New progress in cancer biomarkers and therapy

Edited by
Guohui Sun, Wei Zhang, Jianhua Wang
and Chengwei He

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

New progress in cancer biomarkers and therapy

Topic editors

Guohui Sun − Beijing University of Technology, China
Wei Zhang − University of Southern California, United States
Jianhua Wang − Capital Institute of Pediatrics, China
Chengwei He − University of Macau, China

Table of contents

Editorial: New progress in cancer biomarkers and therapy

Guohui Sun[1]*, Chengwei He[2]* and Jianhua Wang[3]*

[1]Beijing Key Laboratory of Environmental and Viral Oncology, College of Chemistry and Life Science, Beijing University of Technology, Beijing, China, [2]State Key Laboratory of Quality Research in Chinese Medicine, Institute of Chinese Medical Sciences, University of Macau, Taipa, China, [3]Beijing Municipal Key Laboratory of Child Development and Nutriomics, Capital Institute of Pediatrics, Beijing, China

Editorial on the Research Topic
New progress in cancer biomarkers and therapy

As we are aware, cancer poses a significant threat to human health, with projections suggesting it may surpass cardiovascular disease as the primary cause of premature death in many countries in the coming years, as per the latest data from "Global Cancer Statistics 2020" (Sung et al., 2021). Despite that sounds pretty terrible, there's optimism surrounding advancements in tumor biology that could potentially improve this scenario (Nakamura et al., 2021; Bai et al., 2023; Loi et al., 2023). Of notable importance are cancer biomarkers, which carry substantial clinical implications, aiding in early detection, monitoring treatment progress, and predicting cancer prognosis (Sarhadi and Armengol, 2022; Wang and Deng, 2023). Over the past several decades, analytical techniques in tumor biomarker research are essential tools for identifying, quantifying, and characterizing molecules or proteins that can serve as indicators of cancer presence, progression, or response to treatment. These techniques encompass a range of methods, each offering unique advantages and applications, including but not limited (Jayanthi et al., 2017; Japp et al., 2021; Eftekhari et al., 2022; Lee et al., 2023): immunohistochemistry (IHC), enzyme-linked immunosorbent assay (ELISA), flow cytometry, fluorescent probe, biosensor, polymerase chain reaction (PCR), next-generation sequencing (NGS), mass spectrometry (MS), surface enhanced raman spectroscopy (SERS), liquid biopsy, bioinformatics, *etc.* These analytical techniques play a crucial role in the discovery, validation, and clinical application of tumor biomarkers, contributing to improved cancer diagnosis, prognosis, and treatment outcomes. In this Research Topic, our focus lies on the theme of "cancer biomarkers," encompassing various aspects such as: 1) the exploration of new cancer biomarkers, which includes novel analytical techniques, identification of new molecules or proteins, and their association with emerging cancer types; 2) understanding the biological properties inherent to cancer biomarkers; 3) elucidating the mechanisms of action and biological significance of these biomarkers; and 4) exploring how biomarkers can inform chemotherapy and biotherapy strategies. In addition to traditional methods, cutting-edge tools and concepts such as proteomics, bioinformatics, machine learning, artificial intelligence (AI), and single-cell sequencing play pivotal roles in the discovery and comprehension of cancer biomarkers. Particularly noteworthy is the potential of cancer biomarkers in shaping clinical approaches to cancer prevention and treatment, warranting in-depth investigation.

In this Research Topic, we successfully published 12 high-quality papers that will be of interest for researchers in cancer biomarkers. Of the 12 papers, seven are related to the cancer risk prediction, three are related to the potential of small metabolites as cancer biomarkers, the application of fluorescence *in situ* hybridization (FISH) in cancer diagnosis and the influencing factors of caner biomarkers. Another two papers are reviews that summarize the biological implications of potential cancer biomarkers.

The opening paper conducted by Yang et al. highlights the significance of M2 macrophage-related genes in both the treatment and prognosis of pancreatic cancer. Using multiple bioinformatic tools, the authors constructed the risk predictive models and revealed that the risk levels were closely associated with tumor mutational burden, immune checkpoint blockade related genes, and immune cells. Moreover, they predicted the potential associations between different risk models and the efficacy of chemotherapeutic agents (e.g., metformin, paclitaxel and lapatinib), and underscored the utility of WGCNA-based analysis of M2 macrophage-related genes in prognosticating outcomes for pancreatic cancer patients and suggested novel avenues for immunotherapy in this context. In the next work, Deng et al. explored the role of lysine acetylation-related genes (LARGs) in oral squamous cell carcinoma (OSCC) using bioinformatic methods. They developed a lysine acetylation-related prognostic model using TCGA OSCC datasets and revealed that patients with lower risk scores had better prognoses in both the overall cohort and within the subgroups, thus offering a new model for classifying OSCC and determining its prognosis. Xie et al. examined Centrosomal Protein 55 (CEP55) as a cancer-testis antigen, assessing its expression in tumors and its impact on prognosis. They developed a CEP55-based model for hepatocellular carcinoma (HCC), linking high CEP55 levels to increased cell cycle activity, proliferation, and immune pathways. CEP55 correlated with immune modulators and showed promise in predicting responses to immune checkpoint inhibitors (ICIs). The study also associated CEP55 expression with specific HCC molecular subtypes and devised a nomogram for survival prediction. Overall, CEP55 may serve as a prognostic biomarker and predictor of ICI efficacy, potentially influencing tumor immune microenvironments across various cancers. In head and neck squamous cell carcinoma (HNSCC), tumor microenvironment (TME) also plays an important role in tumor progression, however, the relationship between TME characteristics and the prognosis of HNSCC patients remains poorly understood. Wan et al. utilized the "estimate" R package to calculate the immune and stromal cell scores and identify seven new markers. They constructed a risk model categorizing HNSCC samples into low- and high-risk groups, validated for accuracy using Kaplan-Meier survival and ROC analyses. CIBERSORT algorithm revealed significant differences in immune cell infiltration between risk groups. These findings shed light on TME roles and unveil new prognostic biomarkers for HNSCC patients.

In acute myeloid leukemia (AML), ferroptosis offers potential against drug resistance. Using TCGA data, Wu et al. created a prognostic model incorporating eight prognosis-related ferroptosis genes (PRFGs) via LASSO regression. The constructed nomogram, integrating LASSO score, age, and cytogenetic risk, accurately predicts overall survival. Low-risk patients demonstrate significantly improved survival. Gene expression analyses reveal

the relevance of PARPs with different clinical subgroups and the overall survival in AML patients. Immune-related pathways influence prognosis disparities, suggesting a TME role. Combining PARP inhibitors with ferroptosis inducers shows promise as an AML therapy. This comprehensive approach aids in patient stratification and prognosis, offering novel treatment avenues. In another paper, Geng et al. developed an extracellular matrix (ECM)-based prediction model for ovarian serous adenocarcinoma survival using AI techniques. Analyzing TCGA-OV data, 15 key ECM genes were identified, validating the ECM risk score's predictive efficacy. Multivariate COX analysis revealed independent prognostic factors. High ECM risk score patients responded better to thyroglobulin-targeted immunotherapy, while low-risk patients benefitted from RYR2 gene-related treatment. Low-risk patients showed elevated immune checkpoint gene expression and immunophenoscore levels, indicating better immunotherapy response. The ECM risk score serves as a reliable tool for immunotherapy sensitivity assessment and ovarian cancer prognosis prediction. Zhao et al. reported an interesting study aimed to identifying clinical-significant circadian clock (CC)-related genes in ovarian cancer (OC). Using TCGA data, 12 CC genes was analyzed to generate a Circadian Clock Index (CCI). High CCI correlated with poor overall survival (OS) and immune biomarkers. WGCNA identified a CCI-correlated gene module, yielding 15 hub genes significantly associated with OS and immune cell infiltration. Upstream regulators such as transcription factors and miRNAs of key genes were also predicted. These findings reveal 15 crucial CC genes having indicative value for OC prognosis and immune microenvironment, offering insights into OC molecular mechanisms.

Small molecule metabolites may serve as ovarian cancer biomarkers, yet causal links are unclear. Utilizing Mendelian randomization, Chang et al. identified 242 single nucleotide polymorphisms (SNPs) correlated with small molecule metabolites as instrumental variables to elucidate the causal relationship. Six metabolites correlated with reduced ovarian cancer risk, including hexadecenoylcarnitine and methioninesulfoxide. Fifteen metabolites were associated with subtype cancers; methionine sulfoxide and tetradecanoyl carnitine linked to reduced risk in clear cell and high-grade serous cancers, while tryptophan elevated risk in endometrioid cancer. These findings suggest potential biomarkers for early detection and highlight metabolites' etiological roles, warranting further investigation into underlying mechanisms. UroVysionTM FISH, often used for urothelial carcinoma (UC), may also detect carcinoma of non-urothelial lineages (CNUL). Ke et al. found that 64% of CNUL cases showed positive urine FISH results. Histological FISH results aligned with urine FISH in most cases, suggesting FISH's applicability in CNUL diagnosis. Squamous cell carcinoma antigen (SCCA) is a specific biomarker of squamous cell carcinoma, however, the elevation of SCCA in pneumonia patients without malignancy has not been studied. Wang et al. analyzed the influencing factors of SCCA elevation in community-acquired pneumonia patients. Among 309 community-acquired pneumonia (CAP) patients with normal serum indicator levels, 46.3% showed elevated SCCA. Age inversely affects SCCA elevation (OR = 0.97), while higher body temperature significantly increases risk (OR = 3.75). Patients in higher

quartiles of body temperature face substantially elevated SCCA risk. Thus, age and body temperature influence SCCA levels in CAP patients, with higher temperatures indicating heightened SCCA risk.

The last two papers are reviews by Yang et al. and Huang et al., respectively. Tyrosineprotein kinase-1 (ROS1) gene rearrangements occur in 0.9%–2.6% of non-small-cell lung cancers (NSCLCs). Targeting ROS1 can effectively inhibit tumor growth, offering clinical benefits. Yang et al. synthesizes insights into ROS1 rearrangements in NSCLCs, covering their oncogenic mechanisms, prevalence, detection methods, molecular features, therapeutic options, and drug resistance mechanisms. Abnormal translation regulation, crucial in cancer, involves eukaryotic translation initiation factor 4A1 (eIF4A1), an RNA helicase. It is regulated by microRNAs and long non-coding RNAs, impacting tumor cell proliferation and metastasis. Huang et al. summarized that eIF4A1 could serve as a biomarker for tumor diagnosis, staging, and outcome prediction, aiding precision medicine and targeted therapy. Small molecule inhibitors also show promise in clinical practice, supporting eIF4A1's therapeutic potential.

In conclusion, we sincerely hope that the articles included in this Research Topic on cancer biomarkers will contribute to the research on cancer prevention and treatment, particularly by providing valuable insights into early diagnosis, biomarker-based therapy, and effective prognosis for tumors.

Author contributions

GS: Writing–original draft, Writing–review and editing. CH: Writing–review and editing. JW: Writing–review and editing.

Funding

Acknowledgments

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

Bai, P., Fan, T., Wang, X., Zhao, L., Zhong, R., and Sun, G. (2023). Modulating MGMT expression through interfering with cell signaling pathways. *Biochem. Pharmacol.* 215, 115726. doi:10.1016/j.bcp.2023.115726

Eftekhari, A., Maleki Dizaj, S., Sharifi, S., Salatin, S., Khalilov, R., Samiei, M., et al. (2022). "Chapter six - salivary biomarkers in cancer," in *Advances in clinical chemistry*. Editor G. S. Makowski (Elsevier), 171–192. doi:10.1016/bs.acc.2022.06.005

Japp, N. C., Souchek, J. J., Sasson, A. R., Hollingsworth, M. A., Batra, S. K., and Junker, W. M. (2021). Tumor biomarker in-solution quantification, standard production, and multiplex detection. *J. Immunol. Res.* 2021, 9942605. doi:10.1155/2021/9942605

Jayanthi, V. S. P. K., Sankara, A., Das, A. B., and Saxena, U. (2017). Recent advances in biosensor development for the detection of cancer biomarkers. *Biosens. Bioelectron.* 91, 15–23. doi:10.1016/j.bios.2016.12.014

Lee, Y., Ni, J., Beretov, J., Wasinger, V. C., Graham, P., and Li, Y. (2023). Recent advances of small extracellular vesicle biomarkers in breast cancer diagnosis and prognosis. *Mol. Cancer* 22 (1), 33. doi:10.1186/s12943-023-01741-x

Loi, S., Settleman, J., Joyce, J. A., Pramesh, C. S., Bernards, R., Fan, J., et al. (2023). The next big questions in cancer research. *Cell* 186 (8), 1523–1527. doi:10.1016/j.cell.2023.01.037

Nakamura, Y., Kawazoe, A., Lordick, F., Janjigian, Y. Y., and Shitara, K. (2021). Biomarker-targeted therapies for advanced-stage gastric and gastro-oesophageal junction cancers: an emerging paradigm. *Nat. Rev. Clin. Oncol.* 18 (8), 473–487. doi:10.1038/s41571-021-00492-2

Sarhadi, V. K., and Armengol, G. (2022). Molecular biomarkers in cancer. *Biomolecules* 12 (8), 1021. doi:10.3390/biom12081021

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

Wang, Y., and Deng, B. (2023). Hepatocellular carcinoma: molecular mechanism, targeted therapy, and biomarkers. *Cancer Metastasis Rev.* 42 (3), 629–652. doi:10.1007/s10555-023-10084-4

An artificial intelligence prediction model based on extracellular matrix proteins for the prognostic prediction and immunotherapeutic evaluation of ovarian serous adenocarcinoma

Tianxiang Geng[1], Mengxue Zheng[2], Yongfeng Wang[3], Janne Elin Reseland[1] and Athina Samara[1]*

[1]Department of Biomaterials, FUTURE, Center for Functional Tissue Reconstruction, Faculty of Dentistry, University of Oslo, Oslo, Norway, [2]Laboratory of Reproductive Biology, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [3]Department of Obstetrics and Gynecology, Seventh People's Hospital of Shanghai University of Traditional Chinese Medicine, Shanghai, China

Background: Ovarian Serous Adenocarcinoma is a malignant tumor originating from epithelial cells and one of the most common causes of death from gynecological cancers. The objective of this study was to develop a prediction model based on extracellular matrix proteins, using artificial intelligence techniques. The model aimed to aid healthcare professionals to predict the overall survival of patients with ovarian cancer (OC) and determine the efficacy of immunotherapy.

Methods: The Cancer Genome Atlas Ovarian Cancer (TCGA-OV) data collection was used as the study dataset, whereas the TCGA-Pancancer dataset was used for validation. The prognostic importance of 1068 known extracellular matrix proteins for OC were determined by the Random Forest algorithm and the Lasso algorithm establishing the ECM risk score. Based on the gene expression data, the differences in mRNA abundance, tumour mutation burden (TMB) and tumour microenvironment (TME) between the high- and low-risk groups were assessed.

Results: Combining multiple artificial intelligence algorithms we were able to identify 15 key extracellular matrix genes, namely, *AMBN, CXCL11, PI3, CSPG5, TGFBI, TLL1, HMCN2, ESM1, IL12A, MMP17, CLEC5A, FREM2, ANGPTL4, PRSS1, FGF23*, and confirm the validity of this ECM risk score for overall survival prediction. Several other parameters were identified as independent prognostic factors for OC by multivariate COX analysis. The analysis showed that thyroglobulin (TG) targeted immunotherapy was more effective in the high ECM risk score group, while the low ECM risk score group was more sensitive to the RYR2 gene-related immunotherapy. Additionally, the patients with low ECM risk scores had higher immune checkpoint gene expression and immunophenoscore levels and responded better to immunotherapy.

Conclusion: The ECM risk score is an accurate tool to assess the patient's sensitivity to immunotherapy and forecast OC prognosis.

Introduction

Ovarian cancer (OC) is one of the most common gynaecological malignancies. According to the Global Cancer Observatory of the World Health organization (WHO) international agency for research on cancer, a total of 207,252 new fatalities due to ovarian cancer were reported in 2020, placing it 14th out of 36 different types of tissue cancers (World Health Organization International Agency for Research on Cancer, 2020). Most ovarian malignancies originate from epithelial cells, and the most prevalent histological subtype of epithelial ovarian cancer is ovarian serous adenocarcinoma (Heintz et al., 2006). Early OC detection is the best treatment scenario, but as OC presents with nonspecific symptoms and reflects detection, most patients are given a stage III diagnosis, indicating that the disease has spread throughout the peritoneum and/or has involved the lymph nodes (Prat and FIGO Committee on Gynecologic Oncology, 2014). A multi-stage evaluation is necessary to manage OC, to determine personalized treatment, and to predict the presence of distant metastases, tumour stage and prognosis.

As a new treatment option, immunosuppressants, address the tumour microenvironment (TME) (Pitt et al., 2016). For ovarian cancer, this cutting-edge therapeutical approach is recently being studied and applied (Yang et al., 2022, 2023). Despite the fact that many variables have been demonstrated to predict the therapeutic effectiveness of immunosuppressant's, the accuracy of this strategy still needs to be improved (Gibney et al., 2016). Tumour development, spread and invasion are all dependent on the TME (Schreiber et al., 2011; Lei et al., 2020), which contains multiple cell types, including stroma, vasculature, secretory factors, surrounding stroma and the internal environment of the tumour cells. As the TME is primarily determined by the genomic landscape of the tumour, several algorithms have been developed to predict tumour purity and estimate the abundance of tumour-infiltrating immune cells based on gene expression profiles (Tamborero et al., 2018). These include CIBERSORT, MCP, Xcell, EPIC, ESTIMATE, Timer, IPS, and QuantiSeq.

As an essential component of TME, the non-cellular network surrounding the cells, known as the extracellular matrix (ECM), is tightly associated to the pathophysiology of healthy and cancerous tissue (Frantz et al., 2010; Henke et al., 2020; Zhu et al., 2022). This renders ECM a crucial study niche for the initiation, progression, dissemination, and furthermore treatment and prognosis of epithelial ovarian cancer (Ween et al., 2011). The metabolic disruption of various ECM protein-related factors derived from epithelial cells during tumorigenesis leads to the formation of a pro-tumorigenic microenvironment that favors tumor growth and metastasis. This is followed by tumour cell-mediated ECM remodelling, which ultimately promotes the survival of tumour cells at the expense of healthy tissue (Zigrino et al., 2005). Therefore, ECM proteins, which have bidirectional effects on the generation, recurrence and metastasis of tumour cells (Valmiki et al., 2021 should be considered key players to the treatment and prognosis of tumours (Donelan et al., 2022; Zhu et al., 2022).

An artificial intelligence algorithm, Random Forest (RF) has been recently employed to predict disease progression by virtue of its high performance and interpretability (Wu et al., 2021). A convincing predictive model can be constructed by combining analysis of gene expression data with diagnostic and therapeutic data. This model could be effective at forecasting patient survival, the course of the tumour, and recurrence following various types of treatment (Lin et al., 2022; Miao et al., 2022). Big data machine learning may also also be applied. Despite the recent advances in machine learning methods for ovarian cancer survival analysis, integration of multi-omics data with immunotherapy targeting is an approach that has not been thoroughly explored (Henderson et al., 2016; Belotti et al., 2022). This approach could be advantageous for the identification of potential therapeutic targets and may lead to improved outcomes for ovarian cancer patients.

In this study we used artificial intelligence algorithms to integrate multifaceted omics data with immunotherapy targets in ovarian cancer. Specifically, we employed the Random Forest and Lasso algorithms to process gene expression and survival data from the TCGA database. The tumour risk score was calculated to construct features for predicting OC prognosis and immunotherapy efficacy.

Materials and methods

Datasets and data quality control

Transcriptome expression profiles, somatic mutation data and clinical survival data were downloaded from the TCGA database (Supplementary Table S1). FPKM expression data from the UCSC XENA Project (https://xenabrowser.net/datapages/), which included the TCGA cancer gene expression sequencing data, were analysed together to increase the reliability of data analysis. Normal ovary tissue transcriptome sequencing data from the GTEx database (https://www.gtexportal.org/home/) were used as representative normal/healthy tissue data. We utilized the immune cell markers used in the Tumour MicroEnvironment (TME) analysis following the method described at Bindea et al. (2013) and ECM-related gene information was obtained from Naba et al. (2016). Following quality control measures on gene expression data and somatic mutations (SNPs and small INDELs), we filtered out 373 valid sample samples from the pool of 758 valid patient survival datasets of the TCGA-OV collection.

Construction and evaluation of an ECM risk score model related to survival

The TCGA-OV data were randomly partitioned into a training set ($n = 298$) and a test set ($n = 75$). We used the randomForestSRC package (3.1.1) (Ishwaran et al., 2022) to down-size the 1068 ECM genes including survival information of OC patients. Further

dimensionality reduction was performed by the Lasso algorithm in
the glmnet package (4.1–2) (Friedman et al., 2010). Survival analysis
of key genes in OV was performed with multivariate COX regression
in the survival package (3.2–10) (Therneau, 2015).

Differential expression and enrichment analyses

Two groups of patients with high and low-risk scores were generated.
To calculate the differential gene expression between cancer data and
normal tissue data we used DESeq2 v.1.36.0 (Love et al., 2014) in R (4.2.1).
We performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes
and Genomes (KEGG) enrichment analyses using ClusterProfiler v.3.14.3
(Yu et al., 2012) in R (3.6.3). To find BP term enrichment, the Gene Set
Enrichment Analysis (GSEA) (Subramanian et al., 2005) of ranked lists of
differentially expressed genes was carried out. Significant enrichment in
GSEA analysis is achieved when the False discovery rate (FDR) is 0.25 and
an adjusted p-value of 0.05.

Tumour microenvironment (TME) and somatic mutation analyses of the TCGA-OV dataset

We used the "maftools" package (2.12.0) (Mayakonda et al.,
2018) for the calculation and evaluation of somatic mutations for
each patient. The "drugInteractions" function was employed to
analyse the correspondence between mutated genes and currently
available genetic drugs based on the DGIdb database (Griffith et al.,
2013). We further used multiple algorithms built into the IOBR
package (0.99.9) (Zeng et al., 2021) to assess the immune cell
infiltration level, including B cells, CD4+ T cells, CD8+ T cells,
neutrophils, macrophages and dendritic cells. Then we explored the
variations in immune infiltration and somatic mutation between
groups with high and low-risk scores.

Statistical analysis

The differences between the two datasets were determined using the
Mann-Whitney U test (also known as the Wilcoxon rank sum test) and
independent t-test. To evaluate between-group differences, one-way
analysis of variance (ANOVA) with the Kruskal–Wallis test and chi-
square test were utilised. Correlation analysis was conducted using non-
parametric Spearman correlation tests. The connection between potential
genes and overall survival was examined using a single-variable Cox
regression analysis (OS). The difference was shown to be statistically
significant when $p < 0.05$ was used ($p < 0.05$ *; $p < 0.01$ **; $p < 0.001$ ***).

Results

Screening and validation of ECM-related prognostic key genes

The clinicopathological characteristics of 379 OC patients in the
TCGA database, are summarized in Table 1. The random forest

TABLE 1 Characteristics of OC patients; source TCGA database.

Characteristics		N
Age	≤60	208
	>60	171
	Total	379
OS	Alive	147
	Dead	232
	Total	379
FIGO stage	Stage I	1
	Stage II	23
	Stage III	295
	Stage IV	57
	Total	376

algorithm was used to decrease the training set. 147 genes were
screened out of 1068 ECM-related genes, and the accuracy of this
survival prediction model was validated using the test set. The
receiver operating characteristic curve (ROC) for the training set
and test set were plotted separately, with the area under the curve
(AUC) of 0.810 for the training set and 0.684 for the test set
(Figure 1A).

The results of "lambda.min" of the Lasso algorithm were
employed and 15 key genes closely related to prognosis were
obtained (Figures 1B, C). These were Ameloblastin (AMBN),
Chemokine (C-X-C motif) ligand 11 (CXCL11), Peptidase
inhibitor 3 (PI3), Chondroitin sulfate proteoglycan 5 (CSPG5),
Transforming growth factor (TGFBI), Tolloid-like 1 (TLL1),
Endothelial cell-specific molecule 1 (ESM1), Matrix
metallopeptidase 17 (MMP17), Angiopoietin-like 4 (ANGPTL4),
Fibroblast growth factor 23 (FGF23), Hemicentin 2 (HMCN2),
Interleukin 12A (IL12A), C-type lectin domain family 5, member
A (CLEC5A), FRAS1 related extracellular matrix protein 2 (FREM2),
serine protease 1 (PRSS1), and the gen prediction model with the
risk coefficient of 15 genes given by the Lasso algorithm was
constructed:

$$\begin{aligned}
\text{Risk score} = {}& TGFBI * (0.0092) + CSPG5 * (-0.0307) \\
& + PI3 * (0.0481) + CXCL11 * (-0.1219) \\
& + MMP17 * (0.0512) + IL12A * (0.0084) \\
& + ESM1 * (-0.0317) + HMCN2 * (0.0707) \\
& + TLL1 * (0.4759) + FGF23 * (0.0328) \\
& + PRSS1 * (-0.0277) + ANGPTL4 * (0.0245) \\
& + FREM2 * (-0.0014) + CLEC5A * (0.0311) \\
& + AMBN * (1.2734)
\end{aligned}$$

The Cox model was used to verify the predictive ability of the
15 key genes for the 1-year, 3-year and 5-year overall survival (OS),
and the key genes fit well with the ideal line at the three-time points
(Figure 1D). The TCGA-OV sample was divided into two groups
with high and low-risk scores based on the average risk score.
Furthermore, the Kaplan Meier (KM) curves of high/low-risk
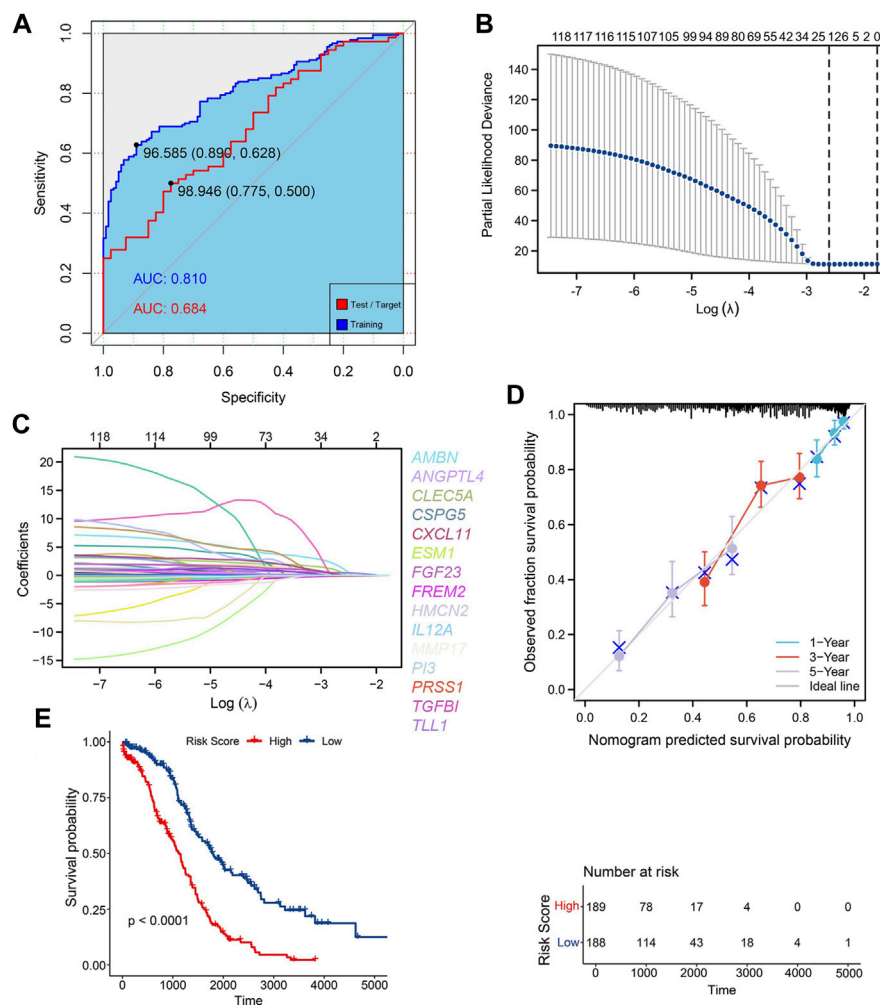
FIGURE 1
Construction of ECM risk score model **(A)** AUC for random forest train model (blue) and test model (red); **(B)** Lasso dimensionality reduction for random forest prognosis model; **(C)** Locus plot of all genes in random forest prognosis model; **(D)** Nomogram survival plot for 15 key prognosis genes; **(E)** KM survival plot for high/low ECM risk score group.

score groups were plotted, showing a significant difference between the high and low-risk score groups (Figure 1E). Additional multifactorial Cox models were used to analyse the relationship between the 15 key genes and ovarian cancer OS (Table 2), and we found that *AMBN, CXCL11, CLEC5A,* CSPG5 *FREM2, MMP17,* and *PI3* were independent prognostic factors for ovarian cancer.

Differential expression analysis and functional enrichment of high and low ECM risk score groups

The ECM risk score, survival information, and one-to-one correspondence to the expression of the 15 key genes for each sample in TCGA-OV are presented in Figure 2A. The results of differential gene expression of high vs. the low-risk score group showed that 1004 genes were significantly upregulated (logFC > 0.4, adj. $p < 0.05$), and 378 genes were significantly downregulated (Figure 2B).

We documented that there is a substantial difference in the extracellular matrix-related processes in the enrichment of GOKEGG functions (Supplementary Figure S1A). Additionally, it was shown that biological processes associated with immune cells differed dramatically (GO:0071621, GO:0043030 et al.). Two immune-related pathways were revealed to be blocked in the high ECM risk score group in the GSEA results (Supplementary Figure S1C). Based on the 15 key prognostic genes in the TCGA-OV at the principal component analysis (PCA cluster), there was little difference between the high and low-risk subgroups in the PC1 and PC2 dimensions (Figure 2C).

We also examined the expression of the 15 key genes in normal ovarian tissue vs. the TCGA-OV collection, and in low/high ECM risk score groups (Figures 2D, E). In the expression analysis of normal vs. tumour tissues, only *PRSS1* was not significantly differentially expressed in normal versus tumour tissues. The expression levels of *TLL1, HMCN2, FREM2 and MMP17* were significantly higher in normal ovarian tissues than in tumour

TABLE 2 Multifactorial Cox survival analysis of the 15 key genes in TCGA-OV patients.

Characteristics	High	Low (Reference)	Univariate analysis		Multivariate analysis	
			Hazard ratio (95% CI)	*p* value	Hazard ratio (95% CI)	*p* value
TGFBI	186	187	1.118 (0.863–1.448)	0.398	0.822 (0.587–1.150)	0.252
CSPG5	186	187	0.754 (0.582–0.978)	0.033	0.811 (0.611–1.077)	**0.047**
P13	186	187	1.470 (1.133–1.908)	**0.004**	1.379 (1.044–1.820)	**0.023**
CXCL11	186	187	0.614 (0.472–0.798)	**<0.001**	0.567 (0.425–0.756)	**<0.001**
MMP17	186	187	1.689 (1.301–2.192)	**<0.001**	1.553 (1.150–2.098)	**0.004**
IL12A	186	187	1.010 (0.780–1.308)	0.939	0.942 (0.716–1.238)	0.666
ESM1	186	187	0.936 (0.723–1.213)	0.618	1.035 (0.783–1.369)	0.809
HMCN2	186	187	1.503 (1.157–1.953)	**0.002**	1.172 (0.871–1.578)	0.294
TLL1	186	187	1.400 (1.080–1.815)	**0.011**	1.285 (0.966–1.711)	0.085
FGF23	186	187	1.191 (0.919–1.543)	0.185	1.085 (0.821–1.436)	0.565
PRSS1	186	187	0.818 (0.630–1.061)	0.131	0.979 (0.738–1.299)	0.885
ANGPTL4	186	187	1.377 (1.061–1.786)	0.016	1.064 (0.794–1.424)	0.679
FREM2	186	187	0.724 (0.558–0.940)	0.015	0.633 (0.473–0.847)	**0.002**
CLEC5A	186	187	1.577 (1.213–2.051)	**<0.001**	1.444 (1.044–1.998)	**0.026**
AMBN			6.630 (2.508–17.526)	**<0.001**	8.544 (3.045–23.976)	**<0.001**

*Total number of patients 373.
Statistically significant values are indicated in bold.

tissues. *AMBN, TGFBI, CSPG5, PI3, CXCL11, ESM1, FGF23, ANGPTL4, CLEC5A* and *IL12A* all showed significantly higher expression in tumour tissues samples.

In the differential expression analysis of low/high ECM risk score groups, *AMBN, IL12A* and *FREM2* were not statistically different, whereas *PI3, TGFBI, TLL1, HMCN2, MMP17, CLEC5A, ANGPTL4, FGF23* had higher expression in the high-risk group and *CXCL11, CSPG5, ESM1, PRSS1* were highly expressed in the low-risk group.

The genes *BCRA2, MUC1* and *MUC16* (in the red-framed rectangles in Figures 1C, D, and the heat map), were also assessed to supplement our analysis with three genes from the same dataset, previously functionally characterized for their prognostic role in OC (Wang et al., 2007; Zhai et al., 2020; Custódio et al., 2022). The genes are also.

Assessing the role of ECM risk score in tumour immune cell infiltration and immunotherapy response

The analysis using almost all algorithms, documented that CD8 T cells showed a significant difference, with lower levels of infiltration in the high-risk group than in the low-risk group. In the high-risk group, the signature score of CD4 T memory resting cells was higher, and lower in all other T cells (Figure 3A). Neutrophils scored variable results among the four algorithms: there were group differences in the infiltration levels of neutrophils in the CIBERSORT and MCPcounter algorithms, with both algorithms

showing higher levels in the high-risk group (Figure 3B). According to the EPIC, MCPcounter, and xCell algorithms, the high-risk group had higher numbers of cancer-associated fibroblasts (CAFs) (Figure 3C).

Both the xCell and ESTIMATE algorithms indicated a lower immune microenvironment score for the high-risk group when computing the immune microenvironment score. The high-risk group displayed a higher stromal score in the ESTIMATE algorithm, indicating the presence of more stromal cells (Figure 3D).

The analysis of B cells also showed high variation among the algorithms used: significant differences between groups were only documented by the xCell algorithm; naive B cells and plasma cells showed group differences under the CIBERSORT algorithm but not when the xCell algorithm was employed (Supplementary Figure S2). There was no discernible difference between the two groups in monocytes (Supplementary Figure S2). Only xCell revealed group differences in DC cells (Supplementary Figure S2). NK cells only showed between-group differences under the *MCPcounter* and *quantiseq* algorithms, but there was an opposite trend: NK cells showed relatively low levels in the high-risk group under the *MCPcounter* algorithm but relatively high levels in the high-risk group under the quantiseq algorithm (Supplementary Figure S2). The high-risk group's Macrophage M1 levels were only marginally different to the low-risk group according to the CIBERSORT and xCell results, and there was no difference between groups when the EPIC method was used (Supplementary Figure S2).

We also extracted the expression levels of eight immunological checkpoint genes (Figure 3E). The expression
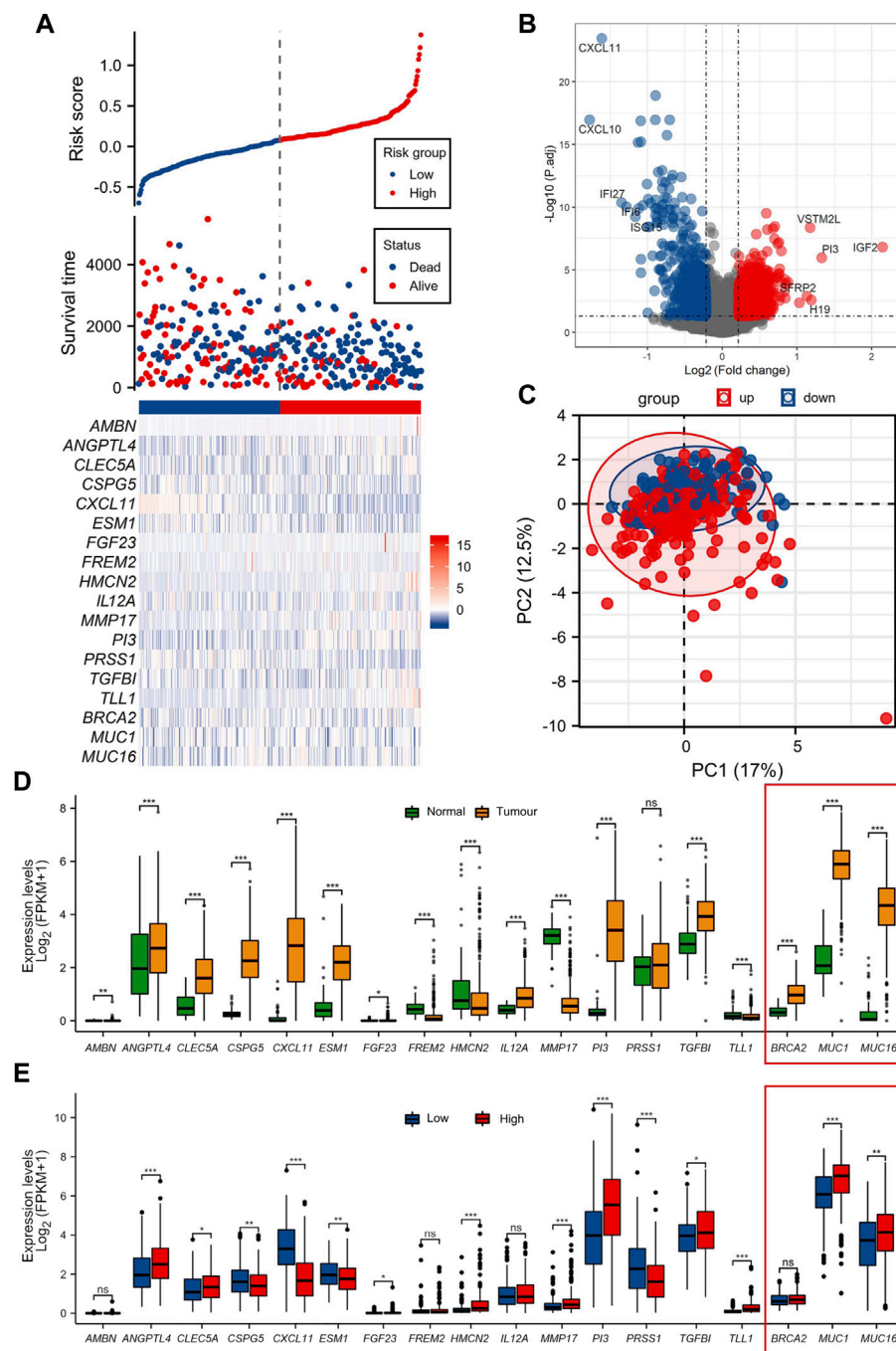
FIGURE 2
Differential analysis of function between high/low ECM risk score group **(A)** Information of sample group based on ECM risk score and 15 key prognosis gene expression heatmap. **(B)** Volcano map of differential gene expression analysis in TCGA-OV. **(C)** Principal component analysis (PCA cluster) based on the 15 key prognostic genes in TCGA-OV. **(D)** 15 key genes expressed in normal ovary and TCGA-OV. **(E)** 15 key genes expression in ECM high/low-risk score groups. The genes *BCRA2*, *MUC1* and *MUC16* (in red frames) have been functionally characterized in other studies, for their prognostic role in OC, and used as a reference.

levels of six immune checkpoint genes (*TIGIT, CD274, PDCD1, CTLA4, LAG3, and PDCD1LG2*) was higher in the low-risk group than in the high-risk group, with the exception of *SIGLEC15* and *HAVCR2*. The IPS score for the major histocompatibility complex (MHC), and for senescent cells (SC) was greater in the low-risk group than in the high-risk

group. Endothelial cells (EC) IPS score did not differ statistically significantly between the two groups. However, the high-risk group had a higher Classical Pathway (CP) IPS score than the low-risk group. Both the aggregated z-score (AZ) and the weighted total IPS showed that the low-risk group was higher than the high-risk group (Figure 3F).
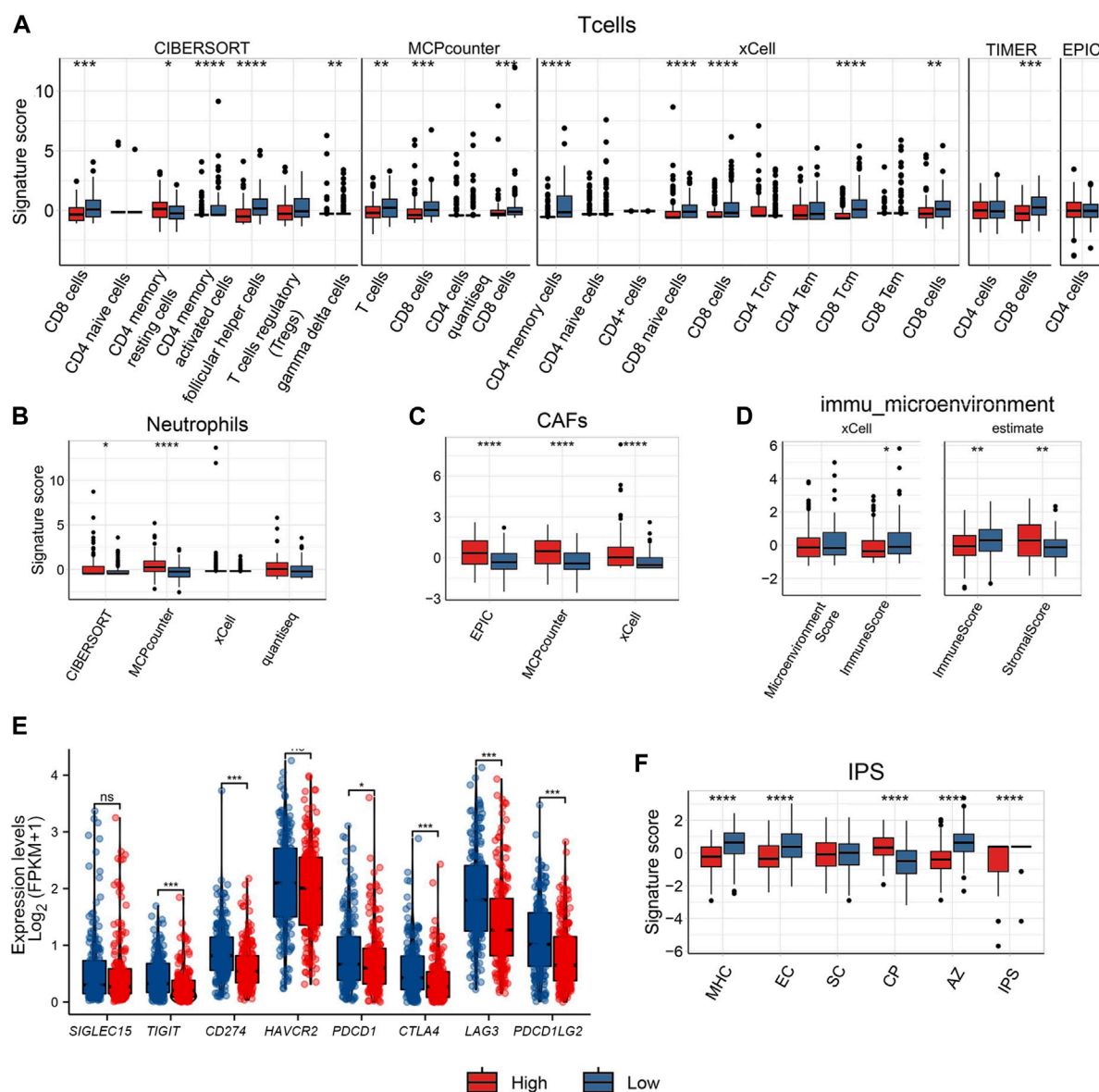
FIGURE 3
Comparison of immune cell infiltration between high/low ECM risk score group groups in TCGA-OV. Comparison with several algorithms for T cell
(A) infiltration, Neutrophil **(B),** cancer-associated fibroblast (CAFs) **(C),** Immune microenvironment score **(D)** infiltration, expression levels of 8 immune
checkpoint genes in high/low ECM risk score group in TCGA-OV **(E),** and IPS score in high/low ECM risk score group in TCGA-OV **(F).**

Validation of the prognostic function of ECM risk score in the TCGA-pan-cancer dataset

TCGA pan-cancer data with survival information were used to validate the ECM risk score. A total of 9162 "Primary Solid Tumour" data with both gene expression data and survival data were included in the analysis. We screened all adenocarcinoma expression data and survival data as a validation dataset. 2084 samples meeting the criteria were extracted, of which 1580 samples carried information on initial treatment outcome. We found that the low-risk group had a higher initial treatment Complete Response (CR) and Partial Response (PR), there was no significant difference in the number of patients with Progressive Disease (PD) between the low- and

high-risk groups, and the number of patients with Stable Disease (SD) was significantly higher in the high-risk group than in the low-risk group (Figure 4A). We also analysed the expression levels of eight immune checkpoint genes in the high/low ECM risk score group (Figure 4B). Only four immune checkpoint genes showed significant differences between the two groups. Fifteen ECM' key genes were extracted from all gene expression data, and the ECM risk score was calculated for each patient. The expression levels of 15 ECM key genes were differentially expressed in both the high and low-risk groups (Figure 4C). Computes the predicted survivor function for a Cox proportional hazards model and plots the KM curve (Figure 4D). The ECM risk score was found to be a good predictive tool for overall survival prognosis in the adenocarcinoma
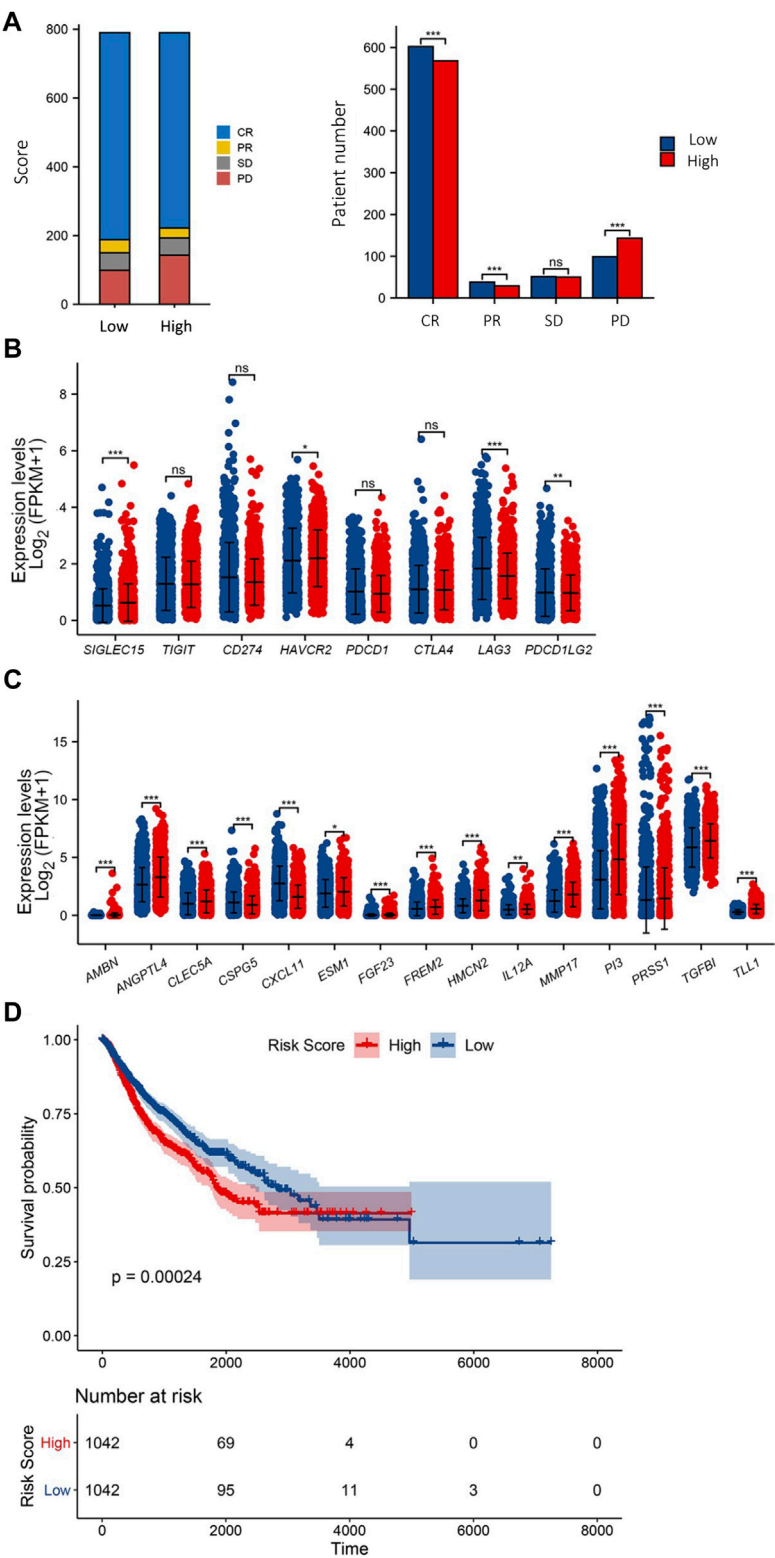
FIGURE 4
Validation of ECM risk key genes in TCGA Adenocarcinoma. First-course treatment outcome of high/low ECM risk score group in TCGA-Adenocarcinoma **(A)**. Expression levels of 8 immune checkpoint genes **(B)** and 15 ECM key genes **(C)** in high/low ECM risk score group in TCGA-Adenocarcinoma. KM survival plot for high/low ECM risk score group in TCGA-Adenocarcinoma **(D)**.
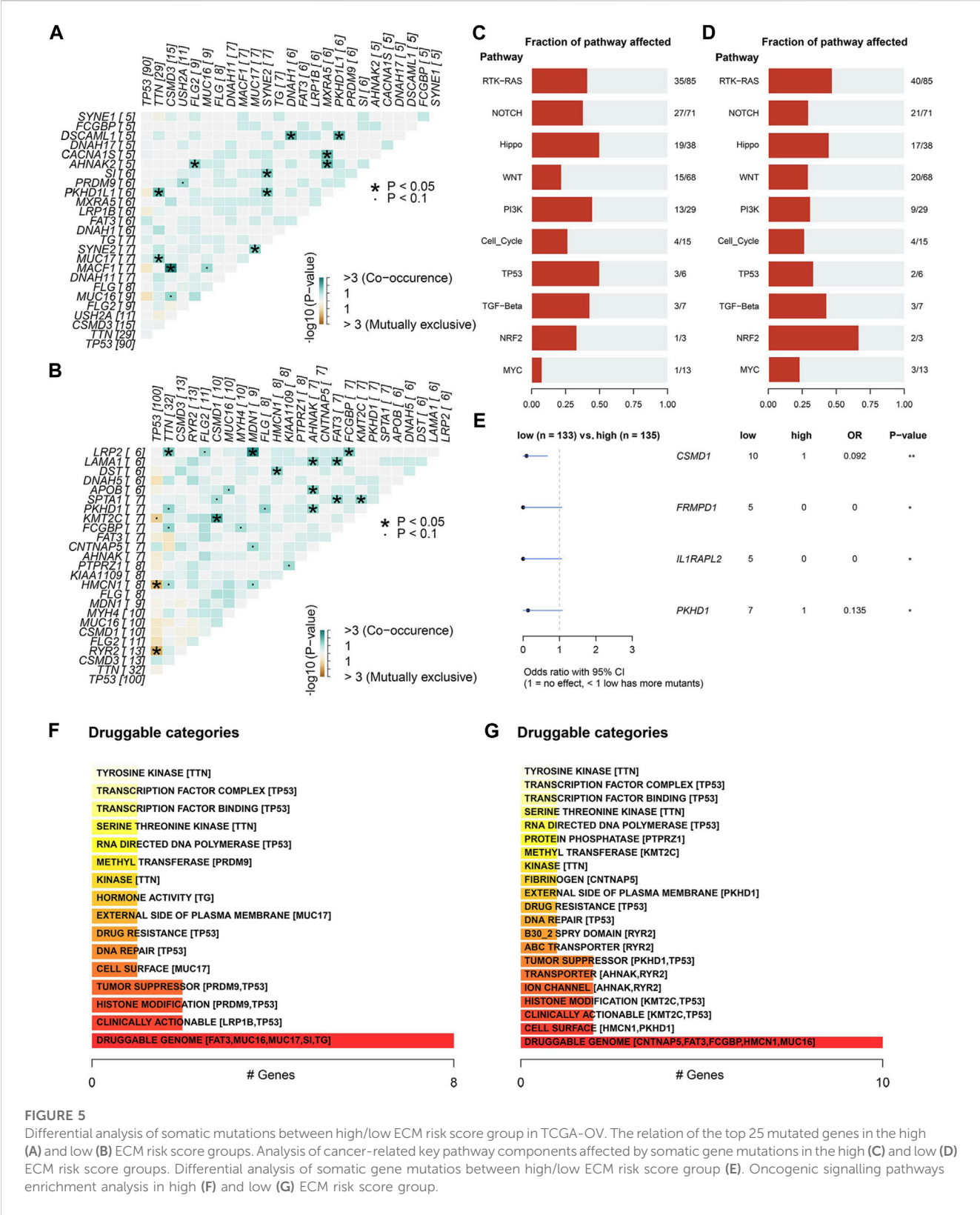
FIGURE 5
Differential analysis of somatic mutations between high/low ECM risk score group in TCGA-OV. The relation of the top 25 mutated genes in the high
(A) and low **(B)** ECM risk score groups. Analysis of cancer-related key pathway components affected by somatic gene mutations in the high **(C)** and low **(D)**
ECM risk score groups. Differential analysis of somatic gene mutatios between high/low ECM risk score group **(E)**. Oncogenic signalling pathways
enrichment analysis in high **(F)** and low **(G)** ECM risk score group.

data. However, it was not adequate in effectively predicting for OS between the high and low-risk groups after 4000 days.

Relationship between ECM risk score and tumour mutation burden (TMB)

We further investigated the connection between the ECM risk score and TMB because OC is marked by a high degree of somatic mutation. Missense Mutation was the main component of somatic mutation in the high/low ECM risk score group. In the single nucleotide variant (SNV) analysis showed that the highest rate of C > T was observed in both groups. Variants per sample of the high ECM risk score group were lower than in the low ECM risk score group (Supplementary Figures S3A, B). In both high/low ECM risk score groups, *CSMD3, TTN, TP53, FLG2, MUC16, FLG* and *FAT3* were found in the top 20 mutated genes (Supplementary Figures S3C, D). In comparison to the high-risk group, the mutation rates in *TP53, TTN* and *RYR2* were higher in the low-risk group, while those of *CSMD3, USH2A* and *FLG2* were nearly identical (Supplementary Figures S3E, F).

In the mutation Exclusive/Co-occurring analysis of the top 20 mutation genes, we found 11 pairs of genes with Co-occurring relationships in each of the high/low ECM risk score groups but 2 pairs of Exclusive in the low ECM risk score group (Figures 5A, B). We also examined the mutation status of the elements of eight signalling pathways that have been demonstrated to be crucial in the development of tumours (Sanchez-Vega et al., 2018). We found that the number of genes affected by somatic mutations in most signalling pathways was approximately the same between the two groups, except for *RTK-RAS, NOTCH, WNT,* and *PIK2* (Figures 5C, D). The expression levels of all somatic mutations were analysed between the high/low ECM risk score groups, and we found statistically significant differences in *CSMD1, FRMPD1, IL1PARL2* and *PKHD1*. All four genes showed a higher mutation rate in the low ECM risk score group (Figures 5D, E). Based on the differences in somatic mutations between the two groups, we enriched the analysis for drug-gene interactions, and the "Druggable Genome: was found to be highly enriched within both groups (Figure 5F, G).

Additionally, we documented that the low ECM risk score group was more responsive to RYR2 gene-related immunotherapy, whereas the high ECMs risk score group may be more responsive to TG-related immunotherapeutics (Supplementary Table S2). We also analysed the mutations in 15 key genes (Supplementary Figure S4).

Discussion

Ovarian cancer is a malignant tumour with a high mortality rate. Because of its insidious onset, it is usually late-stage when obvious clinical symptoms appear (Prat J and FIGO Committee on Gynecologic Oncology, 2014). Patient prognosis is thus based on an accurate and reliable assessment. As the extracellular matrix is closely linked to epithelial cells, we aimed to construct a prognostic prediction model based on extracellular matrix proteins to evaluate the prognostic survival of patients with ovarian serous adenocarcinomas. We successfully filtered 15 key genes out of 1068 extracellular matrix-associated proteins, with excellent predictive ability on ovarian serous adenocarcinoma prognosis using the Random Forest and the Lasso algorithms.

We used the random forest algorithm as is a popular machine learning technique with documented outstanding performance in a wide range of predictive modeling tasks, including cancer prognosis prediction (Toth et al., 2019; Li Y et al., 2020). When analyzing tumor sequencing data, the relationship between variables and outcomes can be complex due to the volume of data. Random forest can handle a large number of input variables without overfitting, which has been a challenge for traditional models such as logistic regression or decision trees (Maroco et al., 2011; Lan et al., 2020). Additionally, traditional linear models are unable to handle the nonlinear relationship between a large amount of sequencing data and survival outcomes, but random forest can address this issue (Lee and Lim, 2019). Compared to support vector machines (SVM), random forest is less sensitive to outliers and missing data, which is important in cancer prognosis prediction, where data quality may vary and missing data is common (Pelckmans et al., 2005; Lee and Lim, 2019).

Among the 15 key genes generated by our our analysis, *CSPG5, CXCL11* and *ESM1* mRNAs were abundantly expressed in OC tissue compared to normal tissue. Cancer cells, fibroblasts, endothelial cells, and immune cells such as leukocytes, monocytes, and dendritic cells are primarily responsible for *CXCL11* production (Gao and Zhang, 2021) *CXCL11* is an effector chemokine regulating T cell recruitment that promotes effector immune cells (e.g., CD8T cells, Th1 cells, TH17 cells and antigen-presenting cells). Studies have shown that induction of *CXCL10* and *CXCL11* expression in breast cancer cells enhances the infiltration of CD8 T cells (Liu et al., 2011; Gao et al., 2019). In addition to its influence on tumour progression through its angiostatic effects (Strieter et al., 2006), CXCL11 is part of the CXCR7/CXCL11 axis that was shown to induce the epithelial–mesenchymal transition and metastatic behaviour of OC cells under ERα control (Benhadjeba et al., 2018). *ESM1* is a soluble proteoglycan expressed by vascular endothelial cells and is associated with inflammatory cell recruitment (Hung et al., 2020). Vascular endothelial dysfunction can be brought on by high levels of ESM1 (Kalantaridou et al., 2006; Rocha et al., 2014; Sun et al., 2019), whereas it has also been shown to be closely correlated with OC development and progression (Li et al., 2023). Our analysis showed that the tumour tissue had significantly higher *ESM1* expression levels than normal ovarian tissue, which may be associated with abnormally elevated cell proliferation and tumour tissue revascularization. Furthermore, CSPG5, also known as Neuroglycan C (NGC), is a protein originally associated with extracellular matrix production in the nervous system (Pintér et al., 2020) and shown to decrease first and then increase following ischemic and hypoxic injury, presumably associated with ECM damage repair (Matsui et al., 2005). Of note, two recent studies showed that expression of CSPG5 was significantly correlated with the prognosis of patients with epithelial OC (Su et al., 2021; Wang et al., 2023). In our study, higher *CSPG5, CXCL11* and *ESM1* expression and inflammatory cell infiltration, especially of CD8 T cells, were present in the low-risk group. The high level of immune cell environment may explain the better prognosis for overall survival in the low-risk group.

Interestingly, two proteins commonly associated with bone biology and development were identified among the ECM prognostic markers.

AMBN was discovered as a tooth enamel matrix protein, playing an important role in enamel, cranial and long bone development. It was however demonstrated that AMBN was among four factors that were potential independent prognostic factors for prostate cancer (Xu et al., 2018). In our multifactorial Cox analysis, AMBN was shown to be one of the independent prognostic factors for ovarian serous adenocarcinomas. Furthermore, the analysis revealed a prognostic role for *FGF23*, the bone-derived hormone secreted by osteoblasts and osteocytes (Guo and Yuan, 2015). Previous studies have shown its expression alterations in breast cancer (Cekin et al., 2020) and identified that serum or plasma FGF23 concentrations are elevated in patients with advanced stage epithelial ovarian cancer (Tebben et al., 2005). According to our data, FGF23 was expressed in the tumours, and its expression levels were higher in the high-risk group.

MMP17, PI3, TLL1, ANGPTL4, and *TGFBI* have all been previously associated with cancer. Our analysis found that transcripts of all these five genes were expressed at higher levels in the high-risk group than in the low-risk group. *MMP17* has been associated with the maintenance of normal physiological function in vascular smooth muscle (Martín-Alonso et al., 2015) and a promotive effect on tumour cells (Paye et al., 2014). Additionally, it was shown that its expression was much higher in EOC patients than in pericarcinomatous tissues (Xiao et al., 2022). *PI3*, also known as elafin, is an elastase-specific inhibitor that directly affects tumour suppression by inhibiting elastase (Hunt et al., 2013). High levels of *PI3* are associated with severe disease severity in various cancers (Hunt et al., 2013; Longatto-Filho et al., 2021), while another TCGA-OV analysis showed its prognostic value in OC (Li J et al., 2020). Additionally, high elafin expression has been associated with unfavorable OS but better immunotherapy responses (Lu et al., 2023). Recent studies have found that *TGFBI* CpG islands are hypermethylated in adjacent normal colon tissue, with the corresponding sequences showing hypomethylation in colon cancer tissue, and that higher *TGFBI* levels are associated with poorer prognosis (Zhang H. et al., 2019). In mammals, Tolloid-like (mTLL)-1 is a BMP-1-associated protease, and *BMP1/TLL1* is involved in the process of tissue remodelling in the ovary, assisting in the maturation of pre-collagen molecules and the deposition of collagen fibres (Ohnishi et al., 2005). In hepatocellular carcinoma (Matsuura et al., 2017), S*NP* may impact the splicing of *TLL1* mRNA and result in short variants with high catalytic activity, speeding up the development of liver fibrosis and cancer. In a recent study, *TGFBI, PI3, TLL1* and *MMP17* were predicted to be among the 19 genes that comprise the TME-related high grade serous ovarian carcinoma prognostic genetic panel (Belotti et al., 2022). *ANGPTL4* is regulated by peroxisome proliferator-activated receptor γ (PPARγ) (Aryal et al., 2019), who has been observed to be significantly increased in malignant ovarian tumours (grade 1, 2 and 3) compared to benign and borderline tumours (Zhang et al., 2005). It was also recently identified in a scRNA-seq study of ovarian cancer CAF ligands to epithelial cells (Carvalho et al., 2022). Finally, *ANGPTL4* and *TGFBI* were identified both in a hypoxia risk model constructed to reflect the OC immune microenvironment in and predict prognosis (Wei et al., 2021), and among the genes that comprise an OC glycolysis-related gene signature (Zhang et al., 2021).

To further assess the relationship between the ECM risk score and the immune microenvironment, we assessed the abundance of multiple infiltrating cells in the immune microenvironment of these patients using multiple algorithms. Patients with more infiltrating and activated immune cells in TME may have better immunotherapeutic

outcomes (Li et al., 2016). We observed higher levels of CD8 T cell infiltration in the low-risk group, suggesting enhanced immune surveillance via CD8[+] T cells, while this implies a possible enhanced susceptibility to PD-1/PD-L1-targeted immune checkpoint therapies (Strickland et al., 2016; Iyer et al., 2021). Furthermore, the degree of macrophage infiltration was similar to both CD8 and CD4 T cell infiltration, with both showing high levels in the low-risk group. This is consistent with previous studies describing a positive correlation between T cells and macrophage infiltration levels (Desbois et al., 2020).

Previous studies have also shown that cancer cells regulate their local microenvironment to promote tumour survival, chemoresistance and evasion of immune surveillance (Kim et al., 2007) and that there is a tight association between malignant tumour cells and CAFs in promoting tumour growth and survival (Xing et al., 2010; Karagiannis et al., 2012; Chen et al., 2021). scRNA-seq analysis of high grade serous ovarian cancers, also showed that CAFs induce epithelial-mesenchymal transition (EMT) of tumor cells via TGFβ signaling, with consequent effects on chemoresistance and metastasis (Kan et al., 2020). In accordance with these, our results also showed that higher levels of CAFs were found in high-risk groups with poor prognosis.

It is known that the immune microenvironment immune cells, immunomodulating factors and immune checkpoint molecules are crucial for the immune escape of tumour cells (Charoentong et al., 2017; Zhang Y. et al., 2019). We thus developed an immunophenoscore (IPS) based on immune subpopulation infiltration and expression of immune regulatory molecules using the random forest to identify determinants of immunogenicity. Among several IPS subtypes tested, we found that the low-risk group had higher IPS and could benefit during treatment with immune checkpoint inhibitors. As the immunophenoscore is a surrogate to patients' immunotherapeutic outcomes, our IPS results of the ECMs risk scores may only be considered of predictive value, and future studies will confirm their clinical importance.

We also explored the tumour mutation burden (TMB) changes in the TCGA-OV cohort. The ovarian cancer genome exhibits high levels of instability, as evidenced by functional cells (Stewart et al., 2011), copy number changes (Schwarz et al., 2015), and status of somatic mutations (Bashashati et al., 2013). TMB is the total number of substitutions and insertion/deletion mutations per megabase in the exon-coding region of the gene under evaluation in the tumour cell genome (Stenzinger et al., 2019). Somatic mutations may result in tumourigenesis and many somatic mutations can generate neoantigens facilitating anti-tumour immunity (Gubin et al., 2015). In a study on immunotherapy for lung cancer, researchers discovered that patients with PD-L1 1% but a subgroup of 10 mutations/Mb in the combination chemotherapy group had a better objective response rate (ORR) and median progression-free survival (median PFS, mPFS) with the immune combination regimen CheckMate 227 (Hellmann et al., 2018). This suggests that in the higher TMB population, PFS was better in the combination immunotherapy group than in the chemotherapy alone group, irrespective of PD-L1 expression. In our study, the mean TMB values were higher in the low-risk group than in the high-risk group, implying that the low-risk group may have more potential for immunotherapy. *TP53*, the gene encoding the tumour suppressor protein p53, is one of the most commonly mutated genes in human cancers, and driver mutations are prevalent in high-grade ovarian plasmacytoma (Ahmed et al., 2010). Chalmers and coworkers have shown that *TP53* mutations were associated with high TMB (Chalmers et al., 2017). Our study similarly confirmed that in the TCGA-OV datasets *TP53* mutations were the most frequent in the high-

and low-risk groups, and that *TP53* mutations were higher in the low-risk group than in the high-risk group. However, there are limitations to cohort-based studies. Most mutated genes are unique to each case, and in clinical treatment, patients should be treated based on their mutation sequencing results. Our analysis may provide theoretical support for the selection of immunological agents.

Our study aimed to investigate the role of the matrisome and the gene changes in the ECM in a widely studied, publicly available ovarian cancer transcriptomic and clinicopathological collection of patients. The role of matrisome has been regretfully understudied in this type or cancer that commonly goes undetected till it reaches high grades, as the ECM can influence drug resistance. We used bioinformatics and machine learning methods to investigate the TCGA-OV collection and identified several prognostic genes, some of which have also been identified by previous studies. Given the current cost-effectiveness of biotechnological approaches, rapid genetic testing tools are commonly promoted and widely applied in clinical diagnostics and treatment (Young and Argáez, 2019). These tools have improved accuracy and testing times have significantly shortened. Targeted multigene tests and genetic screening can be thus rapidly employed to assist diagnosis postoperatively and develop more effective treatment plans.

Conclusion

In conclusion, this study developed an ECM risk score prediction model to enable prognosis of patients with ovarian serous adenocarcinoma. We further identified the tumour microenvironment and somatic mutations using the TCGA-OV collection datasets. These results should be further validated with targeted future studies to evaluate their real predictive function and their use in personalized immunotherapy applications.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://xenabrowser.net/datapages/?cohort=GDC%20TCGA%20Ovarian%20Cancer%20(OV)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443.

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

Conceptualization, TG, MZ and JER, methodology and data visualization TG and YW, software and data curation TG, formal analysis TG, MZ, and AS, draft preparation TG, MZ and AS, review, editing, supervision, project administration, and funding acquisition AS

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

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1200354/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Enrichment analysis for different expression genes **(A)** GOKEGG results of TCGA-OV. Top 5 Head **(B)** and Tail **(C)** aggregation of GSEA results in TCGA-OV DEGs.

SUPPLEMENTARY FIGURE S2
Comparison of immune cell infiltration between high/low ECM risk score group groups in TCGA-OV.

SUPPLEMENTARY FIGURE S3
Summary of somatic mutation analysis in high/low ECM risk score group in TCGAOV Summary of mutations in the high **(A)** and low **(B)** ECM risk score groups. Mutation profile of TCGAOV patients in high **(C)** and low **(D)** ECM risk score group and percentage of top 6 mutated genes in high/low ECM risk score group **(E)**. Comparison of mutation load among high/low ECM risk score group and TCGA cohorts **(F)**.

SUPPLEMENTARY FIGURE S4
Details of ECM risk score in the 15 key gene mutations.

SUPPLEMENTARY TABLE S1
ECM-related gene list.

SUPPLEMENTARY TABLE S2
Drug-gene interactions.

References

Ahmed, A. A., Etemadmoghadam, D., Temple, J., Lynch, A. G., Riad, M., Sharma, R., et al. (2010). Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J. Pathol.* 221, 49–56. doi:10.1002/path.2696

Aryal, B., Price, N. L., Suarez, Y., and Fernández-Hernando, C. (2019). ANGPTL4 in metabolic and cardiovascular disease. *Trends Mol. Med.* 25, 723–734. doi:10.1016/j.molmed.2019.05.010

Bashashati, A., Ha, G., Tone, A., Ding, J., Prentice, L. M., Roth, A., et al. (2013). Distinct evolutionary trajectories of primary high-grade serous ovarian cancers revealed through spatial mutational profiling. *J. Pathol.* 231, 21–34. doi:10.1002/path.4230

Belotti, Y., Lim, E. H., and Lim, C. T. (2022). The role of the extracellular matrix and tumor-infiltrating immune cells in the prognostication of high-grade serous ovarian cancer. *Cancers* 14, 404. doi:10.3390/cancers14020404

Benhadjeba, S., Edjekouane, L., Sauvé, K., Carmona, E., and Tremblay, A. (2018). Feedback control of the CXCR7/CXCL11 chemokine axis by estrogen receptor α in ovarian cancer. *Mol. Oncol.* 12 (10), 1689–1705. Epub 2018 Aug 23. doi:10.1002/1878-0261.12362

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, 782–795. doi:10.1016/j.immuni.2013.10.003

Carvalho, R. F., do Canto, L. M., Abildgaard, C., Aagaard, M. M., Tronhjem, M. S., Waldstrøm, M., et al. (2022). Single-cell and bulk RNA sequencing reveal ligands and receptors associated with worse overall survival in serous ovarian cancer. *Cell Commun. Signal* 20 (1), 176. doi:10.1186/s12964-022-00991-4

Cekin, R., Arici, S., Atci, M. M., Secmeler, S., and Cihan, S. (2020). The clinical i? mportance of fibroblast growth factor 23 on breast cancer patients. *EJMI* 4 (4), 471–476.

Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., et al. (2017). Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* 9, 34. doi:10.1186/s13073-017-0424-2

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. *Cell Rep.* 18, 248–262. doi:10.1016/j.celrep.2016.12.019

Chen, Y., McAndrews, K. M., and Kalluri, R. (2021). Clinical and therapeutic relevance of cancer-associated fibroblasts. *Nat. Rev. Clin. Oncol.* 18 (12), 792–804. doi:10.1038/s41571-021-00546-5

Custódio, N., Savisaar, R., Carvalho, C., Bak-Gordon, P., Ribeiro, M. I., Tavares, J., et al. (2022). Expression profiling in ovarian cancer reveals coordinated regulation of *BRCA1/2* and homologous recombination genes. *Biomedicines* 10 (2), 199. doi:10.3390/biomedicines10020199

Desbois, M., Udyavar, A. R., Ryner, L., Kozlowski, C., Guan, Y., Dürrbaum, M., et al. (2020). Integrated digital pathology and transcriptome analysis identifies molecular mediators of T-cell exclusion in ovarian cancer. *Nat. Commun.* 11, 5583. doi:10.1038/s41467-020-19408-2

Donelan, W., Dominguez-Gutierrez, P. R., and Kusmartsev, S. (2022). Deregulated hyaluronan metabolism in the tumor microenvironment drives cancer inflammation and tumor-associated immune suppression. *Front. Immunol.* 13, 971278. doi:10.3389/fimmu.2022.971278

Frantz, C., Stewart, K. M., and Weaver, V. M. (2010). The extracellular matrix at a glance. *J. Cell Sci.* 123, 4195–4200. doi:10.1242/jcs.023820

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

Gao, Q., Wang, S., Chen, X., Cheng, S., Zhang, Z., Li, F., et al. (2019). Cancer-cell-secreted CXCL11 promoted CD8+ T cells infiltration through docetaxel-induced-release of HMGB1 in NSCLC. *J. Immunother. Cancer* 7, 42. doi:10.1186/s40425-019-0511-6

Gao, Q., and Zhang, Y. (2021). CXCL11 signaling in the tumor microenvironment. *Adv. Exp. Med. Biol.* 1302, 41–50. doi:10.1007/978-3-030-62658-7_4

Gibney, G. T., Weiner, L. M., and Atkins, M. B. (2016). Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol.* 17, e542–e551. doi:10.1016/S1470-2045(16)30406-5

Griffith, M., Griffith, O. L., Coffman, A. C., Weible, J. V., McMichael, J. F., Spies, N. C., et al. (2013). DGIdb: Mining the druggable genome. *Nat. Methods* 10, 1209–1210. doi:10.1038/nmeth.2689

Gubin, M. M., Artyomov, M. N., Mardis, E. R., and Schreiber, R. D. (2015). Tumor neoantigens: Building a framework for personalized cancer immunotherapy. *J. Clin. Invest.* 125, 3413–3421. doi:10.1172/JCI80008

Guo, Y.-C., and Yuan, Q. (2015). Fibroblast growth factor 23 and bone mineralisation. *Int. J. Oral Sci.* 7, 8–13. doi:10.1038/ijos.2015.1

Heintz, A., Odicino, F., Maisonneuve, P., Quinn, M. A., Benedet, J. L., Creasman, W. T., et al. (2006). Carcinoma of the ovary. *Int. J. Gynaecol. Obstet. Off. Organ Int. Fed. Gynaecol. Obstet.* 95 (1), S161–S192. doi:10.1016/S0020-7292(06)60033-7

Hellmann, M. D., Ciuleanu, T.-E., Pluzanski, A., Lee, J. S., Otterson, G. A., Audigier-Valette, C., et al. (2018). Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N. Engl. J. Med.* 378, 2093–2104. doi:10.1056/NEJMoa1801946

Henderson, M. P. A., Hirte, H., Hotte, S. J., and Kavsak, P. A. (2016). Cytokines and cell adhesion molecules exhibit distinct profiles in health, ovarian cancer, and breast cancer. *Heliyon* 2, e00059. doi:10.1016/j.heliyon.2015.e00059

Henke, E., Nandigama, R., and Ergün, S. (2020). Extracellular matrix in the tumor microenvironment and its impact on cancer therapy. *Front. Mol. Biosci.* 6, 160. doi:10.3389/fmolb.2019.00160

Hung, T.-W., Chu, C.-Y., Yu, C.-L., Lee, C.-C., Hsu, L.-S., Chen, Y.-S., et al. (2020). Endothelial cell-specific molecule 1 promotes endothelial to mesenchymal transition in renal fibrosis. *Toxins* 12, 506. doi:10.3390/toxins12080506

Hunt, K. K., Wingate, H., Yokota, T., Liu, Y., Mills, G. B., Zhang, F., et al. (2013). Elafin, an inhibitor of elastase, is a prognostic indicator in breast cancer. *Breast Cancer Res. BCR* 15, R3. doi:10.1186/bcr3374

Ishwaran, H., Kogalur, U. B., and Kogalur, M. U. B. (2022). Package 'randomForestSRC. ' *breast* 6, 1.

Iyer, S., Zhang, S., Yucel, S., Horn, H., Smith, S. G., Reinhardt, F., et al. (2021). Genetically defined syngeneic mouse models of ovarian cancer as tools for the discovery of combination immunotherapy. *Cancer Discov.* 11, 384–407. doi:10.1158/2159-8290.CD-20-0818

Kalantaridou, S. N., Naka, K. K., Bechlioulis, A., Makrigiannakis, A., Michalis, L., and Chrousos, G. P. (2006). Premature ovarian failure, endothelial dysfunction and estrogen-progestogen replacement. *Trends Endocrinol. Metab. Tem.* 17, 101–109. doi:10.1016/j.tem.2006.02.003

Kan, T., Wang, W., Ip, P. P., Zhou, S., Wong, A. S., Wang, X., et al. (2020). Single-cell EMT-related transcriptional analysis revealed intra-cluster heterogeneity of tumor cell clusters in epithelial ovarian cancer ascites. *Oncogene* 39 (21), 4227–4240. doi:10.1038/s41388-020-1288-2

Karagiannis, G. S., Poutahidis, T., Erdman, S. E., Kirsch, R., Riddell, R. H., and Diamandis, E. P. (2012). Cancer-associated fibroblasts drive the progression of metastasis through both paracrine and mechanical pressure on cancer tissue. *Mol. Cancer Res.* 10 (11), 1403–1418. doi:10.1158/1541-7786.MCR-12-0307

Kim, R., Emi, M., and Tanabe, K. (2007). Cancer immunoediting from immune surveillance to immune escape. *Immunology* 121, 1–14. doi:10.1111/j.1365-2567.2007.02587.x

Kim, S.-H., Park, Y.-Y., Kim, S.-W., Lee, J.-S., Wang, D., and DuBois, R. N. (2011). ANGPTL4 induction by prostaglandin E2 under hypoxic conditions promotes colorectal cancer progression. *Cancer Res.* 71, 7010–7020. doi:10.1158/0008-5472.CAN-11-1262

Lan, T., Hu, H., Jiang, C., Yang, G., and Zhao, Z. (2020). A comparative study of decision tree, random forest, and convolutional neural network for spread-F identification. *Adv.Space Res.* 65, 2052–2061. doi:10.1016/j.asr.2020.01.036

Lee, S., and Lim, H. (2019). Review of statistical methods for survival analysis using genomic data. *Genomics Inf.* 17, e41. doi:10.5808/GI.2019.17.4.e41

Lei, X., Lei, Y., Li, J.-K., Du, W.-X., Li, R.-G., Yang, J., et al. (2020). Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett.* 470, 126–133. doi:10.1016/j.canlet.2019.11.009

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

Li J, J., Tian, Y., Zhu, Y., Zhou, T., Li, J, Ding, K., et al. (2020). A multicenter random forest model for effective prognosis prediction in collaborative clinical research network. *Artif. Intell. Med.* 103, 101814. doi:10.1016/j.artmed.2020.101814

Li Y, Y., Li, H., Yang, B., Wei, J., Zhen, C., and Feng, L. (2020). Clinical significance of PI3 and HLA-DOB as potential prognostic predicators for ovarian cancer. *Transl. Cancer Res.* 9 (2), 466–476. doi:10.21037/tcr.2019.11.30

Li, Y. K., Zeng, T., Guan, Y., Liu, J., Liao, N. C., Wang, M. J., et al. (2023). Validation of ESM1 related to ovarian cancer and the biological function and prognostic significance. *Int. J. Biol. Sci.* 19 (1), 258–280. doi:10.7150/ijbs.66839

Lin, J., Yin, M., Liu, L., Gao, J., Yu, C., Liu, X., et al. (2022). The development of a prediction model based on random survival forest for the postoperative prognosis of pancreatic cancer: A SEER-based study. *Cancers* 14, 4667. doi:10.3390/cancers14194667

Liu, M., Guo, S., and Stiles, J. K. (2011). The emerging role of CXCL10 in cancer (Review). *Oncol. Lett.* 2, 583–589. doi:10.3892/ol.2011.300

Longatto-Filho, A., Fregnani, J. H., Mafra da Costa, A., de Araujo-Souza, P. S., Scapulatempo-Neto, C., Herbster, S., et al. (2021). Evaluation of elafin immunohistochemical expression as marker of cervical cancer severity. *Acta Cytol.* 65, 165–174. doi:10.1159/000512010

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi:10.1186/s13059-014-0550-8

Lu, W., Xie, B., Tan, G., Dai, W., Ren, J., Pervaz, S., et al. (2023). Elafin is related to immune infiltration and could predict the poor prognosis in ovarian cancer. *Front. Endocrinol. (Lausanne)* 14, 1088944. doi:10.3389/fendo.2023.1088944

Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., and de Mendonça, A. (2011). Data mining methods in the prediction of dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests. *BMC Res. Notes* 4, 299. doi:10.1186/1756-0500-4-299

Martín-Alonso, M., García-Redondo, A. B., Guo, D., Camafeita, E., Martínez, F., Alfranca, A., et al. (2015). Deficiency of MMP17/MT4-MMP proteolytic activity predisposes to aortic aneurysm in mice. *Circ. Res.* 117, e13–e26. doi:10.1161/CIRCRESAHA.117.305108

Matsui, F., Kakizawa, H., Nishizuka, M., Hirano, K., Shuo, T., Ida, M., et al. (2005). Changes in the amounts of chondroitin sulfate proteoglycans in rat brain after neonatal hypoxia-ischemia. *J. Neurosci. Res.* 81, 837–845. doi:10.1002/jnr.20603

Matsuura, K., Sawai, H., Ikeo, K., Ogawa, S., Iio, E., Isogawa, M., et al. (2017). Genome-wide association study identifies TLL1 variant associated with development of hepatocellular carcinoma after eradication of hepatitis C virus infection. *Gastroenterology* 152, 1383–1394. doi:10.1053/j.gastro.2017.01.041

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

Miao, Y., Liu, J., Liu, X., Yuan, Q., Li, H., Zhang, Y., et al. (2022). Machine learning identification of cuproptosis and necroptosis-associated molecular subtypes to aid in prognosis assessment and immunotherapy response prediction in low-grade glioma. *Front. Genet.* 13, 951239. doi:10.3389/fgene.2022.951239

Naba, A., Clauser, K. R., Ding, H., Whittaker, C. A., Carr, S. A., and Hynes, R. O. (2016). The extracellular matrix: Tools and insights for the "omics" era. *Matrix Biol. J. Int. Soc. Matrix Biol.* 49, 10–24. doi:10.1016/j.matbio.2015.06.003

Ohnishi, J., Ohnishi, E., Shibuya, H., and Takahashi, T. (2005). Functions for proteinases in the ovulatory process. *Biochim. Biophys. Acta* 1751, 95–109. doi:10.1016/j.bbapap.2005.05.002

Paye, A., Truong, A., Yip, C., Cimino, J., Blacher, S., Munaut, C., et al. (2014). EGFR activation and signaling in cancer cells are enhanced by the membrane-bound metalloprotease MT4-MMP. *Cancer Res.* 74, 6758–6770. doi:10.1158/0008-5472.CAN-13-2994

Pelckmans, K., De Brabanter, J., Suykens, J. A. K., and De Moor, B. (2005). Handling missing values in support vector machine classifiers. *Neural Netw.* 18, 684–692. doi:10.1016/j.neunet.2005.06.025

Pintér, A., Hevesi, Z., Zahola, P., Alpár, A., and Hanics, J. (2020). Chondroitin sulfate proteoglycan-5 forms perisynaptic matrix assemblies in the adult rat cortex. *Cell. Signal.* 74, 109710. doi:10.1016/j.cellsig.2020.109710

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. Off. J. Eur. Soc. Med. Oncol.* 27, 1482–1492. doi:10.1093/annonc/mdw168

Prat, J.FIGO Committee on Gynecologic Oncology (2014). Staging classification for cancer of the ovary, fallopian tube, and peritoneum. *Int. J. Gynaecol. Obstet. Off. Organ Int. Fed. Gynaecol. Obstet.* 124, 1–5. doi:10.1016/j.ijgo.2013.10.001

Rocha, S. F., Schiller, M., Jing, D., Li, H., Butz, S., Vestweber, D., et al. (2014). Esm1 modulates endothelial tip cell behavior and vascular permeability by enhancing VEGF bioavailability. *Circ. Res.* 115, 581–590. doi:10.1161/CIRCRESAHA.115.304718

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., et al. (2018). Oncogenic signaling pathways in the cancer genome Atlas. *Cell* 173, 321–337.e10. e10. doi:10.1016/j.cell.2018.03.035

Schreiber, R. D., Old, L. J., and Smyth, M. J. (2011). Cancer immunoediting: Integrating immunity's roles in cancer suppression and promotion. *Science* 331, 1565–1570. doi:10.1126/science.1203486

Schwarz, R. F., Ng, C. K. Y., Cooke, S. L., Newman, S., Temple, J., Piskorz, A. M., et al. (2015). Spatial and temporal heterogeneity in high-grade serous ovarian cancer: A phylogenetic analysis. *PLoS Med.* 12, e1001789. doi:10.1371/journal.pmed.1001789

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

Stewart, J. M., Shaw, P. A., Gedye, C., Bernardini, M. Q., Neel, B. G., and Ailles, L. E. (2011). Phenotypic heterogeneity and instability of human ovarian tumor-initiating cells. *Proc. Natl. Acad. Sci. U. S. A.* 108, 6468–6473. doi:10.1073/pnas.1005529108

Strickland, K. C., Howitt, B. E., Shukla, S. A., Rodig, S., Ritterhouse, L. L., Liu, J. F., et al. (2016). Association and prognostic significance of BRCA1/2-mutation status with neoantigen load, number of tumor-infiltrating lymphocytes and expression of PD-1/PD-L1 in high grade serous ovarian cancer. *Oncotarget* 7, 13587–13598. doi:10.18632/oncotarget.7277

Strieter, R. M., Burdick, M. D., Mestas, J., Gomperts, B., Keane, M. P., and Belperio, J. A. (2006). Cancer CXC chemokine networks and tumour angiogenesis. *Eur. J. Cancer* 42, 768–778. doi:10.1016/j.ejca.2006.01.006

Su, T., Zhang, P., Zhao, F., and Zhang, S. (2021). A novel immune-related prognostic signature in epithelial ovarian carcinoma. *Aging (Albany NY)* 13 (7), 10289–10311. doi:10.18632/aging.202792

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, H., Fang, F., Li, K., Zhang, H., Zhang, M., Zhang, L., et al. (2019). Circulating ESM-1 levels are correlated with the presence of coronary artery disease in patients with obstructive sleep apnea. *Respir. Res.* 20, 188. doi:10.1186/s12931-019-1143-6

Tamborero, D., Rubio-Perez, C., Deu-Pons, J., Schroeder, M. P., Vivancos, A., Rovira, A., et al. (2018). Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* 10, 25. doi:10.1186/s13073-018-0531-8

Tebben, P. J., Kalli, K. R., Cliby, W. A., Hartmann, L. C., Grande, J. P., Singh, R. J., et al. (2005). Elevated fibroblast growth factor 23 in women with malignant ovarian tumors. *Mayo Clin. Proc.* 80 (6), 745–751. doi:10.1016/S0025-6196(11)61528-0

The Cancer Genome Atlas Research Network (2011). Integrated genomic analyses of ovarian carcinoma. *Nature* 474, 609–615. doi:10.1038/nature10166

Therneau, T. (2015). *A package for survival analysis in S. R Package Version 2.*

Toth, R., Schiffmann, H., Hube-Magg, C., Büscheck, F., Höflmayer, D., Weidemann, S., et al. (2019). Random forest-based modelling to detect biomarkers for prostate cancer progression. *Clin. Epigenetics* 11, 148. doi:10.1186/s13148-019-0736-8

Valmiki, S., Aid, M. A., Chaitou, A. R., Zahid, M., Valmiki, M., Fawzy, P., et al. (2021). Extracellular matrix: A treasure trove in ovarian cancer dissemination and chemotherapeutic resistance. *Cureus* 13, e13864. doi:10.7759/cureus.13864

Wang, H., Li, S., Wang, Q., Jin, Z., Shao, W., Gao, Y., et al. (2021). Tumor immunological phenotype signature-based high-throughput screening for the discovery of combination immunotherapy compounds. *Sci. Adv.* 7, eabd7851. doi:10.1126/sciadv.abd7851

Wang, L., Ma, J., Liu, F., Yu, Q., Chu, G., Perkins, A. C., et al. (2007). Expression of MUC1 in primary and metastatic human epithelial ovarian cancer and its therapeutic significance. *Gynecol. Oncol.* 105 (3), 695–702. doi:10.1016/j.ygyno.2007.02.004

Wang, X., Zhao, J., Zhang, Y., Liu, Y., Wang, J., Shi, R., et al. (2023). Molecular mechanism of Wilms' tumor (*Wt1*) (+/−KTS) variants promoting proliferation and migration of ovarian epithelial cells by bioinformatics analysis. *J. Ovarian Res.* 16, 46. doi:10.1186/s13048-023-01124-2

Ween, M. P., Oehler, M. K., and Ricciardelli, C. (2011). Role of versican, hyaluronan and CD44 in ovarian cancer metastasis. *Int. J. Mol. Sci.* 12, 1009–1029. doi:10.3390/ijms12021009

Wei, C., Liu, X., Wang, Q., Li, Q., and Xie, M. (2021). Identification of hypoxia signature to assess the tumor immune microenvironment and predict prognosis in patients with ovarian cancer. *Int. J. Endocrinol.* 2021, 4156187. doi:10.1155/2021/4156187

World Health Organization International Agency for Research on Cancer (WHO-IARC) (2020). Globocan 2020: Estimated cancer incidence, mortality and prevalence worldwide in 2020. Ovary Avaialable at: https://gco.iarc.fr/today/data/factsheets/cancers/25-Ovary-fact-sheet.pdf (Accessed May 5, 2023).

Wu, Q.-W., Xia, J.-F., Ni, J.-C., and Zheng, C.-H. (2021). Gaerf: Predicting lncRNA-disease associations by graph auto-encoder and random forest. *Brief. Bioinform* 22, bbaa391. doi:10.1093/bib/bbaa391

Xiao, C., Wang, Y., Cheng, Q., and Fan, Y. (2022). Increased expression of MMP17 predicts poor clinical outcomes in epithelial ovarian cancer patients. *Med. Baltim.* 101 (34), e30279. doi:10.1097/MD.0000000000030279

Xing, F., Saidou, J., and Watabe, K. (2010). Cancer associated fibroblasts (CAFs) in tumor microenvironment. *Front. Biosci. (Landmark Ed.)* 15 (1), 166–179. doi:10.2741/3613

Xu, N., Wu, Y.-P., Yin, H.-B., Xue, X.-Y., and Gou, X. (2018). Molecular network-based identification of competing endogenous RNAs and mRNA signatures that predict survival in prostate cancer. *J. Transl. Med.* 16, 274. doi:10.1186/s12967-018-1637-x

Yang, J., Wang, C., Zhang, Y., Cheng, S., Wu, M., Gu, S., et al. (2023). A novel autophagy-related gene signature associated with prognosis and immune microenvironment in ovarian cancer. *J. Ovarian Res.* 16 (1), 86. doi:10.1186/s13048-023-01167-5

Yang, Y., Zhao, T., Chen, Q., Li, Y., Xiao, Z., Xiang, Y., et al. (2022). Nanomedicine strategies for heating "cold" ovarian cancer (OC): Next evolution in immunotherapy of OC. *Adv. Sci. Weinh. Baden-Wurtt. Ger.* 9, e2202797. doi:10.1002/advs.202202797

Young, C., and Argáez, C. (2019). Rapid genome-wide testing: A review of clinical utility, cost-effectiveness, and guidelines. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health Available at: http://www.ncbi.nlm.nih.gov/books/NBK549546/ (Accessed May 2, 2023).

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118

Zeng, D., Ye, Z., Shen, R., Yu, G., Wu, J., Xiong, Y., et al. (2021). Iobr: Multi-Omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front. Immunol.* 12, 687975. doi:10.3389/fimmu.2021.687975

Zhai, Y., Lu, Q., Lou, T., Cao, G., Wang, S., and Zhang, Z. (2020). *MUC16*affects the biological functions of ovarian cancer cells and induces an antitumor immune response by activating dendritic cells. *Ann. Transl. Med.* 8 (22), 1494. doi:10.21037/atm-20-6388

Zhang, D., Li, Y., Yang, S., Wang, M., Yao, J., Zheng, Y., et al. (2021). Identification of a glycolysis-related gene signature for survival prediction of ovarian cancer patients. *Cancer Med.* 10 (22), 8222–8237. doi:10.1002/cam4.4317

Zhang, G. Y., Ahmed, N., Riley, C., Oliva, K., Barker, G., Quinn, M. A., et al. (2005). Enhanced expression of peroxisome proliferator-activated receptor gamma in epithelial ovarian carcinoma. *Br. J. Cancer* 92 (1), 113–119. doi:10.1038/sj.bjc.6602244

Zhang, H., Dong, S., and Feng, J. (2019a). Epigenetic profiling and mRNA expression reveal candidate genes as biomarkers for colorectal cancer. *J. Cell. Biochem.* 120, 10767–10776. doi:10.1002/jcb.28368

Zhang, Y., Xu, J., Zhang, N., Chen, M., Wang, H., and Zhu, D. (2019b). Targeting the tumour immune microenvironment for cancer therapy in human gastrointestinal malignancies. *Cancer Lett.* 458, 123–135. doi:10.1016/j.canlet.2019.05.017

Zhu, Y., Li, X., Wang, L., Hong, X., and Yang, J. (2022). Metabolic reprogramming and crosstalk of cancer-related fibroblasts and immune cells in the tumor microenvironment. *Front. Endocrinol.* 13, 988295. doi:10.3389/fendo.2022.988295

Zigrino, P., Löffek, S., and Mauch, C. (2005). Tumor–stroma interactions: Their role in the control of tumor cell invasion. *Biochimie* 87, 321–328. doi:10.1016/j.biochi.2004.10.025

Check for updates

Integrative analysis of circadian clock with prognostic and immunological biomarker identification in ovarian cancer

Lianfang Zhao[1†], Yuqin Tang[2†], Jiayan Yang[1†], Fang Lin[1], Xiaofang Liu[1], Yongqiang Zhang[3]* and Jianhui Chen[1]*

[1]Prenatal Diagnosis Center, Suining Central Hospital, Suining, Sichuan, China, [2]Clinical Bioinformatics Experimental Center, Henan Provincial People's Hospital, Zhengzhou University, Zhengzhou, China, [3]Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou, China

Objective: To identify circadian clock (CC)-related key genes with clinical significance, providing potential biomarkers and novel insights into the CC of ovarian cancer (OC).

Methods: Based on the RNA-seq profiles of OC patients in The Cancer Genome Atlas (TCGA), we explored the dysregulation and prognostic power of 12 reported CC-related genes (CCGs), which were used to generate a circadian clock index (CCI). Weighted gene co-expression network analysis (WGCNA) and protein-protein interaction (PPI) network were used to identify potential hub genes. Downstream analyses including differential and survival validations were comprehensively investigated.

Results: Most CCGs are abnormally expressed and significantly associated with the overall survival (OS) of OC. OC patients with a high CCI had lower OS rates. While CCI was positively related to core CCGs such as *ARNTL*, it also showed significant associations with immune biomarkers including CD8[+] T cell infiltration, the expression of *PDL1* and *CTLA4*, and the expression of interleukins (*IL-16*, *NLRP3*, *IL-1β,* and *IL-33*) and steroid hormones-related genes. WGCNA screened the green gene module to be mostly correlated with CCI and CCI group, which was utilized to construct a PPI network to pick out 15 hub genes (*RNF169*, *EDC4*, *CHCHD1*, *MRPL51*, *UQCC2*, *USP34*, *POM121*, *RPL37*, *SNRPC*, *LAMTOR5*, *MRPL52*, *LAMTOR4*, *NDUFB1*, NDUFC1, *POLR3K*) related to CC. Most of them can exert prognostic values for OS of OC, and all of them were significantly associated with immune cell infiltration. Additionally, upstream regulators including transcription factors and miRNAs of key genes were predicted.

Conclusion: Collectively, 15 crucial CC genes showing indicative values for prognosis and immune microenvironment of OC were comprehensively identified. These findings provided insight into the further exploration of the molecular mechanisms of OC.

Abbreviations: CC, circadian clock; OC, ovarian cancer; CCI, circadian clock index; WGCNA, weighted gene co-expression network analysis; TCGA, The Cancer Genome Atlas; PPI, protein-protein interaction; TME, tumor microenvironment; TIME, tumor immune microenvironment; OS, overall survival; DSS, disease-specific survival; MAD, median absolute deviation; TOM, topological overlap matrix; GS, gene significance; MM, module membership; STRING, the Search Tool for the Retrieval of Interacting Genes; MTIs, miRNA-target interactions; CCGs, CC-related genes; GO, Gene ontology; AUCs, area under the curves.

1 Introduction

With 207,252 deaths worldwide in 2020 (Sung et al., 2021), the mortality rate of ovarian cancer is ranked first among gynecological malignant tumors. The characteristics of insidious symptoms, high degree of malignancy, and easy metastasis of ovarian cancer make it a great challenge for early screening. Nearly 70% of OC patients are diagnosed at an advanced stage (Cho and Shih Ie, 2009). Moreover, the pathogenesis and metastasis mechanism of OC remain elusive, so it is particularly urgent to identify new biomarkers and potential drug targets for improving the clinical outcomes of OC patients.

Circadian rhythm is a physiological cycle phenomenon of about 24 h in mammalian body. Accumulating evidence has elucidated the significance of circadian rhythm composed of central and peripheral clocks in harmony with the external environment (LeGates et al., 2014; Jagannath et al., 2017). *BMAL1* and *CLOCK* form a dimer, which promotes the expression of cryptochrome (*CRY1* and *CRY2*) and period (*PER1*, *PER2*, and *PER3*) genes. *CRY* and *PER* form a complex that enters the nucleus and suppresses the *CLOCK–BMAL1* complex (Etchegaray et al., 2003; Hirayama et al., 2007; Nader et al., 2009). This forms the transcription-translation loop of the central clock. Nuclear receptors *REV-ERBα* and retinoic acid-like orphan receptor α (*RORα*) regulate the ROR element on the *BMAL1* promoter, providing a stable oscillation cycle (Badiu, 2003). The circadian clock gene is the molecular basis for the circadian clock to produce and maintain circadian rhythms. At present, increasing evidence has found that more than ten circadian clock genes form transcriptional translation loops, including *ARNTL*, *CLOCK*, *CRY1*, *CRY2*, *NR1D1*, *NR1D2*, *PER1*, *PER2*, *PER3*, *RORA*, *RORB*, *RORC*, etc. (Yang et al., 2017). Target genes regulated by circadian clock genes are called clock control genes, which are effector molecules of the circadian clock. In mammals, clock control genes play an important role in maintaining physiological homeostasis (e.g., hormone secretion, cell growth, proliferation, and cell metabolism) (Udo et al., 2004).

A growing number of studies strongly support the existence of crosstalk between cancer and the circadian clock (Battaglin et al., 2021). Increasing data have suggested that circadian clocks exerted a vital function in the regulation of cancer-related physiological systems, such as cell proliferation, cell apoptosis, DNA injury and repair, and metabolism (Peyric et al., 2013; Dakup and Gaddameedhi, 2017). The imbalance of the circadian clock particularly affects the progression of endocrine-related cancers including cervical, prostate, and ovarian cancers by dysregulating key hormone levels (Morales-Santana et al., 2019; Hadadi and Acloque, 2021). Aberrant expression of *ARNTL* and *CRY1* was found in OC cell lines (Tokunaga et al., 2008). Moreover, the low expression of *CRY1* and *BMAL1* was reported to be associated with the poor survival of OC patients (Tokunaga et al., 2008), and the overexpression of *BMAL1* could inhibit the cell growth of OC (Yeh et al., 2014). But controversy still exists. For existence, a recent study did not find the increased risk of ovarian cancer for night-shift work (Dun et al., 2020). Thus, a comprehensive understanding of circadian clock is required to better evaluate its critical role in the carcinogenesis of OC, which will enable the identification of clinical biomarkers and molecular targets.

Moreover, previous studies indicated that circadian clock gene disruption contributes to independent risk factors for tumor microenvironment (TME) (Aiello et al., 2020; Xuan et al., 2021). However, the specific functions of circadian clock in the prognosis and therapy of OC are still unclear. Therefore, this study aimed to elucidate the vital role of the circadian clock in OC using multi-omics methods. We integratively identified 15 potential key genes related to CC, which showed great correlations to the infiltration levels of multiple tumor immune microenvironment (TIME) cell types. These data strengthened the linkage of CC and tumor immune status of OC, and extended the understanding of its molecular mechanism, and survival analysis results suggested their good potential in the future development of new prognostic biomarkers and immunotherapy targets. The workflow of this study is shown in Figure 1.

2 Materials and methods

2.1 Data collection and preprocessing

We downloaded the gene-expression dataset (TCGA-OV) and the complete OS data of 378 OC in the TCGA database from the UCSC XENA project (https://xena.ucsc.edu). The disease-specific survival (DSS) and corresponding clinicopathologic characteristics including age, stage, grade, tumor status, and lymphatic invasion status were also achieved. The RNA-seq data were preprocessed as we previously reported (Zhang et al., 2022). 12 literature-derived CCGs in basic mammalian central feedback loop were collected (*ARNTL*, *CLOCK*, *PER1*, *PER2*, *PER3*, *CRY1*, *CRY2*, *NR1D1*, *NR1D2*, *RORA*, *RORB*, and *RORC*) (Shen et al., 2020; Cash et al., 2021). Gene expression levels between the tumor and adjacent normal tissues were compared using the online GEPIA database (Tang et al., 2017).

2.2 Correlations of CCI with the prognosis and TIME of OC

Based on the TCGA-OV dataset, we generated a circadian clock index to comprehensively represent the overall activity of the CC status, which was computed as the average RNA-seq z-scores of the 12 core CCGs for each sample. All patients were designated into the high- and low-CCI groups depending on the optimal cutoff for the Kaplan-Meire survival analysis. The single sample gene set enrichment analysis (ssGSEA) algorithm was applied to infer the activated CD8 T cell infiltration by the "GSVA" package (Bindea et al., 2013; Charoentong et al., 2017; Zhang et al., 2022; Guo et al., 2023). Scatter plots were drawn to determine the correlations of CCI with TIME signatures including activated CD8 T cell infiltration, the expression of well-known immunotherapeutic targets such as *PDL1* and *CTLA4*, and steroid hormones-related genes. The correlations between CCI and multiple interleukins including the *IL-1*

FIGURE 1
Workflow of this study.
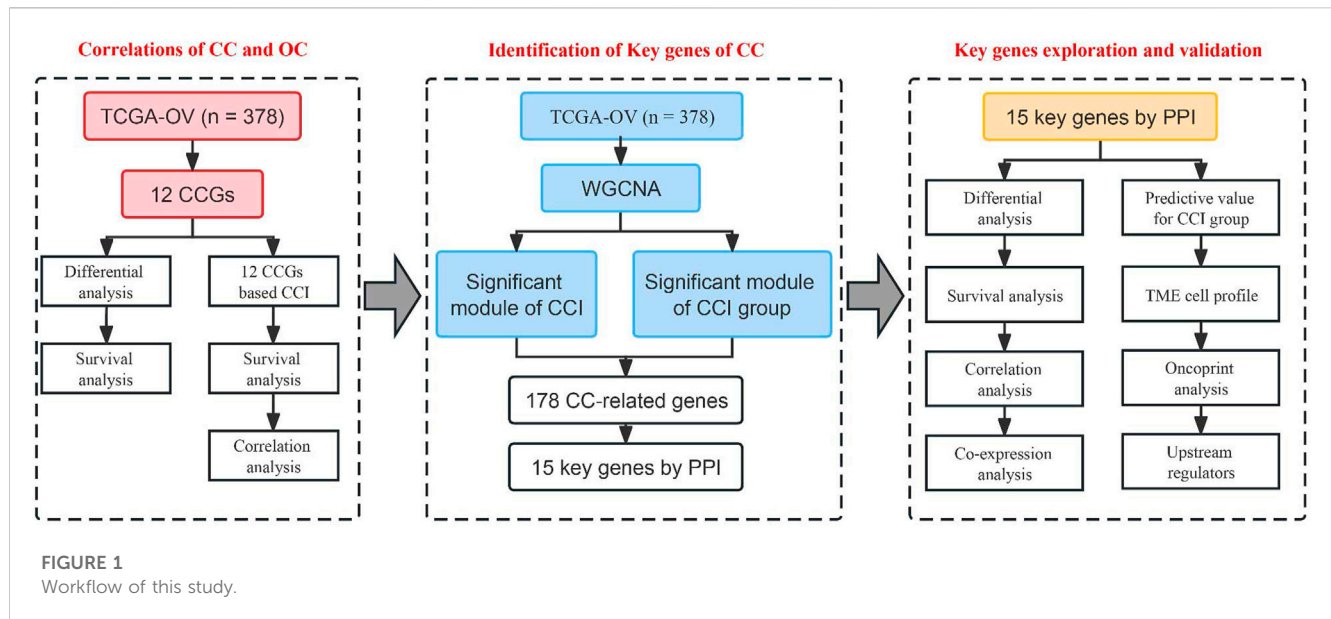
superfamily, *IL-6* family, *IL-10* family, *IL-17* family, and other interleukins were also evaluated.

2.3 WGCNA and key module identification

We construct a co-expression network with the "WGCNA" package in R (Langfelder and Horvath, 2008; Tang et al., 2022b). First, we filtered the median absolute deviation (MAD) top 5,000 genes for sample clustering (Tang et al., 2022b). Next, we select the best soft-thresholding power to ensure the scale-free topology ($R^2 > 0.9$). The adjacency was then transformed into a topological overlap matrix (TOM) by using TOM similarity and the corresponding dissimilarity (dissTOM). Finally, at least 30 co-expressed genes were aggregated into different MEs by the dynamic tree-cut method. We decided on a cut line (0.3) for merging highly similar modules to make the modules more compact (Zhang et al., 2022). To determine their relevance to clinical traits, Pearson correlations between module eigengenes and clinical phenotypes were calculated and were shown with a correlation heatmap. In this study, we chose the most significant module that correlated with CCI and CCI group for further analysis, and gene significance (GS) and module membership (MM) were also calculated. The CC-related genes were determined as those with MM > 0.5 and GS > 0.4 in the most significant module.

2.4 PPI network construction and key genes identification

The PPI network is a useful tool to explore molecular interactions as well as to identify potential key biomarkers. Based on the CC-related genes of OC, we constructed a PPI network by the Search Tool for the Retrieval of Interacting Genes (STRING) database with a combined score of higher

than 0.7 (high confidence) (Zhang et al., 2021). The Cytoscape software (version 3.8.2, http://www.cytoscape.org) was then utilized for the visualization. Next, we performed sub-module analysis with the MCODE plugin. Potential key genes of CC were defined as those with a clustering coefficient of 1 in the entire PPI network.

2.5 Function enrichment

The DAVID database was used to conduct the function enrichment for the CC-related genes and the key genes in the most significant gene module. The Gene ontology (GO) database, which is comprised of MF, BP, and CC, is used for performing gene annotation, and the Rectome database was used to analyze the major pathways. The "ggplot2" package was used for data visualization.

2.6 Assessment of prognostic and clinical significance

For survival evaluation, we depicted the Kaplan-Meire plots with log-rank tests for OS or DSS by the optimized expression value, which was carried out on the 378 OC patients (Guo et al., 2021; Tang et al., 2022b), who were subsequently classified into different risk groups. To unveil the correlations between the expression of these key genes and clinicopathological traits, we compared their expression levels between different subgroups divided by clinicopathological features including age, stage, grade, tumor status, and lymphatic invasion status. Pearson correlation matrices were further computed to evaluate the co-expression of these crucial genes of CC. Besides, we use boxplots and ROC curves to assess the predictive capabilities of these key genes for discriminating the distinct CCI groups. The online cBioPortal database (http://www.cbioportal.org/) was used to verify their genetic alterations with an oncoprint diagram.

2.7 Immune infiltration profile

To explore the relationships between the potential key genes and TIME, we utilized the CIBERSORT algorithm (Newman et al., 2015; Tang et al., 2022a) to estimate the relative abundance of 22 immune cell types in each OC sample based on the TCGA-OV dataset. The relative abundance of immune cells in different CCI groups was computed and presented by a heatmap plot. Spearman correlation analysis was used to assess the relevance of the key genes' expression and immune cell infiltration.

2.8 Prediction of upstream regulations-transcript factor and miRNA

MiRNet 2.0, an up-to-date online platform that illustrates "multiple-to-multiple" relationships and functional interpretation was used to examine the upstream transcription factors of key genes. We selected the JASPAR database to screen the possible transcription factor-key gene interactions in the present study. Moreover, the upstream miRNAs of the key genes were putatively predicted with miRTarBase 8.0, which is a powerful tool to predict miRNA-target interactions (MTIs) that were verified by multiple cell-based experiments. The transcription factor-key gene interactions and miRNA-key genes interactions were visualized by the Cytoscape software.

3 Results

3.1 Survival and differential analysis of 12 CCGs in OC

To achieve a systematic understanding of the circadian clocks in OC, we first selected 12 well-known circadian genes in this study, namely, *ARNTL*, *CLOCK*, *PER1*, *PER2*, *PER3*, *CRY1*, *CRY2*, *NR1D1*, *NR1D2*, *RORA*, *RORB*, and *RORC*. Based on the GEPIA online database, we preliminarily compared the mRNA levels of these genes in OC tissues and adjacent ovary tissues (Supplementary Figure S1A). As a result, the mRNA level of *RORC* was upregulated, while the expression of *ARNTL*, *CRY2*, *NR1D1*, *PER1*, *PER3*, and *RORA* was decreased in OC tissues versus normal tissues. This evidence demonstrates that core circadian clock genes are widely altered at the mRNA level in OC. To analyze the prognostic value of circadian clock genes in OC, we performed Kaplan-Meier analysis and log-rank tests of the 12 CCGs for OS of OC. As shown in Supplementary Figure S1B, we found that *ARNTL* ($p = 0.021$), *CRY1* ($p = 0.0074$), *CRY2* ($p = 0.045$), *NR1D1* ($p = 0.0035$), *NR1D2* ($p = 0.0062$), *PER1* ($p = 0.014$) and *PER2* ($p = 0.025$) were significantly associated with the OS of OC. These results suggest a possible link between CC and the progression and prognosis in OC.

3.2 CCI was significantly associated with the prognosis and immune status in OC

Firstly, we computed CCIs for all of the OC patients and performed heatmap visualization to show the expression pattern

of CCI genes with CCI group annotation in OC (Figure 2A). To evaluate the prognostic power of CCI, we carried out Kaplan-Meier analysis for OS. As a result, OC patients with high expression CCI ($p = 0.009$) had much poorer OS rates than the low CCI group (Figure 2B). Subsequently, we confirmed the correlations between CCI and two core CCGs (*ARNTL*, *CLOCK*) by scatter plots, and found strong positive correlations for both *ARNTL* ($p = 1.28e-42$, rpearson = 0.63) and *CLOCK* ($p = 3.61e-19$, rpearson = 0.44) were linked to CCI in OC (Figures 2C, D). Numerous studies have elaborated on the close correlation between the disruption of CC and the activity of tumor immune status as well as therapeutic effects in various cancers (Yang et al., 2017; Sulli et al., 2019; Kinouchi and Sassone-Corsi, 2020; Xuan et al., 2021; He et al., 2022). To explore the correlations of CCI and immune signatures, we clarified the correlation between CCI and CD8$^+$ T cells in OC and found that the CCI was negatively associated with the abundance of CD8$^+$ T cells ($p = 5.94e-04$, rpearson = -0.18). Considering the well-known indicative powers of *PDL1* and *CTLA4* for immunotherapeutic response in cancers and their prognostic values in OC (Huang et al., 2017; Huang and Odunsi, 2017), we also examined the correlations between CCI and the expression of them. Interestingly, CCI was positively correlated with the expression of *PDL1* ($p = 1.20e-05$, rpearson = 0.22) and *CTLA4* ($p = 0.01$, rpearson = 0.13) (Figures Figure2E–G). Given the particular relevance and pleiotropic role of interleukins in the progression of cancer, which has attracted great interest of interleukins in translational cancer research recently (Briukhovetska et al., 2021), we further examined the correlations between CCI and several interleukin families. As a result, *IL-16* was with the strongest correlation with CCI, and significant associations were also observed for *NLRP3* and *IL-1* superfamily (including *IL-1β* and *IL-33*) (Figure 2H; Supplementary Figure S1). Besides, we also found significant correlations between CCI and several steroid hormones-related genes (*SRD5A2*, *HSD17B12*, and *NR3C1*) (Supplementary Figure S2). Collectively, these data demonstrated CCI may provide indicative value for future immunotherapy in OC.

3.3 Identification of key biomarkers and CC-related genes via WGCNA

WGCNA is a useful approach to construct a co-expression network of genes and to identify significant gene modules or key biomarkers from multiple samples in cancer. In this study, we conducted WGCNA to disclose the most important module associated with CCI and the CCI group. We chose the optimal soft-thresholding power of 4 (scale-free $R^2 > 0.90$) as the soft-thresholding to construct a scale-free network (Supplementary Figure S3), followed by the hierarchical clustering of samples using the average linkage method (Figure 3A). Next, the adjacency matrix was produced and transformed into a TOM, which was used to calculate the dissTOM (1-TOM) for evaluating the distance of genes. The dissTOM was subsequently used to conduct hierarchical clustering and to generate dynamic gene modules. After
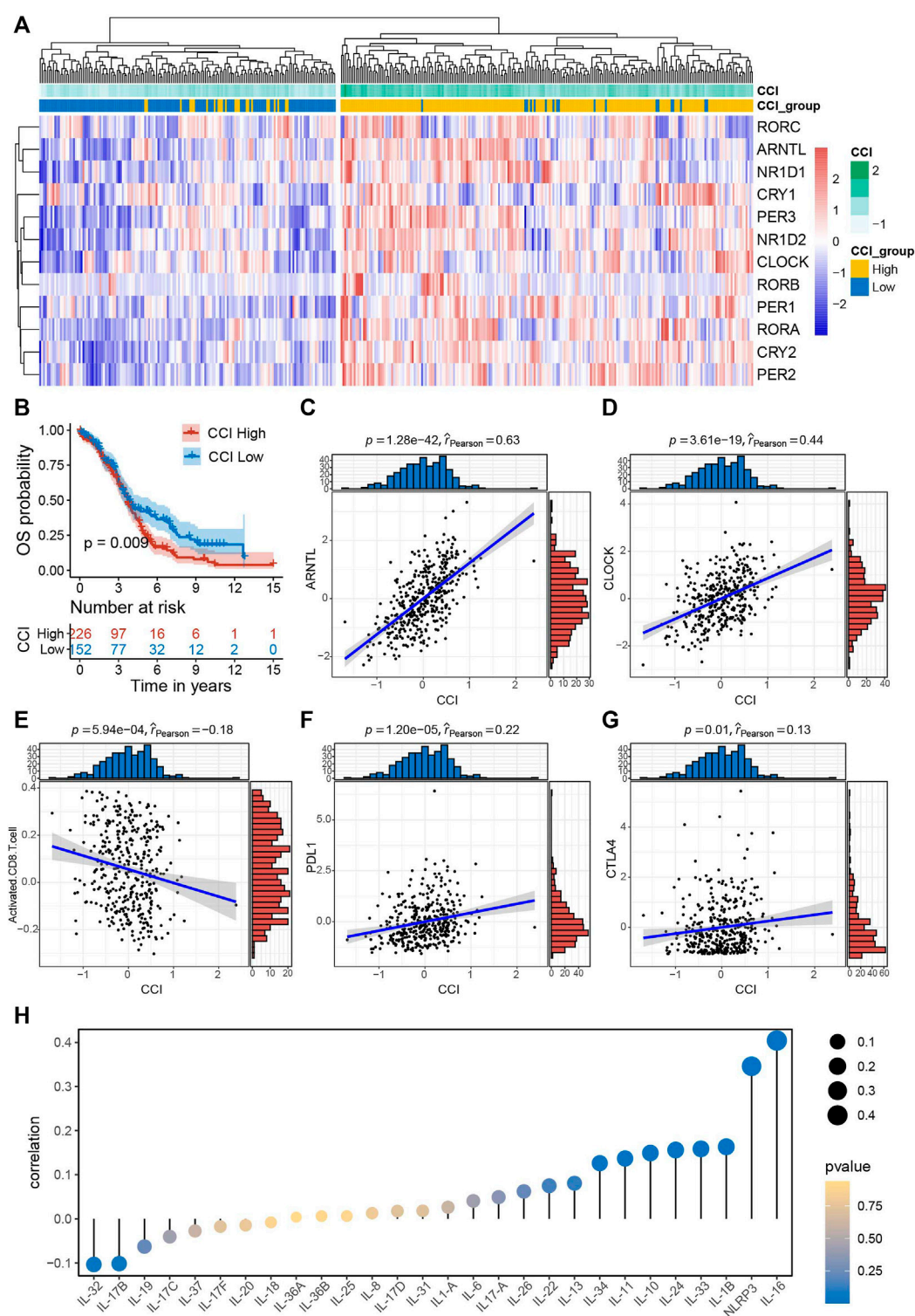
FIGURE 2
Survival significance of circadian clock index (CCI) and the correlations between CCI and common immune signatures. **(A)** Clustering heatmap of the expression pattern of CCI genes with CCI group annotation in OC. **(B)** Kaplan-Meire plots for OS in the high- and low-CCI group of OC patients. **(C,D)** The strong correlations between CCI and two core CCGs in OC. **(E–G)** The significant correlations between CCI and important immune signatures. **(H)** correlations between CCI and interleukin families.

merging highly similar modules by the cut line of 0.3, a total of twelve modules were identified (Figure 3B). The Pearson correlation heatmap showed the green module has the most

significant correlation with both the CCI group and CCI ($R^2 = 0.61$ and p = 3e−39 for the CCI group, $R^2 = 0.7$ and $p = 13e−57$ for CCI, respectively). Thus the green gene
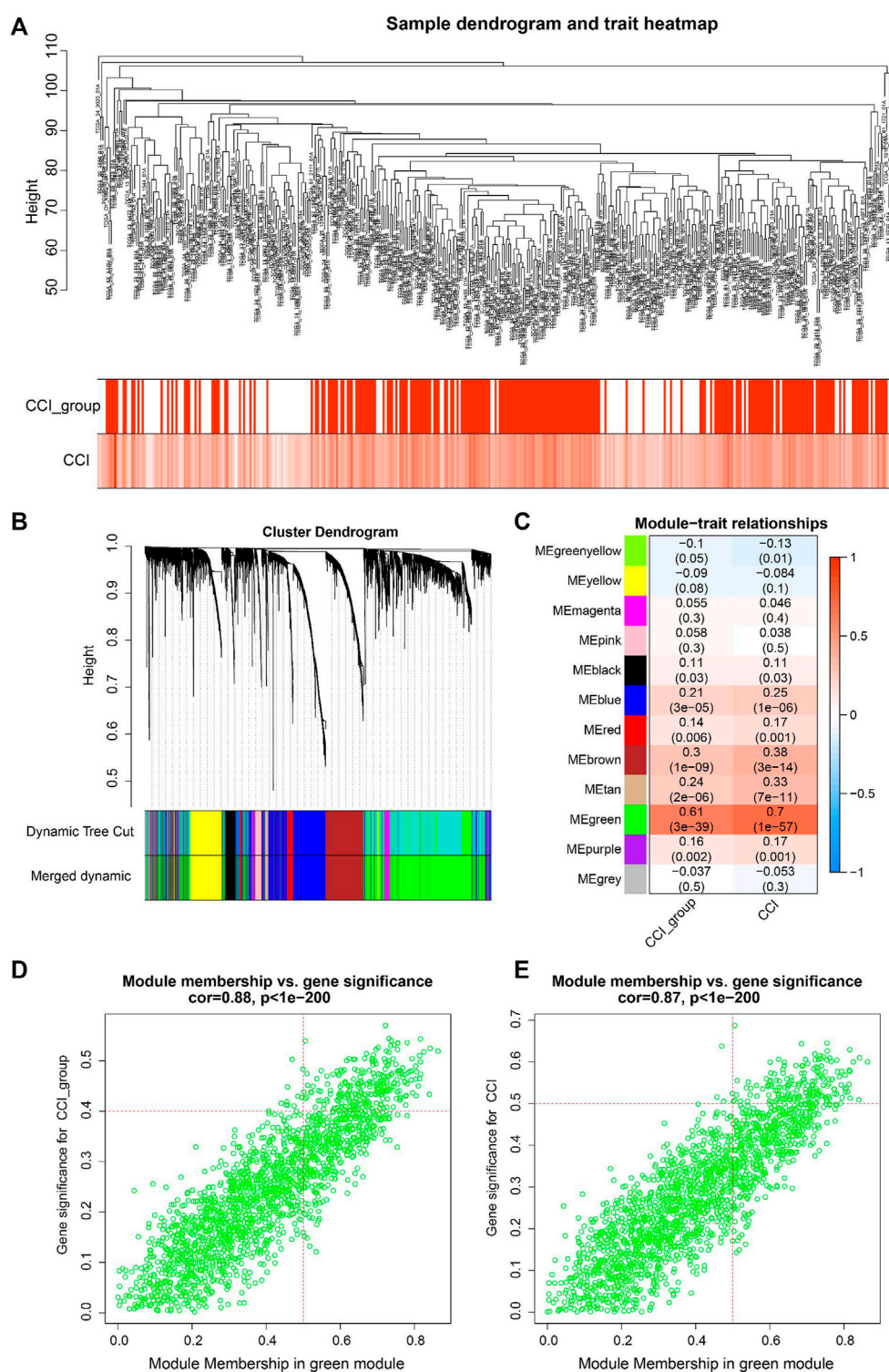
FIGURE 3
WGCNA for identification of key module and circadian clock-related genes in OC. **(A)** Sample clustering tree with CCI and CCI group annotations. **(B)** Gene clustering dendrogram with the topological overlap matrix (TOM) based dissimilarity. **(C)** Pearson correlation analysis between module eigengenes and CCI and CCI group. **(D,E)** Scatter plots visualizing the gene significance (GS) vs. module membership (MM) of each gene for CCI and the CCI group in the green gene module. CC-related genes were regarded as those with MM > 0.5 and GS > 0.4 for CCI group and GS > 0.5 for CCI in the green module.

module was selected for further study (Figure 3C). Besides, the GS and MM of each gene for CCI and the CCI group in the green module were presented in Figures 3D, E and 178CC-

related genes in the green module were picked out with the MM > 0.5 and GS > 0.4 for the CCI group and GS > 0.5 for CCI.
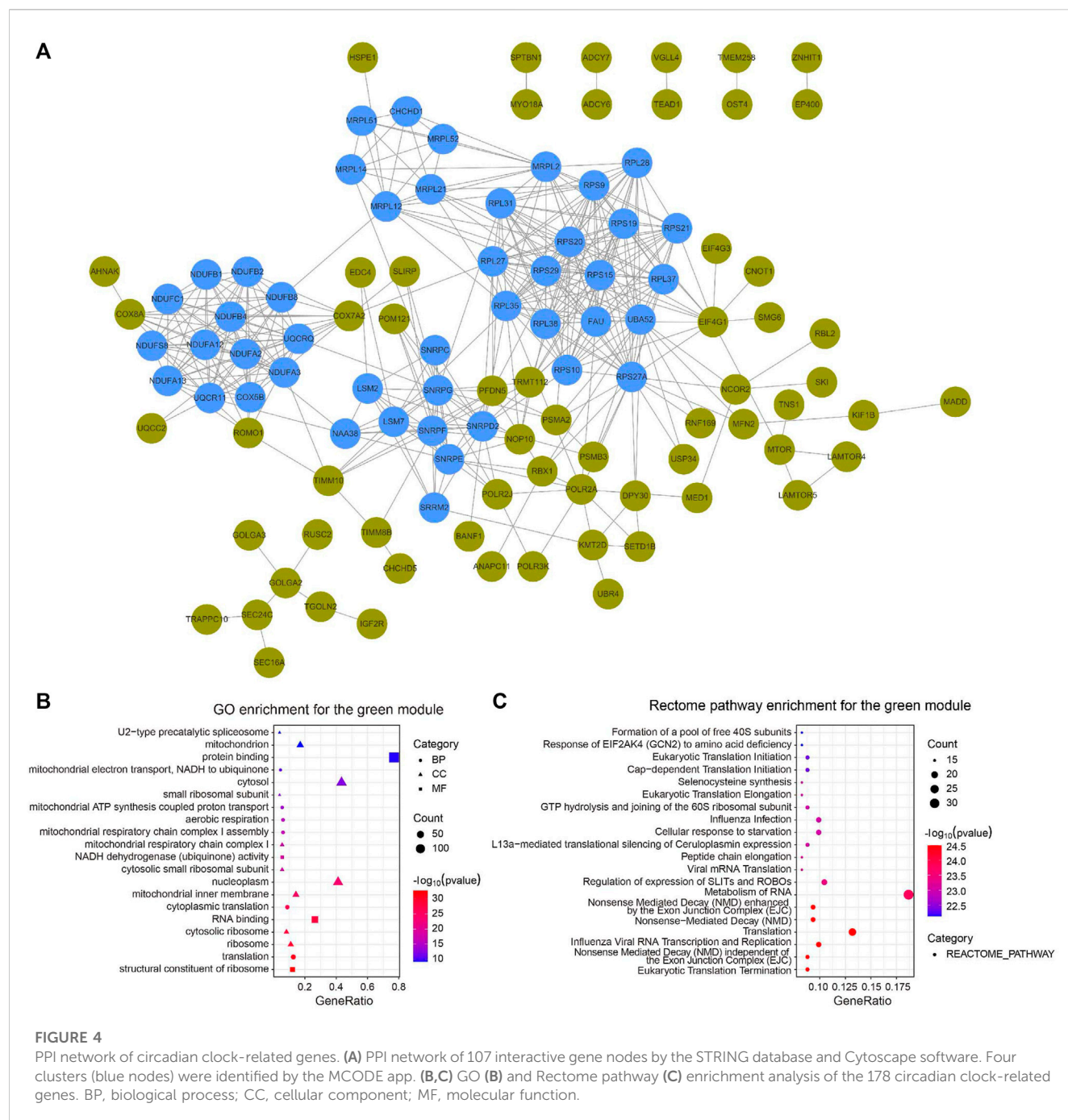
FIGURE 4
PPI network of circadian clock-related genes. **(A)** PPI network of 107 interactive gene nodes by the STRING database and Cytoscape software. Four clusters (blue nodes) were identified by the MCODE app. **(B,C)** GO **(B)** and Rectome pathway **(C)** enrichment analysis of the 178 circadian clock-related genes. BP, biological process; CC, cellular component; MF, molecular function.

3.4 PPI network of CC-related genes

We constructed a PPI network with the CC-related genes in the green module using the STRING online database and the Cytoscape software (Tang et al., 2020), which contained 107 interactive nodes and 424 edges (Figure 4A). The clustering coefficient was 0.554 and the average number of neighbors was 9.258. Four clusters (blue nodes) were identified by the MCODE app with a high network score (>6). To explore the involved biological functions and pathways of the 178 circadian clock-related genes, we conducted GO and Rectome pathway analysis. GO analysis demonstrated these CC-related genes were mostly enriched in the GO terms of protein

binding, cytosol, nucleoplasm, RNA binding, etc. (Figure 4B). Meanwhile, pathway analysis suggested the green gene module was mainly related to the Rectome pathway of Metabolism of RNA and so on (Figure 4C).

3.5 Identification and function enrichment of key genes of the circadian clock

For the identification of the key genes, we selected the top 15 nodes in the PPI network with a clustering coefficient of 1, i.e., *RNF169*, *EDC4*, *CHCHD1*, *MRPL51*, *UQCC2*, *USP34*,

FIGURE 5
Identification and function enrichment of key genes of the circadian clock. **(A)** 15 key genes of CC with connected genes in the PPI network. **(B,C)** GO **(B)** and Rectome pathway **(C)** enrichment analysis of 15 key genes of CC (Clustering coefficient = 1).
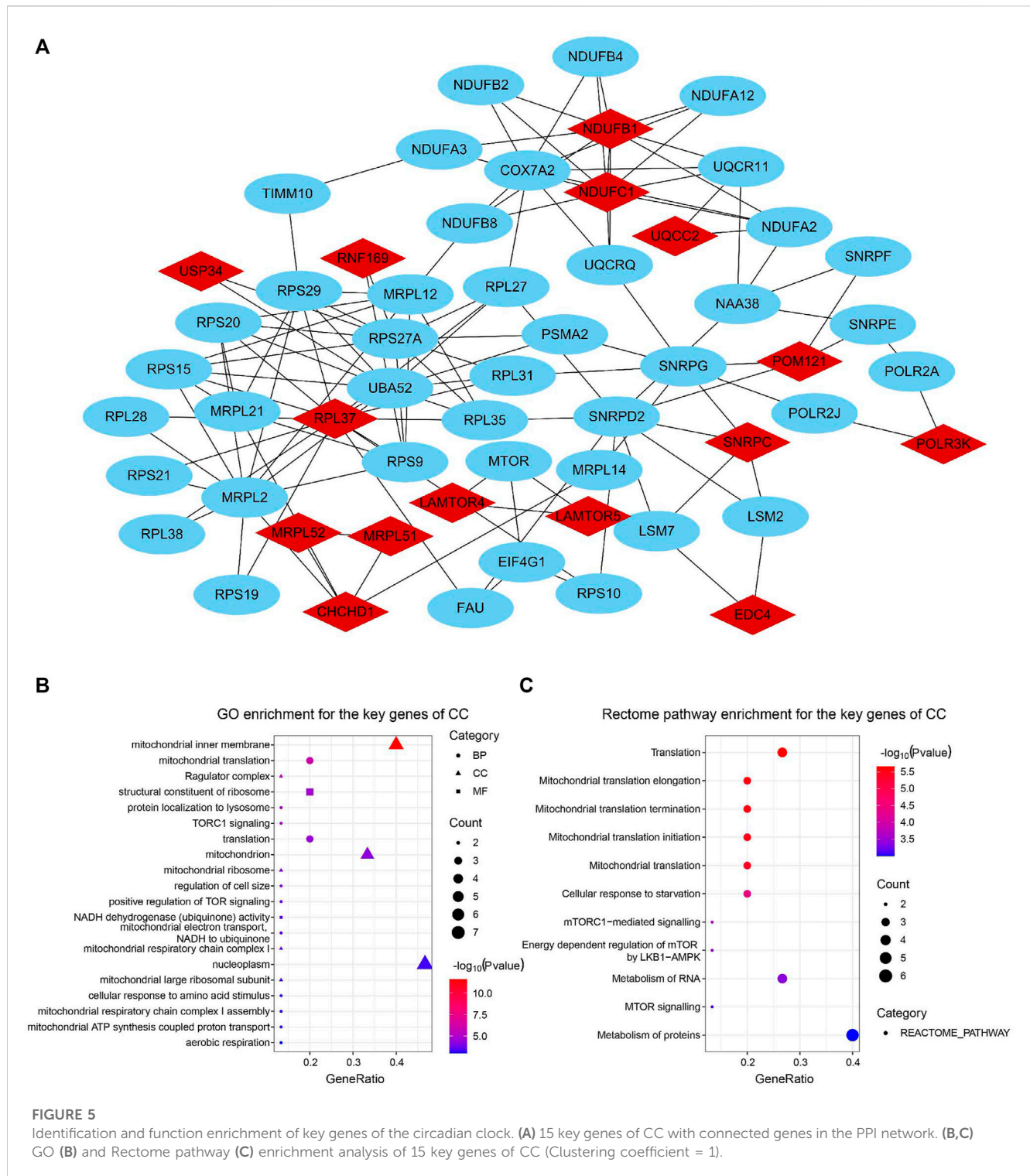
POM121, *RPL37*, *SNRPC*, *LAMTOR5*, *MRPL52*, *LAMTOR4*, *NDUFB1*, *NDUFC1*, and *POLR3K*. Figure 5A showed the 15 key genes (red) of CC with connected neighbors (blue) in the PPI network. We also employed GO enrichment to elucidate their biological functions and found them remarkably correlated with the GO term of the mitochondrial inner membrane, mitochondrion, and nucleoplasm (Figure 5B). Meanwhile,

Rectome pathway analysis showed that the key genes of CC played roles in Mitochondrial translation initiation, Mitochondrial translation elongation, Mitochondrial translation termination, Cellular response to starvation, and Metabolism of RNA and proteins pathways (Figure 5C). The above results indicate their potential hub roles for tumorigenesis.

FIGURE 6
Kaplan−Meier survival curves of the 15 key genes characterizing OS difference with log-rank tests in OC (only significant genes were shown). OS, overall survival.
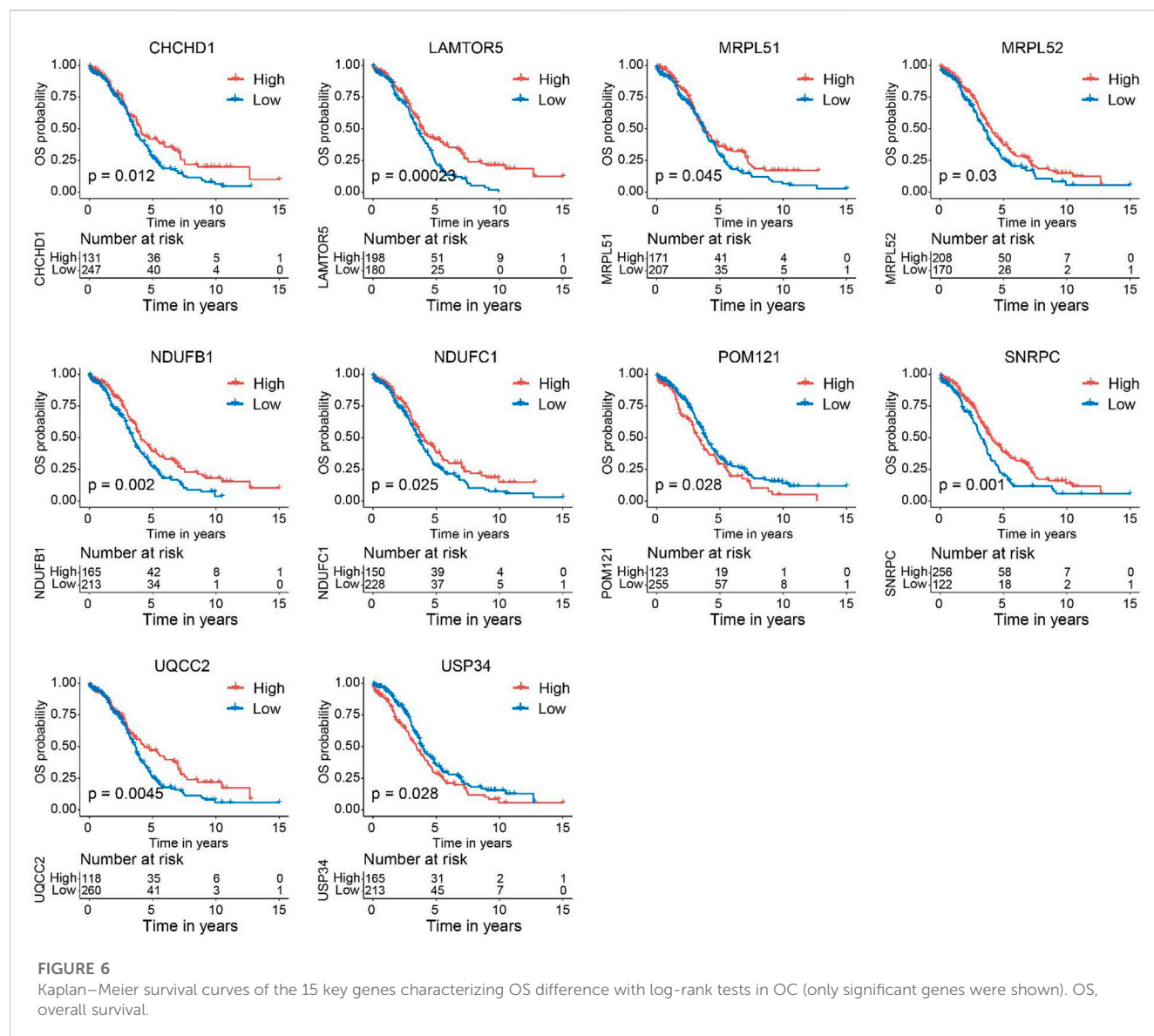
3.6 Aberrant expression and prognostic values of the key genes

Next, we verified the expression of the filtered key genes in the GEPIA database. *EDC4*, *SNRPC*, and *UQCC2* were found significantly differentially expressed in OC vs. normal tissue ($p <$ 0.05). All of them were elevated in OC tissues compared with the normal tissues (Supplementary Figure S4). To evaluate the prognostic powers of these key genes, we examined the 15 key genes in perspective of OS and DSS using Kaplan-Meier analysis and log-rank tests. Among these key genes, the expression levels of *CHCHD1*, *LAMTOR5*, *MRPL51*, *MRPL52*, *NDUFB1*, *NDUFC1*, *POM121*, *SNRPC*, *UQCC2*, and *USP34* were significantly linked to OS of OC (Figure 6). For DSS survival analysis, We found that low expression of *CHCHD1* ($p = 0.048$), *LAMTOR5* ($p = 0.00048$), *MRPL51* ($p = 0.016$), *MRPL52* ($p =$ 0.036), *NDUFB1* ($p = 0.013$), *NDUFC1* ($p = 0.035$), *SNRPC* ($p =$

3e-04), *UQCC2* ($p = 0.0029$) was significantly associated with worse prognosis (Figure 7). These data showed the good potential of the 15 key genes for the development of prognostic indicators in OC.

3.7 Correlation analysis of the key genes

To seek the clinical relevance of these key genes, we compared the expression levels of all 15 key genes between subgroups by clinical variables including age, stage, grade, tumor status, and lymphatic invasion status, but all showed no significant differences for most of these features except age and tumor stage (Figures 8A, B). Moreover, Pearson correlation analysis coupled with statistical significance demonstrated strong correlations between the expression of these key genes, implying their tight connections (Figure 8C).

FIGURE 7
Kaplan–Meier survival curves of the 15 key genes characterizing DSS difference with log-rank tests in OC (only significant genes were shown). DSS, disease-specific survival.
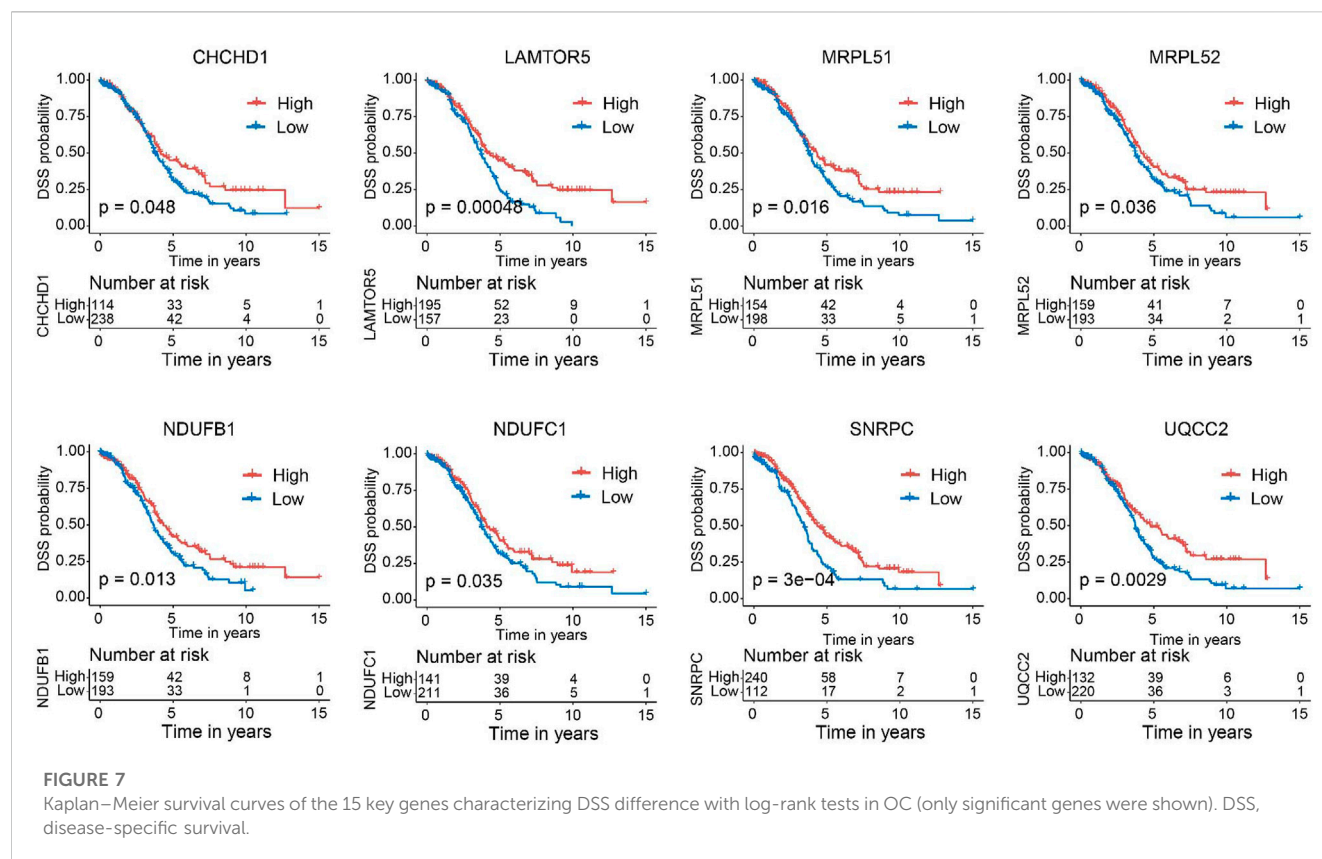
3.8 All of the key genes were highly indicative of the CCI group in OC

To evaluate the discriminating capacities of the 15 key genes in different CCI groups, we compared the transcription expression levels of the 15 key genes between the high-CCI group and the low-CCI group. Consequently, all of the 15 key genes showed statistical differences between the high CCI group and the low CCI group (Figure 9A). The above findings prompted us to speculate that these key genes may have good discriminating capacities for different CCI groups, which was verified by plotting ROC curves. The area under the curves (AUCs) of individual key genes exceeds 0.7 (Figures 9B–D). The above findings proved their great potential for discriminating between different CCI groups.

3.9 The key genes were significantly correlated to the immune infiltration of OC

To further explore the correlations between the CCI group and immune infiltration, the relative proportions of 22 immune cell types in OC were calculated by the CIBERSORT algorithm (Figure 10A). Then, we examined the correlation between immune cell infiltration and the expression of the 15 key genes, which indicated that all 15 key genes of the CCI group were significantly associated with most of the immune cells (Figure 10B). At least 8 of them were negatively associated with the infiltration of CD8 cells. And interestingly, we found 11 kinds of immune cells were positively correlated to T cells CD4 memory

resting, while they were all negatively correlated to B cells memory in OC. The above results indicated that these key genes might play an important role in the immune status of OC.

3.10 Other downstream validations of the key genes

Subsequently, we focused on the potential upstream regulations, including transcription factor and regulative miRNA of the 15 key genes of CC. We predicted the upstream transcription factors of the 15 key genes by the JASPAR database that was integrated into the web-based application of miRNet. A transcription factor-key gene network was constructed and visualized by the Cytoscape software, which includes all 15 key genes and 44 putative upstream TFs (Supplementary Figure S5). Obviously, well-known oncogenic transcription factors such as *JUN*, and *E2F1* were involved in the expression regulation of the key genes. Considering the critical role of miRNA in regulating gene expression and in tumorigenesis (Zhang et al., 2016), the upstream miRNAs of the 15 key genes of CC were predicted by the miRTarBase 8.0 database. The miRNA-target network was next constructed and visualized by the Cytoscape software, which includes 13 of the 15 genes and 242 putative upstream miRNAs (Supplementary Figure S6). Hence, these data could lay a firm basis for the role of the CC-related key genes in the further exploration of the molecular mechanisms of OC.

To obtain a deeper underlying of the 15 key genes from the DNA layer, we analyzed the gene mutation landscape of the 15 key genes of CCI with the aid of cBioPortal. Supplementary Figure S7 showed
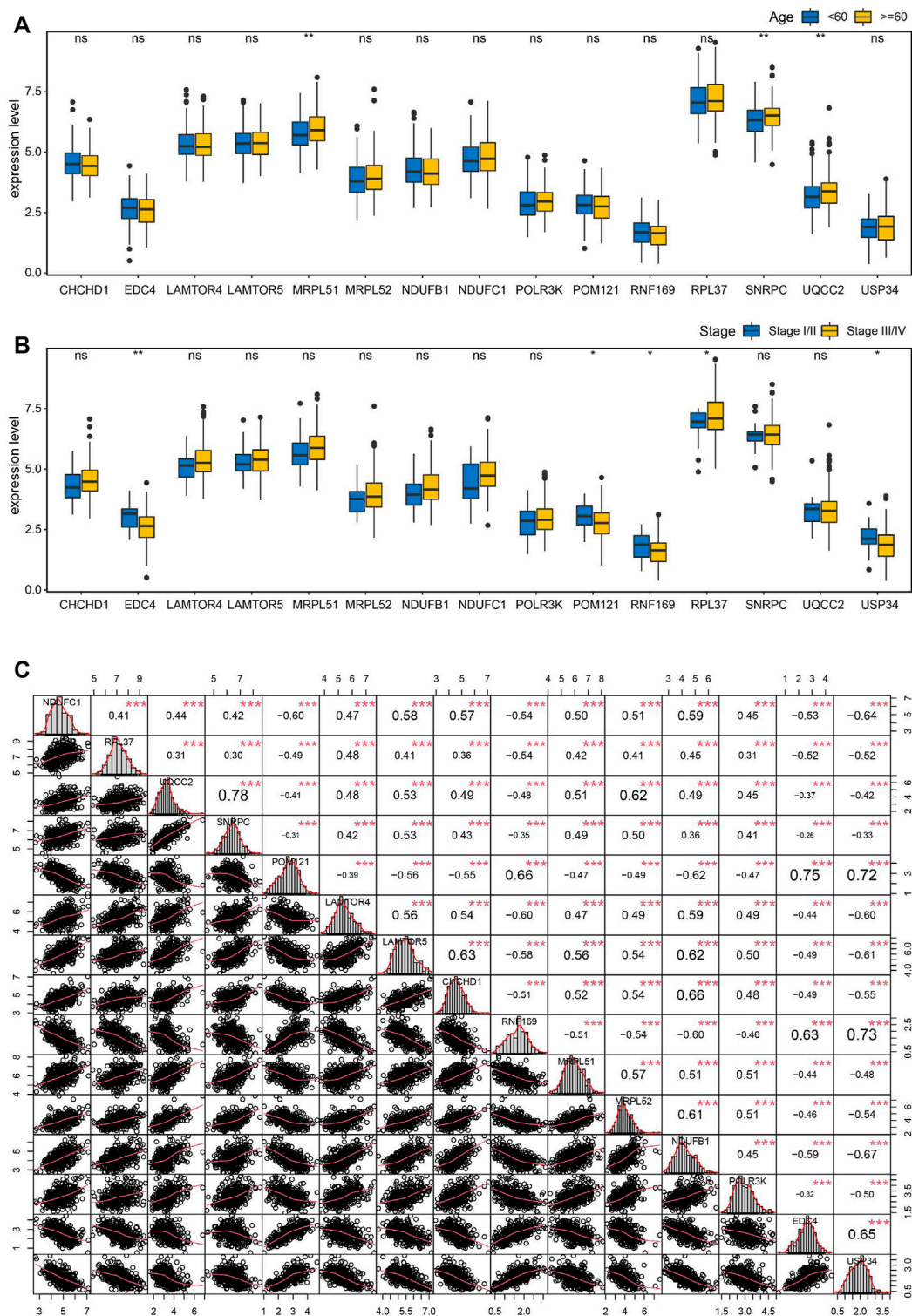
FIGURE 8
Correlation analysis of the key genes in OC. **(A)** Boxplots showing the correlations of the key genes with age **(A)** and stage **(B)**. **(C)** Pearson correlation matrices of expression values of the key genes. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ns, no significance.

all kinds of common genetic mutations of the 15 key genes including inframe-mutation, Missense-mutation, truncating-mutation, amplification, deep deletion, mRNA high, and mRNA low. Among these genes, *RNF169* was the most mutated one, which

exhibited a genetic mutation rate of 14% in the TCGA-OV cohort, other frequently mutated key genes including *EDC4* (11%), *CHCHD1* (10%), *MRPL51* (10%), *UQCC2* (10%), and *USP34*(10%), respectively. These results may disclose the
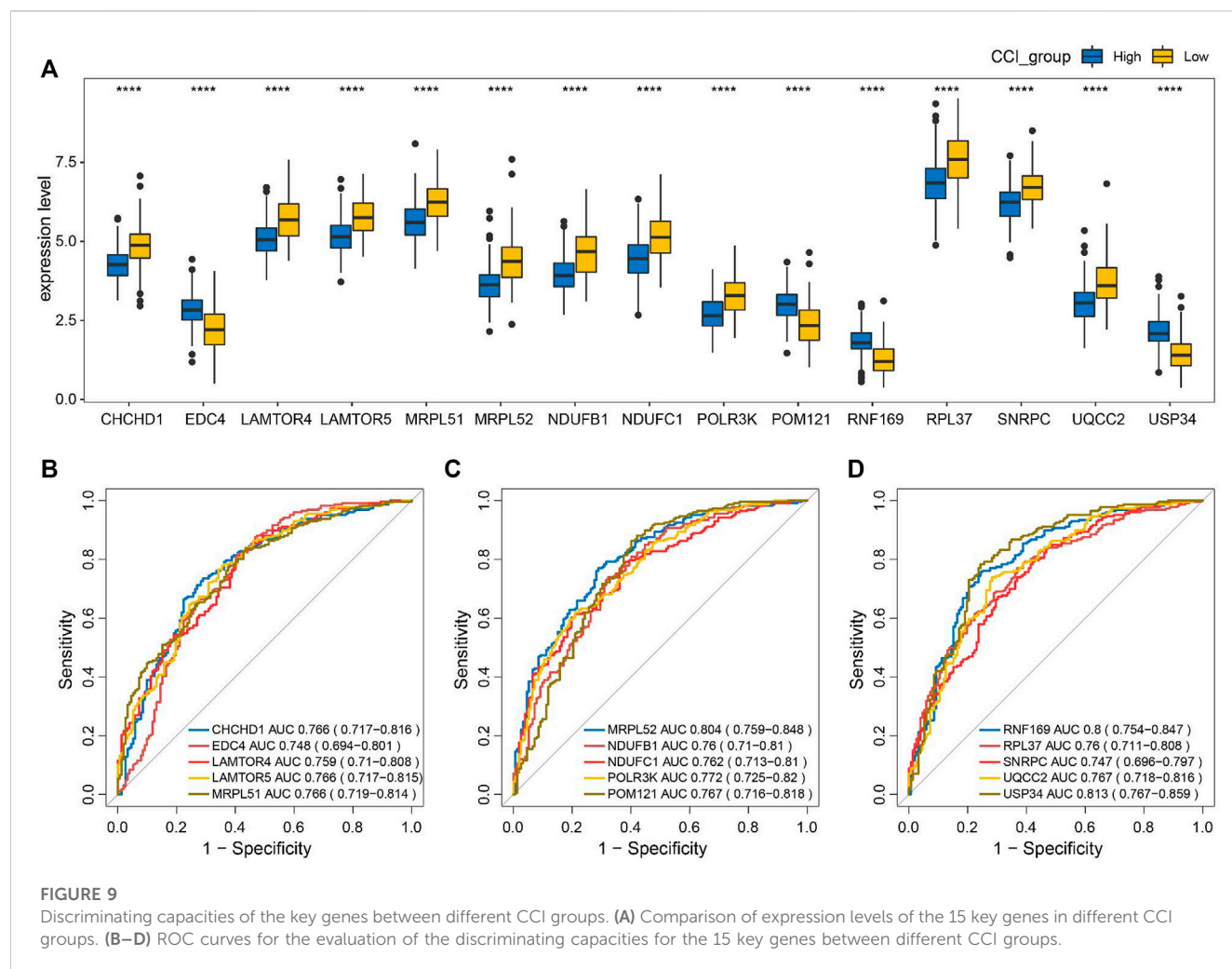
FIGURE 9
Discriminating capacities of the key genes between different CCI groups. **(A)** Comparison of expression levels of the 15 key genes in different CCI groups. **(B–D)** ROC curves for the evaluation of the discriminating capacities for the 15 key genes between different CCI groups.

4 Discussions

Ovarian cancer is the most common malignancy among all gynecological cancers, and its histopathological classification mainly includes epithelial ovarian cancer, sex cord-stromal tumors, and germ cell tumors, among which epithelial tumors are the most common, accounting for more than 90%. Ovarian cancer is a malignant tumor of the female reproductive system with the highest mortality rate, and the current treatment is mainly surgical treatment supplemented by radiotherapy, chemotherapy, immunotherapy, etc., but the overall survival rate of ovarian cancer is still poor. Biorhythms play an important role in regulating numerous physiological activities of mammals, including cell growth, secretion, and metabolism (Reppert and Weaver, 2002). Numerous studies have confirmed a close link between biorhythm and cell cycle and apoptosis (Fu et al., 2002). Abnormal expression of cell cycle-related genes will cause cell cycle disorder, resulting in an imbalance of cell growth and apoptosis, which can lead to the occurrence of cancer (Viallard et al., 2001). The ovaries have specific

cyclical activities, which are closely related to human reproductive health, but the biorhythm and the occurrence and regulation mechanism of ovarian cancer are not clear.

The change in the gene dose of the circadian clock and loss of control of gene doses in the linked transcription-translation feedback loop lead to disruption of the circadian clock, and thus develop into tumors (Lee et al., 2011). The ovary is one of the important reproductive organs of mammals, and the rhythmic expression of circadian clock genes is found in the intact ovary (Fahrenkrug et al., 2006). In infertility studies, it has been found that the reproductive cycle of the ovaries is affected by circadian rhythms, and their rhythms are coordinated and synchronized by neurological and endocrine tissues (Gallego et al., 2006; Etchegaray et al., 2009). This coordination is facilitated by gene expression and cellular physiology at all levels of the hypothalamic-pituitary-ovary (HPO) axis. The expression of circadian clock genes has been observed in the endocrine regulatory axis of the ovaries, and the circadian clocks at all levels coordinated and synchronized with each other to maintain normal reproductive behavior (Alvarez et al., 2008). In peripheral ovarian tissues, changes in the timing of circadian clock gene expression may be the result of hormonal imbalances associated with polycystic ovary syndrome (PCOS) (Amaral et al., 2014). The reproductive function of the ovaries is
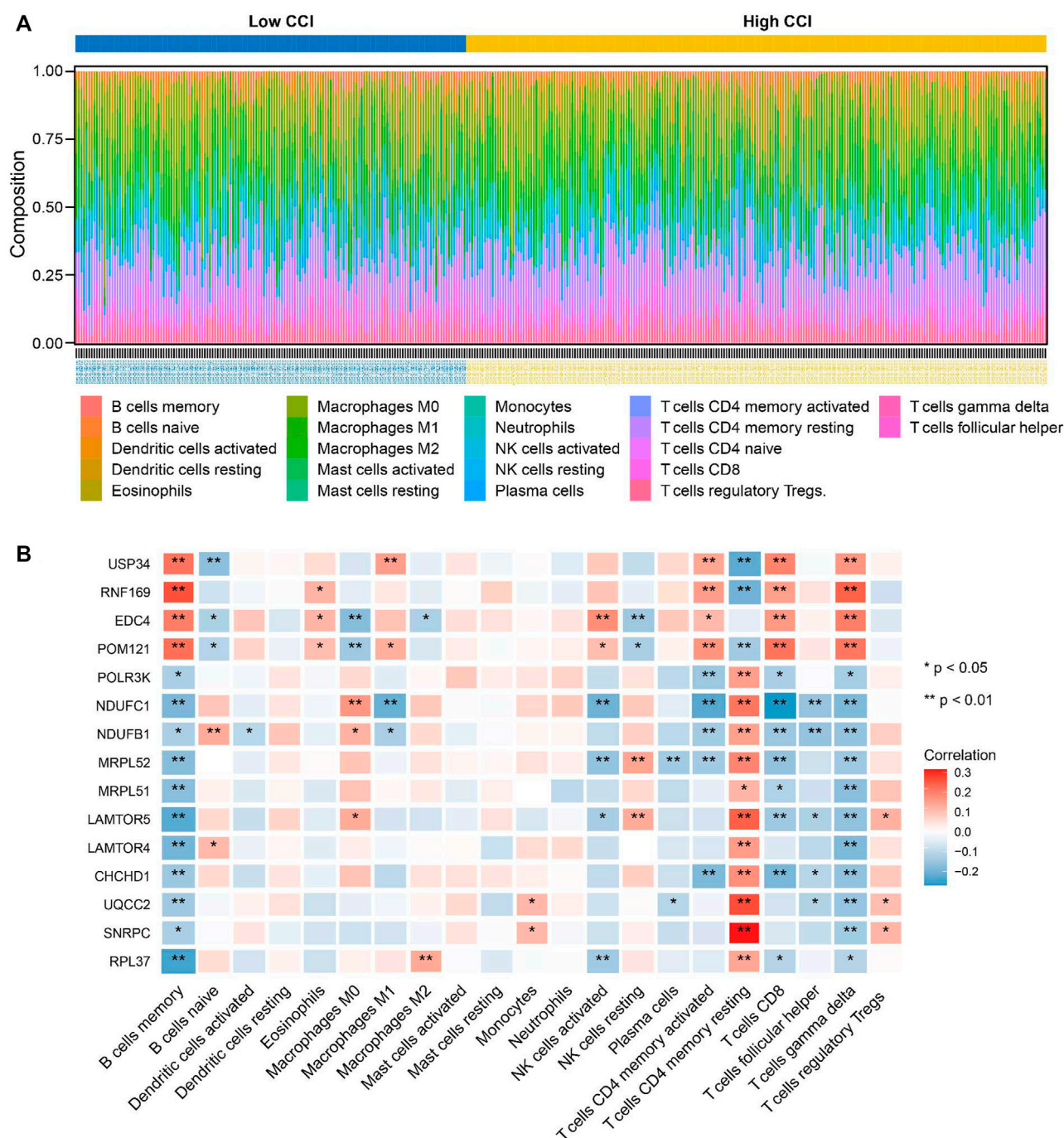
FIGURE 10
Relationship between key genes of CC and the abundance of immune infiltration cells. **(A)** The landscape of tumor immune infiltration regarding the CCI group in OC. **(B)** Heatmap showing the correlation matrix of the 15 key genes and relative abundance of 22 immune cell types. Red denotes the positive correlation, while blue denotes the negative correlation.

richly related to biorhythms, so the relationship between ovarian cancer and biorhythms is of great interest to us.

We comprehensively identified 15 key genes by elucidating the relationship between circadian clock control genes and survival, tumor stage, and subtype in ovarian cancer patients using various statistical and bioinformatics methods. These genes play a role in

translation, mitochondrial translation initiation, extension, termination, and protein metabolism pathways that influence tumor progression and development. For example, *POM121* forms the core component of the nuclear pore complex, mediating the transport of molecules in and out of the nucleus, and blocking the function of *POM121* can inhibit the nuclear

localization of transcription factors including *MYC* and *E2F1*, thereby reducing the growth rate of prostate cancer tumors (Rodriguez-Bravo et al., 2018). The expression levels of mRNA and protein of *POM121* in colon cancer, oral squamous cell carcinoma (Ma et al., 2019), lung cancer (Zhang et al., 2020), and laryngeal cancer tissues are also found to be higher than in adjacent tissues (Zhao et al., 2020). *EDC4* is a well-known regulator of mRNA decapping, which was related to mRNA decapping, genome stability, and sensitivity of drugs. *EDC4* plays a key role in homologous recombination by stimulating end resection at double-strand breaks. Lack-of-function mutations in *EDC4* were detected in breast cancer (Gudkova et al., 2012). *SNRPC* is one of the specific protein components encoding U1 small ribonucleoprotein (snRNP) granules and is upregulated and prognostically related in liver cancer patients (Hernandez et al., 2018). The same trend of changes in the *POM121*, *EDC4*, and *SNRPC* genes was also found in ovarian cancer.

The linkage between mitochondrial alteration and cancer has been uncovered in multiple cancer types including liver (Huang et al., 2016), lung (Qi et al., 2020), colon (Tailor et al., 2014), breast (Zhao et al., 2013), and ovarian (Kingnate et al., 2018) cancers. In this study, several key genes of CC such as *CHCHD1*, *MRPL52*, *MRPL51*, *UQCC2*, *NDUFB1*, and *NDUFC1* are all important regulators in the mitochondrial respiratory chain, which had significant effects on the expression of mitochondrially encoded proteins. For example, MRPs play an important role in the synthesis of the basic subunits of the oxidative phosphorylation (OXPHOS) complex, *MRPL51*, and *MRPL52* may interact with Mhr1 to regulate mtDNA repair (Cai et al., 2021). The protein encoded by *UQCC2* affects insulin secretion and mitochondrial ATP production by regulating mitochondrial respiratory chain activity. *NDUFB1* and *NDUFC1* are auxiliary subunits of NADH dehydrogenase (complex I), responsible for transporting electrons from NADH to the respiratory chain necessary for oxidative phosphorylation. The downregulation of *NDUFC1* expression significantly inhibits the proliferation of hepatoma cells (Sahu et al., 2019) and increases the number of apoptotic cells in liver cancer (Han et al., 2022). We suspect that circadian clock genes may regulate the function of these genes through a transcription-translation feedback loop, thus having a potentially pivotal role in tumorigenesis and development.

Given the circadian rhythm regulation of cancer-related physiological systems such as immune response, cell cycle, and apoptosis, immune therapy may become a promising trend in tumor treatment (Dumbell et al., 2016). The immune system plays a vital role in immune surveillance, with immune cells penetrating the tumor microenvironment and helping to regulate tumor progression. Immune cells are the cellular basis of immunotherapy, and a deeper understanding of immune infiltration in the tumor immune microenvironment can reveal the underlying molecular mechanism and provide new strategies for improving the efficacy of immunotherapy (Wu J. et al., 2019). Immunoinvasive studies have shown that the tumor immune microenvironment plays a key role in cancer progression and influences clinical outcomes in cancer patients (Zhang and Zhang, 2020). Studies have shown that the induction of immune response and the regulation of autoimmunity are affected by the

regulation time of the immune system (Sutton et al., 2017). Studies have also shown that the body's immune system fluctuates rhythmically with circadian rhythms (Wang et al., 2023), and the body's internal clock (biorhythm) has an important impact on the ability of immune cells to recognize cancer cells and promote their clearance, a discovery that may be used to improve the effectiveness of cancer treatment.

The composition of immune cells in the tumor microenvironment also affects cancer prognosis (Ock et al., 2016; Guo et al., 2022). Studies have shown a strong correlation between the circadian clock and immune cells (Zhou et al., 2020), and in many tumors, and downregulation of core clock genes (PER1, PER2, PER3, CRY1, and CRY2) expression is significantly associated with T cell failure and upregulation of immunosuppressive molecules (Wu Y. et al., 2019). Our study found that CCI in OC was negatively correlated with CD8$^+$ T cell abundance, which coincided with the change in response intensity of CD8 T cells during vaccination with time (Nobis et al., 2019). At the same time, melanoma studies have found dendritic cells and CD8$^+$ T cells to exert circadian rhythm anti-tumor functions and control the volume of melanoma (Qian et al., 2021). The above observations suggest that CCI may be a candidate for future immunotherapy in OC. Interestingly, further analysis found that the selected key genes were closely linked to immune infiltration. Studies have shown that *POM121* inhibits macrophage inflammatory response by reducing NF-κB phosphorylated P65 nucleation, which is associated with tumor lymph node metastasis staging (Zhang et al., 2020), and the expression of *USP34* in diffuse large B-cell lymphoma is significantly higher than that in reactive lymphoid hyperplasia (Li et al., 2018). *USP34* can also promote the proliferation and migration of pancreatic cancer cells by upregulating p-AKT and p-PKC proteins (Gu et al., 2019). Similarly, the *EDC4* and Dcp1a complexes are involved in the post-transcriptional regulation of *IL-6*, thereby affecting the function of immune cells (Seto et al., 2015). In this study, the expression levels of 15 key genes were analyzed concerning the proportion of immune cell types. As a result, 11 key genes were negatively associated with T cell gamma delta, most of which were also negatively correlated with the infiltration of CD8 T cell. The left four genes (*USP34*, *RNF169*, *EDC4*, and *POM121* were positively correlated with the infiltration of CD8 T cell, T cell gamma delta, B cells memory, and T cells CD4 memory activated). These data in turn strengthened the utility of CCI as a potential immunotherapy target. Consequently, the combination of drug treatment timing and circadian rhythm may be a new strategy to improve the therapeutic responses and improve the survival rate of patients with OC. Therefore, the circadian rhythm of cancer immune surveillance is not only critical for controlling tumor size but can also be used to guide scientists and doctors to administer cancer immunotherapy to patients at the right point in time, pointing to a new direction for cancer treatment. The expression of CCI-related genes and the identified key CC genes are expected to be taken into consideration in clinical practice for determining a personalized treatment regimen in patients with OC, and the CCI or key CC genes-based drugs and small compounds might be further designed for targeted therapy.

Upstream transcription factors or miRNAs of the key CC genes were comprehensively predicted, and well-known oncogenic transcription factors were involved. Genomic alterations revealed frequent somatic mutations of *RNF169*, *EDC4*, *CHCHD1*, *MRPL51*, *UQCC2*, and *USP34*, which were poorly studied in OC, thus providing new insights into the molecular regulation of these genes from the genomic layer.

In summary, we comprehensively identified 15 key CC genes with clinical implications, which not only improve the understanding of the critical role of CC in tumor initiation and progression, as well as the tumor immune microenvironment but also provide novel insight for future biomarkers or molecular classification development.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://portal.gdc.cancer.gov/.

Author contributions

LZ, YT, and JY: methodology, formal analysis, visualization, writing-original draft. FL and XL: investigation, and validation. YZ and JC: conceptualization, software, methodology, writing review and editing. All authors contributed to the article and approved the submitted version.

Acknowledgments

We acknowledge the Cancer Genome Atlas (TCGA) database for data sharing.

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.2023.1208132/full#supplementary-material

References

Aiello, I., Fedele, M. L. M., Roman, F., Marpegan, L., Caldart, C., Chiesa, J. J., et al. (2020). Circadian disruption promotes tumor-immune microenvironment remodeling favoring tumor cell proliferation. *Sci. Adv.* 6 (42), eaaz4530. doi:10.1126/sciadv.aaz4530

Alvarez, J. D., Hansen, A., Ord, T., Bebas, P., Chappell, P. E., Giebultowicz, J. M., et al. (2008). The circadian clock protein BMAL1 is necessary for fertility and proper testosterone production in mice. *J. Biol. Rhythms* 23 (1), 26–36. doi:10.1177/0748730407311254

Amaral, F. G., Castrucci, A. M., Cipolla-Neto, J., Poletini, M. O., Mendez, N., Richter, H. G., et al. (2014). Environmental control of biological rhythms: Effects on development, fertility and metabolism. *J. Neuroendocrinol.* 26 (9), 603–612. doi:10.1111/jne.12144

Badiu, C. (2003). Genetic clock of biologic rhythms. *J. Cell. Mol. Med.* 7 (4), 408–416. doi:10.1111/j.1582-4934.2003.tb00243.x

Battaglin, F., Chan, P., Pan, Y., Soni, S., Qu, M., Spiller, E. R., et al. (2021). Clocking cancer: The circadian clock as a target in cancer therapy. *Oncogene* 40 (18), 3187–3200. doi:10.1038/s41388-021-01778-6

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

Briukhovetska, D., Dorr, J., Endres, S., Libby, P., Dinarello, C. A., and Kobold, S. (2021). Interleukins in cancer: From biology to therapy. *Nat. Rev. Cancer* 21 (8), 481–499. doi:10.1038/s41568-021-00363-z

Cai, J., Zhou, M., and Xu, J. (2021). N6-methyladenosine (m6A) RNA methylation regulator SNRPC is a prognostic biomarker and is correlated with immunotherapy in hepatocellular carcinoma. *World J. Surg. Oncol.* 19 (1), 241. doi:10.1186/s12957-021-02354-8

Cash, E., Sephton, S., Woolley, C., Elbehi, A. M., R, I. A., Ekine-Afolabi, B., et al. (2021). The role of the circadian clock in cancer hallmark acquisition and immune-based cancer therapeutics. *J. Exp. Clin. Cancer Res.* 40 (1), 119. doi:10.1186/s13046-021-01919-5

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. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019

Cho, K. R., and Shih Ie, M. (2009). Ovarian cancer. *Annu. Rev. Pathol.* 4, 287–313. doi:10.1146/annurev.pathol.4.110807.092246

Dakup, P., and Gaddameedhi, S. (2017). Impact of the circadian clock on UV-induced DNA damage response and photocarcinogenesis. *Photochem Photobiol.* 93 (1), 296–303. doi:10.1111/php.12662

Dumbell, R., Matveeva, O., and Oster, H. (2016). Circadian clocks, stress, and immunity. *Front. Endocrinol. (Lausanne)* 7, 37. doi:10.3389/fendo.2016.00037

Dun, A., Zhao, X., Jin, X., Wei, T., Gao, X., Wang, Y., et al. (2020). Association between night-shift work and cancer risk: Updated systematic review and meta-analysis. *Front. Oncol.* 10, 1006. doi:10.3389/fonc.2020.01006

Etchegaray, J. P., Lee, C., Wade, P. A., and Reppert, S. M. (2003). Rhythmic histone acetylation underlies transcription in the mammalian circadian clock. *Nature* 421 (6919), 177–182. doi:10.1038/nature01314

Etchegaray, J. P., Machida, K. K., Noton, E., Constance, C. M., Dallmann, R., Di Napoli, M. N., et al. (2009). Casein kinase 1 delta regulates the pace of the mammalian circadian clock. *Mol. Cell Biol.* 29 (14), 3853–3866. doi:10.1128/MCB.00338-09

Fahrenkrug, J., Georg, B., Hannibal, J., Hindersson, P., and Gras, S. (2006). Diurnal rhythmicity of the clock genes Per1 and Per2 in the rat ovary. *Endocrinology* 147 (8), 3769–3776. doi:10.1210/en.2006-0305

Fu, L., Pelicano, H., Liu, J., Huang, P., and Lee, C. (2002). The circadian gene Period2 plays an important role in tumor suppression and DNA damage response *in vivo*. *Cell* 111 (1), 41–50. doi:10.1016/s0092-8674(02)00961-3

Gallego, M., Kang, H., and Virshup, D. M. (2006). Protein phosphatase 1 regulates the stability of the circadian protein PER2. *Biochem. J.* 399 (1), 169–175. doi:10.1042/BJ20060678

Gu, Z., Lin, C., Hu, J., Xia, J., Wei, S., and Gao, D. (2019). USP34 regulated human pancreatic cancer cell survival via AKT and PKC pathways. *Biol. Pharm. Bull.* 42 (4), 573–579. doi:10.1248/bpb.b18-00646

Gudkova, D., Panasyuk, G., Nemazanyy, I., Zhyvoloup, A., Monteil, P., Filonenko, V., et al. (2012). EDC4 interacts with and regulates the dephospho-CoA kinase activity of CoA synthase. *FEBS Lett.* 586 (20), 3590–3595. doi:10.1016/j.febslet.2012.08.033

Guo, C., Tang, Y., Li, Q., Yang, Z., Guo, Y., Chen, C., et al. (2023). Deciphering the immune heterogeneity dominated by natural killer cells with prognostic and therapeutic implications in hepatocellular carcinoma. *Comput. Biol. Med.* 158, 106872. doi:10.1016/j.compbiomed.2023.106872

Guo, C., Tang, Y., Yang, Z., Li, G., and Zhang, Y. (2022). Hallmark-guided subtypes of hepatocellular carcinoma for the identification of immune-related gene classifiers in the prediction of prognosis, treatment efficacy, and drug candidates. *Front. Immunol.* 13, 958161. doi:10.3389/fimmu.2022.958161

Guo, C., Tang, Y., Zhang, Y., and Li, G. (2021). Mining TCGA data for key biomarkers related to immune microenvironment in endometrial cancer by immune score and weighted correlation network analysis. *Front. Mol. Biosci.* 8, 645388. doi:10.3389/fmolb.2021.645388

Hadadi, E., and Acloque, H. (2021). Role of circadian rhythm disorders on EMT and tumour-immune interactions in endocrine-related cancers. *Endocr. Relat. Cancer* 28 (2), R67–R80. doi:10.1530/ERC-20-0390

Han, F., Liu, J., Chu, H., Cao, D., Wu, J., Fu, H., et al. (2022). Knockdown of NDUFC1 inhibits cell proliferation, migration, and invasion of hepatocellular carcinoma. *Front. Oncol.* 12, 860084. doi:10.3389/fonc.2022.860084

He, L., Fan, Y., Zhang, Y., Tu, T., Zhang, Q., Yuan, F., et al. (2022). Single-cell transcriptomic analysis reveals circadian rhythm disruption associated with poor prognosis and drug-resistance in lung adenocarcinoma. *J. Pineal Res.* 73 (1), e12803. doi:10.1111/jpi.12803

Hernandez, G., Ramirez, M. J., Minguillon, J., Quiles, P., Ruiz de Garibay, G., Aza-Carmona, M., et al. (2018). Decapping protein EDC4 regulates DNA repair and phenocopies BRCA1. *Nat. Commun.* 9 (1), 967. doi:10.1038/s41467-018-03433-3

Hirayama, J., Sahar, S., Grimaldi, B., Tamaru, T., Takamatsu, K., Nakahata, Y., et al. (2007). CLOCK-mediated acetylation of BMAL1 controls circadian function. *Nature* 450 (7172), 1086–1090. doi:10.1038/nature06394

Huang, Q., Zhan, L., Cao, H., Li, J., Lyu, Y., Guo, X., et al. (2016). Increased mitochondrial fission promotes autophagy and hepatocellular carcinoma cell survival through the ROS-modulated coordinated regulation of the NFKB and TP53 pathways. *Autophagy* 12 (6), 999–1014. doi:10.1080/15548627.2016.1166318

Huang, R.-Y., and Odunsi, K. (2017). Abstract 606: Targeting multiple immune checkpoints and their ligands using a humanized mouse model of ovarian cancer. *Cancer Res.* 77 (13), 606. doi:10.1158/1538-7445.AM2017-606

Huang, R. Y., Francois, A., McGray, A. R., Miliotto, A., and Odunsi, K. (2017). Compensatory upregulation of PD-1, LAG-3, and CTLA-4 limits the efficacy of single-agent checkpoint blockade in metastatic ovarian cancer. *Oncoimmunology* 6 (1), e1249561. doi:10.1080/2162402X.2016.1249561

Jagannath, A., Taylor, L., Wakaf, Z., Vasudevan, S. R., and Foster, R. G. (2017). The genetics of circadian rhythms, sleep and health. *Hum. Mol. Genet.* 26 (R2), R128–R138. doi:10.1093/hmg/ddx240

Kingnate, C., Charoenkwan, K., Kumfu, S., Chattipakorn, N., and Chattipakorn, S. C. (2018). Possible roles of mitochondrial dynamics and the effects of pharmacological interventions in chemoresistant ovarian cancer. *EBioMedicine* 34, 256–266. doi:10.1016/j.ebiom.2018.07.026

Kinouchi, K., and Sassone-Corsi, P. (2020). Metabolic rivalry: Circadian homeostasis and tumorigenesis. *Nat. Rev. Cancer* 20 (11), 645–661. doi:10.1038/s41568-020-0291-9

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559

Lee, H. M., Chen, R., Kim, H., Etchegaray, J. P., Weaver, D. R., and Lee, C. (2011). The period of the circadian oscillator is primarily determined by the balance between casein kinase 1 and protein phosphatase 1. *Proc. Natl. Acad. Sci. U. S. A.* 108 (39), 16451–16456. doi:10.1073/pnas.1107178108

LeGates, T. A., Fernandez, D. C., and Hattar, S. (2014). Light as a central modulator of circadian rhythms, sleep and affect. *Nat. Rev. Neurosci.* 15 (7), 443–454. doi:10.1038/nrn3743

Li, C., Huang, L., Lu, H., Wang, W., Chen, G., Gu, Y., et al. (2018). Expression and clinical significance of ubiquitin-specific-processing protease 34 in diffuse large B-cell lymphoma. *Mol. Med. Rep.* 18 (5), 4543–4554. doi:10.3892/mmr.2018.9447

Ma, H., Li, L., Jia, L., Gong, A., Wang, A., Zhang, L., et al. (2019). POM121 is identified as a novel prognostic marker of oral squamous cell carcinoma. *J. Cancer* 10 (19), 4473–4480. doi:10.7150/jca.33368

Morales-Santana, S., Morell, S., Leon, J., Carazo-Gallego, A., Jimenez-Lopez, J. C., and Morell, M. (2019). An overview of the polymorphisms of circadian genes associated with endocrine cancer. *Front. Endocrinol. (Lausanne)* 10, 104. doi:10.3389/fendo.2019.00104

Nader, N., Chrousos, G. P., and Kino, T. (2009). Circadian rhythm transcription factor CLOCK regulates the transcriptional activity of the glucocorticoid receptor by acetylating its hinge region lysine cluster: Potential physiological implications. *FASEB J.* 23 (5), 1572–1583. doi:10.1096/fj.08-117697

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Nobis, C. C., Dubeau Laramee, G., Kervezee, L., Maurice De Sousa, D., Labrecque, N., and Cermakian, N. (2019). The circadian clock of CD8 T cells modulates their early response to vaccination and the rhythmicity of related signaling pathways. *Proc. Natl. Acad. Sci. U. S. A.* 116 (40), 20077–20086. doi:10.1073/pnas.1905080116

Ock, C. Y., Keam, B., Kim, S., Lee, J. S., Kim, M., Kim, T. M., et al. (2016). Pan-cancer immunogenomic perspective on the tumor microenvironment based on PD-L1 and CD8 T-cell infiltration. *Clin. Cancer Res.* 22 (9), 2261–2270. doi:10.1158/1078-0432.CCR-15-2834

Peyric, E., Moore, H. A., and Whitmore, D. (2013). Circadian clock regulation of the cell cycle in the zebrafish intestine. *PLoS One* 8 (8), e73209. doi:10.1371/journal.pone.0073209

Qi, M., Dai, D., Liu, J., Li, Z., Liang, P., Wang, Y., et al. (2020). AIM2 promotes the development of non-small cell lung cancer by modulating mitochondrial dynamics. *Oncogene* 39 (13), 2707–2723. doi:10.1038/s41388-020-1176-9

Qian, D. C., Kleber, T., Brammer, B., Xu, K. M., Switchenko, J. M., Janopaul-Naylor, J. R., et al. (2021). Effect of immunotherapy time-of-day infusion on overall survival among patients with advanced melanoma in the USA (MEMOIR): A propensity score-matched analysis of a single-centre, longitudinal study. *Lancet Oncol.* 22 (12), 1777–1786. doi:10.1016/S1470-2045(21)00546-5

Reppert, S. M., and Weaver, D. R. (2002). Coordination of circadian timing in mammals. *Nature* 418 (6901), 935–941. doi:10.1038/nature00965

Rodriguez-Bravo, V., Pippa, R., Song, W. M., Carceles-Cordon, M., Dominguez-Andres, A., Fujiwara, N., et al. (2018). Nuclear pores promote lethal prostate cancer by increasing pom121-driven E2F1, MYC, and AR nuclear import. *Cell* 174 (5), 1200–1215. doi:10.1016/j.cell.2018.07.015

Sahu, P. K., Salim, S., Pp, M., Chauhan, S., and Tomar, R. S. (2019). Reverse genetic analysis of yeast YPR099C/MRPL51 reveals a critical role of both overlapping ORFs in respiratory growth and MRPL51 in mitochondrial DNA maintenance. *FEMS Yeast Res.* 19 (6), foz056. doi:10.1093/femsyr/foz056

Seto, E., Yoshida-Sugitani, R., Kobayashi, T., and Toyama-Sorimachi, N. (2015). The assembly of EDC4 and Dcp1a into processing bodies is critical for the translational regulation of IL-6. *PLoS One* 10 (5), e0123223. doi:10.1371/journal.pone.0123223

Shen, H., Cook, K., Gee, H. E., and Hau, E. (2020). Hypoxia, metabolism, and the circadian clock: New links to overcome radiation resistance in high-grade gliomas. *J. Exp. Clin. Cancer Res.* 39 (1), 129. doi:10.1186/s13046-020-01639-2

Sulli, G., Lam, M. T. Y., and Panda, S. (2019). Interplay between circadian clock and cancer: New Frontiers for cancer treatment. *Trends Cancer* 5 (8), 475–494. doi:10.1016/j.trecan.2019.07.002

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

Sutton, C. E., Finlay, C. M., Raverdeau, M., Early, J. O., DeCourcey, J., Zaslona, Z., et al. (2017). Loss of the molecular clock in myeloid cells exacerbates T cell-mediated CNS autoimmune disease. *Nat. Commun.* 8 (1), 1923. doi:10.1038/s41467-017-02111-0

Tailor, D., Hahm, E. R., Kale, R. K., Singh, S. V., and Singh, R. P. (2014). Sodium butyrate induces DRP1-mediated mitochondrial fusion and apoptosis in human colorectal cancer cells. *Mitochondrion* 16, 55–64. doi:10.1016/j.mito.2013.10.004

Tang, Y., Guo, C., Chen, C., and Zhang, Y. (2022a). Characterization of cellular senescence patterns predicts the prognosis and therapeutic response of hepatocellular carcinoma. *Front. Mol. Biosci.* 9, 1100285. doi:10.3389/fmolb.2022.1100285

Tang, Y., Guo, C., Yang, Z., Wang, Y., Zhang, Y., and Wang, D. (2022b). Identification of a tumor immunological phenotype-related gene signature for predicting prognosis, immunotherapy efficacy, and drug candidates in hepatocellular carcinoma. *Front. Immunol.* 13, 862527. doi:10.3389/fimmu.2022.862527

Tang, Y., Zhang, Y., and Hu, X. (2020). Identification of potential hub genes related to diagnosis and prognosis of hepatitis B virus-related hepatocellular carcinoma via integrated bioinformatics analysis. *Biomed. Res. Int.* 2020, 4251761. doi:10.1155/2020/4251761

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 (1), W98–W102. doi:10.1093/nar/gkx247

Tokunaga, H., Takebayashi, Y., Utsunomiya, H., Akahira, J., Higashimoto, M., Mashiko, M., et al. (2008). Clinicopathological significance of circadian rhythm-related gene expression levels in patients with epithelial ovarian cancer. *Acta Obstet. Gynecol. Scand.* 87 (10), 1060–1070. doi:10.1080/00016340802348286

Udo, R., Hamada, T., Horikawa, K., Iwahana, E., Miyakawa, K., Otsuka, K., et al. (2004). The role of Clock in the plasticity of circadian entrainment. *Biochem. Biophys. Res. Commun.* 318 (4), 893–898. doi:10.1016/j.bbrc.2004.04.113

Viallard, J. F., Lacombe, F., Belloc, F., Pellegrin, J. L., and Reiffers, J. (2001). Molecular mechanisms controlling the cell cycle: Fundamental aspects and implications for oncology. *Cancer Radiother.* 5 (2), 109–129. doi:10.1016/s1278-3218(01)00087-7

Wang, C., Barnoud, C., Cenerenti, M., Sun, M., Caffa, I., Kizil, B., et al. (2023). Dendritic cells direct circadian anti-tumour immune responses. *Nature* 614 (7946), 136–143. doi:10.1038/s41586-022-05605-0

Wu, J., Chen, J., Feng, Y., Tian, H., and Chen, X. (2019a). Tumor microenvironment as the "regulator" and "target" for gene therapy. *J. Gene Med.* 21 (7), e3088. doi:10.1002/jgm.3088

Wu, Y., Tao, B., Zhang, T., Fan, Y., and Mao, R. (2019b). Pan-cancer analysis reveals disrupted circadian clock associates with T cell exhaustion. *Front. Immunol.* 10, 2451. doi:10.3389/fimmu.2019.02451

Xuan, W., Khan, F., James, C. D., Heimberger, A. B., Lesniak, M. S., and Chen, P. (2021). Circadian regulation of cancer cell and tumor microenvironment crosstalk. *Trends Cell Biol.* 31 (11), 940–950. doi:10.1016/j.tcb.2021.06.008

Yang, S. L., Ren, Q. G., Wen, L., Hu, J. L., and Wang, H. Y. (2017). Research progress on circadian clock genes in common abdominal malignant tumors. *Oncol. Lett.* 14 (5), 5091–5098. doi:10.3892/ol.2017.6856

Yeh, C. M., Shay, J., Zeng, T. C., Chou, J. L., Huang, T. H., Lai, H. C., et al. (2014). Epigenetic silencing of ARNTL, a circadian gene and potential tumor suppressor in ovarian cancer. *Int. J. Oncol.* 45 (5), 2101–2107. doi:10.3892/ijo.2014.2627

Zhang, S., Zheng, C., Li, D., Bei, C., Zhang, H., Tian, R., et al. (2020). Clinical significance of POM121 expression in lung cancer. *Genet. Test. Mol. Biomarkers* 24 (12), 819–824. doi:10.1089/gtmb.2020.0053

Zhang, Y. Q., Wang, W. Y., Xue, J. X., Xu, Y., Fan, P., Caughey, B. A., et al. (2016). MicroRNA expression profile on solid subtype of invasive lung adenocarcinoma reveals a panel of four miRNAs to Be associated with poor prognosis in Chinese patients. *J. Cancer* 7 (12), 1610–1620. doi:10.7150/jca.14923

Zhang, Y., Tang, Y., Guo, C., and Li, G. (2021). Integrative analysis identifies key mRNA biomarkers for diagnosis, prognosis, and therapeutic targets of HCV-associated hepatocellular carcinoma. *Aging (Albany NY)* 13, 12865–12895. doi:10.18632/aging.202957

Zhang, Y., Yang, Z., Tang, Y., Guo, C., Lin, D., Cheng, L., et al. (2022). Hallmark guided identification and characterization of a novel immune-relevant signature for prognostication of recurrence in stage I–III lung adenocarcinoma. *Genes Dis.* 2022. doi:10.1016/j.gendis.2022.07.005

Zhang, Y., and Zhang, Z. (2020). The history and advances in cancer immunotherapy: Understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell Mol. Immunol.* 17 (8), 807–821. doi:10.1038/s41423-020-0488-6

Zhao, J., Zhang, J., Yu, M., Xie, Y., Huang, Y., Wolff, D. W., et al. (2013). Mitochondrial dynamics regulates migration and invasion of breast cancer cells. *Oncogene* 32 (40), 4814–4824. doi:10.1038/onc.2012.494

Zhao, R., Tang, G., Wang, T., Zhang, L., Wang, W., Zhao, Q., et al. (2020). POM121 is a novel marker for predicting the prognosis of laryngeal cancer. *Histol. Histopathol.* 35 (11), 1285–1293. doi:10.14670/HH-18-267

Zhou, L., Luo, Z., Li, Z., and Huang, Q. (2020). Circadian clock is associated with tumor microenvironment in kidney renal clear cell carcinoma. *Aging (Albany NY)* 12 (14), 14620–14632. doi:10.18632/aging.103509

Frontiers in Molecular Biosciences

Analysis of M2 macrophage-associated risk score signature in pancreatic cancer TME landscape and immunotherapy

Dashuai Yang[1†], Fangrui Zhao[2†], Yang Su[3†], Yu Zhou[1†], Jie Shen[1], Kailiang Zhao[1]* and Youming Ding[1]*

[1]Department of Hepatobiliary Surgery, Renmin Hospital of Wuhan University, Wuhan, China, [2]Department of Oncology, Renmin Hospital of Wuhan University, Wuhan, China, [3]Department of Gastrointestinal Surgery, Tongji Hospital, Tongji Medical College in Huazhong University of Science and Technology, Wuhan, Hubei, China

Background: M2 macrophages perform an influential role in the progression of pancreatic cancer. This study is dedicated to explore the value of M2 macrophage-related genes in the treatment and prognosis of pancreatic cancer.

Methods: RNA-Seq and clinical information were downloaded from TCGA, GEO and ICGC databases. The pancreatic cancer tumour microenvironment was revealed using the CIBERSORT algorithm. Weighted gene co-expression network analysis (WGCNA) was used to detect M2 macrophage-associated gene modules. Univariate Cox regression, Least absolute shrinkage and selection operator (LASSO) regression analysis and multivariate Cox regression were applied to develop the prognostic model. The modelling and validation cohorts were divided into high-risk and low-risk groups according to the median risk score. The nomogram predicting survival was constructed based on risk scores. Correlations between risk scores and tumour mutational load, clinical variables, immune checkpoint blockade, and immune cells were further explored. Finally, potential associations between different risk models and chemotherapeutic agent efficacy were predicted.

Results: The intersection of the WGCNA results from the TCGA and GEO data screened for 317 M2 macrophage-associated genes. Nine genes were identified by multivariate COX regression analysis and applied to the construction of risk models. The results of GSEA analysis revealed that most of these genes were related to signaling, cytokine receptor interaction and immunodeficiency pathways. The high and low risk groups were closely associated with tumour mutational burden, immune checkpoint blockade related genes, and immune

cells. The maximum inhibitory concentrations of metformin, paclitaxel, and rufatinib lapatinib were significantly differences on the two risk groups.

Conclusion: WGCNA-based analysis of M2 macrophage-associated genes can help predict the prognosis of pancreatic cancer patients and may provide new options for immunotherapy of pancreatic cancer.

Background

Pancreatic cancer is one of the worst prognoses of all malignant parenchymal tumours-the 5-year survival rate is only around 9% (Bray et al., 2018; Christenson et al., 2020). Age has been identified as a risk factor for pancreatic cancer.

With the global trend of ageing, the incidence of pancreatic cancer is increasing every year (Siegel et al., 2018; Siegel et al., 2020; Huang and Setiawan, 2022). In recent years, advances have been made in the comprehensive treatment of pancreatic cancer, such as immune checkpoint blockade therapy, which has provided new treatment options for patients with

FIGURE 1
Analysis of immune infiltration in pancreatic cancer. **(A)** 22 immune cell subpopulation in TCGA pancreatic cancer samples. **(B)** Proportional heat map of 22 immune cell in TCGA pancreatic cancer samples. **(C)** Correlation analysis of 22 immune cells.

FIGURE 2
WGCNA Analysis. **(A, C, E)** TCGA database aggregation of gene modules with similar expression models based on the WGNCA algorithm and correlation analysis of modules with immune infiltrating cells. **(B, D, F)** GEO database aggregation of gene modules with similar expression models based on the WGNCA algorithm and correlation analysis of modules with immune infiltrating cells.

chemotherapy-resistant pancreatic cancer (Le et al., 2015; Ullman et al., 2022). However, immunotherapy still requires large randomised prospective studies to confirm its role in improving the prognosis of patients with pancreatic cancer (Fan et al., 2020; Ostios-Garcia et al., 2021; Ullman et al., 2022).

The tumour microenvironment is the internal environment upon which tumour cell genesis, growth and metastasis depend (Yang et al., 2021). The tumour microenvironment in pancreatic cancer consists of a large number of tissue interstitial, immune cell

infiltrates and other components, of which tumour-associated macrophages are the main component. M2-type macrophages are the main type of tumour-associated macrophages, which play an irreplaceable role in functions such as trophic competition, inflammatory response, metabolic changes, tumour metastasis and immunosuppression (Feig et al., 2012; Noy and Pollard, 2014; Cohen et al., 2015; Davies and Taylor, 2015).

However, few existing studies have investigated the potential role of tumour-associated macrophages in the

tumour microenvironment of pancreatic cancer as a mechanism for chemoresistance and immunotherapy in pancreatic cancer (Ip et al., 2017). Consequently, this study was based on the construction of co-expression networks through The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) and International Cancer Genome Consortium (ICGC) databases employing a WGCNA analysis approach to identify prognostic models of macrophage-associated genes in the pancreatic cancer microenvironment. This study systematically investigated the potential mechanisms of the genes in the model and the response of patients in different risk groups to chemotherapy and immunotherapy to provide a practical reference model for individualised clinical treatment of pancreatic cancer.

Data downloading and processing

RNA sequencing (RNA-seq) data, clinical profiles and tumour mutation burden data for pancreatic cancer patients were obtained from the Cancer Genome Atlas data (https://portal.gdc.cancer.gov/repository). Meanwhile, clinical profiles and RNA expression data for pancreatic cancer patients from the GSE71729 database were downloaded from the Gene Expression Omnibus (GEO) repository (nlm.nih.gov/gds/). Gene expression data and prognostic data from the International Cancer Genome Consortium (ICGC) database of pancreatic cancer patients were utilized for model validation (https://dcc.icgc.org/projects/LIRI-JP). Inclusion criteria (Bray et al., 2018): survival time >0 and (Christenson et al., 2020) complete clinical information.

Evaluation of immune cell infiltration

Immune cells in the tumour microenvironment affect tumour progression and treatment efficacy. CIBERSORT predicts the proportion of 22 immune cells in tumour sample expression data based on linear support vector regression principles. Based on results at $p < 0.05$, the proportion of immune cells in patients with pancreatic cancer samples from the TCGA and GEO databases was calculated and the results presented by the ggplot2 R package.

WGCNA

Genes associated with M2 macrophages in pancreatic cancer are identified by an algorithm using weighted gene co-expression network analysis (WGCNA), which is a common analytical method for exploring the relationship between gene sets and the phenotype of interest. The R-based "WGCNA" package was built for co-expression networks of genes in TCGA and GEO, respectively. A proximity matrix was constructed based on the best soft threshold β from 1–20 to match the gene distribution to a connection-based scale-free network. Neighbourhood relationships were then converted into a topological overlap matrix (TOM) and clustered in a chain hierarchy based on the mean of different TOM-based metrics. Similar genes are introduced into the same candidate

modules using a "dynamic tree cutting" algorithm. Correlations between the module signature genes and the phenotypes of interest were analysed using Pearson's correlation test ($p < 0.05$). Finally, the expression of genes in the co-expression modules of WGCNA was performed to correlate the proportion of immune cell infiltration in patients.

Building the model

The results of the WGCNA analysis of the TCGA and GEO databases were used to select the set of genes most relevant to M2 macrophages in the module and to take the intersecting genes of both. The intersecting genes were first integrated with patient survival data from TCGA; then univariate COX regression was used to identify the genes that affected patient survival. Next a penalty function was generated using lasso regression to compress the coefficients of the variables to prevent overfitting of the model. Finally, the results of the multifactorial COX regression analysis were confirmed for M2 macrophage-related genes affecting survival in pancreatic cancer patients.

Risk score = βgene A × expr gene A + βgene B × expr gene B+. . .+ βgene N × expr gene N, expr is the mRNA expression of the pivotal gene and β is the corresponding regression coefficient in multivariate genetic Cox regression analysis.

Model validation

Results based on TCGA multifactorial COX regression analysis were screened from the ICGC database for the appropriate genes and combined with survival times to validate the data from the modelling group. Time-dependent ROC curves were employed to validate the accuracy of risk scores in predicting patient prognosis.

Nomogram

The nomogram provided a visual representation of a patient's prognosis. Based on the risk score and the patient's clinical data a nomogram was constructed to predict the patient's prognosis at 1 year, 2 years and 3 years. Calibration curves (by bootstrap method with 500 resamples) and receiver operating characteristic (ROC) curves were applied to evaluate the nomogram.

Sample tumour mutation burden analysis

The TCGA database provides the raw tumour mutation burden data for the samples. The study first downloaded the original tumour mutation burden for each sample of pancreatic cancer samples and calculated the value of the tumour mutation burden for each sample. Waterfall plots for the high-risk and low-risk groups were plotted by "maftools". In addition, survival curves were plotted between the four subgroups based on the median mutational load of pancreatic cancer patients.

FIGURE 3
Building a risk model **(A)** Venn diagram. **(B)** Log (λ) change curves of regression coefficients. **(C)** Tenfold cross-validation of adjusted parameter choices in lasso regression. Vertical lines are plotted from the best data according to the minimum criterion and 1 standard error criterion. **(D, E)** Results of univariate and multivariate COX regression analyses.
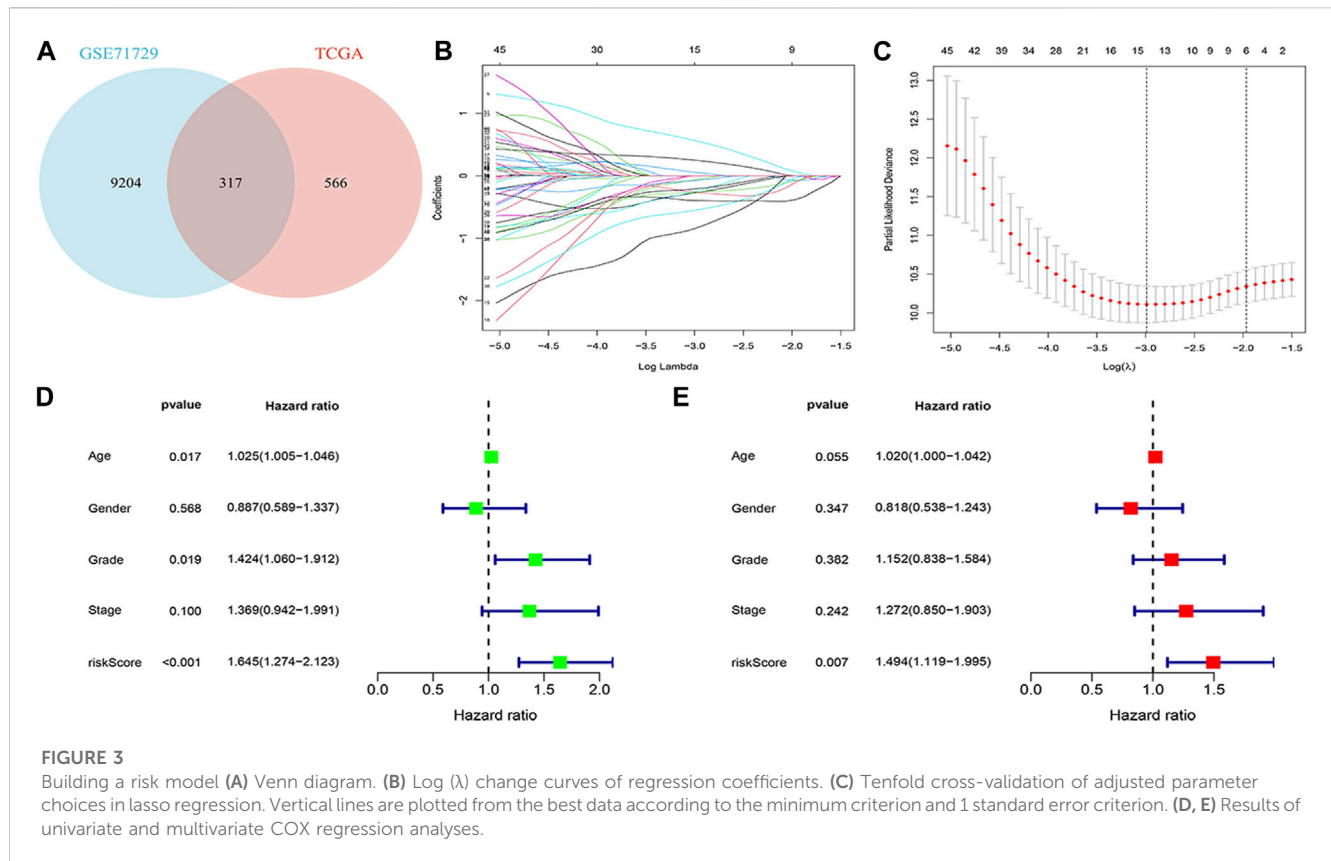
GSEA enrichment analysis

To explore the ranking of genes in the model that lie in the correlation of different phenotypes, functional annotations were explored utilising the c2. cp.kegg.v7.4. symbol and c5. go.v7.4. symbol collections against the Gene Set Enrichment Analysis (GSEA) software. The first six of the annotated results were selected for display and defined as statistically significant with a two-sided *p*-value of <0.05.

GSVA enrichment analysis

To explore the pathways by which genes in the M2 macrophage-associated model may influence the pancreatic cancer tumour microenvironment. The MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb) was used for pathway analysis of M2 macrophage-associated genes.

The relationship between risk models and the tumour microenvironment

To further explore the relationship between the role of M2 macrophage-associated risk models in the immune microenvironment, XCELL, timer, quantitative, MCPcer, EPIC, Sibe sorting and Sibe sort-abs were employed to explore the relationship between risk scores and patient immune function.

Scores for each sample were first assessed based on gene expression using the ESTIMATE algorithm. Secondly, Spearman correlation analysis was applied to evaluate the relationship between risk scores and tumour immune function.

Immunological target prediction

Immunotherapeutic targets play a decisive role in immunotherapy and immune tolerance. Expression of M2 macrophage-associated genes and 47 and immunotherapeutic targets between high and low risk groups were systematically analysed. The immune round of cancer cells determines the efficacy of immunotherapy. The immune panel score (IPS) is an important measure of the immune prototype. The immune score of a sample was integrated by calculating the scores for antigen presentation, effector, suppressor and checkpoint separately.

Drug sensitivity prediction

M2 macrophage-associated models may influence the effectiveness of chemotherapy in patients. Differences in drug sensitivity between high- and low-risk groups were explored based on the "pRRophetic" "ggplot2". Differences in half-maximal inhibitory concentrations (half-inhibitory concentrations) of various chemotherapeutic agents were evaluated between high- and low-risk groups of patients with pancreatic cancer using the Wilcoxon signed-rank test.

Real-time PCR

20 Total RNA from pancreatic cancer tissue and paired paracancer tissue samples was treated with an RNA separator total RNA extraction reagent (Vazyme). The cDNA was synthesized from total RNA using NovoScript® plus an all-in-one first strand cDNA synthesis kit (Novo protein). GAPDH was appllied as an internal control. Primers used for RT-PCR assay are shown in **Additional File 1: Table 1**.

Statistical analysis

The Wilcoxon rank sum test was used as a backup to compare differences between two groups. kruskalWallis test was used to compare differences between three groups and more. Kaplan-Meier method and log-rank test were used for prognostic analysis. All data analysis was done by R (4.1.2, https://www.r-project.org/) software. Bilateral $p < 0.05$ was considered statistically significant.

Results

Patient data

After collating and screening the clinical and expression data of the patients, 172 samples from the TCGA database, 79 from the GEO database GSE71729 and 80 from the ICGC database were included in the study. The median follow-up time for patients with pancreatic cancer in the TCGA, GEO and ICGC databases were 15.61 [interquartile range (IQR): 8.98–22.49] months, 13 (IQR: 6.00–22.00) months, 15.20 (IQR: 8.66–26.46) months respectively.

Tumour microenvironment analysis

The proportions of 22 immune cells were calculated for each sample of pancreatic cancer patients in the TCGA and GEO databases were calculated based on the CIBERSORT algorithm, respectively (Additional File 1: Supplementary Tables S2, S3). As shown in Figure 1A, the row names represent each sample and the different colours of the cylindrical plot represent the proportion of different immune cells. The heat map (Figure 1B) demonstrates the difference in immune infiltration between normal and tumour tissue. The correlation heat map suggests a potential relationship between the 22 immune cells (Figure 1C).

WGCNA

A WGCNA co-expression network was built based on gene expression files (TCGA: 19,819 genes, GEO: 19,014 genes) and immune cell infiltration results. The optimal soft threshold power (TCGA: b = 8,GEO: b = 11) when the scale-free topology index first reached 0.9 was set as the first set of

power values to build the scale-free network (Figures 2A, B). Genes with similar expression patterns were grouped into the same gene module to form a hierarchical clustering tree based on a "dynamic tree cutting" algorithm (module size = 60). Finally, a weighted hierarchical clustering analysis was performed to obtain the clustered gene modules (Figures 2C, D). The highest correlations with M2 macrophages in the TCGA and GEO databases were green and turquoise respectively. The intersecting genes of the two modules were finally identified as the set of M2 macrophage-associated genes for the next analysis(Figures 2E, F).

Building the model

The 317 genes in the TCGA and GEO databases were finally recognised as M2 macrophage-associated genes (Figure 3A). Clinical data and follow-up information of patients were extracted from the TCGA database and merged with the expression of the 317 genes. Sixty genes were screened for association with patient prognosis after univariate COX regression analysis. The results of the Lasso regression were used in the multifactorial COX regression analysis (Figures 3B, C), and the final 9 genes (ABCB4, APOBEC3C, ENPP6, FGFBP2, LIPE, MT2A, OXER1, PLD4, ZNF589) were selected for model construction (Additional File 1: Supplementary Tables S4, S5). Risk scores were calculated for each sample using the risk score formula and the samples were divided into low and high risk groups depending on the median score. Risk score is an independent prognostic element for patients (Figures 3D, E). Protein expression levels in pancreatic cancer patients were explored based on the HPA database. The results suggested that the protein of the target gene is differentially expressed in normal tissues and pancreatic cancer tissues. Meanwhile, model genes were differentially expressed in both tumor tissues and normal tissues (Additional File 2).

Validation of the model

Survival curves classifying each of the nine genes into high and low risk groups based on median expression levels indicated that the expression levels of all nine genes correlated with patient prognosis (Figure 4). The heat map clearly demonstrated the difference in expression of the model genes between the high- and low-risk groups. Scatterplot of risk scores and patient survival revealed a higher proportion of patients with higher risk scores. The results were validated in the validation group (Figure 5).

Correlation of clinical variables

High and low risk groups were correlated with clinical variables. The correlation between high and low risk groups and age, gender, pathological grade and tumour stage, respectively, is shown in Figure 6.

Predicted prognosis nomogram

The risk score was combined with clinical information to construct the nomogram predicting the prognosis of patients at 1, 2 and 3 years to enhance the functionality of the risk score in clinical practice. For example, if the total patient score for the example in the nomogram is 274, the overall probability of patient survival at 1, 2 and 3 years is 0.875,0.606 and 0.539 respectively (Figure 7A). Calibration curves showed stable predictive power of the nomogram (Figure 7B). Time dependent ROC curves indicated 1, 2 and 3 years AUC values of 0.760, 0.781 and 0.802 for the modelling group and 0.759, 0.673 and 0.767 for the validation group (Figures 7C, E), indicating that the model has high predictive ability of the model. Simultaneously, the AUC values for risk scores were higher than for other clinical variables (Figures 7D, F).

GSEA enrichment analysis

GSEA enrichment analysis was performed to explore the possible pathways through which M2 macrophage-associated genes affect the immune microenvironment. The genes were divided into high and low expression groups according to their median expression and the differences in signalling pathways between the two groups were investigated. The KEGG enrichment project indicated that ABCB4 was involved in signalling, cytokine receptor interaction, and cellular value-added signalling pathways, APOBEC3C was linked to immune rejection, cytokine receptor interaction, and gastrointestinal immune signalling pathways, ENPP6 mainly affected cell adhesion, cytokine receptor interaction signalling pathways, FGFBP2 was related to academic signalling, drug metabolism, haematopoietic cell pathways, LIPE affected signaling pathways of calcium signaling, biosynthesis, leucine isoleucine synthesis, MT2A was concerned with signaling pathways of chemical signaling, hematopoietic cell lineage, gastrointestinal immunity, OXER1 was associated with chemical signaling, steroid synthesis signaling pathways, PLD4 was engaged in cell adhesion, chemical signaling, cytokine receptor interaction signaling pathways, ZNF58 impacted signaling pathways of chemical signaling, immunodeficiency, and taste perception (Figure 8).
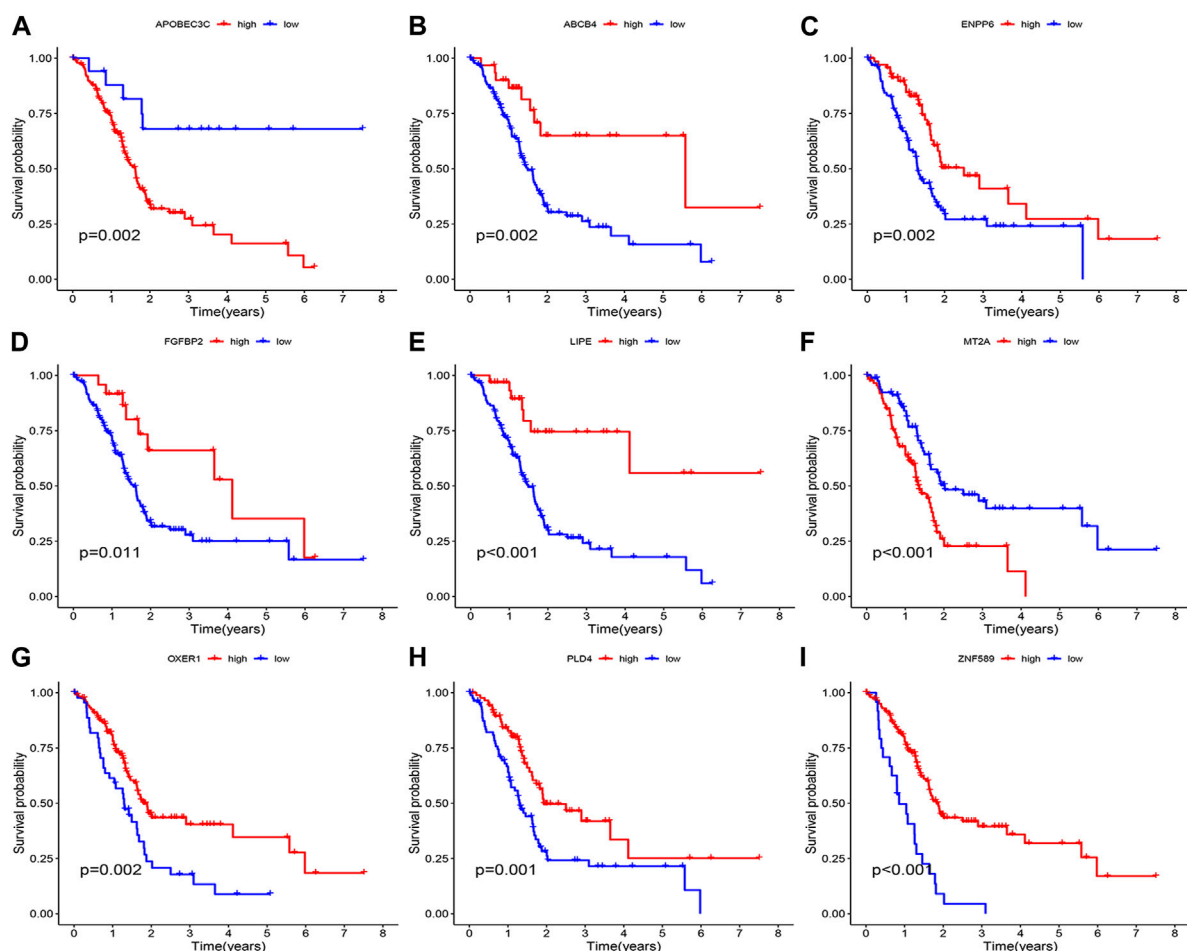
FIGURE 4
Prognostic analysis of the model gene. **(A)** APOBEC3C. **(B)** ABCB4. **(C)** ENPP6. **(D)** FGFBP2. **(E)** LIPE. **(F)** MT2A. **(G)** OXER1. **(H)** PLD4. **(I)** ZNF589.
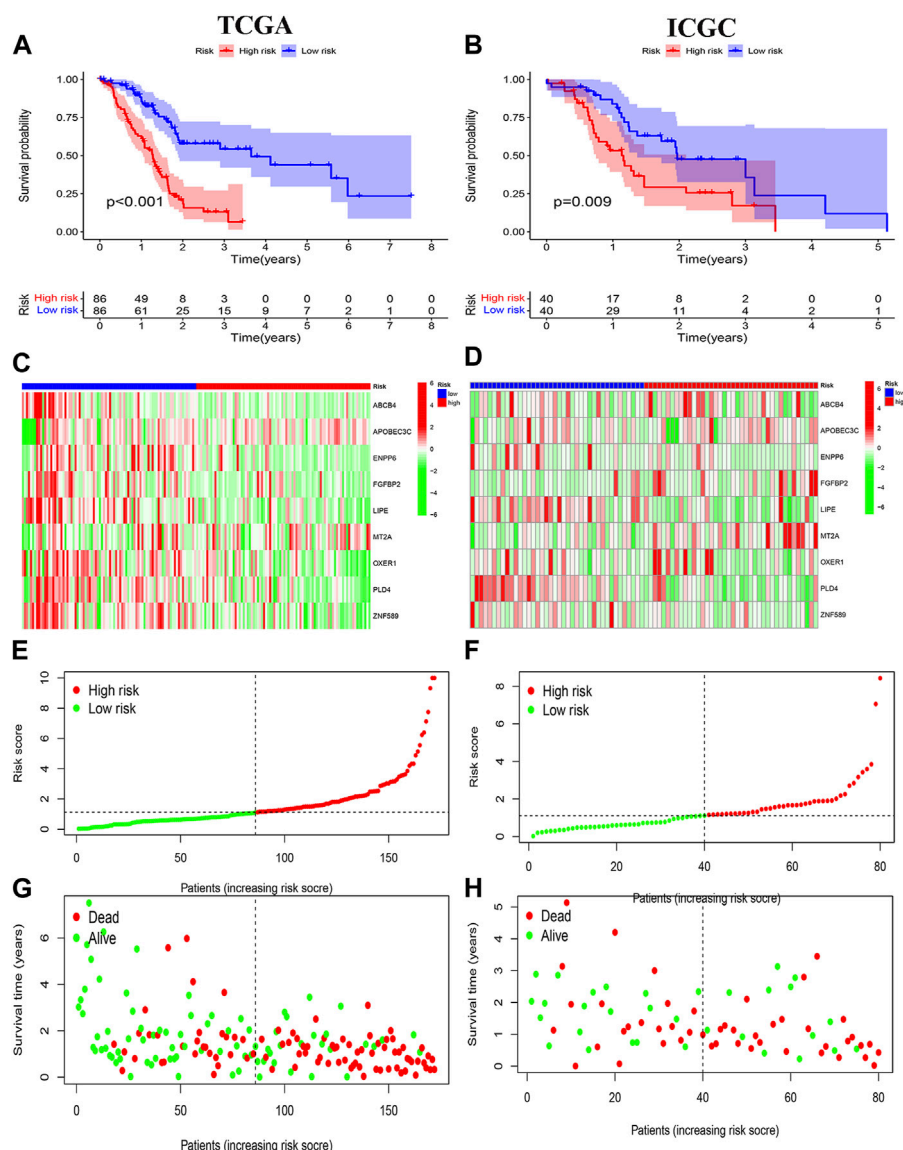
FIGURE 5
Prognostic analysis. **(A, B)** Risk signature survival analysis in TCGA and ICGC databases. **(C, E, G)** Heat plot, risk score plot and scatter plot based on TCGA dataset. **(D, F, H)** Heat plot, risk score plot and scatter plot based on ICGC dataset.

Tumour mutation burden analysis

The tumour mutation burden was first calculated for all samples. Statistically significant differences in tumour mutation burden levels between the two groups (Figure 9A). Correlation analysis between risk score and tumour mutation burden indicated that a higher risk score implied a higher tumour mutation burden (Figure 9B). The sample was immediately divided into a high and low tumour mutation group by median tumour mutation value. There was a significant difference in survival between the two groups ($p < 0.05$). Figures 9E, F showed that this finding was validated in the high-risk and low-risk groups.

The relationship between risk models and the tumour microenvironment

The potential relationship between model genes and the tumour microenvironment was investigated based on Spearman correlation analysis (Figure 10A). Figure 10B indicated that immune scores, stromal score, and ESTIMATE scores were different between the high-risk and low-risk groups ($p < 0.05$). The results of the correlation between risk scores and tumour microenvironment analysed by the four methods CIBERSORT-ABS (Figure 10C), CIBERSORT (Figure 10D), QUANTISEQ (Figure 10E), and XCELL (Figure 10F) immediately afterwards demonstrated the potential of
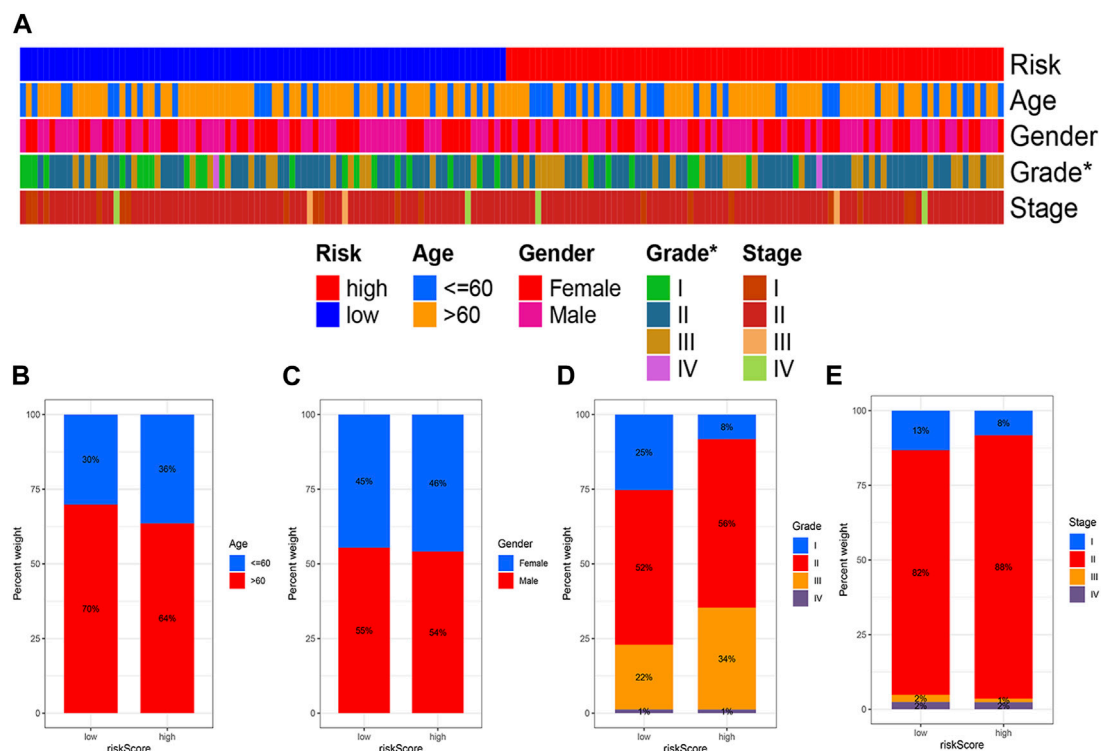
FIGURE 6
Correlation analysis of risk models and clinical variables. **(A)** Heat map showing clinical characteristics and risk scores for each sample. **(B)** Age. **(C)** Gender. **(D)** Grade. **(E)** Stage. (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).
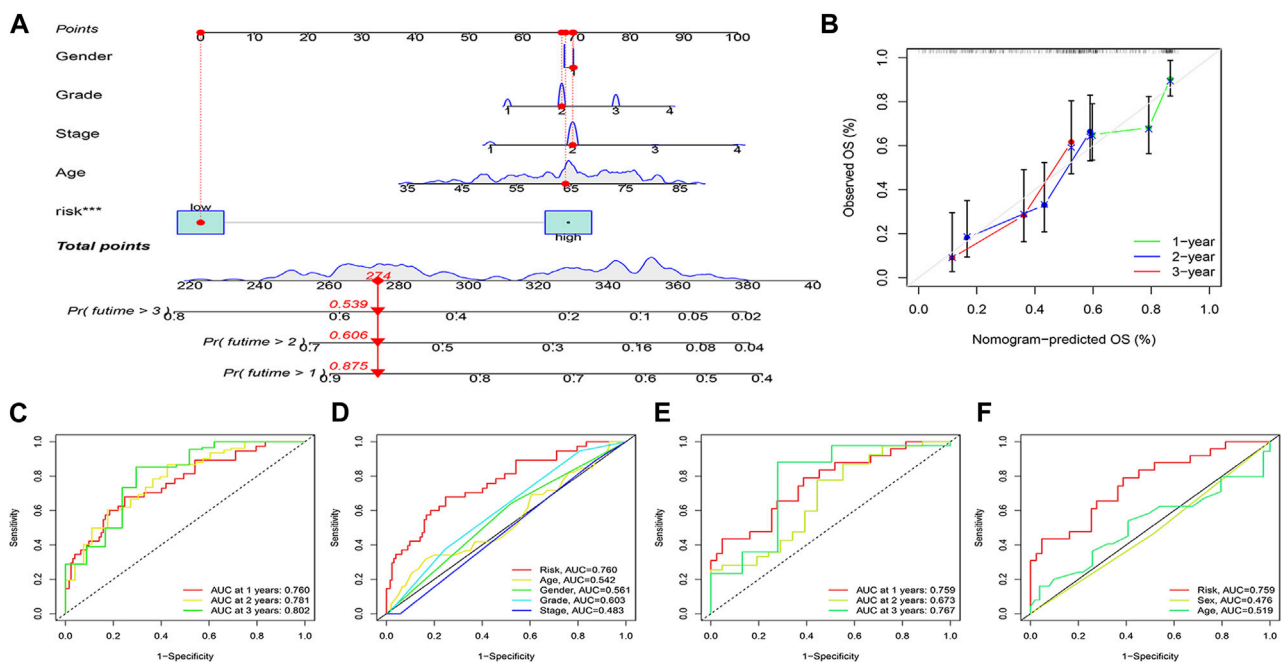
FIGURE 7
Nomogram and its verification. **(A)** Nomogram of patients' prognosis at 1–3 years. **(B)** Calibration curves. **(C, D)** ROC analysis based on TCGA dataset. **(E, F)** ROC analysis based on ICGC dataset.

FIGURE 8
GSEA enrichment analysis. **(A)** ABCB4. **(B)** APOBEC3C. **(C)** ENPP6. **(D)** FGFBP2. **(E)** LIPE. **(F)** MT2A. **(G)** OXER1. **(H)** PLD4. **(I)** ZNF589.
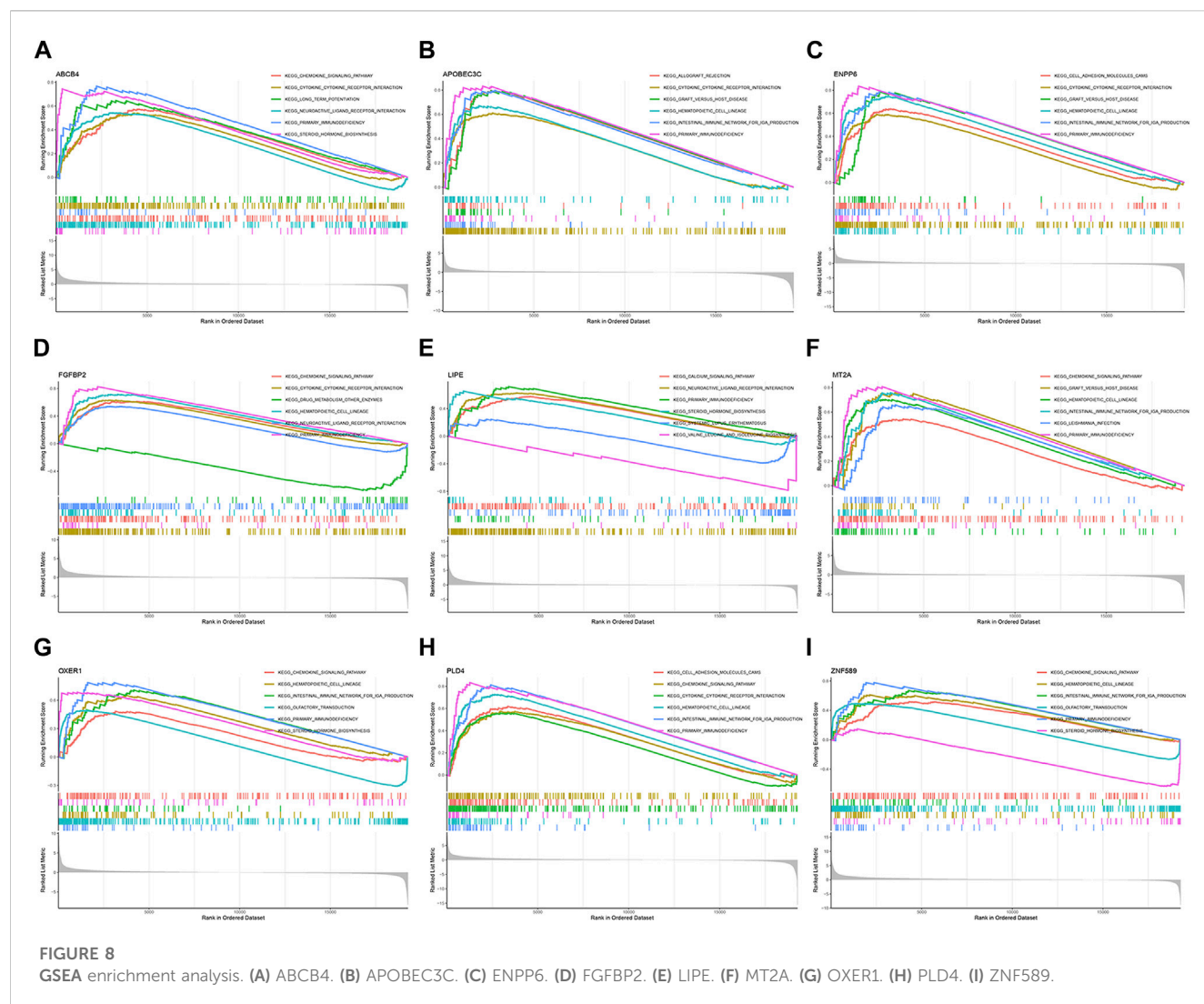
M2 macrophage-related model genes to influence the pancreatic cancer tumour microenvironment.

GSVA enrichment analysis

GSVA enrichment analysis revealed a negative correlation between ABCB4, ENPP6, FGFBP2, LIPE, OXER1, PLD4, ZNF589 and the p53 signaling pathway. ABCB4, APOBEC3C, ENPP6, FGFBP2, MT2A, PLD4 and the MAPK signaling pathway were positively correlated. ABCB4, ENPP6, PLD4 and the calcium signaling pathway were positively correlated. FGFBP2, PLD4 and calcium signaling pathway were positively correlated (Figure 11A).

Immunotherapy predictions

The prognostic model for M2 macrophage-associated genes was negatively correlated with most immune checkpoint blockage-associated genes (CD40, IDO2, TNFRSF8, CD48,

CD28, PDCD1) and a few immune checkpoint blockage genes (TNFSF9, TNFSF4, CD44, CD276) were positively correlated with the risk score model (Figure 11B). Higher IPS scores in the low-risk group (pd1-negative and ctla4-negative) indicated that the low-risk group was better treated with the new immune checkpoint inhibitors (ICI) (Figures 11C–F). These results demonstrated the potential role of M2 macrophage-related risk groups in predicting the outcome of immunotherapy in patients.

Predicting chemotherapy drug efficacy

Analysis of the chemotherapeutic drugs' semi-inhibitory concentrations identified that paclitaxel, rafatinib and lapatinib had a higher drug sensitivity in the low-risk group than in the high-risk group, while metformin had a higher drug sensitivity in the high-risk group. The results of the study showed a correlation between the effect of chemotherapeutic drugs and the prognostic model of M2 macrophage-associated genes (Figures 11G, H).

FIGURE 9
Tumour mutation load analysis, **(A)** Differential counting of tumour mutation burden between high and low risk groups. **(B)** Correlation analysis of risk score and mutation burden. **(C, D)** OncoPrint between high and low risk groups. **(E)** Prognostic analysis of tumour mutation load, **(F)** Prognostic analysis of tumour mutation load in high and low risk groups.
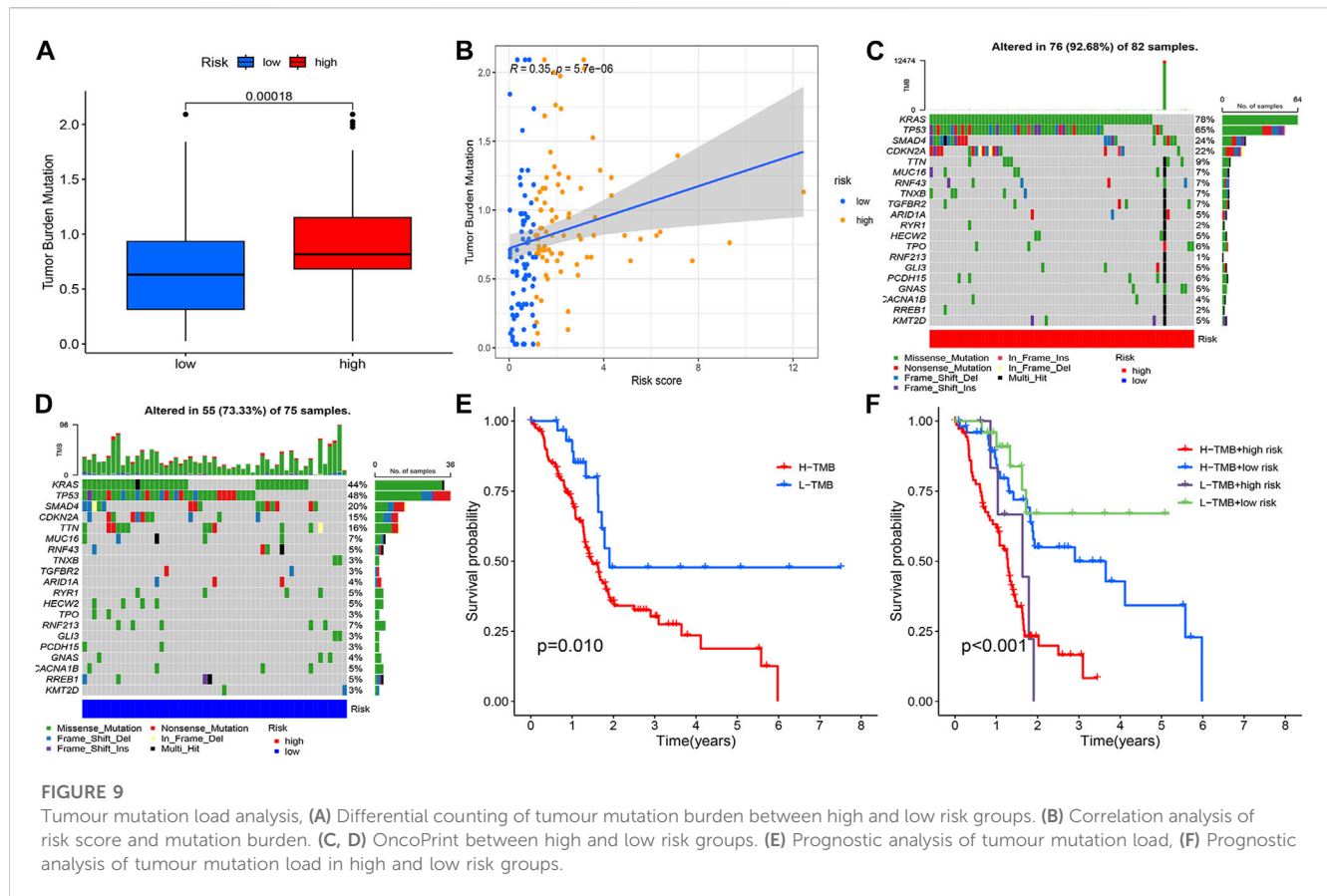
Independent sample validation

The gene expression difference was verified by RT-PCR detection of 20 samples of pancreatic cancer patients from Renmin Hospital of Wuhan University. The results showed that APOBEC3C and LIPE were highly expressed in pancreatic cancer tissues. However, ABCB4, ENPP6, FGFBP2, MT2A, OXER1, PLD4 and ZNF589 were low expressed in pancreatic cancer tissues (Figure 12).

Discussion

Pancreatic cancer is highly aggressive and most patients are diagnosed at an advanced stage and are deprived of effective treatment options (Hidalgo et al., 2015; Karamitopoulou, 2019). The immune system of the body is the last barrier to kill tumour cells. The low immunogenicity and immune escape characteristics of pancreatic cancer reduce the therapeutic efficacy of patients with pancreatic cancer (Liu et al., 2022). Tumour-associated macrophages account for a substantial proportion of the pancreatic cancer tumour microenvironment, and the major part of pancreatic cancer-associated macrophages differentiate into M2-type tumour-associated macrophages (Velasco et al., 2023). Recent studies have demonstrated that M2-type macrophages are involved in immune escape from pancreatic cancer (Campbell et al., 2010; Evan et al., 2022). Therefore, further study to uncover M2-related

genes in the tumour microenvironment macrophage-associated genes and the mechanisms of action between the tumour microenvironment may increase new horizons for immune tolerance in pancreatic cancer therapy.

In recent years, there have been some advances in the treatment of solid tumours. For example, immunotherapy has been applied in the treatment protocols for solid tumours such as breast cancer, lung cancer and liver cancer (Link et al., 2018; Locati et al., 2020). However, only a few tumours have achieved favourable clinical outcomes. There is an urgent clinical need for new treatment options to stimulate the patient's immune system to kill tumour cells. The tumour microenvironment provides a supportive ecological environment for cancer cell development and metastasis. It has been found that in solid tumours macrophages occupy a predominant component of the tumour microenvironment. However, macrophages have a dual role in cancer (Hinshaw and Shevde, 2019; Pittet et al., 2022). In different settings macrophages exhibit different forms of activation. In the classical pathway macrophages differentiate into M1 macrophages in response to stimulation by bacterial products and interferons (Cao et al., 2022). M2 macrophages are produced in the type 2 immune response by factors such as IL-4 and IL-13 via the alternative pathway (Biswas and Mantovani, 2010). The M1 type of macrophage possesses the function of killing tumour cells (Schlundt et al., 2021). In contrast, M2 is involved in the entire process of tumourigenesis and metastasis. It has been shown that m2 macrophages can be recruited by individual tumour initiating

FIGURE 10
Estimated abundance of tumor-infiltrating cells. Patients in the **(A)** high-risk group had a stronger correlation with tumor-infiltrating immune cells, as shown by the Spearman correlation analysis. **(B)** Association between prognostic risk signatures and central immune checkpoint genes. The asterisks represented the statistical *p*-value. The correlations predicted by the four methods CIBERSORT–ABS **(C)**, CIBERSORT **(D)**, QUANTISEQ **(E)**, and XCELL **(F)** were validated. (**p* < 0.05; ***p* < 0.01; ****p* < 0.001).

cells and thus provided a culture ecology for early tumourigenesis (Gordon and Plüddemann, 2019). Meanwhile, the pro-angiogenic and remodelling matrix of m2 cells can promote tumour growth and metastasis (Raghavan et al., 2021). The immune tolerance that occurs during tumour immunotherapy may be related to the overexpression of suppressive counter-receptors (e.g., PDL1 and PDL2) by m2 macrophages that suppress the body's immune function (Pushalkar et al., 2018). Therefore, blocking macrophage-associated immunosuppressive targets may be a way to suppress adaptive immune responses. Blocking macrophage-associated immunosuppressive targets may therefore be a potential therapeutic option to suppress adaptive immune responses and enhance the efficacy of immunotherapy (Stakheyeva et al., 2017; Riquelme et al., 2018).

The M2 macrophage-associated genes identified in this study have been reported in the existing pancreatic cancer tumour microenvironment (Mazarico et al., 2016; Saito et al., 2022). Mazarico et al. (2016) discovered that ABCB4 was overexpressed in pancreatic cancer-resistant patients treated with gemcitabine, indicating that ABCB4 may enhance immune escape of tumour cells by affecting macrophage function, leading to resistance to chemotherapeutic agents. Qian (Qian et al., 2020; Qian et al., 2022) revealed that overexpression of APOBEC3C induces genomic instability in pancreatic cancer, increases tumour cell heterogeneity and participates in the remodelling of the tumour immune microenvironment by influencing the function of immune cells. In the tumor microenvironment, ENPP can inhibit the aggregation of immune cells by reducing cGAMP, resulting in
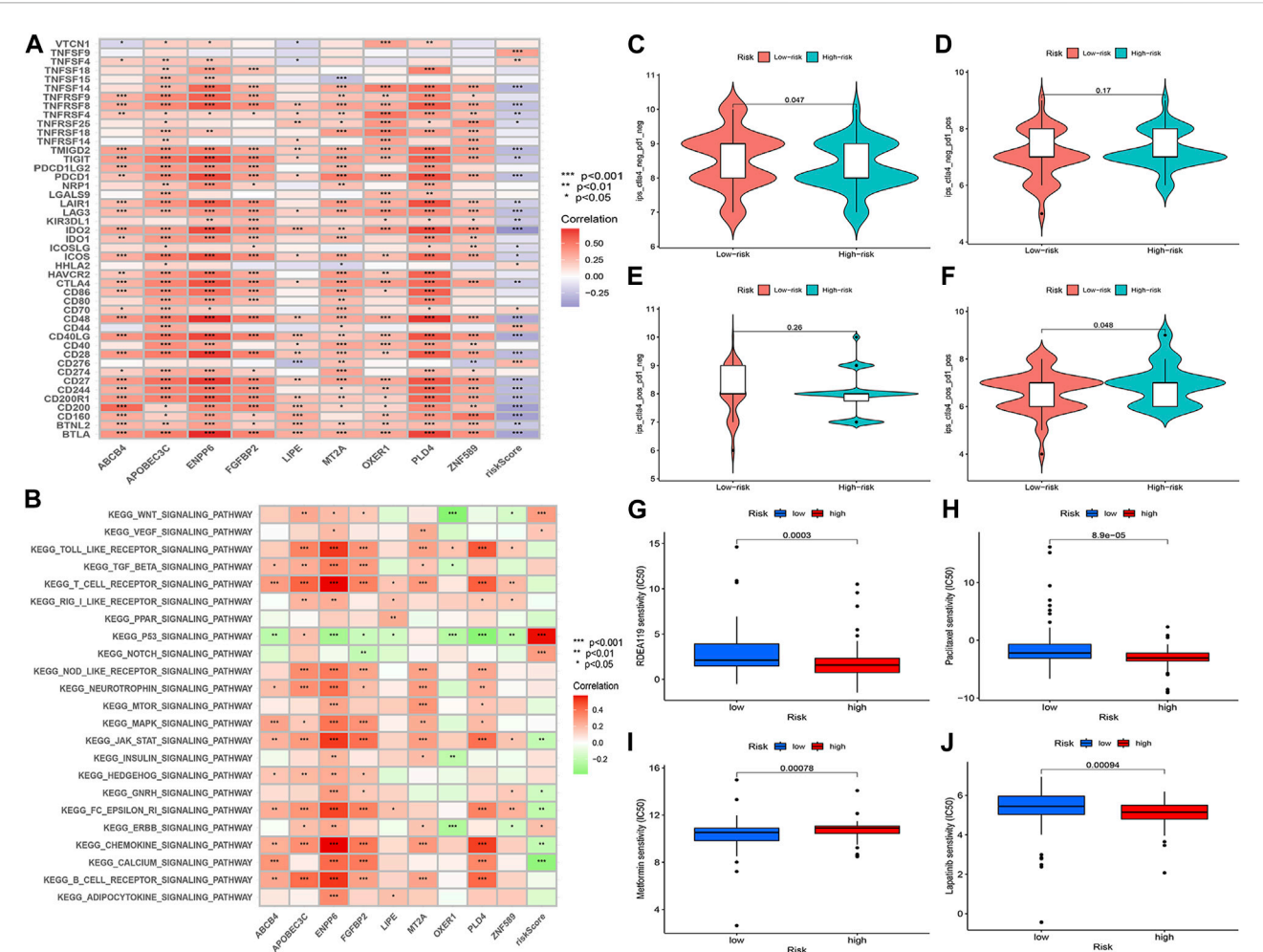
FIGURE 11
(A) GSVA enrichment analysis. (B) Correlation analysis of immune checkpoint blockade gene expression levels and risk scores. (C–F) IPS score distribution map. Estimates of chemotherapy effect risk scores. (G) Metformin. (H) Rafatinib. (I) Paclitaxel. (J) Lapatinib.

FIGURE 12
Analysis of expression differences. Verifying the expression of genes that constitute the risk model through RT qPCR.
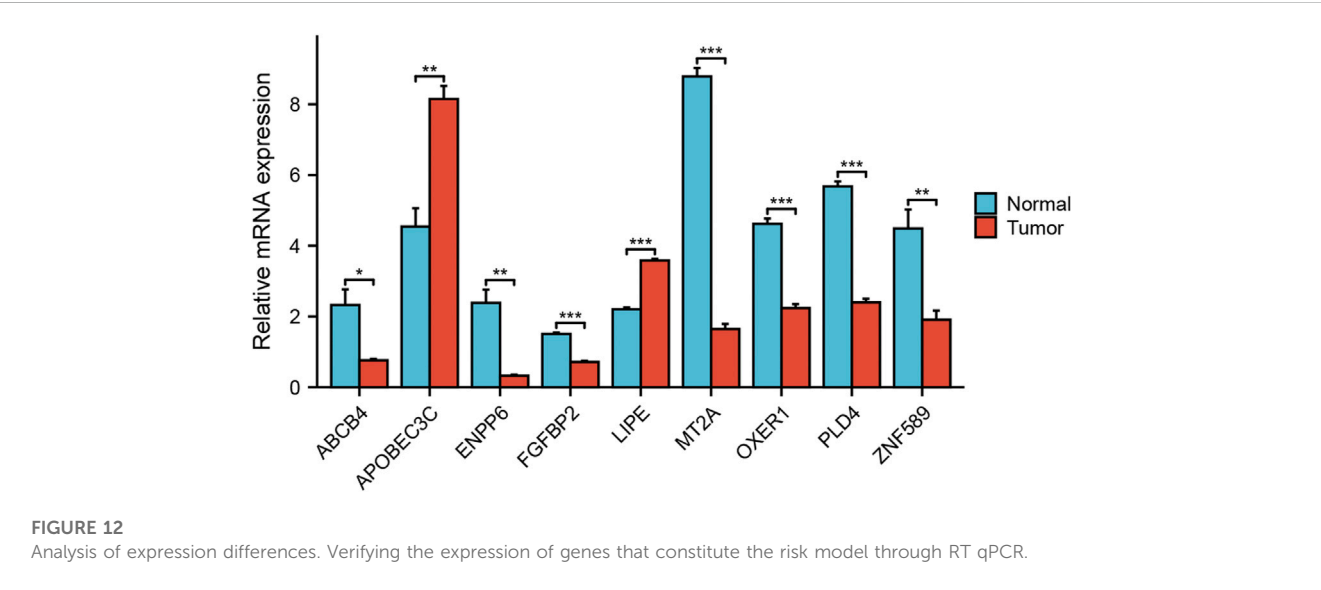
enhanced immune escape of tumor cells (Matas-Rico et al., 2021; Borza et al., 2022). Böker (Böker et al., 2022) showed a large number of single nucleotide variants in FGFBP2 in pancreatic tumour cells, and these changes affect the growth and migration of tumour cells. Masi revealed that OXER1 may be involved in the remodelling of the tumour immune microenvironment through multiple pathways and could be a potential target for immunotherapy (Masi et al., 2021). Although LIPE, MT2A, PLD4 and ZNF589 have been studied in other tumours, their relationship with tumour-associated macrophages in pancreatic cancer remains to be investigated. These findings not only provide new insights into the pathogenesis and immune tolerance mechanisms of pancreatic cancer in the future, but also may be potential new therapeutic targets for pancreatic cancer.

Undoubtedly, there are still some limitations in the present study. Firstly, the difference in mRNA expression was verified in tumor and normal tissues. However, further validation in cells and animals should proceed. Secondly, the results of the study need more work before they can be applied clinically.

Conclusion

The M2 macrophage-associated prognostic model for pancreatic cancer performed excellently in patient prognosis, tumour mutation load analysis, immune checkpoint prediction, and chemotherapy drug sensitivity prediction. Meanwhile, M2 macrophage-related genes may be involved in the targeting of immunotherapy in pancreatic cancer patients.

Data availability statement

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

Ethics statement

Statement of Institutional Review Board: This study was approved by the Ethics Committee of Renmin Hospital of Wuhan University and was conducted in accordance with the principles of the Declaration of Helsinki. Informed Consent: Written informed consent was obtained from all patients to participate in the study.

Author contributions

Conceptualization: DY, FZ, YS, and YZ data curation: JS, FZ formal analysis: DY, YZ writing-original draft: DY, FZ, and KZ writing-review and editing: YD, KZ. All authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1184708/full#supplementary-material

References

Biswas, S. K., and Mantovani, A. (2010). Macrophage plasticity and interaction with lymphocyte subsets: Cancer as a paradigm. *Nat. Immunol.* 11 (10), 889–896. doi:10.1038/ni.1937

Böker, V., Häußler, J., Baumann, J., Sunami, Y., Trojanowicz, B., Harwardt, B., et al. (2022). Analysis of genomic alterations in cancer associated human pancreatic stellate cells. *Sci. Rep.* 12 (1), 13532. doi:10.1038/s41598-022-17748-1

Borza, R., Salgado-Polo, F., Moolenaar, W. H., and Perrakis, A. (2022). Structure and function of the ecto-nucleotide pyrophosphatase/phosphodiesterase (ENPP) family: Tidying up diversity. *J. Biol. Chem.* 298 (2), 101526. doi:10.1016/j.jbc.2021.101526

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

Campbell, P. J., Yachida, S., Mudie, L. J., Stephens, P. J., Pleasance, E. D., Stebbings, L. A., et al. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467 (7319), 1109–1113. doi:10.1038/nature09460

Cao, X., Lai, S. W. T., Chen, S., Wang, S., and Feng, M. (2022). Targeting tumor-associated macrophages for cancer immunotherapy. *Int. Rev. Cell Mol. Biol.* 368, 61–108. doi:10.1016/bs.ircmb.2022.02.002

Christenson, E. S., Jaffee, E., and Azad, N. S. (2020). Current and emerging therapies for patients with advanced pancreatic ductal adenocarcinoma: A bright future. *Lancet Oncol.* 21 (3), e135–e145. doi:10.1016/S1470-2045(19)30795-8

Cohen, R., Neuzillet, C., Tijeras-Raballand, A., Faivre, S., de Gramont, A., and Raymond, E. (2015). Targeting cancer cell metabolism in pancreatic adenocarcinoma. *Oncotarget* 6 (19), 16832–16847. doi:10.18632/oncotarget.4160

Davies, L. C., and Taylor, P. R. (2015). Tissue-resident macrophages: Then and now. *Immunology* 144 (4), 541–548. doi:10.1111/imm.12451

Evan, T., Wang, V. M., and Behrens, A. (2022). The roles of intratumour heterogeneity in the biology and treatment of pancreatic ductal adenocarcinoma. *Oncogene* 41 (42), 4686–4695. doi:10.1038/s41388-022-02448-x

Fan, J. Q., Wang, M. F., Chen, H. L., Shang, D., Das, J. K., and Song, J. (2020). Current advances and outlooks in immunotherapy for pancreatic ductal adenocarcinoma. *Mol. cancer* 19 (1), 32. doi:10.1186/s12943-020-01151-3

Feig, C., Gopinathan, A., Neesse, A., Chan, D. S., Cook, N., and Tuveson, D. A. (2012). The pancreas cancer microenvironment. *Clin. cancer Res. official J. Am. Assoc. Cancer Res.* 18 (16), 4266–4276. doi:10.1158/1078-0432.CCR-11-3114

Gordon, S., and Plüddemann, A. (2019). The mononuclear phagocytic system. Generation of diversity. *Front. Immunol.* 10, 1893. doi:10.3389/fimmu.2019.01893

Hidalgo, M., Cascinu, S., Kleeff, J., Labianca, R., Löhr, J. M., Neoptolemos, J., et al. (2015). Addressing the challenges of pancreatic cancer: Future directions for improving outcomes. *Pancreatol. official J. Int. Assoc. Pancreatol. (IAP) [et al]* 15 (1), 8–18. doi:10.1016/j.pan.2014.10.001

Hinshaw, D. C., and Shevde, L. A. (2019). The tumor microenvironment innately modulates cancer progression. *Cancer research* 79 (18), 4557–4566. doi:10.1158/0008-5472.CAN-18-3962

Huang, B. Z., and Setiawan, V. W. (2022). Changes in metabolic syndrome and its implications on the risk and racial/ethnic disparities of pancreatic cancer. *Gastroenterology* 162 (2), 392–394. doi:10.1053/j.gastro.2021.11.012

Ip, W. K. E., Hoshi, N., Shouval, D. S., Snapper, S., and Medzhitov, R. (2017). Anti-inflammatory effect of IL-10 mediated by metabolic reprogramming of macrophages. *Science (New York, NY)* 356 (6337), 513–519. doi:10.1126/science.aal3535

Karamitopoulou, E. (2019). Tumour microenvironment of pancreatic cancer: Immune landscape is dictated by molecular and histopathological features. *British journal of cancer* 121 (1), 5–14. doi:10.1038/s41416-019-0479-5

Le, D. T., Uram, J. N., Wang, H., Bartlett, B. R., Kemberling, H., Eyring, A. D., et al. (2015). PD-1 blockade in tumors with mismatch-repair deficiency. *The N. Engl. journal of medicine* 372 (26), 2509–2520. doi:10.1056/NEJMoa1500596

Link, V. M., Duttke, S. H., Chun, H. B., Holtman, I. R., Westin, E., Hoeksema, M. A., et al. (2018). Analysis of genetically diverse macrophages reveals local and domain-wide mechanisms that control transcription factor binding and function. *Cell* 173 (7), 1796–1809.e17. doi:10.1016/j.cell.2018.04.018

Liu, Y., Wang, X., Zhu, Y., Cao, Y., Wang, L., Li, F., et al. (2022). The CTCF/LncRNA-PACERR complex recruits E1A binding protein p300 to induce pro-tumour macrophages in pancreatic ductal adenocarcinoma via directly regulating PTGS2 expression. *Clinical and translational medicine* 12 (2), e654. doi:10.1002/ctm2.654

Locati, M., Curtale, G., and Mantovani, A. (2020). Diversity, mechanisms, and significance of macrophage plasticity. *Annual review of pathology* 15, 123–147. doi:10.1146/annurev-pathmechdis-012418-012718

Masi, M., Racchi, M., Travelli, C., Corsini, E., and Buoso, E. (2021). Molecular characterization of membrane steroid receptors in hormone-sensitive cancers. *Cells* 10 (11), 2999. doi:10.3390/cells10112999

Matas-Rico, E., Frijlink, E., van der Haar Àvila, I., Menegakis, A., van Zon, M., Morris, A. J., et al. (2021). Autotaxin impedes anti-tumor immunity by suppressing chemotaxis and tumor infiltration of CD8(+) T cells. *Cell reports* 37 (7), 110013. doi:10.1016/j.celrep.2021.110013

Mazarico, J. M., Sánchez-Arévalo Lobo, V. J., Favicchio, R., Greenhalf, W., Costello, E., Carrillo-de Santa Pau, E., et al. (2016). Choline kinase alpha (CHKα) as a therapeutic target in pancreatic ductal adenocarcinoma: Expression, predictive value, and sensitivity to inhibitors. *Molecular cancer therapeutics* 15 (2), 323–333. doi:10.1158/1535-7163.MCT-15-0214

Noy, R., and Pollard, J. W. (2014). Tumor-associated macrophages: From mechanisms to therapy. *Immunity* 41 (1), 49–61. doi:10.1016/j.immuni.2014.06.010

Ostios-Garcia, L., Villamayor, J., Garcia-Lorenzo, E., Vinal, D., and Feliu, J. (2021). Understanding the immune response and the current landscape of immunotherapy in pancreatic cancer. *World journal of gastroenterology* 27 (40), 6775–6793. doi:10.3748/wjg.v27.i40.6775

Pittet, M. J., Michielin, O., and Migliorini, D. (2022). Clinical relevance of tumour-associated macrophages. *Nature reviews Clinical oncology* 19 (6), 402–421. doi:10.1038/s41571-022-00620-6

Pushalkar, S., Hundeyin, M., Daley, D., Zambirinis, C. P., Kurz, E., Mishra, A., et al. (2018). The pancreatic cancer microbiome promotes oncogenesis by induction of innate and adaptive immune suppression. *Cancer discovery* 8 (4), 403–416. doi:10.1158/2159-8290.CD-17-1134

Qian, Y., Gong, Y., Fan, Z., Luo, G., Huang, Q., Deng, S., et al. (2020). Molecular alterations and targeted therapy in pancreatic ductal adenocarcinoma. *Journal of hematology & oncology* 13 (1), 130. doi:10.1186/s13045-020-00958-3

Qian, Y., Gong, Y., Zou, X., Liu, Y., Chen, Y., Wang, R., et al. (2022). Aberrant APOBEC3C expression induces characteristic genomic instability in pancreatic ductal adenocarcinoma. *Oncogenesis* 11 (1), 35. doi:10.1038/s41389-022-00411-9

Raghavan, S., Winter, P. S., Navia, A. W., Williams, H. L., DenAdel, A., Lowder, K. E., et al. (2021). Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. *Cell* 184 (25), 6119–6137.e26. doi:10.1016/j.cell.2021.11.017

Riquelme, E., Maitra, A., and McAllister, F. (2018). Immunotherapy for pancreatic cancer: More than just a gut feeling. *Cancer discovery* 8 (4), 386–388. doi:10.1158/2159-8290.CD-18-0123

Saito, R. F., Andrade, L. N. S., Bustos, S. O., and Chammas, R. (2022). Phosphatidylcholine-derived lipid mediators: The crosstalk between cancer cells and immune cells. *Frontiers in immunology* 13, 768606. doi:10.3389/fimmu.2022.768606

Schlundt, C., Fischer, H., Bucher, C. H., Rendenbach, C., Duda, G. N., and Schmidt-Bleek, K. (2021). The multifaceted roles of macrophages in bone regeneration: A story of polarization, activation and time. *Acta biomaterialia* 133, 46–57. doi:10.1016/j.actbio.2021.04.052

Siegel, R. L., Miller, K. D., and Jemal, A. Cancer statistics, 2018. *CA a cancer journal for clinicians.* 68(1):7–30. doi:10.3322/caac.21442

Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA a cancer journal for clinicians* 70 (1), 7–30. doi:10.3322/caac.21590

Stakheyeva, M., Riabov, V., Mitrofanova, I., Litviakov, N., Choynzonov, E., Cherdyntseva, N., et al. (2017). Role of the immune component of tumor microenvironment in the efficiency of cancer treatment: Perspectives for the personalized therapy. *Current pharmaceutical design* 23 (32), 4807–4826. doi:10.2174/1381612823666170714161703

Ullman, N. A., Burchard, P. R., Dunne, R. F., and Linehan, D. C. (2022). Immunologic strategies in pancreatic cancer: Making cold tumors hot. *Journal of clinical oncology official journal of the American Society of Clinical Oncology* 40 (24), 2789–2805. doi:10.1200/JCO.21.02616

Velasco, R. M., García, A. G., Sánchez, P. J., Sellart, I. M., and Sánchez-Arévalo Lobo, V. J. (2023). Tumour microenvironment and heterotypic interactions in pancreatic cancer. *Journal of physiology and biochemistry* 79 (1), 179–192. doi:10.1007/s13105-022-00875-8

Yang, J., Li, Y., Sun, Z., and Zhan, H. (2021). Macrophages in pancreatic cancer: An immunometabolic perspective. *Cancer letters* 498, 188–200. doi:10.1016/j.canlet.2020.10.029

Check for updates

Cancer-testis antigen CEP55 serves as a prognostic biomarker and is correlated with immune infiltration and immunotherapy efficacy in pan-cancer

Xiaodong Xie[1,2], Hongyin Liang[1,2], Wushuang Jiangting[3], Yu Wang[4], Xiao Ma[1,2], Zhen Tan[1,2], Long Cheng[1,2,5], Zhulin Luo[1,2,5]* and Tao Wang[1,2,5]*

[1]Department of General Surgery, The General Hospital of Western Theater Command, Chengdu, Sichuan, China, [2]Department of General Surgery and Pancreatic Injury and Repair Key Laboratory of Sichuan Province, The General Hospital of Western Theater Command, Chengdu, Sichuan, China, [3]Department of Anesthesiology, The General Hospital of Western Theater Command, Chengdu, Sichuan, China, [4]Department of Microbiology, Zunyi Medical University, Zunyi, Guizhou, China, [5]College of Medicine, The Southwest Jiaotong University, Chengdu, Sichuan, China

Background: Centrosomal Protein 55 (CEP55) was initially described as a main participant in the final stage of cytokinesis. Further research identified CEP55 as a cancer-testis antigen (CTA) that is aberrantly expressed in different malignancies and a cancer vaccination candidate. The current study aimed to disclose the complete expression of CEP55, its effect on various malignancy prognoses, and its role in the tumor microenvironment.

Methods: Transcriptional information regarding tumor and normal tissues, as well as externally validated and protein expression data were gathered from the Cancer Genome Atlas, Genotype-Tissue Expression project, Gene Expression Omnibus, and Human Protein Atlas. We examined the effect of CEP55 on tumor prognosis using Kaplan-Meier (KM) and univariate Cox regression analyses. In addition, we investigated the connections between CEP55 expression and hallmark cancer pathways, immune cell infiltration, and immune regulator expression across malignancies. We constructed and validated a CEP55-related risk model for hepatocellular carcinoma (HCC) and explored the correlations between CEP55 expression and HCC molecular subtypes. Finally, we investigated putative small-molecule drugs targeting CEP55 using a connectivity map (CMap) database and validated them using molecular docking analysis.

Findings: CEP55 was aberrantly expressed in most cancers and revealed a prognostic value for several malignancies. Cancers with high CEP55 expression showed significantly enhanced cell cycle, proliferation, and immune-related pathways. For most malignancies, elevated CEP55 expression was associated with the infiltration of myeloid-derived suppressor cells (MDSCs) and Th2 cells. In addition, CEP55 expression was linked to immunomodulators and the potential prediction of immune checkpoint inhibitor (ICI) responses, and strongly associated with distinct molecular HCC subtypes, whereby the CEP55-based nomogram performed well in predicting short- and long-term HCC survival.

Finally, we used connectivity map (CMap) and molecular docking analyses to discover three candidate small-molecule drugs that could directly bind to CEP55.

Conclusion: CEP55 affected the occurrence and development of various cancers and possibly the regulation of the tumor immune microenvironment. Our findings suggest that CEP55 is a potential biomarker for prognosis and a powerful biomarker for ICI efficacy prediction.

1 Introduction

The considerable increase in cancer incidence and mortality in recent decades has caused this disease to become the primary cause of death and a main factor in the decrease in life expectancies worldwide (Sung et al., 2021). Approximately 1,958,310 million new cancer cases and 609,820 cancer-related deaths are expected in the United States in 2023 (Siegel et al., 2023). Although targeted therapy and immunotherapy advances have greatly improved tumor treatment, the long-term survival rates of patients with various forms of tumors are unsatisfactory (Hughes et al., 2016; Sharma et al., 2017). Novel, sensitive diagnostic markers and suitable therapeutic targets are critical for improving prognoses for cancer patients.

Various proteins are involved in the pathogenesis and progression of cancers, including survival-associated signaling kinases with relatively common mutants (e.g., BRAF V600E and KRAS G12D) (Cisowski et al., 2016), DNA damage repair response molecules (e.g., p53 and Rad51) (Bonilla et al., 2020), and cell cycle progression-associated compounds (e.g., Rb1 and CDK4/6) (Fassl et al., 2022). In particular, the targeting of cell cycle progression has advanced greatly with the development of CDK4/6 inhibitors that are now in clinical use for HER2+ breast cancer and are under investigation for the treatment of a variety of Rb1+ cancers (Xu et al., 2017; Pesch et al., 2022). However, these treatments are ineffective for the large proportion of cancers that have CDK4/6 or Rb1 mutations. Therefore, cell cycle inhibitors must be sourced that can be used in these contexts.

Centrosomal Protein 55 (CEP55, also known as c10orf3 and FLJ10540) was initially described in physiological studies as a midbody- and centrosome-associated coiled-coil protein with a size of approximately 55 kDa (Fabbro et al., 2005). Aberrant expression of CEP55 has been strongly associated with clinical features and prognoses in cancer patients, which has led to an increasing interest in this molecule as a possible therapeutic target (Jeffery et al., 2016). In recent years, this interest has led to the identification of multiple independent molecular mechanisms that explain CEP55 activity. CEP55 interacts with endosomal sorting components to recruit the endosomal sorting complex required for transport (ESCRT) to the midbody, and may play a significant role in the constriction of the intracellular bridge and the promotion of abscission and cytokinesis (Zhao et al., 2006). Defects in cytokinesis can cause the generation of aneuploid cells, which is a key step in tumorigenesis; therefore, CEP55 is suggested to be involved in tumor initiation (Lens and Medema, 2019). In addition, CEP55 binds and stabilizes the PI3K catalytic subunit and facilitates the activation of the PI3K/Akt pathway, thus promoting the survival and proliferation of cancer cells (Montero-Conde et al., 2008). This association offers an alternative mechanism for CEP55 to participate in the later stages of tumor progression and metastasis. Beyond these cell division and signaling roles, CEP55 has recently been identified as an immunogenic tumor-associated antigen (TAA) and cancer-testis antigen (CTA), which recommend it as a potential candidate for cancer vaccine therapies. CEP55 peptides are naturally present in breast cancer cells and can induce an increase in antigen-specific cytotoxic T lymphocytes (CTLs) (Inoda et al., 2009). CEP55-specific CTLs could be capable of recognizing and killing chemotherapy-resistant colon cancer stem cells (Inoda et al., 2011; Gao and Wang, 2015). A recent study revealed that CEP55 is expressed in exosomes derived from malignant cells and has the potential to be a non-invasive diagnostic marker for tumors (Qadir et al., 2018). Collectively, the results of these studies demonstrate that CEP55 could have various functions in multiple cancers. However, a comprehensive picture of CEP55 in the tumor immune microenvironment has not yet been reported, and small therapeutic compounds targeting CEP55 remain elusive.

Through an integrated analysis of the genomic and expression data of multiple independent databases, this study aimed to thoroughly characterize the functions of CEP55 in the regulation of tumor biological processes and the immune microenvironment and its potential as an immunotherapy and chemotherapy target at the pan-cancer level. CEP55 was aberrantly expressed in most cancers and showed prognostic value for several malignancies. High CEP55 expression was strongly correlated with the infiltration of myeloid-derived suppressor cells (MDSCs) and Th2 cells in most cancers, and was substantially related to distinct molecular subtypes of HCC, whereby the CEP55-based nomogram performed well in predicting short- and long-term HCC survival. Finally, using a connectivity map (CMap) and molecular docking analyses, we identified three potential small molecules targeting CEP55, which may mitigate the immunosuppressive microenvironment and enhance the anti-tumor effect of ICIs.

2 Materials and methods

2.1 Data acquisition and processing

Gene expression data and pathological and clinical information from The Cancer Genome Atlas (TCGA) project and normal

samples from the Genotype-Tissue Expression (GTEx) dataset were downloaded from UCSC Xena (https://xenabrowser.net/). This study comprehensively analyzed 33 types of cancer from 9,807 tumor tissues from TCGA database and 7,873 normal tissues (727 from TCGA database and 7,146 from the GTEx database). Batch effects were corrected using the UCSC TOIL RNA-seq Recompute workflow as previously described (Vivian et al., 2017). Transcripts per kilobase million (TPM) normalized expression data were used for subsequent analyses. Microarray expression data of 29 cancers were downloaded from the GEO database (http://www.ncbi.nih.gov/geo/) for external validation. HCC gene expression data from the ICGC (https://dcc.icgc.org/) database (including 197 normal tissues and 240 HCC tumor tissues) and the GSE14520 dataset (including 239 normal tissues and 247 HCC tumor tissues) were included in the validation of CEP55 in HCC. Information on the datasets used in this study is detailed in Supplementary Table S1.

2.2 CEP55 protein expression, localization, and interaction

The Human Protein Atlas (HPA; www.proteinatlas.org) initiative uses multiple omics methods to focus on protein expression in cells, tissues, and organs (Uhlén et al., 2015). The expressions of CEP55 in normal tissues and cell types were obtained using the "Tissue" and "Single Cell Type" sections of this atlas. The "Pathology" section was used to validate the protein levels of CEP55 (HPA023430) in various cancers. Immunofluorescence staining images from the "Subcellular" section were used to explore the subcellular localization of CEP55. The protein-protein interactions of CEP55 were analyzed using the ComPPI website (https://comppi.linkgroup.hu/), which provides integrated information regarding protein-protein interactions and their localization (Veres et al., 2015).

2.3 Genomic alteration analysis of CEP55

The cBioPortal site (http://www.cbioportal.org/) is a public open-access link to multidimensional cancer genomic features, including genetics, epigenetics, gene expression, and proteomics (Cerami et al., 2012; Gao et al., 2013). The CEP55 genome alterations and their relationships with survival were determined using the "Mutation," "Plots" and "Cancer Types Summary" modules of this portal. Since microsatellite instability and tumor mutation burden (MSI and TMB, respectively) can predict responses to immunotherapy, MSI scores were obtained from the somatic mutation data. TMB scores were calculated using a Perl script and modified by dividing the scores by the total length of the exons. Spearman's correlation coefficient was used to assess the relationship between CEP55 expression and TMB or MSI.

2.4 Tumor immune infiltration analysis

The "Estimate" method infers the fraction of infiltrating stromal and immune cells in tumor samples (Yoshihara et al., 2013). The R

package "estimate" was utilized to compute the "Immune Score," "Stromal Score," "Tumor Purity" and "Estimate Score" in this study. The Tumor Immune Estimation Resource (TIMER) database provides a comprehensive resource for analyzing the abundances of infiltrated immune cells across cancers using various immune deconvolution methods (Li et al., 2017; Li et al., 2020). The correlations between CEP55 expression and pan-cancer immune infiltration levels were accessed from the "Immune Association" section of the TIMER2.0 database (http://timer.cistrome.org/).

2.5 Functional and pathway enrichment analysis

The patients with CEP55 mRNA expression levels in the highest 30% were referred to as the high-CEP55 group and those with the lowest 30% were referred to as the low-CEP55 group. The "limma" R package was used to determine the different expression genes (DEGs) of the high- and low-CEP55 subgroups (Ritchie et al., 2015). The cancer HALLMARK geneset (h.all.v7.2. symbols) was used for the functional and pathway enrichment analyses. The normalized enrichment score (NES) and false discovery rate (FDR) for each cancer type were determined using the R package "clusterProfiler" (Yu et al., 2012).

2.6 Immunotherapy prediction analysis

Immunomodulators, which contain various immunoregulatory genes, are critical components of immunotherapy treatments (Thorsson et al., 2018). Spearman's correlation coefficients between CEP55 expression and the immunomodulators were calculated for each cancer type. To explore the influence of CEP55 in the response to ICIs, CEP55 expression differences between responders and non-responders were calculated for nine independent immunotherapy cohorts. Detailed information on each immunotherapy cohort is provided in Supplementary Table S2.

2.7 Connectivity map for specific inhibitor analysis

The Connectivity Map (Cmap; https://clue.io/) database contains millions of gene expression profiles from different cell types that have been treated with perturbagens and is commonly used to predict cellular responses to chemical stimuli (Lamb et al., 2006). The 150 main up- or downregulated DEGs for each cancer type were subjected to Cmap analysis. Detailed lists of the query results were acquired and the scores were used to perform a heatmap analysis.

2.8 Molecular docking analysis

The 3D structure of CEP55 was downloaded from the PDB Data Bank (http://www.rcsb.org/), and the structural formulae of the small compounds were obtained from PubChem (https://
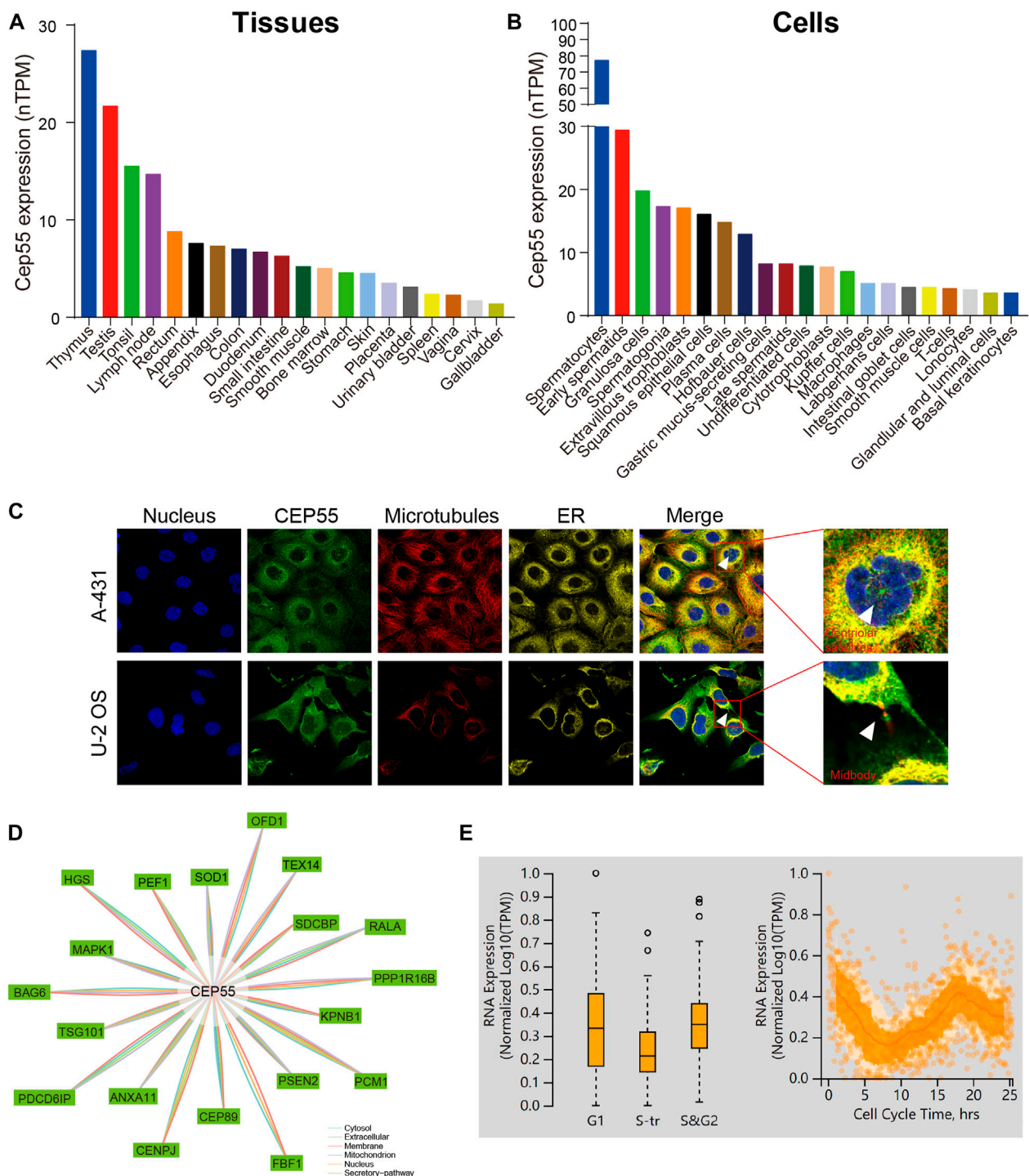
FIGURE 1
CEP55 expression and localization in normal tissues and cells. **(A)** The expression of CEP55 in normal tissues. **(B)** The expression of CEP55 in normal cells. **(C)** Subcellular localization of CEP55 in A-431 and U251 cell lines. The images were retrieved from the HPA (www.proteinatlas.org) database. **(D)** Protein-protein interaction (PPI) network of CEP55. **(E)** The correlation between CEP55 mRNA expression and cell cycle progression. The picture was obtained from the HPA (www.proteinatlas.org) database.

pubchem.ncbi.nlm.nih.gov/). AutoDockTools (v1.5.7) software was used for the molecular docking analysis. A docking free energy of ≤ −7.0 kcal/mol was considered stable binding. The molecular docking results were visualized using PyMOL software (v2.4.0) and the Protein Plus online server (https://proteins.plus/) (Schöning-Stierand et al., 2020).

2.9 Survival analysis

Overall survival (OS) and disease-specific survival (DSS) were used to determine the prognostic value of CEP55 expression across the various cancer types. Patients were separated into low- and high-CEP55 expression groups using the median CEP55 expression as the cutoff value. The "survival" and "forestplot" R packages were used to conduct a Cox regression analysis for pan-cancer. The Kaplan-Meier method and log-rank test were used for the survival analysis.

2.10 Statistical analyses

All computational procedures and statistical analyses were conducted using R software (https://www.r-project.org). Two normally distributed groups were analyzed for statistical significance using the unpaired Student's t-test, and non-normally distributed groups were compared using the Wilcoxon rank-sum test. The chi-squared or Fisher's exact test was used to compare the contingency table variables, and the log-rank test was used to compare the results of the survival analysis using Kaplan-Meier method. Statistical significance was defined as $p < 0.05$.

3 Results

3.1 The expression and localization of CEP55 in normal tissues and cells

We first analyzed CEP55 expression in tissues and cells using the HPA database and enhanced expression was observed in the testis and lymphoid tissues such as the thymus, tonsils, and lymph nodes. At the cellular level, CEP55 was highly expressed in germ cells (spermatocytes, early spermatids, and spermatogonia), epithelial cells (granulosa cells and squamous epithelial cells), and immune-related cells (plasma cells and Hofbauer cells) (Figures 1A,B). The immunofluorescence (IF) images of the HPA database showed that CEP55 was mainly located in the plasma membrane, centriolar satellites, and midbody region (Figure 1C). CEP55 interacted proteins were localized in the cytosol, extracellular space, membrane, mitochondria, nucleus, and secretory pathway (Figure 1D). Based on the fluorescent ubiquitination-based cell cycle indicator (FUCCI) assay and single-cell RNA sequences from the HPA database, CEP55 showed a strong correlation with the cell cycle and variable peak expression across cell cycle phases, which are consistent with its function in cytokinesis (Figure 1E).

3.2 Aberrant expression of CEP55 in cancer tissues

To further investigate the critical role of CEP55, we analyzed its expression in various cancers. Through integrated GTEx-TCGA data analysis, CEP55 showed significantly higher expression in tumor tissues than normal tissues in approximately 30 cancers, 23 of which were verified by external datasets: adrenocortical cancer (ACC), breast invasive carcinoma (BRCA), bladder urothelial

carcinoma (BLCA), colon adenocarcinoma (COAD), cervical and endocervical cancer (CESC), cholangiocarcinoma (CHOL), diffuse large B-cell lymphoma (DLBC), glioblastoma multiforme (GBM), esophageal carcinoma (ESCA), kidney chromophobe (KICH), kidney papillary cell carcinoma (KIRP), kidney clear cell carcinoma (KIRC), liver hepatocellular carcinoma (LIHC), lung squamous cell carcinoma (LUSC), lung adenocarcinoma (LUAD), ovarian serous cystadenocarcinoma (OV), rectum adenocarcinoma (READ), pancreatic adenocarcinoma (PAAD), skin cutaneous melanoma (SKCM), sarcoma (SARC), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), and uterine corpus endometrioid carcinoma (UCEC) (Figure 2A; Supplementary Figure S1). Furthermore, CEP55 expression was considerably higher in cancer tissues than in the corresponding tumor-adjacent normal tissues in 17 cancers: BRCA, BLCA, CHOL, ESCA, COAD, HNSC, KICH, KIRP, KIRC, LIHC, LUSC, LUAD, prostate adenocarcinoma (PRAD), STAD, READ, THCA, and UCEC (Figure 2B). At the protein level, immunohistochemistry (IHC) data from the HPA database revealed that CEP55 expression was considerably greater in cancer tissues than in normal tissues (Figure 2C). Therefore, we assessed the correlation between CEP55 expression and clinical characteristics and found that CEP55 expression was greater in high-stage tumors (Stages III and IV) than in low-stage tumors (Stages I and II) of ACC, HNSC, KICH, KIRP, KIRC, LIHC, LUAD, and LUSC, whereas CEP55 expression was lower in high-stage than low-stage COAD tumors (Figure 2D; Supplementary Figure S2A). CEP55 was elevated in high-grade (Grades III and IV) malignancies in BLCA, KIRC, HNSC, LIHC, LGG, OV, and UCEC, but showed reduced expression in high-grade CESC and STAD tumors (Figure 2E; Supplementary Figure S2B). Moreover, patients with recurrences during follow-up showed increased expression of CEP55 in ACC, KIRP, KICH, LIHC, LGG, PRAD, pheochromocytoma and paraganglioma (PCPG), UCEC, SARC, and uveal melanoma (UVM), and reduced expression in STAD (Figure 2F). These findings indicate that CEP55 is frequently overexpressed in multiple cancers and is associated with clinical characteristics.

3.3 CEP55 correlates with prognosis in pan-cancer

To assess the relationship between CEP55 expression and pan-cancer prognosis, we estimated overall survival (OS) and disease-specific survival (DSS) using univariate Cox regression with TCGA dataset. Increased expression of CEP55 showed significant correlations with poor OS rates in UVM (HR = 1.56, $p = 0.037$), PAAD (HR = 1.64, $p < 0.001$), MESO (HR = 1.75, $p < 0.001$), LUAD (HR = 1.15, $p = 0.002$), LIHC (HR = 1.12, $p = 0.014$), LGG (HR = 1.47, $p < 0.001$), KIRP (HR = 1.66, $p < 0.001$), KIRC (HR = 1.36, $p < 0.001$), KICH (HR = 2.35, $p < 0.001$), and ACC (HR = 2.17, $p < 0.001$). Higher CEP55 expressions predicted better OS rates in LUSC (HR = 0.90, $p = 0.008$) and THYM (HR = 0.69, $p = 0.033$) (Figure 3A). Higher expression of CEP55 was also related to poor DSS in UVM (HR = 1.59, $p = 0.035$), PRAD (HR = 2.18, $p < 0.043$), PAAD (HR = 1.69, $p < 0.001$), MESO (HR = 2.03, $p < 0.001$), LUAD (HR = 1.18 $p = 0.004$), LIHC (HR = 1.22, $p = 0.002$), LGG (HR = 1.48, $p < 0.001$), KIRP (HR = 2.08, $p < 0.001$), KIRC
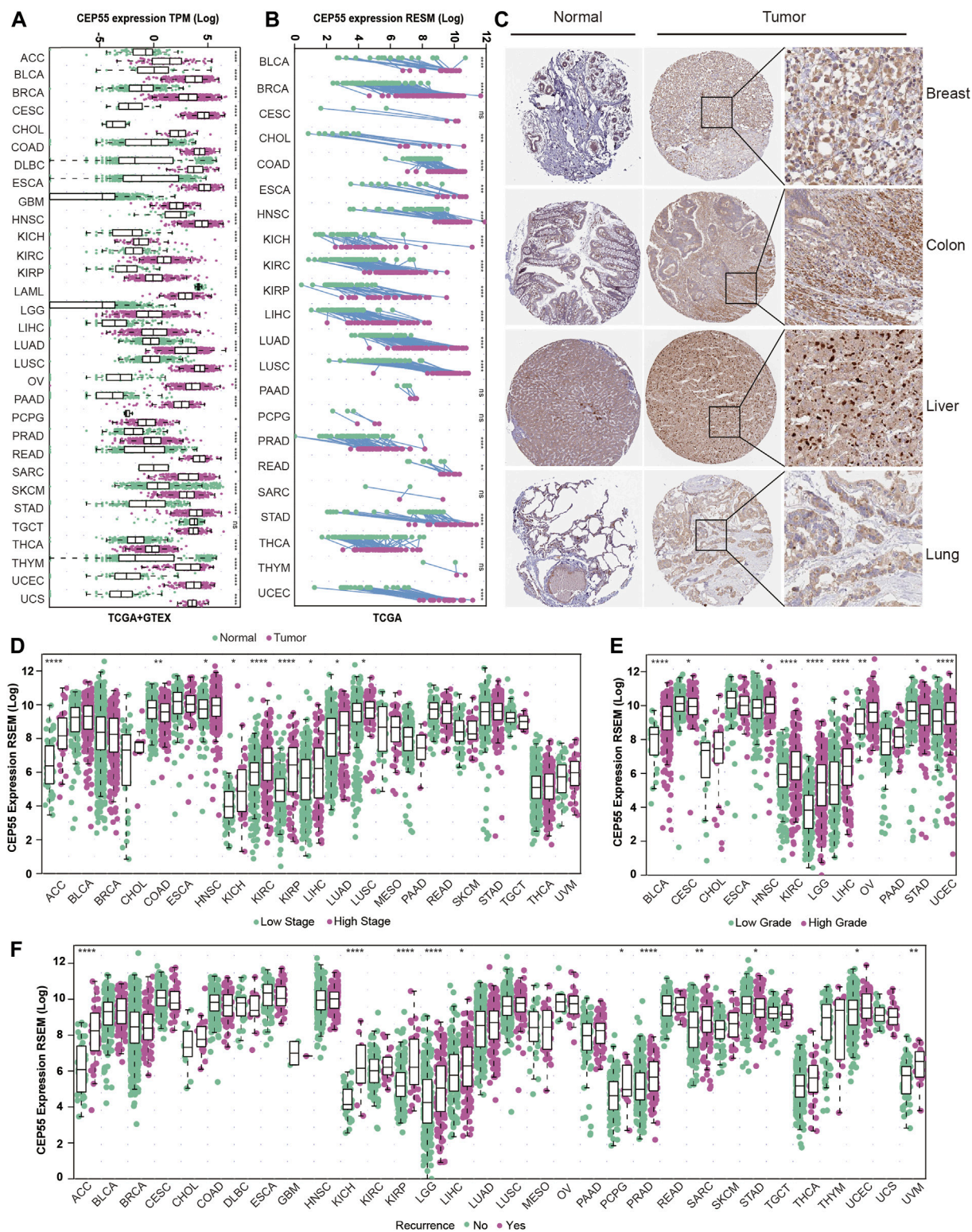
FIGURE 2
CEP55 expression in human cancers. **(A)** CEP55 expression in tumor and normal tissues based on the integrated data from TCGA and GTEx datasets. **(B)** The expression of CEP55 in tumor and paired tumor-adjacent normal tissues based on the TCGA dataset. **(C)** Immunohistochemical staining of CEP55 in cancers. The images are from the HPA (www.proteinatlas.org) database. **(D)** The expression of CEP55 in high- and low-stage tumors. **(E)** The expression of CEP55 in high- and low-grade tumors. **(F)** The expression of CEP55 in recurrent and non-recurrent tumors. Asterisks indicate statistical p-values (ns, $p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and $****p < 0.0001$).
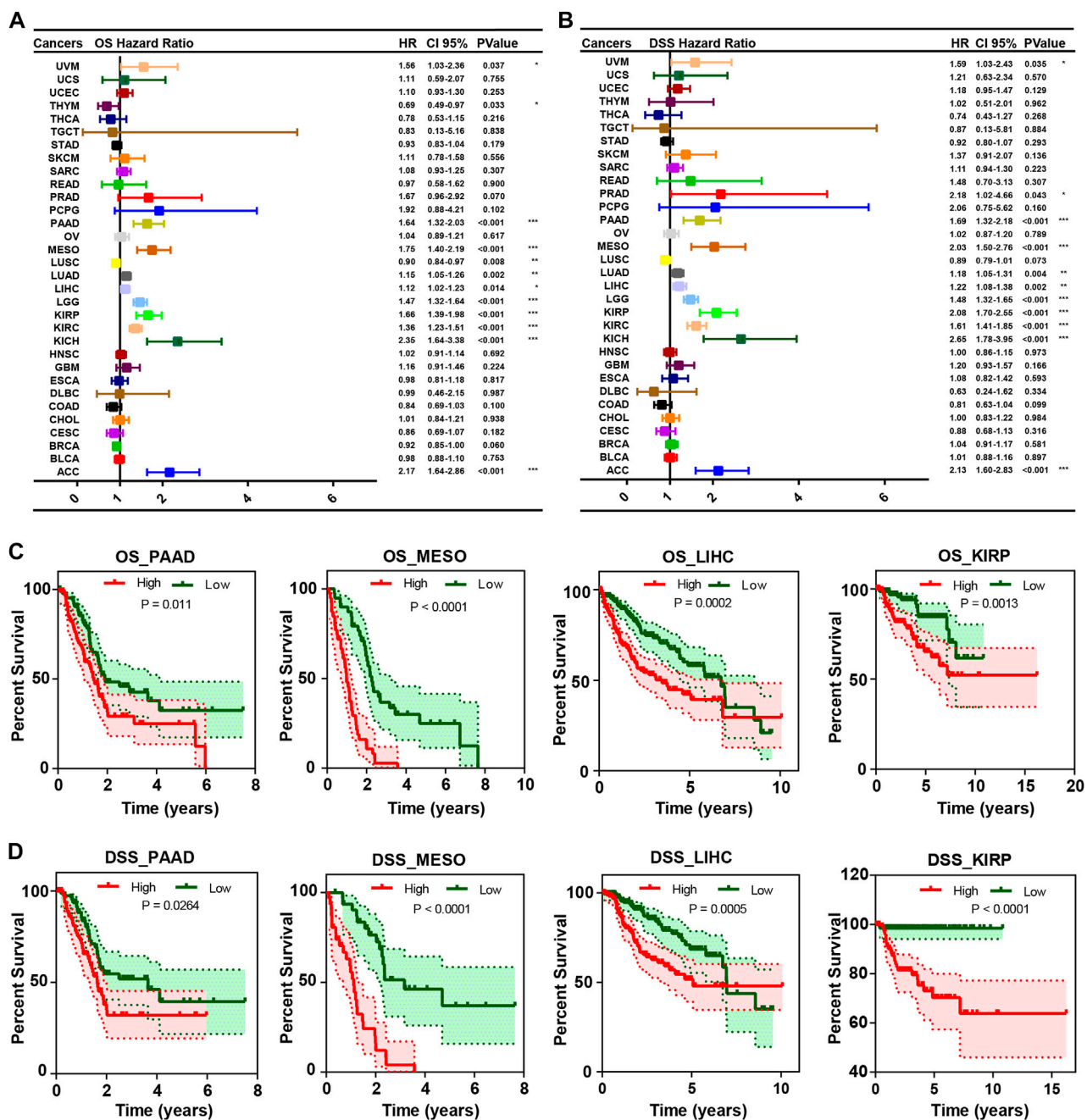
FIGURE 3
The correlation between CEP55 expression and overall survival (OS) and disease-specific survival (DSS). **(A)** The correlation between CEP55 expression and overall survival (OS). **(B)** The correlation between CEP55 expression and disease-specific survival (DSS). **(C)** Kaplan-Meier overall survival curves of CEP55 in PAAD, MESO, LIHC, and KIRP. **(D)** Kaplan-Meier disease-specific survival curves of CEP55 in PAAD, MESO, LIHC, and KIRP. Asterisks indicate statistical *p*-values (ns, *p* > 0.05, **p* < 0.05, ***p* < 0.01, and ****p* < 0.001).

(HR = 1.61, *p* < 0.001), KICH (HR = 2.66, *p* < 0.001), and ACC (HR = 2.13, *p* < 0.001) (Figure 3B; Supplementary Table S3). KM survival curves showed better OS in the low-CEP55 expression group for PAAD (HR = 1.71, *p* = 0.011), MESO (HR = 4.86, *p* < 0.001), LIHC (HR = 2.00, *p* = 0.001), KIRP (HR = 2.64, *p* = 0.0013), ACC (HR = 8.63, *p* < 0.001), LGG (HR = 3.07, *p* < 0.001), KIRC (HR = 1.91, *p* < 0.001), LUAD (HR = 1.51, *p* = 0.003), and PRAD (HR = 7.07, *p* = 0.031); however, high OS was observed in the high-

CEP55 group for THYM (HR = 0.21, *p* = 0.030) and STAD (HR = 0.67, *p* = 0.009) (Figure 3C; Supplementary Figure S3A). In PAAD (HR = 1.70, *p* = 0.026), MESO (HR = 6.5, *p* < 0.001), LIHC (HR = 2.21, *p* < 0.001), KIRP (HR = 6.26, *p* < 0.001), ACC (HR = 8.01, *p* < 0.001), BRCA (HR = 1.54, *p* = 0.038), KIRC (HR = 2.78, *p* < 0.001), LGG (HR = 3.30, *p* < 0.001), LUAD (HR = 1.68, *p* = 0.003), PRAD (HR = 6.48, *p* = 0.038), and UVM (HR = 2.56, *p* = 0.034), the low-CEP55 expression group had higher DSS, while for STAD (HR =

FIGURE 4
Mutation characteristics of CEP55 in cancers. **(A)** The alteration frequencies of CEP55 in various cancers. **(B)** The mutation counts of CEP55 in various cancers. **(C)** The mutation sites of CEP55. **(D)** Kaplan−Meier curve for OS in the altered and unaltered groups. **(E)** Kaplan−Meier curve for DSS in the altered and unaltered groups. **(F)** The correlations between CEP55 expression and TMB (purple) and MSI (green). Asterisks indicate statistical p-values (ns, $p > 0.05$, *$p < 0.05$, and **$p < 0.01$).

0.59, $p = 0.012$), the high expression group showed better DSS rates (Figure 3D; Supplementary Figure S3B). These results indicate that higher CEP55 expression is generally associated with worse prognoses for the majority of tumors.

3.4 Mutation landscape of CEP55 in pan-cancer

We inspected the mutation profiles of CEP55 using the cBioPortal database to investigate genomic alterations. Genetic variations in CEP55 were detected in 20 of the 33 cancer types. The highest alteration frequency was observed with UCEC, with 6.24% of all UCEC cases (Figure 4A). "Mutation" was the primary

variation type in most of the cancers, such as UCEC, BLCA, and SKCM, while "amplification" and "deep deletion" were predominant in other cancers, such as UCS, PRAD, and SARC. The relationships between alteration type and CEP55 expression in pan-cancer are shown in Figure 4B. A total of 86 putative non-synonymous mutation sites were detected between amino acids 0 and 464, including 68 missenses, 16 truncating, and 2 splice variants (Figure 4C). Moreover, patients with CEP55 variation showed a trend toward better prognosis in terms of OS and disease-free survival (DFS) (Figures 4D,E). Tumor mutation burden (TMB) and microsatellite instability (MSI) are associated with the occurrence and development of malignancies and are regarded as predictive markers for responses to immunotherapy. We then analyzed the relationships between CEP55 expression and TMB/
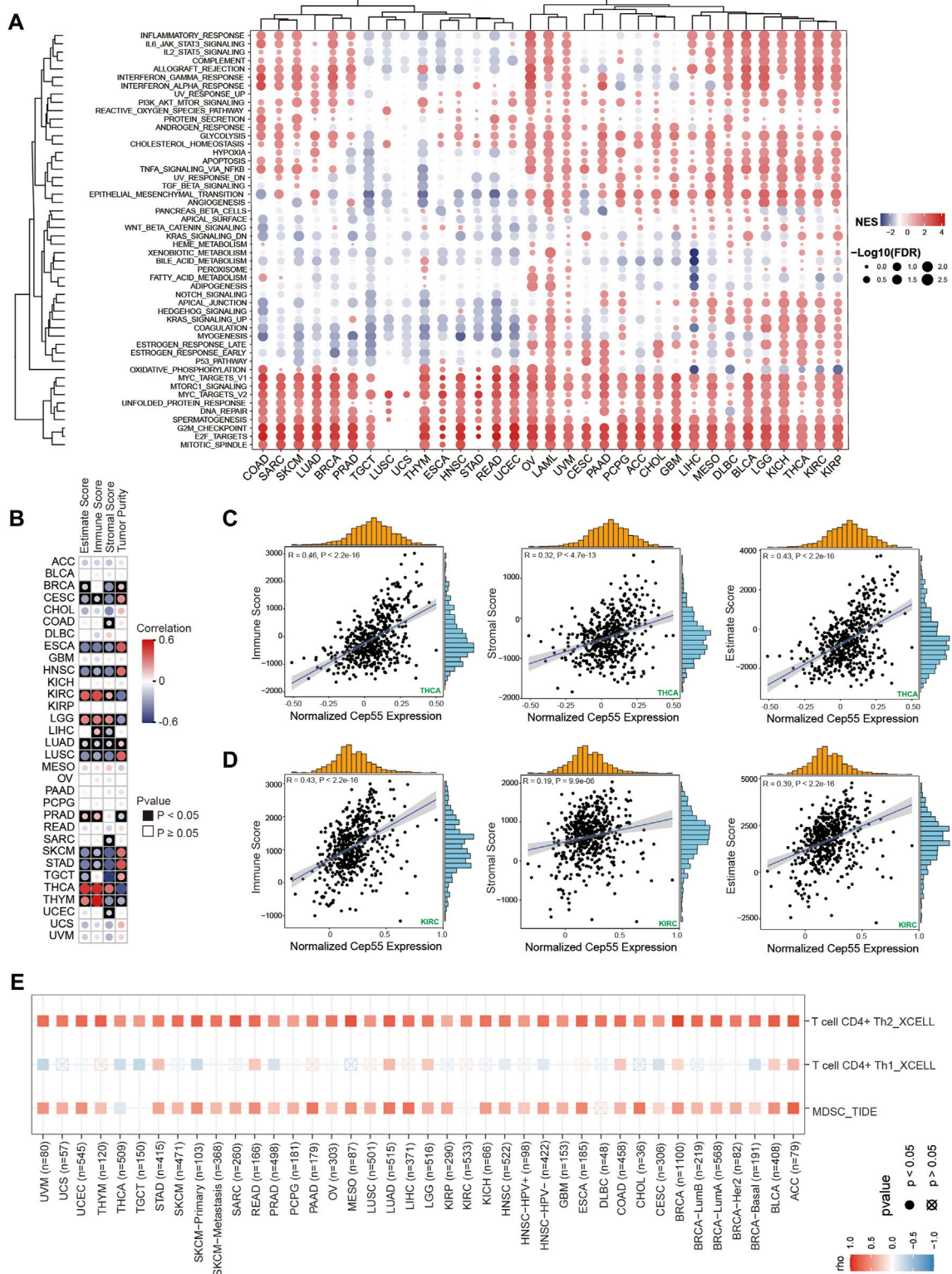
FIGURE 5
Gene set enrichment analysis (GSEA) and immune infiltration analysis in pan-cancer. **(A)** Heatmap of enrichment scores. The size of the circle represents the value of the false discovery rate (FDR), and the color represents the normalized enrichment score (NES). **(B)** The correlations between CEP55 expression and estimated scores. Red, positive correlation; Blue, negative correlation. **(C–D)** Scatter plots of correlations between CEP55 expression and immune, stromal, and estimated scores in THCA and KIRC. **(E)** Correlations between CEP55 expression and infiltration abundances of Th2 cells, Th1 cells, and MDSCs.

MSI in all cancers. CEP55 expression was positively correlated with TMB in UCEC, STAD, SARC, SKCM, READ, PAAD, PRAD, LUSC, LGG, LUAD, KICH, KIRC, COAD, HNSC, CESC, BLCA, BRCA, and ACC, and negatively correlated with TMB in THYM. In addition, the expression of CEP55 was positively associated with MSI in UCS, UCEC, STAD, COAD, SARC, READ, LUSC, and DLBC, but these were negatively correlated in SKCM and DLBC (Figure 4F, *p* < 0.05; Supplementary Table S4).

3.5 Functional analysis based on CEP55 expression

To further elucidate the biological role of CEP55 in cancer, a differential expression analysis was performed between high- and low-CEP55 expression groups for each cancer type. Gene set enrichment analysis (GSEA) was then performed using the different expression genes (DEGs) for each malignancy. Cancers with high CEP55 levels were considerably enriched in cell cycle- and proliferation-related pathways, such as the mitotic spindle, E2F targets, and G2M checkpoints, indicating that the malignancies were in a high proliferation state. Furthermore, we discovered that immune-related pathways, such as the IL6-JAK-STAT3 signaling, inflammatory response, IFNα response, and IFNγ response pathways, were significantly enriched in high-CEP55 malignancies, particularly for KICH, LGG, THCA, and LIHC (Figure 5A; Supplementary Table S5). These findings suggest that CEP55 is involved in tumor invasion and progression, as well as immunological responses.

3.6 CEP55 expression correlates with tumor immune cell infiltration

Tumor-infiltrating immune cells are directly involved in the occurrence, progression, and metastasis of malignancies. The GSEA revealed a correlation between CEP55 expression and immune-related pathways. We further explored the relationships between the expression of CEP55 and the abundance of tumor-infiltrating immune cells. First, we calculated the immune and stromal scores using the ESTIMATE algorithm. CEP55 expression was significantly correlated with immune scores in ESCA, CESC, HNSC, LGG, KIRC, LIHC, LUSC, LUAD, PRAD, SKCM, THCA, STAD, and THYM (Figures 5B–D, *p* < 0.05; Supplementary Table S6). Furthermore, to investigate the immune cells that were correlated with CEP55, we used the TIMER2.0 database, which estimates the abundance of diverse immune cells using multiple methods. The pan-cancer immune cell invasion is shown in Supplementary Figure S4. The majority of the malignancies showed substantial correlations between the expression of CEP55 and the presence of myeloid-derived suppressor cells (MDSCs) and T helper 2 (Th2) cells (Figure 5E). CEP55 expression was positively correlated with the number of MDSCs infiltrating the malignancies, including ACC (Cor = 0.769, *p* = 2.07e-15), LIHC (Cor = 0.69, *p* = 5.48e-50), and LUAD (Cor = 0.656, *p* = 5.05e-62). Moreover, CEP55 expression was positively correlated with the infiltration abundance of Th2 cells in ACC (Cor = 0.809, *p* = 4.86e-18), BLCA (Cor = 0.778, *p* = 9.62e-76),

BRCA (Cor = 0.873, *p* = 1.45e-311), BRCA-LumA (Cor = 0.745, *p* = 1.17e-92), BRCA-LumB (Cor = 0.715, *p* = 2.31e-31), LUAD (Cor = 0.762, *p* = 6.73e-95), MESO (Cor = 0.821, *p* = 6.21e-22), READ (Cor = 0.703, *p* = 1.11e-14), SARC (Cor = 0.79, *p* = 2.86e-53), SKCM-Primary (Cor = 0.773, *p* = 1.91e-21), SARC (Cor = 0.764, *p* = 3.26e-23), and UCEC (Cor = 0.716, *p* = 5.60e-15). These profiles show that CEP55 expression is selectively correlated with the infiltration of immune cell populations into tumors and may serve as a key regulator in the tumor microenvironment.

3.7 CEP55 correlates with immunomodulators and predicts response to cancer immunotherapy

To better characterize the mechanistic relationship between CEP55 and the tumor immunological milieu, we assessed the correlation between CEP55 expression and a collection of immunomodulators (Thorsson et al., 2018). CEP55 expression was positively correlated with most immunomodulatory variables in OV, THCA, KIRC, LGG, LIHC, MESO, BLCA, BRCA, PRAD, and UVM, but negatively correlated with STAD, PAAD, ESCA, and THYM (Figure 6A; Supplementary Table S7). We further studied whether CEP55 expression affected patient response to ICIs by detecting the differences in CEP55 expression between ICI response (CR/PR) and non-response (SD/PD) patients. For urinary tumors treated with PDL1 inhibitors (the IMvigor210 cohort), CEP55 expression differed significantly between the CR/PR and SD/PD groups (Figure 6B), with the CR/PR group exhibiting higher CEP55 expression than the SD/PD group (Figure 6C). When patients with urinary malignancies were treated with a PDL1 inhibitor, those with high CEP55 expression outlived the patients with low CEP55 expression (Figure 6D). Patients with high CEP55 expression responded to the PDL1 inhibitor at a rate of 28.2%, which was considerably greater than the 17.4% shown in patients with low CEP55 expression (Figure 6E). CEP55 and PDL1 expression showed a strong positive correlation (Figure 6F). Moreover, in the GSE91061 melanoma cohort treated with PD1 and CTLA4 inhibitors, CEP55 expression decreased considerably after ICI treatment in the CR/PR group (Figure 6G). These findings indicate that, in addition to its role in carcinogenesis, as previously documented, CEP55 may also be involved in the modulation of the tumor immune microenvironment; therefore, it has the potential as a biomarker to predict immunotherapy response.

3.8 Construction and validation of CEP55-Related risk model in HCC

Two recent studies have identified certain functions of CEP55 in HCC. Li et al. demonstrated that CEP55 regulated the JAK2-STAT3-MMPs signaling pathway and promoted HCC cell migration and invasion (Li et al., 2018). Yang et al. revealed that CEP55 participated in SPAG5-mediated proliferation and migration of HCC cells and provided a viable therapeutic target for the clinical treatment of HCC (Yang et al., 2018). These findings indicate that increased CEP55 expression could be

FIGURE 6

CEP55 in association with immunomodulators predicts cancer immunotherapy response. **(A)** Heatmap of correlation efficiency of CEP55 with immunoregulators. Red, positive correlation; Blue, negative correlation. **(B)** Differences in CEP55 expression between responders and non-responders in multiple immunotherapy cohorts. **(C)** CEP55 expression in CR/PR and SD/PD patients of the IMvigor210 cohort. **(D)** Kaplan-Meier curve of the low and high-CEP55 subgroups in the IMvigor210 cohort. **(E)** The proportion of responders and non-responders in the high and low-CEP55 expression groups of the IMvigor210 cohort. **(F)** The correlation between CEP55 expression and PDL1 expression in the IMvigor210 cohort. **(G)** CEP55 expression changes after ICI treatment in the GSE91061 cohort. Asterisks indicate statistical p-values (ns, $p > 0.05$, $*p < 0.05$, $**p < 0.01$, and $***p < 0.001$).

particularly important in the context of HCC. Therefore, we investigated the effects of CEP55 on HCC. Consistent with previous findings, CEP55 was documented to be more abundant in tumor tissues than in non-tumor tissues in both the ICGC and GSE14520 datasets (Figures 7A,B). In addition, the group with high CEP55 expression had a lower survival rate (Figures 7C,D). The prognostic effect of CEP55 in HCC was further examined using univariate and multivariate Cox
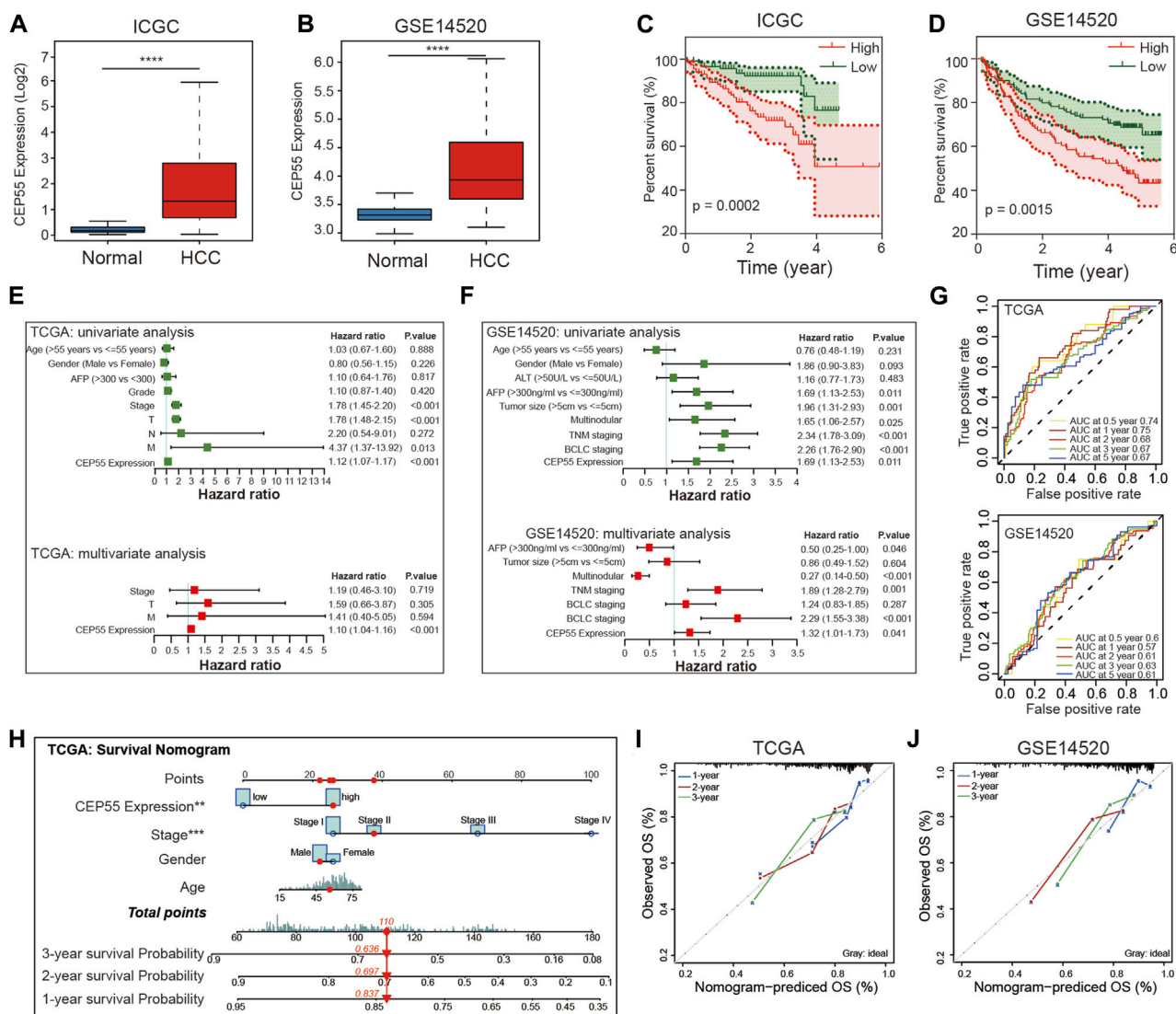
FIGURE 7
Validation of CEP55 expression and construction of prognostic risk model for HCC. **(A-B)** CEP55 expression in normal and tumor tissues in the ICGC and GSE14520 datasets. **(C-D)** Kaplan-Meier curve of the high and low CEP55 expression groups in the ICGC and GSE14520 datasets. **(E-F)** Univariate and multivariate Cox regression analyses for CEP55 expression and clinical features in TCGA and GSE14520 datasets. **(G)** Time-dependent ROC curve analysis to assess the predictive efficacy of CEP55 in TCGA and GSE14520 datasets. **(H)** Nomogram for quantitatively predicting the probability of 1-, 2-, and 3-year OS for HCC patients. **(I-J)** Calibration plots of nomogram in TCGA and GSE14520 datasets. Asterisks indicate statistical p-values (ns, $p > 0.05$, *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$).

regression analyses of HCC datasets. The univariate (HR = 1.12, $p < 0.001$) and multivariate (HR = 1.10, $p < 0.001$) analyses of TCGA LIHC dataset showed that CEP55 expression was significantly correlated with prognosis. Similar outcomes were found using the GSE14520 dataset with univariate (HR = 1.69, $p = 0.011$) and multivariate (HR = 1.32, $p = 0.041$) Cox regression analyses (Figures 7E,F). Moreover, a time-dependent ROC curve analysis was performed to evaluate the predictive classification efficiency of CEP55 expression in HCC. In TCGA dataset, the area under the curve (AUC) values for 0.5-, 1-, 2-, 3-, and 5-year overall survival were 0.74, 0.75, 0.68, 0.67, and 0.67, respectively, whereas the corresponding values of the GSE14520 dataset were

0.6, 0.57, 0.61, 0.63, and 0.61 (Figure 7G). These results indicate that CEP55 expression may help predict the short- and long-term survival status of patients with HCC. We then merged the separate prognostic markers to create a nomogram that doctors can use as a quantitative technique to predict death in HCC patients (Figure 7H). Each patient can be assigned a total score by adding the scores for each prognostic criterion. Higher overall scores would indicate a worse prognosis for the patient. In addition, the calibration graphs of several cohorts (TCGA and GSE14520) revealed that the nomogram performed similarly to the ideal model and was capable of explaining a large portion of patient outcomes (Figures 7I,J).

FIGURE 8
Relationships between CEP55 expression and HCC molecular subtypes in **(A)** TCGA and **(B)** GSE14520 datasets. Asterisks indicate statistical *p*-values
(ns, *p* > 0.05, **p* < 0.05, ***p* < 0.01, and ****p* < 0.001).
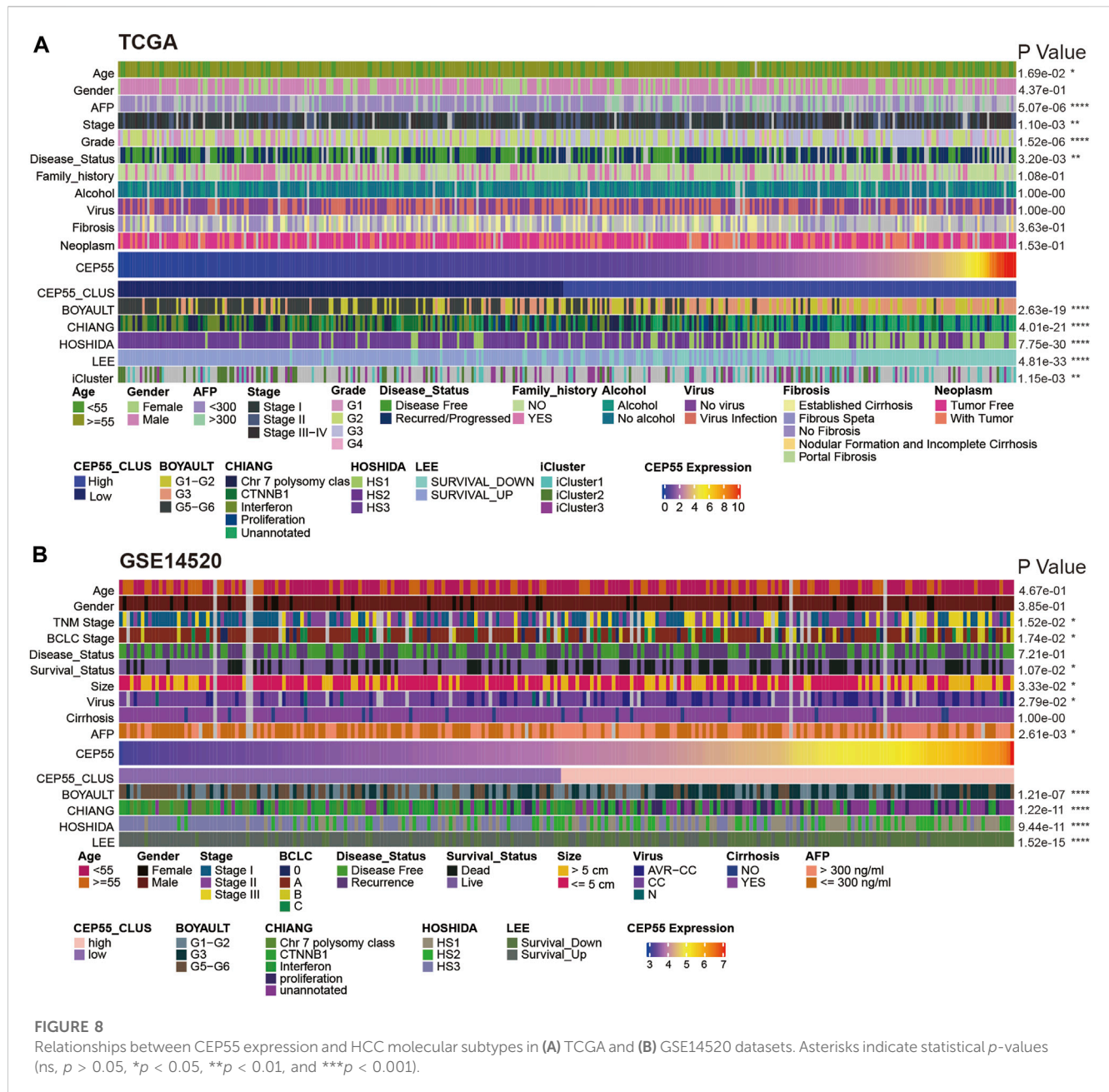
3.9 CEP55 expression is linked to HCC subclass

Subsequently, using data from TCGA LIHC cohort and the GSE14520 dataset, we examined the association of tumor-associated clinicopathological characteristics with CEP55 expression (Figures 8A,B). These components were significantly correlated in HCC, based on the chi-squared test. In TCGA cohort, CEP55 expression was correlated with age, serum AFP level, pathologic stage, histologic grade, and tumor recurrence (*p* < 0.05). Similarly, in the GEO cohort, CEP55 expression was associated with tumor stage, survival status, tumor size, viral infection, and serum AFP level (*p* < 0.05). Several studies have reported that molecular-based stratification of HCC could identify disease subtypes with

differential outcomes (Lee et al., 2004; Boyault et al., 2007; Hoshida et al., 2009), and the relationships between CEP55 expression and these subtypes were analyzed. In TCGA cohort, increased CEP55 expression was substantially related to Boyault's G3 subclass (*p* < 0.001), Chiang's unannotated subclass (*p* < 0.001), Hoshida's1 subclass (*p* < 0.001), Lee's Survival_Down subclass (*p* < 0.001), and TCGA iCluster1 subclass (*p* < 0.001). Low CEP55 expression was associated with Boyault's G5/G6 subclass (*p* < 0.001), Chiang's CTNNB1 subclass (*p* < 0.001), Hoshida's2 subclass (*p* < 0.001), Lee's Survival_Up subclass (*p* < 0.001), and TCGA iCluster2 subclass (*p* < 0.001). Similarly, increased CEP55 expression in the GEO cohort was substantially related to Boyault's G3 subclass (*p* < 0.001), Chiang's unannotated subclass (*p* < 0.001), Hoshida's1 subclass (*p* < 0.001), and Lee's Survival_

FIGURE 9

Prediction and validation of potential therapeutic compounds. **(A)** Heatmap of scores for each drug in the CMAP database for various cancers. Red, negative score; Blue, positive score. **(B)** The mechanisms of action (MoA) that were shared by the compounds. **(C–E)** Docking position of the CEP55 predicted active pocket with cytochalasin B, palbociclib, and bisbenzimide.
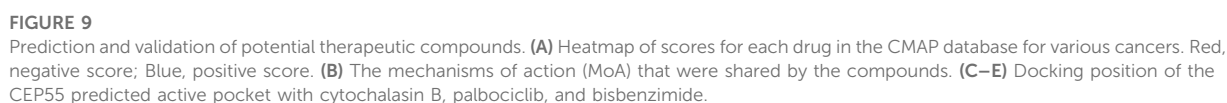
Down subclass ($p < 0.001$). Low CEP55 expression was associated with Boyault's G5/G6 subclass ($p < 0.001$), Chiang's CTNNB1 subclass ($p < 0.001$), Hoshida's2 subclass ($p < 0.001$), and Lee's Survival_Up subclass ($p < 0.001$). Based on these results, we showed that the expression of CEP55 may predict an HCC prognosis and that CEP55 exhibits diverse patterns of expression in various molecular subtypes of this disease.

3.10 Prediction and validation of potential therapeutic compounds

We screened CEP55-associated inhibitors and components for each cancer type using the connectivity map (CMap) database. Various substances, including purvalanol-A, NCH-51, and ISOX, demonstrated therapeutic potential against CEP55 expression (Figure 9A; Supplementary Table S8). Topoisomerase, CDK, HDAC, and PI3K inhibitors were common mechanisms among small-molecule drugs (Figure 9B). We used a molecular docking approach to determine whether these small molecules could attach to CEP55. Cytochalasin B and CEP55 had a binding free energy of −7.16, and cytochalasin B formed hydrogen bonds with CEP55 via GLU-179 and GLN-183 (Figure 9C). Palbociclib and CEP55 had a binding free energy of −7.38 and hydrogen bonds were formed between these through GLU-206, ARG-191, and TRY-804 (Figure 9D). Bisbenzimide and CEP55 showed a binding free energy of −7.00, and these hydrogen bonds were created via LYS-196 (Figure 9E). Thus, these clinically available compounds could bind CEP55, and further studies should be conducted to evaluate the outcomes of CEP55 inhibition in a tumor context. The design of new molecular inhibitors, particularly proteolysis-targeting chimeras, may also be able to use these as reference scaffolds.

4 Discussion

CEP55 has been identified as a critical factor in cellular abscission, and further research revealed that it is abnormally expressed in various cancers and is implicated in cancer cell proliferation, invasion, and migration (Jeffery et al., 2016). CEP55 knockdown arrests the cell cycle at the G2/M phase and suppresses gastric cancer cell expansion (Tao et al., 2014). Furthermore, CEP55 drives the migration and invasion of oral cavity squamous cell carcinoma by increasing FOXM1 and MMP-2 activity (Chen et al., 2009). Recently, researchers demonstrated that CEP55 promoted the epithelial-mesenchymal transition (EMT) and activated the PI3K/AKT/mTOR pathway in renal cell cancer (Chen et al., 2019). Collectively, these reports show that apart from regulating the cell cycle, CEP55 may also actively participate in the central cell processes required for the occurrence of malignancies. The possible involvement of CEP55 in the tumor microenvironment warrants further investigation.

In the present study, we investigated CEP55 expression and its relationship with pan-cancer prognoses. CEP55 expression was revealed to be considerably higher in tumor tissues than non-cancerous tissues in the following cancers: ACC, BRCA, BLCA, COAD, CESC, CHOL, DLBC, GBM, ESCA, KICH, KIRP, KIRC, LIHC, LUSC, LUAD, OV, READ, PAAD, SKCM, SARC, STAD,

THCA, and UCEC. In addition, overexpression of this protein has been reported in association with numerous malignancies, including gastric carcinoma, lung cancer, renal cell cancer, breast cancer, and esophageal squamous cell carcinoma (Ma et al., 2003; Jones et al., 2005; Sakai et al., 2006; Yan et al., 2021). CEP55 is involved in multiple malignancy risk signatures for predicting cancer prognosis, progression, and chemotherapy response, and is within the main 70 overexpressed genes in cancers with chromosomal instability (the CIN70 signature) that are frequently used to predict clinical prognoses and distant metastases (Carter et al., 2006). CEP55 is also part of the PAM50 signature, which is used for breast cancer categorization and prognostic prediction. Our findings show that CEP55 may be a carcinogenic indicator and a viable prognostic biomarker for a variety of malignancies.

In addition to its involvement in cell cycle and proliferation-related pathways, this study showed that CEP55 was closely associated with immune-related pathways, including the IL6-JAK-STAT3 signaling and the IFNα/γ response pathways. In addition, this protein is highly expressed in the lymph nodes, tonsils, and other immune-associated tissues. This suggests that CEP55 could have a more active role in immune modulation than those previously reported. We discovered that abnormal CEP55 expression was substantially linked to the invasion of MDSCs and Th2 cells in most malignancies. MDSCs have a diverse population of cells that are descended from myeloid cells, and activated MDSCs produce the immune-suppressive factors arginase 1, iNOS, and ROS (Gabrilovich and Nagaraj, 2009). T-helper 1 (Th1) and Th2 cells are the major functional subsets of CD4+ T cells (Chatzileontiadou et al., 2020; Zhu and Zhu, 2020), and the two subgroups differ in terms of cytokine production and immunological responses. Th1 cells boost the cytotoxic effects of NK and CD8+ T cells by secreting Th1-type cytokines such as TNFα, IL2, and IFNγ (Mosmann and Coffman, 1989; Murphy et al., 2000; Alspach et al., 2019), whereas Th2 cells produce immunosuppressive cytokines, including IL-4, IL-5, IL-10, and IL-13. Th2-type cytokines inhibit Th1 cell development and the immune response, thereby limiting antitumor immunity (Protti and De Monte, 2012; Nakayama et al., 2017; Basu et al., 2021). Balancing Th1/Th2 cells is essential for maintaining immune homeostasis, and a change in this ratio toward increased Th2 cell infiltration could promote cancer development and weaken the immune system (Kidd, 2003; Ruterbusch et al., 2020). According to the results of this investigation, CEP55 expression was positively linked to Th2 cell infiltration, suggesting that CEP55 may be the catalyst or promoter of this Th1/Th2 balance shift. In addition, the current study indicated that CEP55 could be implicated in the modulation of the immunosuppressive microenvironment in malignancies and can be a marker for predicting responses to ICIs. However, the fundamental mechanisms by which CEP55 participates in immune modulation and immunotherapy responses remain unknown.

HCC is the fourth leading cause of cancer-related mortality globally and the most common primary liver cancer (Villanueva, 2019). The elevated CEP55 expression in tumor tissue and its association with poor survival in HCC identified in this study corroborate previously published results. Overexpression of this protein causes AKT phosphorylation and activation, which promotes survival and tumor formation in HCC (Chen et al., 2007). HCC is a heterogeneous malignancy of hepatocytes that is characterized by the accumulation of many genomic and epigenomic changes that have undergone Darwinian selection. Effective precision medicine for HCC therapy

can be achieved by the accurate molecular subtyping of HCC (Rebouissou and Nault, 2020). We demonstrated that distinct hepatocellular carcinoma molecular subtypes can be identified based on CEP55 expression. Owing to the high levels of CEP55 expression in exosomes, CEP55 may be a viable biomarker for liquid biopsies to predict survival and direct precision medicine for HCC patients. Moreover, CEP55 expression was significantly inversely associated with specific metabolic pathways in HCC, such as xenobiotic and bile/fatty acid metabolism. Metabolic reprogramming has been identified as one of the variables driving cancer aggression and affecting neoadjuvant chemotherapy response (Wu et al., 2022). Therefore, a greater understanding of the function of CEP55 in metabolic regulation is required.

We detected possible inhibitors for the regulation of CEP55 expression or the targeting of CEP55-related pathways and confirmed these using molecular docking analyses. Three small-molecule compounds were identified that had a strong affinity for CEP55: cytochalasin B, which is a G-actin superimposed inhibitor that can suppress the growth of many types of cancer (Croop and Holtzer, 1975); palbociclib, which is a CDK4/6 inhibitor licensed by the FDA for the treatment of ER+, HER2-breast cancer (Konar et al., 2022), and given that CDK4/6 are critical cell cycle regulators, off-target effects of palbociclib on CEP55 may contribute to its efficacy in triggering tumor cell cycle arrest; and bisbenzimide, which is involved in numerous pharmacological actions, including anticancer, antiparasitic, antibacterial, antifungal, antiviral, and chemosensor activities (Verma et al., 2020). However, the exact effects of these three small compounds on CEP55 remain to be further investigated. We expect that by binding to CEP55, these three molecules will alleviate the immunosuppressive microenvironment and enhance the antitumor activity of ICIs.

Several studies have investigated the specific roles of CEP55 in malignancies. Fu et al. demonstrated that CEP55 was a diagnostic and predictive factor in patients with LUAD and LUSC (Fu et al., 2020). Yang et al. discovered that CEP55 was overexpressed in liver cancer tumor tissue and linked to poor prognoses and enhanced immune infiltration (Yang et al., 2020). However, these studies did not thoroughly analyze the role of CEP55 in the tumor immune microenvironment. Yang et al. found that CEP55 was correlated with the infiltration of B cells, CD4+ T cells, CD8+ T cells, macrophages, neutrophils, and dendritic cells in liver cancer. Through our pan-cancer analysis, we found that CEP55 was significantly correlated with MDSCs and Th1 cells in most cancers and positively correlated with the expression of checkpoints, such as PD1, CTLA4, LAG3, and TIGIT in certain cancer types. This indicates that CEP55 overexpression may be a driving factor for CD8+ T cell exhaustion. Furthermore, we used molecular docking analysis to predict and validate chemicals that could bind to CEP55, which has not been previously reported.

This study had some limitations. Although we validated the results of this work using a large number of datasets, additional experiments and clinical investigations are required to identify the precise function of CEP55 in carcinogenesis and development. Furthermore, we discovered that the expression of CEP55 is associated with an imbalance in Th1/Th2 cells; however, it is unclear how CEP55 controls the differentiation of Th1 and Th2 cells and additional research is required to elucidate the underlying processes. In addition, the potential of tumor-expressing CEP55 as a TAA for modulating immune activity

should be explored, possibly through the identification and clonal tracing of CEP55 antigen-specific T cells in the context of human tumors.

In conclusion, our analysis provides a comprehensive assessment of CEP55 expression in various cancers. CEP55 is a powerful tumor prognostic marker that is involved in tumor immune modulation. Therefore, CEP55 has potential as a therapeutic target.

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

Conception and design: XX, ZL, and TW; acquisition and organization of data: XX, HL, WJ, XM, and YW; analysis and interpretation of data: XX, ZL, and TW; technical and material support: XX, HL, and ZT; drafting of the manuscript: XX, ZL, and TW; All authors contributed to the article and approved the submitted version.

Funding

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.2023.1198557/full#supplementary-material

References

Alspach, E., Lussier, D. M., and Schreiber, R. D. (2019). Interferon Γ and its important roles in promoting and inhibiting spontaneous and therapeutic cancer immunity. *Cold Spring Harb. Perspect. Biol.* 11 (3), a028480. doi:10.1101/cshperspect.a028480

Basu, A., Ramamoorthi, G., Albert, G., Gallen, C., Beyer, A., Snyder, C., et al. (2021). Differentiation and regulation of T(H) cells: A balancing act for cancer immunotherapy. *Front. Immunol.* 12, 669474. doi:10.3389/fimmu.2021.669474

Bonilla, B., Hengel, S. R., Grundy, M. K., and Bernstein, K. A. (2020). Rad51 gene family structure and function. *Annu. Rev. Genet.* 54, 25–46. doi:10.1146/annurev-genet-021920-092410

Boyault, S., Rickman, D. S., de Reyniès, A., Balabaud, C., Rebouissou, S., Jeannot, E., et al. (2007). Transcriptome classification of hcc is related to gene alterations and to new therapeutic targets. *Hepatol. Baltim. Md)* 45 (1), 42–52. doi:10.1002/hep.21467

Carter, S. L., Eklund, A. C., Kohane, I. S., Harris, L. N., and Szallasi, Z. (2006). A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat. Genet.* 38 (9), 1043–1048. doi:10.1038/ng1861

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

Chatzileontiadou, D. S. M., Sloane, H., Nguyen, A. T., Gras, S., and Grant, E. J. (2020). The many faces of Cd4(+) T cells: Immunological and structural characteristics. *Int. J. Mol. Sci.* 22 (1), 73. doi:10.3390/ijms22010073

Chen, C. H., Chien, C. Y., Huang, C. F., Chuang, H. C., Fang, F. M., et al. (2009). Expression of Flj10540 is correlated with aggressiveness of oral cavity squamous cell carcinoma by stimulating cell migration and invasion through increased Foxm1 and mmp-2 activity. *Oncogene* 28 (30), 2723–2737. doi:10.1038/onc.2009.128

Chen, C. H., Lu, P. J., Chen, Y. C., Fu, S. L., Wu, K. J., Tsou, A. P., et al. (2007). Flj10540-Elicited cell transformation is through the activation of pi3-kinase/akt pathway. *Oncogene* 26 (29), 4272–4283. doi:10.1038/sj.onc.1210207

Chen, H., Zhu, D., Zheng, Z., Cai, Y., Chen, Z., and Xie, W. (2019). Cep55 promotes epithelial-mesenchymal transition in renal cell carcinoma through pi3k/akt/mtor pathway. *Clin. Transl. Oncol. official Publ. Fed. Span. Oncol. Soc. Natl. Cancer Inst. Mexico* 21 (7), 939–949. doi:10.1007/s12094-018-02012-8

Cisowski, J., Sayin, V. I., Liu, M., Karlsson, C., and Bergo, M. O. (2016). Oncogene-induced senescence underlies the mutual exclusive nature of oncogenic kras and braf. *Oncogene* 35 (10), 1328–1333. doi:10.1038/onc.2015.186

Croop, J., and Holtzer, H. (1975). Response of myogenic and fibrogenic cells to cytochalasin B and to colcemid. I. Light microscope observations. *J. Cell Biol.* 65 (2), 271–285. doi:10.1083/jcb.65.2.271

Fabbro, M., Zhou, B. B., Takahashi, M., Sarcevic, B., Lal, P., Graham, M. E., et al. (2005). Cdk1/Erk2-and plk1-dependent phosphorylation of a centrosome protein, Cep55, is required for its recruitment to midbody and cytokinesis. *Dev. Cell* 9 (4), 477–488. doi:10.1016/j.devcel.2005.09.003

Fassl, A., Geng, Y., and Sicinski, P. (2022). Cdk4 and Cdk6 kinases: From basic science to cancer therapy. *Science* 375, eabc1495. doi:10.1126/science.abc1495

Fu, L., Wang, H., Wei, D., Wang, B., Zhang, C., Zhu, T., et al. (2020). The value of Cep55 gene as a diagnostic biomarker and independent prognostic factor in luad and lusc. *PloS one* 15 (5), e0233283. doi:10.1371/journal.pone.0233283

Gabrilovich, D. I., and Nagaraj, S. (2009). Myeloid-derived suppressor cells as regulators of the immune system. *Nat. Rev. Immunol.* 9 (3), 162–174. doi:10.1038/nri2506

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cbioportal. *Sci. Signal.* 6 (269), pl1. doi:10.1126/scisignal.2004088

Gao, X. Y., and Wang, X. L. (2015). An adoptive T cell immunotherapy targeting cancer stem cells in a colon cancer model. *J. BUON official J. Balkan Union Oncol.* 20 (6), 1456–1463.

Hoshida, Y., Nijman, S. M., Kobayashi, M., Chan, J. A., Brunet, J. P., Chiang, D. Y., et al. (2009). Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res.* 69 (18), 7385–7392. doi:10.1158/0008-5472.Can-09-1089

Hughes, P. E., Caenepeel, S., and Wu, L. C. (2016). Targeted therapy and checkpoint immunotherapy combinations for the treatment of cancer. *Trends Immunol.* 37 (7), 462–476. doi:10.1016/j.it.2016.04.010

Inoda, S., Hirohashi, Y., Torigoe, T., Nakatsugawa, M., Kiriyama, K., Nakazawa, E., et al. Cep55/C10orf3, a tumor antigen derived from a centrosome residing protein in breast carcinoma. *J. Immunother. Hagerst. Md 1997)* (2009) 32(5):474–485. doi:10.1097/CJI.0b013e3181a1d109

Inoda, S., Morita, R., Hirohashi, Y., Torigoe, T., Asanuma, H., Nakazawa, E., et al. (2011). The feasibility of Cep55/C10orf3 derived peptide vaccine therapy for colorectal carcinoma. *Exp. Mol. pathology* 90 (1), 55–60. doi:10.1016/j.yexmp.2010.10.001

Jeffery, J., Sinha, D., Srihari, S., Kalimutho, M., and Khanna, K. K. (2016). Beyond cytokinesis: The emerging roles of Cep55 in tumorigenesis. *Oncogene* 35 (6), 683–690. doi:10.1038/onc.2015.128

Jones, J., Otu, H., Spentzos, D., Kolia, S., Inan, M., Beecken, W. D., et al. (2005). Gene signatures of progression and metastasis in renal cell cancer. *Clin. Cancer Res.* 11 (16), 5730–5739. doi:10.1158/1078-0432.Ccr-04-2225

Kidd, P. (2003). Th1/Th2 balance: The hypothesis, its limitations, and implications for health and disease. *Altern. Med. Rev. a J. Clin. Ther.* 8 (3), 223–246.

Konar, D., Maru, S., Kar, S., and Kumar, K. (2022). Synthesis and clinical development of palbociclib: An overview. *Med. Chem. (Shariqah (United Arab. Emir.* 18 (1), 2–25. doi:10.2174/1573406417666201204161243

Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., et al. (2006). The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Sci. (New York, NY)* 313 (5795), 1929–1935. doi:10.1126/science.1132939

Lee, J. S., Chu, I. S., Heo, J., Calvisi, D. F., Sun, Z., Roskams, T., et al. (2004). Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. *Hepatol. Baltim. Md)* 40 (3), 667–676. doi:10.1002/hep.20375

Lens, S. M. A., and Medema, R. H. (2019). Cytokinesis defects and cancer. *Nat. Rev. Cancer* 19 (1), 32–45. doi:10.1038/s41568-018-0084-6

Li, M, Gao, J., Li, D., and Yin, Y. (2018). Cep55 promotes cell motility via Jak2⁻Stat3⁻Mmps cascade in hepatocellular carcinoma. *Cells* 7 (8), 99. doi:10.3390/cells7080099

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). Timer: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.Can-17-0307

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

Ma, X. J., Salunga, R., Tuggle, J. T., Gaudet, J., Enright, E., McQuary, P., et al. (2003). Gene expression profiles of human breast cancer progression. *Proc. Natl. Acad. Sci. U. S. A.* 100 (10), 5974–5979. doi:10.1073/pnas.0931261100

Montero-Conde, C., Martín-Campos, J. M., Lerma, E., Gimenez, G., Martínez-Guitarte, J. L., Combalía, N., et al. (2008). Molecular profiling related to poor prognosis in thyroid carcinoma. Combining gene expression data and biological information. *Oncogene* 27 (11), 1554–1561. doi:10.1038/sj.onc.1210792

Mosmann, T. R., and Coffman, R. L. (1989). Th1 and Th2 cells: Different patterns of lymphokine secretion lead to different functional properties. *Annu. Rev. Immunol.* 7, 145–173. doi:10.1146/annurev.iy.07.040189.001045

Murphy, K. M., Ouyang, W., Farrar, J. D., Yang, J., Ranganath, S., Asnagli, H., et al. (2000). Signaling and transcription in T helper development. *Annu. Rev. Immunol.* 18, 451–494. doi:10.1146/annurev.immunol.18.1.451

Nakayama, T., Hirahara, K., Onodera, A., Endo, Y., Hosokawa, H., Shinoda, K., et al. (2017). Th2 cells in health and disease. *Annu. Rev. Immunol.* 35, 53–84. doi:10.1146/annurev-immunol-051116-052350

Pesch, A. M., Hirsh, N. H., Michmerhuizen, A. R., Jungles, K. M., Wilder-Romans, K., Chandler, B. C., et al. (2022). Rb expression confers sensitivity to cdk4/6 inhibitor-mediated radiosensitization across breast cancer subtypes. *JCI Insight* 7 (3), e154402. doi:10.1172/jci.insight.154402

Protti, M. P., and De Monte, L. (2012). Cross-talk within the tumor microenvironment mediates Th2-type inflammation in pancreatic cancer. *Oncoimmunology* 1 (1), 89–91. doi:10.4161/onci.1.1.17939

Qadir, F., Aziz, M. A., Sari, C. P., Ma, H., Dai, H., Wang, X., et al. (2018). Transcriptome reprogramming by cancer exosomes: Identification of novel molecular targets in matrix and immune modulation. *Mol. Cancer* 17 (1), 97. doi:10.1186/s12943-018-0846-5

Rebouissou, S., and Nault, J. C. (2020). Advances in molecular classification and precision oncology in hepatocellular carcinoma. *J. Hepatol.* 72 (2), 215–229. doi:10.1016/j.jhep.2019.08.017

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 (7), e47. doi:10.1093/nar/gkv007

Ruterbusch, M., Pruner, K. B., Shehata, L., and Pepper, M. (2020). *In vivo* CD4[+] T cell differentiation and function: Revisiting the Th1/Th2 paradigm. *Annu. Rev. Immunol.* 38, 705–725. doi:10.1146/annurev-immunol-103019-085803

Sakai, M., Shimokawa, T., Kobayashi, T., Matsushima, S., Yamada, Y., Nakamura, Y., et al. (2006). Elevated expression of C10orf3 (chromosome 10 open reading frame 3) is involved in the growth of human colon tumor. *Oncogene* 25 (3), 480–486. doi:10.1038/sj.onc.1209051

Schöning-Stierand, K., Diedrich, K., Fährrolfes, R., Flachsenberg, F., Meyder, A., Nittinger, E., et al. (2020). Proteinsplus: Interactive analysis of protein-ligand binding interfaces. *Nucleic Acids Res.* 48 (W1), W48–w53. doi:10.1093/nar/gkaa235

Sharma, P., Hu-Lieskovan, S., Wargo, J. A., and Ribas, A. (2017). Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* 168 (4), 707–723. doi:10.1016/j.cell.2017.01.017

Siegel, R. L., Miller, K. D., Wagle, N. S., and Jemal, A. (2023). Cancer statistics, 2023. *CA Cancer J. Clin.* 73 (1), 17–48. doi:10.3322/caac.21763

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, J., Zhi, X., Tian, Y., Li, Z., Zhu, Y., Wang, W., et al. (2014). Cep55 contributes to human gastric carcinoma by regulating cell proliferation. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* 35 (5), 4389–4399. doi:10.1007/s13277-013-1578-1

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 (4), 812–830.e14. doi:10.1016/j.immuni.2018.03.023

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Sci. (New York, NY)* 347 (6220), 1260419. doi:10.1126/science.1260419

Veres, D. V., Gyurkó, D. M., Thaler, B., Szalay, K. Z., Fazekas, D., Korcsmáros, T., et al. (2015). Comppi: A cellular compartment-specific database for protein-protein interaction network analysis. *Nucleic Acids Res.* 43 (Database issue), D485–D493. doi:10.1093/nar/gku1007

Verma, S., Ravichandiran, V., Ranjan, N., and Flora, S. J. S. (2020). Recent advances in therapeutic applications of bisbenzimidazoles. *Med. Chem. (Shariqah (United Arab. Emir.* 16 (4), 454–486. doi:10.2174/1573406415666190416120801

Villanueva, A. (2019). Hepatocellular carcinoma. *N. Engl. J. Med.* 380 (15), 1450–1462. doi:10.1056/NEJMra1713263

Vivian, J., Rao, A. A., Nothaft, F. A., Ketchum, C., Armstrong, J., Novak, A., et al. (2017). Toil enables reproducible, open source, big biomedical data analyses. *Nat. Biotechnol.* 35 (4), 314–316. doi:10.1038/nbt.3772

Wu, Y., Yang, S., Ma, J., Chen, Z., Song, G., Rao, D., et al. (2022). Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. *Cancer Discov.* 12 (1), 134–153. doi:10.1158/2159-8290.Cd-21-0316

Xu, H., Yu, S., Liu, Q., Yuan, X., Mani, S., Pestell, R. G., et al. (2017). Recent advances of highly selective cdk4/6 inhibitors in breast cancer. *J. Hematol. Oncol.* 10 (1), 97. doi:10.1186/s13045-017-0467-2

Yan, S. M., Liu, L., Gu, W. Y., Huang, L. Y., Yang, Y., Huang, Y. H., et al. (2021). Cep55 positively affects tumorigenesis of esophageal squamous cell carcinoma and is correlated with poor prognosis. *J. Oncol.* 2021, 8890715. doi:10.1155/2021/8890715

Yang, L., He, Y., Zhang, Z., and Wang, W. (2020). Upregulation of Cep55 predicts dismal prognosis in patients with liver cancer. *BioMed Res. Int.* 2020, 4139320. doi:10.1155/2020/4139320

Yang, Y. F., Zhang, M. F., Tian, Q. H., Fu, J., Yang, X., Zhang, C. Z., et al. (2018). Spag5 interacts with Cep55 and exerts oncogenic activities via pi3k/akt pathway in hepatocellular carcinoma. *Mol. Cancer* 17 (1), 117. doi:10.1186/s12943-018-0872-3

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). Clusterprofiler: An R package for comparing biological themes among gene clusters. *Omics a J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zhao, W. M., Seki, A., and Fang, G. (2006). Cep55, a microtubule-bundling protein, associates with centralspindlin to control the midbody integrity and cell abscission during cytokinesis. *Mol. Biol. Cell* 17 (9), 3881–3896. doi:10.1091/mbc.e06-01-0015

Zhu, X., and Zhu, J. (2020). Cd4 T helper cell subsets and related human immunological disorders. *Int. J. Mol. Sci.* 21 (21), 8011. doi:10.3390/ijms21218011

Frontiers in Molecular Biosciences

Effective TME-related signature to predict prognosis of patients with head and neck squamous cell carcinoma

Lingfei Wan[1,2†], Yuanshuai Li[1†*], Wenting Pan[1,2†], Yuting Yong[1,2†], Chao Yang[3†], Chen Li[1], Xingxing Zhao[1,2], Ruihong Li[1], Wen Yue[2,4*] and Xinlong Yan[1*]

[1]College of Life Science and Bioengineering, Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing, China, [2]Stem Cell and Regenerative Medicine Lab, Beijing Institute of Radiation Medicine, Beijing, China, [3]Department of Nucleus Radiation-Related Injury Treatment, PLA Rocket Force Characteristic Medical Center, Beijing, China, [4]South China Research Center for Stem Cell and Regenerative Medicine, Guangzhou, China

Introduction: The tumor microenvironment (TME) is crucial for the development of head and neck squamous cell carcinoma (HNSCC). However, the correlation of the characteristics of the TME and the prognosis of patients with HNSCC remains less known.

Methods: In this study, we calculated the immune and stromal cell scores using the "estimate" R package. Kaplan-Meier survival and CIBERSORT algorithm analyses were applied in this study.

Results: We identified seven new markers: FCGR3B, IGHV3-64, AC023449.2, IGKV1D-8, FCGR2A, WDFY4, and HBQ1. Subsequently, a risk model was constructed and all HNSCC samples were grouped into low- and high-risk groups. The results of both the Kaplan-Meier survival and receiver operating characteristic curve (ROC) analyses showed that the prognosis indicated by the model was accurate (0.758, 0.756, and 0.666 for 1-, 3- and 5-year survival rates). In addition, we applied the CIBERSORT algorithm to reveal the significant differences in the infiltration levels of immune cells between the two risk groups.

Discussion: Our study elucidated the roles of the TME and identified new prognostic biomarkers for patients with HNSCC.

1 Introduction

With approximately 600,000 new cases diagnosed annually, squamous cell carcinoma of the head and neck (HNSCC) is the sixth most common malignancy worldwide. More than 50% of HNSCC cases develop to an advanced stage with a 5-year overall survival (OS) rate of approximately 50% (Miyauchi et al., 2019; Yi et al., 2020; Siegel et al., 2021).

Immunotherapy has revolutionized the treatment of cancer, and the clinical application of immune checkpoint inhibitors (ICIs) has provided benefits to patients with various malignant tumors. A known characteristic of HNSCC is severe immunosuppression (Romano and Romero, 2015); therefore, therapy with ICIs play an important role in the treatment of HNSCC patients (Chen et al., 2021). Although many studies have suggested that patients with recurrent and

metastatic HNSCC may benefit from ICI immunotherapy, most have shown limited success in the clinical setting, with a 13%–18% overall response rate (Solomon et al., 2018; von Witzleben et al., 2020). The role of immune infiltration in the TME is important for tumorigenesis and tumor progression, both of which affect the clinical prognosis of patients with tumors (Ferris, 2015; Gavrielatou et al., 2020). Furthermore, there is increasing evidence that the tumor mutation burden (TMB) is associated with immunotherapy response (Liu et al., 2019).

Here, we comprehensively analyzed the relationship between the TME, prognosis, TMB, and ICIs in patients with HNSCC. We then established a risk model based on the TME to improve prognostic risk stratification, facilitating better treatment and decision–making for patients. Differentially expressed genes (DEGs) identified here could facilitate a more in-depth understanding of tumor progression and immunotherapy treatment. In addition, this study may help elucidate the mechanism of tumor escape and establish a framework for the development of new prognostic markers.

2 Materials and methods

2.1 Data download and processing

From The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/), we downloaded the mRNA expression, clinical information, and somatic mutation data of HNSCC samples. After obtaining the somatic mutation data, we used Perl scripts based on the JAVA 8 platform to determine the mutation frequency with number of variants/the length of exons (38 million). Meanwhile, the tumor mutation burden (TMB) value for each sample was calculated.

2.2 TME analysis

Using the "estimate" R package, we estimated the infiltration levels of immune and stromal cells, in the form of two scores, immune score and stromal score (Yoshihara et al., 2013). Meanwhile, the sum of immune and stromal score was reflected by the ESTIMATE score. We then explored the correlation between the expression levels of model genes and these scores by performing the Spearman's rank correlation coefficient test. Additionally, we employed the CIBERSORT algorithm to assess the 22 types of infiltrating immune cells of each sample (Newman et al., 2015).

2.3 Identification of differentially expressed genes (DEGs) based on the stromal and immune scores

According to the median stromal and immune scores, we divided 502 HNSCC patients into high- and low-score groups. To identify DEGs between the two score groups, we applied the "limma" R package, with a false-discovery rate (FDR) ≤ 0.05 and |log2 fold change (FC)| ≥ 1.

2.4 Construction and validation of the prognostic prediction model in HNSCC

By taking the intersection of the DEGs from the both score groups, the univariate Cox analysis was conducted to primarily

screen out immune- and stromal-related genes with prognostic value, using the "survival" R package. A least absolute shrinkage and selection operator (LASSO) analysis was further applied to narrow these prognostic genes. Finally, a multivariate Cox regression model was utilized to select candidate genes related to survival and to construct the prediction model. The risk score was then calculated as follows: risk score = (0.2086 × expression level of FCGR3B) + (−0.0550 × expression level of IGHV3-64) + (−1.8215 × expression level of AC023449.2) + (0.0075 × expression level of IGKV1D-8) + (0.0582 × expression level of FCGR2A) + (−0.5416 × expression level of WDFY4) + (0.0914 × expression level of HBQ1).

Based on the median risk score, we classified all HNSCC patients into low- and high-risk groups. The Kaplan-Meier (KM) survival analysis and the receiver operating characteristic (ROC) curve analyses were used to analyze the OS between the two risk groups and assess the sensitivity and specificity of the signature using the "survivalROC" and "timeROC" R packages.

2.5 Functional enrichment analysis

We carried out the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for the DEGs between the two risk groups, using the "clusterProfiler," "enrichplot," and "org.Hs.eg.db" R packages. Furthermore, we used the "GSVA" R package to perform a gene set variation analysis (GSVA) with the purpose of estimating the variation of pathway between the low- and high-risk groups, based on the "c2.cp.kegg.v7.4.symbols.gmt" database, which was downloaded from the Molecular Signatures Database (v7.4, http://www.gsea-msigdb.org/gsea/msigdb/index.jsp) (Hanzelmann et al., 2013).

2.6 TMB calculation and visualization

The somatic mutation data were obtained from the TCGA database. The TMB was defined as the total number of somatic gene coding errors, base substitutions, insertions, or deletions detected per megabyte bases of tumor tissue. The value of it for each patient was defined as the total mutation frequency/the length of the human exon (38 Mb) (Lv et al., 2020; Jiang et al., 2021). When calculating TMB, we excluded all synonymous mutations. At the same time, we further studied the mutation status under different risk groups.

2.7 Construction of the protein-protein interaction (PPI) network and the competitive endogenous RNA (ceRNA) network

We performed differential analysis for patients between high- and low-risk groups and used the differential genes to construct the PPI network by using the Search Tool for the Retrieval of Interacting Genes (STRING) database.

Construct protein-protein interaction (PPI) network. In addition, we used model genes in the Starbase database (http://

FIGURE 1
Determination of TME-related DEGs and functional analysis. Distribution of immune and stromal score by clinical characteristics, including **(A)** tumor grades, **(B)** sex, **(C)** T classification, **(D)** N classification, and **(E)** tumor stages. **(F)** Common upregulated and downregulated genes based on immune and stromal scores. **(G)** GO and **(H)** KEGG analyses of 395 common DEGs. **(I)** Correlation between the proportions of 22 types of immune cells in the TME of HNSCC patients. **(J)** Violin plots were used to display the differential infiltration of 22 types of immune cells between tumor and normal samples.

starbase.sysu.edu.cn/). The ceRNA regulatory network of model genes was screened and constructed in the database. When predicting the miRNA binding to the model gene through this database, we first ensured that there should be a negative correlation between the expression of miRNA and mRNA. At the same time, miRNA was differentially expressed in normal and tumor. In addition, using the median value of candidate miRNAs, we divided patients into high and low expression groups, and screened miRNAs with survival differences between the two groups through km database. Subsequently, we screened lncRNAs through the Starbase database. According to the theory of ceRNA, there was a positive correlation between lncRNA expression and mRNA. At the same time, the candidate lncRNA should be differentially expressed in normal and tumor tissues, and have survival differences in different expression groups based on the median expression level. According to the theory, the ceRNA networks related to the important model genes was screened and constructed.

2.8 Statistical analysis

All statistical analyses were accomplished using the R software (v4.1.1). We followed the methods of Ai-Min Jiang, Yue Zhao, and Ke-Wei Bi et al. (Bi et al., 2020; Jiang et al., 2021; Zhao et al., 2021). To compare the expression level of model genes between the tumor and normal samples, we conducted the Wilcoxon test. To explore the correlation between model gene expression levels and the OS of patients, we used the log-rank test and KM curve analysis. Meanwhile, we performed the univariate and multivariate Cox regression analyses to explore the independent prognostic value of the risk mode. p-value ≤ 0.05 was regarded as significant.

3 Results

3.1 Acquisition of DEGs based on immune and stromal scores

To elucidate the relationship between the immune and stromal scores and clinical features of HNSCC, we used the Wilcoxon test to analyze the differences among patients with different statuses. We found significant differences in immune scores according to tumor grade (Figure 1A), sex (Figure 1B), and T and N stage (Figures 1C, D). Furthermore, stromal scores were significantly different between tumor stage I and III (Figure 1E). These results showed that the immune- and stromal-related activities were associated with HNSCC progression.

Based on the median immune score, we identified 1,558 DEGs, including 1,255 upregulated and 303 downregulated genes (Supplementary Table S1). There were 1,307 DEGs, including 1,191 upregulated and 116 downregulated genes, based on the stromal score (Supplementary Table S2). At the intersection of these two sets of DEGs, 365 upregulated and 30 downregulated genes were identified (Figure 1F; Supplementary Table S3).

We then performed a gene ontology (GO) enrichment analysis on the 395 genes that may be the key factors in the TME. We found that these genes were predominantly associated with the immune

responses, such as phagocytosis, activating cell surface receptor signaling pathways, and B cell-mediated immunity (Figure 1G). The Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis gave similar results, with responses such as phagosome, NF-kappa B signaling pathway, and B cell receptor signaling pathway (Figure 1H). These results indicated that immune-related activities were important characteristics in the TME of HNSCC.

Furthermore, to identify the proportion of the 22 kinds of immune cells in the TME of patients with HNSCC, we conducted a CIBERSORT analysis, using the "CIBERSORT" R package. Using the Pearson analysis, we found that M0 macrophages negatively correlated with CD8$^+$ T cells. However, CD8$^+$ T cells positively correlated with activated memory CD4$^+$ T cells (Figure 1I). These results indicated that there were significant differences between the normal and tumor groups. The normal samples had a higher proportion of native B, resting memory CD4$^+$ T, resting mast, and resting dendritic cells than the tumor samples. Moreover, in the tumor patient group, the proportion of resting NK cells and M0 macrophages was higher than that in normal group (Figure 1J).

3.2 Establishment of the prognostic prediction model with TCGA cohort

The univariate Cox analysis of the 395 DEGs identified 50 genes significantly related to OS (Supplementary Table S4). We then used the LASSO regression analysis to screen these genes, and 13 genes were finally identified (Figures 2A, B). All HNSCC samples were then randomly divided into training and validation cohorts, at a ratio of 1:1. The 13 genes were further screened using the multivariate Cox regression analysis. Finally, a set of seven genes, FCGR3B, IGHV3-64, AC023449.2, IGKV1D-8, FCGR2A, WDFY4, and HBQ1, was selected to construct the prognostic model and calculate the risk score (Figure 2C).

Based on the median risk score, 126 and 125 patients were in the low- and high-risk groups, respectively. The results of the principal component analysis (PCA) indicated that patients at different risks were clearly separated into two groups (Figure 2D). Moreover, patients in the high-risk group had a higher death rate and shorter survival time than those in the low-risk group (Supplementary Figure S1A). Based on the KM analysis, we found that patients in the low-risk group had a significantly better OS than those in the high-risk group ($p < 0.001$; Figure 2E). The model had a good predictability for OS, with the area under the curve (AUC) being 0.758, 0.756, and 0.666 for 1, 3, and 5-year OS rates through the time-dependent ROC analysis, respectively (Figure 2F).

To test whether the risk score was independent of other clinical features, such as age, sex, tumor stage, and tumor grade, we performed univariate and multivariable Cox regression analyses. The results showed that it was independent (Figures 2G, H).

Furthermore, we established a nomogram to predict the 1-, 3-, and 5-year survival rates in patients with HNSCC, according to the expression levels of the model genes (Figure 2I). Using the calibration curve, we found that the nomogram had a good predictive value compared to the ideal model (Figure 2J). In the training cohort, the concordance index (C-index) was 0.687.
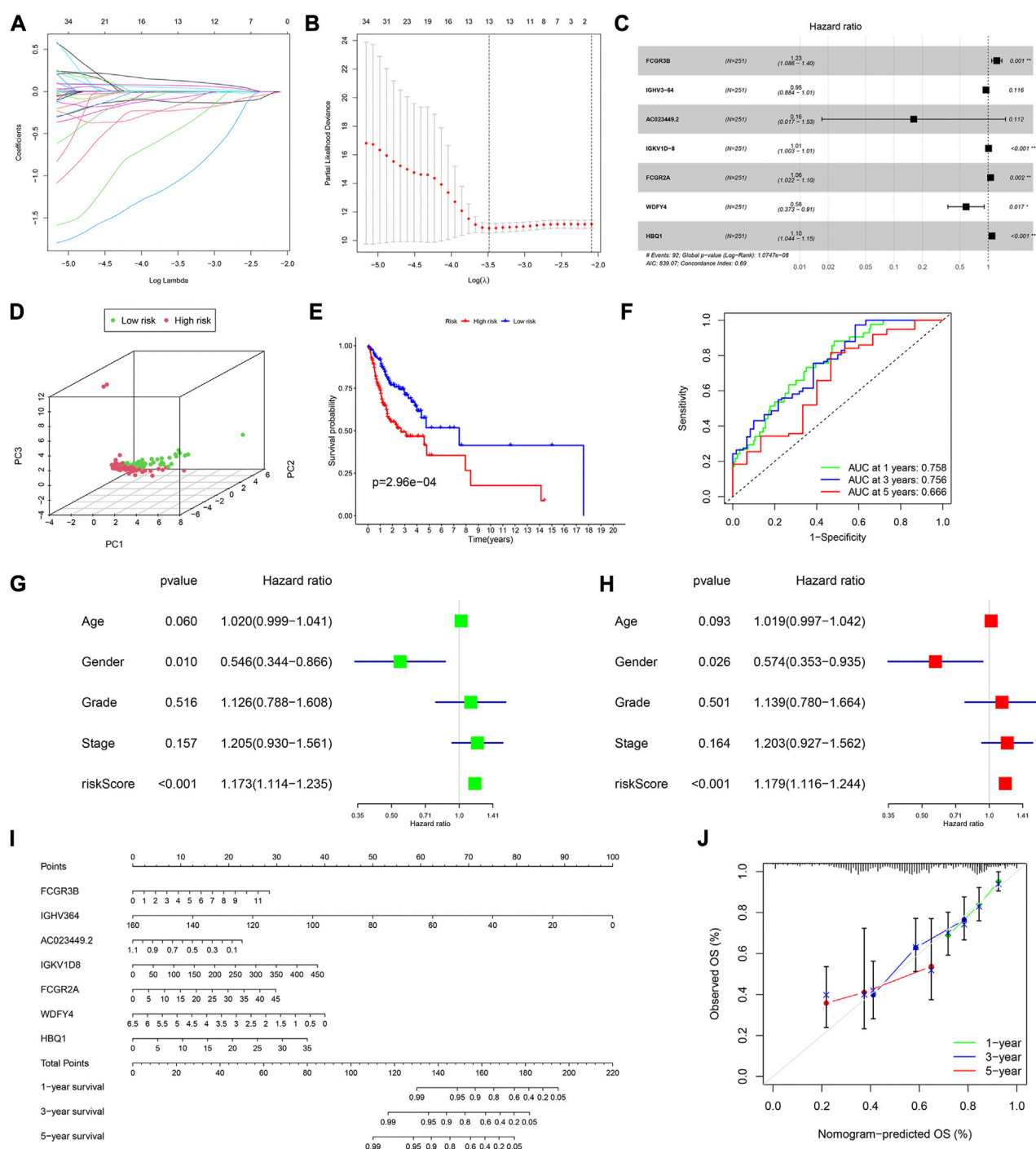
FIGURE 2
Construction of a prognostic model in the training cohort. **(A,B)** LASSO regression algorithm. **(C)** A prognostic model was constructed by the multivariate Cox regression analysis. **(D)** Principal component analysis. **(E)** Kaplan–Meier (KM) curves of OS for patients in the high- and low-risk groups, respectively. **(F)** Time-dependent ROC curve analysis of the prognostic model. **(G)** Univariate and **(H)** multivariate cox regression analyses to evaluate the prognostic signature. **(I)** Nomogram predicting the survival of HNSCC patients. **(J)** Calibration plot based on the 1-, 3-, and 5-year OS rates of the nomogram. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$.

3.3 Validation of the prognostic model

According to the median risk score, there were 133 and 115 HNSCC patients in the low and high-risk groups, respectively. The PCA showed a good separation between the risk groups (Figure 3A). The analysis of survival time and patient status in both risk groups showed consistent results (Supplementary Figure S1B). Furthermore, the p-value of the KM analysis was 0.01283 (Figure 3B), and the AUC values were 0.709, 0.647, and 0.629 for 1-,3-,5- y survival rates in HNSCC patients, respectively (Figure 3C).
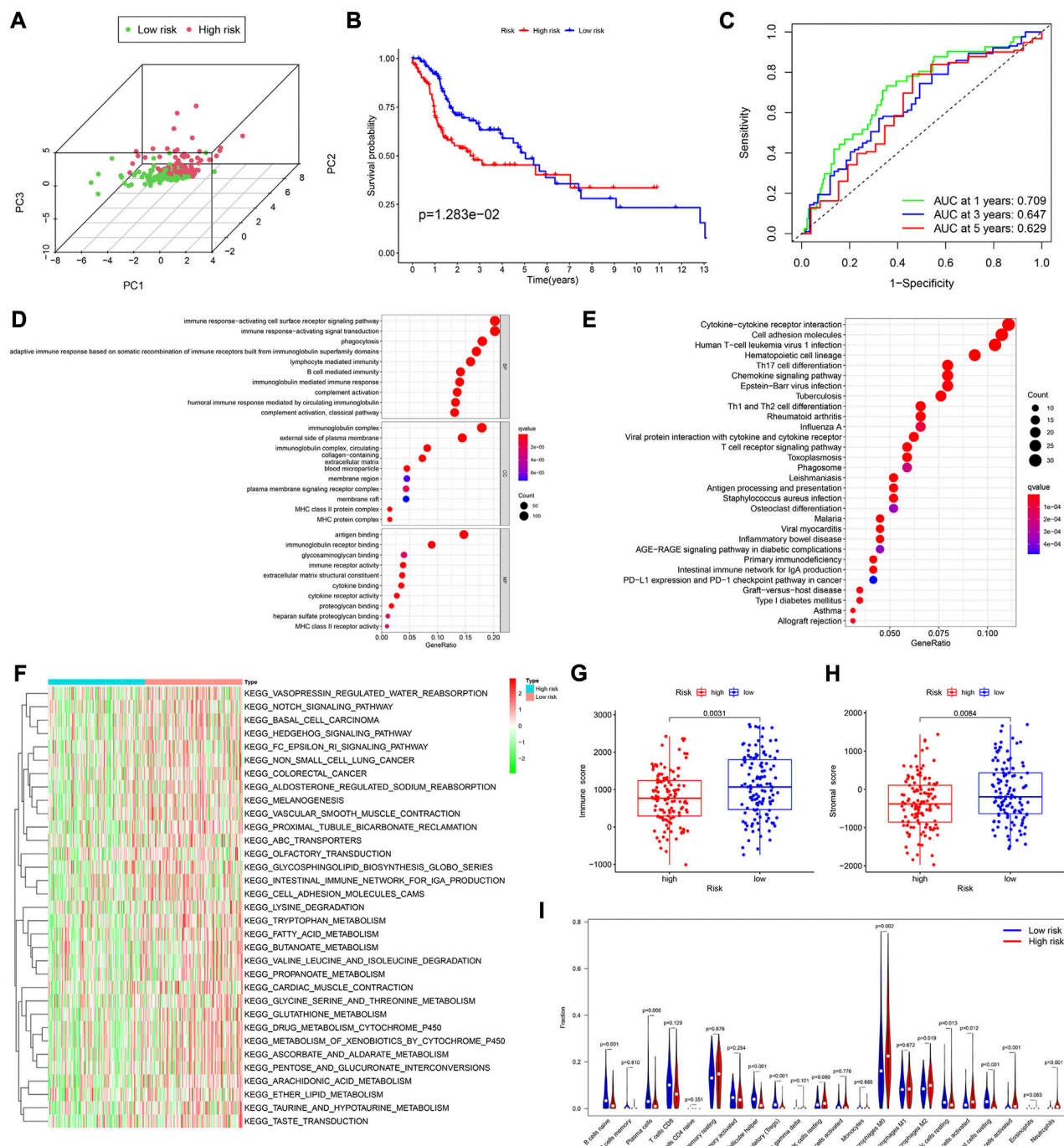
FIGURE 3
Assessment of the prognostic model in the validation cohort. **(A)** Plot of principal component analysis. **(B)** KM curve of OS for patients in the high-
and low-risk groups. **(C)** Time-dependent ROC curve analysis of the prognostic model. **(D,E)** GO and KEGG analyses of DEGs between the high- and low-
risk groups in the training cohort. **(F)** GSVA analyses to estimate the variation of pathway between the low- and high-risk groups in the training cohort.
(G,H) Distribution of immune and stromal scores between the high- and low-risk groups in the training cohort. **(I)** Comparison of infiltration levels of
22 immune cells between the high- and low-risk groups in the training cohort.

3.4 Gene set variation analysis and functional analysis based on the risk model

Using the "limma" R package, we performed a differential analysis of the two risk groups, using the following criteria:

FDR ≤0.05 and |log2FC | ≥ 1. In the training and validation cohorts, we identified 750 (Supplementary Table S5), and 755 DEGs (Supplementary Table S7), respectively. Based on these DEGs, GO enrichment and KEGG pathway analyses were performed. We found that in both cohorts, the DEGs
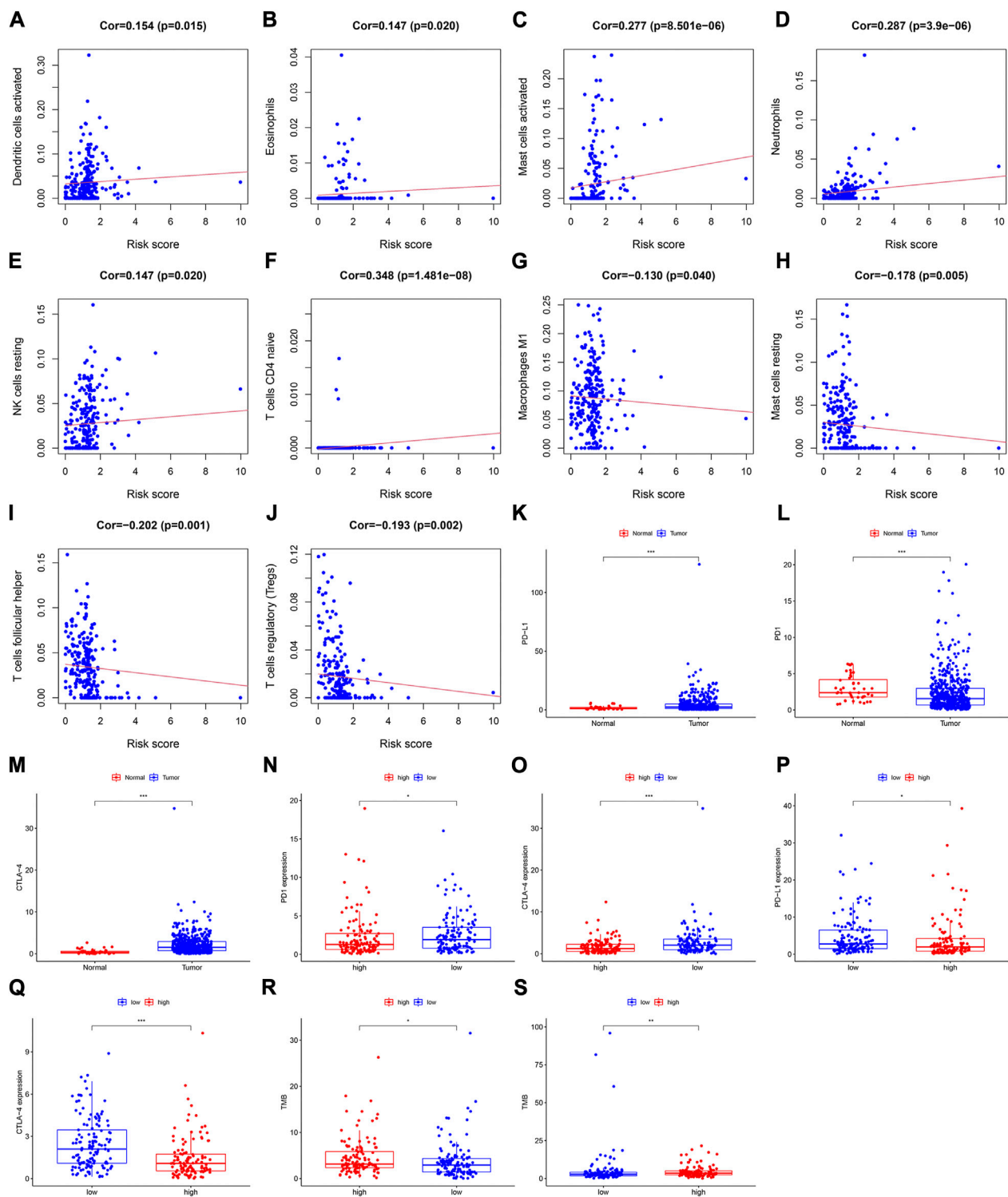
FIGURE 4
Assessment of the relationships between the risk score and infiltration levels of 22 immune cell types in the training cohort. **(A)** Activated dendritic cells, **(B)** eosinophils, **(C)** activated mast cells, **(D)** neutrophils, **(E)** resting NK cells, **(F)** naive CD4⁺ T cells, **(G)** M1 macrophages, **(H)** resting mast cells, **(I)** follicular helper T cells, and **(J)** regulatory T cells. Differential expression of **(K)** PD-L1, **(L)** PD1, and **(M)** CTLA-4 between the tumor and normal samples. Expression levels of **(N)** PD1, and **(O)** CTLA-4 between the high- and low-risk groups in the training cohort. The differential expression of **(P)** PD-L1, and **(Q)** CTLA-4 between the high- and low-risk groups in the validation cohort. The TMB value in both risk groups in the **(R)** training cohort, and **(S)** validation cohort.

were mainly associated with immune-related activities, such as humoral immune response, immune response-activating signal transduction, and immune response-activating cell surface receptor signaling pathway (Figures 3D, E; Supplementary Figures S1C–S1F).

Subsequently, the Gene set variation analysis (GSVA) was used to explore the different biological activities between the two risk groups, with FDR ≤0.05 as the criterion. The results showed that pathways related to metabolism, such as fatty acid metabolism, glycine, serine and threonine metabolism, and ascorbate and aldarate metabolism, were significantly enriched (Figure 3F; Supplementary Figure S2C).

3.5 Analysis of immune cell infiltration between the two risk groups

We found that in both cohorts, patients in the low-risk group had a stronger immune response (Figure 3G; Supplementary Figure S2B) and higher stromal scores (Figure 3H; Supplementary Figure S2C) than those in the high-risk groups, using the Wilcoxon signed-rank test. To explore the differences in immune cells, we used the deconvolution algorithm CIBERSORT. The results indicated that in the training cohort, native B ($p < 0.001$), plasma ($p = 0.006$), follicular helper T ($p < 0.001$), regulatory T ($p < 0.001$), resting mast ($p < 0.001$), and resting dendritic cells ($p < 0.001$) were significantly more abundant in the low-risk group than those in the high-risk group, whereas M0 macrophages ($p = 0.002$), M2 macrophages ($p = 0.019$), activated mast cells ($p < 0.001$), neutrophils ($p < 0.001$), and activated dendritic cells ($p = 0.012$) were less abundant (Figure 3I). In the validation cohort, similar results about the immune status were obtained (Supplementary Figure S2D).

A Pearson analysis was used to analyze the relationship between the risk score and infiltration levels of the 22 immune cell types. In the training cohort, the risk scores had a significantly positive correlation with activated dendritic cells (Figure 4A), eosinophils (Figure 4B), activated mast (Figure 4C), neutrophils (Figure 4D), resting NK (Figure 4E), and naïve CD4+ T cells (Figure 4F). However, the risk score was negatively correlated with M1 macrophages (Figure 4G), resting mast (Figure 4H), follicular helper T (Figure 4I), and regulatory T cells (Figure 4J).

In the validation cohort, the risk score was positively correlated with eosinophils (Supplementary Figure S2E), activated mast cells (Supplementary Figure S2F), neutrophils (Supplementary Figure S2G), and resting NK cells (Supplementary Figure S2H), whereas it was negatively correlated with naïve B (Supplementary Figure S2I), resting mast (Supplementary Figure S2J), CD8+ T (Supplementary Figure S2K), and regulatory T cells (Supplementary Figure S2L).

3.6 Association of immune checkpoint molecules with the prognosis prediction model

To explore the relationship between the immune checkpoint molecules and the prognostic model, we evaluated the differential expression of checkpoint molecules in the two risk groups. Compared with the normal tissues, PD-L1 and CTLA-4 expression levels were upregulated in HNSCC tissues ($p < 0.001$; Figures 4K, M), whereas PD1 expression levels were downregulated ($p < 0.001$); (Figure 4L). In both cohorts, the expression level of CTLA-4 in the low-risk group was significantly higher than that in the high-risk group (Figures 4O, Q). In the training cohort, the expression levels of PD1 in the low-risk group were significantly higher than those in the high-risk group (Figure 4N), whereas PD-L1 was more highly expressed in the low-risk group than the high-risk group in the validation cohort (Figure 4P).

These results indicated that the expression levels of immune checkpoint molecules were higher in the low-risk group than those in the high-risk group. Therefore, the prognostic model may provide effective predictive biomarkers, which will enable the optimization of immune checkpoint therapies.

3.7 Mutation analysis and visualization

In the different risk groups of both cohorts, we found that there was a difference in TMB. Namely, in both cohorts, the TMB of the high-risk group was higher than that of the low-risk group (Figures 4R, S).

We utilized the "maftools" R package to analyze and visualize the somatic mutation profiles of 478 HNSCC patients. The detailed mutation information of each sample was illustrated via a waterfall plot, and different mutation types were distinguished by various color annotations. We found that missense mutations, single-nucleotide polymorphism (SNP), and C > T mutations comprised the vast majority of the classification categories. Additionally, the median value of mutations in the samples was 78, and the maximum mutations was 2,393 (Figure 5A). We then presented the number of variant classifications in different samples using box plots. The top 10 mutated genes were TP53 (66%), TTN (35%), FAT1 (21%), CDKN2A (20%), MUC16 (17%), CSMD3 (16%), NOTCH1 (16%), PIK3CA (16%), SYNE1 (15%), and LRP1B (14%) (Figure 5B).

We also investigated the somatic mutation status of different risk groups in the two cohorts. The results showed that the top 10 mutated genes in the four groups differed. In the high-risk group of the training cohort, the top 10 mutated genes were TP53 (72%), TTN (36%), FAT1 (25%), CDKN2A (25%), CSMD3 (18%), PIK3CA (16%), MUC16 (15%), KMT2D (15%), NOTCH1 (15%), and PCLO (15%) (Figure 5C). In the low-risk group, they were TP53 (63%), TTN (34%), FAT1 (21%), CDKN2A (19%), PIK3CA (18%), CSMD3 (17%), MUC16 (16%), SYNE1 (14%), DNAH5 (13%), and NOTCH1 (12%) (Figure 5D). In the high-risk group of the validation cohort, 114 patients possessed somatic mutations, and the top 10 mutated genes were TP53 (68%), TTN (36%), FAT1 (24%), CDKN2A (23%), NOTCH1 (22%), MUC16 (19%), SYNE1 (18%), KMT2D (17%), LRP1B (17%), and CSMD3 (16%) (Figure 5E). In the low-risk group, they were TP53 (62%), TTN (32%), MUC16 (17%), FAT1 (16%), PIK3CA (16%), CSMD3 (16%), LRP1B (16%), SYNE1 (16%), CDKN2A (15%), and NOTCH1 (14%) (Figure 5F).
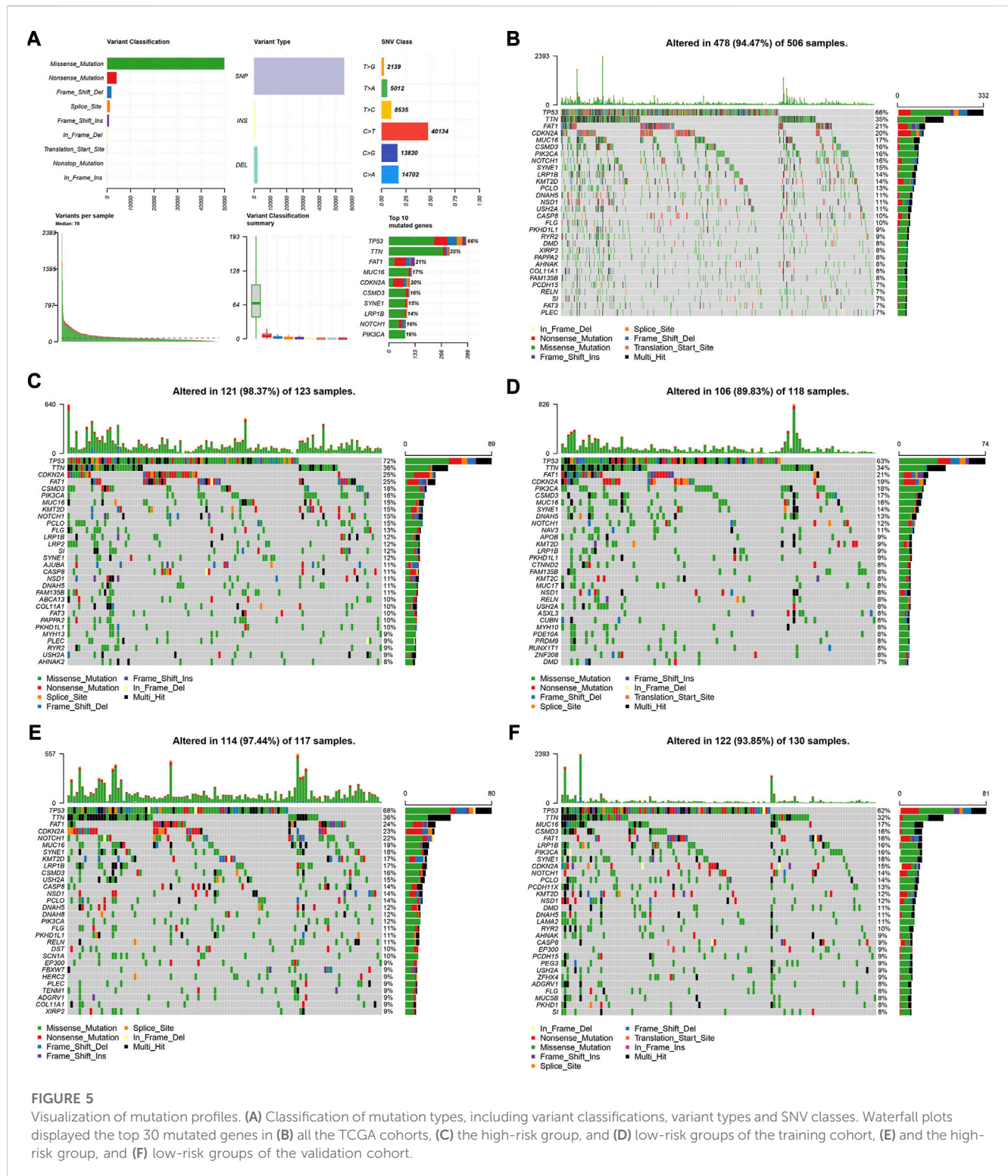
FIGURE 5
Visualization of mutation profiles. **(A)** Classification of mutation types, including variant classifications, variant types and SNV classes. Waterfall plots displayed the top 30 mutated genes in **(B)** all the TCGA cohorts, **(C)** the high-risk group, and **(D)** low-risk groups of the training cohort, **(E)** and the high-risk group, and **(F)** low-risk groups of the validation cohort.

3.8 Acquisition of core genes and establishment of the competitive endogenous RNA (ceRNA) network

To elucidate the biological relationships among the 395 DEGs, we used the Search Tool for the Retrieval of Interacting Genes (STRING) database (https://www.string-db.org/), based on genes

with co-expression coefficients higher than 0.9 (Figure 6A). We identified 130 genes with strong mutual correlations, and also identified the top 30 genes according to the number of degrees between the two pairs (Figure 6B).

We identified two common genes, FCGR2A and FCGR3B, which were located at the intersection of the 130 STRING-identified and model genes (Figure 6C). Using the Wilcoxon test,
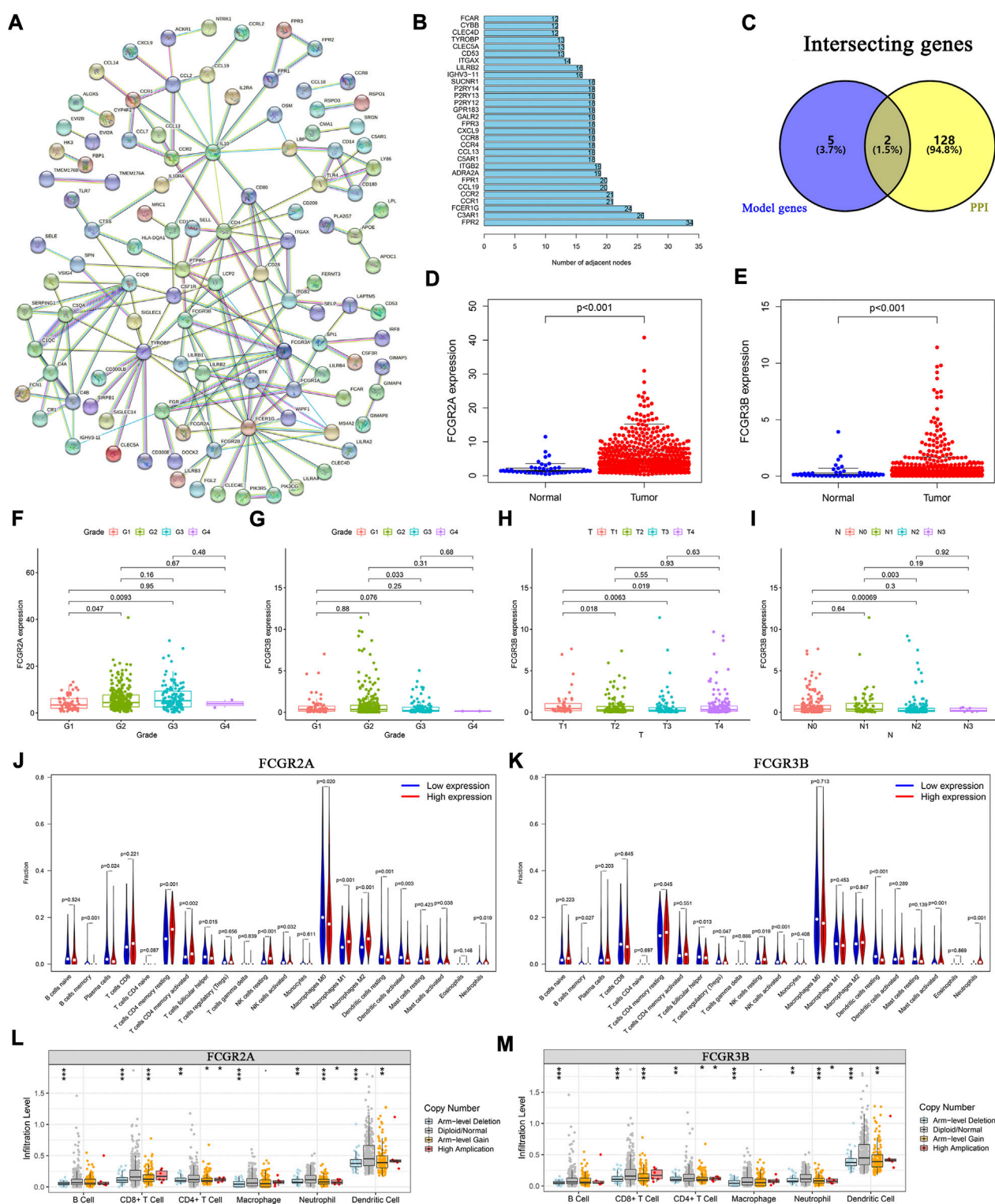
FIGURE 6
Protein-protein interaction network and the landscape of core genes. **(A)** Interaction network of 395 DEGs. **(B)** The top 30 core elements calculated by the number of degrees. **(C)** Common genes at the intersection of 130 genes and model genes. The differential expression of **(D)** FCGR2A, and **(E)** FCGR3B in tumor and normal samples. Distribution of FCGR2A and FCGR3B in **(F, G)** tumor grades, **(H)** T and **(I)** N classifications. Comparison of infiltration levels of 22 immune cells in the high- and low-expression groups based on the median expression levels of **(J)** FCGR2A, and **(K)** FCGR3B. Analysis of immune cell infiltration levels and somatic copy number alterations in **(L)** FCGR2A, and **(M)** FCGR3B.

we found that the expression levels of FCGR2A and FCGR3B were significantly higher in HNSCC samples than those in normal samples (Figures 6D, E). Furthermore, using the paired-sample test analysis, we found that the expression of FCGR2A significantly differed between normal and tumor tissue (Supplementary Figure S3A). However, no significant survival differences were observed between the high- and low-expression groups based on the median expression levels of FCGR2A (Supplementary Figure S3B) and FCGR3B (Supplementary Figure S3C).

We then examined the expression of the two genes under different clinical conditions. The results showed that FCGR2A was significantly differentially expressed between tumor grade I and II (Figure 6F). Moreover, the expression level of FCGR3B was significantly different between tumor grades (Figure 6G), T (Figure 6H) and N stages (Figure 6I). In both cohorts, the results indicated that the expression of FCGR3B in the high-risk group was significantly higher than that in the low-risk group (Supplementary Figures S3E, S3F). In the training cohort, the expression of FCGR2A in the high-risk group was higher than that in the low-risk group (Supplementary Figure S3D).

We then used two methods to identify the immune cells associated with FCGR2A and FCGR3B. First, HNSCC samples were divided into low- and high-expression groups based on the median expression of FCGR2A. The Wilcoxon test was used to compare the different infiltration levels of the 22 immune cells in the two groups. The results indicated that the infiltration levels of the resting memory CD4$^+$ T cells, activated resting memory CD4$^+$ T, resting NK cells, and M1 and M2 macrophages in the high-expression group were higher than those in the low-expression group, whereas plasma, activated dendritic, resting dendritic, follicular helper T cells, and M0 macrophages showed the opposite trend (Figure 6J). Subsequently, using Spearman's rank correlation analysis, we found that the infiltration levels of plasma, naive CD4$^+$ T, activated memory CD4$^+$ T, follicular helper T cells, resting dendritic, activated dendritic, activated mast cells, M0 and M1 macrophages, eosinophils, and neutrophils were closely correlated with the expression of FCGR2A. Considering the intersection of the immune cells from the two sets of results, memory B, plasma, resting memory and activated memory CD4$^+$ T, follicular helper T, resting dendritic, activated dendritic, activated mast, resting NK, activated NK cells, M0, M1, M2 macrophages, and neutrophils were correlated with the expression of FCGR2A (Supplementary Figure S3G). The immune cells that closely associated with FCGR3B were memory B, resting memory CD4$^+$ T, follicular helper T, resting NK, activated NK, resting dendritic, activated mast cells, and neutrophils (Figure 6K; Supplementary Figure S3H).

To analyze the effects of somatic cell copy number alternations (CNAs) of these two genes on infiltration of immune cells, such as B, CD4$^+$ T, CD8$^+$ T, dendritic cells, neutrophils, and macrophages, we applied the Tumor Immune Estimation Resource (TIMER, https://cistrome.shinyapps.io/timer/). The results showed that the six immune cells were significantly affected by the arm-level deletion and gain of the two genes in HNSCC (Figures 6L, M). It has been widely acknowledged that miRNAs are short noncoding RNAs that can induce mRNA silencing and instability by binding to specific target sites. We predicted that the upstream miRNAs might bind to FCGR2A. These upstream miRNAs, including hsa-miR-124-3p, hsa-miR-145-5p, hsa-

miR-299-3p, hsa-miR-513a-5p, hsa-miR-506-3p, and hsa-miR-671-5p, were found through the ENCORI (https://starbase.sysu.edu.cn/) database, which predicted target genes using PITA, RNA22, miRmap, DIANA-microT, miRanda, PicTar, and TargetScan programs. We performed the following analysis only for the predicted miRNAs that appeared in more than two programs. Based on the ceRNA hypothesis, hsa-miR-506-3p was finally chosen (Figures 7A–C). Next, we predicted the upstream lncRNAs. The results showed that there were 33 possible lncRNAs upstream of hsa-miR-506-3p. LncRNAs can competitively bind to shared miRNAs to increase mRNA expression. Therefore, there should be a negative correlation between lncRNAs and miRNAs or a positive correlation between lncRNAs and mRNAs. Based on expression, survival and correlation analysis, we found that AC110048.2 may potentially be the upstream lncRNA of the miR-506-3p/FCGR2A axis in HNSCC (Figures 7D–F).

Finally, we established a transcription factor regulatory network for the model genes, using the Cistrome website (http://cistrome.org/). From this database, 314 transcription factors were identified. There were 59 differentially expressed transcription factors between normal and tumor samples under the criteria FDR <0.05 and $|\log_2 FC| \geq 1$ (Supplementary Table S7). According to the criteria of $|$correlation coefficient$| > 0.3$ and p-value <0.001 using Pearson's correlation analyses, we identified 10 transcription factors associated with the model genes, and constructed the transcription factor regulatory network (Figure 7G; Supplementary Table S8).

4 Discussion

Cancer immunotherapy, which differently regulates the immune system, has been widely used in the field of oncology (Baxevanis et al., 2009; Yang, 2015). The TME is closely connected with immunotherapy and plays an important role in tumor genesis and development (Quail and Joyce, 2013). Therefore, it is necessary to explore potential therapeutic targets for early diagnosis and therapy. Thus, immune-based prognostic characteristics have become the focus of cancer risk prediction (Belli et al., 2018; Roma-Rodrigues et al., 2019; Shi et al., 2021).

In this study, based on the transcriptome data of HNSCC, we calculated the scores of immune and stromal cells in the TME, and found that they were significantly different in each phase of tumor development, suggesting that TME played a significant role in tumor growth. Based on the median scores, we obtained 395 DEGs related to the TME. The GO and KEGG enrichment analysis showed that these genes were significantly enriched in immune- and metabolism-related pathways, which preliminarily suggested that immune-related genes and pathways had significant association with the occurrence and development of HNSCC.

Based on these DEGs, a prognostic model consisting of FCGR3B, IGHV3-64, AC023449.2, IGKV1D-8, FCGR2A, WDFY4, and HBQ1 was constructed. Moreover, based on the model genes, a clinical prediction nomogram was constructed and verified to have good predictability. Based on the literature, FCGR3B is a gene that encodes FcγRIIIb and plays an important role in the immune system. Therefore, the biological function of FCGR3B in head and neck squamous cell carcinoma (HNSCC) may be related to the immune system. Another study found that copy number variations of FCGR3B were associated with susceptibility to autoimmune diseases, suggesting that FCGR3B may be involved in regulating immune responses
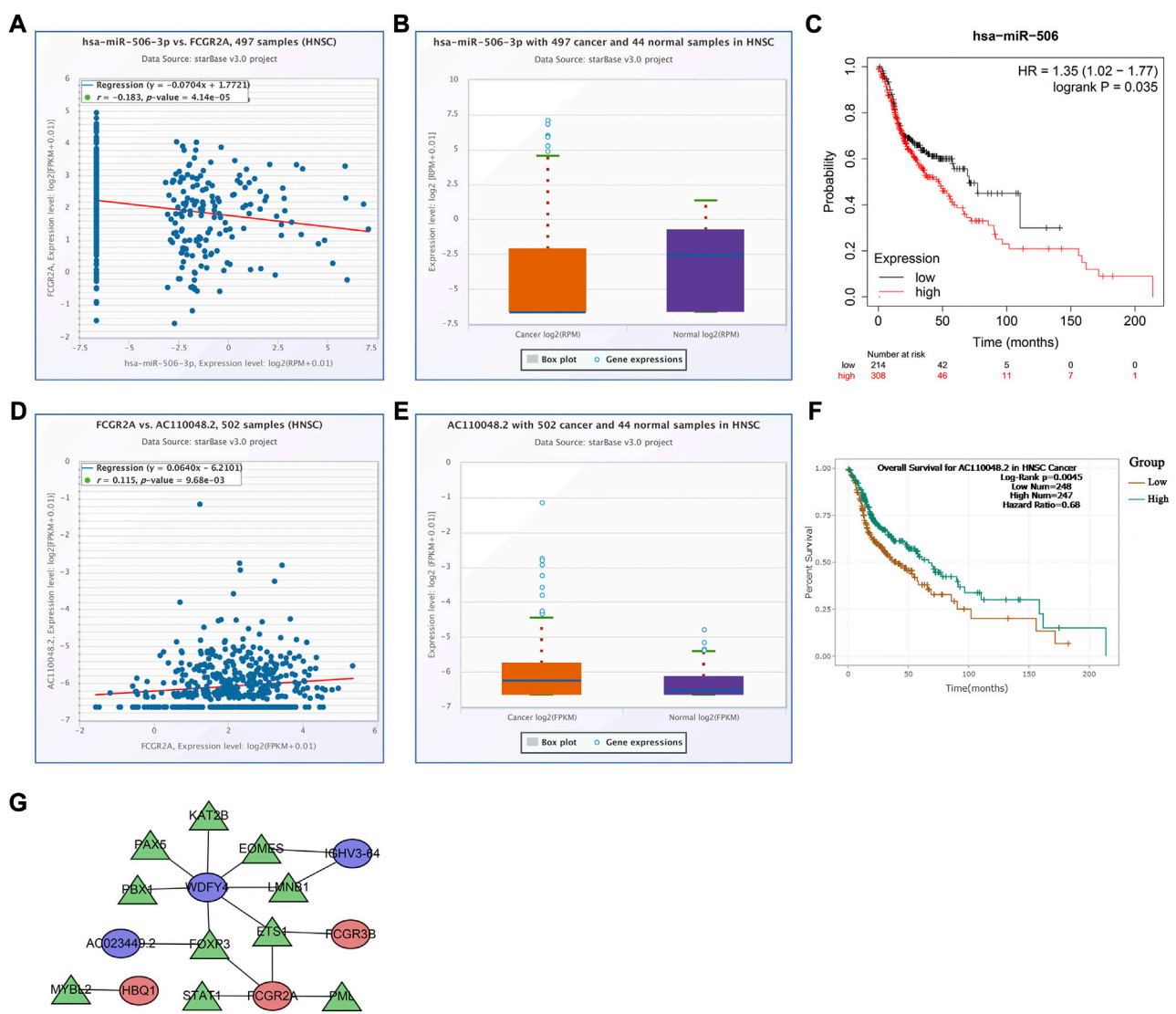
FIGURE 7
Establishment of the competitive endogenous RNA network constructed using the starBase database. Correlation of expression of **(A)** hsa-miR-506-3p with FCGR2A, and **(D)** AC110048.2 with FCGR2A. The expression levels of **(B)** hsa-miR-506-3p, and **(E)** AC110048.2 in HNSCC and normal samples. KM curves of the prognostic value of **(C)** hsa-miR-506 in the KM plot database, and **(F)** AC110048.2 in the StarBase database. **(G)** The transcription factor regulatory networks associated with the model genes. Red ellipse indicates upregulated model genes; blue ellipse indicates downregulated model genes; green triangle indicates transcription factors.

(Leemans et al., 2011; Molokhia et al., 2011; Alberici et al., 2020). The IGH family is involved in the development of B-cell malignancies. Somatic hypermutation of IGHV genes is characteristic of many B-cell lymphomas (Sahota et al., 1997; Ghia et al., 2007; Varettoni et al., 2013). One member of the IGH family, IGHV3-64, was found to be involved in the regulation of immune cells, particularly the positive regulation of B cell activation. Currently, some studies have explored the biological functions of IGHV3-64 in other cancers. For example, in chronic lymphocytic leukemia (CLL), the expression level of IGHV3-64 is closely related to clinical prognosis (Crombie and Davids, 2017). In addition, there are studies suggesting that IGHV3-64 may be associated with the development and prognosis of gastrointestinal (Guan et al., 2020). However, the results of these studies are inconsistent, and more research is needed to determine the biological functions of IGHV3-64 in different cancers. Previous studies have shown that WDFY4 is involved

in the function of various immune cells, and it can modulate B cells through noncanonical autophagy, and participates in the regulation of systemic lupus erythematosus (Zhao et al., 2012; Yuan et al., 2018). Furthermore, the deficiency of WDFY4 results in a decrease in $CD8^+$ T cells (Li et al., 2021). Hemoglobin subunit theta 1 (HBQ1) is often used as an indicator related to tumor metabolism. When patients with breast cancer were treated with the combination of bevacizumab and doxorubicin, HBQ1 was often differentially expressed (Borgan et al., 2013; Bae et al., 2022) and IGKV1D-8 was primarily involved in immune response (Gaudet et al., 2011). However, there have been a few reports on AC023449.2 and IGKV1D-8 (Alberici et al., 2018; Treffers et al., 2018; Dai et al., 2021). FCGR2A is closely associated with immunity and is considered a cell-surface receptor on phagocytic cells. Although FCGR2A has rarely been reported in HNSCC, our study showed that immune cells, such as memory B, plasma, resting memory

CD4[+] T, activated memory CD4[+] T, follicular helper T, resting and activated dendritic, activated mast, resting and activated NK cells, M0, M1, and M2 macrophages, and neutrophils, were closely associated with FCGR2A expression in HNSCC. As a product of immune cells, FCGR3B plays an important role in the connection and clearance of neutrophils and other immune complexes (Coxon et al., 2001; Fanciulli et al., 2007). We found that immune cells including memory B, resting memory CD4[+] T, follicular helper T, resting and activated NK, resting dendritic, activated mast cells, and neutrophils, were closely associated with FCGR3B.

Following GO and KEGG analyses, we found that the DEGs were strongly associated with immunity in the two risk groups. The GSVA results indicated that metabolism-related pathways, such as fatty acid, butanoate, glycine, serine, and threonine metabolism, were significantly different between the two risk groups. Changes in cell metabolism affected tumor progression. Fatty acid metabolism plays a crucial role in tumorigenesis and Epithelial–mesenchymal transition (EMT) regulation.

Furthermore, the infiltration of CD8[+] T cells was higher in the low-risk group than that in the high-risk group. We also found that in both cohorts, the low-risk group had a higher expression of PD1, PD-L1, and CTLA-4. We therefore speculated that the low-risk group may benefit the most from antibody therapies targeting the PD1, PD-L1, and CTLA-4 immune checkpoints. However, in both cohorts, the TMB was higher in the high-risk group than that in the low-risk group.

5 Conclusion

In conclusion, our study highlights the importance of the tumor microenvironment (TME) in the development and prognosis of head and neck squamous cell carcinoma (HNSCC). By analyzing gene expression data from the TCGA database, we identified seven new markers that were found to be associated with HNSCC prognosis. We also constructed a risk model based on the TME that accurately predicted patient outcomes. Our study further revealed significant differences in the infiltration levels of immune cells between low- and high-risk groups. These findings provide a better understanding of the mechanisms of tumor progression and immune infiltration in HNSCC and offer potential biomarkers for prognosis and treatment. Our study may also facilitate the development of new therapeutic strategies for HNSCC patients.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The TCGA data in the study have been downloaded from https://portal.gdc.cancer.gov/.

Author contributions

LW: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing–original draft. YL: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. WP: Conceptualization, Data curation, Methodology, Software, Validation. YY: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Validation. CY: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Validation. CL: Methodology. XZ: Methodology. RL: Methodology, Software. WY: Data curation, Funding acquisition, Supervision, Validation, Writing–original draft, Writing–review and editing. XY: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing–original draft, Writing–review and editing. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1232875/full#supplementary-material

References

Alberici, F., Bonatti, F., Adorni, A., Daminelli, G., Sinico, R. A., Gregorini, G., et al. (2020). FCGR3B polymorphism predicts relapse risk in eosinophilic granulomatosis with polyangiitis. *Rheumatol. Oxf.* 59, 3563–3566. doi:10.1093/rheumatology/keaa134

Alberici, F., Bonatti, F., Maritati, F., Urban, M. L., Moroni, G., Emmi, G., et al. (2018). Association of a polymorphism of the Fcγ-receptor 2A (FCGR2A) gene with chronic periaortitis. *Clin. Exp. Rheumatol.* 37, 222–226.

Bae, S. W., Berlth, F., Jeong, K. Y., Park, J. H., Choi, J. H., Park, S. H., et al. (2022). Glucose metabolic profiles evaluated by PET associated with molecular characteristic landscape of gastric cancer. *Gastric Cancer* 25, 149–160. doi:10.1007/s10120-021-01223-3

Baxevanis, C. N., Perez, S. A., and Papamichail, M. (2009). Cancer immunotherapy. *Crit. Rev. Clin. Lab. Sci.* 46, 167–189. doi:10.1080/10408360902937809

Belli, C., Trapani, D., Viale, G., D'Amico, P., Duso, B. A., Della Vigna, P., et al. (2018). Targeting the microenvironment in solid tumors. *Cancer Treat. Rev.* 65, 22–32. doi:10.1016/j.ctrv.2018.02.004

Bi, K.-W., Wei, X.-G., Qin, X.-X., and Li, B. (2020). BTK has potential to Be a prognostic factor for lung adenocarcinoma and an indicator for tumor microenvironment remodeling: a study based on TCGA data mining. *Front. Oncol.* 10, 424. doi:10.3389/fonc.2020.00424

Borgan, E., Lindholm, E. M., Moestue, S., Maelandsmo, G. M., Lingjaerde, O. C., Gribbestad, I. S., et al. (2013). Subtype-specific response to bevacizumab is reflected in the metabolome and transcriptome of breast cancer xenografts. *Mol. Oncol.* 7, 130–142. doi:10.1016/j.molonc.2012.10.005

Chen, L., Cai, Z., Lyu, K., Cai, Z., and Lei, W. (2021). A novel immune-related long non-coding RNA signature improves the prognosis prediction in the context of head and neck squamous cell carcinoma. *Bioengineered* 12, 2311–2325. doi:10.1080/21655979.2021.1943284

Coxon, A., Cullere, X., Knight, S., Sethi, S., Wakelin, M. W., Stavrakis, G., et al. (2001). Fc gamma RIII mediates neutrophil recruitment to immune complexes. a mechanism for neutrophil accumulation in immune-mediated inflammation. *Immunity* 14, 693–704. doi:10.1016/s1074-7613(01)00150-9

Crombie, J., and Davids, M. S. (2017). IGHV mutational status testing in chronic lymphocytic leukemia. *Am. J. Hematol.* 92, 1393–1397. doi:10.1002/ajh.24808

Dai, Y., Chen, W., Huang, J., and Cui, T. (2021). FCGR2A could function as a prognostic marker and correlate with immune infiltration in head and neck squamous cell carcinoma. *BioMed Res. Int.* 2021, 8874578–8874615. doi:10.1155/2021/8874578

Fanciulli, M., Norsworthy, P. J., Petretto, E., Dong, R., Harper, L., Kamesh, L., et al. (2007). FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. *Nat. Genet.* 39, 721–723. doi:10.1038/ng2046

Ferris, R. L. (2015). Immunology and immunotherapy of head and neck cancer. *J. Clin. Oncol.* 33, 3293–3304. doi:10.1200/JCO.2015.61.1509

Gaudet, P., Livstone, M. S., Lewis, S. E., and Thomas, P. D. (2011). Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Briefings Bioinforma.* 12, 449–462. doi:10.1093/bib/bbr042

Gavrielatou, N., Doumas, S., Economopoulou, P., Foukas, P. G., and Psyrri, A. (2020). Biomarkers for immunotherapy response in head and neck cancer. *Cancer Treat. Rev.* 84, 101977. doi:10.1016/j.ctrv.2020.101977

Ghia, P., Stamatopoulos, K., Belessi, C., Moreno, C., Stilgenbauer, S., Stevenson, F., et al. (2007). ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. *Leukemia* 21, 1–3. doi:10.1038/sj.leu.2404457

Guan, X., Xu, Z. Y., Chen, R., Qin, J. J., and Cheng, X. D. (2020). Identification of an immune gene-associated prognostic signature and its association with a poor prognosis in gastric cancer patients. *Front. Oncol.* 10, 629909. doi:10.3389/fonc.2020.629909

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

Jiang, A. M., Ren, M. D., Liu, N., Gao, H., Wang, J. J., Zheng, X. Q., et al. (2021). Tumor mutation burden, immune cell infiltration, and construction of immune-related genes prognostic model in head and neck cancer. *Int. J. Med. Sci.* 18, 226–238. doi:10.7150/ijms.51064

Leemans, C. R., Braakhuis, B. J., and Brakenhoff, R. H. (2011). The molecular biology of head and neck cancer. *Nat. Rev. Cancer* 11, 9–22. doi:10.1038/nrc2982

Li, Y., Li, J., Yuan, Q., Bian, X., Long, F., Duan, R., et al. (2021). Deficiency in WDFY4 reduces the number of CD8(+) T cells via reactive oxygen species-induced apoptosis. *Mol. Immunol.* 139, 131–138. doi:10.1016/j.molimm.2021.08.022

Liu, L., Bai, X., Wang, J., Tang, X. R., Wu, D. H., Du, S. S., et al. (2019). Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. *Clin. Cancer Res.* 25, 7413–7423. doi:10.1158/1078-0432.CCR-19-0558

Lv, J., Zhu, Y., Ji, A., Zhang, Q., and Liao, G. (2020). Mining TCGA database for tumor mutation burden and their clinical significance in bladder cancer. *Biosci. Rep.* 40. doi:10.1042/BSR20194337

Miyauchi, S., Kim, S. S., Pang, J., Gold, K. A., Gutkind, J. S., Califano, J. A., et al. (2019). Immune modulation of head and neck squamous cell carcinoma and the tumor microenvironment by conventional therapeutics. *Clin. Cancer Res.* 25, 4211–4223. doi:10.1158/1078-0432.CCR-18-0871

Molokhia, M., Fanciulli, M., Petretto, E., Patrick, A. L., McKeigue, P., Roberts, A. L., et al. (2011). FCGR3B copy number variation is associated with systemic lupus erythematosus risk in Afro-Caribbeans. *Rheumatol. Oxf.* 50, 1206–1210. doi:10.1093/rheumatology/keq456

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

Quail, D. F., and Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* 19, 1423–1437. doi:10.1038/nm.3394

Roma-Rodrigues, C., Mendes, R., Baptista, P. V., and Fernandes, A. R. (2019). Targeting tumor microenvironment for cancer therapy. *Int. J. Mol. Sci.* 20, 840. doi:10.3390/ijms20040840

Romano, E., and Romero, P. (2015). The therapeutic promise of disrupting the PD-1/PD-L1 immune checkpoint in cancer: unleashing the CD8 T cell mediated anti-tumor activity results in significant, unprecedented clinical efficacy in various solid tumors. *J. Immunother. Cancer* 3, 15. doi:10.1186/s40425-015-0059-z

Sahota, S. S., Leo, R., Hamblin, T. J., and Stevenson, F. K. (1997). Myeloma VL and VH gene sequences reveal a complementary imprint of antigen selection in tumor cells. *Blood* 89, 219–226. doi:10.1182/blood.v89.1.219

Shi, R., Bao, X., Unger, K., Sun, J., Lu, S., Manapov, F., et al. (2021). Identification and validation of hypoxia-derived gene signatures to predict clinical outcomes and therapeutic responses in stage I lung adenocarcinoma patients. *Theranostics* 11, 5061–5076. doi:10.7150/thno.56202

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

Solomon, B., Young, R. J., and Rischin, D. (2018). Head and neck squamous cell carcinoma: genomics and emerging biomarkers for immunomodulatory cancer treatments. *Semin. Cancer Biol.* 52, 228–240. doi:10.1016/j.semcancer.2018.01.008

Treffers, L. W., van Houdt, M., Bruggeman, C. W., Heineke, M. H., Zhao, X. W., van der Heijden, J., et al. (2018). FcγRIIIb restricts antibody-dependent destruction of cancer cells by human neutrophils. *Front. Immunol.* 9, 3124. doi:10.3389/fimmu.2018.03124

Varettoni, M., Zibellini, S., Capello, D., Arcaini, L., Rossi, D., Pascutto, C., et al. (2013). Clues to pathogenesis of Waldenstrom macroglobulinemia and immunoglobulin M monoclonal gammopathy of undetermined significance provided by analysis of immunoglobulin heavy chain gene rearrangement and clustering of B-cell receptors. *Leuk. Lymphoma* 54, 2485–2489. doi:10.3109/10428194.2013.779689

von Witzleben, A., Wang, C., Laban, S., Savelyeva, N., and Ottensmeier, C. H. (2020). HNSCC: tumour antigens and their targeting by immunotherapy. *Cells* 9, 2103. doi:10.3390/cells9092103

Yang, Y. (2015). Cancer immunotherapy: harnessing the immune system to battle cancer. *J. Clin. Invest.* 125, 3335–3337. doi:10.1172/JCI83871

Yi, L., Wu, G., Guo, L., Zou, X., and Huang, P. (2020). Comprehensive analysis of the PD-L1 and immune infiltrates of m(6)A RNA methylation regulators in head and neck squamous cell carcinoma. *Mol. Ther. Nucleic Acids* 21, 299–314. doi:10.1016/j.omtn.2020.06.001

Yoshihara, K., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Yuan, Q., Li, Y., Li, J., Bian, X., Long, F., Duan, R., et al. (2018). WDFY4 is involved in symptoms of systemic lupus erythematosus by modulating B cell fate via noncanonical autophagy. *J. Immunol.* 201, 2570–2578. doi:10.4049/jimmunol.1800399

Zhao, H., Yang, W., Qiu, R., Li, J., Xin, Q., Wang, X., et al. (2012). An intronic variant associated with systemic lupus erythematosus changes the binding affinity of Yinyang1 to downregulate WDFY4. *Genes Immun.* 13, 536–542. doi:10.1038/gene.2012.33

Zhao, Y., Xu, L., Wang, X., Niu, S., Chen, H., and Li, C. (2021). A novel prognostic mRNA/miRNA signature for esophageal cancer and its immune landscape in cancer progression. *Mol. Oncol.* 15, 1088–1109. doi:10.1002/1878-0261.12902

Integrative analysis of lysine acetylation-related genes and identification of a novel prognostic model for oral squamous cell carcinoma

Shi-Zhou Deng [ID] [1†], Xuechen Wu [ID] [2†], Jiezhang Tang [ID] [3†], Lin Dai[4]* and Bo Cheng[2]*

[1]Department of Hepatobiliary Surgery, Xi-Jing Hospital, The Fourth Military Medical University, Xi'an, China, [2]Department of Stomatology, Zhongnan Hospital of Wuhan University, Wuhan, China, [3]Department of Burn and Plastic Surgery, Tangdu Hospital, Fourth Military Medical University, Xi'an, China, [4]Department of Stomatology, The First Hospital of Wuhan, Wuhan, China

Introduction: Oral squamous cell carcinoma (OSCC), which accounts for a high proportion of oral cancers, is characterized by high aggressiveness and rising incidence. Lysine acetylation is associated with cancer pathogenesis. Lysine acetylation-related genes (LARGs) are therapeutic targets and potential prognostic indicators in various tumors, including oral squamous cell carcinoma. However, systematic bioinformatics analysis of the Lysine acetylation-related genes in Oral squamous cell carcinoma is still unexplored.

Methods: We analyzed the expression of 33 Lysine acetylation-related genes in oral squamous cell carcinoma and the effects of their somatic mutations on oral squamous cell carcinoma prognosis. Consistent clustering analysis identified two lysine acetylation patterns and the differences between the two patterns were further evaluated. Least absolute shrinkage and selection operator (LASSO) regression analysis was used to develop a lysine acetylation-related prognostic model using TCGA oral squamous cell carcinoma datasets, which was then validated using gene expression omnibus (GEO) dataset GSE41613.

Results: Patients with lower risk scores had better prognoses, in both the overall cohort and within the subgroups These patients also had "hot" immune microenvironments and were more sensitive to immunotherapy.

Disscussion: Our findings offer a new model for classifying oral squamous cell carcinoma and determining its prognosis and offer novel insights into oral squamous cell carcinoma diagnosis and treatment.

KEYWORDS

OSCC, oral squamous cell carcinoma, lysine acetylation, prognostic model, TCGA, GEO

1 Introduction

Oral squamous cell carcinoma (OSCC) is one of the most common cancers worldwide and accounted for around 369,000 new in 2012. OSCC incidence has continued to grow, with two-thirds of the cases occurring in developing countries. OSCC has a wide range of clinical patterns (Ghantous and Abu Elnaaj, 2017), and the majority of the cases are associated with

lifestyle habits like smoking, excessive alcohol consumption, and betel nut chewing. According to the National Comprehensive Cancer Network (NCCN) clinical practice guidelines in oncology, all OSCC is primarily treated through surgery in combination with radiotherapy and chemotherapy, and the use of targeted treatments is recommended for advanced cases (stages III–IV) (Warnakulasuriya, 2009). Following initial surgery and proper adjuvant treatment, the pathologic nodal stage is the main predictor of the malignant degree of OSCC patients (Zanoni et al., 2019). However, OSCC recurrence is common after the first R0 resection, resulting in a low survival rate (Warnakulasuriya, 2009), with an inadequate quality of life (Lin et al., 2022).Moreover, survival rates decline with increasing time before treatment initiation (Jensen et al., 2021). These factors emphasize the need for early OSCC diagnosis as well as novel molecular targets for treatment. For instance, although cetuximab, which targets the epidermal growth factor receptor, was approved for OSCC treatment in 2006 and anti-PD1 therapy has recently been used to treat patients with metastatic disease following relapse or progression during or after chemotherapy (Ferris et al., 2016), their efficacies have not been significant. Thus, understanding the molecular changes that underlie OSCC pathogenesis and the factors that contribute to OSCC patient prognosis is an unmet medical need.

Cell transporter functional expression has been demonstrated to be modulated by post-translational modification (PTM) via a variety of molecular pathways. These changes are made by adding These changes are made by adding specific chemical groups to certain amino acid residues (Czuba et al., 2018). Acetylation is a common PTM initiated by specific enzymes that transfer acetyl groups to the amino side chain of lysine. Recent studies show that acetylation can also occur non-enzymatically and is influenced by the availability of acetyl-CoA (Narita et al., 2019). Although acetylation was previously thought to be specific to histones, thousands of non-histone proteins have been shown to contain lysine acetylation, including nuclear, mitochondrial, and cytoplasmic proteins. Non-histone acetylation regulates several cellular processes, including transcription, DNA damage repair, and cell signaling. Lysine acetylation drives tumorigenesis by actively modifying the expression and function of oncogenic or tumor-suppressive factors (O'Garro et al., 2021; Hu et al., 2022). The acetylation process can influence tumor formation and progression by modulating immune activity and response in a variety of ways. Several immune-related acetylation/deacetylation modification targets are mentioned below (Ding et al., 2022). For example, p300 can acetylates PD-L1 and inhibits its translocation into the nucleus (Gao et al., 2020). And in non-small cell lung cancer HDAC3 can be suppressed by the decreased COP1, which increases PD-L1 expression (Wang H. et al., 2020).

Histone and non-histone acetylation, have double-edged roles in tumor metastasis and metabolism (Hu et al., 2022). Four human histone deacetylase inhibitors (HDACi) with the potential to trigger tumor suppressor genes, have emerged as epigenome-targeting drugs that can improve the chemotherapeutic and radiosensitivity of cancer cells, and have received FDA approval for use in clinical settings (Ding et al., 2022). DLUE1 is reported to be overexpressed in early OSCC tumors, and its knockdown suppresses OSCC cell proliferation, migration, and invasion, implying that DLEU1 drives the expression of several genes during OSCC carcinogenesis (Hatanaka et al., 2021). The expression of the deacetylase genes, HDAC6 and

HDAC9, is markedly elevated in OSCC (Sakuma et al., 2006; Rastogi et al., 2016). Antitumor effects of novel HDACi in OSCC have also been reported (Bai et al., 2011). For instance, HDACi target cancer stem cells by inhibiting tumor growth and inducing cytotoxicity and intracellular reactive oxygen species and are potential OSCC treatments (Marques et al., 2020). Impairment of lysine acetylation is thought to impair ribosome biogenesis and might contribute to OSCC pathogenesis (Dong et al., 2022).

In this study, we used bioinformatics to analyze the expression of 33 lysine acetylation-related genes (LARGs) as well as their mutations in OSCC tissues vs. normal tissues and then validated their expression using RT-qPCR. Based on the expression of "HDAC3" and "SIRT5", OSCC patients were divided into two groups, and their correlation with clinical characteristics examined. Univariate and LASSO regression analyses were used to develop an OSCC prognostic model. The efficacies of immunotherapy and chemotherapy, as well as the OSCC immune landscape, were analyzed in various risk groups.

2 Materials and methods

2.1 Oral squamous cell carcinoma patient datasets

RNA sequencing (RNA-seq) data on tissues from 323 OSCC patients and 32 normal tissues, as well as associated clinical data, were downloaded from TCGA. Gene microarray data and associated clinical data for 97 tumor samples were obtained from dataset GSE41613 from gene expression omnibus (GEO) (Supplementary Table S1). The "limma" package was used for internal standards and then applied to perform difference analysis.

2.2 Identification of differentially expressed lysine acetylation-related genes (LARGs)

Thirty-three LARGs were retrieved from a previous review (Narita et al., 2019) (Supplementary Table S2). The "limma" package was used to identify differentially expressed LARGs with $p < 0.05$. Next, we evaluated gene express variations in the 33 LARGs in each TCGA OSCC sample to identify the LARGs associated with mutagenesis. Data on gene mutations was also gathered from TCGA. The frequency of different mutations was computed. Finally, the R package "maftools" was used for visualization. Waterfall diagrams were used to visualize the status of somatic mutation integration in OSCCs. Univariate analysis was used to identify prognostic LARGs. Protein–protein interaction (PPI) networks for the 30 connected LARGs were constructed on STRING (https://cn.string-db.org/) (von Mering et al., 2005).

2.3 mRNA and protein level analyses of OSCC samples

This study involved patients who underwent routine intraoral examination, followed by oral mucosal biopsy and diagnosis of squamous cell carcinoma of the oral cavity. Ten pairs of OSCC and adjacent normal tissues were collected at Zhongnan Hospital. Patients with a history of systemic illness or with other primary

tumors were excluded from the analysis. OSCC samples and matched adjacent noncancerous tissues were obtained before preoperative radiotherapy or chemotherapy and immediately frozen in liquid nitrogen, followed by storage at −80°C until RNA extraction. Total RNA was extracted using Trizol reagent (Servicebio, China). Ethical approval for the study (No. 2022095K) was granted by Zhongnan Hospital of Wuhan University Medical Ethics Committee. RT-qPCR was done on a BIO-RAD system using a SYBR green dye qPCR mix (Servicebio, China). Primer information is provided in Supplementary Table S3. The paired-T test was used to determine the expression levels of the LARGs and GAPDH. Human Protein Atlas (HPA) immunohistochemistry data were used to identify the protein levels of two patterns, SIRT5 and HDAC3, in paracancerous tissue and malignant tissues.

2.4 Consensus clustering analysis of the LARGs

The "ConsensuClusterPlus" package was used to delimit distinct lysine acetylation-related OSCC patterns (Seiler et al., 2010). Based on different lysine acetylation-associated OSCC patterns, we examined the clinicopathological features and prognosis of the patients. The Kaplan–Meier (KM) analysis of the correlation between the lysine acetylation-associated OSCC patterns was carried out by R packages "survival" and "survminer" (Rich et al., 2010).

2.5 Identification of a LARGs prognostic signature for OSCC

GSE41613 was used as the test cohort, whereas the TCGA dataset was used as the training cohort. The LARGs-associated signature was used to set up the prognostic model in the training cohort. Next, univariate Cox regression analysis was used to identify the prognostic differentially expressed genes (DEGs) between the lysine acetylation-related patterns. LASSO regression analysis was then used to identify prognostic DEGs ($p < 0.05$) using the "glmnet" package (Simon et al., 2011). The risk score of the patients was calculated by the formula as follows: $Risk\ score = \sum_{i=1}^{n} coefi^{*}expi$. The median risk score was used to group the patients. Survival differences between the two groups were comparatively analyzed through KM survival analysis. Based on gene expression, principal component analysis (PCA) was done with the "stats" package. Moreover, t-distributed stochastic neighbor embedding (t-SNE) was conducted to discuss the distribution of different groups via the "Rtsne" package. The receiver operating characteristic (ROC) curve analyses were carried out to estimate the prognostic power of the gene signature by using the "survivalROC" package. The prognostic relationship between risk score and age, gender, grade, clinical stage, and immune score was analyzed. Additionally, we explored the correlation between risk scores and cluster patterns.

2.6 Construction of the OSCC nomogram

We created a nomogram based on the risk scores and the clinical data of the OSCC patients, including age, stage, grade, and genderto

expoit the predictive value of the eight-gene-based signature for clinical application. To this end, the 'rms', 'nomogramEx', and 'regplot' R packages were used to construct the nomogram. Next, ROC curve analysis was used to assess how well the nomogram could predict OSCC prognosis (Pencina and D'Agostino, 2004). Additionally, we used calibration curves to determine if the projected survival outcome (one-, three-, and five-year survival) was close to the actual outcome (Alba et al., 2017). The 45° line shows the best nomogram-predicted survival.

2.7 Validation of grouping efficacy and association analysis of immune cell infiltration

The relationship between risk scores and immune cells infiltration in OSCC samples was analyzed by the Pearson correlation analysis using the GSVA package. Statistical analysis was done using the ssGSEA algorithm (Hänzelmann et al., 2013). Various immune indicators to study the relationship between factors and immune phenotypes. We analyzed the association between risk scores and immune cell infiltration, as well as the expression of immune biomarkers, HLA family, chemokines, and chemokine receptors. Immune checkpoint was examined via Pearson correlation analysis using $p = 0.05$ as the cutoff threshold. The immunophenotype scores (IPS) of the patients were used to predict OSCC response to checkpoint blockade immunotherapy (Charoentong et al., 2017).

2.8 Drug sensitivity analysis

To assess the therapeutic potential of chemotherapy drugs on OSCC, the semi-inhibitory concentration (IC50) of common drugs was determined using the "pRRophetic" package (Geeleher et al., 2014). The sensitivity of the chemotherapeutic agents in different patient groups was also predicted.

2.9 Statistical analysis

Statistics acquired from TCGA were merged and conducted on R then processed and analyzed on R using the indicated packages. Normally distributed continuous variables were expressed as Mean ± standard deviation. Non-normally distributed continuous variables were presented as medians (range). Categorical variables were described as counts and percentages. Two-sided $p < 0.05$ indicated statistically significant differences.

3 Results

3.1 The landscape of lysine acetylation-related genes in OSCC patients

The detailed flowchart of the study is shown in Supplementary Figure S1. Using the TCGA dataset, we identified the expression levels of 33 LARGs in OSCC samples, and normal paracancerous

FIGURE 1
Landscape of (lysine acetylation-related genes) LARGs in OSCC **(A)** Molecular expression of LARGs in normal tissues compared with oral squamous cell carcinomas. **(B)** The genetic alterations of LARGs in OSCC. **(C)** The interactions between the candidate genes were shown by the PPI network. **(D)** The findings of the univariate Cox regression demonstrate the relationship between OS and gene expression.

specimens and found that the 24 of the 33 LARGs (73%) were expressed significantly different in OSCCs (Figure 1A).

Given the importance of gene mutations in carcinogenesis, we investigated the somatic mutations of 33 LARGs in OSCC samples and found that 122 of the 506 (24.11%) OSCC samples we analyzed had genetic changes. Among the LARGs we studied, EP300 was shown to have the highest mutation rate, followed by CREBBP and HDAC9. (Figure 1B). EP300 and CREBBP are both often mutated in squamous cell carcinoma and lymphomas (Attar and Kurdistani, 2017). Most of the changes are missense point mutations. HDAC9 interacts with a variety of transcriptional repressors and oncogenes (Ning et al., 2020) and may influence anticancer immune responses by limiting T-cell infiltration into the tumor microenvironment (TME) (Yang et al., 2021).

A PPI network revealed that 30 LARGs were closely interconnected (Figure 1C), the other 3 genes were eliminated because they do not interact with other lysine acetylation-related genes. Univariate Cox regression analysis

revealed that high HDAC3 and SIRT5 expression was associated with poor OSCC survival (Hazard ratio, HR: >1; Figure 1D).

3.2 HDAC3 and SIRT5 are upregulated in OSCC tissues when compared with normal tissue

We next conducted studies based on the expression of HDAC3 and SIRT5, and it appeared that there were substantial disparities in their overall survival (Figures 2A, D). Analysis of immunohistochemical data on HPA revealed that OSCC tissues exhibited significantly higher SIRT5 and HDAC3 staining when compared with normal tissues (Figures 2B, E). Moreover, RT-qPCR analysis revealed that SIRT5 and HDAC3 expression levels in cancer tissues were significantly higher than in normal tissues ($p < 0.05$; Figures 2C, F).
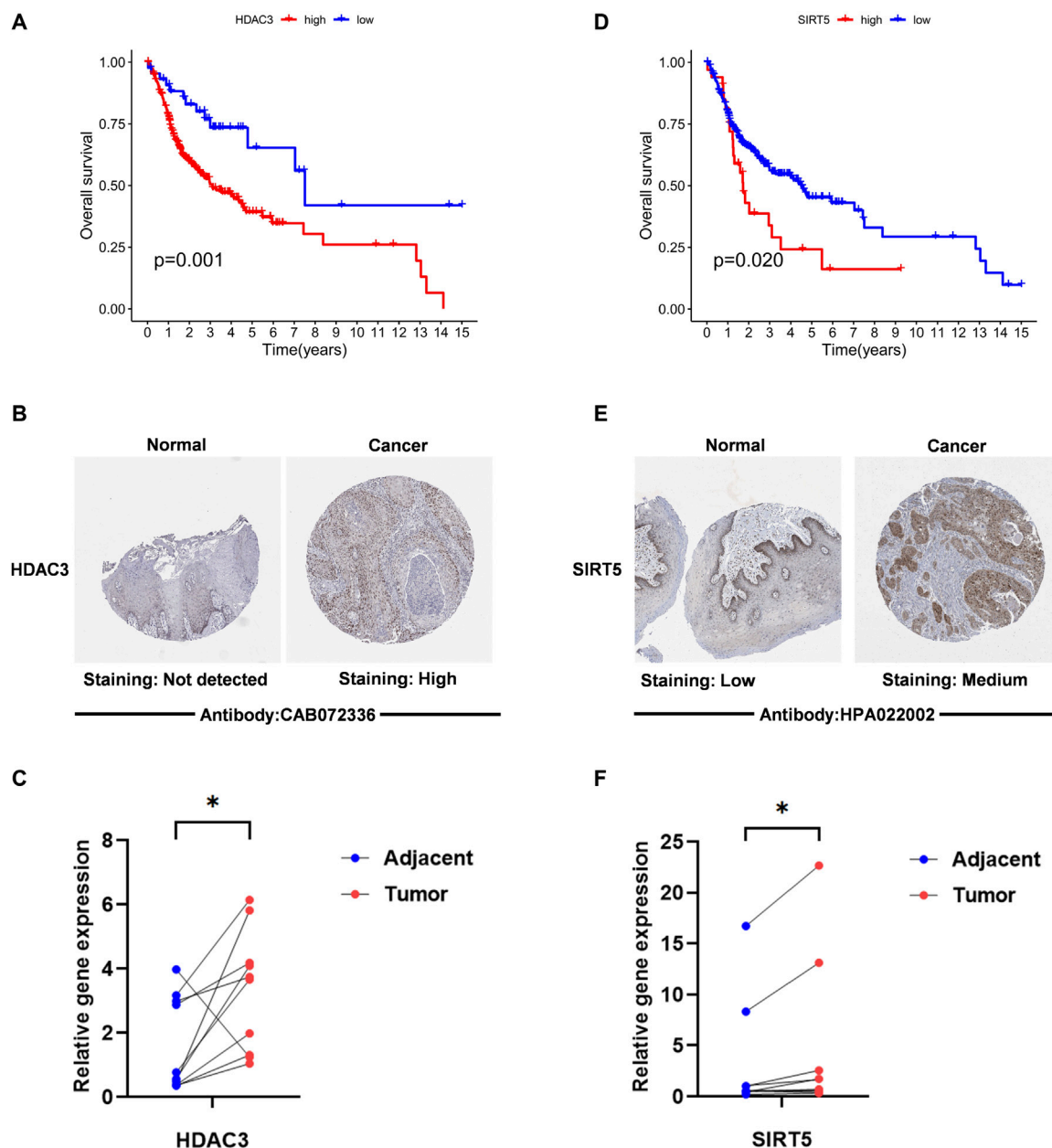
FIGURE 2
The expression level of "HDAC3" "SIRT5" in OSCC. **(A,D)** Kaplan−Meier survival analysis based on the expression of acetylation-related genes The OSCC patient survival curve for those with high and low gene expression was depicted by the red curve and the blue curve. **(B,E)** The HPA immunohistochemistry data were utilized to identify the protein levels of two genes in normal and malignant tissues. **(C,F)** HDAC3 and SIRT5 expression levels in OSCC tissues and surrounding normal tissues are compared. RT-PCR was used to identify the alterations in the expression of 2 LARGs in OSCC and its normal tissue. *if $p < 0.05$, ** if $p < 0.01$, and *** if $p < 0.001$.

3.3 Tumor classification based on the prognostic value of lysine acetylation regulators

Consistent clustering was used to examine SIRT5 and HDAC3 expression in a TCGA dataset of 323 OSCC cases. To this end we grouped the OSCC patients into two clusters based on cumulative distribution function (CDF) values (k = 2; Figure 3A, and k = 3–9; Supplementary Figure S2). PCA analysis found that the two clusters are clearly identifiable (Figure 3B).

This analysis also revealed that overall survival of cluster 1 was worse than that of cluster 2 ($p = 0.014$; Figure 3C). Analysis of whether the variability in survival was caused by differences in infiltration by the 23 immune cells in the 2 clusters revealed that immune cell infiltration differed significantly in 16 of 23 OSCCs (Figure 3D). These findings suggest that in the context of reduced expression of lysine acetylation-associated genes, OSCC patients with immune cell infiltration had better prognosis.

Furthermore, except for stage, other clinical parameters, including grade, gender, age, and TNM did not differ across

FIGURE 3
Consensus clustering is carried out based on the LARGs. **(A)** The consensus clustering was used to divide 323 OSCC patients into two groups (k = 2).
(B) PCA analysis showed a clear distinction between the cluster 1 and 2. **(C)** Survival curves for genes involved in RNA methylation that are linked to overall
survival. Clusters 1 and 2 were shown to be substantially linked to survival (*p* = 0.014). **(D)** 23 immune cell types infiltration is significantly different in two
clusters. **(E)** The clinicopathologic characteristics between the two clusters are shown on a heatmap.

these two clusters In cluster 1, most genes are upregulated, while in
cluster 2, the genes are downregulated, as shown in the heat map
(Figure 3E).

3.4 Developing an independent prognostic risk model based on LARGs clustering

We used "limma" package of R (4.1.1) to conduct, we discovered
323 DEGs between the two clusters (Supplementary Table S4), these
DEGs were then examined via univariate Cox regression analysis.
Twenty-six genes were finally proved that can be employed as
distinct prognostic indicators (Figure 4A). After filtration, LASSO
Cox regression analysis found NKX2-3, SAPCD2, SPINK7, LYNX1,
AKR1C3, SYT17, MASP1, and CTSG to be significantly associated
with overall survival (OS) (Figures 4B, C; adjusted *p* < 0.05).

The genes were used to calculate risk score based on the formula:

$$risk\ score = (-0.207 * NKX2-3) + (0.045 * SAPCD2)$$
$$+ (-0.011 * SPINK7) + (-0.054 * LYNX1)$$
$$+ (0.014 * AKR1C3) + (0.185 * SYT17)$$
$$+ (-0.399 * MASP1) + (-0.185 * CTSG).$$

Next, samples were divided into the high and low survival risk
groups based on the median risk score, as shown using KM survival
curves (*p* < 0.001). These analyses indicate that the multigene
signature had a significant prognostic value (Figure 4D) and that
the risk scores distinguished patients with high and low survival
rates (Figure 4F). The area under the curve (AUC) analysis at one,
three, and 5 years (AUC: 0.655, 0.707, and 0.707, respectively)
showed that the prognostic signature was highly accurate at
predicting OS in OSCC patients (Figure 4E). PCA analysis and
t-SNE analysis suggested that the OSCCs in distinct risk categories
were distributed in two directions (Figures 4G, H).

3.5 Validation of the prognostic value in the subgroups

Next, we split the GEO dataset into two categories based on risk
score (Supplementary Figure S3). KM (*p* = 0.02; Supplementary
Figure S3A) and ROC curve analyses revealed that the low-risk
group had a higher overall survival rate, indicating that the model
was accurate (one-three-, and five-year AUC: 0.736, 0.645, and
0.661, respectively; Supplementary Figure S3B). There were fewer
deaths in the low-risk group, which exhibited lower expression levels
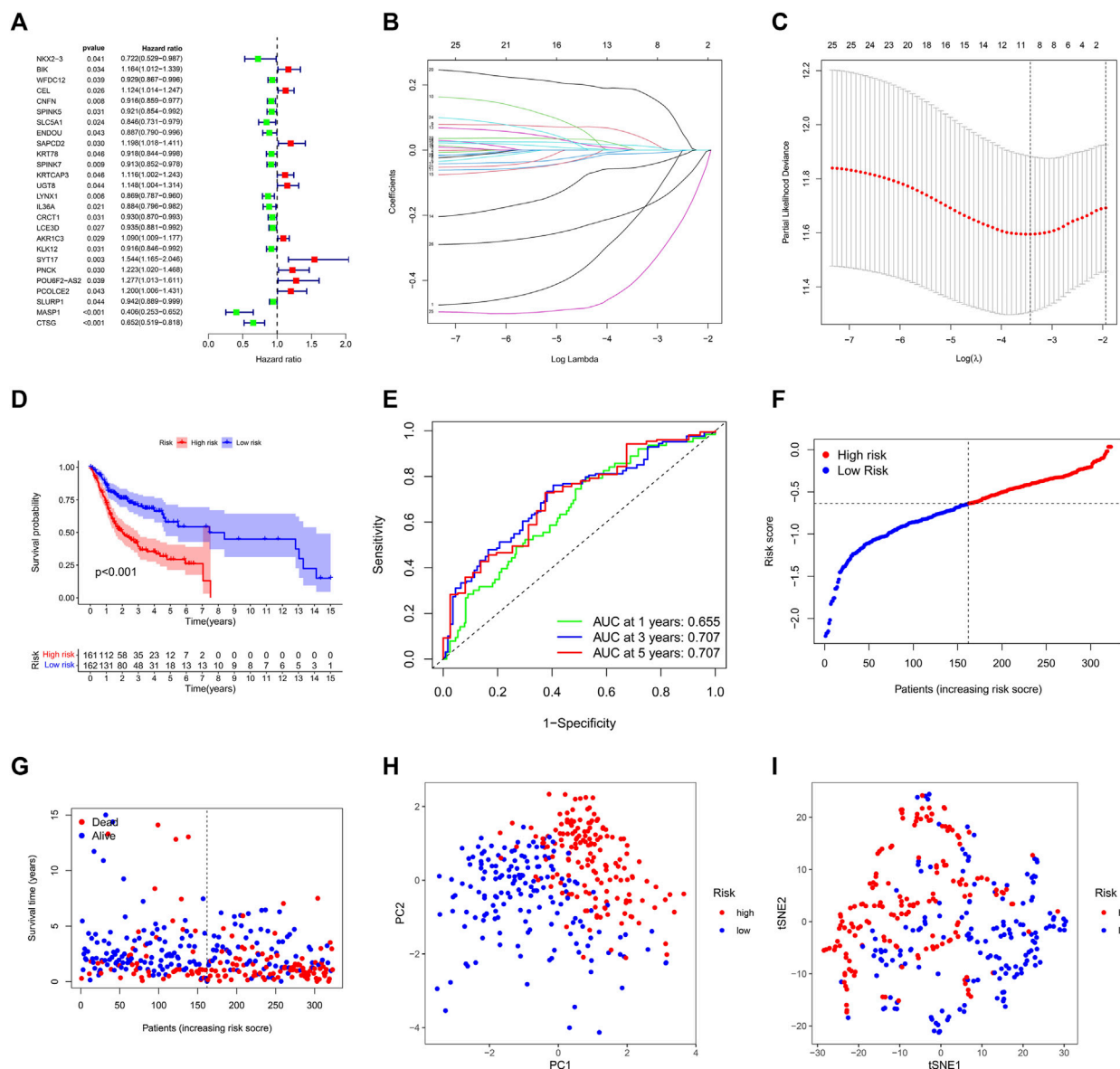
FIGURE 4
Construction of the lysine acetylation-related prognostic signature in the training cohort. **(A)** An investigation of OS for each DEG on clusters 1 and 2 using univariate cox regression. **(B)** LASSO regression of DEGs in OSCC. **(C)** Cross validation in the LASSO regression. **(D)** The OS of OSCC patients in the high-risk group was considerably poorer than that of the low-risk group, according to K-M curves. **(E)** An evaluation of the prognostic signature for OS in OSCC patients using ROC curves. **(F,G)** OSCC patients characteristics by high- and low-risk categories **(H,I)** To demonstrate how the samples of the various risk groups associated with lysine acetylation were dispersed independently, PCA **(H)** and t-SNE **(I)** were used.

of the risk genes (Supplementary Figures S3C, F). Finally, t-SNE analysis and PCA revealed that the risk genes were very effective in differentiating the two risk groups (Supplementary Figures S3D–E).

3.6 Subgroup survival analysis based on clinical parameters

To determine the ability of various clinical parameters to predict OSCC prognosis, we carried out a stratified analysis of clinical parameters in the test cohort by creating multiple

subgroups for the patients in the TCGA dataset using various clinical parameters. KM analysis of the correlation between age (≤65 and >65 years), sex, grade (G1–G2 or G3–G4), stage (I–II or III–IV), and survival indicated that except for G3–G4, high-risk patients had a lower likelihood of survival than low-risk patients (Supplementary Figures S4A–H). We also studied how the clinical parameters and the risk scores correlated with one another. This analysis revealed that high-risk scores and the AJCC stage, clusters, and immune scores differ significantly from each other (Supplementary Figures S5A–F). High-risk scores were mainly observed in patients with lower immune scores when compared
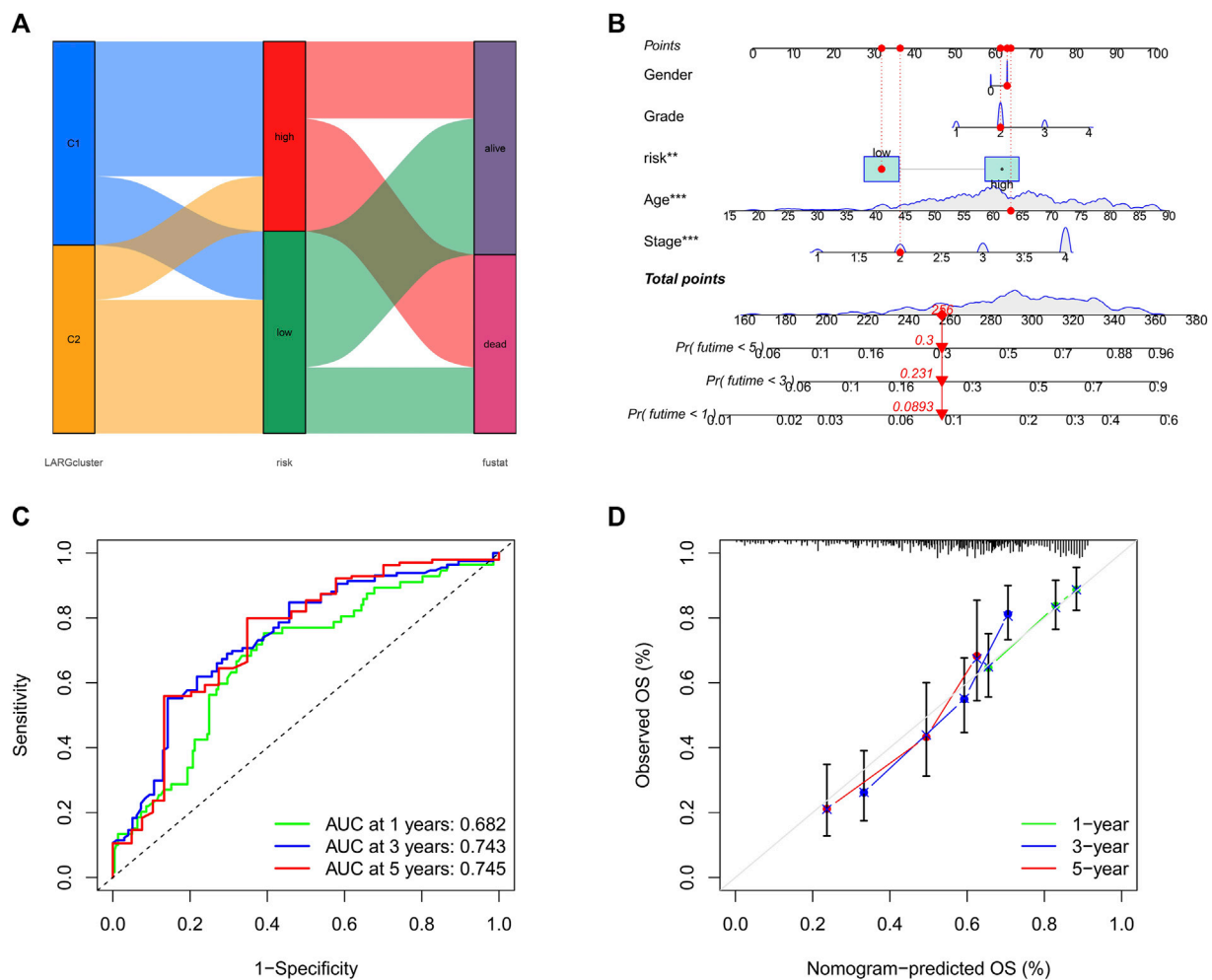
FIGURE 5
(A) Sankey plot shows quantities of patients flow from 2 clusters to the risk score distributions group than to final status. **(B)** Nomogram for OSCC patient survival predictions at 1, 3, and 5 years **(C)** The ROC curves for 1–, 3–, and 5–year OS in OSCC patients. **(D)** Calibration curves of the nomogram measured by Hosmer-Lemeshow test.

with those with high immune scores. Advanced disease stage was also associated with higher risk scores.

3.7 Development of a nomogram and model efficiency prediction

Sankey plot analysis revealed that the patients were distributed into two LARG clusters, two risk score clusters, and two future status clusters (Figure 5A).

Next, we developed a nomogram to illustrate the connection between these independent prognostic markers and survival probabilities (Figure 5B). Clinicians might forecast a patient's prognosis based on their total points. Patients with higher total points had lower survival. Additionally, calibration curves indicated that the nomogram could accurately predict one-, three-, and five-year OS (Figure 5C). A nomogram calibration curve was used to assess consistency between predicted and observed OS outcomes, with red, blue, and green lines indicating how the nomogram performed, whereas the gray line at 45° indicates flawless prediction (Figure 5D).

3.8 Gene set enrichment analysis and immune activity

The ESTIMATE algorithm was applied to generate TME scores. This analysis showed that patients with high-risk scores had significantly lower estimate score, immune score, and stromal score ($p < 0.001$) than those patients with high-risk score ($p < 0.001$) (Figure 6A). Moreover, ssGSEA analysis of the differences in multiple immune cells and signal pathways revealed that the high-risk group had lower immune cell infiltration ($p < 0.05$; Figure 6B). Moreover, these pathways were suppressed in patients with high-risk scores, including APC co-inhabitation, CCR, immune checkpoint, and cytolytic activity ($p < 0.05$; Figure 6C). Chemokines mediate the leukocyte migration to various sites during normal homeostasis and inflammation. Therefore, we investigated the correlation between 19 chemokine receptors and 43 chemokines and risk categories (Figures 6D, F). This analysis revealed that most chemokines, including inflammatory chemokines like CCL2 and CXCL12, which promote the proliferation of B progenitor cells in the bone marrow milieu

FIGURE 6
Two-group TME evaluation and checkpoints analysis. **(A)** Relationship between ESTIMATE score and high and low risk groups **(B)** The aggregation and expression of 16 immune cells were different in OSCC patients. **(C)** In high and low risk groups, 13 immunological functions expressed themselves differently. **(D)** Comparisons between the risk scores and the quantity of chemokines expression. **(E)** Human Leukocyte Antigen (HLA) expression in the high and low risk categories. **(F)** The abundance of chemokine receptors in different risk score groups. **(G)** Expression levels of clinically targetable ICP were measured in two risk groups. ($p < 0.05$ *; $p < 0.01$ **; $p < 0.001$ ***, not significant ns).

where they are produced, were markedly lower in patients with high-risk scores. Indicating that the differences between the immunological microenvironment of the high and low-risk groups were caused by the equivalent reduction in chemokine levels.

We also investigated the correlation between risk scores and HLA complex genes (Figure 6E), including HLA-DRB5 and HLA-DRB1, which are crucial for immune activity because of their antigen-presenting function. The potential of checkpoint
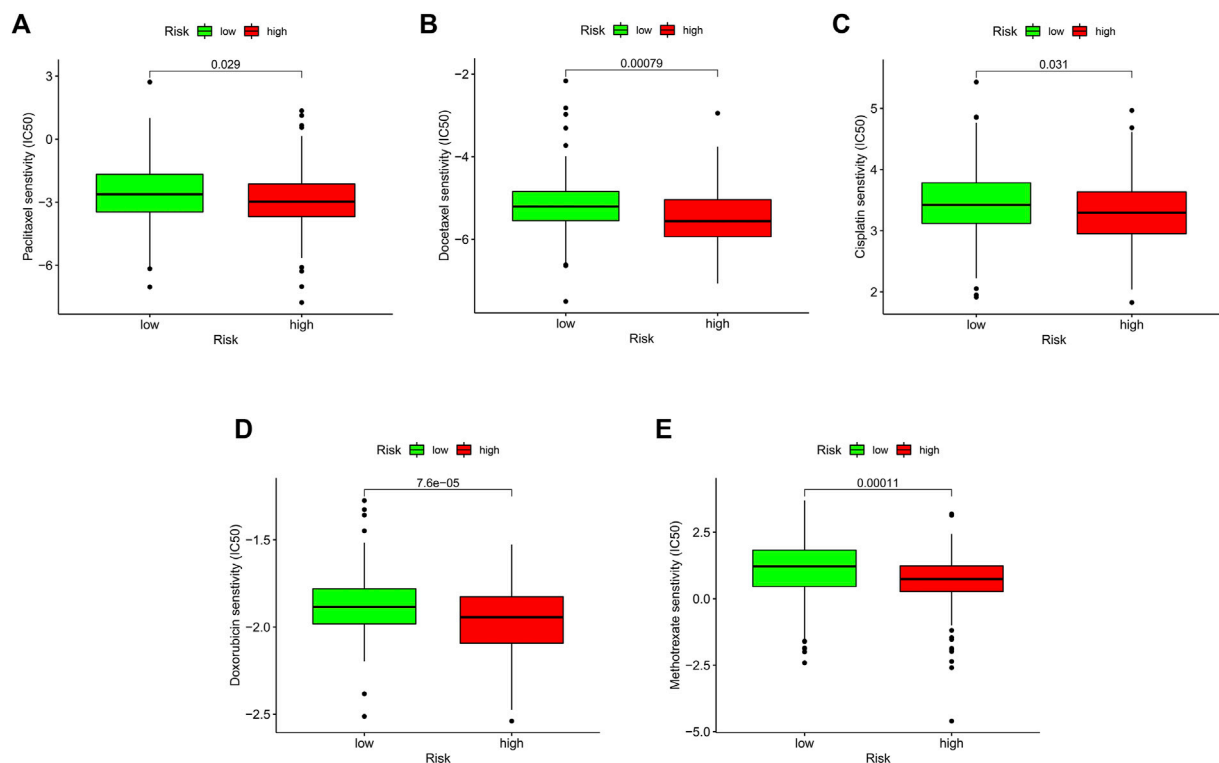
FIGURE 7
Relationship between risk score and therapeutic sensitivity. **(A–E)** Association between risk score and chemotherapeutic sensitivity.

inhibitors to treat cancer has attracted significant interest. Therefore, we deduced that the OSCC inflammatory condition may be associated with a unique expression of immune checkpoint (ICP). The expression levels of eight ICP genes,LAG3, TIGIT, PDCD1, HAVCR2, CTLA4, SIGLEC15, PDCD1LG2, and CD274 were calculated to determine the correlations between immune checkpoints and risk score. This analysis found LAG3, CTLA4, PDCD1, TIGIT, PDCD1LG2, and CD274 to be downregulated in the high-risk group (Figure 6G), suggesting that the low-risk group is sensitive to immunotherapy. To assess the response of patients to immune checkpoint inhibitors, we calculated the IPS scores of each sample and found that the IPS scores(ips_ctla4_pos_pd1_pos) of low-risk groups were higher, indicating that the patients in this group may be more sensitive to the combined PD-1/CTLA4 blockade (Supplementary Figures S6A–D).

3.9 Drug sensitivity analysis

Chemotherapy, targeted therapy, and immunotherapy may slow tumor growth in OSCC patients and enhance patient prognosis. We calculated the IC50 values of various chemotherapies in the test cohort using the "pRRophetic" package on R. This analysis found that paclitaxel, docetaxel, cisplatin, doxorubicin, methotrexate, and several targeted treatment drugs are more effective in patients with high-risk scores (Figures 7A–E). Paclitaxel primarily affects the M phase of mitosis, and disrupts tubulin synthesis, thereby inhibiting

the replication of tumor cells. Docetaxel belongs to the same family as paclitaxel but has a higher affinity for microtubule sites and exhibits higher anticancer activity. Cisplatin is a platinum compound and acts on the chemical structure of DNA. Doxorubicin and methotrexate enter the nucleus, bind to DNA, and inhibit nucleic acid synthesis and mitosis. In summary, these findings suggest that risk scores can predict drug sensitivity.

4 Discussion

Because of its molecular heterogeneity, few treatments are effective against terminal oral cancer. To improve OSCC prognosis, novel biomarkers, and treatment targets are needed. The emergence of high-throughput array technologies presents a chance to investigate the mechanisms underlying OSCC occurrence and progression. Lysine acetylation, a key regulatory mechanism of gene expression, might be associated with OSCC pathophysiology but it is unclear if acetylation-related genes influence OSCC or whether they are associated with OSCC survival.

Here, we first assessed the expression levels of 33 LARGs in OSCC vs. normal tissues and found that most of were differentially expressed, with Sirtuin 5 (SIRT5) and Histone Deacetylase 3 (HDAC3) exhibiting the highest differential expression. Analysis of the correlation between the expression of SIRT5 and HDAC3 and overall survival revealed that both genes were linked to the prognosis of OSCC patients. HPA and RT-qPCR analysis of whether they are aberrantly expressed in OSCC showed that SIRT5 and HDAC3 were

significantly upregulated in tumor tissues when compared with normal samples.

In OSCC, HDACs are thought to have excellent antitumor potential. It is proposed that RNA splicing and HDACs might be linked, with HDACs controlling acetylation and splicing through interaction with ribonucleoprotein complexes and the spliceosomes (Rahhal and Seto, 2019). Thus, we hypothesized that dysregulated acetylation might influence OSCC development by controlling RNA splicing. Lysine acetylation has been associated with the ribosome pathway, especially with the loss of acetylation on RPS6 and RPS3, which might have therapeutic target potential against OSCCs. (Dong et al., 2022). SIRT5 has been implicated in various malignancies. LDHA-K118su, a SIRT5 substrate markedly elevates invasion and migration by prostate cancer cells (Kwon et al., 2022). SIRT5 negatively regulates cancer cell proliferation in pancreatic ductal adenocarcinoma patients and is related to better prognosis. SIRT5 has also been associated with metabolic regulation and changes in the tumor microenvironment (Sun et al., 2022) in promoting hepatocarcinogenesis. SIRT5 deficiency can increase immune cell activity, indicating that it influences immune cell development (Wang K. et al., 2020).Our immune analyses indicate that acetylation influences the OSCC TME composition.

Next, two clusters were generated based on 'HDAC3' and 'SIRT5'. To further evaluate the prognostic value of these acetylation-related regulatory factors, we used univariate and LASSO regression analyses to construct a risk model using eight genes and then validated its performance on an external dataset. We show that in OSCC patients, risk score is a reliable predictor of OS. Next, we developed a nomogram for clinical analysis of individualized prognosis and risk based on a risk score, age and stage. The calibration curve revealed a high fitness between the actual and predicted OS rates. Taken together, these findings indicate that the prognostic risk scoring model based on the eight-gene signature is an effective indicator of OSCC prognosis.

Next, we further investigated the eight genes used to construct the model. NK2 homeobox 3 (NKX2-3) has been reported as a prognostic factor in head and neck squamous cell carcinoma (HNSCC) (Huang L. et al., 2021; Liu et al., 2021). Suppressor APC domain containing neuroblastoma (SAPCD2) (Zhang et al., 2022), has been reported to regulate Yap/Taz, MAPK, and mTOR signaling in various cancers, including colorectal (Luo et al., 2020) and prostate cancer (Sun et al., 2021). Serine peptidase inhibitor Kazal type 7 (SPINK7) has also been proposed as a prognostic factor also a molecular biomarker in HNSCC (Pennacchiotti et al., 2021; Du et al., 2022). Ly6/ neurotoxin 1 (LYNX1) has been suggested as a prognostic factor in ovarian serous cystadenocarcinoma (Liu et al., 2020) and glioblastoma (Ren et al., 2022). A quantitative sequencing study found that LYNX1 expression significantly increased the recurrence of methylation groups in oropharyngeal tumors. Aldo-keto reductase family 1 member C3 (AKR1C3) has been associated with poor prognosis in patients with oropharyngeal cancer, especially in HPV-positive patients (Peraldo-Neia et al., 2021). Synaptotagmin 17 (SYT17) was found to be differentially expressed in non-Hodgkin's lymphoma (Fucà et al., 2021). MBL associated serine protease 1 (MASP1) has also been proposed as a prognostic factor in HNSCC and oral cancer (Belotti et al., 2021; Zhang and Wang, 2022). Cathepsin G (CTSG) overexpression is associated with poor diffuse large B-cell lymphoma survival (Carreras et al., 2021).

Numerous studies have found that the TME significantly influences cancer incidence, development, and metastasis (Belli et al., 2018; Laplane et al., 2019). Our analysis found that higher immune/stromal scores, were associated with lower risk scores, however, tumor purity had the opposite effect. In OSCC patients, higher risk scores predict a worse prognosis, which demonstrated that the higher the number of immune cells in OSCC, the more difficult it is to identify cancer cells (Gandara et al., 2018). The low infiltration level of antitumor immune cells indicates that immune function was impaired in the high-risk group (Li et al., 2017). Comparing the immune cell infiltration in high- and low-risk groups revealed that the number of invading immune cells in the high-risk group was less than in the low-risk group.

Intriguingly, we found that the proportion of Tregs was higher in the low-risk group than in the high-risk group. Tregs have been associated with subpar clinical outcomes and have been shown to downregulate anti-tumor immunity (Wolf et al., 2005; Toker et al., 2018). This might be explained by the need for Tregs in the TME to control excessive acetylation-induced inflammation Additionally, two key Treg subtypes identified in colon cancer have been shown to have competing roles in controlling the TME (Saito et al., 2016). The risk score was negatively associated with B cell infiltration. B cell infiltration in OSCC has not been extensively studied and available literature is inconsistent. B cell infiltration has been shown to enhance immunological function (Ammirante et al., 2010) while impairing T cell-dependent responses (Shalapour et al., 2015). Therefore, the different Treg subtypes in OSCC should be considered. Except for the APC co-stimulation pathway and MHC class I, the activities of other immunological pathways differed significantly between the two cohorts. These data suggest that a decrease in antitumor immunity may cause the low survival rates in high-risk OSCCs.

CCL2 is an important chemokine that is reported to promote the proliferation and metastasis of osteosarcoma cells by activating NF-κB signaling (Lazennec and Richmond, 2010; Chen et al., 2015). In the category of biological processes, the inflammatory reaction had the strongest correlation with risk scores. Inflammatory responses are reported to be crucial for cancer development, growth, malignant transformation, invasion, and metastasis (Tang et al., 2018). By controlling therapeutic response and immunological surveillance, inflammation also affects patient survival (Grivennikov et al., 2010).

Recent advances in bioinformatics have led to the development of powerful tools for identifying new cancer treatment targets, including for OSCC, based on tumor immunotherapy and microarray sequencing (Almangush et al., 2021; Huang G. G. et al., 2021). Although anti-PD-1/PD-L1 immunotherapy has been widely used to treat terminal OSCCs, only a limited number of cases benefit from this therapy (Dong et al., 2021). Hadler-Olsen et al. (Hadler-Olsen and Wirsing, 2019) discovered that CD163+ M2 and CD57+ showed a positive correlation with the outcome OSCC outcomes.

LAG3, TIGIT, PDCD1, CTLA4, PDCD1LG2, and CD274 checkpoints exhibit significant differences between patients with different risk scores. This may offer new immunotherapy strategies for OSCC and raises the possibility that patients in the high-risk category may benefit from ICP inhibitor treatment than patients with low-risk scores. In

conclusion, our data indicate that immunosuppression might underlie poor prognosis in high-risk patients and that acetylation may be important for OSCC immunotherapy. However, this study has some limitations. First, the OSCC samples used are from public databases. Secondly, although our prognostic model has been confirmed in different datasets, the study is retrospective. To validate the clinical utility of the developed model, additional, well-designed studies are required. To determine the pathways involved, the identified genes should undergo experimental validation, either in cancer cells or mouse models. Additionally, we did not perform our own sequencing, and the follow-up data, as well as the sample size, were too small to carry out a similar survival study. We anticipate that the limitations highlighted above will define the scope and depth of our future research.

Few studies have examined the acetylation mechanisms underlying OSCC. Here, we identified two prognostic markers associated with acetylation in OSCC, SIRT5 and HDAC3, which are overexpressed in tumors, and found that their upregulation is associated with poor OS. We conducted a basic study on the prognostic value of these LARGs and built up some theoretical evidences to support future researches. The prognostic value of both genes warrants further validation using clinical data. Importantly, the prognosis model based on univariate Cox and LASSO regression analyses is closely associated with immune cell infiltration.

Data availability statement

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

Ethics statement

The studies involving humans were approved by Zhongnan Hospital of Wuhan University Medical Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The human samples used in this study were acquired from primarily isolated as part of our previous study for which ethical approval was obtained. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

Funding

Acknowledgments

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.2023.1185832/full#supplementary-material

References

Alba, A. C., Agoritsas, T., Walsh, M., Hanna, S., Iorio, A., Devereaux, P. J., et al. (2017). Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *Jama* 318 (14), 1377–1384. doi:10.1001/jama.2017.12126

Almangush, A., Leivo, I., and Makitie, A. A. (2021). Biomarkers for immunotherapy of oral squamous cell carcinoma: current status and challenges. *Front. Oncol.* 11, 616629. doi:10.3389/fonc.2021.616629

Ammirante, M., Luo, J. L., Grivennikov, S., Nedospasov, S., and Karin, M. (2010). B-cell-derived lymphotoxin promotes castration-resistant prostate cancer. *Nature* 464 (7286), 302–305. doi:10.1038/nature08782

Attar, N., and Kurdistani, S. K. (2017). Exploitation of EP300 and CREBBP lysine acetyltransferases by cancer. *Cold Spring Harb. Perspect. Med.* 7 (3), a026534. doi:10.1101/cshperspect.a026534

Bai, L. Y., Chiu, C. F., Pan, S. L., Sargeant, A. M., Shieh, T. M., Wang, Y. C., et al. (2011). Antitumor activity of a novel histone deacetylase inhibitor (S)-HDAC42 in oral squamous cell carcinoma. *Oral Oncol.* 47 (12), 1127–1133. doi:10.1016/j.oraloncology.2011.07.031

Belli, C., Trapani, D., Viale, G., D'Amico, P., Duso, B. A., Della Vigna, P., et al. (2018). Targeting the microenvironment in solid tumors. *Cancer Treat. Rev.* 65, 22–32. doi:10.1016/j.ctrv.2018.02.004

Belotti, Y., Lim, S. B., Iyer, N. G., Lim, W. T., and Lim, C. T. (2021). Prognostic matrisomal gene panel and its association with immune cell infiltration in head and neck carcinomas. *Cancers (Basel)* 13 (22), 5761. doi:10.3390/cancers13225761

Carreras, J., Hiraiwa, S., Kikuti, Y. Y., Miyaoka, M., Tomita, S., Ikoma, H., et al. (2021). Artificial neural networks predicted the overall survival and molecular subtypes of diffuse large B-cell lymphoma using a pancancer immune-oncology panel. *Cancers (Basel)* 13 (24), 6384. doi:10.3390/cancers13246384

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. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019

Chen, Q., Sun, W., Liao, Y., Zeng, H., Shan, L., Yin, F., et al. (2015). Monocyte chemotactic protein-1 promotes the proliferation and invasion of osteosarcoma cells and upregulates the expression of AKT. *Mol. Med. Rep.* 12 (1), 219–225. doi:10.3892/mmr.2015.3375

Czuba, L. C., Hillgren, K. M., and Swaan, P. W. (2018). Post-translational modifications of transporters. *Pharmacol. Ther.* 192, 88–99. doi:10.1016/j.pharmthera.2018.06.013

Ding, P., Ma, Z., Liu, D., Pan, M., Li, H., Feng, Y., et al. (2022). Lysine acetylation/deacetylation modification of immune-related molecules in cancer immunotherapy. *Front. Immunol.* 13, 865975. doi:10.3389/fimmu.2022.865975

Dong, J., He, J., Zhang, Z., Zhang, W., Li, Y., Li, D., et al. (2022). Identification of lysine acetylome of oral squamous cell carcinoma by label-free quantitative proteomics. *J. Proteomics* 262, 104598. doi:10.1016/j.jprot.2022.104598

Dong, Y., Wang, Z., Mao, F., Cai, L., Dan, H., Jiang, L., et al. (2021). PD-1 blockade prevents the progression of oral carcinogenesis. *Carcinogenesis* 42 (6), 891–902. doi:10.1093/carcin/bgab035

Du, P., Chai, Y., Zong, S., Yue, J., and Xiao, H. (2022). Identification of a prognostic model based on fatty acid metabolism-related genes of head and neck squamous cell carcinoma. *Front. Genet.* 13, 888764. doi:10.3389/fgene.2022.888764

Ferris, R. L., Blumenschein, G., Jr., Fayette, J., Guigay, J., Colevas, A. D., Licitra, L., et al. (2016). Nivolumab for recurrent squamous-cell carcinoma of the head and neck. *N. Engl. J. Med.* 375 (19), 1856–1867. doi:10.1056/NEJMoa1602252

Fucà, G., Ambrosini, M., Agnelli, L., Brich, S., Sgambelluri, F., Mortarini, R., et al. (2021). Fifteen-year follow-up of relapsed indolent non-Hodgkin lymphoma patients vaccinated with tumor-loaded dendritic cells. *J. Immunother. cancer* 9 (6), e002240. doi:10.1136/jitc-2020-002240

Gandara, D. R., Paul, S. M., Kowanetz, M., Schleifman, E., Zou, W., Li, Y., et al. (2018). Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. *Nat. Med.* 24 (9), 1441–1448. doi:10.1038/s41591-018-0134-3

Gao, Y., Nihira, N. T., Bu, X., Chu, C., Zhang, J., Kolodziejczyk, A., et al. (2020). Acetylation-dependent regulation of PD-L1 nuclear translocation dictates the efficacy of anti-PD-1 immunotherapy. *Nat. Cell Biol.* 22 (9), 1064–1075. doi:10.1038/s41556-020-0562-4

Geeleher, P., Cox, N. J., and Huang, R. S. (2014). Clinical drug response can be predicted using baseline gene expression levels and *in vitro* drug sensitivity in cell lines. *Genome Biol.* 15 (3), R47. doi:10.1186/gb-2014-15-3-r47

Ghantous, Y., and Abu Elnaaj, I. (2017). Global incidence and risk factors of oral cancer. *Harefuah* 156 (10), 645–649.

Grivennikov, S. I., Greten, F. R., and Karin, M. (2010). Immunity, inflammation, and cancer. *Cell* 140 (6), 883–899. doi:10.1016/j.cell.2010.01.025

Hadler-Olsen, E., and Wirsing, A. M. (2019). Tissue-infiltrating immune cells as prognostic markers in oral squamous cell carcinoma: a systematic review and meta-analysis. *Br. J. cancer* 120 (7), 714–727. doi:10.1038/s41416-019-0409-6

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

Hatanaka, Y., Niinuma, T., Kitajima, H., Nishiyama, K., Maruyama, R., Ishiguro, K., et al. (2021). DLEU1 promotes oral squamous cell carcinoma progression by activating interferon-stimulated genes. *Sci. Rep.* 11 (1), 20438. doi:10.1038/s41598-021-99736-5

Hu, M., He, F., Thompson, E. W., Ostrikov, K. K., and Dai, X. (2022). Lysine acetylation, cancer hallmarks and emerging onco-therapeutic opportunities. *Cancers (Basel)* 14 (2), 346. doi:10.3390/cancers14020346

Huang, G. Z., Chen, Z. Q., Wu, J., Shao, T. R., Zou, C., Ai, Y. L., et al. (2021). Pan-cancer analyses of the tumor microenvironment reveal that ubiquitin-conjugating enzyme E2C might Be a potential immunotherapy target. *J. Immunol. Res.* 2021, 9250207. doi:10.1155/2021/9250207

Huang, L., Yu, X., Jiang, Z., and Zeng, P. (2021). Novel autophagy-related gene signature investigation for patients with oral squamous cell carcinoma. *Front. Genet.* 12, 673319. doi:10.3389/fgene.2021.673319

Jensen, J. S., Jakobsen, K. K., Mirian, C., Ghanizada, M., Hakansson, K., Wessel, I., et al. (2021). Impact of time to treatment initiation for patients with oral cavity squamous cell carcinoma: a population-based, retrospective study. *Acta Oncol.* 60 (4), 491–496. doi:10.1080/0284186X.2020.1863462

Kwon, O. K., Bang, I. H., Choi, S. Y., Jeon, J. M., Na, A. Y., Gao, Y., et al. (2022). LDHA desuccinylase Sirtuin 5 as A novel cancer metastatic stimulator in aggressive prostate cancer. *Genomics, proteomics Bioinforma.* 21, 177–189. doi:10.1016/j.gpb.2022.02.004

Laplane, L., Duluc, D., Bikfalvi, A., Larmonier, N., and Pradeu, T. (2019). Beyond the tumour microenvironment. *Int. J. cancer* 145 (10), 2611–2618. doi:10.1002/ijc.32343

Lazennec, G., and Richmond, A. (2010). Chemokines and chemokine receptors: new insights into cancer-related inflammation. *Trends Mol. Med.* 16 (3), 133–144. doi:10.1016/j.molmed.2010.01.003

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.CAN-17-0307

Lin, X., Wu, W., Ying, Y., Luo, J., Xu, X., Zheng, L., et al. (2022). MicroRNA-31: a pivotal oncogenic factor in oral squamous cell carcinoma. *Cell Death Discov.* 8 (1), 140. doi:10.1038/s41420-022-00948-z

Liu, C., Wu, W., Xu, M., Mi, J., Xu, L., and Wang, R. (2021). Development and Validation of an autophagy-related signature for head and neck squamous cell carcinoma. *BioMed Res. Int.* 2021, 1028158. doi:10.1155/2021/1028158

Liu, H., Wang, A., and Ma, Y. (2020). Increased expression of LYNX1 in ovarian serous cystadenocarcinoma predicts poor prognosis. *BioMed Res. Int.* 2020, 1392674. doi:10.1155/2020/1392674

Luo, Y., Wang, L., Ran, W., Li, G., Xiao, Y., Wang, X., et al. (2020). Overexpression of SAPCD2 correlates with proliferation and invasion of colorectal carcinoma cells. *Cancer Cell Int.* 20, 43. doi:10.1186/s12935-020-1121-6

Marques, A. E. M., do Nascimento Filho, C. H. V., Marinho Bezerra, T. M., Guerra, E. N. S., Castilho, R. M., and Squarize, C. H. (2020). Entinostat is a novel therapeutic agent to treat oral squamous cell carcinoma. *J. Oral Pathol. Med.* 49 (8), 771–779. doi:10.1111/jop.13039

Narita, T., Weinert, B. T., and Choudhary, C. (2019). Functions and mechanisms of non-histone protein acetylation. *Nat. Rev. Mol. Cell Biol.* 20 (3), 156–174. doi:10.1038/s41580-018-0081-3

Ning, Y., Ding, J., Sun, X., Xie, Y., Su, M., Ma, C., et al. (2020). HDAC9 deficiency promotes tumor progression by decreasing the CD8(+) dendritic cell infiltration of the tumor microenvironment. *J. Immunother. cancer* 8 (1), e000529. doi:10.1136/jitc-2020-000529

O'Garro, C., Igbineweka, L., Ali, Z., Mezei, M., and Mujtaba, S. (2021). The biological significance of targeting acetylation-mediated gene regulation for designing new mechanistic tools and potential therapeutics. *Biomolecules* 11 (3), 455. doi:10.3390/biom11030455

Pencina, M. J., and D'Agostino, R. B. (2004). Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Statistics Med.* 23 (13), 2109–2123. doi:10.1002/sim.1802

Pennacchiotti, G., Valdés-Gutiérrez, F., González-Arriagada, W. A., Montes, H. F., Parra, J. M. R., Guida, V. A., et al. (2021). SPINK7 expression changes accompanied by HER2, P53 and RB1 can be relevant in predicting oral squamous cell carcinoma at a molecular level. *Sci. Rep.* 11 (1), 6939. doi:10.1038/s41598-021-86208-z

Peraldo-Neia, C., Ostano, P., Mello-Grand, M., Guana, F., Gregnanin, I., Boschi, D., et al. (2021). AKR1C3 is a biomarker and druggable target for oropharyngeal tumors. *Cell. Oncol. Dordr.* 44 (2), 357–372. doi:10.1007/s13402-020-00571-z

Rahhal, R., and Seto, E. (2019). Emerging roles of histone modifications and HDACs in RNA splicing. *Nucleic acids Res.* 47 (10), 4911–4926. doi:10.1093/nar/gkz292

Rastogi, B., Raut, S. K., Panda, N. K., Rattan, V., Radotra, B. D., and Khullar, M. (2016). Overexpression of HDAC9 promotes oral squamous cell carcinoma growth, regulates cell cycle progression, and inhibits apoptosis. *Mol. Cell Biochem.* 415 (1-2), 183–196. doi:10.1007/s11010-016-2690-5

Ren, W., Jin, W., and Liang, Z. (2022). Construction and Validation of an immune-related risk score model for survival prediction in glioblastoma. *Front. neurology* 13, 832944. doi:10.3389/fneur.2022.832944

Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C., Nussenbaum, B., and Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves. *official J. Am. Acad. Otolaryngology-Head Neck Surg.* 143 (3), 331–336. doi:10.1016/j.otohns.2010.05.007

Saito, T., Nishikawa, H., Wada, H., Nagano, Y., Sugiyama, D., Atarashi, K., et al. (2016). Two FOXP3(+)CD4(+) T cell subpopulations distinctly control the prognosis of colorectal cancers. *Nat. Med.* 22 (6), 679–684. doi:10.1038/nm.4086

Sakuma, T., Uzawa, K., Onda, T., Shiiba, M., Yokoe, H., Shibahara, T., et al. (2006). Aberrant expression of histone deacetylase 6 in oral squamous cell carcinoma. *Int. J. Oncol.* 29 (1), 117–124. doi:10.3892/ijo.29.1.117

Seiler, M., Huang, C. C., Szalma, S., and Bhanot, G. (2010). ConsensusCluster: a software tool for unsupervised cluster discovery in numerical data. *Omics a J. Integr. Biol.* 14 (1), 109–113. doi:10.1089/omi.2009.0083

Shalapour, S., Font-Burgada, J., Di Caro, G., Zhong, Z., Sanchez-Lopez, E., Dhar, D., et al. (2015). Immunosuppressive plasma cells impede T-cell-dependent immunogenic chemotherapy. *Nature* 521 (7550), 94–98. doi:10.1038/nature14395

Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011). Regularization paths for cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* 39 (5), 1–13. doi:10.18637/jss.v039.i05

Sun, R., Zhang, Z., Bao, R., Guo, X., Gu, Y., Yang, W., et al. (2022). Loss of SIRT5 promotes bile acid-induced immunosuppressive microenvironment and hepatocarcinogenesis. *J. hepatology* 77 (2), 453–466. doi:10.1016/j.jhep.2022.02.030

Sun, Z., Mao, Y., Zhang, X., Lu, S., Wang, H., Zhang, C., et al. (2021). Identification of ARHGEF38, NETO2, GOLM1, and SAPCD2 associated with prostate cancer progression by bioinformatic analysis and experimental Validation. *Front. Cell Dev. Biol.* 9, 718638. doi:10.3389/fcell.2021.718638

Tang, F., Wang, Y., Hemmings, B. A., Rüegg, C., and Xue, G. (2018). PKB/Akt-dependent regulation of inflammation in cancer. *Semin. Cancer Biol.* 48, 62–69. doi:10.1016/j.semcancer.2017.04.018

Toker, A., Nguyen, L. T., Stone, S. C., Yang, S. Y. C., Katz, S. R., Shaw, P. A., et al. (2018). Regulatory T cells in ovarian cancer are characterized by a highly activated phenotype distinct from that in melanoma. *Clin. cancer Res.* 24 (22), 5685–5696. doi:10.1158/1078-0432.CCR-18-0554

von Mering, C., Jensen, L. J., Snel, B., Hooper, S. D., Krupp, M., Foglierini, M., et al. (2005). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic acids Res.* 33, D433–D437. doi:10.1093/nar/gki005

Wang, H., Fu, C., Du, J., Wang, H., He, R., Yin, X., et al. (2020). Enhanced histone H3 acetylation of the PD-L1 promoter via the COP1/c-Jun/HDAC3 axis is required for PD-L1 expression in drug-resistant cancer cells. *J. Exp. Clin. Cancer Res.* 39 (1), 29. doi:10.1186/s13046-020-1536-x

Wang, K., Hu, Z., Zhang, C., Yang, L., Feng, L., Yang, P., et al. (2020). SIRT5 contributes to colorectal cancer growth by regulating T cell activity. *J. Immunol. Res.* 2020, 3792409. doi:10.1155/2020/3792409

Warnakulasuriya, S. (2009). Global epidemiology of oral and oropharyngeal cancer. *Oral Oncol.* 45 (4-5), 309–316. doi:10.1016/j.oraloncology.2008.06.002

Wolf, D., Wolf, A. M., Rumpold, H., Fiegl, H., Zeimet, A. G., Muller-Holzner, E., et al. (2005). The expression of the regulatory T cell-specific forkhead box transcription factor FoxP3 is associated with poor prognosis in ovarian cancer. *Clin. cancer Res.* 11 (23), 8326–8331. doi:10.1158/1078-0432.CCR-05-1244

Yang, C., Croteau, S., and Hardy, P. (2021). Histone deacetylase (HDAC) 9: versatile biological functions and emerging roles in human cancer. *Cell. Oncol. Dordr.* 44 (5), 997–1017. doi:10.1007/s13402-021-00626-9

Zanoni, D. K., Montero, P. H., Migliacci, J. C., Shah, J. P., Wong, R. J., Ganly, I., et al. (2019). Survival outcomes after treatment of cancer of the oral cavity (1985-2015). *Oral Oncol.* 90, 115–121. doi:10.1016/j.oraloncology.2019.02.001

Zhang, L., and Wang, X. (2022). An immune-related gene signature can predict clinical outcomes and immunotherapeutic response in oral squamous cell carcinoma. *Front. Genet.* 13, 870133. doi:10.3389/fgene.2022.870133

Zhang, Z. M., Cao, H. B., Li, Z. H., Zhuo, R., Tao, Y. F., Li, X. L., et al. (2022). SAPCD2 promotes neuroblastoma progression by altering the subcellular distribution of E2F7. *Cell death Dis.* 13 (2), 174. doi:10.1038/s41419-022-04624-z

UroVysion™ fluorescence *in situ* hybridization (FISH) possibly has a high positive rate in carcinoma of non-urothelial lineages

Chunjin Ke[1], Xuguang Liu[2], Jie Wan[2], Zhiquan Hu[1]* and Chunguang Yang[1]*

[1]Department of Urology, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology (HUST), Wuhan, China, [2]Department of Pathology, Tongji Hospital of Tongji Medical College, Huazhong University of Science and Technology (HUST), Wuhan, China

Background: Positive UroVysion™ fluorescence *in situ* hybridization (FISH) is generally considered as urothelial carcinoma (UC). We clarify if UroVysion™ FISH can be positive in carcinoma of non-urothelial lineages (CNUL), and verify the consistency of urine FISH and histological FISH in CNUL.

Methods: All CNUL subjects detected by urine FISH assay due to haematuria from Tongji Hospital were screened. Meanwhile, 2 glandular cystitis and 2 urothelial carcinoma were served as negative or positive control. Paraffin-embedded tissue sections of all subjects were sent to the pathology department for histological FISH detection.

Results: A total of 27 patients were included in this study, including 9 with adenocarcinomas, 11 with squamous cell carcinomas, and 7 with other tumour types. The overall positive rate in urine FISH was 64.00% (16/25) in patients with CNUL, 77.78% (7/9) in those with adenocarcinoma and 54.55% (6/11) in those with squamous carcinoma. There was a significant difference in the GLP p16 gene deletion rate between UC and CNUL (100% vs. 8.00%, $p = 0.017$). Histological FISH results showed that the histological results of 19 patients were consistent with their urine FISH results, and only one patient with stage IIIa urachal carcinoma had inconsistent histological FISH results (positive) and urine FISH (negative) results.

Conclusion: We demonstrated for the first time the application value of FISH in CNUL on urine samples. Positive urine FISH tests indicate not only UC, but also CNUL. UroVysion™ FISH possibly has a high positive rate in CNUL. CNUL and UC have different genetic changes shown by FISH.

KEYWORDS

fluorescence *in situ* hybridization, chromosome, CNUL, squamous cell carcinoma, adenocarcinoma

1 Introduction

Fluorescence *in situ* hybridization (FISH) detects chromosomal or genetic abnormalities in cell and tissue samples by detecting fluorescence signals through fluorescence microscopy after hybridization between the probe and the DNA of the sample through the complementarity of DNA base pairs (Wiegant et al., 1991; Sokolova et al., 2000; Levsky and Singer, 2003; National Library of Medicine, 2020). The U.S. Food and Drug

TABLE 1 Basic clinical data of 27 patients.

No	Age (years)	Sex	Diagnosis	Urine FISH (+/−)	Abnormal cell ratio (%)				Histological FISH (+/−)	Genetic material changes (+/−)			
					CSP3	CSP7	GLP p16	CSP17		CSP3	CSP7	GLP p16	CSP17
1	69	male	Urothelial carcinoma	+	47	47	45	51	+	+	+	+	+
2	64	female	Urothelial carcinoma	+			19	15	+	-	-	+	+
3	66	male	Small cell carcinoma of the bladder	+	83	86		85	+	+	+	-	+
4	34	male	Bladder paraganglioma	+	65			75					
5	50	female	Cystitis glandularis	-					-	-	-	-	-
6	53	male	Cystitis glandularis	-					-	-	-	-	-
7	50	male	Renal secondary non-Hodgkin lymphoma	+	63	63		58					
8	25	male	Urachal adenocarcinoma	+	39	47			+	+	+	-	-
9	49	female	Urachal adenocarcinoma	-					-	-	-	-	-
10	50	female	Urachal adenocarcinoma	-					+	-	+	-	+
11	68	male	Urachal adenocarcinoma	+	14	12							
12	46	male	Urachal adenocarcinoma	+	27	30	31	32					
13	54	male	Urachal adenocarcinoma with distant visceral metastasis	+	31	29							
14	30	male	Urachal adenocarcinoma	+	15	17		18					
15	69	male	Prostate cancer invades the bladder	+	33	35		31	+	+	+	-	+
16	72	male	Prostate cancer invades the bladder	+	67	69		67	+	+	+	-	+
17	56	male	Renal squamous cell carcinoma (oesophageal squamous cell carcinoma metastasis)	+	23	23		21	+	+	+	-	+
18	65	male	Renal squamous cell carcinoma	-					-	-	-	-	-
19	58	male	Renal squamous cell carcinoma	-					-	-	-	-	-
20	63	male	Renal squamous cell carcinoma (lung squamous cell carcinoma metastases)	+	37	35		41					
21	51	male	Renal squamous cell carcinoma	-					-	-	-	-	-
22	52	female		+	17	17		31	+	+	+	-	+

TABLE 1 (*Continued*) Basic clinical data of 27 patients.

No	Age (years)	Sex	Diagnosis	Urine FISH (+/−)	Abnormal cell ratio (%)				Histological FISH (+/−)	Genetic material changes (+/−)			
					CSP3	CSP7	GLP p16	CSP17		CSP3	CSP7	GLP p16	CSP17
			Renal squamous cell carcinoma										
23	39	male	Renal squamous cell carcinoma	-					-	-	-	-	-
24	77	male	Renal squamous cell carcinoma	-					-	-	-	-	-
25	64	male	Bladder squamous cell carcinoma	+	82	82		78	+	+	+	-	+
26	61	male	Bladder squamous cell carcinoma	+	73	75	68	71	+	+	+	+	+
27	50	male	Bladder squamous cell carcinoma	+	85	87		79	+	+	+	-	+

Note: CSP, chromosomal centromeric probe; GLP, gene locus-specific probe; "+", a positive FISH assay; "-", a negative FISH, assay.

TABLE 2 Changes in genetic material of different tumour types.

Diagnosis	Chromosomal amplification/gene deletion				P
	3 #	7 #	p16	17 #	
Adenocarcinoma (n = 9)					$P^1 > 0.05$
Urachal adenocarcinoma (n = 7)	5/7 (71.42%)	5/7 (71.42%)	1/7 (14.29%)	2/7 (28.57%)	
prostate cancer (n = 2)	2/2 (100%)	2/2 (100%)	0	2/2 (100%)	
Total	7/9 (77.78%)	7/9 (77.78%)	1/9 (11.11%)	4/9 (44.44%)	
Squamous cell carcinoma (n = 11)					
Primary renal squamous cell carcinoma (n = 6)	1/6 (16.67%)	1/6 (16.67%)	0	1/6 (16.67%)	
Primary bladder squamous cell carcinoma (n = 3)	3/3 (100%)	3/3 (100%)	1/3 (33.33%)	3/3 (100%)	
Secondary renal squamous cell carcinoma (n = 2)	2/2 (100%)	2/2 (100%)	0	2/2 (100%)	
Total	6/11 (54.55%)	6/11 (54.55%)	1/11 (9.09%)	6/11 (54.55%)	
Metastatic tumour (n = 3)					
Renal squamous cell carcinoma (oesophageal squamous cell carcinoma metastasis, n = 1)	1/1 (100%)	1/1 (100%)	0	1/1 (100%)	
Renal squamous cell carcinoma (lung squamous cell carcinoma metastasis, n = 1)	1/1 (100%)	1/1 (100%)	0	1/1 (100%)	
Renal secondary non-Hodgkin lymphoma (haematologic lymphoma metastasis, n = 1)	1/1 (100%)	1/1 (100%)	0	1/1 (100%)	
Total	3/3 (100%)	3/3 (100%)		3/3 (100%)	
Other types (n = 6)					
Urothelial carcinoma (n = 2)	1/2 (50%)	1/2 (50%)	2/2 (100%)	2/2 (100%)	$P^2 = 0.017$
Small cell carcinoma of the bladder (n = 1)	1/1 (100%)	1/1 (100%)	0	1/1 (100%)	
Bladder paraganglioma (n = 1)	1/1 (100%)	0	0	1/1 (100%)	
Cystitis glandularis (n = 2)	0	0	0	0	
Total	3/6 (50%)	2/6 (33.33%)	2/6 (33.33%)	4/6 (66.67%)	

Note: P^1: compared with adenocarcinoma, there was no statistical significance in the number amplification of chromosomes 3, 7 and 17 and GLP p16 gene deletion in squamous cell carcinoma ($p > 0.05$); P^2: GLP p16 gene deletion rate was different between non-urothelial carcinoma and urothelial carcinoma ($p = 0.017$).
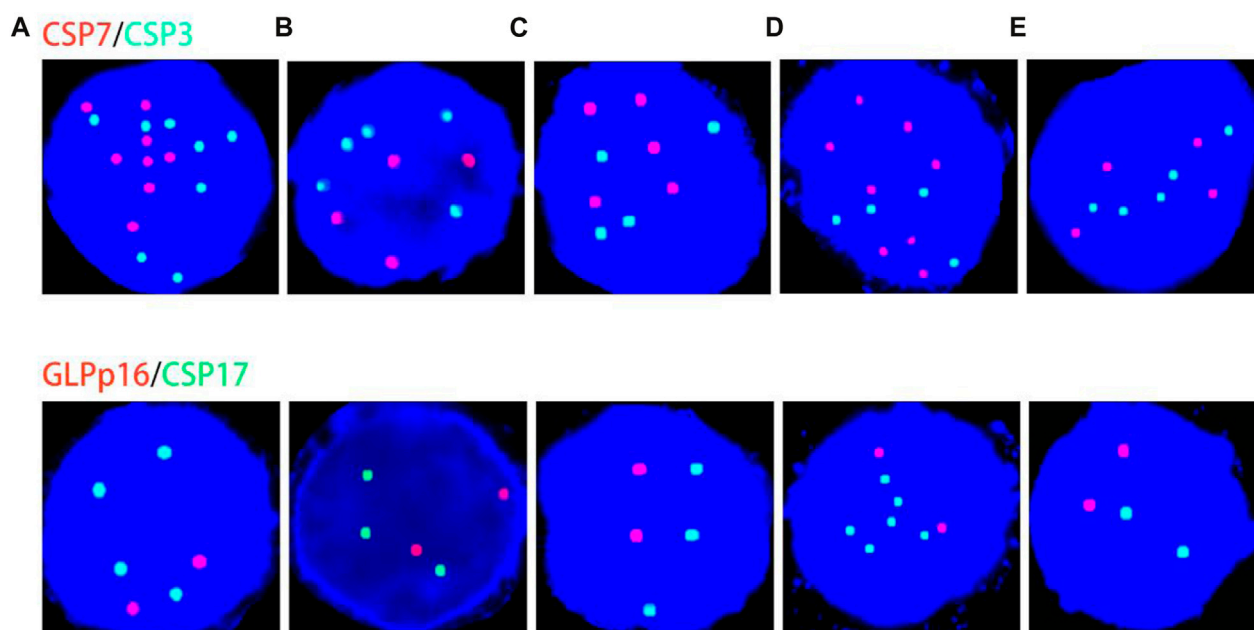
FIGURE 1
Genetic material changes in patients with distant metastases or highly malignant tumors. **(A–E)** correspond to case No. 7 (secondary non-Hodgkin lymphoma of the kidney), case No. 13 (urachal carcinoma with distant visceral metastasis), case No. 17 (renal pelvis squamous cell carcinoma derived from esophageal squamous cell carcinoma), case No. 20 (renal pelvis squamous cell carcinoma derived from lung squamous cell carcinoma metastases), and case No. 25 (primary bladder squamous cell carcinoma), respectively. Red represents CSP7 and GLP p16. Green represents CSP3 and CSP17. Note: Case B cited a case in the previous published articles of our team (Hu et al., 2020). Case A and case C cited the cases from our previous published articles of our team (Hu et al., 2021).

Administration approved UroVysion™ FISH probes (chromosomes 3, 7, and 17 combined with the 9p21 probe) in 2001 and 2005, respectively, for urine detection in patients with suspected bladder cancer and postoperative recurrence monitoring in patients with bladder cancer (Halling et al., 2000). Worldwide, the incidence of bladder cancer ranks 9th among all malignant tumours in the body and 7th among male patients. Regarding histopathological classification, more than 90% of cases are bladder urothelial carcinoma (Jemal et al., 2010; Siegel et al., 2017). In recent years, the incidence of bladder cancer in China has been increasing year by year, with an average growth rate of 68.29% in the past 15 years, due to changes in diet, increased work pressure, harsh environment and other factors (Chen, 2015; Chen et al., 2016). Therefore, FISH positive is usually considered to be urothelial carcinoma (UC).

A review of the national and international literature shows that there are very few studies on the application of FISH in carcinoma of non-urothelial lineages (CNUL). (Reid-Nicholson et al., 2009) performed histological FISH detection on paraffin sections from CNUL patients and found that FISH positivity was common in primary and secondary adenocarcinomas but rare in squamous cell carcinomas. (Kipp et al., 2008) also performed histological FISH on paraffin sections and found that chromosomal abnormalities detected in UC were common in rare histological types of bladder cancer. (Yang et al., 2016) found that preoperative urine FISH were positive in patients with bladder paraganglioma, which showed polyploidy of chromosomes 3 and 17. Urine FISH was performed again after the operation and the result turned negative. In our clinical practice, we successively found that urine FISH showed positive manifestations in urachal carcinoma (Case 8-Case 14),

(Hu et al., 2020), renal secondary non-Hodgkin lymphoma (Case 7) and renal squamous cell carcinoma (Case 17), (Hu et al., 2021), so we did a comprehensive review of all the cases in our center since the FISH technique was introduced.

In summary, it is clear that the positive presentation of urine FISH in CNUL is not coincidental, however, none of the existing studies have been cross-validated by histological FISH with urine FISH, thus causing a lack of studies to demonstrate the relationship between the two specimen types. This study focuses on elucidating that FISH can also show positive in urine or tissue specimens of CNUL, thus suggesting that FISH-positive patients do not always have UC. The second is to confirm the consistency of histological and urine FISH analysis results to fill the gap of previous studies.

2 Methods

2.1 Research objects

With the approval of the Medical Ethics Committee of Tongji Hospital affiliated with Tongji Medical College, Huazhong University of Science and Technology (Approval No. TJ-IRB20210521), we applied to the Department of Pathology to query the information of patients with CNUL admitted to the Department of Urology in the past 10 years, including all squamous carcinomas, adenocarcinomas and other rare types of tumours. Then, we retrospectively searched the FISH database at our Institute of Urology for relevant FISH testing information for these
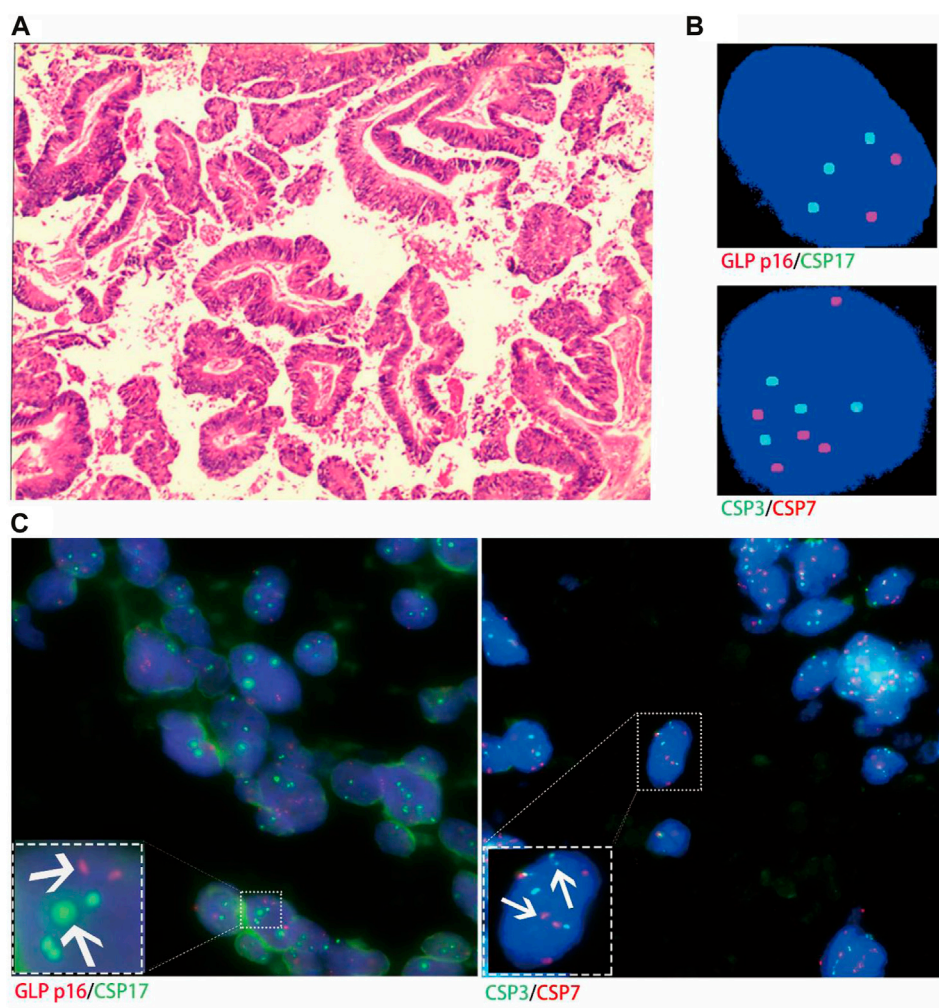
FIGURE 2
Case No.15. Positive validation of histological FISH and urine FISH in adenocarcinoma. **(A)** is the histopathological image of prostate cancer. Microscopy showed moderately and poorly differentiated with Gleason score 4 + 4 = 8(hematoxylin-eosin staining, ×200 magnification). **(B)** is the urine FISH result, showing the amplification of chromosomes 3, 7 and 17, and no GLP p16 gene deletion. **(C)** is histological FISH (×400 magnification), also showing amplification of chromosomes 3, 7 and 17, without GLP p16 gene deletion (indicated by arrows). The mean fluorescence signals of chromosome 3, 7, 17 and GLP p16 locus in each cell were 3.48, 3.94, 2.93 and 2.06, respectively. Red represents CSP7 and GLP p16. Green represents CSP3 and CSP17.

patients. A total of 25 patients who met the requirements were screened, and 2 patients with UC were selected as the control group. Inclusion criteria were as follows: ① can obtain specific clinical data through the electronic medical record system; ② have histological specimens in our hospital and have been pathologically diagnosed as CNUL; ③ have no urinary calculi, chemotherapy, radiotherapy, etc.

2.2 Research methods

2.2.1 Detection method of urine FISH

The specific results of urine FISH in 27 patients were obtained directly from the Institute of Urology of Tongji hospital. Approximately 200 ml of urine was collected in the morning. The volume of urine specimen should not be too small, otherwise it will affect the number of cells in the specimen and cannot meet

the basic requirements of FISH detection technology. Urine specimens should be kept free of contaminants such as prostate fluid, semen, leukorrhea, menstrual blood, etc. After the specimen is collected, it is necessary to send it for examination as soon as possible to prevent the dissolution of cells in the specimen, resulting in changes in the composition of the specimen.

2.2.2 Detection method of histological FISH

Haematoxylin and eosin-stained slides (and immunohistochemistry slides, if applicable) of the paraffin-embedded tissue of the relevant patient were first requested from the Department of Pathology, and then representative paraffin blocks containing ≥ 80% tumour cells were selected for histological FISH. The target areas for hybridization were highlighted on each representative slide. All metastatic tumours were confirmed by immunohistochemistry and/or review of the
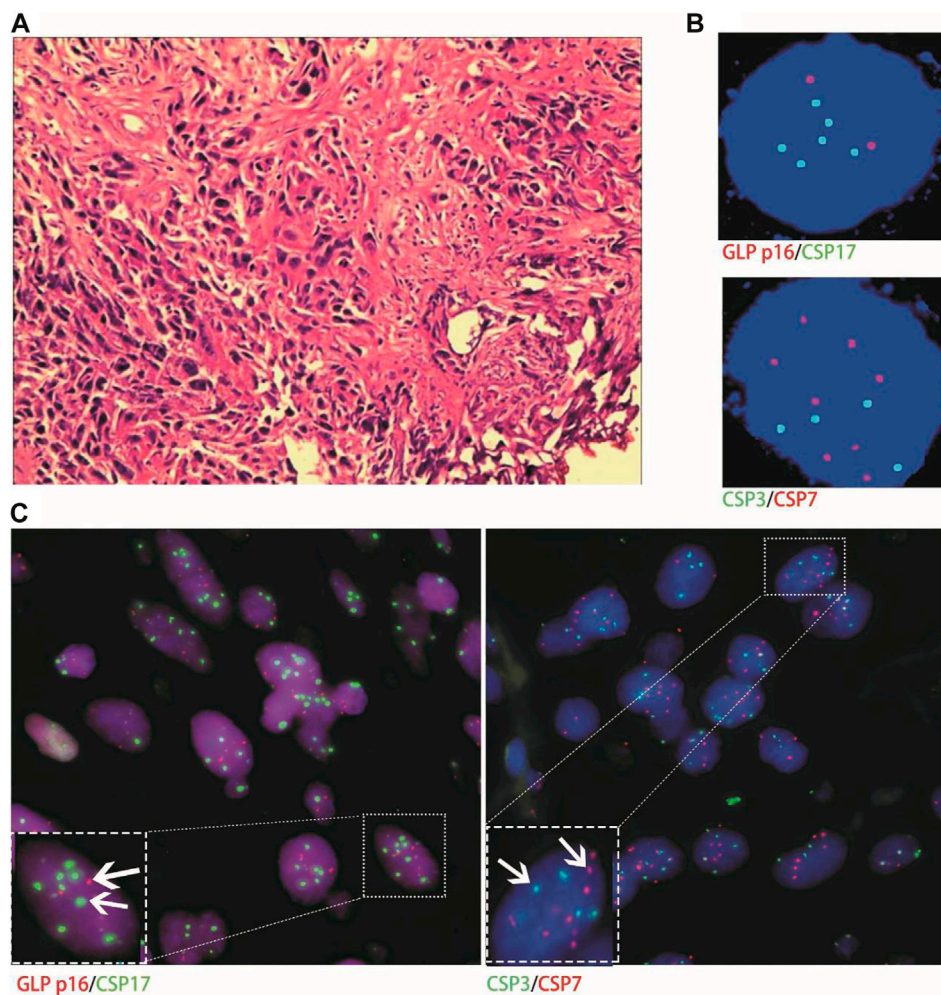
FIGURE 3
Case No.25. Positive validation of histological FISH and urine FISH in squamous cell carcinoma. **(A)** is the histopathological image of a patient with bladder squamous cell carcinoma, showing invasive squamous cell carcinoma under a microscope (hematoxylin-eosin staining, ×200 magnification); **(B)** is the urine FISH result, showing the amplification of chromosomes 3, 7 and 17, and no GLP p16 gene deletion; **(C)** is histological FISH (×400 magnification), also showing amplification of chromosomes 3, 7 and 17, without GLP p16 gene deletion (indicated by arrows). The mean fluorescence signals of chromosome 3, 7, 17 and GLP p16 locus in each cell were 3.64, 4.19, 3.86 and 1.97, respectively. Red represents CSP7 and GLP p16. Green represents CSP3 and CSP17.

primary tumour. The FISH DNA probe kit (Bladder cancer cell chromosome and gene abnormality detection box: China Food and Drug Administration No. 3400251, 2009; Order number F01008-02) was purchased from Beijing Jinpujia Medical Technology Co. Ltd. The FISH DNA probe is labelled with tetramethylrhodamine and fluorescein isothiocyanate and consists of two combinations: CSP3 (green)/CSP7 (red) and GLP p16 (red)/CSP17 (green). For experimental procedures and result interpretation standards (Supplementary Material), refer to the published articles by our team (Hu et al., 2020) and the official website of the kit supplier company (National Library of Medicine, 2020).

2.3 Statistical analysis

Data analysis was performed using IBM SPSS Statistics® version 23 (IBM Corp, Armonk, NY, USA) (Liang et al., 2019). Continuous variables are expressed as the median ± interquartile spacing, and count variables are described as frequencies, ratios and percentages. Categorical variables were compared using the chi-square test and Fisher's exact test when data were limited. Differences with $p < 0.05$ were considered statistically significant.

3 Results

3.1 Urine FISH positive detection rate of CNUL

A total of 27 patients were included in this study: 9 with adenocarcinoma, 11 with squamous carcinoma, and 7 with other types [2 with glandular cystitis, 2 with UC (control group), 1 with small cell carcinoma of the bladder, 1 with renal secondary non-Hodgkin's lymphoma, and 1 with bladder paraganglioma].
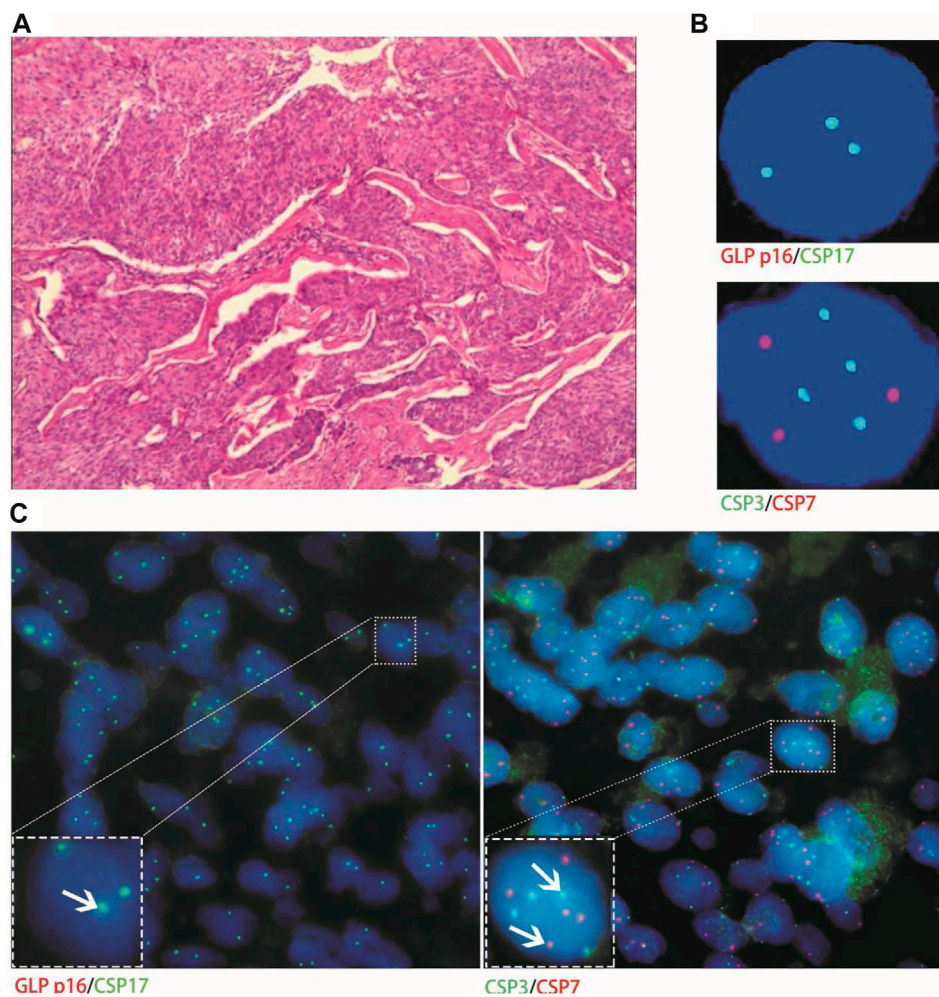
FIGURE 4
Case No.1. Positive validation of histological and urine FISH in urothelial carcinoma. **(A)** is the histopathological image of a patient with urothelial carcinoma. The microscope shows high-grade urothelial carcinoma of the bladder, invading the full thickness of the bladder wall (hematoxylin-eosin staining, ×200 magnification); **(B)** is the urine FISH, showing the amplification of chromosomes 3, 7, 17 and the deletion of the GLP p16 locus; **(C)** is histological FISH (×400 magnification), also showing amplification of chromosomes 3, 7, 17 and deletion of the GLP p16 locus (indicated by arrows). The mean fluorescence signals of chromosome 3, 7, 17 and GLP p16 locus in each cell were 3.59, 3.02, 2.87 and 0.63, respectively. Red represents CSP7 and GLP p16. Green represents CSP3 and CSP17.

According to the pathological results of postoperative specimens and related clinical data, adenocarcinoma was further divided into 7 cases of urachal adenocarcinoma and 2 cases of prostate acinar carcinoma invading the bladder. Squamous cell carcinoma was divided into primary and secondary squamous cell carcinoma. The primary squamous cell carcinoma was divided into 6 cases of renal pelvis squamous cell carcinoma and 3 cases of bladder squamous cell carcinoma. Two secondary renal squamous cell carcinoma cases were derived from oesophageal squamous cell carcinoma and lung squamous cell carcinoma.

The above 27 patients were all subjected to urine FISH assay due to haematuria or suspected UC and other factors, and the collected urine samples all met the testing requirements. Of the 27 patients, 22 were males, and 5 were females, with a median age of 54 (50–65) years. The overall positive rate in urine FISH was 64.00% (16/25) in patients with CNUL. The positive rate of adenocarcinoma was 77.78% (7/9), including 5 cases of urachal carcinoma (71.43%, 5/7) and 2 cases of prostate acinar carcinoma invading the bladder (100%, 2/2). In squamous cell carcinoma, the positive rate was 54.55% (6/11), including 3 cases of primary pure bladder squamous cell carcinoma (100%, 3/3), 1 case of primary renal pelvis squamous cell carcinoma (16.67%, 1/6), and 2 cases of secondary renal squamous cell carcinoma (derived from oesophageal squamous cell carcinoma and lung squamous cell carcinoma).

For advanced or certain rare tumors, such as metastatic tumors (2 cases of metastatic renal squamous carcinoma, 1 case of metastatic non-Hodgkin's lymphoma, and 1 case of urachal carcinoma with distant visceral metastasis), prostate cancer invading the bladder, and small cell carcinoma of the bladder, urine FISH is prone to be positive. Therefore, the more malignant and advanced the tumor is, the more likely it is to result in positive urine FISH (Table 1).
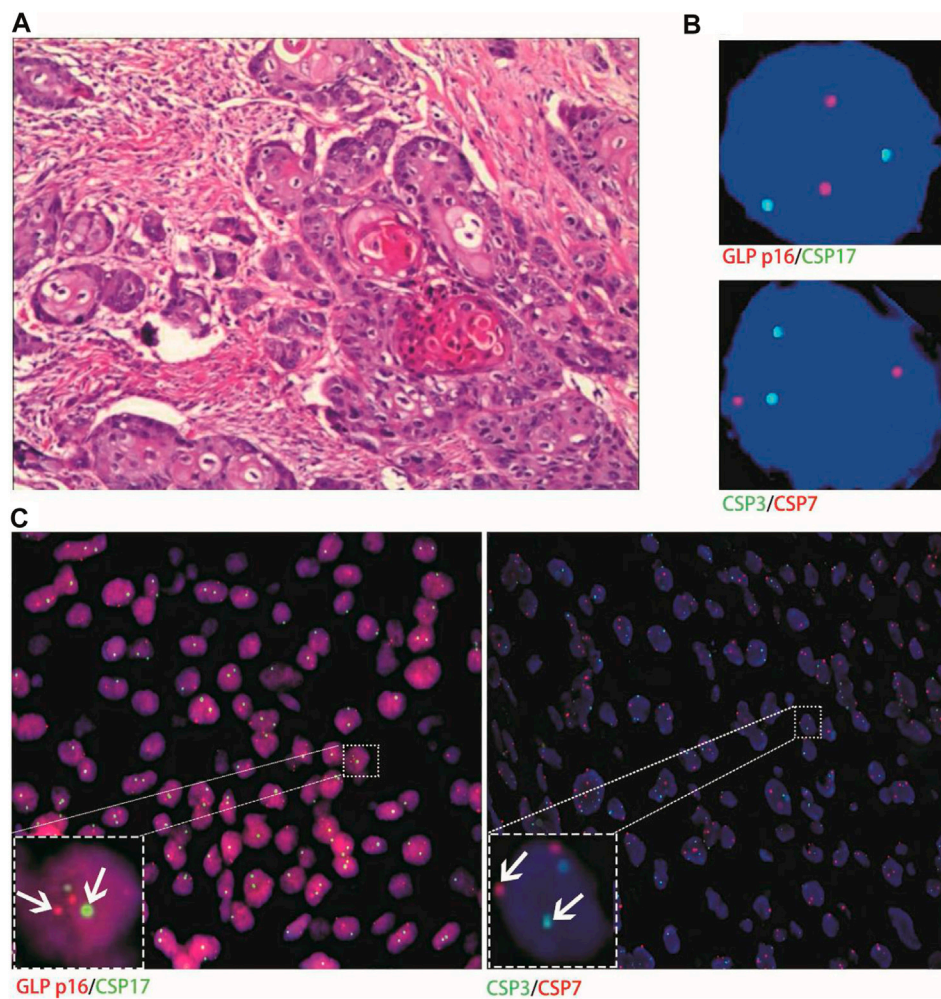
FIGURE 5
Case No.18. Negative validation of histological and urine FISH. **(A)** is the histopathological image of a patient with primary pure renal pelvis squamous cell carcinoma. Microscopically, high-medium-differentiated squamous cell carcinoma invaded perirenal fat, and no atypical hyperplasia, carcinoma *in situ*, or infiltrating carcinoma components of urinary tract epithelium were observed (hematoxylin–eosin staining, ×200 magnification); **(B)** is the urine FISH result of case No. 18, which was negative; **(C)** is histological FISH (×400 magnification), which was also negative (indicated by the arrow). The mean fluorescence signals of chromosome 3, 7, 17 and GLP p16 locus in each cell were 1.93, 2.05, 1.95 and 1.96, respectively. Red represents CSP7 and GLP p16. Green represents CSP3 and CSP17.

3.2 Genetic material changes in CNUL patients with positive urine FISH

Among the 25 patients included in the study, the rates of chromosome 3, 7 and 17 amplifications and GLP p16 gene deletion were 64.00% (16/25), 60.00% (15/25), 52.00% (13/25) and 8.00% (2/25), respectively. In adenocarcinoma, the amplification rates of chromosomes 3, 7, and 17 were 77.78% (7/9), 77.78% (7/9), and 44.44% (4/9), respectively, while the GLP p16 gene deletion rate was only 11.11% (1/9). In squamous cell carcinoma, the amplification rate of chromosomes 3, 7 and 17 was 54.55% (6/11), while the deletion rate of the GLP p16 gene was only 9.09% (1/11). Therefore, the incidence of GLP p16 gene deletion is very low in both adenocarcinoma and squamous cell carcinoma. There was no statistically significant difference in the frequency of chromosome 3, 7, 17 amplification and GLP p16 gene deletion in squamous cell carcinoma compared with adenocarcinoma ($p > 0.05$) (Table 2).

In addition, the proportion of cells with abnormal genetic material in urine samples of patients with bladder squamous cell carcinoma, renal secondary non-Hodgkin's lymphoma and bladder small cell carcinoma were all greater than 65%, indicating that these tumour cells are more likely to shed into the urine (Table 1). During our data collection, we also found that for patients with advanced or distant metastases, such as renal secondary non-Hodgkin's lymphoma, urachal carcinoma with visceral distant, metastatic renal squamous carcinoma (derived from lung squamous carcinoma and esophageal squamous carcinoma), and primary squamous carcinoma of the bladder, their urine exfoliated cells had more frequent chromosomal amplifications, often appearing as 5-ploidy and 6-ploidy (Figure 1).
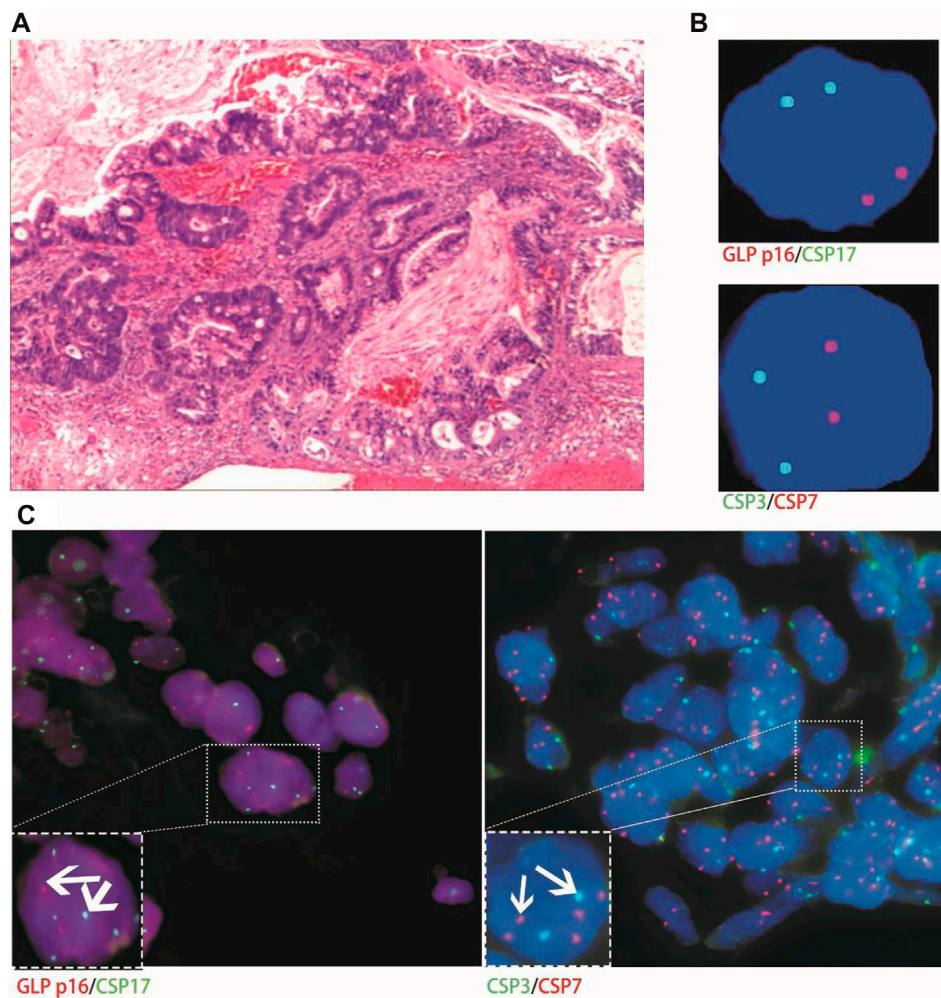
FIGURE 6
Case No.10. The histological FISH (positive) of a patient with stage IIIa urachal carcinoma showed amplification of chromosomes 7 and 17, which was inconsistent with its urine FISH (negative). **(A)** is the histopathological image of a patient with urachal carcinoma, showing intestinal adenocarcinoma under a microscope (hematoxylin-eosin staining, ×200 magnification); **(B)** is negative for urine FISH results; **(C)** is histological FISH (×400 magnification), showing amplification of chromosomes 7 and 17, no amplification of chromosome 3 and deletion of the GLP p16 gene (indicated by the arrow). The mean fluorescence signals of chromosome 3, 7, 17 and GLP p16 locus in each cell were 2.11, 3.88, 3.74 and 2.03, respectively. Red represents CSP7 and GLP p16. Green represents CSP3 and CSP17.

3.3 Mutual validation of histological and urine FISH results

For this study, haematoxylin and eosin staining slides (as well as immunohistochemical slides, if applicable) of paraffin-embedded tissues of 20 patients (Paraffin sections of some patients are obsolete or have too few tissue specimens to perform histological FISH) were applied to the Pathology Department. Representative paraffin blocks containing ≥80% tumour cells were then selected for histological FISH detection, and the hybridized target area was highlighted on each representative slide.

The 20 patients included 5 adenocarcinomas (3 urachal adenocarcinomas, 2 prostate cancer invading the bladder), 10 squamous cell carcinomas (7 renal pelvis squamous cell carcinomas, 3 bladder squamous cell carcinomas), 1 bladder small cell carcinoma, 2 UC (1 bladder UC, 1 renal pelvis UC) and 2 glandular cystitis. Sections were processed in strict accordance with the

instructions of the FISH kit, and *in situ* hybridization was performed using fluorescent dye-labelled GLP p16 gene locus-specific probes and CSP3/CSP7/CSP17 chromosomal centromeric probes. Pathologists with 10 years of work experience read the films.

Histological FISH results showed that the histological results were consistent with urine FISH results in 19 patients. Due to a large number of mutually verified cases, only representative cases are shown here (one case each of adenocarcinoma, squamous cell carcinoma, and UC with positive verification and one case with negative verification of squamous cell carcinoma). Positive verification: The histological FISH of the adenocarcinoma patient in case No. 15 also showed amplification of chromosomes 3, 7 and 17 without the deletion of the GLP p16 gene, which was consistent with his urine FISH (Figure 2). The histological FISH of the squamous cell carcinoma patient in case No. 25 showed amplification of chromosomes 3, 7 and 17 and no GLP p16 gene deletion,

which was consistent with the results of urine FISH (Figure 3). Histological FISH of the patient with UC in case No. 1 showed amplification of chromosomes 3, 7, and 17 and deletion of GLP p16, which was consistent with the urine FISH (Figure 4). Negative verification, such as case No. 10, indicates that histology and cytology FISH is negative (Figure 5).

Only one patient with stage-IIIa urachal carcinoma had histological FISH (positive) findings of chromosome 7 and 17 amplification, which was inconsistent with its urine FISH (negative) results (Figure 6).

4 Discussion

Molecular cytogenetic research in recent years has focused on identifying relevant changes in abnormal cellular DNA in urine specimens, as they often precede the appearance of macroscopic and microscopic lesions, allowing detection of subclinical disease. FISH is a highly sensitive and specific molecular test for detecting urothelial carcinoma (Halling et al., 2000). However, in clinical work, the authors found that FISH also showed positive performance in urine samples of various CNUL, which aroused research interest.

Although numerous studies have evaluated the performance of FISH in typical UC, there is a significant lack of data from studies evaluating FISH in CNUL. This study found that the positive rate of FISH in adenocarcinoma was 77.78% (7/9), and there was no significant difference compared with the sensitivity of 81% in urothelial carcinoma ($p > 0.05$) (Caraway and Katz, 2010). Although the positive rate in squamous carcinoma was 54.55% (6/11), it varied widely, with 100% in both primary squamous carcinoma of the bladder (3 cases) and metastatic renal squamous carcinoma (2 cases), while the positive rate in primary simple renal pelvis squamous carcinoma was only 16.67% (1/6), so it caused us to think about it.

The low positive rate of urine FISH in pure renal pelvis squamous cell carcinoma may be due to insufficient shedding of tumor cells into the urine, or the absence of related genetic material changes such as chromosomes 3, 7, 17, and p16 genes, resulting in negative FISH. This study also confirmed that histological FISH results in patients with simple renal pelvis squamous carcinoma were consistent with urine FISH, thus suggesting that alterations in the heritage material of urine FISH-negative renal pelvic squamous cell carcinoma may not match the probe combination used. Possible reasons for the high rate of urine FISH positivity in bladder squamous cell carcinoma include the following: First, the bladder is an organ for storing urine and controlling urination. Bladder squamous cell carcinoma is a tumour of the lower urinary tract, and tumour cells are easy to shed and collect. Some studies (Gomella et al., 2017; Huang, 2020) have also shown that the sensitivity of urine FISH assay in lower urinary tract tumours is significantly higher than that of upper urinary tract tumours. Second, it is difficult for some postoperative specimens to have only one pathological type, which may or may not be accompanied by a small urothelial carcinoma component. Therefore, we will doubt whether the positive urine FISH assay is caused by UC or bladder squamous cell carcinoma. This doubt can be explained by pathological and histological FISH results. First, the pathological reports of 3 cases of bladder squamous carcinoma included in this study suggested highly differentiated squamous cell carcinoma, and no atypical hyperplasia, carcinoma *in situ* or invasive carcinoma of urinary tract epithelium was

observed. In addition, the histological FISH test results of these 3 cases were consistent with urine FISH assay, suggesting that there were indeed genetic changes related to FISH positivity in bladder squamous cell carcinoma.

In a large data from multiple clinical institutions in China studied by zhou et al. (2019) suggested that UC patients with chromosome 3, 7, and 17 amplification or GLP p16 gene deletion accounted for 71.3% (2941/4125), 72.2% (2978/4125), 67.4% (2780/4125), and 72.9% (3007/4125), respectively. The changes in the genetic material in CNUL and UC patients are indeed different. CNUL patients mainly had amplification of chromosomes 3, 7 and 17 ($p > 0.05$), while the deletion rate of the GLP p16 gene was significantly lower than that of UC (8.0% vs. 72.9%, $p < 0.001$). There are many genetic abnormalities in UC during its occurrence and development. Partial or complete loss of chromosome 9 is one of the most common genetic changes. This abnormality is closely related to the early occurrence of bladder cancer because it contains important tumour suppressor genes related to cell cycle regulation (Halling et al., 2000; Sokolova et al., 2000; Zhang et al., 2004). Mutations in chromosome 9p and fibroblast growth factor receptor 3 in normal urothelium may lead to urothelial hyperplasia or low-grade papillary urothelial carcinoma, (Zhang et al., 2004) which may also explain the difference in genetic material between CNUL and UC.

During the mutual verification of histological and urine FISH, we found only one patient with stage IIIa urachal carcinoma whose histological FISH (positive) was inconsistent with his urine FISH assay (negative). This may be because the tumor cells were not shed in sufficient quantity in the urine, resulting in a negative urine FISH. This study can also rule out operational errors and interpretation errors and verifies that the tumour cells shed in the urine originate from tumour tissue rather than inflammatory proliferative reactions or other lesions.

The following is an analysis of why FISH is positive in urine and tissue specimens of CNUL. The FISH DNA probe used in our hospital is a combination of a centromere probe and a site-specific recognition probe provided by Beijing Jinpujia Medical Technology Co., Ltd., consisting of two combinations, CSP3 (green)/CSP7 (red) and GLP p16 (red)/CSP17 (green). If the tumor cells have chromosome 3, 7, 17 aberrations or (and) GLP p16 locus deletions and the diseased cells can be shed in sufficient quantity into the urine, both urine FISH and histological FISH are likely to be positive. Chromosomal aberrations are a prominent feature of human malignancies. Most solid tumours exhibit complex alterations of genetic material. In this study, positive urine FISH assays were found in patients with metastatic tumours and other rare and highly malignant tumours, which confirmed the findings of Ashley et al. (2006) and Lopez-Beltran et al. (Lopez-Bel et al., 2008) that patients with aggressive and highly malignant rare tumours had many genetic abnormalities. Many studies (Offit, 1992; Atkin et al., 1995; Pycha et al., 1999; Kasahara et al., 2002; Wu et al., 2006; Reid-Nicholson et al., 2009; Schaefer et al., 2010; Collazo-et al., 2016; Haisley et al., 2017; Liu et al., 2017; Reis et al., 2018) have also confirmed that adenocarcinoma (prostate cancer, urachal carcinoma), squamous cell carcinoma (bladder squamous cell carcinoma, oesophageal squamous cell carcinoma, lung squamous cell carcinoma), non-Hodgkin's lymphoma, small cell carcinoma of the bladder, paraganglioma of the bladder, etc., may have chromosome 3, 7, 17 aberrations or (and) deletion of the GLP p16 gene locus. Therefore, FISH may be positive. For FISH-negative patients, it is possible that the genetic material changes in the tumour do not fully match the type of probe used.

This study also has limitations, such as the small number of cases collected and the lack of large multicentre samples to verify the conclusions.

5 Conclusion

We demonstrated for the first time the application value of FISH in CNUL on urine samples. Positive urine FISH tests indicate not only UC, but also CNUL. UroVysion™ FISH possibly has a high positive rate in CNUL. Urine FISH is more likely to be positive for patients with high malignancy or distant metastasis. CNUL and UC may have different genetic material changes. If a sufficient number of tumor cells are shed into the urine, the results of histological and urine FISH tests are consistent. Urologists should combine medical history and imaging information when interpreting FISH results for accurate diagnosis and treatment.

Data availability statement

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

Ethics statement

This study was reviewed and approved by the Medical Ethics Committee of the Tongji Hospital of Huazhong University of Science and Technology (IRB ID: TJ-IRB20210521) and individual consent for this retrospective analysis was waived. All procedures performed in studies involving human participants were in accordance with the ethical standards of the Lifespan institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Author contributions

CK: Data curation; Conceptualization; original draft. XL and JW: Data curation; Supervision. ZH: Project administration; Resources;

Supervision. CY: Conceptualization; Funding acquisition; Writing—review and editing. All authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2023.1250442/full#supplementary-material

References

Ashley, R. A., Inman, B. A., Sebo, T. J., Leibovich, B. C., Blute, M. L., Kwon, E. D., et al. (2006). Urachal carcinoma: clinicopathologic features and long-term outcomes of an aggressive malignancy. *Cancer* 107, 712–720. doi:10.1002/cncr.22060

Atkin, N. B., Baker, M. C., and Wilson, G. D. (1995). Chromosome abnormalities and p53 expression in a small cell carcinoma of the bladder. *Cancer Genet. Cytogenet* 79, 111–114. doi:10.1016/0165-4608(94)00114-q

Caraway, N. P., and Katz, R. L. (2010). A review on the current state of urine cytology emphasizing the role of fluorescence *in situ* hybridization as an adjunct to diagnosis. *Cancer Cytopathol.* 118, 175–183. doi:10.1002/cncy.20080

Chen, W. (2015). Cancer statistics: updated cancer burden in China. *Chin. J. Cancer Res.* 27, 1. doi:10.3978/j.issn.1000-9604.2015.02.07

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

Collazo-Lorduy, A., Castillo-Martin, M., Wang, L., Patel, V., Iyer, G., Jordan, E., et al. (2016). Urachal carcinoma shares genomic alterations with colorectal carcinoma and

may respond to epidermal growth factor inhibition. *Eur. Urol.* 70, 771–775. doi:10.1016/j.eururo.2016.04.037

Gomella, L. G., Mann, M. J., Cleary, R. C., Hubosky, S. G., Bagley, D. H., Thumar, A. B., et al. (2017). Fluorescence *in situ* hybridization (FISH) in the diagnosis of bladder and upper tract urothelial carcinoma: the largest single-institution experience to date. *Can. J. Urol.* 24, 8620–8626.

Haisley, K. R., Dolan, J. P., Olson, S. B., Toledo-Valdovinos, S. A., Hart, K. D., Bakis, G., et al. (2017). Sponge sampling with fluorescent *in situ* hybridization as a screening tool for the early detection of esophageal cancer. *J. Gastrointest. Surg.* 21, 215–221. doi:10.1007/s11605-016-3239-3

Halling, K. C., King, W., Sokolova, I. A., Meyer, R. G., Burkhardt, H. M., Halling, A. C., et al. (2000). A comparison of cytology and fluorescence *in situ* hybridization for the detection of urothelial carcinoma. *J. Urol.* 164, 1768–1775. doi:10.1016/s0022-5347(05)67104-2

Hu, Z., Ke, C., Liu, Z., Zeng, X., Li, S., Xu, H., et al. (2020). Evaluation of UroVysion for urachal carcinoma detection. *Front. Med. (Lausanne).* 7, 437. doi:10.3389/fmed.2020.00437

Hu, Z., Ke, C., Shen, Y., Zeng, X., and Yang, C. (2021). Renal metastases from esophageal cancer and retroperitoneal lymphoma detected via chromosome duplications identified by fluorescence *in situ* hybridization in urine exfoliated cells: first 2 case reports. *Med. Baltim.* 100, e24010. doi:10.1097/MD.0000000000024010

Huang, J. (2020). *Chinese guidelines for the diagnosis and treatment of Urology and andrology diseases*. Beijing: Science Press, 27–84.

Jemal, A., Siegel, R., Xu, J., and Ward, E. (2010). Cancer statistics. *CA Cancer J. Clin.* 60, 277–300. doi:10.3322/caac.20073

Kasahara, K., Taguchi, T., Yamasaki, I., Kamada, M., Yuri, K., and Shuin, T. (2002). Detection of genetic alterations in advanced prostate cancer by comparative genomic hybridization. *Cancer Genet. Cytogenet* 137, 59–63. doi:10.1016/s0165-4608(02)00552-6

Kipp, B. R., Tyner, H. L., Campion, M. B., Voss, J. S., Karnes, R. J., Sebo, T. J., et al. (2008). Chromosomal alterations detected by fluorescence *in situ* hybridization in urothelial carcinoma and rarer histologic variants of bladder cancer. *Am. J. Clin. Pathol.* 130, 552–559. doi:10.1309/DFJUHY3WPC9GUU2W

Levsky, J. M., and Singer, R. H. (2003). Fluorescence *in situ* hybridization: past, present and future. *J. Cell Sci.* 116, 2833–2838. doi:10.1242/jcs.00633

Liang, G., Fu, W., and Wang, K. (2019). Analysis of t-test misuses and SPSS operations in medical research papers. *Burns Trauma* 7, 31. doi:10.1186/s41038-019-0170-3

Liu, W., Lyu, R., Huang, W. Y., Li, C. W., Liu, H., Li, J., et al. (2017). Characteristics and prognostic significance of cytogenetic abnormalities in diffuse large B-cell lymphoma patients with bone marrow involvement. *Zhongguo Shi Yan Xue Ye Xue Za Zhi* 25, 761–765. doi:10.7534/j.issn.1009-2137.2017.03.022

Lopez-Beltran, A., Requena, M. J., Cheng, L., and Montironi, R. (2008). Pathological variants of invasive bladder cancer according to their suggested clinical significance. *Bju Int.* 101, 275–281. doi:10.1111/j.1464-410X.2007.07271.x

National Library of Medicine (2020). *Beijing jin Pu jia medicial Technology Co. L.*

Offit, K. (1992). Chromosome analysis in the management of patients with non-Hodgkin's lymphoma. *Leuk. Lymphoma* 7, 275–282. doi:10.3109/10428199209049778

Pycha, A., Mian, C., Posch, B., Haitel, A., Mokhtar, A. A., El-Baz, M., et al. (1999). Numerical chromosomal aberrations in muscle invasive squamous cell and transitional cell cancer of the urinary bladder: an alternative to classic prognostic indicators? *Urology* 53, 1005–1010. doi:10.1016/s0090-4295(98)00626-8

Reid-Nicholson, M. D., Ramalingam, P., Adeagbo, B., Cheng, N., Peiper, S. C., and Terris, M. K. (2009). The use of Urovysion fluorescence *in situ* hybridization in the diagnosis and surveillance of non-urothelial carcinoma of the bladder. *Mod. Pathol.* 22, 119–127. doi:10.1038/modpathol.2008.179

Reis, H., van der Vos, K. E., Niedworok, C., Herold, T., Módos, O., Szendrői, A., et al. (2018). Pathogenic and targetable genetic alterations in 70 urachal adenocarcinomas. *Int. J. Cancer* 143, 1764–1773. doi:10.1002/ijc.31547

Schaefer, I. M., Gunawan, B., Fuzesi, L., Blech, M., Frasunek, J., and Loertzer, H. (2010). Chromosomal imbalances in urinary bladder paraganglioma. *Cancer Genet. Cytogenet* 203, 341–344. doi:10.1016/j.cancergencyto.2010.07.131

Siegel, R. L., Miller, K. D., and Jemal, A. (2017). Cancer statistics. *CA Cancer J. Clin.* 67, 7–30. doi:10.3322/caac.21387

Sokolova, I. A., Halling, K. C., Jenkins, R. B., Burkhardt, H. M., Meyer, R. G., Seelig, S. A., et al. (2000). The development of a multitarget, multicolor fluorescence *in situ* hybridization assay for the detection of urothelial carcinoma in urine. *J. Mol. Diagn* 2, 116–123. doi:10.1016/S1525-1578(10)60625-3

Wiegant, J., Ried, T., Nederlof, P. M., van der Ploeg, M., Tanke, H. J., and Raap, A. K. (1991). *In situ* hybridization with fluoresceinated DNA. *Nucleic Acids Res.* 19, 3237–3241. doi:10.1093/nar/19.12.3237

Wu, Y. P., Yang, Y. L., Yang, G. Z., Wang, X. Y., Luo, M. L., Zhang, Y., et al. (2006). Identification of chromosome aberrations in esophageal cancer cell line KYSE180 by multicolor fluorescence *in situ* hybridization. *Cancer Genet. Cytogenet* 170, 102–107. doi:10.1016/j.cancergencyto.2006.05.006

Yang, C., Liu, Z., Lan, R., Wang, Z., Hu, Z., Chen, Z., et al. (2016). Paraganglioma of the urinary bladder with chromosome duplications detected by fluorescence *in situ* hybridization in urine exfoliated cells: A case report. *Oncol. Lett.* 11, 795–797. doi:10.3892/ol.2015.3941

Zhang, J., Zheng, S., Gao, Y., Rotolo, J. A., Xiao, Z., Li, C., et al. (2004). A partial allelotyping of urothelial carcinoma of bladder in the Chinese. *Carcinogenesis* 25, 343–347. doi:10.1093/carcin/bgh015

Zhou, L., Yang, K., Li, X., Ding, Y., Mu, D., Li, H., et al. (2019). Application of fluorescence *in situ* hybridization in the detection of bladder transitional-cell carcinoma: A multi-center clinical study based on Chinese population. *Asian J. Urol.* 6, 114–121. doi:10.1016/j.ajur.2018.06.001

Eukaryotic translation initiation factor 4A1 in the pathogenesis and treatment of cancers

Jinghong Huang[1†], Lei Zhang[2†], Rui Yang[1†], Lixia Yao[1], Jinming Gou[3], Dongdong Cao[1], Zeming Pan[1], Dongmei Li[1]*, Yuanming Pan[4]* and Wei Zhang[5]*

[1]Key Laboratory of Xinjiang Endemic and Ethnic Diseases, School of Medicine, Shihezi University, Shihezi, Xinjiang, China, [2]Clinical Laboratory, First Affiliated Hospital of Shihezi University, Shihezi, Xinjiang, China, [3]Troops of the People's Liberation Army, Urumqi, Xinjiang, China, [4]Cancer Research Center, Beijing Chest Hospital, Beijing Tuberculosis and Thoracic Tumor Research Institute, Capital Medical University, Beijing, China, [5]Shihezi People's Hospital, Shihezi, Xinjiang, China

Abnormal translate regulation is an important phenomenon in cancer initiation and progression. Eukaryotic translation initiation factor 4A1 (eIF4A1) protein is an ATP-dependent Ribonucleic Acid (RNA) helicase, which is essential for translation and has bidirectional RNA unwinders function. In this review, we discuss the levels of expression, regulatory mechanisms and protein functions of eIF4A1 in different human tumors. eIF4A1 is often involved as a target of microRNAs or long non-coding RNAs during the epithelial-mesenchymal transition, associating with the proliferation and metastasis of tumor cells. eIF4A1 protein exhibits the promising biomarker for rapid diagnosis of pre-cancer lesions, histological phenotypes, clinical staging diagnosis and outcome prediction, which provides a novel strategy for precise medical care and target therapy for patients with tumors at the same time, relevant small molecule inhibitors have also been applied in clinical practice, providing reliable theoretical support and clinical basis for the development of this gene target.

1 Introduction

According to the recent statistical analysis of global cancer types and cancer incidence in various countries, cancer is one of the most important diseases affecting human beings, causing serious physical damage and psychological stress to patients (BRAY et al., 2018). Cancer is a disease that occurs in the context of a complex interplay of multiple factors. Cancer occurs under the complex regulation of multiple factors, and the final manifestation *in vivo* is the alteration of the expression of the relevant oncoproteins, which is strictly controlled during the translation of messenger RNA (mRNA).

Abbreviations: EIF4A1, Eukaryotic Translation Initiation Factor 4A1; GEO, Gene Expression Omnibus; GC, Gastric Cancer; G3-MB, G3-type medulloblastoma; HSF1, Heat Shock Factor 1; LncRNA, Long noncoding RNA; mRNA, Messenger RNA; MiRNAs, MicroRNAs; PDCD4, Programmed Cell Death 4; RNA, Ribonucleic Acid; USP15, Ubiquitin-Specific Peptidase 15.

mRNA translation dysregulation is one of the most important factors in the predisposition of cancer (WALDRON et al., 2019). mRNA translation is a complex process that includes an initiation step, an extension step, and a termination step (DEVER and GREEN, 2012). The rate-limiting step of most protein synthesis occurs at the initiation stage, and its aberrant regulation greatly contributes to translational reprogramming, which characterizes cancer cells, such as abnormal proliferation and chemoresistance (JACKSON et al., 2010). The translation initiation program in cancer cells maintains their malignant capacity in tumor progression and metastasis. Downregulation of translational capacity, particularly through inhibition of translation initiation will result in diminished tumor migration and invasiveness (MANIER et al., 2017).

Eukaryotic Translation Initiation Factor (eIF) have multiple functions in these processes: They can act as RNA chaperones, ATP-dependent RNA helicases, as RNPases by mediating RNA-protein association and dissociation or as co-activators and co-repressors of translation (SCHUTZ et al., 2010). There is substantial evidence that eIF is strongly associated with poor prognosis and resistance to chemotherapy and targeted therapy in many cancers, including leukemia (WOLFE et al., 2014).

Members of the eukaryotic translation initiation factor 4 A family are required for the process of translation initiation, and they are also prototypical of members of the dead-box (DEAD-box) family. The DEAD-box family acts as a helicase that unwinds RNAs, separating base paired RNA strands bound by hydrogen bonds from one another and removing secondary structures in RNA, so that they can be ligated by RNA enzymes, a change that is dependent on the supply of energy from ATP. The eIF4A structural domain was the first DEAD-box protein structure to be identified. Box protein structure with RecA-like folding (binding sites for nucleotides) and interactions between conserved motifs within the structural domains (LINDER, 2006). The eIF4A family includes the following isoforms: eIF4A1, eIF4A2, and eIF4A3 (XUE et al., 2021). Generally, eIF4A1 and eIF4A2 are mainly located in the cytoplasm and are more abundant in eIF4A1 relative to eIF4A2, while eIF4A3 is mainly located in the nucleus (LU et al., 2014). eIF4A1 and eIF4A2 are mainly involved in the initiation of mRNA translation, whereas eIF4A3 proteins mainly play a role in RNA metabolism, including mRNA localization, export, and coupling of mRNA's splicing to translation (MAZLOOMIAN et al., 2019).

Recent studies have shown that eIF4A1 as the ATP-dependent RNA helicase has low activity. Its RNA unwinding activity is coupled to a conformational cycle in which eIF4A alternates between an open conformation with a wide cleft between its two RecA domains (SUN et al., 2014), and a closed conformation in the presence of ATP and RNA, in which the two RecA domains interact with each other and with bound ATP and RNA. Formation of the closed state is linked to duplex destabilization (ANDREOU and KLOSTERMEIER, 2014). Upon ATP hydrolysis and phosphate release, eIF4A assumes an open conformation and disengages from the RNA. Binding of free eIF4A to ATP or RNA is not ordered but is coupled (PELLETIER and SONENBERG, 2019). The preferred substrate for eIF4A1 is a single-stranded poly-purine RNA. RNA fragments as short as 4 nucleotides can stimulate the ATPase activity of free eIF4A, but fragments of 15–20 nucleotides are optimal (PECK and HERSCHLAG, 1999) eIF4A1 binds to

PKP1 in the eIF4F cap-binding complex and is stimulated by PKP1 to promote ATPase activity, thereby increasing the translation rate of mRNAs (WOLF et al., 2010). eIF4A1 can also be deubiquitinated by interacting with ubiquitin-specific peptidase 15 (USP15), which promotes the self-stabilization of eIF4A1, thereby increasing the translation efficiency of mRNAs (ZHAO et al., 2020). However, MA et al. (2019) found that inhibition of PHGDH reduces its interaction with eIF4A1, decreases eIF4A1 activity, and blocks the formation of the translation initiation complex eIF4F, thereby preventing the translation of the entire mRNA process.

The ATP-dependent RNA helicases eIF4A1 plays a crucial role in various cancers in humans (WOLF et al., 2010) The expression level of eIF4A1 varies in different types of malignant tumors (LIN et al., 2018). With the understanding and deepening of the regulatory mechanism of eIF4A1, we found that eIF4A1 can be one of the potential points of action and biomarkers for cancer diagnosis and therapy (HAN et al., 2022). In this paper, we will discuss the regulatory mechanism and biological function of eIF4A1 during mRNA translation. In addition, this review also discusses the role of eIF4A1 in the process of tumor proliferation and metastasis, and existing inhibitors, suggests the possibility of using them as a potential point of action and biomarker for cancer diagnosis, treatment and prognosis.

2 Regulatory mechanisms of eIF4A1

eIF4A1 is an important component of eIF4F, and the eIF4F cap-binding complex consists of the eIF4A1 translation initiation factor, eIF4G scaffolding protein, and eIF4E m7 G-binding protein (JACKSON et al., 2010). Many eIFs are assembled with small (40S) ribosomal subunits to form a 43S pre initiation complex, which scans the first AUG codon on mRNA and then binds to the large (60S) subunit to form an active 80S ribosome on this initiation codon (IWASAKI et al., 2019). Through the eIF4F complex, eIF4A1 is involved in two main translation steps: loading mRNA onto the 43S pre-initiation-complex and translocating it along the 5'UTR to the translation initiation point (SCHMIDT et al., 2023). The loading function requires only eIF4A1's ATPase activity (WANG et al., 2022). eIF4A1 relies on ATP binding to single stranded RNA (AMBARU et al., 2022), unlocking the double stranded region in RNA (WILCZYNSKA et al., 2019). In addition, eIF4A1 itself is a weak helicase, and the dissociation efficiency achieved by eIF4A1 is significantly improved by forming a complex of eIF4A1 and cofactor proteins eIF4G, eIF4B, and eIF4H that synergistically regulate the conformational cycle of eIF4A1 (GARCíA-GARCíA et al., 2015). The translation initiation factors eIF4B and eIF4G jointly stimulate the weak intrinsic RNA-dependent ATPase and ATP-dependent RNA helicase activities of yeast and human eIF4A (OZES et al., 2011) through modulation of the eIF4A conformational cycle. In the presence of eIF4G, eIF4A alternates between a half open conformation, stabilized by binding of eIF4G to both RecA domains of eIF4A, and the closed state (HARMS et al., 2014). eIF4B binds to eIF4A through its 7-repeats domain. Binding of eIF4B to eIF4A further accelerates closing when eIF4G is present,

and thus causes an additional shift of the conformational equilibrium of eIF4A toward the closed state (JOYCE et al., 2017). So its unwinds function is largely dependent on the stimulation of its binding partner, whereas the binding of single-stranded RNA can largely stimulate the activity of eIF4A (SHEN and PELLETIER, 2020). eIF4B and eIF4H also stabilize partially unwound substrates and/or prevent mRNA reannealing, activities that further facilitate RNA restructuring during initiation.

As an important translation initiation factor, eIF4A1 loads all mRNAs and binds to ribosomes. However, recent research results indicate that RNA is also involved in influencing the function of eIF4A1. For example, the catalytic activity of the complex eIF4A-eIF4B-eIF4G can be increased by the length of single stranded RNA (ANDREOU et al., 2019), However, the role of RNA in regulating the function of eIF4A1 has not been thoroughly studied.

Previous studies have demonstrated that translational dysregulation is an important step in tumorigenesis and progression that directly controls selective translation and protein synthesis of cancer genes (WALDRON et al., 2019). The eIF4F translation initiation complex, under the regulation of the PI3K/Akt/mTOR signaling pathway, the mitogen-activated protein kinase signaling pathway, and the cysteine-dependent apoptosis pathway, serves as a key node for translation initiation and controls the translation initiation phase of mRNAs of many oncogenes (LIN et al., 2008; BLAGDEN and WILLIS, 2011). As an important component of eIF4F, eIF4A1 plays an important role in the process of gastric carcinogenesis and development and epithelial mesenchymal transition in gastric cancer, and recent studies have shown that the expression of eIF4A1 in cancers such as gastric cancer, colorectal cancer, cervical cancer, breast cancer, and melanoma exhibits abnormalities (LIANG et al., 2014; MODELSKA et al., 2015; JOYCE et al., 2017; GAO et al., 2020; CHEN et al., 2021a; SOYLEMEZ et al., 2021). Mutations in eIF4A1 lead to translational repression (WILCZYNSKA et al., 2019). Free eIF4A1 is regulated by Programmed Cell Death 4 (PDCD4), which binds eIF4A, blocks formation of the closed conformation. And the level of eIF4A1 itself is regulated by mTOR and the carcinogen miR-21 (ASANGANI et al., 2008).

3 Expression of eIF4A1 in different tumors

eIF4A1 is dysregulated and aberrantly expressed in many different tumor tissues (LIN et al., 2018), although the exact role of eIF4A1 in tumorigenesis and development is unclear, it may be associated with abnormal RNA unwinds function and lead to aberrant expression of proteins formed by aberrant RNA translation (LOH et al., 2009).

3.1 Expression of eIF4A1 in gastric cancer

Gao et al. (2020) used Gene Expression Omnibus (GEO) to detect the mRNA level of EIF4A1 expression in gastric cancer (GC) tissues, and the data showed that eIF4A1 expression was significantly upregulated in gastric cancer compared with that in adjacent normal tissues. Furthermore, IHC results of GC patients showed that eIF4A1 protein levels were generally elevated in GC tissues compared with normal ones (56.5%, 108/191) (GAO et al., 2020). Wei et al. also found that the protein levels of eIF4A1 expression were significantly upregulated in 74 clinical cancer samples and its control samples (58.1%, 58/74) (WEI et al., 2019). Furthermore, overexpression of eIF4A1 was significantly correlated with advanced tumor metastasis, epithelial mesenchymal transition, poor tumor differentiation and poor prognosis in GC patients (GAO et al., 2020).

3.2 Expression of eIF4A1 in colorectal cancer

Li et al. found that immunohistochemical staining in colorectal cancer patients showed that eIF4A1 was highly expressed in 86% (44/51) of primary colorectal cancer tissues (LI et al., 2017). However, Zafer et al. found that eIF4A1 was highly expressed in stage II colorectal cancer tissues and lowly expressed in stage I, III, and IV colorectal cancer tissues, and that eIF4A1 was highly expressed in the peripheral blood of patients with stage I, II, and III colorectal cancer, but lowly expressed in patients with stage IV colorectal cancer (SOYLEMEZ et al., 2021). Yang et al. reported that eIF4A1 is recruited by Long noncoding RNA (LncRNA) MAPKAPK5-AS1 to promote translation of MAPK-activated protein kinase 5 in colorectal cancer cells (YANG et al., 2020).

3.3 Expression of eIF4A1 in cervical cancer

Liang et al. found that overexpression of eIF4A1 was detected in 83.9% of cervical cancer tissues and that overexpression of eIF4A1 was associated with advanced tumor proliferation, lymph node metastasis, squamous cell in patients with cervical cancer, before and after brachytherapy by using immunohistochemistry in 35 cases of normal cervical tissues, 87 cases of cervical cancer tissues without surgical treatment, and 50 pairs of cervical cancer tissues histology, deep mesenchymal invasion, and poor prognosis were significantly correlated (LIANG et al., 2014). They also found that silencing eIF4A1 can increase the radiosensitivity of cervical cancer, leading to delayed repair of radiation-induced DNA double strand breaks (LIANG et al., 2014)

3.4 Expression of eIF4A1 in breast cancer

The study by Modelska et al. found that upregulation of eIF4A1 expression in estrogen receptor-negative breast cancers were associated with higher histological grades by immunohistochemical testing of tissue microarrays from approximately 4,000 patients and by post-statistical analysis of the patients' cancer grades (89.5%), and that eIF4A1 is involved in dysregulation of the mRNA translation process through pro-carcinogenic signaling, which contributes to breast cancer in the generation of malignant phenotype, suggesting that eIF4A1 can be used as a biomarker to predict the prognosis of breast cancer patients (MODELSKA et al., 2015).

3.5 Expression of eIF4A1 in melanoma

Eberle et al. found that eIF4A1 was consistently overexpressed in melanoma cells and discovered that eIF4A1 aided melanocytoma proliferation, whereas inhibition of endogenous eIF4A1 expression suppressed the value-added migratory and invasive abilities of melanocytomas (64.2%) (EBERLE et al., 2002).

3.6 Expression of eIF4A1 in other tumors

Zhao et al. reported that low levels of PDED4 and high levels of eIF4A1 predicted poorer differentiation and higher recurrence rates after surgery for oral squamous carcinomas, suggesting that these proteins are significant independent risk factors for such cancers(47.8%, 33/69) (ZHAO et al., 2019). Wang et al. found that for prostate cancer cells, elevated mRNA levels of eIF4A1 correlated with DNA hypomethylation levels on CpG-rich eIF4A1 islands, eIF4A1 translation products were epigenetically regulated through DNA methylation, and eIF4A1 exerted its oncogenic effects through BRD2 signaling (WANG et al., 2022). Similarly, Zhao et al. found that in Myc-amplified G3-type medulloblastoma (G3-MB), eIF4A1 was highly expressed and positively correlated with Myc expression, and that inhibiting eIF4A1 expression could effectively inhibit Myc expression at the translational level, and through this process, promote apoptosis of G3-MB cells and inhibit G3-MG cell proliferation to block the growth of tumor cells (ZHAO et al., 2020). Recently, in a study on brain gliomas (KRASSNIG et al., 2021) and endometrial cancer (LOMNYTSKA et al., 2012) and human cytomegalovirus (QI et al., 2013), there were also showing that eIF4A1 has significantly higher expression levels in different tumors and functions as a tumor promoter.

4 Biological functions of eIF4A1 protein in different tumors

About the biological functions of eIF4A1 protein, Wolfe et al. reported that an eIF4A-dependent mechanism of translational control that is encoded in the 5′-UTR of susceptible transcripts, including many oncogenes and transcriptional regulators (for example, Myc, Myb, Notch, Cdk6, Bcl2, and others). Thereby accelerating the progression of Notch-driven T-cell acute lymphoblastic leukemia (WOLFE et al., 2014). Similarly, Cailin et al. found that eIF4A1 is generated by acting on the coding region and 3′-UTR of mRNA to produce an effect on the translational phase of melanoma cells (JOYCE et al., 2017). Li et al. reported that eIF4A1 is a direct target of miR-133a, and miR-133a inhibit colon cancer cells by inhibiting eIF4A1 expression (LI et al., 2017). Similar to the role of miR-133a, miR-1284 can directly inhibit the expression of EMT related genes c-Myc and MMP12 by inhibiting eIF4A1 in gastric cancer (WEI et al., 2019).

Ritesh et al. found that eIF4A1 stimulation by Raf/MAPK/extracellular signaling pathway-regulated kinase signaling significantly promoted the expression of genes associated with the cell cycle and accelerated tumor size in cutaneous squamous cell carcinoma. Combined inhibition of the Raf/MAPK/extracellular

signaling-regulated kinase axis and eIF4A1 decreased the 5′-capsule-dependent translational process and attenuated the growth, metastasis, and invasiveness of cutaneous squamous cell carcinoma cells (ZHAO et al., 2019; SRIVASTAVA et al., 2021). Nishida et al. found that eIF4A1 promoted the proliferation of Heat Shock Factor 1 (HSF1) in cells, thereby promoting the proliferation of progenitor cells and leukemia-initiating cells and accelerating leukemogenesis (NISHIDA et al., 2021).

Xu et al. found that eIF4A1 was also involved in the critical steps of platelet healing. SiRNA-USP15 was shown to be involved in platelet healing through promoting eIF4A1 de-ubiquitination enhanced the functional properties of platelet cells to promote wound healing (XU et al., 2021). Zhao et al. reported that in pancreatic cancer cells, eIF4A1 elevated the expression of E-cadherin and N-cadherin through the c-myc/miR-9 axis. eIF4A1 and c-myc promoted epithelial mesenchymal transformation and metastatic ability of pancreatic cancer cells, while eIF4A1 alone upregulation reduced the inhibitory effect of c-myc downregulation on epithelial mesenchymal transformation and metastasis. The eIF4A1 inhibitor Rocaglamide (RocA) and the c-Myc inhibitor Mycro3, alone or in combination, significantly reduced the expression levels of markers of epithelial mesenchymal transition in pancreatic cancer cells (ZHAO et al., 2021).

Oblinger et al. reported that inhibition of eIF4A1 with the eIF4A1 inhibitor, Cevistro, consistently reduced the expression levels of several cell cycle proteins, Aurora A kinase, and the mitogen-activating enzymes, AKT and ERKs in nerve sheath tumors (OBLINGER et al., 2016). Cevastro treatment significantly inhibited tumor proliferation in nerve sheath tumors. Joyce et al. found that silencing of eIF4A1 in WM858 cells significantly reduced melanoma proliferation and invasion (JOYCE et al., 2017). (Figure 1)

5 eIF4A1 inhibitors as the promising drugs in cancer treatment

Based on the above results, it can be concluded that the high expression level of eIF4A1 significantly stimulates the malignant phenotype (proliferation, invasion, migration, and epithelial mesenchymal transition) of cancer cells. Therefore, the upregulation of eIF4A1 seems to have an impact on transformed cells through specific information, making eIF4A1 an attractive target for therapeutic interventions. Several natural compounds have been described as inhibiting cap-dependent translation by specifically inhibiting eIF4A1 activity, including Hippuristanol (CHIO et al., 2016), Pateamine A (NAINENI et al., 2021) and Rocaglates (CHEN et al., 2021b). (Table 1)

Hippuristanol is a natural product originally isolated from the Isis hippopotamus (CHAO et al., 2005), A type of bamboo coral that cancels the RNA binding activity of eIF4A1 by locking the helicase in a closed conformation (SUN et al., 2014). Hippuristanol has been shown to interact with amino acids present within or adjacent to the motif in the CTD of eIF4A1, This binding site is not conserved in other DEAD box RNA helicases (HOWARD et al., 2020).

Recently, structural analysis of the Rocaglates: eIF4A1: Polypurine RNA complex has shown that Rocaglates, as an

FIGURE 1
The potential mechanism diagram of eIF4A1. eIF4A1 stimulation by Raf/MAPK/extracellular signaling pathway-regulated kinase signaling significantly promoted cellular behavior such as proliferation, migration, invasion, and EMT of cancer cells through Snail, c-myc, MMPs, AKT and ERK proteins. MiRNAs such as miR-1284 and miR-133a can inhibit the expression of eIF4A1 protein, and these behaviors occur.

eIF4A1 inhibitor, it interacts indispensable with eIF4A1 and two adjacent RNA purine bases (CHU et al., 2019). Rocaglates are a class of artemisinins extracted from plants containing cyclic penta [*b*] benzofuran structures, and are one of the most effective and specific eIF4A inhibitors known (NAINENI et al., 2020). Over 200 natural and synthetic Rocaglates have been described since Rocaglamide A was first isolated from Asian mahogany, the only genus known to produce Rocaglates (GREGER, 2022). Structural analysis of the eIF4A: Rocaglates: Polypurine RNA complex has revealed that Rocaglates specifically bind in a cavity formed between human eIF4A1 (at Phe163, Gln195, Asp198, and Ile199) and adjacent purines (A and G) (IWASAKI et al., 2019).

Similar in function to the complex formed by Rocaglate analogues, the complex formed by Patemine A with eIF4A1, mRNA, and AMPPNP also has the function of inhibiting eIF4A1, although the structures of these two complexes are very diff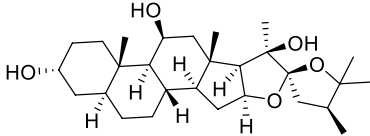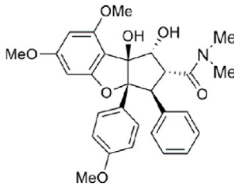erent (NAINENI et al., 2021). In addition to overlapping their binding positions on the eIF4A1 protein, Rocaglates and Patemine A also interact directly with RNA. RNA bases are aromatic rings with flat hydrophobic surfaces. Adjacent bases stack with each other in the double helix of RNA (and DNA), burying hydrophobic bases and providing favorable interactions between π-electrons in their aromatic systems. Stacking also contributes greatly to the conformation of single stranded nucleic acids, and RNA binding proteins can interact with RNA in a similar manner through the π-π stacking of amino

acid side chains and RNA bases. However, at the curvature of the spine where the dead box protein binds, the base accumulation is disrupted. These proteins completely rely on contact with the sugar phosphate backbone to bind RNA, without interacting with bases to compensate for lost stacking interactions. Bending the bases on both sides forms a hydrophobic pocket that can be occupied by PatemineA or Rocaglates, where the conjugated π-system in these drugs interacts with RNA through stacking (INGOLIA, 2021).

Although the interactions between Patemine A and Rocaglates with RNA are similar, their internal structures are not entirely the same. PatemineA has an extended linear π-conjugated system that can interact well with all four bases. In contrast, the aromatic ring in Rocaglate can only interact well with purines. These molecular differences translate into differences in sequence selectivity. Patemine A stabilizes protein drug RNA complexes that include both pyrimidine and mixed purine pyrimidine RNA sequences, although weaker than purine RNA, while Rocaglates only forms such complexes with purine RNA (IWASAKI et al., 2016). This selectivity extends to the cellular effects of these drugs. PatemineA inhibits the translation of reporters resistant to Rocaglates, which lack a purine sequence in their 50 leading sequences. The difference between PatemineA derivatives and Rocaglates highlights the potential of stacking interactions to provide sequence specificity in nucleic acid binding; This specificity is more related to alkali specific hydrogen bonding (NAINENI et al., 2021). The natural

TABLE 1 Structure and action of different eIF4A1 inhibitors.

Name	Structure	Combination method and action	Ref.
Hippuristanol		Interact with amino acids present within or adjacent to the motif in the CTD of eIF4A1 and cancels the RNA binding activity of eIF4A1 by locking the helicase in a closed conformation.	(SUN et al., 2014; HOWARD et al., 2020)
Rocaglamide		Specifically bind in a cavity formed between human eIF4A1 (at Phe163, Gln195, Asp198, and Ile199) and adjacent purines (A and G). Whenever ATP-bound eIF4A binds to RNA and kinks it to induce unwinding, a bimolecular cavity is formed between the eIF4A NTD and the bent single-stranded RNA. When human eIF4A1 binds to consecutive purine residues, the resultant bimolecular cavity can accommodate Rocaglate or Pateamine A. The complex leading to translation repression.	INGOLIA (2021)
Pateamine A			
Elatol		N/A	PETERS et al. (2018)
eFT226		It forming a ternary complex with eIF4A and AGAGAG purine RNA oligonucleotides, preventing eIF4A1 from releasing from the purine RNA motif.	ERNST et al. (2020)

compound Elatol from the ocean also has similar effects to the above two compounds (PETERS et al., 2018).

In addition to natural compounds, some small molecule eIF4A1 inhibitors have also entered our field of vision. EFT226 (Zotatifin) is the first eIF4A inhibitor to enter human clinical trials. This drug was first used in clinical trials of patients with ER$^+$ breast cancer, with an expansion cohort for patients with Cyclin D1 alterations. (ERNST et al., 2020). It promotes the binding of specific mRNA sequences with recognition motifs in eIF4A and 5′-UTRs, and interferes with the assembly of eIF4F complexes downstream of mTOR. Its sensitivity is related to mTOR mediated eIF4A activation (THOMPSON et al., 2021). EFT226 inhibits translation initiation by forming a ternary complex with eIF4A and AGAGAG purine RNA oligonucleotides, preventing eIF4A1 from releasing from the purine RNA motif (ERNST et al., 2020). EFT226 treatment downregulates the protein expression of key translation factors Myc and Bcl6, leading to selective gene expression reprogramming, inhibiting cell proliferation, inducing cell death, and thus producing therapeutic effects on various cancer models (THOMPSON et al., 2021). Also, eIF4A inhibitors repress the protein expression of Cyclin proteins including Cyclin D1 and its binding partners CDK4/6. Similarly,

Kong et al. showed that Rocaglates can suppress Cyclin-induced feedback to CDK4/6 inhibitors used in lung cancer (KONG et al., 2019).

Furthermore, some effective ingredients in traditional Chinese medicine have also been found to have inhibitory effects on eIF4A1. Berberine is a cyclopentane [b] benzofuran compound found in cactus plants, several of which are used in traditional Chinese medicine. These traditional Chinese medicines are used to treat contusions, coughs, diarrhea, fever, and inflammation (NEBIGIL et al., 2020).

6 Conclusion

eIF4A1 is frequently a target of various microRNAs (miRNAs) or LncRNAs and plays a key role in tumor cell proliferation, invasion and metastasis (GINGOLD et al., 2014). Given the importance of the translation process of mRNAs in cancer development, several small molecules have been shown to have antitumor activity by acting on or inhibiting eIF4A1 (STONELEY and WILLIS, 2015). Recent studies have shown that the natural marine products cycloartenol and cycloterpenol can inhibit eIF4A1 and offer promising prospects for cancer therapy (PETERS et al., 2018). In addition, Equine uranol,

Cevistrol and Patamine A are all good inhibitors against eIF4A1 (OBLINGER et al., 2016). Rocagrelor has been shown to have potent antitumor activity *in vivo* and *in vitro* by decreasing the cellular translation rate through enhancing the mRNA binding capacity of eIF4A1 and eIF4A2 (CHU et al., 2019). However, inhibitors of eIF4A1 are still in the preclinical research stage, lacking appropriate clinical trial data and clinical evaluation, and their ability to act as antitumor agents still needs to be further explored. In this paper, we review the differential expression and protein functional role of the eIF4A1 in specific tumor types and the regulatory mechanisms, and discuss the relationship between the eIF4A1 and the large number of immune cell infiltration and tumor malignancy, which will provide clues for the next step of research. Our findings confirm the protein functional role and regulatory mechanism of eukaryotic translation initiation factor 4A1 protein in human cancer. And we propose eukaryotic translation initiation factor 4A1 as a target and biomarker for cancer prognosis, diagnosis and treatment.

Currently, the eIF4A1 inhibitor space is still stagnant in the preclinical stage, clearly defined compounds have complex mechanisms and chemical structures. With the progress of current experiments and the extensive application of artificial intelligence in predicting protein spatial configurations and related fields, it is believed that in the near future, more eIF4A1 related inhibitors will enter clinical trials and applications.

Author contributions

JH: Writing–original draft, Data curation, Formal Analysis, Investigation, Methodology, Resources, Software, Validation, Visualization. LZ: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. RY: Data curation, Investigation, Writing–original draft. LY: Data curation, Investigation, Writing–original draft. JG: Writing–original draft. DC: Writing–original draft. ZP: Writing–original draft. DL: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. YP: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. WZ: Data curation, Formal Analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing–original draft.

Conflict of interest

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

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.2023.1289650/full#supplementary-material

References

Ambaru, B., Gangadharan, G. M., Subramanya, H. S., and Gupta, C. M. (2022). Profilin is involved in G1 to S phase progression and mitotic spindle orientation during Leishmania donovani cell division cycle. *PLoS One* 17 (3), e0265692. doi:10.1371/journal.pone.0265692

Andreou, A. Z., Harms, U., and Klostermeier, D. (2019). Single-stranded regions modulate conformational dynamics and ATPase activity of eIF4A to optimize 5′-UTR unwinding. *Nucleic Acids Res.* 47 (10), 5260–5275. doi:10.1093/nar/gkz254

Andreou, A. Z., and Klostermeier, D. (2014). eIF4B and eIF4G jointly stimulate eIF4A ATPase and unwinding activities by modulation of the eIF4A conformational cycle. *J. Mol. Biol.* 426 (1), 51–61. doi:10.1016/j.jmb.2013.09.027

Asangani, I. A., Rasheed, S. A., Nikolova, D. A., Leupold, J. H., Colburn, N. H., Post, S., et al. (2008). MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* 27 (15), 2128–2136. doi:10.1038/sj.onc.1210856

Blagden, S. P., and Willis, A. E. (2011). The biological and therapeutic relevance of mRNA translation in cancer. *Nat. Rev. Clin. Oncol.* 8 (5), 280–291. doi:10.1038/nrclinonc.2011.16

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492

Chao, C.-H., Huang, L.-F., Yang, Y.-L., Su, J. H., Wang, G. H., Chiang, M. Y., et al. (2005). Polyoxygenated steroids from the gorgonian Isis hippuris. *J. Nat. Prod.* 68 (6), 880–885. doi:10.1021/np050033y

Chen, L., Miao, X., Si, C., Qin, A., Zhang, Y., Chu, C., et al. (2021a). Long non-coding RNA SENP3-eif4a1 functions as a sponge of miR-195-5p to drive triple-negative breast cancer progress by overexpressing CCNE1. *Front. Cell. Dev. Biol.* 9, 647527. doi:10.3389/fcell.2021.647527

Chen, M., Asanuma, M., Takahashi, M., Shichino, Y., Mito, M., Fujiwara, K., et al. (2021b). Dual targeting of DDX3 and eIF4A by the translation inhibitor rocaglamide A. *Cell. Chem. Biol.* 28 (4), 475–486.e8. doi:10.1016/j.chembiol.2020.11.008

Chio, I. I. C., Jafarnejad, S. M., Ponz-Sarvise, M., Park, Y., Rivera, K., Palm, W., et al. (2016). NRF2 promotes tumor maintenance by modulating mRNA translation in pancreatic cancer. *Cell.* 166 (4), 963–976. doi:10.1016/j.cell.2016.06.056

Chu, J., Zhang, W., Cencic, R., Devine, W. G., Beglov, D., Henkel, T., et al. (2019). Amidino-rocaglates: a potent class of eIF4A inhibitors. *Cell. Chem. Biol.* 26 (11), 1586–1593. doi:10.1016/j.chembiol.2019.08.008

Dever, T. E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 4 (7), a013706. doi:10.1101/cshperspect.a013706

Eberle, J., Fecker, L. F., Bittner, J. U., Orfanos, C. E., and Geilen, C. C. (2002). Decreased proliferation of human melanoma cell lines caused by antisense RNA against translation factor eIF-4A1. *Br. J. Cancer* 86 (12), 1957–1962. doi:10.1038/sj.bjc.6600351

Ernst, J. T., Thompson, P. A., Nilewski, C., Sprengeler, P. A., Sperry, S., Packard, G., et al. (2020). Design of development candidate eFT226, a first in class inhibitor of eukaryotic initiation factor 4A RNA helicase. *J. Med. Chem.* 63 (11), 5879–5955. doi:10.1021/acs.jmedchem.0c00182

Gao, C., Guo, X., Xue, A., Ruan, Y., Wang, H., and Gao, X. (2020). High intratumoral expression of eIF4A1 promotes epithelial-to-mesenchymal transition and predicts unfavorable prognosis in gastric cancer. *Acta Biochim. Biophys. Sin. (Shanghai)* 52 (3), 310–319. doi:10.1093/abbs/gmz168

GarcíA-GarcíA, C., Frieda, K. L., Feoktistova, K., Fraser, C. S., and Block, S. M. (2015). RNA BIOCHEMISTRY. Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science* 348 (6242), 1486–1488. doi:10.1126/science.aaa5089

Gingold, H., Tehler, D., Christoffersen, N. R., Nielsen, M. M., Asmar, F., Kooistra, S. M., et al. (2014). A dual program for translation regulation in cellular proliferation and differentiation. *Cell.* 158 (6), 1281–1292. doi:10.1016/j.cell.2014.08.011

Greger, H. (2022). Comparative phytochemistry of flavaglines (= rocaglamides), a group of highly bioactive flavolignans from Aglaia species (Meliaceae). *Phytochem. Rev.* 21 (3), 725–764. doi:10.1007/s11101-021-09761-5

Han, L., Wu, Y., Liu, F., et al. (2022). eIF4A1 inhibitor suppresses hyperactive mTOR-associated tumors by inducing necroptosis and G2/M arrest. *Int. J. Mol. Sci.* (13), 23. doi:10.3390/ijms23136932

Harms, U., Andreou, A. Z., Gubaev, A., and Klostermeier, D. (2014). eIF4B, eIF4G and RNA regulate eIF4A activity in translation initiation by modulating the eIF4A conformational cycle. *Nucleic Acids Res.* 42 (12), 7911–7922. doi:10.1093/nar/gku440

Howard, C. M., Estrada, M., Terrero, D., Tiwari, A. K., and Raman, D. (2020). Identification of cardiac glycosides as novel inhibitors of eIF4A1-mediated translation in triple-negative breast cancer cells. *Cancers (Basel)* 12 (8), 2169. doi:10.3390/cancers12082169

Ingolia, N. T. (2021). eIF4A inhibitors PatA and RocA stack the deck against translation. *Structure* 29 (7), 638–639. doi:10.1016/j.str.2021.06.009

Iwasaki, S., Floor, S. N., and Ingolia, N. T. (2016). Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature* 534 (7608), 558–561. doi:10.1038/nature17978

Iwasaki, S., Iwasaki, W., Takahashi, M., Sakamoto, A., Watanabe, C., Shichino, Y., et al. (2019). The translation inhibitor rocaglamide targets a bimolecular cavity between eIF4A and polypurine RNA. *Mol. Cell.* 73 (4), 738–748. doi:10.1016/j.molcel.2018.11.026

Jackson, R. J., Hellen, C. U., and Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell. Biol.* 11 (2), 113–127. doi:10.1038/nrm2838

Joyce, C. E., Yanez, A. G., Mori, A., Yoda, A., Carroll, J. S., and Novina, C. D. (2017). Differential regulation of the melanoma proteome by eIF4A1 and eIF4E. *Cancer Res.* 77 (3), 613–622. doi:10.1158/0008-5472.CAN-16-1298

Kong, T., Xue, Y., Cencic, R., Zhu, X., Monast, A., Fu, Z., et al. (2019). eIF4A inhibitors suppress cell-cycle feedback response and acquired resistance to CDK4/6 inhibition in cancer. *Mol. Cancer Ther.* 18 (11), 2158–2170. doi:10.1158/1535-7163.MCT-19-0162

Krassnig, S., Wohlrab, C., Golob-Schwarzl, N., Raicht, A., Schatz, C., Birkl-Toeglhofer, A. M., et al. (2021). A profound basic characterization of eIFs in gliomas: identifying eIF3I and 4H as potential novel target candidates in glioma therapy. *Cancers (Basel)* 13 (6), 1482. doi:10.3390/cancers13061482

Li, W., Chen, A., Xiong, L., Chen, T., Tao, F., Lu, Y., et al. (2017). miR-133a acts as a tumor suppressor in colorectal cancer by targeting eIF4A1. *Tumour Biol.* 39 (5), 1010428317698389. doi:10.1177/1010428317698389

Liang, S., Zhou, Y., Chen, Y., Ke, G., Wen, H., and Wu, X. (2014). Decreased expression of EIF4A1 after preoperative brachytherapy predicts better tumor-specific survival in cervical cancer. *Int. J. Gynecol. Cancer* 24 (5), 908–915. doi:10.1097/IGC.0000000000000152

Lin, C. J., Cencic, R., Mills, J. R., Robert, F., and Pelletier, J. (2008). c-Myc and eIF4F are components of a feedforward loop that links transcription and translation. *Cancer Res.* 68 (13), 5326–5334. doi:10.1158/0008-5472.CAN-07-5876

Lin, Y., Zhang, J., Cai, J., Liang, R., Chen, G., Qin, G., et al. (2018). Systematic analysis of gene expression alteration and Co-expression network of eukaryotic initiation factor 4A-3 in cancer. *J. Cancer* 9 (24), 4568–4577. doi:10.7150/jca.27655

Linder, P. (2006). Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Res.* 34 (15), 4168–4180. doi:10.1093/nar/gkl468

Loh, P. G., Yang, H. S., Walsh, M. A., Wang, Q., Wang, X., Cheng, Z., et al. (2009). Structural basis for translational inhibition by the tumour suppressor Pdcd4. *EMBO J.* 28 (3), 274–285. doi:10.1038/emboj.2008.278

Lomnytska, M. I., Becker, S., Gemoll, T., Lundgren, C., Habermann, J., Olsson, A., et al. (2012). Impact of genomic stability on protein expression in endometrioid endometrial cancer. *Br. J. Cancer* 106 (7), 1297–1305. doi:10.1038/bjc.2012.67

Lu, W. T., Wilczynska, A., Smith, E., and Bushell, M. (2014). The diverse roles of the eIF4A family: you are the company you keep. *Biochem. Soc. Trans.* 42 (1), 166–172. doi:10.1042/BST20130161

Ma, X., Li, B., Liu, J., Fu, Y., and Luo, Y. (2019). Phosphoglycerate dehydrogenase promotes pancreatic cancer development by interacting with eIF4A1 and eIF4E. *J. Exp. Clin. Cancer Res.* 38 (1), 66. doi:10.1186/s13046-019-1053-y

Manier, S., Huynh, D., Shen, Y. J., Zhou, J., Yusufzai, T., Salem, K. Z., et al. (2017). Inhibiting the oncogenic translation program is an effective therapeutic strategy in multiple myeloma. *Sci. Transl. Med.* 9 (389), eaal2668. doi:10.1126/scitranslmed.aal2668

Mazloomian, A., Araki, S., Ohori, M., El-Naggar, A. M., Yap, D., Bashashati, A., et al. (2019). Pharmacological systems analysis defines EIF4A3 functions in cell-cycle and RNA stress granule formation. *Commun. Biol.* 2, 165. doi:10.1038/s42003-019-0391-9

Modelska, A., Turro, E., Russell, R., Beaton, J., Sbarrato, T., Spriggs, K., et al. (2015). The malignant phenotype in breast cancer is driven by eIF4A1-mediated changes in the translational landscape. *Cell. Death Dis.* 6, e1603. doi:10.1038/cddis.2014.542

Naineni, S. K., Itoua MaïGA, R., Cencic, R., Putnam, A. A., Amador, L. A., Rodriguez, A. D., et al. (2020). A comparative study of small molecules targeting eIF4A. *RNA* 26 (5), 541–549. doi:10.1261/rna.072884.119

Naineni, S. K., Liang, J., Hull, K., Cencic, R., Zhu, M., Northcote, P., et al. (2021). Functional mimicry revealed by the crystal structure of an eIF4A:RNA complex bound to the interfacial inhibitor, desmethyl pateamine A. *Cell. Chem. Biol.* 28 (6), 825–834.e6. doi:10.1016/j.chembiol.2020.12.006

Nebigil, C. G., Moog, C., Vagner, S., Benkirane-Jessel, N., Smith, D. R., and Désaubry, L. (2020). Flavaglines as natural products targeting eIF4A and prohibitins: from traditional Chinese medicine to antiviral activity against coronaviruses. *Eur. J. Med. Chem.* 203, 112653. doi:10.1016/j.ejmech.2020.112653

Nishida, Y., Zhao, R., Heese, L. E., Akiyama, H., Patel, S., Jaeger, A. M., et al. (2021). Inhibition of translation initiation factor eIF4a inactivates heat shock factor 1 (HSF1) and exerts anti-leukemia activity in AML. *Leukemia* 35 (9), 2469–2481. doi:10.1038/s41375-021-01308-z

Oblinger, J. L., Burns, S. S., Akhmametyeva, E. M., Huang, J., Pan, L., Ren, Y., et al. (2016). Components of the eIF4F complex are potential therapeutic targets for malignant peripheral nerve sheath tumors and vestibular schwannomas. *Neuro Oncol.* 18 (9), 1265–1277. doi:10.1093/neuonc/now032

Ozes, A. R., Feoktistova, K., Avanzino, B. C., and Fraser, C. S. (2011). Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B. *J. Mol. Biol.* 412 (4), 674–687. doi:10.1016/j.jmb.2011.08.004

Peck, M. L., and Herschlag, D. (1999). Effects of oligonucleotide length and atomic composition on stimulation of the ATPase activity of translation initiation factor eIF4A. *RNA* 5 (9), 1210–1221. doi:10.1017/s1355838299990817

Pelletier, J., and Sonenberg, N. (2019). The organizing principles of eukaryotic ribosome recruitment. *Annu. Rev. Biochem.* 88, 307–335. doi:10.1146/annurev-biochem-013118-111042

Peters, T. L., Tillotson, J., Yeomans, A. M., Wilmore, S., Lemm, E., Jiménez-Romero, C., et al. (2018). Target-based screening against eIF4A1 reveals the marine natural product Elatol as a novel inhibitor of translation initiation with *in vivo* antitumor activity. *Clin. Cancer Res.* 24 (17), 4256–4270. doi:10.1158/1078-0432.CCR-17-3645

Qi, M., Qi, Y., Ma, Y., He, R., Ji, Y., Sun, Z., et al. (2013). Over-expression of human cytomegalovirus miR-US25-2-3p downregulates eIF4A1 and inhibits HCMV replication. *FEBS Lett.* 587 (14), 2266–2271. doi:10.1016/j.febslet.2013.05.057

Schmidt, T., Dabrowska, A., Waldron, J. A., Hodge, K., Koulouras, G., Gabrielsen, M., et al. (2023). eIF4A3-dependent mRNAs employ purine-rich 5'UTR sequences to activate localised eIF4A1-unwinding through eIF4A1-multimerisation to facilitate translation. *Nucleic Acids Res.* 51 (4), 1859–1879. doi:10.1093/nar/gkad030

Schutz, P., Karlberg, T., VAN Den Berg, S., Collins, R., Lehtiö, L., Högbom, M., et al. (2010). Comparative structural analysis of human DEAD-box RNA helicases. *PLoS One* 5 (9), e12791. doi:10.1371/journal.pone.0012791

Shen, L., and Pelletier, J. (2020). General and target-specific DExD/H RNA helicases in eukaryotic translation initiation. *Int. J. Mol. Sci.* 21 (12), 4402. doi:10.3390/ijms21124402

Soylemez, Z., Arikan, E. S., Solak, M., Arıkan, Y., Tokyol, Ç., and Şeker, H. (2021). Investigation of the expression levels of CPEB4, APC, TRIP13, EIF2S3, EIF4A1, IFNg, PIK3CA and CTNNB1 genes in different stage colorectal tumors. *Turk J. Med. Sci.* 51 (2), 661–674. doi:10.3906/sag-2010-18

Srivastava, R. K., Khan, J., Arumugam, A., Muzaffar, S., Guroji, P., Gorbatyuk, M. S., et al. (2021). 5'-Cap–Dependent translation as a potent therapeutic target for lethal human squamous cell carcinoma. *J. Investig. Dermatol* 141 (4), 742–753.e10. doi:10.1016/j.jid.2020.08.021

Stoneley, M., and Willis, A. E. (2015). eIF4A1 is a promising new therapeutic target in ER-negative breast cancer. *Cell. Death Differ.* 22 (4), 524–525. doi:10.1038/cdd.2014.210

Sun, Y., Atas, E., Lindqvist, L. M., Sonenberg, N., Pelletier, J., and Meller, A. (2014). Single-molecule kinetics of the eukaryotic initiation factor 4AI upon RNA unwinding. *Structure* 22 (7), 941–948. doi:10.1016/j.str.2014.04.014

Thompson, P. A., Eam, B., Young, N. P., Fish, S., Chen, J., Barrera, M., et al. (2021). Targeting oncogene mRNA translation in B-cell malignancies with eFT226, a potent and selective inhibitor of eIF4A. *Mol. Cancer Ther.* 20 (1), 26–36. doi:10.1158/1535-7163.MCT-19-0973

Waldron, J. A., Tack, D. C., Ritchey, L. E., Gillen, S. L., Wilczynska, A., Turro, E., et al. (2019). mRNA structural elements immediately upstream of the start codon dictate dependence upon eIF4A helicase activity. *Genome Biol.* 20 (1), 300. doi:10.1186/s13059-019-1901-2

Wang, C., Leavenworth, J., Zhang, C., et al. (2022). *Epigenetic regulation of EIF4A1 through DNA methylation and an oncogenic role of eIF4A1 through BRD2 signaling in prostate cancer.* China: Oncogene.

Wei, W., Cao, W., Zhan, Z., Yan, L., Xie, Y., and Xiao, Q. (2019). MiR-1284 suppresses gastric cancer progression by targeting EIF4A1. *Onco Targets Ther.* 12, 3965–3976. doi:10.2147/OTT.S191015

Wilczynska, A., Gillen, S. L., Schmidt, T., Meijer, H. A., Jukes-Jones, R., Langlais, C., et al. (2019). eIF4A2 drives repression of translation at initiation by Ccr4-Not through purine-rich motifs in the 5'UTR. *Genome Biol.* 20 (1), 262. doi:10.1186/s13059-019-1857-2

Wolf, A., Krause-Gruszczynska, M., Birkenmeier, O., Ostareck-Lederer, A., Hüttelmaier, S., and Hatzfeld, M. (2010). Plakophilin 1 stimulates translation by promoting eIF4A1 activity. *J. Cell. Biol.* 188 (4), 463–471. doi:10.1083/jcb.200908135

Wolfe, A. L., Singh, K., Zhong, Y., Drewe, P., Rajasekhar, V. K., Sanghvi, V. R., et al. (2014). RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* 513 (7516), 65–70. doi:10.1038/nature13485

Xu, Y., Lin, Z., He, L., Qu, Y., Ouyang, L., Han, Y., et al. (2021). Platelet-rich plasma-derived exosomal USP15 promotes cutaneous wound healing via deubiquitinating EIF4A1. *Oxid. Med. Cell. Longev.* 2021, 9674809. doi:10.1155/2021/9674809

Xue, C., Gu, X., Li, G., Bao, Z., and Li, L. (2021). Expression and functional roles of eukaryotic initiation factor 4A family proteins in human cancers. *Front. Cell. Dev. Biol.* 9, 711965. doi:10.3389/fcell.2021.711965

Yang, T., Chen, W. C., Shi, P. C., Liu, M. R., Jiang, T., Song, H., et al. (2020). Long noncoding RNA MAPKAPK5-AS1 promotes colorectal cancer progression by cis-regulating the nearby gene MK5 and acting as a let-7f-1-3p sponge. *J. Exp. Clin. Cancer Res.* 39 (1), 139. doi:10.1186/s13046-020-01633-8

Zhao, M., Ding, L., Yang, Y., Chen, S., Zhu, N., Fu, Y., et al. (2019). Aberrant expression of PDCD4/eIF4A1 signal predicts postoperative recurrence for early-stage oral squamous cell carcinoma. *Cancer Manag. Res.* 11, 9553–9562. doi:10.2147/CMAR.S223273

Zhao, Y., Li, T., Tian, S., Meng, W., Sui, Y., Yang, J., et al. (2020). Effective inhibition of MYC-amplified group 3 medulloblastoma through targeting EIF4A1. *Cancer Manag. Res.* 12, 12473–12485. doi:10.2147/CMAR.S278844

Zhao, Y., Wang, Y., Chen, W., Bai, S., Peng, W., Zheng, M., et al. (2021). Targeted intervention of eIF4A1 inhibits EMT and metastasis of pancreatic cancer cells via c-MYC/miR-9 signaling. *Cancer Cell. Int.* 21 (1), 670. doi:10.1186/s12935-021-02390-0

Mendelian randomization analysis to elucidate the causal relationship between small molecule metabolites and ovarian cancer risk

Xin Chang[1,2], Shijia Liu[1,2] and Lu Han[1*]

[1]Department of Gynecology, Dalian Women and Children's Medical Group, Dalian, Liaoning, China,
[2]Department of Graduate, Dalian Medical University, Dalian, Liaoning, China

Background: Small molecule metabolites are potential biomarkers for ovarian cancer. However, the causal relationship between small molecule metabolites and ovarian cancer remains unclear.

Methods: Single nucleotide polymorphisms (SNPs) correlated with 53 distinct small molecule metabolites were identified as instrumental variables (IVs) from comprehensive genome-wide association studies. Aggregate data encompassing 25,509 cases of ovarian cancer and 40,941 controls of European descent were procured from the Ovarian Cancer Association Consortium. To evaluate causative associations, four Mendelian randomization techniques—including inverse-variance weighted, weighted median, maximum likelihood, and MR-Egger regression—were employed.

Results: In total, 242 SNPs were delineated as IVs for the small molecule metabolites under consideration. A significant association with the overarching risk of ovarian cancer was observed for six distinct metabolites. Hexadecenoylcarnitine and methioninesulfoxide were associated with a 32% and 31% reduced risk, respectively. Fifteen metabolites were linked to subtype ovarian cancers. For instance, both methionine sulfoxide and tetradecanoyl carnitine exhibited an inverse association with the risk of clear cell and high-grade serous ovarian cancers. Conversely, tryptophan demonstrated a 1.72-fold elevated risk for endometrioid ovarian cancer.

Conclusion: This study identified several metabolites with putative causal effects on ovarian cancer risk using Mendelian randomization analysis. The findings provide insight into the etiological role of small molecule metabolites and highlight potential early detection biomarkers for ovarian cancer. Subsequent investigations are imperative to corroborate these findings and elucidate the underlying pathophysiological mechanisms.

KEYWORDS

ovarian cancer, Mendelian randomization, single nucleotide polymorphisms, amino acids, biomarkers

1 Introduction

Cancer remains one of the most formidable adversaries in the realm of global health, contributing significantly to the burden of disease and mortality, with its impact felt acutely in developing countries. Within this broader context, ovarian cancer (OC) emerges as a predominant gynecological malignancy. Recent statistics from 2020 have underscored this reality, revealing an estimated 310,000 new cases of OC and a deeply concerning figure of 210,000 deaths associated with the condition. The trajectory of OC is particularly alarming, with projections suggesting that by the year 2040, we may witness the global incidence of this cancer soar to approximately 434,184 cases (1–4). The high mortality rate is largely attributed to the asymptomatic nature and late diagnosis of OC (5). In diseases with a significant burden, such as OC, the underlying etiology and pathogenesis remain largely elusive. Established risk factors for OC encompass age at menarche, age at natural menopause, and age at diagnosis of endometriosis (6). Elevated dietary consumption of fiber and soy has demonstrated potential prophylactic benefits against OC (7, 8). Furthermore, a deficiency in vitamin D levels has been associated with an augmented risk of OC (9). Identification of novel biomarkers that can detect OC at an early stage or predict susceptibility is urgently needed to reduce disease burden.

Emerging evidence suggests that metabolic reprogramming is implicated in ovarian tumorigenesis and progression (10). Metabolomics profiling has revealed aberrant levels of various small molecule metabolites, such as amino acids, biogenic amines, acylcarnitines, and carbohydrates, in OC (10–13). These small molecules are involved in multiple oncogenic signaling pathways and may serve as diagnostic biomarkers or therapeutic targets. Recent comprehensive genome-wide association studies (GWAS) have delineated single nucleotide polymorphisms (SNPs) correlated with metabolic phenotypes. These SNPs can be judiciously employed as instrumental variables to infer putative causal associations between specific metabolites and disease outcomes (14–16). Conversely, a limited number of studies have delved into the relationship between OC and the small molecular derivatives of metabolism.

Mendelian randomization (MR) analysis employs genetic variants as instrumental variables (IVs), enhancing the robustness of causal inference and mitigating biases stemming from reverse causation and confounding (17). This methodology has been extensively employed to assess the putative causal role of alterable exposures in carcinogenesis (18). However, to date, no investigation has probed the potential causal implications of small molecular metabolites on OC via Mendelian randomization. In this study, we performed a two-sample MR analysis to evaluate putative causal

Abbreviations: OC, ovarian cancer; HGSOC, high-grade serous ovarian cancer; LMPOC, low malignant potential ovarian cancer; IMOC, invasive mucinous ovarian cancer; CCOC, clear cell ovarian cancer; EndoOC, endometrioid ovarian cancer; GWAS, genome-wide association studies; SNPs, single nucleotide polymorphisms; MR, Mendelian randomization; IVs, instrumental variables; IVW, inverse-variance weighted; LD, linkage disequilibrium; CPT-1, carnitine palmitoyltransferase-1.

associations of genetically predicted small molecule metabolites with OC risk. Findings from this study could uncover novel etiological mechanisms and guide future development of metabotype-based biomarkers for OC.

2 Materials and methods

2.1 Study design

In our study, we utilized SNPs identified through GWAS as genetic instrumental variables (IVs), aiming to investigate the plausible causal connection between small molecule metabolic products and ovarian cancer. As presented in Figure 1, our two-sample MR study was built upon three principal assumptions (19). 1) Relevance assumption: The IVs had a strong connection to the exposure (19). 2) Independence assumption: There was no correlation between the IVs and any variables that affected both exposure and outcome (19). 3) Exclusion Restriction Assumption: The IVs exclusively influenced the exposure, without introducing any additional causal pathways that could impact the outcome (19). All the summary data utilized in our study were openly accessible to the general public (IEU OpenGWAS project). Additional data can be found in the Supplementary Material (Supplementary Table 1). As our research relied on publicly available GWAS data, no additional ethical approval was necessary.

2.2 Data source and study samples of ovarian cancer

This study considered a total of six frequently observed clinical phenotypes of ovarian cancer, specifically: OC, high-grade serous ovarian cancer (HGSOC), low malignant potential ovarian cancer (LMPOC), invasive mucinous ovarian cancer (IMOC), clear cell ovarian cancer (CCOC), and endometrioid ovarian cancer (EndoOC). The findings presented in this study are based on a genome-wide association studies conducted within the Ovarian Cancer Association Consortium (OCAC) (20). This thorough analysis was conducted using a dataset that encompassed 25,509 cases of ovarian cancer and 40,941 controls of European ancestry, enabling an exploration of the associations between genetic factors and ovarian cancer (21). The dataset encompasses 63 distinct genotyping project/case-control sets (21). Genomic information was acquired through direct genotyping utilizing an Illumina Custom Infinium array, known as OncoArray, featuring around 530,000 SNPs (21). To enhance the dataset's comprehensiveness, imputation was executed utilizing the 1000 Genomes Project Phase 3 dataset as a reference (21). The cases encompassed the subsequent invasive epithelial ovarian cancer types: HGSOC (n = 13,037), LMPOC (n = 3,103), MOC (n = 1,417), CCOC (n = 1,366) and EndoOC (n = 2,810) (21). The majority of individuals were recruited from cancer genetics clinics, which also included some related individuals (21). More specific information regarding the cohorts, genotyping, quality control, and imputation can be viewed in previous studies (21).

FIGURE 1
Flow chart of MR analysis and three assumptions in this study. GWAS, Genome-wide association study; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

2.3 Genetic instruments selection

The genetic components involving small molecule metabolic products, which encompass Acylcarnitines, Amino acids, Biogenic amines, and Hexose, were derived from a GWAS. This study involved a collective cohort of 86,507 adults with European heritage drawn from the Fenland cohort, with synergies established between the EPIC-Norfolk and INTERVAL studies (14). We first selected IVs for each small molecule products based on a strict cutoff of $P < 5 \times 10^{-8}$. Independent SNPs (r2 < 0.01, distance = 250 kb) were preserved after calculating the linkage disequilibrium (LD) of related SNPs. Furthermore, the robustness of the genetic instruments was evaluated through F-statistics to mitigate potential biases from weak instruments. The F-statistics were computed using the formula: F-statistics = $(Beta/Se)^2$, with the mean serving as the comprehensive measure and an F-statistic > 10 signified substantial statistical potency (22, 23). Finally, a total of 242 SNPs associated with 53 small molecule products of metabolism were remained as the instrument variables (IVs). Detailed information of the IVs form small molecule products of metabolism were summarized in Supplementary Table 2, respectively.

2.4 Statistical analyses

Four methods including the inverse-variance-weighting (IVW), weighted median, maximum likelihood-based methods, and MR-Egger regression. maximum likelihood-based methods were performed to assess the causal association between small molecule products of metabolism and OC. The IVW method operates under the assumption of the validity of all instrumental variables, amalgamating their effects to produce an overall weighted outcome (24). Given the potential heterogeneity, the random effect and fixed effect IVW were both calculated and regarded as the main analyses (24). The weighted median estimator can

generate resilient causal estimates, maintaining robustness even when up to 50% of instrumental variables may be invalid (25). Besides, under an assumption of a linear relationship between exposure and outcome, the maximum likelihood-based method offered normal bivariate distribution for the estimated causal association (26). Finally, the MR Egger method introduces an intercept term in the regression model to assess the directional pleiotropy (27). A substantially non-zero intercept term in statistical analysis signals the existence of pleiotropy and a breach in the fundamental Mendelian randomization assumption (27).

We employed the Cochran's Q test to evaluate the heterogeneity among IVs (28). In case of notable heterogeneity being detected ($P < 0.05$), the random-effects model was employed; conversely, if heterogeneity was not significant ($P > 0.05$), the fixed-effects model was utilized (28). A leave-one-out analysis was conducted to pinpoint influential SNPs in the causal estimations. A threshold of statistical significance was set at $P < 0.05$ (two-sided). When the quantity of SNPs is fewer than four, the analysis is confined to using the IVW method. All analyses were performed using "TwoSampleMR", and "gg-plot2" packages in R software (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria).

3 Results

3.1 Causal estimates of genetically predicted small molecule metabolic products on overall ovarian cancer

As shown in Table 1 and Figure 2, we totally found six small molecule metabolic products were associated with overall OC. In brief, genetically predicted hexadecenoylcarnitine as well as methioninesulfoxide dropped a 32% (OR, 0.68; 95% CI =0.51-0.91, $P = 0.010$) and 31% (OR=0.69, 95% CI =0.48-1.00, $P =$

TABLE 1 The main result of small molecule metabolites and ovarian cancer risk.

Outcome	Exposure	Method	Number of SNP	OR	LCI	UCI	P-value	P for heterogeneity	P for pleiotropy
Overall ovarian cancer	Hexadecenoylcarnitine	Wald ratio	1	0.68	0.51	0.91	0.010	/	/
		Wald ratio	1	0.69	0.48	1.00	0.048	/	/
Overall ovarian cancer	Octadecandienylcarnitine	Inverse variance weighted (fixed effects)	4	0.88	0.80	0.97	0.011	0.426	
		Maximum likelihood	4	0.88	0.79	0.97	0.011		
		Simple median	4	0.88	0.76	1.01	0.075		
		Weighted median	4	0.86	0.77	0.96	0.007		
		MR Egger	4	0.84	0.64	1.11	0.341		0.766
Overall ovarian cancer	Octadecenoylcarnitine	Inverse variance weighted (fixed effects)	2	0.83	0.73	0.95	0.008	0.478	/
		Maximum likelihood	2	0.83	0.72	0.96	0.010		
Overall ovarian cancer	Phenylalanine	Inverse variance weighted (fixed effects)	4	1.29	1.03	1.62	0.028	0.894	
		Maximum likelihood	4	1.29	1.03	1.63	0.029		
		Simple median	4	1.32	1.00	1.74	0.050		
		Weighted median	4	1.32	1.02	1.71	0.038		
		MR Egger	4	1.21	0.44	3.32	0.748		0.907
Overall ovarian cancer	Tetradecanoylcarnitine	Inverse variance weighted (fixed effects)	2	0.80	0.66	0.97	0.020	0.168	/
		Maximum likelihood	2	0.80	0.66	0.97	0.022		
High grade serous ovarian cancer	Tetradecanoylcarnitine	Inverse variance weighted (fixed effects)	3	0.81	0.66	0.99	0.041	0.696	
		Maximum likelihood	3	0.81	0.66	0.99	0.042		
		Simple median	3	0.80	0.62	1.03	0.089		
		Weighted median	3	0.79	0.63	0.99	0.039		
		MR Egger	3	0.48	0.13	1.79	0.471		0.574
Low malignant	Arginine	Inverse variance	7	0.63	0.42	0.95	0.028	0.871	

(Continued)

TABLE 1 Continued

Outcome	Exposure	Method	Number of SNP	OR	LCI	UCI	P-value	P for heterogeneity	P for pleiotropy
potential ovarian cancer		weighted (fixed effects)							
		Maximum likelihood	7	0.63	0.42	0.95	0.029		
		Simple median	7	0.69	0.39	1.22	0.201		
		Weighted median	7	0.57	0.34	0.93	0.025		
		MR Egger	7	0.53	0.27	1.04	0.124		0.537
Low malignant potential ovarian cancer	Dodecanoylcarnitine	Wald ratio	1	0.12	0.02	0.74	0.022	/	/
Low malignant potential ovarian cancer	Leucine	Inverse variance weighted (fixed effects)	4	4.25	1.22	14.83	0.023	0.064	
		Maximum likelihood	4	4.45	1.23	16.08	0.023		
		Simple median	4	4.84	0.91	25.83	0.065		
		Weighted median	4	4.35	0.82	23.06	0.084		
		MR Egger	4	0.19	0.00	99.92	0.659		0.416
Low malignant potential ovarian cancer	Octadecenoylcarnitine	Inverse variance weighted (fixed effects)	2	0.61	0.37	1.00	0.049	0.124	/
		Maximum likelihood	2	0.60	0.36	1.00	0.050		
Low malignant potential ovarian cancer	Threonine	Inverse variance weighted (fixed effects)	3	0.42	0.23	0.77	0.005	0.352	
		Maximum likelihood	3	0.41	0.22	0.77	0.005		
		Simple median	3	0.66	0.26	1.70	0.390		
		Weighted median	3	0.44	0.22	0.86	0.017		
		MR Egger	3	0.07	0.00	1.01	0.301		0.406
Invasive mucinous ovarian cancer	alpha-Aminoadipic acid	Inverse variance weighted (fixed effects)	2	2.09	1.02	4.28	0.045	0.811	/

(Continued)

TABLE 1 Continued

| Outcome	Exposure	Method	Number of SNP	OR	LCI	UCI	P-value	P for heterogeneity	P for pleiotropy
		Maximum likelihood	2	2.09	1.00	4.35	0.049		
Invasive mucinous ovarian cancer	Creatinine	Inverse variance weighted (fixed effects)	12	0.45	0.24	0.84	0.012	0.666	
		Maximum likelihood	12	0.45	0.24	0.84	0.013		
		Simple median	12	0.41	0.18	0.95	0.038		
		Weighted median	12	0.43	0.19	0.99	0.046		
		MR Egger	12	0.05	0.00	1.00	0.079		0.175
Invasive mucinous ovarian cancer	Hexose	Wald ratio	1	2.51	1.05	6.02	0.039	/	/
Invasive mucinous ovarian cancer	Methionine	Wald ratio	1	0.23	0.06	0.89	0.033	/	/
Invasive mucinous ovarian cancer	Tetradecenoylcarnitine	Inverse variance weighted (fixed effects)	2	0.39	0.17	0.93	0.034	0.502	
		Maximum likelihood	2	0.39	0.16	0.95	0.038		
Clear cell ovarian cancer	Butyrylcarnitine	Inverse variance weighted (fixed effects)	2	0.62	0.40	0.95	0.029	0.116	/
		Maximum likelihood	2	0.62	0.40	0.95	0.030		
Clear cell ovarian cancer	Methioninesulfoxide	Wald ratio	1	0.28	0.09	0.85	0.025	/	/
Endometrioid ovarian cancer	Citrulline	Inverse variance weighted (fixed effects)	4	1.65	1.17	2.34	0.005	0.259	
		Maximum likelihood	4	1.66	1.17	2.38	0.005		
		Simple median	4	1.64	1.08	2.49	0.021		
		Weighted median	4	1.80	1.16	2.80	0.008		
		MR Egger	4	0.44	0.03	6.32	0.606		0.428
Endometrioid ovarian cancer	Tryptophan	Inverse variance weighted (fixed effects)	2	1.72	1.12	2.64	0.013	0.698	/
		Maximum likelihood	2	1.72	1.11	2.66	0.015		
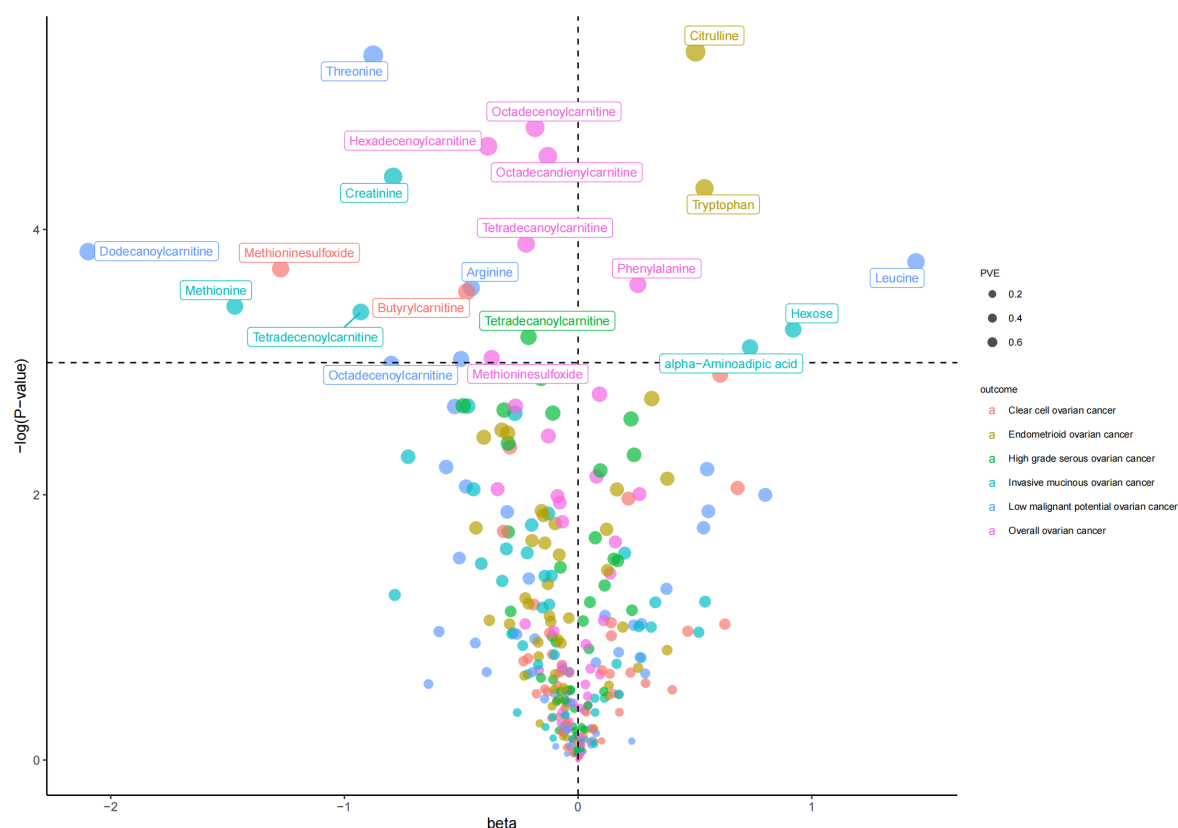
FIGURE 2

The volcano plot for inverse-variance-weighted method in the association between molecule metabolic products and ovarian cancer.

0.048) risk of overall OC by the Wald ratio method, respectively. This decreased risk was also observed in the association between octadecandienylcarnitine and overall OC, replicated by the Maximum likelihood method (OR = 0.88, 95% CI = 0.79-0.97, P =0.011 and weighted median approach (OR=0.86, 95% CI=0.77-0.96, P=0.007). Estimates in size were similar for the association of octadecenoylcarnitine and tetradecanoylcarnitine with overall OC (Supplementary Table 3). The heterogeneity test and pleiotropy test indicated that there no influence for the casual effect of octadecandienyl carnitine on overall ovarian cancer (P>0.05).

3.2 Causal estimates of genetically predicted small molecule metabolic products on subtype ovarian cancers

The results of all small molecule metabolic products on subtype ovarian cancers were presented in the Supplementary Tables 4–8. Figure 2 presented the estimate from the MR analysis and suggested that a total of 15 small molecule metabolic products were related to the subtype ovarian cancers. Methionine sulfoxide was observed that associated clear cell ovarian cancer with dramatically reduced risk (OR=0.28, 95% CI=0.09-0.85, P=0.024). The similar causal association between tetradecanoylcarnitine and high grade serous ovarian cancer was detected. For endometrioid ovarian cancer,

IVW method suggested genetically predicted Tryptophan would climb its 1.72-fold risk (95% CI=1.12-2.64, P =0.013). Five small molecule metabolic products were found that related to low malignant potential ovarian cancer and invasive mucinous ovarian cancer, respectively. For example, genetically predicted creatinine reduced the risk of invasive mucinous ovarian cancer, with estimates of IVW at 0.45 (95% CI=0.24-0.84, P =0.012; Figure 3). This causal relationship also was verified by Maximum likelihood approach and simple median method, while it did not attach a statistical significance in weighted median. Besides, arginine had a negative effect (OR=0.63, 95% CI=0.42-0.95, P =0.028; Figure 4) on low malignant potential ovarian cancer as well as threonine (OR=0.42, 95% CI=0.23-0.77, P =0.004; Figure 5). The pleiotropy test of Egger intercept suggested that there was no pleiotropy (P>0.05).

4 Discussion

OC remains one of the foremost gynecological malignancies with a significant global impact (5). Despite significant advancements in elucidating its etiology, the insidious onset of OC frequently results in advanced-stage diagnoses, underscoring the paramount importance of early detection biomarkers (5). This current study addresses this pressing clinical gap, harnessing the

FIGURE 3
The negative effect of creatinine on the invasive mucinous ovarian cancer risk in the IVW analysis and it verified by Maximum likelihood approach and simple median method (all *P*<0.05).

FIGURE 4
The inverse causal estimate between molecule metabolic products and the low malignant potential ovarian cancer (Increased arginine level may decrease the 46.7% risk of low malignant potential ovarian cancer. *P* =0.004).
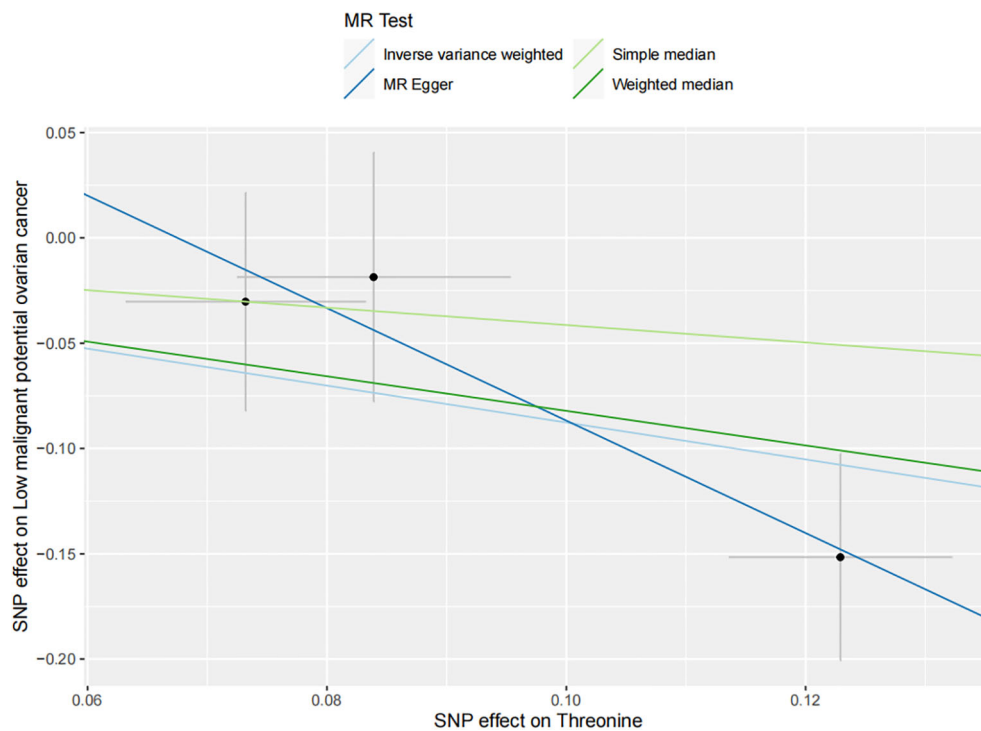
FIGURE 5
The inverse causal estimate between molecule metabolic products and the low malignant potential ovarian cancer. Threonine has a negative impact on the low malignant potential ovarian cancer (OR=0.42, 95% CI=0.23-0.77, *P* =0.004).

capabilities of Mendelian randomization to discern putative causal associations between small molecular metabolites and OC susceptibility. Significantly, this represents the inaugural effort of its nature to apply this methodology on comprehensive genetic datasets to assess these correlations. Overall, our findings provide novel insights into the complex metabolic underpinnings of OC.

In recent studies, several small-molecule metabolites have emerged as potential biomarkers for the early detection, risk stratification, and targeted prevention of OC (29).

In our comprehensive metabolic analysis, we identified a total of six metabolites significantly associated with the incidence of overall OC. Decreased risks were observed in association with the following metabolites: octadecandienylcarnitine, octadecenoylcarnitine, hexadecenoylcarnitine, tetradecanoylcarnitine, and methionine sulfoxide. Conversely, an elevated phenylalanine level was significantly associated with an augmented risk of overall OC. For HGSOC, tetradecanoylcarnitine was indicative of a reduced risk. In the context of LMPOC, the metabolites arginine, octadecenoylcarnitine, and threonine were inversely correlated with risk, while an increase in leucine levels showed a heightened risk. Regarding IMOC, a diminished risk was noted in conjunction with creatinine, decenoylcarnitine, methionine, and tetradecanoylcarnitine. Conversely, the levels of alpha-Aminoadipic acid and hexose were positively correlated with increased risk. For CCOC, butyrylcarnitine was a marker of reduced risk, whereas increased methioninesulfoxide levels were linked to heightened risk. In EndoOC, citrulline and tryptophan were indicative of a reduced risk. However, elevated glycine levels were observed to increase the risk.

Notably, an increase in genetically predicted levels of methionine sulfoxide correlated with a 31% reduction in the risk of overall OC. Methioninesulfoxide is generated via oxidation of methionine residues in proteins and serves as a biomarker of oxidative damage (30, 31). Accumulating evidence suggests that methioninesulfoxide reductases act as antioxidant repair enzymes to revert oxidized methionines and defend against oxidative stress (32). Our results indicate that methioninesulfoxide may confer protection against OC through antioxidant effects. Genetically elevated tryptophan levels were associated with a 1.72-fold increased risk of endometrioid OC in our analysis. Tryptophan is an essential amino acid and precursor for bioactive molecules like serotonin and melatonin (33). Previous studies have found that changes in tryptophan metabolism in tumors are often accompanied by abnormal expression of tryptophan-related enzyme genes. Among the observed alterations, variations in the expression of genes associated with indoleamine 2,3-dioxygenase and tryptophan 2,3-dioxygenase emerge as the most prevalent (34). In the human body, tryptophan has three metabolic pathways. Catabolism of tryptophan through the kynurenine pathway produces immunosuppressive metabolites and has been implicated in facilitating tumor immune evasion (35). We also found that higher predicted arginine and threonine were associated with 37% and 58% decreased risks of low malignant potential ovarian cancer, respectively. Threonine serves several functions. One of its primary roles is in the synthesis of mucin, a substance crucial for maintaining intestinal integrity and function (36). Additionally, threonine plays a significant part in immune function, contributing to the body's defense mechanisms (36). It is also involved in the

phosphorylation and glycosylation of proteins, processes that are essential for protein function and stability (36). Lastly, threonine in the synthesis of glycine, an amino acid that has numerous roles in the body (36). Arginine is a semi-essential amino acid and substrate for nitric oxide synthesis (37). Nitric oxide is a ubiquitous messenger molecule with dichotomous pro- and anti-tumorigenic actions (38). Further research should clarify the role of arginine metabolism in OC.

We found that higher genetically predicted levels of several acylcarnitine species (tetradecanoylcarnitine, hexadecenoylcarnitine, octadecenoylcarnitine, octadecadienyl carnitine) were associated with decreased risks of OC. Acylcarnitines are generated via esterification of fatty acids and shuttle lipids into the mitochondrial matrix for β-oxidation (39). Reduced acylcarnitine levels imply impaired fatty acid oxidation and mitochondrial dysfunction, which are implicated in OC (40, 41). Additionally, an association was observed between tetradecanoylcarnitine and a diminished risk of HGSOC. Carnitine palmitoyltransferase-1 (CPT-1), positioned on the outer membrane of mitochondria, principally catalyzes the conversion of long-chain fatty acyl-CoA and carnitine to fatty acyl carnitine. This conversion represents the preliminary rate-limiting phase in the mitochondrial oxidation of fatty acids (42). CPT-1 downregulation induces a glycolytic shift in cancer cells (43, 44). Targeting CPT-1 may thus restrain HGSOC growth by blocking fatty acid oxidation.

Regarding IMOC, a diminished risk was noted in conjunction with higher genetically predicted levels of creatinine, decenoylcarnitine, methionine, and tetradecenoylcarnitine. This suggests impairments in pathways related to these metabolites, such as fatty acid oxidation, antioxidant defenses, and nitrogen metabolism, may contribute to the development of this ovarian cancer subtype (45). Creatinine is a breakdown product of creatine phosphate in muscle and is usually produced at a fairly constant rate by the body (46). The lower creatinine levels associated with higher ovarian cancer risk may indicate impaired muscle metabolism or greater catabolism in this patient population. This is also consistent with previous studies (47). Decenoylcarnitine is a medium-chain fatty acid derivative involved in transporting fatty acids into the mitochondria for beta-oxidation. The reduced cancer risk with higher decenoylcarnitine hints at a possible role of improved fatty acid metabolism in protecting against ovarian carcinogenesis. Methionine is an essential amino acid that serves as a precursor for protein synthesis and other important biomolecules like cysteine and taurine (48). The inverse association between methionine and IMOC risk is consistent with its known functions in maintaining genomic stability and redox homeostasis through DNA methylation and antioxidant systems (49). Higher methionine levels may suppress ovarian tumorigenesis through these mechanisms. In contrast, elevated levels of alpha-aminoadipic acid and hexose were associated with increased IMOC risk. Alpha-aminoadipic acid is an intermediate in lysine degradation, while hexoses are simple sugars. The positive correlations indicate dysregulated lysine catabolism and carbohydrate metabolism could play pathogenic roles. Alpha-aminoadipic acid is an intermediate in lysine metabolism and a marker of oxidative stress that may accumulate with possible lysine deficiency or dysfunction in this pathway (50). Hexose represents the combined pool of six-carbon sugars including glucose and fructose (51). The increased ovarian cancer risk with higher hexose levels may

stem from greater availability of glycolytic intermediates to fuel rapid tumor growth (52). This fits with existing evidence on the key role of glycolytic metabolism in ovarian cancer progression (53).

For CCOC, elevated levels of butyrylcarnitine were associated with a reduced risk. Butyrylcarnitine is involved in fatty acid metabolism, and previous studies have found fatty acid oxidation pathways to be downregulated in CCOC (42). The reduced butyrylcarnitine levels observed here likely reflect impairments in this metabolic pathway that may promote CCOC pathogenesis.

For EndoOC, lower citrulline and tryptophan levels were indicative of a reduced risk. Citrulline is a key intermediate in the urea cycle, while tryptophan is an essential amino acid. Past work indicates both these metabolites are involved in maintaining immune homeostasis (54). The decreased levels seen here imply EndoOC risk may rise when immune regulation is disrupted. Meanwhile, elevated glycine was tied to heightened EndoOC risk. Glycine serves as a precursor for glutathione, a key antioxidant. The increased glycine levels likely reflect a compensatory response to mitigate oxidative damage that otherwise enables EndoOC pathogenesis (55).

This study has several strengths. We employed mendelian randomization–a powerful genetic epidemiological tool. It utilizes SNPs closely tied to the exposure, serving as IVs to uncover potential causal links between the exposure and the outcome. Genotypes are believed to be randomly distributed during gametogenesis. Thus, using the IVs model addresses many confounding challenges in observational research, especially when biases arise from unmeasured confounders. Thanks to the inherent randomness of genotypic distribution, MR helps counter potential confounding and reverse causality. We drew from the most extensive GWAS dataset on OC, enhancing our statistical validity.

This study presents several limitations. Primarily, the cohort was confined to individuals of European descent, which minimizes potential bias from population stratification but may not adequately capture the diversity of SNP redundancy, particularly given the unavailability of the original dataset. Moreover, while our results suggest a potential causal linkage between small molecular metabolites and OC, the clarity of data presentation regarding the relationships between different metabolites could be improved for the reader's comprehension and comparative analysis. Recognizing these issues, we assert the necessity for subsequent investigations, including experimental validation in a broader array of populations and in-depth exploration of the underlying biological mechanisms. Such research will not only corroborate our findings but also illuminate the complex metabolic interactions associated with OC, offering substantial contributions to the oncological community's understanding of this disease.

5 Conclusion

In this MR analysis, we observed putatively causal associations between certain small molecule metabolites and the risk of OC. Our observations underscore the potential for metabolic profiling in risk stratification, early diagnosis, and individualized preventive strategies for OC. These findings not only enhance our etiological understanding but also pave the way for subsequent investigations into targeting anomalous metabolic pathways in OC.

Data availability statement

Author contributions

XC: Conceptualization, Investigation, Methodology, Writing – original draft. SL: Formal analysis, Project administration, Validation, Visualization, Writing – original draft. LH: Formal analysis, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1291033/full#supplementary-material

References

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

2. Yang Y, Li X, Qian H, Di G, Zhou R, Dong Y, et al. C-reactive protein as a prognostic biomarker for gynecologic cancers: A meta-analysis. *Comput Intell Neurosci* (2022) 2022:6833078. doi: 10.1155/2022/6833078

3. Zara R, Rasul A, Sultana T, Jabeen F, Selamoglu Z. Identification of macrolepiota procera extract as a novel G6PD inhibitor for the treatment of lung cancer. *Saudi J Biol Sci* (2022) 29:3372–9. doi: 10.1016/j.sjbs.2022.02.018

4. Salehi B, Selamoglu Z S, Mileski K, Pezzani R, Redaelli M, WC C, et al. Liposomal cytarabine as cancer therapy: From chemistry to medicine. *Biomolecules* (2019) 9:773. doi: 10.3390/biom9120773

5. Torre LA, Trabert B, DeSantis CE, Miller KD, Samimi G, Runowicz CD, et al. Ovarian cancer statistics, 2018. *CA: A Cancer J Clin* (2018) 68:284–96. doi: 10.3322/caac.21456

6. Mallen AR, Townsend MK, Tworoger SS. Risk factors for ovarian carcinoma. *Hematol Oncol Clin North Am* (2018) 32:891–902. doi: 10.1016/j.hoc.2018.07.002

7. Lheureux S, Braunstein M, Oza AM. Epithelial ovarian cancer: Evolution of management in the era of precision medicine. *CA Cancer J Clin* (2019) 69:280–304. doi: 10.3322/caac.21559

8. Huang X, Wang X, Shang J, Lin Y, Yang Y, Song Y, et al. Association between dietary fiber intake and risk of ovarian cancer: a meta-analysis of observational studies. *J Int Med Res* (2018) 46:3995–4005. doi: 10.1177/0300060518792801

9. Guo H, Guo J, Xie W, Yuan L, Sheng X. The role of vitamin d in ovarian cancer: epidemiology, molecular mechanism and prevention. *J Ovarian Res* (2018) 11:71. doi: 10.1186/s13048-018-0443-7

10. Sun L, Suo C, Li S-T, Zhang H, Gao P. Metabolic reprogramming for cancer cells and their microenvironment: Beyond the warburg effect. *Biochim Biophys Acta Rev Cancer* (2018) 1870:51–66. doi: 10.1016/j.bbcan.2018.06.005

11. Fang F, He X, Deng H, Chen Q, Lu J, Spraul M, et al. Discrimination of metabolic profiles of pancreatic cancer from chronic pancreatitis by high-resolution magic angle spinning 1H nuclear magnetic resonance and principal components analysis. *Cancer Sci* (2007) 98:1678–82. doi: 10.1111/j.1349-7006.2007.00589.x

12. Feng Y, Tang Y, Mao Y, Liu Y, Yao D, Yang L, et al. PAX2 promotes epithelial ovarian cancer progression involving fatty acid metabolic reprogramming. *Int J Oncol* (2020) 56:697–708. doi: 10.3892/ijo.2020.4958

13. Li N, Li N, Wen S, Li B, Zhang Y, Liu Q, et al. HSP60 regulates lipid metabolism in human ovarian cancer. *Oxid Med Cell Longev* (2021) 2021:6610529. doi: 10.1155/2021/6610529

14. Lotta LA, Pietzner M, Stewart ID, Wittemans LBL, Li C, Bonelli R, et al. A cross-platform approach identifies genetic regulators of human metabolism and health. *Nat Genet* (2021) 53:54–64. doi: 10.1038/s41588-020-00751-5

15. Long T, Hicks M, Yu H-C, Biggs WH, Kirkness EF, Menni C, et al. Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nat Genet* (2017) 49:568–78. doi: 10.1038/ng.3809

16. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet* (2012) 90:7–24. doi: 10.1016/j.ajhg.2011.11.029

17. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ* (2005) 330:1076–9. doi: 10.1136/bmj.330.7499.1076

18. Morris TT, Heron J, Sanderson ECM, Davey Smith G, Didelez V, Tilling K. Interpretation of mendelian randomization using a single measure of an exposure that varies over time. *Int J Epidemiol* (2022) 51:1899–909. doi: 10.1093/ije/dyac136

19. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. *Nat Rev Methods Primers* (2022) 2:6. doi: 10.1038/s43586-021-00092-5

20. Pearce CL, Near AM, Van Den Berg DJ, Ramus SJ, Gentry-Maharaj A, Menon U, et al. Validating genetic risk associations for ovarian cancer through the international ovarian cancer association consortium. *Brit J Cancer* (2009) 100:412–20. doi: 10.1038/sj.bjc.6604820

21. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, et al. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet* (2017) 49:680–91. doi: 10.1038/ng.3826

22. Georgakis MK, Gill D, Rannikmäe K, Traylor M, Anderson CD, Lee J-M, et al. Genetically determined levels of circulating cytokines and risk of stroke. *Circulation* (2019) 139:256–68. doi: 10.1161/CIRCULATIONAHA.118.035905

23. Chen L, Yang H, Li H, He C, Yang L, Lv G. Insights into modifiable risk factors of cholelithiasis: A mendelian randomization study. *Hepatol (Baltimore Md)* (2022) 75:785–96. doi: 10.1002/hep.32183

24. Brion M-JA, Shakhbazov K, Visscher PM. Calculating statistical power in mendelian randomization studies. *Int J Epidemiol* (2013) 42:1497–501. doi: 10.1093/ije/dyt179

25. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* (2016) 40:304–14. doi: 10.1002/gepi.21965

26. Xue H, Shen X, Pan W. Constrained maximum likelihood-based mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. *Am J Hum Genet* (2021) 108:1251–69. doi: 10.1016/j.ajhg.2021.05.014

27. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. *Int J Epidemiol* (2015) 44:512–25. doi: 10.1093/ije/dyv080

28. Greco M FD, Minelli C, Sheehan NA, Thompson JR. Detecting pleiotropy in mendelian randomisation studies with summary data and a continuous outcome. *Stat Med* (2015) 34:2926–40. doi: 10.1002/sim.6522

29. Kuk C, Kulasingam V, Gunawardana CG, Smith CR, Batruch I, Diamandis EP. Mining the ovarian cancer ascites proteome for potential ovarian cancer biomarkers. *Mol Cell Proteomics* (2009) 8:661–9. doi: 10.1074/mcp.M800313-MCP200

30. Stadtman ER, Moskovitz J, Levine RL. Oxidation of methionine residues of proteins: biological consequences. *Antioxid Redox Sign* (2003) 5:577–82. doi: 10.1089/152308603770310239

31. Levine RL, Mosoni L, Berlett BS, Stadtman ER. Methionine residues as endogenous antioxidants in proteins. *Proc Natl Acad Sci U.S.A.* (1996) 93:15036–40. doi: 10.1073/pnas.93.26.15036

32. Kim H-Y. The methionine sulfoxide reduction system: selenium utilization and methionine sulfoxide reductase enzymes and their functions. *Antioxid Redox Sign* (2013) 19:958–69. doi: 10.1089/ars.2012.5081

33. Araos P, Vidal R, O'Shea E, Pedraz M, García-Marchena N, Serrano A, et al. Serotonin is the main tryptophan metabolite associated with psychiatric comorbidity in abstinent cocaine-addicted patients. *Sci Rep-uk* (2019) 9:16842. doi: 10.1038/s41598-019-53312-0

34. Duarte D, Amaro F, Silva I, Silva D, Fresco P, Oliveira JC, et al. Carbidopa alters tryptophan metabolism in breast cancer and melanoma cells leading to the formation of indole-3-Acetonitrile, a pro-proliferative metabolite. *Biomolecules* (2019) 9:409. doi: 10.3390/biom9090409

35. Opitz CA, Litzenburger UM, Sahm F, Ott M, Tritschler I, Trump S, et al. An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* (2011) 478:197–203. doi: 10.1038/nature10491

36. Silva MDFG, Dionísio AP, Abreu FAPD, Brito ES, Wurlitzer NJ, Silva LMAE, et al. Evaluation of nutritional and chemical composition of yacon syrup using 1H NMR and UPLC-ESI-Q-TOF-MSE. *Food Chem* (2018) 245:1239–47. doi: 10.1016/j.foodchem.2017.11.092

37. Wu G, Morris SM. Arginine metabolism: nitric oxide and beyond. *Biochem J* (1998) 336(Pt 1):1–17. doi: 10.1042/bj3360001

38. Fukumura D, Kashiwagi S, Jain RK. The role of nitric oxide in tumour progression. *Nat Rev Cancer* (2006) 6:521–34. doi: 10.1038/nrc1910

39. An Z, Zheng D, Wei D, Jiang D, Xing X, Liu C. Correlation between acylcarnitine and peripheral neuropathy in type 2 diabetes mellitus. *J Diabetes Res* (2022) 2022:8115173. doi: 10.1155/2022/8115173

40. Huang Y, Du Y, Zheng Y, Wen C, Zou H, Huang J, et al. Ct-OATP1B3 promotes high-grade serous ovarian cancer metastasis by regulation of fatty acid beta-oxidation and oxidative phosphorylation. *Cell Death Dis* (2022) 13:556. doi: 10.1038/s41419-022-05014-1

41. Heiserman JP, Nallanthighal S, Gifford CC, Graham K, Samarakoon R, Gao C, et al. Heat shock protein 27, a novel downstream target of collagen type XI alpha 1, synergizes with fatty acid oxidation to confer cisplatin resistance in ovarian cancer cells. *Cancers* (2021) 13:4855. doi: 10.3390/cancers13194855

42. Chen L-Y, Yang B, Zhou L, Ren F, Duan Z-P, Ma Y-J. Promotion of mitochondrial energy metabolism during hepatocyte apoptosis in a rat model of acute liver failure. *Mol Med Rep* (2015) 12:5035–41. doi: 10.3892/mmr.2015.4029

43. Zaugg K, Yao Y, Reilly PT, Kannan K, Kiarash R, Mason J, et al. Carnitine palmitoyltransferase 1C promotes cell survival and tumor growth under conditions of metabolic stress. *Gene Dev* (2011) 25:1041–51. doi: 10.1101/gad.1987211

44. Sidorkiewicz M. Hepatitis c virus uses host lipids to its own advantage. *Metabolites* (2021) 11:273. doi: 10.3390/metabo11050273

45. Liu X, Liu G, Chen L, Liu F, Zhang X, Liu D, et al. Untargeted metabolomic characterization of ovarian tumors. *Cancers* (2020) 12:3642. doi: 10.3390/cancers12123642

46. Baxmann AC, Ahmed MS, Marques NC, Menon VB, Pereira AB, Kirsztajn GM, et al. Influence of muscle mass and physical activity on serum and urinary creatinine and serum cystatin c. *Clin J Am Soc Nephrology: CJASN* (2008) 3:348–54. doi: 10.2215/CJN.02870707

47. Palviainen M, Laukkanen K, Tavukcuoglu Z, Velagapudi V, Kärkkäinen O, Hanhineva K, et al. Cancer alters the metabolic fingerprint of extracellular vesicles. *Cancers (Basel)* (2020) 12:3292. doi: 10.3390/cancers12113292

48. Kalhan SC, Marczewski SE. Methionine, homocysteine, one carbon metabolism and fetal growth. *Rev Endocr Metab Dis* (2012) 13:109–19. doi: 10.1007/s11154-012-9215-7

49. Cavuoto P, Fenech MF. A review of methionine dependency and the role of methionine restriction in cancer growth control and life-span extension. *Cancer Treat Rev* (2012) 38:726–36. doi: 10.1016/j.ctrv.2012.01.004

50. Narvik J, Vanaveski T, Innos J, Philips M-A, Ottas A, Haring L, et al. Metabolic profile associated with distinct behavioral coping strategies of 129Sv and Bl6 mice in repeated motility test. *Sci Rep* (2018) 8:3405. doi: 10.1038/s41598-018-21752-9

51. Armitage EG, Barbas C. Metabolomics in cancer biomarker discovery: current trends and future perspectives. *J Pharm BioMed Anal* (2014) 87:1–11. doi: 10.1016/j.jpba.2013.08.041

52. Zhao Y, Butler EB, Tan M. Targeting cellular metabolism to improve cancer therapeutics. *Cell Death Dis* (2013) 4:e532. doi: 10.1038/cddis.2013.60

53. Li J, Condello S, Thomes-Pepin J, Ma X, Xia Y, Hurley TD, et al. Lipid desaturation is a metabolic marker and therapeutic target of ovarian cancer stem cells. *Cell Stem Cell* (2017) 20:303–314.e5. doi: 10.1016/j.stem.2016.11.004

54. Zhou Y, Peng J, Li P, Du H, Li Y, Li Y, et al. Discovery of novel indoleamine 2,3-dioxygenase 1 (IDO1) inhibitors by virtual screening. *Comput Biol Chem* (2019) 78:306–16. doi: 10.1016/j.compbiolchem.2018.11.024

55. Sekhar RV. GlyNAC (Glycine and n-acetylcysteine) supplementation improves impaired mitochondrial fuel oxidation and lowers insulin resistance in patients with type 2 diabetes: Results of a pilot study. *Antioxidants (Basel Switzerland)* (2022) 11:154. doi: 10.3390/antiox11010154

*CORRESPONDENCE
Ruiming Ou,
✉ ouruiming@126.com
Shuang Liu,
✉ liush@gd2h.org.cn

†These authors have contributed equally
to this work

A prognostic model based on prognosis-related ferroptosis genes for patients with acute myeloid leukemia

Feima Wu[1†], Guosheng Xu[2†], Guangchao Li[1†], Zhao Yin[1], Huijuan Shen[1], Kaiheng Ye[1], Yangmin Zhu[1], Qing Zhang[1], Ruiming Ou[1]* and Shuang Liu[1]*

[1]Department of Hematology, Guangdong Second Provincial General Hospital, Guangzhou, Guangdong, China, [2]Department of Blood Transfusion, Guangdong Second Provincial General Hospital, Guangzhou, Guangdong, China

Background: Acute myeloid leukemia (AML) is a heterogeneous disorder with an unpredictable prognosis. Ferroptosis, the iron-dependent cell death program, could serve as an alternative for overcoming drug resistance. However, its effect on AML remains largely unclear.

Methods: We collected RNA sequencing data and relevant clinical information of AML patients from The Cancer Genome Atlas to construct a prognosis prediction model. Risk score was calculated with eight prognosis-related ferroptosis genes (PRFGs) discovered through univariate analysis and Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression. A nomogram was constructed by incorporating LASSO risk score, age, and cytogenetic risk based on univariate/ multivariate Cox regression.

Results: Of the 33 AML PRFGs identified from the TCGA-derived dataset, 8 genes were used to construct a gene signature to predict AML prognosis. Principal component analysis and heatmap showed significant differences between the low and high risk score groups. Next, LASSO risk score, age, and cytogenetic risk were incorporated into the nomogram to predict the overall survival (OS) of AML patients. According to survival analysis, patients with a low risk score had markedly increased OS as compared to those with a high risk score. Based on the results of Gene Ontology and Kyoto Encyclopedia of Genes and Genomes, the differences between the two risk groups showed a close relationship with immune-related pathways and membrane transportation. The analysis of tumor-infiltrating immune cells and immune checkpoints revealed that the immunosuppressive tumor microenvironment possibly facilitated different prognostic outcomes between the two groups. Gene expression analyses showed that the mRNA expression levels of PARP1 and PARP3 (PARPs) were closely related to the different clinical subgroups and the analyzed OS in AML patients. Finally, the PARP inhibitor talazoparib and the ferroptosis inducer erastin exerted a synergistic anti-proliferative effect on AML cells.

Conclusion: We constructed a nomogram by incorporating PRFGs, and the constructed nomogram showed a good performance in AML patient stratification and prognosis prediction. The combination of PARP inhibitors with ferroptosis inducers could be a novel treatment strategy for treating AML patients.

Introduction

Acute myeloid leukemia (AML) is a heterogeneous hematologic tumor with the features of abnormal growth of myeloblasts or pro-granulocytes without physical differentiation. Over the past 10 years, tremendous progress has been made in the development of approaches for treating AML, such as hematopoietic stem cell transplantation; inhibitors targeting MCL-1, IDH2, and NPM1/FLT3-ITD mutations, epigenetic agents, antibody-based treatments, and cellular therapies. Consequently, the survival of patients with AML has significantly improved (Newell and Cook, 2021). Between 2017 and 2019, nine new drugs were approved by the FDA for AML patients. Despite these advances, older patients with AML show poor prognostic outcomes, with a long-term survival rate of <15% (Short et al., 2018). Because AML is highly heterogeneous, the treatment of this disease is challenging due to varying clinical features, including abnormal genetic and cytogenetic characteristics, and isolated factors such as coexisting diseases and physical conditions (Dohner et al., 2017). Hence, a prognosis prediction model that incorporates these well-recognized factors is important to precisely stratify the pretreatment risk and to implement clinical treatment decision-making.

Ferroptosis is a specific cell death program triggered by iron-dependent phospholipid peroxidation, and it is regulated by different cellular metabolic pathways such as iron metabolism, redox homeostasis, amino acid/lipid/sugar metabolism, and mitochondrial activity (Stockwell et al., 2017; Jiang et al., 2021). Recently, ferroptosis has received increasing attention, and considerable progress has been achieved in developing drugs targeting the regulatory mechanisms of ferroptosis in cancer cells (Lei et al., 2022). Several studies have investigated the involvement of ferroptosis in AML. Yu et al. (2015) showed that the ferroptosis inducer erastin not only inhibits AML cell growth but also enhances their sensitivity to chemotherapeutic drugs. Du et al. (2019) revealed that dihydroartemisinin specifically induced AML cells ferroptosis by modulating the activation of the AMPK/mTOR/p70S6k autophagy pathway activation. Yusuf et al. (2020) found that leukemic cells, rather than healthy myeloid cells, were dependent on the aldehyde dehydrogenase 3a2 enzyme for oxidizing long-chain aliphatic aldehydes to prevent cellular oxidative injury and synthetic lethality of ferroptosis inducers. Moreover, glutathione peroxidase-4 (GPX4) (Liu et al., 2023), TP53 (Cui et al., 2022), reactive oxygen species (ROS) metabolism (Du et al., 2020), and glutathione (GSH) metabolism (Wei et al., 2020) were found to be closely associated with the ferroptosis process in AML cells and with prognostic outcomes in AML

patients. Additionally, numerous ferroptosis-related genes (FRGs) have been discovered, but these genes have shown inconsistent functions. Recently, the association of FRGs with the prognostic outcome of AML patients has been investigated. However, a prognosis prediction model incorporating the prognosis-related FRGs together with the clinical characteristics of AML patients is yet to be established. In the present study, machine learning was used to assess data from AML patients collected from public databases. We then established a prognostic signature by incorporating prognosis-related FRGs and used this signature to generate a prognosis prediction model for AML patients.

Materials and methods

Cell lines

AML cell lines MOLM-13, U937 and KG-1a were acquired from American Type Culture Collection (ATCC). Cells were cultured at 37°C and 5% CO_2 in humidified incubator. Culture medium for MOLM-13 and U937 consisted of RPMI-1640 with 10% v/v fetal bovine serum and 1% v/v penicillin/streptomycin, while medium for KG-1a was IMDM with 20% v/v fetal bovine serum and 1% v/v penicillin/streptomycin.

Data collection

RNA sequencing data and the related clinical information of AML patients were obtained from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/repository/). Expression profiles and clinical information of AML patients from two datasets (GSE71014 and GSE37642) were obtained from the Gene Expression Omnibus (GEO) database (https://www.genecards.org/).

PRFG screening

Prognosis-related genes (PRGs) of AML patients were analyzed with "survival" package in R software (version 4.2.1). PRGs were identified based on the following criteria: hazard ratio (95% CI) > 1.0 and p-value ≤ 0.05. The "limma" package was used to analyze differentially expressed genes (DEGs). The threshold was set as log2 fold change > 1 and false discovery rate < 0.05. FRGs were collected based on FerrDB (http://www.zhounan.org/ferrdb/current/). The intersection genes of PRGs, DEGs, and FRGs were defined as PRFGs.

Protein-protein interaction (PPI) network construction

PPI network of PRFGs was constructed with STRING web tool (https://string-db.org/) and visualized with Cytoscape (version 3.9.1). The parameters were set as follows: network type: full STRING network, meaning of network edges: confidence, active interaction sources: experiments, text mining, databases, co-expression, neighborhood, gene fusion, co-occurrence, minimum required interaction score = 0.4, max number of interactors to show: Query proteins only. The node scores of the PPI network were calculated with cytoHubba module of Cytoscape. The top 10 nodes rank by node score were defined as hub genes.

Establishment and verification of the PFRG-based prediction model for AML patients

The "glmnet" package in R software was used to establish a PRFG signature by least absolute shrinkage and selection operator (LASSO)-penalized Cox regression (Liu et al., 2018). The lowest partial likelihood of deviance was used to determine the model penalty parameter (λ). The regression coefficient (β) of the LASSO model was linearly combined with gene expression to determine the prognostic risk score. AML patients were assigned to the high- and low-risk score groups according to the threshold. Principal component analysis (PCA) was performed using the "prcomp" function in R software according to the risk scores of the identified genes. The effect of the prognostic PRFG-based signature on prediction was analyzed through Kaplan–Meier survival analysis and time-dependent receiver operating characteristic (ROC) curves.

Construction of the prognostic nomogram for AML patients

Based on the univariate/multivariate Cox regression analysis of the clinical features of AML patients, a nomogram was constructed using the R packages "survival" and "rms." Based on the median nomogram risk scores, AML patients were assigned to the high or low-risk score groups. The accuracy of the nomogram was evaluated based on ROC curves and the concordance index (C-index).

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses of DEGs between the low- and high-risk score groups were conducted with "clusterProfiler" and "ggplot2" packages in R software (Subramanian et al., 2005; Yu et al., 2012). $p <$ 0.05 was considered the significance level for the enriched pathways.

Analysis of immune profiles

To analyze the immune status of each sample, we used Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts (CIBERSORT) to calculate 22 tumor-infiltrating immune cell (TIIC) proportions in AML patients (Newman et al., 2015; Becht et al., 2016). We also used CIBERSORT to convert mRNA data to tumor-infiltrating non-cancer cell proportions in the tumor microenvironment with standard annotation files for organizing gene expression profiles. The list of immune checkpoint molecules was derived from (Fang et al., 2022). By querying the PubMed, thirty-two molecules out of the forty-seven immune checkpoint list that related to AML were further analyzed.

CCK-8 assay and treatment combination analysis

Cell viability was determined by cell counting kit-8 (CCK-8) assay (Beyotime Technology, China). After the cells culturing for 48 h, 20 μL CCK-8 solution was added to each well, and the absorbance (OD) value was measured at 450 nm. The concentrations of erastin used in this study were set as around 40% inhibitory activity of AML cell lines. For synergize assay, the concentrations of talazoparib were set as serial concentration less than EC50 of AML cell lines. The combination effect of talazoparib and erastin at indicated concentration, combination index (CI), fraction affected (FA) levels were calculated by CompuSyn software using Chou-Talalay method with constant-ratio combinations. CI values less than 1, equal to 1, greater than 1 indicate synergistic, additive, or antagonistic effects, respectively.

Cell migration assay

A total of 1×10^5 cells were resuspended with 200 μL serum-free medium and seeded into the Transwell chamber (8 μm in diameter, Corning, United States), then the chambers were insert into a well with 500 μL culture medium containing corresponding concentrations of drugs. The plate was placed for 72 h incubation at 37°C. The migrated cells from the chambers were imaged with a microscope and the number was calculated.

Statistical analysis

Continuous variables that exhibited a normal distribution were presented as the mean ± standard deviation, and comparisons between groups were examined using Student's t-test. Kaplan–Meier survival analysis was used to estimate overall survival and Cox regression was used to compare survival differences between patient groups. Survival analysis was carried out with "survminer" and "survival" R packages. R software (Version 4.2.1) and GraphPad Prism 9 were adopted for data analysis. p-value < 0.05 stood for statistically significant.

Results

Discovery of PRFGs

We obtained 3436 DEGs and 1613 PRGs by comparing dead and alive AML patients derived from the TCGA database. Heatmap showed a total of 78 significant differential expressed FRGs in alive and dead patients and they were further divided into up- and

FIGURE 1
Prognosis-related gene (PRG) analysis in AML patients. **(A)** Heatmap of the differential expressed PRGs in AML patients. **(B)** Venn plots showing prognosis-related ferroptosis genes (PRFGs). **(C)** PPI network diagram of PRFGs.
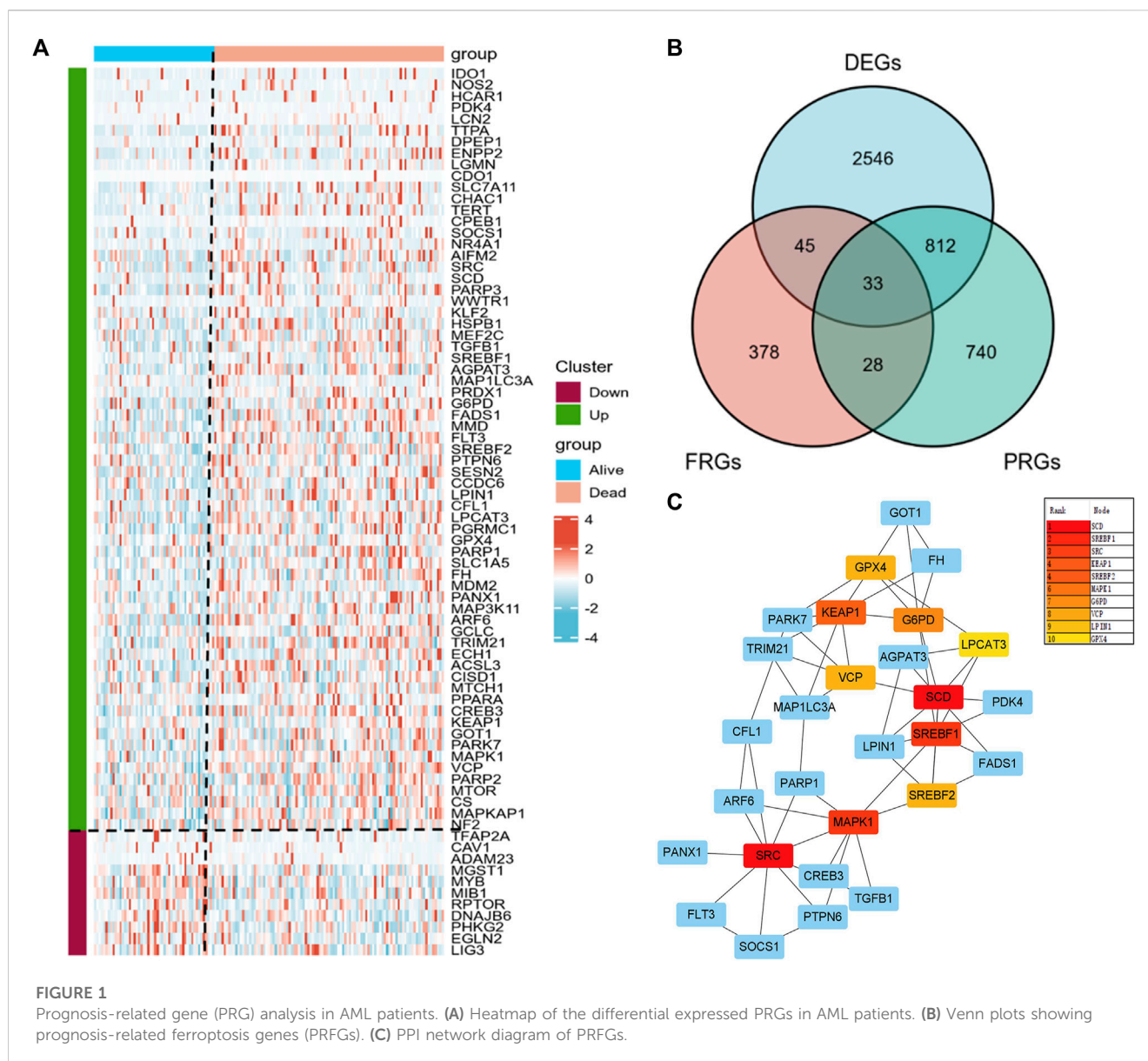
downregulated groups according to their log2 fold-change (Figures 1A, B). By intersecting DEGs, PRGs, and FRGs, we obtained 33 PRFGs (Figure 1B). We then constructed protein–protein interaction (PPI) networks based on the STRING database to analyze and predict protein interactions and protein functional connectivity. Rank by node score of the PPI network, as shown in Figure 1C, SCD, SREBF1, SRC, SREBF2, KEAP1, etc. were identified as the hub genes.

Establishment and verification of the prognostic gene signature

We incorporated the 33 PRFGs in LASSO Cox regression (Figure 2A) and constructed the 8-gene signature according to the optimum λ value (Figure 2B). These 8 genes included *SOCS1, PARP1, TGFB1, AGAPT3, PARP3, FH, ARF6,* and *CREB3.* To clarify the association of the selected genes with patient survival, univariate Cox

regression was performed (Figure 2C). According to the β-value of every gene discovered based on LASSO Cox regression, the prognostic risk score was calculated as follows: (0.0893* SOCS1 expression) + (0.0815 * PARP1 expression) + (0.0014 * TGFB1 expression) + (0.00159 * AGPAT3 expression) + (0.0065 * PARP3 expression) + (0.0217 * FH expression) + (0.0067 * ARF6 expression) + (0.1914 * CREB3 expression). Based on the median risk score, we assigned the patients to the high-risk score group ($n = 74$) or the low-risk score group ($n = 56$). Furthermore, based on the PCA results, patients in both subgroups showed a discrete distribution (Figure 2D). The expression levels of the selected genes also showed a significant difference between both groups (Figure 2E; Supplementary Figure S3). Moreover, Kaplan–Meier survival analysis indicated a significantly increased OS in the low-risk score group as compared to that in the high-risk score group (Figure 2F). Furthermore, the area under the ROC curve (AUC) values for 1-, 2-, and 3-year OS were 0.867, 0.855, and 0.810, respectively, which indicated good predictive performances of LASSO analysis (Figures 2G–I).
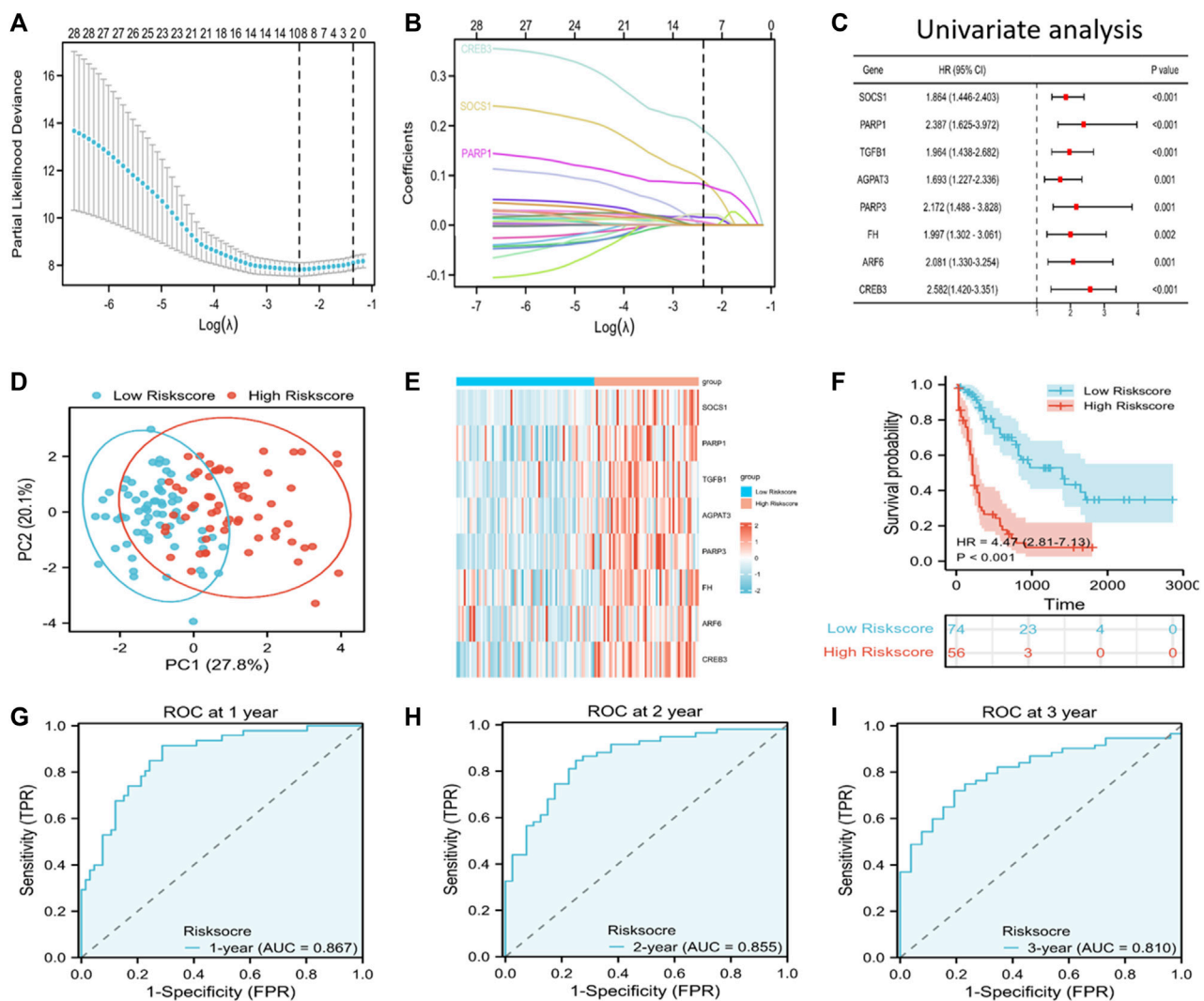
FIGURE 2
Discovery of prognosis-related ferroptosis genes (PRFGs) for establishing the prognosis prediction signature for AML patients derived from the TCGA database. **(A)** LASSO Cox regression of the PRFGs. **(B)** LASSO coefficients of the PRFGs. **(C)** Univariate Cox regression confirmed the relationship between PRFGs and the prognostic outcome for AML patients. **(D)** PCA plot showing AML cases based on the expression levels of the signature genes in both risk score groups. **(E)** Heatmap showing the mRNA levels of the eight chosen PRFGs in the low-risk score and high-risk score groups. **(F)** Kaplan−Meier survival curves suggest increased OS in the low-risk score group as compared to that in the high-risk score group. **(G−I)** ROC curves of the LASSO model to predict 1-, 2- and 3-year OS of AML patients.

Next, GEO-derived data were analyzed for model validation. The AML patients from the GEO cohort were classified into the high or low-risk score group according to the risk score. A significant difference was noted between both subgroups, with a markedly increased survival rate in the low-risk score group compared to that in the high-risk score group ($p < 0.001$) (Supplementary Figures S1A–F, S2A–F).

Risk score independently predicts the prognosis of AML

To predict the OS of AML patients, the risk score and clinical features, including age, gender, BM blast percentage, FLT3 mutation, and cytogenetic risk were incorporated into the univariate/multivariate Cox regression. The results revealed that the risk score independently predicted

patient survival. Furthermore, age, cytogenetic risk (intermediate), and cytogenetic risk (poor) also independently predicted the prognosis of $p \leq 0.05$ (Figures 3A, B). All patients were then classified according to age, cytogenetic risk (intermediate), and cytogenetic risk (poor). Based on the Kaplan–Meier survival curves, the low-risk score group showed increased survival as compared to the high-risk score group (Figures 3C–F).

Establishment and verification of the prognosis prediction nomogram

A nomogram that can visually represent the prognosis prediction model was constructed by incorporating age, cytogenetic risk, and the risk score for illustrating patient survival (Figure 4A). The nomogram showed that the Lasso risk score most significantly affected 1-, 2-, and

FIGURE 3
Independent prognostic factors of risk scores and clinical features. **(A,B)** Univariate/multivariate Cox regression confirmed that age, cytogenetic risk (poor), and LASSO risk score could independently predict prognosis. **(C−F)** Kaplan−Meier survival curves suggested that OS was associated with age < 60 **(C)**, age ≥ 60 **(D)**, cytogenetic risk score (intermediate) **(E)**, and cytogenetic risk score (poor) **(F)**.

3-year survival of AML patients, followed by cytogenetic risk and age. ROC curves and C-index were then used to evaluate the discrimination performance of the nomogram. The C-index value for predicting 1-, 2-, and 3-year patient OS was 0.785 (0.762–0.808). The area under the ROC curve (AUC) values for 1-, 2-, and 3-year OS were 0.872, 0.891, and 0.863, respectively, which were superior to those of the Lasso risk score model (Figure 4B). The model calibration performance was analyzed with a calibration curve. We found that our predicted results were consistent with the observed results (Figure 4C). The patients were assigned to high- or low-risk score groups in accordance with the median risk score value. Compared to the high-risk AML patients, low-risk AML patients showed markedly distinct dispersion direction with superior OS ($p < 0.001$) (Figures 4D, E).
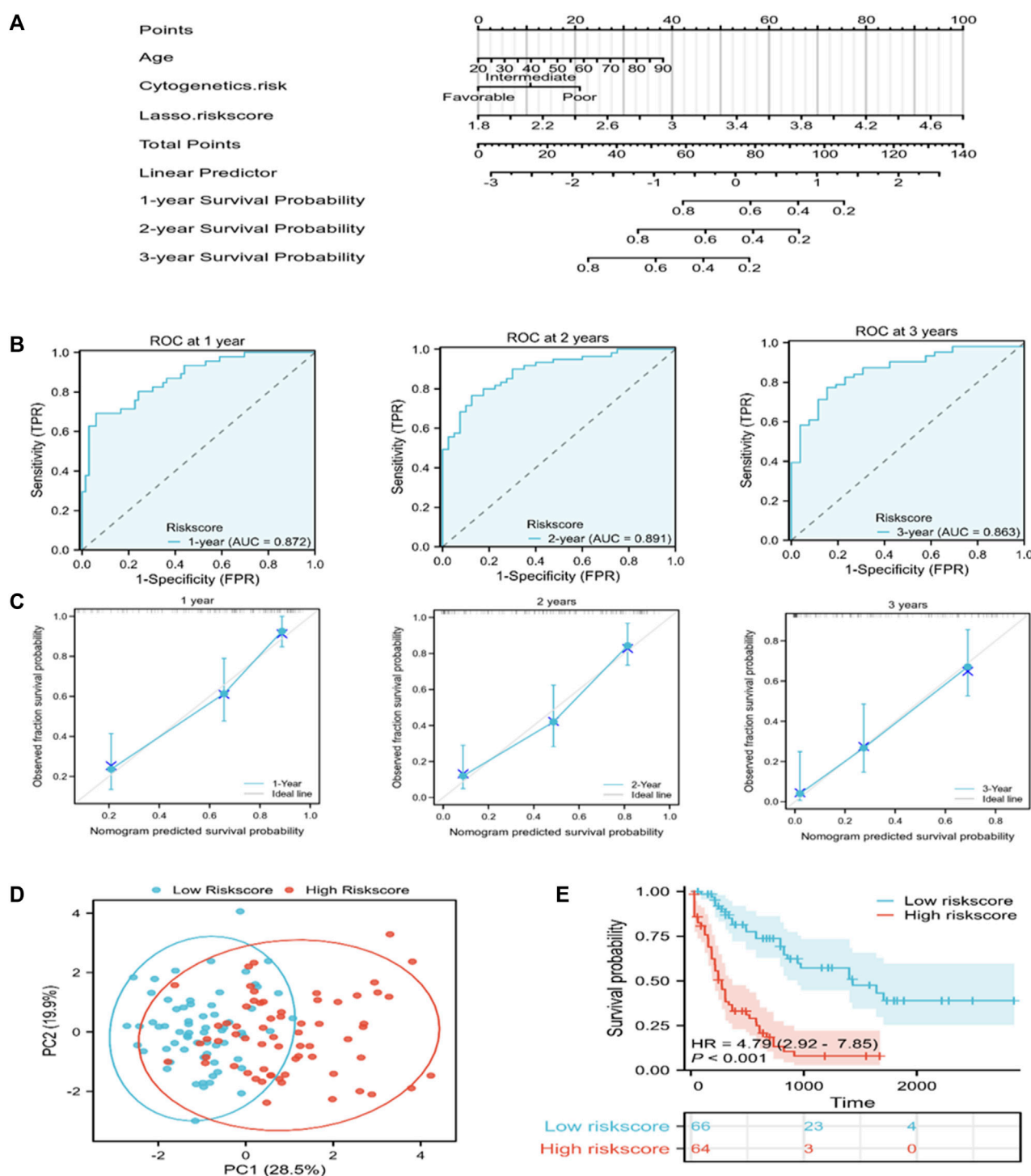
FIGURE 4
A nomogram for predicting the prognosis of TCGA-derived AML patients. **(A)** A nomogram was constructed to predict 1-, 2-, and 3-year OS in AML patients. **(B)** ROC curves for the nomogram for predicting 1-, 2-, and 3-year survival of AML patients. **(C)** Nomogram calibration curves showing survival probabilities at 1-, 2-, and 3-year. **(D)** PCA plot for AML cases according to the mRNA levels of the signature genes in both risk groups. **(E)** Kaplan-Meier survival curves suggest increased OS of patients in the low-risk score group as compared to that of patients in the high-risk score group.

Functional annotation

To elucidate the pathways related to the nomogram risk score, KEGG pathway enrichment analysis was performed for DEGs in both risk score groups. We observed that DEGs were mostly associated with "neuroactive ligand-receptor interaction," "cytokine-cytokine receptor interaction," "PI3K-Akt pathway," and "Phagosome" (Figure 5A). GO functional enrichment

FIGURE 5
GO and KEGG analyses for DEGs in both risk score groups. **(A)** KEGG analysis of DEGs in both risk score groups. **(B)** GO annotation of DEGs in both risk score groups. **(C)** Bubble plots of top 3 GO and KEGG enrichment terms in both risk score groups. **(D)** Chord diagram showing the related genes of the GO term 0050900 (leukocyte migration) and the KEGG term hsa0460 (cytokine–receptor interaction).

analyses showed that the biological process (BP) terms were mostly "leukocyte migration," "regionalization," and "regulation of cell-cell adhesion" (Figures 5B–D). The cellular component (CC) terms were mainly "collagen-containing extracellular matrix," "synaptic membrane," and "external side of plasma membrane" (Figures 5B–D). The molecular functions (MF) terms were "receptor ligand activity," "passive transmembrane transporter," and "channel activity" (Figures 5B–D).

Immune status of AML patients based on the nomogram risk score groups

Because the enrichment analysis highlighted immune-related terms like "cytokine-cytokine receptor interaction" (hsa04060) and "leukocyte migration" (GO005090) (Figures 5A, B, D), we examined the correlation between the nomogram and TIICs. The differences

of 22 types of TIICs in AML patients between the two risk score groups were assessed by the CIBERSORT algorithm. Figure 6A shows the similarities and differences in immune cell infiltration between the AML subgroups. "T cells CD4+ memory resting," "Mast cell resting," and "Monocyte" exhibited significant differences between both risk score groups (Figure 6B). Immune checkpoint molecules are the indicators for prognosis and serve as immunotherapeutic targets for AML patients. The expression levels of *LAIR1, LAG3, CTLA4, CD200R1, CD276, KIR3DL1, CD80, PDCD1, LGALS9, TNFSF14, PDCD1LG2, CD86, and CD274* in high-risk patients were markedly increased as compared to those in low-risk patients (Figure 6C). We further evaluated the association between immune checkpoint molecules and OS in AML patients. Higher expression levels of *LAIR1, CD276, LGALS9, PDCD1, PDCD1L2G,* and *TNFSF14* levels were correlated with poor OS (Supplementary Figure S4). Taken together, these data suggest that the poor prognostic outcome in the high-risk score

FIGURE 6
Immune profiles between the nomogram risk groups. **(A)** Heatmap showing immune cell expression between the risk score groups. **(B)** Comparison of diverse immune cell subtypes of both risk score groups. **(C)** Expression levels of immune checkpoint molecules in both risk score groups. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; and ****$p < 0.0001$.
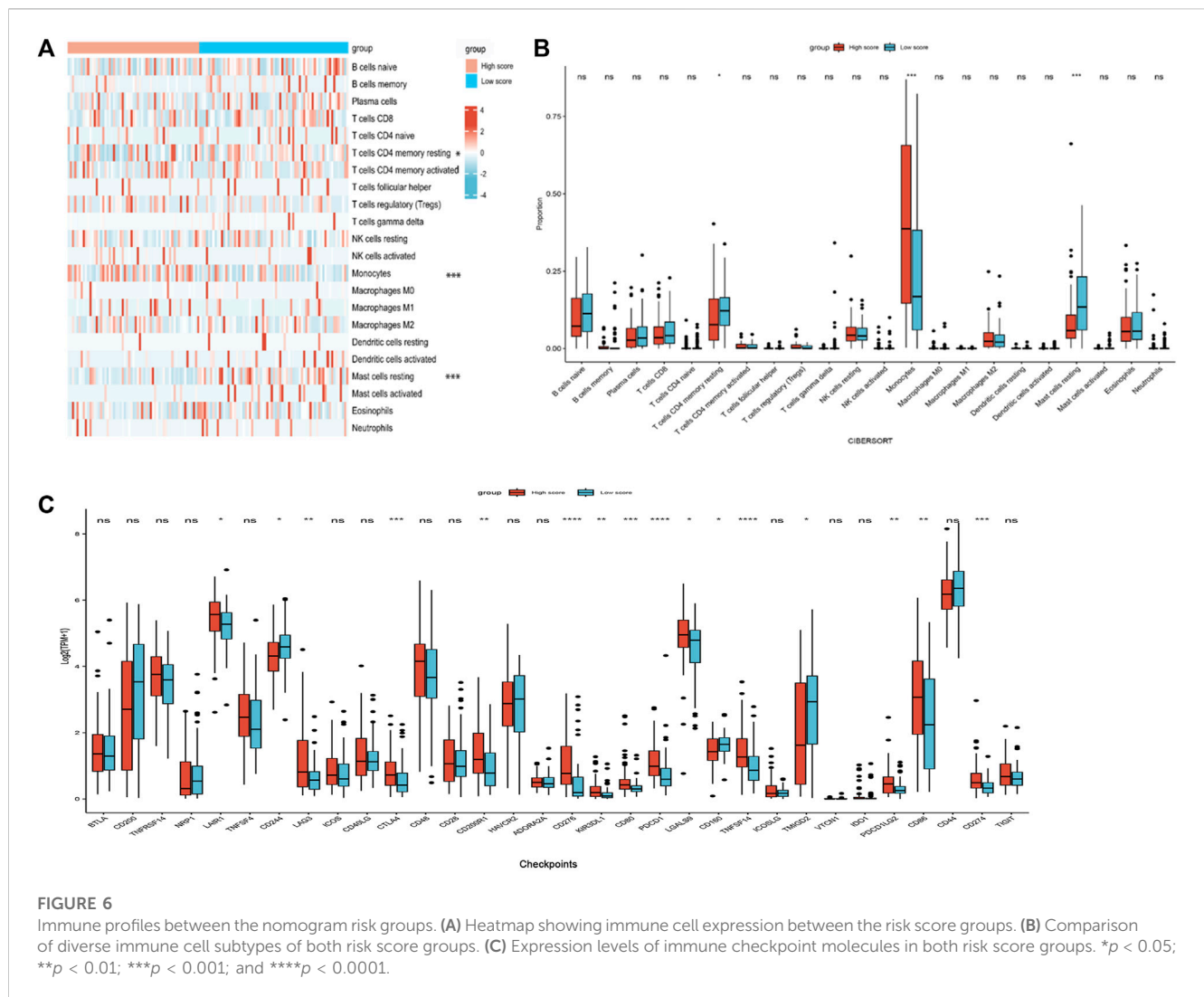
group may be partly associated with the tumor immune microenvironment.

Correlation between the expression levels of the risk-associated genes and clinicopathological subgroups of AML patients
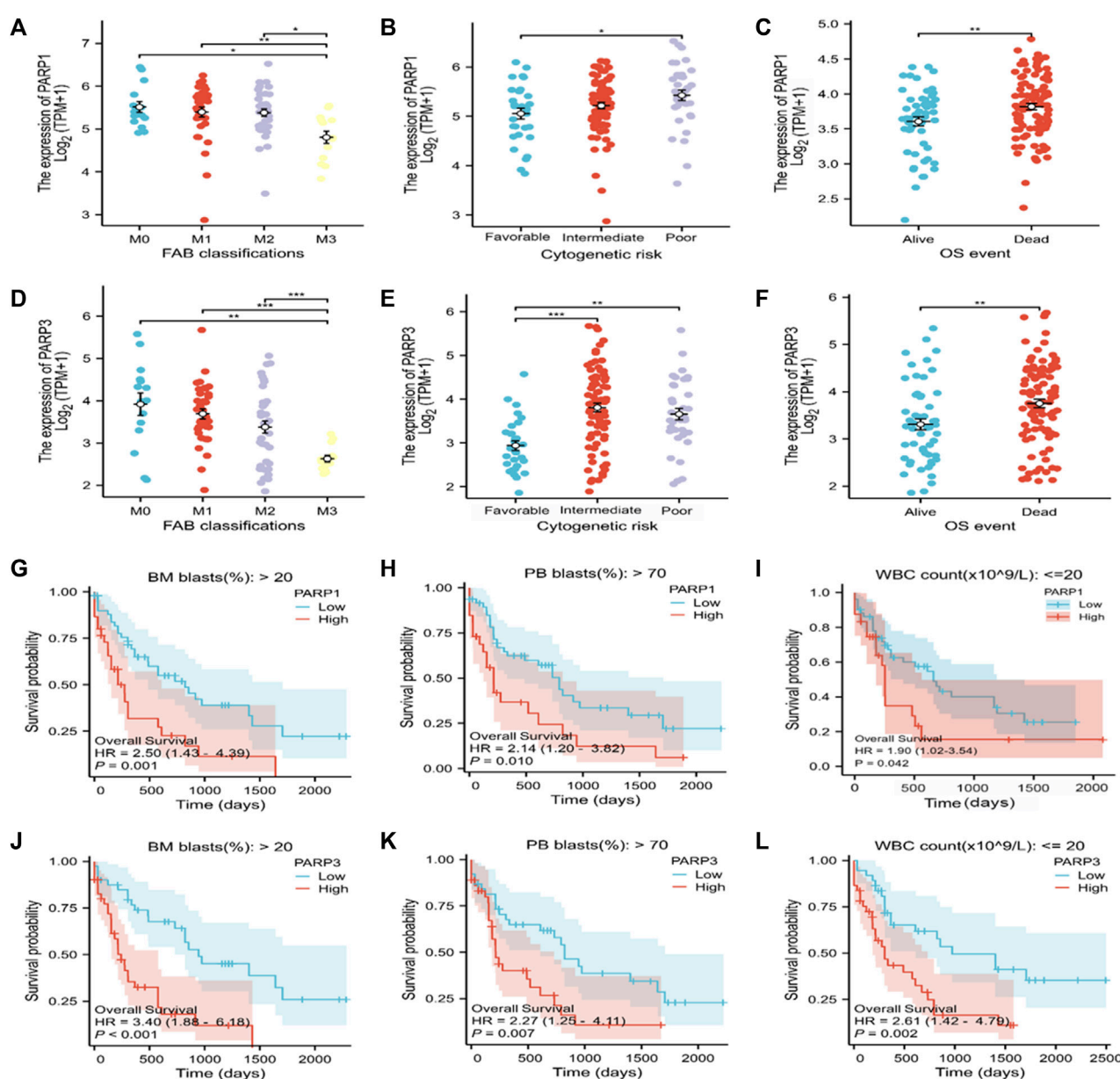
PARPs are potential therapeutic targets for AML. However, the correlation between PARPs expression levels and AML clinicopathologic features, as well as the prognosis of patients with specific clinical variables remained to be explored. PARP1 and PARP3 were selected for further investigations as they showed significantly higher expression levels in the high Lasso risk score groups in all the analyzed datasets (Figure 2E; Supplementary Figures S1B, S2B, S3). According to the subgroup classification, the association between the clinicopathologic features of TCGA AML patients and the expression level of PARPs was analyzed. As shown in Figures 7A–F, the expression levels of PARP1 and PARP3 were significantly associated with the clinicopathologic features of FAB classification, cytogenetic risk, and OS event. Moreover, a higher

expression level of PARP1 or PARP3 also predicted poor prognosis in the clinical subgroups of BM (bone marrow) blasts (%) > 20, PB (peripheral blood) blasts (%) > 70, and WBC (white blood cell) count ($\times 10^9$/L) ≤ 20 (Figures 7G–L).

The PARP1 inhibitor talazoparib shows a synergistic effect with the ferroptosis inducer erastin on AML cells

Antitumor drug combinations can effectively prevent resistance and provide novel treatments. Previous studies have shown that the PARP inhibitor shows a synergistic effect with ferroptosis inducers on BRCA-proficient ovarian cancer. We tested whether this drug synergy also affected the survival of AML cells. Consistent with the results of previous studies, talazoparib inhibited the growth of AML cells (Figures 8A–C). Furthermore, every combination index at specific talazoparib and erastin doses was <1 for MOLM-13, U937, and KG-1a cells (Figures 8D–F). Talazoparib also showed a synergistic effect with erastin to inhibit AML cell migration (Figures 8G–I; Supplementary Figure S5). These findings suggest that the PARP1 inhibitor shows a synergistic effect with the
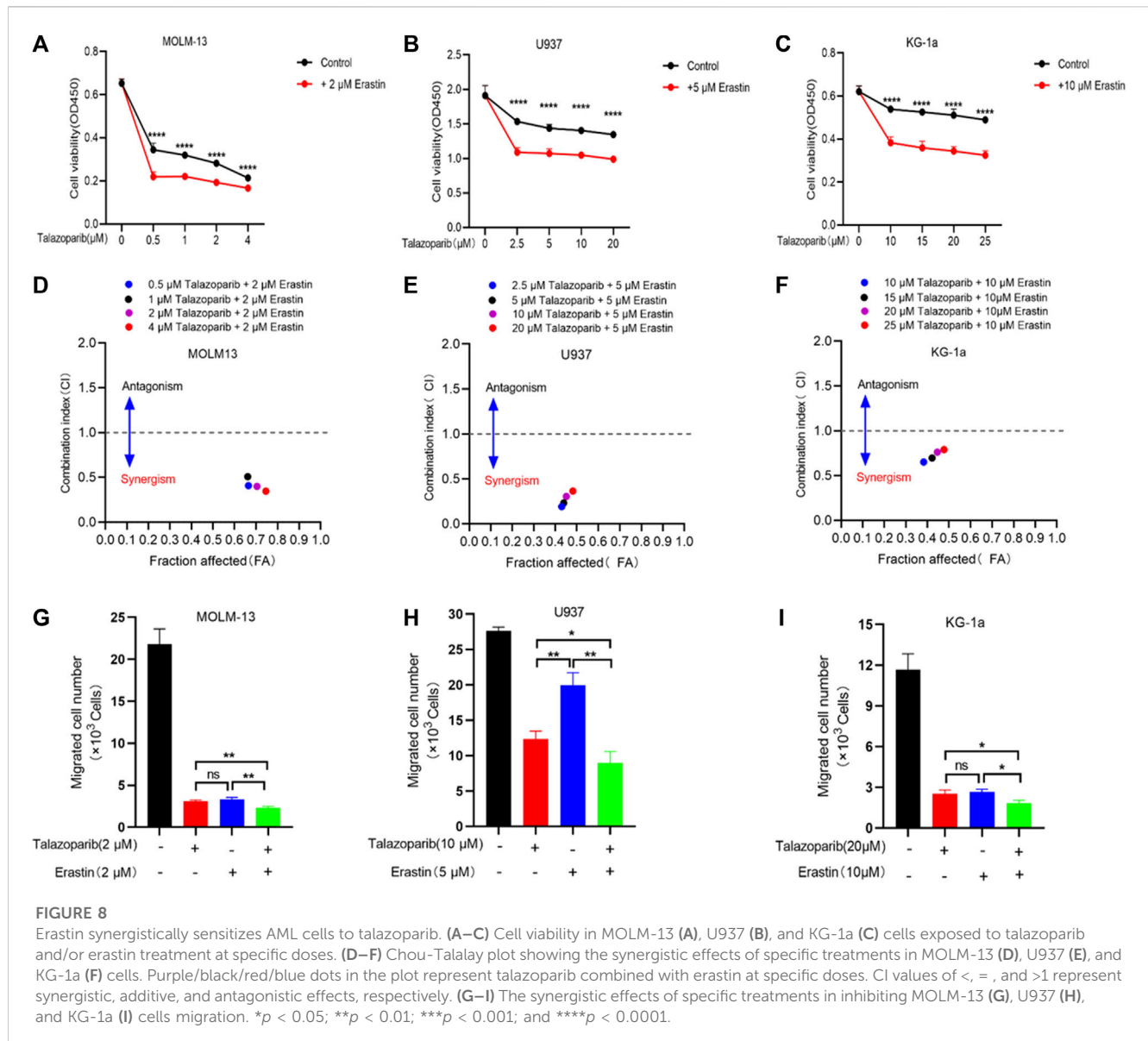
FIGURE 7
Association of the expression of PARPs with different clinical features of AML. **(A–C)** Correlation between PARP1 expression and FAB classification **(A)**, cytogenetic risk **(B)**, and OS event **(C)**. **(D–F)** Correlation between PARP3 expression and FAB classification **(D)**, cytogenetic risk **(E)**, and OS event **(F)**. **(G–I)** Correlation between PARP1 expression level and BM blasts (20%) > 20% clinical subgroup, PB blasts (20%) > 70% clinical subgroup, and WBC count ($\times 10^9$/L) ≤ 20. **(J–L)** Correlation between PARP3 expression level and BM blasts (20%) > 20% clinical subgroup, PB blast (20%) > 70% clinical subgroup, WBC count ($\times 10^9$/L): ≤20. *$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$. OS, overall survival; BM, bone marrow; WBC, white blood cell; PB, peripheral blood.

ferroptosis inducer erastin for inhibiting the growth and migration of AML cells.

Discussion

In the present study, to predict the survival of AML patients, a prognosis prediction model was constructed based on the combination of 8 prognosis-related FRGs and clinical characteristics. The established model exhibited favorable calibration and discrimination performance for predicting patient survival. The association of the model with TIICs

and immune checkpoint molecules was also partially investigated. The PARP inhibitor talazoparib showed a synergistic effect with the ferroptosis inducer erastin to enhance anti-proliferation efficacy for AML cells. AML has the highest occurrence frequency among acute leukemias during adulthood. Presently, cytogenetic markers play a critical role in stratifying the associated risk and treatment of AML patients. Although several studies have been conducted to find appropriate prognostic biomarkers, AML remains the disease with substantially different prognostic outcomes. The 5-year OS rate of AML is <50%, and the 2-year survival rate is only 20% among old people with AML (Gregory et al., 2009; Riva et al., 2012). Ferroptosis is a novel cell death

FIGURE 8
Erastin synergistically sensitizes AML cells to talazoparib. **(A–C)** Cell viability in MOLM-13 **(A)**, U937 **(B)**, and KG-1a **(C)** cells exposed to talazoparib and/or erastin treatment at specific doses. **(D–F)** Chou-Talalay plot showing the synergistic effects of specific treatments in MOLM-13 **(D)**, U937 **(E)**, and KG-1a **(F)** cells. Purple/black/red/blue dots in the plot represent talazoparib combined with erastin at specific doses. CI values of <, = , and >1 represent synergistic, additive, and antagonistic effects, respectively. **(G–I)** The synergistic effects of specific treatments in inhibiting MOLM-13 **(G)**, U937 **(H)**, and KG-1a **(I)** cells migration. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; and ****$p < 0.0001$.

program (Dixon et al., 2012), which is closely associated with AML (Yin et al., 2022). Although some ferroptosis-related prediction models for AML have been reported (Cui et al., 2022; Jinghua Wang et al., 2022; Kai Zhu et al., 2022; Yu et al., 2022), a model that incorporates well-recognized factors is highly important to precisely stratify AML patients.

By analyzing DEGs in AML patients, 33 prognosis-related ferroptosis DEGs were obtained based on 3436 DEGs in dead and alive patients. We chose eight genes following LASSO Cox regression and univariate analysis. Most of these genes were verified or predicted to be closely associated with cancers. The increased SOCS1 level in the bone marrow of AML patients was closely associated with advanced age, mutations in *FLT3-ITD*, *NPM1*, and *DNMT3A*, and SOCS1 overexpression in zebrafish mimic leukemia phenotype (Hou et al., 2017). PARPs, including PARP1 and PARP3, can mediate the early stage in DNA damage response. The inhibition of these proteins shows varying degrees of antitumor activity in AML, which mainly depend on the rearrangement of the genes (Padella et al., 2022). TGFB1 induces

ALDH+ stem cell-like phenotype in AML cells and contributes to leukemogenesis and chemotherapy resistance (Yuan et al., 2020). ARF6 belongs to the small GTPase ADP-ribosylation factor (Arf) family, and the upregulation and activation of ARF6 are markedly associated with the migration and invasion of multiple cancers (Li et al., 2017). CREB3 has been identified as an HDAC3-interacting protein that enhances NF-κB activation and promotes the migration of breast cancer cells (Kim et al., 2010).

The clinical features of AML patients are closely associated with their prognostic outcomes. To optimize the model and improve its survival prediction performance, we selected age and cytogenetic risk based on univariate/multivariate Cox regression. Age and cytogenetic risk are significantly associated with AML performance status, multidrug resistance, and prognosis outcome (Appelbaum et al., 2006; Dohner et al., 2017); hence, we incorporated these two clinical features and a gene signature to construct the nomogram. Additionally, the ROC curve and C-index were used to evaluate the discrimination performance of the

nomogram. The C-index for 1-year, 2-year, and 3-year OS of TCGA-derived AML patients was 0.785 (0.762–0.808), while the AUC values were 0.872, 0.891, and 0.863, respectively. These findings suggest the favorable discrimination performance of the model for survival prediction. Furthermore, according to the calibration curve, the constructed nomogram demonstrated good calibration. Based on the nomogram risk score, the patients were assigned to the high-risk score group or the low-risk score group. Functional annotation of the DEGs in both risk groups revealed that the enriched functional terms mainly involved leukocyte cell migration, T cell activation, and transmembrane transportation. We postulated that this finding might be correlated with the tumor-infiltrating cells in both subgroups. Hence, the proportions of immune cell in both groups were determined. CD4[+] memory resting T cells and resting mast cells showed an evidently higher proportion in the low-risk group, while monocytes showed a higher abundance in the high-risk score group. These findings suggest the association of the poor prognostic outcome of AML patients with immune cell infiltration. FRGs also possibly affect cancer cells through immune cells. We compared the expression levels of the immune checkpoint molecules such as *LAIR1*, *LAG3*, *CTLA4*, *CD200R1*, *CD276*, *KIR3DL*, *CD80*, *PDCD1*, *LGALS9*, TNFSF14, *PDCD1L2G*, CD86, and *CD274* in high-risk and low-risk AML patients. AML patients showing a higher expression level of *LAIR1*, *CD276*, *LGALS9*, *PDCD1*, *PDCD1L2G*, and *TNFSF14* experienced worse prognosis; thus, indicating that targeting these immune checkpoints may be beneficial for treating high-risk AML patients.

PARP1 and PARP3 belong to the Ploy (ADP-ribose) polymerase superfamily and are associated with DNA damage response (DDR). Based on our computational results, PARP1 and PARP3 were found to be associate with certain clinicopathologic features and overall survival of clinical subgroups. These data may facilitate the risk grouping and treatment of AML patients. Moreover, inhibition of PARPs is suggested to enhance the antitumor effect by regulating ferroptosis. Hong et al. (2021) reported that PARP inhibitor olaparib synergizes with erastin via repressing SLC7A11 in BRCA-proficient ovarian cancer cells. The mechanism underlying this synergistic effect is that the repression of SLC7A11 by olaparib may potently enhance lipid peroxidation and ferroptosis. Another study conducted by Tang et al. (2022) showed that PARPs inhibitor olaparib enhances the arsenic trioxide induces ferroptosis by suppressing the expression levels of stearoyl-CoA desaturase1 (SCD1) in platinum-resistant ovary cancer cells. It is, therefore, rational to postulate that PARP1 inhibitors may show a synergistic effect with ferroptosis inducers. Our preliminary results show that the PARP1 inhibitor talazoparib exhibited a synergistic effect with the ferroptosis inducer erastin in suppressing the growth and migration of AML cells. To the best of our knowledge, the combination of talazoparib and erastin in AML has not been previously reported. Since PARPs inhibitors were reported to enhance the oxidative stress of cancer cells (Giovannini et al., 2019) and oxidative stress enhance the lipid peroxidative-dependent ferroptosis (Lei et al., 2022), the synergistic effect of talazoparib and erastin may be achieved by enhancing the ferroptotic cell death pathway.

In conclusion, we constructed a new prognosis prediction model involving 8 PFRGs and clinical characteristics. The model showed favorable calibration and discrimination performance. Our study also provided preliminary evidence that ferroptosis inducers sensitize AML cells to PARP inhibitors, which may benefit AML treatment.

Data availability statement

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

Ethics statement

The studies involving humans were approved by the Ethics Committee of Guangdong Second Provincial General Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

FW: Conceptualization, Funding acquisition, Methodology, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. GX: Formal Analysis, Methodology, Validation, Visualization, Writing–review and editing. GL: Resources, Visualization, Writing–review and editing, Investigation. ZY: Resources, Visualization, Writing–review and editing, Validation. HS: Methodology, Validation, Writing–review and editing. KY: Methodology, Validation, Writing–review and editing. YZ: Visualization, Writing–review and editing. QZ: Funding acquisition, Supervision, Writing–review and editing. RO: Conceptualization, Funding acquisition, Supervision, Writing–review and editing, Project administration. SL: Conceptualization, Funding acquisition, Supervision, Writing–review and editing, Project administration.

Funding

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.2023.1281141/full#supplementary-material

References

Appelbaum, F. R., Gundacker, H., Head, D. R., Slovak, M. L., Willman, C. L., Godwin, J. E., et al. (2006). Age and acute myeloid leukemia. *Blood* 107 (9), 3481–3485. doi:10.1182/blood-2005-09-3724

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 (1), 218. doi:10.1186/s13059-016-1070-5

Cui, Z., Fu, Y., Yang, Z., Gao, Z., Feng, H., Zhou, M., et al. (2022). Comprehensive analysis of a ferroptosis pattern and associated prognostic signature in acute myeloid leukemia. *Front. Pharmacol.* 13, 866325. doi:10.3389/fphar.2022.866325

Dixon, S. J., Lemberg, K. M., Lamprecht, M. R., Skouta, R., Zaitsev, E. M., Gleason, C. E., et al. (2012). Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* 149 (5), 1060–1072. doi:10.1016/j.cell.2012.03.042

Dohner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Buchner, T., et al. (2017). Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 129 (4), 424–447. doi:10.1182/blood-2016-08-733196

Du, J., Wang, T., Li, Y., Zhou, Y., Wang, X., Yu, X., et al. (2019). DHA inhibits proliferation and induces ferroptosis of leukemia cells through autophagy dependent degradation of ferritin. *Free Radic. Biol. Med.* 131, 356–369. doi:10.1016/j.freeradbiomed.2018.12.011

Du, Y., Bao, J., Zhang, M. J., Li, L. L., Xu, X. L., Chen, H., et al. (2020). Targeting ferroptosis contributes to ATPR-induced AML differentiation via ROS-autophagy-lysosomal pathway. *Gene* 755, 144889. doi:10.1016/j.gene.2020.144889

Fang, H., Sheng, S., Chen, B., Wang, J., Mao, D., Han, Y., et al. (2022). A pan-cancer analysis of the oncogenic role of cell division cycle-associated protein 4 (CDCA4) in human tumors. *Front. Immunol.* 13, 826337. doi:10.3389/fimmu.2022.826337

Giovannini, S., Weller, M. C., Repmann, S., Moch, H., and Jiricny, J. (2019). Synthetic lethality between BRCA1 deficiency and poly(ADP-ribose) polymerase inhibition is modulated by processing of endogenous oxidative DNA damage. *Nucleic Acids Res.* 47 (17), 9132–9143. doi:10.1093/nar/gkz624

Gregory, T. K., Wald, D., Chen, Y., Vermaat, J. M., Xiong, Y., and Tse, W. (2009). Molecular prognostic markers for adult acute myeloid leukemia with normal cytogenetics. *J. Hematol. Oncol.* 2, 23. doi:10.1186/1756-8722-2-23

Hong, T., Lei, G., Chen, X., Li, H., Zhang, X., Wu, N., et al. (2021). PARP inhibition promotes ferroptosis via repressing SLC7A11 and synergizes with ferroptosis inducers in BRCA-proficient ovarian cancer. *Redox Biol.* 42, 101928. doi:10.1016/j.redox.2021.101928

Hou, H. A., Lu, J. W., Lin, T. Y., Tsai, C. H., Chou, W. C., Lin, C. C., et al. (2017). Clinico-biological significance of suppressor of cytokine signaling 1 expression in acute myeloid leukemia. *Blood Cancer J.* 7 (7), e588. doi:10.1038/bcj.2017.67

Jiang, X., Stockwell, B. R., and Conrad, M. (2021). Ferroptosis: mechanisms, biology and role in disease. *Nat. Rev. Mol. Cell Biol.* 22 (4), 266–282. doi:10.1038/s41580-020-00324-8

Jinghua Wang, Z. Z., Wang, Y., Yang, S., Chen, J., Wang, Y., Geng, S., et al. (2022). Identification and validation of a prognostic risk-scoring model based on ferroptosis-associated cluster in acute myeloid leukemia. *Front. Cell Dev. Biol.* 9, 800267. doi:10.3389/fcell.2021.800267

Kai Zhu, Z. L., Zhan, Y., Tao, Q., Yu, Z., Chen, L., Fan, C., et al. (2022). A novel 10-gene ferroptosis related prognostic signature in acute myeloid leukemia. *Front. Oncol.* 12, 1023040. doi:10.3389/fonc.2022.1023040

Kim, H. C., Choi, K. C., Choi, H. K., Kang, H. B., Kim, M. J., Lee, Y. H., et al. (2010). HDAC3 selectively represses CREB3-mediated transcription and migration of metastatic breast cancer cells. *Cell Mol. Life Sci.* 67 (20), 3499–3510. doi:10.1007/s00018-010-0388-5

Lei, G., Zhuang, L., and Gan, B. (2022). Targeting ferroptosis as a vulnerability in cancer. *Nat. Rev. Cancer* 22 (7), 381–396. doi:10.1038/s41568-022-00459-0

Li, R., Peng, C., Zhang, X., Wu, Y., Pan, S., and Xiao, Y. (2017). Roles of Arf6 in cancer cell invasion, metastasis and proliferation. *Life Sci.* 182, 80–84. doi:10.1016/j.lfs.2017.06.008

Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., et al. (2018). An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 173 (2), 400–416. doi:10.1016/j.cell.2018.02.052

Liu, X., Zhong, S., Qiu, K., Chen, X., Wu, W., Zheng, J., et al. (2023). Targeting NRF2 uncovered an intrinsic susceptibility of acute myeloid leukemia cells to ferroptosis. *Exp. Hematol. Oncol.* 12 (1), 47. doi:10.1186/s40164-023-00411-4

Newell, L. F., and Cook, R. J. (2021). Advances in acute myeloid leukemia. *BMJ* 375, n2026. doi:10.1136/bmj.n2026

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Padella, A., Ghelli Luserna Di Rora, A., Marconi, G., Ghetti, M., Martinelli, G., and Simonetti, G. (2022). Targeting PARP proteins in acute leukemia: DNA damage response inhibition and therapeutic strategies. *J. Hematol. Oncol.* 15 (1), 10. doi:10.1186/s13045-022-01228-0

Riva, L., Luzi, L., and Pelicci, P. G. (2012). Genomics of acute myeloid leukemia: the next generation. *Front. Oncol.* 2, 40. doi:10.3389/fonc.2012.00040

Short, N. J., Rytting, M. E., and Cortes, J. E. (2018). Acute myeloid leukaemia. *Lancet* 392 (10147), 593–606. doi:10.1016/S0140-6736(18)31041-9

Stockwell, B. R., Friedmann Angeli, J. P., Bayir, H., Bush, A. I., Conrad, M., Dixon, S. J., et al. (2017). Ferroptosis: a regulated cell death nexus linking metabolism, redox biology, and disease. *Cell* 171 (2), 273–285. doi:10.1016/j.cell.2017.09.021

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

Tang, S., Shen, Y., Wei, X., Shen, Z., Lu, W., and Xu, J. (2022). Olaparib synergizes with arsenic trioxide by promoting apoptosis and ferroptosis in platinum-resistant ovarian cancer. *Cell Death Dis.* 13 (9), 826. doi:10.1038/s41419-022-05257-y

Wei, R., Qiu, H., Xu, J., Mo, J., Liu, Y., Gui, Y., et al. (2020). Expression and prognostic potential of GPX1 in human cancers based on data mining. *Ann. Transl. Med.* 8 (4), 124. doi:10.21037/atm.2020.02.36

Yin, Z., Li, F., Zhou, Q., Zhu, J., Liu, Z., Huang, J., et al. (2022). A ferroptosis-related gene signature and immune infiltration patterns predict the overall survival in acute myeloid leukemia patients. *Front. Mol. Biosci.* 9, 959738. doi:10.3389/fmolb.2022.959738

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yu, T., Wei, L., and You, H. (2022). Ferroptosis-related gene signature predicts the clinical outcome in pediatric acute myeloid leukemia patients and refines the 2017 ELN classification system. *Front. Mol. Biosci.* 9, 954524. doi:10.3389/fmolb.2022.954524

Yu, Y., Xie, Y., Cao, L., Yang, L., Yang, M., Lotze, M. T., et al. (2015). The ferroptosis inducer erastin enhances sensitivity of acute myeloid leukemia cells to chemotherapeutic agents. *Mol. Cell Oncol.* 2 (4), e1054549. doi:10.1080/23723556.2015.1054549

Yuan, B., El Dana, F., Ly, S., Yan, Y., Ruvolo, V., Shpall, E. J., et al. (2020). Bone marrow stromal cells induce an ALDH+ stem cell-like phenotype and enhance therapy resistance in AML through a TGF-beta-p38-ALDH2 pathway. *PLoS One* 15 (11), e0242809. doi:10.1371/journal.pone.0242809

Yusuf, R. Z., Saez, B., Sharda, A., van Gastel, N., Yu, V. W. C., Baryawno, N., et al. (2020). Aldehyde dehydrogenase 3a2 protects AML cells from oxidative death and the synthetic lethality of ferroptosis inducers. *Blood* 136 (11), 1303–1316. doi:10.1182/blood.2019001808

Progress of non-small-cell lung cancer with *ROS1* rearrangement

Xin Yang[1], Zhe Tang[2], Jing Li[1], Jizong Jiang[1] and Yue Liu[3]*

[1]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China, [2]Department of Thoracic Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China, [3]Department of Thyroid and Breast Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

ROS1 rearrangement is found in 0.9%−2.6% of people with non-small-cell lung cancers (NSCLCs). Tyrosine kinase inhibitors (TKIs) target *ROS1* and can block tumor growth and provide clinical benefits to patients. This review summarizes the current knowledge on *ROS1* rearrangements in NSCLCs, including the mechanisms of *ROS1* oncogenicity, epidemiology of *ROS1*-positive tumors, methods for detecting rearrangements, molecular characteristics, therapeutic agents, and mechanisms of drug resistance.

1 Introduction

Lung cancer remains the most fatal malignant tumor, with approximately 85% of cases being non-small-cell lung cancer (NSCLC). Of those with NSCLC, about 25% carry positive-driven gene changes that can benefit from the corresponding molecular-targeted therapy (Siegel et al., 2022). Compared with epidermal growth factor receptor (*EGFR*) mutations and anaplastic lymphoma kinase (*ALK*) rearrangements, the genetic proto-oncogene tyrosine-protein kinase-1 (*ROS1*) is less prevalent in NSCLC, accounting for approximately 0.9%–2.6% of cases (Bergethon et al., 2012; Cai et al., 2013). The results of prospective Phase I/II clinical trials have confirmed the effectiveness of crizotinib in *ROS1*-positive NSCLC (Shaw et al., 2019a), and in recent years, several targeted drugs, including entrectinib, ceritinib, and lorlatinib, have also shown excellent antitumor activity (Shaw et al., 2017; Peters et al., 2020; Dziadziuszko et al., 2021). This article provides an overview of the progress regarding research on NSCLC with the *ROS1* rearrangement.

2 *ROS1* gene

The *ROS1* gene was discovered in the 1980s in the products of bird myeloma virus RNA UR2 (Balduzzi et al., 1981). The human *ROS1* gene is located on chromosome 6q21 (Nagarajan et al., 1986), which belongs to the family insulin receptor genes of receptor tyrosine kinases (RTKs), which encode intermembrane proteins consisting of 2,347 amino acids, including hydrophobic extracellular domains, a transmembrane region, and intracellular parts of the tyrosine kinase domain (Roskoski, 2017). Rikova et al. first reported the role of the *ROS1* oncogene in NSCLC in 2007 and identified two new protein fusion transcription factors, *SLC34A2* and *CD74* (Rikova et al., 2007). With the continuous improvement in modern sequencing technology, an increasing number of fusion

species have been discovered, and their role as carcinogenic genes in multiple cancers has been gradually confirmed (Dagogo-Jack et al., 2019; Zhang et al., 2019).

ROS1 plays an important role in activating multiple signaling pathways, including those involved in cell differentiation, proliferation, growth, and survival (Figure 1). *ROS1* rearrangement causes disorders in enzyme-active proteins and the abnormal activation of the associated signaling pathways by forming phosphate thyroxine recruitment spots at the end of the ROS, including tyrosine phosphatase tumor suppressor SHP1/SHP2, pro-mitotic protein extracellular-signal-regulated kinase (ERK)-1/2, insulin-receptor substrate (IRS)-1, phosphatidylinositol 3-kinase (PI3K), protein kinase B (AKT), mitogen-activated protein kinases (MAPKs), signal transducer and activator of transcription (STAT)-3, and the VAV3-related signaling pathway (Huang et al., 2020).

3 Epidemiological and clinical features

Among the people with NSCLC in China, approximately 2.59% carry the *ROS1* fusion gene, and approximately 17,000 new cases of *ROS1*-positive NSCLC are estimated to occur annually in China. *ROS1* rearrangements are more common in young, female, and nonsmoking patients (Fu et al., 2015; Zhu et al., 2015). The main pathological types are adhesive, vesicular, or solid glandular cancers; a few are squamous-cell, multicellular, or large-cell cancer (Park et al., 2019), more than

90% of which express thyroid transcription factor-1 (TTF-1), mostly diagnosed as phase III–IV, high incidence of brain transfusions (Drilon et al., 2021). Compared to other types of NSCLC, ROS1-positive NSCLC has a significantly increased risk of developing thromboembolic diseases (Shah et al., 2021; Woodford et al., 2021; Zhu et al., 2021), but the underlying mechanism remains unclear.

4 Molecular characteristics

4.1 Fusion partners

The most common fusion partners of the *ROS1* gene include *CD74* (38%–54%), *EZR* (13%–24%), *SDC4* (9%–13%), and *SLC34A2* (5%–10%) (Cui et al., 2020; Huang et al., 2021a). With the continuous improvement in DNA- and RNA-sequencing technology, new fusion partners have been discovered, such as *CCCKC6, TFG, SLMAP, MYO5C, FIG, LIMA1, CLTC, GOPC, ZZCCHC8, CEP72, MLL3, KDELR2, LRIG3, MSN, MPRIP, WNK1, SLC6A17, TMEM106B, FAM135B, TPM3,* and *TDP52L1* (Li et al., 2018). People with NSCLC with the *CD74–ROS1* rearrangement have longer progression-free survival (PFS) and overall survival (OS) than those with other types of rearrangements (Li et al., 2018), but no similar conclusions have been found in other studies (Cui et al., 2020). Currently, the relationship between the fusion partners and prognosis remains unclear.
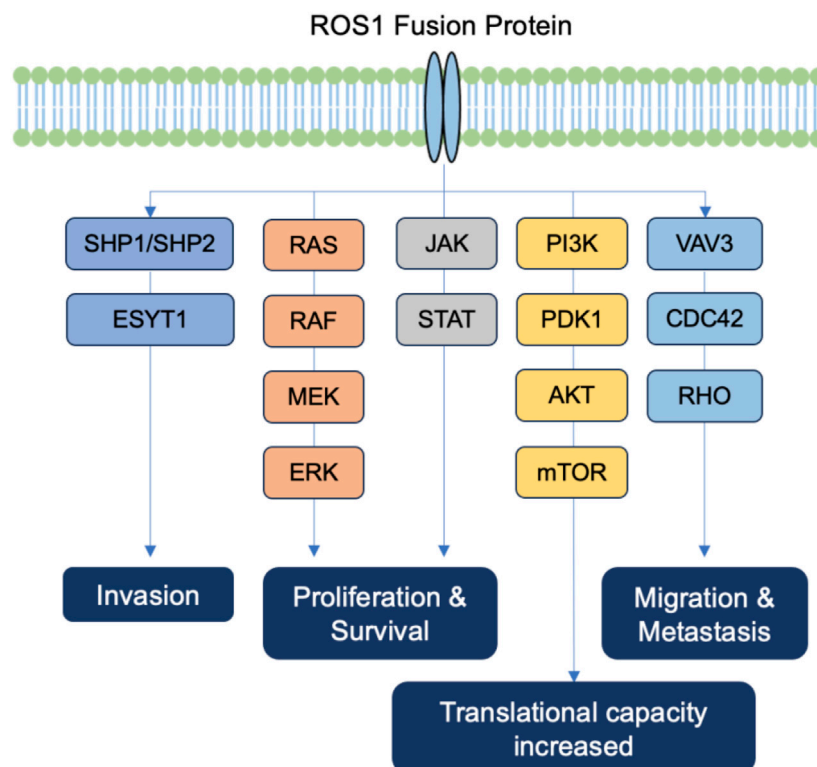

FIGURE 1
The ROS1 signaling pathway.

TABLE 1 Summary of co-occurring genetic mutation of ROS1.

Co-occurring genetic mutation of ROS1	Effect
ROS1-EGFR	ROS1 rearrangement may be one of the mechanisms of resistance to EGFR inhibitors; Combination of EGFR inhibitors with crizotinib may be effective
ROS1-ALK	Both mutations are sensitive to crizotinib, suggesting crizotinib as a potential treatment
ROS1-KRAS	Limited response to crizotinib; Variable treatment outcomes
ROS1-MET	Variable treatment outcomes; Combined use of highly selective MET inhibitors with crizotinib is necessary
ROS1-BRAF	Limited information on treatment outcomes; No published reports on combined treatment with ROS1 and BRAF inhibitors
ROS1-TP53	Associated with a shorter survival time

4.2 Co-occurring genetic mutations

Approximately 36% of people with NSCLC with *ROS1* rearrangements have co-occurring genetic mutations. Those with co-mutation have a worse prognosis than non-co-mutant patients (PFS 8.5 months versus 15.5 months, $p = 0.0213$) (Zeng et al., 2018). *ROS1* rearrangement typically does not simultaneously occur with mutations in other genes (such as *EGFR, ALK*, and *KRAS*). However, in recent years, occasional cases of *ROS1* rearrangements associated with mutations in other driving genes have been reported (Zhu et al., 2016; Uguen et al., 2017). Lambros et al. reported 15 cases of NSCLC with an *ROS1* mutation in conjunction with an *EGFR* mutation, including 9 cases of 19 deletion, 5 cases of L858R mutation, and 1 case of 20 insertion. *ROS1* rearrangement may be one of the mechanisms of resistance to EGFR inhibitors, and the combination of EGFR inhibitors with crizotinib may be an effective treatment (Lambros et al., 2018). Crizotinib, as a tyrosine kinase inhibitor, binds to the ATP-binding site of the *ROS1* kinase domain, inhibiting its enzymatic activity. By blocking *ROS1* signaling, crizotinib disrupts the intracellular pathways that drive cancer cell proliferation and survival.

Co-mutations in both ROS1 and ALK are rare but occasionally reported. Both mutations were found to be sensitive to crizotinib, which may be a more suitable treatment option. Uguen et al. and Song et al. respectively reported a lung adenocarcinoma patient with concurrent ALK/ROS1 rearrangements confirmed by fluorescence *in situ* hybridization (FISH) analyses, and both patients showed response to crizotinib (Song et al., 2017; Uguen et al., 2017).

ROS1 and *KRAS* co-mutations are rare and can be primary or successive. In a study involving six patients with the *KRAS–ROS1* comutation, only one patient benefitted from crizotinib treatment, while KRASG13D- or G12V-mutation carriers showed no response (Lin et al., 2017).

ROS1–MET comutations are extremely rare. Tang et al. reported one *ROS-1–MET* co-mutation in a series of 15 patients, however, no treatment information was provided (Tang et al., 2018). Rihawi et al. noted one case of NSCLC with a *ROS1–MET* co-mutation, and MET inhibitor capmatinib treatment failed, followed by crizotinib treatment, and a PFS of up to 11 months (Rihawi et al., 2018). Zeng et al. reported a case of NSCLC with *ROS1* rearrangement and MET expansion, with disease progression occurring 1.5 months after crizotinib treatment (Zeng et al., 2018). Therefore, the combined use of highly selective MET inhibitors based on crizotinib treatment is necessary for patients with this type of comutation. Therefore, further clinical research and practice are required.

Comutations in *ROS1* and *BRAF* have been reported, but no reports on combined treatment with an *ROS1* inhibitor with *BRAF* inhibitors have been published (Wiesweg et al., 2017). Furthermore, it was reported that comutations with *TP53* may be associated with a shorter survival time (Lindeman et al., 2018) (Table 1).

5 Methods of *ROS1* detection

The detection of *ROS1* is currently recommended in all patients with non-squamous-cell cancer to determine the indication for the selection of the corresponding targeted drug.

Fluorescence *in situ* hybridization (FISH) is the gold standard for diagnosing *ROS1* rearrangements. FISH uses a two-probe (3′ and 5′) separation design and can detect ≥50 tumor cells; the result is considered positive when more than 15% of the cells show 3′- and 5′-probe separation or separate 3′ signals. The disadvantages of FISH include its high cost, technical difficulty, and time consumption. At present, probes that can detect both *ROS1* and *ALK* rearrangements have been put into clinical use as they have less strict requirements for tumor samples (Ginestet et al., 2018; Zito Marino et al., 2020).

The immunohistochemistry (IHC) technique is commonly used to screen for *ROS1* arrangements owing to its convenience, low cost, and ease of operation; IHC has a sensitivity of approximately 90%–100% and a specificity of approximately 70%–90%. False-positive outcomes may occur in one-third of patients, especially in those with adhesive or adenomatous EGFR-mutant glandular cancer (Wang et al., 2020; Makarem et al., 2021). Therefore, positive or suspicious IHC results require further confirmation by FISH, reverse -transcription quantitative polymerase chain reaction (RT-qPCR), or next-generation sequencing (NGS) (Hofman et al., 2019; Cheung et al., 2021; Fielder et al., 2022).

Two situations should be noted: 1) IHC positivity and FISH negativity, indicating the presence of another carcinogen-driven mutation not included in FISH detection and requiring further confirmation by RT-qPCR or NGS testing (Kim et al., 2021); 2) FISH may result in false-negative results for some fusion partners, mainly *GOPC–ROS1* or *EZR–ROS1*. In the latter, the 5′-*ROS1* gene is usually lacking, and the corresponding FISH detection uses a separate 3′ probe (Capizzi et al., 2019). In a recent study, the authors detected a rare case of NSCLC with *ROS1* fusion

(*SQSTM1*), *ROS1* mutation, and *ROS1* expansion with positive IHC expression using NGS technology (Huang et al., 2021b). Therefore, if the results remain uncertain after both IHC and FISH testing, NGS should be performed to confirm the presence of unusual fusion genotypes.

With RT-qPCR, the unique primers detecting *ROS1* rearrangements are used; the method has a sensitivity of up to 100% and a specificity of 85.1%. The disadvantage of the method is its technical difficulty, requiring several steps, including RNA extraction, complementary DNA synthesis, quantitative PCR, and data analysis, which are currently more commonly applied in laboratories: few clinical applications are lacking (Shan et al., 2015).

NGS can be used to theoretically identify all fusion partners, including new variation types and other carcinogenic genetic variations. The requirements for the samples are no strict, and tumor tissues or blood plasma can be used as test samples. Recent advances in NGS technology for detecting ROS1 rearrangements involve both DNA and RNA approaches. DNA-based methods, such as targeted sequencing, whole exome sequencing (WES), and whole genome sequencing (WGS), target specific genomic regions or the entire genome to identify structural alterations. RNA-based techniques, including targeted RNA sequencing and RNA-Seq, directly detect fusion transcripts, providing valuable information on gene rearrangements. Hybrid approaches integrating DNA and RNA analyses enhance sensitivity and specificity. Improved bioinformatics tools and the use of single-molecule sequencing technologies contribute to increased accuracy, while emerging liquid biopsy methods offer less invasive options. Combining these approaches allows for a comprehensive and precise assessment of ROS1 rearrangements in lung cancer genomes (Clave et al., 2019). NGS was used to detect many novel uncommon ROS1 fusions, most of which were reported to be sensitive to matched targeted therapy, similar to the canonical fusions. The clinical significance of some genomic breakpoints remained unclear and could be explored further via NGS technology (Li et al., 2022). The disadvantages are that the cost is higher and the results cannot be quickly obtained (Mosele et al., 2020).

6 Treatment of NSCLC with *ROS1* rearrangement

The U.S. Food and Drug Administration (FDA) has approved crizotinib and entrectinib as first-line treatments for patients with unresectable NSCLC with *ROS1* rearrangements. Other tyrosine kinase inhibitors (TKIs), including ceritinib and lorlatinib, also exhibit excellent antitumor activity.

6.1 Crizotinib

Crizotinib is a small molecule inhibitor with a complex molecular structure. It contains various functional groups, including pyrazole, pyridine, and piperidine rings. The three-dimensional structure allows it to bind to the ATP-binding site of the ROS1 kinase domain, inhibiting its activity. The Phase I PROFILE 1001 study included 50 people with *ROS1*-positive advanced NSCLC receiving crizotinib (250 mg twice daily). The

objective remission rate (ORR) was 72%; disease control rate (DCR) was 90%; median duration of response (DOR) was 24.7 months; and median PFS and OS were 19.3 months and 51.4 months, respectively. The most common adverse events included visual impairment (82%), diarrhea (44%), nausea (40%), edema (40%), constipation (34%), vomiting (34%), elevated transaminase (22%), fatigue (20%), and taste disturbance (18%). Most adverse reactions were grade 1 to 2[4].

Compared with the PROFILE 1001 study, in the AcSé Phase I/II study based on 37 patients, the effectiveness of crizotinib was relatively poor, with an ORR of 47.2%, and median PFS and OS of 5.5 and 17.2 months, respectively. The researchers presumed that more patients with a higher performance status (PS) score of two points (25% vs. 2%) were grouped in the AcSé study (Wu et al., 2018). The results of the EUCROSS and METROS studies from Europe were similar to those of the PROFILE 1001 study, with ORRs of 70% and 65%, and median PFS values of 20 and 22.8 months, respectively (Michels et al., 2019).

Although crizotinib showed excellent anti-tumor activity in the treatment of *ROS1*-positive NSCLC, its blood–brain barrier penetration rate was low, and brain metastases (47%) became the main area of disease progression. In addition, approximately 36% those with *ROS1*-positive NSCLCs also experienced brain metastases at the baseline level. Therefore, ROS1–TKIs that can better cross the blood–brain barrier should be developed in the future.

6.2 Entrectinib

Entrectinib is a multitarget inhibitor of ROS1, ALK, and pan-tropomyosin receptor kinase (TRK). Its molecular structure consists of various cyclic and aromatic structures, including a tetrahydropyrrolopyrazine ring. Entrectinib is designed to penetrate the blood-brain barrier, making it effective against central nervous system metastases. In *in vitro* experiments, its anti-ROS1 activity was 40 times stronger than that of crizotinib (Rolfo et al., 2015). The results of two Phase I/II studies showed antitumor activity and good tolerance to entrectinib (Drilon et al., 2017). The most common side effects included discomfort in taste (41.4%), fatigue (27.9%), dizziness (25.4%), constipation (23.7%), diarrhea (22.8%), nausea (20.8%), and weight gain (19.4%). The results of the STARTRK-2 study confirmed the efficacy of entrectinib in *ROS1*-positive NSCLC involving a total of 161 patients who had not previously received anti-*ROS1* treatment. Of these, 34.8% people also experienced baseline brain metastases. The ORR was 67.1%, PFS median was 15.7 months, and 1-year OS rate was 81%. For 24 patients at the baseline with measurable brain metastases, the intracerebral ORR was 79.2%, and intracerebral PFS was 12 months (Drilon et al., 2020).

6.3 Lorlatinib

Lorlatinib is a third-generation *ROS1* inhibitor designed to overcome resistance mutations that may develop during treatment with earlier-generation inhibitors. It has a more intricate structure compared to crizotinib, with multiple fused rings and functional groups, increasing the permeability of the blood–brain barrier by

reducing P-glucose-1-mediated exudation. Lorlatinib exhibits activity against *ROS1* as well as *ALK*. In a Phase I/II study, 61 people with *ROS1*-positive NSCLC were included, including 21 patients treated with primary TKIs and 40 patients previously treated with crizotinib. The ORR of the primary TKI patients was 62%, their median PFS was 21 months, their intracerebral ORR was 64%, and intracerebral PFS was not achieved. The ORR, median PFS, and intracerebral ORR of the patients treated with crizotinib were 35%, 8.5 months, and 50%, respectively. The most common adverse events included hypocholesterolemia (65%), hypoglycemia (42%), peripheral edema (39%), surrounding neuropathy (35%), cognitive changes (26%), weight gain (16%), and mood disorders (16%). The incidences of grade 3 and 4 adverse reactions were 43% and 6%, respectively (Shaw et al., 2019b). Monitoring the plasma concentration of lorlatinib may help control adverse events without altering the effectiveness of this antitumor therapy (Chen et al., 2021). When resistance mutations, such as $ROS1^{K1991E}$ or $ROS1^{S1986F}$, appear after treatment with crizotinib, lorlatinib may have a stronger effect; however, it has a minor effect on the resistance mutation type $ROS1^{G2032R}$.

6.4 Ceritinib

Ceritinib is an ALK inhibitor that exhibits antitumor activity and intracerebral effects in those with ALK-positive NSCLC (Shaw et al., 2017). The molecular structure of ceritinib includes various aromatic rings and functional groups. *In vitro* experiments showed that ceritinib has potential anti-*ROS1* rearrangement activity. In a Phase II study involving 32 patients with *ROS1*-positive NSCLC, the ORR for ceritinib treatment was 62%, and the median PFS was 9.3 months. In the subgroup (n = 30) that was not treated with crizotinib, the median PFS was 19.3 months, and its effectiveness was comparable to that of other TKIs. The ORR in the brain metastasis subgroup (n = 8) was 63%. Its tolerance was similar to that of other TKIs, with an incidence of 37% of adverse events of grade 3 or above (Drilon et al., 2016).

6.5 Cabozantinib

Cabozantinib is a small-molecule TKI and consists of a pyridine ring with a fluorine atom and a 3-(morpholin-4-yl) propoxy group attached to it, as well as a 3-aminopyridine-2-carboxamide group and a 4-(6-(propan-2-yl) pyridin-3-yl) benzoic acid group (Maroto et al., 2022). It targets ROS1, MET, VEGFR-2, RET, and AXL and has a strong ability to penetrate the blood–brain barrier. The results of preclinical studies and case reports have indicated that cabozantinib is effective for treating patients with resistance to other TKIs (crizotinib, entrectinib, and ceritinib) and resistance mutations (such as *D2033N* or *G2032R*) (Sun et al., 2019). Therefore, cabozantinib is clinically used as a treatment option after the development of resistance to other TKIs.

6.6 Brigatinib

Brigatinib is a kinase inhibitor with a complex structure. It contains pyrimidine, pyridine, and aniline moieties, among

others, arranged in a way that allows it to inhibit the activity of certain tyrosine kinases, it is a multitarget inhibitor of ROS1 and ALK and has antitumor activity against EGFR-mutant NSCLC. Brigatinib exhibits antitumor activity against several drug-resistant *ROS1* mutations. In one study, researchers assessed the efficacy and tolerance of brigatinib in eight patients with *ROS1*-positive NSCLC, one of whom did not receive TKI treatment, and seven of whom developed disease progression after crizotinib treatment (Dudnik et al., 2020). The ORRs for the total population and the crizotinib-treated subgroups were 37% and 29%, respectively. No Grade 3 or 4 adverse events were observed. In one case report, the disease progressed after prior treatment with several ROS1-TKIs, and brigatinib therapy remained effective (Hegde et al., 2019). *In vitro* experiments showed that Bugatti has strong anti-tumor activity against NSCLC carrying the *L2026M* mutation but is ineffective against the *G2032R* mutation (Camidge et al., 2018).

6.7 Repotrectinib

Repotrectinib (TPX-0005) is also a multitarget TKI that can target ROS1, TRK, and ALK and effectively cross the blood–brain barrier. It contains various functional groups, including cyclic structures and heteroatoms like nitrogen and oxygen. The specific arrangement of atoms in Repotrectinib allows it to interact with the ATP-binding sites of these kinases, inhibiting their activity and disrupting the signaling pathways that contribute to cancer cell growth. In preclinical studies, repotrectinib showed strong antitumor activity in NSCLC models with *ROS1*-positive brain metastases, prior ROS1-TKI treatment, and $ROS1^{G2032R}$ mutations (Yun et al., 2020). A Phase I/II clinical trial (NCT03093116) to evaluate the safety and efficacy of repotrectinib in ROS1-positive NSCLC is currently underway.

6.8 Taletrectinib

Taletrectinib (DS-6051B) is an inhibitor targeting *ROS1* and *NTRK* and shows antitumor activity against crizotinib-resistant NSCLC, including in patients carrying the *G2032R* mutation. In *in vitro* experiments, taletrectinib showed strong activity against resistance mutations, such as *G2032R, L1951R, S1986F,* and *L2026M.* Two Phase I clinical trials conducted in the United States and Japan evaluated the effectiveness of taletrectinib in patients with *ROS1*-positive NSCLC. The former included 46 patients with an ORR of 33% in patients with crizotinib resistance; the latter included 15 patients with ORRs of 58.3% and 66.7% in all patients and in patients with no prior crizotinib treatment, respectively (Papadopoulos et al., 2020).

6.9 Ensartinib

Ensartinib is an ALK-TKI that demonstrated 10-times higher anti-ALK activity than crizotinib in *in vitro* experiments. A Phase II trial of *ROS1*-positive NSCLC (NCT03608007) showed a certain therapeutic effectiveness, with an ORR of 27%, and intracerebral

disease control was achieved in three out of four patients with brain metastases (Ai et al., 2021).

6.10 Other treatments

Chemotherapy remains the recommended second-line treatment after the failure of crizotinib treatment. The combination of antivascular therapy with ROS1-TKIs is another potentially effective treatment strategy. In *in vitro* experiments, the combined use of vascular endothelial growth factor (VEGF) blockers and ROS1-TKI increased antitumor activity (Watanabe et al., 2021). In a clinical trial involving 14 patients, those with NSCLC that was *ALK*-positive, *ROS1*-positive, or had *MET* expansion showed an ORR was 58.3% with good tolerance; however, 3 patients discontinued treatment due to hepatotoxicity or hemorrhage (Saito et al., 2019).

The effectiveness of immunotherapy in *ROS1*-positive NSCLC has not yet been fully elucidated. In *in vitro* experiments, ROS1 was found to regulate the expression of programmed death protein ligand-1 (PD-L1) by activating the MEK–ERK and ROS1–SHP2 signaling pathways. Most ROS1-positive tumors do not express PD-L1 and have a low tumor mutation burden (TMB) (Choudhury et al., 2021). The results of small-sample studies have shown that the ORR of people with *ROS1*-positive NSCLC receiving mono-immunotherapy was 13%–17% (Mazieres et al., 2019), and the ORR of immunotherapy combined with chemotherapy was 83% (Guisier et al., 2020). Choudhury et al. (Choudhury et al., 2021) found no noticeable differences in the expression levels of PD-L1 between patients who received effective and ineffective immunotherapies. In patients with TKI resistance, combined immunotherapy has a high clinical application value, but

potential toxic reactions to sequential immunotherapy with TKIs must be monitored.

7 Resistance mechanisms of ROS1 inhibitors

7.1 Resistance mechanisms of crizotinib

7.1.1 Structural domain mutations

ROS1 kinase structural domain mutation is the most common resistance mechanism to crizotinib, accounting for approximately 40%–55% of the total. G2032R is the most common type of mutation, occurring in the solvent area of the ATP binding site, accounting for approximately 33%–41% of cases (Awad et al., 2013). In *in vitro* studies, the $ROS1^{G2032R}$ mutation increased the expression of *TWIST1*, promoting epithelial–mesenchymal transition (EMT), cell migration, and resistance to ROS1-TKIs by modifying the combination of location points and spatial blockages (Gou et al., 2018). Currently, repotrectinib, topotrectinib, and cabozantinib have shown improved anti-$ROS1^{G2032R}$ activity (Gou et al., 2018).

Other common mutations include *D2033N* (2.4%–6%), *S1986Y/F* (2.4%–6%), *L2026M* (1%), *L2155S* (1%), *L1951R* (1%), and *S1886* (1%) (Katayama et al., 2015) (Table 2). $ROS1^{D2033N}$ induces the modification of ATP-binding pockets, resulting in the weakening of the ability of tumor cells to bind to ROS1-TKIs. In *in vitro* experiments, $ROS1^{D2033N}$ led to resistance to crizotinib, entrectinib, and ceritinib, but remained sensitive to lorlatinib, repotrectinib, and cabozantinib. $ROS1^{S1986F}$ results in resistance to crizotinib, entrectinib, and ceritinib by changing the position of a ring structure rich in glycine at the end of the C spiral (Facchinetti et al., 2016).

TABLE 2 Common mutation types of crizotinib resistance and effective inhibitors.

Type of ROS1 fusion	Mutation site	Mechanisms	Effective TKI
CD74-ROS1	G2032R	Altered spatial structure of ROS1 domain interferes with drug binding and leads to resistance to ROS1 inhibitors	cabozantinib
			repotrectinib
	D2033N	Alteration in the electrostatic force on the outer surface of the ATP-binding site and a rearrangement of ATP-binding site	lorlatinib
			cabozantinib
			repotrectinib
	L2026M	Leucine to methionine substitution	ceritinib
			lorlatinib
			cabozantinib
			repotrectinib
SLC34A2-ROS1	L2155S	Protein dysfunction	lorlatinib
			cabozantinib
			repotrectinib
EZR-ROS1	S1986F/Y	Increased kinase activity	lorlatinib
			cabozantinib
			repotrectinib

7.1.2 Activation of other signaling pathways

Crizotinib resistance may also be associated with the activation of other signaling pathways downstream of *ROS1*. The activation of these downstream signaling pathways can lead to the following: 1) stimulation of the signaling pathway that resists ROS1-TKIs and 2) the production of new mutations or expansions at the level of other oncogenes. For example, SHP2 activation of the MAPK/MEK/ERK pathway can lead to TKI resistance. In *in vitro* experiments, a combination of SHP2 inhibitors and ROS1-TKI increased the inhibition of tumor growth (Li et al., 2021). Additionally, mutations or proliferation of *ALK* (Li et al., 2021), *BRAF* (Ren et al., 2021), *KRAS*, and *MET* (Wang et al., 2021) can lead to resistance to ROS1-TKIs.

7.1.3 Phenotype transformation

In a related case report, small-cell cancer transformation may occur after the resistance to ROS1-TKIs (Yang et al., 2021). According to these findings, this phenotype transformation may be associated with the inactivation of the retinoblastoma 1 (*RB1*) and *TP53* genes (Fares et al., 2020; Lin et al., 2020).

7.2 Resistance mechanisms of lorlatinib

The mechanisms underlying the resistance to TKIs other than crizotinib are not fully understood. Lin et al. analyzed 28 cases of post-lorlatinib progressive tumor tissue samples and found mutations in the kinase structural domain, especially *G2032K* and *L2086F* mutations (Lin et al., 2021). The results of *in vitro* experiments showed that the $ROS1^{G2032K}$ mutation conferred resistance to crizotinib, entrectinib, and lorlatinib. For the $ROS1^{L2086F}$ mutation, the related models showed that it was equally resistant to crizotinib, entrectinib, and lorlatinib, whereas cabozantinib may have a therapeutic effect (Mazieres et al., 2015). Other resistance mechanisms to lorlatinib include *MET* expansion (4%), $KRAS^{G12C}$ mutation (4%), *KRAS* expansion (4), and *NRAS* extension (4%) (Papadopoulos et al., 2020).

8 Conclusion

Targeted therapy is the foundation for the treatment of people with unresectable NSCLC with *ROS1* rearrangements. Crizotinib and entrectinib are currently recommended as the standard first-line

therapeutic drugs. In future drug development, antitumor activity and brain permeation are important indicators for measuring the effectiveness of ROS1-TKIs. The new TKIs, entrectinib and lorlatinib, have a higher rate of blood–brain barrier permeation and are expected to provide increased control of brain metastases. For the primary mutants of $ROS1^{G2302R}$, repotrectinib and taletrectinib also showed high clinical efficacy.

In cases of disease progression after crizotinib treatment, the choice of secondary treatment should depend on the type of progression and specific resistance mechanisms. In the case of oligometastasis, topical treatment, represented by radiation or surgery, should be quickly administered, especially in patients with brain metastases. When multiorgan progression occurs, a whole-body treatment scheme based on chemotherapy containing platinum is still the current standard treatment. Targeted drugs may be more suitable treatment options for patients who have undergone genetic testing to clearly identify resistance mechanisms.

Author contributions

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

Ai, X., Wang, Q., Cheng, Y., Liu, X., Cao, L., Chen, J., et al. (2021). Safety but limited efficacy of ensartinib in ROS1-positive NSCLC: a single-arm, multicenter phase 2 study. *J. Thorac. Oncol.* 16 (11), 1959–1963. doi:10.1016/j.jtho.2021.06.023

Awad, M. M., Engelman, J. A., and Shaw, A. T. (2013). Acquired resistance to crizotinib from a mutation in CD74-ROS1. *N. Engl. J. Med.* 369 (12), 1173. doi:10.1056/NEJMc1309091

Balduzzi, P. C., Notter, M. F., Morgan, H. R., and Shibuya, M. (1981). Some biological properties of two new avian sarcoma viruses. *J. Virol.* 40 (1), 268–275. doi:10.1128/JVI.40.1.268-275.1981

Bergethon, K., Shaw, A. T., Ou, S. H., Katayama, R., Lovly, C. M., McDonald, N. T., et al. (2012). ROS1 rearrangements define a unique molecular class of lung cancers. *J. Clin. Oncol.* 30 (8), 863–870. doi:10.1200/JCO.2011.35.6345

Cai, W., Li, X., Su, C., Fan, L., Zheng, L., Fei, K., et al. (2013). ROS1 fusions in Chinese patients with non-small-cell lung cancer. *Ann. Oncol.* 24 (7), 1822–1827. doi:10.1093/annonc/mdt071

Camidge, D. R., Kim, D. W., Tiseo, M., Langer, C. J., Ahn, M. J., Shaw, A. T., et al. (2018). Exploratory analysis of brigatinib activity in patients with anaplastic lymphoma kinase-positive non-small-cell lung cancer and brain metastases in two clinical trials. *J. Clin. Oncol.* 36 (26), 2693–2701. doi:10.1200/JCO.2017.77.5841

Capizzi, E., Dall'Olio, F. G., Gruppioni, E., Sperandi, F., Altimari, A., Giunchi, F., et al. (2019). Clinical significance of ROS1 5' deletions in non-small cell lung cancer. *Lung Cancer* 135, 88–91. doi:10.1016/j.lungcan.2019.07.017

Chen, J., Ruiz-Garcia, A., James, L. P., Peltz, G., Thurm, H., Clancy, J., et al. (2021). Lorlatinib exposure-response analyses for safety and efficacy in a phase I/II trial to

support benefit-risk assessment in non-small cell lung cancer. *Clin. Pharmacol. Ther.* 110 (5), 1273–1281. doi:10.1002/cpt.2228

Cheung, C. C., Smith, A. C., Albadine, R., Bigras, G., Bojarski, A., Couture, C., et al. (2021). Canadian ROS proto-oncogene 1 study (CROS) for multi-institutional implementation of ROS1 testing in non-small cell lung cancer. *Lung Cancer* 160, 127–135. doi:10.1016/j.lungcan.2021.08.003

Choudhury, N. J., Schneider, J. L., Patil, T., Zhu, V. W., Goldman, D. A., Yang, S. R., et al. (2021). Response to immune checkpoint inhibition as monotherapy or in combination with chemotherapy in metastatic ROS1-rearranged lung cancers. *JTO Clin. Res. Rep.* 2 (7), 100187. doi:10.1016/j.jtocrr.2021.100187

Clave, S., Rodon, N., Pijuan, L., Díaz, O., Lorenzo, M., Rocha, P., et al. (2019). Next-generation sequencing for ALK and ROS1 rearrangement detection in patients with non-small-cell lung cancer: implications of FISH-positive patterns. *Clin. Lung Cancer* 20 (4), e421–e429. doi:10.1016/j.cllc.2019.02.008

Cui, M., Han, Y., Li, P., Zhang, J., Ou, Q., Tong, X., et al. (2020). Molecular and clinicopathological characteristics of ROS1-rearranged non-small-cell lung cancers identified by next-generation sequencing. *Mol. Oncol.* 14 (11), 2787–2795. doi:10.1002/1878-0261.12789

Dagogo-Jack, I., Rooney, M., Nagy, R. J., Lin, J. J., Chin, E., Ferris, L. A., et al. (2019). Molecular analysis of plasma from patients with ROS1-positive NSCLC. *J. Thorac. Oncol.* 14 (5), 816–824. doi:10.1016/j.jtho.2019.01.009

Drilon, A., Jenkins, C., Iyer, S., Schoenfeld, A., Keddy, C., and Davare, M. A. (2021). ROS1-dependent cancers - biology, diagnostics and therapeutics. *Nat. Rev. Clin. Oncol.* 18 (1), 35–55. doi:10.1038/s41571-020-0408-9

Drilon, A., Siena, S., Dziadziuszko, R., Barlesi, F., Krebs, M. G., Shaw, A. T., et al. (2020). Entrectinib in ROS1 fusion-positive non-small-cell lung cancer: integrated analysis of three phase 1-2 trials. *Lancet Oncol.* 21 (2), 261–270. doi:10.1016/S1470-2045(19)30690-4

Drilon, A., Siena, S., Ou, S. I., Patel, M., Ahn, M. J., Lee, J., et al. (2017). Safety and antitumor activity of the multitargeted pan-TRK, ROS1, and ALK inhibitor entrectinib: combined results from two phase I trials (ALKA-372-001 and STARTRK-1). *Cancer Discov.* 7 (4), 400–409. doi:10.1158/2159-8290.CD-16-1237

Drilon, A., Somwar, R., Wagner, J. P., Vellore, N. A., Eide, C. A., Zabriskie, M. S., et al. (2016). A novel crizotinib-resistant solvent-front mutation responsive to cabozantinib therapy in a patient with ROS1-rearranged lung cancer. *Clin. Cancer Res.* 22 (10), 2351–2358. doi:10.1158/1078-0432.CCR-15-2013

Dudnik, E., Agbarya, A., Grinberg, R., Cyjon, A., Bar, J., Moskovitz, M., et al. (2020). Clinical activity of brigatinib in ROS1-rearranged non-small cell lung cancer. *Clin. Transl. Oncol.* 22 (12), 2303–2311. doi:10.1007/s12094-020-02376-w

Dziadziuszko, R., Krebs, M. G., De Braud, F., Siena, S., Drilon, A., Doebele, R. C., et al. (2021). Updated integrated analysis of the efficacy and safety of entrectinib in locally advanced or metastatic ROS1 fusion-positive non-small-cell lung cancer. *J. Clin. Oncol.* 39 (11), 1253–1263. doi:10.1200/JCO.20.03025

Facchinetti, F., Loriot, Y., Kuo, M. S., Mahjoubi, L., Lacroix, L., Planchard, D., et al. (2016). Crizotinib-resistant ROS1 mutations reveal a predictive kinase inhibitor sensitivity model for ROS1- and ALK-rearranged lung cancers. *Clin. Cancer Res.* 22 (24), 5983–5991. doi:10.1158/1078-0432.CCR-16-0917

Fares, A. F., Lok, B. H., Zhang, T., Cabanero, M., Lau, S. C. M., Stockley, T., et al. (2020). ALK-rearranged lung adenocarcinoma transformation into high-grade large cell neuroendocrine carcinoma: clinical and molecular description of two cases. *Lung Cancer* 146, 350–354. doi:10.1016/j.lungcan.2020.06.005

Fielder, T., Butler, J., Tierney, G., Holmes, M., Lam, K. Y., Satgunaseelan, L., et al. (2022). ROS1 rearrangements in lung adenocarcinomas are defined by diffuse strong immunohistochemical expression of ROS1. *Pathology* 54 (4), 399–403. doi:10.1016/j.pathol.2021.07.012

Fu, S., Liang, Y., Lin, Y. B., Wang, F., Huang, M. Y., Zhang, Z. C., et al. (2015). The frequency and clinical implication of ROS1 and RET rearrangements in resected stage IIIA-N2 non-small cell lung cancer patients. *PLoS One* 10 (4), e0124354. doi:10.1371/journal.pone.0124354

Ginestet, F., Lambros, L., Le Flahec, G., Marcorelles, P., and Uguen, A. (2018). Evaluation of a dual ALK/ROS1 fluorescent *in situ* hybridization test in non-small-cell lung cancer. *Clin. Lung Cancer* 19 (5), e647–e653. doi:10.1016/j.cllc.2018.04.016

Gou, W., Zhou, X., Liu, Z., Wang, L., Shen, J., Xu, X., et al. (2018). CD74-ROS1 G2032R mutation transcriptionally up-regulates Twist1 in non-small cell lung cancer cells leading to increased migration, invasion, and resistance to crizotinib. *Cancer Lett.* 422, 19–28. doi:10.1016/j.canlet.2018.02.032

Guisier, F., Dubos-Arvis, C., Vinas, F., Doubre, H., Ricordel, C., Ropert, S., et al. (2020). Efficacy and safety of anti-PD-1 immunotherapy in patients with advanced NSCLC with BRAF, HER2, or MET mutations or RET translocation: GFPC 01-2018. *J. Thorac. Oncol.* 15 (4), 628–636. doi:10.1016/j.jtho.2019.12.129

Hegde, A., Hong, D. S., Behrang, A., Ali, S. M., Juckett, L., Meric-Bernstam, F., et al. (2019). Activity of brigatinib in crizotinib and ceritinib-resistant ROS1- rearranged non-small-cell lung cancer. *JCO Precis. Oncol.* 3, 1–6. doi:10.1200/po.18.00267

Hofman, V., Rouquette, I., Long-Mira, E., Piton, N., Chamorey, E., Heeke, S., et al. (2019). Multicenter evaluation of a novel ROS1 immunohistochemistry assay (SP384) for detection of ROS1 rearrangements in a large cohort of lung adenocarcinoma patients. *J. Thorac. Oncol.* 14 (7), 1204–1212. doi:10.1016/j.jtho.2019.03.024

Huang, L., Jiang, S., and Shi, Y. (2020). Tyrosine kinase inhibitors for solid tumors in the past 20 years (2001-2020). *J. Hematol. Oncol.* 13 (1), 143. doi:10.1186/s13045-020-00977-0

Huang, R. S. P., Gottberg-Williams, A., Vang, P., Yang, S., Britt, N., Kaur, J., et al. (2021b). Correlating ROS1 protein expression with ROS1 fusions, amplifications, and mutations. *JTO Clin. Res. Rep.* 2 (2), 100100. doi:10.1016/j.jtocrr.2020.100100

Huang, R. S. P., Haberberger, J., Sokol, E., Schrock, A. B., Danziger, N., Madison, R., et al. (2021a). Clinicopathologic, genomic and protein expression characterization of 356 ROS1 fusion driven solid tumors cases. *Int. J. Cancer* 148 (7), 1778–1788. doi:10.1002/ijc.33447

Katayama, R., Kobayashi, Y., Friboulet, L., Lockerman, E. L., Koike, S., Shaw, A. T., et al. (2015). Cabozantinib overcomes crizotinib resistance in ROS1 fusion-positive cancer. *Clin. Cancer Res.* 21 (1), 166–174. doi:10.1158/1078-0432.CCR-14-1385

Kim, S. W., Do, S. I., and Na, K. (2021). External validation of ALK and ROS1 fusions detected using an oncomine comprehensive assay. *Anticancer Res.* 41 (9), 4609–4617. doi:10.21873/anticanres.15274

Lambros, L., Guibourg, B., and Uguen, A. (2018). ROS1-rearranged non-small cell lung cancers with concomitant oncogenic driver alterations: about some rare therapeutic dilemmas. *Clin. Lung Cancer* 19 (1), e73–e74. doi:10.1016/j.cllc.2017.08.005

Li, N., Chen, Z., Huang, M., Zhang, D., Hu, M., Jiao, F., et al. (2022). Detection of ROS1 gene fusions using next-generation sequencing for patients with malignancy in China. *Front. Cell Dev. Biol.* 10, 1035033. doi:10.3389/fcell.2022.1035033

Li, Z., Lin, Y., Chi, X., Xu, M., and Wang, H. (2021). Appearance of an ALK mutation conferring resistance to crizotinib in non-small cell lung cancer harboring oncogenic ROS1 fusion. *Lung Cancer* 153, 174–175. doi:10.1016/j.lungcan.2021.01.001

Li, Z., Shen, L., Ding, D., Huang, J., Zhang, J., Chen, Z., et al. (2018). Efficacy of crizotinib among different types of ROS1 fusion partners in patients with ROS1-rearranged non-small cell lung cancer. *J. Thorac. Oncol.* 13 (7), 987–995. doi:10.1016/j.jtho.2018.04.016

Lin, J. J., Choudhury, N. J., Yoda, S., Zhu, V. W., Johnson, T. W., Sakhtemani, R., et al. (2021). Spectrum of mechanisms of resistance to crizotinib and lorlatinib in ROS1 fusion-positive lung cancer. *Clin. Cancer Res.* 27 (10), 2899–2909. doi:10.1158/1078-0432.CCR-21-0032

Lin, J. J., Langenbucher, A., Gupta, P., Yoda, S., Fetter, I. J., Rooney, M., et al. (2020). Small cell transformation of ROS1 fusion-positive lung cancer resistant to ROS1 inhibition. *NPJ Precis. Oncol.* 4, 21. doi:10.1038/s41698-020-0127-9

Lin, J. J., Ritterhouse, L. L., Ali, S. M., Bailey, M., Schrock, A. B., Gainor, J. F., et al. (2017). ROS1 fusions rarely overlap with other oncogenic drivers in non-small cell lung cancer. *J. Thorac. Oncol.* 12 (5), 872–877. doi:10.1016/j.jtho.2017.01.004

Lindeman, N. I., Cagle, P. T., Aisner, D. L., Arcila, M. E., Beasley, M. B., Bernicker, E. H., et al. (2018). Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the college of American pathologists, the international association for the study of lung cancer, and the association for molecular pathology. *Arch. Pathol. Lab. Med.* 142 (3), 321–346. doi:10.5858/arpa.2017-0388-CP

Makarem, M., Ezeife, D. A., Smith, A. C., Li, J. J. N., Law, J. H., Tsao, M. S., et al. (2021). Reflex ROS1 IHC screening with FISH confirmation for advanced non-small cell lung cancer-A cost-efficient strategy in a public healthcare system. *Curr. Oncol.* 28 (5), 3268–3279. doi:10.3390/curroncol28050284

Maroto, P., Porta, C., Capdevila, J., Apolo, A. B., Viteri, S., Rodriguez-Antona, C., et al. (2022). Cabozantinib for the treatment of solid tumors: a systematic review. *Ther. Adv. Med. Oncol.* 14, 17588359221107112. doi:10.1177/17588359221107112

Mazieres, J., Drilon, A., Lusque, A., Mhanna, L., Cortot, A. B., Mezquita, L., et al. (2019). Immune checkpoint inhibitors for patients with advanced lung cancer and oncogenic driver alterations: results from the IMMUNOTARGET registry. *Ann. Oncol.* 30 (8), 1321–1328. doi:10.1093/annonc/mdz167

Mazieres, J., Zalcman, G., Crino, L., Biondani, P., Barlesi, F., Filleron, T., et al. (2015). Crizotinib therapy for advanced lung adenocarcinoma and a ROS1 rearrangement: results from the EUROS1 cohort. *J. Clin. Oncol.* 33 (9), 992–999. doi:10.1200/JCO.2014.58.3302

Michels, S., Massuti, B., Schildhaus, H. U., Franklin, J., Sebastian, M., Felip, E., et al. (2019). Safety and efficacy of crizotinib in patients with advanced or metastatic ROS1-rearranged lung cancer (EUCROSS): a European phase II clinical trial. *J. Thorac. Oncol.* 14 (7), 1266–1276. doi:10.1016/j.jtho.2019.03.020

Mosele, F., Remon, J., Mateo, J., Westphalen, C. B., Barlesi, F., Lolkema, M. P., et al. (2020). Recommendations for the use of next-generation sequencing (NGS) for patients with metastatic cancers: a report from the ESMO Precision Medicine Working Group. *Ann. Oncol.* 31 (11), 1491–1505. doi:10.1016/j.annonc.2020.07.014

Nagarajan, L., Louie, E., Tsujimoto, Y., Balduzzi, P. C., Huebner, K., and Croce, C. M. (1986). The human c-ros gene (ROS) is located at chromosome region 6q16----6q22. *Proc. Natl. Acad. Sci. U. S. A.* 83 (17), 6568–6572. doi:10.1073/pnas.83.17.6568

Papadopoulos, K. P., Borazanci, E., Shaw, A. T., Katayama, R., Shimizu, Y., Zhu, V. W., et al. (2020). U.S. Phase I first-in-human study of taletrectinib (DS-6051b/AB-106), a ROS1/TRK inhibitor, in patients with advanced solid tumors. *Clin. Cancer Res.* 26 (18), 4785–4794. doi:10.1158/1078-0432.CCR-20-1630

Park, E., Choi, Y. L., Ahn, M. J., and Han, J. (2019). Histopathologic characteristics of advanced-stage ROS1-rearranged non-small cell lung cancers. *Pathol. Res. Pract.* 215 (7), 152441. doi:10.1016/j.prp.2019.152441

Peters, S., Shaw, A. T., Besse, B., Felip, E., Solomon, B. J., Soo, R. A., et al. (2020). Impact of lorlatinib on patient-reported outcomes in patients with advanced ALK-positive or ROS1-positive non-small cell lung cancer. *Lung Cancer* 144, 10–19. doi:10.1016/j.lungcan.2020.02.011

Ren, S., Huang, S., Ye, X., Feng, L., Lu, Y., Zhou, C., et al. (2021). Crizotinib resistance conferred by BRAF V600E mutation in non-small cell lung cancer harboring an oncogenic ROS1 fusion. *Cancer Treat. Res. Commun.* 27, 100377. doi:10.1016/j.ctarc.2021.100377

Rihawi, K., Cinausero, M., Fiorentino, M., Salvagni, S., Brocchi, S., and Ardizzoni, A. (2018). Co-Alteration of c-met and ROS1 in advanced NSCLC: ROS1 wins. *J. Thorac. Oncol.* 13 (3), e41–e43. doi:10.1016/j.jtho.2017.10.024

Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., et al. (2007). Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131 (6), 1190–1203. doi:10.1016/j.cell.2007.11.025

Rolfo, C., Ruiz, R., Giovannetti, E., Gil-Bazo, I., Russo, A., Passiglia, F., et al. (2015). Entrectinib: a potent new TRK, ROS1, and ALK inhibitor. *Expert Opin. Investig. Drugs* 24 (11), 1493–1500. doi:10.1517/13543784.2015.1096344

Roskoski, R., Jr. (2017). ROS1 protein-tyrosine kinase inhibitors in the treatment of ROS1 fusion protein-driven non-small cell lung cancers. *Pharmacol. Res.* 121, 202–212. doi:10.1016/j.phrs.2017.04.022

Saito, H., Fukuhara, T., Furuya, N., Watanabe, K., Sugawara, S., Iwasawa, S., et al. (2019). Erlotinib plus bevacizumab versus erlotinib alone in patients with EGFR-positive advanced non-squamous non-small-cell lung cancer (NEJ026): interim analysis of an open-label, randomised, multicentre, phase 3 trial. *Lancet Oncol.* 20 (5), 625–635. doi:10.1016/S1470-2045(19)30035-X

Shah, A. T., Bernardo, R. J., Berry, G. J., Kudelko, K., and Wakelee, H. A. (2021). Two cases of pulmonary tumor thrombotic microangiopathy associated with ROS1-rearranged non-small cell lung cancer. *Clin. Lung Cancer* 22 (2), e153–e156. doi:10.1016/j.cllc.2020.09.020

Shan, L., Lian, F., Guo, L., Qiu, T., Ling, Y., Ying, J., et al. (2015). Detection of ROS1 gene rearrangement in lung adenocarcinoma: comparison of IHC, FISH and real-time RT-PCR. *PLoS One* 10 (3), e0120422. doi:10.1371/journal.pone.0120422

Shaw, A. T., Kim, T. M., Crino, L., Gridelli, C., Kiura, K., Liu, G., et al. (2017). Ceritinib versus chemotherapy in patients with ALK-rearranged non-small-cell lung cancer previously given chemotherapy and crizotinib (ASCEND-5): a randomised, controlled, open-label, phase 3 trial. *Lancet Oncol.* 18 (7), 874–886. doi:10.1016/S1470-2045(17)30339-X

Shaw, A. T., Riely, G. J., Bang, Y. J., Kim, D. W., Camidge, D. R., Solomon, B. J., et al. (2019a). Crizotinib in ROS1-rearranged advanced non-small-cell lung cancer (NSCLC): updated results, including overall survival, from PROFILE 1001. *Ann. Oncol.* 30 (7), 1121–1126. doi:10.1093/annonc/mdz131

Shaw, A. T., Solomon, B. J., Chiari, R., Riely, G. J., Besse, B., Soo, R. A., et al. (2019b). Lorlatinib in advanced ROS1-positive non-small-cell lung cancer: a multicentre, open-label, single-arm, phase 1-2 trial. *Lancet Oncol.* 20 (12), 1691–1701. doi:10.1016/S1470-2045(19)30655-2

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *CA Cancer J. Clin.* 72 (1), 7–33. doi:10.3322/caac.21708

Song, Z., Zheng, Y., Wang, X., Su, H., Zhang, Y., and Song, Y. (2017). ALK and ROS1 rearrangements, coexistence and treatment in epidermal growth factor receptor-wild type lung adenocarcinoma: a multicenter study of 732 cases. *J. Thorac. Dis.* 9 (10), 3919–3926. doi:10.21037/jtd.2017.09.79

Sun, T. Y., Niu, X., Chakraborty, A., Neal, J. W., and Wakelee, H. A. (2019). Lengthy progression-free survival and intracranial activity of cabozantinib in patients with crizotinib and ceritinib-resistant ROS1-positive non-small cell lung cancer. *J. Thorac. Oncol.* 14 (2), e21–e24. doi:10.1016/j.jtho.2018.08.2030

Tang, Z., Zhang, J., Lu, X., Wang, W., Chen, H., Robinson, M. K., et al. (2018). Coexistent genetic alterations involving ALK, RET, ROS1 or MET in 15 cases of lung adenocarcinoma. *Mod. Pathol.* 31 (2), 307–312. doi:10.1038/modpathol.2017.109

Uguen, A., Schick, U., and Quere, G. (2017). A rare case of ROS1 and ALK double rearranged non-small cell lung cancer. *J. Thorac. Oncol.* 12 (6), e71–e72. doi:10.1016/j.jtho.2017.02.007

Wang, W., Cheng, G., Zhang, G., and Song, Z. (2020). Evaluation of a new diagnostic immunohistochemistry approach for ROS1 rearrangement in non-small cell lung cancer. *Lung Cancer* 146, 224–229. doi:10.1016/j.lungcan.2020.06.019

Wang, Y., Chen, Z., Han, X., Li, J., Guo, H., and Shi, J. (2021). Acquired MET D1228N mutations mediate crizotinib resistance in lung adenocarcinoma with ROS1 fusion: a case report. *Oncologist* 26 (3), 178–181. doi:10.1002/onco.13545

Watanabe, H., Ichihara, E., Kayatani, H., Makimoto, G., Ninomiya, K., Nishii, K., et al. (2021). VEGFR2 blockade augments the effects of tyrosine kinase inhibitors by inhibiting angiogenesis and oncogene signaling in oncogene-driven non-small-cell lung cancers. *Cancer Sci.* 112 (5), 1853–1864. doi:10.1111/cas.14801

Wiesweg, M., Eberhardt, W. E. E., Reis, H., Ting, S., Savvidou, N., Skiba, C., et al. (2017). High prevalence of concomitant oncogene mutations in prospectively identified patients with ROS1-positive metastatic lung cancer. *J. Thorac. Oncol.* 12 (1), 54–64. doi:10.1016/j.jtho.2016.08.137

Woodford, R., Lu, M., Beydoun, N., Cooper, W., Liu, Q., Lynch, J., et al. (2021). Disseminated intravascular coagulation complicating diagnosis of ROS1-mutant non-small cell lung cancer: a case report and literature review. *Thorac. Cancer* 12 (17), 2400–2403. doi:10.1111/1759-7714.14071

Wu, Y. L., Yang, J. C., Kim, D. W., Lu, S., Zhou, J., Seto, T., et al. (2018). Phase II study of crizotinib in east asian patients with ROS1-positive advanced non-small-cell lung cancer. *J. Clin. Oncol.* 36 (14), 1405–1411. doi:10.1200/JCO.2017.75.5587

Yang, J., Zhou, P., Yu, M., and Zhang, Y. (2021). Case report: high-level MET amplification as a resistance mechanism of ROS1-tyrosine kinase inhibitors in ROS1-rearranged non-small cell lung cancer. *Front. Oncol.* 11, 645224. doi:10.3389/fonc.2021.645224

Yun, M. R., Kim, D. H., Kim, S. Y., Joo, H. S., Lee, Y. W., Choi, H. M., et al. (2020). Repotrectinib exhibits potent antitumor activity in treatment-naive and solvent-front-mutant ROS1-rearranged non-small cell lung cancer. *Clin. Cancer Res.* 26 (13), 3287–3295. doi:10.1158/1078-0432.CCR-19-2777

Zeng, L., Li, Y., Xiao, L., Xiong, Y., Liu, L., Jiang, W., et al. (2018). Crizotinib presented with promising efficacy but for concomitant mutation in next-generation sequencing-identified ROS1-rearranged non-small cell lung cancer. *Onco Targets Ther.* 11, 6937–6945. doi:10.2147/OTT.S176273

Zhang, Q., Wu, C., Ding, W., Zhang, Z., Qiu, X., Mu, D., et al. (2019). Prevalence of ROS1 fusion in Chinese patients with non-small cell lung cancer. *Thorac. Cancer* 10 (1), 47–53. doi:10.1111/1759-7714.12899

Zhu, Q., Zhan, P., Zhang, X., Lv, T., and Song, Y. (2015). Clinicopathologic characteristics of patients with ROS1 fusion gene in non-small cell lung cancer: a meta-analysis. *Transl. Lung Cancer Res.* 4 (3), 300–309. doi:10.3978/j.issn.2218-6751.2015.05.01

Zhu, V. W., Zhao, J. J., Gao, Y., Syn, N. L., Zhang, S. S., Ou, S. H. I., et al. (2021). Thromboembolism in ALK+ and ROS1+ NSCLC patients: a systematic review and meta-analysis. *Lung Cancer* 157, 147–155. doi:10.1016/j.lungcan.2021.05.019

Zhu, Y. C., Xu, C. W., Ye, X. Q., Yin, M. X., Zhang, J. X., Du, K. Q., et al. (2016). Lung cancer with concurrent EGFR mutation and ROS1 rearrangement: a case report and review of the literature. *Onco Targets Ther.* 9, 4301–4305. doi:10.2147/OTT.S109415

Zito Marino, F., Rossi, G., Cozzolino, I., Montella, M., Micheli, M., Bogina, G., et al. (2020). Multiplex fluorescence *in situ* hybridisation to detect anaplastic lymphoma kinase and ROS proto-oncogene 1 receptor tyrosine kinase rearrangements in lung cancer cytological samples. *J. Clin. Pathol.* 73 (2), 96–101. doi:10.1136/jclinpath-2019-206152

Check for updates

*CORRESPONDENCE
Jing Zhang
✉ zjingmo@163.com

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

Analysis of influencing factors of serum SCCA elevation in 309 CAP patients with normal CEA, NSE and CYFRA21-1

Jinghan Wang[1†], Xiao Tang[2†], Xin Liu[1] and Jing Zhang[3*]

[1]Department of Clinical Laboratory, the Second Hospital of Dalian Medical University, Dalian, China,
[2]Department of Health Statistics, School of Public Health, Dalian Medical University, Dalian, China,
[3]Respiratory Department, the Second Hospital of Dalian Medical University, Dalian, China

Introduction: Squamous cell carcinoma antigen (SCCA) is one of the auxiliary diagnostic indicators of lung squamous cell carcinoma, and an increase in serum SCCA can predict the occurrence of lung squamous cell carcinoma. However, whether SCCA is also elevated in pneumonia patients without malignancy is still not clear. Therefore, we studied influencing factors of elevated serum SCCA in patients with community-acquired pneumonia.

Methods: We retrospectively enrolled 309 patients who were admitted to the Respiratory department with normal serum Carcinoembryonic antigen (CEA), Neuron specific enolase (NSE), and Cytokeratin 19 fragment (CYFRA21-1) level and were diagnosed with community-acquired pneumonia (CAP). The patients' serum SCCA level, body temperature, age, sex, white blood cell (WBC) count, hypersensitive C-reactive protein (Hs-CRP) level, and serum amyloid A (SAA) were recorded. Logistic regression models were used to analyze the risk factors of SCCA elevation. The dose-response relationship between temperature and risk of SCCA increase was analyzed using Restricted cubic splines (RCS).

Results: Of the 309 patients, 143(46.3%) showed elevated SCCA levels. The logistic regression analysis revealed a significant influence of age and body temperature on elevated SCCA (*P*<0.05) levels. For every one-year increase in age, the probability of elevated SCCA decreased by 3% [OR=0.97,95% CI:0.95,0.99].For every 1°C increase in body temperature, the risk of elevated SCCA increased by 2.75 times [OR=3.75,95%CI:2.55,5.49].The patients were sorted into quartiles based on body temperature. Compared with patients in the Q1 of body temperature group, patients in the Q3 group were at 7.92 times higher risk [OR=7.92, 95%CI:3.27,19.16].and the risk of elevated SCCA was increased by 22.85 times in the Q4 group [OR=23.85,95% CI:8.38,67.89] after adjusting for age, gender, Hs-CRP, SAA, and WBC. RCS analysis showed there was a linear relationship between temperature index and risk of elevated SCCA.

Conclusion: In summary, for CAP patients with normal CEA,NSE and CYFRA21-1 level, age and body temperature are influencing factors of SCCA elevation. Higher body temperature has a strong association with the occurrence of SCCA elevation.

Introduction

Lung cancer is one of the most common malignant tumors in the world. As the early symptoms of lung cancer are not obvious, most patients are diagnosed at an advanced stage, and the prognosis is poor (1). Pneumonia is a common benign disease in the respiratory department, and patients often have symptoms of fever, cough, sputum, hemoptysis and other clinical manifestations. Pneumonia can be cured with active anti-infective treatment. The early clinical symptoms of lung cancer patients are similar to pneumonia and lack specificity. For hospitalized pneumonia patients, in addition to anti-infective treatment, the exclusion of lung cancer has become one of the most important purposes of hospitalization for them.

SCCA is a tumor-specific antigen that was first discovered in the 1970s by Kato and Torigoe from cervical squamous cell carcinoma tissues (2). It is widely found in the cytoplasm of squamous cell carcinomas, such as the uterus, cervix, lung, head and neck, especially in nonkeratinized cancer cells, in which the content of SCCA is more abundant (3–6). SCCA exists in squamous epithelial cells is involved in the differentiation of squamous epithelium and tumor growth of tumor cells, and is often used in the diagnosis of squamous epithelium-derived carcinoma (7). SCCA is a specific marker of squamous cell carcinoma and is an independent prognostic factor for cervical squamous cell carcinoma (1).

As one of the tumor markers, SCCA alone does not confirm the tumor, but it can provide valuable information for the diagnosis and prognosis assessment of various malignant tumors, such as in the preconditioning evaluation of the tumor scope, evaluation of the treatment response and prediction of prognosis (4). As a tumor marker of various squamous cell cancers (esophageal cancer, lung cancer, head and neck cancer, anal canal cancer, cervical cancer, etc.),SCCA is of great significance in reflecting tumor stage, tumor size, interstitial infiltration, status of the lymphatic vascular space, and status of lymph nodes (8). Therefore, SCCA is particularly suitable for detecting squamous cell carcinoma (6, 7). Meanwhile, its elevation is also a poor prognostic factor for squamous cell carcinoma (4). Sun et al. (3) has shown that SCCA can initially evaluate the radiotherapy effect of lung cancer patients and has a certain predictive effect on prognosis. Yang et al. (1) has also pointed out that SCCA is important for the assessment of prognosis and survival assessment of lung cancer. An earlier article pointed out (9) that a high serum SCCA level is an independent poor prognostic factor in patients with peripheral squamous cell carcinoma.

For the past few years, there have been an increasing number of studies on SCCA in nonneoplastic diseases. In addition to tumors, elevated serum SCCA levels have also been detected in chronic liver disease, pulmonary infiltration with eosinophilia, renal insufficiency, and chronic inflammatory skin diseases (such as psoriasis, pemphigus, or eczema) (4, 8, 10–12). The serum SCCA level of diabetic nephropathy patients is significantly higher than that of normal proteinuria patients and healthy controls (10). A study has shown that the SCCA level is closely related to patients with varus papilloma, and SCCA may have the potential to be a useful biomarker for patients with varus papilloma (5).More articles have shown that SCCA is also elevated in some nonmalignant pulmonary diseases (2, 5). For example, there has been a case report showing that influenza B virus infection can lead to increased SCCA (13). Many studies have shown that the SCCA level also increases in bronchial asthma (13–15).

However, the elevation of SCCA in patients with pneumonia has not been studied. In clinical work, we found that some patients with pneumonia also had elevated serum SCCA, whereas other tumor markers, such as CEA, NSE, and CYFRA21-1,were normal, and a series of subsequent clinical examination could rule out the diagnosis of lung cancer. Therefore, in this study, we selected some patients whose CEA,NSE and CYFRA21-1 levels were normal and who had lung cancer excluded. We detected their SCCA levels, analyzed the relationship between SCCA concentration and body temperature, age and inflammatory indicators. The aim of the present study is to investigate the influencing factors of serum

---

Abbreviations: SCCA, Squamous cell carcinoma antigen; CAP, Community acquired pneumonia; CEA, Carcinoembryonic antigen; NSE, Neuron specific enolase ; Cyfra21-1, Cytokeratin 19 fragment; WBC, White blood cell; Hs-CRP, Hypersensitive C-reactive protein; SAA, Serum amyloid A; RCS, Restricted cubic splines; ILCs, Innate lymphoid cells; TRPM8, Transient receptor potential cation channel subfamily M member 8; SERPINB, the Clade B of serine protease inhibitors family.

SCCA elevation in CAP patients with normal CEA,NSE and CYFRA21-1. Identification of these factors, which are helpful to improve the health education for CAP patients and reduce unnecessary examinations, is highly warranted.

Methods

Study subjects and exclusion criteria

A total of 1,300 patients diagnosed with pneumonia after admission to the Respiratory Department of the Second Affiliated Hospital of Dalian Medical University from January 2019 to December 2020 were considered for this study.

The patients included in the study were selected based on the following exclusion criteria:

(1) age ≥80 years and ≤15 years old;

(2) bronchial asthma;

(3) acute/chronic bronchitis, interstitial pneumonia, tuberculosis, bronchiectasis, and chronic obstructive pulmonary disease;

(4) patients with lung cancer and undiagnosed lung nodules;

(5) other malignant tumors;

(6) CEA,NSE, and CYFR21-1 tests were not performed during hospitalization;

(7) any of the serum CEA,NSE and CYFR21-1 test results was abnormal;

(8) pregnancy.

The diagnostic criteria of pneumonia were based on the Chinese Guidelines for Diagnosis and Treatment of Community-acquired Pneumonia (2016 edition) (16): 1. community disease onset; 2. showing the following related clinical manifestations of pneumonia: (i) recent cough, expectoration, or existing respiratory disease symptoms with or without purulent sputum/chest pain/dyspnea/hemoptysis; (ii) fever; (iii) pulmonary consolidation signs and/or wet rales; (iv) peripheral WBCs>10×109/L,OR<4×109/L with or without a neutrophilic left shift; and 3. chest imaging revealing new patchy infiltration, leaf/segment contrast, ground glass shadow, or interstitial changes with or without pleural effusion. Patients showing criteria 1 and 3 and any one of the 2 criteria were diagnosed with pneumonia, except for those with pulmonary tuberculosis, pulmonary tumor, noninfectious interstitial disease, pulmonary edema, atelectasis, pulmonary embolism, pulmonary eosinophil infiltration, and pulmonary vasculitis.

Ultimately, a total of 309 cases were enrolled, including 135 males and 174 females,143 with normal SCCA and 166 with elevated SCCA. According to the statistical rule, the number of cases with elevated SCCA was at least 10 times that of the included variables in the regression analysis. A total of 9 independent variables were included in this study, so the sample size requirement for statistical analysis was met.

Measurement

For quantitative detection of serum SCCA, 3 mL of fasting venous blood was drawn, and a chemiluminescence immunoassay was performed using the MAGLUMI 2000 automatic chemiluminescence instrument. The immunoassay was performed with the reagents from the instrument's supporting kit and operated in strict accordance with the manufacturer's instructions.

The serum SCCA reference interval was 0 ng/mL–2.5 ng/mL. "SCCA elevation" refers to a serum SCCA level >2.5 ng/mL.

WBC count was measured by a Sysmex XN 9000 automatic hematology analyzer using flow cytometry, and other inflammation factors, such as Hs-CRP and SAA levels, were detected with a chemiluminescence immunoassay.

In this study, body temperature was measured by a thermometer reading, and "body temperature increase" refers to a body temperature >37.2°C. "Wheezing" refers to the auscultation of both lungs with rhonchi.

Statistical methods

Patients with pneumonia were divided into two groups according to the SCCA detection value: the normal SCCA group (SCCA ≤ 2.5 ng/mL) and the elevated SCCA group (SCCA>2.5 ng/mL). The chi-square test was used to compare the differences between the enumeration data of two groups,i.e., age, sex, body temperature (normal/increased),and wheezing (yes/no). WBC, hs-CRP,and SAA data showed abnormal distribution expressed as the median M (P25,P75) and a comparison between the two groups was performed with a rank sum test of two independent samples.

Patients with pneumonia were grouped into quartiles from small to large according to SCCA levels. We used the chi-square test for categorical variables (age, sex, fever) and a multigroup rank sum test for group comparisons of continuous-type variables. Logistic regression was used twice to analyze the risk factors for SCCA elevation. First, the independent variables selected were all the observed indicators to determine risk factors. Second, body temperature was divided into quartiles, and risks of SCCA elevation for each quartile were evaluated by setting the lowest quartiles of body temperature as the reference group. Model 1 was adjusted for age and gender and BMI, and model 2 was further adjusted for Hs-CRP, SAA, and WBC. Restricted cubic spline (RCS) analysis was applied to analyze the dose–response relationship between the temperature and the risk of elevated SCCA.

Statistical analysis was performed using SPSS 26.0. $P<0.05$ was considered significant. The RCS analysis was performed in R 4.0.3.

Results

Patient demographics

Ultimately, 309 CAP patients,135 males and 174 females, whose serum CEA, NSE, and CYFRA21-1 levels were normal met our enrolment criteria.

For analysis, the patients were grouped based on the level of serum SCCA and body temperature. The serum SCCA level was ≤2.5 ng/mL in 143 (46.28%) patients and >2.5 ng/mL in 166 (53.72%) patients. Among the 309 patients,125 (40.45%) had a normal body temperature, and 184 (59.55%) had a fever.

Comparison of basic conditions and inflammatory indicators of pneumonia patients in two groups

Patients in the elevated SCCA group were younger than those in the normal SCCA group but had a higher probability of wheezing and increased body temperature ($P<0.05$). Moreover, the inflammation indices, such as Hs-CRP and SAA, were significantly higher than those in the normal group ($P<0.05$) (Table 1). There was no significant difference between sex and WBC count in the two groups.

Comparison of patient characteristics according to serum SCCA levels

The patient groups were ranked from small to large according to SCCA levels. The top 25% of patients were in Group 1,the patients in the top 25%–50% were Group 2,the patients in the top 50%–75% were in Group 3,and the patients in the top >75% were in Group 4.

Our analysis showed statistically significant differences in age, fever, Hs-CRP, and SAA levels among the different groups. Patients in Group 4 had the highest serum SCCA level, while their age was significantly lower. In contrast, the proportions of patients with fever and the Hs-CRP, and SAA levels were significantly higher

than those in the other three groups ($P<0.05$) (Table 2). The trend increased with the SCCA level. No significant difference was found among the four groups between wheezing ($P=0.075$) and WBC count ($P=0.103$).

Factors impacting elevated SCCA levels

Logistic regression was used for the first analysis of risk factors for increased SCCA. The results showed that age and body temperature were associated with elevated SCCA levels in all the observed measures ($P<0.05$). Age was a protective factor against SCCA elevation, while body temperature was a risk factor for SCCA elevation. The probability of SCCA increasing decreased by 3% with an age increase of one year. In contrast, the risk of SCCA increased by 2.75 times with a temperature increase of 1°C (Table 3).

Effects of body temperature on serum SCCA elevation

For the second step of regression analysis, body temperature was divided into four groups as an independent variable for analysis. The logistic regression results showed that the risk of SCCA value elevation was 7.92 times higher in Q3 patients than that in Q1 patients ($P<0.05$) and the risk of SCCA elevation in Q4 patients was 22.85 times higher than that in Q1 patients ($P<0.05$). For every 1°C increase in body temperature, the risk of elevated SCCA increased by 2.75 times[OR=3.75,95%CI:2.55,5.49] (Table 4). The RCS analysis showed a linear relationship between the temperature index and the risk of elevated SCCA (Pnonlinear =0.7371, Figure 1).

TABLE 1 General Characteristics of participants.

Indices	Normal SCCA (N=166)	Elevated SCCA (N=143)	Z value	P value
Age	60 (46,65)	50 (33,63)	3.96	*<0.05*
Gender N(%)				
Male	65 (39.2%)	70 (49.0%)	2.99	0.083
Female	101 (60.8%)	73 (51.0%)		
Wheezing N(%)				
Yes	135 (81.3%)	130 (90.9%)	5.78	*<0.05*
No	31 (18.7%)	13 (9.1%)		
Body temperature N(%)				
Normal	107 (64.5%)	18 (12.6%)	85.81	*<0.05*
Increase	59 (35.5%)	125 (87.4%)		
WBC (*10^9/L)	7.70 (6.09, 10.24)	7.83 (5.82, 9.44)	1.23	0.220
Hs-CRP (mg/L)	11.69 (2.03, 44.98)	33.70 (13.40, 74.55)	4.82	*<0.05*
SAA (mg/L)	31.46 (9.33, 177.10)	179.38 (47.95, 237.27)	4.97	*<0.05*

Data are the mean number (percentage) or median (P_{25},P_{75}).

TABLE 2 Comparison of participants' characteristics with the serum SCCA levels.

Indices	SCCA,ng/ml				Z value	P value
	Q1 (<33.00)	Q2 (33.00-50.00)	Q3 (50.00-63.00)	Q4 (>63.00)		
Age	60 (48, 67)	60 (45, 65)	60 (45, 65)	50 (32, 64)	18.36	<0.001
Wheezing N(%)						
Yes	64 (80%)	61 (81.3%)	60 (89.6%)	80 (92.0%)	6.89	0.075
Temperature (°C)	36.8 (36.7, 37.7)	36.8 (36.7, 38.0)	38.4 (37.4, 39.0)	38.9 (38.0, 39.3)	85.64	<0.05
Fever N (%)						
Yes	24 (30.0%)	30 (40.0%)	51 (76.1%)	79 (90.8%)	83.82	<0.05
WBC(*10^9/L)	7.46 (5.85, 9.71)	8.31 (6.00, 11.29)	7.87 (6.52, 9.86)	7.52 (4.98, 8.98)	61.19	0.103
Hs-CRP (mg/L)	8.57 (1.56, 31.55)	15.64 (2.12, 55.07)	31.38 (8.81, 72.65)	34.50 (13.61, 74.58)	22.06	<0.05
SAA (mg/L)	29.29 (9.22, 163.03)	45.66 (9.07, 176.05)	160.00 (23.02, 222.31)	192.02 (79.47, 255.13)	26.98	<0.05

Data are the mean number (percentage) or median (P_{25},P_{75}).

Discussion

SCCA is originally purified from cervical squamous cell carcinoma (17, 18) and it is widely expressed in tongue, tonsil, esophagus, cervix, vagina, trachea, skin and other normal tissues (14, 17, 19). SCCA can be used as a diagnostic marker for cervical cancer, lung cancer, esophageal cancer, head and neck cancer and other squamous cell carcinomas (8, 14, 17, 18). Studies have shown that the sensitivity of SCCA in the diagnosis of non-small cell lung cancer (NSCLC) is 17%,and the specificity can reach 95%. The sensitivity and specificity of diagnosis in lung squamous cell carcinoma are 95% and 32% (20). Ando et al. measured seven serum biomarkers in 312 NonSmall Cell Lung Cancer (NSCLC) patients and found that SCCA had the highest positive rate at 55.4%,followed by CYFRA 21-1 (48.2%) (4). Clinically, SCCA is often used in combination with CYFRA 21-1,CEA,and NSE to screen for early stage of lung cancer. In recent years, a few studies and case reports have found that SCCA levels are elevated in benign lung diseases, such as asthma (8, 14, 15, 21), influenza virus infection (13), respiratory syncytial virus infection (22),

TABLE 3 The risks factors of elevated SCCA levels by a logistic model.

Indices	B value	OR (95% CI)	P value
Age	-0.028	0.97 (0.95, 0.99)	<0.05
Gender (male/female)	-0.594	0.52 (0.29, 1.05)	0.071
Wheezing (Yes/No)	-2.400	0.79 (0.31, 1.99)	0.613
Temperature (°C)	1.320	3.75 (2.55, 5.49)	<0.05
WBC (*10^9/L)	-0.088	0.92 (0.82, 1.02)	0.104
Hs-CRP (mg/L)	-0.005	0.99 (0.99, 1.01)	0.368
SAA (mg/L)	0.001	1.00 (0.99, 1.00)	0.682

tuberculosis (23), pulmonary sarcoidosis (23), pulmonary fibrosis (24), critically ill patients with COVID-19 (25), eosinophilic pulmonary infiltration (4, 26) and other diseases. The mechanism of the serum SCCA increased in patients with allergic diseases has been extensively studied (8, 27). For example, in atopic dermatitis and bronchial asthma patients, increased SCCA is caused by IL-4 and IL-13, which are secreted by Th2 cells (15, 17, 28, 29). These cytokines act on skin keratinocytes or bronchial epithelial cells (8), inducing high expression of SCCA in airway epithelial cells and/ or keratinocytes.

The present study is a retrospectively clinical investigation of the serum SCCA levels of 309 CAP patients with normal serum CEA,CYFRA21-1 and NSE. The impact factors of SCCA elevation was analyzed by grouping the patients into quartiles according to SCCA and body temperature, respectively. To the best of our knowledge, this study may be the first to report that body temperature and age influence the elevation of serum SCCA in patients with community-acquired pneumonia (CAP). The main finding of the study was that age was a protective factor against SCCA increase, while body temperature was a risk factor for SCCA value elevation. There was a linear relationship between temperature index and risk of elevated SCCA.

Why does body temperature elevation lead SCCA increase in CAP patients? It is well known that fever, as an important host defense mechanism, is accomplished by integrated physiological and neural circuits (30). Transient receptor potential cation channel subfamily M member 8 (TRPM8) is a kind of cold-sensing neuron expressed in the nerve endings of sensory neurons and keratinocytes in the epidermis of skin. Its activation induces a series of cold defenses, such as brown adipose tissue thermogenesis, shivering thermogenesis and skin vasoconstriction (31). As the outermost barrier tissue of human body, skin is composed of epidermis, dermis and subcutis (32). Meanwhile, as a major immune organ, skin contains a large number of type 2 innate

TABLE 4   The association of body temperatures with SCCA elevation by logistic regression analyses.

| Indices	Model 1			Model 2		
	B value	OR (95% CI)	P value	B value	OR (95% CI)	P value
Body temperatures (continuous)	1.29	3.63 (2.69, 4.88)	<0.05	1.32	3.75 (2.55, 5.49)	**<0.05**
Quartiles of body temperatures						
Q1 (≤36.7°C)	–	–	–	–	–	–
Q2 (36.8°C-37.9°C)	-0.01	0.99 (0.43, 2.26)	0.98	0.38	1.46 (0.60, 3.56)	0.40
Q3 (38.0°C-38.8°C)	2.01	7.42 (3.55, 15.53)	<0.05	2.07	7.92 (3.27, 19.16)	<0.05
Q4 (38.9°C-42.0°C)	3.09	21.86 (9.22, 51.81)	<0.05	3.17	23.85 (8.38,67.89)	<0.05

Model 1: Adjusted for: age and gender.
Model 2: Adjusted for: age, gender, Hs-CRP, SAA, and WBC.

lymphoid cells (ILC2s) (32, 33). ILC2s in the dermal skin are activated by signals from cold-sensing neurons (31), which respond to the changes in ambient temperature to help regulate thermal homeostasis in the skin. It has been reported that the activation of TRPM8 can promote thermogenesis, and dermal ILC2s are activated by stimulating TRPM8 (31). The study also found that TRPM8 expressed in dermis was partially responsible for the activation of skin-resident ILC2s,and TRPM8,which is expressed in epidermal keratinocytes, may also be involved in sensing ambient temperature to promote local ILC2s activation (31). ILC2s express the transcription factors GATA 3 and RORα during development which are capable of producing type 2 cytokines such as IL-4,IL-5,and IL-13 (31, 34). Stimulated by

interleukin-4 (IL-4) or interleukin-13 (IL-13), keratinocytes can secrete SCCA1 and SCCA2 (17, 21).

Thus, we speculate when a CAP patient had a fever, the skin cold receptors could activate skin-resident ILC2s to secrete IL-4,IL-5 and IL-13.Then,the cytokines trigger keratinocytes to secrete SCCA.

Furthermore, does SCCA directly mediate the inflammatory response to pneumonia? Studies showed that SCCA could also be detected in the serum of patients with lung diseases such as bronchitis and pneumonia (35). These findings raise the possibility that SCCA may act as a marker for certain inflammation. A paper also mentioned that the Clade B of serine protease inhibitors family 3(SERPINB3)and the Clade B of serine

FIGURE 1
Dose-response relationship between temperature index and SCCA.
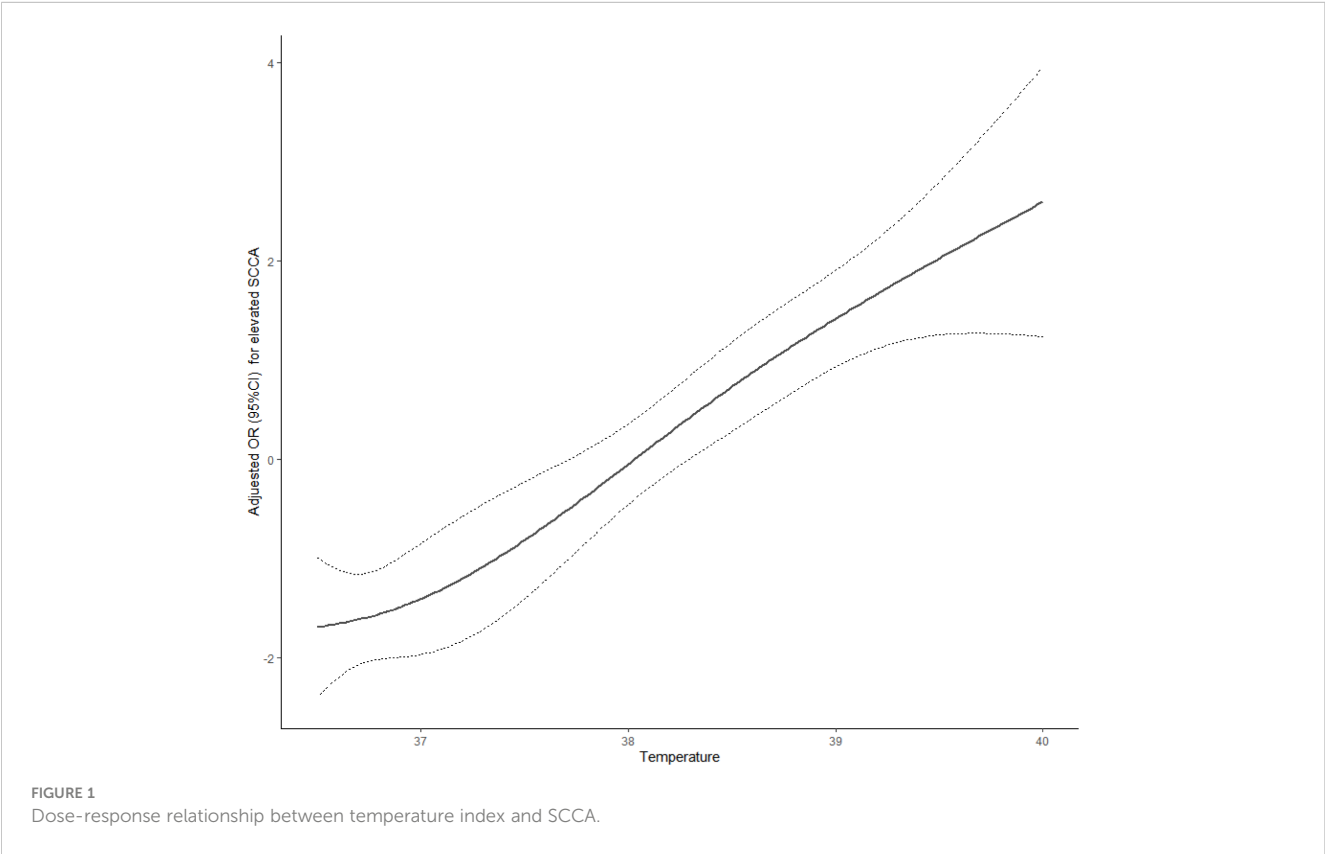
protease inhibitors family 4 (SERPINB4), also known as squamous cell carcinoma antigen-1 and -2(SCCA1/2). Elevated levels of these inhibitors were detected during inflammation, which may indicate that they are upregulated to help suppress the inflammatory response. Similarly, their overexpression may indicate that they can cause a pro-inflammatory response (36). In our study, the inflammation indicators, such as Hs-CRP and SAA, were significantly higher than those in the normal group (P<0.05).However, in the Logistic regression analysis of risk factors for increased SCCA, the P values of them did not reach the statistical significance (P= 0.065).Therefore, we consider that SCCA might be involved in mediating the inflammatory response in patients with pneumonia, but we have not conducted relevant experiments and in-depth research analysis, and more studies are needed to confirm this situation in the future.

Interestingly, this study also found that, unlike the effect body temperature, increase of age was a protective factor for SCCA elevation. It means that the risk of SCCA elevation decreases with increase of age. We speculate that it might be related with skin aging. The skin exerts its barrier function through a multilayered structure comprised of three distinct anatomical compartments: epidermis, dermis, and subcutis (32). Keratinocytes exist in the epidermis (37) and ILC2s are distributed in the dermis (32, 33). Considered both physiologic and inevitable, skin aging is a degenerative phenomenon (38). Studies (39) have shown there is a functional difference in the stratum corneum of young versus old skin because recovery of aged skin from insults to this layer are significantly slower than those seen in young skin, and permeability to certain substances is altered. We suggested that the increased risk of SCCA decreased associated with age may be due to the decline in secretory function of skin cells with age increasing. That's just our conjecture, we need to expand the sample size and data sources, reduce confounding bias, and conduct more in-depth studies.

There are some limitations in our study. First, 309 patients were enrolled in this study, of which 143 were in the SCCA elevated group. Although the sample size was small, the sample size met the statistical requirements. It is undeniable that insufficient sample size may introduce significant interference in the use of statistical models with numerous covariates. Since this is a single-center study, it is necessary to conduct multi-center studies in the future to expand the sample size and make the conclusions more reliable. Second, for CAP patients, we did not further group them by pathogen to analyze whether they were associated with increased SCCA. It is because the positive rate of pathogenic examination is generally low, and the amount of pathogen examination performed in our patients is small. This information cannot be obtained. Third, we do not conduct the in-depth studies on SCCA whether or not mediates inflammation. We can design more experiments in the future to explain this question.

In summary, we find that for CAP patients with normal serum CEA, NSE, and CYFRA21-1, body temperature and age are significantly correlated with increased SCCA. Increased body temperature is a risk factor for SCCA elevation, while increasing age is a protective factor for SCCA elevation. For CAP patients with elevated SCCA, appropriate explanations can be done to reduce the anxiety of patients and their families about lung cancer.

Data availability statement

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

Ethics statement

The studies involving humans were approved by the Second Affiliated Hospital of Dalian Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

JZ and JW collected data, reviewed literature, and wrote manuscripts. XT was responsible for the statistical analysis. XL was in charge of data collection, literature review. All authors read and approved the final manuscript.

Funding

Conflict of interest

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

Publisher's note

References

1. Yang Q, Zhang P, Wu R, Lu K, Zhou H. Identifying the best marker combination in CEA, CA125, CY211, NSE, and SCC for lung cancer screening by combining ROC curve and logistic regression analyses: is it feasible? *Dis Markers* (2018) 2018:2082840. doi: 10.1155/2018/2082840

2. Zhu H. Squamous cell carcinoma antigen: clinical application and research status. *Diagnostics* (2022) 12(5):1065. doi: 10.3390/diagnostics12051065

3. Sun L, Shao Q. Expression changes and clinical significance of serum neuron-specific enolase and squamous cell carcinoma antigen in lung cancer patients after radiotherapy. *Clinics* (2023) 78:100135. doi: 10.1016/j.clinsp.2022.100135

4. Nakamura H, Nishimura T. History, molecular features, and clinical importance of conventional serum biomarkers in lung cancer. *Surg Today* (2017) 47:1037–59. doi: 10.1007/s00595-017-1477-y

5. Promsopa C, Suwansri S, Khuntikij P. The serum squamous cell carcinoma antigen level in inverted sinonasal papilloma and nasal polyps patients. *World J otorhinolaryngology-head Neck Surg* (2021) 7(1):23–7. doi: 10.1016/j.wjorl.2020.02.002

6. Su H, Liu K, Zhao Y, Shi F, Li Y, Wu J, et al. High serum squamous cell carcinoma antigen level associated with remission of mild/moderate dysplasia of the esophagus: A nested case–control study. *Gastroenterol Res Pract* (2022) 2022:2961337. doi: 10.1155/2022/2961337

7. Li J, Chen Y, Wang X, Wang C, Xiao M. The value of combined detection of CEA, CYFRA21-1, SCC-Ag, and pro-GRP in the differential diagnosis of lung cancer. *Trans Cancer Res* (2021) 10(4):1900. doi: 10.21037/tcr-21-527

8. Izuhara K, Yamaguchi Y, Ohta S, Nunomura S, Nanri Y, Azuma Y, et al. Squamous cell carcinoma antigen 2 (SCCA2, SERPINB4): an emerging biomarker for skin inflammatory diseases. *Int J Mol Sci* (2018) 19(4):1102. doi: 10.3390/ijms19041102

9. Kinoshita T, Ohtsuka T, Hato T, Goto T, Kamiyama I, Tajima A, et al. Prognostic factors based on clinicopathological data among the patients with resected peripheral squamous cell carcinomas of the lung. *J Thorac Oncol* (2014) 9(12):1779–87. doi: 10.1097/JTO.0000000000000338

10. Chen J, Tao F, Zhang B, Chen Q, Qiu Y, Luo Q, et al. Elevated squamous cell carcinoma antigen, cytokeratin 19 fragment, and carcinoembryonic antigen levels in diabetic nephropathy. *Int J Endocrinol* (2017) 2017:5304391. doi: 10.1155/2017/5304391

11. Nomura F, Koyama A, Ishijima M, Takano S, Narita M, Nakai T. Serum levels of five tumor markers for lung cancer in patients with chronic renal failure. *Oncol Rep* (1998) 5(2):389–481. doi: 10.3892/or.5.2.389

12. Shimomura M, Okura Y, Takahashi Y, Kobayashi I. A serum level of squamous cell carcinoma antigen as a real-time biomarker of atopic dermatitis. *Pediatr Allergy Immunol Pulmonol* (2021) 34(3):102–5. doi: 10.1089/ped.2021.0049

13. Sano A. Transient elevation of squamous cell carcinoma antigen levels with influenza virus infection. *Respirol Case Rep* (2018) 6(8):e00362. doi: 10.1002/rcr2.362

14. Sun Z, Shi X, Wang Y, Zhao Y. Serum squamous cell carcinoma antigen in psoriasis: a potential quantitative biomarker for disease severity. *Dermatology* (2018) 234(3-4):120–6. doi: 10.1159/000488672

15. Suzuki K, Inokuchi A, Miyazaki J, Kuratomi Y, Izuhara K. Relationship between squamous cell carcinoma antigen and the clinical severity of allergic rhinitis caused by Dermatophagoides farinae and Japanese cedar pollen. *Ann Otol Rhinol Laryngol* (2010) 119(1):22–6. doi: 10.1177/000348941011900104

16. Association RSoCM. Guidelines for the diagnosis and treatment of community-acquired pneumonia in Chinese adults. *Clin J Tuberc Respir Dis* (2016) 39(4):253–79. doi: 10.3760/cma.j.issn.1001-0939.2016.03.000

17. Izuhara K, Ohta S, Kanaji S, Shiraishi H, Arima K. Recent progress in understanding the diversity of the human ov-serpin/clade B serpin family. *Cell Mol Life Sci* (2008) 65:2541–53. doi: 10.1007/s00018-008-8049-7

18. Charakorn C, Thadanipon K, Chaijindaratana S, Rattanasiri S, Numthavaj P, Thakkinstian A. The association between serum squamous cell carcinoma antigen and recurrence and survival of patients with cervical squamous cell carcinoma: a systematic review and meta-analysis. *Gynecologic Oncol* (2018) 150(1):190–200. doi: 10.1016/j.ygyno.2018.03.056

19. Holdenrieder S, Molina R, Qiu L, Zhi X, Rutz S, Engel C, et al. Technical and clinical performance of a new assay to detect squamous cell carcinoma antigen levels for the differential diagnosis of cervical, lung, and head and neck cancer. *Tumor Biol* (2018) 40(4):1010428318772202. doi: 10.1177/1010428318772202

20. Nooreldeen R, Bach H. Current and future development in lung cancer diagnosis. *Int J Mol Sci* (2021) 22(16):8661. doi: 10.3390/ijms22168661

21. Ray R, Choi M, Zhang Z, Silverman GA, Askew D, Mukherjee AB. Uteroglobin suppresses SCCA gene expression associated with allergic asthma. *J Biol Chem* (2005) 280(11):9761–4. doi: 10.1074/jbc.C400581200

22. Nakamura H, Akashi K, Watanabe M, Ohta S, Ono J, Azuma Y, et al. Up-regulation of serum periostin and squamous cell carcinoma antigen levels in infants with acute bronchitis due to respiratory syncytial virus. *Allergol Int* (2018) 67(2):259–65. doi: 10.1016/j.alit.2017.10.003

23. Torre GC. SCC antigen in Malignant and nonmalignant squamous lesions. *Tumour Biol* (1998) 19(6):517–26. doi: 10.1159/000030045

24. Calabrese F, Lunardi F, Giacometti C, Marulli G, Gnoato M, Pontisso P, et al. Overexpression of squamous cell carcinoma antigen in idiopathic pulmonary fibrosis: clinicopathological correlations. *Thorax* (2008) 63(9):795–802. doi: 10.1136/thx.2007.088583

25. Wei X, Su J, Yang K, Wei J, Wan H, Cao X, et al. Elevations of serum cancer biomarkers correlate with severity of COVID-19. *J Med Virol* (2020) 92(10):2036–41. doi: 10.1002/jmv.25957

26. Sakito O, Kadota J-i, Kohno S, Itoh N, Takahara O, Hara K. Pulmonary infiltration with eosinophilia and increased serum levels of squamous cell carcinoma-related antigen and neuron specific enolase. *Internal Med* (1994) 33(9):550–3. doi: 10.2169/internalmedicine.33.550

27. Hirayama J, Fujisawa T, Nagao M, Kuwabara Y, Kainuma K, Azuma Y, et al. Squamous cell carcinoma antigens are sensitive biomarkers for atopic dermatitis in children and adolescents: a cross-sectional study. *Asia Pacific Allergy* (2021) 11(4). doi: 10.5415/apallergy.2021.11.e42

28. Kanaji S, Tanaka Y, Sakata Y, Takeshita K, Arima K, Ohta S, et al. Squamous cell carcinoma antigen 1 is an inhibitor of parasite-derived cysteine proteases. *FEBS Lett* (2007) 581(22):4260–4. doi: 10.1016/j.febslet.2007.07.072

29. Nishi N, Miyazaki M, Tsuji K, Hitomi T, Muro E, Zaitsu M, et al. Squamous cell carcinoma-related antigen in children with acute asthma. *Ann Allergy Asthma Immunol* (2005) 94(3):391–7. doi: 10.1016/S1081-1206(10)60993-3

30. Evans SS, Repasky EA, Fisher DT. Fever and the thermal regulation of immunity: the immune system feels the heat. *Nat Rev Immunol* (2015) 15(6):335–49. doi: 10.1038/nri3843

31. Xu M, Li C, Yang J, Ye A, Yan L, Yeoh BS, et al. Activation of CD81+ skin ILC2s by cold-sensing TRPM8+ neuron-derived signals maintains cutaneous thermal homeostasis. *Sci Immunol* (2022) 7(72):eabe0584. doi: 10.1126/sciimmunol.abe0584

32. Kobayashi T, Moro K. Tissue-specific diversity of group 2 innate lymphoid cells in the skin. *Front Immunol* (2022) 13:885642. doi: 10.3389/fimmu.2022.885642

33. Streilein JW. Skin-associated lymphoid tissues (SALT): origins and functions. *J Invest Dermatol* (1983) 80:12s–16s. doi: 10.1111/1523-1747.ep12536743

34. Xiong L, Nutt SL, Seillet C. Innate lymphoid cells: More than just immune cells. *Front Immunol* (2022) 13:1033904. doi: 10.3389/fimmu.2022.1033904

35. Cataltepe S, Gornstein ER, Schick C, Kamachi Y, Chatson K, Fries J, et al. Co-expression of the squamous cell carcinoma antigens 1 and 2 in normal adult human tissues and squamous cell carcinomas. *J Histochem Cytochem* (2000) 48(1):113–22. doi: 10.1177/002215540004800112

36. Kelly-Robinson GA, Reihill JA, Lundy FT, McGarvey LP, Lockhart JC, Litherland GJ, et al. The serpin superfamily and their role in the regulation and dysfunction of serine protease activity in COPD and other chronic lung diseases. *Int J Mol Sci* (2021) 22(12):6351. doi: 10.3390/ijms22126351

37. Gruber F, Kremslehner C, Eckhart L, Tschachler E. Cell aging and cellular senescence in skin aging—Recent advances in fibroblast and keratinocyte biology. *Exp Gerontol* (2020) 130:110780. doi: 10.1016/j.exger.2019.110780

38. Russell-Goldman E, Murphy GF. The pathobiology of skin aging: new insights into an old dilemma. *Am J Pathol* (2020) 190(7):1356–69. doi: 10.1016/j.ajpath.2020.03.007

39. Harvell JD, Maibach HI. Percutaneous absorption and inflammation in aged skin: a review. *J Am Acad Dermatol* (1994) 31(6):1015–21. doi: 10.1016/S0190-9622(94)70273-X

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 | Research Topics