mutually exclusive trend. In the WHO-2016 CNS classification system, primary GBM is characterized by *IDH* wildtype, and secondary GBM by *IDH* and *ATRX* mutations. If we only use the *IDH* and *TERT*-p mutation statuses to select the high-risk samples, a significant proportion of samples will be ignored, which can

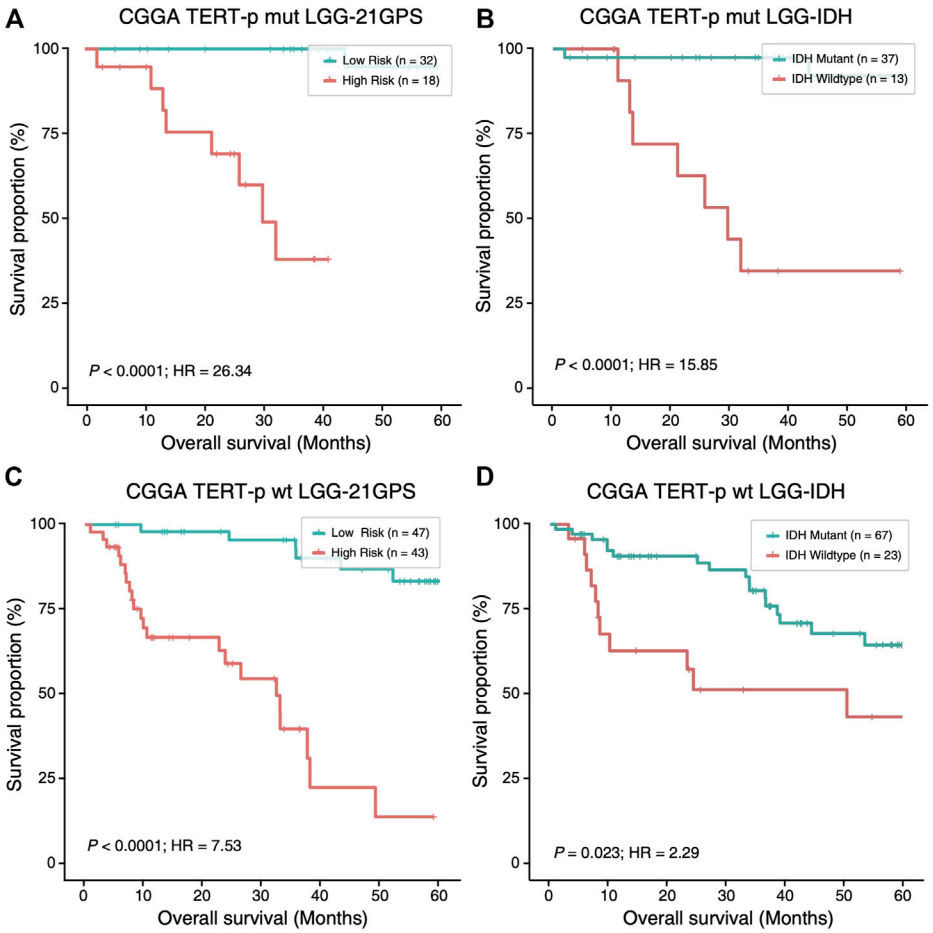



**FIGURE 5 |** Comparison of the survival curves of the 21-GPS classification with the *IDH*-based classification in the *TERT*-p mutant samples **(A,B)** and wildtype samples **(C,D)**. The survival of the high-risk samples is significantly shorter than the low-risk samples in both the *TERT*-p mutated **(A)** and wildtype **(C)** LGGs. The survival of the *IDH* wildtype samples is also shorter than the *IDH* mutant samples in the *TERT*-p mutant LGGs or in the *TERT*-p wildtype LGGs **(D)**. But the hazard ratios and the significance levels are higher in the 21-GPS classification than the *IDH* classification.

be identified by 21-GPS. Additionally, the landscape also revealed some interesting features in the high-risk samples. For example, *IDH* wildtype samples are frequently accompanied by *EGFR* mutations, and *PTEN* and *TP53* mutations tend to be mutually exclusive.

The mutation frequency plots of the high-risk group and the low-risk group (**Figure 6A**) are similar, respectively, with the high-grade and low-grade subsets. For example, in the high-grade or high-risk groups, the frequency of malignant mutations such as those of *PTEN*, *EGFR* and neurofibromin 1 (*NF1*), are higher while in the low-grade or low-risk groups the frequency of *IDH* mutations is higher. In the classification system based on the *IDH* and *TERT*-p mutation statuses, the samples with *ATRX* and *TP53* mutations are ignored, which represent a significant proportion of LGGs. *EGFR* mutations appear only in the *IDH* wildtype samples, but a proportion of *EGFR* mutated samples were found in the low-grade or low-risk groups. The prognostic value of these less-frequent mutations is ignored by the *IDH* and *TERT*-p mutation classification system, but not by our signature.

## DISCUSSION AND CONCLUSION

Diffuse glioma is a highly heterogeneous malignant tumor, which sees no much progress in its classification for more than 10 years (Filbin and Suva, 2016). With the advancement of high-throughput sequencing technologies, molecular markers have become important complements to histological features which make clinical classification of gliomas more accurate. However, we found that some low-grade gliomas have poor survival (less than 20 months), which may be related to inappropriate treatment. Previous study found that *TERT*-p mutation has a "bivalent" prognostic effect in diffuse gliomas (Arita et al., 2016). The low-grade gliomas especially oligodendrogliomas have the best prognosis and the mutation rate of *TERT*-p is close to 80%, while GBMs, which have the worst survival, also have a high mutation rate of more than 80% (Arita et al., 2013; Killela et al., 2014; Eckel-Passow et al., 2015). The mechanism of this opposite prognostic impact remains unexplained. One possible reason is due to high heterogeneity in gliomas and the classification based on the mutation status of a single-gene could lead wrong results.

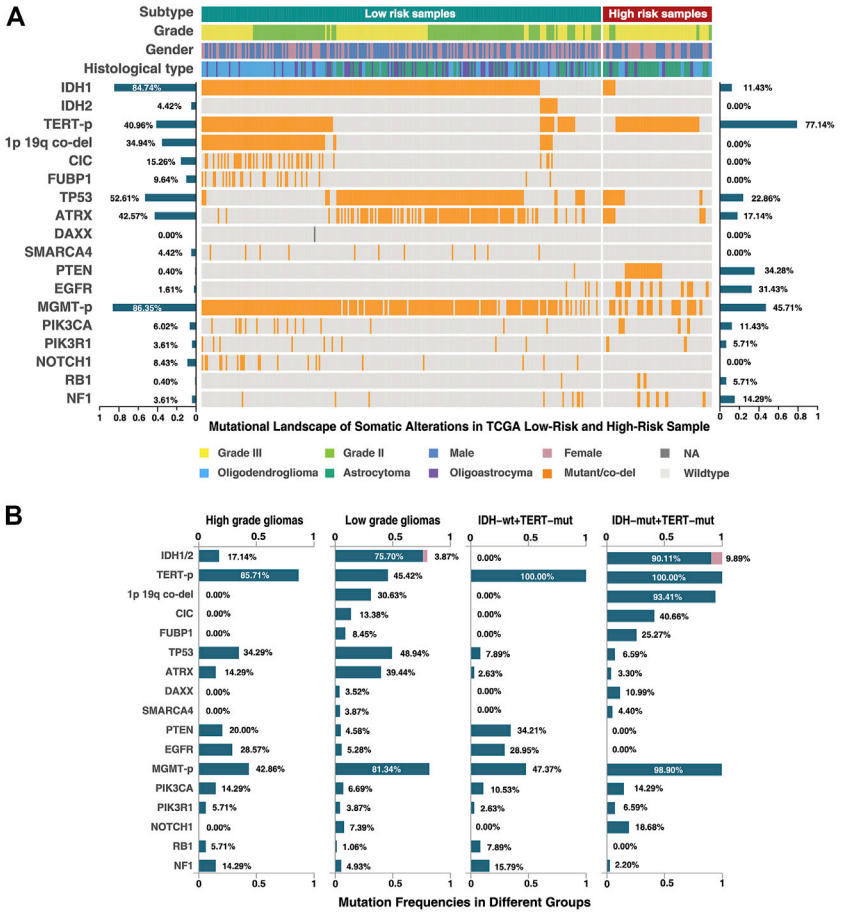



**FIGURE 6 |** Mutational landscape of the low-risk and high-risk samples of the TCGA dataset classified by 21-GPS. The mutation frequencies are shown on the left-side for the low-risk group and on the right-side for the high-risk group in panel **(A)**. The mutation frequencies are also shown for the high-grade (GBM) and low-grade gliomas (left-two panels in **(B)**) and for the *IDH* wildtype and mutant groups in the *TERT*-p mutant panels (right two panels in **(B)**).

For example, if a tumor is mixed with *IDH*-mutated oligodendrogliomas and *IDH*-wildtype GBM, it will be classified as *IDH*-mutated oligodendrogliomas because the high sensitivity of next-generation sequencing technique enables us to detect the mutation at a very low tumor cell purity. But GBMs usually progress at a much higher rate and the tumor should be treated as GBM instead of oligodendroglioma. Our signature is based on the relative expression orderings of 21 gene pairs and uses the principle of majority voting to classify a sample. It has more robust performance than the classification based on genetic mutation statuses against tumor heterogeneity and sampling uncertainty.

Mutant (mut) or wildtype (wt) statuses of *IDH* and *TERT*-p have been used to classify gliomas into four categories (both mut, both wt, *IDH* mut and *TERT*-p wt, and *IDH* wt and *TERT*-p mut) (Zhang et al., 2015; Arita et al., 2016; Akyerli et al., 2018), which were shown significantly associated with prognosis. But this simple molecular classification does not consider the influence of tumor grades and cannot explain the biological mechanism of these two molecules. We show that the classification efficiency of *IDH* is not significant in the *TERT*-p wt samples. But our signature can classify both *TERT*-p wt samples and the mutation-status unknown samples robustly.

Using our signature to analyze the TCGA cohort, we found that the proportion of grade III samples is significantly greater than that of grade II in the high-risk group (**Figure 6**) and the *IDH* mutation frequency is significantly lower (15.9%) in the high-risk group than in the low-risk group while the frequencies of other malignant mutations are significantly higher. Among the high-risk samples, we found no samples with co-mutations of *IDH* and *TERT*-p, and those *IDH* mutated and *TERT*-p wildtype samples are usually accompanied by *ATRX* and *TP53* mutations. This is consistent with our understanding on GBM. In the *IDH* mutated high-risk samples, most of them are co-mutated with *ATRX*, which represents a typical molecular state of recurrent glioblastoma. In *IDH* wildtype high-risk samples, most of them carry the mutations of *PTEN* and/or *EGFR*, which are the characteristics of primary GBMs. The result is consistent with the telomere replacement extension mechanism discussed in previous studies (Heaphy et al., 2011; Akter and Kamijo, 2021; Rachakonda et al., 2021; Valenzuela et al., 2021). This indicates that our signature is capable of identifying the high-risk samples with either primary or recurrent malignant GBM characteristics.

In the fifth edition of the World Health Organization (WHO) Classification of Tumors of the Central Nervous System (WHO CNS5), which was released recently (Louis et al., 2021; Wen and Packer, 2021), *IDH* wildtype diffuse and astrocytic gliomas are classified as "GBM (*IDH* wildtype subtype)" with the presence of *TERT*-p mutation or *EGFR* gene amplification. This change reflects that there exist high-risk samples similar to GBM in grade II/III gliomas, even if they do not show the histological characteristics of GBM. In **Figure 6**, 77% of the reclassified high-risk group have *TERT*-p mutations, and 31.43% have *EGFR* mutations (*EGFR* mutation rate is only 1.61% in the low-risk group). This result is consistent with the new classification of WHO CNS5. In addition, in the low-risk group, almost all samples with *TERT*-p mutations are oligodendrogliomas, which characterizes the "oligodendroglioma, *IDH* mutant, *TERT*-p mutant" group in CNS5. In the reclassified TCGA cohort, approximately 90% of the high-risk group are *IDH* wildtype, and most of them were labeled as astrocytoma. In the other 10% of samples with *IDH* mutation did not include oligodendrogliomas, and they are likely to be classified as "Astrocytoma, *IDH*-mutant, grade IV" in CNS5. Therefore, our signature can screen out high-risk samples with GBM-like characteristics from all glioma samples, which has high potential in clinical application.

In conclusion, we developed a molecular signature based on the relative expression orderings of 21 gene pairs which can identify high-risk samples from gliomas, without relying on WHO grade and *IDH* status. These high-risk tumors show GBM-like characteristics and should receive more aggressive treatment. The 21-GPS has great potential to be applied to clinical diagnosis in the future.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The datasets GSE16011, GSE61374, GSE68848, GSE43388 for this study can be found in the Gene Expression Omnibus, https://www.ncbi.nlm.nih.gov/geo/. The datasets TCGA for this study can be found in The Cancer Genome Atlas, https://cancergenome.nih.gov/. The datasets CGGA for this study can be found in the Chinese Glioma Genome Atlas, http://www.cgga.org.cn/. The datasets E-MTAB-3892 for this study can be found in ArrayExpress, https://www.ebi.ac.uk/arrayexpress/.

## AUTHOR CONTRIBUTIONS

WZ designed the experiments; RZ and ZH performed the analysis and drafted the manuscript; JL, HW, YZ and JZ contributed to data collection, analysis and results discussion; and X.W. supervised the project and revised the manuscript.

## FUNDING

XW acknowledges the financial support from Fujian Medical University (Grant No. XRCZX2017001) and the Natural Science Foundation of Fujian Province (Grant No. 2019J01294).

## ACKNOWLEDGMENTS

Professor Zhen Guo from Harbin Medical University supervised the project initially but passed away in 2019 unfortunately.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.806727/full#supplementary-material

## REFERENCES

Akter, J., and Kamijo, T. (2021). How Do Telomere Abnormalities Regulate the Biology of Neuroblastoma? *Biomolecules* 11 (8), 1112. doi:10.3390/biom11081112

Akyerli, C. B., Yüksel, Ş., Can, Ö., Erson-Omay, E. Z., Oktay, Y., Coşgun, E., et al. (2018). Use of Telomerase Promoter Mutations to Mark Specific Molecular Subsets with Reciprocal Clinical Behavior in IDH Mutant and IDH Wild-type Diffuse Gliomas. *J. Neurosurg.* 128 (4), 1102–1114. doi:10.3171/2016.11.JNS16973

Arita, H., Narita, Y., Fukushima, S., Tateishi, K., Matsushita, Y., Yoshida, A., et al. (2013). Upregulating Mutations in the TERT Promoter Commonly Occur in Adult Malignant Gliomas and Are Strongly Associated with Total 1p19q Loss. *Acta Neuropathol.* 126 (2), 267–276. doi:10.1007/s00401-013-1141-6

Arita, H., Yamasaki, K., Matsushita, Y., Nakamura, T., Shimokawa, A., Takami, H., et al. (2016). A Combination of TERT Promoter Mutation and MGMT Methylation Status Predicts Clinically Relevant Subgroups of Newly Diagnosed Glioblastomas. *Acta Neuropathol. Commun.* 4 (1), 79. doi:10.1186/s40478-016-0351-2

Bao, Z.-S., Chen, H.-M., Yang, M.-Y., Zhang, C.-B., Yu, K., Ye, W.-L., et al. (2014). RNA-seq of 272 Gliomas Revealed a Novel, Recurrent PTPRZ1-MET Fusion Transcript in Secondary Glioblastomas. *Genome Res.* 24 (11), 1765–1773. doi:10.1101/gr.165126.113

Barrett, T., Troup, D. B., Wilhite, S. E., Ledoux, P., Rudnev, D., Evangelista, C., et al. (2007). NCBI GEO: Mining Tens of Millions of Expression Profiles-Ddatabase and Tools Update. *Nucleic Acids Res.* 35, D760–D765. doi:10.1093/nar/gkl887

Bollam, S. R., Berens, M. E., and Dhruv, H. D. (2018). When the Ends Are Really the Beginnings: Targeting Telomerase for Treatment of GBM. *Curr. Neurol. Neurosci. Rep.* 18 (4), 15. doi:10.1007/s11910-018-0825-7

Bruner, J. M., Inouye, L., Fuller, G. N., and Langford, L. A. (1997). Diagnostic Discrepancies and Their Clinical Impact in a Neuropathology Referral Practice. *Cancer* 79 (4), 796–803. doi:10.1002/(sici)1097-0142(19970215)79:4<796::aid-cncr17>3.0.co;2-v

Buckner, J., Giannini, C., Eckel-Passow, J., Lachance, D., Parney, I., Laack, N., et al. (2017). Management of Diffuse Low-Grade Gliomas in Adults - Use of Molecular Diagnostics. *Nat. Rev. Neurol.* 13 (6), 340–351. doi:10.1038/nrneurol.2017.54

Cairncross, G., Wang, M., Shaw, E., Jenkins, R., Brachman, D., Buckner, J., et al. (2013). Phase III Trial of Chemoradiotherapy for Anaplastic Oligodendroglioma: Long-Term Results of RTOG 9402. *Jco* 31 (3), 337–343. doi:10.1200/JCO.2012.43.2674

Cancer Genome Atlas Research, N., Brat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., et al. (2015). Comprehensive, Integrative Genomic

Analysis of Diffuse Lower-Grade Gliomas. *N. Engl. J. Med.* 372 (26), 2481–2498. doi:10.1056/NEJMoa1402121

Chan, A. K.-Y., Yao, Y., Zhang, Z., Chung, N. Y.-F., Liu, J. S.-M., Li, K. K.-W., et al. (2015). TERT Promoter Mutations Contribute to Subset Prognostication of Lower-Grade Gliomas. *Mod. Pathol.* 28 (2), 177–186. doi:10.1038/modpathol.2014.94

Eckel-Passow, J. E., Lachance, D. H., Molinaro, A. M., Walsh, K. M., Decker, P. A., Sicotte, H., et al. (2015). Glioma Groups Based on 1p/19q,IDH, andTERTPromoter Mutations in Tumors. *N. Engl. J. Med.* 372 (26), 2499–2508. doi:10.1056/NEJMoa1407279

Filbin, M. G., and Suvà, M. L. (2016). Gliomas Genomics and Epigenomics: Arriving at the Start and Knowing it for the First Time. *Annu. Rev. Pathol. Mech. Dis.* 11, 497–521. doi:10.1146/annurev-pathol-012615-044208

Foote, M. B., Papadopoulos, N., and Diaz, L. A., Jr. (2015). Genetic Classification of Gliomas: Refining Histopathology. *Cancer Cell* 28 (1), 9–11. doi:10.1016/j.ccell.2015.06.014

Heaphy, C. M., de Wilde, R. F., Jiao, Y., Klein, A. P., Edil, B. H., Shi, C., et al. (2011). Altered Telomeres in Tumors with ATRX and DAXX Mutations. *Science* 333 (6041), 425. doi:10.1126/science.1207313

Hegi, M. E., Diserens, A.-C., Gorlia, T., Hamou, M.-F., de Tribolet, N., Weller, M., et al. (2005). MGMTGene Silencing and Benefit from Temozolomide in Glioblastoma. *N. Engl. J. Med.* 352 (10), 997–1003. doi:10.1056/NEJMoa043331

Heidenreich, B., and Kumar, R. (2017). TERT Promoter Mutations in Telomere Biology. *Mutat. Research/Reviews Mutat. Res.* 771, 15–31. doi:10.1016/j.mrrev.2016.11.002

Heidenreich, B., Rachakonda, P. S., Hosen, I., Volz, F., Hemminki, K., Weyerbrock, A., et al. (2015). TERT Promoter Mutations and Telomere Length in Adult Malignant Gliomas and Recurrences. *Oncotarget* 6 (12), 10617–10633. doi:10.18632/oncotarget.3329

Hoadley, K. A., Yau, C., Hinoue, T., Wolf, D. M., Lazar, A. J., Drill, E., et al. (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* 173 (2), 291–e6. doi:10.1016/j.cell.2018.03.022

Hu, H., Mu, Q., Bao, Z., Chen, Y., Liu, Y., Chen, J., et al. (2018). Mutational Landscape of Secondary Glioblastoma Guides MET-Targeted Trial in Brain Tumor. *Cell* 175 (6), 1665–1678. doi:10.1016/j.cell.2018.09.038

Jones, M. P., and Crowley, J. (1989). A General Class of Nonparametric Tests for Survival Analysis. *Biometrics* 45 (1), 157–170. doi:10.2307/2532042

Kamoun, A., Idbaih, A., Idbaih, A., Dehais, C., Elarouci, N., Carpentier, C., et al. (2016). Integrated Multi-Omics Analysis of Oligodendroglial Tumours Identifies Three Subgroups of 1p/19q Co-deleted Gliomas. *Nat. Commun.* 7, 11263. doi:10.1038/ncomms11263

Killela, P. J., Pirozzi, C. J., Healy, P., Reitman, Z. J., Lipp, E., Rasheed, B. A., et al. (2014). Mutations in IDH1, IDH2, and in the TERT Promoter Define Clinically Distinct Subgroups of Adult Malignant Gliomas. *Oncotarget* 5 (6), 1515–1525. doi:10.18632/oncotarget.1765

Killela, P. J., Reitman, Z. J., Jiao, Y., Bettegowda, C., Agrawal, N., Diaz, L. A., Jr., et al. (2013). TERT Promoter Mutations Occur Frequently in Gliomas and a Subset of Tumors Derived from Cells with Low Rates of Self-Renewal. *Proc. Natl. Acad. Sci. U.S.A.* 110 (15), 6021–6026. doi:10.1073/pnas.1303607110

Li, X., Cai, H., Zheng, W., Tong, M., Li, H., Ao, L., et al. (2016). An Individualized Prognostic Signature for Gastric Cancer Patients Treated with 5-Fluorouracil-Based Chemotherapy and Distinct Multi-Omics Characteristics of Prognostic Groups. *Oncotarget* 7 (8), 8743–8755. doi:10.18632/oncotarget.7087

Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., et al. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a Summary. *Acta Neuropathol.* 131 (6), 803–820. doi:10.1007/s00401-016-1545-1

Louis, D. N., Perry, A., Wesseling, P., Brat, D. J., Cree, I. A., Figarella-Branger, D., et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: a Summary. *Neuro Oncol.* 23 (8), 1231–1251. doi:10.1093/neuonc/noab106

Lu, C., Ward, P. S., Kapoor, G. S., Rohle, D., Turcan, S., Abdel-Wahab, O., et al. (2012). IDH Mutation Impairs Histone Demethylation and Results in a Block to Cell Differentiation. *Nature* 483 (7390), 474–478. doi:10.1038/nature10860

Masui, K., Komori, T., Kato, Y., Masutomi, K., Ichimura, K., Ogasawara, S., et al. (2018). Elevated TERT Expression in TERT-Wildtype Adult Diffuse Gliomas: Histological Evaluation with a Novel TERT-specific Antibody. *Biomed. Res. Int.* 2018, 1–12. doi:10.1155/2018/7945845

O'Sullivan, R. J., and Karlseder, J. (2010). Telomeres: Protecting Chromosomes against Genome Instability. *Nat. Rev. Mol. Cel Biol* 11 (3), 171–181. doi:10.1038/nrm2848

Olar, A., and Sulman, E. P. (2015). Molecular Markers in Low-Grade Glioma-Toward Tumor Reclassification. *Semin. Radiat. Oncol.* 25 (3), 155–163. doi:10.1016/j.semradonc.2015.02.006

Ostrom, Q. T., Cioffi, G., Gittleman, H., Patil, N., Waite, K., Kruchko, C., et al. (2019). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016. *Neuro Oncol.* 21 (Suppl. 5), v1–v100. doi:10.1093/neuonc/noz150

Rachakonda, S., Hoheisel, J. D., and Kumar, R. (2021). Occurrence, Functionality and Abundance of the TERT Promoter Mutations. *Intl J. Cancer* 149 (11), 1852–1862. doi:10.1002/ijc.33750

Reitman, Z. J., Pirozzi, C. J., and Yan, H. (2013). Promoting a New Brain Tumor Mutation: TERT Promoter Mutations in CNS Tumors. *Acta Neuropathol.* 126 (6), 789–792. doi:10.1007/s00401-013-1207-5

Simon, M., Hosen, I., Gousias, K., Rachakonda, S., Heidenreich, B., Gessi, M., et al. (2015). TERT Promoter Mutations: a Novel Independent Prognostic Factor in Primary Glioblastomas. *Neuro Oncol.* 17 (1), 45–52. doi:10.1093/neuonc/nou158

Valenzuela, M., Amato, R., Sgura, A., Antoccia, A., and Berardinelli, F. (2021). The Multiple Facets of ATRX Protein. *Cancers* 13 (9), 2211. doi:10.3390/cancers13092211

van den Bent, M. J., Brandes, A. A., Taphoorn, M. J. B., Kros, J. M., Kouwenhoven, M. C. M., Delattre, J.-Y., et al. (2013). Adjuvant Procarbazine, Lomustine, and Vincristine Chemotherapy in Newly Diagnosed Anaplastic Oligodendroglioma: Long-Term Follow-Up of EORTC Brain Tumor Group Study 26951. *Jco* 31 (3), 344–350. doi:10.1200/JCO.2012.43.2229

Wen, P. Y., and Packer, R. J. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: Clinical Implications. *Neuro Oncol.* 23 (8), 1215–1217. doi:10.1093/neuonc/noab120

Williams, E. A., Miller, J. J., Tummala, S. S., Penson, T., Iafrate, A. J., Juratli, T. A., et al. (2018). TERT Promoter Wild-type Glioblastomas Show Distinct Clinical Features and Frequent PI3K Pathway Mutations. *Acta Neuropathol. Commun.* 6 (1), 106. doi:10.1186/s40478-018-0613-2

Zhang, Z.-Y., Chan, A. K.-Y., Ding, X.-J., Qin, Z.-Y., Hong, C. S., Chen, L.-C., et al. (2015). TERTpromoter Mutations Contribute toIDHmutations in Predicting Differential Responses to Adjuvant Therapies in WHO Grade II and III Diffuse Gliomas. *Oncotarget* 6 (28), 24871–24883. doi:10.18632/oncotarget.4549

Zhao, Z., Meng, F., Wang, W., Wang, Z., Zhang, C., and Jiang, T. (2017). Comprehensive RNA-Seq Transcriptomic Profiling in the Malignant Progression of Gliomas. *Sci. Data* 4, 170024. doi:10.1038/sdata.2017.24

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

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

frontiers | Frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 26 April 2022
doi: 10.3389/fmolb.2022.843234

# Identification of Prognostic Biomarkers in Patients With Malignant Rhabdoid Tumor of the Kidney Based on mTORC1 Signaling Pathway-Related Genes

*Chenghao Zhanghuang[1†], Zhigang Yao[1†], Haoyu Tang[1], Kun Zhang[1], Chengchuang Wu[1], Li Li[2], Yucheng Xie[3], Zhen Yang[4] and Bing Yan[1]**

[1]Department of Urology, Kunming Children's Hospital, Kunming, China, [2]Key Laboratory of Pediatric Major Diseases, Kunming Children's Hospital, Kunming, China, [3]Department of Pathology, Kunming Children's Hospital, Kunming, China, [4]Department of Oncology, Kunming Children's Hospital, Kunming, China

**Background:** Malignant rhabdoid tumor of the kidney (MRTK) is an infrequent malignant tumor in childhood, accounting for approximately 2% of all childhood kidney tumors. Although the development of current treatments, the overall survival (OS) rate of MRTK patients is only 25%. The aim of this research was to explore the prognostic value of genes associated with the mTORC1 signaling pathway in MRTK.

**Methods:** The transcriptome data of MRTK samples were downloaded from the TARGET database. The 200 genes of HALLMARK_MTORC1_SIGNALING were downloaded from the Molecular Signatures Database (MSigDB). Furthermore, we applied gene set variation analysis (GSVA) to screen differentially expressed gene sets between the MRTK and normal samples. The 200 genes were combined with differentially expressed genes (DEGs) identified from differentially expressed gene sets. Then, a gene signature of mTORC1 pathway-related genes (mTRGs) was constructed in MRTK. The molecular mechanism of prognostic factors in MRTK was further analyzed using gene set enrichment analysis (GSEA). The target drugs based on these prognostic factors were explored from The Comparative Toxicogenomics Database (CTD). Moreover, six paired fresh tumor tissues and paraneoplastic tissues from children with MRTK were collected to validate the expressions of P4HA1, MLLT11, AURKA, and GOT1 in clinical samples via real-time fluorescence quantitative PCR and Western blot.

**Results:** A four-gene signature (P4HA1, MLLT11, AURKA, and GOT1) related to the mTORC1 pathway was developed in MRTK, which divided the MRTK patients into high-risk and low-risk groups. The patients with high-risk scores were strongly associated with reduced OS. Receiver operating characteristic (ROC) analysis indicated a good prediction performance of the four biomarker signatures. GSEA revealed that the mTOR signaling pathway was significantly enriched. The risk score was demonstrated to be an independent predictor for MRTK outcome. According to the correlation of tumor stem cell index and prognostic factors, the target drugs were obtained for the treatment of

OPEN ACCESS

**Edited by:**
Pedro Borralho,
Squad Lead—Novartis Oncology, Portugal

**Reviewed by:**
Daotai Nie,
Southern Illinois University Carbondale, United States
Chen Li,
Free University of Berlin, Germany
Udayan Bhattacharya,
NewYork-Presbyterian, United States
Xiaoyi Huang,
Harbin Medical University Cancer Hospital, China

***Correspondence:**
Bing Yan
ybwcy@163.com

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

**Specialty section:**
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

**Received:** 25 December 2021
**Accepted:** 18 March 2022
**Published:** 26 April 2022

**Citation:**
Zhanghuang C, Yao Z, Tang H, Zhang K, Wu C, Li L, Xie Y, Yang Z and Yan B (2022) Identification of Prognostic Biomarkers in Patients With Malignant Rhabdoid Tumor of the Kidney Based on mTORC1 Signaling Pathway-Related Genes. Front. Mol. Biosci. 9:843234. doi: 10.3389/fmolb.2022.843234

MRTK patients. Furthermore, the expressions of RT-qPCR and Western blot were consistent with RNA-sequencing data such that their expressions were significantly elevated in tumor tissues.

**Conclusion:** A total of four genes (P4HA1, MLLT11, AURKA, and GOT1) were screened as prognostic markers, further providing a new understanding for the treatment of patients with MRTK.

**Keywords: mTOR signaling pathway, malignant rhabdoid tumor of the kidney, therapeutic, prognosis, target drugs**

## BACKGROUND

Malignant rhabdoid tumor of the kidney (MRTK) is a kind of malignant rhabdoid tumor (MRT), which often occurs in infanthood. It has high invasiveness, short survival rate, and fast metastasis, and up to 80% of patients have metastatic disease (Wang et al., 2020). It has been reported that the overall survival rate of patients with MRTK was only 25% (Yanagisawa et al., 2009). Therefore, MRTK has the worst prognosis of all renal tumors. Moreover, the prognosis was much worse for children younger than 6 months with distant brain metastasis (Tahir et al., 2019). Our previous studies have shown that the PI3K-AKT signaling pathway and microRNA-related proteins have a very high potential value for the diagnosis and treatment of MRTK. (Zhanghuang et al., 2021). However, the prognosis prediction of MRTK has not been fully clarified. Therefore, exploring new targets for MRTK is urgent.

mTOR is a serine/threonine protein kinase belonging to the phosphoinositide 3-kinase (PI3K)-related family, which interacts with several proteins to form two different complexes, called mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2) (Zoncu et al., 2011). The core mTOR is mainly through insulin receptor (IR), insulin receptor substrate (IRS), type I phosphoinositide 3-kinase (PI3K), phosphoinositide-dependent protein kinase 1 (PDK1), and AKT-inducing kinase. The signaling cascade is mediated (Laplante and Sabatini, 2012). The loss of p53, a common event in cancer, promotes the activation of mTORC1 (Feng et al., 2005). The inhibition of mTORC1/mTORC2 blocked growth and induced catastrophic macropinocytosis in tumor cells (Kim et al., 2017; Srivastava et al., 2019). Significantly, there are no reports on the relationship between mTOR signaling pathway and MRTK.

In the present study, we successfully constructed a gene signature of mTORC1 pathway-related genes (mTRGs) and screened four genes (P4HA1, MLLT11, AURKA, and GOT1) as prognostic markers. The correlation between the tumor stem cell index and target drug was further explored, providing a new insight into further study of MRTK treatment.

## METHODS

### Ethic Statement

All human studies performed in the current study were approved by the Medical Ethics Committee of Kunming Children's Hospital before collection (approval number: 20190822002), and the guardians of the children signed informed consent forms.

### Data Collection

We retrieved the transcriptome data of MRTK from the TARGET database (https://ocg.cancer.gov/programs/target).

A total of 200 genes in the gene set of HALLMARK_MTORC1_SIGNALING were retrieved from the Molecular Signatures Database (MSigDB).

Clinical tissue specimen: The tissue specimens of six children with MRTK cancer and adjacent tissues in our hospital were selected. The relevant studies have been approved by the relevant ethics committee, and informed consent was signed by their families.

### Gene Set Variation Analysis

A total of 50 MRTK samples and six normal samples were applied to conduct GSVA, of which h.all.v7.2.symbols.gmt was considered to be the preset pathway. Then, we used the "limma" package to analyze the differences of GSVA scores between the MRTK and control groups. The screening conditions were $|t\text{-value}| > 2$ and $p\text{-value} < 0.05$. In the GSVA, the screened pathways with t-value $> 0$ were considered to be activated in the MRTK group, while screened pathways with $t\text{-value} < 0$ were considered to be activated in the normal group.

### Differentially Expressed Analysis of Activated Pathway in the Malignant Rhabdoid Tumor of the Kidney Group

The genes were extracted from the identified HALLMARK_MTORC1_SIGNALING pathway. Thereafter, we used the "limma" R package to perform the differential analysis for identifying the differentially expressed genes (DEGs) between the MRTK samples and normal samples. The screening conditions of DEGs were $|\log_2 FC| > 1$ and $p\text{-value} < 0.05$. We combined the DEGs with 200 genes of the HALLMARK_MTORC1_SIGNALING gene set, obtaining the differentially expressed mTORC1 pathway-related genes (DE-mTRGs).

### Construction of an mTRG-Related Gene Signature

A total of 49 MRTK samples containing complete clinical information were applied to construct an mTRG-related gene signature. The univariate Cox regression analysis and

Kaplan–Meier (K-M) survival curves were performed by employing the survival R package to identify the DE-mTRGs related to the overall survival (OS) of MRTK ($p < 0.05$). DE-mTRGs with $p < 0.05$ were input into a multivariate Cox regression analysis for the construction of an mTRG-related gene signature. The risk score of each MRTK sample is obtained as the following formula: Risk score = h0(t)*exp ($\beta$1*X*1+$\beta$2*X*2+. . .+$\beta$n*X*n). In this formula, $\beta$ is the regression coefficient, and h0(t) is the benchmark risk function. A total of 49 MRTK samples were classified into high-risk and low-risk groups with the boundary of the median risk score. The K-M survival curve was generated via the "survdiff" functions of the R package to compare the OS of two risk groups. Thereafter, a time-independent receiver operating characteristic (ROC) curve was conducted to demonstrate the effectiveness of this mTRG-related gene signature for OS in MRTK patients.

The relationship between risk score and clinicopathological features was evaluated using the Chi-square test. Then, we executed univariate and multivariate Cox regression analyses to evaluate whether these clinicopathological features were independent predictors for RKT prognosis. A nomogram containing independent prognostic factors was generated using the "rms" R package, and the corresponding calibration plot was further established to evaluate the efficiency of the nomogram.

## Functional Enrichement Analysis

Gene set enrichment analysis (GSEA) of prognostic mTRGs was analyzed by the "clusterProfiler" R package (version 3.18.0) to explore the potential signaling pathways that were involved. The "limma" package was executed to screen the DEGs in high- and low-risk groups with the screening conditions being $|\log_2 FC| > 1$ and $p$-value < 0.05. In addition, functional enrichment analyses including Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were further conducted using Metascape software based on the DEGs.

## Tumor Stem Cell Index Calculation and Drug Prediction for Malignant Rhabdoid Tumor of the Kidney

The tumor stem cell index of MRTK samples was calculated using the OCLR algorithm. The correlations of tumor stem cell index and the prognostic factors were analyzed by the Spearman correlation analysis. Additionally, we searched the target drugs of these prognostic factors from The Comparative Toxicogenomics Database (CTD). The interactions of target drugs and prognostic genes were visualized by Cytoscape software.

## Clinical Samples

This study has been approved by the Ethics Committee of Kunming Children's Hospital, and all patients and their parents signed informed consent before joining the study. A total of six cases of fresh tumor and paratumor tissues were collected from children with MRTK admitted to Kunming Children's Hospital from July 2014 to June 2020. There were 2 male and 4 female cases aging from 4 months to 4 years and were not treated with radiotherapy before surgery. The tumor tissues were excised as soon as possible after isolation in 0.5 $cm^3$ volume, and relatively normal kidney tissues were taken from the distal cut edge of pathologically confirmed paratumor-infiltrated tissues. The samples were snap-frozen with liquid nitrogen and transferred to −80 °C for storage prior to use.

## Real-Time Reverse Transcription Polymerase Chain Reaction

Total RNA was extracted from MRTK tumor and paratumoral tissues using a TRIzol (T9424, Sigma, Germany) reagent following the manufacturer's instructions. The RNA concentration and purity were determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). Then, mRNA was reversely transcribed into cDNA using a PrimeScript RT reagent kit with gDNA Eraser (RR047A, Takara, Japan). qRT-PCR was performed using a GoTaq qPCR and RT-qPCR kit (A6001, Promega, United States) using an ABI ViiATM7 real-time PCR system (Applied Biosystems, United States). The PCR program was started with an initial step of 95°C for 3 min, followed by 40 cycles of 15 s denaturation at 95°C and 30 s extension at 60°C. The mRNA expression levels were calculated by the $2^{-\Delta\Delta t}$ method [△△ Ct = (Ct target gene-Ct internal reference gene) experimental group-(Ct target gene-Ct internal reference gene) control group] and normalized to GAPDH. The experiments were performed in triplicate for at least three samples of each group. The primers were synthesized by Sangon Biotech (Shanghai) and are listed in **Table 1**.

## Western Blot

The protein was extracted from tumor and paratumor tissues using RIPA buffer (Beyotime, Shanghai, China). BCA protein quantification kit (Shanshan Jinqiao Company, Beijing, China) was used for the quantification of protein samples. About 30 μg protein of each sample was loaded on SDS-PAGE gels. After transferring to the polyvinylidene fluoride membrane (Millipore, United States), the blots were probed with the appropriate primary antibody (anti-P4HA1, 1: 1000, ProteinTech; anti-MLLT11, 1:1000, ProteinTech; anti-AURKA, 1:1000, Wanleibio; anti-GOT1, 1:1000, Wanleibio; anti-β-actin, 1: 1000, Zen-Bioscience). After incubation with the primary antibodies, the membrane was incubated with the corresponding secondary antibodies and monitored with an enhanced chemiluminescence (ECL, Affinity, United States) substrate kit (Amersham, Biosciences, United States). The protein bands were photographed using an Image Lab system (Bio-Rad Laboratories, Inc., United States). After developing, the image was saved, and the Image Lab (Bio-Rad Laboratories, Inc., United States) analyzed the gray value.

## Statistical Analysis

All bioinformatics analyses involved in this research were conducted using the R software (version 3.6.3). The relationship between risk score and clinicopathological features was demonstrated using the Chi-square test. Moreover, the independent prognostic factors for MRTK identified the clinicopathological features by univariate and multivariate Cox regression analyses. GraphPad8.0 statistical software was used for data analysis of mRNA and protein expression. The t-test was used to compare the difference between the tumor and paratumor groups. $p < 0.05$ was considered statistically significant unless specified.

**TABLE 1 |** Primer sequences of P4HA1, MLLT11, AURKA, GOT1, and GAPDH.

| | Primer sequence (5′-3′) |
|---|---|
| P4HA1 | Forward:AGTACATGACCCTGAGACTGGAAAA |
| | Reverse:ATCTGGCTCATCTTTCCGTGC |
| MLLT11 | Forward:CCATCTTTGGAACACGCCAG |
| | Reverse:AACGCTGCTGTCTTTGACCT |
| AURKA | Forward:TGGAAGACTTGGGTCCTTGG |
| | Reverse:AAATATCCCCGCACTCTGGC |
| GOT1 | Forward:GCCAAGTGGTCGAATCAACG |
| | Reverse:CTGGACGGGTGGTGTTTCTT |
| GAPDH | Forward:AGGTCGGTGTGAACGGATTTG |
| | Reverse:GGGGTCGTTGATGGCAACA |

# RESULTS

## Screening of Differentially Expressed mTORC1 Pathway-Related Genes in Malignant Rhabdoid Tumor of the Kidney

Through GSVA analysis, we screened 9 activated pathways in the MRTK group and 15 activated pathways in the normal group. Interestingly, we observed that the mTORC1 signaling pathway was activated in the MRTK samples (**Figure 1A**; **Supplementary Table S1**). Subsequently, we extracted the genes from the mTORC1 signaling pathway to screen the DEGs between the MRTK and normal samples. As a consequence, a total of 4787 DEGs were screened in the MRTK samples compared with normal samples, of which 1294 were upregulated and 3493 were downregulated (**Figure 1B**; **Supplementary Table S2**). After intersection with the 200 genes of the HALLMARK_MTORC1_SIGNALING gene set obtained from the MsigDB, we identified 70 DE-mTRGs in MRTK (**Figure 1C**). The expressions of 70 DE-mTRGs between the MRTK and normal samples were displayed in a heatmap. Among the 70 DE-mTRGs, 22 genes were downregulated and 48 genes were upregulated in patients with MRTK (**Figure 1D**).

## Development of a DE-mTRG-Related Signature in Malignant Rhabdoid Tumor of the Kidney

To assess whether these 70 DE-mTRGs were related to survival, we first performed the K-M survival curves and univariate Cox regression analysis. As a result, four DE-mTRGs were demonstrated to be related to the OS of MRTK (**Figure 2A**, **Supplementary Table S3**, $p < 0.05$). A total of four DE-mTRGs including P4HA1, MLLT11, AURKA, and GOT1 were obtained from the multivariate Cox regression analysis, which were used to develop an mTRG-related gene signature (**Figure 2B**). Moreover, the survival analyses of each gene uncovered that patients with high expression of these genes were all related to a poor prognosis (all with $p < 0.05$, **Supplementary Figure S1**). The risk score of each MRTK patient was obtained as the formula mentioned in Materials and Methods. The MRTK patients were stratified into a high-risk group ($n = 25$) and a low-risk group ($n = 24$) by the boundary of the median risk score. The risk curves and patient survival scatter plots showed that the number of patient deaths increased with the risk score (**Figure 2C**). The K-M curves demonstrated that the risk score significantly differentiated clinical outcomes in MRTK patients ($p < 0.0001$), with a high-risk score implying a poor prognosis (**Figure 2D**). The area under curve (AUC) was 0.811 and 0.878 at 1 and 3 years, respectively, indicating the validity of the 4-gene-based prognostic signature (**Figure 2E**). Furthermore, the heatmap showed that all 4 model genes possessed relatively high expression levels in the high-risk group (**Figure 2F**).

## Independent Prognostic Analysis in Malignant Rhabdoid Tumor of the Kidney

Next, a Chi-square test was implemented to assess the distribution of clinical characteristics of MRTK patients in the high- and low-risk groups. The results showed that the age distribution of patients in the high- and low-risk groups was significantly different ($p = 0.032$), with 72.0% of patients in the

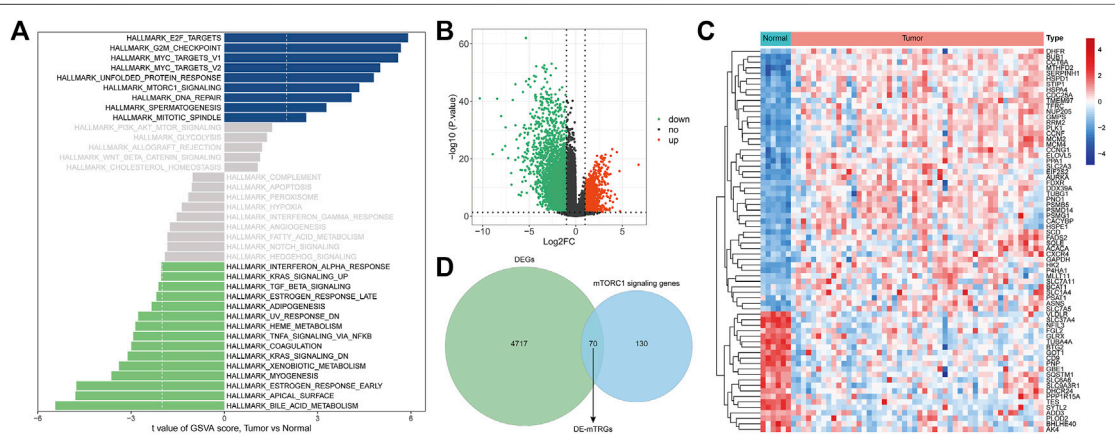



**FIGURE 1 |** Identification of differentially expressed mTORC1 pathway-related genes (DE-mTRGs) in MRTK. **(A)** GSVA analysis; **(B)** volcano map. Green: downregulated (Down), red: upregulated (Up), black: unchanged (no); **(C)** Heatmap of DE-mTRGs in MRTK; **(D)** Venn diagram. Green: differentially expressed genes; blue: mTORC1 signaling genes;intersecting section: DE-mTRGs.

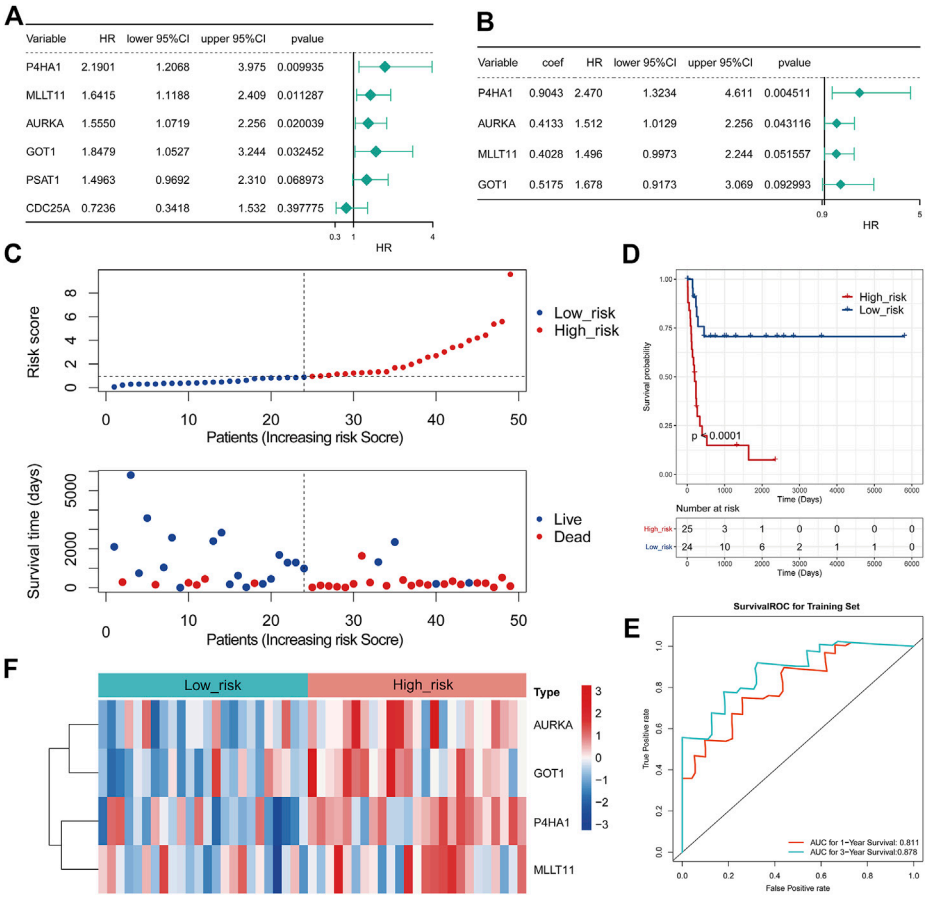



**FIGURE 2 |** Development of an mTRG-related gene signature in MRTK. **(A)** Forest plot of hazard ratios (HRs) demonstrating the prognostic values of DE-mTRG. **(B)** Forest plot illustrating the multivariable Cox model results of each gene in the 4-mTRG risk signature. The dashed line was used to mark the location of HR = 1. **(C)** Relationship between the survival status/risk score rank and survival time (days)/risk score rank. **(D)** Kaplan–Meier (K-M) plot for overall survival (OS) based on the risk score of the four gene-based signature of patients with MRTK. **(E)** Receiver operating characteristic (ROC) curve for OS of the TARGET-MRTK set. The area under curve (AUC) was assessed at 1 and 3 years. **(F)** Heatmap demonstrating the distribution of the four DE-mTRG expressions in the TARGET-MRTK cohort.

**TABLE 2 |** High and low risk groups association with clinical characteristics.

| | Expression | | | |
|---|---|---|---|---|
| | **Total (*N* = 49)** | **High_risk (*N* = 25)** | **Low_risk (*N* = 24)** | **p-value** |
| Gender | | | | |
| Female | 23 (46.9%) | 13 (52.0%) | 10 (41.7%) | 0.661 |
| Male | 26 (53.1%) | 12 (48.0%) | 14 (58.3%) | |
| Age(Days) | | | | |
| ≥365(Days) | 27 (55.1%) | 18 (72.0%) | 9 (37.5%) | 0.032 |
| <365(Days) | 22 (44.9%) | 7 (28.0%) | 15 (62.5%) | |
| Pathologic stage | | | | |
| I-II | 13(26.5%) | 4 (16.0%) | 9 (37.5%) | 0.167 |
| III-IV | 36 (73.5%) | 21 (84.0%) | 15(62.5%) | |

high-risk group being ≥365 days old and 62.5% of patients in the low-risk group being less than 365 days old (**Table 2**). Furthermore, the Wilcoxon rank sum test indicated that the level of the risk score was significantly higher in MTRK patient's ≥365 days than in MTRK patients' <365 days ($p < 0.01$); however, the risk score was not statistically significant in the gender and stage subgroups (**Figure 3A**). Subsequently, we performed an independent prognostic analysis by Cox regression to assess whether the risk score could affect the clinical outcome of OS in MRTK patients independent of the clinicopathological characteristics (age, sex, and stage). The univariate Cox regression analysis showed that the risk score and stage were significantly associated with OS in patients with MRTK ($p < 0.05$; **Table 3**). Finally, multivariate Cox regression analysis pointed out that the risk score and stage were independent prognostic factors for MRTK patients (**Table 4**). Thereafter, we constructed a nomogram based on two independent prognostic factors that predicted patients' 1-year and 3-year OS (**Figure 3B**), which had a C-index of 0.767. The calibration curve indicated that the predictive value of the nomogram model for OS in MRTK patients was similar to the actual observed value (**Figure 3C**), implying that the combined model of the two independent prognostic factors may be more clinically applicable.

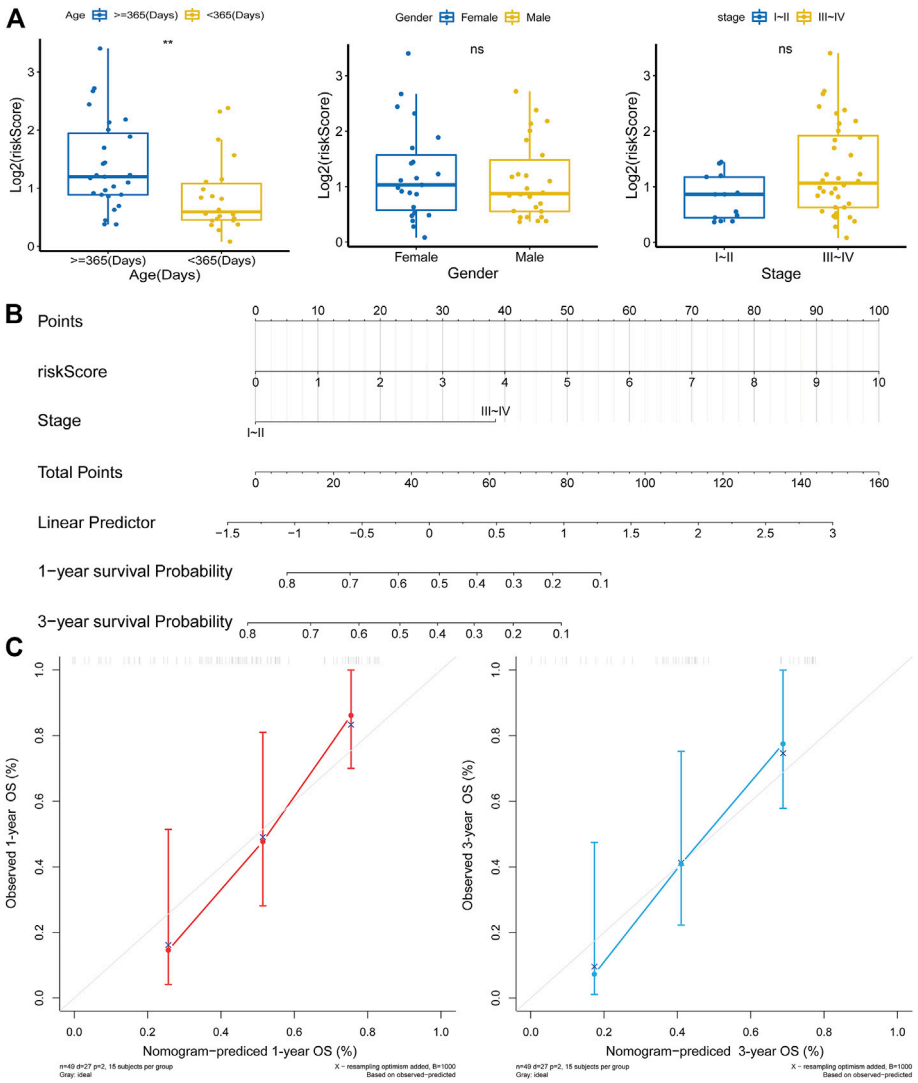



**FIGURE 3 |** Independent prognostic analysis in MRTK. **(A)** Correlation analysis between risk score and clinicopathological characteristics. **(B)** Nomogram that integrated the risk score and stage predicted the probability of the 1- and 3-year OS. To calculate the survival probability, identification of the patient values on each axis and then for a vertical line upward to the point axis was drawn for each value. The points for all variables were added, and this sum on the total point line was located. **(C)** Calibration plot of the nomogram for predicting probabilities of 1-year and 3-year OS of patients.

**TABLE 3 |** Univariate Cox regression analysis of high-risk and low-risk groups.

| ID | HR | HR.95L | HR.95H | p-value |
|---|---|---|---|---|
| riskScore | 1.39 | 1.19 | 1.61 | 0.00 |
| Stage(Reference:StageI~II) | 3.90 | 1.17 | 12.98 | 0.03 |
| Gender(Reference:Female) | 0.63 | 0.29 | 1.36 | 0.24 |
| Age_at_Diagnosis_in_Days | 1.00 | 1.00 | 1.00 | 0.25 |

**TABLE 4 |** Multivariate Cox regression analysis of high-risk and low-risk groups.

| ID | HR | HR.95L | HR.95H | p-value |
|---|---|---|---|---|
| riskScore | 1.34 | 1.15 | 1.56 | 0.00 |
| Stage(Reference:StageI~II) | 3.05 | 0.89 | 10.42 | 0.08 |

## Gene Set Variation Analysis of the Prognostic Biomarkers

We conducted GSEA to further explore the involved signaling pathways of the four prognostic biomarkers. The results indicated that AURKA was mainly involved in cell cycle, DNA replication, non-alcoholic fatty liver disease, oxidative phosphorylation, proteasome, ribosome, and ribosome biogenesis in eukaryotes, RNA transport, spliceosome, and thermogenesis (**Figure 4A**; **Supplementary Table S4**). Similarly, GOT1 was significantly associated with various signaling pathways such as ECM–receptor interaction, carbon metabolism, human papillomavirus infection, AGE–RAGE signaling pathway in diabetic complications, focal adhesion, insulin resistance, lysosome, MAPK signaling pathway, peroxisome, and tight

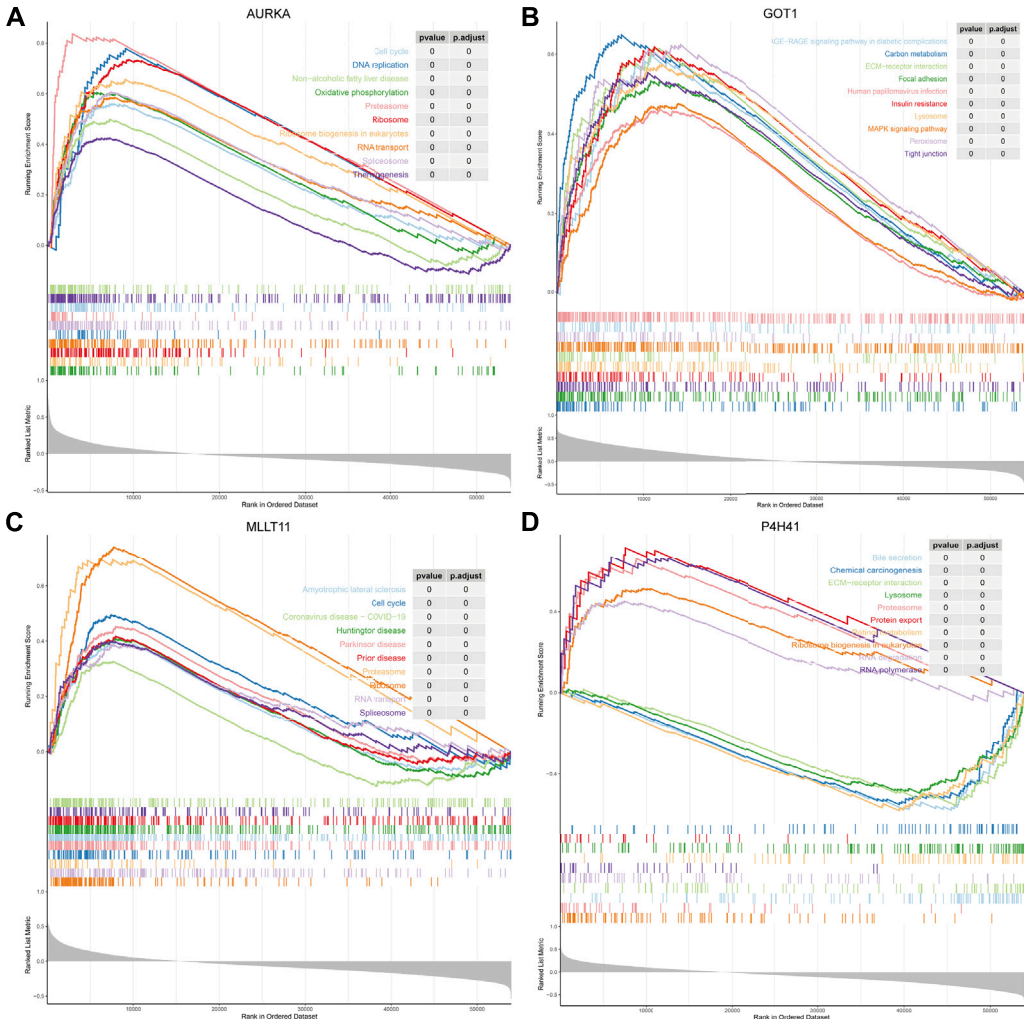



**FIGURE 4 |** GSEA of the prognostic biomarkers. GSEA explores related signal pathways of **(A)** AURKA; **(B)** GOT1; **(C)** MLLT11; and **(D)** P4H41.

junction (**Figure 4B**; **Supplementary Table S5**). Interestingly, the expression of MLLT11 was correlated with several diseases, including amyotrophic lateral sclerosis, coronavirus disease-COVID-19, Huntington's disease, Parkinson's disease, and prion disease (**Figure 4C**; **Supplementary Table S6**). In addition, P4H41 was also mainly involved in bile secretion, chemical carcinogenesis, ECM–receptor interaction, lysosome, proteasome, RNA degradation, and RNA polymerase (**Figure 4D**; **Supplementary Table S7**).

## Functional Enrichment Analysis

To elucidate the molecular mechanism that was involved in the risk score, we conducted GO and KEGG functional enrichment analyses on DEGs of the two risk groups (**Figures 5A,B**). The results showed that the DEGs mainly focused on several important biological processes including central nervous system neuron differentiation and developmental process involved in reproduction, forebrain development, forebrain neuron differentiation, positive regulation of cell development, reproductive structure development, and reproductive system development (**Figure 5C**). KEGG analysis indicated that these DEGs were mainly associated with basal cell carcinoma, pathways in cancer, proteoglycans in cancer, melanogenesis, PI3K–AKT signaling pathway, gastric cancer, signaling pathways, mTOR signaling pathway, regulating pluripotency of stem cells, hepatocellular carcinoma, human papillomavirus infection, Hippo signaling pathway, cGMP-PKG signaling pathway, breast cancer, Cushing syndrome, and Wnt signaling pathway (**Figure 5D**).

## Target Drug Prediction Based on the Tumor Stem Cell Index

It has been reported that mTOR exerts important functions in cancer stem cells through the specific functions related to stemness (Matsubara et al., 2013). We calculated the tumor stem cell index (mRNAsi) of MRTK samples based on the OCLR algorithm (Zhang et al., 2020) (**Figure 6A**). The

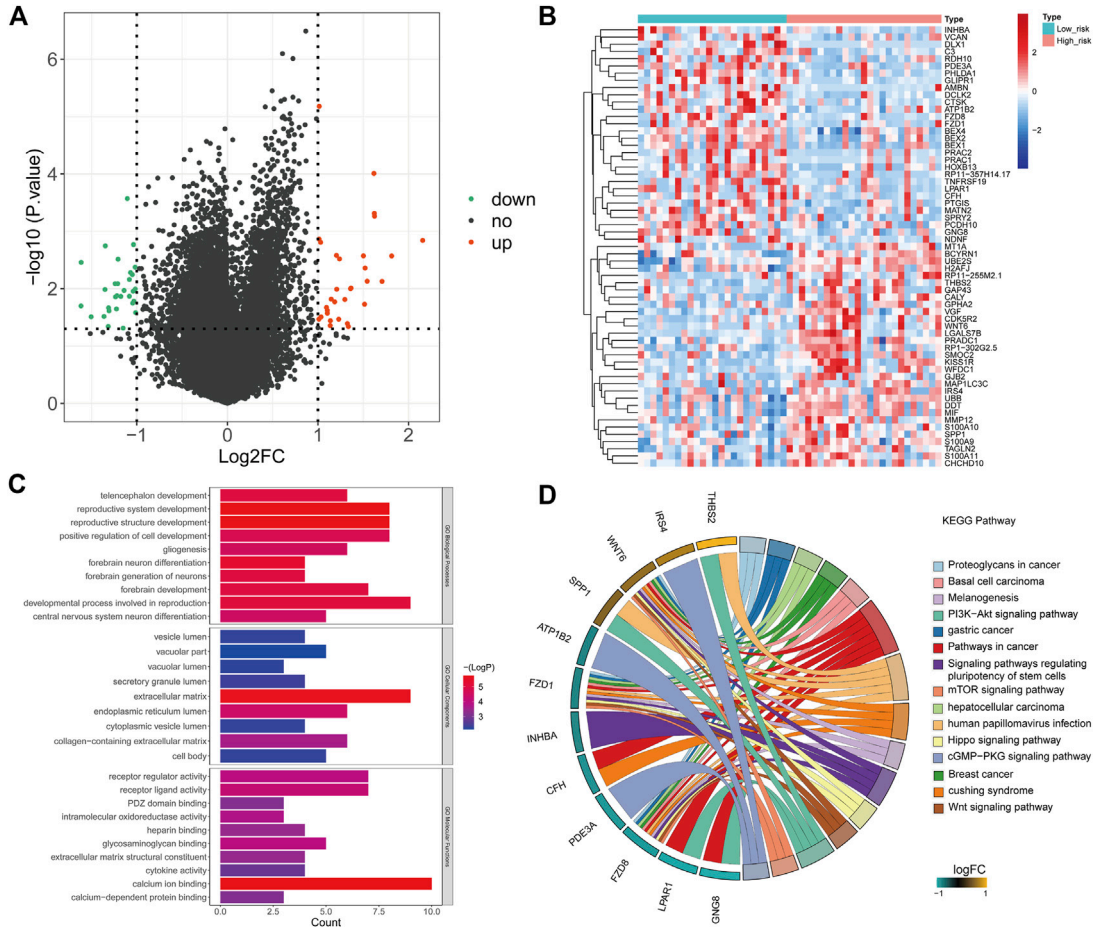



**FIGURE 5 |** Functional enrichment analysis. **(A)** Volcano plot showing the DEGs of the risk score (high vs. low). Green: downregulated (Down), red: upregulated (Up), black: unchanged (no). **(B)** Heatmap based on 61 DEGs that were differentially expressed between high- and low-risk score samples. Each column represents a sample, and each row represents the mRNA level. Red indicates that the gene is upregulated, and green indicates downregulated genes. **(C)** Top 10 Gene Ontology terms enriched by the DEGs. **(D)** KEGG enrichment analysis of DEGs. The left side of the circle is the gene, and the right side is the pathway. Different colors represent different pathways, and the color of each pathway is annotated in right of the circle. If the gene belongs to a pathway, there will be a line between the gene and the pathway.

correlation analysis showed that the expressions of AURKA and GOT1 were positively correlated with the mRNAsi in MRTK (**Figure 6B**). Then, we searched the corresponding target drugs of AURKA and GOT1 from the CTD. A network containing 262 nodes and 349 edges showed the interactions between the target drugs and prognostic factors that were positively correlated with mRNAsi (**Figure 6C**). These target drugs predicted by AURKA and GOT1 may play an essential role in the therapy of MRTK.

## Expression Validation of Prognostic Biomarkers in Clinical Samples

RT-qPCR results showed that the abundance of P4HA1, MLLT11, AURKA, and GOT1 were significantly elevated in tumor tissues (**Figure 7A**). Consistent with the RT-qPCR results, Western blot also showed that the protein levels of P4HA1, MLLT11, AURKA, and GOT1 were markedly higher in the tumor tissues than those in the paratumor tissues (**Figure 7B**). These results were in accordance with the RNA-sequencing results, indicating the reliability of our analysis.

## DISCUSSION

In this study, we first determined that the mTORC1 signaling pathway was activated in the MRTK samples through GSVA analysis. Tuberous sclerosis, a disease in which mTORC1 is over-activated by TSC1 or TSC2 deletion, which leads to the formation of a wide range of benign tumors, is also direct evidence that mTORC1 plays a key role in tumorigenesis. The negative feedback of insulin receptor substrate (IRS)-1 mediated by mTORC1 may be the reason for the limited progression of these tumors. It can effectively downregulate PI3K signaling downstream of most receptor tyrosine kinases (RTKs) (Harrington et al., 2004; Shah et al., 2004; Um et al.,

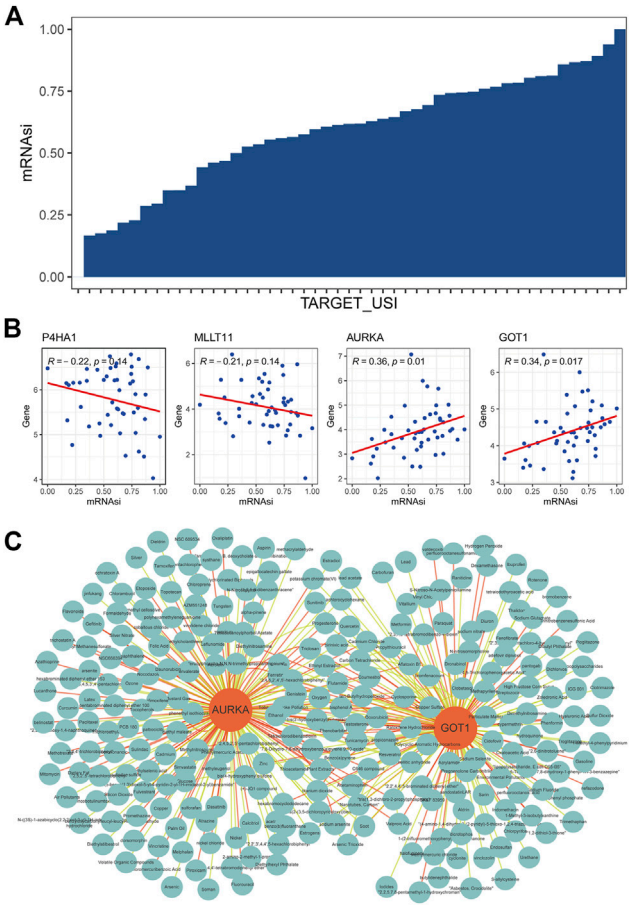



**FIGURE 6 |** Target drug prediction based on the tumor stem cell index. **(A)** Tumor stem cell index (mRNAsi) of MRTK samples based on the OCLR algorithm. TARGET_USI: unique sample identity document (USI) for 50 MRTK samples from the TARGET database; **(B)** correlation analysis between mRNAsi and P4HA1, MLLT11, AURKA, and GOT1; **(C)** Interaction between target drugs and prognostic factors that are positively correlated with mRNAsi effect.

2004). In addition, mTORC1 can directly phosphorylate a linker called Grb10 that directly binds to RTK (Hsu et al., 2011; Yu et al., 2011). The patients with tuberous sclerosis also proved that mTORC1 signaling as a single molecular mutation is a powerful driving force for cell proliferation. Under the premise of genetic and molecular changes, mTORC1 signaling aggravates tumor progression through a variety of molecular mechanisms.

Autophagy is an intracellular process that leads to the orderly degradation and recycling of cellular components. In addition to regulating cell growth and metabolism, mTORC1 also regulates this process of autophagy. mTORC1 can negatively regulate autophagy inhibition to initiate autophagy through the phosphorylation of ULK and VPS34 (Kim and Guan, 2015). The knockout of Beclin1 can block autophagy and promote tumor formation, so autophagy is considered to be an effective tumor suppressor (Qu et al., 2003; Yue et al., 2003; White, 2015). However, there is a lot of evidence that tumor cells can also use autophagy to enhance cell activity under metabolic stress (Degenhardt et al., 2006). In this case, the inhibition of mTORC1 will greatly enhance autophagy, which may cause cells to increase the synthesis of nutrient molecules and enhance tumor cell survival.

Based on mTORC1-related genes, we constructed a prognostic model by 4 gene signatures, including P4HA1, MLLT11, AURKA, and GOT1. Collagen prolyl 4-hydroxylase alpha polypeptide I (P4HA1) is essential for the collagen prolyl hydroxylation of normal cells. Previous studies on P4HA1 mainly focused on the congenital connective tissue disorder caused by gene mutation, and the correlation between P4HA1 and tumors has rarely been reported. In recent years, the relationship between P4HA1 and tumors has attracted people's attention. Gilkes et al. (2013) found that the hypoxia-inducible factor (HIF1) was widely involved in ECM remodeling by regulating P4HA1 and P4HA2, thereby regulating the occurrence and development of breast cancer. *In vivo* and *in vitro* studies by Chakravarthi et al. (2014) showed that the over-expression of P4HA1 improved the proliferation and invasion ability of prostate cancer cells through the miR-124/P4HA1/MMP1 axis. A new study indicated that miR-122 inhibited the epithelial–mesenchymal transformation process by targeting P4HA1, thereby regulating the invasion and abdominal metastasis of ovarian cancer (Zhou et al., 2017; Duan et al., 2018). Aurora kinase A (AURKA) is a member of the serine/threonine kinase family necessary for the regulation of mitosis and the process of cell division. Aurora kinase A (AURKA) can regulate the process of mitosis, centrosome maturation and separation, and spindle mitosis. (Marumoto et al., 2005). Due to the extremely high expression of AURKA in cancer, whether AURKA can be used as a potential therapeutic target has aroused great interest in the academic community (Marumoto et al., 2005). Aurora kinase inhibitors such as MLN8237 and PHA-739358 have been developed and applied (Kollareddy et al., 2012), but only moderate effects have been found in preclinical and clinical studies (Kollareddy et al., 2012; Goldberg et al., 2014). In addition, new evidence shows that AURKA also promotes the occurrence and development of cancer through other mechanisms unrelated to its kinase activity (Otto et al., 2009). Glutamine metabolism is also very important for the proliferation of cancer cells. Recent studies have shown that hyperproliferative cells can use glutamine derived from glutamine to produce non-essential amino acids (NEAA) through glutamate-oxaloacetate transaminase (*GOT1*), but resting cells can pass through glutamate dehydrogenase 1 (*GLUD1*) and the subsequent decarboxylation reaction in the TCA cycle of metabolizing glutamate (Hensley et al., 2013). Research by Fenja M Feld et al. proved that *GOT1* could be used as a prognostic biomarker for pancreatic ductal cancer (Yang, 2016). Yang Yong et al. have shown that inhibiting *GOT1* can enhance the efficacy of adriamycin against triple-negative breast cancer (Yang, 2016). The MLLT11 gene, located on chromosome 1 band q21, was initially identified as a mixed-lineage leukemia (MLL) fusion partner from acute myeloid leukemia (AML) patients whose leukemic cells carried a t(1; 11) (q21; q23) chromosomal abnormality (Tse et al., 1995).

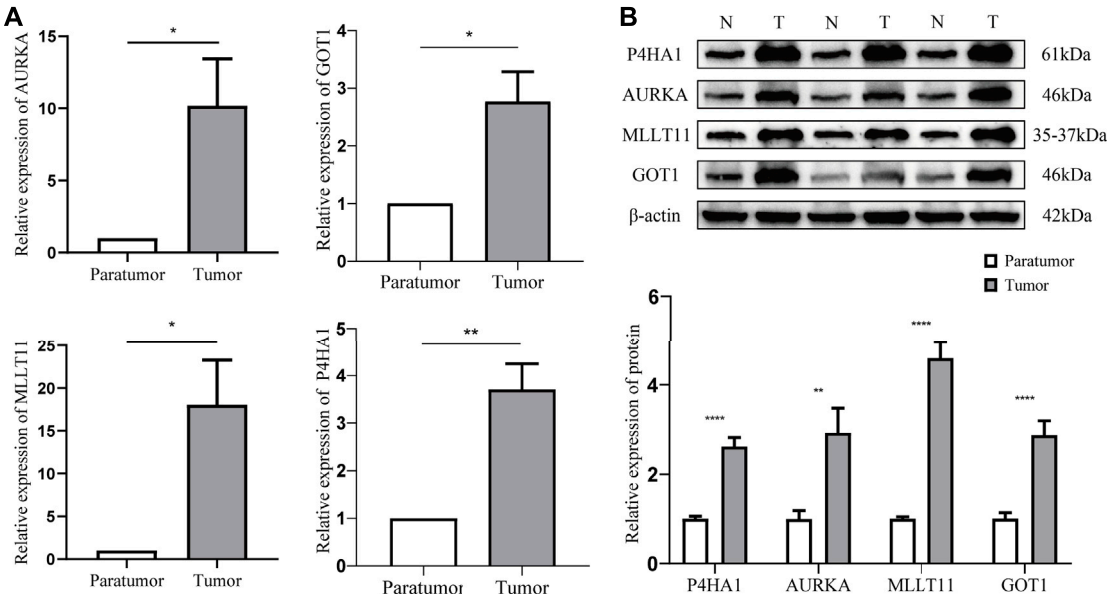



**FIGURE 7 |** Expression validation of prognostic biomarkers in clinical samples. **(A)** Expressions of P4HA1, MLLT11, AURKA, and GOT1 in tumor and paratumor samples by RT-qPCR. **(B)** Expressions and quantification result of P4HA1, MLLT11, AURKA, and GOT1 in tumor and paratumor samples by Western blot. *, $p < 0.05$ vs. control; **, $p < 0.01$ vs. control; ***, $p < 0.001$; ****, $p < 0.0001$.

The expression of MLLT11 is strictly regulated in normal lineage-directed hematopoietic progenitor cells (HPCs) (Tse et al., 1995). However, the expression of MLLT11 is increased widely in acute myeloid and lymphoid leukemia (Tse et al., 1995). Research by Yin Xiong et al. proved that AML children with low MLLT11 expression have poorer overall survival (Xiong et al., 2011). However, Elisabeth S Gruber proved that the high expression of AF1q (MLLT11 corresponding gene name) could independently be used as a prognostic indicator for patients with esophageal cancer (Gruber et al., 2019).

Subsequently, we performed a functional enrichment analysis on the DEGs of the high- and low-risk groups. Interestingly, the mTOR signaling pathway was significantly enriched. Previous studies have shown the importance of the mTOR pathway in the pathogenesis of cancer. mTOR complex 1 (mTORC1) can increase mRNA translation, protein synthesis, and cell proliferation (Tasian et al., 2014). The activation of the second mTOR complex (mTORC2) involved in the regulation of the cytoskeleton may be a feedback effect of the AKT loop (Tasian et al., 2014). Balsara et al. found that 74% of specimens from patients with non-small cell lung cancer (NSCLC) stained positive for mTOR through the use of tissue microarrays (TMAs) (Balsara et al., 2004). Rictor is a subunit of mTORC2, which promotes the assembly and activity of mTORC2, and endows glioma cells with proliferation and invasion potential (Fan and Weiss, 2012).

Cancer stem cells (CSCs) are a subgroup of tumor tissues that are highly immune to traditional cancer treatments. At present, slow-circulating CSC is the main obstacle to the eradication of most tumors (Begicevic and Falasca, 2017). Traditionally, the mTOR pathway is over-activated in CSC. Furthermore, it has been found that targeted inhibition of the mTORC1 signaling pathway is 30 times more powerful than non-CSC in inhibiting the proliferation and survival of CSC in solid tumor cell populations. Transforming growth factor-$\beta$ (TGF-$\beta$) can promote epithelial–mesenchymal transition (EMT), thereby increasing the production and activation of cancer stem cells. mTOR is a key node in the TGF-$\beta$ signaling pathway to enhance cancer stemness and drug resistance (Katsuno et al., 2019). Some mTOR inhibitors have shown inhibitory function on CSC (Francipane and Lagasse, 2016), for example, rapamycin, everolimus, and PF-04691502 can inhibit the activation of breast cancer stem cells induced by tamoxifen (Karthik et al., 2015). Also, inhibiting mTOR can restore the resistance of breast cancer cells to tamoxifen (DeGraffenried et al., 2004). In addition, Torin1 (ATP competitive mTOR inhibitor) and VS-5584 (PI3K/mTOR inhibitor) can significantly reduce tumor CSC levels in a variety of human cancer transplant models (Francipane and Lagasse, 2013; Kolev et al., 2015). Increasing studies have demonstrated the role of the mTOR signaling pathway in maintaining CSC. Chang et al. (https://www.ncbi.nlm.nih.gov/pubmed/24157869/) found that the radiation resistance of prostate cancer was related to the enhanced CSC phenotype through the activation of the PI3K/AKT/mTOR signaling pathway (Chang et al., 2013). Inhibition of mTORC2 leads to a reduction in the expression of liver CSC markers (epithelial cell adhesion molecule, EpCAM) in hepatocellular carcinoma stem cells (Sunayama et al., 2010; Nishitani et al., 2013). Bleau et al. found that the cross-inhibitory regulation between MEK/ERK and PI3K/mTOR pathways maintained the self-renewal and tumorigenic ability of glioblastoma cancer stem-like cells (Bleau et al., 2009). Corominas-Faja et al. found that AKT regulated the

activity of the ATP-binding cassette transporter (ABCG2) in glioma tumor stem-like cells (Corominas-Faja et al., 2013).

Our study only analyzed the data of TCGA database, and the results have several limitations. In our subsequent studies, we will collect clinical specimens combining cells and animal experiments to further verify the results in our study, which provides a reliable theoretical basis for the treatment of MRTK. Also, the reliability of the results was demonstrated by RT-PCR and WB validation of fresh tumor tissue and paratumor samples from six children.

The US Food and Drug Administration (FDA) approved an mTOR inhibitor for the treatment of renal cell carcinoma, named temsirolimus (Bergmann et al., 2014). So far, humans have developed three generations of compounds targeting the PI3K/mTOR signaling pathway. As the first generation of PI3K inhibitors, pan-inhibitors can be used to bind to all PI3K targets (Martelli et al., 2012). The second-generation inhibitors with isoform-specific selective activity have higher specificity (Martelli et al., 2012). The dual PI3K/mTOR inhibitor is the third-generation inhibitor. It can inhibit not only all PI3K I subtypes but also mTORC1 and mTORC2 (Zhu et al., 2012).

## CONCLUSION

In conclusion, we found that the mTORC1 signaling pathway was activated in MRTK samples and successfully constructed a risk model using differentially expressed mTORC1 pathway genes. A total of four biomarker genes, P4HA1, MLLT11, AURKA, and GOT1, were selected. The potential molecular mechanisms of the four biomarker genes involved in MRTK and their correlation with the tumor stem cell index were explored. Target drugs of AURKA and GOT1 were predicted, which need further experimental screening. Our study provides a reference for further understanding the possible pathogenesis in the prognosis of MRTK.

## DATA AVAILABILITY STATEMENT

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

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Kunming Children's Hospital before collection (approval number: 2020-03-106-K01). Written informed consent to participate in this study was provided by the participants' legal guardian/next-of-kin.

## AUTHOR CONTRIBUTIONS

CH organized the article writing and critically modified the manuscript. ZG modified the manuscript. HY drafted the manuscript and was responsible for the acquisition of data. KZ participated in the data analysis. CC and ZY contributed to the literature search. YC and BY checked and corrected the language. All authors read and approved the manuscript and agreed to be accountable for all aspects of the research in ensuring that the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## FUNDING

This study was supported by the Yunnan Education Department of Science Research Fund (No. 2020 J0228), Kunming City Health Science and Technology Talent "1000" Training Project [No. 2020-SW (Reserve)-112], Kunming Health and Health Commission Health Research Project (No. 2020-0201-001), and Kunming Medical Joint Project of Yunnan Science and Technology Department (No. 202001 AY070001-271). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## ACKNOWLEDGMENTS

The authors would like to thank the parents and children who enrolled in the study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.843234/full#supplementary-material

## REFERENCES

Balsara, B. R., Pei, J., Mitsuuchi, Y., Page, R., Klein-Szanto, A., Wang, H., et al. (2004). Frequent Activation of AKT in Non-small Cell Lung Carcinomas and Preneoplastic Bronchial Lesions. *Carcinogenesis* 25 (11), 2053–2059. doi:10.1093/carcin/bgh226

Begicevic, R.-R., and Falasca, M. (2017). ABC Transporters in Cancer Stem Cells: Beyond Chemoresistance. *Ijms* 18 (11), 2362. doi:10.3390/ijms18112362

Bergmann, L., Maute, L., and Guschmann, M. (2014). Temsirolimus for Advanced Renal Cell Carcinoma. *Expert Rev. Anticancer Ther.* 14 (1), 9–21. doi:10.1586/14737140.2014.864562

Bleau, A.-M., Hambardzumyan, D., Ozawa, T., Fomchenko, E. I., Huse, J. T., Brennan, C. W., et al. (2009). PTEN/PI3K/Akt Pathway Regulates the Side Population Phenotype and ABCG2 Activity in Glioma Tumor Stem-like Cells. *Cell Stem Cell* 4 (3), 226–235. doi:10.1016/j.stem.2009.01.007

Chakravarthi, B. V. S. K., Pathi, S. S., Goswami, M. T., Cieślik, M., Zheng, H., Nallasivam, S., et al. (2014). The miR-124-Prolyl Hydroxylase P4HA1-MMP1 axis Plays a Critical Role in Prostate Cancer Progression. *Oncotarget* 5 (16), 6654–6669. doi:10.18632/oncotarget.2208

Chang, L., Graham, P. H., Hao, J., Ni, J., Bucci, J., Cozzi, P. J., et al. (2013). Acquisition of Epithelial-Mesenchymal Transition and Cancer Stem Cell Phenotypes Is Associated with Activation of the PI3K/Akt/mTOR Pathway

in Prostate Cancer Radioresistance. *Cell Death Dis* 4, e875. doi:10.1038/cddis.2013.407

Corominas-Faja, B., Cufí, S., Oliveras-Ferraros, C., Cuyàs, E., López-Bonet, E., Lupu, R., et al. (2013). Nuclear Reprogramming of Luminal-like Breast Cancer Cells Generates Sox2-Overexpressing Cancer Stem-like Cellular States Harboring Transcriptional Activation of the mTOR Pathway. *Cell Cycle* 12 (18), 3109–3124. doi:10.4161/cc.26173

Degenhardt, K., Mathew, R., Beaudoin, B., Bray, K., Anderson, D., Chen, G., et al. (2006). Autophagy Promotes Tumor Cell Survival and Restricts Necrosis, Inflammation, and Tumorigenesis. *Cancer Cell* 10 (1), 51–64. doi:10.1016/j.ccr.2006.06.001

DeGraffenried, L. A., Friedrichs, W. E., Russell, D. H., Donzis, E. J., Middleton, A. K., Silva, J. M., et al. (2004). Inhibition of mTOR Activity Restores Tamoxifen Response in Breast Cancer Cells with Aberrant Akt Activity. *Clin. Cancer Res.* 10 (23), 8059–8067. doi:10.1158/1078-0432.CCR-04-0035

Duan, Y., Dong, Y., Dang, R., Hu, Z., Yang, Y., Hu, Y., et al. (2018). MiR-122 Inhibits Epithelial Mesenchymal Transition by Regulating P4HA1 in Ovarian Cancer Cells. *Cell Biol Int* 42 (11), 1564–1574. doi:10.1002/cbin.11052

Fan, Q.-W., and Weiss, W. A. (2012). Inhibition of PI3K-Akt-mTOR Signaling in Glioblastoma by mTORC1/2 Inhibitors. *Methods Mol. Biol.* 821, 349–359. doi:10.1007/978-1-61779-430-8_22

Feng, Z., Zhang, H., Levine, A. J., and Jin, S. (2005). The Coordinate Regulation of the P53 and mTOR Pathways in Cells. *Proc. Natl. Acad. Sci. U.S.A.* 102 (23), 8204–8209. doi:10.1073/pnas.0502857102

Francipane, M. G., and Lagasse, E. (2013). Selective Targeting of Human colon Cancer Stem-like Cells by the mTOR Inhibitor Torin-1. *Oncotarget* 4 (11), 1948–1962. doi:10.18632/oncotarget.1310

Francipane, M. G., and Lagasse, E. (2016). Therapeutic Potential of mTOR Inhibitors for Targeting Cancer Stem Cells. *Br. J. Clin. Pharmacol.* 82 (5), 1180–1188. doi:10.1111/bcp.12844

Gilkes, D. M., Bajpai, S., Chaturvedi, P., Wirtz, D., and Semenza, G. L. (2013). Hypoxia-inducible Factor 1 (HIF-1) Promotes Extracellular Matrix Remodeling under Hypoxic Conditions by Inducing P4HA1, P4HA2, and PLOD2 Expression in Fibroblasts. *J. Biol. Chem.* 288 (15), 10819–10829. doi:10.1074/jbc.M112.442939

Goldberg, S. L., Fenaux, P., Craig, M. D., Gyan, E., Lister, J., Kassis, J., et al. (2014). An Exploratory Phase 2 Study of Investigational Aurora A Kinase Inhibitor Alisertib (MLN8237) in Acute Myelogenous Leukemia and Myelodysplastic Syndromes. *Leuk. Res. Rep.* 3 (2), 58–61. doi:10.1016/j.lrr.2014.06.003

Gruber, E., Oberhuber, G., Birner, P., Schlederer, M., Kenn, M., Schreiner, W., et al. (2019). The Oncogene AF1Q Is Associated with WNT and STAT Signaling and Offers a Novel Independent Prognostic Marker in Patients with Resectable Esophageal Cancer. *Cells* 8 (11), 1357. doi:10.3390/cells8111357

Harrington, L. S., Findlay, G. M., Gray, A., Tolkacheva, T., Wigfield, S., Rebholz, H., et al. (2004). The TSC1-2 Tumor Suppressor Controls Insulin-Pi3k Signaling via Regulation of IRS Proteins. *J. Cel Biol* 166 (2), 213–223. doi:10.1083/jcb.200403069

Hensley, C. T., Wasti, A. T., and DeBerardinis, R. J. (2013). Glutamine and Cancer: Cell Biology, Physiology, and Clinical Opportunities. *J. Clin. Invest.* 123 (9), 3678–3684. doi:10.1172/JCI69600

Hsu, P. P., Kang, S. A., Rameseder, J., Zhang, Y., Ottina, K. A., Lim, D., et al. (2011). The mTOR-Regulated Phosphoproteome Reveals a Mechanism of mTORC1-Mediated Inhibition of Growth Factor Signaling. *Science* 332 (6035), 1317–1322. doi:10.1126/science.1199498

Karthik, G.-M., Ma, R., Lövrot, J., Kis, L. L., Lindh, C., Blomquist, L., et al. (2015). mTOR Inhibitors Counteract Tamoxifen-Induced Activation of Breast Cancer Stem Cells. *Cancer Lett.* 367 (1), 76–87. doi:10.1016/j.canlet.2015.07.017

Katsuno, Y., Meyer, D. S., Zhang, Z., Shokat, K. M., Akhurst, R. J., Miyazono, K., et al. (2019). Chronic TGF-β Exposure Drives Stabilized EMT, Tumor Stemness, and Cancer Drug Resistance with Vulnerability to Bitopic mTOR Inhibition. *Sci. Signal.* 12 (570), eaau8544. doi:10.1126/scisignal.aau8544

Kim, L. C., Cook, R. S., and Chen, J. (2017). mTORC1 and mTORC2 in Cancer and the Tumor Microenvironment. *Oncogene* 36 (16), 2191–2201. doi:10.1038/onc.2016.363

Kim, Y. C., and Guan, K.-L. (2015). mTOR: a Pharmacologic Target for Autophagy Regulation. *J. Clin. Invest.* 125 (1), 25–32. doi:10.1172/JCI73939

Kolev, V. N., Wright, Q. G., Vidal, C. M., Ring, J. E., Shapiro, I. M., Ricono, J., et al. (2015). PI3K/mTOR Dual Inhibitor VS-5584 Preferentially Targets Cancer Stem Cells. *Cancer Res.* 75 (2), 446–455. doi:10.1158/0008-5472.CAN-14-1223

Kollareddy, M., Zheleva, D., Dzubak, P., Brahmkshatriya, P. S., Lepsik, M., and Hajduch, M. (2012). Aurora Kinase Inhibitors: Progress towards the Clinic. *Invest. New Drugs* 30 (6), 2411–2432. doi:10.1007/s10637-012-9798-6

Laplante, M., and Sabatini, D. M. (2012). mTOR Signaling in Growth Control and Disease. *Cell* 149 (2), 274–293. doi:10.1016/j.cell.2012.03.017

Martelli, A. M., Chiarini, F., Evangelisti, C., Cappellini, A., Buontempo, F., Bressanin, D., et al. (2012). Two Hits Are Better Than One: Targeting Both Phosphatidylinositol 3-kinase and Mammalian Target of Rapamycin as a Therapeutic Strategy for Acute Leukemia Treatment. *Oncotarget* 3 (4), 371–394. doi:10.18632/oncotarget.477

Marumoto, T., Zhang, D., and Saya, H. (2005). Aurora-A - a Guardian of Poles. *Nat. Rev. Cancer* 5 (1), 42–50. doi:10.1038/nrc1526

Matsubara, S., Ding, Q., Miyazaki, Y., Kuwahata, T., Tsukasa, K., and Takao, S. (2013). mTOR Plays Critical Roles in Pancreatic Cancer Stem Cells through Specific and Stemness-Related Functions. *Sci. Rep.* 3, 3230. doi:10.1038/srep03230

Nishitani, S., Horie, M., Ishizaki, S., and Yano, H. (2013). Branched Chain Amino Acid Suppresses Hepatocellular Cancer Stem Cells through the Activation of Mammalian Target of Rapamycin. *Plos One* 8 (11), e82346. doi:10.1371/journal.pone.0082346

Otto, T., Horn, S., Brockmann, M., Eilers, U., Schüttrumpf, L., Popov, N., et al. (2009). Stabilization of N-Myc Is a Critical Function of Aurora A in Human Neuroblastoma. *Cancer Cell* 15 (1), 67–78. doi:10.1016/j.ccr.2008.12.005

Qu, X., Yu, J., Bhagat, G., Furuya, N., Hibshoosh, H., Troxel, A., et al. (2003). Promotion of Tumorigenesis by Heterozygous Disruption of the Beclin 1 Autophagy Gene. *J. Clin. Invest.* 112 (12), 1809–1820. doi:10.1172/JCI20039

Shah, O. J., Wang, Z., and Hunter, T. (2004). Inappropriate Activation of the TSC/Rheb/mTOR/S6K Cassette Induces IRS1/2 Depletion, Insulin Resistance, and Cell Survival Deficiencies. *Curr. Biol.* 14 (18), 1650–1656. doi:10.1016/j.cub.2004.08.026

Srivastava, R. K., Li, C., Khan, J., Banerjee, N. S., Chow, L. T., and Athar, M. (2019). Combined mTORC1/mTORC2 Inhibition Blocks Growth and Induces Catastrophic Macropinocytosis in Cancer Cells. *Proc. Natl. Acad. Sci. U.S.A.* 116 (49), 24583–24592. doi:10.1073/pnas.1911393116

Sunayama, J., Matsuda, K.-I., Sato, A., Tachibana, K., Suzuki, K., Narita, Y., et al. (2010). Crosstalk between the PI3K/mTOR and MEK/ERK Pathways Involved in the Maintenance of Self-Renewal and Tumorigenicity of Glioblastoma Stem-like Cells. *Stem Cells* 28 (11), 1930–1939. doi:10.1002/stem.521

Tahir, F., Majid, Z., Qadar, L. T., Abbas, A., and Raza, M. (2019). Co-existent Rhabdoid Tumor of the Kidney and Brain in a Male Infant: A Rare Case. *Cureus* 11 (8), e5423. doi:10.7759/cureus.5423

Tasian, S. K., Teachey, D. T., and Rheingold, S. R. (2014). Targeting the PI3K/mTOR Pathway in Pediatric Hematologic Malignancies. *Front. Oncol.* 4, 108. doi:10.3389/fonc.2014.00108

Tse, W., Zhu, W., Chen, H., and Cohen, A. (1995). A Novel Gene, AF1q, Fused to MLL in T(1;11) (Q21;q23), Is Specifically Expressed in Leukemic and Immature Hematopoietic Cells. *Blood* 85 (3), 650–656. doi:10.1182/blood.v85.3.650.bloodjournal853650

Um, S. H., Frigerio, F., Watanabe, M., Picard, F., Joaquin, M., Sticker, M., et al. (2004). Absence of S6K1 Protects against Age- and Diet-Induced Obesity while Enhancing Insulin Sensitivity. *Nature* 431 (7005), 200–205. doi:10.1038/nature02866

Wang, H., Tang, D., Wu, D., Tao, C., Chen, G., Ru, W., et al. (2020). Renal Malignant Tumor with the Loss of INI1 Expression and WT1 Positive in a 3-Year-Old Boy: a Case Report. *Transl Androl. Urol.* 9 (5), 2275–2280. doi:10.21037/tau-20-834

White, E. (2015). The Role for Autophagy in Cancer. *J. Clin. Invest.* 125 (1), 42–46. doi:10.1172/JCI73941

Xiong, Y., Li, Z., Ji, M., Tan, A.-C., Bemis, J., Tse, J.-V., et al. (2011). MIR29B Regulates Expression of MLLT11 (AF1Q), an MLL Fusion Partner, and Low MIR29B Expression Associates with Adverse Cytogenetics and Poor Overall Survival in AML. *Br. J. Haematol.* 153 (6), 753–757. doi:10.1111/j.1365-2141.2011.08662.x

Yanagisawa, S., Kadouchi, I., Yokomori, K., Hirose, M., Hakozaki, M., Hojo, H., et al. (2009). Identification and Metastatic Potential of Tumor-Initiating Cells in Malignant Rhabdoid Tumor of the Kidney. *Clin. Cancer Res.* 15 (9), 3014–3022. doi:10.1158/1078-0432.CCR-08-2237

Yang, Y. (2016). Enhancing Doxorubicin Efficacy through Inhibition of Aspartate Transaminase in Triple-Negative Breast Cancer Cells. *Biochem. Biophysical Res. Commun.* 473 (4), 1295–1300. doi:10.1016/j.bbrc.2016.04.061

Yu, Y., Yoon, S.-O., Poulogiannis, G., Yang, Q., Ma, X. M., Villén, J., et al. (2011). Phosphoproteomic Analysis Identifies Grb10 as an mTORC1 Substrate that Negatively Regulates Insulin Signaling. *Science* 332 (6035), 1322–1326. doi:10.1126/science.1199484

Yue, Z., Jin, S., Yang, C., Levine, A. J., and Heintz, N. (2003). Beclin 1, an Autophagy Gene Essential for Early Embryonic Development, Is a Haploinsufficient Tumor Suppressor. *Proc. Natl. Acad. Sci. U.S.A.* 100 (25), 15077–15082. doi:10.1073/pnas.2436255100

Zhang, Y., Tseng, J. T.-C., Lien, I.-C., Li, F., Wu, W., and Li, H. (2020). mRNAsi Index: Machine Learning in Mining Lung Adenocarcinoma Stem Cell Biomarkers. *Genes* 11 (3), 257. doi:10.3390/genes11030257

Zhanghuang, C., Chen, S., Li, L., Yang, Z., Xie, Y., Li, J., et al. (2021). Clinical and Molecular Differentiation between Malignant Rhabdoid Tumor of the Kidney and Normal Tissue: A Two-Case Report. *Front. Oncol.* 11, 659709. doi:10.3389/fonc.2021.659709

Zhou, Y., Jin, G., Mi, R., Zhang, J., Zhang, J., Xu, H., et al. (2017). Knockdown of P4HA1 Inhibits Neovascularization via Targeting Glioma Stem Cell-Endothelial Cell Transdifferentiation and Disrupting Vascular Basement Membrane. *Oncotarget* 8 (22), 35877–35889. doi:10.18632/oncotarget.16270

Zhu, Y., Zhang, X., Liu, Y., Zhang, S., Liu, J., Ma, Y., et al. (2012). Antitumor Effect of the mTOR Inhibitor Everolimus in Combination with Trastuzumab on Human Breast Cancer Stem Cells *In Vitro* and *In Vivo*. *Tumor Biol.* 33 (5), 1349–1362. doi:10.1007/s13277-012-0383-6

Zoncu, R., Efeyan, A., and Sabatini, D. M. (2011). mTOR: from Growth Signal Integration to Cancer, Diabetes and Ageing. *Nat. Rev. Mol. Cel Biol* 12 (1), 21–35. doi:10.1038/nrm3025

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

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

frontiers | Frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 15 June 2022
doi: 10.3389/fmolb.2022.816102

# Expression of ALG3 in Hepatocellular Carcinoma and Its Clinical Implication

*Zhen Zhao*[1,2†], *Zehao Zheng*[2,3†], *Jianfeng Huang*[1,2], *Jianxi Wang*[2,4], *Tianyi Peng*[2,3], *Ye Lin*[2*] *and Zhixiang Jian*[2*]

[1]School of Medicine, South China University of Technology, Guangzhou, China, [2]Department of General Surgery, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China, [3]Shantou University of Medical College, Shantou, China, [4]The Second School of Clinical Medicine, Southern Medical University, Guangzhou, China

**Background:** Recent studies have shown that alpha-1,3-mannosyltransferase (ALG3) promoted tumorigenesis and progression in multiple cancer types. Our study planned to explore the clinical implication and potential function of ALG3 in hepatocellular carcinoma.

**Materials and Methods:** Data from public databases were used to analyze the ALG3 expression and its impact on the clinical significance of patients with HCC. The ALG3 expression was confirmed by qRT-PCR and Western blot. Immunohistochemistry was used to confirm the ALG3 expression and explore its clinical implication in HCC. KEGG, GO, and GSEA enrichment analyses were utilized to explore the biological pathways related to ALG3 in HCC. TIMER2.0 was applied to assess the association between ALG3 and immune infiltration. CCK8, MTT, and transwell assays were used to investigate the role of ALG3 downregulation in HCC cell lines.

**Results:** qRT-PCR, WB, and IHC proved ALG3 was highly overexpressed in HCC tissues. The Kaplan–Meier analysis verified the overexpression of ALG3 was related to poor overall survival ($p < 0.001$). Multivariate cox regression analysis showed that the high ALG3 expression was an independent risk prognostic factor. GSEA and TIMER2.0 predicted that ALG3 participates in cell differentiation and cycle and correlates with immune cell infiltration. Transwell assay results showed that ALG3 silencing also impaired the invasion ability of HCC cells.

**Conclusion:** ALG3 was overexpressed and considered a potential indicator of survival in HCC, and our findings provided a novel therapeutic target for HCC.

**Keywords: ALG3, hepatocellular carcinoma, immune infiltration, prognosis, clinical implication**

OPEN ACCESS

***Edited by:***
*Liang Wang,*
*Moffitt Cancer Center & Research Institute, United States*

***Reviewed by:***
*Xiaoyi Huang,*
*Harbin Medical University Cancer Hospital, China*
*Udayan Bhattacharya,*
*Weill Cornell Medical Center, NewYork-Presbyterian, United States*

***Correspondence:**
*Ye Lin*
*linye@gdph.org.cn*
*Zhixiang Jian*
*jzx_118@163.com*

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

***Specialty section:***
*This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences*

***Received:*** *16 November 2021*
***Accepted:*** *09 May 2022*
***Published:*** *15 June 2022*

***Citation:***
*Zhao Z, Zheng Z, Huang J, Wang J, Peng T, Lin Y and Jian Z (2022) Expression of ALG3 in Hepatocellular Carcinoma and Its Clinical Implication. Front. Mol. Biosci. 9:816102. doi: 10.3389/fmolb.2022.816102*

## INTRODUCTION

Liver cancer is the second most common cause of cancer-related mortality (Bray et al., 2018), and the highest incidence rates are found in Asia and Africa (McGlynn et al., 2021). As one of the top five deadliest cancers in the world, the incidence of liver cancer has increased annually and is an increasingly frequent cause of cancer death (Siegel et al., 20192019). Hepatocellular carcinoma (HCC) is the primary type of liver cancer, accounting for about 75% of the total cancer deaths (Petrick et al., 2020). At an early stage, surgical resection and organ transplantation remain the only curative options for HCC, but the 5-year recurrence rate is around 50% (Pinato et al., 2017) due to the high proportion of metastasis (Lu et al., 2019). The lack of survival benefits from conventional drugs

and sorafenib suggests the urgent need to find new biological therapeutic strategies (Faivre et al., 2020).

Following the successful application of immune checkpoint blockers (ICBs) in several tumor types, lots of studies are researching ICBs alone or in combination with other treatments in patients with HCC (Pinter et al., 2021). Currently, another four targeted therapies have been approved for HCC, except for sorafenib: lenvatinib (Zhao et al., 2020), regorafenib (Grothey et al., 2020), cabozantinib (Shang et al., 2021), and ramucirumab (Zhu et al., 2019), but immunotherapy is only effective in a subset of HCC patients (Ruf et al., 2021). Hence, a deeper understanding of the molecular mechanisms and biological pathways involved in the development of HCC will be conducive to exploring new biomarkers and improving survival time in HCC patients.

Glycosylation is involved in cell–matrix interaction, recognition, signal recognition, and synaptic signaling and plays an important role in many cellular processes (Varki, 2017). Targeting glycosylation is considered as a new way of cancer drug discovery (Costa et al., 2020). The different expressions of glycosyltransferases in several cancer types have been verified as a potential indicator of survival and treatment targets (Song et al., 2020). Alpha-1,3-mannosyltransferase (ALG3) encodes an endoplasmic reticulum-localized ALG3 family member that participates in N-glycan synthesis (Bloch et al., 2020). Researchers have observed that ALG3 contributes to high mannose-type N-glycans in several cancers, and mannose-type N-glycan upregulation has been shown to be related to cancer progression (Silva et al., 2020). Studies have demonstrated that ALG3 is overexpressed in oral squamous cell carcinoma (Shao et al., 2021), non–small cell lung cancer (Ke et al., 2020), and breast cancer (Yang et al., 2018; Sun et al., 2021) and contributes to drug resistance in the acute myeloid leukemia (Liu et al., 2020a). However, the expression of ALG3 in HCC is unknown, and its prognosis values and correlations with clinicopathologic features of HCC patients need to be further explored.

In our study, public databases were first utilized to identify the mRNA expression of ALG3 and its relation to the prognosis. Then, clinical cases, tissues, and cell lines were applied to prove the ALG3 expression and the relationship between ALG3 and clinicopathologic characteristics. Functional enrichment analysis and TIMER2.0 were utilized to explore the potential mechanisms and their relations with immune infiltrations.

# MATERIALS AND METHODS

## Cell Culture

The human liver cell line LO2 and hepatocellular cell lines SNU398, SNU449, MHCC97H, Hep3B, PLC/PRF/5, and HepG2 were purchased from the ATCC (Manassas, VA, United States). PLC/PRF/5, Hep3B, MHCC97H, and HepG2 cell lines were cultured in complete Dulbecco's modified Eagle's medium (Gibco, United States), and SNU398 and SNU449 cells were cultured in RPMI1640 medium (Gibco, United States); all cells were supplemented with 10% fetal bovine serum (FBS) at 37°C in a humidified incubator with 5% $CO_2$ in the air.

HCC cell lines (Hep3B, HepG2, and SNU398) were transfected with siRNAs targeting ALG3 or control siRNA (HIPPOBIO, China) using Lipofectamine 2000 Reagent (Invitrogen, United States). The efficiency of the knockdown was confirmed using qRT-PCR analysis. The siRNA used were as follows: si-NC: 5′-UUCUUCGAAGGUGUCACGUTT-3′, ALG3 siRNA#1: 5′-GGUUUCGUGUACAUCUUUAUG-3′, and ALG3 siRNA#2: 5′-GGACCUGAGUCUACCCUCAGG-3'.

## Patients and Specimens

Eight fresh tissue specimens of HCC and 115 paraffin-embedded specimens, which were pathologically diagnosed with HCC and obtained from Guangdong Provincial People's Hospital (GDPH) (Guangdong, China), were enrolled in this study. The intrahepatic recurrence and extrahepatic metastasis were both considered as recurrences. The use of these clinical materials for our study was approved by the Research Ethics Committee of (GDPH) and by each patient. The clinicopathologic features are summarized in **Supplementary Table S1**.

## Tumor IMmune Estimation Resource Database

The online database "Tumor IMmune Estimation Resource" (TIMER2.0; https://cistrome.shinyapps.io/timer/) was used to assess the expression levels of ALG3 mRNA in different cancer patients and the relationship between the ALG3 expression and immune cells in HCC (Li et al., 2020a).

## Human Protein Atlas Database

The HPA (https://www.proteinatlas.org/) is a public server, including patients' proteomic and RNA-seq data in every single-cell, organ, and tissue from human normal or abnormal tissues. Thus, ALG3 was analyzed in single-cell types and single-cell lines by using the search term "ALG3" (Thul and Lindskog, 2018).

## Data Acquisition

Gene expression data were acquired from The Cancer Genome Atlas (TCGA) database up to May 2021 (https://portal.gdc.cancer.gov). The GSE14520 and GSE124535 datasets were downloaded from the Gene Expression Omnibus (GEO) database (https://ncbi.nlm.nih.gov/geo). The **Supplementary Table S3** dataset was obtained from HCCDB (Lian et al., 2018) for expression validation.

## RNA Extraction, Reverse Transcription, and Quantitative Real-Time PCR

All RNA samples from the HCC cell lines and clinical patient's tissues were extracted using the TRIzol reagent (Invitrogen, Carlsbad, CA, United States). For the procedure, 1 μg RNA was used for cDNA synthesis primed with random hexamers. cDNAs were amplified by SYBR-Green in a CFX96 real-time system C1000 cycler. Results of amplification were downloaded and analyzed with EXCEL, and GAPDH was used for internal control. Each sample was tested in triplicate at least three times. All fold changes were calculated through relative quantification 2^ [(GAPDHCq)—(ALG3Cq)]. The detailed primers sequences are provided in **Supplementary Table S2**.

## Western Blot

Tissue sample and cell lines (LO2, SNU449, SNU398, PLC/PRF/5, MHC97-H, Hep3B, and HepG2) of total protein were obtained in lysis buffer. Protein density was measured using the BCA test kit (Fdbio Science, China), according to the manufacturer's suggested protocols. Equal amounts (40 μg) of protein were loaded to SDS–PAGE and transferred to the NC membrane. After being blocked with 5% milk for 45 min, the membranes were incubated overnight at 4° with antibodies against ALG3 (1: 1,000 dilution; NO. 20290–1-AP, Proteintech, Manchester, United Kingdom) and GAPDH (1:1,000 dilution; CST, Shanghai, China); next day, the membranes were washed five times for 5 min with TBST (2% Tween) and then incubated with anti-rabbit IgG for 45 min at room temperature. The expression of the target protein was detected by chemiluminescence.

## Immunohistochemistry Analysis

IHC staining was applied on 115 paraffin sections of clinical HCC specimens using the ALG3 antibody (Sigma, HPA045130, 1:200). Two pathologists independently assessed the specimens, who were blinded to the details, on the basis of staining degree and the proportion of positive tumor cells. The percentage of positive cells was scored as follows: 0: <5%; 1: 5%–25%; 2: 5%–50%; 3: 50%–75%; and 4: >75%. The immunostaining intensity grade was as follows: 0 (negative), 1 (weak), 2(moderate), and 3(strong). The staining degree was multiplied by the percentage of positive cells as the staining score. In this study, an ALG3 score ≥8 was defined as a high expression.

## Functional and Pathway Enrichment Analysis

The STRING (https://string-db.org/) database was applied to construct a protein–protein interaction (PPI) network for ALG3. R language was used to visualize the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. $p < 0.05$ was considered significant.

## Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (GSEA) (http://www.gsea-msigdb.org/gsea/index.jsp) was used to evaluate the potential biological mechanisms. $p < 0.05$ and FDR<0.25 were the criteria to determine a significant differential expression.

## Cell Proliferation Assay

The cellular proliferation was evaluated by the CCK-8 (Beyotime, China) and MTT (Sigma-Aldrich, United States) assays, according to the manufacturer's instructions. Briefly, the cells successfully transfected with ALG3 siRNA or control siRNA were seeded in 96-well plates, with three replicate wells in each group. The optical density (OD) values of the cells were measured every 24 h. Before the measurement, 10 μL of CCK8 or MTT (5 mg/ml) solution was added to the medium and incubated for 2 h. The OD values were measured under the absorbance of 450 or 490 nm by using a microplate reader (Bio-Rad Laboratories, Hercules, CA, United States). Finally, a growth curve was made by GraphPad Prism 8.

## Cell Invasion Assay

Transwell chambers (8 μM, Corning Costar, China) were used to detect the invasion abilities of Hep3B cells. The upper chamber was pre-coated with Matrigel to form a matrix barrier, and $2 \times 10^4$ cells suspended in 200 μl serum-free DMEM medium were seeded in triplicate into the upper chamber of 24-well chambers. The under compartment included a 600 μl DMEM medium with 10% FBS. After cultivation in a 37°C incubator for 24 h, the chambers were rinsed with PBS gently to remove the cells glued to the upper surface. After being fixed and stained, the cells were photographed in three fields randomly under a microscope.

## Statistical Analysis

SPSS version 25.0 and GraphPad Prism 8 (San Diego, CA, United States) were used for statistical analyses. Student's test, Kaplan–Meier survival curves were plotted by GraphPad Prism 8.0.1 to analyze the difference between the two groups. Cox regression analysis was performed to identify the independent prognostic factors of patients with hepatocellular carcinoma. $p < 0.05$ indicates a significant difference.

# RESULTS

## Elevated Expression of Alpha-1,3-Mannosyltransferase in Variety of Malignant Tumors

We chose TIMER2.0 and HPA cancer databases to investigate the expression of ALG3 mRNA. As shown in **Figure 1A**, our result showed that ALG3 was significantly expressed in multiple cancer types compared with normal samples, including BLCA ($p$ = 8.82E-12), BRCA ($p$ = 2.72E-55), CHOL ($p$ = 4.51E-09), COAD ($p$ = 4.29E-23), ESCA ($p$ = 6.73E-08), HNSC ($p$ = 5.54E-25), KIRC ($p$ = 0.007), KIRP ($p$ = 1.95E-13), LIHC ($p$ = 1.55E-26), LUAD ($p$ = 8.58E-32), LUSC ($p$ = 1.02E-30), PRAD ($p$ = 7.86E-12), READ ($p$ = 5.17E-07), STAD ($p$ = 2.14E-19), and UCEC ($p$ = 2.29E-19). In the HPA public database, the mRNA expression of ALG3 was upregulated in germ cells, trophoblast cells, blood and immune cells, and epithelial cells. Importantly, ALG3 was highly expressed in hepatocytes, as shown in **Figure 1B**. In short, our results suggested that ALG3 was overexpressed in different cancers including liver hepatocellular carcinoma tissues and liver cell lines.

## Alpha-1,3-Mannosyltransferase Is Markedly Upregulated in Hepatocellular Carcinoma

To research the role of ALG3 in hepatocellular carcinoma, first, we explored the ALG3 expression on online public human hepatocellular carcinoma datasets of the GEO and TCGA databases. Obviously, the result revealed that compared with normal controls, ALG3 was indeed upregulated in HCC tissues and in paired HCC tissues (**Figures 2A,B**). Other data sets from HCDDB also demonstrated the differential expression of ALG3 in HCC (**Supplementary Table S3**). To further confirm both mRNA and protein expression levels of the ALG3 expression in HCC, qRT-

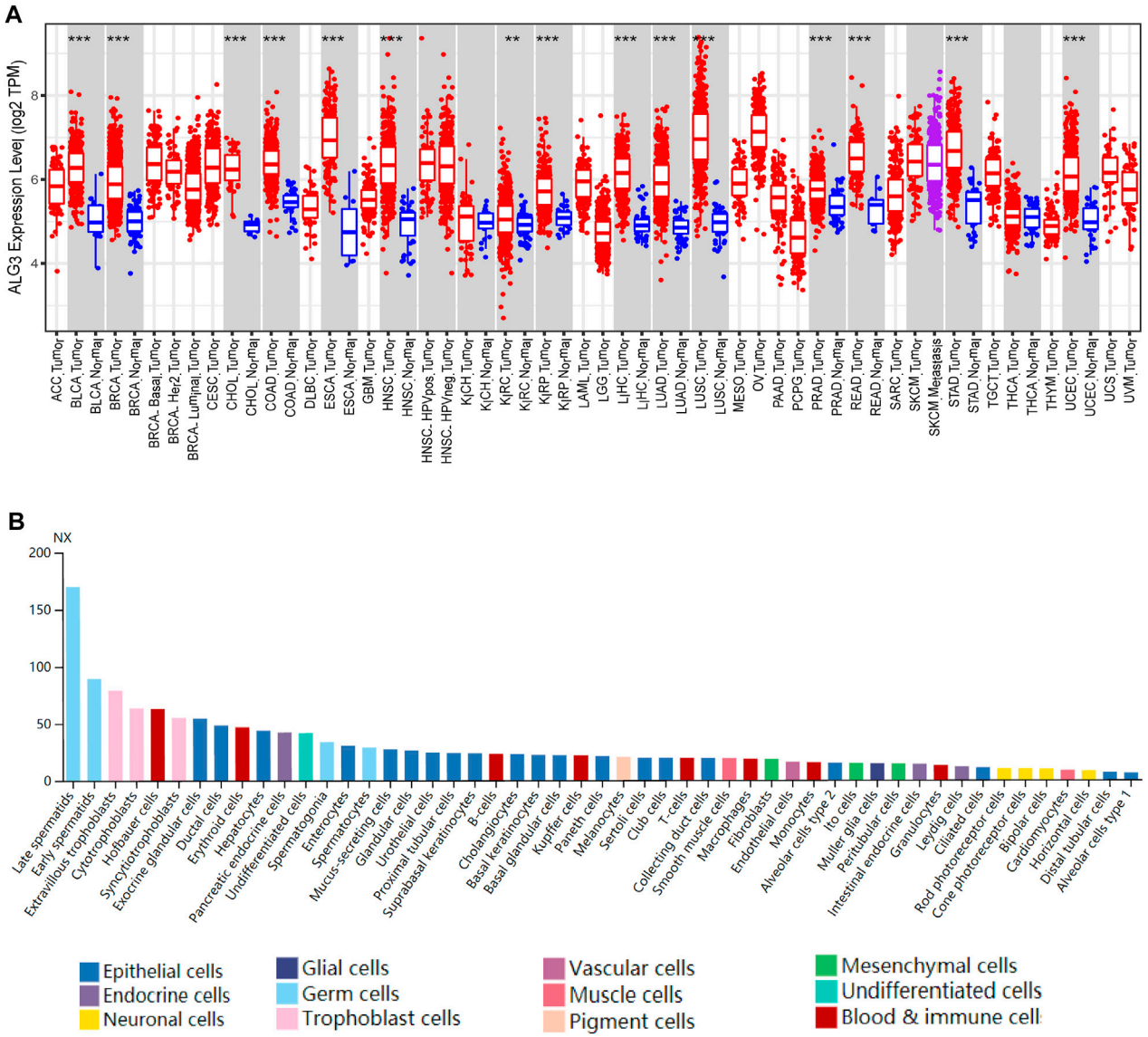



**FIGURE 1 |** Expression of ALG3 in different cancer types. **(A)** ALG3 mRNA expression levels across different cancer types in the TIMER database. **(B)** ALG3 expression in single cell type lines analyzed by HPA. (**$p < 0.01$ and ***$p < 0.001$) ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBCL, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe carcinoma; KIRC, clear cell renal cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PRAD, prostate adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; READ, rectal adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumor; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma.

PCR and Western blot were conducted for the specimens from our hospital. The results showed that ALG3 was significantly overexpressed in HCC tissues compared with non-tumor adjacent tissues (**Figures 2C,D**). In addition, as shown in **Figures 2E,F**, the differential expression of ALG3 has been further verified in liver cell lines. Therefore, these results indicate that the ALG3 expression greatly increases in HCC.

## Relationships Between Alpha-1,3-Mannosyltransferase and Clinicopathologic Features of Hepatocellular Carcinoma Patients

We performed a subgroup analysis of TCGA data based on clinicopathologic characteristics. Importantly, the ALG3

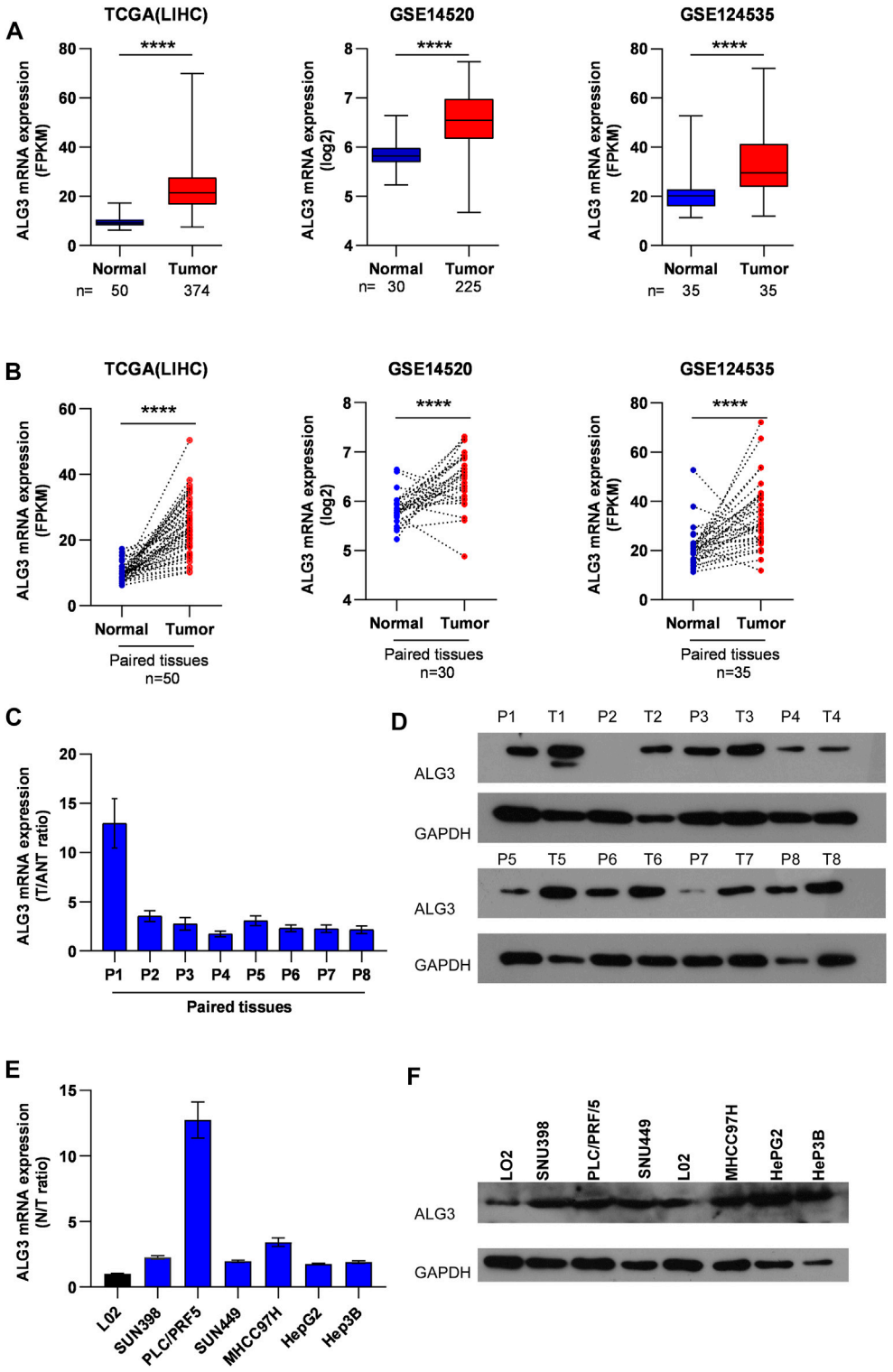



**FIGURE 2 |** Upregulation of ALG3 in HCC. **(A)** ALG3 mRNA expression levels were markedly overexpressed in unpaired HCC tissues, as indicated by the TCGA LIHC dataset and GEO HCC datasets GSE14520 and GSE124535. **(B)** ALG3 mRNA was upregulated in paired HCC tissues and adjacent normal tissues in the TCGA LIHC dataset and GSE14520 and GSE142535 datasets. RT-PCR **(C,E)** and WB **(D,F)** analysis of ALG3 expression in paired HCC tissues and non-tumor adjacent tissues and in the liver cancer cell line. GAPDH was used as a loading control. (**$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$)

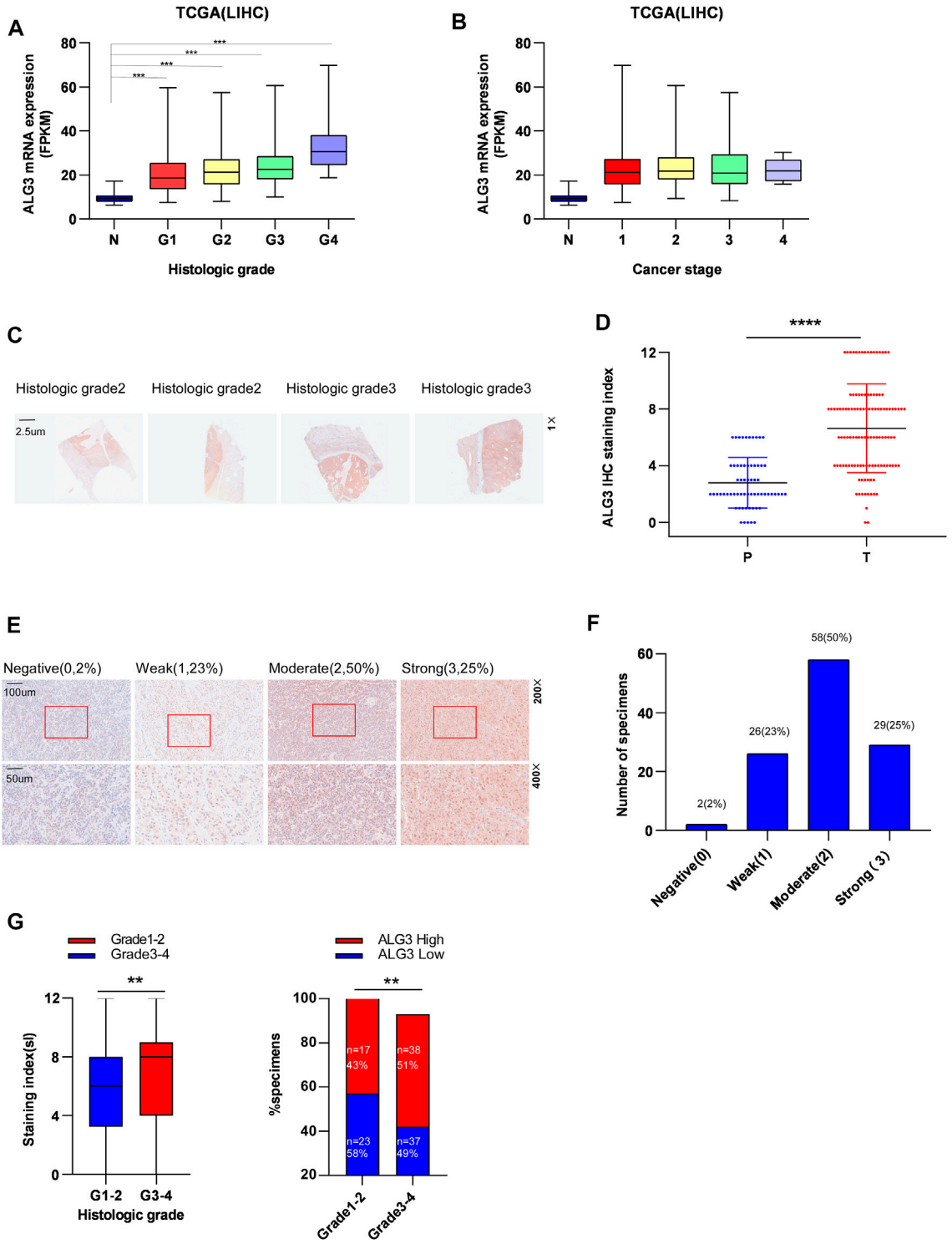



**FIGURE 3 |** Relationship between ALG3 and clinicopathologic features of HCC patients. The expression of ALG3 is ordered by histologic grade **(A)** and cancer stage **(B)** in TCGA. **(C)** Representative images of ALG3 staining in paired HCC tissues. **(D)** Detailed IHC staining index. **(E)** IHC analysis of ALG3 protein expression in HCC tissues of different representative images. The bottom shows a double magnification of part of the aforementioned image. **(F)** Number of cases with different staining intensities. **(G)** ALG3 expression levels in different histologic grade specimens. (**$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$).

**TABLE 1 |** Association between the ALG3 expression and other clinicopathologic features.

| Feature of HCC | No. of patients (%) | ALG3 expression | | *p*-values |
|---|---|---|---|---|
| | | Low expression (60) | High expression (55) | |
| Age (years) | | | | 0.047# |
| ≤60 | 60 | 26 | 34 | |
| >60 | 55 | 34 | 21 | |
| Gender | | | | 0.930# |
| Male | 79 | 41 | 38 | |
| Female | 36 | 19 | 17 | |
| Histology classification | | | | 0.404# |
| I–II | 40 | 23 | 17 | |
| III–IV | 75 | 37 | 38 | |
| TNM stage (AJCC) | | | | 0.112# |
| T1–T2 | 89 | 50 | 39 | |
| T3–T4 | 26 | 10 | 16 | |
| Vital status | | | | 0.002# |
| Alive | 64 | 25 | 39 | |
| Dead | 51 | 35 | 16 | |
| Recurrent status | | | | 0.473# |
| Yes | 79 | 43 | 36 | |
| No | 36 | 17 | 19 | |
| Tumor size (cm) | | | | 0.827# |
| <5 | 43 | 23 | 20 | |
| ≥5 | 72 | 37 | 35 | |
| Tumor number | | | | 0.181# |
| Single | 90 | 44 | 46 | |
| Multiple | 25 | 16 | 9 | |
| Tumor capsular | | | | 0.585# |
| Complete | 96 | 49 | 47 | |
| Incomplete | 19 | 11 | 8 | |
| Cut edge | | | | 1* |
| Negative | 107 | 56 | 51 | |
| Positive | 8 | 4 | 4 | |
| Vascular tumor emboli | | | | 0.343# |
| Present | 45 | 21 | 24 | |
| Absent | 70 | 39 | 31 | |
| Serum AFP (ng/ml) | | | | 0.797# |
| <400 | 76 | 39 | 37 | |
| ≥400 | 39 | 21 | 18 | |
| HBV DNA | | | | 0.473# |
| <500 | 79 | 43 | 36 | |
| ≥500 | 36 | 17 | 19 | |
| HBsAg | | | | 0.160# |
| Negative | 41 | 25 | 16 | |
| Positive | 74 | 35 | 39 | |
| Postoperative complication | | | | 0.335* |
| No | 102 | 51 | 51 | |
| Postoperative hemorrhage | 4 | 3 | 1 | |
| Bile leakage | 3 | 3 | 0 | |
| Liver failure | 1 | 0 | 1 | |
| Others | 5 | 3 | 2 | |
| Postoperative adjuvant therapy | | | | 0.582* |
| HIPEC | 8 | 6 | 2 | |
| TACE | 39 | 19 | 20 | |
| Sorafenib | 3 | 2 | 1 | |
| No | 65 | 33 | 32 | |

*TBX3 immunohistochemical score ≥8 was regarded as high expression.*
#p *values and* *p *values were calculated with the chi-squared test and Fisher's exact test, respectively.*

expression based on tumor grade and cancer stage performed significant differences (**Figures 3A,B**). To better understand how ALG3 protein affected the prognosis of HCC patients, we further performed IHC staining in 115 paraffins obtained from GDPH. The baseline clinicopathologic information of 155 HCC patients is showed in **Supplementary Table S1**, and the association of the ALG3 expression with other pathological features is showed in **Table 1**. As shown in **Figures 3C,D**, the tumor and non-tumor

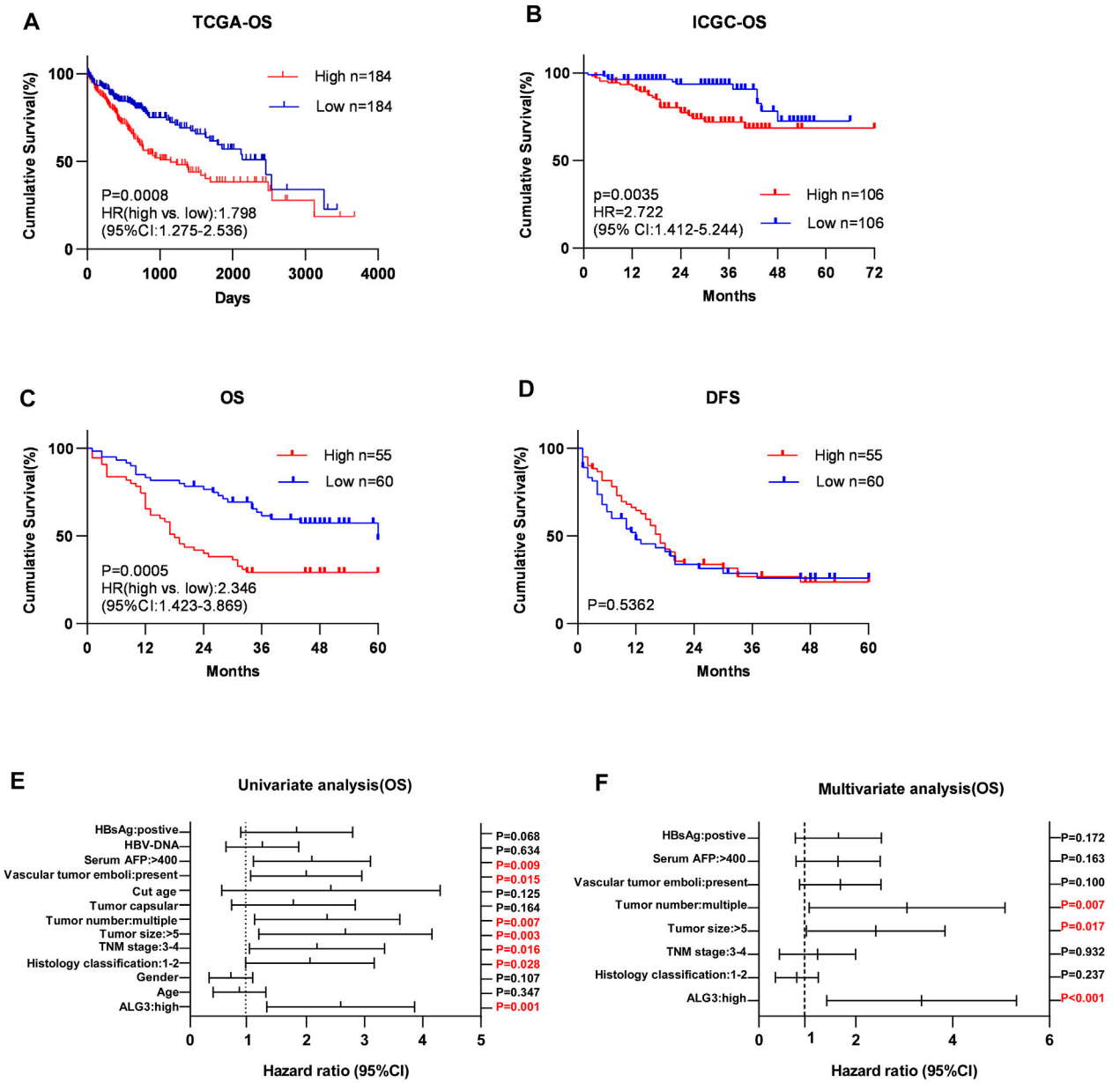



**FIGURE 4 |** Over-expression of ALG3 is correlated with poor prognosis in HCC. Kaplan–Meier analysis of 5-year overall survival for HCC in the TCGA-LIHC dataset **(A)** and ICGC dataset **(B)**. Kaplan–Meier analysis of OS **(C)** and DFS **(D)** for GHPD stratified by low and high ALG3 expressions ($n$ = 115, log-rank test). Univariate**(E)** and multivariate**(F)** Cox regression analyses to assess the significance of the association between ALG3 and 5-year overall survival in the presence of other clinical variables.

tissues were notably in different degrees of staining, and ALG3 staining was stronger in advanced-grade versus early-grade carcinoma. The representative images of different ALG3 protein expressions are shown in **Figure 3E**. The detailed case information with all staining intensities is displayed in **Figures 3D,F**. IHC results revealed that ALG3 protein was mainly localized in the cytoplasm. Subgroup analysis showed that the ALG3 expression was higher in histologic grades III- IV than in histologic grades I and II (**Figure 3G**) and was obviously associated with the histologic grade ($p$ = 0.0097).

## Overexpression of Alpha-1,3-Mannosyltransferase Is Correlated With Poor Prognosis in Hepatocellular Carcinoma

Kaplan–Meier survival curves for HCC in TCGA, LIHC, and ICGC datasets indicated that the overall survival of HCC patients with a low expression of ALG3 was longer than that of high expression (**Figures 4A,B** ). According to IHC staining results and the grouping criteria, the specimens were divided into high-

**TABLE 2 |** Univariate and multivariate analysis of overall survival between the ALG3 expression and other clinicopathologic features.

| Feature | No. | Univariate analysis (OS) | | | Multivariate analysis (OS) | | |
|---|---|---|---|---|---|---|---|
| | | EXP(B) | 95% CI | *p* | EXP(B) | 95% CI | *p* |
| ALG3 | | 2.380 | 1.434–3.949 | 0.001 | 2.905 | 1.669–5.508 | <0.001 |
| High expression | 55 | | | | | | |
| Low expression | 60 | | | | | | |
| Age (years) | | 0.766 | 0.439–1.336 | 0.347 | | | |
| ≤50 | 29 | | | | | | |
| >50 | 86 | | | | | | |
| Gender | | 0.633 | 0.363–1.103 | 0.107 | | | |
| Male | 79 | | | | | | |
| Female | 36 | | | | | | |
| Histology classification | | 1.863 | 1.068–3.250 | 0.028 | 0.701 | 0.389–1.263 | 0.237 |
| I–II | 40 | | | | | | |
| III–IV | 75 | | | | | | |
| TNM stage | | 1.975 | 1.137–3.431 | 0.016 | 1.031 | 0.531–2.072 | 0.932 |
| T1–T2 | 89 | | | | | | |
| T3–T4 | 26 | | | | | | |
| Tumor size (cm) | | 2.394 | 1.339–4.278 | 0.003 | 2.129 | 1.143–3.965 | 0.017 |
| <5 | 43 | | | | | | |
| ≥5 | 72 | | | | | | |
| Tumor number | | 2.136 | 1.233–3.701 | 0.007 | 2.614 | 1.297–5.270 | 0.007 |
| Single | 90 | | | | | | |
| Multiple | 25 | | | | | | |
| Tumor capsular | | 1.562 | 0.833–2.930 | 0.164 | | | |
| Complete | 96 | | | | | | |
| Incomplete | 19 | | | | | | |
| Cut edge | | 1.934 | 0.832–4.496 | 0.125 | | | |
| Negative | 107 | | | | | | |
| Positive | 8 | | | | | | |
| Vascular tumor emboli | | 1.844 | 1.126–3.021 | 0.015 | 1.542 | 0.920–2.582 | 0.100 |
| Present | 45 | | | | | | |
| Absent | 70 | | | | | | |
| Serum AFP (ng/ml) | | 1.932 | 1.177–3.171 | 0.009 | 1.482 | 0.853–2.573 | 0.163 |
| <400 | 76 | | | | | | |
| ≥400 | 39 | | | | | | |
| HBV DNA | | 1.135 | 0.673–1.913 | 0.634 | | | |
| <500 | 79 | | | | | | |
| ≥500 | 36 | | | | | | |
| HBsAg | | 1.662 | 0.964–2.867 | 0.068 | 1.481 | 0.843–2.602 | 0.172 |
| Negative | 41 | | | | | | |
| Positive | 75 | | | | | | |

expression (55 patients) and low-expression groups (60 patients). Patients with high ALG3 expression had shorter overall survival ($p = 0.0005$; hazard ratio (95% CI) = 2.346 (1.423–3.869) but had the same disease-free survival than the low ALG3 expression group ($p = 0.5362$) (**Figures 4C,D**). Furthermore, univariate and multivariate analyses were applied to estimate the influence of ALG3 and pathological characteristics on HCC patients (**Table 2**). Univariate analysis showed that ALG3 expression, histology classification, TNM stage (AJCC), tumor size, tumor number, vascular tumor emboli, serum AFP, and HBsAg were the potential independent prognostic factors for HCC patients (**Figure 4E**). Multivariate analysis indicated that ALG3 expression, tumor number, and tumor size were the independent factors of 5-year overall survival (**Figure 4F**). Collectively, these findings indicated that a higher expression of ALG3 might promote HCC progression, leading to poor clinical prognosis.

## Functional and Pathway Enrichment Analyses of Alpha-1,3-Mannosyltransferase

We intend to figure out the potential mechanism on which ALG3 promotes the progression of HCC. First, we used STRING tools to construct a PPI network to identify genes associated with ALG3 and found that ALG3 was connected with DIBD1, DPM1-3, ALG5, ALG6, ALG8, ALG9, ALG11, and ALG12 (**Figure 5A**). Next, the volcano plot depicted the analysis of the differentially expressed proteins between low and high ALG3 expression groups (**Figure 5B**). KEGG and GO analysis were utilized to explore the potential biological mechanism of ALG3. Based on GO biological process analysis, the DEGs were significantly enriched in response to the xenobiotic stimulus, G protein-coupled receptor signaling pathway, hormone metabolic process, adenylate cyclase-modulating G protein-coupled receptor signaling pathway, collagen-containing extracellular matrix, synaptic membrane, and receptor ligand activity (**Figure 5D**).

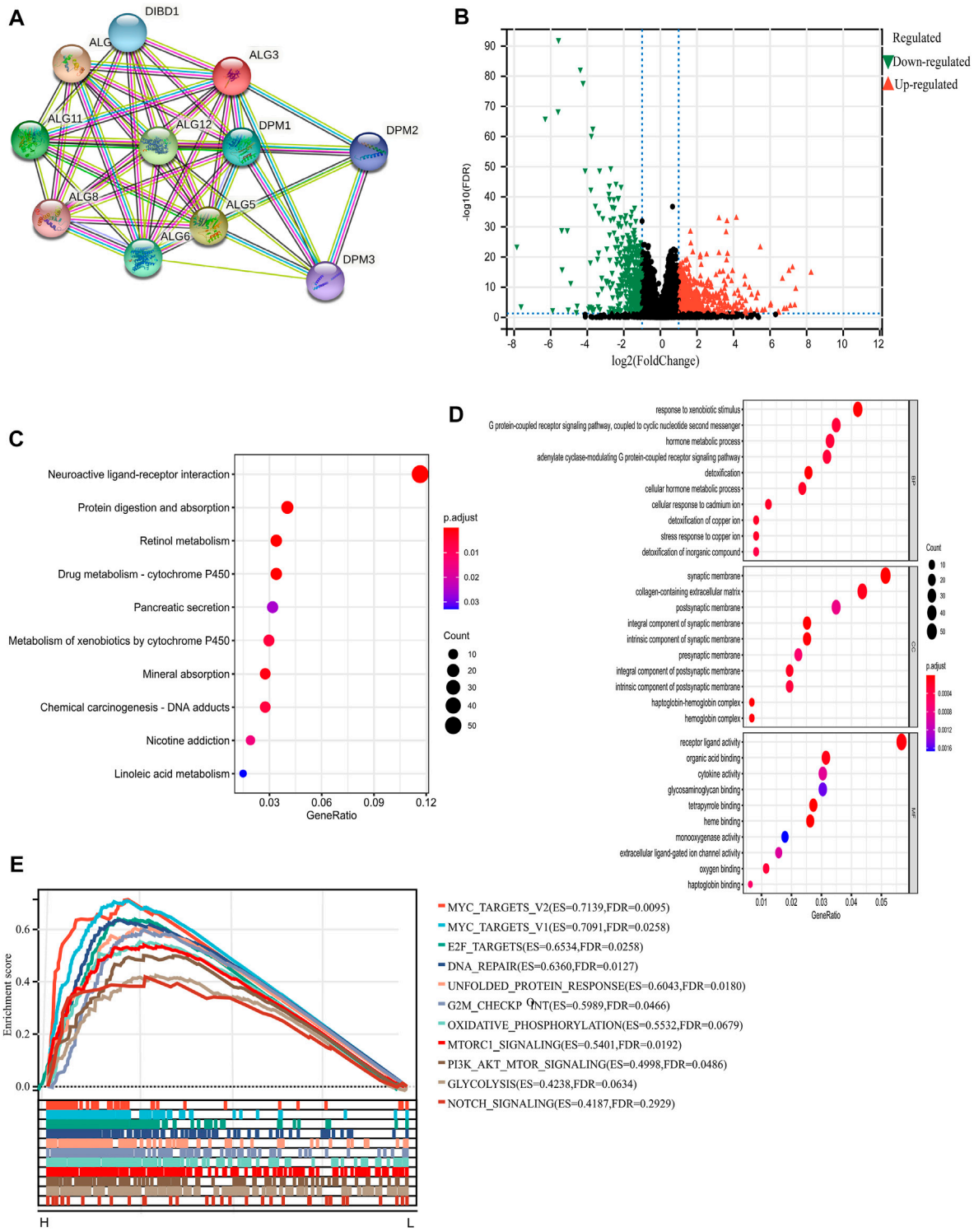



**FIGURE 5 |** PPI and functional enrichment analyses of ALG3 in HCC. **(A)** STRING tool to construct a network of the interaction between ALG3 and other proteins. **(B)** Volcano plot depicting the DEGs. **(C)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, **(D)** Gene Ontology (GO) enrichment, and **(E)** gene set enrichment analysis (GSEA) of ALG3.

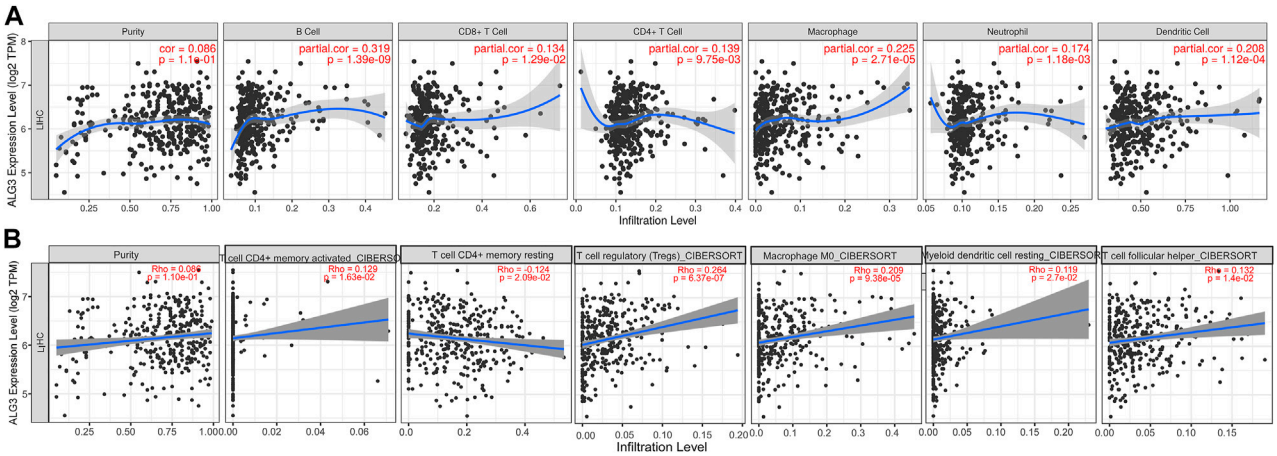



**FIGURE 6 |** Correlations of ALG3 expression with TILs in HCC. Correlations between the ALG3 expression and TILs were verified with TIMER web server **(A)** and CIBERSORT **(B)**.

KEGG pathway analysis results indicated that these DEGs were mainly enriched in neuroactive ligand-receptor interaction, protein digestion and absorption, retinol metabolism, and drug metabolism-cytochrome P450 (**Figure 5C**). In addition, we chose the "hallmark gene set" for GSEA based on the expression of ALG3 to perform the different enriched pathways between upregulated and downregulated ALG3 expressions in HCC samples. The positively enriched biological pathways include MYC targets, E2F targets, DNA repair, G2M checkpoint, and MTORC1 signaling (**Figure 5E**).

## Relationship Between Alpha-1,3-Mannosyltransferase Expression and Immune Infiltration in Hepatocellular Carcinoma

As we all know, immune infiltration has a huge impact on the occurrence and development of tumors (Sun et al., 2020). To better understand the roles of ALG3 in immune infiltration, the TIMER web server was used to evaluate the relationships between ALG3 and the immune cells of HCC, including dendritic cells. As shown in **Figure 6A**, the high ALG3 expression was significantly related to the infiltration of neutrophils (R = 0.174 and $p$ = 1.18e-03), B cells (R = 0.319 and $p$ = 1.39e-09), dendritic cells (R = 0.208 and $p$ = 1.12e-04), $CD8^+$ T cells (R = 0.134 and $p$ = 1.29e-02), macrophages (R = 0.225 and $p$ = 2.71e-05), and $CD4^+$ T cells (R = 0.139 and $p$ = 9.75e-03). CIBERSORT's data also validated the high expression of ALG3 was mainly correlated with T-cell subsets: $CD4^+$ T-cell memory activated and rest, Tregs, and T helper cells (**Figure 6B**). However, more detailed experimental research is needed to prove this conclusion.

## Knockdown of Alpha-1,3-Mannosyltransferase Impeded the Invasion of Hepatocellular Carcinoma Cells

First, we verified the mRNA and protein expression of ALG3 in HCC cell lines. According to **Figures 2E,F**, we have selected HepG2, Hep3B, and SNU398 to construct transient cell lines for the function test. As shown in **Figure 7A**, we verified that ALG3 was successfully silenced by siRNA#1 in Hep3B cells ($p < 0.01$). Next, we used the CCK-8 and MTT assays to explore the effects of ALG3 silencing on the proliferation of Hep3B cells. Our results showed that siRNA-induced silencing did not significantly inhibit the proliferation of Hep3B cells (**Figure 7B**), but the overall trend is slightly downward. In addition, transwell assay results showed that ALG3 silencing impaired the invasion and migratory ability of Hep3B cells (**Figure 7C**). These results suggested that high expression of ALG3 was associated with the proliferation and invasion of HCC cell lines.

# DISCUSSION

Immunotherapy and chemotherapy have a huge impact on hepatocellular carcinoma therapy with the development of technology (Kennedy and Salama, 2020). Currently, studies have shown that the anti-PD-1 antibody combined with locoregional treatments or other molecular-targeted agents is an effective therapy for HCC (Xu et al., 2018), but the precise efficacy and benefit populations have yet to be assessed. Additional reliable novel therapeutic targets need to be found. Previous studies have shown that ALG3 is a high correlation with congenital disorders of glycosylation (Marques-da-Silva et al., 2017; Bian et al., 2020; Ferrer et al., 2020; Alsharhan et al., 2021). In recent years, within the tumor cells, different expressions of glycosyltransferases have been proved as promising biomarkers and potential treatment targets, and glycosylated proteins have attracted increasing attention from researchers. Currently, alpha-fetoprotein (AFP) for liver cancer (Liu et al., 2017), prostate-specific antigen (PSA) for prostate cancer (Wilt and Dahm, 2015), and carcinoembryonic antigen (CEA) for colon cancer (Konishi et al., 2018) are all glycosylated proteins that have been used as biomarkers in clinical application. According to previous studies, ALG3 promotes the invasion, migration, and proliferation of oral squamous cell carcinoma (Shao et al., 2021), non-small cell lung cancer (Ke et al., 2020), and breast cancer (Yang et al., 2018; Sun et al.,

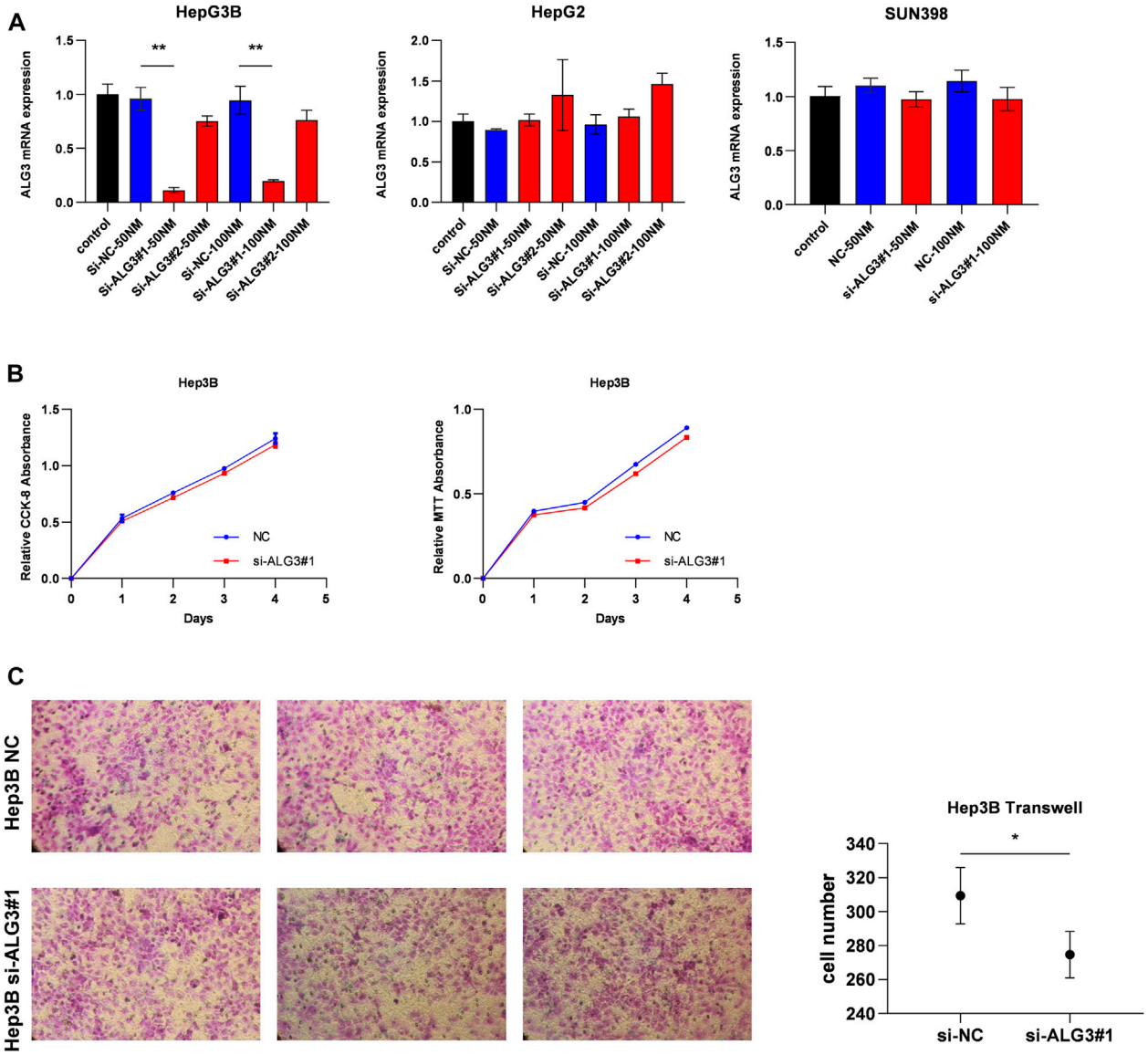



**FIGURE 7 |** ALG3 promotes HCC cell invasion. **(A)** Quantitation of the levels of ALG3 mRNA in HepG2, Hep3B, and SNU398 cells by qRT-PCR after transfection for 48 h with different siRNAs targeting ALG3 (si#1 and si#2) or a control siRNA. **(B)** CCK-8 and MTT assays evaluated the proliferation of Hep3B cells after the knockdown of ALG3. **(C)** Invasion number of Hep3B cells was decreased with ALG3 down-regulation. (*$p < 0.05$ and **$p < 0.01$)

2021), but the expression of ALG3 and its relation with the prognosis of HCC patients remains unknown.

First, we discovered that ALG3 was highly expressed in several malignant tumors, and high expressions can affect the prognosis of patients. In order to find out the role of ALG3 in HCC, we used the TCGA, GEO, and HCDDB datasets to investigate its expression in HCC patients and accordingly found that ALG3 was overexpressed in HCC. Then, we verified the mRNA and protein expression in eight fresh tissue specimens and L02, SNU398, SNU449, PLC/PRF/5, MHCC-97H, HepG2, and Hep3B cell lines. We further proved that both mRNA and protein were higher than in the control group. Next, we conducted the IHC analysis in HCC specimens' slices. IHC results showed the abnormal expression of ALG3 was positively correlated with the advanced histologic grade and cancer stage. Moreover, Kaplan–Meier analyses proved higher expression of ALG3 led to the shorter OS, and multivariate cox regression analysis identified the ALG3 as an independent prognostic factor for HCC patients. To sum up, our findings showed that ALG3 is upregulated in HCC patients and could be regarded as a potential prognosis biomarker.

Then, we constructed PPI networks for ALG3 to search for the relevant genes which play a great role in HCC progression. DPM1, one of the relevant genes, is a vital sensor molecule related to the regulation of cytosolic $Ca^{2+}$ levels in the maintenance of B-cell functions and a potential prognostic tumor marker of HCC (Li et al., 2020b). ALG5 was reported as a prognostic biomarker for early tumor progression in advanced high-grade serous ovarian cancer

(HGSOC) ALG5 (Lu et al., 2021). A previous study indicated that ALG9 was regulated by lncRNA MEG3 and participated in the drug resistance mechanism in acute myeloid leukemia (Yu et al., 2020). Moreover, ALG11 was confirmed as the independent risk factor for the prognosis of HCC in Liu et al. (2021). These results indicated that co-expressed genes played an important role in various cancers, but the synergistic pathways of ALG3 and its correlated genes in HCC remained elusive. Next, we used GO and KEGG to analyze the pathways which were affected by DEGs. The result showed that ALG3 is mainly related to the neuroactive ligand-receptor interaction and G protein-coupled receptor signaling pathway. GSEA enrichment analysis indicated that ALG3 may be regulated with "MYC targets, E2F targets, DNA repair, G2M checkpoint, and MTORC1 signaling". Disorders in the MYC expression are not conducive to patient survival and are involved in cell proliferation, apoptosis, differentiation, and metabolism (Chen et al., 2018). Increased E2F activity led to an abnormal cell cycle (Kent and Leone, 2019). Considering our results, we could predict that ALG3 may participate in crucial biological processes.

In addition, the tumor microenvironment has been confirmed to play an important role in tumorigenesis and progression of immune cells and is a key role in this process (Hinshaw and Shevde, 2019). Therefore, we used TIMER2.0 and CIBERSORT to assess the relationship between ALG3 and immune cell infiltration in LIHC. The result indicated that ALG3 expression was highly related to the enrichment of TILs, including B cells, $CD8^+$ T cells, dendritic cells, neutrophils, $CD4^+$ T cells, and macrophages in HCC. It has been reported that the interaction between tumor-infiltrating B cells and T cells promotes the progression of HCC (Garnelo et al., 2017). Single-cell analysis of primary HCC and recurrent HCC showed that $CD8^+$ T cells and DCs were increased in relapsed HCC, and the aberrant expression led to the compromised antitumor immunity (Zhang et al., 2019). Based on our studies, we could speculate that ALG3 may participate in the occurrence and development of liver cancer by regulating and recruiting the expression of immuno-infiltrating cells.

In order to further explore whether ALG3 affects the biological behavior of HCC cells, we have established transient cell lines. As shown in **Figure 7B**, CCK-8 and MTT results showed that ALG3 had no effect on the proliferation of HCC cells. However, the results of the transwell assay showed that knockdown of ALG3 inhibited migration in Hep3B (**Figure 7C**). Consistent with the results of previous studies, two inferences can be drawn as follows. First, ALG3 may only affect the survival prognosis of patients by affecting the invasion of HCC cells, and the effect on the proliferation of HCC cells is not obvious. Second, a previous study has reported that ALG3 can affect the proliferation of BRCA, OSCC, and NSCLC (Shao et al., 2021; Ke et al., 2020; Sun et al., 2021), indicating that ALG3 can affect the proliferation of some cancer cells. In **Figure 7A**, only Hep3B was successfully constructed among the three cell lines, suggesting that the expected results could not be achieved due to the lack of quality of si-ALG3. In addition, researchers can now assess the glycoprotein secreted on the cell surface to reflect the overall cellular state in health and disease. It has been suggested that altered glycosylation of proteins that participate in the immunological synapse affects the signaling processes and cell proliferation, as well as exacerbation of the effector mechanisms of T cells that trigger systemic damage and autoimmunity (Gómez-Henao et al., 2021). Current research indicates that changes in some glycosylated proteins can cause the immune escape of viruses and promote metastasis of cancer cells, but some glycosylated proteins can also regulate the cell cycle of cancer cells and induce apoptosis of tumor cells (Reily et al., 2019). New insights into the structure and function of glycosylates can be applied to therapeutic development, improving our ability to regulate immune response and inflammation. Studies on the correlation between glycosylation and the regulation of immune response (Li et al., 2018), PD-1 receptor pathway (Liu et al., 2020b), and tumor microenvironment (Chandler et al., 2019) suggest that new tumor therapeutic targets may be found by paying attention to glycosylation, so we planned to construct stable cell lines to continue to study deeper mechanisms and explore the possible pathways of ALG3 affecting hepatocellular carcinoma.

To sum up, our findings proved that ALG3 overexpression resulted in shorter overall survival in HCC patients and may promote tumor progression and immune cell infiltration. Collectively, our study provided a potential independent unfavorable prognostic biomarker for HCC patients and new insights into the mechanism of HCC tumorigenesis and progression.

## DATA AVAILABILITY STATEMENT

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

## ETHICS STATEMENT

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

ZZa and ZZe contributed to the study's conception and design and performed the RT-PCR and IHC. ZZa, TP, and JW collected data. ZZa and JH analyzed data. ZZa and ZZe wrote the first draft of the manuscript. All authors contributed to the manuscript revision and approved the submitted version.

## FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 81972792), the Natural Science Foundation of Guangdong Province, People's Republic of China (Grant No. 2020A1515010149), the National key Clinical Specialty Construction Project (2021-2024, No. 2022YW030009), the Science and Technology Project of Guangzhou city (No. 202102020973) and the NSFC Incubation Program of GDPH (No. KY012021166).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.816102/full#supplementary-material

## REFERENCES


Alsharhan, H., Ng, B. G., Daniel, E. J. P., Friedman, J., Pivnick, E. K., Al-Hashem, A., et al. (2021). Expanding the Phenotype, Genotype and Biochemical Knowledge of ALG3-CDG. *J Inher Metab Disea* 44 (4), 987–1000. doi:10.1002/jimd.12367

Bian, Y., Qiao, C., Zheng, S., Qiu, H., Li, H., Zhang, Z., et al. (2020). ALG3-CDG: Lethal Phenotype and Novel Variants in Chinese Siblings. *J. Hum. Genet.* 65 (12), 1129–1134. doi:10.1038/s10038-020-0798-7

Bloch, J. S., Pesciullesi, G., Boilevin, J., Nosol, K., Irobalieva, R. N., Darbre, T., et al. (2020). Structure and Mechanism of the ER-Based Glucosyltransferase ALG6. *Nature* 579 (7799), 443–447. doi:10.1038/s41586-020-2044-z

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

Chandler, K. B., Costello, C. E., and Rahimi, N. (2019). Glycosylation in the Tumor Microenvironment: Implications for Tumor Angiogenesis and Metastasis. *Cells* 8 (6). doi:10.3390/cells8060544

Chen, H., Liu, H., and Qing, G. (2018). Targeting Oncogenic Myc as a Strategy for Cancer Treatment. *Sig Transduct. Target Ther.* 3, 5. doi:10.1038/s41392-018-0008-7

Costa, A. F., Campos, D., Reis, C. A., and Gomes, C. (2020). Targeting Glycosylation: A New Road for Cancer Drug Discovery. *Trends cancer* 6 (9), 757–766. doi:10.1016/j.trecan.2020.04.002

Faivre, S., Rimassa, L., and Finn, R. S. (2020). Molecular Therapies for HCC: Looking outside the Box. *J. hepatology* 72 (2), 342–352. doi:10.1016/j.jhep.2019.09.010

Ferrer, A., Starosta, R. T., Ranatunga, W., Ungar, D., Kozicz, T., Klee, E., et al. (2020). Fetal Glycosylation Defect Due to ALG3 and COG5 Variants Detected via Amniocentesis: Complex Glycosylation Defect with Embryonic Lethal Phenotype. *Mol. Genet. metabolism* 131 (4), 424–429. doi:10.1016/j.ymgme.2020.11.003

Garnelo, M., Tan, A., Her, Z., Yeong, J., Lim, C. J., Chen, J., et al. (2017). Interaction between Tumour-Infiltrating B Cells and T Cells Controls the Progression of Hepatocellular Carcinoma. *Gut* 66 (2), 342–351. doi:10.1136/gutjnl-2015-310814

Gómez-Henao, W., Tenorio, E. P., Sanchez, F. R. C., Mendoza, M. C., Ledezma, R. L., and Zenteno, E. (2021). Relevance of Glycans in the Interaction between T Lymphocyte and the Antigen Presenting Cell. *Int. Rev. Immunol.* 40 (4), 274–288. doi:10.1080/08830185.2020.1845331

Grothey, A., Blay, J.-Y., Pavlakis, N., Yoshino, T., and Bruix, J. (2020). Evolving Role of Regorafenib for the Treatment of Advanced Cancers. *Cancer Treat. Rev.* 86, 101993. doi:10.1016/j.ctrv.2020.101993

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

Ke, S.-b., Qiu, H., Chen, J.-m., Shi, W., Han, C., Gong, Y., et al. (2020). ALG3 Contributes to the Malignancy of Non-small Cell Lung Cancer and Is Negatively Regulated by MiR-98-5p. *Pathology - Res. Pract.* 216 (3), 152761. doi:10.1016/j.prp.2019.152761

Kennedy, L. B., and Salama, A. K. S. (2020). A Review of Cancer Immunotherapy Toxicity. *CA A Cancer J. Clin.* 70 (2), 86–104. doi:10.3322/caac.21596

Kent, L. N., and Leone, G. (2019). The Broken Cycle: E2F Dysfunction in Cancer. *Nat. Rev. Cancer* 19 (6), 326–338. doi:10.1038/s41568-019-0143-7

Konishi, T., Shimada, Y., Hsu, M., Tufts, L., Jimenez-Rodriguez, R., Cercek, A., et al. (2018). Association of Preoperative and Postoperative Serum Carcinoembryonic Antigen and Colon Cancer Outcome. *JAMA Oncol.* 4 (3), 309–315. doi:10.1001/jamaoncol.2017.4420

Li, C. W., Lim, S. O., Chung, E. M., Kim, Y. S., Park, A. H., Yao, J., et al. (2018). Eradication of Triple-Negative Breast Cancer Cells by Targeting Glycosylated PD-L1. *Cancer Cell* 33 (2), 187–e10. doi:10.1016/j.ccell.2018.01.009

Li, M., Xia, S., and Shi, P. (2020). DPM1 Expression as a Potential Prognostic Tumor Marker in Hepatocellular Carcinoma. *PeerJ* 8, e10307. doi:10.7717/peerj.10307

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for Analysis of Tumor-Infiltrating Immune Cells. *Nucleic acids Res.* 48 (W1), W509–W514. doi:10.1093/nar/gkaa407

Lian, Q., Wang, S., Zhang, G., Wang, D., Luo, G., Tang, J., et al. (2018). HCCDB: A Database of Hepatocellular Carcinoma Expression Atlas. *Genomics, proteomics Bioinforma.* 16 (4), 269–275. doi:10.1016/j.gpb.2018.07.003

Liu, B., Ma, X., Liu, Q., Xiao, Y., Pan, S., and Jia, L. (2020). Retraction Note: Aberrant Mannosylation Profile and FTX/miR-342/ALG3-axis Contribute to Development of Drug Resistance in Acute Myeloid Leukemia. *Cell Death Dis.* 11 (2), 122. doi:10.1038/s41419-020-2320-8

Liu, G., Tang, H., Li, C., Zhen, H., Zhang, Z., and Sha, Y. (2021). Prognostic Gene Biomarker Identification in Liver Cancer by Data Mining. *Am. J. Transl. Res.* 13 (5), 4603–4613.

Liu, H., Xu, Y., Xiang, J., Long, L., Green, S., Yang, Z., et al. (2017). Targeting Alpha-Fetoprotein (AFP)-MHC Complex with CAR T-Cell Therapy for Liver Cancer. *Clin. Cancer Res.* 23 (2), 478–488. doi:10.1158/1078-0432.ccr-16-1203

Liu, K., Tan, S., Jin, W., Guan, J., Wang, Q., Sun, H., et al. (2020). N-glycosylation of PD-1 Promotes Binding of Camrelizumab. *EMBO Rep.* 21 (12), e51444. doi:10.15252/embr.202051444

Lu, H., Cunnea, P., Nixon, K., Rinne, N., Aboagye, E. O., and Fotopoulou, C. (2021). Discovery of a Biomarker Candidate for Surgical Stratification in High-Grade Serous Ovarian Cancer. *Br. J. Cancer* 124 (7), 1286–1293. doi:10.1038/s41416-020-01252-2

Lu, J., Zhang, X.-P., Zhong, B.-Y., Lau, W. Y., Madoff, D. C., Davidson, J. C., et al. (2019). Management of Patients with Hepatocellular Carcinoma and Portal Vein Tumour Thrombosis: Comparing East and West. *lancet Gastroenterology hepatology* 4 (9), 721–730. doi:10.1016/s2468-1253(19)30178-5

Marques-da-Silva, D., Dos Reis Ferreira, V., Monticelli, M., Janeiro, P., Videira, P. A., Witters, P., et al. (2017). Liver Involvement in Congenital Disorders of Glycosylation (CDG). A Systematic Review of the Literature. *J. Inherit. Metab. Dis.* 40 (2), 195–207. doi:10.1007/s10545-016-0012-4

McGlynn, K. A., Petrick, J. L., and El-Serag, H. B. (2021). Epidemiology of Hepatocellular Carcinoma. *Hepatol. Baltim. Md)* 73 (Suppl. 1), 4–13. doi:10.1002/hep.31288

Petrick, J. L., Florio, A. A., Znaor, A., Ruggieri, D., Laversanne, M., Alvarez, C. S., et al. (2020). International Trends in Hepatocellular Carcinoma Incidence, 1978-2012. *Int. J. Cancer* 147 (2), 317–330. doi:10.1002/ijc.32723

Pinato, D. J., Sharma, R., Allara, E., Yen, C., Arizumi, T., Kubota, K., et al. (2017). The ALBI Grade Provides Objective Hepatic Reserve Estimation across Each BCLC Stage of Hepatocellular Carcinoma. *J. hepatology* 66 (2), 338–346. doi:10.1016/j.jhep.2016.09.008

Pinter, M., Scheiner, B., and Peck-Radosavljevic, M. (2021). Immunotherapy for Advanced Hepatocellular Carcinoma: a Focus on Special Subgroups. *Gut* 70 (1), 204–214. doi:10.1136/gutjnl-2020-321702

Reily, C., Stewart, T. J., Renfrow, M. B., and Novak, J. (2019). Glycosylation in Health and Disease. *Nat. Rev. Nephrol.* 15 (6), 346–366. doi:10.1038/s41581-019-0129-4

Ruf, B., Heinrich, B., and Greten, T. F. (2021). Immunobiology and Immunotherapy of HCC: Spotlight on Innate and Innate-like Immune Cells. *Cell Mol. Immunol.* 18 (1), 112–127. doi:10.1038/s41423-020-00572-w

Shang, R., Song, X., Wang, P., Zhou, Y., Lu, X., Wang, J., et al. (2021). Cabozantinib-based Combination Therapy for the Treatment of Hepatocellular Carcinoma. *Gut* 70 (9), 1746–1757. doi:10.1136/gutjnl-2020-320716

Shao, P., Wei, C., and Wang, Y. (2021). ALG3 Contributes to the Malignant Properties of OSCC Cells by Regulating CDK-Cyclin Pathway. *Colorectal Dis.* 27 (6), 1426–1434. doi:10.1111/odi.13687

Siegel, R. L., Miller, K. D., and Jemal, A. (20192019). Cancer Statistics, 2019. *CA A Cancer J. Clin.* 69 (1), 7–34. doi:10.3322/caac.21551

Silva, M. C., Fernandes, Â., Oliveira, M., Resende, C., Correia, A., de-Freitas-Junior, J. C., et al. (2020). Glycans as Immune Checkpoints: Removal of Branched N-Glycans Enhances Immune Recognition Preventing Cancer Progression. *Cancer Immunol. Res.* 8 (11), 1407–1425. doi:10.1158/2326-6066.cir-20-0264

Song, X., Zhou, Z., Li, H., Xue, Y., Lu, X., Bahar, I., et al. (2020). Pharmacologic Suppression of B7-H4 Glycosylation Restores Antitumor Immunity in Immune-Cold Breast Cancers. *Cancer Discov.* 10 (12), 1872–1893. doi:10.1158/2159-8290.cd-20-0402

Sun, L., Wang, X., Saredy, J., Yuan, Z., Yang, X., and Wang, H. (2020). Innate-adaptive Immunity Interplay and Redox Regulation in Immune Response. *Redox Biol.* 37, 101759. doi:10.1016/j.redox.2020.101759

Sun, X., He, Z., Guo, L., Wang, C., Lin, C., Ye, L., et al. (2021). ALG3 Contributes to Stemness and Radioresistance through Regulating Glycosylation of TGF-β Receptor II in Breast Cancer. *J. Exp. Clin. Cancer Res.* 40 (1), 149. doi:10.1186/s13046-021-01932-8

Thul, P. J., and Lindskog, C. (2018). The Human Protein Atlas: A Spatial Map of the Human Proteome. *Protein Sci.* 27 (1), 233–244. doi:10.1002/pro.3307

Varki, A. (2017). Biological Roles of Glycans. *Glycobiology* 27 (1), 3–49. doi:10.1093/glycob/cww086

Wilt, T. J., and Dahm, P. (2015). PSA Screening for Prostate Cancer: Why Saying No Is a High-Value Health Care Choice. *J. Natl. Compr. Canc Netw.* 13 (12), 1566–1574. doi:10.6004/jnccn.2015.0182

Xu, F., Jin, T., Zhu, Y., and Dai, C. (2018). Immune Checkpoint Therapy in Liver Cancer. *J. Exp. Clin. Cancer Res.* 37 (1), 110. doi:10.1186/s13046-018-0777-4

Yang, Y., Zhou, Y., Xiong, X., Huang, M., Ying, X., and Wang, M. (2018). ALG3 Is Activated by Heat Shock Factor 2 and Promotes Breast Cancer Growth. *Med. Sci. Monit.* 24, 3479–3487. doi:10.12659/msm.907461

Yu, Y., Kou, D., Liu, B., Huang, Y., Li, S., Qi, Y., et al. (2020). LncRNA MEG3 Contributes to Drug Resistance in Acute Myeloid Leukemia by Positively Regulating ALG9 through Sponging miR-155. *Int. J. Lab. Hem.* 42 (4), 464–472. doi:10.1111/ijlh.13225

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

Zhao, Y., Zhang, Y.-N., Wang, K.-T., and Chen, L. (2020). Lenvatinib for Hepatocellular Carcinoma: From Preclinical Mechanisms to Anti-cancer Therapy. *Biochimica Biophysica Acta (BBA) - Rev. Cancer* 1874 (1), 188391. doi:10.1016/j.bbcan.2020.188391

Zhu, A. X., Kang, Y.-K., Yen, C.-J., Finn, R. S., Galle, P. R., Llovet, J. M., et al. (2019). Ramucirumab after Sorafenib in Patients with Advanced Hepatocellular Carcinoma and Increased α-fetoprotein Concentrations (REACH-2): a Randomised, Double-Blind, Placebo-Controlled, Phase 3 Trial. *Lancet Oncol.* 20 (2), 282–296. doi:10.1016/s1470-2045(18)30937-9

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

frontiers | Frontiers in Molecular Biosciences

ORIGINAL RESEARCH
published: 22 June 2022
doi: 10.3389/fmolb.2022.901829

# DNA Damage Response Gene-Based Subtypes Associated With Clinical Outcomes in Early-Stage Lung Adenocarcinoma

*Yang Zhao[1†], Bei Qing[2†], Chunwei Xu[3†], Jing Zhao[4], Yuchen Liao[4], Peng Cui[4], Guoqiang Wang[4], Shangli Cai[4], Yong Song[3], Liming Cao[5*] and Jianchun Duan[6*]*

[1]Department of Thoracic Surgery, Shanghai Chest Hospital, Shanghai Jiao Tong University, Shanghai, China, [2]Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, Changsha, China, [3]Department of Respiratory Medicine, Jinling Hospital, Nanjing University School of Medicine, Nanjing, China, [4]Burning Rock Biotech, Guangzhou, China, [5]Department of Respiratory Medicine, Xiangya Hospital, Central South University, Changsha, China, [6]CAMS Key Laboratory of Translational Research on Lung Cancer, State Key Laboratory of Molecular Oncology, Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences Peking Union Medical College, Beijing, China

OPEN ACCESS

**Edited by:**
Umberto Malapelle,
University of Naples Federico II, Italy

**Reviewed by:**
Valerio Gristina,
University of Palermo, Italy
Ramya Sivakumar,
University of Washington, United States
Xiawei Cheng,
East China University of Science and Technology, China

***Correspondence:**
Liming Cao
clming@csu.edu.cn
Jianchun Duan
duanjianchun79@163.com

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

**Specialty section:**
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

**Received:** 22 March 2022
**Accepted:** 11 May 2022
**Published:** 22 June 2022

**Citation:**
Zhao Y, Qing B, Xu C, Zhao J, Liao Y, Cui P, Wang G, Cai S, Song Y, Cao L and Duan J (2022) DNA Damage Response Gene-Based Subtypes Associated With Clinical Outcomes in Early-Stage Lung Adenocarcinoma. Front. Mol. Biosci. 9:901829. doi: 10.3389/fmolb.2022.901829

DNA damage response (DDR) pathways play a crucial role in lung cancer. In this retrospective analysis, we aimed to develop a prognostic model and molecular subtype based on the expression profiles of DDR-related genes in early-stage lung adenocarcinoma (LUAD). A total of 1,785 lung adenocarcinoma samples from one RNA-seq dataset of The Cancer Genome Atlas (TCGA) and six microarray datasets of Gene Expression Omnibus (GEO) were included in the analysis. In the TCGA dataset, a DNA damage response gene (DRG)–based signature consisting of 16 genes was constructed to predict the clinical outcomes of LUAD patients. Patients in the low-DRG score group had better outcomes and lower genomic instability. Then, the same 16 genes were used to develop DRG-based molecular subtypes in the TCGA dataset to stratify early-stage LUAD into two subtypes (DRG1 and DRG2) which had significant differences in clinical outcomes. The Kappa test showed good consistency between molecular subtype and DRG ($K = 0.61$, $p < 0.001$). The DRG subtypes were significantly associated with prognosis in the six GEO datasets (pooled estimates of hazard ratio, OS: 0.48 (0.41–0.57), $p < 0.01$; DFS: 0.50 (0.41–0.62), $p < 0.01$). Furthermore, patients in the DRG2 group benefited more from adjuvant therapy than standard-of-care, which was not observed in the DRG1 group. In summary, we constructed a DRG-based molecular subtype that had the potential to predict the prognosis of early-stage LUAD and guide the selection of adjuvant therapy for early-stage LUAD patients.

**Keywords: DNA damage response, signature, prognostic, molecular subtype, early-stage, lung adenocarcinoma**

## INTRODUCTION

Lung cancer is the major cause of global cancer mortality in 2020, with an estimated 1.8 million deaths worldwide (Sung et al., 2021). Non–small cell lung cancer (NSCLC) represents 85% of all lung cancers. Based on histology, NSCLC can be further divided into lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LSCC), large-cell carcinoma, etc. (Bender, 2014). The survival of patients

with NSCLC is largely determined by the tumor stage at diagnosis. Only 15% of patients with late-stage disease (stages III–IV) are alive after 5 years, which makes NSCLC one of the cancers with the worst prognosis (Necchi et al., 2017). Although the 5-year survival rate increases to approximately 60% and 40% for stage I and stage II patients, respectively, around 30–55% of them experienced disease recurrence within 5 years after surgery (Howington et al., 2013; Wang et al., 2017). In recent years, the immuno-oncology (IO)–based strategies, such as immune checkpoint inhibitors (ICIs), the combination of different ICIs, or chemotherapies, have achieved evolutionized improvements in the treatment for a subset of patients with lung cancer (Listì et al., 2019; Passiglia et al., 2021). Besides the breakthrough in cancer therapy, it is also important to improve recurrence prediction and clinical management with the increase of early-stage tumors due to the progress of lung cancer screening.

The rapid development of high-throughput technologies, especially DNA microarrays and RNA-sequencing, has facilitated the exploration of several expression-based gene signatures for risk stratification in NSCLC patients. Beer et al. proposed a 50-gene signature to identify low- and high-risk stage I lung adenocarcinomas using microarray analysis (Beer et al., 2002). The Director's Challenge Consortium validated the performance of several such prognostic models in a large multi-site cohort with 442 lung adenocarcinomas (Chen et al., 2007; Shedden et al., 2008; Sun et al., 2008). In addition, a 14-gene expression signature (RT-PCR–based) has been commercialized to stratify different risk groups for resected non-squamous NSCLC patients (Kratz et al., 2012). A 25–immune gene signature and a 31–proliferation gene signature both have shown promising clinical utility for risk stratification and individualized management in NSCLC patients (Wistuba et al., 2013; Li et al., 2017). However, none of these signatures was further analyzed in patients with and without adjuvant therapy to validate the potential clinical utility in the guidance of adjuvant therapy.

Genomic instability is one of the key hallmarks of cancer, and DNA damage response (DDR) plays a significant role in maintaining genomic integrity (Hanahan and Weinberg, 2011). The DDR system is a complex signaling network which involves eight pathways: base excision repair (BER), mismatch repair (MMR), homologous recombination repair (HRR), nonhomologous end joining (NHEJ), checkpoint factors (CPF), Fanconi anemia (FA), nucleotide-excision repair (NER), and DNA translesion synthesis (TLS) (Scarbrough et al., 2016). These pathways operate collectively to detect diverse types of DNA lesions and activate signaling mechanisms to boost the repair machine (Jackson and Bartek, 2009). Previous studies have demonstrated that the DDR pathways play significant roles in cancer progression and the response to cancer therapies. Several prognostic models, based on DDR genes, have been constructed for glioblastoma, ovarian cancer, and low-grade gliomas (Knijnenburg et al., 2018; Gobin et al., 2019; Sun et al., 2019; Pang et al., 2020). However, the DDR genes identified in these prognostic models vary widely between different cancers, suggesting that DDR genes may exert different molecular effects in different cellular environments. The relationships of various DDR genes with prognosis in lung adenocarcinoma are not well-established.

In this study, we aimed to identify and validate a group of DDR genes to stratify early-stage LUAD patients into different subtypes with different prognoses and guide the use of adjuvant therapy.

# MATERIALS AND METHODS

## Molecular and Clinical Data

The LUAD dataset of The Cancer Genome Atlas (TCGA) and six microarray datasets of Gene Expression Omnibus (GEO) were included in the analysis. For the TCGA dataset, RNA-sequencing data (FPKM format), genetic mutations, copy number variant (CNV), and clinical features, including age, sex, tumor stage, histology subtype, adjuvant treatment, and follow-up information, were obtained from the GDC (https://portal.gdc.cancer.gov/). In addition, normalized microarray data and the corresponding clinical characteristics of patients with early-stage (stages I and II) lung adenocarcinoma from six GEO cohorts (GSE31210, GSE37745, GSE68465, GSE30219, GSE72094, and GSE13213) were obtained for further external validation in this study.

## DDR Gene–Based Signature Construction

A total of 200 DDR-related genes were curated and analyzed to identify prognosis-related markers (Scarbrough et al., 2016). These genes used in the study are listed in **Supplementary Table S1**. Univariable Cox regression and LASSO Cox regression analyses with minimum partial likelihood deviance were performed to select genes associated with OS. We defined the risk score using the following formula

$$riskscore = \sum_{k=0}^{n} (coef\ of\ gene\ k * expr\ of\ gene\ k),$$

where $n$ is the number of markers. The nearest neighbor estimation method was applied to identify the best cutoff point of risk score to stratify patients into high- and low-risk subgroups. Kaplan–Meier (KM) analysis and receiver operating characteristic (ROC) curve were used to assess the performance of the signature.

## Association Between DDR Signature and Genome Features and Gene Expression

In order to explore the potential molecular mechanisms of the DDR-gene–based signature, the associations of the DDR signature with somatic mutation, CNV, genomic scar signature, and gene expression data were analyzed based on TCGA data.

## DDR Molecular Subtype Identification and Validation

Unsupervised clustering with the hierarchical cluster algorithm (based on Euclidean distance and Ward's linkage) of the

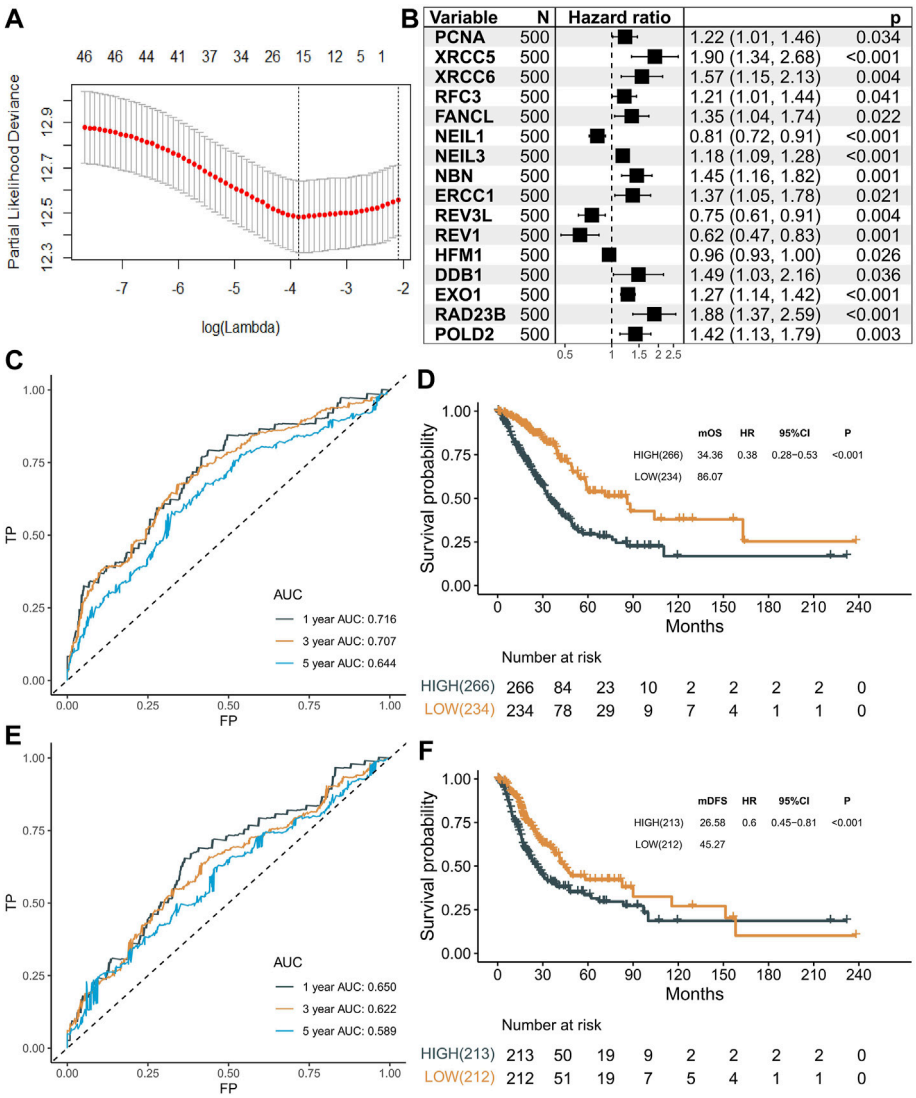



**FIGURE 1 |** Selection of prognostic markers. **(A)** Tuning parameter (λ) selection in the LASSO model using 10-fold cross-validation via minimum criteria. **(B)** Forest plot showing the results of univariable Cox regression analyses. **(C)** Predictive value of 16 genes in the overall survival of patients in the LUAD dataset. **(D)** Kaplan–Meier curves of overall survival for high- and low-risk patient groups in the TCGA–LUAD dataset. Patients were divided into two groups with a cutoff score of 21.96. **(E)** Predictive value of 16 genes in the disease-free survival of patients in the LUAD dataset. **(F)** Kaplan–Meier curves of disease-free survival for high- and low-risk patient groups in the TCGA–LUAD dataset. Patients were divided into two groups with a cutoff score of 21.96.

expression profiles of the genes in the DDR signature was performed to identify molecular subtype in early-stage LUAD. The default parameters of the hclust function were used to perform the classification. The cluster number was selected as 2. It was further validated in six GEO datasets.

## Statistical Analyses

R software v4.0.2 was used for all the bioinformatics and statistical analyses, including data preprocess, LASSO Cox regression, CNV and mutation visualization, and ROC analysis. The KM method and log-rank test were adopted to generate and evaluate the statistical significance of the survival curves between groups. The specificity and sensitivity of the signature were evaluated using the ROC curve, and the area under the curve (AUC) of distinct survival time was quantified using R-package pROC. The Kappa consistency test was used to analyze the consistency between the two group methods. The Cox proportional hazards model was applied to identify the independence of the signature. The prognostic values of single genes in signatures were accessed using the "szcox" function of the ezcox package. R-package SubgrPlots was used for subgroup analysis, which was visualized using the Forester package. A propensity score matching (PSM) analysis was performed according to a 1:1 ratio between the two subgroups (with or without adjuvant therapy) to adjust for clinicopathologic characteristics bias using the MatchIt package. Heat maps of

**TABLE 1 |** Univariable analysis and multivariable Cox regression analyses of OS and DFS in TCGA cohorts.

| Cohort | Characteristics | Sample size | OS | | | | DFS | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Univariable Cox | | Multivariable Cox | | Univariable Cox | | Multivariable Cox | |
| | | | HR (95% CI) | *P*-value | HR (95% CI) | *P*-value | HR (95% CI) | *P*-value | HR (95% CI) | *P*-value |
| TCGA | | | | | | | | | | |
| — | Age (≤60 vs. >60) | 490 | 0.86 (0.62–1.18) | 0.3400 | — | — | 1.03 (0.77–1.38) | 0.8590 | 0.69 (0.49–0.96) | 0.0295 |
| — | Sex (Male vs. Female) | 500 | 1.05 (0.78–1.40) | 0.7530 | — | — | 0.59 (0.41–0.83) | 0.0031 | — | — |
| — | Stage (I_II vs. III_IV) | 492 | 0.39 (0.28–0.53) | <0.001 | 0.43 (0.31–0.58) | <0.001 | 1.00 (0.65–1.53) | 0.9970 | 0.66 (0.46–0.95) | 0.0271 |
| — | Smoking (Never vs. Ever) | 486 | 1.14 (0.75–1.72) | 0.5450 | — | — | 2.19 (1.18–4.06) | 0.0130 | — | — |
| — | EGFR (MUT vs. WT) | 383 | 2.38 (1.31–4.34) | 0.0005 | 3.08 (1.66–5.71) | <0.001 | 0.84 (0.56–1.27) | 0.4130 | 2.55 (1.33–4.92) | 0.0050 |
| — | KRAS (MUT vs. WT) | 383 | 0.87 (0.55–1.36) | 0.5310 | — | — | 0.99 (0.41–2.42) | 0.9850 | — | — |
| — | ALK (MUT vs. WT) | 383 | 0.60 (0.19–1.89) | 0.3820 | — | — | 1.04 (0.73–1.47) | 0.8420 | — | — |
| — | TP53 (MUT vs. WT) | 383 | 1.23 (0.84–1.80) | 0.2790 | — | — | 0.90 (0.49–1.68) | 0.7940 | — | — |
| — | BRAF (MUT vs. WT) | 383 | 0.68 (0.31–1.46) | 0.3200 | — | — | 0.60 (0.45–0.81) | 0.0007 | — | — |
| — | DRG (low vs. high) | 500 | 0.38 (0.28–0.53) | <0.001 | 0.42 (0.30–0.58) | <0.001 | 0.72 (0.52–1.00) | 0.0479 | 0.58 (0.43–0.79) | 0.0005 |

TCGA–LUAD and GEO datasets were generated using the pheatmap package. The maftools package was used to visualize the mutation landscape in the TCGA–LUAD dataset. Two-sided $p < 0.05$ was considered to be statistically significant.

# RESULTS

## Construction of a 16-Gene Signature

A total of 1,785 primary LUAD tumors and their clinicopathological features were downloaded from TCGA and GEO databases, and the baseline characteristics are summarized in **Supplementary Table S2**. To identify the survival-related genes, univariable Cox regression was performed in the 200 DDR-related genes with the TCGA–LUAD dataset (n = 500), and 46 DDR genes were identified to be significantly associated with OS. Then, ten-fold cross-validation of LASSO Cox was implemented using the "glmnet" package, and 16 genes (*PCNA*, *XRCC5*, *XRCC6*, *RFC3*, *FANCL*, *NEIL1*, *NEIL3*, *NBN*, *ERCC1*, *REV3L*, *REV1*, *HFM1*, *DDB1*, *EXO1*, *RAD23B*, and *POLD2*) were identified to be the most informative and were used to construct a risk score (**Figure 1A**). In brief, *NEIL1*, *HFM1*, *REV3L*, and *REV1* genes were protective factors (all HRs < 1, $p < 0.05$), while the others genes were risk factors for the prognosis in patients with LUAD (all HRs > 1, $p < 0.05$) (**Figure 1B**). Then, we established a DNA damage response gene (DRG)–based signature for each patient based on the following formula: DRG = (−0.0618*PCNA)+(0.2175*XRCC5)+(0.1094*XRCC6)+(−0.0929*RFC3)+(0.1894*FANCL)+(−0.0008*NEIL1)+(0.0095*NEIL3)+(0.1651*NBN)+(0.1594*ERCC1)+(−0.0817*REV3L)+(−0.0640*REV1)+(−0.0119*HFM1)+(0.1600*DDB1)+(0.1070*EXO1)+(0.2896*RAD23B)+(0.0548*POLD2). The best cutoff of 21.96 was used to stratify the patients into high- or low-risk groups. The AUCs for 1-, 3-, and 5-year overall survival (OS) rate predictions for the DRG of the TCGA–LUAD dataset were 0.716, 0.707, and 0.644, respectively (**Figure 1C**). The KM curves revealed significantly higher OS with lower DRG (HR = 0.38, 95% CI: 0.28–0.53, $p < 0.001$, **Figure 1D**). Similar results for disease-free survival (DFS) were obtained. The AUCs for 1-, 3-, and 5-year were 0.650, 0.622, and 0.589, respectively, (**Figure 1E**) and the association between the DRG and DFS was significant (log-rank $p < 0.001$; HR = 0.60, 95% CI: 0.45–0.81, **Figure 1F**). Next, we tested the independent prognostic prediction value of the DRG. After adjusting for clinical features, including age, sex, tumor stage, and smoking, as well as the driver gene mutation (EGFR, KRAS, ALK, ROS1, BRAF, and TP53), the DRG served as an independent prognostic biomarker for predicting outcomes (OS, HR: 0.42 (0.30–0.58); DFS, HR: 0.43 (0.31–0.53), **Table 1**).

## Association Between the 16 Genes and Clinicopathological Factors

To further study the underlying mechanism of the DRG, we explored the molecular function and the association with prognosis of genes in the DRG. Most of them had positive coefficients in this regression equation with HR > 1, indicating poor prognostic genes, while genes (*REV3L*, *REV1*, *NEIL1*, and *HFM1*) had negative coefficients with HR < 1 (**Figure 2A**). To depict the genomic and expression alterations of the 16 DDR genes, we further described the prevalence of somatic mutations, CNV, and mRNA expression of the 16 genes in LUAD patients (**Figure 2B**). Of the 486 LUAD patients, 63 (13.0%) patients harbored at least one mutation of the pattern genes. Among them, *HFM1* had the highest mutation frequency (4%) followed by *EOX1* and *REV3L*, while there were no mutations in *ERCC1*, *NEIL1*, and *PCNA*. Meanwhile, CNV analysis showed that *EXO1*,

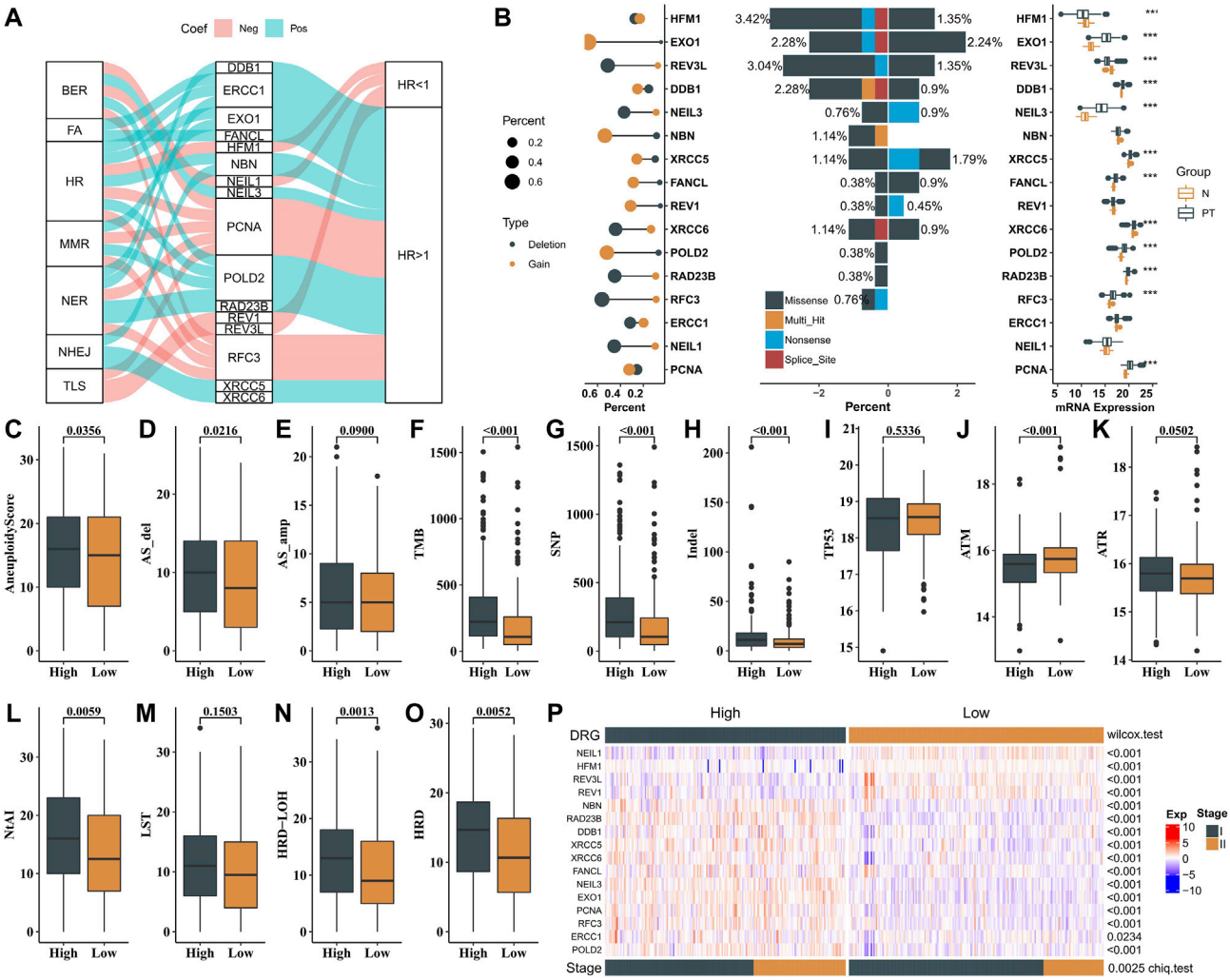



**FIGURE 2 |** Mechanism of validation in mutation, CNV, mRNA expression, and genome instability. **(A)** Sankey plot showing the correlations among 16 genes, DDR pathways, and prognostic value. **(B)** Left panel is the CNV variation frequency of 16 genes, and the deletion frequency is shown by gray dots; the middle panel is the mutation frequency of 16 genes between high- and low-DDR score samples. The right panel is the expression of 16 genes between normal tissues and tumor tissues. Tumor, gray; normal, yellow. **(C–O)** Comparison of the genome instability (NtAI, LST, LOH, HRD, Aneuploidy Score (AS), AS_del, AS_amp, TMB, SNP, and indel) and expression pattern of TP53, ATM, and ATR between high-risk and low-risk patients in the TCGA dataset. High, gray; low, yellow. **(P)** Expression profiles of 16 genes between high- and low-risk groups in the stages I and II TCGA–LAUD dataset. *p*-value of continuous variables was tested using Wilcoxon rank-sum test. Pearson's chi-square test was used to test the categorical variables.

*NBN*, and *POLD2* had a widespread frequency of CNV gain. Furthermore, the mRNA expressions for these genes were significantly higher in patients with CNV gain, suggesting CNV alteration may be a vital contributor to the altered mRNA expression of these genes. Moreover, protein expression levels of 13 genes were obtained from The Human Protein Atlas (THPA). Representative IHC images revealed that these proteins had upregulated expression in lung adenocarcinoma tissues and downregulated expression in normal lung tissues (**Supplementary Figure S1** and **Supplementary Table S3**).

To identify the biological significance of the genes in the DRG signature, GO analysis was conducted, and the results revealed that these genes were enriched in DNA-dependent DNA replication, nucleotide-excision repair, DNA recombination, and DNA geometric change. Furthermore, the outcomes of KEGG pathway analysis illustrated that these genes were mainly enriched in the Fanconi anemia pathway, base excision repair, and homologous recombination (**Supplementary Figure S2**).

Next, we investigated the associations between the two groups and various genomic features. The high-risk group was associated with higher aneuploidy score (AS), tumor mutational burden (TMB), SNP, and indel burden than the low-risk group (**Figure 2C–H**). Higher mRNA expression of ATM was observed in the low-risk group than in the high-risk group,

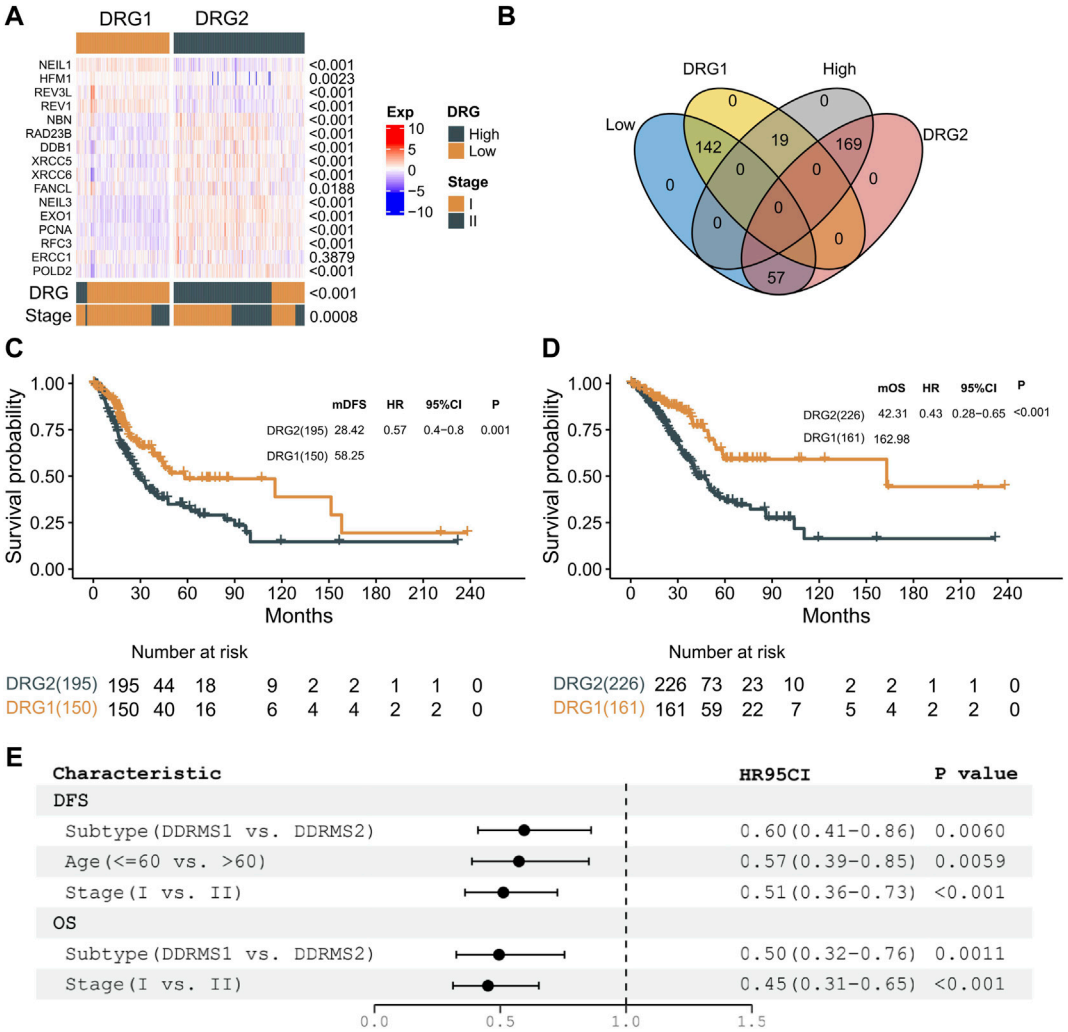



**FIGURE 3 |** Molecular subtype identification. **(A)** Expression profiles of 16 genes between DRG1 and DRG2 in the stages I and II TCGA–LAUD dataset. *p*-value of continuous variables was tested using the Wilcoxon rank-sum test. The consistency of DRG and molecular subtype was tested using the Kappa consistency test. Pearson's chi-square test was used to test the categorical variables. **(B)** Venn plot presenting the intersection of patient share by molecular subtype and DRG. **(C)** Kaplan–Meier curves showing DFS between DRG1 (yellow) and DRG2 (gray) in patients with early-stage LUAD. **(D)** Kaplan–Meier curves showing OS between DRG1 (yellow) and DRG2 (gray) in patients with early-stage LUAD. **(E)** Multivariable analysis of DFS and OS with a Cox proportional hazards model in early-stage lung carcinoma.

while not for TP53 and ATR (**Figure 2I–K**). We also observed that samples in the high-risk group exhibited higher genomic instability—telomeric allelic imbalance (TAI), large-scale state transitions (LST), loss of heterozygosity (LOH), and an incorporated homologous recombination deficiency (HRD) score (**Figure 2L-O**). These results showed the heterogeneity in genomic scar and DDR checkpoint gene expression between the two groups.

## Molecular Subtype Identification

As shown in **Figure 2P**, two expression patterns of the 16 genes were identified from the expression heat map of these signature genes in patients with stages I and II LUAD from the TCGA cohort. Patients in the low-risk group had better clinical outcomes (OS and DFS) and showed significantly higher expressions of *REV1*, *REV3L*, *HFM1*, and *NEIL1*, while the other genes had significantly lower expressions in this group. Meanwhile, the low-risk group had a higher percentage of patients with stage I than in the high-risk group (Chi test, $p = 0.0025$). The abovementioned results demonstrated that the DRG-related genes could be used to classify the early-stage LUAD patients.

Unsupervised hierarchical clustering (based on Euclidean distance and Ward's linkage) of the expression profiles of DDR genes was used to identify molecular subtype instead of the formula derived from the TCGA cohort. The expression profile of the 16 genes was used to develop a DRG-related molecular subtype to stratify early-stage LUAD into two

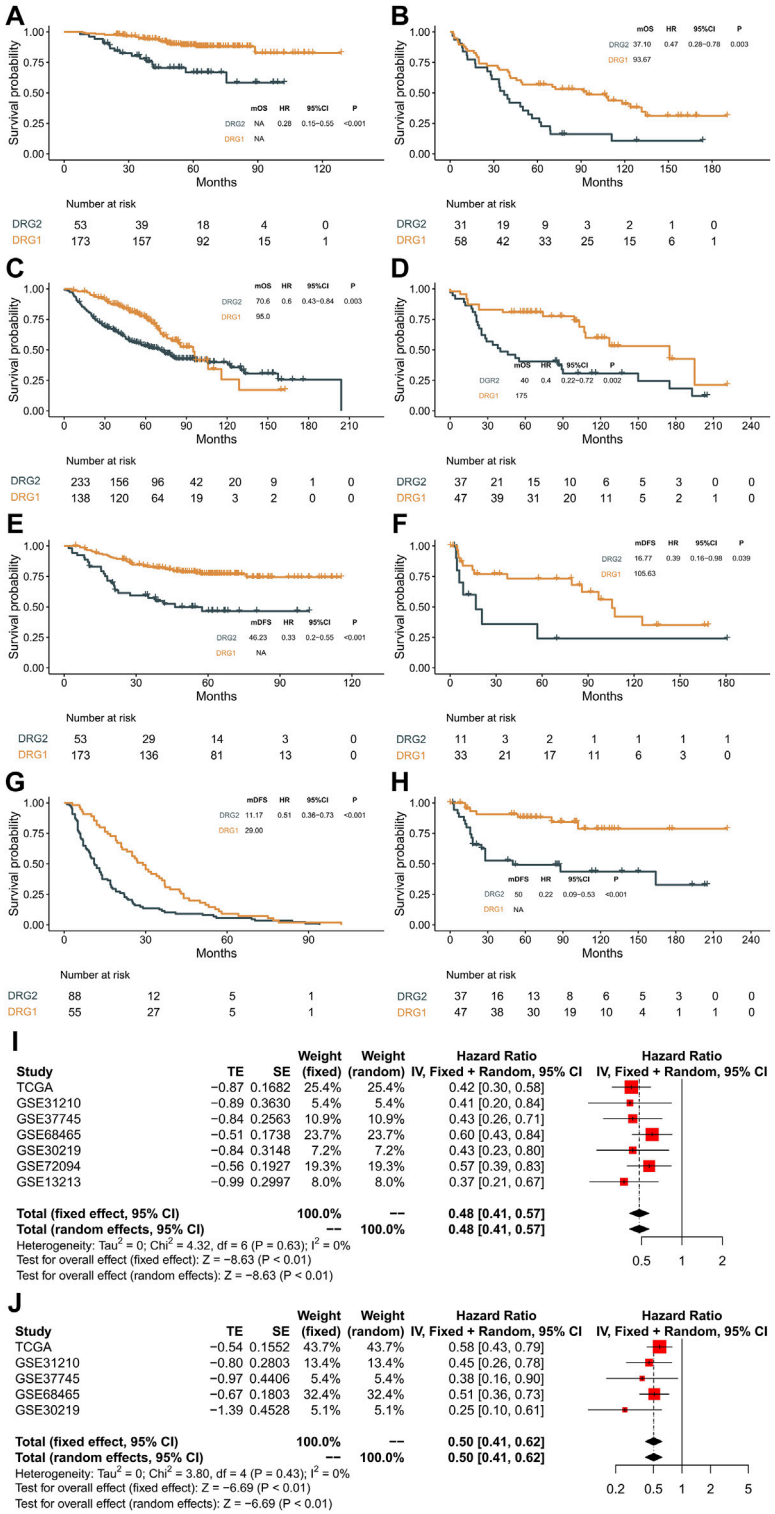



**FIGURE 4 |** Validation in the GEO datasets and meta-analysis. **(A–D)** Kaplan–Meier curves showing overall survival between DRG1 (yellow) and DRG2 (gray) in GSE31210, GSE37745, GSE68465, and GSE30219. **(E–H)** Kaplan–Meier curves showing disease-free survival between DRG1 (yellow) and DRG2 (gray) in GSE31210, GSE37745, GSE68465, and GSE30219. **(I)** Pooled estimates of overall survival. **(J)** Pooled estimates of disease-free survival.

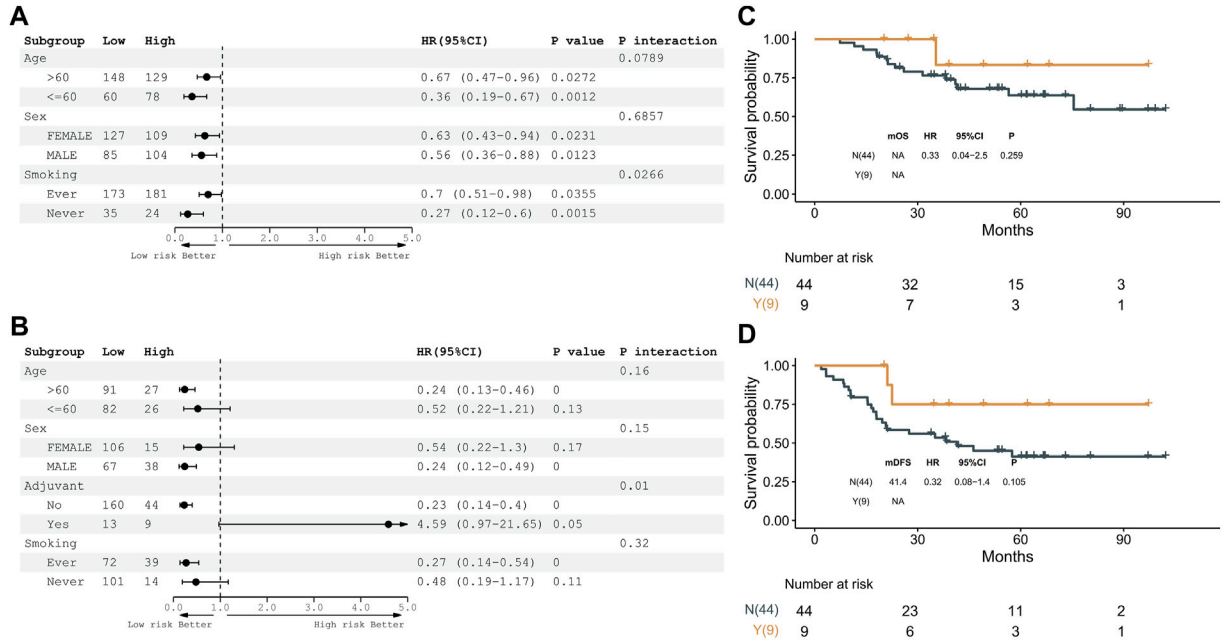



**FIGURE 5 |** Expression pattern of 16 genes is a prognostic biomarker and predicts adjuvant therapy benefits in the GSE31210 dataset. **(A)** Subgroup analyses of overall survival to estimate the clinical prognostic value between DRG1 and DRG2 as independent clinical factors. **(B)** Subgroup analyses of disease-free survival to estimate the clinical prognostic value between DRG1 and DRG2 in independent clinical factors. **(C)** Kaplan–Meier curves of overall survival between patients treated with or without adjuvant therapy. **(D)** Kaplan–Meier curves of disease-free survival between patients treated with or without adjuvant therapy.

subtypes (DRG1 and DRG2) with statistically significant differences in clinical outcomes. A clustering heat map was generated to illustrate that the expressions of DRG-related genes were significantly different between the two subtypes (**Figure 3A**). The Kappa consistency test revealed the consistency of the two methods (DRG and molecular subtype, K = 0.61, $p < 0.001$, **Figure 3A**). As shown in **Figure 3B**, 71.4% (142/199) of the low-risk DRG patients were grouped into DRG1 subtype, and 89.9% (169/188) of the high-risk DRG patients were grouped into DRG2 subtype. Similar results were also documented in the Kaplan–Meier analysis (DFS, log-rank $p$ = 0.001; HR = 0.57, 95% CI: 0.40–0.80; OS, log-rank $p < 0.001$; HR = 0.43, 95% CI: 0.28–0.65, **Figure 3C,D**). After adjusting for clinical factors, the molecular subtype remained an independent prognostic molecular classifier for DFS and OS (DFS, HR = 0.60, 95% CI: 0.41–0.86, $p$ = 0.006; OS, HR = 0.50, 95% CI: 0.32–0.76, $p$ = 0.011, **Figure 3E**). These results indicated that DRG-related genes could stratify early-stage LUAD into two molecular subtypes with distinct prognosis.

## Validation in GEO Datasets and Meta-analysis

In order to validate the molecular subtype and prognostic prediction of the DRG-related genes, a total of 1,285 stage III LUAD patient RNA expression microarray data were collected. The expression patterns of these genes and the survival status of patients in each GEO dataset are shown in **Figure 4** and **Supplementary Figure S3**. The patients in the DRG1 subtype had a longer OS and DFS than those in the DRG2 subtype (GSE31210: OS, log-rank $p < 0.001$, HR = 0.28, 95% CI: 0.15–0.55; DFS, log-rank $p < 0.001$, HR = 0.33, 95% CI: 0.20–0.55. GSE37745: OS, log-rank $p$ = 0.003, HR = 0.47, 95% CI: 0.28–0.78; DFS, log-rank $p$ = 0.039, HR = 0.39, 95% CI: 0.16–0.98. GSE68465: OS, log-rank $p$ = 0.003, HR = 0.60, 95% CI: 0.43–0.84; DFS, log-rank $p < 0.001$, HR = 0.50, 95% CI: 0.37–0.68. GSE30219: OS, log-rank $p$ = 0.002, HR = 0.40, 95% CI: 0.22–0.72; DFS, log-rank $p < 0.001$, HR = 0.22, 95% CI: 0.09–0.53. GSE72094: OS, log-rank $p$ = 0.002, HR = 0.49, 95% CI: 0.32–0.77. GSE13213: OS, log-rank $p < 0.001$, HR = 0.22, 95% CI: 0.01–0.49).

A meta-analysis was performed with a fixed-effects model, and the results indicated that compared with the DRG2 subtype, patients with the DRG1 subtype exhibited higher OS (HR = 0.48, 95% CI: 0.41–0.57, $p < 0.01$, **Figure 4I**) and DFS (HR = 0.5, 95% CI: 0.41–0.62, $p < 0.01$, **Figure 4J**) in the overall dataset. Heterogeneities were not significant in all pooled analyses (OS, $p$ = 0.63; DFS, $p$ = 0.43).

Then, we tested whether the molecular subtype could serve as an independent prognostic factor for early-stage lung adenocarcinoma. In multivariable analysis, the associations of DDR subtypes and prognosis were still significant (**Tables 2**, **3**), which confirmed that the selected DDR genes could stratify patients with different prognoses.

## Subgroup Analysis

A stratification analysis was conducted to assess whether clinical factors had interaction effects on the DRG subtypes. Patients in TCGA and GSE31210 datasets were artificially stratified based on

**TABLE 2 |** Univariable analysis and multivariable Cox regression analyses of OS in six validation cohorts.

| Cohort | Characteristics | Sample size | Univariable Cox | | Multivariable Cox | |
|---|---|---|---|---|---|---|
| | | | HR (95% CI) | *P*-value | HR (95% CI) | *P*-value |
| GSE31210 | | | | | | |
| — | Age (≤60 vs. >60) | 226 | 0.79 (0.4–1.54) | 0.4860 | — | — |
| — | Sex (Male vs. Female) | 226 | 1.52 (0.78–2.96) | 0.2190 | — | — |
| — | Stage (II vs. I) | 226 | 4.23 (2.17–8.24) | <0.001 | 3.23 (1.59–6.55) | 0.0012 |
| — | Smoking (Never vs. Ever) | 226 | 0.61 (0.31–1.19) | 0.1500 | — | — |
| — | Adjuvant (Y vs. N) | 226 | 2.04 (0.79–5.27) | 0.1420 | — | — |
| — | Subtype (DRG1 vs. DRG2) | 226 | 0.28 (0.15–0.55) | 0.0002 | 0.41 (0.2–0.83) | 0.0139 |
| GSE37745 | | | | | | |
| — | Age (≤60 vs. >60) | 106 | 0.82 (0.52–1.3) | 0.3980 | — | — |
| — | Sex (Male vs. Female) | 106 | 1.26 (0.8–1.97) | 0.3160 | — | — |
| — | Stage (III_IV vs. I_II) | 106 | 1.79 (1.01–3.17) | 0.0449 | 2.35 (1.28–4.3) | 0.0056 |
| — | Subtype (DRG1 vs. DRG2) | 106 | 0.51 (0.32–0.82) | 0.0053 | 0.43 (0.26–0.71) | 0.0009 |
| GSE68465 | | | | | | |
| — | Age (≤60 vs. >60) | 371 | 0.53 (0.37–0.74) | 0.0002 | 0.49 (0.34–0.7) | 0.0001 |
| — | Sex (Male vs. Female) | 371 | 1.43 (1.06–1.93) | 0.0198 | 1.52 (1.11–2.09) | 0.0097 |
| — | Smoking (Never vs. Ever) | 297 | 1.13 (0.67–1.91) | 0.6360 | — | — |
| — | Subtype (DRG1 vs DRG2) | 371 | 0.6 (0.43–0.84) | 0.0030 | 0.6 (0.43–0.85) | 0.0036 |
| GSE30219 | | | | | | |
| — | Age (≤60 vs. >60) | 84 | 0.66 (0.36–1.21) | 0.1780 | — | — |
| — | Sex (Male vs Female) | 84 | 1.05 (0.51–2.19) | 0.8870 | — | — |
| — | Stage (II vs I) | 84 | 2.04 (1.04–3.97) | 0.0370 | 1.68 (0.85–3.34) | 0.1370 |
| — | Subtype (DRG1 vs DRG2) | 84 | 0.4 (0.22–0.72) | 0.0027 | 0.43 (0.23–0.79) | 0.0068 |
| GSE72094 | | | | | | |
| — | Age (≤60 vs >60) | 398 | 0.72 (0.42–1.23) | 0.2300 | — | — |
| — | Sex (Male vs Female) | 398 | 1.55 (1.07–2.25) | 0.0198 | 1.78 (1.22–2.6) | 0.0028 |
| — | Stage (III_IV vs I_II) | 393 | 2.61 (1.74–3.91) | <0.001 | — | — |
| — | Smoking (Never vs Ever) | 331 | 0.73 (0.32–1.68) | 0.4590 | 2.81 (1.86–4.24) | <0.001 |
| — | Subtype (DRG1 vs DRG2) | 398 | 0.56 (0.39–0.81) | 0.0023 | 0.57 (0.39–0.83) | 0.0030 |
| GSE13213 | | | | | | |
| — | Age (≤60 vs >60) | 117 | 0.96 (0.55–1.69) | 0.8900 | — | — |
| — | Sex (Male vs Female) | 117 | 1.36 (0.77–2.39) | 0.2860 | — | — |
| — | Stage (II vs I) | 92 | 1.57 (0.64–3.82) | 0.3240 | 1.38 (0.57–3.38) | 0.4780 |
| — | Stage (III vs I) | 104 | 3.2 (1.74–5.88) | 0.0002 | 2.73 (1.48–5.05) | 0.0014 |
| — | Smoking (Never vs Ever) | 117 | 1.36 (0.77–2.39) | 0.2840 | — | — |
| — | Subtype (DRG1 vs DRG2) | 117 | 0.33 (0.18–0.6) | 0.0002 | 0.37 (0.21–0.68) | 0.0012 |

clinical factors, such as age (≤60/>60), sex (female/male), stage (I/II), smoking (ever/never), and adjuvant treatment (no/yes). As shown in **Figure 5A,B** and Supplementary Figure S4A-B, patients in the DRG1 subtype had higher OS and DFS than the DRG2 subtype irrespective of their age, sex, and smoking status. Meanwhile, a significant interaction ($p$ = 0.01) between adjuvant treatment and DRG subtypes was observed in early-stage LUAD patients. Furthermore, we examined the association between adjuvant treatment and prognosis in DRG1 and DRG2 subtypes. We found that in the DRG2 subtype, patients with adjuvant treatment tended to have longer OS and DFS than patients without adjuvant treatment (OS, log-rank $p$ = 0.259, HR = 0.33, 95% CI: 0.04–2.5; DFS, log-rank $p$ = 0.105, HR = 0.32, 95% CI: 0.08–1.4), while in the DRG1 subtype, the results were opposite (OS, log-rank $p$ = 0.001, HR = 5.3, 95% CI: 1.7–16; DFS, log-rank $p$ < 0.001, HR = 6.7, 95% CI: 3.0–15). The abovementioned observation was not statistically significant because the sample size was limited (**Supplementary Figure S4C,D** and **Figure 5C,D**). After the patients were matched by propensity score, similar results were observed (**Supplementary Figure S5**).

## DISCUSSION

In this study, we trained and validated 16 DDR genes with prognostic values and classification effects in early-stage lung adenocarcinoma and classified patients into two subtypes, DRG1 and DRG2. Furthermore, we found that the DRG1 patients without adjuvant therapy and the DRG2 patients with adjuvant therapy tended to have prolonged survival than other patients in the corresponding subtypes.

The DDR system comprised eight pathways with diverse biological functions to maintain genomic integrity. In this study, we discovered that the 16 identified DDR genes were mainly involved in TLS, NER, and BER pathways. *REV3L* and *REV1* were involved in TLS whose lower expressions were associated with worse prognoses. In human cells, when the expressions of TLS genes decrease, the DNA replication stress escalates the accumulative fork stalling and double-strand breaks (DSBs), resulting in genome instability and poor survival (Ghosal and Chen, 2013). These two genes are important DNA polymerase and deoxycytidyl transferase, which play significant roles in maintaining genome stability in the advent of DNA damage (Sasatani et al., 2020; Zhou et al., 2020). It has been reported that lower *REV3L* expression was also shown to

**TABLE 3 |** Univariable analysis and multivariable Cox regression analyses of DFS in four validation cohorts.

| Cohort | Characteristics | Sample size | Univariable Cox | | Multivariable Cox | |
|---|---|---|---|---|---|---|
| | | | HR (95% CI) | *P*-value | HR (95% CI) | *P*-value |
| GSE31210 | | | | | | |
| — | Age (≤60 vs >60) | 226 | 0.61 (0.37–1.02) | 0.0575 | — | — |
| — | Sex (Male vs Female) | 226 | 1.27 (0.78–2.07) | 0.3380 | — | — |
| — | Stage (II vs I) | 226 | 3.16 (1.92–5.21) | 0.0000 | 2.4 (1.36–4.21) | 0.0024 |
| — | Smoking (Never vs Ever) | 226 | 0.75 (0.46–1.23) | 0.2520 | — | — |
| — | Adjuvant (Y vs N) | 226 | 2.37 (1.2–4.66) | 0.0127 | 1.07 (0.51–2.26) | 0.8540 |
| — | Subtype (DRG1 vs DRG2) | 226 | 0.33 (0.2–0.55) | 0.0000 | 0.45 (0.26–0.78) | 0.0044 |
| GSE37745 | | | | | | |
| — | Age (≤60 vs >60) | 53 | 1.23 (0.56–2.71) | 0.6140 | — | — |
| — | Sex (Male vs Female) | 53 | 1.07 (0.48–2.38) | 0.8680 | — | — |
| — | Stage (III_IV vs I_II) | 53 | 1.72 (0.64–4.61) | 0.2800 | 2.06 (0.75–5.64) | 0.1610 |
| — | Subtype (DRG1 vs DRG2) | 53 | 0.42 (0.18–0.98) | 0.0441 | 0.38 (0.16–0.9) | 0.0282 |
| GSE68465 | | | | | | |
| — | Age (≤60 vs >60) | 143 | 1.01 (0.7–1.44) | 0.9700 | — | — |
| — | Sex (Male vs Female) | 143 | 1.05 (0.76–1.47) | 0.7570 | — | — |
| — | Smoking (Never vs Ever) | 130 | 1.02 (0.63–1.66) | 0.9210 | — | — |
| — | Subtype (DRG1 vs DRG2) | 143 | 0.51 (0.36–0.73) | 0.0002 | 0.51 (0.36–0.73) | 0.0002 |
| GSE30219 | | | | | | |
| — | Age (≤60 vs >60) | 84 | 0.83 (0.38–1.79) | 0.6280 | — | — |
| — | Sex (Male vs Female) | 84 | 1.17 (0.44–3.11) | 0.7480 | — | — |
| — | Stage (II vs I) | 84 | 3.35 (1.49–7.57) | 0.0036 | 2.71 (1.19–6.19) | 0.0180 |
| — | Subtype (DRG1 vs DRG2) | 84 | 0.22 (0.09–0.53) | 0.0007 | 0.25 (0.1–0.59) | 0.0018 |

be associated with lower DFS and OS (Agulló-Ortuño et al., 2020), which was consistent with our findings.

Furthermore, we discovered that genes with higher expression in the DRG2 subtype were mainly involved in NER and BER pathways. RAD23B, DDB1, and ERCC family genes (ERCC1, ERCC5, and ERCC6) are key genes in the NER pathway. Many studies have reported that they are significantly correlated with prognosis in different cancer types, such as colorectal cancer, pancreatic cancer, gastric cancer, and so on (Luo et al., 2018; Zhang et al., 2019; Li et al., 2021). *NEIL3*, *PCNA*, *RFC3*, and *POLD2* play important roles in the BER pathway which are recruited to DNA lesions and cleave and repair the damaged bases cooperatively (Robertson et al., 2009; Hurst et al., 2021; Wang et al., 2021). Zhao et al. found that *NEIL3* activated cell cycle progression, leading to poor prognosis (Zhao et al., 2021). Zhang et al. discovered that RFC3 was involved in the epithelial–mesenchymal transition in lung adenocarcinoma, resulting in worse survival (Gong et al., 2019). In tumorigenesis, increased DNA replication stress results in the increased generation of reactive oxygen species (ROS), leading to DNA damage (Jackson and Bartek, 2009). Accumulating evidence supports that NER and BER pathways are involved in the repair of oxidative DNA lesions. Therefore, high expressions of NER and BER genes suggest that more oxidative DNA lesions are being generated, which lead to genome instability and poor prognosis (Melis et al., 2013). Therefore, the imbalance of DNA damage and repair can increase the genome instability and promote tumor cell proliferation which might contribute to worse survival.

The use of adjuvant therapy in early-stage (IA–IIB) lung adenocarcinoma is controversial in NCCN guidelines and mainly depends on the physician's experience (Zheng and Bueno, 2015). Although several studies have constructed various gene expression signatures to stratify LUAD patients, none of them have provided sufficient evidence about whether the high-risk patients could benefit from adjuvant therapy (Chen et al., 2007; Shedden et al., 2008; Sun et al., 2008). In our study, we classified LUAD patients into DRG1 and DRG2 subtypes and explored the interaction between these subtypes and adjuvant therapy in the GSE31210 dataset, which was a relatively rigorous clinical trial with clear inclusion and exclusion criteria. The patients in GSE31210 received no neoadjuvant therapies before surgery, whose stages were pathologically defined. Based on the GSE31210 cohort, we found that in the DRG2 subtype, the prognosis of patients who received adjuvant therapy had prolonged survival than those who did not, whereas in the DRG1 subtype, the patients without adjuvant therapy had better prognosis. Several previous studies also revealed that the low activity of TLS including low expression of *REV3L* enhanced the chemosensitivity of cancer (Wang et al., 2015; Yang et al., 2015; Agulló-Ortuño et al., 2020), which supports our findings that patients in the DRG2 subtype may benefit from adjuvant chemotherapy. In summary, the different clinical benefits of adjuvant therapy in various subtypes suggest that DRG subtypes have the potential to guide the selection of adjuvant therapy for early-stage LUAD patients.

In the present study, we identified novel DDR-gene expression subtypes and explored the association with prognosis and adjuvant therapy. However, there are still some limitations in our study. The current study was a retrospective analysis with a limited sample size in a public database. In addition, the mRNA expression in our study was based on RNA-seq or microarray whose results were less stable than those of RT-PCR or IHC, so the evidence would be more solid if it was validated by RT-PCR or IHC, as well as more cost-effective in clinical application scenarios (Pisapia et al., 2022). However, our findings have been validated in six independent cohorts to reduce false-positive results, and they were further validated in the adjuvant therapy subgroups to confirm the role in guiding therapy selection. In the future, prospective studies with large sample sizes are required to confirm the clinical utility of the 16 DDR-gene expression subtypes on the platform of RT-PCR or IHC.

In summary, we explored the association between DDR gene expression and prognosis in patients with stage I or II LUAD. Sixteen DDR gene–related subtypes were constructed to predict prognosis and guide the use of adjuvant therapy. More research studies are warranted to further confirm the clinical utility of the 16 DDR-gene classifiers.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

Conceptualized and designed by JD and LC; administrative support provided by JD and LC; provision of study materials or patients by YZ, BQ, CX, JZ, YL, PC, GW, SC, and YS; collection and assembly of data by BQ and JZ. Data analysis and interpretation conducted by YZ, CX, and YL. All authors contributed to the writing of manuscript. All authors have provided final approval of the manuscript.

## FUNDING

This work was supported by the General Program of National Natural Science Foundation of China (81972905).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.901829/full#supplementary-material.

## REFERENCES

Agulló-Ortuño, M. T., García-Ruiz, I., Díaz-García, C. V., Enguita, A. B., Pardo-Marqués, V., Prieto-García, E., et al. (2020). Blood mRNA Expression of REV3L and TYMS as Potential Predictive Biomarkers from Platinum-Based Chemotherapy Plus Pemetrexed in Non-small Cell Lung Cancer Patients. *Cancer Chemother. Pharmacol.* 85, 525–535. doi:10.1007/s00280-019-04008-9

Beer, D. G., Kardia, S. L. R., Huang, C.-C., Giordano, T. J., Levin, A. M., Misek, D. E., et al. (2002). Gene-Expression Profiles Predict Survival of Patients with Lung Adenocarcinoma. *Nat. Med.* 8, 816–824. doi:10.1038/nm733

Bender, E. (2014). Epidemiology: The Dominant Malignancy. *Nature* 513, S2–S3. doi:10.1038/513S2a

Chen, H.-Y., Yu, S.-L., Chen, C.-H., Chang, G.-C., Chen, C.-Y., Yuan, A., et al. (2007). A Five-Gene Signature and Clinical Outcome in Non-small-cell Lung Cancer. *N. Engl. J. Med.* 356, 11–20. doi:10.1056/nejmoa060096

Ghosal, G., and Chen, J. (2013). DNA Damage Tolerance: a Double-Edged Sword Guarding the Genome. *Transl. Cancer Res.* 2, 107–129. doi:10.3978/j.issn.2218-676X.2013.04.01

Gobin, M., Nazarov, P. V., Warta, R., Timmer, M., Reifenberger, G., Felsberg, J., et al. (2019). A DNA Repair and Cell-Cycle Gene Expression Signature in Primary and Recurrent Glioblastoma: Prognostic Value and CLinical Implications. *Cancer Res.* 79, 1226–1238. doi:10.1158/0008-5472.CAN-18-2076

Gong, S., Qu, X., Yang, S., Zhou, S., Li, P., and Zhang, Q. (2019). RFC3 Induces Epithelial-mesenchymal Transition in Lung Adenocarcinoma Cells through the Wnt/β-catenin Pathway and Possesses Prognostic Value in Lung Adenocarcinoma. *Int. J. Mol. Med.* 44. doi:10.3892/ijmm.2019.4386

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

Howington, J. A., Blum, M. G., Chang, A. C., Balekian, A. A., and Murthy, S. C. (2013). Treatment of Stage I and II Non-Small Cell Lung Cancer: Diagnosis and Management of Lung Cancer, 3rd ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 143, e278S–e313S. doi:10.1378/chest.12-2359

Hurst, V., Challa, K., Shimada, K., and Gasser, S. M. (2021). Cytoskeleton Integrity Influences XRCC1 and PCNA Dynamics at DNA Damage. *Mol. Biol. Cell.* 32, br6. doi:10.1091/mbc.E20-10-0680

Jackson, S. P., and Bartek, J. (2009). The DNA-Damage Response in Human Biology and Disease. *Nature* 461, 1071–1078. doi:10.1038/nature08467

Knijnenburg, T. A., Wang, L., Zimmermann, M. T., Chambwe, N., Gao, G. F., Cherniack, A. D., et al. (2018). Genomic and Molecular Landscape of DNA Damage Repair Deficiency across the Cancer Genome Atlas. *Cell Rep.* 23, 239–e6. doi:10.1016/j.celrep.2018.03.076

Kratz, J. R., He, J., Van Den Eeden, S. K., Zhu, Z. H., Gao, W., Pham, P. T., et al. (2012). A Practical Molecular Assay to Predict Survival in Resected Non-squamous, Non-small-cell Lung Cancer: Development and International Validation Studies. *Lancet* 379, 823–832. doi:10.1016/S0140-6736(11)61941-7

Li, B., Cui, Y., Diehn, M., and Li, R. (2017). Development and Validation of an Individualized Immune Prognostic Signature in Early-Stage Nonsquamous Non-Small Cell Lung Cancer. *JAMA Oncol.* 3, 1529. doi:10.1001/jamaoncol.2017.1609

Li, J., Tian, L., Jing, Z., Guo, Z., Nan, P., Liu, F., et al. (2021). Cytoplasmic RAD23B Interacts with CORO1C to Synergistically Promote Colorectal Cancer Progression and Metastasis. *Cancer Lett.* 516, 13–27. doi:10.1016/j.canlet.2021.05.033

Listì, A., Barraco, N., Bono, M., Insalaco, L., Castellana, L., Cutaia, S., et al. (2018). Immuno-targeted Combinations in Oncogene-Addicted Non-small Cell Lung Cancer. *Transl. Cancer Res.* 8, S55–S63. doi:10.21037/tcr.2018.10.04

Luo, S. S., Liao, X. W., and Zhu, X. D. (2018). Prognostic Value of Excision Repair Cross-Complementing mRNA Expression in Gastric Cancer. *Biomed. Res. Int.* 2018. doi:10.1155/2018/6204684

Melis, J. P. M., Van Steeg, H., and Luijten, M. (2013). Oxidative DNA Damage and Nucleotide Excision Repair. *Antioxidants Redox Signal.* 18, 2409–2419. doi:10.1089/ars.2012.5036

Necchi, A., Joseph, R. W., Loriot, Y., Hoffman-Censits, J., Perez-Gracia, J. L., Petrylak, D. P., et al. (2017). Atezolizumab in Platinum-Treated Locally Advanced or Metastatic Urothelial Carcinoma: Post-Progression Outcomes from the Phase II IMvigor210 Study. *Ann. Oncol.* 28, 3044–3050. doi:10.1093/annonc/mdx518

Pang, F.-M., Yan, H., Mo, J.-L., Li, D., Chen, Y., Zhang, L., et al. (2020). Integrative Analyses Identify a DNA Damage Repair Gene Signature for Prognosis Prediction in Lower Grade Gliomas. *Future Oncol.* 16, 367–382. doi:10.2217/fon-2019-0764

Passiglia, F., Galvano, A., Gristina, V., Barraco, N., Castiglia, M., Perez, A., et al. (2021). Is There Any Place for PD-1/CTLA-4 Inhibitors Combination in the First-Line Treatment of Advanced NSCLC?-a Trial-Level Meta-Analysis in PD-L1 Selected Subgroups. *Transl. Lung Cancer ResLung Cancer Res.* 10, 3106–3119. doi:10.21037/TLCR-21-52

Pisapia, P., Pepe, F., Baggi, A., Barberis, M., Galvano, A., Gristina, V., et al. (2022). Next Generation Diagnostic Algorithm in Non-Small Cell Lung Cancer Predictive Molecular Pathology: The KWAY Italian Multicenter Cost Evaluation Study. *Crit. Rev. Oncology/Hematology* 169, 103525. doi:10.1016/j.critrevonc.2021.103525

Robertson, A. B., Klungland, A., Rognes, T., and Leiros, I. (2009). DNA Repair in Mammalian Cells: Base Excision Repair: the Long and Short of it. *Cell. Mol. Life Sci.* 66, 981–993. doi:10.1007/s00018-009-8736-z

Sasatani, M., Zaharieva, E. K., and Kamiya, K. (2020). The *In Vivo* Role of Rev1 in Mutagenesis and Carcinogenesis. *Genes Environ* 42, 9. doi:10.1186/s41021-020-0148-1

Scarbrough, P. M., Weber, R. P., Iversen, E. S., Brhane, Y., Amos, C. I., Kraft, P., et al. (2016). A Cross-Cancer Genetic Association Analysis of the DNA Repair and DNA Damage Signaling Pathways for Lung, Ovary, Prostate, Breast, and Colorectal Cancer. *Cancer Epidemiol. Biomarkers Prev.* 25, 193–200. doi:10.1158/1055-9965.EPI-15-0649

Shedden, K., Taylor, J. M. G., Enkemann, S. A., Tsao, M. S., Yeatman, T. J., Gerald, W. L., et al. (2008). Gene Expression-Based Survival Prediction in Lung Adenocarcinoma: A Multi-Site, Blinded Validation Study. *Nat. Med.* 14, 822–827. doi:10.1038/nm.1790

Sun, Z., Wigle, D. A., and Yang, P. (2008). Non-Overlapping and Non-Cell-Type-Specific Gene Expression Signatures Predict Lung Cancer Survival. *J. Clin. Oncol* 26, 877–883. doi:10.1200/JCO.2007.13.1516

Sun, H., Cao, D., Ma, X., Yang, J., Peng, P., Yu, M., et al. (2019). Identification of a Prognostic Signature Associated with DNA Repair Genes in Ovarian Cancer. *Front. Genet.* 10, 839. doi:10.3389/fgene.2019.00839

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Wang, W., Sheng, W., Yu, C., Cao, J., Zhou, J., Wu, J., et al. (2015). REV3L Modulates Cisplatin Sensitivity of Non-small Cell Lung Cancer H1299 Cells. *Oncol. Rep.* 34, 1460–1468. doi:10.3892/or.2015.4121

Wang, X., Janowczyk, A., Zhou, Y., Thawani, R., Fu, P., Schalper, K., et al. (2017). Prediction of Recurrence in Early Stage Non-Small Cell Lung Cancer Using Computer Extracted Nuclear Features from Digital H&E Images. *Sci. Rep.* 7, 13543. doi:10.1038/s41598-017-13773-7

Wang, W., Yin, Q., Guo, S., and Wang, J. (2021). NEIL3 Contributes toward the Carcinogenesis of Liver Cancer and Regulates PI3K/Akt/mTOR Signaling. *Exp. Ther. Med.* 22, 1053. doi:10.3892/etm.2021.10487

Wistuba, I. I., Behrens, C., Lombardi, F., Wagner, S., Fujimoto, J., Raso, M. G., et al. (2013). Validation of a Proliferation-Based Expression Signature as Prognostic Marker in Early Stage Lung Adenocarcinoma. *Clin. Cancer Res.* 19, 6261–6271. doi:10.1158/1078-0432.CCR-13-0596

Yang, L., Shi, T., Liu, F., Ren, C., Wang, Z., Li, Y., et al. (2015). REV3L, a Promising Target in Regulating the Chemosensitivity of Cervical Cancer Cells. *Plos One* 10, e0120334. doi:10.1371/journal.pone.0120334

Zhang, Y., Lei, Y., Xu, J., Hua, J., Zhang, B., Liu, J., et al. (2019). Role of Damage DNA-Binding Protein 1 in Pancreatic Cancer Progression and Chemoresistance. *Cancers* 11, 1998. doi:10.3390/cancers11121998

Zhao, C., Liu, J., Zhou, H., Qian, X., Sun, H., Chen, X., et al. (2021). NEIL3 May Act as a Potential Prognostic Biomarker for Lung Adenocarcinoma. *Cancer Cell Int.* 21, 228. doi:10.1186/s12935-021-01938-4

Zheng, Y., and Bueno, R. (2015). Commercially Available Prognostic Molecular Models in Early-Stage Lung Cancer: a Review of the Pervenio Lung RS and Myriad myPlan Lung Cancer Tests. *Expert Rev. Mol. Diagnostics* 15, 589–596. doi:10.1586/14737159.2015.1028371

Zhou, Y.-K., Li, X.-P., Yin, J.-Y., Zou, T., Wang, Z., Wang, Y., et al. (2020). Association of Variations in Platinum Resistance-Related Genes and Prognosis in Lung Cancer Patients. *J. Cancer* 11, 4343–4351. doi:10.7150/jca.44410

**Conflict of Interest:** JZ, YL, PC, GW, and SC are employees of Burning Rock Biotech.

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

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

frontiers | Frontiers in Molecular Biosciences

TYPE Review
PUBLISHED 19 August 2022
DOI 10.3389/fmolb.2022.903065



OPEN ACCESS

EDITED BY
Umberto Malapelle,
University of Naples Federico II, Italy

REVIEWED BY
Alessandro Rizzo,
National Cancer Institute Foundation (IRCCS), Italy
Concetta Saponaro,
National Cancer Institute Foundation (IRCCS), Italy
Jabed Iqbal,
Singapore General Hospital, Singapore

*CORRESPONDENCE
João Nuno Moreira,
jmoreira@ff.uc.pt
João Goncalves,
joao.goncalves@ff.ulisboa.pt

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 23 March 2022
ACCEPTED 13 July 2022
PUBLISHED 19 August 2022

CITATION
Ribeiro R, Carvalho MJ, Goncalves J and Moreira JN (2022), Immunotherapy in triple-negative breast cancer: Insights into tumor immune landscape and therapeutic opportunities. *Front. Mol. Biosci.* 9:903065. doi: 10.3389/fmolb.2022.903065



# Immunotherapy in triple-negative breast cancer: Insights into tumor immune landscape and therapeutic opportunities

Rita Ribeiro[1,2,3], Maria João Carvalho[3,4,5,6,7], João Goncalves[2]* and João Nuno Moreira[1,3]*

[1]CNC—Center for Neurosciences and Cell Biology, Center for Innovative Biomedicine and Biotechnology (CIBB), University of Coimbra, Faculty of Medicine (Polo 1), Coimbra, Portugal, [2]iMed.ULisboa—Research Institute for Medicines, Faculty of Pharmacy, University of Lisbon, Lisbon, Portugal, [3]Univ Coimbra—University of Coimbra, CIBB, Faculty of Pharmacy, Coimbra, Portugal, [4]CHUC—Coimbra Hospital and University Centre, Department of Gynaecology, Coimbra, Portugal, [5]Univ Coimbra—University Clinic of Gynaecology, Faculty of Medicine, University of Coimbra, Coimbra, Portugal, [6]iCBR—Institute for Clinical and Biomedical Research Area of Environment Genetics and Oncobiology (CIMAGO), Faculty of Medicine, University of Coimbra, Coimbra, Portugal, [7]CACC—Clinical Academic Center of Coimbra, Coimbra, Portugal

Triple-negative breast cancer (TNBC) is a clinically aggressive subtype of breast cancer that represents 15–20% of breast tumors and is more prevalent in young pre-menopausal women. It is the subtype of breast cancers with the highest metastatic potential and recurrence at the first 5 years after diagnosis. In addition, mortality increases when a complete pathological response is not achieved. As TNBC cells lack estrogen, progesterone, and HER2 receptors, patients do not respond well to hormone and anti-HER2 therapies, and conventional chemotherapy remains the standard treatment. Despite efforts to develop targeted therapies, this disease continues to have a high unmet medical need, and there is an urgent demand for customized diagnosis and therapeutics. As immunotherapy is changing the paradigm of anticancer treatment, it arises as an alternative treatment for TNBC patients. TNBC is classified as an immunogenic subtype of breast cancer due to its high levels of tumor mutational burden and presence of immune cell infiltrates. This review addresses the implications of these characteristics for the diagnosis, treatment, and prognosis of the disease. Herein, the role of immune gene signatures and tumor-infiltrating lymphocytes as biomarkers in TNBC is reviewed, identifying their application in patient diagnosis and stratification, as well as predictors of efficacy. The expression of PD-L1 expression is already considered to be predictive of response to checkpoint inhibitor therapy, but the challenges regarding its value as biomarker are described. Moreover, the rationales for different formats of immunotherapy against TNBC currently under clinical research are discussed, and major clinical trials are highlighted. Immune checkpoint inhibitors have demonstrated clinical benefit, particularly in early-stage tumors and when administered in combination with chemotherapy, with several regimens approved by the regulatory authorities. The success of antibody–drug conjugates and research on other emerging

approaches, such as vaccines and cell therapies, will also be addressed. These advances give hope on the development of personalized, more effective, and safe treatments, which will improve the survival and quality of life of patients with TNBC.

KEYWORDS

triple-negative breast cancer, biomarkers, tumor mutational burden, immune gene signatures, infiltrating T lymphocytes, immunotherapy, immune checkpoint inhibitors

# Introduction

Cancer is the second leading cause of death in the world, with 9.6 million deaths in 2018 and incidence rates growing worldwide (Bray et al., 2018). Breast cancer is the most frequent malignancy in women, accounting for one of four cancer cases, and it represents approximately 15% of all cancer deaths (Bray et al., 2018). It is a particularly stressful condition due to its associated physical, emotional, social, and economic burden (Luengo-Fernandez et al., 2013), despite the major efforts that have been made in treating the disease.

Breast cancer is a set of diseases with distinct pathological features and clinical outcomes that reflect different gene signatures and molecular patterns. Recognizing this diversity, classifications of breast tumors have emerged to aid in the diagnosis, treatment, and prognosis of the disease. Since 1968, the World Health Organization has been publishing up-to-date versions of a Histological Classification of Breast Tumors, a collection of histological and molecular pathology features for breast cancer diagnosis that is provided by experts in the field. Later, as gene expression advanced, Perou and Sorlie classified breast tumors into four subtypes: luminal A and luminal B, normal-like, and human epidermal growth factor receptor 2 (HER2)-enriched (Perou et al., 2000). A surrogate classification based on histological (hormone receptor expression) and molecular (Ki67 proliferation marker index) features is used in clinics, dividing breast cancers into five subtypes: luminal A-like, luminal B-like HER2-, luminal B-like HER2+, HER2 enriched, and triple-negative (Harbeck et al., 2019). The latter constitutes the subject of this review.

# Triple-negative breast cancer

## Epidemiology

Triple-negative breast cancer (TNBC) is a subtype that represents 15–20% of all breast tumors (Diana et al., 2018). It is more common in young pre-menopausal and African-American or Hispanic women (Morris et al., 2007).

Triple-negative tumors are particularly relevant due to their aggressiveness. TNBC is more frequently diagnosed in the advanced stages, with a preference for visceral metastasis and to a lesser extent in the bone (Klimov et al., 2017). Recurrence rates are also higher in TNBC, mostly in the first 5 years of diagnosis (Dent et al., 2007; Reddy et al., 2018). Despite their sensitivity to chemotherapy, with increased pathological complete response (pCR) rates relative to other subtypes of breast cancer (22–45 vs. 7%, respectively (Fornier and Fumoleau, 2012)), TNBC patients have higher mortality (37% mortality within the first 6 months of diagnosis (VazLuis et al., 2017)), worse cause-specific and overall survival (Li et al., 2017a) when a complete response is not achieved. This contradiction between responsiveness and poor survival is now referred to as "the paradox of TNBC," and it drives the imperative need for new and more effective therapeutics against the disease.

## Genetic and molecular features

Triple-negative breast tumors are characterized by low expression or absence of expression of progesterone (PR) and estrogen (ER) receptors and by the lack of overexpression or amplification of human epidermal growth factor receptor-2 (HER2).

The advance in transcriptomic technologies has allowed a deeper characterization of the genetic and molecular features of triple-negative tumors, with implications for diagnosis and the search for new therapies. Gene expression profiling showed that 15–20% of women with TNBC carry BRCA1 or BRCA2 gene mutations or deficiencies, which impair DNA stability and repair, promoting carcinogenesis (Turner et al., 2004; Wong-Brown et al., 2015). Microarray assays demonstrated that 20% of TNBCs are claudin-low, characterized by genomic instability and a propensity for epithelial-to-mesenchymal transition (EMT) and thus more prone to metastization. In triple-negative tumors, 80% are basal-like (Foulkes et al., 2010), featuring the expression of proliferation-related genes (e.g., Ki67), the presence of a high number of mutations of tumor suppressor genes (TP53, RB1, and BRCA1) and of the PIK3CA oncogene (The Cancer Genome Atlas Network, 2012; Shi et al., 2018), along with the expression of proliferation and EMT-related molecules, such as cytokeratines (CK5/6, CK14, and CK17), epidermal growth factor receptor (EGFR), and vimentin (Rakha and Ellis, 2009). These properties corroborate TNBC's aggressiveness and poor clinical prognosis (Rakha et al., 2008; Sabatier et al., 2014).

Based on this genetic and molecular heterogeneity, Lehman et al. suggested a division of TNBCs into four distinct subtypes:

two highly proliferative basal-like (BL1 and BL2), a mesenchymal-like (M) involved in cell motility and EMT, and a luminal androgen receptor (LAR) enriched in hormonally regulated pathways, driven by the androgen receptor (Lehmann et al., 2011; Lehmann et al., 2016).

## Diagnosis

The diagnosis of TNBC relies on the combined data from pathology and mainly immunohistochemistry. The aforementioned subtyping of triple-negative tumors by gene expression analysis is an important tool in better understanding TNBCs' biology, but its predictive value has not yet been established in clinical routine (Lehmann et al., 2016).

Imaging techniques alone (mammography and ultrasound) are not sufficient to distinguish TNBC from other breast cancers. However, it has been demonstrated that certain morphological features, such as circumscribed margins and the absence of calcifications, are common across TNBC tumors but atypical in other subtypes (Dogan and Turnbull, 2012).

Certain pathological attributes (tumor size, lymph infiltrate status, proliferation index, and necrosis) are well documented and enable us to distinguish between TNBC and non-TNBC as well as grade tumors. TNBCs are characterized by large lymphoplasmacytic infiltrate, stromal fibrosis and tumor necrosis, vascular and nerve invasion, and a high rate of proliferation (Rapoport et al., 2014; Abdollahi and Etemadi, 2016).

Immunohistochemistry (IHC) provides the most accurate data for TNBC diagnosis, as it allows the assessment of ER, PR, and HER2 status in patient samples (Penault-Llorca and Viale, 2012). Over the years, guidelines were written to standardize IHC techniques and the respective thresholds across laboratories. Samples are now considered ER/PR-negative when they present <1% of immunoreactive cells in IHC slides (Hammond et al., 2010), and are negative for HER2 with an IHC result of 0 or 1 for membrane protein expression (defined as no staining or weak/incomplete membrane staining) (Wolff et al., 2007). For uncertain HER2 results, the American Society of Clinical Oncology (ASCO) recommends a confirmatory analysis, performed by fluorescence *in situ* hybridization, which detects false positives and false negatives (Wolff et al., 2018). Genetic counseling and BRCA mutation status testing may also be required at the time of diagnosis, as it can influence the choice of therapeutic regimen.

## Triple-negative breast cancer therapy

### Standard therapy and its challenges

Over the years, great advances have been made in breast cancer therapy, with proven efficacy of anti-HER2 antibodies in HER2-positive tumors (trastuzumab and pertuzumab) (Costa and Czerniecki, 2020), and targeted endocrine therapy in hormone-positive cancers (ER modulators—tamoxifen, and aromatase inhibitors—anastrazole) (Tremont et al., 2017). TNBC treatment, however, remains a challenge, especially due to the absence of the therapeutic targets expressed in other breast cancers. The heterogeneous nature of TNBC is also a problem, with different subtypes of triple-negative tumors demonstrating different sensitivities to available treatments (Yin et al., 2020). Consequently, no official recommendations have been created on how to treat TNBC. Nevertheless, the National Comprehensive Cancer Network (NCCN) in the United Stated of America and the European Society for Medical Oncology (ESMO) has published guidelines to help manage the disease (Cardoso et al., 2019; NCCN, 2020; Paluch-Shimon et al., 2020; Gennari et al., 2021).

The primary treatment approach against any localized breast tumor is surgical removal, conservative or radical, according to focality and clinical conditions, followed by adjuvant radiotherapy for locoregional management and chemotherapy. Pharmacological treatment is used as a neoadjuvant to minimize tissue removal and evaluate tumor sensitivity to chemotherapy. Adjuvant therapy can eliminate residual cancer cells and prevent relapse. Neoadjuvant treatment is particularly recommended in triple-negative tumors due to their high sensitivity to chemotherapy, especially in premenopausal women (Houssami et al., 2012).

Anthracyclines and taxanes are the main chemotherapeutic regimens against TNBC. Anthracyclines, such as doxorubicin, are molecules that inhibit topoisomerase II, blocking DNA replication and transcription and, consequently, arresting the cell cycle. Taxanes (e.g., paclitaxel and docetaxel) are antimitotic agents that inhibit cell division by affecting the stabilization of microtubules. Platinum-based compounds, such as carboplatin and cisplatin, interlink DNA strands, causing them to break and leading to cell apoptosis. This is particularly beneficial in the case of tumors that carry BRCA gene mutations, with underlying impaired DNA repairing mechanisms, and prevalent among TNBC patients. Other drugs, such as cyclophosphamide (causing DNA damage), fluorouracil, and capecitabine (blocking DNA synthesis) have also been used, particularly in combination or in sequential regimens with anthracyclines and/or taxanes or when the latter are contraindicated (Cardoso et al., 2019).

In early-stage TNBC, neoadjuvant chemotherapy is the standard of care. Triple-negative patients tend to have better responses to neoadjuvant treatments than non-TNBCs (Von Minckwitz et al., 2012), and patients that achieve pCR after neoadjuvant therapy have better survival outcomes (Huober et al., 2010). Still, only about 33% of patients present a complete response to standard neoadjuvant treatments (Cortazar et al., 2014). In TNBCs, the risk of recurrence and death is increased when the residual disease remains after the first

3 years after neoadjuvant treatment (Liedtke et al., 2008). Patients at an early stage of disease who do not receive neoadjuvant therapy should undergo adjuvant treatment (Cardoso et al., 2019). Platinum agents also have demonstrated efficacy in the neoadjuvant setting of TNBC, either as single agents (70% pCR rate) or in addition to standard neoadjuvant chemotherapy (22–75% pCR) (Sikov et al., 2015; Caramelo et al., 2019).

Adjuvant therapy with anthracyclines and/or taxanes remains the first line for advanced and metastatic TNBC (Al-Mahmood et al., 2018). Capecitabine has also been considered in the management of residual disease after neoadjuvant treatment for its good outcomes in clinical trials in terms of disease-free survival (DFS—69.8%), OS (78.8%) (Masuda et al., 2017), and pCR (33.6%) (Martín et al., 2019). In this setting, platinum therapy has failed to demonstrate a clear benefit against standard first-line anthracycline regimens (Pandy et al., 2018).

Safety remains a concern in currently available TNBC therapies due to anthracycline-induced cardiotoxicity (Barrett-Lee et al., 2009) and paclitaxel-associated hypersensitivity, neutropenia, and neurotoxicity (Patt et al., 2006). Nab-paclitaxel was developed to overcome taxane toxicity and increase the extent of tumor delivery. Nab-paclitaxel consists of a colloidal suspension of albumin-bound paclitaxel nanoparticles, a formulation that allows better pharmacokinetics and safety profile than free (solvent-based) paclitaxel (Schettini et al., 2016). The former presented a pCR of 48% in the neoadjuvant setting, in contrast with the 26% enabled by the latter (Untch et al., 2016).

However, TNBCs lack the benefit resulting from the use of targeted or hormonal systemic therapies in other subtypes. In this regard, a deeper knowledge of the molecular characteristics of TNBCs paved the way for the development of novel targeted therapeutics (Won and Spruck, 2020) and patient stratification.

## The search for targeted therapies

Inhibitors targeting poly (ADP ribose) polymerase (PARP), an enzyme involved in DNA repair pathways, impair DNA repairing mechanisms, leading to tumor cell death. Olaparib is an inhibitor of PARP-1, PARP-2, and PARP-3 enzymes, and has been approved by the European Medicines Agency (EMA) and the Food and Drug Administration (FDA) for the treatment of patients with germline BRCA1/2-mutations, who are HER2-negative and have locally advanced or metastatic breast cancer, and who have been treated with anthracyclines and taxanes (European Medicines Agency, 2014; Food and Drug Administration, 2014). The fact that both EMA and the FDA also approved olaparib monotherapy as first-line treatment for other BRCA-mutated cancers (advanced BRCA-mutated ovarian cancer) gives hope for its use for other BRCA-deficient tumors, such as TNBC, as a first-line option (Montemorano et al., 2019). More recently, talazoparib, an inhibitor of PARP-1 and PARP-2 enzymes, was approved by both FDA (2018) and EMA (2019) as an alternative for patients with germline BRCA mutations and HER2-negative locally advanced or metastatic breast cancer who have been previously treated with anthracycline and/or taxane (Food and Drug Administration, 2018; European Medicines Agency, 2019). Veliparib and niraparib are PARP-1 and PARP-2 inhibitors that are under clinical investigation for the treatment of TNBC (Geenen et al., 2018). Velaparib demonstrated efficacy when used in combination with EGFR-inhibitor lapatinib (Stringer-Reasor et al., 2021) or cisplatin (Sharma et al., 2020) but did not improve pCR when added to a standard neoadjuvant plus carboplatin regimen (Loibl et al., 2018). Niraparib's clinical activity was evaluated in a phase II trial in combination with pembrolizumab immune checkpoint inhibitor (Vinayak et al., 2019) and a phase III trial as monotherapy against standard chemotherapy (Turner et al., 2021). However, due to the small sample sizes in the trials and information-censoring issues, no accurate conclusions could be drawn about its effectiveness in TNBC treatment.

Because a large percentage of triple-negative tumors have been identified as basal-like and thus have a high proliferation index, antimitotic agents (taxanes) are considered a targeted therapy. In this respect, response markers are important for determining which patients will benefit the most from taxane therapy. A microtubule-associated protein (Bauer et al., 2010) and mitotic and ceramide metagenes (Juul et al., 2010) are examples of markers that are associated with higher pCR levels in patients treated with neoadjuvant paclitaxel with basal-like-TNBC than in other triple-negative subtypes or non-triple-negative tumors. Further, as previously mentioned, many basal-like tumors express EGFR and thus are a potential target for EGFR inhibitors such as cetuximab. This has been tried against basal-like TNBC in clinical trials, either alone or in combination with chemotherapy (paclitaxel or carboplatin), although with only modest efficacy (Costa et al., 2017).

Angiogenesis inhibitors are another approach against TNBC, as the overexpression of VEGF in these tumors is higher than it is in non-TNBC (Linderholm et al., 2009). Bevacizumab, an anti-VEGF antibody, has shown moderate results in a neoadjuvant setting, either as first-line (Miles et al., 2013) or in combination with chemotherapy (Bell et al., 2017). In advanced and metastatic tumors, so far, it has failed to demonstrate a robust benefit (Manso et al., 2015). In 2008, FDA approved the use of bevacizumab in combination with paclitaxel for the treatment of metastatic breast cancer. However, in 2010, bevacizumab's indication for breast carcinoma was withdrawn, after it was shown not to be safe and effective in this indication (Food and Drug Administration, 2019). In Europe, the use of bevacizumab is still in place, combined with paclitaxel or capecitabine for the first-line treatment of metastatic breast cancer.

Due to the highly proliferative nature of triple-negative tumors, inhibitors of the PI3K/Akt/mTOR pathway, which regulates the cell cycle, are also considered as potential anti-

TNBC therapies (Khan et al., 2019). This is based on demonstrations that mutations in PI3K are more prevalent in TNBC than in other breast cancers (Koboldt et al., 2012), as well as on the activation of the mTOR pathway and its correlation with poor prognosis (Pelicano et al., 2014). Ipatasertib is a PI3K inhibitor that showed modest but positive efficacy as neoadjuvant therapy in PTEN-mutant patients when used in combination with paclitaxel by demonstrating a progression-free survival (PFS) of 6.2 vs. 3.7 months in the group receiving paclitaxel as monotherapy (Kim et al., 2017), and a pCR of 16 vs. 13%, respectively (Oliveira et al., 2019). Everolimus is an m-TOR inhibitor that has been tested in combination with carboplatin (Singh et al., 2014), liposomal doxorubicin/bevacizumab (Basho et al., 2018), gemcitabine/cisplatin (Park et al., 2018), and cisplatin/paclitaxel (Jovanovic et al., 2017). Despite moderate efficacy, the concerns that emerged regarding hematological toxicity (neutropenia and thrombocytopenia) demand further research.

Patients with the LAR subtype of TNBC, enriched in androgen receptor expression and hormonally regulated pathways, may benefit from anti-androgen therapy. Androgen receptor inhibitors (enzalutamide and bicalutamide) (Gucalp et al., 2013; Traina et al., 2018), androgen synthesis inhibitors (abiraterone acetate) (ClinicalTrials.gov, 1842), and a combination of androgen inhibitors with PI3K inhibitors (Lehmann et al., 2014) are among the approaches currently receiving clinical attention.

## Chemoresistance: A real threat

Chemoresistance is a growing concern in TNBC therapy, with about 30–50% of patients undergoing neoadjuvant therapy evolving to resistant recurrences, resulting in poor outcomes (Kim et al., 2018). Mechanisms of resistance arise when tumor cells are exposed to cytotoxic agents, as a means of maintaining their viability. Some of these mechanisms have been demonstrated for TNBC standard therapies, and strategies to overcome them have been proposed (O'Reilly et al., 2015).

It was previously demonstrated that resistance to anthracyclines is associated with reduced expression or function of the target DNA repair enzyme, topoisomerase II (Nielsen et al., 1996), while taxane resistance has been linked to β-tubulin III overexpression (Tommasi et al., 2007).

ATP-binding cassette (ABC) transporters are a family of transmembrane proteins that promote drug efflux, preventing its action in the cell. In TNBC cells, the three ABC transporters responsible for resistance to anthracyclines and taxanes have been identified: multidrug-resistant protein-1 (MRP1), breast cancer resistance protein (ABCG2), and P-glycoprotein (MDR1) pump (Scharenberg et al., 2002; Leonessa and Clarke, 2003). Attempts to hinder ABC resistance have been studied for several cancers, including breast cancers. The strategies used include inhibition of ABC activity (with non-steroidal anti-inflammatory drugs (O'Connor et al., 2007) or tyrosine kinase inhibitors (Nakai et al., 2016)) and expression (with microRNA (Wang et al., 2015)).

### The stemness phenotype

A strong hypothesis that could explain the emergence of chemoresistant breast tumors is the maintenance of a minor quiescent population, namely, cancer stem cells (CSC). It has been demonstrated that TNBC patient tissues and cell lines present a higher abundance of CD44+/CD24-stem cells, as well as EMT-related gene signatures, than other subtypes, in line with the stemness phenotype (Park et al., 2010; Zhou et al., 2016; Shibue and Weinberg, 2017). Evidence of this phenomenon also comes from an increase in RNA transcripts, the expression of TGF-β1 and TGF-β type 1 receptors, and molecules associated with CSCs and EMT mechanisms in TNBC tumor biopsies after chemotherapy (Bhola et al., 2013). Strategies to circumvent this resistance may include selective inhibitors of CSC (Gupta et al., 2009), antagonists of CSC membrane markers (e.g., CD44), or targeting of TGF-β signaling (O'Conor et al., 2018).

Hypoxia-inducible factors (HIFs) are transcription factors that maintain oxygen homeostasis in the cell. HIF expression is increased in tumors, in response to the insufficient supply of oxygen resulting from the rapid and disordered neo-angiogenesis (Hanahan and Weinberg, 2011). In TNBC, HIFs have been associated with the CSC phenotype, chemoresistance, and poor prognosis (Jin et al., 2016; Xiong et al., 2018). In fact, the co-administration of HIF inhibitors and chemotherapy have been demonstrated to overcome CSC-mediated resistance, as exemplified by the ability of digoxin to reverse resistance to paclitaxel and gemcitabine in mouse models of TNBC (Samanta et al., 2014). In spite of the search for molecular targets involved in hypoxia to overcome chemoresistance, no clinical benefit has been reported so far in this context (Soni and Padwad, 2017).

### Genomic diversity

The genomic diversity characteristic of TNBC could also explain the appearance of chemotherapy-resistant relapses. In 2018, Santonja et al. correlated the different subtypes of TNBC with patient responsiveness and resistance to treatment. The LAR subtype had lower pCR rates, suggesting higher resistance to standard chemotherapy, in line with its predicted responsiveness to endocrine therapy (Santonja et al., 2018). This genomic diversity is also reflected in genes that are often mutated in TNBC patients and correlated with apoptosis avoidance, a drug-resistance mechanism (e.g., the p53 (Lin et al., 2019) and Bcl-2 genes) (Inao et al., 2018). The higher frequency of mutations found in treated TNBCs relative to naive tumors may support the hypothesis that certain gene alterations, such as in tumor suppressor genes (e.g., p53, and PTEN), or those associated with cell proliferation, such as MYC, likely form the basis for the development of chemoresistance (Koboldt et al., 2012; Balko et al., 2014).

Finally, the signaling pathways involved in cell growth and proliferation may be implicated in TNBC chemoresistance. NF-κB regulates the transcription of genes involved in TNBC progression (Poma et al., 2017) and chemoresistance in breast cancers (Fan et al., 2008). As mentioned, the PTEN/PI3K/AKT/mTOR pathway is often hyperactivated in TNBC and has been associated with chemoresistance in breast cancer (Steelman et al., 2008). The JAK/STAT pathway consists of a phosphorylation cascade of proteins that regulates the transcription of genes involved in tumorigenesis, survival, and anti-apoptosis. Because TNBCs enriched in JAK amplifications have been associated with the worst prognosis, it is hypothesized that this pathway may be also involved in chemoresistance (Balko et al., 2014). Different STAT proteins have been associated with distinct outcomes in TNBC, with STAT3 and STAT5 activation being linked to chemoresistance and their inhibition to therapy sensitivity (Furth, 2014; Mumin et al., 2019).

The lack of therapeutic targets, along with the rise of chemoresistance, reinforces the need for the development of new approaches against TNBC.

## Triple-negative breast cancer: An immunogenic subtype

The microenvironment of breast tumors is composed of not only cancer cells but also of a variety of other cellular and non-cellular components, such as endothelial cells, immune cells, fibroblasts and adipocytes, the extracellular matrix, and chemical mediators. The composition of the tumor microenvironment differs among the various breast cancer subtypes and understanding it is important to improve diagnosis and provide more effective treatment of TNBC (Yu and Di, 2017).

The immune component of tumors has already been intensively studied, particularly since 2011, when Hanahan and Weinberg suggested that it is involved in carcinogenesis, proposing the ability of tumors to evade attack and elimination by the immune system as a revised hallmark of cancer (Hanahan and Weinberg, 2011). The work of Dunn et al. contributed to the understanding the crosstalk between tumors and the immune system is much more complex, for both immunodeficient and immunocompetent individuals, by recognizing that, not only do tumors have mechanisms for escaping immunologic defenses but they can also be shaped by their immune surroundings, in a process named immunoediting (Dunn et al., 2004). These findings are particularly important for immunotherapies because it has been observed, although the reason has not been clearly understood, that only a small subset of patients benefit from these therapeutics. For example, in 2018, only about 13% of patients enrolled in checkpoint inhibitors treatments in the United States were responsive to treatment (Haslam and Prasad, 2019).

In the tumor microenvironment, both innate and adaptive immune responses are triggered. Innate immune cells include antigen-presenting cells (APC), macrophages, neutrophils, and monocytes. These constitute the first response to foreign elements and help the adaptive immune cells, i.e., T and B lymphocytes, recognize the neoantigens expressed in tumor cells. T lymphocytes are a set of cells with distinct roles that can be differentiated by their cell surface markers. $CD8^+$ lymphocytes are cytotoxic T cells that recognize tumor antigens and eliminate malignant cells by releasing pro-inflammatory cytokines, such as interferons (INF) and interleukins (IL), as well as granzyme–perforin complexes. $CD4^+$ T lymphocytes are helper cells that can differentiate into Th1 or Th2 cells. Th1 cells secrete INFα, INFγ, and IL-2, which are able to activate macrophages and NK cells against tumor cells and are thus predictive of a good prognosis. On the other hand, Th2 cells produce and release anti-inflammatories IL-4, IL-5, and IL-10, which promote tumor growth and metastasis. Regulatory T cells (Tregs) are a class of lymphocytes that express forkhead box P3 (FOXP3) transcription factor, and CD25 surface marker. They have immunosuppressive activity, thus being correlated with a worse prognosis (Gajewski et al., 2013; Lee et al., 2019). T lymphocytes, like tumor cells, express immune inhibitory programmed cell death receptor-1 (PD-1), its ligand programmed death ligand-1 (PD-L1), and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), which is responsible for suppressing immune activity (He and Xu, 2020). Consequently, the immune system is either a promoter or suppressor of tumor growth, depending on the equilibrium of existing immune cells and cytokines. Its understanding and modulation have changed the cancer treatment paradigm.

In this regard, the gene signatures of immune cells (gene-transcription and proliferation related genes) and tumor-infiltrating lymphocytes (TILs) have major implications for tumor development, clinical response, and prognostic value. This knowledge may be particularly important for tumors that have limited treatment options and worse prognoses, such as TNBC. Although breast cancers are not as immunogenic as other solid tumors that have benefited from immunotherapies, such as renal cancer, non-small cell lung cancer (NSCLC), or melanoma, evidence shows that TNBC and HER2+ are more immunogenic than the hormone-positive subtypes, and the study of the TNBC immunologic landscape has provided valuable information on immunogenicity and immune activity (Liu et al., 2018).

The advancement of sequencing technologies, such as whole genome sequencing (WGS), next generation sequencing (NGS), and RNA-seq profiling has been an important contribution to the growing and clearer perception of the role of tumor immune microenvironment in cancer development, treatment, and prognosis.

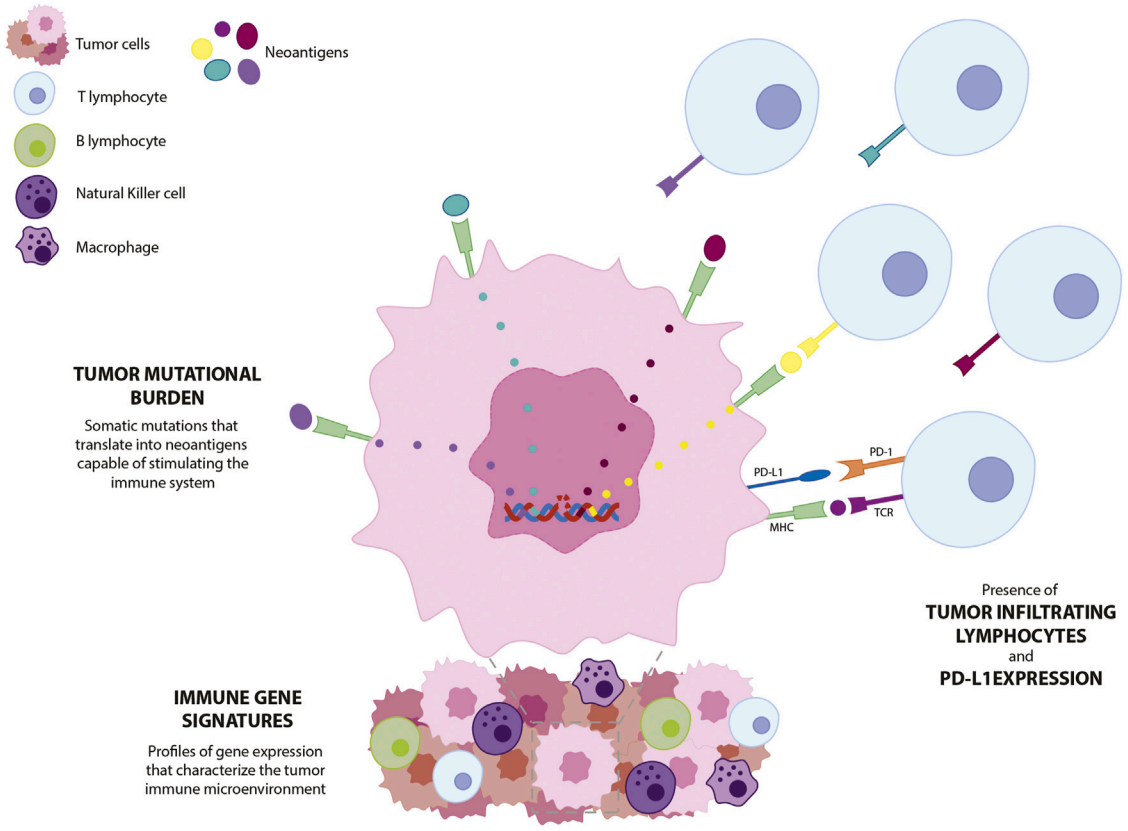



FIGURE 1
Schematic representation of biomarkers currently explored in TNBC. Point mutations in tumor cells translate into expression of tumor-specific neoantigens. Neoantigens induce T lymphocyte infiltration and increased expression of PD-1 and PD-L1 checkpoint molecules. Gene profiling of tumor biopsies allows the characterization of the immune components of the tumor. PD-L1—programmed death-ligand 1; PD-1—programmed cell death protein 1; MHC—major histocompatibility complex: TCR—T-cell receptor.

# The importance of immune biomarkers

## Tumor mutational burden

Tumor immunogenicity refers to the ability of a tumor to elicit an immune response. This depends mostly on the relative and absolute density of the antigens capable of activating the immune system, whether they are shared antigens (already present in normal tissues but overexpressed in tumor cells) or tumor-specific neoantigens (mutated proteins in tumors, not present in normal tissues) (Blankenstein et al., 2012). Neoantigens are a consequence of non-synonymous somatic mutations that result in peptides or proteins expressed at the surface of tumor cells but not of normal cells, which makes them an ideal target for immunotherapy, as represented in Figure 1.

Determination of the number of non-synonymous somatic mutations occurring in a tumor (mutations/megabase), i.e., the tumor mutational burden (TMB), demonstrated that neoantigens generated by point mutations in normal genes may be tumor specific. This is very important for breaking tumor heterogeneity and for clustering patients with similar genetic signatures, thus enabling superior patient stratification to benefit from a certain treatment for use as a predictive of therapy response (Samstein et al., 2019). In fact, TMB is predictive of the response to immunotherapy in a variety of cancers, as different tumors with higher TMB values showed an increased objective response rate (ORR) when treated with immunotherapies (Yarchoan et al., 2017; Loibl et al., 2019). Additionally, the establishment of TMB as a predictive factor in immunotherapy has been recently validated by FDA's approval of checkpoint inhibitor pembrolizumab for patients with solid tumors presenting high TMB (defined to have ≥10 mutations/megabase) (FDA, 2020c). This agnostic approval was based on the results of the KEYNOTE-158 study that concluded that high TMB is predictive of pembrolizumab's efficacy against a variety of advanced solid tumors, such as small cell lung cancer, mesothelioma, and neuroendocrine, biliary, vulvar, and anal cancers (Marabelle et al., 2020).

Although breast cancer is not among the group of cancers with higher TMBs, hormone-negative tumors such as TNBC and

TABLE 1 Immune gene signatures that characterize antitumor immune response, including pan-cancer signature, breast cancer signatures, and TNBC signature. Gene signatures include genes involved in immune cell activation and proliferation, and expression of molecules that regulate the immune response, such as chemokines and interferons.

| Immune gene signature | Gene | Immune function | Reference |
|---|---|---|---|
| Pan-cancer immune gene signature | *CD84* | Cytotoxic T cells | Liu et al. (Liu, 2019) |
| | *CD79A* | B cells | |
| | *EVI2B* | Neutrophils | |
| | *GATA3*<br>*STAT* | T-helper cell regulation | |
| | *FOXP3* | T regulatory cells | |
| | *CD68* | Macrophages | |
| | *KLRC1* | NK cells | |
| | *LILRA4* | Dendritic cells | |
| | *TBX21* | IFNγ production | |
| SDPP signature (good prognosis) | *CD8A*, *CD247*, *CD3D* | T cells | Finak et al. (2008) |
| | *CD8A* | MHC class I protein binding | |
| | *GZMA*, *GZMB* | Granzymes A and B | |
| SDPP (bad prognosis) | *GIMAP5* | T cells | |
| | *ADM* | Adrenomedullin | |
| | *CXCL14* | NK cells activity | |
| | *IL8* | Interleukin-8 | |
| | *EDN1* | Endothelin-1 | |
| | *SPP1* | Osteopontin | |
| | *CXCL1* | Chemokine ligand | |
| Immune response module | *STAT1*, *STAT4* | Lymphocytes activity | Desmedt et al. (2008) |
| | *CD69* | T cells | |
| | *CD48* | B cells | |
| | *CXCL9*, *CXCL10*, *CXCL11*, *CCL5*, *CCL8* | Chemokine ligands | |
| Immune response and regulation signature | *CD27*, *CD52* | T cells | Yang et al. (2018) |
| | *GZMA*, *GZMK* | Granzymes | |
| | *CCR2* | Chemokine receptor | |
| | *CCL5*, *CXCL9* | Chemokine ligands | |
| Four genes immune signature in TNBC | *CXCL13* | Chemokine ligand | Criscitiello et al. (2018) |
| | *GBP1* | Interferon-induced protein | |
| | *SULT1E1* | Immunogenic protein | |
| | *HLF* | Immunogenic cell death | |

CD, Cluster of Differentiation; EVI, Ecotropic Viral Integration Site; GATA, GATA binding protein; STAT, Signal Transducer and Activator of Transcription; FOXP3, Forkhead box P3; KLRC1, Killer Cell Lectin Like Receptor C1; LILRA4, Leukocyte Immunoglobulin Like Receptor A4; TBX21, T-Box Transcription Factor 21; GZM, Granzyme; GIMAP5, GTPase, IMAP Family Member 5; ADM, Adrenomedullin; CXCL, Chemokine Ligand (C-X-C motif); CCL, Chemokine Ligand (C-C motif); IL, Interleukin; EDN, Endothelin; SPP1, Secreted Phosphoprotein 1; CCR, Chemokine Receptor; GBP1, Guanylate Binding Protein 1; SULT1E1, Sulfotransferase Family 1E Member 1; HLF, Hepatic leukemia factor.

HER2+ were observed to have significantly higher TMB than other subtypes (Barroso-Sousa et al., 2020). In this regard, an association between TMB and immune-mediated patient survival was demonstrated for patients with a favorable immune-infiltrate disposition, while patients with high TMB and poor immune-infiltrates showed a worse prognosis (Thomas et al., 2018). In another study, a better prognosis was also correlated with lower TMB and neoantigen counts, possibly due to the lower clonal heterogeneity due to immunosurveillance mechanisms, which eliminate cells with higher neoantigen expression (Karn et al., 2017). Therefore, the potential value of TMB as a biomarker for triple-negative breast tumors is still controversial, and its clinical role is far from being significant.

To improve the value of TMB as a biomarker in cancer therapy, further investigation is needed, including studies that have larger sample sizes, standardized sample treatments and sequencing parameters, and harmonization of clinical output

reporting. The organizations Friends of Cancer Research and the Quality Assurance Initiative Pathology published a set of recommendations to improve TMB estimation and report in the clinics (Stenzinger et al., 2019).

## Immune gene signatures

Immune gene signatures are the profiles of gene expression used to characterize the immune response in tumors. Transcriptomic techniques and in-depth statistical analysis are used in cancer cell lines or patient-derived tumor samples to identify clusters of genes whose expression is involved in the immune response to cancer, as well as to characterize the immune cell composition of tumors (Figure 1).

For breast cancer, the regulatory authorities have approved four commercial multi-gene expression assays with prognostic value, Prosigna® (PAM50), MammaPrint®, Oncotype DX®, and Endopredict® (Cardoso et al., 2019), which are based on gene signatures that are reflected in the tumor phenotype and its response to its environment. Specifically regarding the immune response in breast cancer, no clinical signatures have yet been made available. Nevertheless, the genes involved in the immune response to cancer and the underlying prognostic value have been identified. Finak *et al.* demonstrated that two gene clusters from a set of 163 genes were correlated with good or bad prognosis in breast cancer, termed the stroma-derived prognostic predictor (SDPP; Table 1). Good outcomes were related to the enrichment of gene coding for T lymphocyte and NK cell activation, as well as to granzyme activity, suggesting an immune response to the tumor. Genes associated with bad outcomes (higher risk of recurrence or death from the disease) code for molecules involved in carcinogenic mechanisms, such as angiogenesis, hypoxia, EMT, and tumor-associated macrophages (immunosuppression and metastasis) (Finak et al., 2008).

Some signatures are tumor-specific, which can help differentiate immune mechanisms across cancer subtypes and potentially lead to patient stratification and targeted treatment. For TNBCs, these immune signatures may have particular importance due to the absence of common molecular features in breast cancer (ER, PR, and HER2 expression). Desmedt *et al.* studied immune gene signatures for more than 2000 breast cancers and associated the different subtypes with different biological processes (tumor invasion/metastasis, immune response, angiogenesis, evasion of apoptosis, growth and proliferation, and ER and HER2 signaling), and each of these with a prognosis value (Table 1). In this study, only ER/HER2 tumors presented a clinical outcome that was significantly associated with the immune response module, where the high expression levels of these genes were correlated with increased relapse-free survival (Desmedt et al., 2008). Yang *et al.* used a 17-gene immune signature to identify an immune-enhanced group of breast cancers, with characteristics of ER- and claudin-low tumors, which was correlated with better clinical outcomes (lower risk of recurrence, metastasis, or death) (Yang et al., 2018) (Table 1).

As patients responding to neoadjuvant therapy tend to have longer disease-free survival, and chemotherapy is hypothesized to activate the tumor's immune system in breast cancer (Kroemer et al., 2015), immune gene signatures that are predictive of prognosis after neoadjuvant treatment may become particularly relevant to clinical practice. Regarding the immune response after therapy, studies in samples from patients after neoadjuvant therapy correlated immune signatures with recurrence-free survival (Rody et al., 2011), increased pCR (Ignatiadis et al., 2012; Lee et al., 2015), and prolonged survival (Stoll et al., 2015; Hendrickx et al., 2017). Criscitiello *et al.* assessed gene immune signatures specifically in TNBC patients who were previously treated with neoadjuvant therapy. *CXCL13* (chemokine), *GBP1* (interferon response), *SULT1E1* (estrogen homeostasis, potentially immunogenic), and *HLF* (immunogenic cell death) were correlated with chemotherapy-induced immune response and increased distant relapse-free survival (Criscitiello et al., 2018) (Table 1).

Importantly, many of these immune signatures show little overlap, which hinders their usefulness for prognosis or therapeutic choice in clinical practice. Further research is needed, including large clinical trials designed to identify and validate gene markers, to facilitate the translation from the laboratory to the clinic.

## Tumor infiltrating lymphocytes and PD-L1 expression

Notwithstanding that healthy breast tissue has a low number of immune cells, tumor development is associated with leucocytes infiltrating the area, with B and T lymphocytes, Tregs, and neutrophils representing the main immune populations (Althobiti et al., 2018).

Tumors with highly proliferative characteristics, such as the triple-negative and HER2-positive subtypes, present increased levels of TILs. In TNBC, in particular, this is explained by its increased genomic instability and mutational burden, with consequent stimulation of the immune system for elimination of cells bearing non-self-antigens (Smid et al., 2011) (Figure 1). For these breast cancer subtypes, the presence of immune infiltrates is associated with a good prognosis in patients treated with standard neoadjuvant chemotherapy or trastuzumab (for HER2-overexpressing tumors) (Denkert et al., 2018; Ochi et al., 2019).

Triple-negative tumors are also enriched in both $CD8^+$ T cells and T reg cells relative to other breast cancer subtypes (Stanton et al., 2016). As expected, the presence of cytotoxic T cells is associated with a good prognosis in TNBC in the early stage of

the disease (Blackley and Loi, 2019), and with response prediction to neoadjuvant (Denkert et al., 2018) and adjuvant chemotherapy (Adams et al., 2014; Pruneri et al., 2016). In addition, the low area fraction of stromal TILs and deficiency of $CD8^+$ cells are indicative of an increased risk of mortality (Vihervuori et al., 2019). Unexpectedly, T regulatory cells were found to be predictive of good prognosis in TNBC, despite their role in suppressing the activity of immune cells. TNBC patients who had higher expression of FOXP3 had higher OS and PFS than TNBC patients with lower levels of FOXP3 (Lee et al., 2012; Jiang et al., 2015; Yeong et al., 2017). However, these results may be misleading, as there is a need for optimization of T reg cells identification. This was exemplified by a study correlating FOXP3/CD25+ Treg cells with improved OS in TNBC patients but not Tregs identified by the FOXP3 marker alone (Zhang et al., 2019).

An important source of immune cells in the tumor microenvironment is the presence of tertiary lymphoid structures (TLS). These vascularized clusters of lymphoid cells have a similar structure and function to secondary lymphoid organs, as they are mainly composed of T cells, dendritic cells, plasma cells, and B cells, promoting cellular and humoral anti-tumor response (Wang et al., 2022). Therefore, TLS constitute an opportunity for immunotherapeutic approaches, not as predictors of outcome, but as potentiators of immunotherapies (Sautès-Fridman et al., 2019). In TNBC, the presence of TLS was correlated with higher DFS and OS (Lee et al., 2016), and increased pCR in patients treated with neo-adjuvant chemotherapy (Song et al., 2017).

It is important to emphasize that metastasis from TNBC primary tumors is less immunogenic than the latter, as they present inferior levels of infiltration of $CD8^+$ lymphocytes and PD-1-positive T lymphocytes, the downregulation of immune-activating cytokines, and the upregulation of immunosuppressive molecules (Ogiya et al., 2016; Szekely et al., 2018; Manson et al., 2019; He et al., 2020). This might support the higher efficacy of checkpoint inhibitors in early-stage TNBC than in pretreated metastatic tumors (Adams et al., 2019a). This suggests a benefit to the use of immunotherapeutic options in earlier stages of the disease, associated with higher tumor immunogenic potential, rather than in the metastatic setting.

In fact, the expression of checkpoint molecules by TILs, supports the high expression of PD-1 and PD-L1 in triple-negative tumors (Wang and Liu, 2020) and correlations with good prognosis (increased pCR (Kitano et al., 2017)). The expression of PD-L1 in TILs could also be predictive of the response to checkpoint blockade immunotherapy against TNBC (Rizzo and Ricci, 2022). In fact, positive PD-L1 expression is currently used in clinics for the selection of patients who may benefit from anti-PD-1 (FDA, 2020a) and anti-PD-L1 (Roche, 2019a) treatments. Nonetheless, the results from the Impassion031 (Mittendorf et al., 2020) and KEYNOTE-522 (Schmid et al., 2020a) trials, demonstrating the efficacy of checkpoint inhibitors in combination with chemotherapy in early TNBC regardless of PD-L1 status, should prompt further research on the use of PD-L1 expression as a biomarker for patient stratification.

With the growing understanding of TILs as markers of breast tumors prognosis, in 2019, the WHO acknowledged the importance of stromal TILs as prognostic markers in breast cancers and changed their histological classification to include a category of TIL-rich Invasive Breast Carcinoma of No Special Type (Hoon Tan et al., 2020). Nevertheless, harmonization of TILs and PD-L1 expression determination and scoring is needed, as immunotherapy becomes an increasingly valuable alternative for the treatment of TNBC patients.

## Immunotherapy against triple-negative breast cancer

Immunotherapy entails a set of therapeutic approaches whose mechanism of action implies the stimulation of the immune system, in either an active (e.g., vaccines) or a passive (e.g., antibodies or immune modulators) way.

In breast cancer, immunotherapy has enabled great advances, particularly in the case of HER2-positive tumors. Patients with this subtype of breast cancer benefit from treatment with the anti-HER2 monoclonal antibody trastuzumab, both in the early and advanced stages, enjoying prolonged survival and lower toxicity (Lambertini et al., 2017). Trastuzumab's success has encouraged the study of immunotherapy for other forms of breast cancer, such as TNBC. Higher levels of TNBC immunogenicity that are discussed in this review add to potential approaches such as immune checkpoint inhibitors, antibody-drug conjugates (ADCs), vaccines, and cellular therapies.

## Checkpoint inhibitors: From research to the clinics

Checkpoint inhibitors are monoclonal antibodies that target checkpoint proteins, inhibiting their immune-suppressive functions and prompting immune-mediated tumor cell death. In TNBC, the use of the checkpoint inhibitors pembrolizumab (anti-PD-1), atezolizumab (anti-PD-L1), avelumab (anti-PD-L1), and durvalumab (anti-PD-L1) have been studied in different regimens and for the treatment of different stages of the disease (Table 2). Blockage of CTLA-4 has also been considered for TNBC treatment, as its higher expression in patient tumors has been correlated with a better prognosis (Peng et al., 2020). However, this topic remains within the early stages of clinical research (phase I and II trials), and thus far, no approved anti-CTLA-4 molecules have been approved for use against TNBC. Therefore, this review focuses on the targeting of PD-1/PD-L1 interaction.

TABLE 2 Most relevant clinical trials for checkpoint inhibitors in TNBC, according to regimen and tumor stage.

| | Trial | Phase | Objective | Intervention | Outcome/Results |
|---|---|---|---|---|---|
| **EARLY STAGE** | KEYNOTE-173 NCT02622074 | Phase Ib | To evaluate the safety and efficacy of pembrolizumab in combination with six chemotherapy regimens as neoadjuvant treatment | Pembrolizumab + Nab-paclitaxel or paclitaxel + Carboplatin + Doxorubicin + Cyclophosphamide | Overall pCR = 60% (Schmid et al., 2020b) |
| | KEYNOTE-522 NCT03036488 | Phase III | To evaluate the efficacy and safety of pembrolizumab plus chemotherapy vs. placebo plus chemotherapy as neoadjuvant therapy | Pembrolizumab or Placebo + Carboplatin + Paclitaxel + (Doxorubicin or Epirubicin) + Cyclophosphamide | pCR pembrolizumab + chemo = 64.8%<br>pCR placebo + chemo = 51.2%<br>a difference of 13.6% in pCR<br>PD-L1 positive population did not have a significant increase in pCR (Schmid et al., 2020a) |
| | I-SPY2 NCT01042379 | Phase II | To evaluate the efficacy of the combination of pembrolizumab and chemotherapy vs. chemotherapy alone | (Pembrolizumab +) Paclitaxel + Doxorubicin + Cyclophosphamide | pCR = 60% (vs. 22% with chemotherapy alone) (Nanda et al., 2020) |
| | GeparNuevo NCT02685059 | Phase II | To evaluate response rates of neoadjuvant treatment of sequential chemotherapy and checkpoint inhibitor | Durvalumab or Placebo + Nab-paclitaxel + Epirubicin + Cyclophosphamide | pCR not significantly different between groups (Karn et al., 2020) |
| | Impassion031 NCT03197935 | Phase III | To evaluate the efficacy and safety of neoadjuvant treatment of chemotherapy plus atezolizumab vs. chemotherapy plus placebo | Atezolizumab or Placebo + Doxorubicin + Cyclophosphamide + Nab-paclitaxel | pCR = 58% (vs. 41% with chemotherapy alone)<br>In PD-L1 positive patients: pCR = 53% (vs. 37% with chemotherapy alone) (Mittendorf et al., 2020) |
| | NeoTRIPaPDL1 NCT02620280 | Phase III | To compare the efficacy of chemotherapy plus atezolizumab vs. chemotherapy alone | (Atezolizumab +) Nab-paclitaxel + Carboplatin | pCR not significantly different between groups (ASCO, 2020) |
| **ADVANCED OR METASTATIC** | KEYNOTE-012 NCT01848834 | Phase Ib | To evaluate the efficacy and safety of pembrolizumab in patients with advanced TNBC | Pembrolizumab | ORR = 18.5% (Nanda et al., 2016) |
| | KEYNOTE-086 NCT02447003 | Phase II | To evaluate the efficacy and safety of pembrolizumab monotherapy as first-line or above treatment in patients with metastatic TNBC | Pembrolizumab | First-line: ORR = 21.4% (Adams et al., 2019a)<br>Second-line or above: ORR = 5.3% (Adams et al., 2019b) |
| | PCD4989g NCT01375842 | Phase I | Dose escalation study to evaluate the safety and clinical activity of atezolizumab monotherapy in patients with metastatic TNBC | Atezolizumab | As first-line: OS = 17.6 months; ORR = 24%<br>As second-line or above: OS = 7.3 months; ORR = 6% (Emens et al., 2019) |
| | JAVELIN Solid Tumor NCT01772004 | Phase Ib | Dose escalation study to evaluate safety and clinical activity of avelumab in patients with locally advanced or metastatic TNBC | Avelumab | ORR = 5.2% (Dirix et al., 2017) |
| | KEYNOTE-355 NCT02819518 | Phase III | To compare the safety and efficacy of pembrolizumab plus chemotherapy vs. placebo plus chemotherapy in the treatment of patients with locally recurrent inoperable or metastatic TNBC who had been not previously treated with chemotherapy | Pembrolizumab or Placebo + Nab-paclitaxel or Paclitaxel or Gemcitabine or Carboplatin | PFS = 9.7 months (4.1 months longer than chemotherapy alone) (Cortes et al., 2020) |
| | IMpassion130 NCT02425891 | Phase III | To evaluate safety and efficacy of the combination atezolizumab plus nab-paclitaxel vs. placebo plus nab-paclitaxel in patients with locally advanced or metastatic TNBC who had not received prior therapy for metastatic breast cancer | Atezolizumab or Placebo + Nab-Paclitaxel | OS = 21.3 months (vs. 17.6 months with chemotherapy alone) (Schmid et al., 2018) |
| | IMpassion131 NCT03125902 | Phase III | To evaluate the efficacy and safety of atezolizumab plus paclitaxel vs. placebo plus paclitaxel in patients with previously untreated, locally advanced or metastatic TNBC | Atezolizumab or Placebo + Paclitaxel | No improvement in PFS or OS compared to paclitaxel alone (Miles et al., 2021) |

Due to the relevance of anti-PD-1 and anti-PD-L1 antibodies in the treatment of TNBC, the density of PD-L1 expression as predictive of response has been clinically assessed across trials. Overall, a positive correlation was identified between PD-L1 expression and response outcomes, confirming the importance of PD-L1 assessment as a biomarker in TNBC (Emens et al., 2019; Schmid et al., 2020b; Karn et al., 2020). These results culminated in the approval of two checkpoint inhibitors accompanied by PD-L1 diagnostic tests, as previously mentioned (FDA, 2020a; Roche, 2020). Nonetheless, in the KEYNOTE-522 trial (Table 2), PD-L1 expression was not confirmed as a response predictor. Although PD-L1 status is usually measured through pre-treatment tumor biopsy, the PCD4989g trial demonstrated a significant increase in expression following treatment, suggesting that checkpoint therapy may promote tumor-specific T-cell activation and PD-L1 expression (Emens et al., 2019). This reinforces the limitations in trials designs, particularly the differences in PD-L1 assessment between studies, and the variety of drugs used in chemotherapy regimens, which may influence the tumor immune landscape (Emens et al., 2019). Other biomarkers were recently assessed in the clinical trial GeparNuevo, which was able to demonstrate that TMB and selected immune signatures (*CXCL9*, *CCL5*, *CD8A*, *CD80*, *CXCL13*, *ID O 1*, *PDCD1*, *CD274*, *CTLA4*, and *FOXP3* genes) have independent predictive value for chemotherapy, with and without durvalumab, in early TNBC (Karn et al., 2020).

Neoadjuvant chemotherapy induces tumor cell death, thus increasing antigen presentation, stimulating immunity and decreasing immunosuppression, which leads to changes in the immune cell composition of the tumor microenvironment (Denkert et al., 2018), suggesting that a possible benefit arising from combined regimens of immunotherapy following chemotherapy (Luen et al., 2019). It is then open to investigation whether immune blockage therapy is more beneficial as first-line therapy, or as a second, if it is beyond choice or combination strategies may be more effective than immune monotherapy, and to the identification of the disease setting that could most benefit the most from these approaches.

## First-line therapy vs. beyond

As neoadjuvant chemotherapy is standard for TNBC, most immune regimens are studied as second-line or later options. Nonetheless, the addition of a checkpoint inhibitor to chemotherapy showed significant clinical benefit in previously untreated patients with metastatic TNBC with PFS of 9.7 months for pembrolizumab (Cortes et al., 2020) and 7.5 months for atezolizumab therapies (Necela et al., 2015). The latter results were obtained from the Impassion130 study (Table 2) that prompted accelerated FDA and EMA approval of atezolizumab in combination with nab-paclitaxel for first-line treatment of PD-L1-positive patients with locally advanced and metastatic TNBC, the first immunotherapy approved for this subtype of breast cancer, in 2019 (Roche, 2019b). Recently, a second interim analysis for this study confirmed the clinical benefit of the combination of atezolizumab and nab-paclitaxel in PD-L1-positive tumors and its tolerable safety profile, despite the non-significant difference in OS observed in the intention-to-treat population (Schmid et al., 2020c). However, due to the modest results and the evolution of the landscape in the treatment of TNBC, in August 2021, the company decided to withdraw from accelerated approval of this combination for the treatment of adults with unresectable locally advanced or metastatic TNBC in PD-L1-positive tumors with FDA (Roche, 2020).

Modest results continued to be observed for checkpoint inhibitor monotherapy in first-line metastatic TNBC with the PCD4989g phase I study, which reported an ORR of 24% (Emens et al., 2019), and the KEYNOTE-173 trial, reporting an ORR of 21.4% and a PFS of only 2.1 months (Adams et al., 2019a). The combination of durvalumab with standard chemotherapy for first-line neoadjuvant treatment in early TNBC did not result in a significantly improved pCR relative to chemotherapy alone (Loibl et al., 2019). Overall, the clinical benefit of first-line immunotherapy in TNBC remains unclear, as more information about its effectiveness in previously untreated patients is needed.

## Monotherapy vs. combined regimens

The first study to evaluate the efficacy of checkpoint inhibition in TNBC was the KEYNOTE-012 trial (Table 2), which assessed the clinical activity of single-agent pembrolizumab in heavily pretreated patients with PD-L1-positive metastatic TNBC. An acceptable ORR of 18.5% (Nanda et al., 2016) was demonstrated, later confirmed in the cohort b KEYNOTE-086 trial, with an ORR of 21.4% and a median duration of response of 18 months (Adams et al., 2019a). Atezolizumab monotherapy showed similar results in the PCD4989g trial, with an ORR of 24% and a median OS of 17.6 months as first-line therapy (Emens et al., 2019). Single-agent durvalumab had a modest ORR of 5.2% in advanced or metastatic TNBC (Dirix et al., 2017).

However, the largest benefit of checkpoint inhibitor therapy arises from its combination with chemotherapy. This combination is hypothesized to increase tumor cell death by adding a second mechanism of action to standard cytotoxic agents and by taking advantage of the immune modulating effects of standard therapies (Denkert et al., 2018; Luen et al., 2019; Ji et al., 2021). Combination of atezolizumab with either paclitaxel (Miles et al., 2021) or carboplatin and nab-paclitaxel (ASCO, 2020) did not result in improvements in efficacy or safety outcomes relative to chemotherapy alone. Nonetheless, clinical benefit was achieved when the anti-PD-L1 antibody was

combined with nab-paclitaxel. As mentioned before, this combination is approved for the first-line treatment of PD-L1-positive patients with locally advanced and metastatic cancer (Roche, 2019a; Roche, 2020). The extension of the approval to early and locally advanced stage was proposed after the Impassion-031 study resulted in a 53% pCR in patients in the atezolizumab group compared to 37% in the group treated with chemotherapy alone (Mittendorf et al., 2020). However, in June 2021, this application was withdrawn by the company, after EMA questioned the clinical benefits of the proposed regimen (EMA. Tecentriq, 2021).

As for anti-PD-1 antibody pembrolizumab, its simultaneous combination with the chemotherapeutic regimen resulted in increased benefit (Schmid et al., 2020a; Schmid et al., 2020b; Nanda et al., 2020), with pCR rates up to 38% higher than chemotherapy alone (Nanda et al., 2020). In fact, it has been approved by the FDA since November 2020 for the treatment of patients with locally recurrent unresectable or metastatic TNBC whose tumors express PD-L1 (CPS ≥10) (FDA, 2020a). and as neoadjuvant treatment, in addition to, since July 2021, use as a single agent in an adjuvant setting after surgery, in patients with early-stage high-risk TNBC (FDA, 2019). As for anti-PD-1 antibody pembrolizumab, its simultaneous combination with chemotherapeutic regimens resulted in increased benefit (Schmid et al., 2020a; Schmid et al., 2020b; Nanda et al., 2020), with pCR rates up to 38% higher than for chemotherapy alone (Nanda et al., 2020).

These results demonstrated the benefit of combination regimens and the need for the optimization of such therapeutic strategies that may see their efficacy increased from the use of checkpoint inhibitors together with specific chemotherapeutic agents.

## Early-Stage vs. advanced and metastatic

Early-stage TNBCs are more immunogenic than metastatic or heavily treated tumors (Szekely et al., 2018; Blackley and Loi, 2019; Hutchinson et al., 2020). Higher TMB levels and the increased presence of TILs translate into clear benefit of immunotherapy in early settings of the disease (Table 2). In fact, two major phase III trials, KEYNOTE-522 (Schmid et al., 2020a) and Impassion-031 (Mittendorf et al., 2020), demonstrated that the combination of a checkpoint inhibitor with chemotherapy improved pCR in 13.6 and 17% of patients, respectively, relative to standard chemotherapy. This was the basis for the approval of pembrolizumab in the treatment settings mentioned in the previous section (FDA, 2019). This benefit was corroborated by other studies (Nanda et al., 2017; Schmid et al., 2020b), although the GeparNuevo (Loibl et al., 2019) and NeoTRIPaPDL1 (ASCO, 2020) trials, in which a non-significant difference in pCR was found between groups, did not corroborate it. Nevertheless, patients in the window period cohort of the GeparNuevo trials, before starting chemotherapy, achieved significantly higher pCR after durvalumab treatment, relative to chemotherapy alone, than patients outside the window cohort (51 vs. 37.9%, respectively) (Loibl et al., 2019). This suggests that durvalumab may promote TIL migration to the tumor even before chemotherapy is administered

The study of checkpoint inhibitors in advanced and metastatic TNBC has occupied the majority of clinical trials so far (Rizzo et al., 2021). For previously treated patients, it is important to understand the impact of previous therapy in the immune composition of the tumors, as it may dictate the success or failure of adding an immunotherapeutic agent. Thus, the assessment of immune biomarkers plays a particularly important role in this setting. In this respect, the Impassion130 trial demonstrated that evaluation of PD-L1 positivity should be performed in tissue from the primary tumor instead of metastasis samples (Rugo et al., 2019). PD-L1 expression is mandatory for the selection of patients who may receive a combination of atezolizumab and nab-paclitaxel (Roche, 2019a) or pembrolizumab and chemotherapy (FDA, 2020a), which are currently approved by the regulatory authorities for patients with advanced or metastatic TNBC.

## Safety

Checkpoint inhibition leads to the activation of the immune system, which may not always be tumor specific, culminating in inflammatory toxicities in a variety of tissues, such as the skin, thyroid, liver, pancreas, colon, lung, heart, and central nervous system (Baraibar et al., 2019). In TNBC, the most common adverse events reported in clinical trials with checkpoint inhibitors are similar to those observed with chemotherapy and include anemia, neutropenia, nausea, alopecia, and liver function abnormalities, such as increased alanine aminotransferase (ALT) (Li et al., 2021). The incidence of treatment-related adverse events is clearly higher in combination regimens than in mono-immunotherapy (100% vs. 56–69%) (D'Abreo and Adams, 2019). Combined strategies are also associated with an increased frequency of grade 3/4 adverse events, mainly adrenal insufficiency, hepatitis, stomatitis, neutropenia, pyrexia, neuropathies, and pneumonitis (Li et al., 2021; Sternschuss et al., 2021).

Immune-related adverse events are usually controlled with the use of corticosteroids. Endocrinopathies are a common adverse event that deserves special consideration, as they may imply chronic hormone replacement therapy, thus impacting patients' quality of life (Sternschuss et al., 2021). Infusion-related reactions are frequent both in immune and chemotherapy administration and are usually manageable low-grade events (Schmid et al., 2020b; Mittendorf et al., 2020). Overall, the safety profile of checkpoint blockage therapy is acceptable and manageable, and potential toxicities have to be balanced with the

survival and durable response provided by the use of checkpoint inhibitors.

## Resistance

The emergence of mechanisms of resistance form another challenge in immune treatments, including in TNBC therapy (Luo et al., 2022). Both tumor-intrinsic and tumor-extrinsic mechanisms can lead to tumor resistance, particularly against checkpoint inhibitors (Kim et al., 2022).

The latter is related to the cellular composition of the tumor, where increased T-cell dysfunction (as a consequence of patient aging) (Sceneay et al., 2019) and the emergence of immunosuppressive cell populations after chemotherapeutic regimens (Zhang et al., 2021) play a major role in decreasing the efficacy of checkpoint therapy.

Tumor-intrinsic mechanisms are related to the activation of oncogenic pathways in the tumor microenvironment and they have been extensively studied. The clusters of genes implicated in inflammatory and immune responses are aberrantly expressed, allowing the tumor to escape the immune blockage of checkpoint inhibitors (Xiao et al., 2019). Combinatorial treatments with inhibitors of DNA methyltransferases (Wu et al., 2021) or activators of pathways involved in immune response (Loi et al., 2016) have been suggested to overcome this.

# Radiation: A potential complement to checkpoint inhibition

Radiation is also part of the arsenal available for TNBC management, especially after surgery, as adjuvant therapy addressing the minimal residual disease. Radiation has an immunosuppressive impact on cancer patients' immune systems (Standish et al., 2008). Lymphopenia and decreased immune cell activity are possible consequences of radio-sensitivity of hematopoietic cells and lack of precision of traditional techniques. However, it may also stimulate immunity in the tumor microenvironment, which is now considered a stage in radiobiology principles. This effect depends on the radiation technique employed and dose, and the pre-existing composition of the tumor microenvironment (Demaria and Formenti, 2012). It is hypothesized that radiation increases tumors' mutational load, prompting antigen presentation and T-cell activation, proliferation, and migration into the tumor microenvironment, and maybe even leading to a decrease in immune suppressors. Therefore, radiotherapy can function as an inductor of tumor immunity, complementing immunotherapy (Sharabi et al., 2015; Kang et al., 2016). Several clinical trials have been ongoing for a variety of cancers, with a number of phase I (NCT03366844), phase II (NCT02730130, NCT03464942, NCT03872505), and phase III (NCT02954874) studies testing for its combination with anti-PD-1 therapies against TNBC. These studies may be particularly important for patients with a higher risk of recurrence and lower sensitivity to chemotherapy.

A combination of immune and chemotherapy or radiotherapy has been proposed to enhance the expression of checkpoint molecules in tumors, making the tumor more sensitive to immunotherapy and increasing the efficacy of their inhibitors. In this regard, the phase II TONIC trial evaluated patients with metastatic TNBC who had received palliative chemotherapy and were subjected to induction therapies: irradiation of one metastatic lesion, or low-dose chemotherapy (doxorubicin, cyclophosphamide, or cisplatin). After induction treatment, patients received nivolumab, a monoclonal antibody targeting PD-1. Better responses were obtained in the doxorubicin (ORR 35%) and cisplatin (ORR 23%) groups, while the group without induction presented an ORR of 20%, in accordance with a demonstrated upregulation of PD-L1, T-cell activity, and genes related to an inflammatory response in the former group. Although the irradiation cohort showed an inferior ORR relative to the control no-induction group, the presence of TILs, the diversity of T-cell receptor repertoire, and inflammation-related gene signatures were higher than the ones observed for the no-induction cohort, indicating an immune modulation effect of radiation in the tumor microenvironment (Voorwerk et al., 2019).

Despite great advances arising from the use of checkpoint inhibitors, a considerable proportion of non-responders remain. Thus, alternative immunotherapy formats are under development.

# The quest for other forms of immunotherapy

## Vaccines

Vaccines are another strategy to promote antitumoral immune activity. Traditionally, cancer vaccines are based on tumor-associated antigens (TAA), but because TAAs are self-antigens, immune cells that recognize them are usually eliminated in the body's maturation process. Further, the immunosuppressive mechanisms in the tumor microenvironment represent an additional barrier to this technology. Efforts to overcome these issues have been made, such as aiming at neoantigens instead of TAAs or combining vaccines with checkpoint inhibitors. Therefore, some therapeutic vaccines have demonstrated efficacy, long-term immune memory, and safety in a variety of tumors (Hollingsworth and Jansen, 2019). The strategies in development for TNBC are summarized in Table 3.

TABLE 3 Alternative immunotherapeutic approaches against TNBC.

| Immune approach | Type of strategy | Rationale/Regimen | Clinical trial identifier |
|---|---|---|---|
| VACCINES | Peptide vaccine | Folate receptor alpha (overexpressed in TNBC cells (Necela et al., 2015)) peptide vaccine | NCT03012100 |
| | | AE37 (li-key HER2/neu hybrid) peptide vaccine + pembrolizumab | NCT04024800 |
| | | Neoantigen peptide vaccine + nab-paclitaxel + durvalumab + tremelimumab | NCT03606967 |
| | | PVX-410 multi-peptide vaccine (targeting TAAs XBP1, CD138 and CS1 (OncoPep, 2021)) + pembrolizumab | NCT03362060 |
| | | Galinpepimut-S peptide vaccine (targets the Wilms Tumor 1 protein) + pembrolizumab | NCT03761914 |
| | | P10s-PADRE (carbohydrate mimetic peptide P10s fused to the pan HLA DR-binding epitope - PADRE - peptide) + standard neoadjuvant chemotherapy | NCT02938442 |
| | mRNA vaccine | Nanoparticle-containing mRNA, coding for the tumor antigen MUC1, overexpressed in TNBC (Siroy et al., 2013) | NCT00986609 |
| | | Liposome formulated vaccine based on the identification of individualized tumor-specific mutations by NGS and on-demand RNA manufacturing platform | NCT02316457 |
| | DNA vaccine | Neoantigen DNA vaccine (designed based on advanced sequencing techniques and epitope prediction algorithms (Li et al., 2017b)) + durvalumab | NCT03199040 |
| | Adenoviral cancer vaccine | Vaccine-Based Immunotherapy Regimen (VBIR-2): chimpanzee adenovirus expressing TAAs + tremelimumab + sasanlimab | NCT03674827 |
| | Dendritic cell vaccine | Dendritic cell vaccine against Her2/Her3 + cytokine modulation (CKM) regimen + pembrolizumab | NCT04348747 |
| | | Dendritic cell vaccine loaded with cyclin B1, WT1, and CEF (overexpressed in TNBC (O'Shaughnessy et al., 2016)) + neoadjuvant chemotherapy | NCT02018458 |
| | | Tumor neoantigen autologous dendritic cell | NCT04105582 |
| | Others | BN-Brachyury (transcription factor) vaccine + anti-PD-L1 and anti-TGF-β fusion protein | NCT04296942 |
| | | Adagloxad simolenin vaccine (tumor-associated carbohydrate antigen, covalently linked to the carrier protein KLH) (Huang et al., 2020) | NCT03562637 |
| ADOPTIVE CELL THERAPY | Single therapy | TIL autologous therapy with lifileucel (a centrally manufactured TIL infusion product (Sarnaik et al., 2020)) | NCT04111510 |
| | Dendritic and cytokine-induced killer cells + chemotherapy | Dendritic and cytokine-induced killer cells + cyclophosphamide combined thiotepa (a preparative treatment prior to autologous cell transplantation (European Medicines Agency, 2020)) + carboplatin | NCT01395056 |
| | | Dendritic and cytokine-induced killer cells + cyclophosphamide | NCT01232062 |
| | Natural killer and cytotoxic T lymphocytes + chemotherapy | ALECSAT therapy (Autologous Lymphoid Effector Cells Specific Against Tumor cells) (Cytovac, 2019) + carboplatin + gemcitabine | NCT04609215 |
| | Autologous gene-edited T lymphocytes + checkpoint inhibitor | Autologous CD8 and CD4 T cells engineered to express a T cell receptor specific for a neoantigen from the patient's tumor (Sennino et al., 2019) + nivolumab | NCT03970382 |
| | T cells + chemotherapy | Autologous T cells targeting mesothelin (expressed in tumor cells of TNBC (Tozbikian et al., 2014)) + cyclophosphamide | NCT02792114 |
| CAR-T cells | Allogeneic CAR-T cells | Allogeneic CAR-T cells targeting NKG2DL, a natural killer cells ligand expressed in tumor cells and involved in immunosuppressive mechanisms | NCT04107142 |
| | Anti-MUC1 CAR-T + chemotherapy | Autologous CAR-T cells targeting MUS-1, a mucin glycoprotein overexpressed in TNBC, particularly in the basal-like tumors (Siroy et al., 2013) + cyclophosphamide + fludarabine | NCT04025216 |
| ANTIBODY-DRUG CONJUGATES | Ladiratuzumab vedotin | Anti-LIV-1 humanized IgG1 antibody linked to monomethyl auristatin E, by a cleavable protease + pembrolizumab | NCT03310957 |
| | | Anti-LIV-1 humanized IgG1 antibody linked to monomethyl auristatin E, by a cleavable protease + several regimens of chemo and immunotherapy combinations | NCT03424005 |

(Continued on following page)

TABLE 3 (*Continued*) Alternative immunotherapeutic approaches against TNBC.

| Immune approach | Type of strategy | Rationale/Regimen | Clinical trial identifier |
|---|---|---|---|
| | Glembatumumab vedotin | Glembatumumab antibody linked to monomethyl auristatin E via a protease sensitive linker + capecitabine | NCT01997333 |
| | Sacituzumab govitecan | Anti-Trop-2 (anti-humanized antitrophoblast cell-surface antigen 2) antibody with SN-38, linked by a cleavable CLA2 linker | NCT04595565 NCT02574455 |
| | | Anti-Trop-2 (anti-humanized antitrophoblast cell-surface antigen 2) antibody with SN-38, linked by a cleavable CLA2 linker + pembrolizumab | NCT04230109 NCT04468061 |
| | | Anti-Trop-2 (anti-humanized antitrophoblast cell-surface antigen 2) antibody with SN-38, linked by a cleavable CLA2 linker + atezolizumab | NCT04434040 |

## Cell therapy

Cell therapy in the form of tumor-specific lymphocytes is also being explored, as cytotoxic cells express receptors on their surface that are specific to a given antigen, providing a potent and directed immune response. The main drawback of this approach is, therefore, the difficulty in choosing the correct target, especially in solid tumors, which are less immunogenic. However, some options have been adopted for TNBC, which include the adoptive transfer of both autologous or allogenic cells, and chimeric antigen receptor T-cells (CAR-Ts) (Guevara-hoyer et al., 2020).

Adoptive cell therapy consists of the transfer of activated immune cells to the patient. These cells can be harvested from the patient's blood or tumor tissue, expanded *ex vivo*, and re-administered (autologous), or they can be taken from a different donor (allogeneic). The manipulation of immune cells outside the patients' body may include expansion, activation, or engineering, undertaken with the purpose of prompting these cells to recognize specific tumor antigens and eliminate tumor cells. Despite the challenges associated with the manufacture of adoptive cell therapy, and the delivery and regulatory issues, the positive results found for cancer treatment, including solid tumors, are promising in the potential benefit for immunogenic cancers, such as TNBC (Morotti et al., 2021). In this regard, advances in the use of autologous cells in TNBC followed a 2017 study that analyzed TILs from a patient with a metastatic triple-negative tumor, where an immunogenic mutation capable of T-cell activation was identified and proposed for adaptive cell therapy (Assadipour et al., 2017). Since that time, adoptive cell therapy has been studied in the clinical in a variety of regimens, as summarized in Table 3.

CAR-T cells are T lymphocytes that have been engineered to express a molecule that is specific to a certain antigen, usually a small format antibody, such as single-chain variable fragments (scFv). As they are engineered to recognize such a specific target, CAR-T cells can be rapidly up-scaled to obtain a high number of antigen-specific T cells. In addition, their mechanism of action is independent of MHC (major histocompatibility complex) antigen presentation, which increases the number of target tumor cells. However, safety regarding these therapies is still a concern, with organ damage ensuing due to "on-target off-tumor toxicities" and intense (sometimes deadly) immune reactions (Sterner and Sterner, 2021). Still, clinical trials are ongoing for TNBC, either targeting TNBC antigens or other components of the tumor microenvironment, such as endothelial cells or fibroblasts (Dees et al., 2020; Xie et al., 2020). Currently ongoing or recently concluded clinical trials involving CAR-T cells in TNBC are summarized in Table 3.

## Antibody-drug conjugates

ADCs combine the specificity of monoclonal antibodies with the cytotoxic effect of potent small molecular weight drugs. For TNBC, three ADCs are under study (Nagayama et al., 2020) (Table 3). Ladiratuzumab vedotin (or SGN-LIV1A) is an anti-LIV-1 humanized IgG1 antibody linked to monomethyl auristatin E by a cleavable protease. LIV-1 is a transmembrane protein that is overexpressed in metastatic TNBC tumors, and monomethyl auristatin E is an antimitotic agent that is widely used in ADCs. When used as monotherapy in metastatic TNBC, the conjugated antibody demonstrated favorable efficacy (ORR of 32%) and a manageable safety profile, with the most common adverse events being nausea, fatigue, peripheral sensory neuropathy, and decreased appetite (Modi et al., 2018; Tsai et al., 2021). Because of its demonstrated immune stimulatory properties (Cao et al., 2018), ladiratuzumab vedotin has been studied in combination with pembrolizumab for patients with advanced or metastatic TNBC in two ongoing clinical trials (NCT03310957 and NCT03424005). Preliminary data showed an ORR of 54% in a cohort of 26 patients and a tolerable safety profile (HanHeather et al., 2020).

Glembatumumab vedotin (CDX-011) consists of a glembatumumab antibody that is linked to monomethyl auristatin E via a protease-sensitive linker. Glembatumumab

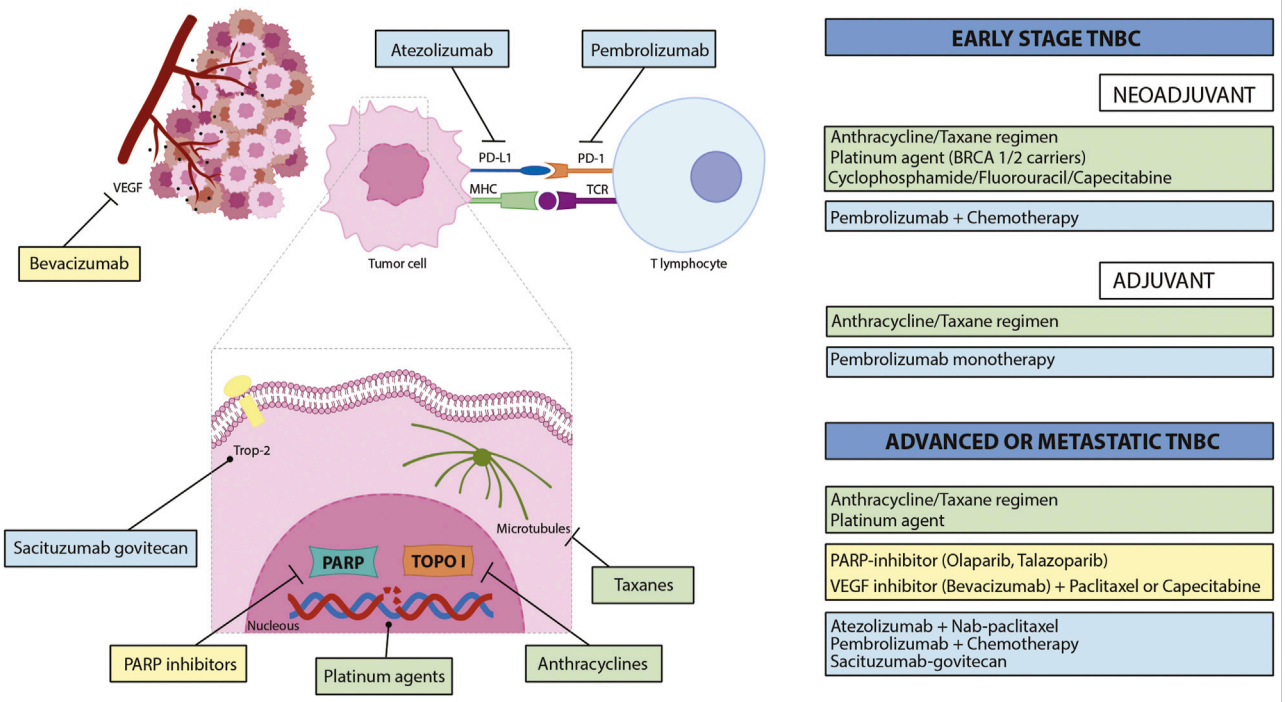



FIGURE 2
Currently approved pharmacological treatments of TNBC and their respective molecular targets. Green rectangles indicate chemotherapeutic regimens, yellow rectangles illustrate targeted therapies, and light blue rectangles refer to immunotherapies. PD-L1—programmed death-ligand 1; PD-1—programmed cell death protein 1; MHC—major histocompatibility complex: TCR—T-cell receptor; PARP—poly(ADP-ribose) polymerase; TOPO I—topoisomerase I.

targets glycoprotein NMB, and it is overexpressed in 40% of triple-negative tumors (Rose et al., 2010). The efficacy and safety of CDX-011 as monotherapy were evaluated in phase II clinical trials in patients with metastatic TNBC, overexpressing glycoprotein NMB (NCT01997333). However, the primary endpoint of the trial was not met (PFS was similar to the control group undergoing capecitabine treatment), and there was no reduced toxicity when compared to capecitabine alone (Vahdat et al., 2021).

Sacituzumab govitecan (IMMU-132) is an ADC that combines an anti-Trop-2 (anti-humanized antitrophoblast cell-surface antigen 2) antibody with SN-38, linked by a cleavable CLA2 linker. Trop-2 is a transmembrane molecule overexpressed in several epithelial cancers, such as TNBC (Huang et al., 2005), and SN-28 is an active metabolite of irinotecan and an inhibitor of topoisomerase I. Due to their demonstrated efficacy in a first phase I/II clinical trial, with an ORR of 33.3% and median response duration of 7.7 months (Bardia et al., 2019), in April 2020, FDA granted accelerated approval to sacituzumab govitecan for patients with metastatic TNBC who received at least two prior treatments for the metastatic setting (FDA, 2020b). Later, a confirmatory phase III trial compared the efficacy of ADC with single-agent chemotherapy in patients with relapsed or refractory metastatic TNBC. Patients treated with sacituzumab govitecan showed a significantly higher OS (12.1 vs. 6.7 months) and OR (35 vs. 5%) than patients treated with chemotherapy (Bardia et al., 2021). These results led to the approval by the FDA of sacituzumab govitecan for patients with unresectable locally advanced or metastatic TNBC, who have received two or more prior systemic therapies, at least one of them for metastatic disease (FDA, 2021). This last approval complements the currently available pharmacological treatments of TNBC, which are summarized in Figure 2.

At present, clinical trials with ADCs are testing their efficacy and safety as monotherapy and in combination with immunotherapy. These are summarized in Table 3.

## Conclusion and future prospects

Triple-negative breast cancer is an aggressive subtype of breast cancer. Advances in its treatment have been sporadic and have had unsatisfactory results, as chemotherapy remains the standard therapy in both neoadjuvant and adjuvant settings. The search for new targeted therapies has been intense, with PARP inhibitors leading the way. In this regard, both FDA and EMA approved olaparib and

talazoparib for patients with germline BRCA mutations and HER2-negative locally advanced or metastatic breast cancer who have been previously treated with anthracycline and/or taxane. Other approaches have been more controversial, such as the use of angiogenesis inhibitors, as the clinical results have not been conclusive so far. Thus, TNBC remains an unmet medical need.

The immunogenic nature of triple-negative tumors represents a relevant opportunity for novel treatment. TNBCs have a higher density of TILs and TMB than other subtypes of breast cancer, encouraging the study of new biomarkers and forms of immunotherapy against the disease. The expression of PD-L1 in triple-negative tumors is already in clinical use as predictive of response to checkpoint inhibitors, but lack of harmonization between assays hampers its routine use. In the future, standardization of such methods, together with combined information from TMB, immune gene signatures, and TIL levels, will allow patients to benefit the most from the potential of immunotherapies against TNBC. The importance of immune biomarkers goes even further, as their role in diagnosis, patient stratification, and prognosis will fulfill the demands of the complexity of personalized medicine.

Regarding immunotherapy, checkpoint inhibitors demonstrated some clinical benefit to TNBC, particularly when used in combination with chemotherapy, by taking advantage of the effects of the latter in the immune landscape of tumors. Their efficacy was also demonstrably greater in the early stages of the disease than in the pretreated metastatic settings, due to the less immunogenic nature of metastasis. In this respect, two checkpoint inhibitors were approved: atezolizumab combined with nab-paclitaxel for first-line treatment of PD-L1 positive tumors with locally advanced and metastatic TNBC (EMA, 2019), pembrolizumab combined with chemotherapy for the treatment of patients with locally recurrent unresectable or metastatic TNBC whose tumors express PD-L1 (FDA, 2020), and pembrolizumab either combined with chemotherapy as neoadjuvant treatment or as a single agent as adjuvant therapy after surgery in patients with early-stage high-risk TNBC (FDA, 2021).

Other forms of immunotherapies are currently being studied, with highlight to ADCs, such as sacituzumab-govitecan, which has recently been approved for patients with unresectable locally advanced or metastatic TNBC who have received two or more prior systemic therapies, at least one of them for metastatic disease (FDA, 2021). Vaccines and cell therapy are under clinical research, with no demonstrated clinical benefit yet.

To improve the use of immunotherapy in the treatment of patients with TNBC, a better understanding of the interplay between the tumor with the immune system is needed, as well as the mechanism of action of the different immune therapies. The design of the clinical trials also requires improvement, as nowadays they tend to include distinct endpoints and different population sizes, which makes it difficult to analyze and compare the data obtained. The study of the combination of different immunotherapies or immune agents with other forms of therapy will need to evolve. Consciousness is growing that the stimulation of the tumor microenvironment causes a more favorable immunogenic state, improvement of the activity of immune therapeutic agents, and subsequent better clinical responses.

## Author contributions

RR and JM designed and wrote this work of review. RR designed the images and tables. MC and JG performed a clinical and scientific review, respectively, of the final manuscript. All authors have revised and agreed to the published version of the manuscript.

## Funding

RR is a student of the Pharmaceutical Sciences PhD program from the Faculty of Pharmacy, University of Coimbra, Coimbra, Portugal, with a fellowship (references SFRH/BD/121935/2016 and COVID/BD/151787/2021) from Fundação para a Ciência e a Tecnologia (FCT). This work was also supported by COMPETE 2020—Operational Program for Competitiveness and Internationalisation and Portuguese national funds via FCT—Fundação para a Ciência e a Tecnologia, I.P., under projects EXPL/MED-FAR/1512/2021, CIBB (FCT references UIDB/04539/2020 and UIDP/04539/2020) and Fundación La Caixa BREAST-BRAIN-N-BBB.

## 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


Abdollahi, A., and Etemadi, M. (2016). Pathological characteristics of triple-negative breast cancer at main referral teaching. *Int. J. Hematol. Stem Cell Res.* 5 (4).

Adams, S., Gray, R. J., Demaria, S., Goldstein, L., Perez, E. A., Shulman, L. N., et al. (2014). Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. *J. Clin. Oncol.* 32 (27), 2959–2966. doi:10.1200/JCO.2013.55.0491

Adams, S., Loi, S., Toppmeyer, D., Cescon, D. W., De LaurentiisM.Nanda, R., et al. (2019). Pembrolizumab monotherapy for previously untreated, PD-L1-positive, metastatic triple-negative breast cancer: Cohort B of the phase II KEYNOTE-086 study. *Ann. Oncol.* 30 (3), 405–411. doi:10.1093/annonc/mdy518

Adams, S., Schmid, P., Rugo, H. S., Winer, E. P., Loirat, D., AwAdA, A., et al. (2019). Pembrolizumab monotherapy for previously treated metastatic triple-negative breast cancer: Cohort A of the phase II KEYNOTE-086 study. *Ann. Oncol.* 30 (3), 397–404. doi:10.1093/annonc/mdy517

Al-Mahmood, S., Sapiezynski, J., Garbuzenko, O. B., and Minko, T. (2018). Metastatic and triple-negative breast cancer: Challenges and treatment options. *Drug Deliv. Transl. Res.* 8 (5), 1483–1507. doi:10.1007/s13346-018-0551-3

Althobiti, M., Aleskandarany, M. A., Joseph, C., Toss, M., Mongan, N., Diez-Rodriguez, M., et al. (2018). Heterogeneity of tumour-infiltrating lymphocytes in breast cancer and its prognostic significance. *Histopathology* 73 (6), 887–896. doi:10.1111/his.13695

ASCO The ASCO post. No improved pathologic complete response with atezolizumab in early triple-negative breast cancer. Availableat: https://ascopost.com/issues/february-25-2020/no-improved-pcr-with-atezolizumab-in-early-triple-negative-breast-cancer/(Accessed September 3, 2021).

Assadipour, Y., Zacharakis, N., Crystal, J. S., Prickett, T. D., Gartner, J. J., Somerville, R. P. T., et al. (2017). Characterization of an immunogenic mutation in a patient with metastatic triple-negative breast cancer. *Clin. Cancer Res.* 4062 (12), 4347–4353. doi:10.1158/1078-0432.CCR-16-1423

Balko, J. M., Giltnane, J. M., Wang, K., Schwarz, L. J., Young, C. D., Cook, R. S., et al. (2014). Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. *Cancer Discov.* 4 (2), 232–245. doi:10.1158/2159-8290.CD-13-0286

Baraibar, I., Melero, I., Ponz-Sarvise, M., and Castanon, E. (2019). Safety and tolerability of immune checkpoint inhibitors (PD-1 and PD-L1) in cancer. *Drug Saf.* 42 (2), 281–294. doi:10.1007/s40264-018-0774-8

Bardia, A., Hurvitz, S. A., Tolaney, S. M., Loirat, D., Punie, K., Oliveira, M., et al. (2021). Sacituzumab govitecan in metastatic triple-negative breast cancer. *N. Engl. J. Med.* 384 (16), 1529–1541. doi:10.1056/NEJMoa2028485

Bardia, A., Mayer, I. A., Vahdat, L. T., Tolaney, S. M., Isakoff, S. J., Diamond, J. R., et al. (2019). Sacituzumab govitecan-hziy in refractory metastatic triple-negative breast cancer. *N. Engl. J. Med.* 380 (8), 741–751. doi:10.1056/NEJMoa1814213

Barrett-Lee, P. J., Dixon, J. M., Farrell, C., Jones, A., LeonaRd, R., MurrayN., et al. (2009). Expert opinion on the use of anthracyclines in patients with advanced breast cancer at cardiac risk. *Ann. Oncol.* 20 (5), 816–827. doi:10.1093/annonc/mdn728

Barroso-Sousa, R., Jain, E., Cohen, O., Kim, D., Buendia-Buendia, J., WinEr, E., et al. (2020). Prevalence and mutational determinants of high tumor mutation burden in breast cancer. *Ann. Oncol.* 31 (3), 387–394. doi:10.1016/j.annonc.2019.11.010

Basho, R. K., Yam, C., Gilcrease, M., Murthy, R. K., Helgason, T., Karp, D. D., et al. (2018). Comparative effectiveness of an mTOR-based systemic therapy regimen in advanced, metaplastic and nonmetaplastic triple-negative breast cancer. *Oncologist* 23 (11), 1300–1309. doi:10.1634/theoncologist.2017-0498

Bauer, J. A., Chakravarthy, A. B., Rosenbluth, J. M., Mi, D., Seeley, E. H., Granja-Ingram, N. M., et al. (2010). Identification of markers of taxane sensitivity using proteomic and genomic analyses of breast tumors from patients receiving neoadjuvant paclitaxel and radiation. *Clin. Cancer Res.* 16 (2), 681–690. doi:10.1158/1078-0432.CCR-09-1091

Bell, R., Brown, J., Parmar, M., Toi, M., SuTer, T., Steger, G. G., et al. (2017). Final efficacy and updated safety results of the randomized phase III BEATRICE trial evaluating adjuvant bevacizumab-containing therapy in triplenegative early breast cancer. *Ann. Oncol.* 28 (4), 754–760. doi:10.1093/annonc/mdw665

Bhola, N. E., Cook, R. S., and Arteaga, C. L. (2013). TGF- b inhibition enhances chemotherapy action against triple-negative breast cancer find the latest version : TGF- β inhibition enhances chemotherapy action against triple-negative breast cancer. *Cancer* 123 (3), 1348–1358. doi:10.1172/JCI65416.1348

Blackley, E. F., and Loi, S. (2019). Targeting immune pathways in breast cancer: Review of the prognostic utility of TILs in early stage triple negative breast cancer (TNBC). *Breast* 48, S44–S48. doi:10.1016/S0960-9776(19)31122-1

Blankenstein, T., Gilboa, E., and Jaffee, E. M. (2012). The determinants of tumour immunogenicity. *Nat. Rev. Cancer* 12 (4), 307–313. doi:10.1038/nrc3246

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

Cao, A. T., Higgins, S., Stevens, N., Gardai, S. J., and Sussman, D. (2018). Abstract 2742: Additional mechanisms of action of ladiratuzumab vedotin contribute to increased immune cell activation within the tumor. *Cancer Res.* 78 (13), 2742. doi:10.1158/1538-7445.AM2018-2742

Caramelo, O., Silva, C., Caramelo, F., Frutuoso, C., and Almeida-Santos, T. (2019). The effect of neoadjuvant platinum-based chemotherapy in BRCA mutated triple negative breast cancers -systematic review and meta-analysis. *Hered. Cancer Clin. Pract.* 17 (1), 11. doi:10.1186/s13053-019-0111-y

Cardoso, F., Kyriakides, S., Ohno, S., Penault-LlorcaF.Poortmans, P., Rubio, I. T., et al. (2019). Early breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 30 (8), 1194–1220. doi:10.1093/annonc/mdz173

ClinicalTrials.gov Abiraterone acetate in molecular apocrine breast cancer (AMA) - NCT01842321. Availableat: https://clinicaltrials.gov/ct2/show/NCT01842321 (Accessed October 22, 2020).

Cortazar, P., Zhang, L., Untch, M., Mehta, K., Costantino, J. P., Wolmark, N., et al. (2014). Pathological complete response and long-term clinical benefit in breast cancer: The CTNeoBC pooled analysis. *Lancet* 384 (9938), 164–172. doi:10.1016/S0140-6736(13)62422-8

Cortes, J., Cescon, D. W., Rugo, H. S., Nowecki, Z., Im, S. A., Yusof, M. M., et al. (2020). Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for previously untreated locally recurrent inoperable or metastatic triple-negative breast cancer (KEYNOTE-355): A randomised, placebo-controlled, double-blind, phase 3 clinical trial. *Lancet* 396, 1817–1828. doi:10.1016/S0140-6736(20)32531-9

Costa, R., Shah, A. N., Santa-Maria, C. A., Cruz, M. R., Mahalingam, D., Carneiro, B. A., et al. (2017). Targeting Epidermal Growth Factor Receptor in triple negative breast cancer: New discoveries and practical insights for drug development. *Cancer Treat. Rev.* 53, 111–119. doi:10.1016/j.ctrv.2016.12.010

Costa, R. L. B., and Czerniecki, B. J. (2020). Clinical development of immunotherapies for HER2+ breast cancer: A review of HER2-directed monoclonal antibodies and beyond. *npj Breast Cancer* 6 (1), 10. doi:10.1038/s41523-020-0153-3

Criscitiello, C., Bayar, M. A., and Curigliano, G. (2018). A gene signature to predict high tumor-infiltrating lymphocytes after neoadjuvant chemotherapy and outcome in patients with triple-negative breast cancer Original article. *annonc*, 162–169. doi:10.1093/annonc/mdx691

Cytovac (2019). Cytovac - what is ALECSAT. Availableat: https://cytovac.com/alecsat/what-is-alecsat/(Accessed June 29, 2021).

D'Abreo, N., and Adams, S. (2019). Immune-checkpoint inhibition for metastatic triple-negative breast cancer: Safety first? *Nat. Rev. Clin. Oncol.* 16 (7), 399–400. doi:10.1038/s41571-019-0216-2

Dees, S., Ganesan, R., Singh, S., and Grewal, I. S. (2020). Emerging CAR-T cell therapy for the treatment of triple-negative breast cancer. *Mol. Cancer Ther.* 19 (12), 2409–2421. doi:10.1158/1535-7163.MCT-20-0385

Demaria, S., and Formenti, S. C. (2012). Radiation as an immunological adjuvant: Current evidence on dose and fractionation. *Front. Oncol.* 2, 153. doi:10.3389/fonc.2012.00153

Denkert, C., von Minckwitz, G., Darb-Esfahani, S., Lederer, B., Heppner, B. I., Weber, K. E., et al. (2018). Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: A pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet. Oncol.* 19 (1), 40–50. doi:10.1016/S1470-2045(17)30904-X

Dent, R., Trudeau, M., Pritchard, K. I., Hanna, W. M., Kahn, H. K., Sawka, C. A., et al. (2007). Triple-negative breast cancer: Clinical features and patterns of recurrence. *Clin. Cancer Res.* 13 (15), 4429–4434. doi:10.1158/1078-0432.CCR-06-3045

Desmedt, C., Haibe-Kains, B., Wirapati, P., Buyse, M., Larsimont, D., Bontempi, G., et al. (2008). Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. *Clin. Cancer Res.* 14 (16), 5158–5165. doi:10.1158/1078-0432.CCR-07-4756

Diana, A., Franzese, E., Centonze, S., Carlino, F., Della Corte, C. M., Ventriglia, J., et al. (2018). Triple-negative breast cancers: Systematic review of the literature on

molecular and clinical features with a focus on treatment with innovative drugs. *Curr. Oncol. Rep.* 20 (10), 76. doi:10.1007/s11912-018-0726-6

Dirix, L. Y., Takacs, I., Jerusalem, G., Nikolinakos, P., Arkenau, H. T., Forero-Torres, A., et al. (2017). Avelumab, an anti-PD-L1 antibody, in patients with locally advanced or metastatic breast cancer: A phase 1b javelin solid tumor study. *Breast Cancer Res. Treat.* 167 (3), 671–686. doi:10.1007/s10549-017-4537-5

Dogan, B. E., and Turnbull, L. W. (2012). Imaging of triple-negative breast cancer. *Ann. Oncol.* 23 (6), vi29. doi:10.1093/annonc/mds191

Dunn, G. P., Old, L. J., and Schreiber, R. D. (2004). The immunobiology of cancer immunosurveillance and immunoediting. *Immunity* 21 (2), 137–148. doi:10.1016/j.immuni.2004.07.017

EMA. Tecentriq (2021). Withdrawal of the application to change the marketing authorisation. Availableat: https://www.ema.europa.eu/en/medicines/human/withdrawn-applications/tecentriq-1 (Accessed September 3, 2021).

Emens, L. A., Cruz, C., Eder, J. P., Braiteh, F., Chung, C., Tolaney, S. M., et al. (2019). Long-term clinical outcomes and biomarker analyses of atezolizumab therapy for patients with metastatic triple-negative breast cancer: A phase 1 study. *JAMA Oncol.* 5 (1), 74–82. doi:10.1001/jamaoncol.2018.4224

European Medicines Agency (2014). *Lynparza - summary of product characteristics*. doi:10.2307/j.ctvdf0dxq.12

European Medicines Agency (2019). *Talzenna - summary of product characteristics*. doi:10.2307/j.ctvdf0dxq.12

European Medicines Agency (2020). *Tepadina - summary of product characteristics*. Published online.

Fan, Y., Dutta, J., Gupta, N., Fan, G., and Gélinas, C. (2008). Regulation of programmed cell death by NF-kappaB and its role in tumorigenesis and therapy. *Adv. Exp. Med. Biol.* 615 (732), 223–250. doi:10.1007/978-1-4020-6554-5_11

FDA (2020c). FDA approves pembrolizumab for adults and children with TMB-H solid tumors. Availableat: https://www.fda.gov/drugs/drug-approvals-and-databases/fda-approves-pembrolizumab-adults-and-children-tmb-h-solid-tumors (Accessed July 2, 2020).

FDA (2019). FDA approves pembrolizumab for high-risk early-stage triple-negative breast cancer. Availableat: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-pembrolizumab-high-risk-early-stage-triple-negative-breast-cancer (Accessed September 5, 2021).

FDA (2020a). FDA grants accelerated approval to pembrolizumab for locally recurrent unresectable or metastatic triple negative breast cancer. Availableat: https://www.fda.gov/drugs/drug-approvals-and-databases/fda-grants-accelerated-approval-pembrolizumab-locally-recurrent-unresectable-or-metastatic-triple (Accessed June 10, 2021).

FDA (2020b). FDA grants accelerated approval to sacituzumab govitecan-hziy for metastatic triple negative breast cancer. Published 2020. Availableat: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-sacituzumab-govitecan-hziy-metastatic-triple-negative-breast-cancer (Accessed June 20, 2021).

FDA (2021). FDA grants regular approval to sacituzumab govitecan for triple-negative breast cancer. Availableat: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-regular-approval-sacituzumab-govitecan-triple-negative-breast-cancer (Accessed June 20, 2021).

Finak, G., Bertos, N., Pepin, F., Sadekova, S., Souleimanova, M., Zhao, H., et al. (2008). Stromal gene expression predicts clinical outcome in breast cancer. *Nat. Med.* 14 (5), 518–527. doi:10.1038/nm1764

Food and Drug Administration (2019). *Avastin (bevacizumab) information*.

Food and Drug Administration (2014). *FDA approves olaparib for germline BRCA-mutated metastatic breast cancer*. doi:10.31525/fda1-ucm592357.htm

Food and Drug Administration (2018). *FDA approves talazoparib for gBRCAm HER2-negative locally advanced or metastatic breast cancer*.

Fornier, M., and Fumoleau, P. (2012). The paradox of triple negative breast cancer: Novel approaches to treatment. *Breast J.* 18 (1), 41–51. doi:10.1111/j.1524-4741.2011.01175.x

Foulkes, W. D., Smith, I. E., and Reis-Filho, J. S. (2010). Triple-negative breast cancer. *N. Engl. J. Med.* 363 (20), 1938–1948. doi:10.1056/NEJMra1001389

Furth, P. A. (2014). STAT signaling in different breast cancer sub-types. *Mol. Cell. Endocrinol.* 382 (1), 612–615. doi:10.1016/j.mce.2013.03.023

Gajewski, T. F., Schreiber, H., and Fu, Y. X. (2013). Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.* 14 (10), 1014–1022. doi:10.1038/ni.2703

Geenen, J. J. J., Linn, S. C., Beijnen, J. H., and Schellens, J. H. M. (2018). PARP inhibitors in the treatment of triple-negative breast cancer. *Clin. Pharmacokinet.* 57 (4), 427–437. doi:10.1007/s40262-017-0587-4

Gennari, A., André, F., Barrios, C. H., Cortes, J., de Azambuja, E., DeMichele, A., et al. (2021). ESMO Clinical Practice Guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer. *Ann. Oncol.* 32 (12), 1475–1495. doi:10.1016/j.annonc.2021.09.019

Gucalp, A., Tolaney, S., Isakoff, S. J., Ingle, J. N., Liu, M. C., Carey, L. A., et al. (2013). Phase II trial of bicalutamide in patients with androgen receptor–positive, estrogen receptor–negative metastatic breast cancer. *Clin. Cancer Res.* 19 (19), 5505–5512. doi:10.1158/1078-0432.CCR-12-3327

Guevara-hoyer, K., Fuentes-Antrás, J., Baliu-pique, M., García-Sáenz, J. A., Pandiella, A., Perez-Segura, P., et al. (2020). Adoptive cell therapy in breast cancer : A current perspective of next-generation medicine. *Front. Oncol.*, 605633. doi:10.3389/fonc.2020.605633

Gupta, P. B., Onder, T. T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R. A., et al. (2009). Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* 138 (4), 645–659. doi:10.1016/j.cell.2009.06.034

Hammond, M. E. H., Hayes, D. F., Dowsett, M., Allred, D. C., Hagerty, K. L., Badve, S., et al. (2010). American society of clinical oncology/college of American pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version). *Arch. Pathol. Lab. Med.* 134 (7), e48–e72. doi:10.1043/1543-2165-134.7.e48

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013

HanHeather, H., Diab, S., Alemany, C., Basho, R., Brown-Glaberman, U., Meisel, J., et al. (2020). Abstract PD1-06: Open label phase 1b/2 study of ladiratuzumab vedotin in combination with pembrolizumab for first-line treatment of patients with unresectable locally-advanced or metastatic triple-negative breast cancer. *Cancer Res.* 80 (4), PD106. doi:10.1158/1538-7445.SABCS19-PD1-06

Harbeck, N., Penault-Llorca, F., Cortes, J., and Gnant, M. (2019). Breast cancer. *Nat. Rev. Dis. Prim.*, 5, 66. doi:10.1038/s41572-019-0111-2

Haslam, A., and Prasad, V. (2019). Estimation of the percentage of US patients with cancer who are eligible for and respond to checkpoint inhibitor immunotherapy drugs. *JAMA Netw. Open* 2 (5), e192535. doi:10.1001/jamanetworkopen.2019.2535

He, T. F., Yost, S. E., Frankel, P. H., Dagis, A., Cao, Y., Wang, R., et al. (2020). Multi-panel immunofluorescence analysis of tumor infiltrating lymphocytes in triple negative breast cancer: Evolution of tumor immune profiles and patient prognosis. *PLoS One* 15 (3), e0229955. doi:10.1371/journal.pone.0229955

He, X., and Xu, C. (2020). Immune checkpoint signaling and cancer immunotherapy. *Cell Res.* 30 (8), 660–669. doi:10.1038/s41422-020-0343-4

Hendrickx, W., Simeone, I., Anjum, S., Mokrab, Y., Bertucci, F., Finetti, P., et al. (2017). Identification of genetic determinants of breast cancer immune phenotypes by integrative genome-scale analysis. *Oncoimmunology* 6 (2), e1253654. doi:10.1080/2162402X.2016.1253654

Hollingsworth, R. E., and Jansen, K. (2019). Turning the corner on therapeutic cancer vaccines. *npj Vaccines* 4 (1), 7. doi:10.1038/s41541-019-0103-y

Hoon Tan, P., Ellis, I., and Allison, K. (2020). *The 2019 WHO classification of tumours of the breast*. London: Histopathology. doi:10.1111/his.14091

Houssami, N., MacAskill, P., Von Minckwitz, G., Marinovich, M. L., and Mamounas, E. (2012). Meta-analysis of the association of breast cancer subtype and pathologic complete response to neoadjuvant chemotherapy. *Eur. J. Cancer* 48 (18), 3342–3354. doi:10.1016/j.ejca.2012.05.023

Huang, C. S., Yu, A. L., Tseng, L. M., Chow, L. W. C., Hou, M. F., Hurvitz, S. A., et al. (2020). Globo H-klh vaccine adagloxad simolenin (OBI-822)/OBI-821 in patients with metastatic breast cancer: Phase II randomized, placebo-controlled study. *J. Immunother. Cancer* 8 (2), e000342. doi:10.1136/jitc-2019-000342

Huang, H., Groth, J., Sossey-Alaoui, K., Hawthorn, L., Beall, S., Geradts, J., et al. (2005). Aberrant expression of novel and previously described cell membrane markers in human breast cancer cell lines and tumors. *Clin. Cancer Res.* 11 (12), 4357–4364. doi:10.1158/1078-0432.CCR-04-2107

Huober, J., Von Minckwitz, G., Denkert, C., Tesch, H., Weiss, E., Zahm, D. M., et al. (2010). Effect of neoadjuvant anthracycline-taxane-based chemotherapy in different biological breast cancer phenotypes: Overall results from the GeparTrio study. *Breast Cancer Res. Treat.* 124 (1), 133–140. doi:10.1007/s10549-010-1103-9

Hutchinson, K. E., Yost, S. E., Chang, C. W., Johnson, R. M., Carr, A. R., McAdam, P. R., et al. (2020). Comprehensive profiling of poor-risk paired primary and recurrent triple-negative breast cancers reveals immune phenotype shifts. *Clin. Cancer Res.* 26 (3), 657–668. doi:10.1158/1078-0432.CCR-19-1773

Ignatiadis, M., Singhal, S. K., Desmedt, C., Haibe-Kains, B., Criscitiello, C., Andre, F., et al. (2012). Gene modules and response to neoadjuvant chemotherapy in breast cancer subtypes: A pooled analysis. *J. Clin. Oncol.* 30 (16), 1996–2004. doi:10.1200/JCO.2011.39.5624

Inao, T., Iida, Y., Moritani, T., Okimoto, T., Tanino, R., Kotani, H., et al. (2018). Bcl-2 inhibition sensitizes triple-negative human breast cancer cells to doxorubicin. *Oncotarget* 9 (39), 25545–25556. doi:10.18632/oncotarget.25370

Ji, Q., Ding, J., Hao, M., Luo, N., Huang, J., Zhang, W., et al. (2021). Immune checkpoint inhibitors combined with chemotherapy compared with chemotherapy alone for triple-negative breast cancer: A systematic review and meta-analysis. *Front. Oncol.* 11, 795650. doi:10.3389/fonc.2021.795650

Jiang, D., Gao, Z., Cai, Z., Wang, M., and He, J. (2015). Clinicopathological and prognostic significance of FOXP3+ tumor infiltrating lymphocytes in patients with breast cancer: A meta-analysis. *BMC Cancer* 15 (1), 727. doi:10.1186/s12885-015-1742-7

Jin, M. S., Lee, H., Park, I. A., Chung, Y. R., Im, S. A., Lee, K. H., et al. (2016). Overexpression of HIF1α and CAXI predicts poor outcome in early-stage triple negative breast cancer. *Virchows Arch.* 469 (2), 183–190. doi:10.1007/s00428-016-1953-6

Jovanovic, B., Mayer, I. A., Mayer, E. L., Abramson, V. G., Bardia, A., Sanders, M. E., et al. (2017). A randomized phase II neoadjuvant study of cisplatin, paclitaxel with or without everolimus in patients with stage II/III triple-negative breast cancer (TNBC): Responses and long-term outcome correlated with increased frequency of DNA damage response gene mutations, TNBC subtype, AR status, and Ki67. *Clin. Cancer Res.* 23 (15), 4035–4045. doi:10.1158/1078-0432.CCR-16-3055

Juul, N., Szallasi, Z., Eklund, A. C., Li, Q., Burrell, R. A., Gerlinger, M., et al. (2010). Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: A retrospective analysis of five clinical trials. *Lancet. Oncol.* 11 (4), 358–365. doi:10.1016/S1470-2045(10)70018-8

Kang, J., Demaria, S., and Formenti, S. (2016). Current clinical trials testing the combination of immunotherapy with radiotherapy. *J. Immunother. Cancer* 4 (1), 51. doi:10.1186/s40425-016-0156-7

Karn, T., Denkert, C., Weber, K. E., Holtrich, U., HanusCh, C., Sinn, B. V., et al. (2020). Tumor mutational burden and immune infiltration as independent predictors of response to neoadjuvant immune checkpoint inhibition in early TNBC in GeparNuevo. *Ann. Oncol.* 31 (9), 1216–1222. doi:10.1016/j.annonc.2020.05.015

Karn, T., Jiang, T., Hatzis, C., Sanger, N., El-Balat, A., Rody, A., et al. (2017). Association between genomic metrics and immune infiltration in triple-negative breast cancer. *JAMA Oncol.* 3 (12), 1707–1711. doi:10.1001/jamaoncol.2017.2140

Khan, M. A., Jain, V. K., Rizwanullah, M., Ahmad, J., and Jain, K. (2019). PI3K/AKT/mTOR pathway inhibitors in triple-negative breast cancer: A review on drug discovery and future challenges. *Drug Discov. Today* 24 (11), 2181–2191. doi:10.1016/j.drudis.2019.09.001

Kim, C., Gao, R., Sei, E., Brandt, R., Hartman, J., Hatschek, T., et al. (2018). Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing. *Cell* 173 (4), 879–893. doi:10.1016/j.cell.2018.03.041

Kim, H., Choi, J-M., and Lee, K. (2022). Immune checkpoint blockades in triple-negative breast cancer: Current state and molecular mechanisms of resistance. *Biomedicines* 10 (5), 1130. doi:10.3390/biomedicines10051130

Kim, S. B., Maslyar, D. J., Dent, R., Espie, M., Blau, S., Tan, A. R., et al. (2017). Ipatasertib plus paclitaxel versus placebo plus paclitaxel as first-line therapy for metastatic triple-negative breast cancer (LOTUS): A multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet. Oncol.* 18 (10), 1360–1372. doi:10.1016/S1470-2045(17)30450-3

Kitano, A., Ono, M., Yoshida, M., Noguchi, E., Shimomura, A., Shimoi, T., et al. (2017). Tumour-infiltrating lymphocytes are correlated with higher expression levels of PD-1 and PD-L1 in early breast cancer. *ESMO Open* 2 (2), e000150. doi:10.1136/esmoopen-2016-000150

Klimov, S., Rida, P. C. G., Aleskandarany, M. A., Green, A. R., Ellis, I. O., Janssen, E. A., et al. (2017). Novel immunohistochemistry-based signatures to predict metastatic site of triple-negative breast cancers. *Br. J. Cancer* 117 (6), 826–834. doi:10.1038/bjc.2017.224

Koboldt, D. C., Fulton, R. S., and McLellan, M. D. (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490 (7418), 61–70. doi:10.1038/nature11412

Kroemer, G., Senovilla, L., Galluzzi, L., André, F., and Zitvogel, L. (2015). Natural and therapy-induced immunosurveillance in breast cancer. *Nat. Med.* 21 (10), 1128–1138. doi:10.1038/nm.3944

Lambertini, M., Pondé, N. F., Solinas, C., and de Azambuja, E. (2017). Adjuvant trastuzumab: A 10-year overview of its benefit. *Expert Rev. Anticancer Ther.* 17 (1), 61–74. doi:10.1080/14737140.2017.1264876

Lee, H. J., Lee, J. J., Song, I. H., Park, I. A., Kang, J., Yu, J. H., et al. (2015). Prognostic and predictive value of NanoString-based immune-related gene signatures in a neoadjuvant setting of triple-negative breast cancer: Relationship to tumor-infiltrating lymphocytes. *Breast Cancer Res. Treat.* 151 (3), 619–627. doi:10.1007/s10549-015-3438-8

Lee, H. J., Park, I. A., Song, I. H., Shin, S. J., Kim, J. Y., Yu, J. H., et al. (2016). Tertiary lymphoid structures: Prognostic significance and relationship with tumour-infiltrating lymphocytes in triple-negative breast cancer. *J. Clin. Pathol.* 69 (5), 422–430. doi:10.1136/jclinpath-2015-203089

Lee, H. L., Jang, J. W., Lee, S. W., Yoo, S. H., Kwon, J. H., Nam, S. W., et al. (2019). Inflammatory cytokines and change of Th1/Th2 balance as prognostic indicators for hepatocellular carcinoma in patients treated with transarterial chemoembolization. *Sci. Rep.* 9 (1), 3260. doi:10.1038/s41598-019-40078-8

Lee, S., Cho, E., Park, Y., Ahn, J., and Im, Y. (2012). *Prognostic impact of FOXP3 expression in triple-negative breast cancer*, 1–9. doi:10.3109/0284186X.2012.731520

Lehmann, B. D., Bauer, J. A., Schafer, J. M., Pendleton, C. S., Tang, L., Johnson, K. C., et al. (2014). PIK3CA mutations in androgen receptor-positive triple negative breast cancer confer sensitivity to the combination of PI3K and androgen receptor inhibitors. *Breast Cancer Res.* 16 (4), 406. doi:10.1186/s13058-014-0406-x

Lehmann, B. D., Jovanović, B., Chen, X., Estrada, M. V., Johnson, K. N., Shyr, Y., et al. (2016). Refinement of triple-negative breast cancer molecular subtypes: Implications for neoadjuvant chemotherapy selection. *PLoS One* 11 (6), e0157368. doi:10.1371/journal.pone.0157368

Lehmann, B. D. B., Bauer, J., Chen, X., Sanders, M. E., Chakravarthy, A. B., Shyr, Y., et al. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* 121 (7), 2750–2767. doi:10.1172/JCI45014

Leonessa, F., and Clarke, R. (2003). ATP binding cassette transporters and drug resistance in breast cancer. *Endocr. Relat. Cancer* 10 (1), 43–73. doi:10.1677/erc.0.0100043

Li, L., Goedegebuure, S. P., and Gillanders, W. E. (2017). Preclinical and clinical development of neoantigen vaccines. *Ann. Oncol.* 28 (12), xii11–xii17. doi:10.1093/annonc/mdx681

Li, X., Yang, J., Peng, L., Sahin, A. A., Huo, L., Ward, K. C., et al. (2017). Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. *Breast Cancer Res. Treat.* 161 (2), 279–287. doi:10.1007/s10549-016-4059-6

Li, Y., Xing, L., Li, F., Liu, H., Gan, L., Yang, D., et al. (2021). Efficacy and safety of adding immune checkpoint inhibitors to neoadjuvant chemotherapy against triple-negative breast cancer: A meta-analysis of randomized controlled trials. *Front. Oncol.* 11, 657634. doi:10.3389/fonc.2021.657634

Liedtke, C., Mazouni, C., Hess, K. R., Andre, F., Tordai, A., Mejia, J. A., et al. (2008). Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* 26 (8), 1275–1281. doi:10.1200/JCO.2007.14.4147

Lin, F., Xie, Y. J., Zhang, X. K., Huang, T. J., Xu, H. F., Mei, Y., et al. (2019). GTSE1 is involved in breast cancer progression in p53 mutation-dependent manner. *J. Exp. Clin. Cancer Res.* 38 (1), 152. doi:10.1186/s13046-019-1157-4

Linderholm, B. K., Hellborg, H., Johansson, U., Elmberger, G., Skoog, L., Lehtio, J., et al. (2009). Significantly higher levels of vascular endothelial growth factor (VEGF) and shorter survival times for patients with primary operable triple-negative breast cancer. *Ann. Oncol.* 20 (10), 1639–1646. doi:10.1093/annonc/mdp062

Liu, Y. (2019). A global immune gene expression signature for human cancers. *Oncotarget* 10 (20), 1993–2005. doi:10.18632/oncotarget.26773

Liu, Z., Li, M., Jiang, Z., and Wang, X. (2018). A comprehensive immunologic portrait of triple-negative breast cancer. *Transl. Oncol.* 11 (2), 311–329. doi:10.1016/j.tranon.2018.01.011

Loi, S., Dushyanthen, S., Beavis, P. A., Salgado, R., Denkert, C., Savas, P., et al. (2016). RAS/MAPK activation is associated with reduced tumor-infiltrating lymphocytes in triple-negative breast cancer: Therapeutic cooperation between MEK and PD-1/PD-L1 immune checkpoint inhibitors. *Clin. Cancer Res.* 22 (6), 1499–1509. doi:10.1158/1078-0432.CCR-15-1125

Loibl, S., O'Shaughnessy, J., Untch, M., Sikov, W. M., Rugo, H. S., McKee, M. D., et al. (2018). Addition of the PARP inhibitor veliparib plus carboplatin or carboplatin alone to standard neoadjuvant chemotherapy in triple-negative breast cancer (BrighTNess): A randomised, phase 3 trial. *Lancet. Oncol.* 19 (4), 497–509. doi:10.1016/S1470-2045(18)30111-6

Loibl, S., Untch, M., Burchardi, N., Huober, J., Sinn, B. V., Blohmer, J. U., et al. (2019). A randomised phase II study investigating durvalumab in addition to an anthracycline taxane-based neoadjuvant therapy in early triple-negative breast cancer: Clinical results and biomarker analysis of GeparNuevo study. *Ann. Oncol.* 30 (8), 1279–1288. doi:10.1093/annonc/mdz158

Luen, S. J., Salgado, R., and Loi, S. (2019). Residual disease and immune infiltration as a new surrogate endpoint for TNBC post neoadjuvant chemotherapy. *Oncotarget* 10 (45), 4612–4614. doi:10.18632/oncotarget.27081

Luengo-Fernandez, R., Leal, J., Gray, A., and Sullivan, R. (2013). Economic burden of cancer across the European union: A population-based cost analysis. *Lancet. Oncol.* 14 (12), 1165–1174. doi:10.1016/S1470-2045(13)70442-X

Luo, C., Wang, P., He, S., Zhu, J., Shi, Y., Wang, J., et al. (2022). Progress and prospect of immunotherapy for triple-negative breast cancer. *Front. Oncol.* 12, 919072. doi:10.3389/fonc.2022.919072

Manso, L., Moreno, F., Márquez, R., Castelo, B., ArcediAno, A., ArroyoM., et al. (2015). Use of bevacizumab as a first-line treatment for metastatic breast cancer. *Curr. Oncol.* 22 (2), e51–e60. doi:10.3747/co.22.2210

Manson, Q. F., Schrijver, W. A. M. E., ter Hoeve, N. D., Moelans, C. B., and van Diest, P. J. (2019). Frequent discordance in PD-1 and PD-L1 expression between primary breast tumors and their matched distant metastases. *Clin. Exp. Metastasis* 36 (1), 29–37. doi:10.1007/s10585-018-9950-6

Marabelle, A., Le, D. T., Ascierto, P. A., Di Giacomo, A. M., De Jesus-Acosta, A., Delord, J. P., et al. (2020). Efficacy of pembrolizumab in patients with noncolorectal high microsatellite instability/mismatch repair–deficient cancer: Results from the phase II KEYNOTE-158 study. *J. Clin. Oncol.* 38 (1), 1–10. doi:10.1200/JCO.19.02105

Martín, M., Barrios, C. H., Torrecillas, L., Ruiz-Borrego, M., Bines, J., Segalla, J., et al. . Abstract GS2-04: Efficacy results from CIBOMA/2004-01_geicam/2003-11 study: A randomized phase III trial assessing adjuvant capecitabine after standard chemotherapy for patients with early triple negative breast cancer. *Cancer Res.* 2019; 79, GS204. doi:10.1158/1538-7445.SABCS18-GS2-04

Masuda, N., Lee, S. J., Ohtani, S., Im, Y. H., Lee, E. S., Yokota, I., et al. (2017). Adjuvant capecitabine for breast cancer after preoperative chemotherapy. *N. Engl. J. Med.* 376 (22), 2147–2159. doi:10.1056/NEJMoa1612645

Miles, D., Gligorov, J., André, F., Cameron, D., Schneeweiss, A., Barrios, C., et al. (2021). Primary results from IMpassion131, a double-blind, placebo-controlled, randomised phase III trial of first-line paclitaxel with or without atezolizumab for unresectable locally advanced/metastatic triple-negative breast cancer. *Ann. Oncol.* 32 (8), 994–1004. doi:10.1016/j.annonc.2021.05.801

Miles, D. W., Diéras, V., Cortés, J., Duenne, A. A., Yi, J., O'Shaughnessy, J., et al. (2013). First-line bevacizumab in combination with chemotherapy for HER2-negative metastatic breast cancer: Pooled and subgroup analyses of data from 2447 patients. *Ann. Oncol.* 24 (11), 2773–2780. doi:10.1093/annonc/mdt276

Mittendorf, E. A., Zhang, H., Barrios, C. H., Saji, S., Jung, K. H., Hegg, R., et al. (2020). Neoadjuvant atezolizumab in combination with sequential nab-paclitaxel and anthracycline-based chemotherapy versus placebo and chemotherapy in patients with early-stage triple-negative breast cancer (IMpassion031): A randomised, double-blind, phase 3 trial. *Lancet* 396 (10257), 1090–1100. doi:10.1016/S0140-6736(20)31953-X

Modi, S., Pusztai, L., Forero, A., Mita, M., Miller, K., Weise, A., et al. (2018). Abstract PD3-14: Phase 1 study of the antibody-drug conjugate SGN-LIV1A in patients with heavily pretreated triple-negative metastatic breast cancer. *Cancer Res.* 78, PD314. doi:10.1158/1538-7445.SABCS17-PD3-14

Montemorano, L., Michelle, D. S., and Bixel, L. K. (2019). Role of olaparib as maintenance treatment for ovarian cancer: The evidence to date. *Onco. Targets. Ther.* 12, 11497–11506. doi:10.2147/OTT.S195552

Morotti, M., Albukhari, A., Alsaadi, A., Artibani, M., Brenton, J. D., Curbishley, S. M., et al. (2021). Promises and challenges of adoptive T-cell therapies for solid tumours. *Br. J. Cancer* 124 (11), 1759–1776. doi:10.1038/s41416-021-01353-6

Morris, G. J., Naidu, S., Topham, A. K., Guiles, F., Xu, Y., McCue, P., et al. (2007). Differences in breast carcinoma characteristics in newly diagnosed african-American and caucasian patients: A single-institution compilation compared with the national cancer institute's surveillance, epidemiology, and end results database. *Cancer* 110 (4), 876–884. doi:10.1002/cncr.22836

Mumin, N. H., Drobnitzky, N., Patel, A., Lourenco, L. M., Cahill, F. F., Jiang, Y., et al. (2019). Overcoming acquired resistance to HSP90 inhibition by targeting JAK-STAT signalling in triple-negative breast cancer. *BMC Cancer* 19 (1), 102. doi:10.1186/s12885-019-5295-z

Nagayama, A., Vidula, N., Ellisen, L., and Bardia, A. (2020). Novel antibody–drug conjugates for triple negative breast cancer. *Ther. Adv. Med. Oncol.* 12, 1758835920915980. doi:10.1177/1758835920915980

Nakai, K., Hung, M. C., and Yamaguchi, H. (2016). A perspective on anti-EGFR therapies targeting triple-negative breast cancer. *Am. J. Cancer Res.* 6 (8), 1609–1623.

Nanda, R., Chow, L. Q. M., Dees, E. C., Berger, R., Gupta, S., Geva, R., et al. (2016). Pembrolizumab in patients with advanced triple-negative breast cancer: Phase Ib keynote-012 study. *J. Clin. Oncol.* 34 (21), 2460–2467. doi:10.1200/JCO.2015.64.8931

Nanda, R., Liu, M. C., Yau, C., Asare, S., Hylton, N., Veer, L. V., et al. (2017). Pembrolizumab plus standard neoadjuvant therapy for high-risk breast cancer (BC): Results from I-SPY 2. *J. Clin. Oncol.* 35 (15), 506. doi:10.1200/JCO.2017.35.15_suppl.506

Nanda, R., Liu, M. C., Yau, C., Shatsky, R., Pusztai, L., Wallace, A., et al. (2020). Effect of pembrolizumab plus neoadjuvant chemotherapy on pathologic complete response in women with early-stage breast cancer: An analysis of the ongoing phase 2 adaptively randomized I-SPY2 trial. *JAMA Oncol.* 6 (5), 676–684. doi:10.1001/jamaoncol.2019.6650

NCCN (2020). National comprehensive cancer network. NCCN guidelines - breast cancer. Availableat: https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1419 (Accessed May 2, 2021).

Necela, B. M., Crozier, J. A., Andorfer, C. A., Lewis-Tuffin, L., Kachergus, J. M., Geiger, X. J., et al. (2015). Folate receptor- α ( FOLR1 ) expression and function in triple negative tumors. *PLoS One* 10, e0122209. doi:10.1371/journal.pone.0122209

Nielsen, D., Maare, C., and Skovsgaard, T. (1996). Cellular resistance to anthracyclines. *Gen. Pharmacol.* 27 (2), 251–255. doi:10.1016/0306-3623(95)02013-6

Ochi, T., Bianchini, G., Ando, M., Nozaki, F., Kobayashi, D., Criscitiello, C., et al. (2019). Predictive and prognostic value of stromal tumour-infiltrating lymphocytes before and after neoadjuvant therapy in triple negative and HER2-positive breast cancer. *Eur. J. Cancer* 118, 41–48. doi:10.1016/j.ejca.2019.05.014

O'Connor, R., O'Leary, M., Ballot, J., Collins, C. D., Kinsella, P., Mager, D. E., et al. (2007). A phase I clinical and pharmacokinetic study of the multi-drug resistance protein-1 (MRP-1) inhibitor sulindac, in combination with epirubicin in patients with advanced cancer. *Cancer Chemother. Pharmacol.* 59 (1), 79–87. doi:10.1007/s00280-006-0240-7

O'Conor, C. J., Chen, T., González, I., Cao, D., and Peng, Y. (2018). Cancer stem cells in triple-negative breast cancer: A potential target and prognostic marker. *Biomark. Med.* 12 (7), 813–820. doi:10.2217/bmm-2017-0398

Ogiya, R., Niikura, N., Kumaki, N., Bianchini, G., Kitano, S., Iwamoto, T., et al. (2016). Comparison of tumor-infiltrating lymphocytes between primary and metastatic tumors in breast cancer patients. *Cancer Sci.* 107 (12), 1730–1735. doi:10.1111/cas.13101

Oliveira, M., Saura, C., Nuciforo, P., Calvo, I., Andersen, J., Passos-Coelho, J. L., et al. (2019). FAIRLANE, a double-blind placebo-controlled randomized phase II trial of neoadjuvant ipatasertib plus paclitaxel for early triple-negative breast cancer. *Ann. Oncol.* 30 (8), 1289–1297. doi:10.1093/annonc/mdz177

OncoPep Pipeline - OncoPep. Availableat: https://www.oncopep.com/pipeline-index (Accessed June 24, 2021).

O'Reilly, E. A., Gubbins, L., Sharma, S., Tully, R., Guang, M. H. Z., Weiner-Gorzel, K., et al. (2015). The fate of chemoresistance in triple negative breast cancer (TNBC). *BBA Clin.* 3, 257–275. doi:10.1016/j.bbacli.2015.03.003

O'Shaughnessy, J., Roberts, L. K., Smith, J. L., Levin, M. K., Timis, R., Finholt, J. P., et al. (2016). Safety and initial clinical efficacy of a dendritic cell (DC) vaccine in locally advanced, triple-negative breast cancer (TNBC) patients (pts). *J. Clin. Oncol.* 34 (15), 1086. doi:10.1200/JCO.2016.34.15_suppl.1086

Paluch-Shimon, S., Cardoso, F., Partridge, A. H., Abulkhair, O., Azim, H. A., Bianchi-Micheli, G., et al. (2020). ESO–ESMO 4th international consensus guidelines for breast cancer in young women (BCY4). *Ann. Oncol.* 31 (6), 674–696. doi:10.1016/j.annonc.2020.03.284

Pandy, J. G. P., Balolong-Garcia, J. C., Cruz-Ordinario, M. V. B., and Que, F. V. F. (2018). Triple negative breast cancer and platinum-based systemic treatment: Meta-analysis and systematic review. *Ann. Oncol.* 29, ix2. doi:10.1093/annonc/mdy426.006

Park, I. H., Kong, S. Y., Kwon, Y., Kim, M. K., Sim, S. H., Joo, J., et al. (2018). Phase I/II clinical trial of everolimus combined with gemcitabine/cisplatin for metastatic triple-negative breast cancer. *J. Cancer* 9 (7), 1145–1151. doi:10.7150/jca.24035

Park, S. Y., Lee, H. E., Li, H., Shipitsin, M., Gelman, R., Polyak, K., et al. (2010). Heterogeneity for stem cell-related markers according to tumor subtype and histologic stage in breast cancer. *Clin. Cancer Res.* 16 (3), 876–887. doi:10.1158/1078-0432.CCR-09-1532

Patt, D., Gauthier, M., and Giordano, S. (2006). Paclitaxel in breast cancer. *Womens Health* 2 (1), 11–21. doi:10.2217/17455057.2.1.11

Pelicano, H., Zhang, W., Liu, J., Hammoudi, N., Dai, J., Xu, R. H., et al. (2014). Mitochondrial dysfunction in some triple-negative breast cancer cell lines: Role of mTOR pathway and therapeutic potential. *Breast Cancer Res.* 16 (5), 434. doi:10.1186/s13058-014-0434-6

Penault-Llorca, F., and Viale, G. (2012). Pathological and molecular diagnosis of triple-negative breast cancer: A clinical perspective. *Ann. Oncol.* 23 (6), vi19–vi22. doi:10.1093/annonc/mds190

Peng, Z., Su, P., Yang, Y., Yao, X., Zhang, Y., Jin, F., et al. (2020). Identification of CTLA-4 associated with tumor microenvironment and competing interactions in

triple negative breast cancer by co-expression network analysis. *J. Cancer* 11 (21), 6365–6375. doi:10.7150/jca.46301

Perou, C. M., Sùrlie, T., Eisen, M. B., van de RijnM.Jeffrey, S. S., Rees, C. A., et al. (2000). Molecular portraits of human breast tumours. *Nature* 533 (5), 747–752. doi:10.1038/35021093

Poma, P., Labbozzetta, M., D'Alessandro, N., and Notarbartolo, M. (2017). NF-κB is a potential molecular drug target in triple-negative breast cancers. *Omi A J. Integr. Biol.* 21 (4), 225–231. doi:10.1089/omi.2017.0020

Pruneri, G., Gray, K. P., Vingiani, A., Viale, G., Curigliano, G., Criscitiello, C., et al. (2016). Tumor-infiltrating lymphocytes (TILs) are a powerful prognostic marker in patients with triple-negative breast cancer enrolled in the IBCSG phase III randomized clinical trial 22-00. *Breast Cancer Res. Treat.* 158 (2), 323–331. doi:10.1007/s10549-016-3863-3

Rakha, E. A., and Ellis, I. O. (2009). Triple-negative/basal-like breast cancer: Review. *Pathology* 41 (1), 40–47. doi:10.1080/00313020802563510

Rakha, E. A., Reis-Filho, J. S., and Ellis, I. O. (2008). Basal-like breast cancer: A critical review. *J. Clin. Oncol.* 26 (15), 2568–2581. doi:10.1200/JCO.2007.13.1748

Rapoport, B. L., Nayler, S., Demetriou, G. S., Moodley, S. D., and Benn, C. A., Specialist Physician and Medical Oncologist, The Medical Oncology Center of Rosebank, Johannesburg, South Africa, Triple negative breast cancer pathologic diagnosis and current chemotherapy treatment options. *Eur. Oncol. Haematol.* (2014). 10 (1):35. doi:10.17925/eoh.2014.10.1.35

Reddy, S. M., Barcenas, C. H., Sinha, A. K., Hsu, L., Moulder, S. L., Tripathy, D., et al. (2018). Long-term survival outcomes of triple-receptor negative breast cancer survivors who are disease free at 5 years and relationship with low hormone receptor positivity. *Br. J. Cancer* 118 (1), 17–23. doi:10.1038/bjc.2017.379

Rizzo, A., and Ricci, A. (2022). Biomarkers for breast cancer immunotherapy: PD-L1, TILs, and beyond. *Expert Opin. Investig. Drugs* 31, 549–555. doi:10.1080/13543784.2022.2008354

Rizzo, A., Ricci, A. D., Lanotte, L., Lombardi, L., Di Federico, A., Brandi, G., et al. (2021). Immune-based combinations for metastatic triple negative breast cancer in clinical trials: Current knowledge and therapeutic prospects. *Expert Opin. Investig. Drugs* 22, 557–565. doi:10.1080/13543784.2022.2009456

Roche (2019a). European Commission approves Roche's Tecentriq in combination with Abraxane for people with PD-L1-positive, metastatic triple-negative breast cancer. Availableat: https://www.roche.com/media/releases/med-cor-2019-08-29.htm (Accessed January 3, 2021).

Roche (2019b). FDA grants Roche's Tecentriq in combination with Abraxane accelerated approval for people with PD-L1-positive, metastatic triple-negative breast cancer. Availableat: https://www.roche.com/media/releases/med-cor-2019-03-11.htm (Accessed January 3, 2021).

Roche (2020). Roche provides update on Tecentriq US indication for PD-L1-positive, metastatic triple-negative breast cancer. Availableat: https://www.roche.com/media/releases/med-cor-2021-08-27.htm (Accessed September 7, 2021).

Rody, A., Karn, T., Liedtke, C., Pusztai, L., Ruckhaeberle, E., Hanker, L., et al. (2011). A clinically relevant gene signature in triple negative and basal-like breast cancer. *Breast Cancer Res.* 13 (5), R97. doi:10.1186/bcr3035

Rose, A. A. N., Grosset, A. A., Dong, Z., Russo, C., Macdonald, P. A., Bertos, N. R., et al. (2010). Glycoprotein nonmetastatic b is an independent prognostic indicator of recurrence and a novel therapeutic target in breast cancer. *Clin. Cancer Res.* 16 (7), 2147–2156. doi:10.1158/1078-0432.CCR-09-1611

Rugo, H. S., Loi, S., Adams, S., Schmid, P., Schneeweiss, A., Barrios, C., et al. (2019). Performance of PD-L1 immunohistochemistry (IHC) assays in unresectable locally advanced or metastatic triple-negative breast cancer (mTNBC): Post-hoc analysis of IMpassion130. *Ann. Oncol.* 30, v858–v859. doi:10.1093/annonc/mdz394.009

Sabatier, R., Finetti, P., Guille, A., Adelaide, J., Chaffanet, M., Viens, P., et al. (2014). Claudin-low breast cancers: Clinical, pathological, molecular and prognostic characterization. *Mol. Cancer* 13 (1), 228. doi:10.1186/1476-4598-13-228

Samanta, D., Gilkesa, D. M., Chaturvedia, P., Xiang, L., and Semenza, G. L. (2014). Hypoxia-inducible factors are required for chemotherapy resistance of breast cancer stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 111 (50), E5429–E5438. doi:10.1073/pnas.1421438111

Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51 (2), 202–206. doi:10.1038/s41588-018-0312-8

Santonja, A., Sánchez-Muñoz, A., Lluch, A., Chica-Parrado, M. R., Albanell, J., Chacon, J. I., et al. (2018). Triple negative breast cancer subtypes and pathologic complete response rate to neoadjuvant chemotherapy. *Oncotarget* 9 (41), 26406–26416. doi:10.18632/oncotarget.25413

Sarnaik, A., Khushalani, N., Chesney, J., Kluger, H., and Curti, B. (2020). Safety & efficacy of lifileucel (ln-144) tumor infiltrating lymphocyte therapy in metastatic. *J. Immunother. Cancer* 8 (1), 2020.

Sautès-Fridman, C., Petitprez, F., Calderaro, J., and Fridman, W. H. (2019). Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* 19 (6), 307–325. doi:10.1038/s41568-019-0144-6

Sceneay, J., Goreczny, G. J., Wilson, K., Morrow, S., DeCristo, M. J., Ubellacker, J. M., et al. (2019). Interferon signaling is diminished with age and is associated with immune checkpoint blockade efficacy in triple-negative breast cancer. *Cancer Discov.* 9 (9), 1208–1227. doi:10.1158/2159-8290.CD-18-1454

Scharenberg, C. W., Harkey, M. A., and Torok-Storb, B. (2002). The ABCG2 transporter is an efficient Hoechst 33342 efflux pump and is preferentially expressed by immature human hematopoietic progenitors. *Blood* 99 (2), 507–512. doi:10.1182/blood.V99.2.507

Schettini, F., Giuliano, M., De Placido, S., and Arpino, G. (2016). Nab-paclitaxel for the treatment of triple-negative breast cancer: Rationale, clinical data and future perspectives. *Cancer Treat. Rev.* 50, 129–141. doi:10.1016/j.ctrv.2016.09.004

Schmid, P., Adams, S., Rugo, H. S., Schneeweiss, A., Barrios, C. H., Iwata, H., et al. (2018). Atezolizumab and nab-paclitaxel in advanced triple-negative breast cancer. *N. Engl. J. Med.* 379 (22), 2108–2121. doi:10.1056/nejmoa1809615

Schmid, P., Cortes, J., Pusztai, L., McArthur, H., Kummel, S., Bergh, J., et al. (2020). Pembrolizumab for early triple-negative breast cancer. *N. Engl. J. Med.* 382 (9), 810–821. doi:10.1056/NEJMoa1910549

Schmid, P., Rugo, H. S., Adams, S., Schneeweiss, A., Barrios, C. H., Iwata, H., et al. (2020). Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): Updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet. Oncol.* 21 (1), 44–59. doi:10.1016/S1470-2045(19)30689-8

Schmid, P., Salgado, R., Park, Y. H., Munoz-CousElo, E., Kim, S. B., Sohn, J., et al. (2020). Pembrolizumab plus chemotherapy as neoadjuvant treatment of high-risk, early-stage triple-negative breast cancer: Results from the phase 1b open-label, multicohort KEYNOTE-173 study. *Ann. Oncol.* 31 (5), 569–581. doi:10.1016/j.annonc.2020.01.072

Sennino, B., Conroy, A., Purandare, B., Litterman, A., Jacoby, K., Moot, R., et al. (2019). Abstract 1433: NeoTCR-P1, a novel neoepitope-specific adoptive cell therapy, consists of T cells with 'younger' phenotypes that rapidly proliferate and kill target cells upon recognition of cognate antigen. *Cancer Res.* 79 (13), 1433. doi:10.1158/1538-7445.AM2019-1433

Sharabi, A. B., Lim, M., DeWeese, T. L., and Drake, C. G. (2015). Radiation and checkpoint blockade immunotherapy: Radiosensitisation and potential mechanisms of synergy. *Lancet. Oncol.* 16 (13), e498–e509. doi:10.1016/S1470-2045(15)00007-8

Sharma, P., Rodler, E., Barlow, W. E., Gralow, J., Huggins-Puhalla, S. L., Anders, C. K., et al. (2020). Results of a phase II randomized trial of cisplatin +/- veliparib in metastatic triple-negative breast cancer (TNBC) and/or germline BRCA-associated breast cancer (SWOG S1416). *J. Clin. Oncol.* 38 (15), 1001. doi:10.1200/JCO.2020.38.15_suppl.1001

Shi, Y., Jin, J., Ji, W., and Guan, X. (2018). Therapeutic landscape in mutational triple negative breast cancer. *Mol. Cancer* 17 (1), 99. doi:10.1186/s12943-018-0850-9

Shibue, T., and Weinberg, R. A. (2017). EMT, CSCs, and drug resistance: The mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* 14 (10), 611–629. doi:10.1038/nrclinonc.2017.44

Sikov, W. M., Berry, D. A., Perou, C. M., Singh, B., Cirrincione, C. T., Tolaney, S. M., et al. (2015). Impact of the addition of carboplatin and/or bevacizumab to neoadjuvant once-per-week paclitaxel followed by dose-dense doxorubicin and cyclophosphamide on pathologic complete response rates in stage II to III triple-negative breast cancer: CALGB 40603 (alliance). *J. Clin. Oncol.* 33 (1), 13–21. doi:10.1200/JCO.2014.57.0572

Singh, J. C., Novik, Y., Stein, S., Volm, M., Meyers, M., Smith, J., et al. (2014). Phase 2 trial of everolimus and carboplatin combination in patients with triple negative metastatic breast cancer. *Breast Cancer Res.* 16 (2), R32. doi:10.1186/bcr3634

Siroy, A., Abdul-Karim, F. W., Miedler, J., Fong, N., Fu, P., Gilmore, H., et al. (2013). MUC1 is expressed at high frequency in early-stage basal-like triple negative breast cancer. *Hum. Pathol.* 44 (10), 2159–2166. doi:10.1016/j.humpath.2013.04.010

Smid, M., Hoes, M., Sieuwerts, A. M., Sleijfer, S., Zhang, Y., Wang, Y., et al. (2011). Patterns and incidence of chromosomal instability and their prognostic relevance in breast cancer subtypes. *Breast Cancer Res. Treat.* 128 (1), 23–30. doi:10.1007/s10549-010-1026-5

Song, I., Lee, H., Park, I., Yu, J., Ahn, J-H., and Gong, G. (2017). Predictive value of tertiary lymphoid structures assessed by high endothelial venule counts in the

neoadjuvant setting of triple-negative breast cancer. *Cancer Res. Treat.* 49 (2), 399–407. doi:10.4143/crt.2016.215

Soni, S., and Padwad, Y. S. (2017). HIF-1 in cancer therapy: Two decade long story of a transcription factor. *Acta Oncol.* 56 (4), 503–515. doi:10.1080/0284186X.2017.1301680

Standish, L. J., Torkelson, C., Hamill, F. A., Yim, D., Hill-Force, A., Fitzpatrick, A., et al. (2008). Immune defects in breast cancer patients after radiotherapy. *J. Soc. Integr. Oncol.* 6 (3), 110–121. doi:10.2310/7200.2008.0018

Stanton, S. E., Adams, S., and Disis, M. L. (2016). Variation in the incidence and magnitude of tumor-infiltrating lymphocytes in breast cancer subtypes: A systematic review. *JAMA Oncol.* 2 (10), 1354–1360. doi:10.1001/jamaoncol.2016.1061

Steelman, L. S., Navolanic, P. M., Sokolosky, M. L., Taylor, J. R., Lehmann, B. D., Chappell, W. H., et al. (2008). Suppression of PTEN function increases breast cancer chemotherapeutic drug resistance while conferring sensitivity to mTOR inhibitors. *Oncogene* 27 (29), 4086–4095. doi:10.1038/onc.2008.49

Stenzinger, A., Allen, J. D., Maas, J., Stewart, M. D., Merino, D. M., Wempe, M. M., et al. (2019). Tumor mutational burden standardization initiatives: Recommendations for consistent tumor mutational burden assessment in clinical samples to guide immunotherapy treatment decisions. *Genes Chromosom. Cancer* 58 (8), 578–588. doi:10.1002/gcc.22733

Sterner, R. C., and Sterner, R. M. (2021). CAR-T cell therapy: Current limitations and potential strategies. *Blood Cancer J.* 11 (4), 69. doi:10.1038/s41408-021-00459-7

Sternschuss, M., Yerushalmi, R., Saleh, R. R., Amir, E., and Goldvaser, H. (2021). Efficacy and safety of neoadjuvant immune checkpoint inhibitors in early-stage triple-negative breast cancer: A systematic review and meta-analysis. *J. Cancer Res. Clin. Oncol.* 147 (11), 3369–3379. doi:10.1007/s00432-021-03591-w

Stoll, G., Enot, D., Mlecnik, B., and Zitvogel, L. (2015). Immune-related gene signatures predict the outcome of neoadjuvant chemotherapy. *Oncoimmunology*, 37–41. doi:10.4161/onci.27884

Stringer-Reasor, E. M., May, J. E., Olariu, E., Caterinicchia, V., Li, Y., Chen, D., et al. (2021). An open-label, pilot study of veliparib and lapatinib in patients with metastatic, triple-negative breast cancer. *Breast Cancer Res.* 23 (1), 30. doi:10.1186/s13058-021-01408-9

Szekely, B., Bossuyt, V., Li, X., Wali, V. B., Patwardhan, G. A., FrederiCk, C., et al. (2018). Immunological differences between primary and metastatic breast cancer. *Ann. Oncol.* 29 (11), 2232–2239. doi:10.1093/annonc/mdy399

The Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70. doi:10.1038/nature11412

Thomas, A., Routh, E. D., Pullikuth, A., Jin, G., Su, J., Chou, J. W., et al. (2018). Tumor mutational burden is a determinant of immune-mediated survival in breast cancer. *Oncoimmunology* 7 (10), e1490854. doi:10.1080/2162402X.2018.1490854

Tommasi, S., Mangia, A., Lacalamita, R., Bellizzi, A., Fedele, V., Chiriatti, A., et al. (2007). Cytoskeleton and paclitaxel sensitivity in breast cancer: The role of β-tubulins. *Int. J. Cancer* 120 (10), 2078–2085. doi:10.1002/ijc.22557

Tozbikian, G., Brogi, E., Kadota, K., Catalano, J., Akram, M., Patil, S., et al. (2014). Mesothelin expression in triple negative breast carcinomas correlates significantly with basal-like phenotype, distant metastases and decreased survival. *PLoS One* 9 (12), e114900. doi:10.1371/journal.pone.0114900

Traina, T. A., Miller, K., Yardley, D. A., Eakle, J., Schwartzberg, L. S., O'Shaughnessy, J., et al. (2018). Enzalutamide for the treatment of androgen receptor-expressing triple-negative breast cancer. *J. Clin. Oncol.* 36 (9), 884–890. doi:10.1200/JCO.2016.71.3495

Tremont, A., Lu, J., and Cole, J. T. (2017). Endocrine therapy for early breast cancer: Updated review. *Ochsner J.* 17 (4), 405–411.

Tsai, M., Han, H. S., Montero, A. J., Tkaczuk, K., Assad, H., Pusztai, L., et al. (2021). 259P Weekly ladiratuzumab vedotin monotherapy for metastatic triple-negative breast cancer. *Ann. Oncol.* 32, S474–S475. doi:10.1016/j.annonc.2021.08.542

Turner, N., Tutt, A., and Ashworth, A. (2004). Hallmarks of "BRCAness" in sporadic cancers. *Nat. Rev. Cancer* 4 (10), 814–819. doi:10.1038/nrc1457

Turner, N. C., Balmaña, J., Poncet, C., Goulioti, T., Tryfonidis, K., Honkoop, A. H., et al. (2021). Niraparib for advanced breast cancer with germline BRCA1 and BRCA2 mutations: The EORTC 1307-BCG/BIG5–13/TESARO PR-30–50–10-C BRAVO study. *Clin. Cancer Res.* 27 (20), 5482–5491. doi:10.1158/1078-0432.CCR-21-0310

Untch, M., Jackisch, C., Schneeweiss, A., Conrad, B., Aktas, B., Denkert, C., et al. (2016). Nab-paclitaxel versus solvent-based paclitaxel in neoadjuvant chemotherapy for early breast cancer (GeparSepto-GBG 69): A randomised, phase 3 trial. *Lancet. Oncol.* 17 (3), 345–356. doi:10.1016/S1470-2045(15)00542-2

Vahdat, L. T., Schmid, P., Forero-Torres, A., Blackwell, K., Telli, M. L., Melisko, M., et al. (2021). Glembatumumab vedotin for patients with metastatic, gpNMB overexpressing, triple-negative breast cancer ("METRIC"): A randomized multicenter study. *npj Breast Cancer* 7 (1), 57. doi:10.1038/s41523-021-00244-6

VazLuis, I., Lin, N. U., Keating, N. L., Barry, W. T., Winer, E. P., Freedman, R. A., et al. (2017). Factors associated with early mortality among patients with de novo metastatic breast cancer: A PopulationBased study. *Oncologist* 22 (4), 386–393. doi:10.1634/theoncologist.2016-0369

Vihervuori, H., Autere, T. A., Repo, H., Kurki, S., KaLLio, L., Lintunen, M. M., et al. (2019). Tumor-infiltrating lymphocytes and CD8+ T cells predict survival of triple-negative breast cancer. *J. Cancer Res. Clin. Oncol.* 145 (12), 3105–3114. doi:10.1007/s00432-019-03036-5

Vinayak, S., Tolaney, S. M., Schwartzberg, L., Mita, M., McCann, G., Tan, A. R., et al. (2019). Open-label clinical trial of niraparib combined with pembrolizumab for treatment of advanced or metastatic triple-negative breast cancer. *JAMA Oncol.* 5 (8), 1132–1140. doi:10.1001/jamaoncol.2019.1029

Von Minckwitz, G., Untch, M., Blohmer, J. U., Costa, S. D., Eidtmann, H., Fasching, P. A., et al. (2012). Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J. Clin. Oncol.* 30 (15), 1796–1804. doi:10.1200/JCO.2011.38.8595

Voorwerk, L., Slagter, M., Horlings, H. M., Sikorska, K., van de Vijver, K. K., de Maaker, M., et al. (2019). Immune induction strategies in metastatic triple-negative breast cancer to enhance the sensitivity to PD-1 blockade: The TONIC trial. *Nat. Med.* 25 (6), 920–928. doi:10.1038/s41591-019-0432-4

Wang, B., Liu, J., Han, Y., Deng, Y., Li, J., Jiang, Y., et al. (2022). The presence of tertiary lymphoid structures provides new insight into the clinicopathological features and prognosis of patients with breast cancer. *Front. Immunol.*, 868155. doi:10.3389/fimmu.2022.868155

Wang, X., and Liu, Y. (2020). PD-L1 expression in tumor infiltrated lymphocytes predicts survival in triple-negative breast cancer. *Pathol. Res. Pract.* 216 (3), 152802. doi:10.1016/j.prp.2019.152802

Wang, Y., Zhao, L., Xiao, Q., Jiang, L., He, M., Bai, X., et al. (2015). miR-302a/b/c/d cooperatively inhibit BCRP expression to increase drug sensitivity in breast cancer cells. *Gynecol. Oncol.* 141 (3), 592–601. doi:10.1016/j.ygyno.2015.11.034

Wolff, A. C., Hammond, E. H., Allison, K. H., Harvey, B. E., Mangu, P. B., and Bartlett, J. M. S. (2018). Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline focused update. *Arch. Pathol. Lab. Med.* 142, 1364–1382. doi:10.5858/arpa.2018-0902-SA

Wolff, A. C., Hammond, M. E. H., Schwartz, J. N., Hagerty, K. L., Allred, D. C., Cote, R. J., et al. (2007). American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J. Clin. Oncol.* 25 (1), 118–145. doi:10.1200/JCO.2006.09.2775

Won, K. A., and Spruck, C. (2020). Triple negative breast cancer therapy: Current and future perspectives (Review). *Int. J. Oncol.* 57 (6), 1245–1261. doi:10.3892/ijo.2020.5135

Wong-Brown, M. W., Meldrum, C. J., Carpenter, J. E., Clarke, C. L., Narod, S. A., Jakubowska, A., et al. (2015). Prevalence of BRCA1 and BRCA2 germline mutations in patients with triple-negative breast cancer. *Breast Cancer Res. Treat.* 150 (1), 71–80. doi:10.1007/s10549-015-3293-7

Wu, S. Y., Xiao, Y., Wei, J. L., Xu, X. E., Jin, X., Hu, X., et al. (2021). MYC suppresses STING-dependent innate immunity by transcriptionally upregulating DNMT1 in triple-negative breast cancer. *J. Immunother. Cancer* 9 (7), e002528. doi:10.1136/jitc-2021-002528

Xiao, Y., Ma, D., Zhao, S., Suo, C., Shi, J., Xue, M. Z., et al. (2019). Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms of triple-negative breast cancer. *Clin. Cancer Res.* 25 (16), 5002–5014. doi:10.1158/1078-0432.CCR-18-3524

Xie, Y., Hu, Y., Zhou, N., Yao, C., Wu, L., Liu, L., et al. (2020). CAR T-cell therapy for triple-negative breast cancer: Where we are. *Cancer Lett.* 491 (8), 121–131. doi:10.1016/j.canlet.2020.07.044

Xiong, G., Stewart, R. L., Chen, J., Gao, T., Scott, T. L., Samayoa, L. M., et al. (2018). Collagen prolyl 4-hydroxylase 1 is essential for HIF-1α stabilization and TNBC chemoresistance. *Nat. Commun.* 9 (1), 4456. doi:10.1038/s41467-018-06893-9

Yang, B., Chou, J., Tao, Y., Wu, D., Wu, X., Li, X., et al. (2018). An assessment of prognostic immunity markers in breast cancer. *npj Breast Cancer* 4 (1), 35. doi:10.1038/s41523-018-0088-0

Yarchoan, M., Hopkins, A., and Jaffee, E. M. (2017). Tumor mutational burden and response rate to PD-1 inhibition. *N. Engl. J. Med.* 377 (25), 2500–2501. doi:10.1056/nejmc1713444

Yeong, J., Thike, A. A., Lim, J. C. T., Lee, B., Li, H., Wong, S. C., et al. (2017). Higher densities of Foxp3+ regulatory T cells are associated with better prognosis in triple-negative breast cancer. *Breast Cancer Res. Treat.* 163 (1), 21–35. doi:10.1007/s10549-017-4161-4

Yin, L., Duan, J. J., Bian, X. W., and Yu, S. C. (2020). Triple-negative breast cancer molecular subtyping and treatment progress. *Breast Cancer Res.* 22 (1), 61. doi:10.1186/s13058-020-01296-5

Yu, T., and Di, G. (2017). Role of tumor microenvironment in triple-negative breast cancer and its prognostic significance. *Chin. J. Cancer Res.* 29 (3), 237–252. doi:10.21147/j.issn.1000-9604.2017.03.10

Zhang, L., Wang, X. I., Ding, J., Sun, Q., and Zhang, S. (2019). The predictive and prognostic value of Foxp3+/CD25+ regulatory T cells and PD-L1 expression in triple negative breast cancer. *Ann. Diagn. Pathol.* 40, 143–151. doi:10.1016/j.anndiagpath.2019.04.004

Zhang, Y., Chen, H., Mo, H., Hu, X., Gao, R., Zhao, Y., et al. (2021). Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. *Cancer Cell* 39 (12), 1578–1593.e8. e8. doi:10.1016/j.ccell.2021.09.010

Zhou, C., Chen, H-Y., Chen, H., and Rui, W. (2016). Targeting cancer stem cells in cancer therapy. *Cancer*, 1–30. doi:10.2991/icmmbe-16.2016.17

frontiers | Frontiers in Molecular Biosciences

TYPE Methods
PUBLISHED 26 August 2022
DOI 10.3389/fmolb.2022.967310



OPEN ACCESS

EDITED BY
Fernando Schmitt,
University of Porto, Portugal

REVIEWED BY
Zisis Kozlakidis,
International Agency For Research On Cancer (IARC), France
Serena Bonin,
University of Trieste, Italy

*CORRESPONDENCE
Nicola Fusco,
nicola.fusco@ieo.it

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 12 June 2022
ACCEPTED 01 August 2022
PUBLISHED 26 August 2022

CITATION
Bonizzi G, Zattoni L, Capra M, Cassi C, Taliento G, Ivanova M, Guerini-Rocco E, Fumagalli M, Monturano M, Albini A, Viale G, Orecchia R and Fusco N (2022), Standard operating procedures for biobank in oncology.
*Front. Mol. Biosci.* 9:967310.
doi: 10.3389/fmolb.2022.967310



# Standard operating procedures for biobank in oncology

Giuseppina Bonizzi[1], Lorenzo Zattoni[1,2], Maria Capra[1], Cristina Cassi[1], Giulio Taliento[1], Mariia Ivanova[3], Elena Guerini-Rocco[2,3], Marzia Fumagalli[4], Massimo Monturano[5], Adriana Albini[6], Giuseppe Viale[1,2,3], Roberto Orecchia[6] and Nicola Fusco[1,2,3]*

[1]Biobank for Translational and Digital Medicine, IEO, European Institute of Oncology IRCCS, Milan, Italy, [2]Department of Oncology and Hemato-Oncology, University of Milan, Milan, Italy, [3]Division of Pathology, IEO, European Institute of Oncology IRCCS, Milan, Italy, [4]Technology Transfer Office, IEO, European Institute of Oncology IRCCS, Milan, Italy, [5]Patient Safety and Risk Management Service, IEO, European Institute of Oncology IRCCS, Milan, Italy, [6]Scientific Directorate, IEO, European Institute of Oncology IRCCS, Milan, Italy

Biobanks are biorepositories that collect, process, store, catalog, and distribute human biological samples, and record the associated data. The role and action field of these strategic infrastructures for implementing precision medicine in translational research is continuously evolving. To ensure the optimal quality at all stages of biobanking, specific protocols are required and should be elaborated according to updated guidelines, recommendations, laws, and rules. This article illustrates the standard operating procedures, including protocols, troubleshooting, and quality controls, of a fully certified biobank in a referral Cancer Center. This model involves all clinical departments and research groups to support the dual mission of academic cancer centers, i.e. to provide high-quality care and high-quality research. All biobanking activities based on the type of biological specimens are detailed and the most tricky methodological aspects are discussed, from patients' informed consent to specimen management.

KEYWORDS

**biobank, translational research, biomarkers, cancer research, tissue samples, liquid biopsy, standard operating procedures, quality control**

## Introduction

Modern oncologic research requires that high-quality biological samples and the associated data are collected, tracked, processed, stored, cataloged, and distributed to research groups and collaborating partners (Kinkorová, 2015). This integrated biobanking approach has led to breakthroughs in both biomarker discovery and drug development (Pagni et al., 2019). Biobanks thus represent essential resources for basic, translational, and clinical research but they also act as key players linking academic research and the pharmaceutical biotechnology industry (Vaught et al., 2011; Hewitt and Watson, 2013; Jose et al., 2018; Coppola et al., 2019). Moreover, the ability to integrate not only clinical information but also biospecimens into big data repositories has intensified

the centrality of biobanks (Margolis et al., 2014; Kinkorová, 2015; Drosou et al., 2017; Kinkorová and Topolčan, 2020). This is especially important as biorepositories have begun to incorporate patient information with comprehensive clinicopathologic, epidemiologic, and demographic data, together with multi-omics molecular information (Braun et al., 2014; Luo et al., 2014; Leff and Yang, 2015; Saifuddin et al., 2017; Bycroft et al., 2018; Hulsen et al., 2019; Bonnechère et al., 2021). The collection of this increasing amount of data requires strict quality controls and standard operating procedures (SOPs). The genomic and post-genomic field area has generated a high demand for high-quality biospecimen and data. Biorepositories in cancer research support scientists and clinicians to obtain disease-specific insights. For these reasons, biobanks should be established and updated following international guidelines, such as those from the International Agency for Research on Cancer (IARC), U.S. National Cancer Institute, United Kingdom. Confederation of Cancer Biobanks, and International Society for Biological and Environmental Repositories (ISBER), recommendations, laws, and rules (Vaught et al., 2009; Sanchini et al., 2016; Mendy et al., 2017). In this respect, networking is essential for sharing materials and data among institutions and research groups, particularly for the study of rare diseases (Montserrat and Taruscio, 2019).

Here, we present the organization of a fully certified (UNI EN ISO 9001:2015 - Certiquality) biobank in a referral Cancer Center, which is an integral part of the Italian node of the European Research Infrastructure on Biobanking (BBMRI-ERIC) (Salvaterra and Corfield, 2017). This facility works in compliance with the new standard ISO 20387:2018 "Biotechnology - Biobanking - General requirements for Biobanks". All SOPs herein reported fulfilled the BBMRI-ERIC quality control and audits (https://www.bbmri-eric.eu/services/quality-management/). Protocols and best practices for the collection of surgical tissue samples, as well as biofluids (e.g., plasma, serum, blood, urine), cell cultures, and peripheral blood mononuclear cells (PBMC), are described in detail (Kanof et al., 2001; Elliott and Peakman, 2008; Guerin et al., 2010; Mallone et al., 2011; Fisher et al., 2018; Hojat et al., 2019; Rolfo et al., 2021). This model enables collaboration among research groups and industry, allowing patients to be an integral part of translational research.

## Scientific/ethical approval and patients' recruitment

All procedures involving biobanks must be approved by the local Scientific and Technical Committee, the Ethics Committee, and the directors of the involved clinical Units and surgical programs, according to the 1964 Declaration of Helsinki, the 2018 General Data Protection Regulation (GDPR), and subsequent amendments (Sanchini et al., 2016). In this prototype, the GDPR is represented by the Scientific Research Participation Agreement (SRPA), which is the standard informed consent that patients sign to donate biological samples, sensitive data, or genetic data at the European Institute of Oncology in Milan, Italy. The SRPA should be obtained from all patients for the storage, processing, and use of the data obtained for scientific purposes. Only the signed SRPA allows for biospecimen collection. Through this agreement, each patient can express his or her will and modalities of scientific research participation. Given the complexity of the concepts described in the SRPA, it is advisable to share educational material with the patients. For example, as a reminder of their first visit, patients can receive a text message whereby the SRPA information is provided. In addition, short engaging videos broadcasted in the waiting rooms can be employed to inform patients about the importance and implications of the SRPA. An example of a cartoon on biobank used at the European Institute of Oncology, Milan, Italy is freely available online (https://vimeo.com/679070846). During each phase of the hospitalization, SRPA can be administered to the patient by qualified professionals, including biobank staff, nurses, physicians, and biomedical personnel. If patients agree to participate in any study, they should receive all the specific study information approved by the Ethics Committee. Informed consent in the form of SRPA is obtained from all patients for their material to be stored in the biobank and used for further studies. Hence, SOPs, guidelines, and recommendations do not permit the collection and storage of biospecimens in the absence of patients' consent. Therefore, SRPA should be continuously updated to inform patients properly and comprehensively. To solve any patient's withdrawal from the previous SRPA, a patient sample take-out methodology should be implemented (Schmanski et al., 2021). To obtain the collected samples, researchers, or external collaborators (for-profit or non-profit) should apply to the Biobank Scientific-Technical Committee and/or the Institutional Ethics Committee. A specific form has to be filled by the applicant for evaluation and linked to an approved evaluation of the project.

## Management of biological samples

After previous verification of the patients' SRPA, each biological sample can be collected and treated. Samples may include: fresh and frozen tissue samples related to the patients who underwent surgery; fresh and frozen tissue biopsies; cytological samples (e.g., needle aspirations, excreted, ascites, pleural fluids) from needle aspiration or brushing/scraping, and from affixing to surgically removed tissues for small lesions; biofluids (blood, serum, plasma, PBMC, oral swab, urine, feces, ascites) of patients in pre-hospitalization, patients enrolled in clinical trials, and any other subjects involved in screening projects. Each phase of samples collection should be compliant with the latest ISO standards (e.g. ISO 20387:2018).

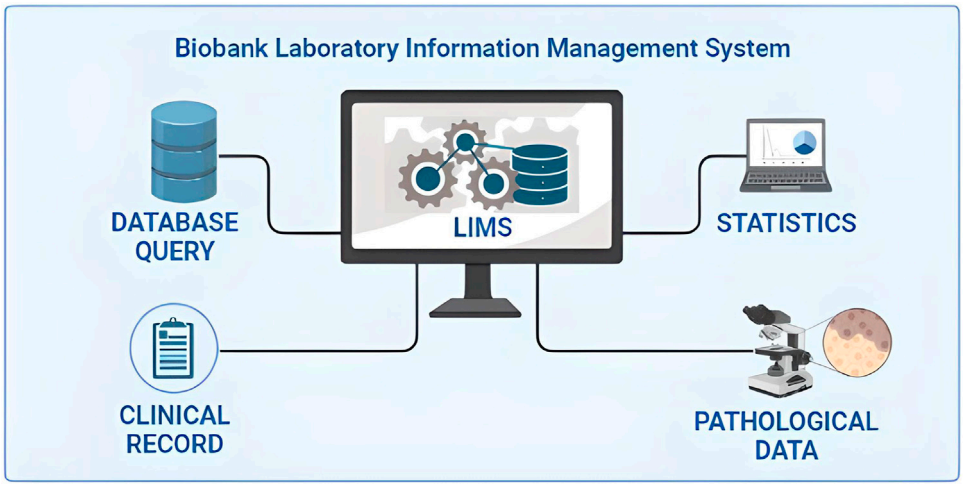



FIGURE 1
Integration of the laboratory information management system (LIMS) of biobanks in the critical junction of data from different sources.

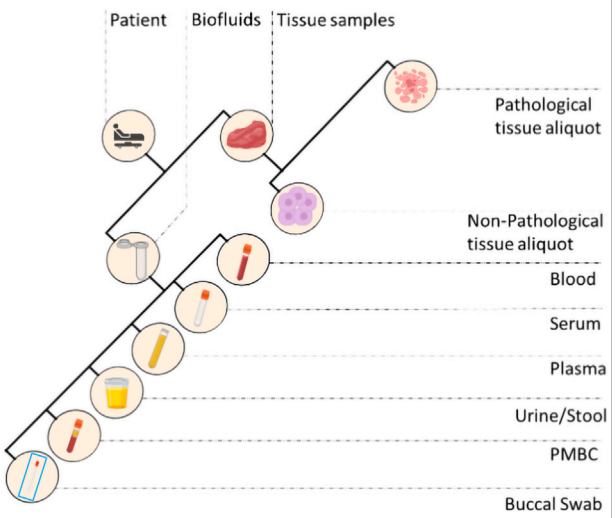



FIGURE 2
Different types of biospecimens collected in a standard biobank. PMBC, peripheral blood mononuclear cells.

TABLE 1 Examples of non-tissue sample types and corresponding biobank codes.

| Matrix | Code |
|---|---|
| ASCITES | AS |
| WHOLE BLOOD | BL |
| BRONCHOSCOPY | BS |
| BUCCAL SWAB | BU |
| BUFFY COAT | BC |
| CYTOLOGICAL SAMPLE, NOS | CY |
| FECES | F |
| PERIPHERAL BLOOD MONONUCLEAR CELLS (PBMCs) | PB |
| BLOOD PLASMA | PL |
| BLOOD SERUM | SE |
| TUBE BRUSHING | TB |
| URINE | U |

## Check-in, aliquoting, and distribution

When surgical intervention is scheduled, the Biobank data manager should check in the Surgery Plan if the patient is eligible for samples collection by the signed SRPA presence. If eligible, a surgery plan for biobank must be prepared to check the correspondence of the patient's inclusion/exclusion criteria, and the patient's consent for clinical trial or research project. To store and track the significant amount of data generated from the processing and analysis of patients' biological samples, the use a Laboratory Information Management System (LIMS) software named is highly recommended. This software is considered the biobank neural network because it should ideally interact bidirectionally with all the other softwares and applications used in the Institution, as portrayed in Figure 1. All processing and analysis are tracked by the LIMS, allowing the biobank operators to minimize possible errors, as each type of aliquot is electronically recorded using barcodes (Figure 2). Further, all samples should be divided into subcategories according to the processing, such as the fresh sample, and the frozen sample. The aliquots' ID is a sequential series of numbers and letters generated by LIMS, according to SOPs, as exemplified for biofluids in Table 1 and tissue samples in Table 2. An example

TABLE 2 Examples of tissue sample types and corresponding biobank codes.

| Atrix | Code |
|---|---|
| ABDOMEN, NOS | A |
| ADRENAL GLAND | AG |
| BONE TISSUE | BO |
| BONE MARROW | BM |
| BRAIN | BR |
| BREAST, NOS | B |
| BREAST - LEFT | B1 |
| BREAST - RIGHT | B2 |
| MESENTERY | M |
| CERVIX | CE |
| COLON | C |
| ESOPHAGUS | E |
| FIMBRIA, NOS | FI |
| FIMBRIA - LEFT | FI1 |
| FIMBRIA - RIGHT | FI2 |
| KIDNEY | K |
| ILEUM | I |
| LARYNX | LA |
| LIVER | LI |
| LUNG, NOS | L |
| LUNG - LEFT | L1 |
| LUNG - RIGHT | L2 |
| LYMPH NODE | LN |
| LYMPH NODE-ABDOMINAL | LNA |
| MESENTERY | M |
| NASOPHARYNX | NA |
| OMENTUM | OM |
| ORAL CAVITY | OR |
| OROPHARYNX | OP |
| OVARY, NOS | O |
| OVARY - LEFT | O1 |
| OVARY - RIGHT | O2 |
| PANCREAS | PA |
| PHARYNX | PH |
| PERITONEUM | PE |
| PROSTATE | *p* |
| RECTUM | R |
| SKIN | SK |
| SOFT TISSUES | ST |
| STOMACH | S |
| TESTIS | TE |
| THYMUS | TH |
| THYROID | T |
| TONGUE | TO |
| FALLOPIAN TUBE | TU |
| URINARY BLADDER | UB |
| UTERUS CORPUS | UC |

of an ID code is: "12-B1-00100-01" where "12" indicates the year (2012), "B1" indicates the anatomical origin (left breast), and "00100-01" identifies the sample. Once an aliquot is requested, a database query allows the retrieval of the requested samples. This process needs a tracked form with information, a short description of the approved research project, and information on the principal investigator (PI). Consequently, the LIMS generates the requested ID for each aliquot sorted into a picklist. Later, the biobank technicians can check the ID list to retrieve the requested aliquots. Last, barcoded tubes (e.g. Nunc™ Coded Cryobank Vial Systems, Thermo Scientific, Waltham, Massachusetts, US) and relative data can be delivered.

## Sample types

### Tissue

During the gross examination of the surgical sample, the pathologist determines whether there is sufficient material (i.e. exceeding diagnosis) for research purposes. When available/possible, the non-pathological counterpart is also collected. Normal and tumor samples are labelled and placed in sterile Petri dishes on ice and divided into fresh and/or optimal cutting temperature (OCT) compound aliquots. Samples are then frozen for 3 min at −120°C in isopentane before transfer to cryopreservation rooms and stored at −80°C. It is important to perform a quality check for each frozen aliquot before distribution and use it to obtain a histological assessment of the cellularity on hematoxylin and eosin (H&E) cryosections (Table 3).

### Whole blood

For whole blood samples collection, 6 ml labeled vacutainer tubes, containing anti-coagulant Na2 ethylenediaminetetraacetic acid (EDTA), are used by the nurses at the assessment centers. Following the collection, the biobank technicians process the sample using its medical record number and the episode code and register each aliquot in the biobank software. For clinical studies, or project-specific requirements, the blood is prepared for shipment, otherwise it is firstly collected using a blood amount of 900 µL in 1 ml barcoded cryotubes for one or more aliquots, and stored at −80°C.

### Blood serum

For blood serum samples collection, 6 ml labeled vacutainer tubes containing a thixotropic barrier gel at the bottom of the tube, are initially used. Tubes are left for clotting for 3 h at room temperature (RT), and then centrifugated on a refrigerated centrifuge at 828 x g for 10 min. Depending on the initial amount of whole blood taken, the biobank technicians' rate 450 µL of serum in the 0.5 ml barcoded cryotubes for the maximum amount of aliquots, and stored at −80°C. When the

TABLE 3 Representative quality Control Form of tissue sections included in OCT and frozen.

| **Label with biobank code** | — |
|---|---|
| Tumor Tissue (%) | — |
| Tumor Tissue Description | — |
| Normal Tissue Counterpart (%) | — |
| Normal Tissue Counterpart Description | — |
| Necrotic Tissue (%) | — |
| Adipose Tissue (%) | — |
| Stromal Tissue (%) | — |
| Inflammation | ☐ Absent |
| | ☐ Sparse |
| | ☐ Intermediate |
| | ☐ Extensive |
| Diameter of Section (mm) | |
| Conclusion | ☐ Insufficient |
| ☐ Sufficient | |
| Notes | — |
| Technical Operator | — |
| Pathologist Operator | — |

last aliquot serum volume is < 450 µL, the sample is registered as "leftover".

### Blood plasma and cf-DNA/RNA

To separate the plasma from the whole blood, blood-filled vacutainers need to be centrifuged at 2,000 x g for 10 min at RT. After centrifugation, the upper plasma layer should be removed and transferred to a sterile 15 ml conical tube for a second centrifugation at 16,000 x g for 10 min at RT to remove contaminating blood cells. Then, the obtained plasma can be transferred to a barcoded cryotube for the maximum amount of aliquots, depending on the volume. The remaining blood is collected for subsequent cf-DNA and/or cf-RNA purification. Aliquots should be stored at −80°C.

### Peripheral blood mononuclear cells

For peripheral blood mononuclear cells (PBMC) isolation, the blood is drawn in 7.5 ml labeled vacutainer tubes, containing $Na_2$ EDTA as an anticoagulant reagent. The blood is subsequently transferred into an empty 50 ml conical tube and diluted in a 1:1 ratio using phosphate buffered saline (PBS) 1X (e.g., 15 ml of blood +15 ml of PBS 1X). Again, the diluted blood is layered on the top of a clean 50 ml conical tube containing 15 ml of Ficoll, without mixing the two solutions. After centrifugation at 400 x g for 30 min at RT, the white layer containing PBMC is collected and placed in a new sterile 50 ml conical tube. PBMCs are washed by adding 45 ml of PBS, mixed, and centrifuged at 400 x g for 10 min at 4°C. After discarding the supernatant, the pellet containing PBMCs is resuspended in PBS and counted using Tuerks solution and single-use slide for counting cells (e.g. Biosigma S.P.A. Cat. no. 347143/001). Cells are washed once more with PBS and resuspended at 2-3x10$^6$ cells/mL of FSB +10% dimethyl sulfoxide (DMSO) to be frozen. Finally, 1 ml of resuspended cells are transferred in cryotubes for storage at −80°C.

### Stool and buccal swab

Feces and buccal swabs are collected in 15 ml tubes (e.g. Stool Sample Collection and Stabilization Kit Canvax Cat. no 0013), containing 50 mM Tris-HCl, 10 mM NaCl, and 10 mM EDTA pH 7.5 and stored at −80°C.

## Disaster recovery plan

Biobanks are dedicated to managing valuable and possibly irreplaceable biological specimens. Therefore, biomaterials and associated data must be managed and protected carefully as their loss can destroy years of research efforts and costs, and potentially result in damage to the Institution (Eng and Tan, 2019). For this reason, risk management and practical crisis management plans must be established for any biobank (Parry-Jones et al., 2017). It is essential to define a data protection program that must satisfy various needs that may range from remote data only (backup) to disaster recovery (as a set of technological measures and organizational processes aimed at restoring systems, data, and infrastructures necessary to provide biobank services during emergencies) and to ensure continuity of service and recovery of materials and data during emergencies (Cicek and Olson, 2020).

## Staff training programs

The quality and quantity of samples and data stored in a biobank directly depend on the biobank personnel, including data managers and technicians (Hartung et al., 2021). Modern biobanking must rely on high-level training programs for biobank employees not only to allow harmonization of correct sample handling but also to ensure safety and quality (Kinkorová and Topolčan, 2020). Not surprisingly, training certificates of biobank employees are needed for the accreditation process (Williams et al., 2019). Types of training programs include master's programs, certificate courses, and workshops. Due to the paucity of available formal training programs, biobanks often train most of their new staff on site (Castellanos-Uribe et al., 2020). Learning about teamwork, personnel safety, patient privacy, biospecimen quality, and SOPs is crucial not only for efficiency and productivity but also for the personnel's career success. A well-designed training program should include helpful

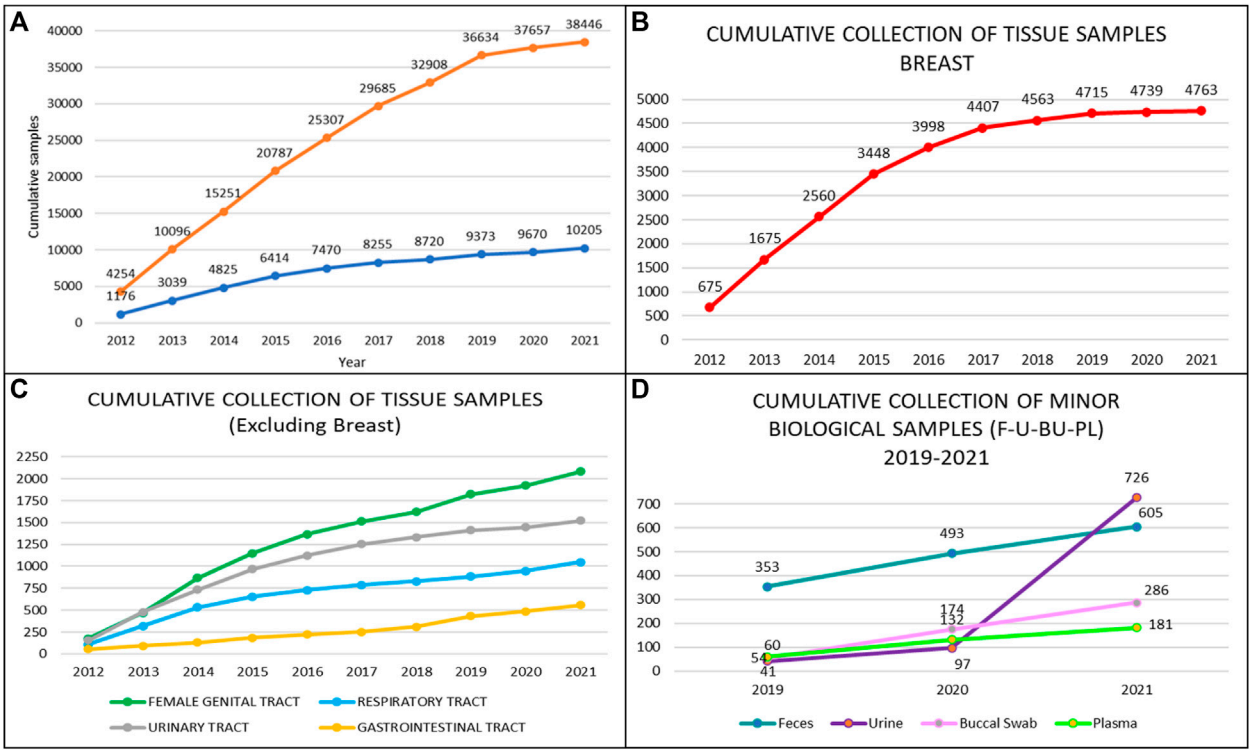



FIGURE 3
Types and number of samples collected by the biobank of the European Institute of Oncology by year. **(A)** Cumulative collection of tissue and blood/serum samples; **(B)** Cumulative collection of breast tissue samples; **(C)** Cumulative collection of tumor samples from the ovary, prostate, lung, and colon; **(D)** Cumulative collection of non-tissue samples, i.e. feces, saliva/swab, plasma, and urine.

tips, tricks, and troubleshooting. International collaboration and exchange programs might facilitate the process of creating next-generation biobanking staff.

## Representative results

A total of 38,446 annotated biofluids and a total of 10,205 tissue samples were collected by the Biobank for Translational and Digital Medicine Unit at the IEO, European Institute of Oncology, Milan, Italy from April 2012 to December 2021 (Figure 3A). The highest number of samples were related to breast cancer, urological malignancies, tumors of the female genital tract, head and neck carcinomas, and lung cancer (Figures 3B,C). The cumulative analysis of plasma, buccal swab, urine, and stool samples revealed a significant increase in the number of collected samples, particularly for the urine in patients with urological malignancies, reaching a total number of 726 urine samples (Figure 3D). These samples have been divided into multiple aliquots related to specific projects or clinical trials and to the institutional universal collection of samples. Taken together, 28 different projects were responsible for the vast majority of aliquot distribution, for a total number of 28,852 aliquots, as detailed in Table 4. The total number of tissue samples whose aliquots were employed for research purposes were 8,383/10,205 (82%). These data confirm the fundamental role of certified biobanks not only for samples collection but also for samples distribution and use by research groups.

## Discussion

The transversal role of biobanks in scientific research, particularly in oncologic pathology, and basic and clinical sciences, has put these important infrastructures on the front line of personalized medicine evolution (Kinkorová, 2015; Coppola et al., 2019). Indeed, cancer is still a leading cause of morbidity and mortality worldwide (Sung et al., 2021). For these reasons, the understanding of cancer pathogenesis, mechanisms of disease, and biomarkers discovery at the multi-omics level is becoming an urgent clinical need, akin the support in drug discovery (Kinkorová, 2015; Yan et al., 2018; Szustakowski et al., 2021). When a biobank is established several challenges, from methodological to operative and ethical issues need to be

10.3389/fmolb.2022.967310

TABLE 4 Distribution of samples aliquots from the biobank of the European Institute of Oncology.

| Project ID | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| T-CELL | — | — | — | — | — | — | 46 | 182 | 144 | 137 | 509 |
| BLADDER | — | — | — | — | 5 | 16 | — | — | — | — | 21 |
| BREAST-1 | 1,168 | 1,611 | 1,340 | 756 | 258 | 65 | 217 | 160 | 12 | 40 | 5,627 |
| BREAST-2 | — | — | — | — | — | — | 26 | 79 | 100 | 65 | 270 |
| BREAST-3 | — | — | — | — | — | — | — | 122 | 122 | 90 | 334 |
| COLON-1 | — | — | — | — | — | — | 22 | — | 185 | — | 207 |
| COLON-2 | — | — | — | 46 | 85 | 97 | 105 | 37 | — | — | 370 |
| COLON-3 | — | — | — | — | — | — | 20 | 14 | 13 | — | 47 |
| COLON-4 | — | — | — | — | — | — | — | 563 | — | — | 563 |
| COSMOS | 5 | 14 | 4 | 3,000 | 363 | — | — | — | — | — | 3,386 |
| FAM | — | — | — | — | 164 | — | — | — | — | — | 164 |
| H&N-1 | — | — | — | 6 | 46 | 22 | 56 | 51 | 20 | 134 | 335 |
| H&N-2 | — | — | — | — | — | — | — | — | — | 35 | 35 |
| H&N-3 | — | — | — | — | — | — | — | — | — | 65 | 65 |
| LUNG-1 | 246 | 311 | 338 | 103 | 26 | 20 | — | — | — | — | 1,044 |
| LUNG-2 | — | — | — | — | — | — | — | 80 | 498 | 342 | 920 |
| LUNG-3 | — | — | — | — | — | — | 186 | — | — | — | 186 |
| MEL-1 | — | — | — | — | — | — | — | — | — | 2 | 2 |
| MEL-2 | — | — | — | — | — | — | — | 628 | 752 | 417 | 1,797 |
| miRNA | 120 | 117 | 206 | 214 | 50 | — | — | — | — | — | 707 |
| OVARY-1 | 320 | 569 | 696 | 561 | 402 | 312 | 204 | 664 | 146 | 166 | 4,040 |
| OVARY-2 | | | | | | | 26 | 113 | 102 | 77 | 318 |
| PROSTATE | 275 | 545 | 299 | 200 | 127 | 128 | 181 | 66 | 23 | | 1,844 |
| SARCOMA | — | — | — | — | — | — | — | — | — | 9 | 9 |
| SKIN | — | — | — | — | — | — | 23 | 10 | — | 47 | 80 |
| STOMACH | — | — | — | — | — | — | — | — | — | 193 | 193 |
| TEST COVID | — | — | — | — | — | — | — | | 692 | | 692 |
| THYMUS | — | — | — | — | 2 | 19 | 24 | 18 | 11 | 13 | 87 |
| **TOTAL** | 2,134 | 3,167 | 2,883 | 4,886 | 1,528 | 679 | 1,136 | 2,787 | 2,820 | 1,832 | 23,852 |

assessed. The first important point is to manage the existing institution data. This can be done by using a laboratory information management system software able to receive and integrate different types of information, from clinical to pathological and digital data. In literature, several valuable softwares have been used to implement biobank databases (Tukacs et al., 2012; Paul et al., 2017; Fthenou et al., 2019; Im et al., 2019), and some freeware can be obtained for biobank management (Voegele et al., 2013; Willers et al., 2021). A biobank consent form is another critical step in data acquisition and management (Beskow et al., 2017; Kinkorová et al., 2019; Kasperbauer et al., 2021; Schmanski et al., 2021). The application of an adequate SRPA is a necessary agreement on legal and ethical aspects of the patient's data storage and usage (D'Abramo et al., 2015; Sanchini et al., 2016). All SOPs described in this work are continuously under evaluation and improvement and they are currently compliant with the ISO 20387: 2018 standards. It should be noted, however, that standardization and improvement of pre-analytical procedures for *in-vitro* diagnostics is a continuous process. The most updated high-priority pre-analytical CEN and ISO standard documents as well as corresponding External Quality Assessment (EQA) schemes and implementation tools are detailed in Table 5. Not only following adequate SOPs is essential to secure research achievements, but also have qualified personnel, aware of the biobank's role and potential (Caixeiro et al., 2016; Kintossou et al., 2020). Another important aspect related to the multidisciplinary collaboration in biobanking is represented by the pathologist (Angerilli et al., 2021). Pathologists are the only professionals able to ensure both the tissue sampling for diagnosis and the biobank. Finally, it should be mentioned that the efforts and resources invested to set up and sustain a biobank are significant and such work should be traced and, most importantly, recognized in scientific publications (Howard et al., 2018). In this respect, the Bioresource

TABLE 5 European Committee for Standardization Technical Committee (CEN/TC) 140 *in vitro* diagnostic medical devices published standards. All projects are sorted by date and available at https://www.spidia.eu/projects/standard-documents (Accessed 28 July 2022).

| References | Date | Title |
|---|---|---|
| CEN/TS 17811:2022 | 22 June 2022 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for urine and other body fluids - Isolated cell free DNA |
| CEN/TS 17747:2022 | 20 April 2022 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for exosomes and other extracellular vesicles in venous whole blood - DNA, RNA and proteins |
| CEN/TS 17742:2022 | 30 March 2022 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for venous whole blood - Isolated circulating cell free RNA from plasma |
| EN ISO 20776-2:2022 | 19 January 2022 | Clinical laboratory testing and *in vitro* diagnostic test systems - Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices - Part 2: Evaluation of performance of antimicrobial susceptibility test devices against References broth micro-dilution (ISO 20776-2:2021) |
| CEN/TS 17688-2:2021 | 22 December 2021 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for Fine Needle Aspirates (FNAs) - Part 2: Isolated proteins |
| CEN/TS 17688-1:2021 | 22 December 2021 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for Fine Needle Aspirates (FNAs) - Part 1: Isolated cellular RNA |
| CEN/TS 17688-3:2021 | 22 December 2021 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for Fine Needle Aspirates (FNAs) - Part 3: Isolated genomic DNA |
| EN ISO 4307:2021 | 3 November 2021 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for saliva - Isolated human DNA (ISO 4307:2021) |
| EN ISO 16256:2021 | 27 October 2021 | Clinical laboratory testing and *in vitro* diagnostic test systems - Broth micro-dilution References method for testing the *in vitro* activity of antimicrobial agents against yeast fungi involved in infectious diseases (ISO 16256:2021) |
| EN ISO 6717:2021 | 8 September 2021 | *In vitro* diagnostic medical devices - Single-use containers for the collection of specimens from humans other than blood (ISO 6717:2021) |
| EN ISO 20166-4:2021 | 28 July 2021 | Molecular *in vitro* diagnostic examinations - Specifications for preexamination processes for formalin-fixed and paraffin-embedded (FFPE) tissue - Part 4: *In situ* detection techniques (ISO 20166-4:2021) |
| EN ISO 23162:2021 | 14 July 2021 | Basic semen examination - Specification and test methods (ISO 23162:2021) |
| EN ISO 17511:2021 | 2 June 2021 | *In vitro* diagnostic medical devices - Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples (ISO 17511:2020) |
| EN ISO 23118:2021 | 2 June 2021 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes in metabolomics in urine, venous blood serum and plasma (ISO 23118:2021) |
| EN ISO 20184-3:2021 | 26 May 2021 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for frozen tissue - Part 3: Isolated DNA (ISO 20184-3:2021) |
| CEN/TS 17626:2021 | 5 May 2021 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for human specimen - Isolated microbiome DNA |
| EN ISO 20776-1:2020 | 1 July 2020 | Susceptibility testing of infectious agents and evaluation of performance of antimicrobial susceptibility test devices - Part 1: Broth micro-dilution References method for testing the *in vitro* activity of antimicrobial agents against rapidly growing aerobic bacteria involved in infectious diseases (ISO 20776-1:2019, including Corrected version 2019-12) |
| EN ISO 22367:2020 | 11 March 2020 | Medical laboratories - Application of risk management to medical laboratories (ISO 22367:2020) |
| CEN/TS 17390-1:2020 | 22 January 2020 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for circulating tumor cells (CTCs) in venous whole blood - Part 1: Isolated RNA |
| CEN/TS 17390-2:2020 | 22 January 2020 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for circulating tumor cells (CTCs) in venous whole blood - Part 2: Isolated DNA |
| CEN/TS 17390-3:2020 | 22 January 2020 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for circulating tumor cells (CTCs) in venous whole blood - Part 3: Preparations for analytical CTC staining |
| EN ISO 20186-3:2019 | 23 October 2019 | Molecular *in-vitro* diagnostic examinations - Specifications for pre-examination processes for venous whole blood - Part 3: Isolated circulating cell free DNA from plasma (ISO 20186-3:2019) |
| EN ISO 20186-1:2019 | 27 March 2019 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for venous whole blood - Part 1: Isolated cellular RNA (ISO 20186-1:2019) |
| EN ISO 20186-2:2019 | 27 March 2019 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for venous whole blood - Part 2: Isolated genomic DNA (ISO 20186-2:2019) |
| EN ISO 15195:2019 | 6 February 2019 | Laboratory medicine - Requirements for the competence of calibration laboratories using References measurement procedures (ISO 15195:2018) |
| EN ISO 20166-3:2019 | 23 January 2019 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for formalin-fixed and paraffin-embedded (FFPE) tissue - Part 3: Isolated DNA (ISO 20166-3:2018) |
| EN ISO 20166-2:2018 | 19 December 2018 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examinations processes for formalin-fixed and paraffin-embedded (FFPE) tissue - Part 2: Isolated proteins (ISO 20166-2:2018) |

(Continued on following page)

TABLE 5 (*Continued*) European Committee for Standardization Technical Committee (CEN/TC) 140 *in vitro* diagnostic medical devices published standards. All projects are sorted by date and available at https://www.spidia.eu/projects/standard-documents (Accessed 28 July 2022).

| References | Date | Title |
|---|---|---|
| EN ISO 20166-1:2018 | 19 December 2018 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for formalin-fixed and paraffin-embedded (FFPE) tissue - Part 1: Isolated RNA (ISO 20166-1:2018) |
| EN ISO 20184-1:2018 | 19 December 2018 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for frozen tissue - Part 1: Isolated RNA (ISO 20184-1:2018) |
| EN ISO 20184-2:2018 | 12 December 2018 | Molecular *in vitro* diagnostic examinations - Specifications for pre-examination processes for frozen tissue - Part 2: Isolated proteins (ISO 20184-2:2018) |
| EN ISO 6710:2017 | 6 September 2017 | Single-use containers for human venous blood specimen collection (ISO 6710:2017) |
| EN ISO 22870:2016 | 30 November 2016 | Point-of-care testing (POCT) - Requirements for quality and competence (ISO 22870:2016) |
| EN ISO 15197:2015 | 10 June 2015 | *In vitro* diagnostic test systems - Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus (ISO 15197:2013) |
| EN ISO 23640:2015 | 10 June 2015 | *In vitro* diagnostic medical devices - Evaluation of stability of *in vitro* diagnostic reagents (ISO 23640:2011) |
| EN ISO 19001:2013 | 20 March 2013 | *In vitro* diagnostic medical devices - Information supplied by the manufacturer with *in vitro* diagnostic reagents for staining in biology (ISO 19001:2013) |
| EN ISO 15189:2012 | 1 November 2012 | Medical laboratories - Requirements for quality and competence (ISO 15189:2012, Corrected version 2014-08-15) |
| EN ISO 18113-5:2011 | 19 October 2011 | *In vitro* diagnostic medical devices - Information supplied by the manufacturer (labelling) - Part 5: *In vitro* diagnostic instruments for self-testing (ISO 18113-5:2009) |
| EN ISO 18113-2:2011 | 19 October 2011 | *In vitro* diagnostic medical devices - Information supplied by the manufacturer (labelling) - Part 2: *In vitro* diagnostic reagents for professional use (ISO 18113-2:2009) |
| EN ISO 18113-3:2011 | 19 October 2011 | *In vitro* diagnostic medical devices - Information supplied by the manufacturer (labelling) - Part 3: *In vitro* diagnostic instruments for professional use (ISO 18113-3:2009) |
| EN ISO 18113-4:2011 | 19 October 2011 | *In vitro* diagnostic medical devices - Information supplied by the manufacturer (labelling) - Part 4: *In vitro* diagnostic reagents for self-testing (ISO 18113-4:2009) |
| EN ISO 18113-1:2011 | 19 October 2011 | *In vitro* diagnostic medical devices - Information supplied by the manufacturer (labelling) - Part 1: Terms, definitions and general requirements (ISO 18113-1:2009) |
| EN ISO 15193:2009 | 1 May 2009 | *In vitro* diagnostic medical devices - Measurement of quantities in samples of biological origin - Requirements for content and presentation of References measurement procedures (ISO 15193:2009) |
| EN ISO 15194:2009 | 1 May 2009 | *In vitro* diagnostic medical devices - Measurement of quantities in samples of biological origin - Requirements for certified References materials and the content of supporting documentation (ISO 15194:2009) |
| EN 14136:2004 | 19 May 2004 | Use of external quality assessment schemes in the assessment of the performance of *in vitro* diagnostic examination procedures |
| EN ISO 18153:2003 | 15 August 2003 | *In vitro* diagnostic medical devices - Measurement of quantities in biological samples - Metrological traceability of values for catalytic concentration of enzymes assigned to calibrators and control materials (ISO 18153:2003) |
| EN 13975:2003 | 19 March 2003 | Sampling procedures used for acceptance testing of *in vitro* diagnostic medical devices - Statistical aspects |
| EN 13612:2002/AC:2002 | 18 December 2002 | Performance evaluation of *in vitro* diagnostic medical devices |
| EN 13641:2002 | 8 May 2002 | Elimination or reduction of risk of infection related to *in vitro* diagnostic reagents |
| EN 13532:2002 | 17 April 2002 | General requirements for *in vitro* diagnostic medical devices for self-testing |
| EN 13612:2002 | 20 March 2002 | Performance evaluation of *in vitro* diagnostic medical devices |
| EN 12322:1999/A1:2001 | 24 October 2001 | *In vitro* diagnostic medical devices - Culture media for microbiology - Performance criteria for culture media |
| EN 12322:1999 | 21 April 1999 | *In vitro* diagnostic medical devices - Culture media for microbiology - Performance criteria for culture media |
| EN 1659:1996 | 20 November 1996 | *In vitro* diagnostic systems - Culture media for microbiology - Terms and definitions |

Research Impact Factor/Framework (BRIF) initiative was proposed for transparency and to promote the responsible and effective use of biomaterials (Cambon-Thomsen, 2003). Another point that is worth mentioning is related to the integration of artificial intelligence (AI) and machine learning into modern biobanks (Kinkorová and Topolčan, 2020; Eccher et al., 2021; Narita et al., 2021; Rizzo et al., 2022). This would allow for the integration of a digitalized database with digital pathology and high throughput molecular data, potentially representing a quantum leap in biobanking.

## Data availability statement

Requests to access the datasets should be directed to B4M=ED@ieo.it.

## Ethics statement

The studies involving human participants were reviewed and approved by IEO IRB. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

Methodology, GB, MC, GT, CC, EG-R, MM, NF, Investigation, NF, GB, MC Writing—original draft, GB, LZ, Writing—Review and Editing, GB, MI, MF, AA, RO, GV, and NF.

## Acknowledgments

The authors would like to thank all the patients that during the last decade actively participated in our research programs through the donation of their biospecimens. Without them, this research would not be possible. We are also grateful to all the personnel working at IEO, nurses, technicians, biologists, doctors, and the Directors of all the clinical and research units. The authors are grateful to Pier Paolo Di Fiore and Giancarlo Pruneri for their guidance since the very beginning of our Biobank Unit. Finally, we dedicate this work to Umberto Veronesi, the founder of IEO, and his pioneering approach to integrating cancer research and patient care.

## Conflict of interest

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

## Publisher's note

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

## References

Angerilli, V., Galuppini, F., Pagni, F., Fusco, N., Malapelle, U., and Fassan, M. (2021). The role of the pathologist in the next-generation era of tumor molecular characterization. *Diagnostics* 11 (2), 339. doi:10.3390/diagnostics11020339

Beskow, L. M., Lin, L., Dombeck, C. B., Gao, E., and Weinfurt, K. P. (2017). Improving biobank consent comprehension: A national randomized survey to assess the effect of a simplified form and review/retest intervention. *Genet. Med.* 19 (5), 505–512. doi:10.1038/gim.2016.157

Bonnechère, B., Liu, J., Thomson, A., Amin, N., and Van Duijn, C. M. (2021). Ethnicity influences risk of dementia in the UK Biobank. *Alzheimer's. Dementia* 17 (10), e056077. doi:10.1002/alz.056077

Braun, K. L., Tsark, J. U., Powers, A., Croom, K., Kim, R., Gachupin, F. C., et al. (2014). Cancer patient perceptions about biobanking and preferred timing of consent. *Biopreserv. Biobank.* 12 (2), 106–112. doi:10.1089/bio.2013.0083

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562 (7726), 203–209. doi:10.1038/s41586-018-0579-z

Caixeiro, N. J., Byun, H. L., Descallar, J., Levesque, J. V., de Souza, P., and Soon Lee, C. (2016). Health professionals' opinions on supporting a cancer biobank: Identification of barriers to combat biobanking pitfalls. *Eur. J. Hum. Genet.* 24 (5), 626–632. doi:10.1038/ejhg.2015.191

Cambon-Thomsen, A. (2003). Assessing the impact of biobanks. *Nat. Genet.* 34 (1), 25–26. doi:10.1038/ng0503-25b

Castellanos-Uribe, M., Gormally, E., Zhou, H., Matzke, E., and Watson, P. H. (2020). Biobanking education. *Biopreserv. Biobank.* 18 (1), 1–3. doi:10.1089/bio.2019.29062.mjc

Cicek, M. S., and Olson, J. E. (2020). Mini-Review of laboratory operations in biobanking: Building biobanking resources for translational research. *Front. Public Health* 8, 362. doi:10.3389/fpubh.2020.00362

Coppola, L., Cianflone, A., Grimaldi, A. M., Incoronato, M., Bevilacqua, P., Messina, F., et al. (2019). Biobanking in health care: Evolution and future directions. *J. Transl. Med.* 17 (1), 172. doi:10.1186/s12967-019-1922-3

D'Abramo, F., Schildmann, J., and Vollmann, J. (2015). Research participants' perceptions and views on consent for biobank research: A review of empirical data and ethical analysis. *BMC Med. Ethics* 16, 60. doi:10.1186/s12910-015-0053-5

Drosou, M., Jagadish, H. V., Pitoura, E., and Stoyanovich, J. (2017). Diversity in big data: A review. *Big Data* 5 (2), 73–84. doi:10.1089/big.2016.0054

Eccher, A., Fontanini, G., Fusco, N., Girolami, I., Graziano, P., Rocco, E. G., et al. (2021). Digital slides as an effective tool for programmed death ligand 1 combined positive score assessment and training: Lessons learned from the "programmed death ligand 1 key learning program in head-and-neck squamous cell carcinoma. *J. Pathol. Inf.* 12, 1. doi:10.4103/jpi.jpi_63_20

Elliott, P., and Peakman, T. C. (2008). The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. *Int. J. Epidemiol.* 37 (2), 234–244. doi:10.1093/ije/dym276

Eng, C. B., and Tan, W. L. (2019). Disaster prevention and recovery. *Methods Mol. Biol.* 1897, 31–41. doi:10.1007/978-1-4939-8935-5_4

Fisher, W. E., Cruz-Monserrate, Z., McElhany, A. L., Lesinski, G. B., Hart, P. A., Ghosh, R., et al. (2018). Standard operating procedures for biospecimen collection, processing, and storage: From the consortium for the study of chronic pancreatitis, diabetes, and pancreatic cancer. *Pancreas* 47 (10), 1213–1221. doi:10.1097/mpa.0000000000001171

Fthenou, E., Al Emadi, A., Mahal, F. F., Chettupuzhakaran, L. T., Al Thani, A., and Afifi, N. (2019). Conception, implementation, and integration of heterogenous information technology infrastructures in the Qatar biobank. *Biopreserv. Biobank.* 17 (6), 494–505. doi:10.1089/bio.2019.0067

Guerin, J. S., Murray, D. W., McGrath, M. M., Yuille, M. A., McPartlin, J. M., and Doran, P. P. (2010). Molecular medicine Ireland guidelines for standardized biobanking. *Biopreserv. Biobank.* 8 (1), 3–63. doi:10.1089/bio.2010.8101

Hartung, M. L., Baber, R., Herpel, E., Specht, C., Brucker, D. P., Schoneberg, A., et al. (2021). Harmonization of biobank education for biobank technicians: Identification of learning objectives. *BioTech.* 10 (2), 7. doi:10.3390/biotech10020007

Hewitt, R., and Watson, P. (2013). Defining biobank. *Biopreserv. Biobank.* 11 (5), 309–315. doi:10.1089/bio.2013.0042

Hojat, A., Wei, B., Olson, M. G., Mao, Q., and Yong, W. H. (2019). Procurement and storage of surgical biospecimens. *Methods Mol. Biol.* 1897, 65–76. doi:10.1007/978-1-4939-8935-5_7

Howard, H. C., Mascalzoni, D., Mabile, L., Houeland, G., Rial-Sebbag, E., and Cambon-Thomsen, A. (2018). How to responsibly acknowledge research work in the era of big data and biobanks: Ethical aspects of the Bioresource research

impact factor (BRIF). *J. Community Genet.* 9 (2), 169–176. doi:10.1007/s12687-017-0332-6

Hulsen, T., Jamuar, S. S., Moody, A. R., Karnes, J. H., Varga, O., Hedensted, S., et al. (2019). From big data to precision medicine. *Front. Med.* 6, 34. doi:10.3389/fmed.2019.00034

Im, K., Gui, D., and Yong, W. H. (2019). An introduction to hardware, software, and other information technology needs of biomedical biobanks. *Methods Mol. Biol.* 1897, 17–29. doi:10.1007/978-1-4939-8935-5_3

Jose, R., Rooney, R., Nagisetty, N., Davis, R., and Hains, D. (2018). Biorepository and integrative genomics initiative: Designing and implementing a preliminary platform for predictive, preventive and personalized medicine at a pediatric hospital in a historically disadvantaged community in the USA. *Epma J.* 9 (3), 225–234. doi:10.1007/s13167-018-0141-y

Kanof, M. E., Smith, P. D., and Zola, H. (2001). Isolation of whole mononuclear cells from peripheral blood and cord blood. *Curr. Protoc. Immunol.* 7, 1. doi:10.1002/0471142735.im0701s19

Kasperbauer, T. J., Schmidt, K. K., Thomas, A., Perkins, S. M., and Schwartz, P. H. (2021). Incorporating biobank consent into a healthcare setting: Challenges for patient understanding. *AJOB Empir. Bioeth.* 12 (2), 113–122. doi:10.1080/23294515.2020.1851313

Kinkorová, J., and Topolčan, O. (2020). Biobanks in the era of big data: Objectives, challenges, perspectives, and innovations for predictive, preventive, and personalised medicine. *EPMA J.* 11 (3), 333–341. doi:10.1007/s13167-020-00213-2

Kinkorová, J., Topolčan, O., and Kučera, R. (2019). Informed consent in the newly established biobank. *Int. J. Environ. Res. Public Health* 16 (20), 3943. doi:10.3390/ijerph16203943

Kinkorová, J. (2015). Biobanks in the era of personalized medicine: Objectives, challenges, and innovation: Overview. *Epma J.* 7 (1), 4. doi:10.1186/s13167-016-0053-7

Kintossou, A. K., N'dri, M. K., Money, M., Cissé, S., Doumbia, S., Soumahoro, M.-K., et al. (2020). Study of laboratory staff' knowledge of biobanking in Côte d'Ivoire. *BMC Med. Ethics* 21 (1), 88. doi:10.1186/s12910-020-00533-y

Leff, D. R., and Yang, G.-Z. (2015). Big data for precision medicine. *Engineering* 1 (3), 277–279. doi:10.15302/J-ENG-2015075

Luo, J., Guo, X. R., Tang, X. J., Sun, X. Y., Yang, Z. S., Zhang, Y., et al. (2014). Intravital biobank and personalized cancer therapy: The correlation with omics. *Int. J. Cancer* 135 (7), 1511–1516. doi:10.1002/ijc.28632

Mallone, R., Mannering, S. I., Brooks-Worrell, B. M., Durinovic-Belló, I., Cilio, C. M., Wong, F. S., et al. (2011). Isolation and preservation of peripheral blood mononuclear cells for analysis of islet antigen-reactive T cell responses: Position statement of the T-cell workshop committee of the immunology of diabetes society. *Clin. Exp. Immunol.* 163 (1), 33–49. doi:10.1111/j.1365-2249.2010.04272.x

Margolis, R., Derr, L., Dunn, M., Huerta, M., Larkin, J., Sheehan, J., et al. (2014). The national institutes of health's big data to knowledge (BD2K) initiative: Capitalizing on biomedical big data. *J. Am. Med. Inf. Assoc.* 21 (6), 957–958. doi:10.1136/amiajnl-2014-002974

Mendy, M., Caboux, E., Lawlor, R. T., Wright, J., and Wild, C. P. (2017). *Common minimum technical standards and protocols for biobanks dedicated to cancer research*. Lyon FR: IARC Technical Publications International Agency for Research on Cancer.

Montserrat, A., and Taruscio, D. (2019). Policies and actions to tackle rare diseases at European level. *Ann. Ist. Super. Sanita* 55 (3), 296–304. doi:10.4415/ann_19_03_17

Narita, A., Ueki, M., and Tamiya, G. (2021). Artificial intelligence powered statistical genetics in biobanks. *J. Hum. Genet.* 66 (1), 61–65. doi:10.1038/s10038-020-0822-y

Pagni, F., Guerini-Rocco, E., Schultheis, A. M., Grazia, G., Rijavec, E., Ghidini, M., et al. (2019). Targeting immune-related biological processes in solid tumors: We do need biomarkers. *Int. J. Mol. Sci.* 20 (21), E5452. doi:10.3390/ijms20215452

Parry-Jones, A., Hansen, J., Simeon-Dubach, D., and Bjugn, R. (2017). Crisis management for biobanks. *Biopreserv. Biobank.* 15 (3), 253–263. doi:10.1089/bio.2016.0048

Paul, S., Gade, A., and Mallipeddi, S. (2017). The state of cloud-based biospecimen and biobank data management tools. *Biopreserv. Biobank.* 15 (2), 169–172. doi:10.1089/bio.2017.0019

Rizzo, P. C., Girolami, I., Marletta, S., Pantanowitz, L., Antonini, P., Brunelli, M., et al. (2022). Technical and diagnostic issues in whole slide imaging published validation studies. *Front. Oncol.* 12, 918580. doi:10.3389/fonc.2022.918580

Rolfo, C., Mack, P., Scagliotti, G. V., Aggarwal, C., Arcila, M. E., Barlesi, F., et al. (2021). Liquid biopsy for advanced nsclc: A consensus statement from the international association for the study of lung cancer. *J. Thorac. Oncol.* 16 (10), 1647–1662. doi:10.1016/j.jtho.2021.06.017

Saifuddin, S. R., Devlies, W., Santaolalla, A., Cahill, F., George, G., Enting, D., et al. (2017). King's health partners' prostate cancer biobank (KHP PCaBB). *BMC Cancer* 17 (1), 784. doi:10.1186/s12885-017-3773-8

Salvaterra, E., and Corfield, J. (2017). *Advances in biobanking practice through public and private collaborations*. Sharjah United Arab Emirates: Bentham Science Publishers.

Sanchini, V., Bonizzi, G., Disalvatore, D., Monturano, M., Pece, S., Viale, G., et al. (2016). A trust-based pact in research biobanks. From theory to practice. *Bioethics* 30 (4), 260–271. doi:10.1111/bioe.12184

Schmanski, A., Roberts, E., Coors, M., Wicks, S. J., Arbet, J., Weber, R., et al. (2021). Research participant understanding and engagement in an institutional, self-consent biobank model. *J. Genet. Couns.* 30 (1), 257–267. doi:10.1002/jgc4.1316

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

Szustakowski, J. D., Balasubramanian, S., Kvikstad, E., Khalid, S., Bronson, P. G., Sasson, A., et al. (2021). Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. *Nat. Genet.* 53 (7), 942–948. doi:10.1038/s41588-021-00885-0

Tukacs, E., Korotij, A., Maros-Szabo, Z., Molnar, A. M., Hajdu, A., and Torok, Z. (2012). Model requirements for biobank software systems. *Bioinformation* 8 (6), 290–292. doi:10.6026/97320630008290

Vaught, J., Kelly, A., and Hewitt, R. (2009). A review of international biobanks and networks: Success factors and key benchmarks. *Biopreserv. Biobank.* 7 (3), 143–150. doi:10.1089/bio.2010.0003

Vaught, J., Rogers, J., Myers, K., Lim, M. D., Lockhart, N., Moore, H., et al. (2011). An NCI perspective on creating sustainable biospecimen resources. *J. Natl. Cancer Inst. Monogr.* 2011 (42), 1–7. doi:10.1093/jncimonographs/lgr006

Voegele, C., Bouchereau, B., Robinot, N., McKay, J., Damiecki, P., and Alteyrac, L. (2013). A universal open-source electronic laboratory notebook. *Bioinformatics* 29 (13), 1710–1712. doi:10.1093/bioinformatics/btt253

Willers, C., Lynch, T., Chand, V., Islam, M., Lassere, M., and March, L. (2021). A versatile, secure, and sustainable all-in-one biobank-registry data solution: The A3BC REDCap model. *Biopreserv. Biobank.* 20 (3), 244–259. doi:10.1089/bio.2021.0098

Williams, R. R., Gupta, D., and Yong, W. H. (2019). Orientation and training of new biobank personnel. *Methods Mol. Biol.* 1897, 51–63. doi:10.1007/978-1-4939-8935-5_6

Yan, H. H. N., Siu, H. C., Law, S., Ho, S. L., Yue, S. S. K., Tsui, W. Y., et al. (2018). A comprehensive human gastric cancer organoid biobank captures tumor subtype heterogeneity and enables therapeutic screening. *Cell Stem Cell* 23 (6), 882–897. doi:10.1016/j.stem.2018.09.016

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 06 October 2022
DOI 10.3389/fmolb.2022.994440



OPEN ACCESS

EDITED BY
Umberto Malapelle,
University of Naples Federico II, Italy

REVIEWED BY
Sachchida Nand Rai,
University of Allahabad, India
Nagarjun Vijay,
Indian Institute of Science Education and Research, India

*CORRESPONDENCE
Xiaoqun Ye,
511201663@qq.com

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 14 July 2022
ACCEPTED 20 September 2022
PUBLISHED 06 October 2022

CITATION
Wu J, Lu F, Yu B, Wang W and Ye X (2022), The oncogenic role of SNRPB in human tumors: A pan-cancer analysis. *Front. Mol. Biosci.* 9:994440. doi: 10.3389/fmolb.2022.994440



# The oncogenic role of SNRPB in human tumors: A pan-cancer analysis

Juan Wu[1], Feng Lu[2], Bin Yu[1], Wenjun Wang[1] and Xiaoqun Ye[1]*

[1]Department of Respiratory Diseases, The Second Affiliated Hospital of Nanchang University, Nanchang, China, [2]Department of Cardiothoracic Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, China

**Purpose:** The purpose of this study was to explore the oncogenic role of small nuclear ribonucleoprotein polypeptides B and B1 (SNRPB) in human tumors.

**Materials and methods:** Study cases were acquired from The Cancer Genome Atlas database, the Gene Expression Omnibus database, The Human Protein Atlas, and the Clinical Proteomic Tumor Analysis Consortium. We then used the R package and several online tools to analyze and visualize the role of SNRPB across tumors.

**Results:** We found that the expression of SNRPB was significantly increased in 28 of 33 tumors, and higher expression was observed in late pathological and TNM stages. Significantly decreased levels of SNRPB promoter methylation were observed in 12 tumors. SNRPB was found to be a risk factor for decreased overall survival in 10 tumors ($p < 0.05$), a risk factor for decreased disease-specific survival in 8 tumors ($p < 0.05$), and a risk factor for decreased progression-free interval in 7 tumors ($p < 0.05$). The PPI network of SNRPB and the top 100 coexpressed genes revealed that CDK1, CDC6, AURKB, CCNB1, CCNA2, and CDC45 were the most closely interacting genes across tumors. The GO and KEGG enrichment analyses revealed that SNRPB and the above genes were mainly enriched with respect to functions in cell cycle-related genetic material replication, assembly, and distribution. SNRPB was significantly associated with immune cell infiltration and the expression of immunomodulation-related genes in several but not all tumors.

**Conclusion and limitations:** The expression of SNRPB was significantly elevated in almost all tumors, and the decreased promoter methylation level may contribute to the elevated expression of SNRPB. SNRPB may facilitate the progression of pathological and TNM stages and is a risk factor for unfavorable prognosis across tumors. However, our research was based on data obtained from public databases, without further validation of our findings at the cellular and animal levels. Therefore, further studies are needed to clarify the oncogenic mechanism of SNRPB and its potential as a therapeutic target.

KEYWORDS

**SNRPB, small nuclear ribonucleoprotein polypeptides B and B1, pan-cancer, human tumors, oncogenic role, oncogene**

# Introduction

As one of the deadliest diseases in the world, cancer is the leading cause of premature death (Bray et al., 2021). Along with the transition of cells from normal to neoplastic states, they acquire various functions necessary for their malignancy, including persistent proliferation signals, genome instability and mutation, immune escape, invasion and metastasis, unlocked phenotypic plasticity, nonmutational epigenetic reprogramming, and the activation of oncogenes (Hanahan and Weinberg, 2011; Hanahan, 2022). The identification of oncogenes and their roles across the spectrum of human cancers is of great importance to better understand the complex pathological mechanism of cancers.

The alternative splicing (AS) of most human genes gives rise to transcript "isoforms" (Pan et al., 2008; Wang et al., 2008), increases the diversity of mRNA expression, and results in functional diversity of the encoded proteins based on enzymatic activity, subcellular localization, and protein–ligand, protein–protein, and protein–DNA physical interactions (Kelemen et al., 2013), which have profound effects on the proliferation and survival of cells (Kelemen et al., 2013). Spliceosomes are required for AS (Scofield and Lynch, 2008; Liu et al., 2019).

Small nuclear ribonucleoprotein polypeptide B and B1 (SNRPB) is a core component of the spliceosome and thus plays a critical role in pre-mRNA splicing (Gray et al., 1999). Dysregulation of SNRPB influences the splicing of pre-mRNA and generates unexpected mRNA variants. The protein translated by these new mRNA variants may play an important role in tumorigenesis (Correa et al., 2016; Liu et al., 2019; Peng et al., 2020; Zhan et al., 2020). The role of SNRPB in promoting tumorigenesis and progression has been observed in tumors such as non-small cell lung cancer (NSCLC) (Liu et al., 2019; Liu et al., 2021), hepatocellular carcinoma (LIHC) (Peng et al., 2020; Zhan et al., 2020; Li et al., 2022), glioblastoma (GBM) (Correa et al., 2016), cervical cancer (CESC) (Zhu et al., 2020), and thyroid carcinoma (THCA) (Deng et al., 2022). However, the role of SNRPB in other tumors remains unclear.

To comprehensively explore the oncogenic role of SNRPB in human tumors, we acquired study cases from The Cancer Genome Atlas database (TCGA), the Gene Expression Omnibus database (GEO) (Barrett et al., 2013), The Human Protein Atlas (HPA) (Uhlen et al., 2017), and the Clinical Proteomic Tumor Analysis Consortium (CPTAC) dataset from the University of ALabama at Birmingham CANcer data analysis portal (Chandrashekar et al., 2017) (UALCAN). We then performed an in-depth pan-cancer analysis of SNRPB on the mRNA and protein expression, prognostic value, genetic variation, promoter methylation, the possible oncogenic mechanism, and the immunological role, which would provide us with new ideas and a theoretical basis for cancer diagnosis and treatment.

# Materials and methods

## Data source and preparation

Case information about mRNA expression (normalized as transcripts per million reads, TPM) and clinical features was obtained from TCGA and Genotype-Tissue Expression project (Consortium, 2013) (GTEx) and downloaded from the University of California Santa Cruz Xena (Goldman et al., 2020) (UCSC Xena, https://xena.ucsc.edu/) platform. Microarray data were downloaded from GEO database (Barrett et al., 2013) (http://www.ncbi.nlm.nih.gov/geo/). The raw data were downloaded as MINiML files and were normalized by log2 transformation. We used the normalize quantiles function of the preprocessCore package in R to normalize the microarray data. Probes were converted to gene symbols according to the annotation of the normalized data in the platform. RemoveBatchEffect function of the limma package in R was used to remove batch effect of samples in different batches. Boxplot was used to assess the result of the data preprocessing. All data in Supplementary Figure S1 were comparable after normalization (the boxplots of data preprocessing results were not shown). Protein expression data were obtained from UALCAN data analysis portal (Chandrashekar et al., 2017) (http://ualcan.path.uab.edu/analysis-prot.html, CPTAC dataset). Immunohistochemistry (IHC)-based protein expression patterns were acquired from HPA (Uhlen et al., 2017) (Human Protein Atlas proteinatlas.org). Genetic alteration data were obtained from cBioPortal (Cerami et al., 2012) (http://www.cbioportal.org). Promoter methylation data were obtained from UALCAN data analysis portal (TCGA dataset). The top 100 coexpressed genes were obtained from Gene Expression Profiling Interactive Analysis (Tang et al., 2019) (GEPIA, version 2, http://gepia2.cancer-pku.cn/#index). Protein–protein interaction (PPI) network were obtained from STRING (Szklarczyk et al., 2021) (https://string-db.org/). Data regarding the relationship between SNRPB and immune cell infiltration as well as immunomodulation-related gene expression were obtained from the Tumor Immune Estimation Resource (Li et al., 2020) (TIMER, version 2, timer.cistrome.org).

## mRNA and protein expression of SNRPB across tumors

We first used the Wilcoxon test to compare the expression level of SNRPB in tumors and the corresponding paracancerous tissues in TIMER2. Since there were no corresponding paracancerous tissue data for some tumors in TCGA, we used data obtained from TCGA and GTEx to compare the expression level of SNRPB in tumors and corresponding normal tissues. The "ggplot2" R package (version 3.3.3) was used to analyze and

visualize the results, and an unpaired samples *t* test was used to compare the expression level of SNRPB between the normal and tumor groups. Normalized SNRPB expression data from GEO were compared using the Wilcoxon test.

Protein expression data of primary tumors and the corresponding normal tissues in the CPTAC dataset from UALCAN data analysis portal were compared.

## IHC-based protein expression of SNRPB across tumors

To verify the protein expression of SNRPB at the histological level, IHC-based protein expression patterns in bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), CESC, colon adenocarcinoma (COAD), LIHC, lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), ovarian serous cystadenocarcinoma (OV), pancreatic adenocarcinoma (PAAD), prostate adenocarcinoma (PRAD), skin cutaneous melanoma (SKCM), stomach adenocarcinoma (STAD), testicular germ cell tumors (TGCT), uterine corpus endometrial carcinoma (UCEC), and the corresponding normal tissues were obtained from HPA.

## mRNA and protein expression of SNRPB in different pathological stages across tumors

GEPIA2 was used to assess the correlation between SNRPB mRNA expression and pathological stages; CPTAC samples from UALCAN data analysis portal were used to assess the correlation between SNRPB protein expression and pathological stages.

## SNRPB expression and TNM stages

The clinical datasets obtained from TCGA were used to explore the effect of SNRPB expression on TNM stages. With 50% as the cutoff value, samples were divided into low and high groups.

## Survival and prognostic analysis

The clinical datasets obtained from TCGA were used to perform the survival and prognostic analysis. With 50% as the cutoff value, samples were divided into low and high groups. Overall survival (OS), disease-specific survival (DSS), and progression-free interval (PFI) were used for the evaluation of survival and prognostic outcomes. We performed Kaplan–Meier (KM) analysis with Cox regression using the "survminer" and "survival" packages in R.

## Genetic alteration

The "TCGA PanCancer Atlas Studies" in cBioPortal were used to analyze and visualize the genetic alteration of SNRPB and the impact of genetic alteration on OS, disease-free survival (DFS), DSS, and progression-free survival (PFS) across tumors.

## Promoter methylation level of SNRPB across tumors

TCGA samples from UALCAN data analysis portal were used to compare the promoter methylation level of SNRPB between primary tumors and the corresponding normal tissues.

## Coexpressed genes and PPI network

The top 100 coexpressed genes of SNRPB were obtained from GEPIA2 based on the datasets of all TCGA tumors. The "Correlation Analysis" module in GEPIA2 was used to analyze and visualize the correlation between SNRPB and the top 6 coexpressed genes, and Pearson correlation was used in the above analysis. The correlation heatmap of SNRPB and the top 10 coexpressed genes was analyzed and plotted using TIMER2. To explore the proteins closely interacting with SNRPB across tumors, we used STRING (https://cn.string-db.org; main parameters: network type: full STRING network, meaning of network edges: evidence, active interaction source: Textmining, Experiments, Databases, Co-expression, Neighborhood, Gene Fusion, and Co-occurrence, minimum required interaction score: Medium confidence [0.400], max number of interactors to show: 1st shell [no more than 50 interactors], 2nd shell [none/query proteins only]) to analyze the relationship between the proteins expressed by SNRPB and the top 100 coexpressed genes and visualized them using Cytoscape (Shannon et al., 2003).

## Functional annotation

SNRPB and the top 100 coexpressed genes obtained from GEPIA2 (101 genes in total) were used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses in R using the "ggplot2", "clusterProfiler", and "GOplot" packages.

## Immune infiltration analysis

We first used the TIMER algorithm to assess the relationship between SNRPB and the infiltration of $CD8^+$ T cells, $CD4^+$ T cells, B cells, macrophages, neutrophils, and dendritic cells. Then, the xCell algorithm was used to evaluate the relationship

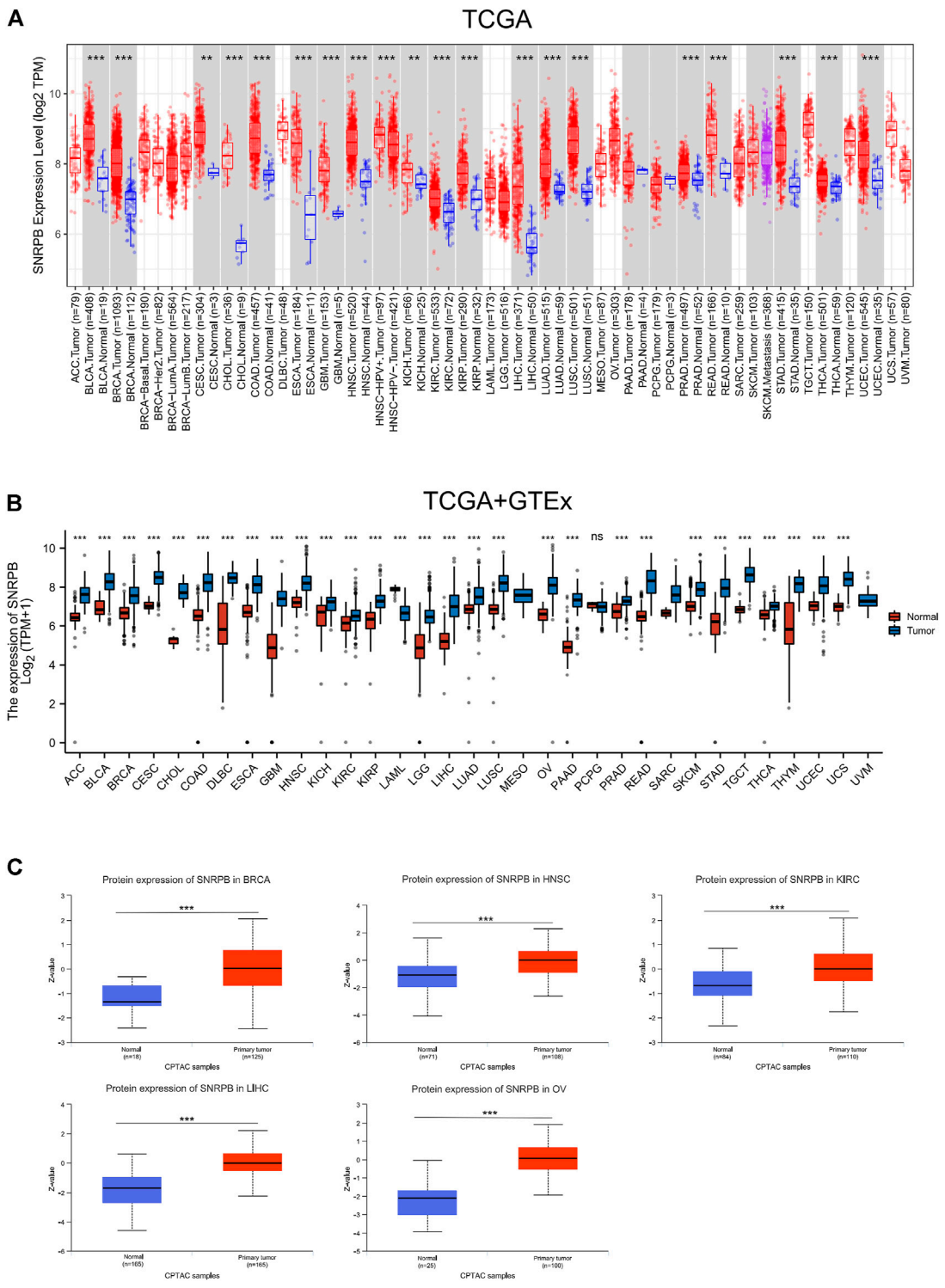



FIGURE 1
mRNA and protein expression levels of SNRPB across tumors. **(A)** mRNA expression level of SNRPB in tumors and adjacent paracancerous tissues in TCGA from TIMER2. **(B)** mRNA expression level of SNRPB in tumors and the corresponding normal tissues in TCGA and GTEx. **(C)**. Protein expression level of SNRPB in primary tumors and the corresponding normal tissues in the CTPAC dataset from UALCAN. ns, no significant difference; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

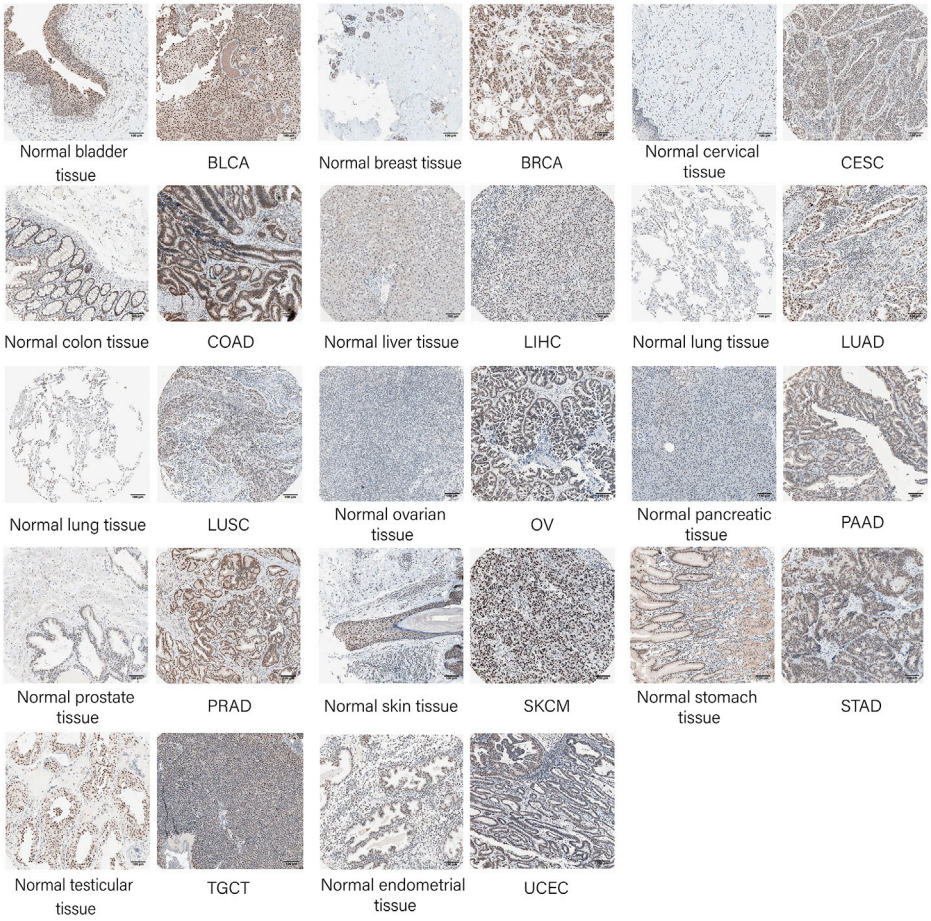



FIGURE 2
Validation of SNRPB protein expression in tumors and the corresponding normal tissues in HPA.

between SNRPB and the infiltration of immune cell subtypes in TIMER2. The correlation between SNRPB and the expression of immunomodulation-related genes across tumors was also analyzed in TIMER2. Finally, we visualized them in R using the "ggplot2" package.

# Results

## SNRPB expression was significantly elevated in almost all tumors

A total of 21 tumors in the TCGA database had data on corresponding paracancerous tissues, and the expression of SNRPB was significantly elevated in 19 of them ($p < 0.05$, Figure 1A). Since there were no corresponding paracancerous tissue data for some tumors in TCGA, we used data obtained from TCGA and GTEx to compare the expression level of SNRPB in cancer and corresponding normal tissues, and the expression of SNRPB was found to be significantly increased in 28 of 33 tumors ($p < 0.05$, Figure 1B). To further clarify the expression of SNRPB across tumors, we obtained the expression levels of SNRPB in 21 tumors from the GEO database, and the expression of SNRPB was found to be significantly increased in 17 of 21 tumors ($p < 0.05$, Supplementary Figure S1).

To verify the protein expression level of SNRPB across tumors, we first acquired protein expression data of primary tumors and the corresponding normal tissues in the CPTAC dataset from UALCAN data analysis portal. Only limited protein expression data of SNRPB were available, including BRCA, head and neck squamous cell carcinoma (HNSC), kidney renal clear cell carcinoma (KIRC), LIHC, and OV, and the protein expression level of SNRPB was significantly increased in all five tumors (Figure 1C, $p < 0.05$). Then, IHC-based protein expression patterns in BLCA, BRCA, CESC, COAD, LIHC, LUAD, LUSC, OV, PAAD, PRAD, SKCM, STAD, TGCT, UCEC, and the corresponding normal tissues were obtained

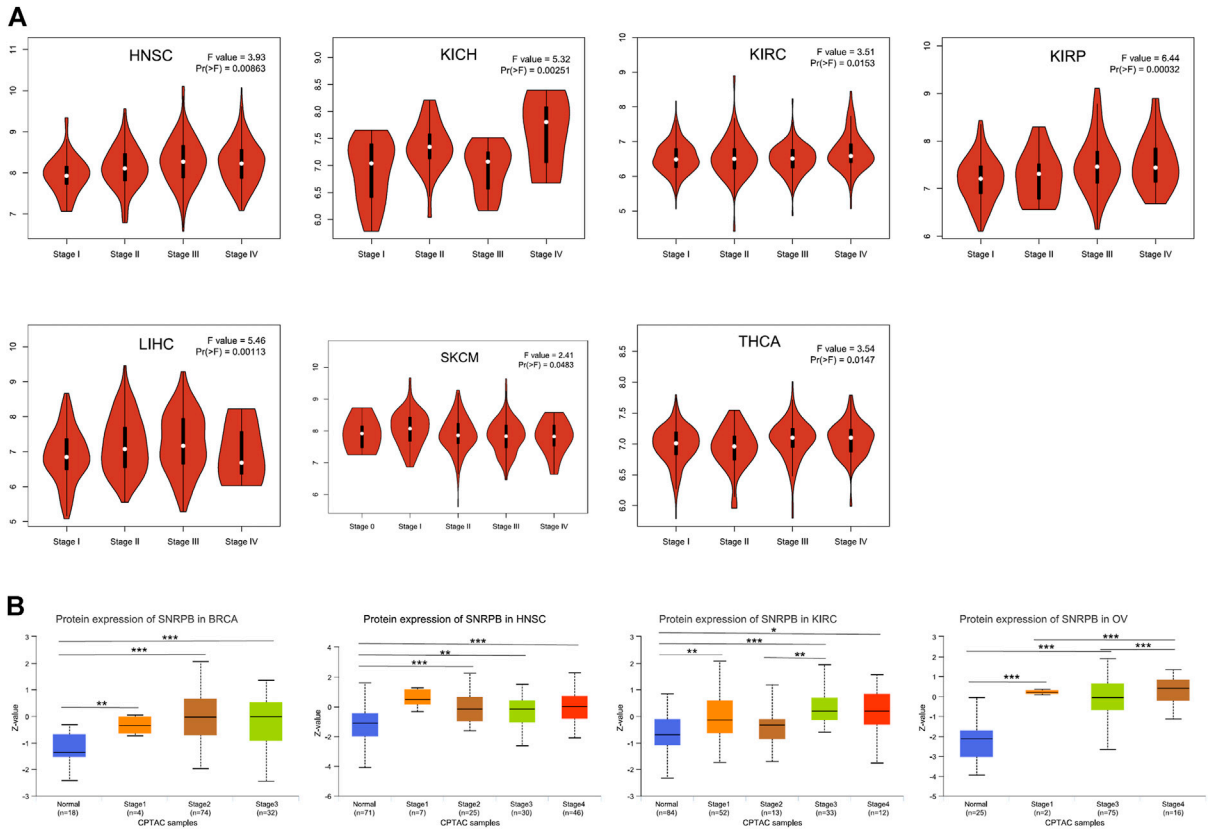



FIGURE 3
mRNA and protein expression levels of SNRPB in different pathological stages across tumors. **(A)** mRNA expression level of SNRPB in different pathological stages. **(B)**. Protein expression level of SNRPB in different pathological stages.*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

from HPA. Compared with normal tissues, the IHC-based SNRPB expression level was significantly increased in the above 14 tumors (Figure 2).

## The expression of SNRPB was higher in late pathological stages

To identify whether SNRPB is differentially expressed among pathological stages, we first analyzed the correlations between the mRNA expression of SNRPB and the pathological stages across tumors using GEPIA2. The results showed that SNRPB was differentially expressed among pathological stages of HNSC, kidney chromophobe (KICH), KIRC, kidney renal papillary cell carcinoma (KIRP), LIHC, SKCM, and THCA, and the expression levels were generally higher in late pathological stages ($p < 0.05$, Figure 3A). Then, SNRPB protein expression data of the CPTAC dataset from UALCAN data analysis portal were obtained, and significantly differential expression of SNRPB was observed in BRCA, HNSC, KIRC, and OV. Additionally, the protein expression levels were generally higher in late pathological stages ($p < 0.05$, Figure 3B). Thus, SNRPB may promote the progression of pathological stages across tumors.

## SNRPB may facilitate the progression of TNM stages

The differential expression pattern of SNRPB in pathological stages suggests that SNRPB may also promote the progression of TNM stages. To explore the relationship between SNRPB and TNM stages, we compared the differences in T (T1, T2, T3, T4), N (N0, N1, N2, N3), and M (M0, M1) stages between the low and high groups. As shown in Table 1, compared with the low group, the proportion of T1 (T2 in PRAD) stage was lower, whereas those of T2 (in BRCA, KIRP, LIHC), T3 (in BRCA, KIRP, LIHC, and PRAD) and T4 (in LIHC and PRAD) stages were higher in the high group. The same phenomenon was also observed for the N and M stages. As shown in Table 2, compared with the low group, the proportion of the N0 stage was lower, whereas N1, N2, and N3 stages were higher in the high group in HNSC, KIRP (only N1 and N2 stages), and LUAD. As shown in Table 3,

TABLE 1 The relationship between SNRPB expression and tumor T stages.

| Tumor | SNRPB expression level | T1 (n [%]) | T2 (n [%]) | T3 (n [%]) | T4 (n [%]) | *p* Value | Method |
|---|---|---|---|---|---|---|---|
| BRCA | | | | | | 0.005 | Chi-square test |
| | Low (*n* = 541) | 160 (14.8%) | 296 (27.4%) | 62 (5.7%) | 22 (2%) | | |
| | High (*n* = 542) | 117 (10.8%) | 333 (30.8%) | 77 (7.1%) | 13 (1.2%) | | |
| KIRP | | | | | | 0.020 | Fisher's exact test |
| | Low (*n* = 144) | 107 (37.3%) | 13 (4.5%) | 21 (7.3%) | 1 (0.3%) | | |
| | High (*n* = 145) | 86 (30%) | 20 (7%) | 38 (13.2%) | 1 (0.3%) | | |
| LIHC | | | | | | 0.010 | Chi-square test |
| | Low (*n* = 187) | 107 (28.8%) | 40 (10.8%) | 34 (9.2%) | 4 (1.1%) | | |
| | High (*n* = 187) | 76 (20.5%) | 55 (14.8%) | 46 (12.4%) | 9 (2.4%) | | |
| PRAD | | | | | | 0.004 | Chi-square test |
| | Low (*n* = 249) | - | 111 (22.6%) | 131 (26.6%) | 3 (0.6%) | | |
| | High (*n* = 250) | - | 78 (15.9%) | 161 (32.7%) | 8 (1.6%) | | |

TABLE 2 The relationship between SNRPB expression and tumor N stages.

| Tumor | SNRPB expression level | N0 (n [%]) | N1 (n [%]) | N2 (n [%]) | N3 (n [%]) | *p* Value | Method |
|---|---|---|---|---|---|---|---|
| HNSC | | | | | | 0.002 | Fisher's exact test |
| | Low (*n* = 251) | 138 (28.7%) | 37 (7.7%) | 63 (13.1%) | 1 (0.2%) | | |
| | High (*n* = 251) | 101 (21%) | 43 (9%) | 91 (19%) | 6 (1.2%) | | |
| KIRP | | | | | | 0.017 | Fisher's exact test |
| | Low (*n* = 144) | 26 (33.8%) | 5 (6.5%) | 1 (1.3%) | - | | |
| | High (*n* = 145) | 23 (29.9%) | 19 (24.7%) | 3 (3.9%) | - | | |
| LUAD | | | | | | 0.017 | Fisher's exact test |
| | Low (*n* = 267) | 185 (35.6%) | 36 (6.9%) | 33 (6.4%) | 0 (0%) | | |
| | High (*n* = 268) | 163 (31.4%) | 59 (11.4%) | 41 (7.9%) | 2 (0.4%) | | |

TABLE 3 The relationship between SNRPB expression and tumor M stages.

| Tumor | SNRPB expression level | M0 (n [%]) | M1 (n [%]) | *p* Value | Method |
|---|---|---|---|---|---|
| BLCA | | | | 0.011 | Fisher's exact test |
| | Low (*n* = 207) | 118 (55.4%) | 2 (0.9%) | | |
| | High (*n* = 207) | 84 (39.4%) | 9 (4.2%) | | |
| KIRP | | | | 0.012 | Fisher's exact test |
| | Low (*n* = 144) | 54 (51.9%) | 1 (1%) | | |
| | High (*n* = 145) | 41 (39.4%) | 8 (7.7%) | | |
| LUAD | | | | 0.013 | Chi-square test |
| | Low (*n* = 267) | 172 (44.6%) | 5 (1.3%) | | |
| | High (*n* = 268) | 189 (49%) | 20 (5.2%) | | |
| READ | | | | 0.013 | Chi-square test |
| | Low (*n* = 83) | 66 (44.3%) | 5 (3.4%) | | |
| | High (*n* = 83) | 60 (40.3%) | 18 (12.1%) | | |

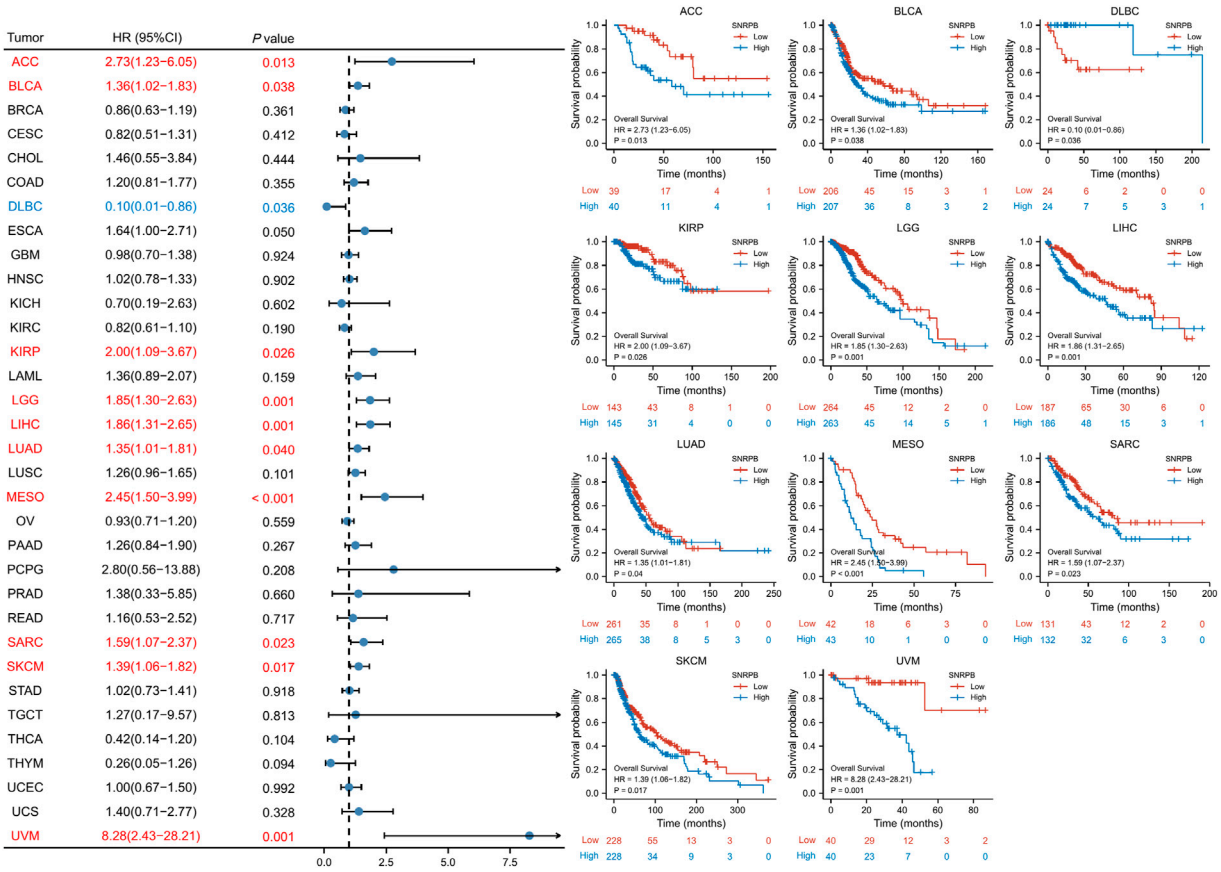



FIGURE 4
The relationship between SNRPB expression and OS across tumors.

compared with the low group, the proportion of M0 (except for LUAD) stage was lower, whereas the M1 stage was higher in the high group in BLCA, KIRP, LUAD, and rectum adenocarcinoma (READ). The above results suggest that the expression of SNRPB may promote local tumor progression, lymph node metastasis, and distant metastasis, thus facilitating the progression of TNM stages.

## SNRPB was a risk factor for unfavorable prognosis across tumors

We used OS, DSS and PFI as outcomes to assess the survival and prognostic value of SNRPB across tumors. To show the results more clearly and concisely, we first generated a forest plot of the prognostic value of SNRPB across tumors and then displayed the KM curve plot of tumors in which SNRPB had a significant impact on prognosis on the right side of the forest plot.

The results showed that high expression of SNRPB was a risk factor for unfavorable prognosis in several tumors. As shown in Figure 4, SNRPB was found to be a risk factor for decreased OS in 10 tumors, including adrenocortical carcinoma (ACC), BLCA, KIRP, brain lower grade glioma (LGG), LIHC, LUAD, mesothelioma (MESO), sarcoma (SARC), SKCM, and uveal melanoma (UVM) ($p < 0.05$). As shown in Figure 5, SNRPB was found to be a risk factor for decreased DSS in 8 tumors, including ACC, BLCA, KIRP, LGG, LIHC, MESO, SKCM, and UVM ($p < 0.05$). As shown in Figure 6, SNRPB was found to be a risk factor for decreased PFI in 7 tumors, including ACC, KIRP, LIHC, MESO, PAAD, pheochromocytoma and paraganglioma (PCPG), and UVM ($p < 0.05$). Therefore, high SNRPB expression could be a risk factor for unfavorable prognosis across tumors.

## Genetic alteration of SNRPB across tumors

Genetic alterations may affect gene expression levels and their prognostic role in tumors (Yang et al., 2019; Zhang et al.,

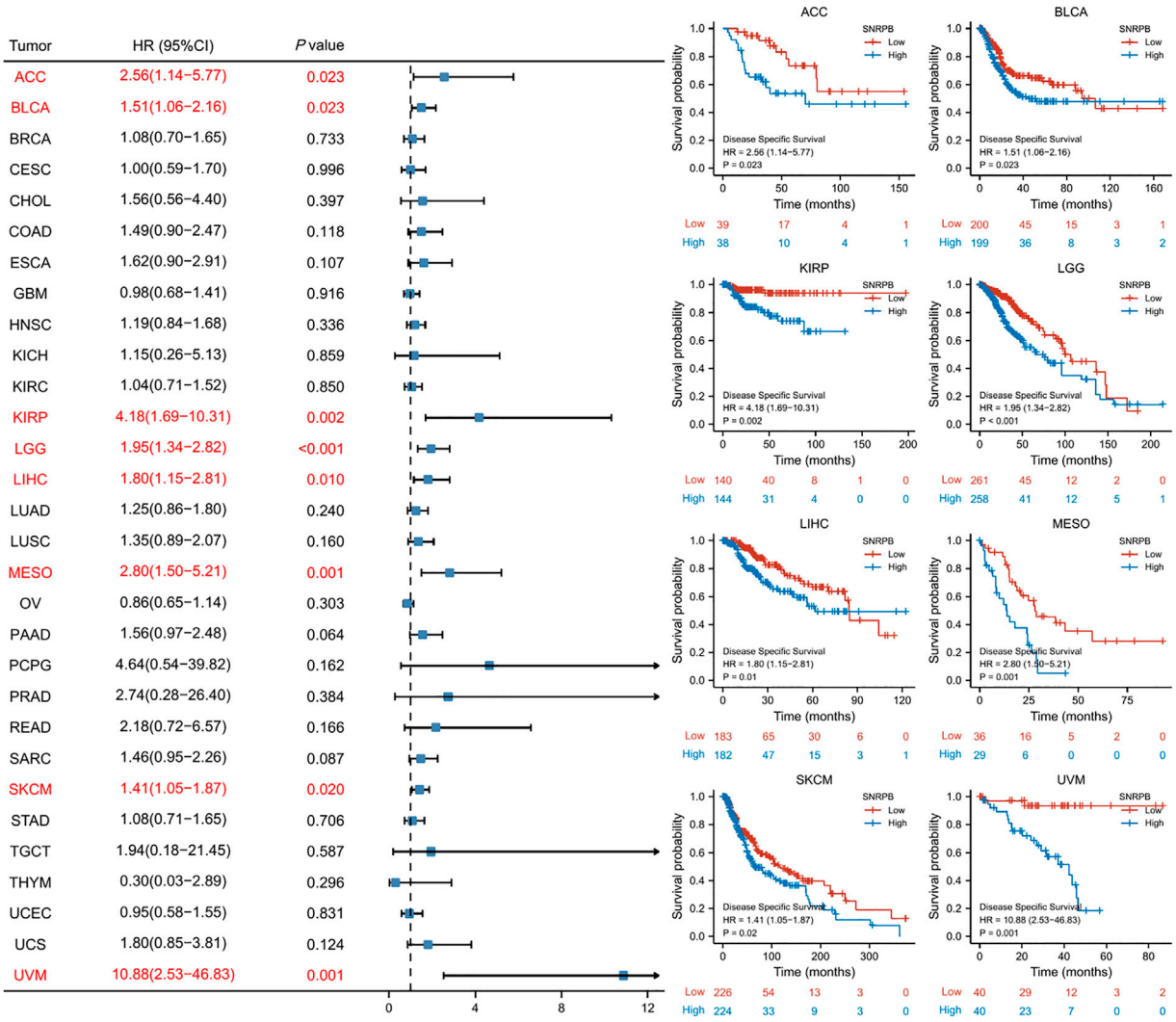



FIGURE 5
The relationship between SNRPB expression and DSS across tumors.

2020). To explore the genetic alteration and its prognostic role across tumors, we obtained SNRPB gene variation data from cBioPortal. The results showed that amplification and mutation are the most frequent alterations across tumors. OV was the tumor with the highest mutation rate, and 4.79% (28/584) of the OV cases harbored genetic variations (Figure 7A). We then investigated the main mutation types and their locations within SNRPB. Missense mutations were the most common mutation type and were randomly distributed within SNRPB, and R236H was the most mutated site, where 3 missense mutations occurred (Figure 7B). Finally, we evaluated the effect of SNRPB genetic alterations on prognosis across tumors (Figure 7C) and found that DFS in the altered group was significantly lower than that in the unaltered group ($p < 0.05$). However, no significant effect of genetic alteration was observed on OS, DSS, or PFS across tumors ($p > 0.05$).

## Decreased promoter methylation levels may contribute to the elevated expression of SNRPB across tumors

Methylation is one of the ways in which nucleobases are chemically modified, and genes can be silenced and reactivated by the methylation and demethylation of cytosines in the promoter region (Traube and Carell, 2017). To explore the mechanism of elevated SNRPB expression across tumors, we obtained data on SNRPB promoter methylation in the TCGA dataset from UALCAN. As shown in Figure 8, the SNRPB

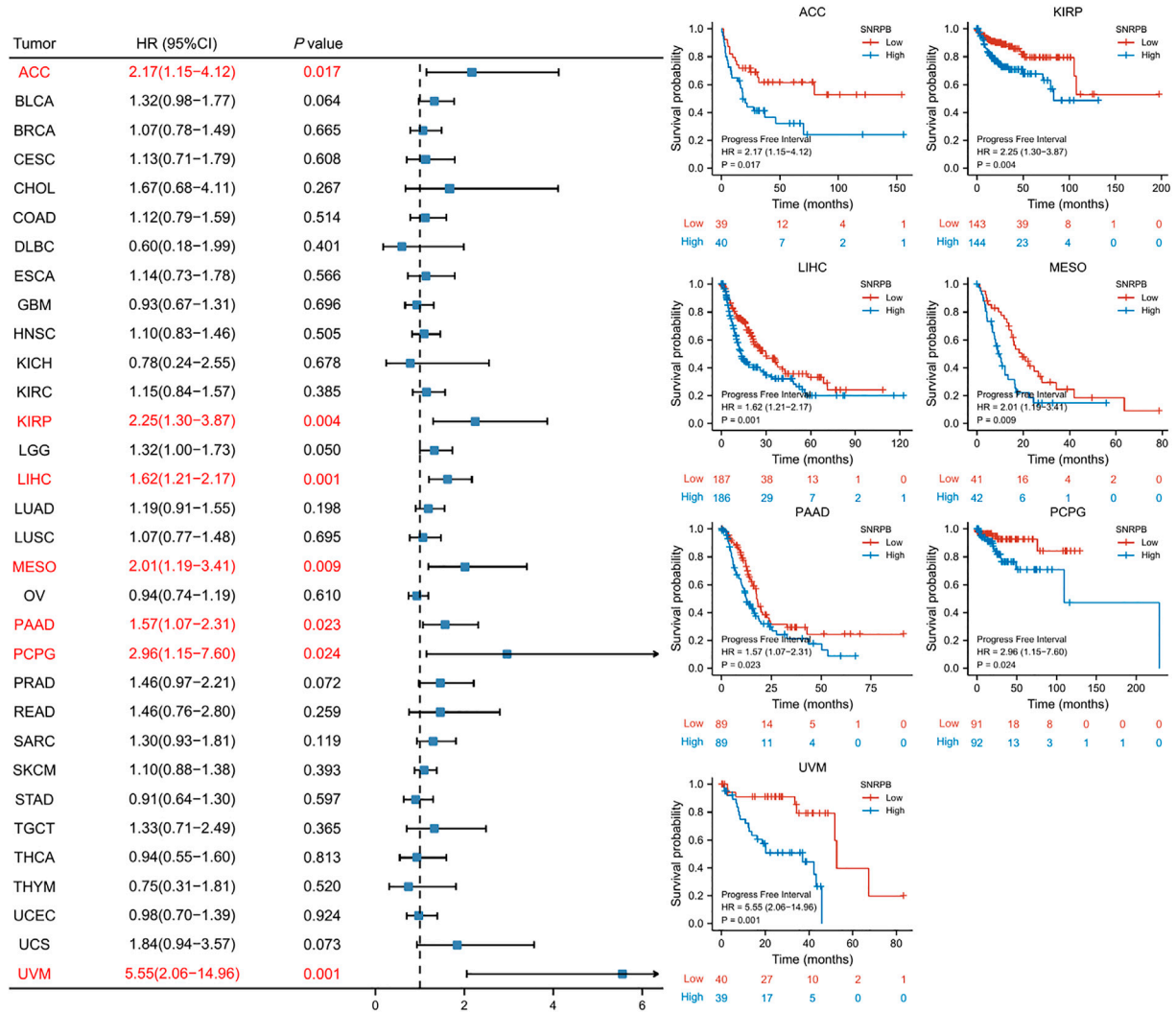



FIGURE 6
The relationship between SNRPB expression and PFI across tumors.

promoter methylation level was significantly altered in 16 tumors ($p < 0.05$), of which 12 tumors (including BLCA, BRCA, CHOL, COAD, ESCA, LIHC, LUSC, PAAD, PCPG, PRAD, READ, and UCEC) showed significantly decreased levels of SNRPB promoter methylation ($p < 0.05$). The above results suggest that the decreased methylation level of the SNRPB promoter may be the reason for its increased expression across tumors.

## Coexpressed genes and PPI network

We obtained the top 100 coexpressed genes from GEPIA2 to explore genes closely associated with SNRPB expression across tumors. The top 10 coexpressed genes were PCNA, NOP56, NXT1, UBE2C, TROAP, BCL2L12, GINS1, AURKB, CDC20, and CCNB2 (Figure 9A). We showed the relationship between SNRPB and the top 6 coexpressed genes across tumors in Figure 9B. Then, we used STRING and Cytoscape to plot the PPI network of SNRPB and the top 100 coexpressed genes. As shown in Figure 9C, CDK1, CDC6, AURKB, CCNB1, CCNA2, and CDC45 were the most closely interacting genes across tumors.

## Functional annotation

We performed GO and KEGG enrichment analyses of SNRPB and the top 100 coexpressed genes (101 genes in total) to explore the roles of these genes across tumors. As shown in Figure 10A, the most enriched biological processes

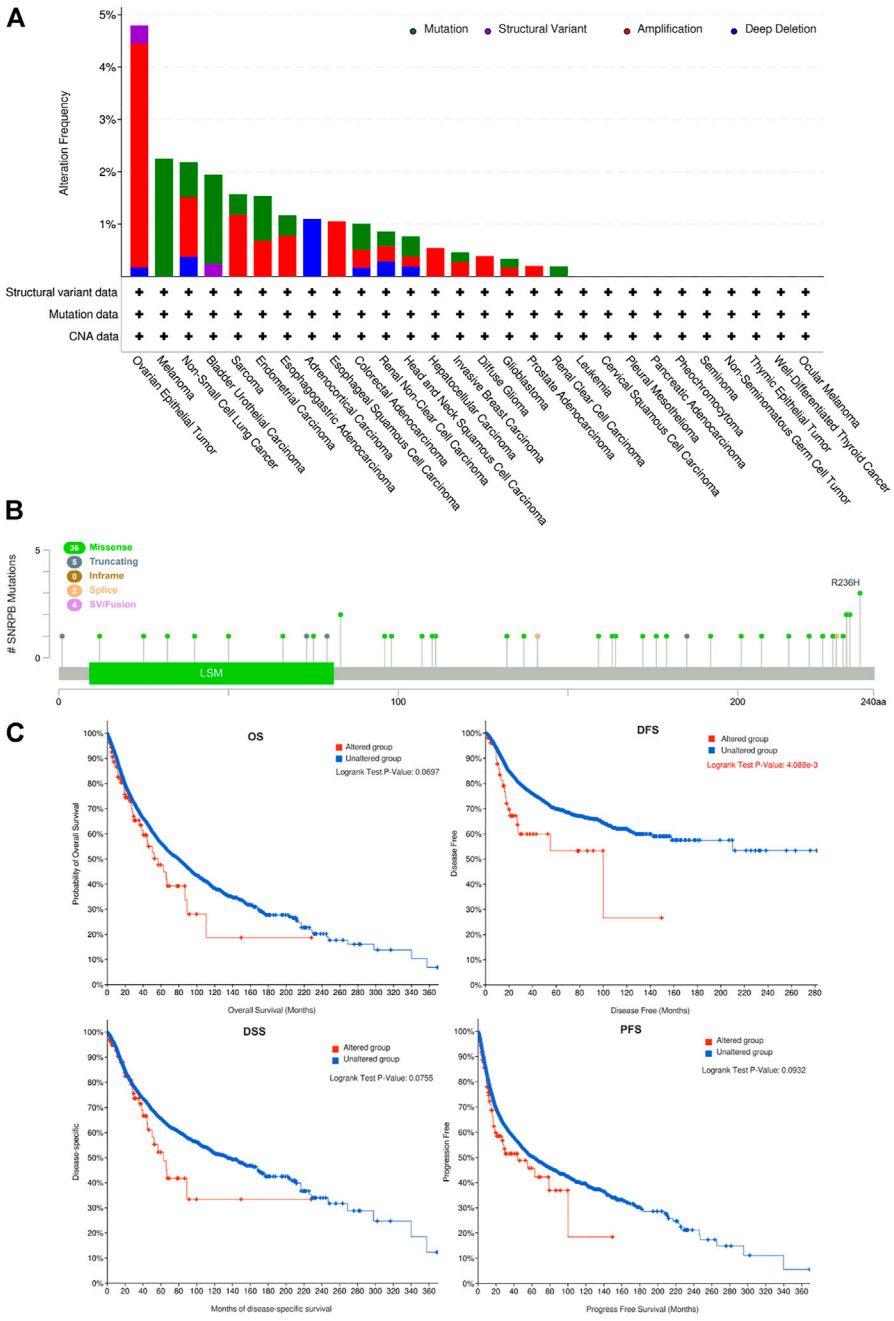



FIGURE 7
Genetic alteration of SNRPB across tumors. **(A)**. Summary of genetic alterations of SNRPB across tumors. **(B)**. Mutated SNRPB sites across tumors. **(C)**. The effect of genetic alteration on OS, DFS, DSS, and PFS across tumors.

(BP) were related to organelle fission, nuclear division, chromosome segregation, microtubule cytoskeleton organization involved in mitosis, and positive regulation of the cell cycle; the most enriched cellular components (CC) were related to the chromosomal region, centromeric region, kinetochore, spindle, cyclin-dependent protein kinase holoenzyme complex, and serine/threonine protein kinase complex; the most enriched molecular functions (MF) were related to catalytic activity, DNA replication origin binding, Ran GTPase binding, and DNA polymerase binding; and the most enriched KEGG pathways were related to cell cycle, oocyte maturation, spliceosome, DNA replication, human

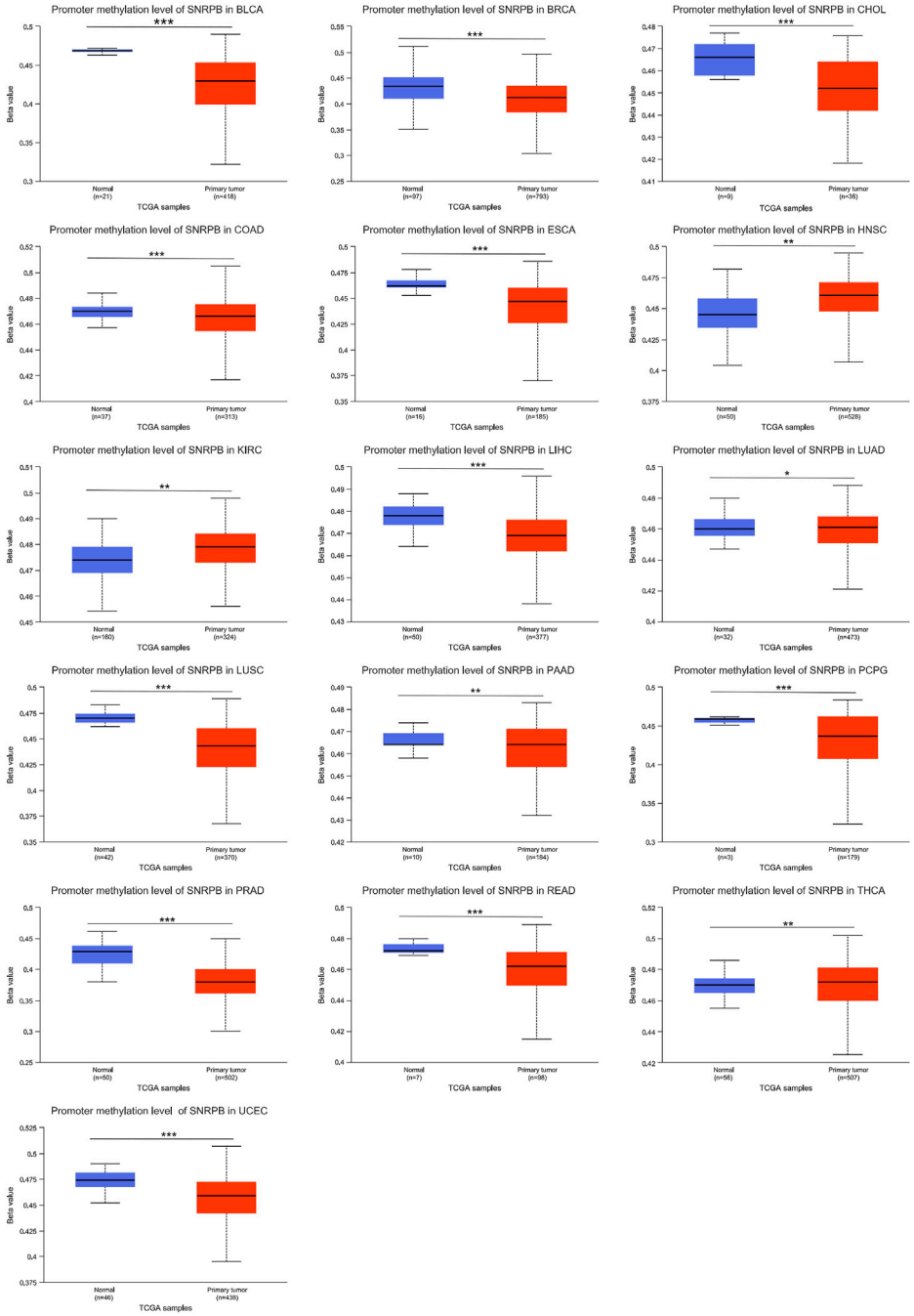



FIGURE 8
The promoter methylation level of SNRPB across tumors. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

T-cell leukemia virus 1 infection, p53 signaling pathway, cellular senescence, mRNA surveillance pathway, and ribosome biogenesis in eukaryotes. The network plot showed the crosstalk between enriched BP, CC, MF, and KEGG functions and genes (Figure 10B). The above research results suggest that related genes are mainly enriched in the processes of cell cycle-related genetic material replication, assembly, and distribution, which indicates that the oncogenic role of SNRPB across tumors may be primarily achieved by affecting the cell cycle.

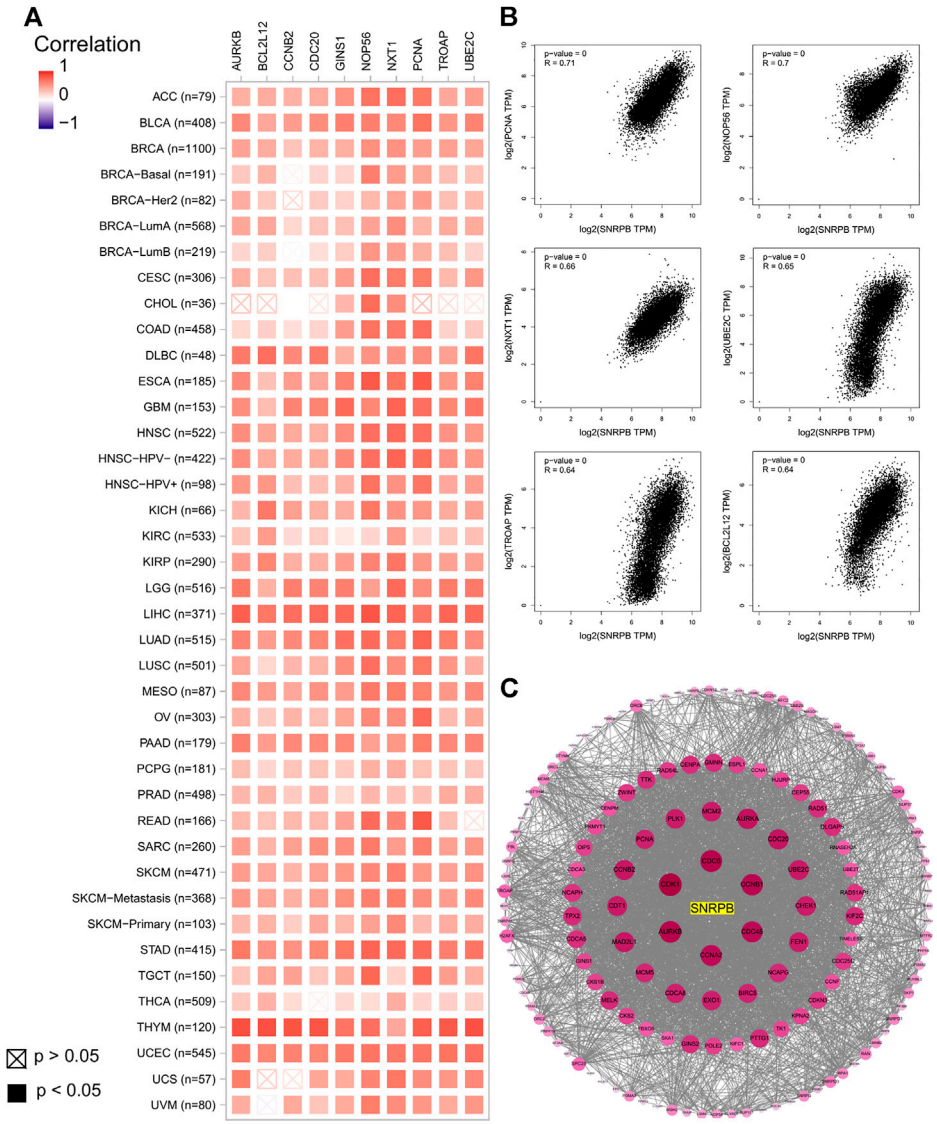



FIGURE 9
Coexpressed genes of SNRPB across tumors. **(A)**. Top 10 coexpressed genes of SNRPB across tumors. **(B)**. Representative figures of the relationship between SNRPB and the top 6 coexpressed genes across tumors. **(C)**. PPI network of SNRPB and the top 100 coexpressed genes across tumors.

## Relationship between SNRPB and immune cell infiltration across tumors

To further clarify the role of SNRPB across tumors, we assessed the relationship between SNRPB and tumor-infiltrating immune cells using TIMER2 (Figure 11). As shown in Figure 11A (TIMER algorithm, purity adjustment), SNRPB was significantly associated with immune infiltration across tumors, including $CD8^+$ T cells, $CD4^+$ T cells, and dendritic cells in 14 tumors, B cells in 7 tumors, macrophages in 12 tumors, and neutrophils in 8 tumors. We then used the xCell algorithm (purity adjustment) to evaluate the relationship between SNRPB and the infiltration of immune cell subtypes (Figure 11B). The results showed that SNRPB was significantly positively correlated with the infiltration of CD4+Th1 and CD4+Th2 cells across tumors.

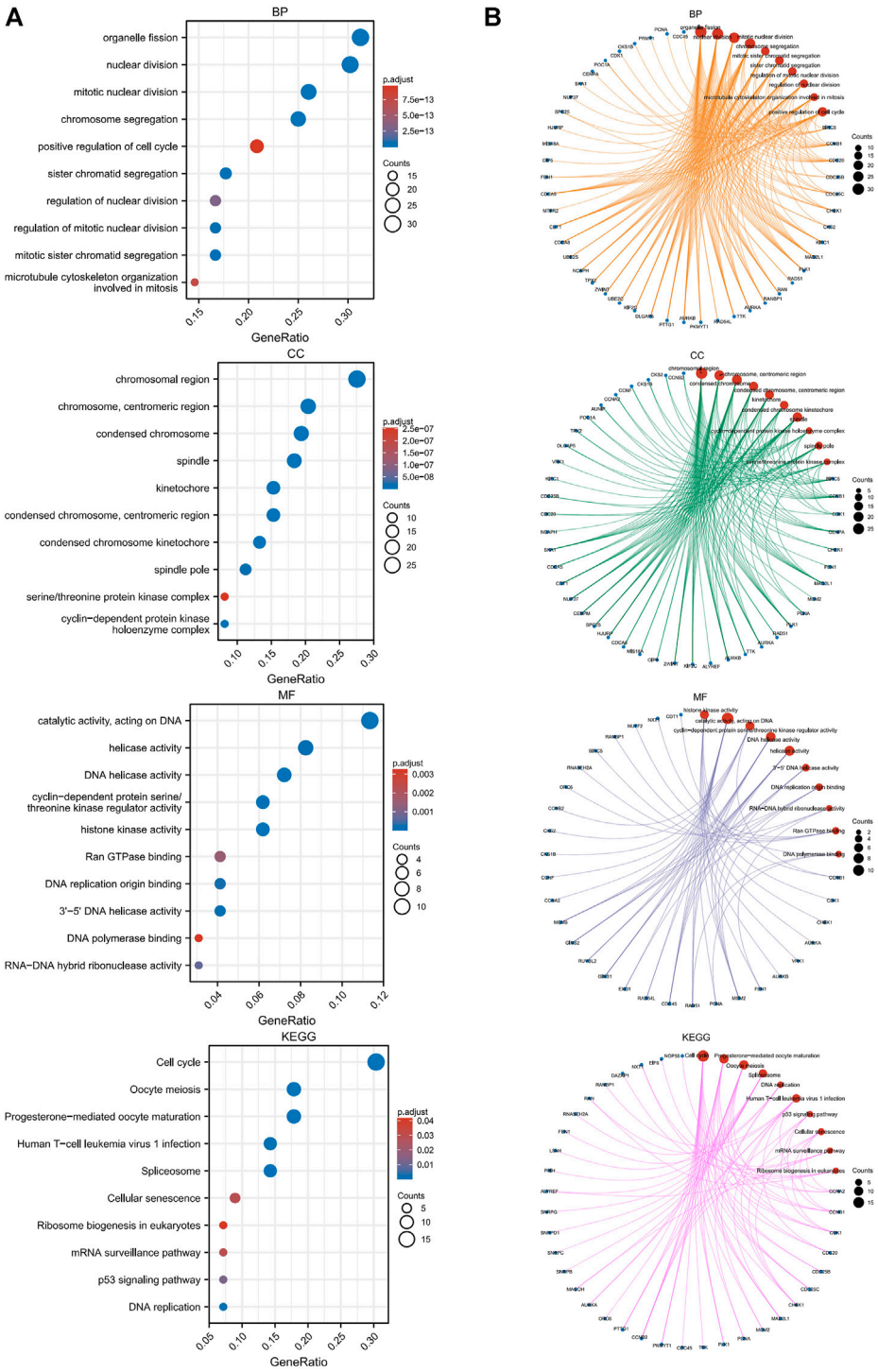



FIGURE 10
GO and KEGG enrichment analyses of SNRPB and the top 100 coexpressed genes across tumors. **(A)**. GO [in biological process (BP), cellular component (CC), molecular function (MF)] and KEGG enrichment analysis. **(B)**. Crosstalk between enriched BP, CC, MF, and KEGG functions and genes.

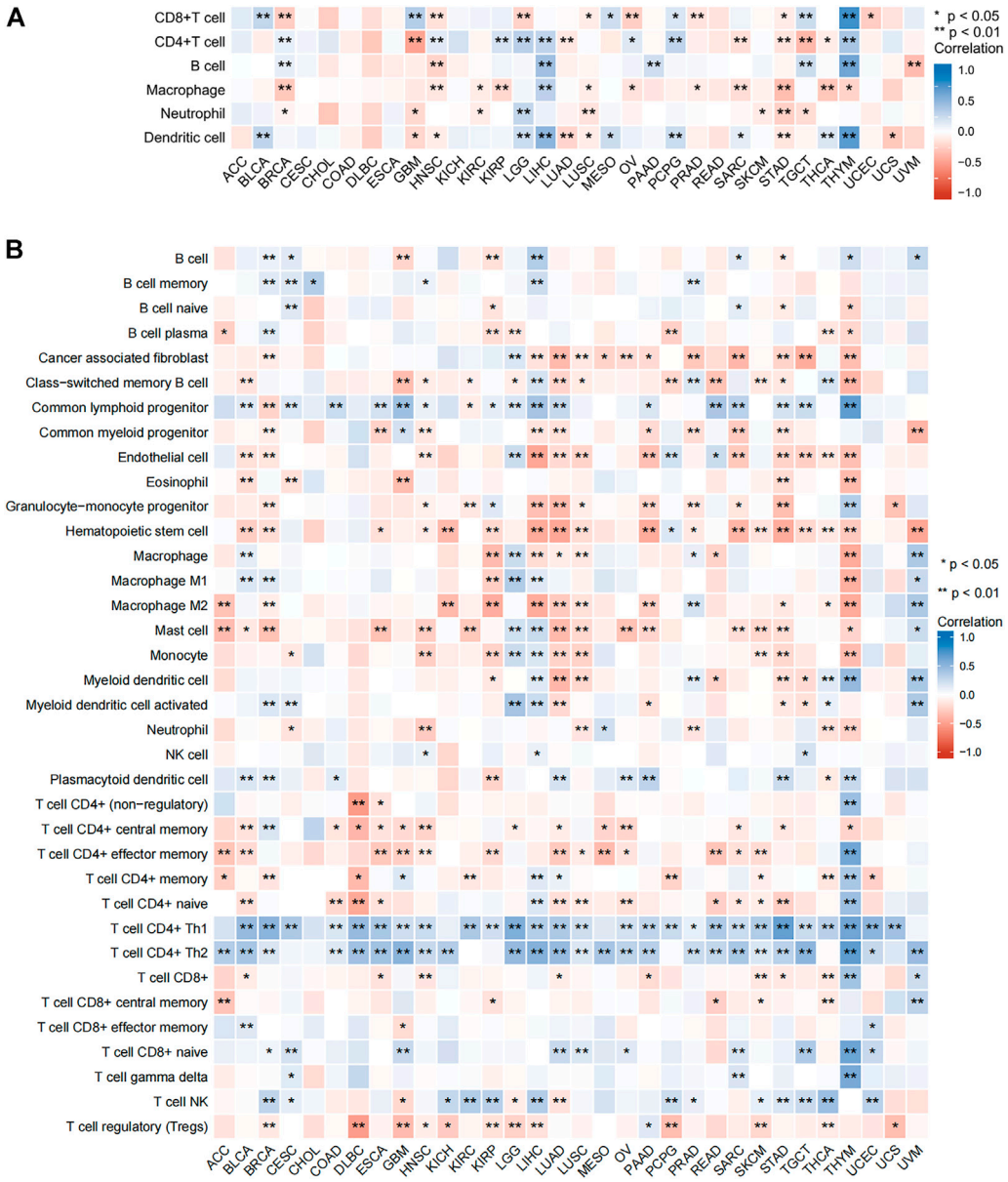



FIGURE 11
Relationship between SNRPB and immune cell infiltration. **(A)**. Relationship between SNRPB and immune cell infiltration based on the TIMER algorithm. **(B)**. Relationship between SNRPB and the infiltration of immune cell subtypes based on the xCell algorithm.

## Relationship of SNRPB with the expression of immunomodulation-related genes across tumors

In order to further explore the oncogenic effects of SNRPB on tumor immune surveillance, immune escape, and immune infiltration, we downloaded data from TIMER2 pertaining to the relationship between SNRPB and the expression of immune activation-related, immunosuppression-related, chemokine, and chemokine receptor genes across tumors. As shown in Supplementary Figure S2, SNRPB was significantly positively correlated with the expression of almost all immune activation-related and immunosuppression-related genes in LGG, LIHC, and UVM and significantly negatively correlated in LUAD and LUSC. As shown in Supplementary Figure S3, SNRPB was significantly positively correlated with the expression of almost all chemokines and chemokine receptor genes in LGG, LIHC, and UVM and significantly negatively correlated in LUSC and thymoma (THYM). Moreover,

SNRPB was also significantly negatively correlated with the expression of almost all chemokine receptor genes in LUAD, OV, PRAD, SKCM, STAD, UCEC, and uterine carcinosarcoma (UCS).

## Discussion

In the present study, we used multiple public database platforms to conduct an in-depth pan-cancer exploration of the role of SNRPB with respect to mRNA and protein expression, clinical outcome, genetic variation, promoter methylation, functional enrichment analysis, and tumor-infiltrating immune cells. Our results indicate that SNRPB expression was elevated in 28 tumors, correlated with late pathology stages and high TNM stages and was a risk factor for decreased OS, DSS, and PFI across tumors. Genetic variation in SNRPB may be associated with poor prognosis, and promoter methylation may be one of the mechanisms by which SNRPB was elevated across tumors. The PPI network revealed that CDK1, CDC6, AURKB, CCNB1, CCNA2, and CDC45 were the most closely interacting genes across tumors. Enrichment analysis of SNRPB and the most associated coexpressed genes were closely related to the cell cycle pathway. Moreover, SNRPB was also closely related to immune cell infiltration and the expression of immunomodulation-related genes in several tumors.

The pre-mRNA produced by DNA translation contains protein-coding exons and noncoding introns; introns are excised in the subsequent process, and exons are combined in different ways under the action of the spliceosome to produce various structurally and functionally distinct proteins (Sciarrillo et al., 2020). SNRPB is an important component of the spliceosome and thus is essential for the diversity of expressed proteins (Saltzman et al., 2008). New proteins produced by SNRPB dysregulation may be involved in tumorigenesis and progression (Correa et al., 2016; Liu et al., 2019; Peng et al., 2020; Zhan et al., 2020). Factors affecting gene expression levels include genetic alteration, epigenetic modifications, noncoding RNAs, mRNA transcriptional stability, and upstream transcription factors (Gu et al., 2017). Our results indicate that SNRPB expression was elevated in almost all tumors, and this change may be related to genetic alterations and decreased promoter methylation levels (epigenetic modifications). Posttranscriptional regulation is another important way to regulate gene expression, and microRNAs (miRNAs) are key posttranscriptional gene regulators (Cantini et al., 2019). Correa, B.R. et al. found that downregulation of tumor suppressor miRNAs could significantly trigger the overexpression of SNRPB in GBM (Correa et al., 2016). Another study in LIHC found that SNRPB is the downstream target of c-Myc, and the elevated expression of SNRPB was caused by overexpression of c-Myc (Peng et al., 2020). Therefore, the elevated expression of SNRPB in tumors may involve multiple mechanisms, including genetic alterations, epigenetic modifications, posttranscriptional modifications, and changes in the expression levels of upstream molecules.

Proliferation and migration are the basis of the progression and metastasis of tumors (Bucay et al., 2017). Both *in vitro* and *in vivo* experiments show that SNRPB can significantly promote the proliferation and migration of various tumor cells. *In vitro* experiments showed that overexpression of SNRPB could significantly promote cell growth, the formation of microspheres, and the migration of cells, while silencing SNRPB reduced proliferation, colony formation, and the number of migrated cells (Peng et al., 2020; Zhan et al., 2020). A similar phenomenon has been observed in NSCLC cells (Liu et al., 2019), CESC cells (Zhu et al., 2020), GBM cells (Correa et al., 2016), and THCA cells (Deng et al., 2022). A xenograft tumor assay indicated that transplanted tumors derived from SNRPB-transfected cells exhibit larger volumes and higher weights, and knockdown of SNRPB inhibits tumor growth (Zhan et al., 2020). The effect of SNRPB on the invasion, metastasis, and progression of tumors has also been confirmed (Liu et al., 2019; Zhu et al., 2020). The results of our study indicate that SNRPB can promote the progression of pathological and TNM stages, which is supported by the above studies.

The ultimate goal of oncology research is to cure the disease, improve the quality of life and prolong survival (Ogino et al., 2011). Therefore, the survival of patients is an important prognostic indicator. Our study found that SNRPB was a risk factor for decreased OS in 10 tumors, decreased DSS in 8 tumors, and decreased PFI in 7 tumors, suggesting that SNRPB could be a prognostic factor across tumors. At present, almost all studies on the effect of SNRPB on tumor prognosis are conducted based on public databases (Liu et al., 2019; Peng et al., 2020; Zhan et al., 2020). The basic findings provide clues regarding the mechanisms of the adverse effects of SNRPB on tumor prognosis. SNRPB not only promotes tumor proliferation, migration, and metastasis but also maintains the stemness of tumor cells (Zhan et al., 2020) and affects tumor responsiveness to therapy (Liu et al., 2021). Therefore, SNRPB may be involved in many aspects, such as tumorigenesis, progression, metastasis, maintenance of stemness, and treatment responsiveness, acting with adverse effects on the prognosis of tumors.

Cancer is a disease with extremely complex mechanisms and phenotypes. During the process of transforming normal cells into cancer cells, cells acquire shared functions necessary for their malignancy, such as sustained proliferation (Hanahan, 2022). Sustained proliferation represents the continuous entry of cells into the cell cycle, which means sustained activation of the cell cycle pathway. The cell cycle is a collection of events involving the replication of genetic material and distribution of replicated genetic material and cytoplasmic components to daughter cells (Zou and Lin, 2021). Through the PPI network, we identified the most closely interacting genes across tumors, including CDK1, CDC6, AURKB, CCNB1, CCNA2, and CDC45. All of these identified genes are associated with cell cycle processes such as

cell cycle transitions, DNA replication, DNA damage, and mitosis. KEGG enrichment analysis revealed that the cell cycle was the most enriched pathway, while GO enrichment analysis revealed that almost all enriched BP, CC, and MF functions were related to the cell cycle. These results indicate that the oncogenic role of SNRPB across tumors may be mainly achieved by affecting the cell cycle, and it would be very interesting to investigate the effect of SNRPB on cell cycle-targeted therapies across tumors in the future. AS is an important way to regulate gene expression and protein diversity, and 95% of multiexon transcripts undergo AS (Pan et al., 2008). The spliceosome pathway, which was significantly enriched in KEGG enrichment analysis, is required for AS. p53 is a well-established tumor suppressor gene, and the loss of p53 function is associated with tumor progression (Hanahan, 2022). Cellular senescence has long been recognized as a tumor suppression mechanism (Lee and Schmitt, 2019); however, recent studies have found that senescent cancer cells can promote proliferative signals, avoid apoptosis, induce angiogenesis, stimulate invasion and metastasis, and suppress tumor immunity in different ways (He and Sharpless, 2017; Faget et al., 2019; Lee and Schmitt, 2019; Wang et al., 2020). The p53 signaling pathway and cellular senescence were also significantly enriched in KEGG enrichment analysis, and the inhibitory effect of SNRPB on the p53 signaling pathway has been confirmed in CESC (Zhu et al., 2020). Therefore, we speculate that the oncogenic role of SNRPB across tumors may be mainly achieved by affecting cell cycle-related processes. In addition, the spliceosome, p53 signaling pathway, and cell senescence may also contribute to its oncogenic effect.

The tumor microenvironment (TME) is an integral part of cancer that significantly affects treatment response and clinical outcomes. As part of the TME, immune cells exert important impacts on tumor progression and prognosis (Pitt et al., 2016). Our study found that SNRPB was strongly associated with immune cell infiltration in some tumors and to a lesser extent or not at all in others. Immune cell subtype analysis revealed that SNRPB was significantly positively correlated with the infiltration of CD4+Th1 and CD4+Th2 in almost all tumors. These results suggest that the relationship between SNRPB and the infiltration of immune cells other than CD4+Th1 and CD4+Th2 cells is tumor specific, and this specificity may be related to the effect of SNRPB on the expression of immunomodulation-related genes. At present, there are few studies on the relationship between SNRPB and tumor immune cell infiltration, and more research is needed to confirm the relationship and mechanism in the future.

Here, for the first time, we systematically analyzed SNRPB expression, genetic variation, promoter methylation, and the relationship with prognosis, immune cell infiltration, and immunomodulation-related genes across tumors and preliminarily explored the oncogenic effect and mechanism of SNRPB. However, our study also has some limitations. First, our research was based on data obtained from public databases, without further validation of our findings at the cellular and animal levels. Furthermore, we failed to systematically and deeply explore the pathophysiological mechanisms underlying the results. Therefore, further studies are needed to clarify the oncogenic mechanism of SNRPB and its potential as a therapeutic target.

## Conclusion

The expression of SNRPB was significantly elevated in almost all tumors, and the decreased promoter methylation level may contribute to the elevated expression of SNRPB. SNRPB may facilitate the progression of pathological and TNM stages and is a risk factor for unfavorable prognosis across tumors.

## Data availability statement

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

## Author contributions

JW conceived the study and drafted the manuscript. JW performed the analyses with the help of FL, WW, and BY. XY supervised the study. All authors discussed the results and contributed to the final manuscript.

## Funding

This study was supported by the Natural Science Foundation of Jiangxi Province (Grant No. 20202ACBL206019) and the National Clinical Research Center for Geriatrics–JiangXi branch center (Grant No. 2021ZDG02001).

## 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/fmolb.2022.994440/full#supplementary-material

**SUPPLEMENTARY FIGURE 1**
validation of SNRPB expression in tumors and the corresponding normal tissues using data from the GEO database. ns, no significant difference; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

**SUPPLEMENTARY FIGURE 2**
Relationship of SNRPB with the expression of immunomodulation-related genes across tumors. **(A)** The relationship of SNRPB and the expression of immune activation-related genes across tumors. **(B)** The relationship of SNRPB and the expression of immunosuppression-related genes across tumors.

**SUPPLEMENTARY FIGURE 3**
Relationship of SNRPB with the expression of immunomodulation-related genes across tumors. **(A)** The relationship of SNRPB and the expression of chemokine genes across tumors. **(B)** The relationship of SNRPB and the expression of chemokine receptor genes across tumors.

## References

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: Archive for functional genomics data sets--update. *Nucleic Acids Res.* 41, D991–D995. (Database issue. doi:10.1093/nar/gks1193

Bray, F., Laversanne, M., Weiderpass, E., and Soerjomataram, I. (2021). The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer* 127 (16), 3029–3030. doi:10.1002/cncr.33587

Bucay, N., Bhagirath, D., Sekhon, K., Yang, T., Fukuhara, S., Majid, S., et al. (2017). A novel microRNA regulator of prostate cancer epithelial-mesenchymal transition. *Cell Death Differ.* 24 (7), 1263–1274. doi:10.1038/cdd.2017.69

Cantini, L., Bertoli, G., Cava, C., Dubois, T., Zinovyev, A., Caselle, M., et al. (2019). Identification of microRNA clusters cooperatively acting on epithelial to mesenchymal transition in triple negative breast cancer. *Nucleic Acids Res.* 47 (5), 2205–2215. doi:10.1093/nar/gkz016

Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2 (5), 401–404. doi:10.1158/2159-8290.CD-12-0095

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B. V. S. K., et al. (2017). Ualcan: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Consortium, G. (2013). The genotype-tissue expression (GTEx) project. *Nat. Genet.* 45 (6), 580–585. doi:10.1038/ng.2653

Correa, B. R., de Araujo, P. R., Qiao, M., Burns, S. C., Chen, C., Schlegel, R., et al. (2016). Functional genomics analyses of RNA-binding proteins reveal the splicing regulator SNRPB as an oncogenic candidate in glioblastoma. *Genome Biol.* 17 (1), 125. doi:10.1186/s13059-016-0990-4

Deng, Y., Wu, J., and Li, X. (2022). lncRNA rundc3a-AS1 regulates proliferation and apoptosis of thyroid cancer cells *via* the miR-151b/SNRPB Axis. *Int. J. Endocrinol.* 2022, 9433434. doi:10.1155/2022/9433434

Faget, D. V., Ren, Q., and Stewart, S. A. (2019). Unmasking senescence: Context-dependent effects of SASP in cancer. *Nat. Rev. Cancer* 19 (8), 439–453. doi:10.1038/s41568-019-0156-2

Goldman, M. J., Craft, B., Hastie, M., Repecka, K., McDade, F., Kamath, A., et al. (2020). Visualizing and interpreting cancer genomics data *via* the Xena platform. *Nat. Biotechnol.* 38 (6), 675–678. doi:10.1038/s41587-020-0546-8

Gray, T. A., Smithwick, M. J., Schaldach, M. A., Martone, D. L., Graves, J. A., McCarrey, J. R., et al. (1999). Concerted regulation and molecular evolution of the duplicated SNRPB'/B and SNRPN loci. *Nucleic Acids Res.* 27 (23), 4577–4584. doi:10.1093/nar/27.23.4577

Gu, M., Shao, N. Y., Sa, S., Li, D., Termglinchan, V., Ameen, M., et al. (2017). Patient-specific iPSC-derived endothelial cells uncover pathways that protect against pulmonary hypertension in BMPR2 mutation carriers. *Cell Stem Cell* 20 (4), 490–504. doi:10.1016/j.stem.2016.08.019

Hanahan, D. (2022). Hallmarks of cancer: New dimensions. *Cancer Discov.* 12 (1), 31–46. doi:10.1158/2159-8290.CD-21-1059

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013

He, S., and Sharpless, N. E. (2017). Senescence in health and disease. *Cell* 169 (6), 1000–1011. doi:10.1016/j.cell.2017.05.015

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., et al. (2013). Function of alternative splicing. *Gene* 514 (1), 1–30. doi:10.1016/j.gene.2012.07.083

Lee, S., and Schmitt, C. A. (2019). The dynamic nature of senescence in cancer. *Nat. Cell Biol.* 21 (1), 94–101. doi:10.1038/s41556-018-0249-2

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 48 (W1), W509–W514. doi:10.1093/nar/gkaa407

Li, Y. R., Guo, D., Chen, D. F., Lu, G. F., Ren, M. D., and He, S. X. (2022). [Regulatory effect of small nuclear ribonucleoprotein-associated protein B on proliferation and metastasis of liver cancer cells]. *Zhonghua Gan Zang Bing Za Zhi* 30 (1), 63–68. doi:10.3760/cma.j.cn501113-20200411-00174

Liu, N., Chen, A., Feng, N., Liu, X., and Zhang, L. (2021). SNRPB is a mediator for cellular response to cisplatin in non-small-cell lung cancer. *Med. Oncol.* 38 (5), 57. doi:10.1007/s12032-021-01502-0

Liu, N., Wu, Z., Chen, A., Wang, Y., Cai, D., Zheng, J., et al. (2019). SNRPB promotes the tumorigenic potential of NSCLC in part by regulating RAB26. *Cell Death Dis.* 10 (9), 667. doi:10.1038/s41419-019-1929-y

Ogino, S., Galon, J., Fuchs, C. S., and Dranoff, G. (2011). Cancer immunology--analysis of host and tumor factors for personalized medicine. *Nat. Rev. Clin. Oncol.* 8 (12), 711–719. doi:10.1038/nrclinonc.2011.122

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40 (12), 1413–1415. doi:10.1038/ng.259

Peng, N., Li, J., He, J., Shi, X., Huang, H., Mo, Y., et al. (2020). c-Myc-mediated SNRPB upregulation functions as an oncogene in hepatocellular carcinoma. *Cell Biol. Int.* 44 (5), 1103–1111. doi:10.1002/cbin.11307

Pitt, J. M., MarAbelle, A., Eggermont, A., Soria, J. C., Kroemer, G., and ZitvogeL, L. (2016). Targeting the tumor microenvironment: Removing obstruction to anticancer immune responses and immunotherapy. *Ann. Oncol.* 27 (8), 1482–1492. doi:10.1093/annonc/mdw168

Saltzman, A. L., Kim, Y. K., Pan, Q., Fagnani, M. M., Maquat, L. E., and Blencowe, B. J. (2008). Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Mol. Cell. Biol.* 28 (13), 4320–4330. doi:10.1128/MCB.00361-08

Sciarrillo, R., Wojtuszkiewicz, A., Assaraf, Y. G., Jansen, G., Kaspers, G. J. L., Giovannetti, E., et al. (2020). The role of alternative splicing in cancer: From oncogenesis to drug resistance. *Drug resist. updat.* 53, 100728. doi:10.1016/j.drup.2020.100728

Scofield, D. G., and Lynch, M. (2008). Evolutionary diversification of the Sm family of RNA-associated proteins. *Mol. Biol. Evol.* 25 (11), 2255–2267. doi:10.1093/molbev/msn175

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* 49 (D1), D605–D612. doi:10.1093/nar/gkaa1074

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* 47 (W1), W556–W560. doi:10.1093/nar/gkz430

Traube, F. R., and Carell, T. (2017). The chemistries and consequences of DNA and RNA methylation and demethylation. *RNA Biol.* 14 (9), 1099–1107. doi:10.1080/15476286.2017.1318241

Uhlen, M., Zhang, C., Lee, S., Sjostedt, E., Fagerberg, L., Bidkhori, G., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* 357 (6352), eaan2507. doi:10.1126/science.aan2507

Wang, B., Kohli, J., and Demaria, M. (2020). Senescent cells in cancer therapy: Friends or foes? *Trends Cancer* 6 (10), 838–857. doi:10.1016/j.trecan.2020.05.004

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., et al. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456 (7221), 470–476. doi:10.1038/nature07509

Yang, H., Cui, W., and Wang, L. (2019). Epigenetic synthetic lethality approaches in cancer therapy. *Clin. Epigenetics* 11 (1), 136. doi:10.1186/s13148-019-0734-x

Zhan, Y. T., Li, L., Zeng, T. T., Zhou, N. N., Guan, X. Y., and Li, Y. (2020). SNRPB-mediated RNA splicing drives tumor cell proliferation and stemness in hepatocellular carcinoma. *Aging (Albany NY)* 13 (1), 537–554. doi:10.18632/aging.202164

Zhang, S., Su, M., Sun, Z., Lu, H., and Zhang, Y. (2020). The signature of pharmaceutical sensitivity based on ctDNA mutation in eleven cancers. *Exp. Biol. Med.* 245 (8), 720–732. doi:10.1177/1535370220906518

Zhu, L., Zhang, X., and Sun, Z. (2020). SNRPB promotes cervical cancer progression through repressing p53 expression. *Biomed. Pharmacother.* 125, 109948doi:10.1016/j.biopha.2020.109948

Zou, T., and Lin, Z. (2021). The involvement of ubiquitination machinery in cell cycle regulation and cancer progression. *Int. J. Mol. Sci.* 22 (11), 5754. doi:10.3390/ijms22115754

frontiers | Frontiers in Molecular Biosciences

TYPE Mini Review
PUBLISHED 26 October 2022
DOI 10.3389/fmolb.2022.983102



OPEN ACCESS

EDITED BY
Hem Chandra Jha,
Indian Institute of Technology Indore, India

REVIEWED BY
Pedro José Carlos Rondot Radío, University of Buenos Aires, Argentina
Jabed Iqbal,
Singapore General Hospital, Singapore
José Manuel Cameselle-Teijeiro, University of Santiago de Compostela, Spain

*CORRESPONDENCE
Fernando Schmitt,
fschmitt@ipatimup.pt

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 30 June 2022
ACCEPTED 13 October 2022
PUBLISHED 26 October 2022

CITATION
Souza da Silva R, Pinto R, Cirnes L and Schmitt F (2022), Tissue management in precision medicine: What the pathologist needs to know in the molecular era.
*Front. Mol. Biosci.* 9:983102.
doi: 10.3389/fmolb.2022.983102



# Tissue management in precision medicine: What the pathologist needs to know in the molecular era

Ricella Souza da Silva[1], Regina Pinto[1], Luis Cirnes[1] and Fernando Schmitt[1,2,3]*

[1]IPATIMUP Diagnostics, IPATIMUP—Institute of Molecular Pathology and Immunology of Porto University, University of Porto, Porto, Portugal, [2]Department of Pathology, Faculty of Medicine of the University of Porto, University of Porto, Porto, Portugal, [3]CINTESIS@RISE (Health Research Network), Porto, Portugal

Precision medicine is "an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person." Among many medical specialists involved in precision medicine, the pathologists play an important and key role in the implementation and development of molecular tests that are in the center of decision of many therapeutic choices. Besides many laboratory procedures directly involved in the molecular tests, is fundamental to guarantee that tissues and cells collected for analysis be managed correctly before the DNA/RNA extraction. In this paper we explore the pivotal and interconnected points that can influence molecular studies, such as pre-analytical issues (fixation and decalcification); diagnosis and material selection, including the calculation of nuclei neoplastic fraction. The standardization of sample processing and morphological control ensures the accuracy of the diagnosis. Tissue or cytological samples constitutes the main foundation for the determination of biomarkers and development of druggable targets. Pathology and precision oncology still have a long way to go in terms of research and clinical practice: improving the accuracy and dissemination of molecular tests, learning in molecular tumor boards for advanced disease, and knowledge about early disease. Precision medicine needs pathology to be precise.

KEYWORDS

**molecular pathology, morphological control, oncology precision, personalized medicine, tissue management, solid tumors**

# 1 Introduction

Pathology is evolving to meet patient needs and be a central driver of personalized healthcare. The main objective of personalized medicine is to identify patients who are candidates for specific treatments. The individualized approach has transformed the oncological therapeutic scenario (Russo et al., 2022) and prognostic evolution (Siegel et al., 2021).

Traditionally, pathologists have been at the forefront of cancer diagnostics with knowledge of cancer biology, morphology and prognosis (Ooft et al., 2021). In precision medicine, target-based classification is progressively used to integrate the histology-based classification of tumors, which remains the pillar of cancer diagnosis and management (Danesi et al., 2021).

Precision oncology reveals an accelerated pace, with great financial investment and a rapidly growing pharmacological arsenal. The dynamic and changing environment has allowed gaps in the standardization and validation of molecular diagnostic procedures. Pathology plays an important role in the implementation and development of molecular tests—precision medicine needs pathology to be precise.

The involvement of the pathologist in genomic medicine and biomarkers starts with the management of the primary sample (Moore et al., 2018). Therefore, it is critical that the training pathologist understand the framework and basic technical elements of the molecular tests. Its performance, in accordance with current practices and guidelines for the molecular diagnostics workflow, also encompasses the main following phases: pre-analytical; diagnosis and appropriate selection of tissue sample (quality and quantity); use of "*in situ*" based techniques; integrate the molecular data in the original diagnostic report (translation to the clinicians); and educational activity (Schmitt, 2011; Cree et al., 2014; Rekhtman and Roy-Chowdhuri, 2016; Fassan, 2018; Roh, 2019; Gullo et al., 2021; Ooft et al., 2021; Vranic and Gatalica, 2021).

In this paper, we explore the pivotal and interconnected points that can influence the molecular study, such as pre-analytical issues (fixation and decalcification); diagnosis and material selection, including the calculation of nuclei neoplastic fraction. The standardization of sample processing and morphological control ensures the accuracy of the molecular diagnosis.

# 2 Pre-analytical

Suboptimal processing can alter morphological, immunohistochemical and molecular characteristics of histological and cytological samples (Bass et al., 2014; Gullo et al., 2021). The magnitude and direction of effects associated with a given pre-analytical factor are dependent on the object of analysis, whether DNA, RNA or protein (Penault-Llorca et al., 2022).

For all samples, the primary pre-analytical factors whose thresholds somehow affect molecular results can be grouped in: cold ischemia, fixation, decalcification and time of paraffin block storage (Bass et al., 2014; Cree et al., 2014).

Cold ischemia is the time that occurs between the removal and fixation of tissue and is especially problematic for large surgical specimens. It affects gene expression at RNA and protein levels and mutation analysis at DNA level, thus being a major factor in molecular pathology (Cree et al., 2014). The recommended thresholds for maximum cold ischemia time are: DNA <1 h for FISH, ≤24 h for PCR; RNA <12 h; Protein <12 h; Morphology <6 h (Bass et al., 2014).

The fixative worldwide accepted for tissue sample preservation is neutral buffered formalin (NBF). NBF penetrates tissue at around 1 mm/h, and fixation will only start when penetration occurs. NBF should only be used 24 h after dilution to 4% w/v, in order to reduce the effect of polymerization, and guarantee a stable 4% concentration. For cytological samples, fixation is commonly achieved with alcoholic-based fixatives (Cree et al., 2014). As the cold ischemia time, also the duration of fixation influences downstream nucleic acid, protein and morphological analyses (Penault-Llorca et al., 2022). The thresholds recommended for fixation times are: DNA <72 h; RNA 8–48 h; Protein 6–24 h; Morphology <1 year (Bass et al., 2014). An optimal fixation window of 6–48 h is recommended, based on findings that minimal nucleic acid degradation is observed before 72 h (Penault-Llorca et al., 2022). This period of time is also recommended to preserve the protein expression detected by immunohistochemistry. Sometimes, immunohistochemistry is used to detect anomalous expression or absence of expression as result of a molecular alteration (for example: P53 expression or absence of one of the mismatch repair proteins).

Decalcification is frequently not mentioned in the pathology report, though it can severely affect molecular analysis. To preserve the integrity of the nucleic acids and proteins, areas suspected of neoplasia in bone pieces should be processed without decalcification whenever they can be isolated. For smaller specimens, as bone biopsies, it is recommended to use Ethylenediaminetetraacetic (EDTA) as a decalcifying agent (Bass et al., 2014; Penault-Llorca et al., 2022).

Regarding the duration of paraffin block storage, the following thresholds were demonstrated: DNA ≤5 years; RNA ≤1 year; Protein ≤25 years for IHC, < 10 years for platforms requiring protein extraction (Bass et al., 2014).

# 3 Diagnosis and material selection

During the microscopic examination the pathologist should correlate the clinical information with morphology and recognize if there is the need for molecular analysis. The use of tissue in

small specimens should be rational, avoiding unnecessary sections and complementary studies (histochemical immunohistochemical, or hybridization). In some countries, as in our experience, molecular studies are done in central laboratories that received material from different hospitals. In this situation is advisable to have a dedicated molecular pathology that evaluates the available material before proceeding to any molecular study.

The following criteria are relevant in this phase: 1) Slide/sample for molecular analysis must accurately represent the diagnosis of the pathological report; 2) Most representative slide/sample: show all the characteristics of the neoplasm such as morphological type and staging; 3) Avoid samples with extensive areas of tumor necrosis, inflammatory infiltrate and fibrosis; 4) It is recommended that the pathologist marks the area of the section containing neoplasia on the hematoxylin and eosin (H&E) slide at the time of diagnosis.

To avoid gross sampling errors, the pathologist must directly certify that the block sent for analysis corresponds to the patient and the diagnosis. In patients with multiple samples, the most recent tissue should be used (Russo et al., 2022). It is also important to avoid choosing samples with scarce number of tumor cells or material previously used for a large number of stainings and/or molecular tests. The goal is to ensure that the sample that will be submitted to molecular tests be the most representative qualitatively and quantitatively, since the morphological control depends on these factors.

In our molecular laboratory, a specific and validated protocol of cutting sections has been optimized and is currently in use. First we perform sections that will be used for extraction of nucleic acids—at 10 μm thickness, and only after we do a 3 μm section and stain with H&E for morphological evaluation. The quantity of extraction sections obtained varies according to the molecular assay and/or biomarker: nine sections for Next Generation Sequencing (NGS) panels requiring DNA and RNA isolation (e.g., EGFR, ALK, ROS1, NTRK, BRAF, RET, MET, ERBB2/HER2, NRG1, and FGFR1); 4–6 sections for Sanger sequencing studies requiring only DNA isolation (e.g., BRAF, PIK3CA, MSI or methylation studies) four sections for real-time PCR based studies (e.g., RAS); 3–4 sections for studies requiring only RNA isolation (e.g., NTRK, PAM50).

## 4 Diagnosis and select of correct material: Morphological control for molecular testing

The percentage of neoplastic cells present in the selected sample must be estimated for morphological control (Cree et al., 2014; Dufraing et al., 2019; Gullo et al., 2021). Tissue quality is expressed as the percentage of neoplastic cells to the total number of nucleated cells in a sample.

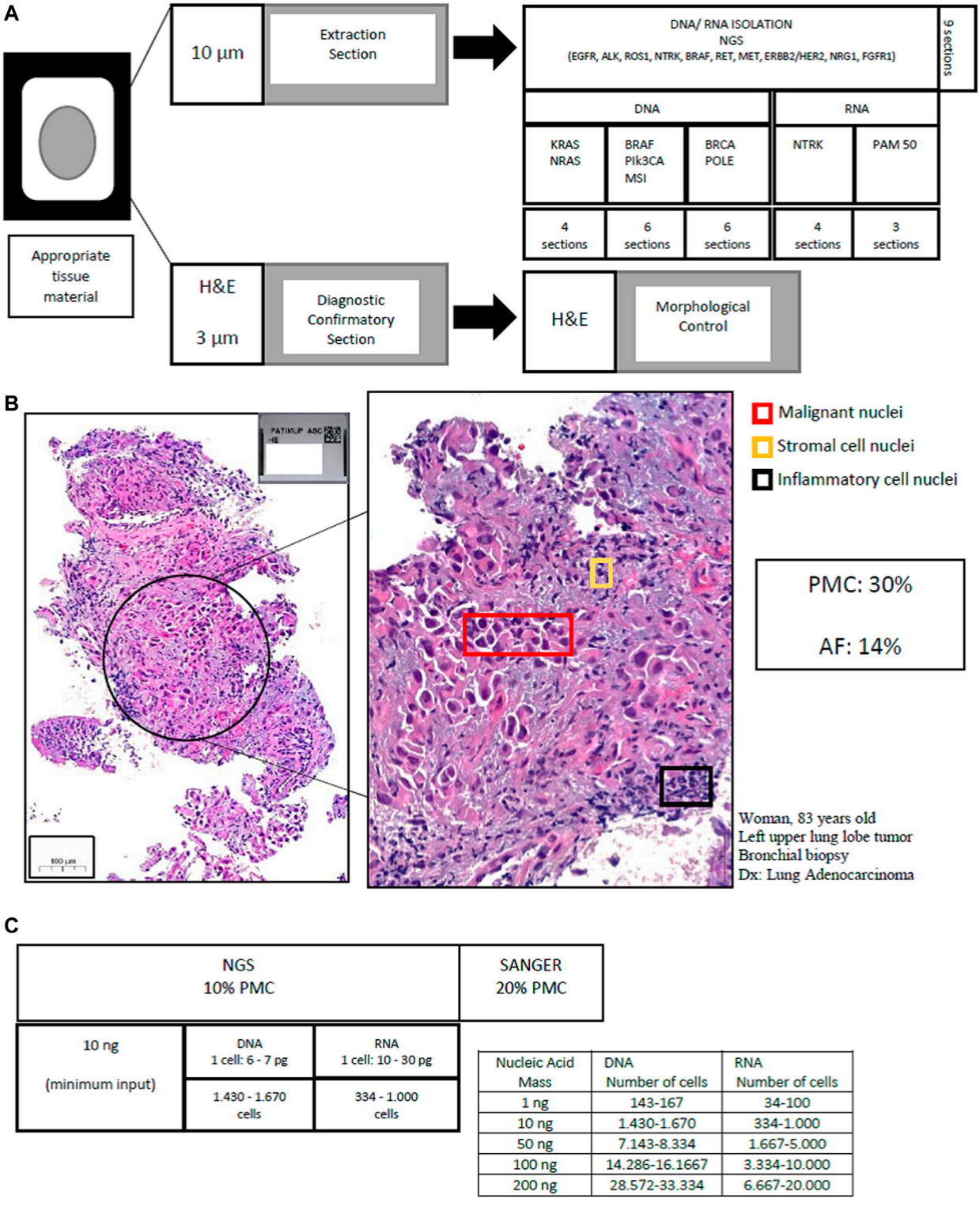
Recommendations for morphological control: 1) The evaluation of the percentage of malignant cells corresponds to the percentage of malignant nuclei (Gullo et al., 2021); 2) The evaluation of the percentage of malignant cells does not correspond to the size/area of the neoplasm; 3) It is recommended to make the estimation in deciles (e.g., 10%, 20%, . . ., 50%, . . . 100%) (Pei et al., 2019).

In estimating the percentage of viable malignant nuclei, regions with necrosis or inflammatory cells or desmoplastic stroma or mucus should be avoided and excluded for molecular analysis. If this is not possible, it is necessary to consider that these non-neoplastic nuclei are viable and dilute the percentage amount of malignant neoplastic DNA (Cree et al., 2014; Dufraing et al., 2019; Gullo et al., 2021; Kotoula et al., 2021). Tumor heterogeneity also needs to be taken into account when evaluating the percentage of neoplastic cells. These different areas, which represent cancer-specific growth patterns and tumor grading variation, must be present in the sample that will be submitted for molecular testing (Cree et al., 2014; Jennings et al., 2017; Nicholson et al., 2022).

The macrodissection is an important tumor cell enrichment technique, highlighting the critical role of the pathologist. This technique improving the accuracy of the molecular analysis, for direct sequencing or next generation sequencing (Gullo et al., 2021). It is considered necessary in the majority of sections from large tumor surgical specimens and occasionally in biopsy sections. The procedure is performed on unstained deparaffinized sections containing tissue fragments that have been directly processed for DNA/RNA extraction (Kotoula et al., 2021). The laser capture microdissection is largely a research tool and not necessary for routine molecular pathology (Cree et al., 2014).

Small samples with limited tumor cell content (<30%) may permit morphological classification; however, the quantity of tumor tissue is not always sufficient for biomarker testing (Penault-Llorca et al., 2022). The minimal amount of tumor DNA/RNA, as well as the minimal malignant cells required for molecular testing, are variable and dependent on the analytic sensitivity of a particular molecular assay (Dufraing et al., 2019). In general, a fraction of malignant cells greater than 10%–20% is considered a lower acceptable limit for molecular methods (e.g., 20% for Sanger sequencing or ~10% for next-generation sequencing).

If samples have lower than acceptable levels of malignant nuclei or in the absence of viable tumour tissue, histological slides used for morphological diagnosis (H&E and immunostained slides) can also be an additional alternative, to extract material for molecular analysis. From our experience, we can extract DNA and RNA from previous H&E slides, while the use of previous immunostained slides is indicated only for DNA extraction. RNA obtained from previous immunostained slides, shows degradation and does not give viable results. The utilization of the archive slides must be reserved and discussed case by case.




| Nucleic Acid Mass | DNA Number of cells | RNA Number of cells |
|---|---|---|
| 1 ng | 143-167 | 34-100 |
| 10 ng | 1.430-1.670 | 334-1.000 |
| 50 ng | 7.143-8.334 | 1.667-5.000 |
| 100 ng | 14.286-16.1667 | 3.334-10.000 |
| 200 ng | 28.572-33.334 | 6.667-20.000 |

FIGURE 1
Workflow to sample processing and morphological control in molecular pathology. **(A)** Tissue cutting protocol (FFPE or cytological preparations): DNA/RNA extraction and H&E stain for morphological control. **(B)** Morphological Control in lung adenocarcinoma: The percentage of neoplastic cells corresponds to the percentage of malignant nuclei. Pathological assessment that 30% of the malignant cells were viable. The allele frequency verified was 14%, confirming the estimate. **(C)** Number of cells needed by nucleic acid quantification in single-multi-gene assays. PMC (percentage malignant cell), AF (allele frequency), NGS (next generation sequencing).

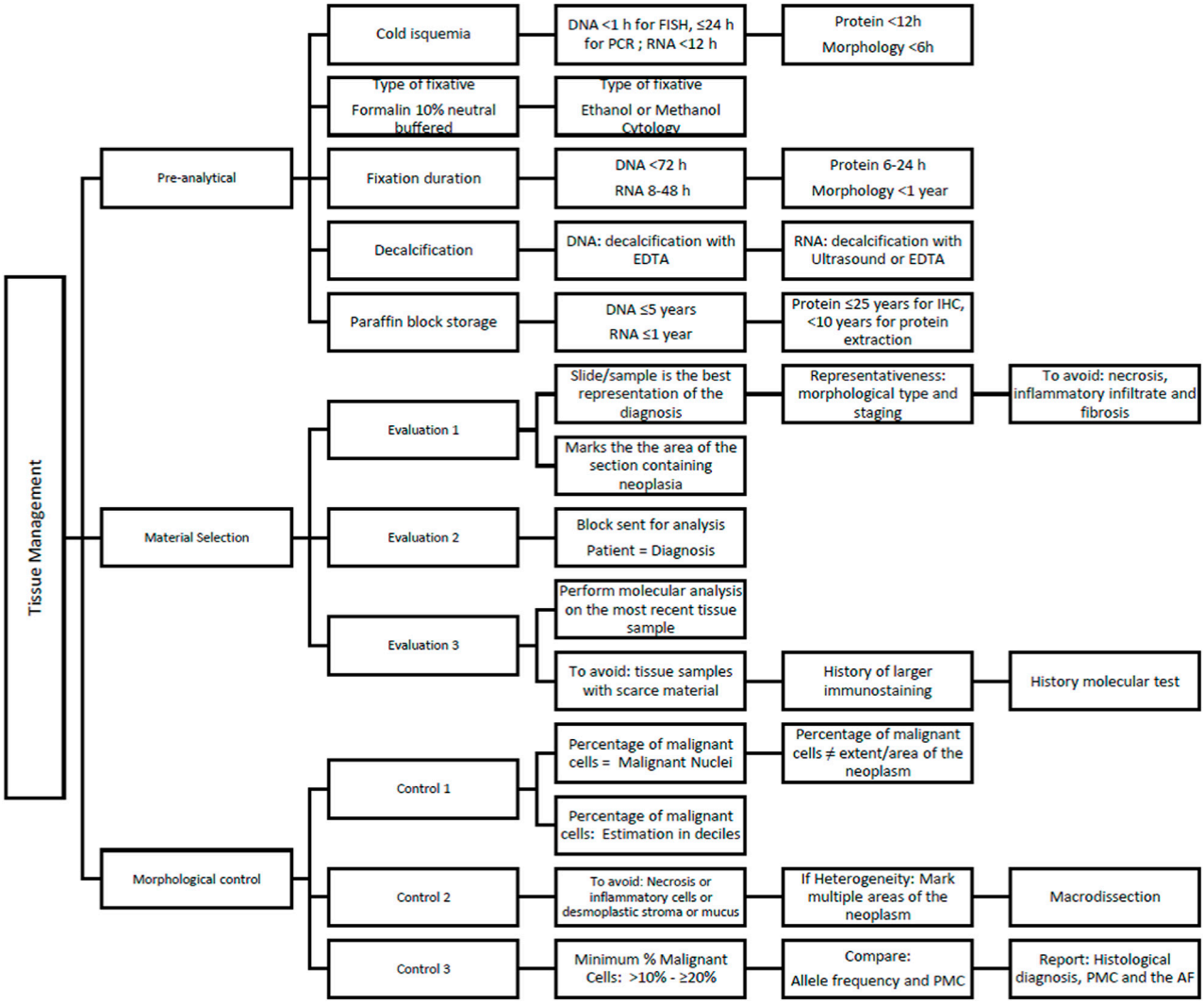



FIGURE 2
Recommendations for Solid Tumors Tissue Management.

Furthermore, the scanning and digital archiving of the slides is mandatory to ensure medical-legal issues.

When DNA is purified from a pure tumor cell population (100% tumor cellularity), the mutant allele frequency would be expected to be 50%. If a sample harbored 50% tumor cells and 50% non-tumor cells, the latter population harboring two copies of the wild type allele, the mutant allele frequency would be 25% (Roh, 2019). If in the purified nucleic acid sample, the mutant allele represent 15%–20% of the allelic population overall (mutant plus wild type), it means that the sample has approximately 30%–40% tumor cells.

It is important to note that an intact diploid cell produces 6–7 pg of DNA and 10–30 pg of RNA. Input nucleic acid mass requirements for molecular testing are variable, with minimum recommendations ranging from of 1–10 ng (typical minimum input for most NGS platforms), 50 ng, 100 ng, and 200 ng (Benayed et al., 2019; Roh, 2019; Qu et al., 2020; Penault-Llorca et al., 2022).

Cytopathology samples ethanol-fixed, as smears and touch preparations, usually contain higher-quality nucleic acids than formalin-fixed, paraffin-embedded (FFPE) samples and are useful for molecular testing (Gan and Roy-Chowdhuri, 2020). For example, to isolate 10 ng of nucleic acids approximately three– fourfold more cells are required from a FFPE sample than from ethanol-fixed material. It is beneficial if the pathologist indicates the areas containing the highest proportion of neoplastic cells on the slide because it helps to avoid contamination of the material for nucleic acid extraction by

non-neoplastic elements (Dietel et al., 2016). Particularly in cytology material, preliminary experiences have demonstrated that samples with <100 cells are not suitable for NGS, samples between 100–2000 represent low levels and >5000 cells are suitable for any NGS including large panels.

Figure 1 shows the approach from sample processing to morphological control in molecular assays.

The interobserver variability in estimating the percentage of neoplastic cells is 20% (Dufraing et al., 2019) and pathologists more accurately estimate the percentage of malignant nuclei in cases containing low amounts of tumor cells (Viray et al., 2013). The variability in assessment can be minimized through feedback from the sequence variants with a review of the histomorphology and assessment of biomarkers. The tumor percentage estimation and allele frequency (AF) should be compared (Ooft et al., 2021). When AF is higher than expected, the variant might be a germline, or the gene might be affected by heterozygosity or tumor cell aneuploidy might have occurred. If a low AF is detected in a sample containing a high percentage of neoplastic cell, the result could be interpreted as a testing artefact or as an indicator for subclonality of a given gene in the context of tumor heterogeneity (Dufraing et al., 2019).

The pathology molecular report should include the histopathological diagnosis, the estimated percentage of malignant cells present in the tissue used for DNA/RNA extraction, and the allele frequency. Figure 2 describes the recommendations for solid tumors tissue management for molecular analysis.

## 5 Discussion

Diagnosis, classification, prognosis (Ooft et al., 2021) and even study of mechanisms of cancer development (Echejoh et al., 2021) are the core of pathology. The development of precise medicine brings molecular diagnosis for the prime time. Pathologists are probably the most appropriate for connecting morphology, clinical setting, mutational status, and the reflection of these findings in therapeutics. Previous tumor categorization criteria are being replaced by a new approach based on specific genetic abnormalities, in which tissue or cytological samples constitutes the main foundation for the determination of biomarkers and development of druggable targets.

The standardization of tissue sample handling at each step—sample processing and morphological control—ensures the accuracy of the diagnosis. Estimation of percent neoplastic cellularity can also affect care in other diagnostic settings, including assessment of copy number variation, chromosomal translocation, determination of residual cancer burden after neoadjuvant therapy, and other genomic aberrations that can be affected by contaminating normal DNA (Viray et al., 2013).

Best practice guidelines aim to management tissue and support a complete molecular diagnosis, so that eligible patients may benefit from targeted therapy. A biomarker negative sample with a tumor cell content below the thresholds for analysis should be determined as inconclusive, requiring further assessment. Consequently, the proportion of cells may also inform the choice of molecular testing.

Molecular therapy demonstrates remarkable response rates, making advances in reducing cancer mortality (Siegel et al., 2021), although the response to genome-targeted therapy has been modest (2.73% in 2006 to 7.04% in 2020) and more trials are needed to determine the impact on survival, which currently stands at 4.7 months (Haslam et al., 2021; Haslam et al., 2022).

In this scenario, it is necessary to continue researching the complex interaction of tumor biology, microenvironment, and immune response; to disseminate molecular tests in clinical practice; to encourage molecular tumor board for advanced diseases; and to learn about early diseases. Pathology and oncology in precision medicine still have a long way to go in research and clinical practice.

## Author contributions

FS, RS, and RP contributed to the study conception and design. LC and RP contribute with laboratory data. The first draft of the manuscript was written by RS, supervised by FS. A final revision was made by FS and RS. All authors read and approved the final manuscript.

## Acknowledgments

FCT Fundação para a Ciência e a Tecnologia, I. P., within the scope of the project “RISE—LA/P/0053/2020”.

## 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


Bass, B. P., Engel, K. B., Greytak, S. R., and Moore, H. M. (2014). A review of preanalytical factors affecting molecular, protein, and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: How well do you know your FFPE specimen? *Arch. Pathol. Lab. Med.* 138, 1520–1530. doi:10.5858/arpa.2013-0691-RA

Benayed, R., Offin, M., Mullaney, K., Sukhadia, P., Rios, K., Desmeules, P., et al. (2019). High yield of RNA sequencing for targetable kinase fusions in lung adenocarcinomas with No mitogenic driver alteration detected by DNA sequencing and low tumor mutation burden. *Clin. Cancer Res.* 25, 4712–4722. doi:10.1158/1078-0432.CCR-19-0225

Cree, I. A., Deans, Z., Ligtenberg, M. J. L., Normanno, N., Edsjö, A., Rouleau, E., et al. (2014). Guidance for laboratories performing molecular pathology for cancer patients. *J. Clin. Pathol.* 67, 923–931. doi:10.1136/jclinpath-2014-202404

Danesi, R., Fogli, S., Indraccolo, S., Del Re, M., Dei Tos, A. P., Leoncini, L., et al. (2021). Druggable targets meet oncogenic drivers: Opportunities and limitations of target-based classification of tumors and the role of molecular tumor boards. *ESMO Open* 6, 100040. doi:10.1016/j.esmoop.2020.100040

Dietel, M., Bubendorf, L., Dingemans, A.-M. C., Dooms, C., Elmberger, G., García, R. C., et al. (2016). Diagnostic procedures for non-small-cell lung cancer (NSCLC): Recommendations of the European expert group. *Thorax* 71, 177–184. doi:10.1136/thoraxjnl-2014-206677

Dufraing, K., van Krieken, J. H., De Hertogh, G., Hoefler, G., Oniscu, A., Kuhlmann, T. P., et al. (2019). Neoplastic cell percentage estimation in tissue samples for molecular oncology: Recommendations from a modified delphi study. *Histopathology* 75, 312–319. doi:10.1111/his.13891

Echejoh, G., Liu, Y., Chung-Faye, G., Charlton, J., Moorhead, J., Clark, B., et al. (2021). Validity of whole genomes sequencing results in neoplasms in precision medicine. *J. Clin. Pathol.* 74, 718–723. doi:10.1136/jclinpath-2020-206998

Fassan, M. (2018). Molecular diagnostics in pathology: Time for a next-generation pathologist? *Arch. Pathol. Lab. Med.* 142, 313–320. doi:10.5858/arpa.2017-0269-RA

Gan, Q., and Roy-Chowdhuri, S. (2020). Small but powerful: The promising role of small specimens for biomarker testing. *J. Am. Soc. Cytopathol.* 9, 450–460. doi:10.1016/j.jasc.2020.05.001

Gullo, I., Marques, A., Pinto, R., Cirnes, L., and Schmitt, F. (2021). Morphological control for molecular testing: A practical approach. *J. Clin. Pathol.* 74, 331–333. doi:10.1136/jclinpath-2020-206890

Haslam, A., Kim, M. S., and Prasad, V. (2022). Overall survival for oncology drugs approved for genomic indications. *Eur. J. Cancer* 160, 175–179. doi:10.1016/j.ejca.2021.10.028

Haslam, A., Kim, M. S., and Prasad, V. (2021). Updated estimates of eligibility for and response to genome-targeted oncology drugs among US cancer patients, 2006-2020. *Ann. Oncol.* 32, 926–932. doi:10.1016/j.annonc.2021.04.003

Jennings, L. J., Arcila, M. E., Corless, C., Kamel-Reid, S., Lubin, I. M., Pfeifer, J., et al. (2017). Guidelines for validation of next-generation sequencing-based oncology panels: A joint consensus recommendation of the association for molecular pathology and college of American pathologists. *J. Mol. Diagn.* 19, 341–365. doi:10.1016/j.jmoldx.2017.01.011

Kotoula, V., Chatzopoulos, K., Papadopoulou, K., Giannoulatou, E., Koliou, G.-A., Karavasilis, V., et al. (2021). Genotyping data of routinely processed matched primary/metastatic tumor samples. *Data Brief.* 34, 106646. doi:10.1016/j.dib.2020.106646

Moore, D. A., Young, C. A., Morris, H. T., Oien, K. A., Lee, J. L., Jones, J. L., et al. (2018). Time for change: A new training programme for morpho-molecular pathologists? *J. Clin. Pathol.* 71, 285–290. doi:10.1136/jclinpath-2017-204821

Nicholson, A. G., Tsao, M. S., Beasley, M. B., Borczuk, A. C., Brambilla, E., Cooper, W. A., et al. (2022). The 2021 WHO classification of lung tumors: Impact of advances since 2015. *J. Thorac. Oncol.* 17, 362–387. doi:10.1016/j.jtho.2021.11.003

Ooft, M. L., Boissiere, J., and Sioletic, S. (2021). The histopathologist is essential in molecular pathology quality assurance for solid tumours. *Virchows Arch.* 479, 1263–1265. doi:10.1007/s00428-021-03226-y

Pei, Z., Cao, S., Lu, L., and Chen, W. (2019). Direct cellularity estimation on breast cancer histopathology images using transfer learning. *Comput. Math. Methods Med.* 2019, 3041250. doi:10.1155/2019/3041250

Penault-Llorca, F., Kerr, K. M., Garrido, P., Thunnissen, E., Dequeker, E., Normanno, N., et al. (2022). Expert opinion on NSCLC small specimen biomarker testing — Part 1: Tissue collection and management. *Virchows Arch.* 481, 335–350. doi:10.1007/s00428-022-03343-2

Qu, X., Yeung, C., Coleman, I., Nelson, P. S., and Fang, M. (2020). Comparison of four next generation sequencing platforms for fusion detection: Oncomine by ThermoFisher, AmpliSeq by illumina, FusionPlex by ArcherDX, and QIAseq by QIAGEN. *Cancer Genet.* 243, 11–18. doi:10.1016/j.cancergen.2020.02.007

Rekhtman, N., and Roy-Chowdhuri, S. (2016). Cytology specimens: A goldmine for molecular testing. *Arch. Pathol. Lab. Med.* 140, 1189–1190. doi: doi:10.5858/arpa.2016-0379-ED

Roh, M. H. (2019). The utilization of cytologic and small biopsy samples for ancillary molecular testing. *Mod. Pathol.* 32, 77–85. doi:10.1038/s41379-018-0138-z

Russo, A., Incorvaia, L., Capoluongo, E., Tagliaferri, P., Galvano, A., Del Re, M., et al. (2022). The challenge of the molecular tumor board empowerment in clinical oncology practice: A position paper on behalf of the aiom- SIAPEC/IAP-SIBioC-SIC-SIF-SIGU-SIRM Italian scientific societies. *Crit. Rev. Oncol. Hematol.* 169, 103567. doi:10.1016/j.critrevonc.2021.103567

Schmitt, F. C. (2011). Molecular cytopathology and flow cytometry: Pre-analytical procedures matter. *Cytopathology* 22, 355–357. doi:10.1111/j.1365-2303.2011.00941.x

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer statistics, 2021. *Ca. Cancer J. Clin.* 71, 7–33. doi:10.3322/caac.21654

Viray, H., Li, K., Long, T. A., Vasalos, P., Bridge, J. A., Jennings, L. J., et al. (2013). A prospective, multi-institutional diagnostic trial to determine pathologist accuracy in estimation of percentage of malignant cells. *Arch. Pathol. Lab. Med.* 137, 1545–1549. doi:10.5858/arpa.2012-0561-CP

Vranic, S., and Gatalica, Z. (2021). The role of pathology in the era of personalized (precision) medicine: A brief review. *Acta Med. Acad.* 50, 47–57. doi:10.5644/ama2006-124.325

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 28 November 2022
DOI 10.3389/fmolb.2022.1074277



OPEN ACCESS

EDITED BY
Umberto Malapelle,
University of Naples Federico II, Italy

REVIEWED BY
Abilash Gangula,
University of Missouri, United States
Dirk Roggenbuck,
Brandenburg University of Technology Cottbus-Senftenberg, Germany
Francesco Pepe,
University of Naples Federico II, Italy

*CORRESPONDENCE
María José Serrano,
mjose.serrano@genyo.es
M. Carmen Garrido-Navas,
carmen.garrido@genyo.es

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 19 October 2022
ACCEPTED 16 November 2022
PUBLISHED 28 November 2022

CITATION
Roa-Colomo A, López Garrido MÁ, Molina-Vallejo P, Rojas A, Sanchez MG, Aranda-García V, Salmeron J, Romero-Gomez M, Muntane J, Padillo J, Alamo JM, Lorente JA, Serrano MJ and Garrido-Navas MC (2022), Hepatocellular carcinoma risk-stratification based on ASGR1 in circulating epithelial cells for cancer interception.
*Front. Mol. Biosci.* 9:1074277.
doi: 10.3389/fmolb.2022.1074277



# Hepatocellular carcinoma risk-stratification based on ASGR1 in circulating epithelial cells for cancer interception

Amparo Roa-Colomo[1,2], María Ángeles López Garrido[3], Pilar Molina-Vallejo[4], Angela Rojas[5,6], Mercedes González Sanchez[3], Violeta Aranda-García[7], Javier Salmeron[2], Manuel Romero-Gomez[5,6], Jordi Muntane[6,8,9], Javier Padillo[10], Jose María Alamo[10], Jose A. Lorente[4,11], María José Serrano[4,12,13]* and M. Carmen Garrido-Navas[4]*

[1]Clinical Medicine and Public Health Doctoral Program, University of Granada, Granada, Spain, [2]Gastroenterology and Hepatology Department, San Cecilio University Hospital, Granada, Spain, [3]Gastroenterology and Hepatology Department, Virgen De Las Nieves University Hospital, Granada, Spain, [4]Genyo-Centro Pfizer-Universidad De Granada-Junta De Andalucía De Genómica e Investigación Oncológica, Granada, Spain, [5]Seliver Group, Institute of Biomedicine of Seville (IBiS)/Hospital Universitario Virgen Del Rocío/CSIC/Universidad De Sevilla, Seville, Spain, [6]Spanish Network for Biomedical Research in Hepatic and Digestive Diseases (CIBERehd), Carlos III Health Institute (ISCIII), Madrid, Spain, [7]Statistician at Fundación para la Investigación Biosanitaria Andalucía Oriental Alejandro Otero (FIBAO), Hospital Virgen de las Nieves, Granada, Spain, [8]Institute of Biomedicine of Seville (IBiS), Hospital University Virgen del Rocío/CSIC/University of Seville, Seville, Spain, [9]Department of Medical Physiology and Biophysics, University of Seville, Seville, Spain, [10]General and Gastrointestinal Surgery Division, Virgen del Rocío University Hospital, Sevilla, Spain, [11]Legal Medicine Department, Medicine School, University of Granada, Granada, Spain, [12]Comprehensive Oncology Division, Clinical University Hospital, Virgen de las Nieves-IBS, Granada, Spain, [13]Department of Pathological Anatomy, Faculty of Medicine, University of Granada, Granada, Spain

**Purpose:** Lack of diagnostic and prognostic biomarkers in hepatocellular carcinoma impedes stratifying patients based on their risk of developing cancer. The aim of this study was to evaluate phenotypic and genetic heterogeneity of circulating epithelial cells (CECs) based on asialoglycoprotein receptor 1 (ASGR1) and miR-122-5p expression as potential diagnostic and prognostic tools in patients with hepatocellular carcinoma (HCC) and liver cirrhosis (LC).

**Methods:** Peripheral blood samples were extracted from LC and HCC patients at different disease stages. CECs were isolated using positive immunomagnetic selection. Genetic and phenotypic characterization was validated by double immunocytochemistry for cytokeratin (CK) and ASGR1 or by *in situ* hybridization with miR-122-5p and CECs were visualized by confocal microscopy.

**Results:** The presence of CECs increased HCC risk by 2.58-fold, however, this was only significant for patients with previous LC ($p$ = 0.028) and not for those without prior LC ($p$ = 0.23). Furthermore, the number of CECs lacking ASGR1 expression correlated significantly with HCC incidence and absence of miR-122-5p expression ($p$ = 0.014; $r$ = 0.23). Finally, overall survival was significantly greater for patients at earlier cancer stages ($p$ = 0.018), but this

difference was only maintained in the group with the presence of CECs ($p$ = 0.021) whereas progression-free survival was influenced by the absence of ASGR1 expression.

**Conclusion:** Identification and characterization of CECs by ASGR1 and/or miR-122-5p expression may be used as a risk-stratification tool in LC patients, as it was shown to be an independent prognostic and risk-stratification marker in LC and early disease stage HCC patients.

KEYWORDS

**hepatocellular carcinoma, liver cirrhosis, circulating tumor cells, precision medicine, cancer interception**

# 1 Introduction

Liver cancer affected more than 900,000 individuals worldwide in 2020 and the estimated incidence is expected to rise to 1.4 million individuals in 2040 (Cancer Tomorrow, 2022). Hepatocellular carcinoma (HCC) is the most common liver cancer, accounting for 75%–85% of all primary liver tumors (Sung et al., 2021). The risk of HCC is known to be increased by external factors (such as excessive alcohol consumption, viral infections, aflatoxin exposure) inducing liver inflammation and hepatic fibrosis progression; however, the impact of internal factors (beyond fatty liver disease) increasing the risk of developing HCC is still poorly studied (Forner et al., 2018). The known HCC risk factors promote liver cirrhosis (LC), which is usually a prior finding in 80% HCC patients (Llovet et al., 2021). In fact, 2%–5% of cirrhotic patients will develop HCC annually (Fateen and Ryder, 2017), so surveillance by liver ultrasound screening is performed in LC for early diagnosis of HCC (Forner et al., 2018; Reig et al., 2021). Unfortunately, curative treatments (based on surgery or liver transplant) are available in less than 60% of cases due to late diagnosis (Cadier et al., 2017). There is an absence of tissue biopsies because diagnosis of HCC in LC patients is mainly based on dynamic imaging tests, what reduces the ability to molecularly characterize the tumor (Reig et al., 2021). In addition to imaging tests, serum markers such as alpha-fetoprotein (AFP) that is linked to the evolutionary stage of the tumor, have been used for diagnosis and prognosis of HCC (Fateen and Ryder, 2017). However, the fact that 1) AFP may be elevated in chronic hepatitis without HCC (Force et al., 2022) 2) many small-sized HCCs have normal AFP levels (Carr et al., 2018), 3) 30% of HCC patients have normal AFP levels at diagnosis (Lee et al., 2019) and 4) high AFP levels were found in HCC-free individuals (Kobeisy et al., 2012), highlights the lack of sensitivity and specificity of this marker (Bialecki and di Bisceglie, 2005). Although AFP levels greater than 400 ng/ml are considered diagnostic of HCC as well as a marker of bad prognosis marker (Bialecki and di Bisceglie, 2005), it is not recommended to use AFP as a sole marker for HCC surveillance (Hanif et al., 2022) and the combination with other blood-based biomarkers is suggested to improve HCC diagnosis (Wang and Zhang, 2020).

Liquid biopsies have the potential to improve sensitivity/specificity as they are non-invasive, represent better the tumor heterogeneity and can be used to monitor disease evolution. The most widely studied type of liquid biopsy with demonstrated clinical utility are Circulating Tumor Cells or CTCs. The evaluation of CTCs in terms of occurrence, cell count and phenotypic characterization has demonstrated its prognostic value in several solid tumors, including lung (Bayarri-Lara et al., 2016), (de Miguel-Pérez et al., 2019), breast (Nadal et al., 2013), (Ye et al., 2019) and colorectal (Delgado-Ureña et al., 2018) cancers. With respect to HCC, several meta-analyses demonstrated the diagnostic and prognostic utility of detecting CTCs (Fan et al., 2015), (Sun et al., 2017), (Cui et al., 2020), although different isolation and detection technologies may undermine their clinical utility. Isolation methodologies based on EpCAM such as CellSearch® are widely used, although elevated EpCAM expression levels were only found in metastatic HCC lesions compared to primary and vascular invaded tumor (Tsuchiya et al., 2019). Furthermore, elevated EpCAM expression levels were also linked to poor prognosis (Shimada et al., 2019). Thus, using alternative methodologies for CTC isolation, such as those based on size (Wang et al., 2018), (Qi et al., 2018) or including other biomarkers such as glypican-3 (Court et al., 2018) among others could overcome CellSearch® limitations. However, Wang et al. (2018) reported lack of correlation between CTCs and recurrence after liver transplantation even though CTCs were detected in more than two-thirds of HCC patients. This suggests that not only improvements for CTC isolation are needed, but also that CTC characterization might be useful for diagnostic and prognostic purposes. Furthermore, presence of Circulating Epithelial Cells (CECs) in patients with pre-malignant diseases such as chronic obstructive pulmonary disease (COPD) had prognostic and diagnostic value (Romero-Palacios et al., 2019) in the context of Cancer Interception (Serrano et al., 2020), although it has been poorly studied in the context of liver cirrhosis (Chen et al., 2020). In fact, in most studies including LC patients, CTCs/CECs were not detected (Vona et al., 2004) or reported percentages were very low (Takahashi et al., 2021), possibly due to either patient selection or isolation methodology.

One liver-specific biomarker used to isolate CTCs from HCC patients is the asialoglycoprotein receptor 1 (ASGR1) (Li et al., 2014; Court et al., 2018), that represents the human lectin subunit 1 of the asialoglycoprotein receptor. ASGR1 heterodimerizes with ASGR2 (human lectin subunit 2) to produce a transmembrane protein primarily present in sinusoidal and basolateral hepatocellular membranes. The main role of ASGR1 is to bind galactosyl residues, facilitating glycoproteins turnover (Muramatsu, 2007). In fact, it was recently studied in the context of cardiovascular diseases (Xie et al., 2021; Wang et al., 2022a) demonstrating that its inhibition reduces hypercholesterolemia and atherosclerosis (Xie et al., 2021) by increasing cholesterol excretion (Wang et al., 2022a). In the context of liver cancer, the role of ASGR1 as tumor suppressor was first suggested *via* its interaction with LASS2 (longevity assurance homolog 2 of yeast LAG1) (Gu et al., 2016) and more recently by its association with DNA methylation (Zhu et al., 2022), confirming the prior findings of loss of ASGR1 expression in HCC tissue (Shi et al., 2013; Witzigmann et al., 2016). Another potential biomarker of HCC is miR-122-5p, a microRNA involved in multiple physiological processes in the liver that was found to suppress cell proliferation and malignant transformation of hepatocytes (Hu et al., 2012). Downregulation of miR-122-5p in HCC patients as well as HCC-derived cell lines was demonstrated together with an inverse correlation with cyclin G1 expression (Gramantieri et al., 2007) and upregulation of miR-122-5p was shown to repress the epithelial to mesenchymal transition (EMT) through the WNT/β-cadherin signaling pathway *via* Snail 1/2 (Jin et al., 2017). In plasma, miR-122-5p together with other four miRNAs were described as a diagnostic tool (Jin et al., 2019) and recently, its prognostic role was demonstrated when survival of HCC patients was greater in those with higher miR-122-5p expression (Wang et al., 2022b).

In this proof-of-concept study, we evaluated the potential diagnostic and prognostic role of circulating epithelial cells (CECs) using an isolation methodology based on immunomagnetic selection with cytokeratins 7/8 (CK) followed by ASGR1 and miR-122-5p characterization in a cohort of patients suffering from LC, with and without subsequent HCC.

# 2 Material and methods

## 2.1 Study design and sample collection

This prospective cohort study included 113 patients aged between 32 and 86 years suffering from hepatocellular carcinoma (HCC, $N$ = 71) or cancer-free liver cirrhosis (LC; $N$ = 42). Inclusion criterion for cirrhotic patients was diagnosis of LC by dynamic imaging tests (computerized axial tomography and/or magnetic resonance imaging). Tumor staging for HCC patients followed the BCLC (Barcelona Clinic Liver Cancer) classification (Forner et al., 2018). All patients were aged over 18 and had no other liver disease beyond liver cirrhosis and/or HCC (exclusion criteria). For analytical purposes, the HCC cohort was divided into those diagnosed at an earlier stage (eHCC = BCLC 0-A; $N$ = 30) and those diagnosed at a later stage (aHCC = BCLC B-C-D; $N$ = 41). Of the 30 eHCC patients, only 11 were subjected to liver transplant. Patients were recruited between 2017 and 2020 at the Gastroenterology and Hepatology Units of three Andalusian University hospitals: San Cecilio (Granada), Virgen de las Nieves (Granada) and Virgen del Rocio (Seville). Informed consent was obtained from all patients before blood extraction and the study protocol was approved by the Hospital´s Ethics Committee following the ethical guidelines of the 1975 Declaration of Helsinki. Not signing the informed consent and not fulfilling the inclusion criteria were exclusion criteria. Samples were anonymized upon blood collection to ensure patients' privacy and clinicians involved in the project updated the database for clinical information. Patients' clinical and pathological characteristics are shown in Table 1.

## 2.2 Circulating epithelial cell isolation

Peripheral blood samples (15 ml) were collected in K2-EDTA Vacutainer tubes at the time of diagnosis. Blood samples were enriched in peripheral mononuclear blood cells using gradient centrifugation with Ficoll-Histopaque®-1119 and CECs were isolated using the Carcinoma Cell Enrichment Kit (130-060-301, Miltenyi Biotec) based on pan-anti-cytokeratin, as previously described (Nadal et al., 2013). Isolated CECs were then spun down onto Poly-L-lysine-coated glass slides using a cytospin (Hettich) for subsequent phenotypic and genetic characterization (one glass slide was prepared per 7.5 ml peripheral blood).

## 2.3 CECs enumeration and phenotypic characterization by ASGR1 detection

Before isolating and characterizing CECs from patients, antibody specificity for ASGR1 was assessed on the hepatocellular tumor cell line HEPG2 (reference 85011430, lot. 2440) obtained by the Centre of Scientific Instrumentation at the University of Granada. The primary alveolar epithelial cell line hAELVi (InSCREENeX GmbH) was selected as a negative control for ASGR1 expression as stated in the human protein atlas and peripheral blood mononuclear cells (PBMCs) from a healthy donor were used as negative control for both antibodies (Supplementary Figure S1). Both cell lines were tested with STR assay for cell authenticity and for mycoplasma contamination. All experiments were done in duplicates. Subsequently, isolated CECs from patients were enumerated and characterized using double immunocytochemistry with chromogenic staining for cytokeratin (CK) followed by fluorescent detection (for ASGR1). CK detection was done using the Carcinoma Cell

10.3389/fmolb.2022.1074277

TABLE 1 Risk factors for HCC. Abbreviations are: HCC, hepatocellular carcinoma; n, number of individuals; SD, standard deviation; Dx, diagnosis; INR, international normalized ratio; AFP, alpha-fetoprotein; HCV, hepatitis C virus; CEC, circulating epithelial cells; IQ, interquartile; CK, cytokeratin; ASGR1, asialoglycoprotein receptor 1. For quantitative analyses *t*-student and Mann-Whitney tests were used depending on data normality. For qualitative data, Chi-Square was used except for those cases with less than 20% of the data with lower than 5 expected frequencies, in which case, Fisher test was used. *p* values are: ***$p < 0.001$ and **$p < 0.05$. Bold is the mean for each value as a whole.

| | HCC-free $N = 42$ | HCC-affected $N = 71$ | *p* values |
|---|---|---|---|
| Sex [*n* (%)] | | | <0.001*** |
| Female | 16 (38.1%) | 7 (9.9%) | |
| Male | 26 (61.9%) | 64 (90.1%) | |
| Age (mean ± SD, years) | **66 ± 9** | **66 ± 10** | 0.784 |
| Female | 67 ± 9 | 63 ± 11 | |
| Male | 64 ± 9 | 66 ± 10 | |
| Cirrhosis Dx (mean ± SD, years) | **61 ± 10** | **62 ± 11** | 0.529 |
| Female | 63 ± 10 | 63 ± 13 | |
| Male | 59 ± 10 | 62 ± 11 | |
| Bilirubin levels (mean mg/dL) | **0.96 ± 0.38** | **1.70 ± 2.82** | 0.034** |
| Female | 0.90 ± 0.44 | 1.16 ± 1.89 | |
| Male | 1.0 ± 0.34 | 1.76 ± 2.91 | |
| Albumin levels (mean g/dL) | **4.10 ± 0.40** | **3.70 ± 0.60** | <0.001*** |
| Female | 3.90 ± 0.40 | 3.70 ± 0.70 | |
| Male | 4.20 ± 0.30 | 3.70 ± 0.60 | |
| Platelets number (mean $10^6$/L) | **146,524 ± 78,641** | **152,052 ± 103,980** | 0.767 |
| Female | 151,125 ± 69,214 | 173,957 ± 172,546 | |
| Male | 143,692 ± 85,125 | 149,656 ± 95,491 | |
| INR (mean ± SD) | **1.13 ± 0.35** | **1.23 ± 0.32** | 0.116 |
| Female | 1.26 ± 0.54 | 1.12 ± 0.18 | |
| Male | 1.05 ± 0.12 | 1.25 ± 0.33 | |
| Prothrombin activity (%) | **90 ± 21** | **77 ± 18** | <0.001*** |
| Female | 84 ± 27 | 83 ± 77 | |
| Male | 94 ± 16 | 77 ± 18 | |
| AFP baseline (mean ng/mL) | **6.13 ± 16.65** | **1,111.62 ± 4,685.08** | <0.001*** |
| Female | 11.15 ± 27.64 | 5,333.91 ± 13,330.00 | |
| Male | 3.32 ± 1.38 | 619.03 ± 2,030.34 | |
| Alcohol, n (%) | | | 0.17 |
| Female | 4 (33.3%) | 4 (14.33%) | |
| Male | 8 (66.7%) | 24 (85.7%) | |
| HCV, *n* (%) | | | 0.045** |
| Female | 7 (35%) | 1 (6.3%) | |
| Male | 13 (65%) | 15 (93.8%) | |
| Child Pugh-Turcotte, *n* (%) | | | <0.0001*** |
| A (5–6) | 40 (100) | 48 (71.6) | |
| B (7–9) | 0 | 15 (22.4) | |
| C (10–15) | 0 | 4 (6.0) | |
| CEC phenotype, *n* (median; IQ range) | | | 0.044** |
| CK+/ASGR1+ | 17 (2; 4) | 29 (2; 2) | |
| CK+/ASGR1− | 7 (1; 4) | 25 (2; 7) | |
| Negative | 18 (0) | 17 (0) | |
| Total sum of CECs | | | 0.0002*** |
| CK+/ASGR1+ | 36 | 48 | |
| CK+/ASGR1− | 15 | 74 | |

Detection Kit (130-060-301, Miltenyi Biotec) as previously described (Nadal et al., 2013). CECs from HCC and LC patients were then blocked with 10% goat serum and 1.5% FcR blocking in cell stain solution for 45 min and incubated with 1/100 diluted anti-ASGR1 rabbit antibody (Atlas Antibodies Cat# HPA012852, Merck) in 1% goat serum overnight. For visualization of ASGR1, 10 µg of goat anti-rabbit Alexa Fluor 633 antibody (A-21070, Thermo Fisher Scientific) were added for 20 min. Finally, slides were mounted using 4 µg of Hoechst 33342 (Thermo Fisher Scientific) and SlowFade™ Gold Antifade Mounting Media (S36936, Thermo Fisher Scientific). Cell enumeration and characterization was performed in a laser confocal microscope (Zeiss LSM 710) and pictures were taken using a ×60 oil objective. CECs were reported as either $CEC^{CK+/ASGR1+}$ (for positive ASGR1 staining) or $CEC^{CK+/ASGR1-}$ (for negative ASGR1 staining).

## 2.4 CECs genetic characterization by immunoFISH

As a proof-of-concept, we selected 5 individuals positive for CECs belonging to each patient cohort (5 LC, 5 eHCC and 5 aHCC) and a second glass slide for each patient allowed genetic characterization by immunoFISH. This procedure was not performed in all patients due to: availability of enough sample, difficulties of the combination of chromogenic staining with fluorescence detection as well as financial reasons. In patients positive for CECs, *fluorescence in situ hybridization* (FISH) for the miR-122-5p in combination with immunofluorescence with cytokeratin was performed to determine liver-origin of CECs. Samples were treated with filtered 0.1% pepsin (Merck) in 10 mM HCL during 1 min at 37°C, washed twice with PBS and then hybridized for 1 h at 57°C with miRCURY LNA miRNA Detection Probe for miR-122-5p (Qiagen, Cat#339453) according to the manufacturer's protocol. After increasing stringency washes with SSC samples were blocked (PBS 1x, 0.1% Tween-20, 2% goat serum, 1% BSA) and incubated with anti-digoxigenin antibody (Roche, Cat#11093274910) for 1 h at room temperature. After several PBS-Tween washes, samples were stained using the pre-filtered substrate FastRed for 1.5 h at 30°C. Staining was terminated by addition of KTBT buffer and subsequently, immunofluorescence for cytokeratin-FITC was performed. Supplementary Figure S2 shows internal controls for specificity of probes and methodology.

## 2.5 Statistical analysis

Descriptive analysis of variables was performed using SPSS, calculating measures of central trend and dispersion for the numerical variables; absolute frequencies and percentages for qualitative variables were also calculated. Normality of the data was studied with the Shapiro-Wilks test. A bivariate analysis was carried out to analyze possible factors related to the main variables. For numerical variables, the Student's t test was applied for independent samples or Mann-Whitney in non-parametric cases. For qualitative variables, Pearson's Chi-square test or Fisher's exact test were applied. In addition, odds ratio and its 95% confidence interval were calculated for each variable. With those that were statistically significant, a multivariate logistic regression model was proposed to jointly predict which factors influenced tolerance to treatment. The variable selection method was performed by successive steps backwards, eliminating in each step those variables that did not significantly influence the model, applying the likelihood ratio test. To evaluate the goodness of fit of the model, the Hosmer-Lemeshow statistic was calculated and to calculate survival rates, overall survival (OS) and progression-free survival (PFS) were plotted as Kaplan Meier curves. Statistical significance was considered for $p < 0.05$. Graphs were created using GraphPad.

# 3 Results

## 3.1 Clinical and pathological characteristics of the study cohort

Our study cohort included 113 individuals, 42 cancer-free LC patients and 71 HCC patients (of which only 9 did not have prior LC). Univariate analysis of factors associated with HCC included sex, albumin levels and prothrombin activity ($p < 0.001$), bilirubin ($p = 0.034$), etiology including hepatitis virus C ($p = 0.045$) and Child Pugh-Turcotte stage (Table 1). No relevant information was obtained using multivariate analysis (data not shown).

## 3.2 Circulating epithelial cells characterization in LC and HCC patients

CECs were detected in 79 patients (69.9%) with significantly greater frequencies in HCC (54/71; 76.1%) than in HCC-free patients (24/42; 57.1%) ($p = 0.023$). Phenotypic heterogeneity of CECs was identified both intra and inter individual based on ASGR1. Particularly, two phenotypes were identified: ASGR1 positive (CK+/ASGR1+) and negative (CK+/ASGR1−), with varying intensities and sizes (Figure 1).

No significant differences were observed between the ASGR1 positive (46/78; 58.9%) and the ASGR1 negative (32/78; 41.1%) phenotype ($p = 0.33$), being ASGR1 expression more prevalent in cirrhotic (17/24; 70.8%) than in HCC (29/54; 53.7%) patients. Importantly, all CECs with the ASGR1 positive phenotype also showed positivity for the miR-122-5p liver-specific marker (Figure 2, rows 1,3 and 5), demonstrating

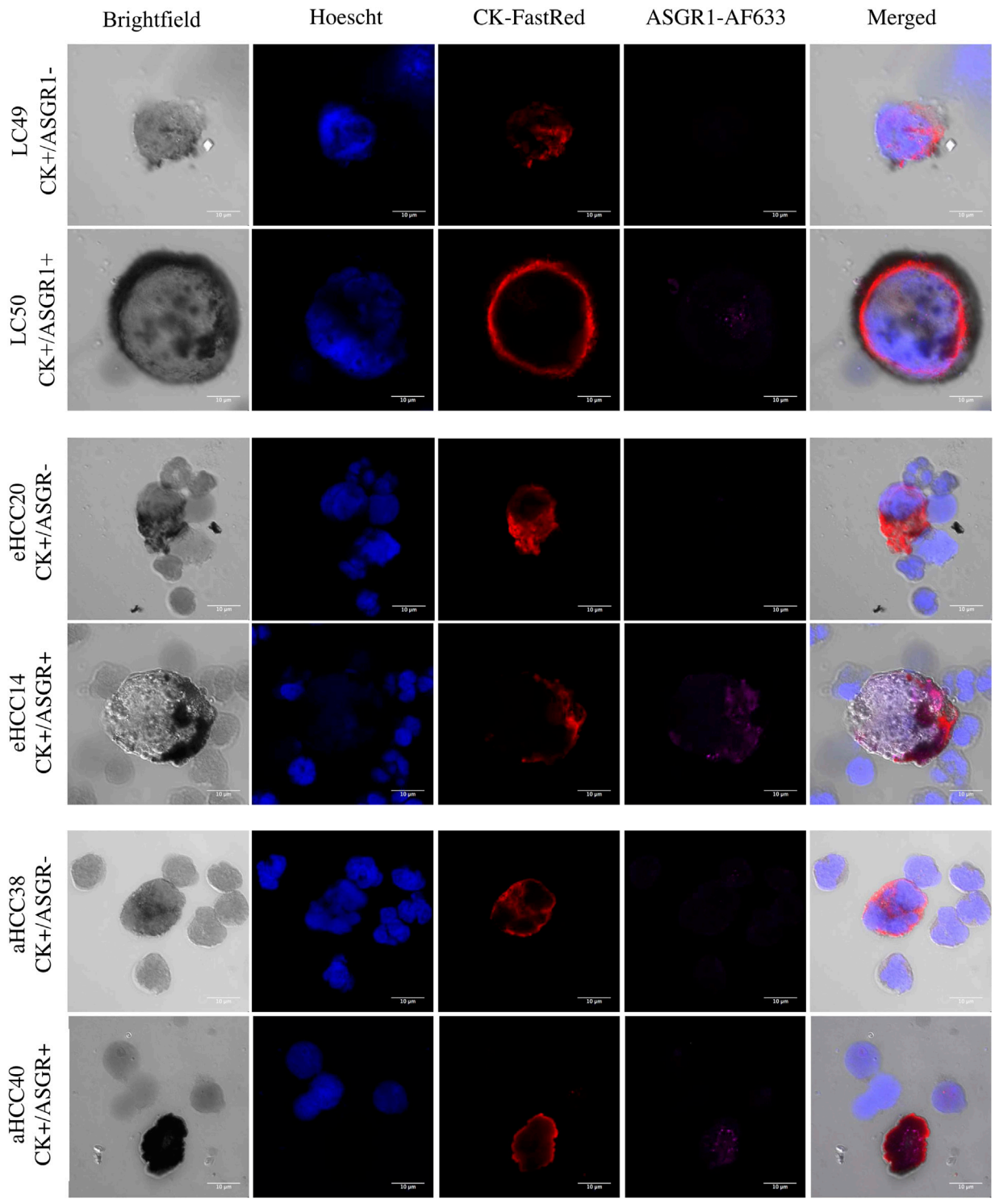



FIGURE 1
CECs heterogeneity in patients with liver cirrhosis or HCC. Heterogeneity is shown with respect to Cytokeratin (CK) and ASGR1 expression in two LC (liver cirrhosis) patients (top), two eHCC (early HCC) patients (middle) and two aHCC (advanced HCC) patients (bottom). In each case, one CEC positive for the two markers (CK and ASGR1) and one CEC negative for ASGR1 are shown. Hoechst is used as nuclear staining.

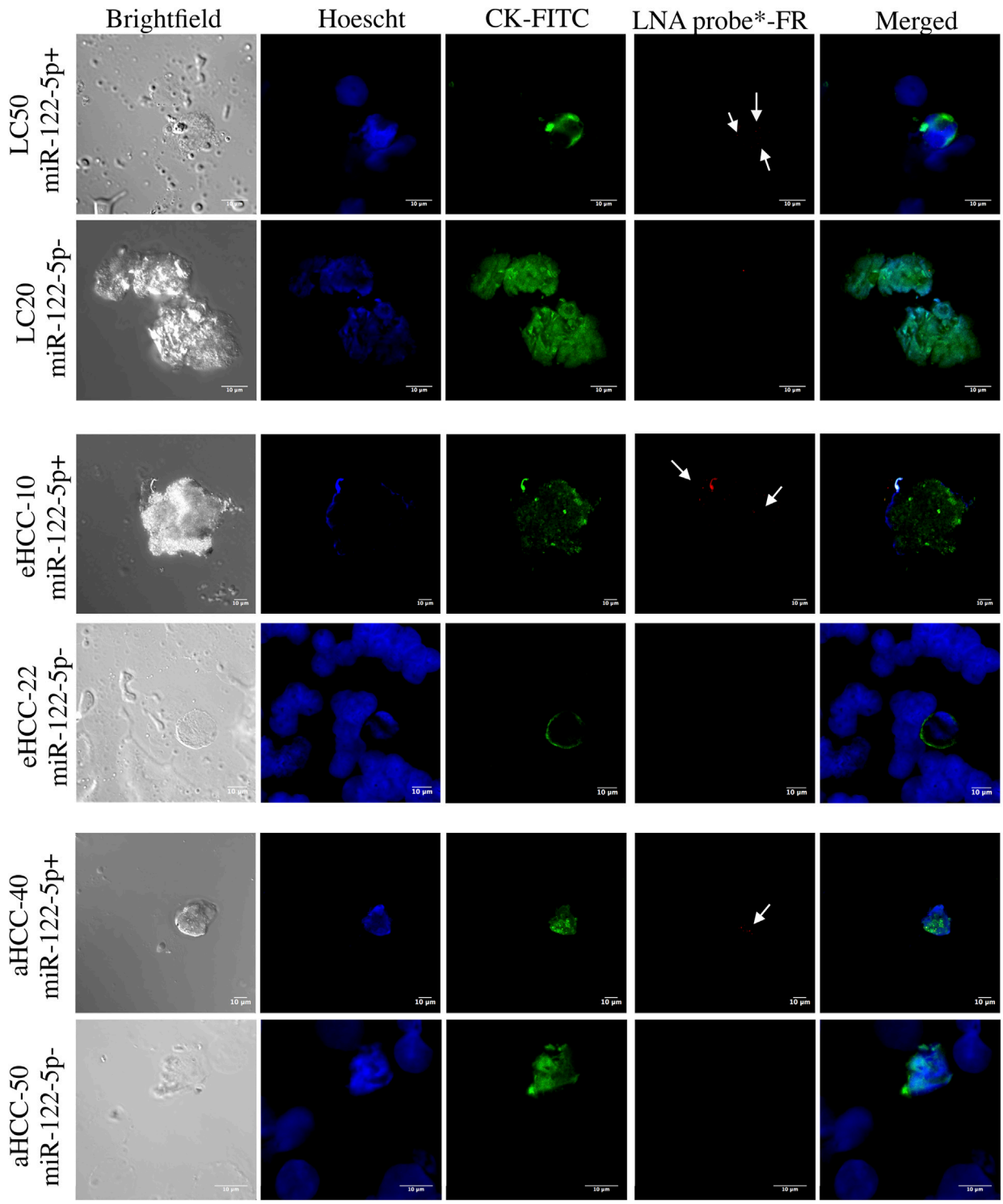



FIGURE 2
Confocal microscopy of immunoFISH for miR-122-5p in patients with liver cirrhosis or HCC. Cytokeratin (CK) is shown in green (FITC) and miR122-5p probe is shown in red. Data from two patients for each group: liver cirrhosis (LC; top), early HCC (eHCC; middle) and advanced HCC (aHCC; bottom) are shown. For each patient group, a positive miR-122-5p signal (coinciding with positive ASGR1 staining) and a negative miR-122-5p signal (coinciding with negative ASGR1 staining) are shown.

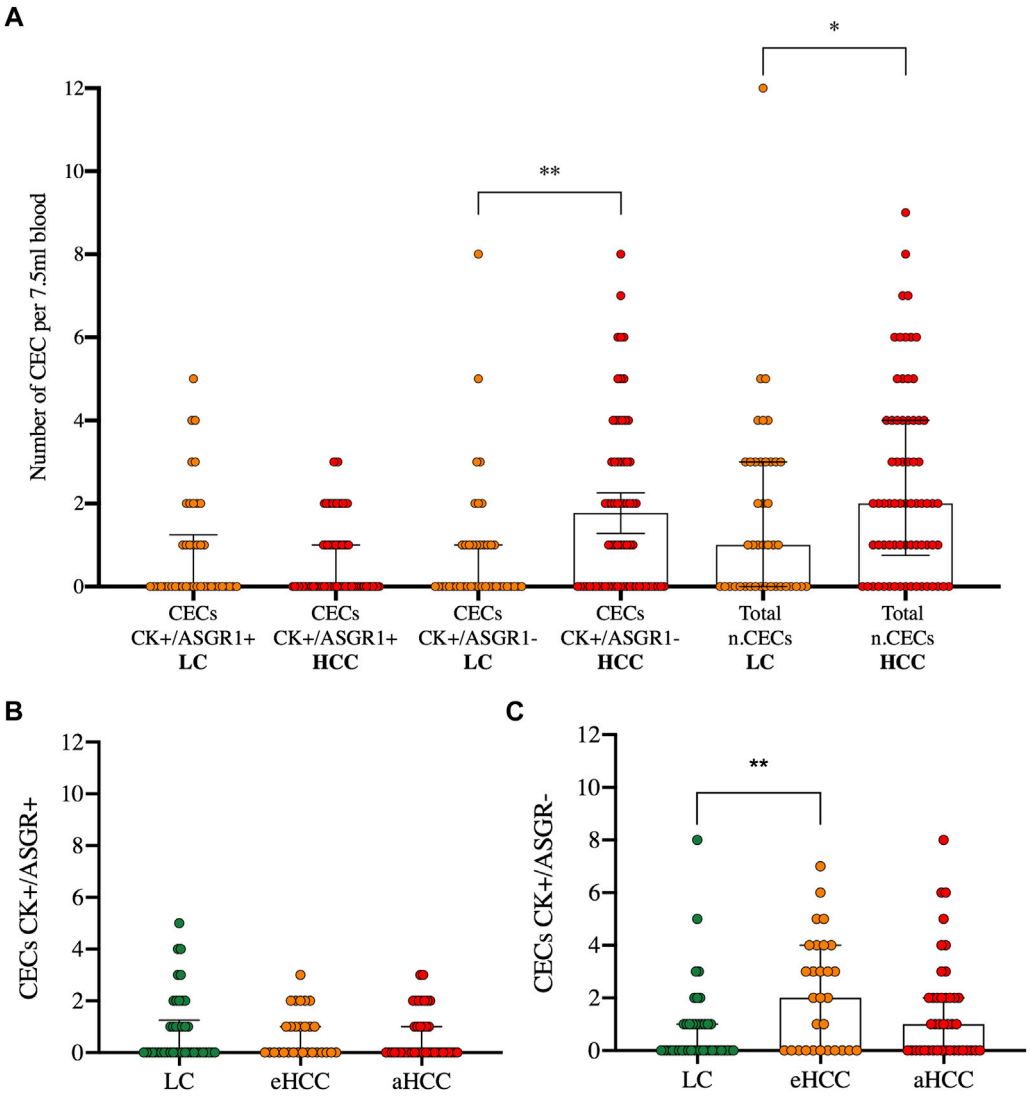



FIGURE 3
CECs enumeration in blood in patients with liver cirrhosis or HCC. CECs are characterized by either double staining (CK+/ASGR1+) or single staining (CK+/ASGR1-) in liver cirrhosis (LC) and hepatocellular carcinoma (HCC) patients **(A)**. The latter are divided into early (eHCC) and advanced (aHCC) disease for both, double **(B)** and single staining **(C)**. *p* values for Kruskal Wallis test and multiple comparisons are:***p* = 0.0096 and **p* = 0.035 **(A)** and ***p* = 0.0051 **(C)**.

their liver origin. In contrast, absence of miR-122-5p was observed in all patients with CECs showing ASGR1 negative staining (Figure 2, rows 2,4 and 6).

## 3.3 Correlation of CEC phenotypes with clinical and pathological data

The amount of $CEC^{CK+/ASGR1-}$ was significantly greater in HCC than in HCC-free patients ($p$ = 0.0096), significantly increasing the total CEC count in HCC compared to LC patients ($p$ = 0.035) (Figure 3A). Furthermore, when accounting for disease stage: early (BCLC 0-A) or advanced (BCLC B-C-D), no differences were observed for the ASGR1 positive population (Figure 3B). However, Kruskal-Wallis tests revealed that the number of CECs with absence of ASGR1 expression was significantly different among disease stages ($p$ = 0.017). There was a significant increase of $CEC^{CK+/ASGR1-}$ between LC and eHCC ($p$ = 0.0051) but not between LC and aHCC ($p$ = 0.08) or eHCC and aHCC ($p$ = 0.23) (Figure 3C).

Our results suggest that both presence of CECs and absence of ASGR1 expression in CECs are risk biomarkers of HCC. In fact, presence of CECs increased HCC risk by 2.58 -fold ($p$ =

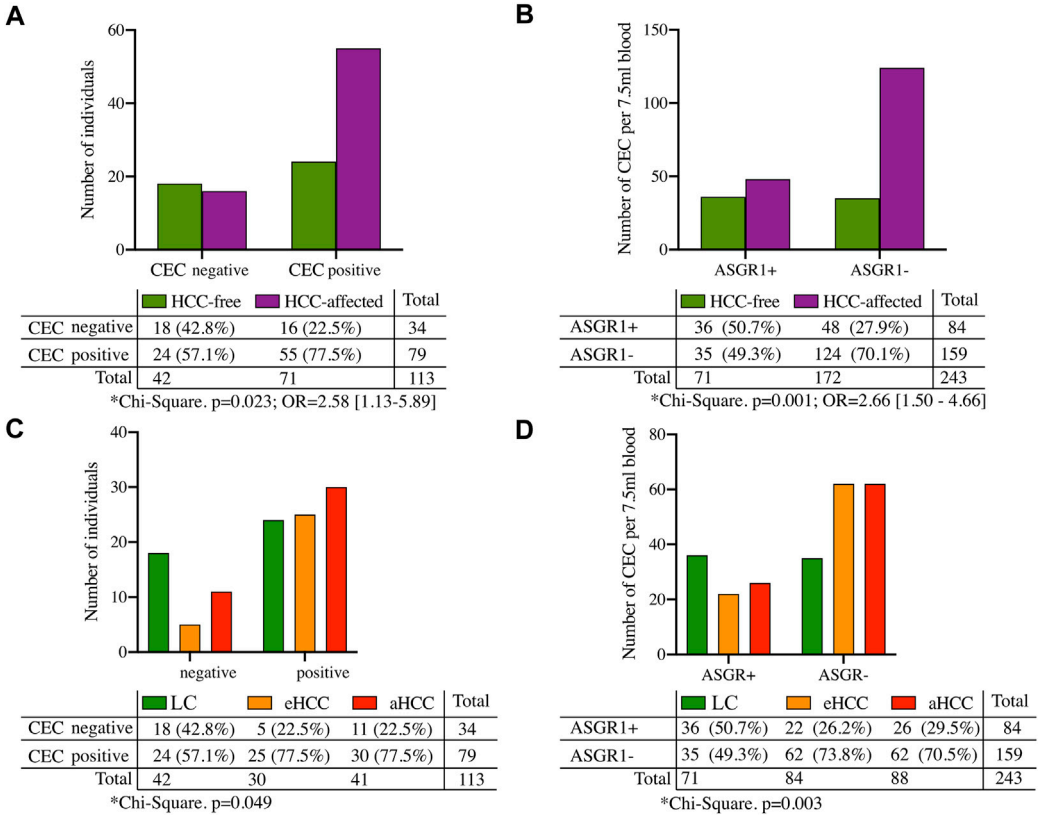



FIGURE 4
Risk factors for HCC. **(A)** Shows presence/absence of CECs and **(B)** shows number of CECs expressing (or not) ASGR1 in liver cirrhosis (LC) or hepatocellular carcinoma patients (HCC). **(C,D)** show same information as in **(A,B)** respectively, but dividing HCC individuals by either early (eHCC) or advanced (aHCC) disease stages.

0.023) (Figure 4A) and the number of CECs lacking ASGR1 expression significantly ($p$ = 0.001) increased risk of developing HCC by 2.66-fold (Figure 4B). A significantly greater proportion of aHCC and eHCC patients had CECs (including ASGR+ and ASGR-) compared with LC ($p$ = 0.049) (Figure 4C) and this increase came with a significant reduction on the number of CECs expressing ASGR1 in HCC compared to LC ($p$ = 0.003) (Figure 4D).

When pathological AFP levels were considered, 13/13 patients (100%) with AFP greater than 400 ng/ml showed HCC (11 of which had prior LC). However, there were a significant ($p$ = 0.0117) proportion of individuals with lower AFP levels (54/93; 58.1%) that also developed HCC, highlighting the poor performance of AFP for identifying high-risk individuals (Figure 5A). Interestingly, there was a significant difference between ASGR1 expression when analyzing liver cirrhosis and cancer occurrence together ($p$ = 0.033). Thus, of the total of individuals with CECs negative for ASGR1 expression, the majority were HCC patients with prior LC (65.6%; 21/32), while there was no difference between cancer-free LC (18.8%; 6/32) and HCC without prior LC (5/32; 15.6%). Contrarily, positive expression of ASGR in CECs was more frequent in LC (42.9%; 18/42) than in HCC without prior LC (22.2%; 2/9), suggesting that LC patients (independently on their cancer status) were more likely to have ASGR1 positive CECs (Figure 5B). In fact, HCC patients without previous LC had the greatest percentage of CECs without ASGR1 expression (median 100%) compared with HCC with LC (median 71.0%) and HCC-free LC (median 51%) patients although this difference was not significantly different ($p$ = 0.21) (Figure 5C). Furthermore, we found that the percentage of CECs without ASGR1 expression increased upon liver dysfunction (Child Pugh-Turcotte stage) (Figure 5D), although this increase was not significant ($p$ = 0.74). Presence of CECs without ASGR1 expression significantly correlated with cancer occurrence ($p$ = 0.012). Indeed, ASGR1 expression in CECs negatively correlated with INR ($p$ = 0.025).

Finally, as expected, early HCC patients showed significantly greater overall survival (OS) than advanced HCC ($p$ = 0.018) (Figure 6A). However, this difference was observed only in patients showing CECs (Figure 6B; $p$ = 0.021) and not in patients without CECs (Figure 6C; $p$ = 0.250). Furthermore, a

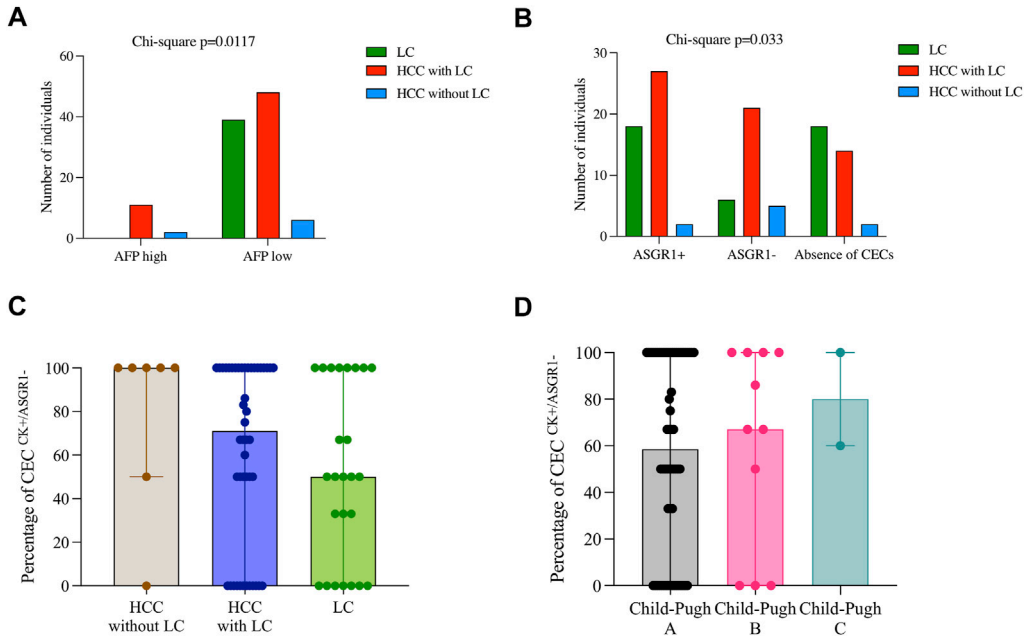



FIGURE 5
Relationship between CEC characterization and clinical and pathological features in patients with liver cirrhosis or HCC. **(A)** Shows AFP levels divided by high (>400 ng/ml) and low (<400 ng/ml) for three groups of patients: LC = liver cirrhosis; HCC of cirrhosis etiology and HCC of other a etiology. **(B)** Shows frequency distribution of CECs for the same groups of patients. **(C,D)** Represent percentages of CECs with lack of ASGR1 expression for either the three previous group of patients **(C)** or Child-Pugh scores **(D)**. *p* values for Chi-Square tests are shown for graphs **(A,B)**. No significant differences are shown in C and D for Kruskal Wallis tests.

significantly different progression-free survival (PFS) was observed considering the two CEC phenotypes. While no significant differences were observed for CECs positive for ASGR1 (Figure 6D; $p = 0.126$, $p = 0.099$ for LC compared to eHCC or aHCC, respectively), significant PFS were observed for CECs negative for ASGR1 expression (Figure 6E; $p < 0.0001$, $p = 0.002$ for LC compared to eHCC or aHCC, respectively). Thus, PFS was significantly lower in aHCC and eHCC patients compared to LC when considering the CEC phenotype (Figure 6F; $p < 0.0001$).

## 4 Discussion

In this study we have demonstrated the suitability of two liver-specific biomarkers (ASGR1 and miR-122-5p) for the characterization of circulating epithelial cells (CECs) from liver cirrhosis (LC) and hepatocellular carcinoma (HCC) patients as prognostic biomarkers. This proof-of-concept study demonstrates that phenotypic characterization of CECs with either ASGR1 or miR-122-5p may allow risk-stratification in patients at earlier disease stages, which is the principal basis of precision medicine.

Currently there is a lack of diagnostic and risk-stratification tools for HCC. Evaluation of circulating tumor cells (CTCs) in advanced stages is a prognostic marker for high risk of metastasis (Wang et al., 2018) and presence of CTCs after surgery or resection is a prognostic marker for PFS in HCC patients (Cui et al., 2020); however, there are some limitations of CTC evaluation, particularly in early and pre-tumoral disease stages. Lu-Nan, Qi et al. (2018) identified CTCs in more than half of the early-stage (BCLC 0-A) HCC patients suggesting it may be a remarkably useful tool for cancer interception (Serrano et al., 2020). However, there are currently only very few studies addressing identification and characterization of CECs in pre-tumoral diseases such as LC (Chen et al., 2020).

Expression of ASGR1 preferentially in the sinusoidal and basolateral hepatocellular membranes makes this protein an important biomarker for HCC. In fact, it has been explored in the context of targeted therapies for HCC (Chen et al., 2017), (Kim et al., 2019), (Nair et al., 2019). Other roles of ASGR1 such as hepatitis C virus binding allowing viral infection (Saunier et al., 2003) or metastasis promotion by interaction with lectins in the tumor microenvironment through the EGFR-ERK pathway (Ueno et al., 2011) are also described in relation with HCC induction. Interestingly, hepatocytes expressing low levels of ASGR1 were characterized as progenitor-like cells with higher levels of EGFR, β1 and α6 integrins expression (Ise et al., 2004). Furthermore, variable ASGR1 expression was observed between tumor stages. For instance, Shi et al. (2013) demonstrated that

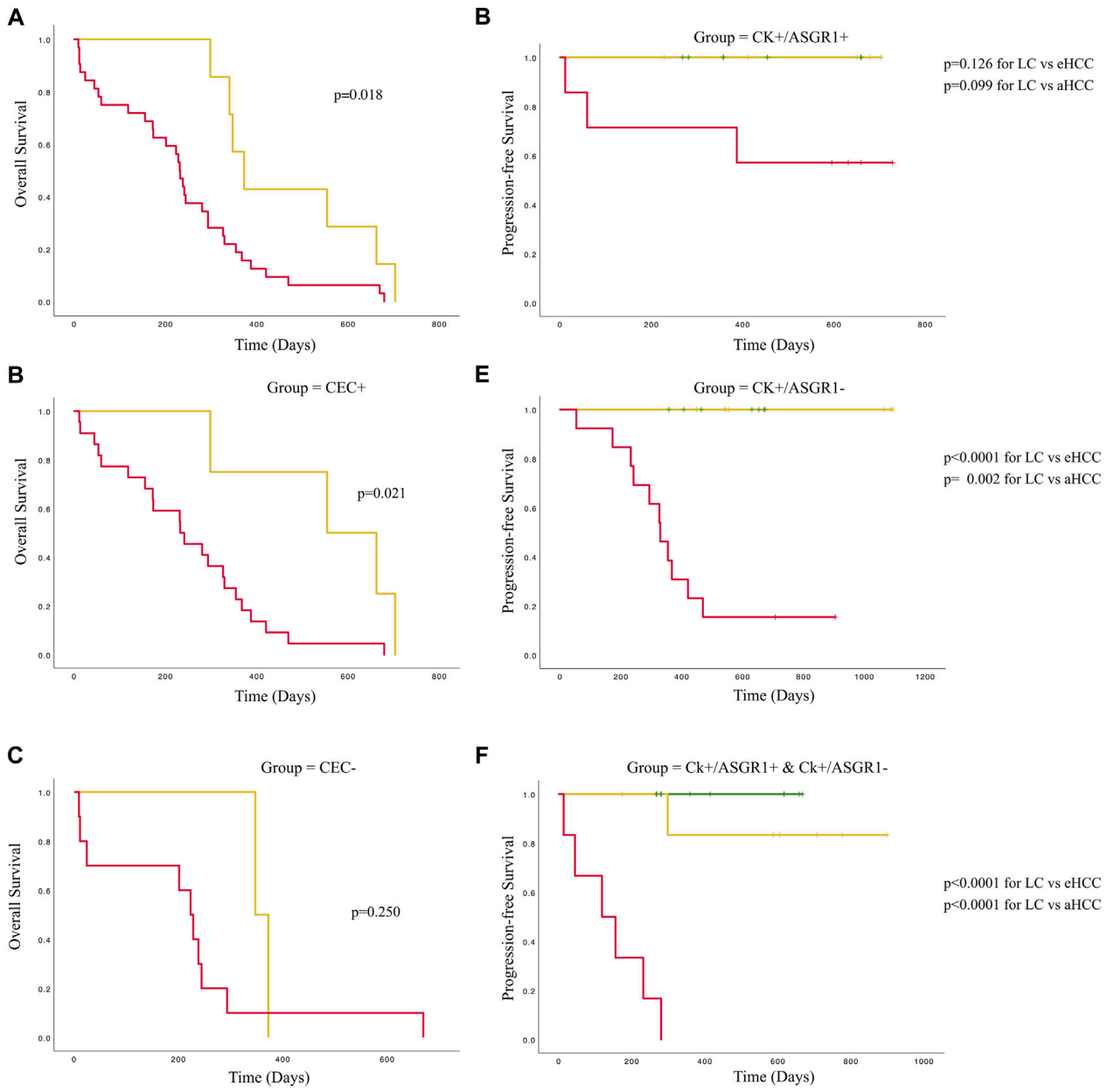



FIGURE 6
Survival curves for patients with liver cirrhosis or HCC accounting for CEC numbers and phenotype. Left part of the figure represents overall survival curves for the whole population **(A)**, only for those positive for CECs **(B)** and for those negative for CECs **(C)**. Colors are Grey = eHCC and Black: aHCC. Right part of the figure represents progression-free survival for patients with all CECs positive for ASGR1 **(D)**, patients with all CECs negative for ASGR1 **(E)** and patients with heterogeneous CEC phenotypes **(F)**. Colors are: Dotted Grey = liver cirrhosis; Grey eHCC and Black: aHCC.

expression of ASGR1 (based on H-scores) in normal adjacent tissue was comparable to that on hepatic cirrhosis and early HCC (grade I or well differentiated) using tissue microarrays, although its expression decreased significantly with increasing tumor stage. This suggested that ASGR1 could potentially be used as an early indicator of the disease status. Likewise, Witzigmann et al. (2016) showed lower levels of ASGR1 mRNA in HCC compared to its adjacent normal tissue, as well as a reduction of mRNA according to increasing HCC stages. Furthermore, a decreased ASGR1 mRNA expression was observed in metastatic or highly proliferative tumors (defined by Ki69 positivity) (Witzigmann et al., 2016) and overexpression

of ASGR1 was shown to inhibit cell migration and invasivity potential both *in vitro* and *in vivo* (Gu et al., 2016), suppressing metastasis and serving a prognostic biomarker.

Together with ASGR, miR-122-5p is another liver-specific biomarker which tumor suppressor role has also been suggested in HCC (Gramantieri et al., 2007; Jin et al., 2017). In fact, miR-122-5p is one of the best diagnostic markers for HCC being usually elevated in circulation of HCC patients (Jin et al., 2019); however, its overexpression in circulation was not correlated with that in tissue, which levels were found to be downregulated in HCC (Gramantieri et al., 2007). Intra-tumor heterogeneity following microenvironment stress cannot be characterized by cytokeratin positive CEC counts alone. Thus, phenotypic characterization with tissue-specific biomarkers, such as ASGR1 or miR-122-5p, may be of great importance. In fact, we identified distinct CEC subpopulations based on ASGR1/miR-122-5p staining within and between patients. Interestingly, a perfect positive correlation was found between the two markers suggesting that both could become excellent prognostic biomarkers for HCC.

Our results show that both the presence of CECs and the absence of ASGR1 expression in CECs are risk factors of developing HCC. Furthermore, ASGR1 staining has been shown to be positively correlated with miR-122-5p expression, suggesting that CECs expressing both markers originated from hepatocytes; contrarily, absence of the two biomarkers suggests a more dedifferentiated state, potentially identifying more aggressive phenotypes. These data agree with those presented by Shi et al. (2013) and Witzigmann et al. (2016), who have shown decreased ASGR1 expression levels in HCC tumor tissues, both at the mRNA and protein level as well as works from Hu et al. (2012) who showed a decreased miR-122-5p expression in HCC compared to controls.

Despite its main expression in liver, ASGR1 was also shown to be expressed in other cell types such as colon (Fang et al., 2009) or peripheral blood mononuclear cells (PBMCs) (Harris et al., 2012), albeit at lower levels compared with hepatocytes. We have also observed some level of ASGR1 expression in PBMCs of some individuals (data not shown); however, any significance of such differences among patients needs to be further investigated. One of the described roles of ASGR1 is the clearance of platelets and other prothrombotic blood components (Sørensen et al., 2009). Therefore, its presence in PBMCs might arise because of this blood hemostasis process. In fact, we showed significant correlations between ASGR1 expression in CECs and INR ($p$ = 0.025) as well as prothrombin activity ($p$ = 0.006), validating the role of ASGR1 maintaining blood homeostasis. Furthermore, the percentage of ASGR1 expression in CECs decreased with Child Pugh-Turcotte stage suggesting it may be used as diagnostic and prognostic biomarker in earlier disease stages as an outstanding cancer interception tool (Serrano et al., 2020). Thus, our data highlights the clinical utility of characterizing CECs using ASGR1/miR-122-5p isolated from patients with chronic liver cirrhosis or HCC to identify potentially more aggressive phenotypes (loss of ASGR1/miR-122-5p) serving as prognostic tool. A larger clinical trial in which clinical utility of these two biomarkers is assessed, might allow to update the current clinical management guidelines for HCC. Incorporation of liquid biopsy based on CTC isolation and characterization using ASGR1 and/or miR-122-5p might become a non-invasive strategy for risk stratification, improving early identification of malignancy in liver cirrhotic patients.

We understand that the cohort size might impact on our ability to reach good diagnostic power for CECs and ASGR1 as biomarkers, although similar (Qi et al., 2018) and even smaller cohort sizes (Vona et al., 2004; Wang et al., 2018; Takahashi et al., 2021) have already been used by other authors. Also, we acknowledge that detection of miR-122-5p on CECs should have been performed in all samples positive for CECs; however due to funding constraints and difficulties of the methodology, we only assessed as a proof-of-concept this biomarker in a small cohort of patients. We are currently working to increase cohort size and to characterize CECs using a more comprehensive biomarker panel including ASGR1 and miR-122-5p simultaneously using our recently developed protocol (Ruiz-Rodríguez et al., 2021) to improve statistical power and to demonstrate the diagnostic utility of characterizing CECs for LC and HCC patients.

## 5 Conclusion

Circulating epithelial cells (CECs) may prove to become a very useful prognostic biomarker for the identification of individuals at risk of developing hepatocellular carcinoma (HCC). Both the presence of CECs and the lack of ASGR1/miR-122-5p expression in CECs were linked with HCC incidence and poorer disease outcomes, highlighting their potential as predictive biomarkers.

## Data availability statement

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

## Ethics statement

The studies involving human participants were reviewed and approved by PEIBA: Portal de Etica de la Investigación Biomédica de Andalucia. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

Conceiving and designing the study: MJS, MG-N, and JL; Analyzing and interpreting the data: AR-C, PM-V, MG-N, and VA-G; Writing the manuscript: MG-N, MJS, AR-C, AR, and MG; Providing critical revisions: JL, MJS, JS, JM, JP, and JA; Approving the final version of the manuscript: MG-N, MJS, and MR-G.

## Funding

The research leading to these results received funding from the Regional Ministry of Health of Andalusia (codes: PC-0522-2016, PC-0267-2017 and PC-0033-2017), principal investigators: MJS, JL, and AR-C respectively. Carmen Garrido's postdoctoral fellowship is funded by the Ministry of Economy, Competitiveness, Enterprises and Universities (DOC_01682); AR-C postdoctoral contract is supported by a Sara Borrell postdoctoral fellowship from Health Institute Carlos III (ISCIII) (CD18/00126).

## Acknowledgments

We would like to acknowledge all patients willing to contribute to this project for their consent and their blood samples and to Cas Kramer for his English proofreading of this manuscript.

## Conflict of interest

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

## Publisher's note

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/fmolb.2022.1074277/full#supplementary-material

## References

Bayarri-Lara, C., Ortega, F. G., Cueto Ladrón de Guevara, A., Puche, J. L., Ruiz Zafra, J., de Miguel-Perez, D., et al. (2016). Circulating tumor cells identify early recurrence in patients with non-small cell lung cancer undergoing radical resection. *PLoS One* 11, e0148659. doi:10.1371/journal.pone.0148659

Bialecki, E. S., and di Bisceglie, A. M. (2005). Diagnosis of hepatocellular carcinoma. *HPB* 7, 26–34. doi:10.1080/13651820410024049

Cadier, B., Bulsei, J., Nahon, P., Seror, O., Laurent, A., Rosa, I., et al. (2017). Early detection and curative treatment of hepatocellular carcinoma: A cost-effectiveness analysis in France and in the United States. *Hepatology* 65, 1237–1248. doi:10.1002/hep.28961

Cancer Tomorrow (2022). Cancer Tomorrow. Available at: https://gco.iarc.fr/tomorrow/en/dataviz/isotype?cancers=11&single_unit=50000&group_cancers=0&multiple_cancers=0. [Accessed 22 Sep 2022].

Carr, B. I., Akkiz, H., Üsküdar, O., Yalcin, K., Guerra, V., Kuran, S., et al. (2018). HCC with low- and normal-serum alpha-fetoprotein levels. *Clin. Pract.* 15, 453–464. doi:10.4172/clinical-practice.1000393

Chen, C., Li, K., Jiang, H., Song, F., Gao, H., Pan, X., et al. (2017). Development of T cells carrying two complementary chimeric antigen receptors against glypican-3 and asialoglycoprotein receptor 1 for the treatment of hepatocellular carcinoma. *Cancer Immunol. Immunother.* 66, 475–489. doi:10.1007/s00262-016-1949-8

Chen, Z., Lin, X., Chen, C., Chen, Y., Zhao, Q., Wu, L., et al. (2020). Analysis of preoperative circulating tumor cells for recurrence in patients with hepatocellular carcinoma after liver transplantation. *Ann. Transl. Med.* 8, 1067. doi:10.21037/atm-20-2751

Court, C. M., Hou, S., Winograd, P., Segel, N. H., Li, Q. W., Zhu, Y., et al. (2018). A novel multimarker assay for the phenotypic profiling of circulating tumor cells in hepatocellular carcinoma. *Liver Transpl.* 24, 946–960. doi:10.1002/lt.25062

Cui, K., Ou, Y., Shen, Y., Li, S., and Sun, Z. (2020). Clinical value of circulating tumor cells for the diagnosis and prognosis of hepatocellular carcinoma (HCC): A systematic review and meta-analysis. *Medicine* 99, e22242. doi:10.1097/MD.0000000000022242

de Miguel-Pérez, D., Bayarri-Lara, C. S., Ortega, F. G., Rodriguez, M., Alvarez-Cubero, M. J., Serrano, E. M., et al. (2019). Post-surgery circulating tumor cells and AXL overexpression as new poor prognostic biomarkers in resected lung adenocarcinoma. *Cancers (Basel)* 11, 1750. doi:10.3390/cancers11111750

Delgado-Ureña, M., Ortega, F. G., de Miguel-Pérez, D., Rodriguez-Martinez, A., Garcia-Puche, J. L., Ilyine, H., et al. (2018). Circulating tumor cells criteria (CyCAR) versus standard RECIST criteria for treatment response assessment in metastatic colorectal cancer patients. *J. Transl. Med.* 16, 251. doi:10.1186/s12967-018-1624-2

Fan, J-L., Yang, Y-F., Yuan, C-H., Chen, H., and Wang, F-B. (2015). Circulating tumor cells for predicting the prognostic of patients with hepatocellular carcinoma: A meta analysis. *Cell. Physiol. biochem.* 37, 629–640. doi:10.1159/000430382

Fang, J., Izawa, R., Gomez-Santos, L., Ueno, S., Sawaguchi, T., Usami, K., et al. (2009). Potentiation of proliferation of some but not all human colon carcinoma cell lines by immobilized hepatic asialoglycoprotein receptor 1. *Oncol. Res.* 17, 437–445. doi:10.3727/096504009789735440

Fateen, W., and Ryder, S. (2017). Screening for hepatocellular carcinoma: Patient selection and perspectives. *J. Hepatocell. Carcinoma* 4, 71–79. doi:10.2147/JHC.S105777

Force, M., Park, G., Chalikonda, D., Roth, C., Cohen, M., Halegoua-Demarzio, D., et al. (2022). Alpha-fetoprotein (AFP) and AFP-L3 is most useful in detection of recurrence of hepatocellular carcinoma in patients after tumor ablation and with low AFP level. *Viruses* 14, 775. doi:10.3390/v14040775

Forner, A., Reig, M., and Bruix, J. (2018). Hepatocellular carcinoma. *Lancet* 391, 1301–1314. doi:10.1016/S0140-6736(18)30010-2

Gramantieri, L., Ferracin, M., Fornari, F., Veronese, A., Sabbioni, S., Liu, C. G., et al. (2007). Cyclin G1 is a target of miR-122a, a MicroRNA frequently down-regulated in human hepatocellular carcinoma. *Cancer Res.* 67, 6092–6099. doi:10.1158/0008-5472.CAN-06-4607

Gu, D., Jin, H., Jin, G., Wang, C., Wang, N., Hu, F., et al. (2016). The asialoglycoprotein receptor suppresses the metastasis of hepatocellular carcinoma via LASS2-mediated inhibition of V-ATPase activity. *Cancer Lett.* 379, 107–116. doi:10.1016/j.canlet.2016.05.030

Hanif, H., Ali, M. J., Khan, I. W., Luna-Cuadros, M. A., Khan, M. M., Tan-Yeung Lau, D., et al. (2022). Update on the applications and limitations of alpha-fetoprotein for hepatocellular carcinoma. *World J. Gastroenterol.* 28, 216–229. doi:10.3748/wjg.v28.i2.216

Harris, R. L., van den Berg, C. W., and Bowen, D. J. (2012). ASGR1 and ASGR2 , the genes that encode the asialoglycoprotein receptor (ashwell receptor), are expressed in peripheral blood monocytes and show interindividual differences in transcript profile. *Mol. Biol. Int.* 1–10, 283974. doi:10.1155/2012/283974

Hu, J., Xu, Y., Hao, J., Wang, S., Li, C., and Meng, S. (2012). MiR-122 in hepatic function and liver diseases. *Protein Cell* 3, 364–371. doi:10.1007/s13238-012-2036-3

Ise, H., Nikaido, T., Negishi, N., Sugihara, N., Suzuki, F., Akaike, T., et al. (2004). Effective hepatocyte transplantation using rat hepatocytes with low asialoglycoprotein receptor expression. *Am. J. Pathol.* 165, 501–510. doi:10.1016/S0002-9440(10)63315-9

Jin, Y., Wang, J., Han, J., Luo, D., and Sun, Z. (2017). MiR-122 inhibits epithelial-mesenchymal transition in hepatocellular carcinoma by targeting Snail1 and Snail2 and suppressing WNT/β-cadherin signaling pathway. *Exp. Cell Res.* 360, 210–217. doi:10.1016/j.yexcr.2017.09.010

Jin, Y., Wong, Y. S., Goh, B. K. P., Chan, C. Y., Cheow, P. C., Chow, P. K. H., et al. (2019). Circulating microRNAs as potential diagnostic and prognostic biomarkers in hepatocellular carcinoma. *Sci. Rep.* 9, 10464. doi:10.1038/s41598-019-46872-8

Kim, Y., Jo, M., Schmidt, J., Luo, X., Prakash, T. P., Zhou, T., et al. (2019). Enhanced potency of GalNAc-conjugated antisense oligonucleotides in hepatocellular cancer models. *Mol. Ther.* 27, 1547–1557. doi:10.1016/j.ymthe.2019.06.009

Kobeisy, M. A., Morsy, K. H., Galal, M., Sayed, S. K., Ashmawy, M. M., and Mohammad, F. M. (2012). Clinical significance of elevated alpha-foetoprotein (AFP) in patients with chronic hepatitis C without hepatocellular carcinoma in upper Egypt. *Arab. J. Gastroenterol.* 13, 49–53. doi:10.1016/j.ajg.2012.06.004

Lee, C. W., Tsai, H. I., Lee, W. C., Huang, S. W., Lin, C. Y., Hsieh, Y. C., et al. (2019). Normal alpha-fetoprotein hepatocellular carcinoma: Are they really normal? *J. Clin. Med.* 8, 1736. doi:10.3390/jcm8101736

Li, J., Chen, L., Zhang, X., Zhang, Y., Liu, H., Sun, B., et al. (2014). Detection of circulating tumor cells in hepatocellular carcinoma using antibodies against asialoglycoprotein receptor, carbamoyl phosphate synthetase 1 and pan-cytokeratin. *PLoS One* 9, e96185. doi:10.1371/journal.pone.0096185

Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular carcinoma. *Nat. Rev. Dis. Prim.* 7, 6. doi:10.1038/s41572-020-00240-3

Muramatsu, T. (2007). "Knockout mice and glycoproteins," in *Comprehensive glycoscience* (Cham: Elsevier), 121–147.

Nadal, R., Ortega, F. G., Salido, M., Lorente, J. A., Rodriguez-Rivera, M., Delgado-Rodriguez, M., et al. (2013). CD133 expression in circulating tumor cells from breast cancer patients: Potential role in resistance to chemotherapy. *Int. J. Cancer* 133, 2398–2407. doi:10.1002/ijc.28263

Nair, A. B., Shah, J., Al-Dhubiab, B. E., Patel, S. S., Morsy, M. A., Patel, V., et al. (2019). Development of asialoglycoprotein receptor-targeted nanoparticles for selective delivery of gemcitabine to hepatocellular carcinoma. *Molecules* 24, E4566. doi:10.3390/molecules24244566

Qi, L-N., Xiang, B-D., Wu, F-X., Ye, J. Z., Zhong, J. H., Wang, Y. Y., et al. (2018). Circulating tumor cells undergoing EMT provide a metric for diagnosis and prognosis of patients with hepatocellular carcinoma. *Cancer Res.* 78, 4731–4744. doi:10.1158/0008-5472.CAN-17-2459

Reig, M., Forner, A., Ávila, M. A., Ayuso, C., Minguez, B., Varela, M., et al. (2021). Diagnosis and treatment of hepatocellular carcinoma. Update of the consensus document of the AEEH, AEC, SEOM, SERAM, SERVEI, and SETH. *Med. Clin.* 156, 463.e1–463463.e30. doi:10.1016/j.medcli.2020.09.022

Romero-Palacios, P. J., Alcázar-Navarrete, B., Díaz Mochón, J. J., de Miguel-Pérez, D., López Hidalgo, J. L., Garrido-Navas, M. del C., et al. (2019). Liquid biopsy beyond of cancer: Circulating pulmonary cells as biomarkers of COPD aggressivity. *Crit. Rev. Oncol. Hematol.* 136, 31–36. doi:10.1016/j.critrevonc.2019.02.003

Ruiz-Rodríguez, A. J., Molina-Vallejo, M. P., Aznar-Peralta, I., Gonzalez Puga, C., Canas Garcia, I., Gonzalez, E., et al. (2021). Deep phenotypic characterisation of CTCs by combination of microfluidic isolation (IsoFlux) and imaging flow cytometry (ImageStream). *Cancers* 13 (24), 6386. doi:10.3390/cancers13246386

Saunier, B., Triyatni, M., Ulianich, L., Maruvada, P., Yen, P., and Kohn, L. D. (2003). Role of the asialoglycoprotein receptor in binding and entry of hepatitis C virus structural proteins in cultured human hepatocytes. *J. Virol.* 77, 546–559. doi:10.1128/jvi.77.1.546-559.2003

Serrano, M. J., Garrido-Navas, M. C., Diaz Mochon, J. J., Cristofanilli, M., Gil-Bazo, I., Pauwels, P., et al. (2020). Precision prevention and cancer interception: The new challenges of liquid biopsy. *Cancer Discov.* 10, 1635–1644. doi:10.1158/2159-8290.CD-20-0466

Shi, B., Abrams, M., and Sepp-Lorenzino, L. (2013). Expression of asialoglycoprotein receptor 1 in human hepatocellular carcinoma. *J. Histochem. Cytochem.* 61, 901–909. doi:10.1369/0022155413503662

Shimada, S., Mogushi, K., Akiyama, Y., Furuyama, T., Watanabe, S., Ogura, T., et al. (2019). Comprehensive molecular and immunological characterization of hepatocellular carcinoma. *EBioMedicine* 40, 457–470. doi:10.1016/j.ebiom.2018.12.058

Sørensen, A. L., Rumjantseva, V., Nayeb-Hashemi, S., Clausen, H., Hartwig, J. H., Wandall, H. H., et al. (2009). Role of sialic acid for platelet life span: Exposure of β-galactose results in the rapid clearance of platelets from the circulation by asialoglycoprotein receptor-expressing liver macrophages and hepatocytes. *Blood* 114, 1645–1654. doi:10.1182/blood-2009-01-199414

Sun, C., Liao, W., Deng, Z., Li, E., Feng, Q., Lei, J., et al. (2017). The diagnostic value of assays for circulating tumor cells in hepatocellular carcinoma: A meta-analysis. *Medicine* 96, e7513. doi:10.1097/MD.0000000000007513

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

Takahashi, K., Ofuji, K., Hiramatsu, K., Nosaka, T., Naito, T., Matsuda, H., et al. (2021). Circulating tumor cells detected with a microcavity array predict clinical outcome in hepatocellular carcinoma. *Cancer Med.* 10, 2300–2309. doi:10.1002/cam4.3790

Tsuchiya, A., Suda, T., Oda, C., Kimura, A., Hosaka, K., Kimura, N., et al. (2019). EpCAM- and/or NCAM-expressing hepatocellular carcinoma in which behavior of hepatic progenitor cell marker-positive cells are followed. *Case Rep. Gastroenterol.* 13, 118–124. doi:10.1159/000498913

Ueno, S., Mojic, M., Ohashi, Y., Higashi, N., Hayakawa, Y., and Irimura, T. (2011). Asialoglycoprotein receptor promotes cancer metastasis by activating the EGFR–ERK pathway. *Cancer Res.* 71, 6419–6427. doi:10.1158/0008-5472.CAN-11-1773

Vona, G., Estepa, L., Béroud, C., Damotte, D., Capron, F., Nalpas, B., et al. (2004). Impact of cytomorphological detection of circulating tumor cells in patients with liver cancer. *Hepatology* 39, 792–797. doi:10.1002/hep.20091

Wang, J. Q., Li, L. L., Hu, A., Deng, G., Wei, J., Li, Y. F., et al. (2022). Inhibition of ASGR1 decreases lipid levels by promoting cholesterol excretion. *Nature* 608 (608), 7922413–7922420. doi:10.1038/s41586-022-05006-3

Wang, S., Zheng, Y., Liu, J., Huo, F., and Zhou, J. (2018). Analysis of circulating tumor cells in patients with hepatocellular carcinoma recurrence following liver transplantation. *J. Investig. Med.* 66, 1–6. doi:10.1136/jim-2017-000655

Wang, T., and Zhang, K-H. (2020). New blood biomarkers for the diagnosis of AFP-negative hepatocellular carcinoma. *Front. Oncol.* 10, 1316. doi:10.3389/fonc.2020.01316

Wang, X., Zhang, P., and Deng, K. (2022). MYC promotes LDHA expression through MicroRNA-122-5p to potentiate glycolysis in hepatocellular carcinoma. *Anal. Cell. Pathol.* 2022, 1435173–1435217. doi:10.1155/2022/1435173

Witzigmann, D., Quagliata, L., Schenk, S. H., Quintavalle, C., Terracciano, L. M., and Huwyler, J. (2016). Variable asialoglycoprotein receptor 1 expression in liver disease: Implications for therapeutic intervention. *Hepatol. Res.* 46, 686–696. doi:10.1111/hepr.12599

Xie, B., Shi, X., Li, Y., Xia, B., Zhou, J., Du, M., et al. (2021). Deficiency of ASGR1 in pigs recapitulates reduced risk factor for cardiovascular disease in humans. *PLoS Genet.* 17, e1009891. doi:10.1371/journal.pgen.1009891

Ye, Z., Wang, C., Wan, S., Mu, Z., Zhang, Z., Abu-Khalaf, M. M., et al. (2019). Association of clinical outcomes in metastatic breast cancer patients with circulating tumour cell and circulating cell-free DNA. *Eur. J. Cancer* 106, 133–143. doi:10.1016/j.ejca.2018.10.012

Zhu, X., Song, G., Zhang, S., Chen, J., Zhu, H., Jia, X., et al. (2022). Asialoglycoprotein receptor 1 functions as a tumor suppressor in liver cancer via inhibition of STAT3. *Cancer Res.* 82, 3987–4000. doi:10.1158/0008-5472.CAN-21-4337

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fmolb.2022.1022056



OPEN ACCESS

EDITED BY
Fernando Schmitt,
University of Porto, Portugal

REVIEWED BY
Yao Lin,
Fujian University of Traditional Chinese Medicine, China
Raheleh Roudi,
Stanford University, United States

*CORRESPONDENCE
Junzhang Zhao,
zhaojzh@mail.sysu.edu.cn
Zunnan Huang,
zn_huang@gdmu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 18 August 2022
ACCEPTED 28 November 2022
PUBLISHED 04 January 2023

CITATION
Hong X, Zhuang K, Xu N, Wang J, Liu Y, Tang S, Zhao J and Huang Z (2023), An integrated analysis of prognostic mRNA signature in early- and progressive-stage gastric adenocarcinoma. *Front. Mol. Biosci.* 9:1022056. doi: 10.3389/fmolb.2022.1022056



# An integrated analysis of prognostic mRNA signature in early- and progressive-stage gastric adenocarcinoma

Xiaoling Hong[1,2,3,4,5†], Kai Zhuang[2,3,4,6†], Na Xu[2,3,4], Jiang Wang[7], Yong Liu[2,3,4], Siqi Tang[2,3,4,5], Junzhang Zhao[8,9*] and Zunnan Huang[1,2,3,4,10*]

[1]The First Dongguan Affiliated Hospital, Guangdong Medical University, Dongguan, China, [2]Key Laboratory of Big Data Mining and Precision Drug Design, Guangdong Medical University, Dongguan, China, [3]Key Laboratory of Computer-Aided Drug Design of Dongguan City, Guangdong Medical University, Dongguan, China, [4]Key Laboratory for Research and Development of Natural Drugs of Guangdong Province, Guangdong Medical University, Dongguan, China, [5]The Second School of Clinical Medicine, Guangdong Medical University, Zhanjiang, China, [6]School of Public Health, Guangdong Medical University, Dongguan, China, [7]School of Biomedical Engineering, Guangdong Medical University, Zhanjiang, China, [8]Department of Gastroenterology, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangdong, Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, National Key Clinical Discipline, Guangzhou, China, [9]Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, Guangzhou, China, [10]Marine Medical Research Institute of Guangdong Zhanjiang, Zhanjiang, China

The pathogenesis and vital factors of early and progressive stages of stomach adenocarcinoma (STAD) have not been fully elucidated. In order to discover novel and potential targets to guide effective treatment strategies, a comprehensive bioinformatics study was performed, and the representative results were then validated by quantitative polymerase chain reaction (qPCR) and immunohistochemical (IMC) staining in clinical samples. A total of 4,627, 4,715, and 3,465 differentially expressed genes (DEGs) from overall-, early-, and progressive-stage STAD were identified, respectively. Prognostic models of 5-year OS were established for overall-, early-, and progressive-stage STAD, and ROC curves demonstrated AUC values for each model were 0.73, 0.87, and 0.92, respectively. Function analysis revealed that mRNAs of early-stage STAD were enriched in chemical stimulus-related pathways, whereas remarkable enrichment of mRNAs in progressive-stage STAD mainly lay in immune-related pathways. Both qPCR and IHC data confirmed the up-regulation of IGFBP1 in the early-stage and CHAF1A in progressive-stage STAD compared with their matched normal tissues, indicating that these two representative targets could be used to predict the prognostic status of the patients in these two distinct STAD stages, respectively. In addition, seven mRNAs (F2, GRID2, TF, APOB, KIF18B, INCENP, and GCG) could be potential novel biomarkers for STAD at different stages from this study. These results contributed to identifying STAD patients at high-risk, thus guiding targeted treatment with efficacy in these patients.

KEYWORDS

**STAD, prognostic model, mRNA, diagnosis, novel biomarkers**

# 1 Introduction

Gastric cancer (GC) is one of the most common malignancies worldwide. Although many advances of systemic treatment in GC had been made over the past decades, it remained a concern that the majority of patients showed strong resistance to many treatment strategies (Rihawi et al., 2021), which resulted in a great health burden.

In GC, stomach adenocarcinoma (STAD) was the most common histological type (about 95%), and clinical guidelines had addressed differences in STAD therapeutic strategies and outcomes within different clinicopathological stages (Ajani et al., 2017). Optimally, patients with early-stage STAD undergo limited resection through endoscopies, while patients with advanced STAD require surgeries and multidisciplinary adjuvant treatments (Ajani et al., 2017). The 5-year survival rate for early-stage STAD (according to TNM malignancy classification) is 95%; however, the median survival time for patients with advanced-stage STAD was only 9–10 months (Ajani et al., 2016). Therefore, spotting high-risk STAD patients and choosing the appropriate treatment at the early time was crucial for prolonging survival time in these patients. Emerging evidence had revealed that biomarkers contributed to molecular classification, predicting prognosis, and driving precision therapy approaches in STAD population (Joshi and Badgwell, 2021). For example, Jiang et al. found that ITGB1-DT was apparently up-regulated in STAD tissues and was connected with the T stage, therapeutic effect, and poor prognosis of STAD patients, while suppression of ITGB1-DT could inhibit cell proliferation, invasion, and migration of STAD cells (Jiang et al., 2022). Furthermore, bioinformatics analysis can be used to screen key immune-related genes (IRGs) and pathways significantly linked to STAD therapy. For example, Xia et al. constructed an immune-related risk signature model consisting of BMP8A, MMP12, NRG4, S100A9, and TUBB3, which were associated with prognosis in patients and could be the potential biomarkers for immunotherapy in STAD (Xia et al., 2022). Nevertheless, the pathogenesis and vital factors of early- and progressive-stage STAD had not been fully highlighted; it is of importance to identify novel and promising targets and Cox model, elucidate the mechanism of STAD, and provide a candidate diagnosis option in patients with distinct STAD stages.

mRNA is single-stranded ribonucleic acid that carries genetic information to guide protein synthesis, and it has a central role in the pathogenesis of various cancers, including STAD. Numerous researchers underlined that mRNA had diagnostic and prognostic values in clinical practice. For example, Huang et al. demonstrated that the expression level of sirtuin-4 (SIRT4) was reduced in STAD tissues compared with normal gastric tissues and was also correlated with pathological differentiation and tumor-infiltrating depth of STAD (Huang et al., 2015). Lu et al. demonstrated that metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) was up-regulated in MGC-803 cells, and the increased MALAT1 could promote the metastasis of cancer cells, while the decreased MALAT1 could suppress the progression and proliferation of STAD (Lu et al., 2019). BicC family RNA-binding protein 1 (BICC1), which codes an RNA-binding protein, proved to be significantly correlated with grade, TNM stage, invasion depth, and even immune infiltrates in STAD (Zhao et al., 2020). Taken together, these results indicated the key roles of mRNAs as the targets helping tumor diagnosis and targeted treatment.

The extensive applications of gene chips and high-throughput sequencing technologies in cancer research had brought omics data explosion, and STAD is no exception. Integrated bioinformatics analysis of publicly available data of STAD improves the insight into the underlying molecular mechanism of tumorigenesis, and it also contributes to identifying potential tumor biomarkers and drug targets for STAD. Depending on these analyses in mRNAs, we distinguished STAD patients from distinct stages and accurately predicted their clinical outcomes by constructing prognostic models. Function annotations and pathway enrichment of mRNA signatures revealed that different STAD stages were dominated by distinct key mechanisms. Representative biomarkers in early and progressive stages were measured in clinical samples by qPCR and IHC detections. The overall bioinformatics analysis procedure is summarized in Figure 1. This study highlighted burgeoning evidence supporting some mRNAs as biomarkers for the diagnosis of patients in different STAD stages, which could form the basis of precision medicine strategies in the future.

# 2 Materials and methods

## 2.1 Data preparation and DEGs identification

mRNAs expression and clinical data of 407 STAD total samples were downloaded from The Cancer Genome Atlas (TCGA) database (http://portal.gdc.cancer.gov/), including 32 paracancer and 375 tumor samples. mRNAs expression and clinical data of early-stage (stages I and II) and progressive-stage (stages III and IV) STAD were classified as two independent datasets. The early-stage STAD samples included 21 paracancers and 164 tumor samples, while the progressive-stage STAD samples included 10 paracancers and 188 tumor samples. These two datasets were obtained from the 407 STAD total samples after the deletion of

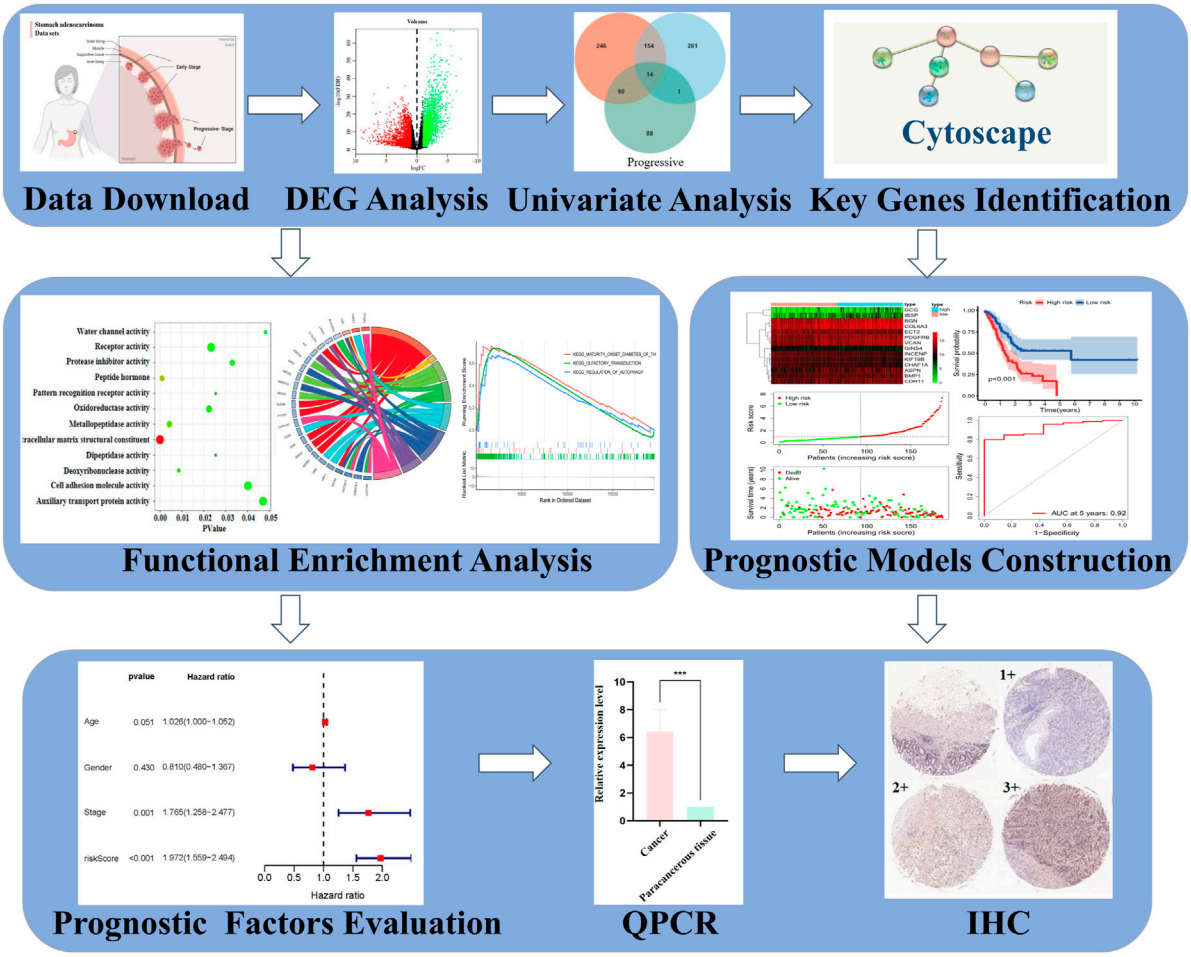



FIGURE 1
Analytic procedures of the study.

24 samples without clinical staging information. All these data of STAD samples were downloaded in March 2021. Based on the three sets of STAD sample groups: overall-, early-, and progressive-stage STAD samples; differentially expressed genes (DEGs) were identified by analyzing the expression profile with edgeR R language package (Version 4.0.2) according to the cutoff criteria of $P_{adj} < 0.05$, $|\log2FC| > 1$. Adjusted $p$-value took into account the false discovery rate (FDR), and the volcano maps of DEGs in these three datasets were generated.

## 2.2 Univariate Cox regression analysis

The survival package of R software was adopted to conduct the univariate Cox proportional hazard regression assessment of DEGs in the overall-, early-, progressive-stage STAD groups. With the criteria $p < 0.05$, overall survival-related mRNAs (mRNAs-OS) in each stage were obtained, which was related to the survival and prognosis of STAD patients.

## 2.3 PPI network construction and key mRNAs identification

To further identify the key mRNAs-OS in all three stages of STAD, the Search Tool for the Retrieval of Interacting Genes (STRING, http://string-db.org) was applied to analyze the mRNAs-OS and obtain protein interaction data. Proteins with a minimum required interaction score of 0.400 or above were selected to create a protein–protein interaction (PPI) network, in which the nodes with network interruption were hidden. The PPI network and its combined scores were then imported into Cytoscape software (Version 3.6.1, https://cytoscape.org/), and potential key mRNAs were identified using CytoHubba, a plug-in in Cytoscape software.

According to the node degree, the top 40 candidate mRNAs selected in each stage were displayed for further analysis.

## 2.4 Establishment of three Cox proportional hazards regression models

The package "surminer" of R was applied to perform multivariate Cox proportional hazard regression analysis on those identified top mRNAs-OS, and subsequently, a prognostic model consisting of mRNAs-OS related to the patients' prognosis (mRNAs-PRO) was constructed for the overall-, early-, progressive-stage STAD. Among them, mRNA-PRO with $p < 0.05$ was regarded as an independent prognostic factor of STAD. Based on the expression of mRNAs-PRO, the risk score of individual patients was computed as follows: risk score = Exp ($mRNA_1$) × $\beta_1$ + Exp ($mRNA_2$) × $\beta_2$ + Exp ($mRNA_3$) × $\beta_3$ +... + Exp ($mRNA_n$) × $\beta_n$. According to the median risk score, the STAD patients were clustered into high-risk and low-risk groups, and 5-year survival rates of the high- and low-risk patients were calculated for each prognostic model. Then, a risk score curve was drawn to distinguish the risk score differences between two groups of patients. A survival status map was drawn to reflect the survival status of each patient. A heatmap was drawn to exhibit the differences of the expression levels of the mRNAs-PRO in the high- and low-risk groups. A survival curve was drawn to display the 5-year survival rate in the high- and low-risk groups. An ROC curve showed by the area under the curve (AUC) of the model was drawn to evaluate its accuracy and reliability of predicting prognosis in each stage.

## 2.5 Functional enrichment and pathway analysis

To understand the underlying biological significance of mRNAs-OS in overall-stage STAD, GO function annotations including cellular component (CC), biological process (BP), and molecular function (MF) based on the FunRich (http://www.funrich.org/) database, and KEGG pathway enrichment based on KOBAS (http://kobas.cbi.pku.edu.cn/kobas3/) database were then analyzed. A $p$-value < 0.05 was set as the threshold to determine the crucial functions or pathways closely related to STAD in the overall stage. Furthermore, gene set enrichment analysis (GSEA) was carried out on the mRNAs of early- and progressive-stages STAD in order to explore the critical functions and pathways of the mRNA signatures in these two stages. The top three terms of BP, CC, MF, and KEGG pathways were presented. The threshold in this step was set based on net enrichment score (NES) and $p$-value. Gene sets with |NES| > 1 and $p < 0.05$ were considered to be statistically significantly enriched.

## 2.6 Diagnostic capability evaluation of prognostic models

In order to compare predictive accuracy of the age, gender, stage, and risk score for individual prognosis in overall-, early-, and progressive-stage STAD, the univariate Cox proportional hazards regression analysis was performed with the criterion of $Ps < 0.05$. Moreover, multivariate Cox proportional hazards regression analysis was applied to identify whether the age, gender, stage, and risk score could be independent prognostic factors in STAD patients. In addition, in order to investigate the prognostic values of mRNAs-PRO in different stratifications of other clinical prognostic variables, the overall-, early-, and progressive-stage STAD patients were clustered into different subgroups of age (≥65 or <65), gender, T stage, M stage, and N stage. The Kaplan–Meier survival curves were used to evaluate the prognostic capacity difference of three prognostic Cox models in STAD patients under different clinical variables.

## 2.7 Quantitative real-time polymerase chain reaction and immunohistochemical staining

A total of 30 pairs of cDNA tissue chips (including 13 early-stage pairs and 17 progressive-stage pairs), 84 pairs of tissue microarrays (including 32 early-stage pairs and 52 progressive-stage pairs), and the related clinicopathological information of these matched STAD and normal samples were obtained from Shanghai OUTDO Biotech Co., Ltd. (Shanghai). All patients were pathologically diagnosed as STAD according to American Joint Committee on Cancer (AJCC) criteria. The samples were obtained following written consent in accordance with an established protocol approved by Institutional Review Board of Biobank in Shanghai Outdo Biotech Co., Ltd.

QPCR was used to detect the expression levels of insulin-like growth factor binding protein-1 (IGFBP1) and chromatin assembly factor 1 subunit A (CHAF1A) in 30 pairs of cancer and normal samples (detailed clinical data are seen in Supplementary Table S1), and GAPDH was used as a reference gene. Total RNA was extracted from tissues and cells by the TRIzol reagent (Sigma, America). The cDNA was obtained by reverse transcription of RNA using PrimeScript™ RT Master Mix (perfect real time) (Takara, Japan) according to the protocol of manufacturer. The expressions of IGFBP1 and CHAF1A were determined by the protocol of the ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China), according to the manufacturer's protocol. The primers used in this study were as follows: IGFBP1, forward: 5′-GCATTTCTG CTCTTCCAAAG-3′, reverse: 5′-GCAACATCACCACAGGTA G-3'; and CHAF1A, forward: 5′-AAAGGAGCAGGACAGTTG GA-3′, reverse: 5′-CTGGAAGGGACTTGATTTGC-3'.

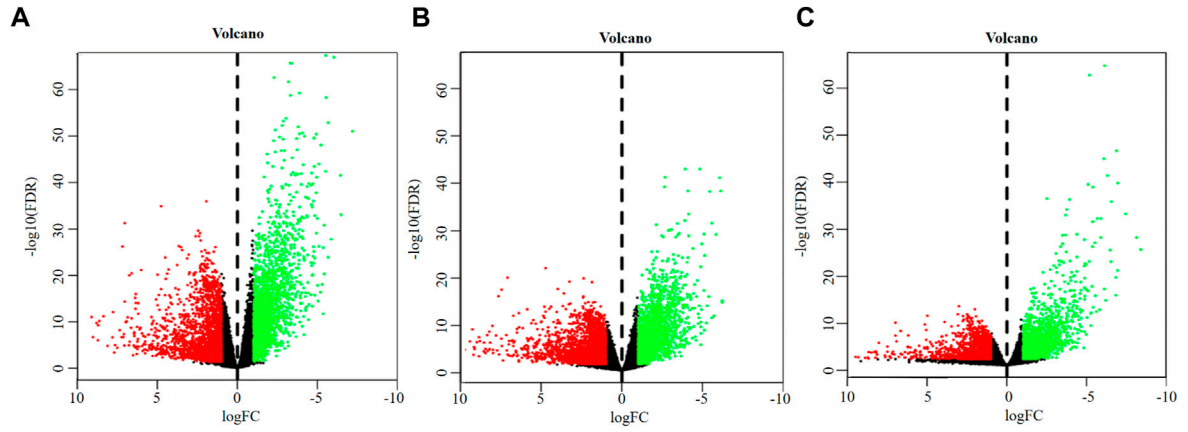



FIGURE 2
Volcano plots of DEGs. **(A)** DEGs of overall-stage STAD group. **(B)** DEGs of the early-stage STAD group. **(C)** DEGs of the progressive-stage STAD group. The abscissa denotes the $\log_2$ transformation value of the differential expression fold change between the STAD samples and the matched paracancerous samples. The ordinate denotes the $-\log_{10}$ transformation value of the $P_{adj}$ (FDR) value. Green dots symbolize significantly downregulated mRNAs. Red dots symbolize significantly upregulated mRNAs.

IHC was conducted to detect the protein levels of IGFBP1 and CHAF1A in 84 pairs of cancer and normal samples (detailed clinical data are seen in Supplementary Table S2), according to standard procedures by two independent pathologists blinded to the study. Based on proportion and staining intensity of positive stained cells, expression levels of IGFBP1 and CHAF1A were accessed semi-quantitatively (Proteintech, China). Proportion was evaluated using semi-quantitative criterion: 0, (no staining); 1, minimal (< 10%); 2, moderate (10–50%); and 3, diffuse (> 50%) staining cells. Staining intensity was also scored as 0 (negative); +1 (weak); +2 (moderate); and +3 (strong). Taken together, the final expression score of each case, considering both proportion and staining intensity, was given as 0+ (0), negative; 1+ (1 or 2), weakly positive; 2+ (3 or 4), moderately positive; and 3+ (5 or 6), strongly positive. The statistical analysis was performed using a software package (SPSS, version 19.0, Chicago, IL, United States). Clinical pathological features and expression data of representative targets were analyzed using Pearson's chi-square and likelihood ratio tests. A level of $p < 0.05$ was considered statistically significant.

# 3 Results

## 3.1 Differential expression analysis

A total of 4,627 DEGs were identified from the overall stage of STAD, composed of 2,445 up-regulated mRNAs and 2,182 down-regulated mRNAs (Figure 2A; Supplementary Table S3). In addition, 4,715 DEGs were distinguished from early-stage STAD, in which 2,542 DEGs were remarkably up-regulated and 2,173 DEGs were notably down-regulated (Figure 2B, Supplementary Table S3). A total of 3,465 DEGs were detected from progressive-stage STAD, composed of 1,493 DEGs up-regulated mRNAs and 1972 down-regulated mRNAs (Figure 2C, Supplementary Table S3).

Intersections between overall- and early-stage, overall- and progressive-stage, and early- and progressive-stage STAD provided 4,059, 2,850, and 2,540 overlapping signatures, respectively. Overlapping of target DEGs of all three stages in STAD obtained 2,526 consensus mRNAs (Figure 3A, Supplementary Table S4).

Interestingly, six mRNAs, including SCGB3A1, SRARP, MUC5B, GABRB1, CNMD, and KRT27, were all up-regulated in early-stage STAD but down-regulated in progressive-stage STAD, while these targets were not significantly different when brought into overall STAD samples (Supplementary Table S4).

## 3.2 Cox proportional hazards regression model of DEGs

Univariate Cox regression analysis identified 504, 430, and 193 mRNAs-OS for the overall-, early-, and progressive-stage STADs (Supplementary Table S5). Intersections analyzes revealed 168, 104, and 15 overlapping mRNAs-OS between overall- and early-stage STAD, overall- and progressive-stage STAD, and early- and progressive-stage STAD, respectively (Figure 3B, Supplementary Table S6). Moreover, after overlapping mRNAs-OS among all three stages of STAD, a total of 14 signature genes were obtained (Figure 3B, Supplementary Table S6). In addition, CytoHubba plug-in analysis selected 40 top mRNAs-OS from the overall-, early-, and progressive- stage STAD (Supplementary Table S7).

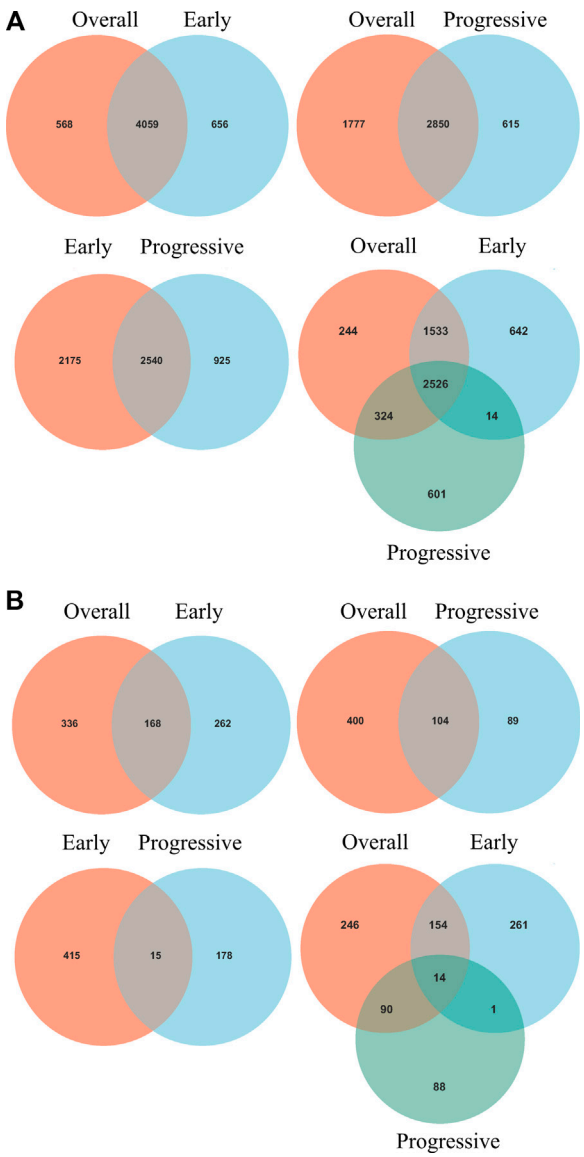



FIGURE 3
Venn diagrams of overlapping DEGs and mRNAs-OS from overall-, early-, and progressive-stage STAD. Venn diagrams are shown for overlapping DEGs **(A)** or mRNAs-OS **(B)**. Terms over the circles represent the groups of DEGs or mRNAs-OS at different stages of STAD.

Further multivariate Cox analysis based on those top mRNAs-OS constructed three prognostic models with 7, 9, and 14 mRNAs-PRO for these three stage sets (Figure 4, Supplementary Table S8). The survival risk score based on mRNAs-PRO for each stage was calculated by the model formula and stratified patients into low- and high-risk groups based on their median risk score. As presented in Figures 4A,D,G, expression heatmaps, risk score curve, and survival status map were plotted between the low- and high-risk groups of 7-, 9-, and 14-mRNA-based prognostic models for the overall-, early-, and progressive-stage STAD, respectively. The Kaplan–Meier survival curve of the high-risk group and the low-risk group for overall-, early-, and progressive-stage STAD analysis is presented in Figures 4B,E,H. The AUC value of the ROC curve was 0.73, 0.87, and 0.92 for overall, early-, and progressive-stage STAD, respectively, indicating that these models could form reliably to accurately predict the prognosis of STAD patients (Figures 4C,F,I).

## 3.3 Functional enrichment analysis

GO functional annotation and KEGG pathway analyses were carried out by the FunRich and KOBAS databases in

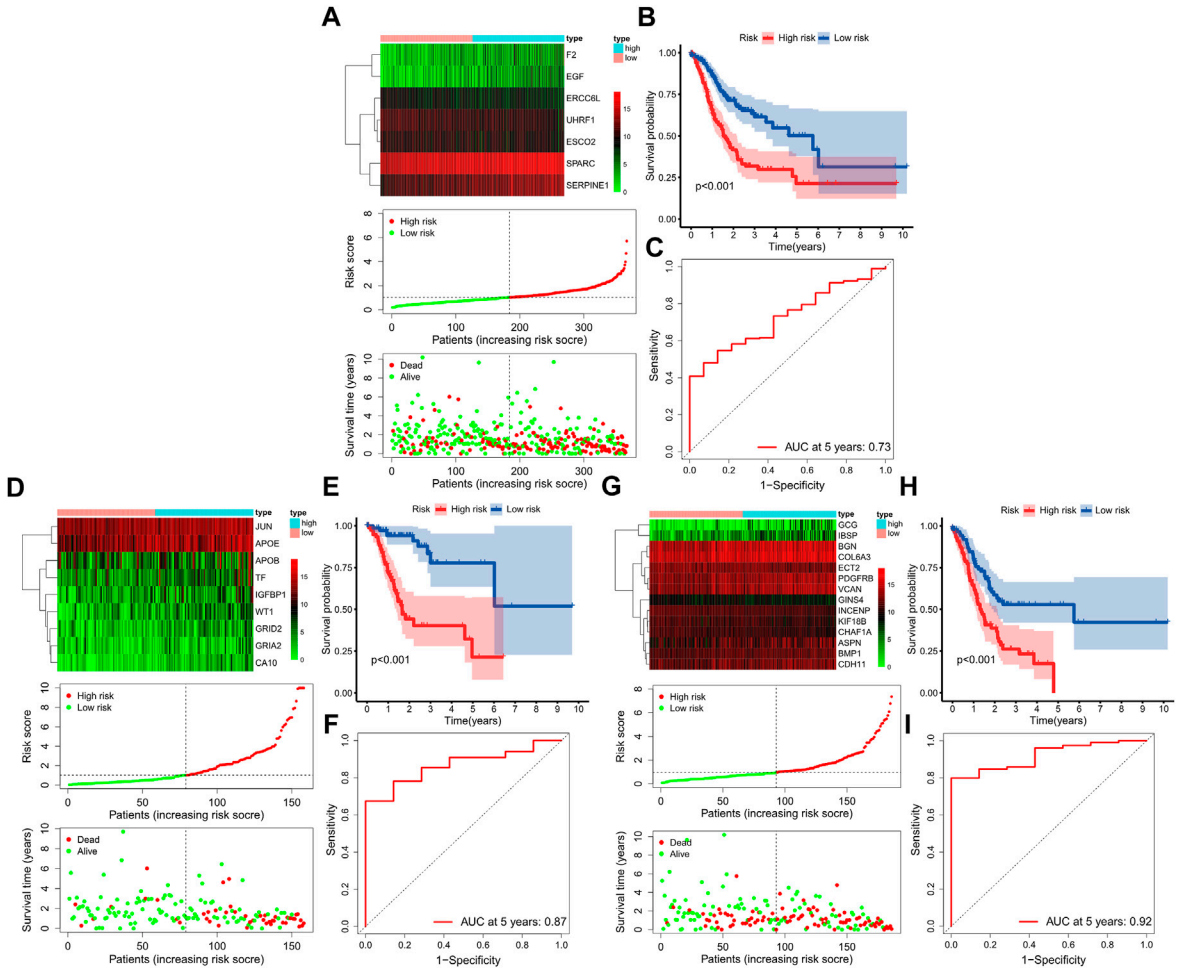



FIGURE 4
Construction of three prognostic models. The top, bottom left, and bottom right corner of this figure (three subgraphs) represent the STAD total samples group **(A–C)**, STAD samples in the early-stage group **(D–F)**, and STAD samples in the progressive-stage group **(G–I)**. **(A,D,G)** From top to bottom of every subgraph are the expression heatmap, risk score curve, and survival status map between the low-risk and high-risk groups. The color bar reveals the relative mRNAs-PRO expression level, with red denoting high expression and green denoting low expression. **(B,E,H)** Survival curve for the low-risk and high-risk groups. **(C,E,I)** ROC curve for survival predictions.

504 mRNAs-OS of the overall-stage STAD. mRNAs-OS of overall stage were enriched in 47 GO terms ($p < 0.05$), composed of seven BP, 12 MF, and 28 CC items. As exhibited in Figure 5A, in BP, the smallest $p$-value lay in the cell communication item ($p = 3.27E-3$), while annotation with the largest count number annotation was signal transduction (count = 122). The annotation with the smallest $p$-value and largest count number in the MF category was the extracellular matrix structural constituent ($p = 2.54E-7$, count = 18). As for CC, mRNAs-OS was remarkably enriched in functions as extracellular ($p = 5.68E-10$), and the largest count number was laid in the plasma membrane (count = 127). KEGG pathway analysis demonstrated that enriched mRNAs-OS were notably involved in 68 pathways, and the top eight pathways with smallest $p$-values are shown in Figure 5B, including neuroactive ligand–receptor interaction, complement and coagulation cascades, ABC transporters, ascorbate and aldarate metabolisms, malaria, cAMP signaling pathway, phospholipase D signaling pathway, and steroid hormone biosynthesis.

In order to explore the functions and pathways of mRNAs in different stages of STAD, gene set enrichment analysis (GSEA) was conducted in early- and progressive-stage STAD, respectively. Three most significantly enriched pathways were exhibited in Figure 6. As for BP, enrichment of mRNAs was remarkably laid in detection of chemical stimulus, detection of stimulus involved in sensory perception, and sensory perception of chemical stimulus in early-stage STAD (Figure 6A), while those mRNAs were significantly enriched in positive regulation

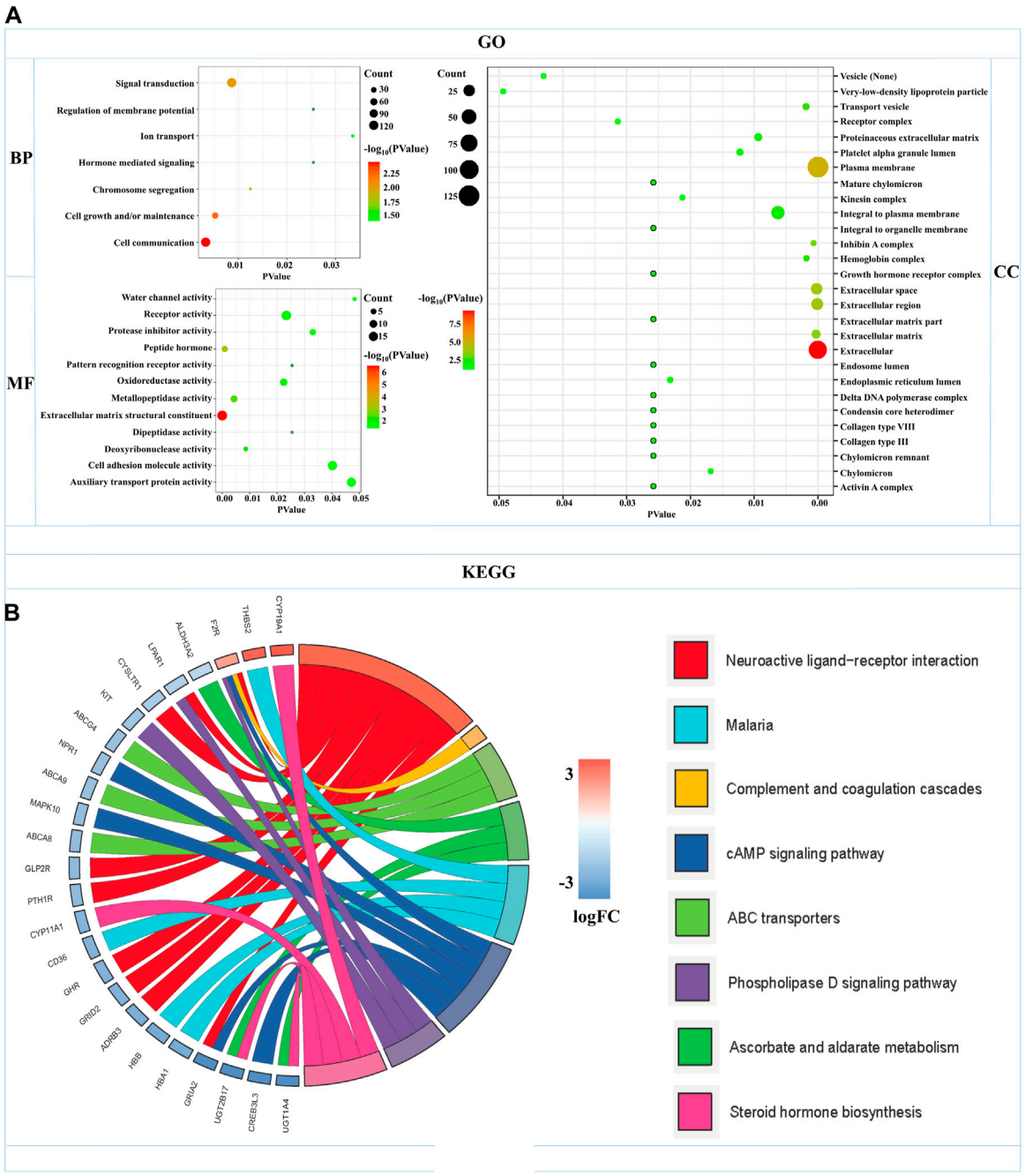



FIGURE 5
Functional enrichments of mRNAs-OS in overall-stage STAD. **(A)** GO function annotations (including BP, CC, and MF). The abscissa represents the *p*-value, and the ordinate represents the name of the functions. The bubble area increases with the number of mRNA-OS increasing. **(B)** KEGG pathways analysis. The color of the columns changes from blue to red as the degree of mRNAs enriched in the pathways increases.

of cell activation, regulation of lymphocyte activation, and regulation of T cell activation in progressive-stage STAD (Figure 6B). mRNAs of CC were markedly enriched in intermediate filament, intermediate filament cytoskeleton, and keratin filament in early-stage STAD (Figure 6C), while they were notably enriched pathways in progressive-stage STAD including external side of plasma membrane, immunological synapse, and presynapse (Figure 6D). The significantly enriched pathways in MF included olfactory receptor activity, odorant binding, and metal cluster binding in early-stage STAD

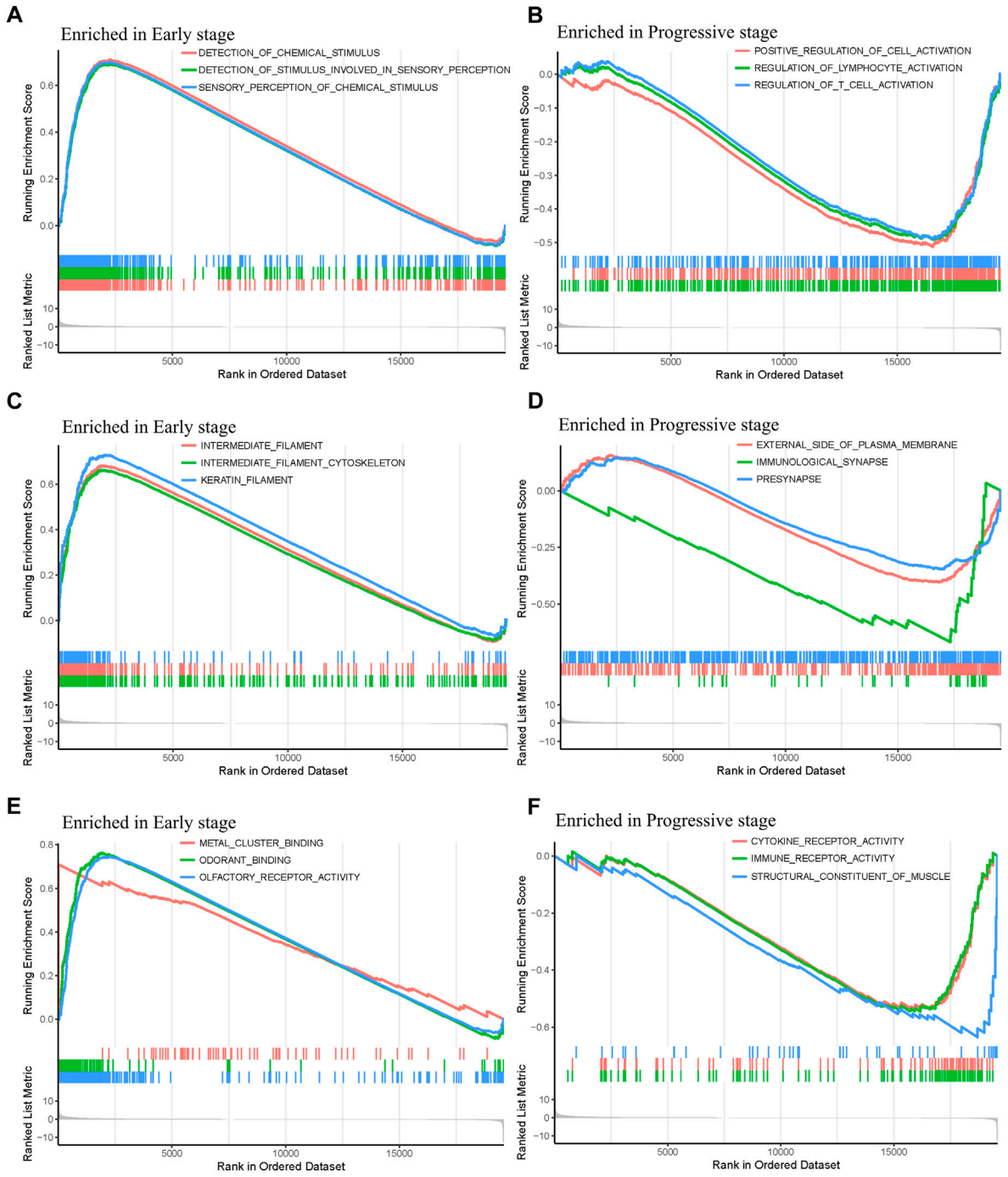



FIGURE 6
GO function annotations of mRNAs in the early- and progressive-stage STAD cohorts *via* GSEA. The left side of the figure refers to the enriched annotations in the early-stage group, while the right side refers to the enriched annotations in the progressive-stage group. From top to bottom of this figure are biological process (BP) **(A,B)**; cellular component (CC) **(C,D)**; and molecular function (MF) **(E,F)**

(Figure 6E), while pathways of progressive-stage STAD enriched in immune receptor activity, cytokine receptor activity, and structural constituent of muscle (Figure 6F). In addition, KEGG pathway analysis of early-stage STAD revealed that mRNAs were enriched significantly in olfactory transduction, maturity onset diabetes of the young, and regulation of autophagy (Figure 7A), while mRNAs were enriched in pathways including cell adhesion molecules cams, primary

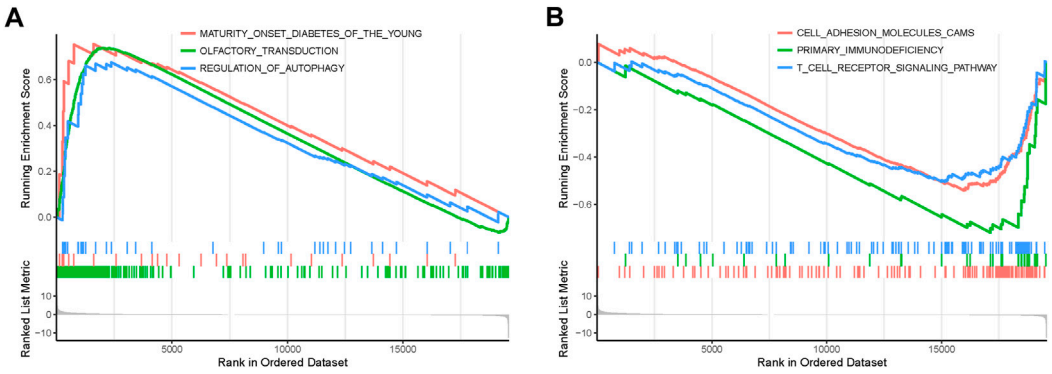



FIGURE 7
KEGG pathways of mRNAs in the early- and progressive-stage STAD cohorts *via* GSEA. **(A)** Enriched pathways in early-stage STAD. **(B)** Enriched pathways in progressive-stage STAD.

immunodeficiency, and T-cell receptor signaling pathway in progressive-stage STAD (Figure 7B).

## 3.4 Prognostic model validation

The predictive association between the age, gender, stage and risk score, and individual prognosis was evaluated by univariate and multivariate Cox regression analyses. Figure 8 shows both univariate and multivariate Cox regression analyses of overall survival time in overall-, early-, and progressive-stage STAD patients. As exhibited in the forest plots, while stage were significantly associated with the overall survival time in overall- and early-stage STAD patients (*Ps* ≤ 0.044) (Figures 8A–D), the risk score was the only independent factor for all the three stages STAD patients by both univariate and multivariate Cox regression models (*Ps* ≤ 0.002) (Figures 8A–F). These demonstrated the rationality of the stage stratification in the overall group and the risk-based prognostic models built for the overall-, early-, and progressive-stage STAD patients in this study. Moreover, the reliability and validity of prognostic models in classifying the high- and low-risk groups were further confirmed under various clinical circumstances by the Kaplan–Meier survival curves, and the overall survival (OS) time in the high-risk group was significantly lower than that of the low-risk group in overall- (Supplementary Figure S1), early- (Supplementary Figure S2), and progressive- (Supplementary Figure S3) stage STADs in all clinical situations, including age (≥ 65 or < 65), gender, T stage, M stage, and N stage (all $p < 0.05$). Thus, the prognostic models of overall-, early-, and progressive-stage STADs could be well used in a variety of clinical practices, indicating these three prognostic models were reliable and stable under different clinical circumstances.

## 3.5 QPCR and IHC of IGFBP1 and CHAF1A in STAD samples

Relative mRNA levels of IGFBP1 and CHAF1A in STAD and matched paracancerous tissues were evaluated *via* qPCR in overall-, early-, and progressive-stage STAD (Figure 9A), and the results showed the expression of IGFBP1 was significantly elevated in early-stage STAD in comparison with matched normal tissues ($p = 0.018$, Figure 9A-top), which was in line with bioinformatics findings. CHAF1A was markedly increased in overall- and progressive-stage STAD in comparison with matched normal tissues (both $p < 0.001$, Figure 9A-bottom), which was also consistent with expected analyses.

In addition, Figure 9B showed the representative IHC staining on the protein levels of IGFBP1 and CHAF1A, and Figure 9C demonstrated the cytoplasmic immunoreactivity of these two representative targets in the paired cancer and paraneoplastic normal tissues of STAD patients. For IGFBP1, cancer samples demonstrated stronger staining than matched normal tissues in early-stage STAD ($p = 0.022$, Figure 9C-top). For CHAF1A, cancer samples showed much stronger staining than matched normal tissues in the overall-, early-, and especially the progressive-stage STAD (all $p < 0.001$, Figure 9C-bottom). The IHC results on the protein levels of these two representative targets were basically lined well with the qPCR results on their mRNA expressions; both were in accordance with the calculation results. Thus, IGFBP1 could serve as a biomarker for early-stage STAD, while CHAF1A could act as a biomarker for progressive-stage STAD.

# 4 Discussion

According to the Global Cancer Statistics 2020, the incidence and mortality of GC ranks the fifth and fourth in all

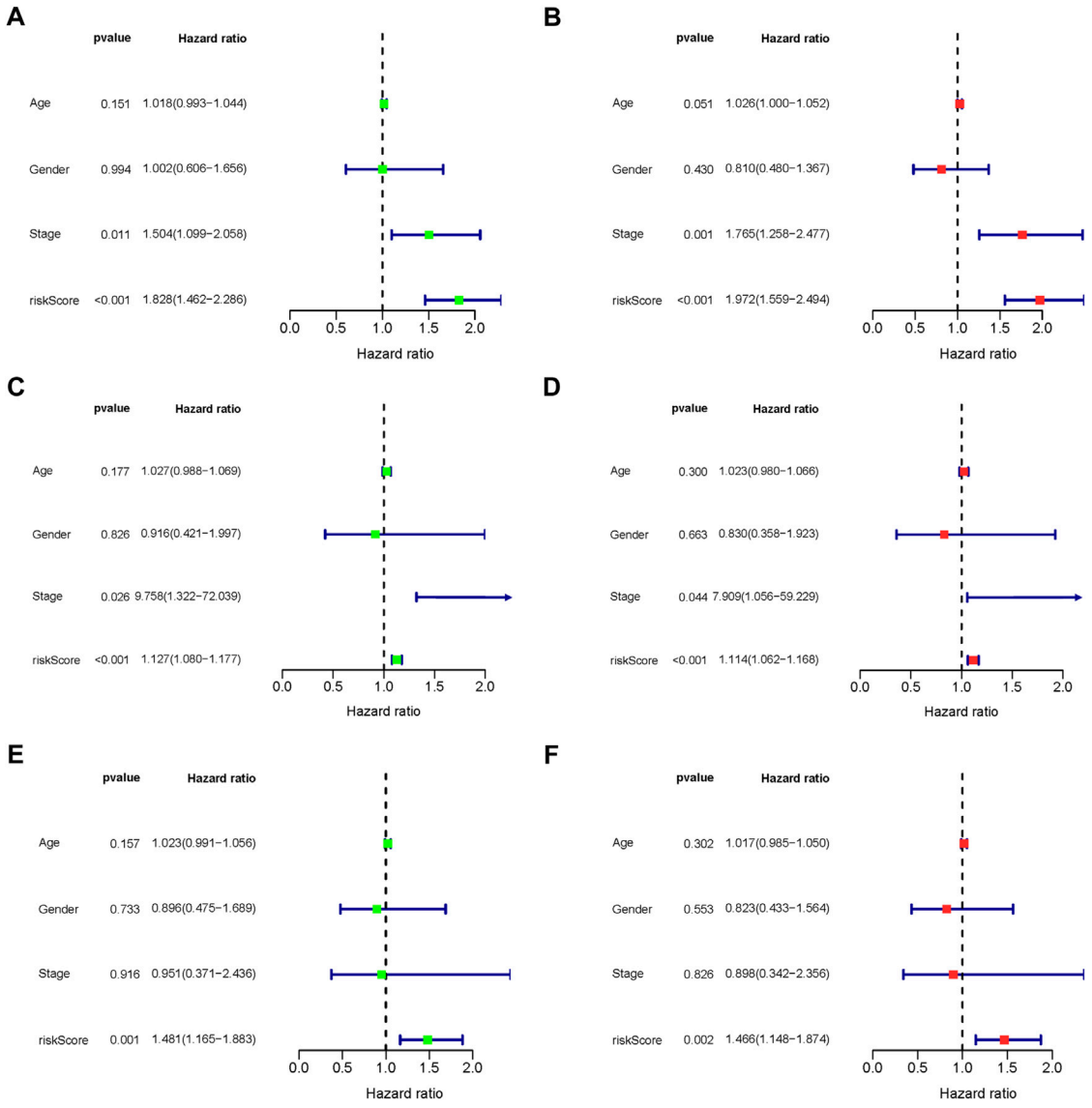



FIGURE 8
Forest plot of clinical factors prognostic analysis. **(A–B)** Overall-stage STAD; **(C–D)** early-stage STAD; and **(E–F)** progressive-stage STAD. **(A,C,E)** Univariate analysis of clinical factors. **(B,D,F)** Multivariate analysis of clinical factors. The dotted line indicates risk ratio (HR) = 1, the green or red rectangle to the left of the dotted line indicates that the clinical factor is a protective factor for STAD (HR < 1), and the green or red rectangle to the right of the dotted line indicates that the clinical factor is a risk factor for STAD (HR > 1).

malignancies, respectively (Sung et al., 2021). It causes a considerable global health burden and requires urgent attention (Yue et al., 2021). STAD accounts for 90% of gastric cancer and is the most common histological type of this malignance (Joshi and Badgwell, 2021). As the onset of STAD is insidious, most patients had already developed an advanced stage of cancer when they were confirmed diagnosis (Joshi and Badgwell, 2021). Current interventions have highlighted advances in therapeutics of STAD, including surgery, systemic chemotherapy, radiotherapy, immunotherapy, and targeted therapy; however, the efficacy does not meet the need for advanced gastric cancer (Jim et al., 2017), and the 5-year survival rate of distant-stage gastric cancer is less than 10% (Li et al., 2022). Under the passive situations nowadays, discovery of novel diagnostic biomarkers and biomarker-driven therapy for STAD is urgently needed.

With the development of high-throughput sequencing technologies, the mRNA molecules have been emerging as a new class of cancer biomarkers. Recent studies have explored associations between mRNAs and cancer features, and

10.3389/fmolb.2022.1022056

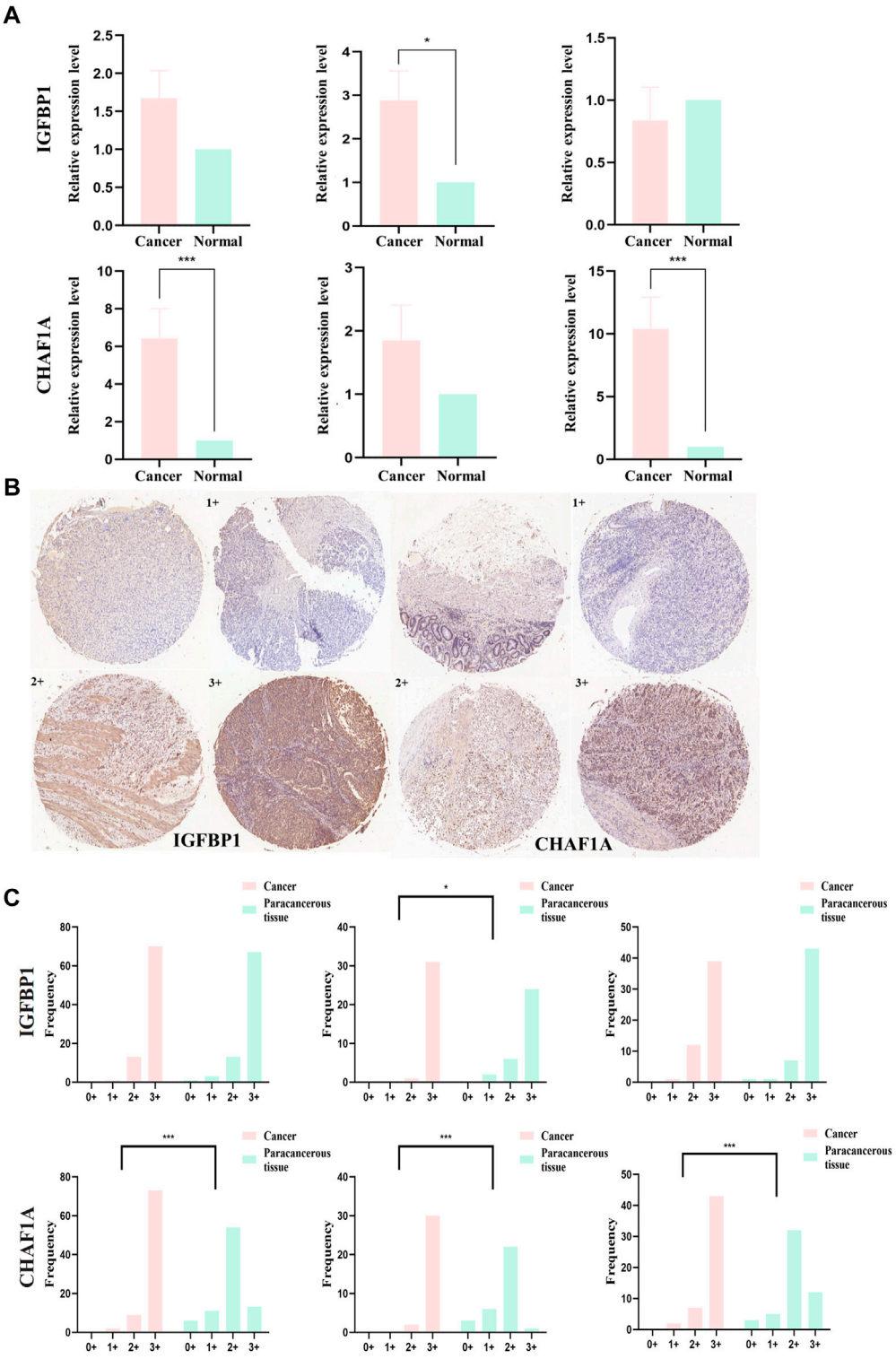



FIGURE 9
mRNA and protein expression levels of IGFBP1 and CHAF1A evaluated from STAD and normal samples by qPCR and IHC, respectively. **(A)** MRNA expression levels of IGFBP1 and CHAF1A detected by qPCR in 30 STAD and paired paracancerous normal tissues (early-stage: 13 pairs and progressive-stage: 17 pairs). **(B)** Representative images of IHC displaying negative (0), low (1+), moderate (2+), and strong (3+) staining for IGFBP1 and CHAF1A proteins from 84 STAD and paired paraneoplastic normal tissues (early-stage: 32 pairs and progressive-stage: 52 pairs). Various staining intensities of matched cancer and normal tissues were demonstrated, respectively, for IGFBP1 and CHAF1A. **(C)** IHC scores of overall, early-stage, and progressive-stage STAD, and paired normal tissues are presented in histograms with frequency. The statistical analysis was performed using Pearson's $\chi 2$ test and likelihood ratio test. $^{*}p < 0.05$ and $^{***}p < 0.001$.

moreover, various mRNA signatures have been reported in STAD for prognostic analysis (Ren et al., 2020; Wu et al., 2020; Zhou et al., 2020). However, some limitations of these studies should be noticed: most of them were reports of key biomarkers in overall STAD, while neglecting STAD stages (Li et al., 2019; Fu et al., 2020; Qiu et al., 2020). Nevertheless, STAD is a cancer characterized with high heterogeneity, and the differences of pathological features in distinct STAD stages were evident (Ren et al., 2021). Therefore, it is inappropriate to use the same mRNAs as biomarkers to classify patients with STAD at all the different stages in the clinical work.

This study identified 4,715 DEGs of early-stage STAD and 3,465 DEGs of progressive-stage STAD, and a total of 2,540 DEGs were overlapping, indicating that many remaining DEGs were evidently different in early- and progressive-stages STAD, and these DEGs might be vital to characterize the two stages. Interestingly, six mRNAs (SCGB3A1, SRARP, MUC5B, GABRB1, CNMD, and KRT27) were up-regulated in early-stage STAD, while they were down-regulated in progressive-stage STAD. This suggested these mRNAs might play a role in different stages of cancers by participating in different mechanisms; meanwhile, these signatures might act as convenient biomarkers to distinguish STAD stages.

Multivariate Cox regression analysis of the age, gender, stage, and risk score demonstrated that the tumor stage was an independent prognostic factor, and the mRNAs-PRO signatures for distinct STAD stages were potential predictors for prognosis.

The prognostic models of overall-, early-, and progressive-stage STAD were, respectively, composed of 7 mRNAs, 9 mRNAs, and 14 mRNAs. There was no consensus mRNAs-PRO in these three prognostic models. The AUC values of ROC curves of the prognostic models for overall, early, and progressive groups were 0.73, 0.87, and 0.92, respectively. The value of the overall stage was the smallest, suggesting a sorting stage might slightly improve the AUC value and the predictive reliability of the prognostic models. We also constructed a prognostic model by analyzing the data of all GC patients with the same procedure as the other three sets of data, and the value of the AUC was only 0.70, which was worst (data were not shown). This re-emphasized that the prognostic model established by stratified data according to influential clinical variables was more effective. As the amount of experimental data increased, novel molecular biomarkers could be identified to facilitate more accurate predictions after stratification. This might indicate that the prognostic model constructed for the overall-stage STAD was problematic due to the stage mixing.

Table 1 compared the abnormal expression of the prognostic mRNAs-PRO predicted in this study with those reported in previous studies. In the prognostic model of overall-stage STAD, all seven mRNAs (EGF, SERPINE1, ESCO2, ERCC6L, UHRF1, SPARC, and F2) were calculated to be up-regulated in this study. After a systematic and careful literature research, we found that the former six mRNAs were all experimentally validated to be overexpressed in either STAD tissues or GC tissues, and the aberrant expression of the remaining F2 in this cancer has not been confirmed by any previous experiments or bioinformatics analysis.

Among the nine mRNAs in the prognostic model of early-stage STAD, four mRNAs were up-regulated (WT1, IGFBP1, APOE, and TF) and five mRNAs were down-regulated (JUN, CA10, GRIA2, GRID2, and APOB). After comparing the outcomes in the related literature with the findings in our results, it was shown that the high expression level of WT1, IGFBP1, and APOE, and the low expression level of JUN was verified by previous clinical cases or experiments on GC, and the down-regulations of CA10 and GRIA2 were demonstrated in other tumor origins, where the results were all matched with our calculations. Although the remaining three mRNAs have not been reported by any cancer experiments before, GRID2 was calculated to be significantly down-regulated in endometrial carcinoma by bioinformatics studies, and TF and APOB have not been mentioned by any cancer studies before and were novel candidate prognostic signatures in early-stage STAD reported for the first time by this study. Notably, Kalantari et al. demonstrated that Lgr5High/DCLK1 high phenotype was associated with early-stage gastric carcinoma specimens (Kalantari et al., 2017). However, in this study, we found that Lgr5 was significantly upregulated, but DCLK1 was significantly downregulated in early-stage STAD (Supplementary Table S3); moreover, DCLK1 was one of the mRNAs-OS (Supplementary Table S5), though it was not in the 9-mRNA prognostic model of this premature stage. The contradiction between these two studies needs to be further explored.

Among the 14 mRNAs in the prognostic model of progressive-stage STAD, 13 mRNAs (CHAF1A, VCAN, CDH11, PDGFRB, BMP1, ASPN, BGN, GINS4, COL6A3, ECT2, IBSP, KIF18B, and INCENP) were upregulated, while one mRNA (GCG) was downregulated. As shown by the summaries of earlier studies in Table 1, we noted that the overexpression of eleven mRNAs (CHAF1A, VCAN, CDH11, PDGFRB, BMP1, ASPN, BGN, GINS4, COL6A3, ECT2, and IBSP) were all verified in GC cell lines; the increased expression level of KIF18B had been confirmed in hepatocellular carcinoma cells, and INCENP and GCG have not been mentioned by any cancer studies before and were novel candidate prognostic signatures in progressive-stage STAD reported for the first time by this study.

Emerging predictive tools for GC had been established based on prognostic genes (Chen J. et al., 2021b), lymph node histopathology (Wang X. et al., 2021b), immune microenvironment or/and tumor microenvironment-

**TABLE 1 Comparison of the abnormal expressions of the mRNAs-PRO in the overall-, early-, and progressive-stage STAD between this study and previous experimental studies.**

| Stage | mRNA-PRO | Feature | Abstract | Citation |
|---|---|---|---|---|
| Overall | EGF | ▲ | Using western blotting and immunocytochemical staining, Han et al. revealed that compared with normal gastric tissues, EGF was significantly higher expressed in STAD tissues and facilitated the migration and invasion of STAD cells by activating ERK1/2 | PMID: 24789460 Han et al. (2014) |
| | SERPINE1 | ▲ | Through the analysis of clinical samples, TCGA database samples, and cell lines *in vitro*, Yang et al. manifested that SERPINEI was significantly upregulated in STAD, and it promoted the invasion, migration, and proliferation of STAD cells by regulating EMT | PMID: 31724495 Yang et al. (2019) |
| | ESCO2 | ≜ | In *in vitro* cell experiments, Chen et al. found that ESCO2 was significantly upregulated in GC cells, and the knockdown of ESCO2 inhibited the proliferation and induced the apoptosis of GC cells | PMID: 29330052 Chen et al. (2018) |
| | ERCC6L | ≜ | Using qRT-PCR and IHC detections, Chen et al. discovered that compared with normal gastric tissues, ERCC6L was significantly upregulated in GC, and it facilitated the growth and metastasis of GC cells by activating NF-κB signaling | PMID: 34425559 Chen et al. (2021a) |
| | UHRF1 | ≜ | Zhang et al. illustrated that UHRF1 was significantly up-regulated in GC by cell experiments, which further promoted the migration and invasion of the GC cells and inhibited the cell apoptosis through ROS signaling | PMID: 30352833 Zhang et al. (2018) |
| | SPARC | ≜ | By the analysis of gene microarray data, Wang et al. revealed that SPARC was significantly upregulated in GC cells, and its high expression was positively correlated to the lymph node metastasis, lymphatic infiltration, and shorter survival of GC patients | PMID: 15558074 Wang et al. (2004) |
| | F2 | △ | Abnormal expression of F2 has not been reported in any cancer experiments. However, our study showed that F2 was significantly upregulated in STAD, implicating that F2 might be a novel biomarker of STAD | NA |
| Early | WT1 | ▲ | Li et al. found that WT1 was significantly upregulated in STAD by qRT-PCR. In addition, the overexpression of BASP1 in cell experiments of STAD could apparently inhibit the activation of the Wnt/β-catenin pathway to restrain the proliferation, migration, and invasion of STAD cells by downregulating the expression of WT1 | PMID: 33426068 Li et al. (2020a) |
| | JUN | ▼ | Using IHC, Jin et al. showed that JUN was down-regulated and had tumor suppressor activity in GC. The loss of JUN expression was correlated to a more advanced stage, lymphatic invasion, lymph node metastasis, and shorter survival of GC patients | PMID: 18158,562 Jin et al. (2007) |
| | IGFBP1 | ≜ | Luo et al. suggested that compared with the control groups, IGFBP1 was significantly upregulated in GC cells infected with *H. pylori* 26695. However, its overexpression could reduce the promoting effect of MMP-9 on the BGC-823 cells migration, indicating the protective role of IGFBP1 in the process of H. pylori-induced GC | PMID: 28864349 Luo et al. (2017) |
| | APOE | ≜ | Sakashita et al. suggested *via* RT-PCR that APOE was significantly highly expressed in GC, and GC samples with high expression of APOE had deeper tumor infiltration, more positive lymph node metastasis, and shorter survival compared with low APOE expression patients | PMID: 19020708 Sakashita et al. (2008) |
| | CA10 | ⩛ | Tao et al. found that CA10 was significantly downregulated in glioma and CA10 secreted by depolarized cultured neurons blocked the neuronal activity-dependent growth and inhibited the invasion of glioma by cell experiments | PMID: 30636076 Tao et al. (2019) |
| | GRIA2 | ⩛ | Through qRT-PCR and IHC, Choi et al. manifested that GRIA2 was significantly downregulated in advanced ovarian serous adenocarcinomas and the upregulation of GRIA2 was associated with the better survival of patients | PMID 22644307 Choi et al. (2012) |
| | TF | △ | The abnormal expression of TF has not been reported in any cancer experiments. However, our study showed that TF was not differentially expressed in overall-stage STAD, but was significantly up-regulated in the early-stage STAD, implying that TF might be a novel biomarker of early-stage STAD | NA |
| | GRID2 | ▽ | Abnormal expression of GRID2 has not been reported in any cancer experiments. However, Chen et al. suggested that GRID2 was significantly downregulated in endometrial carcinoma, and its expression was significantly associated with the proliferation and invasion of cancer cells through bioinformatics analysis. In this study, it was shown that GRID2 was significantly downregulated in the overall- and early-stage STAD, indicating that GRID2 might be a novel biomarker of these two stages of STAD | PMID: 33577492 Chen et al. (2021c) |
| | APOB | ▽ | Abnormal expression of APOB has not been reported in any cancer experiments. However, our study showed that APOB was significantly downregulated in the overall- and early-stage STAD, indicating that APOB might be a novel biomarker of these two stages STAD | NA |

(Continued on following page)

**TABLE 1 (*Continued*) Comparison of the abnormal expressions of the mRNAs-PRO in the overall-, early-, and progressive-stage STAD between this study and previous experimental studies.**

| Stage | mRNA-PRO | Feature | Abstract | Citation |
|---|---|---|---|---|
| Progressive | CHAF1A | ▲ | Zheng et al. found that CHAF1A was highly expressed in GC cells and could promote gastric carcinogenesis by upregulating c-MYC and CCND1. Moreover, the overexpression of CHAF1A in progressive-stage STAD were verified by qRT-PCR and IHC in our study | PMID: 30449701 Zheng et al. (2018) |
| | VCAN | ▲ | Using qRT-PCR of clinical tissue samples, Cheng et al. illustrated that VCAN was significantly upregulated, which promoted the proliferation, invasion, and migration of GC cells | PMID: 33116649 Cheng et al. (2020) |
| | CDH11 | ▲ | By bioinformatics analysis and cell experiments, Liu et al. found that CDH11 was highly expressed in GC and as an oncogene, CDH11 facilitated GC progression *via* transcriptional up-regulation by HEYL, but its overexpression could reversely promote the malignant behavior of HEYL-knockdown GC cells | PMID: 32463580 Liu et al. (2020) |
| | PDGFRB | ▲ | Through the RT-qPCR analysis of clinical tissue samples, Higuchi et al. revealed that PDGFRB was significantly upregulated in stage II/III gastric cancer and was closely related to the poor 5-year survival rate and proliferation of cancer cells | PMID: 28356977 Higuchi et al. (2017) |
| | BMP1 | ▲ | Through targeted RNA sequencing of clinical specimens, Hsieh et al. found that BMP1 was highly expressed in late-stage GC and significantly related to the poor prognosis of GC patients. In addition, *in vitro* suppression of BMP1 led to the reduction of the mobility of the GC cell lines, implying an important role of BMP1 in metastasis | PMID: 29720137 Hsieh et al. (2018) |
| | ASPN | ▲ | Via IHC assay, Zhang et al. illustrated that ASPN was significantly upregulated in GC, which was related to the poor prognosis of GC patients. The upregulation of ASPN promoted the growth of GC cells and inhibited apoptosis *via* deactivating LEF1-gene transcription independent of β-catenin *in vitro* and *in vivo* | PMID: 34127813 Zhang et al. (2021b) |
| | BGN | ▲ | Via *in vitro* and *in vivo* experiments, Pinto demonstrated that BGN was significantly overexpressed in GC tissues, which was related to disease recurrence and poor prognosis of patients with advanced GC. The over-expressed BGN promoted the cell migration, invasion, and angiogenesis of GC | PMID: 33809543 Pinto et al. (2021) |
| | GINS4 | ▲ | Using the qRT-PCR and IHC detection of clinical tissue samples, Zhu et al. illustrated that GINS4 was significantly overexpressed in GC, and the up-regulation of GINS4 was associated with poor differentiation, advanced stage, depth of invasion, and lymph node metastasis of GC tissues | PMID: 31754397 Zhu et al. (2019) |
| | COL6A3 | ▲ | Sun et al. showed that COL6A3 was highly expressed in GC by using RT-PCR, and the overexpression of COL6A3 facilitated the proliferation, migration, and inhibited the apoptosis of GC cells. | PMID: 31122696 Sun et al. (2019) |
| | ECT2 | ▲ | Through RT-qPCR and Western blotting, Zhang et al. suggested that compared with normal control, the expression level of ECT2 in GC increased significantly and was positively correlated to histological differentiation, lymph node metastasis, and TNM stage of GC | PMID: 34367280 Zhang et al. (2021a) |
| | IBSP | ⩓ | Using the qRT-PCR and IHC detection of clinical tissue samples, Wang et al. suggested that IBSP was overexpressed in esophageal squamous cell carcinoma and the upregulation of IBSP was significantly associated with the lymph node metastasis, clinicopathological stage, and poor disease survival | PMID: 31709184 Wang et al. (2019) |
| | KIF18B | ⩓ | With the qRT-PCR and IHC detection of clinical tissue samples, Yang et al. illustrated that KIF18B was significantly upregulated in hepatocellular carcinoma, thus promoting the progression of hepatocellular carcinoma by activating the Wnt/β-catenin pathway | PMID: 32052444 Yang et al. (2020) |
| | INCENP | △ | Abnormal expression of INCENP has not been reported in any cancer experiments. However, our study showed that INCENP was not differentially expressed in overall-stage STAD but was significantly upregulated in progressive-stage STAD, indicating that INCENP might be a novel biomarker of progressive-stage STAD | NA |
| | GCG | ▽ | Abnormal expression of GCG has not been reported in any cancer experiments. However, our study showed that GCG was not differentially expressed in overall-stage STAD but was significantly downregulated in the progressive-stage STAD, indicating that GCG might be a novel biomarker of progressive-stage STAD | NA |

▲mRNAs experimentally up-regulated in STAD in accordance with our calculated result. ▼mRNAs experimentally downregulated in STAD in accordance with our calculated result. ⩓mRNAs experimentally upregulated in GC in accord with our calculated result. ⩔mRNAs experimentally downregulated in GC in accordance with our calculated result. ⩓mRNAs experimentally upregulated in other cancers in accordance with our calculated result. ⩔mRNAs experimentally down-regulated in other cancers in accordance with our calculated result. △mRNAs upregulated in our calculated results not experimentally verified in cancers. ▽mRNAs downregulated in our calculated results not experimentally verified in cancers.

relevant genes (Cai et al., 2020; Wan et al., 2020), deep learning radiomic nomogram (Dong et al., 2020), lncRNAs (Ma et al., 2020), and other scoring systems (Zhang et al., 2019). Although these prognostic models mentioned earlier could predict the outcomes of certain GC patients in diverse clinical practices, they all have shortcomings: they did not be classified as histopathology types or stages, and AUC values of ROC curves for early- and progressive-stage STAD in this study were superior. This demonstrated better survival prediction, probably owing to the more refined and representative data after stratification. Moreover, three models were further confirmed reliably in various clinical circumstance, including different ages (≤65 years or > 65 years), genders (males or females), T stages (T1, 2 or T3, 4), N stages (N0 or N1, 2), and M stage (M0) (Supplementary Figures S1–3), indicating the applicability of prognostic models established by stratified data according to vital clinical variables. As experimental data were enhanced after stratification, novel molecular biomarkers would be identified to facilitate more accurate predictions.

Enrichment analysis of the overall STAD stage was based on mRNAs-OS, whereas the GSEA of stratification stages was based on all mRNAs. As for GO analysis, most of mRNAs-OS were enriched in cell communication, cell growth and/or maintenance, signal transduction, receptor activity, cell adhesion molecule activity, and auxiliary transport protein activity (Figure 5A), which were basic, common, and essential for survival of cancer cells. Both GO and KEGG analyses in stratified STAD stages revealed notable heterogeneity in early- and progressive-stage STAD. In early-stage STAD, GSEA-revealed chemical stimulus, including detection, and sensory were markedly enriched (Figure 6A), but few reports focused on these respects. MRNAs of progressive-stage STAD from GSEA were mainly enriched in immune-related functions, including regulation of lymphocyte activation, regulation of T-cell activation, immunological synapse, cytokine receptor activity, and immune receptor activity (Figures 6B,D,F). Immune-related signatures took part in formation of the tumor microenvironment and could predict prognosis in GC (Dai et al., 2021). Wang et al. performed pathway enrichment analysis based on single-cell sequencing in GC, and 12 pathways were immune-related, such as defensins, IL-7 signaling, and IL6/JAK/STAT3 signaling; meanwhile, these pathways were all associated with longer survival, suggesting certain immune-related biological processes contribute to distinct molecular consequences and patient survival (Wang R. et al., 2021a). As for KEGG analyses in early-stage STAD, regulation of autophagy is of note (Figure 7A). Autophagy is a double-edged sword in GC. On the one hand, it inhibits tumor initiation at early-stage by clearing damaged mitochondria, peroxisome, and also meets the high metabolic needs from enhanced proliferating tumors (Kongara and Karantza, 2012). On the other hand, autophagy protects some tumor cells against nutrient and oxygen deprivation, and harsh tumor microenvironments; meanwhile, autophagy is shown to facilitate resistance to cisplatin in GC cells, and thus, it is a cause of tumor metastasis, recurrence, and chemoresistance (Maes et al., 2013; Katheder et al., 2017). As for KEEG analysis in progressive-stage STAD, cell adhesion molecules (CAMs) were of note (Figure 7B). Cell adhesion molecules have multifaceted roles, including signaling molecules, and key constituents of the cell migration machinery, which are involved in virtually each step of tumor progression from primary cancer development to metastasis (Hamidi and Ivaska, 2018). Altered expression of cell adhesion molecules is frequently detected in tumors, and meanwhile, these molecules contribute to supporting the oncogenic growth factor receptor (GFR) signaling, and GFR-dependent cell invasion and migration, which are main features of progressive tumors (Hamidi and Ivaska, 2018).

Further qPCR and IHC detections were performed to estimate mRNA and protein levels of two representative targets in clinical samples. IGFBP-1, which belongs to the insulin-like growth factor system, plays an essential role in the pathophysiology of various tumors (Lin et al., 2021). The role of IGFBP1 had been explored by previous experiments. Yuya Sato et al. reported IGFBP1 could predict hematogenous metastasis in patients with gastric cancer (Sato et al., 2019), while another research demonstrated that the role of IL35 in GC angiogenesis was altering TIMP1, PAI1, and IGFBP1 (Li X. et al., 2020b). In our study, the experimental results of IGFBP1 in both transcriptional and protein levels were in accordance with the bioinformatics outcome, which illustrated it could predict early-stage STAD. CHAF1A, as a known histone chaperone, is also upregulated in tumors of many origins, including GC. It has been reported CHAF1A could upregulate the c-MYC and CCND1 expressions and further promote gastric carcinogenesis (Zheng et al., 2018). In this study, both qPCR and IHC results of CHAF1A were in line with bioinformatics analysis, which illustrated that it could predict progressive-stage STAD. Thus, these two targets could serve as potential signatures for identification of STAD at different risks.

This study stratified STAD into early and progressive stages, which is closely related to distinct treatment and prognosis, and by combining TCGA data, conclusions drawn are more general and representative. In addition, through research validation in clinical cases, we believed our results were stable and reliable. Meanwhile, it is worth noting that limited to TCGA data, the number of paraneoplastic tissues was small, and cancer and paraneoplastic normal cases were evidently not one-to-one matching. Therefore, selection bias is hard to avoid in bioinformatics results when based on these data. Also, since there were only 30 cases in the qPCR validation

cohort, and most of which were progressive-stage STAD, and for IHC, though validated by more than 80 paired STAD and paraneoplastic normal tissues, it was a semi-quantitative method, so subjectivity cannot be ruled out when evaluation of targets. Therefore, further investigations by large and matched cohorts were strongly suggested to be performed in the future.

In conclusion, our study combined bioinformatics analysis, clinical parameters, qPCR, and IHC detections in clinical samples to draw the conclusion that mRNA-based models by stratification could predict outcomes of patients at different risks, and selected signatures could serve as novel biomarkers for STAD patients at varied stages.

## Data availability statement

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

## Ethics statement

The studies involving human participants were reviewed and approved by the Clinical Research Ethics Committee, Outdo Biotech (Shanghai, China). The patients/participants provided their written informed consent to participate in this study.

## Author contributions

ZH designed this study. XH, KZ, and NX performed the bioinformatics analysis. JZ carried out the qPCR and IHC analyses. XH, JZ, NX, JW, and ST drafted the manuscript. ZH made the critical revision of the manuscript for important intellectual content and provided the funding support. JZ and YL provided the administrative, technical, and material support.

## Funding

This research was funded by the Key Discipline Construction Project of Guangdong Medical University (4SG22004G), the National Natural Science Foundation of China (81902511), and the Undergraduate Training Program for Innovation and Entrepreneurship of Guangdong Province (S202010571083).

## Acknowledgments

The authors would like to acknowledge all the research participants contributing to The Cancer Genome Atlas (TCGA) resource for providing high-quality data for analysis, and the authors also thank BioRENDER (https://app.biorender.com/) for the graphical plot of Figure 1 (approved by BioRENDER).

## 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/fmolb.2022.1022056/full#supplementary-material

## References

Ajani, J. A., D'Amico, T. A., Almhanna, K., Bentrem, D. J., Chao, J., Das, P., et al. (2016). Gastric cancer, version 3.2016, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Canc. Netw.* 14 (10), 1286–1312. doi:10.6004/jnccn.2016.0137

Ajani, J. A., Lee, J., Sano, T., Janjigian, Y. Y., Fan, D., and Song, S. (2017). Gastric adenocarcinoma. *Nat. Rev. Dis. Prim.* 3, 17036. doi:10.1038/nrdp.2017.36

Cai, W. Y., Dong, Z. N., Fu, X. T., Lin, L. Y., Wang, L., Ye, G. D., et al. (2020). Identification of a tumor microenvironment-relevant gene set-based prognostic signature and related therapy targets in gastric cancer. *Theranostics* 10 (19), 8633–8647. doi:10.7150/thno.47938

Chen, D., Liu, Q., and Cao, G. (2021a). ERCC6L promotes cell growth and metastasis in gastric cancer through activating NF-κB signaling. *Aging (Albany NY)* 13 (16), 20218–20228. doi:10.18632/aging.203387

Chen, H., Zhang, L., He, W., Liu, T., Zhao, Y., Chen, H., et al. (2018). ESCO2 knockdown inhibits cell proliferation and induces apoptosis in human gastric cancer cells. *Biochem. Biophys. Res. Commun.* 496 (2), 475–481. doi:10.1016/j.bbrc.2018.01.048

Chen, J., Wang, A., Ji, J., Zhou, K., Bu, Z., Lyu, G., et al. (2021b). An innovative prognostic model based on four genes in asian patient with gastric cancer. *Cancer Res. Treat.* 53 (1), 148–161. doi:10.4143/crt.2020.424

Chen, X., Zhang, W., Zhu, H., and Lin, F. (2021c). Development and validation of a 5-gene autophagy-based prognostic index in endometrial carcinoma. *Med. Sci. Monit.* 27, e928949. doi:10.12659/msm.928949

Cheng, Y., Sun, H., Wu, L., Wu, F., Tang, W., Wang, X., et al. (2020). VUp-regulation of VCAN promotes the proliferation, invasion and migration and serves as a biomarker in gastric cancer. *Onco. Targets. Ther.* 13, 8665–8675. doi:10.2147/ott.S262613

Choi, C. H., Choi, J. J., Park, Y. A., Lee, Y. Y., Song, S. Y., Sung, C. O., et al. (2012). Identification of differentially expressed genes according to chemosensitivity in advanced ovarian serous adenocarcinomas: Expression of GRIA2 predicts better survival. *Br. J. Cancer* 107 (1), 91–99. doi:10.1038/bjc.2012.217

Dai, S., Liu, T., Liu, X. Q., Li, X. Y., Xu, K., Ren, T., et al. (2021). Identification of an immune-related signature predicting survival risk and immune microenvironment in gastric cancer. *Front. Cell Dev. Biol.* 9, 687473. doi:10.3389/fcell.2021.687473

Dong, D., Fang, M. J., Tang, L., Shan, X. H., Gao, J. B., Giganti, F., et al. (2020). Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: An international multicenter study. *Ann. Oncol.* 31 (7), 912–920. doi:10.1016/j.annonc.2020.04.003

Fu, T., Ji, X., Bu, Z., Zhang, J., Wu, X., Zong, X., et al. (2020). Identification of key long non-coding RNAs in gastric adenocarcinoma. *Cancer Biomark.* 27 (4), 541–553. doi:10.3233/cbm-192389

Hamidi, H., and Ivaska, J. (2018). Every step of the way: Integrins in cancer progression and metastasis. *Nat. Rev. Cancer* 18 (9), 533–548. doi:10.1038/s41568-018-0038-z

Han, J., Xie, Y., Lan, F., Yu, Y., Liu, W., Chen, J., et al. (2014). Additive effects of EGF and IL-1β regulate tumor cell migration and invasion in gastric adenocarcinoma via activation of ERK1/2. *Int. J. Oncol.* 45 (1), 291–301. doi:10.3892/ijo.2014.2401

Higuchi, A., Oshima, T., Yoshihara, K., Sakamaki, K., Aoyama, T., Suganuma, N., et al. (2017). Clinical significance of platelet-derived growth factor receptor-β gene expression in stage II/III gastric cancer with S-1 adjuvant chemotherapy. *Oncol. Lett.* 13 (2), 905–911. doi:10.3892/ol.2016.5494

Hsieh, Y. Y., Tung, S. Y., Pan, H. Y., Yen, C. W., Xu, H. W., Deng, Y. F., et al. (2018). Upregulation of bone morphogenetic protein 1 is associated with poor prognosis of late-stage gastric Cancer patients. *BMC Cancer* 18 (1), 508. doi:10.1186/s12885-018-4383-9

Huang, G., Cui, F., Yu, F., Lu, H., Zhang, M., Tang, H., et al. (2015). Sirtuin-4 (SIRT4) is downregulated and associated with some clinicopathological features in gastric adenocarcinoma. *Biomed. Pharmacother.* 72, 135–139. doi:10.1016/j.biopha.2015.04.013

Jiang, N., Guo, Q., and Luo, Q. (2022). Inhibition of ITGB1-DT expression delays the growth and migration of stomach adenocarcinoma and improves the prognosis of cancer patients using the bioinformatics and cell model analysis. *J. Gastrointest. Oncol.* 13 (2), 615–629. doi:10.21037/jgo-22-233

Jim, M. A., Pinheiro, P. S., Carreira, H., Espey, D. K., Wiggins, C. L., and Weir, H. K. (2017). Stomach cancer survival in the United States by race and stage (2001-2009): Findings from the CONCORD-2 study. *Cancer* 123 (24), 4994–5013. doi:10.1002/cncr.30881

Jin, S. P., Kim, J. H., Kim, M. A., Yang, H. K., Lee, H. E., Lee, H. S., et al. (2007). Prognostic significance of loss of c-fos protein in gastric carcinoma. *Pathol. Oncol. Res.* 13 (4), 284–289. doi:10.1007/bf02940306

Joshi, S. S., and Badgwell, B. D. (2021). Current treatment and recent progress in gastric cancer. *Ca. Cancer J. Clin.* 71 (3), 264–279. doi:10.3322/caac.21657

Kalantari, E., Asadi Lari, M. H., Roudi, R., Korourian, A., and Madjd, Z. (2017). Lgr5High/DCLK1High phenotype is more common in early stage and intestinal subtypes of gastric carcinomas. *Cancer Biomark.* 20 (4), 563–573. doi:10.3233/cbm-170383

Katheder, N. S., Khezri, R., O'Farrell, F., Schultz, S. W., Jain, A., Rahman, M. M., et al. (2017). Microenvironmental autophagy promotes tumour growth. *Nature* 541 (7637), 417–420. doi:10.1038/nature20815

Kongara, S., and Karantza, V. (2012). The interplay between autophagy and ROS in tumorigenesis. *Front. Oncol.* 2, 171. doi:10.3389/fonc.2012.00171

Li, L., Meng, Q., Li, G., and Zhao, L. (2020a). BASP1 suppresses cell growth and metastasis through inhibiting wnt/β-catenin pathway in gastric cancer. *Biomed. Res. Int.* 2020, 8628695. doi:10.1155/2020/8628695

Li, L., Zhu, Z., Zhao, Y., Zhang, Q., Wu, X., Miao, B., et al. (2019). FN1, SPARC, and SERPINE1 are highly expressed and significantly related to a poor prognosis of gastric adenocarcinoma revealed by microarray and bioinformatics. *Sci. Rep.* 9 (1), 7827. doi:10.1038/s41598-019-43924-x

Li, X., Niu, N., Sun, J., Mou, Y., He, X., and Mei, L. (2020b). IL35 predicts prognosis in gastric cancer and is associated with angiogenesis by altering TIMP1, PAI1 and IGFBP1. *FEBS Open Bio* 10 (12), 2687–2701. doi:10.1002/2211-5463.13005

Li, Y., Feng, A., Zheng, S., Chen, C., and Lyu, J. (2022). Recent estimates and predictions of 5-year survival in patients with gastric cancer: A model-based period analysis. *Cancer Control.* 29, 10732748221099227. doi:10.1177/10732748221099227

Lin, Y. W., Weng, X. F., Huang, B. L., Guo, H. P., Xu, Y. W., and Peng, Y. H. (2021). IGFBP-1 in cancer: Expression, molecular mechanisms, and potential clinical implications. *Am. J. Transl. Res.* 13 (3), 813–832.

Liu, H., Ni, S., Wang, H., Zhang, Q., and Weng, W. (2020). Charactering tumor microenvironment reveals stromal-related transcription factors promote tumor carcinogenesis in gastric cancer. *Cancer Med.* 9 (14), 5247–5257. doi:10.1002/cam4.3133

Lu, Z., Luo, T., Pang, T., Du, Z., Yin, X., Cui, H., et al. (2019). MALAT1 promotes gastric adenocarcinoma through the MALAT1/miR-181a-5p/AKT3 axis. *Open Biol.* 9 (9), 190095. doi:10.1098/rsob.190095

Luo, C., Sun, F., Zhu, H., Ni, Y., Fang, J., Liu, Y., et al. (2017). Insulin-like growth factor binding protein-1 (IGFBP-1) upregulated by *Helicobacter pylori* and is associated with gastric cancer cells migration. *Pathol. Res. Pract.* 213 (9), 1029–1036. doi:10.1016/j.prp.2017.08.009

Ma, B., Li, Y., and Ren, Y. (2020). Identification of a 6-lncRNA prognostic signature based on microarray re-annotation in gastric cancer. *Cancer Med.* 9 (1), 335–349. doi:10.1002/cam4.2621

Maes, H., Rubio, N., Garg, A. D., and Agostinis, P. (2013). Autophagy: Shaping the tumor microenvironment and therapeutic response. *Trends Mol. Med.* 19 (7), 428–446. doi:10.1016/j.molmed.2013.04.005

Pinto, F., Santos-Ferreira, L., Pinto, M. T., Gomes, C., and Reis, C. A. (2021). The extracellular small leucine-rich proteoglycan biglycan is a key player in gastric cancer aggressiveness. *Cancers (Basel)* 13 (6), 1330. doi:10.3390/cancers13061330

Qiu, J., Sun, M., Wang, Y., and Chen, B. (2020). Identification of hub genes and pathways in gastric adenocarcinoma based on bioinformatics analysis. *Med. Sci. Monit.* 26, e920261. doi:10.12659/msm.920261

Ren, C., Wu, C., Wang, N., Lian, C., and Yang, C. (2021). Clonal architectures predict clinical outcome in gastric adenocarcinoma based on genomic variation, tumor evolution, and heterogeneity. *Cell Transpl.* 30, 963689721989606. doi:10.1177/0963689721989606

Ren, N., Liang, B., and Li, Y. (2020). Identification of prognosis-related genes in the tumor microenvironment of stomach adenocarcinoma by TCGA and GEO datasets. *Biosci. Rep.* 40 (10), BSR20200980. doi:10.1042/bsr20200980

Rihawi, K., Ricci, A. D., Rizzo, A., Brocchi, S., Marasco, G., Pastore, L. V., et al. (2021). Tumor-associated macrophages and inflammatory microenvironment in gastric cancer: Novel translational implications. *Int. J. Mol. Sci.* 22 (8), 3805. doi:10.3390/ijms22083805

Sakashita, K., Tanaka, F., Zhang, X., Mimori, K., Kamohara, Y., Inoue, H., et al. (2008). Clinical significance of ApoE expression in human gastric cancer. *Oncol. Rep.* 20 (6), 1313–1319.

Sato, Y., Inokuchi, M., Takagi, Y., and Kojima, K. (2019). IGFBP1 is a predictive factor for haematogenous metastasis in patients with gastric cancer. *Anticancer Res.* 39 (6), 2829–2837. doi:10.21873/anticanres.13411

Sun, X., Zhang, X., Zhai, H., Zhang, D., and Ma, S. (2019). A circular RNA derived from COL6A3 functions as a ceRNA in gastric cancer development. *Biochem. Biophys. Res. Commun.* 515 (1), 16–23. doi:10.1016/j.bbrc.2019.05.079

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

Tao, B., Ling, Y., Zhang, Y., Li, S., Zhou, P., Wang, X., et al. (2019). CA10 and CA11 negatively regulate neuronal activity-dependent growth of gliomas. *Mol. Oncol.* 13 (5), 1018–1032. doi:10.1002/1878-0261.12445

Wan, L., Tan, N., Zhang, N., and Xie, X. (2020). Establishment of an immune microenvironment-based prognostic predictive model for gastric cancer. *Life Sci.* 261, 118402. doi:10.1016/j.lfs.2020.118402

Wang, C. S., Lin, K. H., Chen, S. L., Chan, Y. F., and Hsueh, S. (2004). Overexpression of SPARC gene in human gastric carcinoma and its clinic-pathologic significance. *Br. J. Cancer* 91 (11), 1924–1930. doi:10.1038/sj.bjc.6602213

Wang, M., Liu, B., Li, D., Wu, Y., Wu, X., Jiao, S., et al. (2019). Upregulation of IBSP expression predicts poor prognosis in patients with esophageal squamous cell carcinoma. *Front. Oncol.* 9, 1117. doi:10.3389/fonc.2019.01117

Wang, R., Dang, M., Harada, K., Han, G., Wang, F., Pool Pizzi, M., et al. (2021a). Single-cell dissection of intratumoral heterogeneity and lineage diversity in metastatic gastric adenocarcinoma. *Nat. Med.* 27 (1), 141–151. doi:10.1038/s41591-020-1125-8

Wang, X., Chen, Y., Gao, Y., Zhang, H., Guan, Z., Dong, Z., et al. (2021b). Predicting gastric cancer outcome from resected lymph node histopathology images using deep learning. *Nat. Commun.* 12 (1), 1637. doi:10.1038/s41467-021-21674-7

Wu, M., Xia, Y., Wang, Y., Fan, F., Li, X., Song, J., et al. (2020). Development and validation of an immune-related gene prognostic model for stomach adenocarcinoma. *Biosci. Rep.* 40 (10), BSR20201012. doi:10.1042/bsr20201012

Xia, N., Xia, L., Zhang, W. F., and Zhou, F. X. (2022). Immune-related genes and their determined immune cell microenvironment to predict the prognosis of gastric adenocarcinoma. *Zhonghua Yi Xue Za Zhi* 102 (12), 840–846. doi:10.3760/cma.j.cn112137-20211023-02348

Yang, B., Wang, S., Xie, H., Wang, C., Gao, X., Rong, Y., et al. (2020). KIF18B promotes hepatocellular carcinoma progression through activating Wnt/β-catenin-signaling pathway. *J. Cell. Physiol.* 235 (10), 6507–6514. doi:10.1002/jcp.29444

Yang, J. D., Ma, L., and Zhu, Z. (2019). SERPINE1 as a cancer-promoting gene in gastric adenocarcinoma: Facilitates tumour cell proliferation, migration, and invasion by regulating EMT. *J. Chemother.* 31 (7-8), 408–418. doi:10.1080/1120009x.2019.1687996

Yue, T., Zuo, S., Zhu, J., Guo, S., Huang, Z., Li, J., et al. (2021). Two similar signatures for predicting the prognosis and immunotherapy efficacy of stomach adenocarcinoma patients. *Front. Cell Dev. Biol.* 9, 704242. doi:10.3389/fcell.2021.704242

Zhang, D., He, W., Wu, C., Tan, Y., He, Y., Xu, B., et al. (2019). Scoring system for tumor-infiltrating lymphocytes and its prognostic value for gastric cancer. *Front. Immunol.* 10, 71. doi:10.3389/fimmu.2019.00071

Zhang, H., Geng, Y., Sun, C., and Yu, J. (2021a). Upregulation of ECT2 predicts adverse clinical outcomes and increases 5-fluorouracil resistance in gastric cancer patients. *J. Oncol.* 2021, 2102890. doi:10.1155/2021/2102890

Zhang, H., Song, Y., Yang, C., and Wu, X. (2018). UHRF1 mediates cell migration and invasion of gastric cancer. *Biosci. Rep.* 38 (6), BSR20181065. doi:10.1042/bsr20181065

Zhang, Z., Min, L., Li, H., Chen, L., Zhao, Y., Liu, S., et al. (2021b). Asporin represses gastric cancer apoptosis via activating LEF1-mediated gene transcription independent of β-catenin. *Oncogene* 40 (27), 4552–4566. doi:10.1038/s41388-021-01858-7

Zhao, R., Peng, C., Song, C., Zhao, Q., Rong, J., Wang, H., et al. (2020). BICC1 as a novel prognostic biomarker in gastric cancer correlating with immune infiltrates. *Int. Immunopharmacol.* 87, 106828. doi:10.1016/j.intimp.2020.106828

Zheng, L., Liang, X., Li, S., Li, T., Shang, W., Ma, L., et al. (2018). CHAF1A interacts with TCF4 to promote gastric carcinogenesis via upregulation of c-MYC and CCND1 expression. *EBioMedicine* 38, 69–78. doi:10.1016/j.ebiom.2018.11.009

Zhou, L., Huang, W., Yu, H. F., Feng, Y. J., and Teng, X. (2020). Exploring TCGA database for identification of potential prognostic genes in stomach adenocarcinoma. *Cancer Cell Int.* 20, 264. doi:10.1186/s12935-020-01351-3

Zhu, Z., Yu, Z., Rong, Z., Luo, Z., Zhang, J., Qiu, Z., et al. (2019). The novel GINS4 axis promotes gastric cancer growth and progression by activating Rac1 and CDC42. *Theranostics* 9 (26), 8294–8311. doi:10.7150/thno.36256

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 06 January 2023
DOI 10.3389/fmolb.2022.1030057



OPEN ACCESS

EDITED BY
Georgina Gonzalez-Avila,
Instituto Nacional de Enfermedades Respiratorias-México (INER), Mexico

REVIEWED BY
Kazuma Sakamoto,
Nagoya University, Japan
Sachchida Nand Rai,
University of Allahabad, India

*CORRESPONDENCE
Wenqing Gao,
gaowenqing0106@163.com
Tong Li,
litong3zx@sina.com

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

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 28 August 2022
ACCEPTED 14 December 2022
PUBLISHED 06 January 2023

CITATION
Zheng Y, Lang Y, Qi B, Wang Y, Gao W and Li T (2023), TSPAN4 is a prognostic and immune target in Glioblastoma multiforme.
*Front. Mol. Biosci.* 9:1030057.
doi: 10.3389/fmolb.2022.1030057



# TSPAN4 is a prognostic and immune target in Glioblastoma multiforme

Yue Zheng[1,2,3,4,5†], Yuheng Lang[2,3,4,5,6†], Bingcai Qi[3,4,5,6†], Yuchao Wang[1,3,4,5,6], Wenqing Gao[2,3,4,5,6*] and Tong Li[1,2,3,4,5,6*]

[1]School of Medicine, Nankai University, Tianjin, China, [2]Department of Heart Center, The Third Central Hospital of Tianjin, Tianjin, China, [3]Nankai University Affiliated Third Center Hospital, Tianjin, China, [4]Tianjin Key Laboratory of Extracorporeal Life Support for Critical Diseases, Tianjin, China, [5]Artificial Cell Engineering Technology Research Center, Tianjin, China, [6]The Third Central Clinical College of Tianjin Medical University, Tianjin, China

**Background:** Atherosclerosis can impact cancer progression due to the cholesterol and calcium metabolism, illustrating the links between atherosclerosis and cancer metastasis. Tetraspanin 4 (TSPAN4) may help understand migrasomes in diseases and provide novel targets for treatment.

**Methods:** TSPAN4 expression in atherosclerosis Gene Expression Omnibus (EO) dataset and multiple omics data were explored, such as enriched pathways analysis, protein-protein interaction analysis, immune subtypes as well as diagnostic and prognostic value in pan-cancer. The relationship between Glioblastoma multiforme (GBM) and TSPAN4 was further investigated.

**Results:** Compared to control, TSPAN4 expression was upregulated in foam cells from patients with atherosclerosis and survival analysis demonstrated high TSPAN4 expression contributes to poor prognosis. TSPAN4 expression differs significantly in immune subtypes of cancers, which can be a diagnostic and prognostic target of cancers due to the high accuracy. Overall survival analysis of subgroups demonstrated that higher TSPAN4 expression had a worse prognosis and the univariate analysis and multivariate analysis demonstrated age, TSPAN4 expression, WHO grade, IDH status and histological types were independent risk factors of Glioblastoma multiforme.

**Conclusion:** The TSPAN4 expression was associated with atherosclerosis progression and pan-cancer, especially in Glioblastoma multiforme and GBMLGG. Therefore, TSPAN4 may serve as a potential biomarker and the crosstalk between atherosclerosis and tumor progression. The results are not fully validated and further studies are still needed to validate *in vivo* and *in vitro*.

KEYWORDS

**atherosclerosis, GO/KEGG pathways, pan-cancer, GSEA, TSPAN4, macrophages, migrasomes, GBMLGG**

# 1 Introduction

Coronary artery disease (CAD) contributes to considerable mortality and morbidity, leading to over one in every seven deaths all over the world (Naghavi et al., 2017). The mortality of patients with atherosclerosis and acute myocardial infarction (MI) has increased 5.6-fold in the last 3 decades and obesity has become the major cause in patients with some chronic diseases, for instance, diabetes and CAD (Chang et al., 2017; Singh et al., 2020). Young patients with type 2 diabetes and MI develop higher long-term cardiovascular-related mortality and more than one-third die within 10 years (Dutta et al., 2012; Chang et al., 2017; Singh et al., 2020).

Tetraspanin 4 (TSPAN4), as a required protein for migrasome formation (Jiang et al., 2019), can interact with the histamine H4 receptor (H4R) in transfected cells (Ma et al., 2021). Previous reports demonstrated that podocytes release the "injury-related" migrasomes (Liu et al., 2020) and migrasome formation is mediated by the assembly of micron-scale tetraspanin macrodomains, for instance, TSPAN4 (Huang et al., 2019). Therefore, the investigation of TSPAN4 functions may help understand migrasomes in diseases and provide novel targets for treatment.

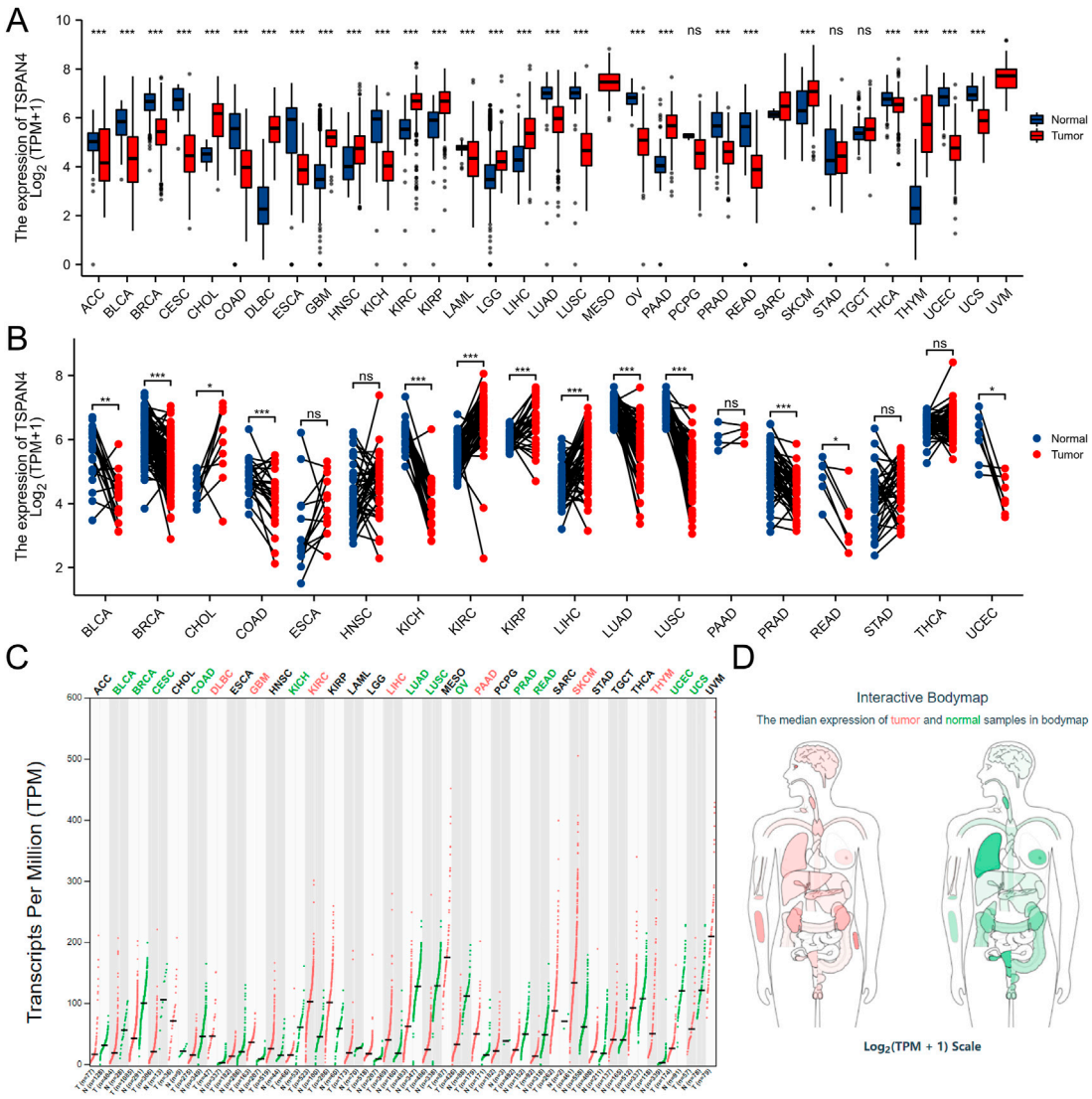



FIGURE 1
TSPAN4 expression in pan-cancer. **(A)** Unpaired samples in pan-cancer demonstrated there were significant differences of TSPAN4 expression in 27 cell lines in TCGA. **(B)** Paired samples in pan-cancer demonstrated there were significant differences of TSPAN4 expression in 13 cell lines in TCGA. **(C)** Using GEPIA, TSPAN4 was highly expressed in 7 tumor cell lines and lowly expressed in 12 tumor cell lines. **(D)** The median expression of tumor and normal samples in bodymap. Red, tumor samples; green, normal samples.

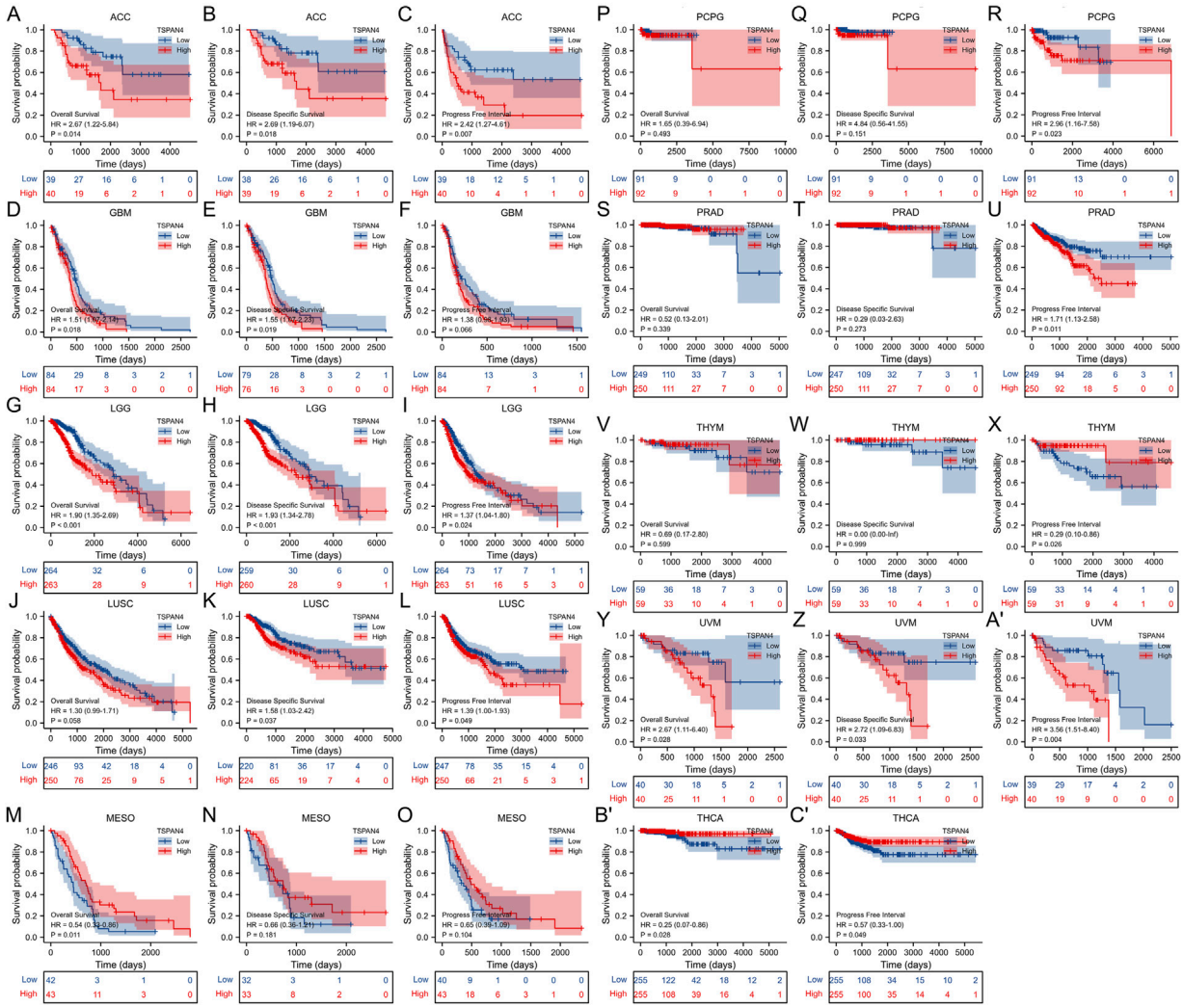



FIGURE 2
Prognostic value of TSPAN4 expression in pan-cancer. **(A–C)** The overall survival analysis **(A)**, diseases specific survival analysis **(B)** and progress free interval analysis **(C)** of TSPAN4 expression in ACC. **(D–F)** The overall survival analysis **(D)**, diseases specific survival analysis **(E)** and progress free interval analysis **(F)** of TSPAN4 expression in GBM. **(G–I)** The overall survival analysis **(G)**, diseases specific survival analysis **(H)** and progress free interval analysis **(I)** of TSPAN4 expression in LGG. **(J–L)** The overall survival analysis **(J)**, diseases specific survival analysis **(K)** and progress free interval analysis **(L)** of TSPAN4 expression in LUSC. **(M–O)** The overall survival analysis **(M)**, diseases specific survival analysis **(N)** and progress free interval analysis **(O)** of TSPAN4 expression in MESO. **(P–R)** The overall survival analysis **(P)**, diseases specific survival analysis **(Q)** and progress free interval analysis **(R)** of TSPAN4 expression in PCPG. **(S–U)** The overall survival analysis **(S)**, diseases specific survival analysis **(T)** and progress free interval analysis **(U)** of TSPAN4 expression in PRAD. **(V–X)** The overall survival analysis **(V)**, diseases specific survival analysis **(W)** and progress free interval analysis **(X)** of TSPAN4 expression in THYM. **(Y-A')** The overall survival analysis **(Y)**, diseases specific survival analysis **(Z)** and progress free interval analysis **(A')** of TSPAN4 expression in UVM. **(B',C')** The overall survival analysis **(B')** and progress free interval analysis **(C')** of TSPAN4 expression in THCA.

Aseptic inflammation can promote neutrophil extracellular traps (NETs) increase in the liver, thus promoting cancer metastasis (Yang et al., 2020). Graeme et al. also reported MI can epigenetically reprogram Ly6C$^{high}$ monocytes in the bone marrow, which were increasingly recruited to the breast cancer microenvironment and promoted MI-induced early-stage breast cancer progression and increased breast cancer patients' mortality and morbidity (Koelwyn et al., 2020). Atherosclerosis can impact cancer progression due to the cholesterol and calcium metabolism, illustrating the links between atherosclerosis and cancer metastasis (DUNGAL and BENEDIKTSSON, 1958; Tapia-Vieyra et al., 2017; Balzan and

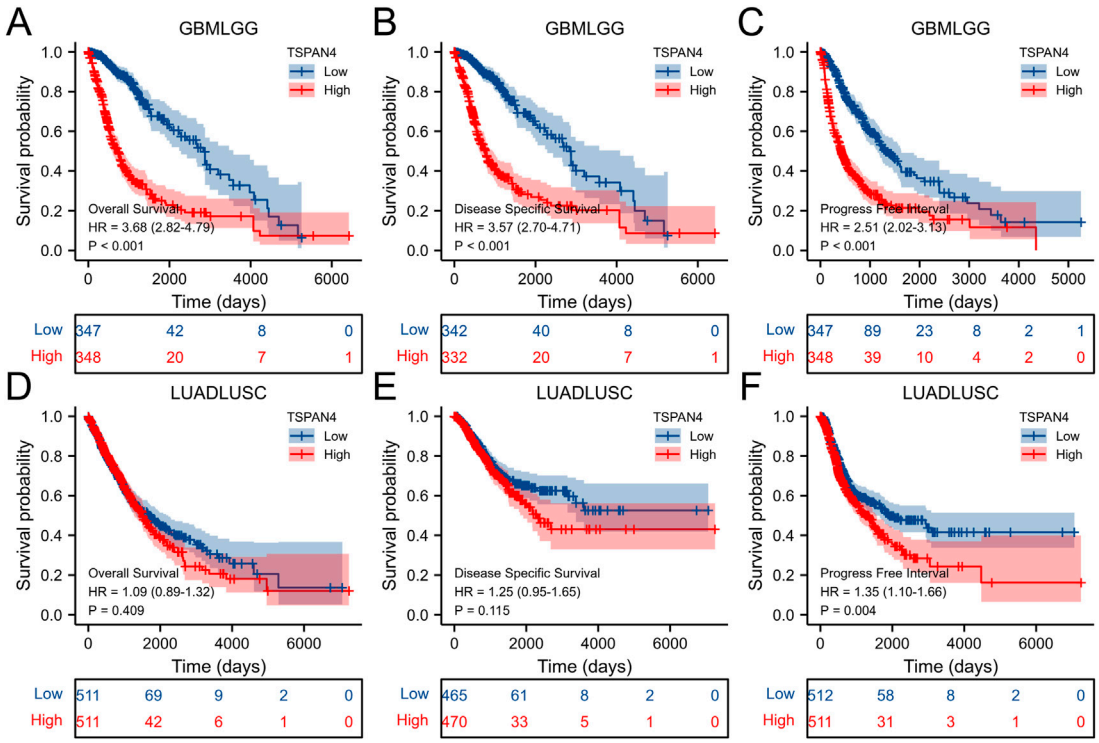



FIGURE 3
Prognostic value of TSPAN4 expression in GBMLGG and LUADLUSC. **(A–C)** The overall survival analysis **(A)**, diseases specific survival analysis **(B)** and progress free interval analysis **(C)** of TSPAN4 expression in GBMLGG. **(D–F)** The overall survival analysis **(D)**, diseases specific survival analysis **(E)** and progress free interval analysis **(F)** of TSPAN4 expression in LUADLUSC.

Lubrano, 2018; Mensah et al., 2021). Therefore, the investigation between CAD, especially MI and atherosclerosis, and tumor progression may provide potential biomarkers to impede tumor progression, thus reducing mortality and morbidity. However, the researches on such crosstalk are little.

In this study, the differentially expressed genes (DEGs) were investigated in foam cells from subjects with atherosclerosis in GSE9874 and enrichment pathways were applied to explore macrophage-related DEGs. Next, we determined the expression of TSPAN4 in atherosclerosis and pan-cancer and the prognostic and diagnostic value of TSPAN4 in pan-cancer, and the correlation between the TSPAN4 expression and immune cells were evaluated as well. DEGs between TSPAN4 high and low expression groups were also explored to validate whether TSPAN4 can be the crosstalk between MI and cancer progression. After that, the subgroup analysis, survival analysis and prognostic value of Glioblastoma multiforme and Brain Lower Grade Glioma (GBMLGG) were determined. Finally, the validation of TSPAN4 expression in macrophages was also carried out.

# 2 Materials and methods

## 2.1 Data source and processing

Using the keywords "atherosclerosis" in "*Homo sapiens*," GSE9874 from the Gene Expression Omnibus (GEO) database was investigated, processed with log2 transformation for normalization and analyzed using the limma package in R (Hägg et al., 2008). The RNA-sequencing of macrophages was based on the Affymetrix Human Genome U133A Array.

The Cancer Genome Atlas (TCGA) database and the Genotype-Tissue Expression (GTEx) database by UCSC XENA (https://portal.gdc.cancer.gov/) were investigated. The gene expression data TCGA pan-cancer, including unpaired samples and paired samples, were analyzed using Xiantao website tool (www.xiantao.love) and GEPIA (cancer-pku.cn) (Tang et al., 2017).

The single-cell sequencing data about TSPAN4 expression in GBMLGG were used from Single Cell Portal (Study: BTSC dependence on GLS reveals a metabolic vulnerability, https://singlecell.broadinstitute.org/) (Restall et al., 2020).

## 2.2 Enrichment pathways analysis

GO/KEGG pathways were explored among macrophage-related DEGs between patients with subclinical atherosclerosis and control (Ashburner et al., 2000; Kanehisa and Goto, 2000; Subramanian et al., 2005a; Yu et al., 2012). A $p$-value < .05 were regarded as the cut-off criterion.

## 2.3 Prognostic value of TSPAN4 expression

Overall survival (OS), Disease Specific Survival (DSS) and Progress Free Interval (PFI) were explored in pan-cancer. In addition, the TSPAN4 expression was also investigated in GBMLGG, Colon adenocarcinoma and Rectum adenocarcinoma Esophageal carcinoma (COADREAD) and Lung adenocarcinoma and Lung squamous cell carcinoma (LUADLUSC).

## 2.4 Diagnostic value of TSPAN4 expression

The receiver operation characteristic (ROC) curve was conducted to investigate the diagnostic performance of TSPAN4 expression and the area under the curve (AUC) was determined using the "pROC" package.

## 2.5 TSPAN4 expression association with immune cells

ssGSEA (GSVA package in R) was used to provide a critical assessment of the relationships between TSPAN4 expression and macrophages in TCGA data (Subramanian et al., 2005b; Yu et al., 2012). A $p$-value < .05 were regarded as significant.

## 2.6 DEGs between TSPAN4 high and low expression groups

The differentially expressed genes (DEGs) between different TSPAN4 expression groups (low expression group: 0%–50%; high expression group: 50%–100%) in LUADLUSC, PRAD, THYM and GBMLGG. A log2(Fold Change) > 1 and a $p$-value < .05 were applied as the cut-off criteria of PRAD and a log2(Fold Change) > 2 and a $p$-value < .05 were applied as the cut-off criteria of LUADLUSC, THYM and GBMLGG.

## 2.7 Analysis of MMR in cancer

The correlation between TSPAN4 expression and several essential mismatch repair (MMR) genes was investigated in LUSC, LUADLUSC, PRAD, THYM, GBM and GBMLGG, including the MutL homologous gene (MLH1), MutS homologous gene (MSH2, MSH6), increased separation after meiosis (PMS2) and epithelial cell adhesion molecule (EPCAM).

## 2.8 Subgroup analysis and survival analysis in GBMLGG

To validate the potential effects of TSPAN4 expressions on GBMLGG progression, the TSPAN4 expressions in subgroups were determined and OS analysis of subgroups was also carried out. Cox regression analysis was used to explore the prognosis of ARL6IP5 expression in subgroups.

## 2.9 Co-expression gene analysis of TSPAN4 expression in GBMLGG

The top 50 co-expression genes positively and negatively related to TSPAN4 expression in GBMLGG were explored. GO/KEGG pathways analysis was used to investigate the related enriched pathways.

## 2.10 Protein-protein interaction (PPI) and the hub genes

To investigate TSPAN4 and its protein interactions, STRING database (https://string-db.org) was used with a combined score >.4 (Szklarczyk et al., 2019). The nodes were analyzed using Cytoscape v.3.7.1 (Shannon et al., 2003). PPI network analysis was used to obtain the hub genes utilizing the Cytoscape plug-in MCODE. The Cytoscape plug-in cytohubba was also used to obtain the genes in top15 MCC, top15 DMNC and top15 Degree.

## 2.11 Prognostic value of TSPAN4 expression in GBM

The nomogram, Calibration curves and time-dependent ROC analysis were used. Lasso regression and risk score were also utilized to explore the functions of TSPAN4 expression on GBM.

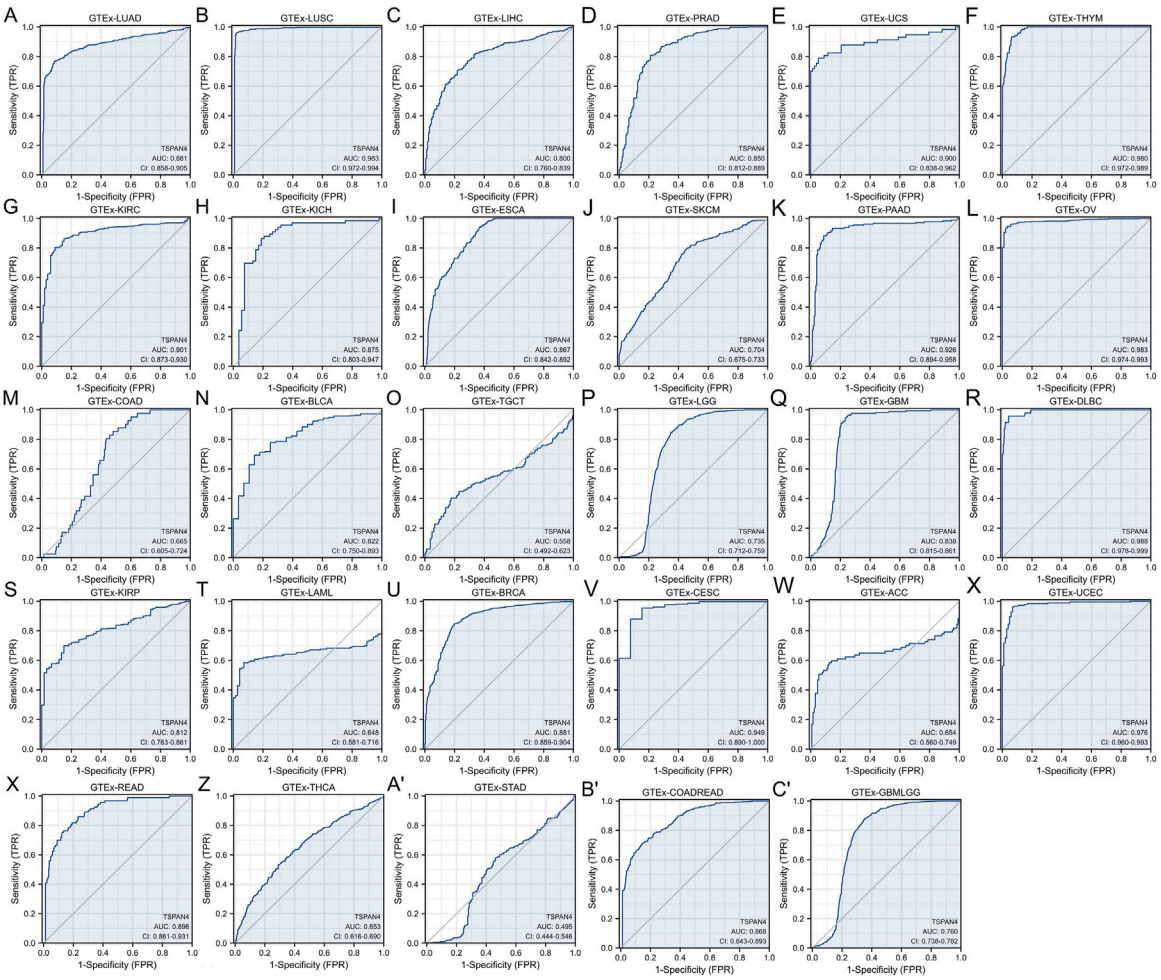



FIGURE 4
Diagnostic value of TSPAN4 expression in pan-cancer and GTEx, including GTEx-LUAD **(A)**, GTEx-LUSC **(B)**, GTEx-LIHC **(C)**, GTEx-PRAD **(D)**, GTEx-UCS **(E)**, GTEx-THYM **(F)**, GTEx-KIRC **(G)**, GTEx-KICH **(H)**, GTEx-ESCA **(I)**, GTEx-SKCM **(J)**, GTEx-PAAD **(K)**, GTEx-OV **(L)**, GTEx-COAD **(M)**, GTEx-BLCA **(N)**, GTEx-TGCT **(O)**, GTEx-LGG **(P)**, GTEx-GBM **(Q)**, GTEx-DLBC **(R)**, GTEx-KIRP **(S)**, GTEx-LAML **(T)**, GTEx-BRCA **(U)**, GTEx-CESC **(V)**, GTEx-ACC **(W)**, GTEx-UCEC **(X)**, GTEx-READ **(Y)**, GTEx-THCA **(Z)**, GTEx-STAD **(A′)**, GTEx-COADREAD **(B′)** and GTEx-GBMLGG **(C′)**.
TSPAN4 expression can diagnose 20 cancer types compared to normal samples (AUC > .8), including LUAD, LUSC, LIHC, PRAD, UCS, THYM, KIRC, KICH, ESCA, PAAD, OV, BLCA, GBM, DLBC, KIRP, BRCA, CESC, UCEC, READ and COADREAD.

## 2.12 Validation of TSPAN4 expression in macrophages

To further validate the TSPAN4 effects, the atherosclerosis mice model and macrophage cell lines were utilized. RAW264.7 and THP-1 macrophage cell lines were purchased from ATCC and adult experimental ApoE$^{-/-}$ male mice were purchased from Charles River. Mice were maintained in an SPF environment (temperature: 23°C–25°C; humidity: 55%–60%) with free access to food and water and a 12/12 light-dark cycle. Protocols were approved by the Institute of Radiation Medicine, the Chinese Academy of Medical Science, which conform to the Guide for the Care and Use of Laboratory Animals.

To construct the atherosclerosis model, the ApoE$^{-/-}$ male mice ($n$ = 6 per group) were fed with a chow diet or a western diet (HFHC100244) for 3 months or 6 months. In the atherosclerosis mice model, the aorta root was collected and qRT-PCR and ELISA were performed. In macrophages hypoxia cell model, the RAW264.7 and THP-1 macrophages were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (FBS, Auxgen) and treated in a hypoxia chamber filled with 5% $CO_2$, 94% $N_2$ and 1% $O_2$ at 37.0°C for 12 h or 24 h. Then the Tspan4 mRNA level was determined using qRT-PCR and the secreted TSPAN4 protein expression was evaluated using ELISA in a cultured medium.

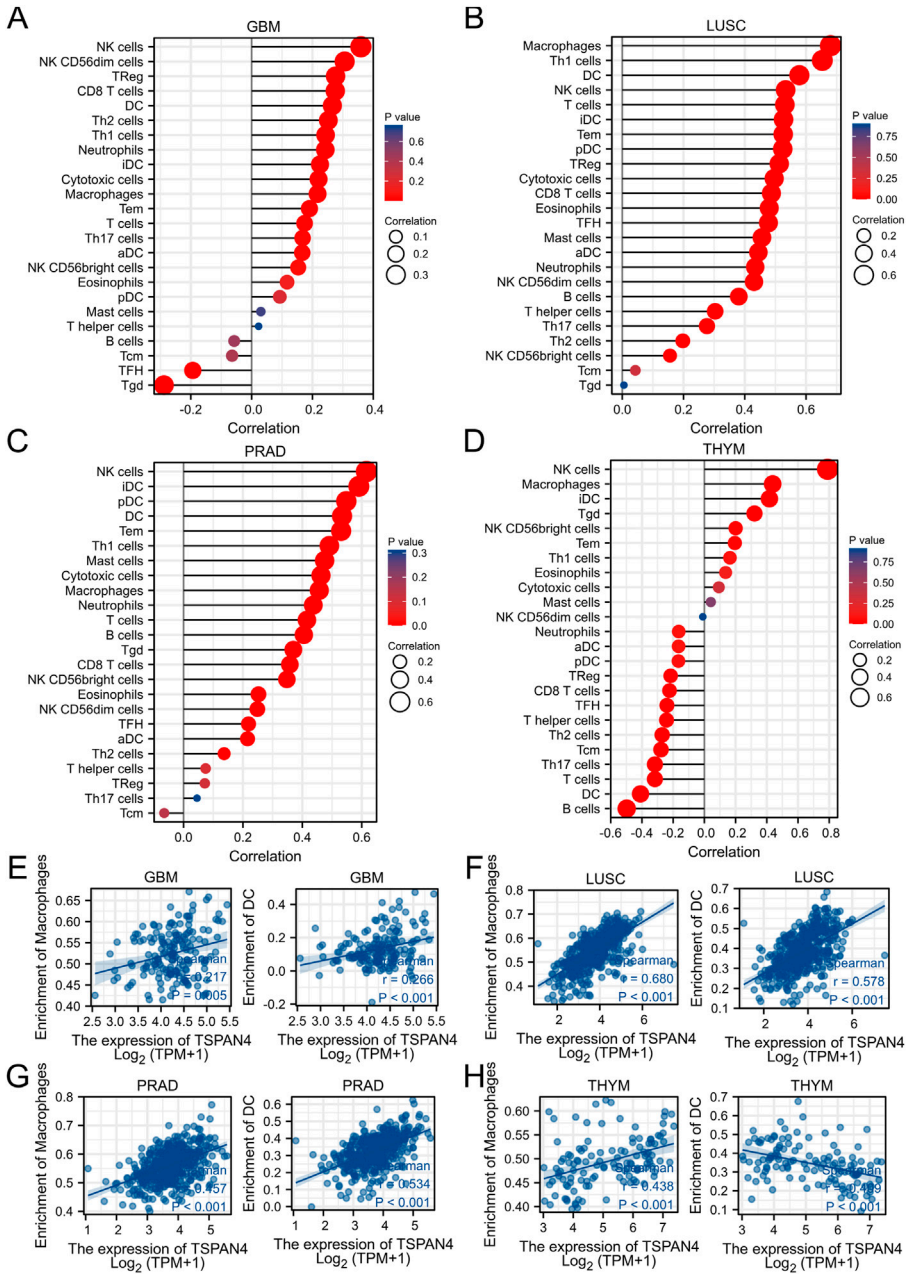



FIGURE 5
TSPAN4 expression and immune infiltration analysis. **(A–D)** The lollipop chart of 24 immune cells in GBM **(A)**, LUSC **(B)**, PRAD **(C)** and THYM **(D)**. **(E–H)** The correlation analysis between the abundance of macrophages or DC cells and TSPAN4 expression in GBM **(E)**, LUSC **(F)**, PRAD **(G)** and THYM **(H)**.

Total RNA was collected using TRIzol and RNA purity was tested. Total RNA was used for reverse transcription reaction and the amplified cDNA samples were mixed using TB Green® Premix Ex Taq™ II (TaKaRa, RR820). Finally, the reaction was conducted on AriaMx HRM. β-actin was used as a positive control and the $2^{-\Delta\Delta CT}$ method was used to calculate each sample. Tspan4: F: TAC CTCATGTTCGCCTTCAAC; R: GATAAGGTGGCAAAGTTT CCCT. β-actin: F: GGCTGTATTCCCCTCCATCG; R: CCAGTT GGTAACAATGCCATGT.

ELISA was performed. The TSPAN4 protein in the aorta root was extracted according to the Kit Instruction (BC3710, Solarbio). For further validation, the TSPAN4 protein levels

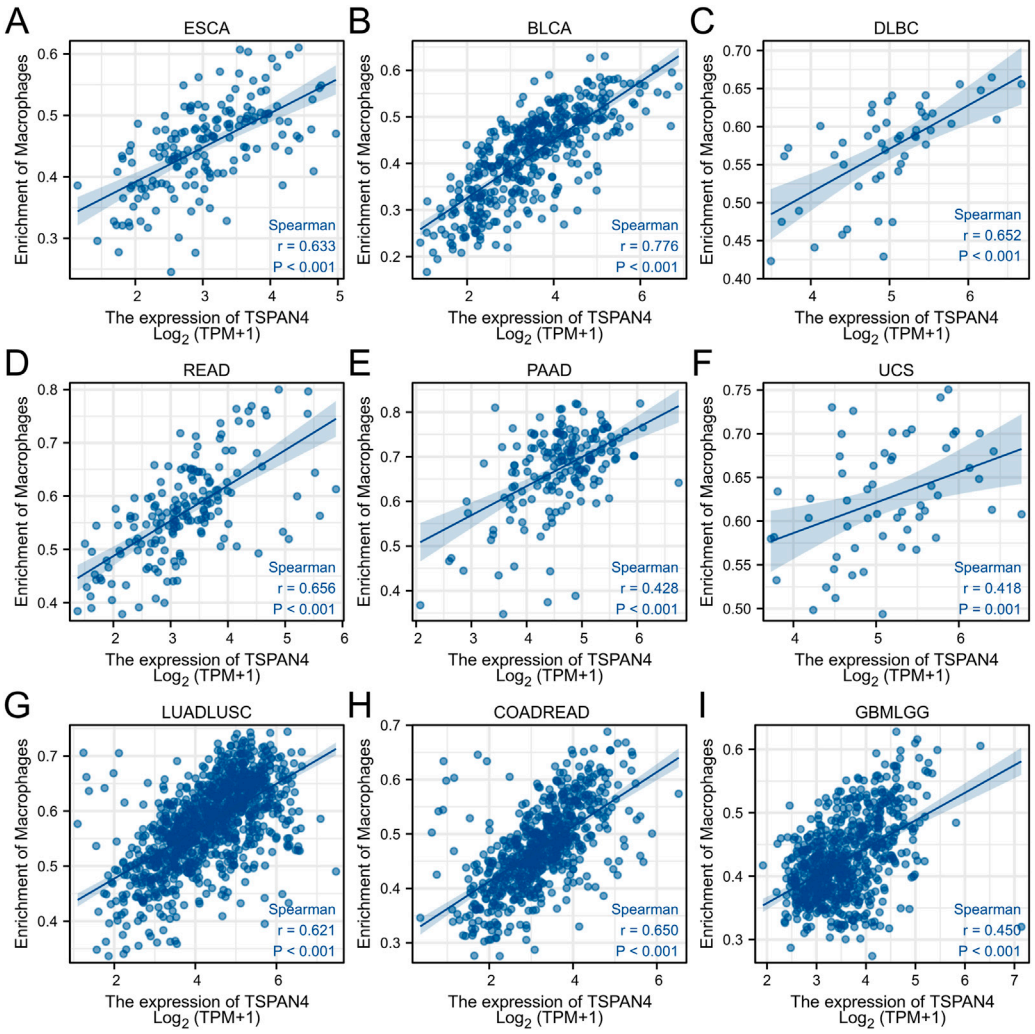



FIGURE 6
The correlation analysis between the abundance of macrophages and TSPAN4 expression in other cell lines, including ESCA **(A)**, BLCA **(B)**, DLBC **(C)**, READ **(D)**, PAAD **(E)**, UCS **(F)**, LUADLUSC **(G)**, COADREAD **(H)** and GBMLGG **(I)**.

in the aorta root and macrophages cultured medium were explored with the 4°C incubation overnight of mouse TSPAN4 primary antibodies (1:600, LS-C676657-50, LSBio). The serum matrix (MILLIPLEX Analyst Kit, Millipore) was used as the positive control in the validation.

## 2.13 Statistical analysis

All data are presented as the mean ± SD. Statistical analyses were performed using SPSS 23.0. Shapiro-Wilk normality test and Wilcoxon rank sum test were used and a value of $p < .05$ was considered statistically significant.

# 3 Results

## 3.1 TSPAN4 was highly expressed in atherosclerosis

TSPAN4 was highly expressed in foam cells from the patients with atherosclerosis, while there was no significant difference in TSPAN4 expression in foam cells from the subjects without atherosclerosis compared to baseline (Supplementary Figure S1). GO/KEGG pathways of DEGs in atherosclerosis were also applied, which were mainly involved in regulation of lipid metabolic process, monooxygenase activity, and NF-kappa B signaling pathway (Supplementary Figure S2; Supplementary Tables S1, S2).

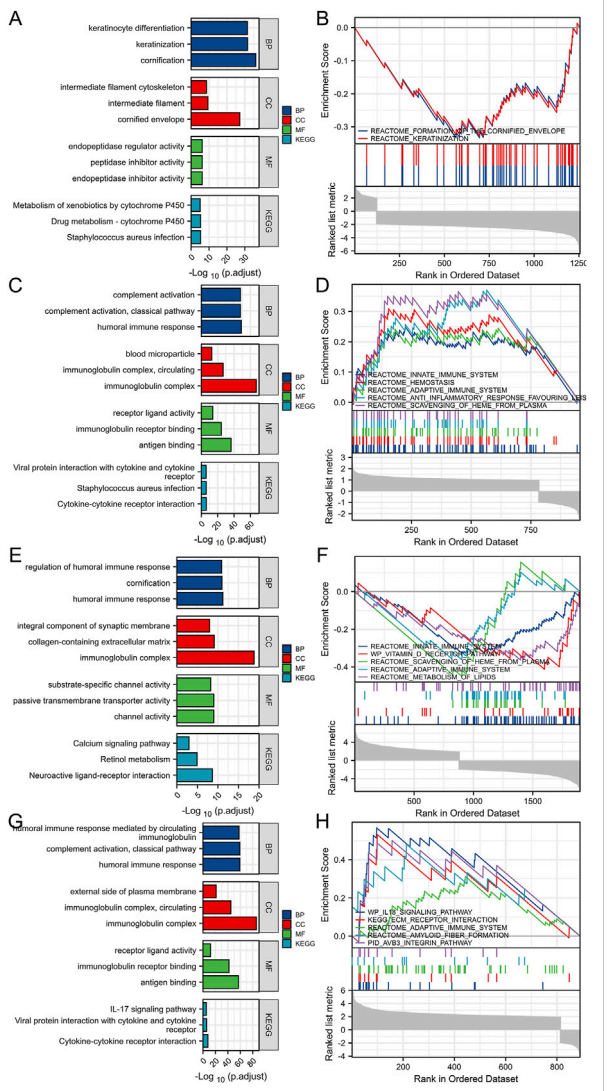



FIGURE 7
DEGs between TSPAN4 high expression and low expression groups in LUADLUSC, PRAD, THYM and GBMLGG. **(A,B)** The bar plot of GO/KEGG pathways **(A)** and GSEA **(B)** enriched by DEGs between TSPAN4 high expression and low expression groups in LUADLUSC. **(C,D)** The bar plot of GO/KEGG pathways **(C)** and GSEA **(D)** enriched by DEGs between TSPAN4 high expression and low expression groups in PRAD. **(E,F)** The bar plot of GO/KEGG pathways **(E)** and GSEA **(F)** enriched by DEGs between TSPAN4 high expression and low expression groups in THYM. **(G, H)** The bar plot of GO/KEGG pathways **(G)** and GSEA **(H)** enriched by DEGs between TSPAN4 high expression and low expression groups in GBMLGG.

## 3.2 TSPAN4 was highly expressed in pan-cancer

The gene expression data in TCGA pan-cancer were extracted, filtered and analyzed (Figure 1). Unpaired samples in pan-cancer demonstrated there were significant differences in TSPAN4 expression in 27 cell lines in TCGA, including Adrenocortical carcinoma (ACC), Bladder Urothelial Carcinoma (BLCA), Breast invasive carcinoma (BRCA), Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), Cholangiocarcinoma (CHOL), Colon adenocarcinoma (COAD), Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC), Esophageal carcinoma (ESCA), GBM, Head and Neck squamous cell carcinoma (HNSC), Kidney Chromophobe (KICH), Kidney renal clear cell carcinoma (KIRC), Kidney renal papillary cell carcinoma (KIRP), Acute Myeloid Leukemia (LAML), Brain Lower Grade Glioma (LGG), Liver hepatocellular carcinoma (LIHC), Lung adenocarcinoma (LUAD), LUSC, Ovarian serous cystadenocarcinoma (OV), Pancreatic adenocarcinoma (PAAD), PRAD, Rectum adenocarcinoma (READ), Skin Cutaneous Melanoma (SKCM), Thyroid carcinoma (THCA), THYM, Uterine Corpus Endometrial Carcinoma (UCEC) and Uveal Melanoma (UCS). Paired samples in pan-cancer demonstrated there were significant differences in TSPAN4 expression in 13 cell lines in TCGA, including BLCA, BRCA, CHOL, COAD, KICH, KIRP, LIHC, LUAD, LUSC, PRAD, READ and UCEC. Using GEPIA, TSPAN4 was highly expressed in 7 tumor cell lines, including DLBC, GBM, KIRC, LIHC, PAAD, SKCM and THYM, and lowly expressed in 12 tumor cell lines, including BLCA, BRCA, CESC, COAD, KICH, LUAD, LUSC, OV, PRAD, READ, UCEC and UCS.

Pan-cancer analysis was also used in TSPAN7 expression (Supplementary Figure S3). Unpaired samples in pan-cancer demonstrated there were significant differences in TSPAN7 expression in 17 cell lines in TCGA. Paired samples in pan-cancer demonstrated there were significant differences in TSPAN7 expression in 13 cell lines in TCGA.

## 3.3 Prognostic value of TSPAN4 expression in pan-cancer

The overall survival analysis, diseases specific survival analysis and progress free interval analysis of TSPAN4 expression demonstrated that high TSPAN4 expression was correlated to poor prognosis of the patients with ACC, GBM, LGG, LUSC, MESO, PCPG, PRAD, THYM, UVM, THCA, GBMLGG or LUADLUSC (Figures 2, 3).

## 3.4 Diagnostic value of TSPAN4 expression in pan-cancer

To explore the diagnostic value of TSPAN4 expression in pan-cancer, ROC analysis was used, suggesting that the TSPAN4 expression can be credible biomarkers in LUSC, UCS, THYM, KIRC, PAAD, OV, DLBC, CESC and UCEC (AUC > .9), and TSPAN4 expression can also be potential

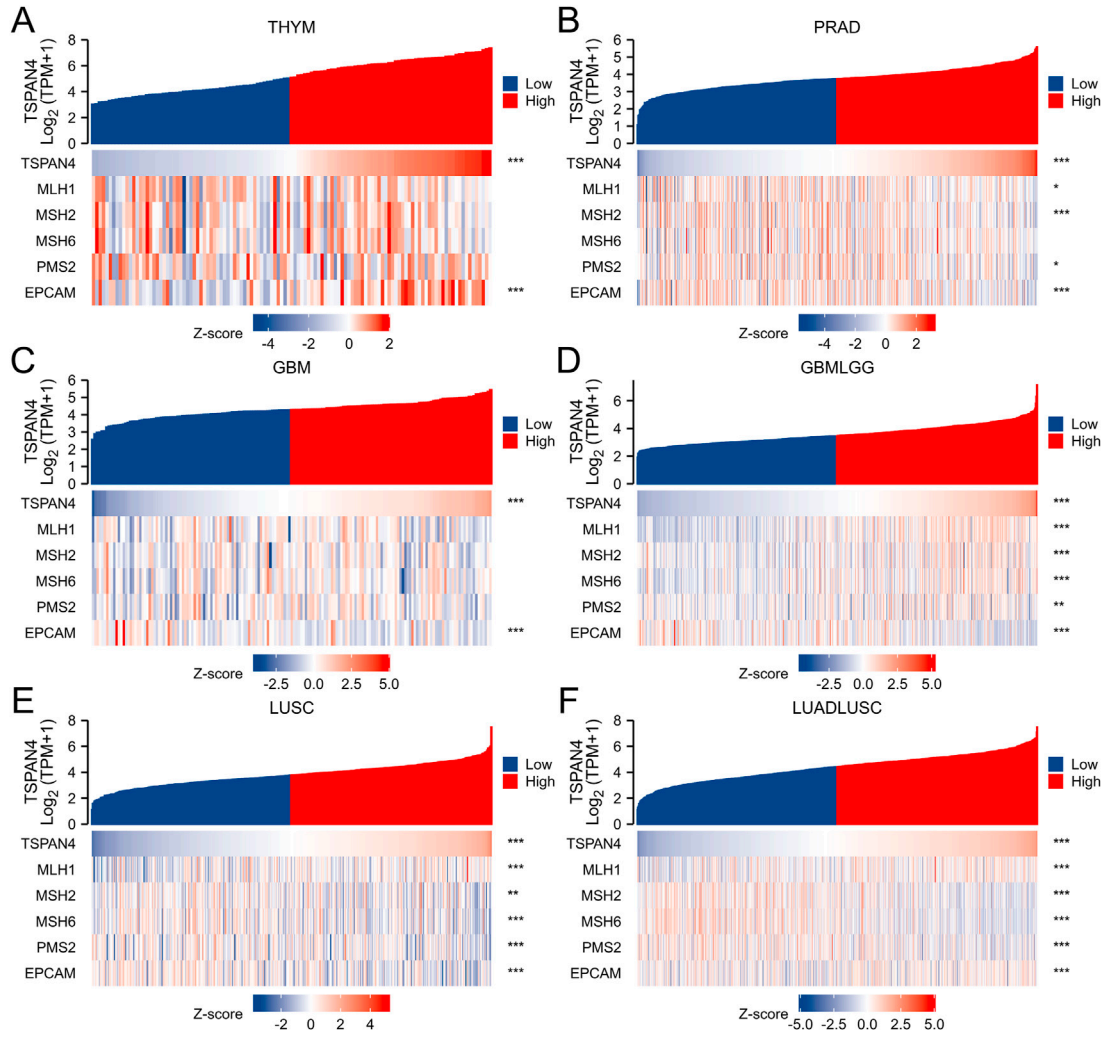



FIGURE 8
Analysis of MMR in cancer. The correlation between TSPAN4 expression and several essential mismatch repair genes, including MLH1, MSH2, MSH6, PMS2 and EPCAM, was explored in THYM **(A)**, PRAD **(B)**, GBM **(C)**, GBMLGG **(D)**, LUSC **(E)** and LUADLUSC **(F)**.

biomarkers in LUAD, LIHC, PRAD, KICH, ESCA, BLCA, GBM, KIRP, BRCA, READ and COADREAD (AUC > .8) (Figure 4).

After evaluation, the TSPAN4 expressions in GBM, LUSC, PRAD and THYM, were further analyzed according to the high prognostic and diagnostic values.

## 3.5 TSPAN4 expression and immune cells analysis

ssGSEA was used in GBM, LUSC, PRAD and THYM, illustrating that TSPAN4 expression was positively correlated to the abundance of macrophages and DC cells in GBM, LUSC and PRAD as well as the abundance of macrophages in THYM. Interestingly, TSPAN4 expression was negatively correlated to the abundance of DC cells (Figure 5). In addition, TSPAN4 expression was highly positively correlated to the abundance of macrophages in ESCA, BLCA, DLBC, READ, PAAD, UCS, LUADLUSC, COADREAD and GBMLGG (Figure 6).

## 3.6 DEGs between TSPAN4 high and low expression groups

After Log2 transformation, DEGs between TSPAN4 high and low expression groups in LUADLUSC were obtained and further GO/KEGG pathway analysis was applied, which were mainly involved in cornification, keratinocyte differentiation,

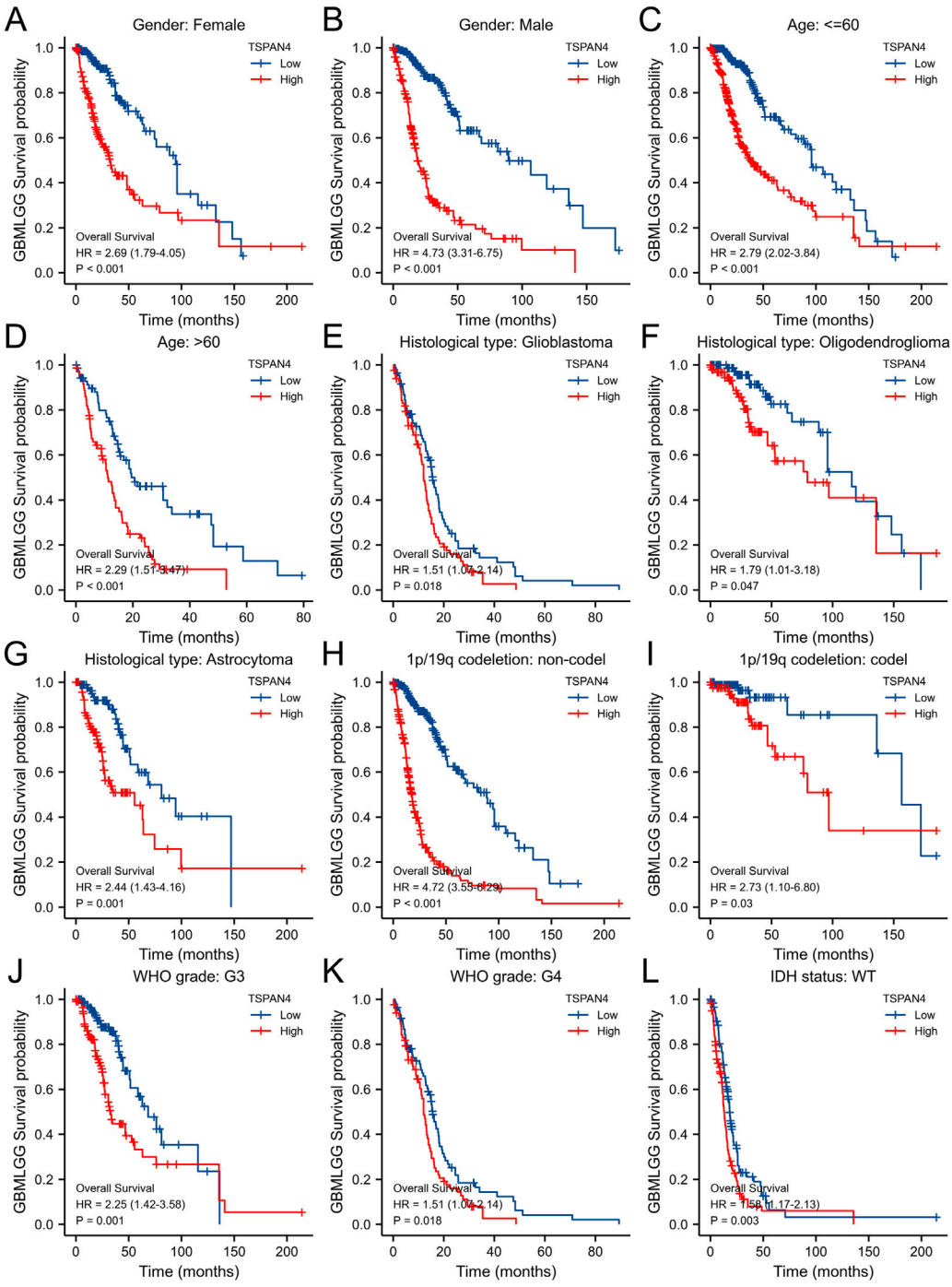



FIGURE 9
The associations between TSPAN4 expression and the overall survival in different clinical subgroups of GBMLGG. **(A)** Gender: Female; **(B)** Gender: Male; **(C)** Age: ≤60; **(D)** Age: >60; **(E)** Histological type: Glioblastoma; **(F)** Histological type: Oligodendroglioma; **(G)** Histological type: Astrocytoma; **(H)** 1p/19q codeletion: non-codel; **(I)** 1p/19q codeletion: codel; **(J)** WHO grade: G3; **(K)** WHO grade: G4; and **(L)** IDH status: WT.

endopeptidase inhibitor activity, and intermediate filament cytoskeleton. GSEA of DEGs between TSPAN4 high and low expression groups was explored, which were mainly enriched in REACTOME_FORMATION_OF_THE_CORNIFIED_ENVELOPE and REACTOME_KERATINIZATION (Figures 7A, B).

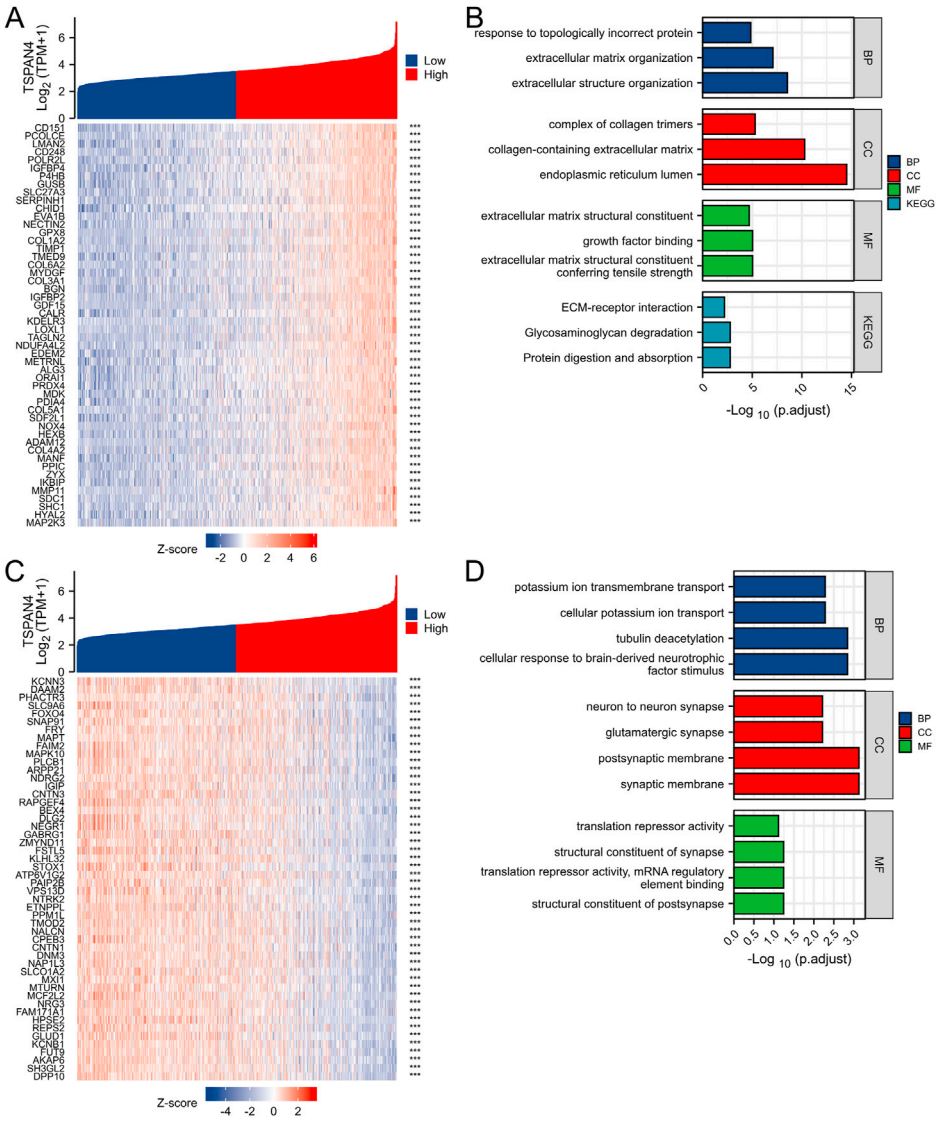



FIGURE 10
Top 50 genes positively and negatively correlated with TSPAN4 expression in GBMLGG. **(A)** The gene co-expression heatmap of the top 50 genes positively correlated with TSPAN4 expression in GBMLGG. **(B)** The bar plot of GO/KEGG pathways enriched by the top 50 genes positively correlated with TSPAN4 expression. **(C)** The gene co-expression heatmap of the top 50 genes negatively correlated with TSPAN4 expression in GBMLGG. **(D)** The bar plot of GO/KEGG pathways enriched by the top 50 genes negatively correlated with TSPAN4 expression.

DEGs between TSPAN4 high and low expression groups in PRAD were obtained and further GO/KEGG pathway analysis was applied, which were mainly involved in humoral immune response, protein activation cascade, and receptor ligand activity. GSEA of DEGs between TSPAN4 high and low expression groups was explored, which were mainly enriched in REACTOME_HEMOSTASIS, REACTOME_ANTI_INFLAMMATORY_RESPONSE_FAVOURING_LEISHMANIA_PARASITE_INFECTION, and REACTOME_VESICLE_MEDIATED_TRANSPORT (Figures 7C, D).

DEGs between TSPAN4 high and low expression groups in THYM were obtained and further GO/KEGG pathway analysis was applied, which were mainly involved in regulation of humoral immune response, protein activation cascade, collagen-containing extracellular matrix, intrinsic component of synaptic membrane, and channel activity. GSEA of DEGs between TSPAN4 high and low expression groups was explored, which were mainly enriched in REACTOME_INNATE_IMMUNE_SYSTEM, REACTOME_INITIAL_TRIGGERING_OF_COMPLEMENT, and REACTOME_CD22_MEDIATED_BCR_REGULATION (Figures 7E, F).

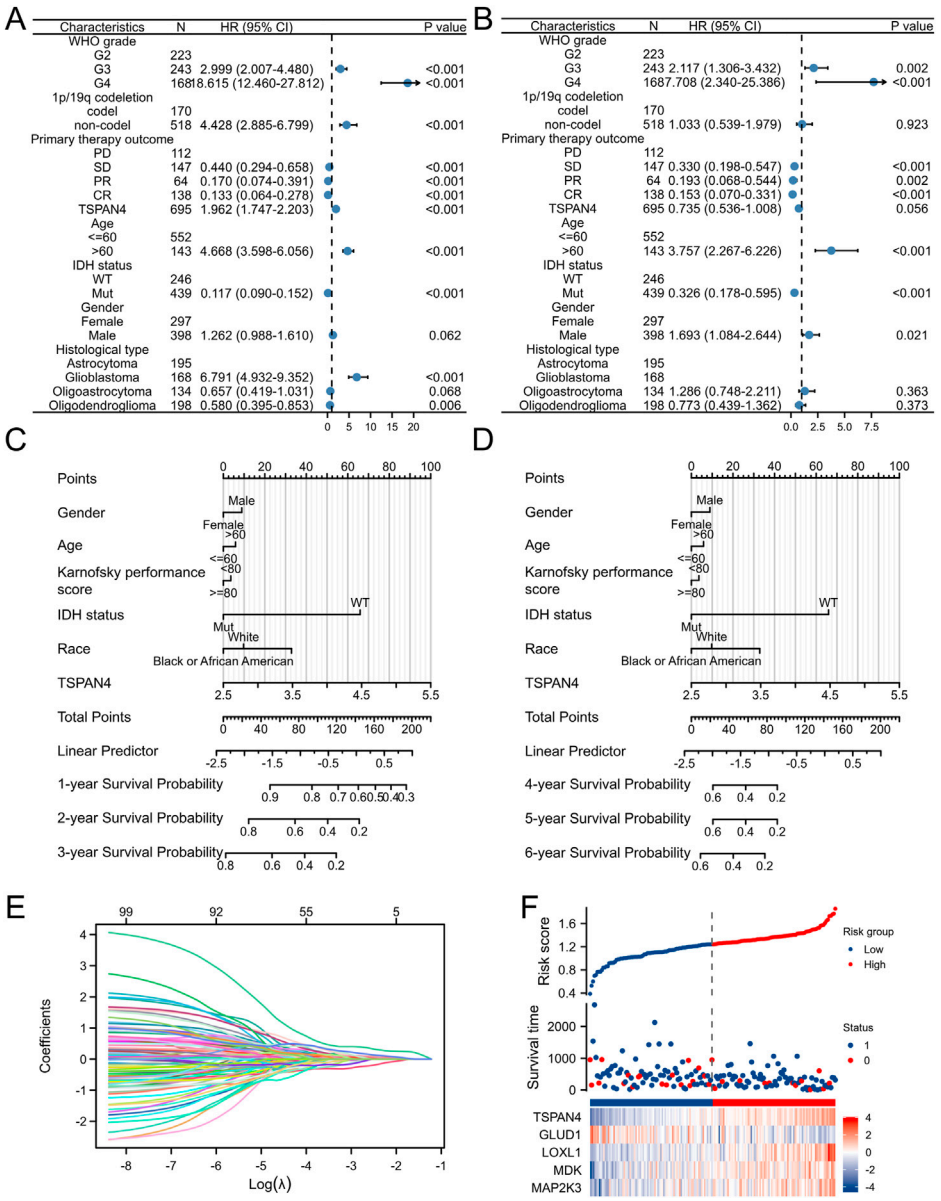



FIGURE 11
Prognostic value of TSPAN4 expression in GBM. **(A,B)** The forest diagrams of univariate analysis **(A)** and multivariate analysis **(B)** of patients' characteristics. **(C)** The nomogram about 1-, 2- and 3-year survival probability of TSPAN4 expression and other clinical characteristics in GBM. **(D)** The nomogram about 4-, 5- and 6-year survival probability of TSPAN4 expression and other clinical characteristics in GBM. **(E,F)** The lasso regression **(E)** and risk score analysis **(F)** of survival time and TSPAN4 expression in GBM. 1, survival; 0, dead.

DEGs between TSPAN4 high and low expression groups in GBMLGG were obtained and further GO/KEGG pathway analysis was applied, which were mainly involved in immunoglobulin complex, external side of plasma membrane, and antigen binding. GSEA of DEGs between TSPAN4 high and low expression groups was explored, which were mainly enriched in REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION, WP_IL18_SIGNALING_PATHWAY, and KEGG_ECM_RECEPTOR_INTERACTION (Figures 7G, H).

## 3.7 Analysis of MMR in cancer

To investigate the correlation between TSPAN4 expression and several essential mismatch repair genes, including MLH1, MSH2, MSH6, PMS2 and EPCAM, the correlation heatmap of several gene expressions was explored in THYM, PRAD, GBM, GBMLGG, LUSC and LUADLUSC (Figure 8). MLH1, MSH2, MSH6, PMS2 and EPCAM expressions were correlated to TSPAN4 expression in GBMLGG, LUSC and LUADLUSC.

MLH1, MSH2, PMS2 and EPCAM expressions were correlated to TSPAN4 expression in PRAD, while only EPCAM expression was correlated to TSPAN4 expression in THYM and GBM.

## 3.8 Subgroup analysis and survival analysis

To validate the potential effects of TSPAN4 expression in GBMLGG, the TSPAN4 expressions in subgroups were determined (Figure 9). Using survival analysis, high TSPAN4 expression contributed to a worse prognosis in OS of GBMLGG subgroup patients, such as patients with Glioblastoma, Oligodendroglioma, Astrocytoma, WHO grade: G3, or WHO grade: G4.

## 3.9 Co-expression gene analysis of TSPAN4 expression in GBMLGG

The top 50 co-expression genes positively related to TSPAN4 expression in GBMLGG were explored, which were mainly involved in extracellular structure organization, endoplasmic reticulum lumen, collagen-containing extracellular matrix, and growth factor binding (Figures 10A, B; Supplementary Figure S4). The top 50 co-expression genes negatively related to TSPAN4 expression in GBMLGG were also investigated, which were mainly involved in cellular potassium ion transport, synaptic membrane, postsynaptic membrane, and glutamatergic synapse (Figures 10C, D; Supplementary Table S5).

## 3.10 PPI network and the hub genes

To investigate TSPAN4 and its protein interactions in GBMLGG, the nodes with a combined score >.4 were analyzed using STRING. The hub genes were obtained utilizing the Cytoscape plug-in MCODE and there were 4 modules in the network (Supplementary Figure S4). Venn diagram of top15 MCC, DMNC and degree demonstrated several targets, including CXCL11, LOX, LUM, CCL20, VEGFA and COL6A3, can be potential targets for GBMLGG progression (Supplementary Figure S5).

## 3.11 Prognostic value of TSPAN4 expression in GBM

To further investigate the prognostic value of TSPAN4 expression in GBM, the univariate analysis and multivariate analysis of patients' characteristics demonstrated age, TSPAN4 expression, WHO grade, IDH status and histological types were independent risk factors of GBM (Figures 11A, B; Supplementary Tables S6, S7). To investigate the application of TSPAN4 expression in GBM progression, a nomogram in 1- to 6-year survival probability was constructed and gender, age, TSPAN4 expression and IDH status were included as prognostic factors in the nomogram (Figures 11C, D). The calibration curve suggested that the nomogram was credible in predicting possibility (Supplementary Figures S6A, B). The risk score analysis of lasso regression also illustrated that TSPAN4 expression in GBM was positively associated with risk scores and both were negatively related to survival time (Figures 11E, F; Supplementary Figure S6C). Time-dependent ROC of TAPAN4 expression in GBM and GBMLGG demonstrated that TSPAN4 can be a credible prognostic biomarker in GBM and GBMLGG (Supplementary Figure S7).

## 3.12 Single-cell sequencing of TSPAN4 expression in GBM

Using the single Cell Portal, single-cell sequencing of TSPAN4 expression in GBM was also investigated, suggesting

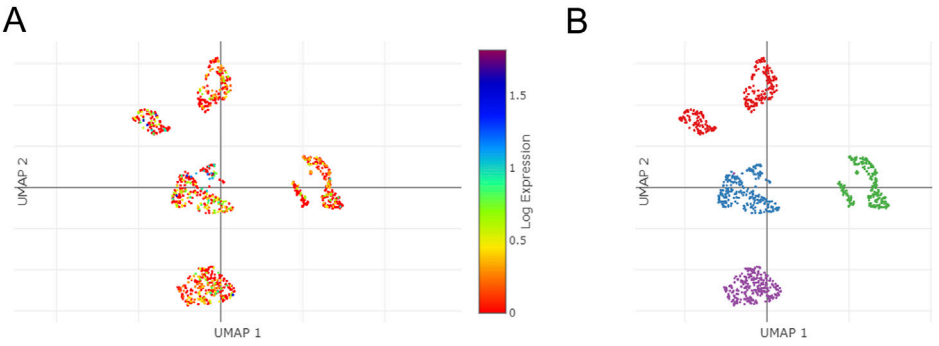



FIGURE 12
The single cells sequencing of TSPAN4 expression in normal and GBM patients. **(A,B)** TSPAN4 expression **(A)** and the overall **(B)** in sequencing cells in GBM patients using Single Cell Portal database.

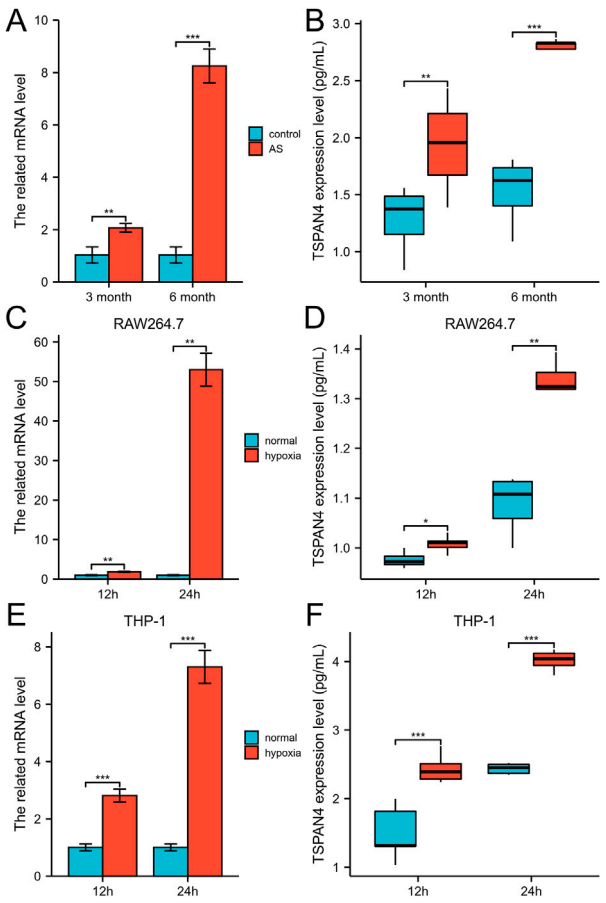



FIGURE 13
The validation of TSPAN4 effects on atherosclerosis and macrophages. **(A,B)** The related mRNA level **(A)** and ELISA result **(B)** of TSPAN4 expression in ApoE$^{-/-}$ mice fed with high fat diet for 3 months and 6 months. **(C,D)** The related mRNA level **(C)** and ELISA result **(D)** of TSPAN4 expression in RAW264.7 macrophages under hypoxia. **(E,F)** The related mRNA level **(E)** and ELISA result **(F)** of TSPAN4 expression in THP-1 macrophages under hypoxia. AS, atherosclerosis. *$p$ < .05; **$p$ < .01; ***$p$ < .001.

TSPAN4 was highly expressed in all GBM cell types compared to control groups (Figure 12).

## 3.13 Validation using mice model and macrophages cell lines

To investigate the TSPAN4 effects on atherosclerosis, the qPCR and ELISA of aorta root in ApoE$^{-/-}$ mice were utilized. The TSPAN4 mRNA and protein levels of the aorta root were highly expressed in atherosclerosis (Figures 13A, B). To explore the TSPAN4 effects on atherosclerosis and GBM progression, RAW264.7 and THP-1 macrophage hypoxia cell lines were utilized. The TSPAN4 mRNA and secreted protein levels were also highly expressed (Figures 13C–F), which may be a potential target for migrasomes and atherosclerosis and GBM progression.

# 4 Discussion

Monocytes and monocyte-derived cells, including macrophages, differ in different phenotypic states from their surrounding microenvironment, which drives T cells activation and function (Padgett et al., 2020). NETs towards atherosclerosis and thrombosis promote breast cancer metastasis to the liver (Moschonas and Tselepis, 2019; Yang et al., 2020). In addition, chemotherapy to treat cancer also induces atherosclerosis plaque formation and progression (Mukai et al., 2018). Therefore, the crosstalk between atherosclerosis and cancer progression needs to be investigated. In this study, TSPAN4 expression in atherosclerosis and pan-cancer was explored and the diagnostic and prognostic value of the TSPAN4 expression as well as the immune analysis were investigated.

TSPANs, such as TSPAN4 and TSPAN7, as well as migrasomes play a critical role in vascular homeostasis (Zhang et al., 2020). Migrasome was demonstrated as an organelle to release cytoplasmic contents during cell migration and cholesterol and TSPAN4 are essential for migrasome formation (Ma et al., 2015; Yu, 2021). Migrating cells, such as monocytes and macrophages, could expel dysfunctional mitochondria in migrasomes to maintain mitochondrial homeostasis and control mitochondrial quality (Jiao et al., 2021; Mehra and Pernas, 2021; Sung et al., 2021). As a novel exosome-like organelle during cell migrations, lateral transfer of mRNA, protein, and even damaged mitochondria can modify the recipient cells (da Rocha-Azevedo and Schmid, 2015; Zhu et al., 2021; Baumann, 2021). The infiltration abundance of migrating cells, such as monocytes and macrophages, may explain the expression levels of TSPAN4 and its impact on prognosis are largely different among tumor types, which was consisted with the previous report by Huang et al., 2022. Huang et al. (2022) reported samples of different tumor types exhibited variable results in immune subtype analysis. Six immune subtypes were correlated tumors, and the expression of TSPAN4 was different among them (Huang et al., 2022).

Migrasomes take center stage in disease development (Tavano and Heisenberg, 2019) and TSPAN4 are essential for migrasome formation. TSPAN4 expression was highly expressed in atherosclerosis and can also be prognostic biomarkers in patients with ACC, GBM, LGG, LUSC, MESO, PCPG, PRAD, THYM, UVM, THCA, GBMLGG or LUADLUSC, while TSPAN4 expression can be diagnostic biomarkers in patients with LUSC, UCS, THYM, KIRC, PAAD, OV, DLBC, CESC and UCEC (AUC > .9) as well as LUAD, LIHC, PRAD, KICH, ESCA, BLCA, GBM, KIRP, BRCA, READ and COADREAD (AUC > .8). Deng et al. (2021) reported that TSPAN4 expression increased,

enhancing proliferation and invasion in gastric cancer tissues (Qi et al., 2018). High TSPAN4 and ELAVL2 expression levels were independent risk factors for poor chemotherapy response in ESCC patients (Zhao et al., 2019). A TSPAN4-CD151 fusion gene was highly expressed in a pediatric infratentorial anaplastic ependymoma (Olsen et al., 2016). In addition, TSPAN4 was one of the top hub genes in patients with fulminant type 1 diabetes (Ye et al., 2020). In this study, TSPAN4 expression was correlated to the abundance of macrophages and DC cells in GBM and GBMLGG and we still need to investigate the functions of TSPAN4 expression on patients with atherosclerosis and cancer progression.

This study provided a critical role of TSPAN4 aberrant expression in the progression of atherosclerosis and pan-cancer. The quality of collected data relies on the source, which could show an impact on the conclusion. Secondly, the result and conclusions are not experimentally validated in the laboratory or clinic. Further studies are still needed to validate *in vivo* and *in vitro*.

## 4 Conclusion

TSPAN4 is differentially expressed in atherosclerosis and also in pan-cancer, which was associated with the progression and immune cell infiltration of the tumor, especially in GBM and GBMLGG. Therefore, TSPAN4 may serve as a potential prognostic and diagnostic biomarker and the crosstalk between atherosclerosis and tumor progression.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The data can be downloaded from GEO datasets and TCGA database.

## Ethics statement

The animal study was reviewed and approved by Protocols were approved by Nankai University, which conform to the Guide for the Care and Use of Laboratory Animals.

## Author contributions

Conceptualization, YZ; Funding acquisition, WG and TL; Investigation, YZ, WG, BQ, and YW; Methodology, YZ and YL; Visualization, YZ; Writing—original draft, YZ; Writing—review and editing, YZ and TL.

## Funding

This work was sponsored by Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-035A), Tianjin biomedical industry chain innovation Project (21ZXSYSY00030), Tianjin Health Research Project (TJWJ2022XK026), Tianjin Health Research Project (TJWJ2022MS020), Tianjin "Project + Team" Key Training Special Project (XC202040), Tianjin "131" Innovative Talent Team Project (201939), Key Project of Tianjin Natural Science Foundation (21JCZDJC00240), the Tianjin Municipal Health and Health Committee Science and Technology Project (ZD20001), Tianjin Health Committee traditional Chinese medicine and integrated traditional Chinese and Western medicine project (2021139), Tianjin Science and Technology Project (21JCYBJC01250), Tianjin Science and Technology Project (21JCYBJC01590), Hebei Provincial Health Commitee Project (20220676), and Hebei Provincial Health and Family Planning Committee Key Medical Science Research Project (20171087).

## Acknowledgments

We would like to express my gratitude to all those who helped us during the writing of this manuscript. Thanks to all the peer reviewers for their opinions and suggestions.

## Conflict of interest

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

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.1030057/full#supplementary-material

# References

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. The gene Ontology consortium. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556

Balzan, S., and Lubrano, V. (2018). LOX-1 receptor: A potential link in atherosclerosis and cancer. *Life Sci.* 198, 79–86. doi:10.1016/j.lfs.2018.02.024

Baumann, K. (2021). Damaged mitochondria are discarded via migrasomes. *Nat. Rev. Mol. Cell Biol.* 22 (7), 442. doi:10.1038/s41580-021-00388-0

Chang, J., Liu, X., and Sun, Y. (2017). Mortality due to acutemyocardial infarction in China from 1987 to 2014: Secular trends and ageperiod-cohort effects. *Int. J. Cardiol.* 227, 229–238. doi:10.1016/j.ijcard.2016.11.130

da Rocha-Azevedo, B., and Schmid, S. L. (2015). Migrasomes: A new organelle of migrating cells. *Cell Res.* 25 (1), 1–2. doi:10.1038/cr.2014.146

Deng, Y., Cai, S., Shen, J., and Peng, H. (2021). Tetraspanins: Novel molecular regulators of gastric cancer. *Front. Oncol.* 11, 702510. doi:10.3389/fonc.2021.702510

Dungal, N., and Benediktsson, T. (1958). Gastric cancer and atherosclerosis. *Lancet* 1 (7027), 931–932. doi:10.1016/s0140-6736(58)91685-4

Dutta, P., Courties, G., Wei, Y., Leuschner, F., Gorbatov, R., Robbins, C. S., et al. (2012). Myocardial infarction accelerates atherosclerosis. *Nature* 487, 325–329. doi:10.1038/nature11260

Hägg, D. A., Jernås, M., Wiklund, O., Thelle, D. S., Fagerberg, B., Eriksson, P., et al. (2008). Expression profiling of macrophages from subjects with atherosclerosis to identify novel susceptibility genes. *Int. J. Mol. Med.* 21 (6), 697–704. doi:10.3892/ijmm.21.6.697

Huang, R., Sun, H., Lin, R., Zhang, J., Yin, H., Xian, S., et al. (2022). The role of tetraspanins pan-cancer. *iScience* 25 (8), 104777. doi:10.1016/j.isci.2022.104777

Huang, Y., Zucker, B., Zhang, S., Elias, S., Zhu, Y., Chen, H., et al. (2019). Migrasome formation is mediated by assembly of micron-scale tetraspanin macrodomains. *Nat. Cell Biol.* 21 (8), 991–1002. doi:10.1038/s41556-019-0367-5

Jiang, D., Jiang, Z., Lu, D., Wang, X., Liang, H., Zhang, J., et al. (2019). Migrasomes provide regional cues for organ morphogenesis during zebrafish gastrulation. *Nat. Cell Biol.* 21 (8), 966–977. doi:10.1038/s41556-019-0358-6

Jiao, H., Jiang, D., Hu, X., Du, W., Ji, L., Yang, Y., et al. (2021). Mitocytosis, a migrasome-mediated mitochondrial quality-control process. *Cell* 184 (11), 2896–2910.e13. e13. doi:10.1016/j.cell.2021.04.027

Kanehisa, M., and Goto, S. (2000). Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

Koelwyn, G. J., Newman, A. A. C., Afonso, M. S., van Solingen, C., Corr, E. M., Brown, E. J., et al. (2020). Myocardial infarction accelerates breast cancer via innate immune reprogramming. *Nat. Med.* 26 (9), 1452–1458. doi:10.1038/s41591-020-0964-7

Liu, Y., Li, S., Rong, W., Zeng, C., Zhu, X., Chen, Q., et al. (2020). Podocyte-released migrasomes in urine serve as an indicator for early podocyte injury. *Kidney Dis. (Basel)* 6 (6), 422–433. doi:10.1159/000511504

Ma, L., Li, Y., Peng, J., Wu, D., Zhao, X., Cui, Y., et al. (2015). Discovery of the migrasome, an organelle mediating release of cytoplasmic contents during cell migration. *Cell Res.* 25 (1), 24–38. doi:10.1038/cr.2014.135

Ma, X., Verweij, E. W. E., Siderius, M., Leurs, R., and Vischer, H. F. (2021). Identification of TSPAN4 as novel histamine H(4) receptor interactor. *Biomolecules* 11 (8), 1127. doi:10.3390/biom11081127

Mehra, C., and Pernas, L. (2021). Move it to lose it: Mitocytosis expels damaged mitochondria. *Dev. Cell* 56 (14), 2014–2015. doi:10.1016/j.devcel.2021.07.001

Mensah, S. A., Nersesyan, A. A., and Ebong, E. E. (2021). Endothelial glycocalyx-mediated intercellular interactions: Mechanisms and implications for atherosclerosis and cancer metastasis. *Cardiovasc Eng. Technol.* 12 (1), 72–90. doi:10.1007/s13239-020-00487-7

Moschonas, I. C., and Tselepis, A. D. (2019). The pathway of neutrophil extracellular traps towards atherosclerosis and thrombosis. *Atherosclerosis* 288, 9–16. doi:10.1016/j.atherosclerosis.2019.06.919

Mukai, M., Komori, K., and Oka, T. (2018). Mechanism and management of cancer chemotherapy-induced atherosclerosis. *J. Atheroscler. Thromb.* 25 (10), 994–1002. doi:10.5551/jat.RV17027

Naghavi, M., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abera, S. F., et al. (2017). Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the global burden of disease study 2016. *Lancet* 390, 1151–1210. doi:10.1016/S0140-6736(17)32152-9

Olsen, T. K., Panagopoulos, I., Gorunova, L., Micci, F., Andersen, K., Kilen Andersen, H., et al. (2016). Novel fusion genes and chimeric transcripts in ependymal tumors. *Genes Chromosom. Cancer* 55 (12), 944–953. doi:10.1002/gcc.22392

Padgett, L. E., Araujo, D. J., Hedrick, C. C., and Olingy, C. E. (2020). Functional crosstalk between T cells and monocytes in cancer and atherosclerosis. *J. Leukoc. Biol.* 108 (1), 297–308. doi:10.1002/JLB.1MIR0420-076R

Qi, W., Sun, L., Liu, N., Zhao, S., Lv, J., and Qiu, W. (2018). Tetraspanin family identified as the central genes detected in gastric cancer using bioinformatics analysis. *Mol. Med. Rep.* 18 (4), 3599–3610. doi:10.3892/mmr.2018.9360

Restall, I. J., Cseh, O., Richards, L. M., Pugh, T. J., Luchman, H. A., and Weiss, S. (2020). Brain tumor stem cell dependence on glutaminase reveals a metabolic vulnerability through the amino acid deprivation response pathway. *Cancer Res.* 80 (24), 5478–5490. doi:10.1158/0008-5472.CAN-19-3923

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Singh, A., Gupta, A., DeFilippis, E. M., Qamar, A., Biery, D. W., Almarzooq, Z., et al. (2020). Cardiovascular mortality after type 1 and type 2 myocardial infarction in young adults. *J. Am. Coll. Cardiol.* 75 (9), 1003–1013. doi:10.1016/j.jacc.2019.12.052

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Sung, B. H., Parent, C. A., and Weaver, A. M. (2021). Extracellular vesicles: Critical players during cell migration. *Dev. Cell* 56 (13), 1861–1874. doi:10.1016/j.devcel.2021.03.020

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47 (D1), D607–D613. doi:10.1093/nar/gky1131

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). Gepia: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45 (W1), W98–W102. doi:10.1093/nar/gkx247

Tapia-Vieyra, J. V., Delgado-Coello, B., and Mas-Oliva, J. (2017). Atherosclerosis and cancer; A resemblance with far-reaching implications. *Arch. Med. Res.* 48 (1), 12–26. doi:10.1016/j.arcmed.2017.03.005

Tavano, S., and Heisenberg, C. P. (2019). Migrasomes take center stage. *Nat. Cell Biol.* 21 (8), 918–920. doi:10.1038/s41556-019-0369-3

Yang, L., Liu, Q., Zhang, X., Liu, X., Zhou, B., Chen, J., et al. (2020). DNA of neutrophil extracellular traps promotes cancer metastasis via CCDC25. *Nature* 583 (7814), 133–138. doi:10.1038/s41586-020-2394-6

Ye, X., Zeng, T., Kong, W., and Chen, L. L. (2020). Integrative analyses of genes associated with fulminant type 1 diabetes. *J. Immunol. Res.* 2020, 1025857. doi:10.1155/2020/1025857

Yu, G. C., Wang, L. G., Han, Y. Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics a J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yu, L. (2021). Migrasomes: The knowns, the known unknowns and the unknown unknowns: A personal perspective. *Sci. China Life Sci.* 64 (1), 162–166. doi:10.1007/s11427-020-1827-8

Zhang, Y., Wang, J., Ding, Y., Zhang, J., Xu, Y., Xu, J., et al. (2020). Migrasome and tetraspanins in vascular homeostasis: Concept, present, and future. *Front. Cell Dev. Biol.* 8, 438. doi:10.3389/fcell.2020.00438

Zhao, W. S., Yan, W. P., Chen, D. B., Dai, L., Yang, Y. B., Kang, X. Z., et al. (2019). Genome-scale CRISPR activation screening identifies a role of ELAVL2-CDKN1A axis in paclitaxel resistance in esophageal squamous cell carcinoma. *Am. J. Cancer Res.* 9 (6), 1183–1200.

Zhu, M., Zou, Q., Huang, R., Li, Y., Xing, X., Fang, J., et al. (2021). Lateral transfer of mRNA and protein by migrasomes modifies the recipient cells. *Cell Res.* 31 (2), 237–240. doi:10.1038/s41422-020-00415-3

## Glossary

**ACC** Adrenocortical carcinoma

**AS** Atherosclerosis

**AUC** Area under curve

**BLCA** Bladder Urothelial Carcinoma

**BRCA** Breast invasive carcinoma

**CAD** Cardiovascular diseases

**CESC** Cervical squamous cell carcinoma and endocervical adenocarcinoma

**CHOL** Cholangiocarcinoma

**CI** Confident interval

**COAD** Colon adenocarcinoma

**COADREAD** Colon adenocarcinoma and Rectum adenocarcinoma esophageal carcinoma

**DEGs** Differentially expressed genes

**DLBC** Lymphoid Neoplasm Diffuse Large B-cell Lymphoma

**DSS** Disease specific survival

**EPCAM** Epithelial cell adhesion molecule

**ESCA** Esophageal carcinoma

**GBM** Glioblastoma multiforme

**GBMLGG** Glioblastoma multiforme and Brain Lower Grade Glioma

**GO** Gene Ontology

**GSEA** Gene Set Enrichment Analysis

**GTEx** Genotype-Tissue Expression

**HR** Hazard ratio

**HNSC** Head and Neck squamous cell carcinoma

**H4R** histamine H4 receptor

**KEGG** Kyoto Encyclopedia of Genes and Genomes

**KICH** Kidney Chromophobe

**KIRC** Kidney renal clear cell carcinoma

**KIRP** Kidney renal papillary cell carcinoma

**LAML** Acute Myeloid Leukemia

**LGG** Brain Lower Grade Glioma

**LIHC** Liver hepatocellular carcinoma

**LUAD** Lung adenocarcinoma

**LUADLUSC** Lung adenocarcinoma and Lung squamous cell carcinoma

**LUSC** Lung squamous cell carcinoma

**MI** Myocardial infarction

**MLH1** MutL homologous gene 1

**MMR** Mismatch repair

**MSH2** MutS homologous gene 2

**MSH6** MutS homologous gene 6

**NETs** Neutrophil extracellular traps

**OS** Overall survival

**OV** Ovarian serous cystadenocarcinoma

**PAAD** Pancreatic adenocarcinoma

**PFI** Progress free interval

**PMS2** Increased separation after meiosis

**PPI** Protein-protein interaction

**PRAD** Prostate adenocarcinoma

**READ** Rectum adenocarcinoma

**ROC** Receiver operation characteristic

**SKCM** Skin Cutaneous Melanoma

**THCA** Thyroid carcinoma

**THYM** Thymoma

**TSPAN4** Tetraspanin 4

**TSPAN7** Tetraspanin 7

**UCEC** Uterine Corpus Endometrial Carcinoma

**UCS** Uveal Melanoma

frontiers | Frontiers in Molecular Biosciences

TYPE Original Research
PUBLISHED 07 February 2023
DOI 10.3389/fmolb.2023.1082915



OPEN ACCESS

EDITED BY
Yunpeng Hua,
The First Affiliated Hospital of Sun Yat-sen University, China

REVIEWED BY
Maria La Mantia,
University of Palermo, Italy
Pathmanathan Rajadurai,
Subang Jaya Medical Centre, Malaysia

*CORRESPONDENCE
Manuel R. Teixeira,
manuel.teixeira@ipoporto.min-saude.pt
Saudade André,
sandre@ipolisboa.min-saude.pt
José Carlos Machado,
josem@i3s.up.pt

SPECIALTY SECTION
This article was submitted to Molecular Diagnostics and Therapeutics, a section of the journal Frontiers in Molecular Biosciences

RECEIVED 28 October 2022
ACCEPTED 05 January 2023
PUBLISHED 07 February 2023

CITATION
Peixoto A, Cirnes L, Carvalho AL, Andrade MJ, Brito MJ, Borralho P, Borralho PM, Carneiro AS, Castro L, Correia L, Dionísio MR, Faria C, Figueiredo P, Gomes A, Paixão J, Pinheiro M, Prazeres H, Ribeiro J, Salgueiro N, Schmitt FC, Silva F, Silvestre AR, Sousa AC, Almeida-Tavares J, Teixeira MR, André S and Machado JC (2023), Evaluation of *PIK3CA* mutations in advanced ER+/HER2-breast cancer in Portugal – U-PIK Project.
*Front. Mol. Biosci.* 10:1082915.
doi: 10.3389/fmolb.2023.1082915



# Evaluation of *PIK3CA* mutations in advanced ER+/HER2-breast cancer in Portugal – U-PIK Project

Ana Peixoto[1], Luís Cirnes[2], Ana Luísa Carvalho[3], Maria João Andrade[4], Maria José Brito[5], Paula Borralho[6,7], Pedro M. Borralho[8], Ana Sofia Carneiro[9], Lisandra Castro[10], Lurdes Correia[9,11], Maria Rita Dionísio[8], Carlos Faria[4], Paulo Figueiredo[12], Ana Gomes[4], Joana Paixão[9], Manuela Pinheiro[1], Hugo Prazeres[12], Joana Ribeiro[5], Natália Salgueiro[10], Fernando C. Schmitt[2,13], Fátima Silva[4,14,15], Ana Rita Silvestre[6], Ana Carla Sousa[16], Joana Almeida-Tavares[9], Manuel R. Teixeira[1,17]*, Saudade André[3]* and José Carlos Machado[2,13]*

[1]Serviço de Genética Laboratorial, Instituto Português de Oncologia do Porto Francisco Gentil (IPO Porto), Porto, Portugal, [2]IPATIMUP - Instituto de Patologia e Imunologia da Universidade do Porto, Porto, Portugal, [3]Serviço de Anatomia Patológica, Instituto Português de Oncologia de Lisboa Francisco Gentil (IPOLFG), Lisboa, Portugal, [4]Centro Hospitalar e Universitário de Coimbra, Coimbra, Portugal, [5]Unidade de Mama, Centro Clínico Champalimaud, Fundação Champalimaud, Lisboa, Portugal, [6]Serviço de Anatomia Patológica, Hospital CUF Descobertas, Lisboa, Portugal, [7]Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal, [8]Novartis Farma - Produtos Farmacêuticos, S.A., Porto Salvo, Portugal, [9]Serviço de Anatomia Patológica, Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte, Lisboa, Portugal, [10]Departamento de Genética Molecular, SYNLAB Genética Médica, S.A., Porto, Portugal, [11]Instituto de Anatomia Patológica, Lisboa, Portugal, [12]Serviço de Anatomia Patológica, IPO Coimbra, Coimbra, Portugal, [13]Faculdade de Medicina da Universidade do Porto, Porto, Portugal, [14]Escola Superior de Tecnologia da Saúde de Coimbra, Coimbra, Portugal, [15]Associação Portuguesa de Técnicas de Anatomia Patológica, Porto, Portugal, [16]GenoMed – Diagnósticos de Medicina Molecular, S.A., Lisboa, Portugal, [17]Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Porto, Portugal

**Background:** Around 40% of ER+/HER2-breast carcinomas (BC) present mutations in the *PIK3CA* gene. Assessment of *PIK3CA* mutational status is required to identify patients eligible for treatment with PI3Kα inhibitors, with alpelisib currently the only approved tyrosine kinase inhibitor in this setting. U-PIK project aimed to conduct a ring trial to validate and implement the *PIK3CA* mutation testing in several Portuguese centers, decentralizing it and optimizing its quality at national level.

**Methods:** Eight Tester centers selected two samples of patients with advanced ER+/HER2- BC and generated eight replicates of each (n = 16). *PIK3CA* mutational status was assessed in two rounds. Six centers used the cobas® *PIK3CA* mutation test, and two used PCR and Sanger sequencing. In parallel, two reference centers (IPATIMUP and the Portuguese Institute of Oncology [IPO]-Porto) performed *PIK3CA* mutation testing by NGS in the two rounds. The quality of molecular reports describing the results was also assessed. Testing results and molecular reports were received and analyzed by U-PIK coordinators: IPATIMUP, IPO-Porto, and IPO-Lisboa.

**Results:** Overall, five centers achieved a concordance rate with NGS results (allele frequency [AF] ≥5%) of 100%, one of 94%, one of 93%, and one of 87.5%, considering the overall performance in the two testing rounds. NGS reassessment of discrepancies in the results of the methods used by the Tester centers and the reference centers identified one probable false positive and two mutations with low AF (1–3%, at the analytical sensitivity threshold), interpreted as subclonal variants with heterogeneous representation in the tissue sections processed by the respective

centers. The analysis of molecular reports revealed the need to implement the use of appropriate sequence variant nomenclature with the identification of reference sequences (HGVS-nomenclature) and to state the tumor cell content in each sample.

**Conclusion:** The concordance rates between the method used by each tester center and NGS validate the use of the *PIK3CA* mutational status test performed at these centers in clinical practice in patients with advanced ER+/HER2- BC.

KEYWORDS

**advanced breast cancer, ER+/HER2–, *PIK3CA* mutations, ring trial, validation, molecular pathology**

## Introduction

Female breast carcinoma surpassed lung carcinoma as the most commonly diagnosed cancer in 2020, with an estimated 2.3 million new cases (11.7%), and it was the fifth leading cause of cancer death (6.9%), according to GLOBOCAN (Sung et al., 2021). Over 70% of breast carcinomas are hormone receptor-positive (HR+) and human epidermal growth factor receptor 2-negative (HER2-), collectively designated as HR+/HER2-breast carcinomas (Setiawan et al., 2009; Howlader et al., 2014). Approximately 40% of estrogen receptor-positive (ER+)/HER2-breast carcinomas have mutations in the *PIK3CA* gene, leading to hyperactivation of the alpha isoform (p110α) of phosphatidylinositol-3-kinase (PI3K) (Cancer Genome Atlas Network, 2012; Goncalves et al., 2018; Mollon et al., 2018). The PI3K-protein kinase B (AKT)-mammalian target of rapamycin (mTOR) cascade is one of the major downstream signaling pathways in human cells (Fruman et al., 2017), and its deregulation has been implicated in breast cancer development and progression (Lee et al., 2015).

*PIK3CA* activating tumor mutations may occur in several domains of the p110α catalytic subunit, but mostly (≈80%) arise in four hotspots of the helical and kinase domains: E542K, E545K, H1047R, and H1047L (Zhao and Vogt, 2008; Kalinsky et al., 2009; Dogruluk et al., 2015). According to the Cancer Genome Atlas Network, *PIK3CA* mutations are more frequent in luminal A (45%) compared to luminal B (29%) subtypes, also occurring in basal-like (9%) and HER2-enriched (39%) subtypes (Cancer Genome Atlas Network, 2012). A recent study including a large number of metastatic breast carcinomas (n = 3871) confirmed these data, reporting the presence of *PIK3CA* mutations in 39% of HR+/HER2- and 37% of HER2-amplified tumors, besides 21% of triple-negative carcinomas (Albanell et al., 2019). The presence of *PIK3CA* mutations represents an independent adverse prognostic factor in breast carcinoma (Sobhani et al., 2018) and has been associated with more aggressive disease and poor outcomes in metastatic disease (Fitzgerald et al., 2019).

Despite the high incidence of *PIK3CA* mutations in breast carcinoma and their putative prognostic role, the therapeutic targeting of the *PIK3CA* gene has fallen short of expectations, with results from clinical trials with pan-PI3K inhibitors in solid tumors largely disappointing, namely due to unfavorable toxicity *versus* clinical benefit ratio (Hanker et al., 2019). The α-selective PI3K inhibitor alpelisib was the first PI3K inhibitor to demonstrate a progression-free survival (PFS) benefit in HR+/HER2-metastatic breast carcinoma with activating *PIK3CA* mutations. In the SOLAR-1 phase III clinical trial, the addition of alpelisib to fulvestrant in patients with disease recurrence/progression on or after prior aromatase inhibitor therapy resulted in an almost double median PFS compared to fulvestrant alone in the *PIK3CA*-mutated cohort (11.0 vs. 5.7 months, HR 0.65; 95% CI 0.50–0.85; $p < 0.001$), with an acceptable safety profile (André et al., 2019). Results from the SOLAR-1 trial led to the approval by the U.S. Food and Drug Administration (FDA) in May 2019 of alpelisib in combination with fulvestrant for postmenopausal women and men with HR+/HER2-, *PIK3CA*-mutated advanced or metastatic breast carcinoma, as detected by an FDA-approved test, following progression on or after an endocrine-based regimen (US Food & Drug Administration (FDA), 2019), and subsequently also to the approval by the European Medicines Agency (EMA) in July 2020 of alpelisib in combination with fulvestrant for postmenopausal women and men with HR+/HER2-, locally advanced or metastatic breast cancer with a *PIK3CA* mutation, after disease progression following endocrine therapy as monotherapy (European Medicines Agency, 2021). The application of Therascreen® companion diagnostic test allowed the detection of 11 *PIK3CA* mutations in patients included in the SOLAR-1 trial: C420R, E542K, E545A, E545D [1635G>T only], E545G, E545K, Q546E, Q546R, H1047L, H1047R, and H1047Y (QIAGEN; André et al., 2019).

In Europe, companion diagnostic tests are not required for market authorization of targeted Oncology drugs, and the prescription of approved targeted therapeutic options relies on evidence of the required mutational status assessed through validated methods. The use of Therascreen® is not widespread in Portugal, but multiple technologies and platforms are available for molecular testing, which can potentially be used to detect *PIK3CA* mutational status in HR+/HER2-breast carcinoma samples. These require local validation and implementation for clinical diagnostic purposes, similar to what has been previously done for several other actionable mutations in different tumor types (e.g., *EGFR* in lung cancer, *BRAF* in melanoma and lung cancer, etc.). Therefore, to enable local assessment of *PIK3CA* mutational status, hospitals and diagnostic centers should select, implement, and validate diagnostic methodologies that allow the molecular detection of all currently but also potentially actionable *PIK3CA* mutations.

Due to the absence of local validation and implementation, *PIK3CA* mutation testing was not routinely performed in HR+/HER2-breast carcinoma and remained largely unavailable in Portugal. This represents an unmet need since the assessment of *PIK3CA* mutational status is crucial for identifying patients eligible for treatment with PI3Kα inhibitors, particularly alpelisib, which is currently FDA- and EMA-approved in this clinical setting. The U-PIK project aimed to conduct a ring trial to validate and implement the *PIK3CA* mutation testing in multiple Portuguese

**TABLE 1 Characteristics of the ER+/HER2-breast carcinoma samples included in the first *PIK3CA* testing round.**

| Sample | A | C | E | G | I | K | M | O |
|---|---|---|---|---|---|---|---|---|
| Patient age (years) | 73 | 63 | 54 | 63 | 71 | 77 | 36 | 46 |
| Tumor content | 75% | 80% | >20% | ≈50% | 30% | 35% | 40% | >20% |
| Necrosis | <1% | 0% | >20% | ≈5% | <5% | <1% | 0% | 0% |
| Histological type | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) |
| Histological grade | G2 | G2 | | | | G1 | G3 | G2 |
| ER (%) | 90–100 | 100 | 80 | 100 | 91–100 | 50 | 90 | 100 |
| PR (%) | 80–90 | 20 | 0 | 40 | 0 | 20 | 90 | 5 |
| HER2 status | 0 | 0 | 0 | 2+ DISH- | 0 | 1+ | 0 | 2+ DISH- |

All samples correspond to surgical specimens (4 primary carcinomas, 1 subcutaneous recurrence [C], and 3 metastases: 1 cervical lymph node [E], 1 hepatic [K], and 1 pleural [M]).
ER, estrogen receptor; PR, progesterone receptor HER2, human epidermal growth factor receptor 2; IBC (NST), invasive breast carcinoma of no special type; G, grade; DISH, dual *in situ* hybridization.

centers, decentralizing the test and optimizing its quality at national level.

## Materials and methods

U-PIK was a research collaboration between 11 partners – 10 centers (seven hospitals and three private laboratories) from the North, Central, and Lisbon regions of Portugal, and Novartis – that consisted of a multicenter, inter-laboratory ring trial conducted between December 2020 and December 2021.

The study was conducted according to the guidelines of the Declaration of Helsinki. Ethical approval for the study was obtained from multiple boards: Ethics Committee of IPOLFG and Research Council, Unidade de Investigação Clínica, Instituto Português de Oncologia de Lisboa Francisco Gentil, EPE; Ethics Committee of Centro Hospitalar e Universitário de Coimbra; Ethics Committee of Centro Hospitalar e Universitário de Lisboa Norte and of Centro Académico de Medicina de Lisboa; Ethics Committee of Hospital CUF Descobertas; and Ethics Council and Ethics Council for Health of Fundação Champalimaud. Written informed consent for study participation or research purposes in the scope of the study was provided by participants or participants' legal guardians/next of kin. In addition, ethical approval from other institutions was waived, as the study is in accordance with Article 19 ("DNA Banks and Other Biological Products") of Portuguese Law No. 12/2005 of 26 January ("Personal genetic information and health information"), which states that in the case of using retrospective samples of human origin or in special situations where the consent of subjects involved cannot be obtained due to the amount of data or subjects, their age, or another similar reason, the material and data can be processed but only for purposes of scientific research or epidemiological and statistical data collection (Law No. 12/2005; Kalokairinou et al., 2018).

Eight centers selected two samples of patients with advanced ER+/HER2-breast carcinoma from routine clinical practice and generated eight replicates of each, in a total of 16 samples. All samples were duly anonymized, and each sample was processed through an initial section stained with hematoxylin and eosin (H&E) followed by 24 sequential sections of 6 µm and a final section stained with H&E with the tumor area demarcated.

All samples were centralized at Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), which generated replicates of each, originating two sample sets with identical sample content comprising the overall group of Test samples. These two sample sets were sent to eight Tester centers, which assessed the *PIK3CA* mutational status in two rounds, 8 weeks apart, using their in-house methodology. Tables 1, 2 depict the characteristics of the ER+/HER2-breast carcinoma samples included in the analysis.

Tester centers were given 15 running days from the date of sample set reception to test *PIK3CA* mutational status and report back on results. As a study requirement, the in-house methodology of each Tester center had to allow the detection of mutations in at least exons 7, 9, and 20 of *PIK3CA* (corresponding to exons 8, 10, and 21 using the MANE Select transcript NM_006218.4), as per SOLAR-1 clinical trial criteria (ClinicalTrials.gov). Six Tester centers used the cobas® *PIK3CA* Mutation Test, and two used PCR and Sanger sequencing (see Supplementary Material). In parallel, two reference centers (IPATIMUP and the Portuguese Institute of Oncology [IPO]-Porto) performed *PIK3CA* mutation testing by NGS in the two rounds (see Supplementary Material).

The concordance rate between the results of *PIK3CA* mutational status obtained by each Tester center and NGS was reported for both testing rounds. The concordance rate was also reported globally, considering all samples tested in both rounds at each Tester center. In addition, the quality of molecular reports describing *PIK3CA* mutation testing results was also assessed. The testing results from Tester sites and the molecular reports were received and analyzed by U-PIK coordinators: IPATIMUP, IPO-Porto, and IPO-Lisboa.

The original data presented in the study will be publicly available. The NGS data was submitted to the European Nucleotide Archive (ENA) at EMBL-EBI, under the project accession number: PRJEB58369 (https://www.ebi.ac.uk/ena).

## Results

### Results of the *PIK3CA* testing rounds

Of the eight participating Tester centers, six assessed the *PIK3CA* mutational status by cobas® *PIK3CA* Mutation Test and two by PCR and Sanger sequencing. In the first testing round, all Testers obtained a

**TABLE 2 Characteristics of ER+/HER2-breast carcinoma samples included in the second *PIK3CA* testing round.**

| Sample | B | D | F | H | J | L | N | P |
|---|---|---|---|---|---|---|---|---|
| Patient age (years) | 42 | 63 | 30 | 59 | 89 | 63 | 55 | 48 |
| Tumor content (%) | 60 | 90 | >20 | ≈40 | 40 | 70 | 60 | >20 |
| Necrosis (%) | <1 | 1 | >10 | 0 | <5 | <1 | 0 | 0 |
| Histological type | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) | IBC (NST) | ILC | IBC (NST) |
| Histological grade | G2 | | G3 | | | G2 | G2 | G2 |
| ER (%) | 60–70 | 100 | 15 | 95 | 91–100 | 95 | 95 | 100 |
| PR (%) | 90–100 | 100 | 0 | 95 | | 30 | 0 | 90 |
| HER2 status | 0 | 0 | 0 | 1+ | 0 | 2+ DISH- | 1+ | 0 |

All samples correspond to surgical specimens (7 primary carcinomas and 1 axillary lymph node recurrence [D]).
ER, estrogen receptor; PR, progesterone receptor HER2, human epidermal growth factor receptor 2; IBC (NST), invasive breast carcinoma of no special type; ILC, invasive lobular carcinoma; G, grade; DISH, dual *in situ* hybridization.

100% concordance between the results of *PIK3CA* mutational status retrieved by their in-house methodology and NGS. Only one sample processed by one Tester center gave an invalid result due to damage of the slide and unavailability of adequate DNA yield, which precluded reanalysis. In the second testing round, four Tester centers achieved 100% concordance, three 88% concordance, and one 63% concordance in the *PIK3CA* mutational status results obtained by their in-house methodology and NGS. Overall, considering the two testing rounds, five centers achieved a concordance rate with NGS results (AF ≥5%) of 100%, one of 94%, one of 93%, and one of 87.5%.

Table 3 shows the discordant results obtained in the two *PIK3CA* testing rounds.

Discordant results were due to discrepancies in the output of three samples (samples F, N, and L), two of which had remaining DNA available and were reanalyzed by NGS (including allele frequencies [AFs] <5%) at IPO-Porto and IPATIMUP reference centers (Tables 4, 5). In one sample (sample F), the H1047X mutation was identified by the Tester center but not by NGS (either in the initial or repeated analysis), being considered a false positive probably resulting from sample contamination. In sample N, the H1047X mutation was detected by three Tester centers but not by NGS analysis using a 5% AF cut-off. However, in NGS reanalysis and integrative genomics viewer (IGV) visualization, the c.3140A>G variant was identified with low AF (1–3%) in all three samples of the three Tester centers, suggesting that this is a low-frequency subclonal variant with heterogeneous representation in the tissue sections processed by the respective centers. Furthermore, in this sample, the c.1258T>C variant (C420R) was detected by one of the Tester centers in addition to the H1047X mutation. This variant was also detected in NGS reanalysis with an AF of 1% and may also correspond to a low-frequency subclonal variant with higher representation in the tissue section processed by that center. The third sample (sample L) for which discrepancies were found could not be reassessed by NGS due to DNA unavailability. In this sample, although the result obtained by the methodology of the Tester center was concordant with NGS for the presence of the H1047X mutation, a second mutation (G1049R) was detected by the cobas® method used by that center (Tables 4, 5), which however was not validated by the Tester center using Sanger sequencing and might represent a known issue of cross-reactivity of the cobas® kit for some variants.

## Quality of molecular reports

The analysis of the molecular reports revealed inconsistencies in the description of sequence variants, highlighting the need to foster the use of the appropriate sequence variant nomenclature (HGVS-nomenclature), the international standard for reporting and exchanging information of DNA, RNA, and protein sequence variants in a consistent and unambiguous way (den Dunnen et al., 2016). It also showed the need to include the identification of reference sequences and state the tumor cell content of each sample in the report.

# Discussion

Despite significant accomplishments in the diagnosis and treatment of advanced breast cancer, it remains largely incurable. In recent years, several clinical studies have sought to identify novel molecular targets and predictive biomarkers that enable tailored management of these patients and improve outcomes. Within this approach, the pharmacologic targeting of *PIK3CA* mutations in ER+/HER2-advanced breast cancer has recently shown significant benefits after the development of endocrine therapy resistance. The U-PIK project aimed to evaluate the analytical performance of *PIK3CA* mutational status testing and contribute to its decentralized implementation in Portuguese centers. Within U-PIK, both testing results and molecular reports were analyzed by U-PIK coordinators IPATIMUP, IPO-Porto, and IPO-Lisboa.

The study found high concordance rates between the results obtained with the methodologies used in each tester center and NGS performed at the two reference centers (AF ≥5%):100% in five centers, and 94%, 93%, and 87.5% in one center each. These results validate the use of the *PIK3CA* mutational status test performed at those centers in the clinical diagnostics of patients with advanced ER+/HER2-breast carcinoma and enable these patients to be selected for an additional treatment option with PI3Kα inhibitors. The selective PI3Kα inhibitor alpelisib brought a renewed interest in *PIK3CA* as a predictive biomarker in HR+/HER2-disease, and the widespread use of the *PIK3CA* mutational status testing in the clinical practice will allow these patients to be

**TABLE 3 Discordant results obtained in the two *PIK3CA* testing rounds.**

| Testing method | *PIK3CA* mutational status | | | | Center 1 | Center 2 | Center 3 | Center 4 | Center 5 | Center 6 | Center 7 | Center 8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | NGS | | | | COBAS | COBAS | COBAS | COBAS | PCR/ Sanger | COBAS | PCR/ Sanger | COBAS |
| Testing sample | HGVS coding region | HGVS protein | VAF IPO-Porto (%) | VAF IPATIMUP (%) | | | | | | | | |
| I | c.3140A>G | p. (His1047Arg) | 29 | 31 | Concordant | Concordant | Concordant | Concordant | Concordant | Invalid | Concordant | Concordant |
| F (P21_225) | Wild type | Wild type | - | - | Concordant | Concordant | H1047X | Concordant | Concordant | Concordant | Concordant | Concordant |
| L | c.3140A>G | p. (His1047Arg) | 33 | 33 | Partially concordant: G1049R | Concordant | Partially concordant: G1049R | Concordant | Concordant | Concordant | Concordant | Concordant |
| N | Wild type | Wild type | - | - | Concordant | Concordant | H1047X | H1047X; C420R | Concordant | H1047X | Concordant | Concordant |
| **Concordance rate (%)** | | | | | **100** | **100** | **87.5** | **94** | **100** | **93** | **100** | **100** |

HGVS, human genome variation society; NGS, next-generation sequencing; VAF, variant allele frequency.

identified and potentially benefit from this targeted agent, now available in their treatment armamentarium.

According to the literature, around 90% of *PIK3CA* mutations cluster in exons 9 and 20, and only ≈5–10% are found in other exons (Campbell et al., 2004; Samuels et al., 2004; Harlé et al., 2013; Arsenic et al., 2015; Martínez-Sáez et al., 2020). Exons 9 and 20 encode the helical and kinase domains of the *PIK3CA* gene, respectively, providing auto-inhibition of the tyrosine kinases, with mutations within these regions triggering the process of constant auto-phosphorylation and resulting in gain of function (Zhao and Vogt, 2008). Exon 4 has also been reported to be often altered in breast cancer, with the p.N345K mutation described with a 5.5% frequency in a recent analysis including 6338 breast cancer patients across 10 publicly available studies (Martínez-Sáez et al., 2020). Although this is a likely pathogenic variant according to the COSMIC and OncoKB databases (Chakravarty et al., 2017; Tate et al., 2019), its sensitivity to PI3K inhibitors has only been demonstrated in preclinical studies so far (Gymnopoulos et al., 2007; Dogruluk et al., 2015). In the present study, from the 16 specimens analyzed, 10 (62.5%) harbored *PIK3CA* mutations, predominantly found in exon 20 (n = 6, 37.5%) and exon 9 (n = 4, 25%), in agreement with most frequently reported *PIK3CA* mutations. Published studies show disparate results regarding the frequency of mutations in exon 9 and 20, with some also reporting that exon 20 is more frequently mutated than exon 9 (Li et al., 2006; Castaneda et al., 2014; Dirican et al., 2014; Loibl et al., 2014; Arsenic et al., 2015; Ahmad et al., 2016), and others showing otherwise (Campbell et al., 2004; Samuels et al., 2004). Nevertheless, the small sample size in this study should always be noted as a caveat and a limitation in the interpretation of the relative *PIK3CA* mutation frequency results in this study.

Four common mutations were identified in this small sample set, at codons 545 and 546 (E545K and Q546K) in the helical domain and at codon 1047 (H1047R and H1047L) in the kinase domain of the *PIK3CA* gene. All these alterations have been shown to respond to alpelisib (Verret et al., 2019; Anderson et al., 2020; Copur, 2020; Martínez-Sáez et al., 2020), although Q546K only in preclinical studies, where it has been associated with increased sensitivity to this agent (Mayer et al., 2017; Martínez-Sáez et al., 2020). In agreement with other reports, the four mutations observed in both exons were of missense type, with a single nucleotide change resulting in a different amino acid at the respective codons: nucleotide change c.1633G>A resulted in amino acid change p.E545K and nucleotide change c.1636C>A resulted in amino acid change p.Q546K at exon 9; nucleotide change c.3140A>G resulted in amino acid change p.H1047R and nucleotide change c.3140A>T resulted in amino acid change p.H1047L at exon 20 (Levine et al., 2005; Nosho et al., 2008; Ross et al., 2015; Ahmad et al., 2016). Both missense mutations E545K and Q546K were equally prevalent in exon 9 (50% each), while the H1047R missense mutation was more prevalent (83.3%) in exon 20 compared to the H1047L mutation (16.7%). These mutations were clustered within the mutational hotspot regions covering nucleotides 1633–1636 (n = 4) and 3140 (n = 6), in agreement with what has been reported by others (Nosho et al., 2008; Wang et al., 2015; Ahmad et al., 2016).

In addition to E545K, Q546K, H1047R, and H1047L, other less common *PIK3CA* mutations have been identified in the samples analyzed in this study. The exon 7 p.C420R mutation was found both in the assessment conducted by one Tester center and by NGS with an AF of 1%. This variant had also been previously reported by Corné et al. and Martínez-Sáez et al. with a frequency of 1.4% and 1.9%, respectively (Martínez-Sáez et al., 2020; Corné et al., 2021), and shown responsiveness to alpelisib in the SOLAR-1 trial (André et al., 2019). On the other hand, the *PIK3CA* exon 20 p.G1049R mutation was detected by the cobas® method used by one Tester center, and had been previously reported in 0.7% of the samples analyzed in the study by Martínez-Sáez and colleagues (Martínez-Sáez et al., 2020). Although rare, this variant is likely pathogenic, with preclinical studies suggesting an increased sensitivity to alpelisib, similar to p.Q546K (Mayer et al., 2017) and p.E542K (Dogruluk et al., 2015) alterations. However, whenever this mutation is found in a tumor also presenting the H1047X mutation, the possibility of a false positive due to a known issue of cross-reactivity of the quantitative PCR technology should be considered.

The detection of *PIK3CA* mutations in liquid biopsy, although not in the scope of the U-PIK study, is also being explored in metastatic breast

**TABLE 4** ***PIK3CA*** **testing results of NGS reanalysis by IPO-Porto reference center.**

| Testing sample | Tester center | *PIK3CA* testing result by in-house methodology (COBAS) | *PIK3CA* testing result by NGS* | H1047X (c.3140A>G/T) | H1047X (c.3139C>T) | C420R (c.1258T>C) |
|---|---|---|---|---|---|---|
| N | Center 4 | H1047X; C420R | wild type | Total count 7023 reads: A 6934; G 78 (**1.1%**); T 7 (≤**0.1%**) | Total count 7149 reads: C 7133; T 1 (≤**0.1%**) | Total count 2397 reads: T 2375; C 18 (**0.8%**) |
| N | Center 3 | H1047X | wild type | Total count 9066 reads: A 8826; G 222 (**2.4%**); T 10 (≤**0.1%**) | Total count 9193 reads: C 9168; T 4 (≤**0.1%**) | |
| N | Center 6 | H1047X | wild type | Total count 8104 reads: A 7996; G 96 (**1.2%**); T 3 (≤**0.1%**) | Total count 8273 reads: C 8225; T 25 (**0.3%**) | |
| F | Center 3 | H1047X | wild type | Total count 7388 reads: A 7381; G 2 (≤**0.1%**); 2 (≤**0.1%**) | Total count 7507 reads: C 7485; T 5 (≤**0.1%**) | |

ID, identification; IGV, integrative genomics viewer; NGS, next-generation sequencing.
*Only VAFs >3% were considered.
Bold values: % read of the total count reads.

**TABLE 5** ***PIK3CA*** **testing results of NGS reanalysis by IPATIMUP reference center.**

| Testing sample | Tester center | *PIK3CA* testing result by in-house methodology (COBAS) | *PIK3CA* testing result by NGS* | H1047X (c.3140A>G/T) | H1047X (c.3139C>T) | C420R (c.1258T>C) |
|---|---|---|---|---|---|---|
| N | Center 4 | H1047X; C420R | wild type | Total count 4545 reads: A 4484; G 60 (**1.3%**); T 0 (0%) | Total count 4546 reads: C 4538; T 8 (≤0.1%) | Total count 3892 reads: T 3865; C 27 (1%) |
| N | Center 3 | H1047X | wild type | Total count 2764 reads: A 2679; G 84 (**3.0%**); T 1 (≤**0.1%**) | Total count 3180 reads: C 3171; T 9 (≤0.1%) | |
| N | Center 6 | H1047X | wild type | Total count 5182 reads: A 5126; G 56 (**1.1%**); T 0 (0%) | Total count 5181 reads: C 5174; T 7 (≤0.1%) | |
| F | Center 3 | H1047X | wild type | Total count 3179 reads: A 3176; G 3 (≤**0.1%**); T 0 (0%) | Total count 2766 reads: C 2765; T 1 (≤0.1%) | |

ID, identification; IGV, integrative genomics viewer; NGS, next-generation sequencing.
*Only VAFs >3% were considered.
Bold values: % read of the total count reads.

cancer setting with encouraging results (Corné et al., 2021; Galvano et al., 2022), and may play a role in identifying patients harboring *PIK3CA* and possibly additional mutations. An individual patient data meta-analysis recently ascertained the accuracy of circulating tumor DNA (ctDNA) for the detection of *PIK3CA* mutations, raising the possibility of replacing tissue for ctDNA tumor sampling in the future as a preferred strategy for metastatic breast cancer patients with low clinical compliance or inaccessible metastatic sites (Galvano et al., 2022). The latest ESO-ESMO guidelines have also acknowledged the assessment of plasma circulating free DNA as a good alternative to metastatic tumor analysis, as it may overcome the challenges (both for clinicians and patients) of obtaining metastatic tissue biopsies, and as an option for the selection of patients eligible for alpelisib (Cardoso et al., 2020). Furthermore, and further attesting to the relevance of identifying actionable *PIK3CA* mutations in breast cancer, major hot spot-activating missense *PIK3CA* mutations (E542K, E545K/A, H1047R/L) have a level IA of evidence for actionability according to the ESMO Scale for Clinical Actionability of molecular Targets (ESCAT; (Condorelli et al., 2019).

Overall, the results obtained in the U-PIK study validate the *PIK3CA* mutational status testing performed through cobas® and PCR/Sanger sequencing (and NGS) at the eight Portuguese centers participating in this study and show the feasibility of adopting these methodologies as part of the clinical routine practice of patients with advanced ER+/HER2-breast carcinoma, decentralizing the analysis from reference laboratories.

## 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 accessed at: https://www.ebi.ac.uk/ena, PRJEB58369.

## Ethics statement

The study was conducted according to the guidelines of the Declaration of Helsinki. Ethical approval for the study was obtained from multiple boards: Ethics Committee of IPOLFG and Research Council, Unidade de Investigação Clínica, Instituto Português de Oncologia de Lisboa Francisco Gentil, EPE; Ethics Committee of Centro Hospitalar e Universitário de Coimbra; Ethics Committee of Centro Hospitalar e Universitário de Lisboa

Norte and of Centro Académico de Medicina de Lisboa; Ethics Committee of Hospital CUF Descobertas; and Ethics Council and Ethics Council for Health of Fundação Champalimaud. Written informed consent for study participation or research purposes in the scope of the study was provided by participants or participants' legal guardians/next of kin. In addition, ethical approval from other institutions was waived, as the study is in accordance with Article 19 ("DNA Banks and Other Biological Products") of Portuguese Law No. 12/2005 of 26 January ("Personal genetic information and health information"), which states that in the case of using retrospective samples of human origin or in special situations where the consent of subjects involved cannot be obtained due to the amount of data or subjects, their age, or another similar reason, the material and data can be processed but only for purposes of scientific research or epidemiological and statistical data collection (Law No. 12/2005; Kalokairinou et al., 2018). The patients/participants provided their written informed consent to participate in this study.

## Author contributions

JCM, SA, MRT, and PMB conceived the U-PIK Project. JCM, SA, MRT, PMB, FCS, MJA, MJB, PB, MRD, HP, NS, ACS, and JAT participated in the study design. AP, LCi, ALC, MJA, MJB, PB, ASC, LCa, LCo, CF, PF, AG, JP, MP, HP, JR, NS, FCS, FS, ARS, ACS, and JAT performed the experimental work involved in sample preparation, DNA extraction and *PIK3CA* testing and reporting. JCM, SA, and MRT were responsible for analyzing *PIK3CA* testing data from U-PIK testers, for reporting anonymized *PIK3CA* testing performance back to U-PIK testers, and also for the analysis of molecular reports. JCM, SA, MRT, PMB, and AP edited the manuscript, drafted and written with editorial assistance of JCS (Acknowledgements). All authors read and approved the final manuscript.

## Funding

The funding for U-PIK project Research Collaboration was provided by Novartis, covering *PIK3CA* testing reagents, sample preparation and shipping, and also medical writing and study publication support.

## Acknowledgments

The authors acknowledge Joana Cavaco-Silva (jo.cvsilva@gmail.com) for editorial assistance in the drafting and writing of the manuscript.

## Conflict of interest

Author PB and MRD are employed by Novartis, authors LC and NS are employed by Synlab, and author ACS is employed by GenoMed, as described in "author affiliations". All the remaining authors in this research collaboration and consortium are employed either by public or private hospitals or research institutions, as described in the "Author affiliations section". This study received funding from Novartis "funder", covering *PIK3CA* testing reagents, sample preparation, and shipping, and also medical writing and article publication support, as described in detail in the "Funding" section. The authors employed by the "funder", together with the investigators, were involved in study conception and design and in manuscript editing, as described in detail under "Author Contributions".

## Publisher's note



## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1082915/full#supplementary-material

## References

Ahmad, F., Badwe, A., Verma, G., Bhatia, S., and Das, B. R. (2016). Molecular evaluation of PIK3CA gene mutation in breast cancer: Determination of frequency, distribution pattern and its association with clinicopathological findings in Indian patients. *Med. Oncol.* 33, 74. doi:10.1007/s12032-016-0788-y

Albanell, J., Casadevall, D., Sokol, E. S., Albacker, L. A., Elvin, J. A., Vergilio, J.-A., et al. (2019). PIK3CA alterations in metastatic breast cancer (mBC). *Ann. Oncol.* 30, v104–v105. doi:10.1093/annonc/mdz242.001

Anderson, E. J., Mollon, L. E., Dean, J. L., Warholak, T. L., Aizer, A., Platt, E. A., et al. (2020). A systematic review of the prevalence and diagnostic workup of PIK3CA mutations in HR+/HER2− metastatic breast cancer. *Int. J. Breast Cancer* 2020, 3759179–3759216. doi:10.1155/2020/3759179

André, F., Ciruelos, E., Rubovszky, G., Campone, M., Loibl, S., Rugo, H. S., et al. (2019). Alpelisib for PIK3CA -mutated, hormone receptor–positive advanced breast cancer. *N. Engl. J. Med.* 380, 1929–1940. doi:10.1056/NEJMoa1813904

Arsenic, R., Treue, D., Lehmann, A., Hummel, M., Dietel, M., Denkert, C., et al. (2015). Comparison of targeted next-generation sequencing and Sanger sequencing for the detection of PIK3CA mutations in breast cancer. *BMC Clin. Pathol.* 15, 20. doi:10.1186/s12907-015-0020-6

Campbell, I. G., Russell, S. E., Choong, D. Y. H., Montgomery, K. G., Ciavarella, M. L., Hooi, C. S. F., et al. (2004). Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res.* 64, 7678–7681. doi:10.1158/0008-5472.CAN-04-2933

Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70. doi:10.1038/nature11412

Cardoso, F., Paluch-Shimon, S., Senkus, E., Curigliano, G., Aapro, M. S., André, F., et al. (2020). 5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5). *Ann. Oncol.* 31, 1623–1649. doi:10.1016/j.annonc.2020.09.010

Castaneda, C. A., Lopez-Ilasaca, M., Pinto, J. A., Chirinos-Arias, M., Doimi, F., Neciosup, S. P., et al. (2014). PIK3CA mutations in Peruvian patients with HER2-amplified and triple negative non-metastatic breast cancers. *Hematology/Oncology Stem Cell. Ther.* 7, 142–148. doi:10.1016/j.hemonc.2014.09.007

Chakravarty, D., Gao, J., Phillips, S., Kundra, R., Zhang, H., Wang, J., et al. (2017). OncoKB: A precision Oncology knowledge base. *JCO Precis. Oncol.* 2017, 1–16. doi:10.1200/PO.17.00011

ClinicalTrials.gov Study (2015). *Assessing the efficacy and safety of alpelisib plus fulvestrant in men and postmenopausal women with advanced breast cancer which progressed on or after aromatase inhibitor treatment (SOLAR-1).* United States: ClinicalTrials.gov Identifier.

Condorelli, R., Mosele, F., Verret, B., Bachelot, T., Bedard, P. L., Cortes, J., et al. (2019). Genomic alterations in breast cancer: Level of evidence for actionability according to ESMO Scale for clinical actionability of molecular targets (ESCAT). *Ann. Oncol.* 30, 365–373. doi:10.1093/annonc/mdz036

Copur, M. S. (2020). Alpelisib to treat breast cancer. *Drugs Today* 56, 357–363. doi:10.1358/dot.2020.56.6.3137526

Corné, J., Le Du, F., Quillien, V., Godey, F., Robert, L., Bourien, H., et al. (2021). Development of multiplex digital PCR assays for the detection of PIK3CA mutations in the plasma of metastatic breast cancer patients. *Sci. Rep.* 11, 17316. doi:10.1038/s41598-021-96644-6

den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., et al. (2016). HGVS recommendations for the description of sequence variants: 2016 update. *Hum. Mutat.* 37, 564–569. doi:10.1002/humu.22981

Dirican, E., Kaya, Z., Gullu, G., Peker, I., Ozmen, T., Gulluoglu, B. M., et al. (2014). Detection of PIK3CA gene mutations with HRM analysis and association with IGFBP-5 expression levels in breast cancer. *Asian Pac. J. Cancer Prev.* 15, 9327–9333. doi:10.7314/APJCP.2014.15.21.9327

Dogruluk, T., Tsang, Y. H., Espitia, M., Chen, F., Chen, T., Chong, Z., et al. (2015). Identification of variant-specific functions of PIK3CA by rapid phenotyping of rare mutations. *Cancer Res.* 75, 5341–5354. doi:10.1158/0008-5472.CAN-15-1654

European Medicines Agency (2021). Piqray (alpelisib). Available at: https://www.ema.europa.eu/en/medicines/human/EPAR/piqray#authorisation-details-section.

Fitzgerald, D., Muzikansky, A., Pinto, C., Henderson, L., Walmsley, C., Allen, R., et al. (2019). Association between PIK3CA mutation status and development of brain metastases in HR+/HER2-metastatic breast cancer. *Ann. Oncol.* 30, v110. v104–v142. doi:10.1093/annonc/mdz242.013

Fruman, D. A., Chiu, H., Hopkins, B. D., Bagrodia, S., Cantley, L. C., and Abraham, R. T. (2017). The PI3K pathway in human disease. *Cell.* 170, 605–635. doi:10.1016/j.cell.2017.07.029

Galvano, A., Castellana, L., Gristina, V., La Mantia, M., Insalaco, L., Barraco, N., et al. (2022). The diagnostic accuracy of PIK3CAmutations by circulating tumor DNA in breast cancer: An individual patient data meta-analysis. *Ther. Adv. Med. Oncol.* 14, 17588359221110162. doi:10.1177/17588359221110162

Goncalves, M. D., Hopkins, B. D., and Cantley, L. C. (2018). Phosphatidylinositol 3-kinase, growth disorders, and cancer. *N. Engl. J. Med.* 379, 2052–2062. doi:10.1056/NEJMra1704560

Gymnopoulos, M., Elsliger, M.-A., and Vogt, P. K. (2007). Rare cancer-specific mutations in PIK3CA show gain of function. *Proc. Natl. Acad. Sci.* 104, 5569–5574. doi:10.1073/pnas.0701005104

Hanker, A. B., Kaklamani, V., and Arteaga, C. L. (2019). Challenges for the clinical development of PI3K inhibitors: Strategies to improve their impact in solid tumors. *Cancer Discov.* 9, 482–491. doi:10.1158/2159-8290.CD-18-1175

Harlé, A., Lion, M., Lozano, N., Husson, M., Harter, V., Genin, P., et al. (2013). Analysis of PIK3CA exon 9 and 20 mutations in breast cancers using PCR-HRM and PCR-ARMS: Correlation with clinicopathological criteria. *Oncol. Rep.* 29, 1043–1052. doi:10.3892/or.2013.2229

Howlader, N., Altekruse, S. F., Li, C. I., Chen, V. W., Clarke, C. A., Ries, L. A. G., et al. (2014). US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. *JNCI J. Natl. Cancer Inst.* 106, dju055. doi:10.1093/jnci/dju055

Kalinsky, K., Jacks, L. M., Heguy, A., Patil, S., Drobnjak, M., Bhanot, U. K., et al. (2009). PIK3CA mutation associates with improved outcome in breast cancer. *Clin. Cancer Res.* 15, 5049–5059. doi:10.1158/1078-0432.CCR-09-0632

Kalokairinou, L., Howard, H. C., Slokenberga, S., Fisher, E., Flatscher-Thöni, M., Hartlev, M., et al. (2018). Legislation of direct-to-consumer genetic testing in Europe: A fragmented regulatory landscape. *J. Community Genet.* 9, 117–132. doi:10.1007/s12687-017-0344-2

Lee, J. J., Loh, K., and Yap, Y.-S. (2015). PI3K/Akt/mTOR inhibitors in breast cancer. *Cancer Biol. Med.* 12, 342–354. doi:10.7497/j.issn.2095-3941.2015.0089

Law No. 12/2005 (2005). Informação genética pessoal e informação de saúde [Personal genetic information and health information law]. Retrieved from: http://dre.pt/pdf1sdip/2005/01/018A00/06060611.pdf.

Levine, D. A., Bogomolniy, F., Yee, C. J., Lash, A., Barakat, R. R., Borgen, P. I., et al. (2005). Frequent mutation of the PIK3CA gene in ovarian and breast cancers. *Clin. Cancer Res.* 11, 2875–2878. doi:10.1158/1078-0432.CCR-04-2142

Li, S. Y., Rong, M., Grieu, F., and Iacopetta, B. (2006). PIK3CA mutations in breast cancer are associated with poor outcome. *Breast Cancer Res. Treat.* 96, 91–95. doi:10.1007/s10549-005-9048-0

Loibl, S., von Minckwitz, G., Schneeweiss, A., Paepke, S., Lehmann, A., Rezai, M., et al. (2014). PIK3CA mutations are associated with lower rates of pathologic complete response to anti–human epidermal growth factor receptor 2 (HER2) therapy in primary HER2-overexpressing breast cancer. *J. Clin. Oncol.* 32, 3212–3220. doi:10.1200/JCO.2014.55.7876

Martínez-Sáez, O., Chic, N., Pascual, T., Adamo, B., Vidal, M., González-Farré, B., et al. (2020). Frequency and spectrum of PIK3CA somatic mutations in breast cancer. *Breast Cancer Res.* 22, 45. doi:10.1186/s13058-020-01284-9

Mayer, I. A., Abramson, V. G., Formisano, L., Balko, J. M., Estrada, M. V., Sanders, M. E., et al. (2017). A phase ib study of alpelisib (BYL719), a PI3kα-specific inhibitor, with letrozole in ER+/HER2− metastatic breast cancer. *Clin. Cancer Res.* 23, 26–34. doi:10.1158/1078-0432.CCR-16-0134

Mollon, L., Aguilar, A., Anderson, E., Dean, J., Davis, L., Warholak, T., et al. (2018). "Abstract 1207: A systematic literature review of the prevalence of PIK3CA mutations and mutation hotspots in hr+/HER2-metastatic breast cancer," in *Epidemiology* (United States: American Association for Cancer Research), 1207. doi:10.1158/1538-7445.AM2018-1207

Nosho, K., Kawasaki, T., Longtine, J. A., Fuchs, C. S., Ohnishi, M., Suemoto, Y., et al. (2008). PIK3CA mutation in colorectal cancer: Relationship with genetic and epigenetic alterations. *Neoplasia* 10, 534–541. doi:10.1593/neo.08336

Qiagen (2019). Qiagen. Available at: https://www.qiagen.com/us/cmp/mdx/pik3ca-rgq-pcr-kit-us.aspx.

Ross, J. S., Ali, S. M., Wang, K., Khaira, D., Palma, N. A., Chmielecki, J., et al. (2015). Comprehensive genomic profiling of inflammatory breast cancer cases reveals a high frequency of clinically relevant genomic alterations. *Breast Cancer Res. Treat.* 154, 155–162. doi:10.1007/s10549-015-3592-z

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., et al. (2004). High frequency of mutations of the PIK3CA gene in human cancers. *Science* 304, 554. doi:10.1126/science.1096502

Setiawan, V. W., Monroe, K. R., Wilkens, L. R., Kolonel, L. N., Pike, M. C., and Henderson, B. E. (2009). Breast cancer risk factors defined by estrogen and progesterone receptor status: The multiethnic cohort study. *Am. J. Epidemiol.* 169, 1251–1259. doi:10.1093/aje/kwp036

Sobhani, N., Roviello, G., Corona, S. P., Scaltriti, M., Ianza, A., Bortul, M., et al. (2018). The prognostic value of PI3K mutational status in breast cancer: A meta-analysis. *J. Cell. Biochem.* 119, 4287–4292. doi:10.1002/jcb.26687

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

Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., et al. (2019). Cosmic: The catalogue of somatic mutations in cancer. *Nucleic Acids Res.* 47, D941–D947. doi:10.1093/nar/gky1015

US Food and Drug Administration (FDA) (2019). Available at: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-alpelisib-metastatic-breast-cancer.

Verret, B., Cortes, J., Bachelot, T., Andre, F., and Arnedos, M. (2019). Efficacy of PI3K inhibitors in advanced breast cancer. *Ann. Oncol.* 30, x12–x20. doi:10.1093/annonc/mdz381

Wang, Y. L., Dai, X., Li, Y. D., Cheng, R. X., Deng, B., Geng, X. X., et al. (2015). Study of PIK3CA, BRAF, and KRAS mutations in breast carcinomas among Chinese women in Qinghai. *Genet. Mol. Res.* 14, 14840–14846. doi:10.4238/2015.November.18.49

Zhao, L., and Vogt, P. K. (2008). Class I PI3K in oncogenic cellular transformation. *Oncogene* 27, 5486–5496. doi:10.1038/onc.2008.244

# Frontiers in
# Molecular Biosciences

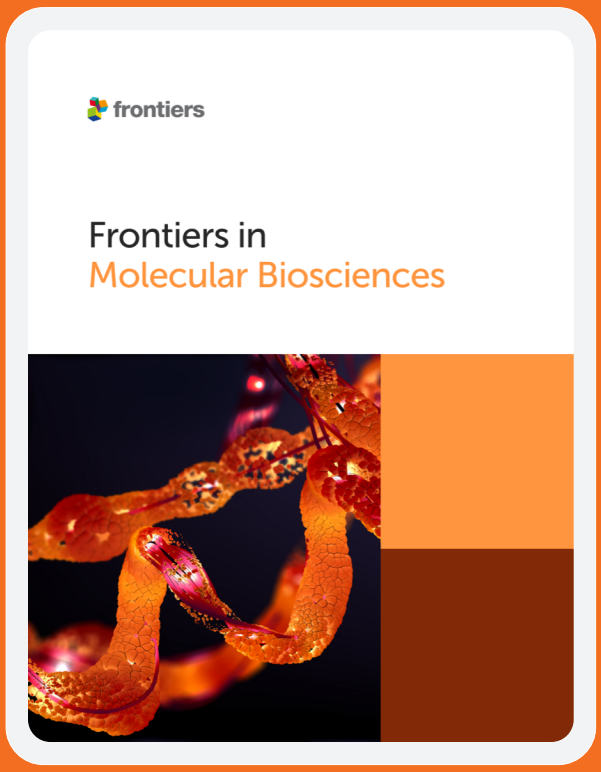
**Explores biological processes in living organisms on a molecular scale**

Focuses on the molecular mechanisms underpinning and regulating biological processes in organisms across all branches of life.

---

## 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