frontiers | Research Topics

Identification of biomarkers for cancer immunotherapy: From bench to bedside, volume I

Edited by
Chang Gu, Qingyuan Huang, Haoran Liu and Amr Ahmed El-Arabey

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-2927-0
DOI 10.3389/978-2-8325-2927-0

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

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

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

Identification of biomarkers for cancer immunotherapy: From bench to bedside, volume I

Topic editors

Chang Gu — Tongji University, China

Qingyuan Huang — Fudan University, China

Haoran Liu — Houston Methodist Research Institute, United States

Amr Ahmed El-Arabey — Al-Azhar University, Egypt

Citation

Gu, C., Huang, Q., Liu, H., El-Arabey, A. A., eds. (2023). *Identification of biomarkers for cancer immunotherapy: From bench to bedside, volume I*.

Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2927-0

Table of contents

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 19 August 2022
DOI 10.3389/fgene.2022.951252

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Shihai Liu,
The Affiliated Hospital of Qingdao University, China
Sen Zhang,
Shanghai Jiao Tong University, China
Bo Feng,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Xiaohui Du,
duxiaohui301@sina.com
Guanglong Dong,
gldong301@163.com

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 23 May 2022
ACCEPTED 08 July 2022
PUBLISHED 19 August 2022

CITATION
Li D, Jiao Y, Gao W, Hu S, Li D, Zhao W, Chen P, Jin L, Zhao Y, Ma Z, Wu X, Yan Y, Sun W, Du X and Dong G (2022), Comprehensive analysis of the prognostic and immunotherapeutic implications of STAT family members in human colorectal cancer.
Front. Genet. 13:951252.
doi: 10.3389/fgene.2022.951252

Comprehensive analysis of the prognostic and immunotherapeutic implications of STAT family members in human colorectal cancer

Dingchang Li[1,2†], Yanan Jiao[1,2†], Wenxing Gao[1,2†], Shidong Hu[2], Dingling Li[3], Wen Zhao[2], Peng Chen[1,2], Lujia Jin[1,2], Yingjie Zhao[1,2], Zhaofu Ma[1,2], Xiansheng Wu[2], Yang Yan[2], Wen Sun[4], Xiaohui Du[2]* and Guanglong Dong[1,2]*

[1]Medical School of Chinese PLA, Beijing, China, [2]Department of General Surgery, The First Medical Centre, Chinese PLA General Hospital, Beijing, China, [3]Medical College of Qinghai University, Xining, China, [4]Department of Anesthesiology, The Second Affiliated Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin, China

Background: Colorectal cancer (CRC) is the third most prevalent cancer worldwide and the second leading cause of cancer mortality. Signal transducer and activator of transcription (STAT) proteins are a group of transcription factors implicated in cell signal transduction and gene transcription in several cancer types. However, the level of expression, genetic alterations, and biological function of different STATs, as well as their prognostic and immunotherapeutic value in CRC remain unclear.

Methods: The mRNA and protein expression levels, genetic alterations, prognostic value, gene–gene and protein–protein interaction networks, and biological function of STATs in CRC were studied using the GEPIA, HPA, cBioPortal, PrognoScan, Kaplan–Meier plotter, GeneMANIA, STRING, and Metascape databases. The expression of STATs in CRC was confirmed using immunohistochemistry (IHC). Finally, the relationship between STAT expression and immune infiltration as well as immunotherapy-associated indicators was also investigated.

Results: The expression levels of *STAT2/5A/5B* are downregulated in CRC, and the *STAT1/3/4/5B* expressions were significantly associated with the tumor stage of patients with CRC. The abnormal expression of *STAT2/4/5B* in patients with CRC is related to the prognosis of patients with CRC. The STATs and their neighboring proteins are primarily associated with lymphocyte activation, cytokine-mediated signaling pathways, positive regulation of immune response, regulation of cytokine production, and growth hormone receptor signaling pathways in cancer. The expression of STATs was significantly associated with immune infiltration and immunotherapy response-associated indicators.

Conclusion: This study may help further understand the molecular mechanism of CRC and provide new prognostic biomarkers and immunotherapy targets in patients with CRC.

KEYWORDS

colorectal cancer, STAT transcription factors, prognostic value, immune infiltration, bioinformatics analysis, tumor immunotherapy

Introduction

Global cancer data in 2020 showed that colorectal cancer (CRC) was the third most prevalent cancer globally and the second leading cause of cancer mortality (Sung et al., 2021). Since the early symptoms of colorectal cancer are not typical (Yoshioka et al., 2014), 35% of patients are often found with metastatic disease when they are diagnosed, and 50% of patients without metastasis ultimately develop metastatic CRC (Zacharakis et al., 2010). Despite advances in chemotherapy, targeted therapy, and immunotherapy, the clinical outcome of CRC remains poor, especially in metastatic CRC (Brenner et al., 2014). Thus, exploring the possible pathogenetic mechanisms of CRC, as well as discovering early diagnostic biomarkers and treatment targets, is crucial for improving patients' prognoses.

STATs are a group of transcription factors encoded by seven members (*STAT1/2/3/4/5A/5B/6*) of the STAT gene family that are involved in cell proliferation, differentiation, apoptosis, and immune system regulation (Verhoeven et al., 2020). As a result, dysregulation of their pathway would result in a variety of diseases, including cancer (Bowman et al., 2000). Extensive studies have already demonstrated that inappropriate activation of specific STAT members contributes to oncogenesis, especially for the Janus kinase (JAK)/*STAT3* pathway, which has been linked to many types of cancer (Johnson et al., 2018). For example, Li et al. reported that long non-coding RNA RP11-468E2.5 could curtail CRC development and promote apoptosis via the JAK/STAT signaling pathway by targeting *STAT5* and *STAT6* (Jiang et al., 2019).

Despite great importance of STATs in malignancies, there has been no study to explore the implications of every STAT factor in CRC, including their expression level, genetic variation, biological function, and potential molecular mechanism. Furthermore, their correlation with the prognosis, immune infiltration, and immunotherapy response in patients with CRC also remains unknown. Thus, it is necessary to comprehensively analyze the significance of each STAT member in CRC development and progression.

Multiple large-scale bioinformatics databases were used in this study for comprehensive bioinformatics analysis of the expression of STATs and their associations with tumor stage in patients with CRC. In addition, immunohistochemistry (IHC) was used to confirm the differential expressions of STATs in CRC and normal tissues. Subsequently, the genetic variation, biological function, and molecular mechanism of each STAT member in CRC were explored. Ultimately, the relationship between the expression of STATs and prognosis, immune infiltration, and immunotherapy response in patients with CRC was analyzed.

Materials and methods

Data acquisition and analysis of differential expression

The Genotype-Tissue Expression (GTEx) database (https://commonfund.nih.gov/GTEx/) collects data from 54 normal human tissues for sequencing, which can be used to compare the differential level of gene expression between normal and diseased tissues (GTEx Consortium, 2013). The Cancer Genome Atlas (TCGA) (https://tcga.xenahubs.net) mainly contains data from 33 different types of tumors. The RNA sequencing data of normal samples from the GTEx database and tumor samples from TCGA were downloaded, and the Wilcoxon rank-sum test method was used to compare the differential mRNA expressions of STATs between 33 different types of cancers and corresponding normal tissues. Threshold values were determined according to the following values: ns, $p \geq 0.05$, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. The "ggplot2" R package was used for the boxplot.

Gene Expression Profiling Interactive Analysis 2 (GEPIA2) dataset

GEPIA2 (http://gepia2.cancer-pku.cn) is the latest version of GEPIA, which analyzes RNA sequencing expression data including 9,736 tumors and 8,587 normal samples from TCGA and GTEx projects using standard processing pipelines (Tang et al., 2019). GEPIA2 offers a variety of functions such as differential gene expression analysis, cancer types and pathological staging, similar gene detection, patient survival analysis, correlation analysis, and dimensionality reduction analysis.

Human Protein Atlas (HPA) dataset and immunohistochemistry (IHC)

HPA database (https://www.proteinatlas.org/) was used to compare the STAT gene protein expression in the CRC tissues and the corresponding normal tissues.

IHC staining was used to further validate the reliability of the above results. Clinical samples were collected from 21 patients with CRC who were undergoing surgical treatment in our hospital, and their clinical information is shown in Supplementary Table S1. Firstly, these samples were made into 3 μm paraffin sections and incubated with rabbit monoclonal antibodies of *STAT1*, *STAT2*, *STAT3*, *STAT4*, *STAT5A*, *STAT5B*, and *STAT6* (1:100, all from Abcam, USA) at 4°C overnight. The sections were then conjugated with horseradish peroxidase (HRP) secondary antibody (Abcam, USA) at 1/500 dilution at room temperature for 2 h. Subsequently, the conjugates were stained with 3,3′-diaminobenzidine (DAB) reagent, and ultimately counterstained with hematoxylin. The IHC score of STATs was assessed manually and calculated by a pathologist. The percentage of positive cells was scored as: 0 (0–10% positive); 1 (10–25% positive); 2 (26–50% positive); 3 (51–75% positive); and 4 (≥76% positive). The staining intensity was scored as: 0 (no staining); 1 (weak); 2 (moderate); and 3 (strong). The overall IHC score was calculated by multiplying the score of positive cells (0–4) by the staining intensity (0–3).

Tumor–Immune System Interactions and Drug Bank (TISIDB) database

TISIDB (http://cis.hku.hk/TISIDB) is a website for the tumor and immune system interaction that integrates multiple types of data in oncoimmunology and reports genes related to antitumor immunity, tumor cell resistance or sensitivity to T cell-mediated killing and immunotherapy, and relationships between genes and immune features of 30 cancer types from TCGA (Ru et al., 2019).

PrognoScan database and the Kaplan–Meier plotter analysis

The prognostic value of STATs mRNA expression in patients with CRC was assessed by the PrognoScan Database (http://www.abren.net/PrognoScan/) (Mizuno et al., 2009). This could be used for evaluating the correlation between gene expression and patient survival including overall survival (OS) and disease-free survival (DFS). Cox $p < 0.05$ was considered statistically significant.

The Kaplan–Meier plotter (www.kmplot.com), an online database containing gene expression data and clinical survival information of cancer patients (Nagy et al., 2018), was further used to validate the relationship between STAT expression in rectal adenocarcinoma (READ) and OS. The hazard ratio (HR) with 95% confidence intervals and log-rank *p*-value were also calculated.

cBioPortal

cBioPortal (http://www.cbioportal.org) is an online database that can conduct multidimensional cancer genomics studies (Gao et al., 2013). A colorectal adenocarcinoma dataset (TCGA, PanCancer Atlas) containing 524 patients was selected to analyze the expression of STATs. The genomic profiles included mutations, putative copy-number alterations from Genomic Identification of Significant Targets in Cancer (GISTIC) scores, and mRNA expression z-scores (RNA Seq V2 RSEM). The z-score threshold was set at ±1.8.

Network analysis

GeneMANIA (www.genemania.org), an online analysis tool that provides protein and genetic co-expression, co-localization, interactions, pathways, and shared protein domains of submitted genes (Franz et al., 2018), was used to perform a gene–gene interaction network for STATs. STRING (https://string-db.org), an online dataset that collects and integrates all publicly available protein–protein interaction (PPI) data and predicts potential functions (Szklarczyk et al., 2019), was used to construct a PPI network for STATs.

Functional enrichment analysis

Firstly, GEPIA was used to identify the top 30 similar genes in CRC for each STAT family member. Metascape was subsequently used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathway analysis of the STATs and similar genes (Zhou et al., 2019). Only terms with $p < 0.01$, minimum count >3, and enrichment factor >1.5 were considered significant.

Tumor Immune Estimation Resource (TIMER) dataset

TIMER (http://timer.cistrome.org/), an online dataset that provides tumor immune infiltrating abundances estimated by multiple immune deconvolution methods (Li et al., 2020), was used in this study to evaluate the correlation between STAT expression levels and immune cell infiltration.

Statistical analysis

The difference between STAT IHC scores in normal and tumor tissues was tested using a two-tailed Student's t-test with unpaired analysis. The correlation between STAT gene expressions and immune infiltration level, tumor purity,

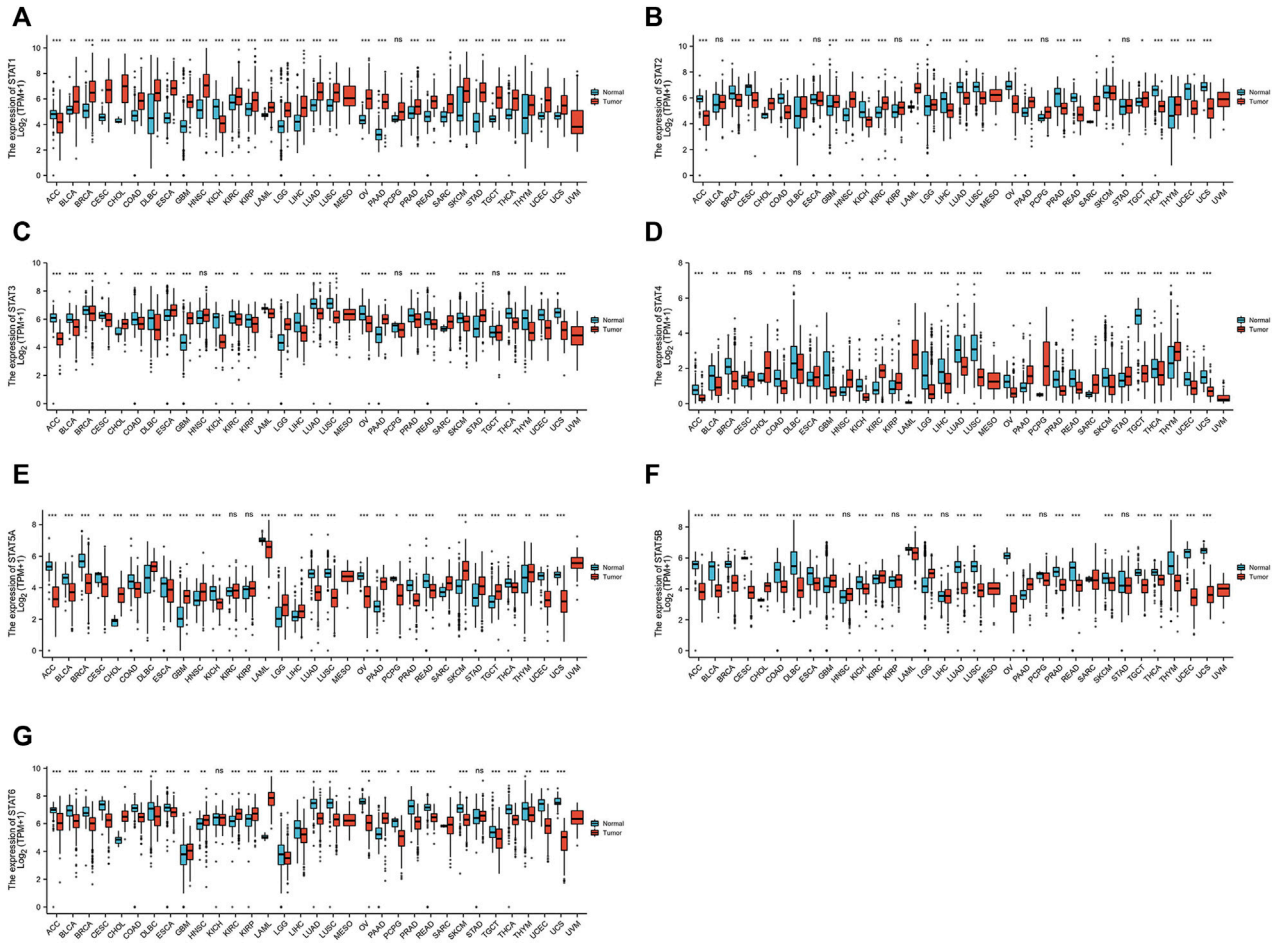


FIGURE 1
Transcription levels of signal transducer and activator of transcription (STAT) factors in different types of cancers from The Cancer Genome Atlas (TCGA) database and the Genotype-Tissue Expression (GTEx) database.

immune checkpoints, tumor mutation burden (TMB), microsatellite instability (MSI), and mismatch repair (MMR) genes in CRC was assessed using Spearman's correlation coefficients, and a $p < 0.05$ was considered statistically significant.

Results

The expression levels of STATs in pan-cancer

STAT expression levels in 33 types of tumors were evaluated using data from TCGA database (n = 9,379) and GTEx database (n = 8,293). In most types of tumors, all STAT family members had significantly abnormal levels of expression compared to normal tissues (Figure 1). Tumors of the digestive system were concentrated on since this study is mainly about CRC. The *STAT1* gene was highly expressed in most tumors, including cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), liver hepatocellular carcinoma (LIHC), pancreatic adenocarcinoma (PAAD), READ, and stomach adenocarcinoma (STAD) (Figure 1A). The *STAT2* gene expression levels were high in CHOL and PAAD but low in COAD and READ (Figure 1B). The *STAT3* gene was highly expressed in CHOL, ESCA, PAAD, and STAD but not in COAD, LIHC, or READ (Figure 1C). The expression of *STAT4* was high in CHOL, ESCA, PAAD, and STAD but low in COAD, LIHC, and READ (Figure 1D). The *STAT5A* gene expression was high in CHOL, LIHC, PAAD, and STAD but low in COAD, ESCA, and READ (Figure 1E). The expression of *STAT5B* was high in CHOL and PAAD but low in COAD, ESCA, and READ (Figure 1F). *STAT6* presented high expression in CHOL and PAAD but low expression in COAD, ESCA, LIHC, and READ (Figure 1G).

10.3389/fgene.2022.951252

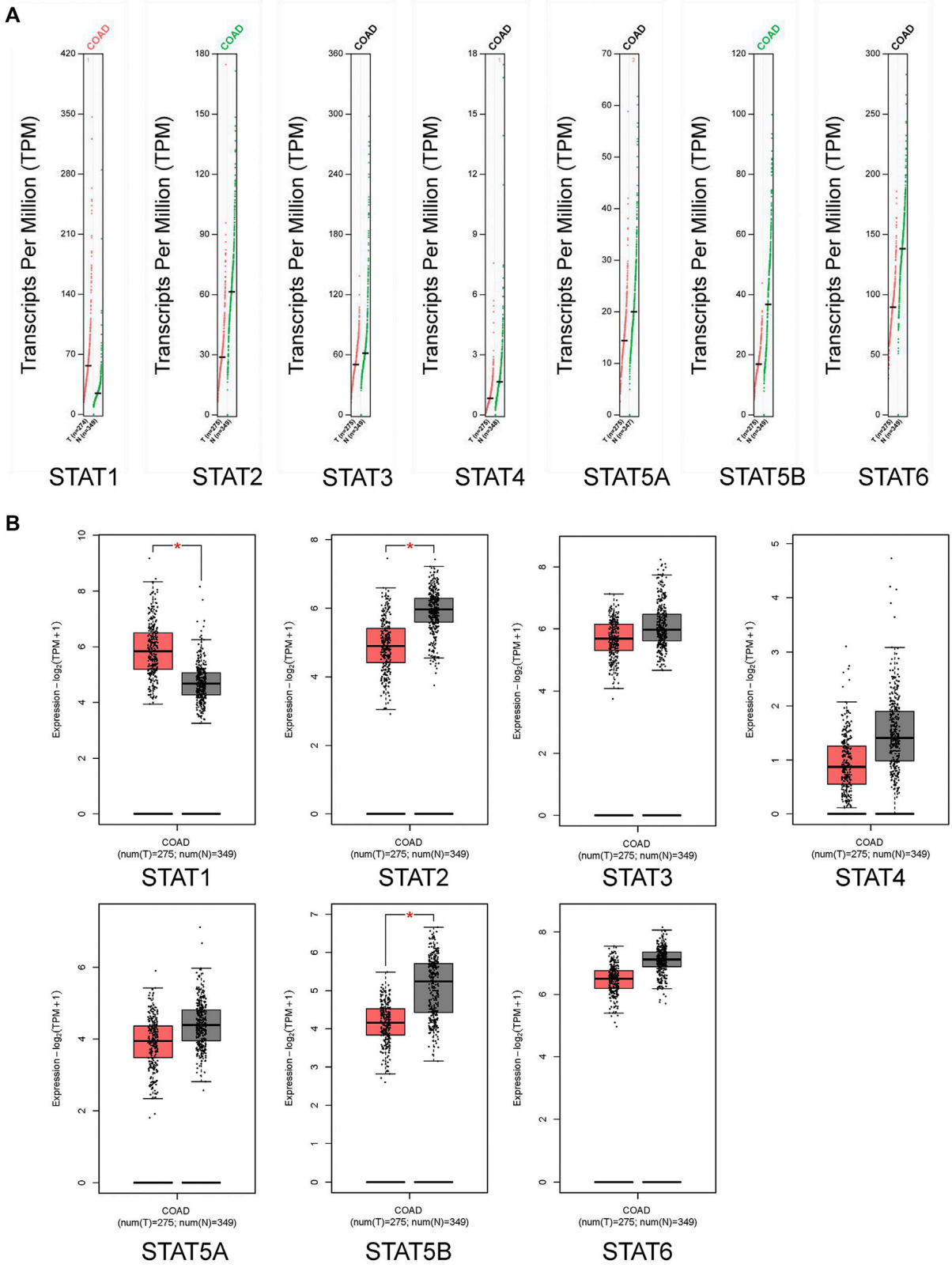


FIGURE 2
Transcription levels of signal transducer and activator of transcription (STAT) factors in colorectal cancer (CRC) from the Gene Expression Profiling Interactive Analysis 2 (GEPIA2) dataset.

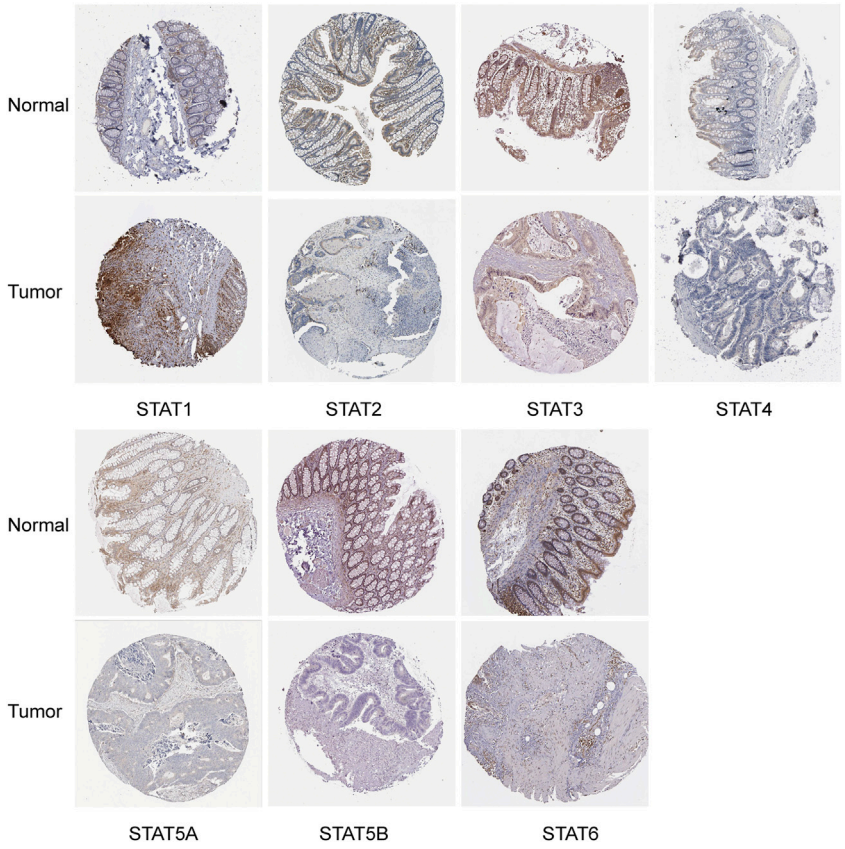


FIGURE 3
Translation levels of signal transducer and activator of transcription (STAT) factors in colorectal cancer (CRC) from the Human Protein Atlas (HPA) dataset.

Transcriptional and translational expression levels of STATs in CRC patients

The GEPIA dataset was used to compare the transcriptional levels of STATs between CRC and normal tissues (Figure 2A and Figure 2B). The results showed that the expression level of *STAT1* in CRC tissues was higher than in normal colon tissues, and the transcriptional levels of *STAT2* and *STAT5B* in CRC were lower than in normal tissues significantly. The expressions of *STAT3/4/5A/6* genes were lower in CRC than in normal samples, although there was no statistical significance.

The HPA database and IHC staining were used to further confirm the protein expression of STATs in CRC and normal tissues. The HPA database results indicated that the *STAT1* protein was more highly expressed in the CRC tissues than in the normal tissues, while *STAT2/5A/5B/6* were significantly less expressed in CRC tissues than in normal tissues (Figure 3). The IHC results and scores from clinical samples showed the protein levels of *STAT1* were higher, and levels of *STAT2/5A/5B* were lower in CRC tissues than in the adjacent normal tissues with great significance (Figure 4 and Supplementary Figure S1).

The prognostic value of STATs in CRC patients

The relationship between transcriptional levels of STATs and CRC stage was investigated using the TISIDB. The results showed that the expression levels of *STAT1/3/4/5B* were significantly associated with the tumor stage of patients with CRC. However, there was no significant correlation between the *STAT2/5A/6* expression and tumor stage (Figure 5).

The correlation between STAT expression and clinical outcome was evaluated using the PrognoScan database and the Kaplan–Meier plotter analysis to assess the value of STATs expression levels in the prognosis of CRC (Figure 6). The PrognoScan database analysis results showed that higher *STAT2/4/5B* mRNA levels were significantly associated with better OS ($p < 0.05$) and increased *STAT2/3/4/5B* transcription levels were significantly associated with longer DFS ($p < 0.05$) (Figure 6A). In addition, the correlation between STAT expression levels and OS in patients with READ was further validated using the Kaplan–Meier plotter, which indicated that high expression of *STAT1/4/5B* favored OS ($p < 0.05$) (Figure 6B).

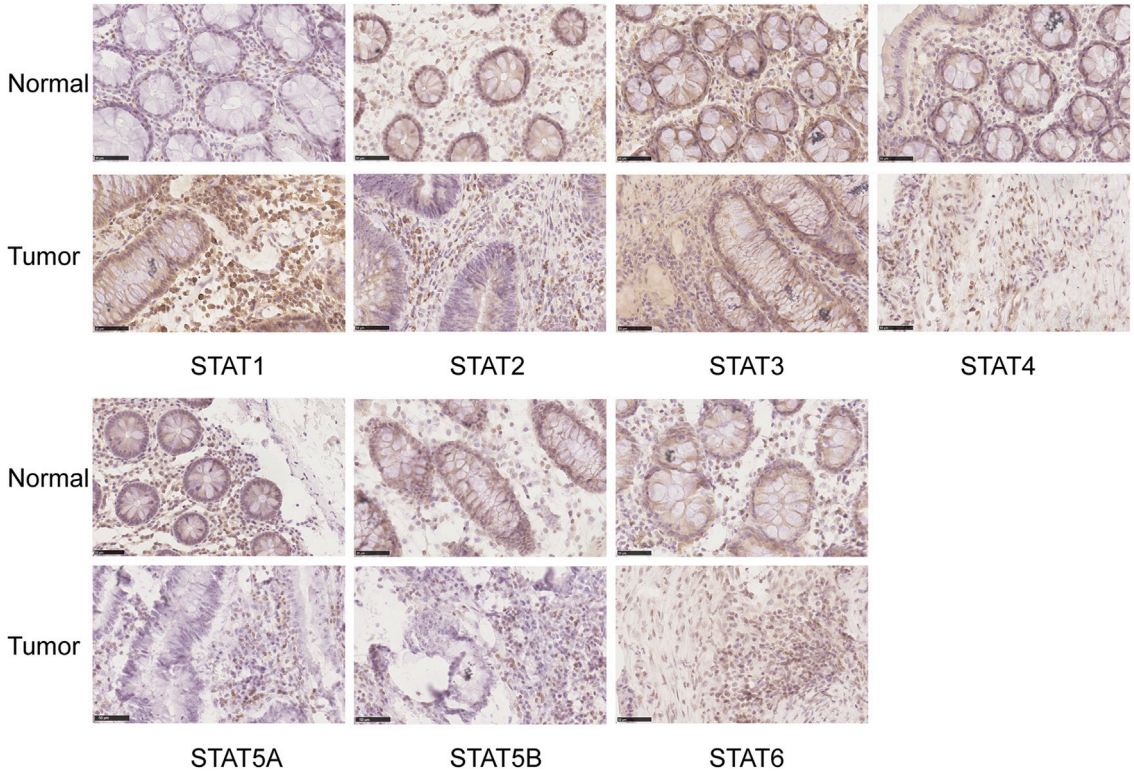


FIGURE 4
Translation levels of signal transducer and activator of transcription (STAT) factors in colorectal cancer (CRC) with immunohistochemistry (IHC). Scale bar = 50 μm.

Gene mutations, co-expression, and interaction analyses of STATs in CRC patients

The cBioPortal online tool was used to evaluate genetic alterations and STAT factor correlations in patients with CRC. STATs were found to be altered in 224 (43%) of 524 patients (Figure 7A). The genes with the highest and lowest mutation rates in STATs are *STAT5B* (15%) and *STAT4* (7%), respectively. The others are *STAT1* (9%), *STAT2* (9%), *STAT3* (12%), *STAT5A* (9%), and *STAT6* (12%) (Figure 7A). It was also found that patients with colorectal mucinous adenocarcinoma were most likely to have STAT gene alterations (53.57% of 56 cases) (Figure 7B). The analysis results from cBioPortal showed that patients in the unaltered group seemed /to have a better prognosis than those in the altered group but without statistical significance (Supplementary Figure S2A). The potential effects of every single STAT factor on prognosis were then evaluated and the results showed that patients with altered *STAT4* had significantly poorer prognostic outcomes compared with unaltered patients (Supplementary Figure S2E). TIMER dataset analysis was used to evaluate the effect of STAT mutations on five types of immune cell infiltration, and the outcomes indicated that mutated *STAT1* correlates with a higher level of neutrophil infiltration; mutated *STAT4* correlates with more B cells, $CD8^+$ T cells, and neutrophil infiltration; mutated *STAT5A* correlates with a higher level of $CD4^+$ T cells and neutrophils; mutated *STAT5B* correlates with more B cells and; mutated *STAT6* correlates with lower B cell infiltration with significance (Supplementary Figure S3).

We further explored the co-expression of STAT members in patients with CRC, and there were strong or moderate positive relationships between *STAT1* and *STAT2*, *STAT3*, *STAT4*, and *STAT5A*; *STAT2* and *STAT3*, *STAT4*, *STAT5A*, *STAT5B*, and *STAT6*; *STAT3* and *STAT4*, *STAT5A*, and *STAT5B* and; *STAT5A* and *STAT5B* and *STAT6* ($p < 0.05$) (Figure 7C).

The gene–gene interaction (GGI) network of STATs was established using the GeneMANIA database (Figure 7D). Based on shared protein domains, co-localization, physical interactions, co-expression, pathways, and genetic interactions, 20 related genes were enriched in this network. These genes are involved in a variety of functions such as receptor tyrosine kinase binding, receptor signaling pathway via STAT, signaling receptor complex adaptor activity, signaling adaptor activity, phosphoprotein binding, protein phosphorylated amino acid binding, and growth hormone receptor signaling pathway. STRING was used to explore the potential interactions between STATs at the protein

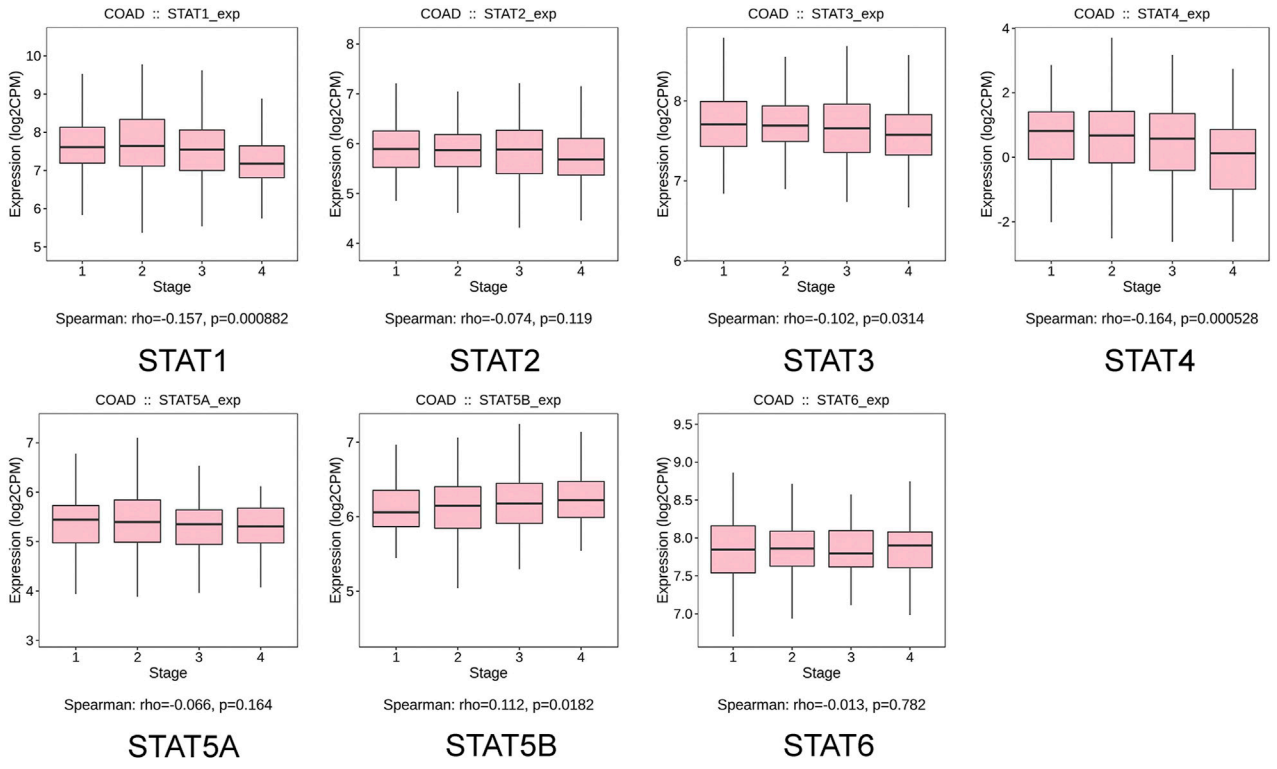


FIGURE 5
Correlation between signal transducer and activator of transcription (STAT) factor expression and tumor stage in patients with colorectal cancer (CRC) from the Tumor–Immune System Interactions and Drug Bank (TISIDB) database.

level, as shown in Figure 7E, where the PPI network diagram had seven nodes and 21 edges. The analysis results of PPI network indicated that each STAT factor has known or predicted interactions with the others, especially *STAT3*, which has experimentally determined interactions with each of the other factors. The top four molecular pairs with strong functional links based on combined scores were *STAT1* and *STAT2*, *STAT1* and *STAT3*, *STAT3* and *STAT5B*, and *STAT1* and *STAT5A*.

Functional enrichment analysis of STATs and the genes similar to them in CRC patients

GEPIA2 datasets were used to identify the top 30 genes that have a similar expression pattern to each STAT family member. The GO and KEGG enrichment pathway analyses of STATs and their similar genes were then performed using Metascape. The top 20 GO enrichment items were composed of 16 biological processes (BP) items, three molecular functions (MF) items, and one cellular component (CC) item (Figure 8A, Figure 8B, and Table 1). The first five projects are all in the BPs, and they are lymphocyte activation, cytokine-mediated signaling pathway, positive regulation of immune response, regulation of cytokine production, and growth hormone receptor signaling pathway via JAK/STAT. MFs that were significantly related to STATs and similar genes were kinase binding, GTPase regulator activity, and CCR5 binding. The only one CC was side of the membrane.

The first 16 KEGG pathways are displayed in Figure 8C, Figure 8D, and Table 2. The results indicated the involvement of STATs in pathways such as Th17 cell differentiation, chemokine signaling pathway, T cell receptor signaling pathway, and cytokine–cytokine receptor interaction.

The relationship between STAT expression levels and immune infiltration levels in CRC

According to the results of GO and KEGG enrichment analysis, it was found that STATs were closely related to immune functions such as lymphocyte activation, positive regulation of immune response, Th17 cell differentiation, and T cell receptor signaling pathway. The results indicated that STATs were involved in the regulation of the tumor immune microenvironment, which is closely related to the initiation and progression of tumors.

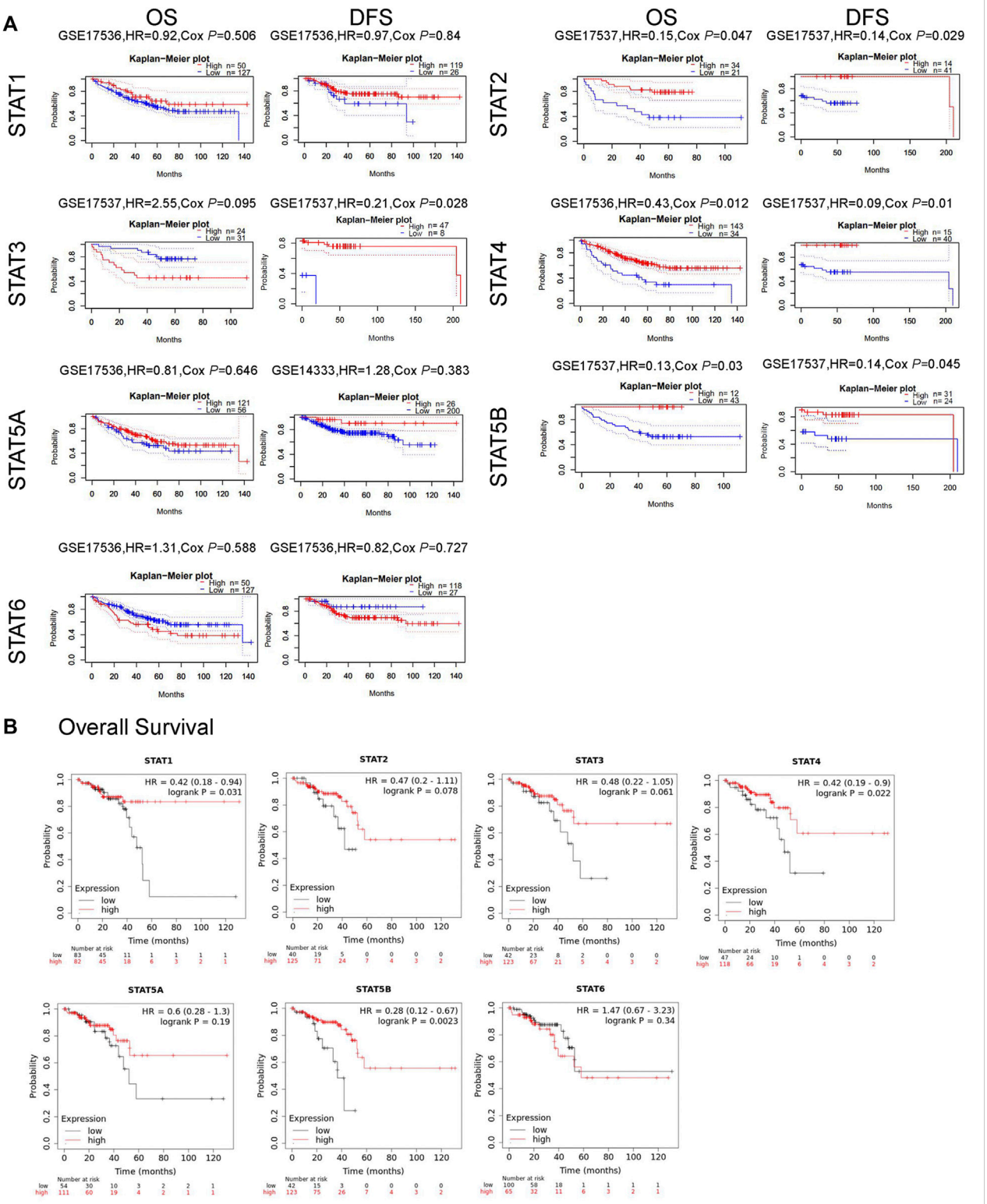


FIGURE 6
Relationship between signal transducer and activator of transcription (STAT) factor expression and prognosis in patients with colorectal cancer (CRC). **(A)** Prognostic value of STATs in patients with CRC in the OS and DFS curves (PrognoScan). **(B)** Prognostic value of STATs in rectal adenocarcinoma (READ) patients in the OS curve (Kaplan–Meier plotter).

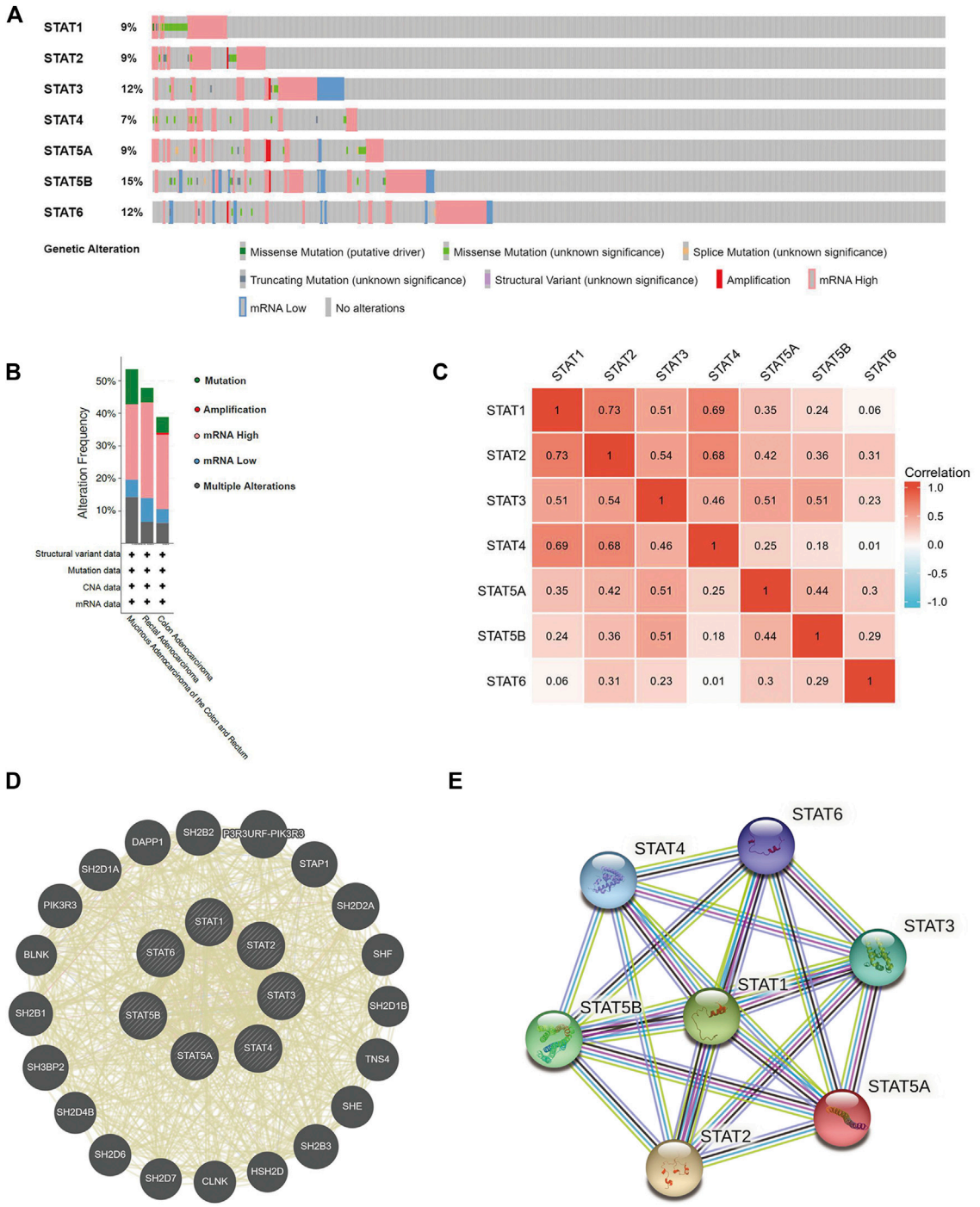


FIGURE 7
Genetic mutations, co-expression, and interaction analysis of signal transducer and activator of transcription (STAT) factors at gene and protein levels in patients with colorectal cancer (CRC) (cBioPortal, GeneMANIA, and STRING). **(A)** Analysis of gene mutations of STAT family members in CRC. **(B)** Summary of alterations in expressed STATs in CRC. **(C)** Correlation heatmap of expressed STATs in CRC. The numbers in the color blocks represent Spearman's correlation coefficient. **(D)** Gene–gene interaction network among STATs predicted by GeneMANIA. **(E)** Protein–protein interaction network among STATs predicted by STRING.

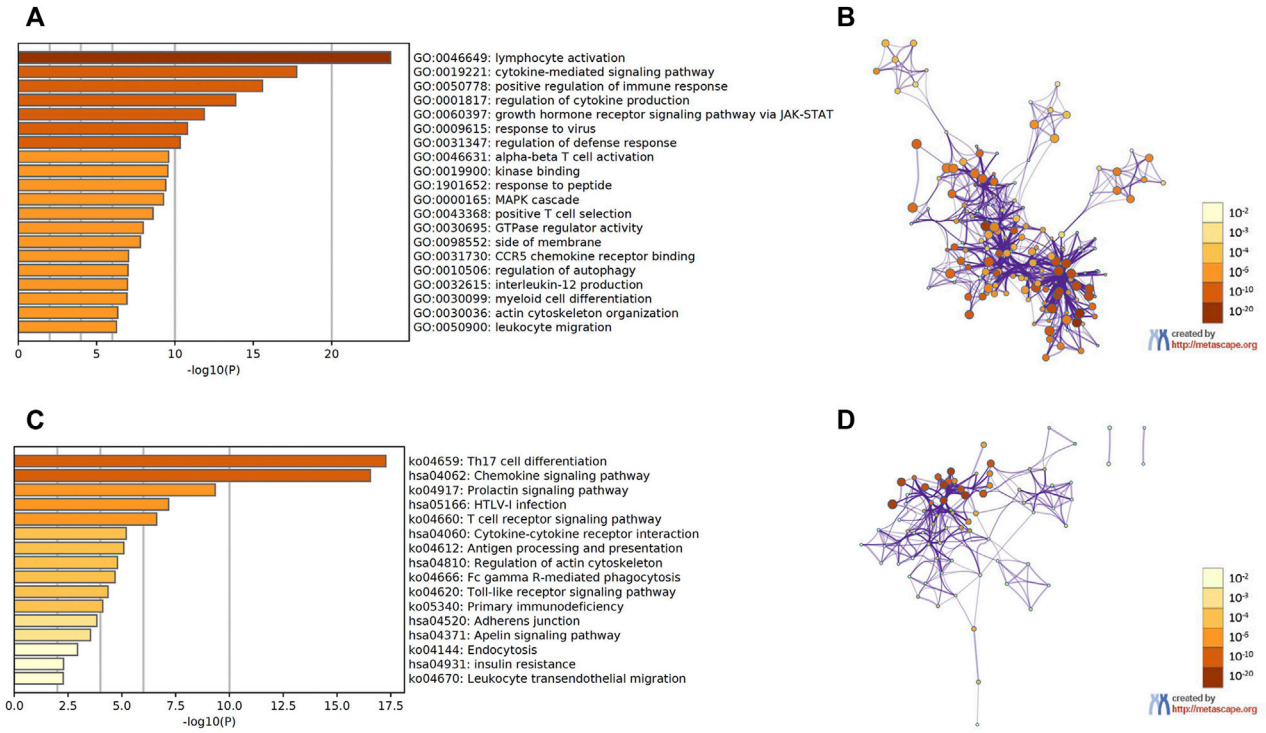


FIGURE 8
Functional enrichment analysis of signal transducer and activator of transcription (STAT) factors and the genes similar to them in patients with colorectal cancer (CRC) (Metascape). **(A)** Heatmap of GO enriched terms colored by *p*-values. **(B)** Network of GO enriched terms colored by *p*-value, where terms containing more genes tend to have a more significant *p*-value. **(C)** Heatmap of KEGG enriched terms colored by *p*-values. **(D)** Network of KEGG enriched terms colored by *p*-value, where terms containing more genes tend to have a more significant *p*-value.

As a result, we used TIMER online dataset to evaluate the relationship between STAT expression and immune cell infiltration in CRC. As shown in Figure 9, the outcomes, as expected, revealed that STATs were involved in many types of immune cell infiltration and influenced the clinical outcome of patients with CRC. *STAT1/2/3/4* expressions had a positive correlation with the infiltration of B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophage, neutrophil, and dendritic cells (Figures 9A–D). The expressions of *STAT5A/5B/6* were positively correlated with the infiltration of $CD4^+$ T cells, macrophage, neutrophil, and dendritic cells, and *STAT5A* expression was also positively related to the infiltration of B cells (Figures 9E–G).

The correlation between STATs and immunotherapy response-related indicators in CRC

The tumor microenvironment (TME) includes tumor cells, stromal cells, immune cells, and extracellular matrix. Stromal cells have been reported to promote tumorigenesis in many ways, and infiltration levels of different immune cells are also linked to tumor prognosis. As a result, the correlation between STAT gene expressions and stromal cells and immune cell content in CRC was subsequently evaluated. The results showed that the expressions of *STAT1*, *STAT2*, *STAT3*, *STAT4*, and *STAT5A* were positively related to stromal score, immune score, and ESTIMATE score with great significance ($p < 0.001$) (Figures 10A–E). Also, the expression of *STAT5B* was positively related to the stromal score and the ESTIMATE score, but there was no relationship between *STAT6* expression and these scores as shown in Figure 10F and Figure 10G. The correlation analysis showed that STAT gene expressions were strongly positively related to eight immune checkpoints, including CD40LG, ADORA2A, TNFSF14, ICOSLG, TNFRSF8, CD27, VSIR, and TNFRSF4 ($r > 0$, $p < 0.001$) (Figure 10H and Table 3). In addition, the expressions of STAT genes were also correlated with most chemokines and their receptors (Figure 11).

TMB is a predictive biomarker of response for cancer patients receiving immune checkpoint blockade. The study results revealed a positive relationship between TMB score and *STAT1*, *STAT2*, and *STAT4* expression ($p < 0.001$), and a negative relationship between TMB score and *STAT5B*

TABLE 1 Gene Ontology (GO) functional enrichment analysis of signal transducer and transcription (STAT) factors and neighbor genes in colorectal cancer (CRC) (Metascape).

GO	Category	Description	Count	%	Log10 (P)	Log10 (q)
GO:0046649	GO biological processes	Lymphocyte activation	39	21.08	-23.76	-19.41
GO:0019221	GO biological processes	Cytokine-mediated signaling pathway	27	14.59	-17.78	-13.90
GO:0050778	GO biological processes	Positive regulation of immune response	27	14.59	-15.59	-11.84
GO:0001817	GO biological processes	Regulation of cytokine production	29	15.68	-13.88	-10.49
GO:0060397	GO biological processes	Growth hormone receptor signaling pathway via JAK/STAT	6	3.24	-11.87	-8.86
GO:0009615	GO biological processes	Response to virus	18	9.73	-10.80	-8.00
GO:0031347	GO biological processes	Regulation of defense response	22	11.89	-10.32	-7.56
GO:0046631	GO biological processes	Alpha-beta T-cell activation	12	6.49	-9.60	-6.96
GO:0019900	GO molecular functions	Kinase binding	23	12.43	-9.55	-6.93
GO:1901652	GO biological processes	Response to peptide	19	10.27	-9.41	-6.79
GO:0000165	GO biological processes	MAPK cascade	23	12.43	-9.27	-6.67
GO:0043368	GO biological processes	Positive T-cell selection	7	3.78	-8.60	-6.06
GO:0030695	GO molecular functions	GTPase regulator activity	17	9.19	-7.98	-5.50
GO:0098552	GO cellular components	Side of membrane	19	10.27	-7.79	-5.34
GO:0031730	GO molecular functions	CCR5 binding	4	2.16	-7.03	-4.64
GO:0010506	GO biological processes	Regulation of autophagy	13	7.03	-7.00	-4.62
GO:0032615	GO biological processes	Interleukin-12 production	7	3.78	-6.98	-4.61
GO:0030099	GO biological processes	Myeloid cell differentiation	14	7.57	-6.94	-4.57
GO:0030036	GO biological processes	Actin cytoskeleton organization	18	9.73	-6.35	-4.06
GO:0050900	GO biological processes	Leukocyte migration	13	7.03	-6.27	-3.99

TABLE 2 Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis of signal transducer and transcription (STAT) factors and neighbor genes in colorectal cancer (CRC) (Metascape).

GO	Category	Description	Count	%	Log10 (P)	Log10 (q)
ko04659	KEGG pathway	Th17 cell differentiation	16	8.65	-17.27	-14.39
hsa04062	KEGG pathway	Chemokine signaling pathway	19	10.27	-16.55	-14.25
ko04917	KEGG pathway	Prolactin signaling pathway	9	4.86	-9.33	-7.76
hsa05166	KEGG pathway	HTLV-I infection	12	6.49	-7.17	-5.78
ko04660	KEGG pathway	T-cell receptor signaling pathway	8	4.32	-6.61	-5.26
hsa04060	KEGG pathway	Cytokine–cytokine receptor interaction	11	5.95	-5.21	-3.96
ko04612	KEGG pathway	Antigen processing and presentation	6	3.24	-5.10	-3.87
hsa04810	KEGG pathway	Regulation of actin cytoskeleton	9	4.86	-4.79	-3.62
ko04666	KEGG pathway	Fc gamma R-mediated phagocytosis	6	3.24	-4.68	-3.53
ko04620	KEGG pathway	Toll-like receptor signaling pathway	6	3.24	-4.35	-3.23
ko05340	KEGG pathway	Primary immunodeficiency	4	2.16	-4.12	-3.04
hsa04520	KEGG pathway	Adherens junction	5	2.70	-3.85	-2.79
hsa04371	KEGG pathway	Apelin signaling pathway	6	3.24	-3.54	-2.54
ko04144	KEGG pathway	Endocytosis	7	3.78	-2.94	-2.01
hsa04931	KEGG pathway	Insulin resistance	4	2.16	-2.28	-1.49
ko04670	KEGG pathway	Leukocyte transendothelial migration	4	2.16	-2.27	-1.48

expression ($p < 0.001$) (Figure 12A). The MSI status is also closely linked to the response to immune checkpoint blockade, especially in patients with CRC. As a result, the correlation between STAT expression and MSI score was also assessed, and the results indicated that *STAT1*, *STAT2*, *STAT3*, and *STAT4* were positively correlated with MSI score ($p < 0.001$),

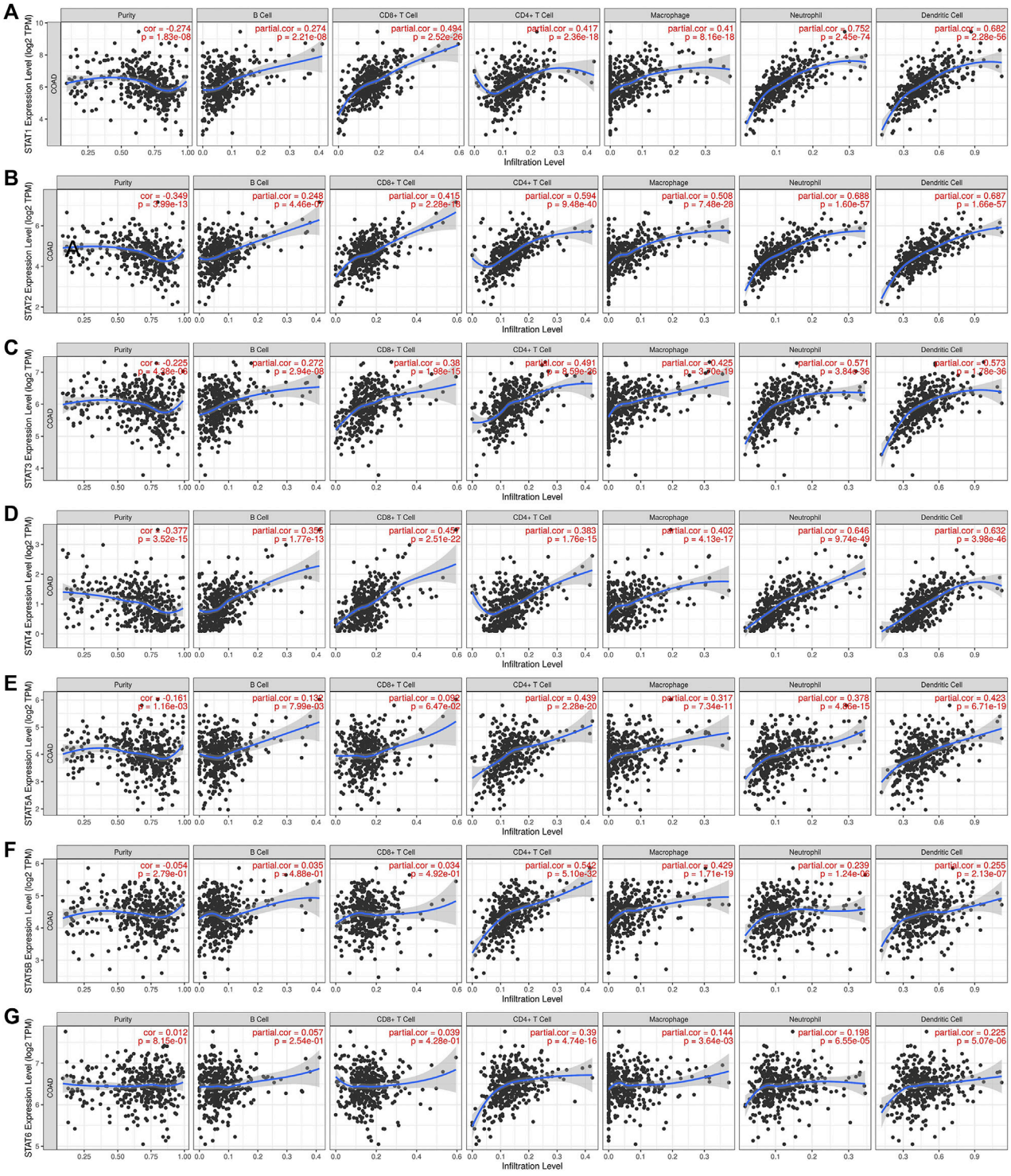


FIGURE 9
Correlation between signal transducer and activator of transcription (STAT) factor expression levels and immune infiltration levels in colorectal cancer (CRC) from the Tumor Immune Estimation Resource (TIMER) dataset. The relationship between the abundance of immune cells and the expression of **(A)** *STAT1*, **(B)** *STAT2*, **(C)** *STAT3*, **(D)** *STAT4*, **(E)** *STAT5A*, **(F)** *STAT5B*, and **(G)** *STAT6* in CRC.

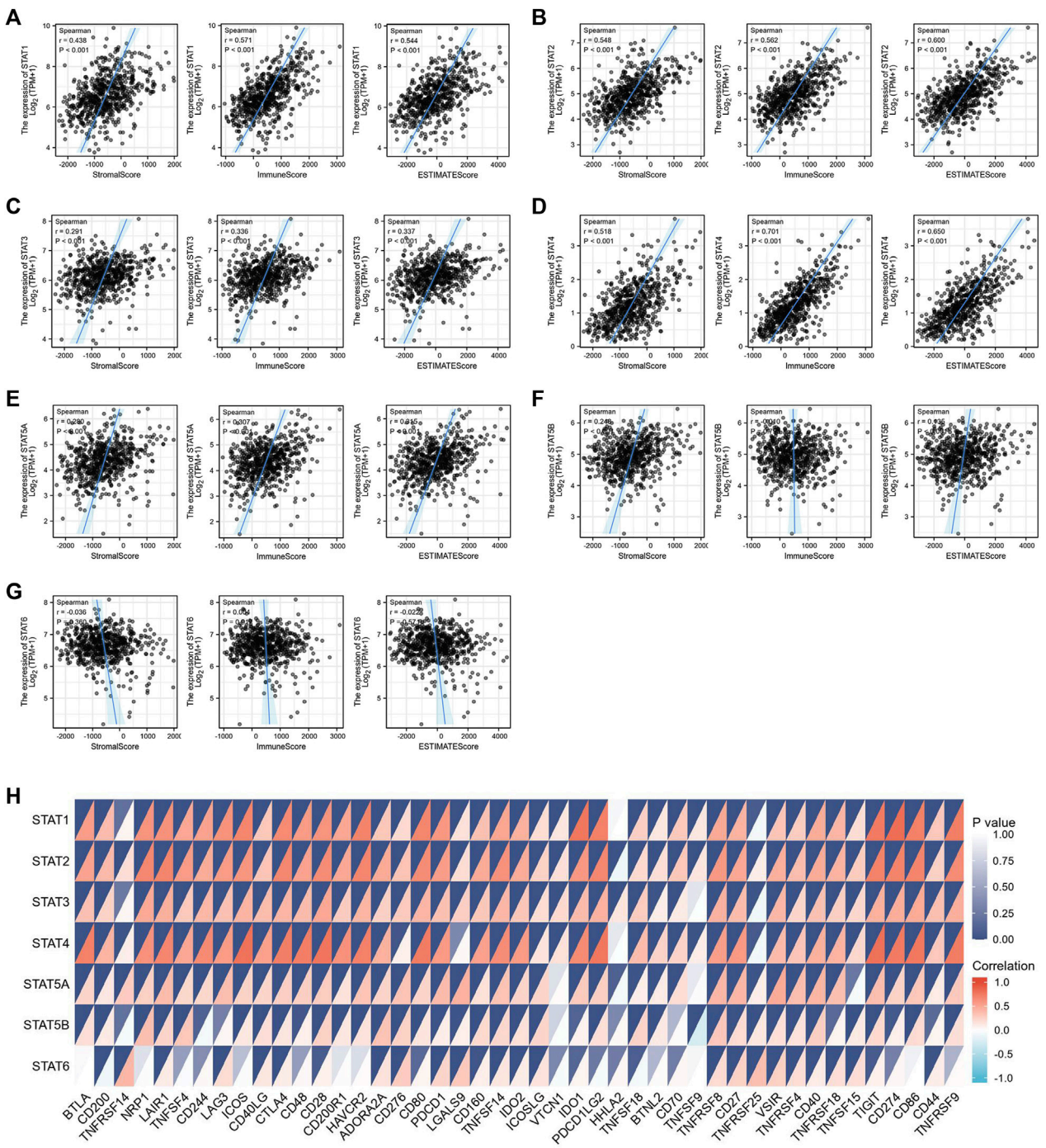


FIGURE 10
Relationship between signal transducer and activator of transcription (STAT) factor expression and stromal score, immune score, ESTIMATE score, and immune checkpoints in colorectal cancer (CRC). The correlation between the expression of **(A)** *STAT1*, **(B)** *STAT2*, **(C)** *STAT3*, **(D)** *STAT4*, **(E)** *STAT5A*, **(F)** *STAT5B*, and **(G)** *STAT6*, and stromal score, immune score, and ESTIMATE score in CRC. **(H)** Correlation heatmap of STATs with immune checkpoints in CRC.

whereas *STAT5B* and *STAT6* were negatively correlated with MSI score ($p < 0.01$) (Figure 12B). The relationship between STAT expression and MMR genes such as MLH1, MSH2, MSH6, PMS2, and EPCAM was evaluated using TCGA expression profile data. The results showed that the expressions of *STAT1*, *STAT2*, *STAT4*, and *STAT5A* were

TABLE 3 Relationship between signal transducer and transcription (STAT) factor expression and immune checkpoints in colorectal cancer (CRC).

R	Immune checkpoint							
	CD40LG	ADORA2A	TNFSF14	ICOSLG	TNFRSF8	CD27	VSIR	TNFRSF4
STAT1	0.311	0.315	0.501	0.189	0.488	0.429	0.325	0.282
STAT2	0.326	0.483	0.618	0.257	0.576	0.497	0.460	0.372
STAT3	0.373	0.261	0.409	0.194	0.433	0.378	0.409	0.234
STAT4	0.536	0.444	0.647	0.152	0.544	0.572	0.383	0.329
STAT5A	0.234	0.241	0.336	0.237	0.409	0.405	0.478	0.366
STAT5B	0.178	0.335	0.224	0.233	0.307	0.199	0.173	0.158
STAT6	0.133	0.163	0.139	0.158	0.199	0.254	0.263	0.187

The *p*-value of immune checkpoints (CD40LG, ADORA2A, TNFSF14, ICOSLG, TNFRSF8, CD27, VSIR, and TNFRSF4) is less than 0.001.

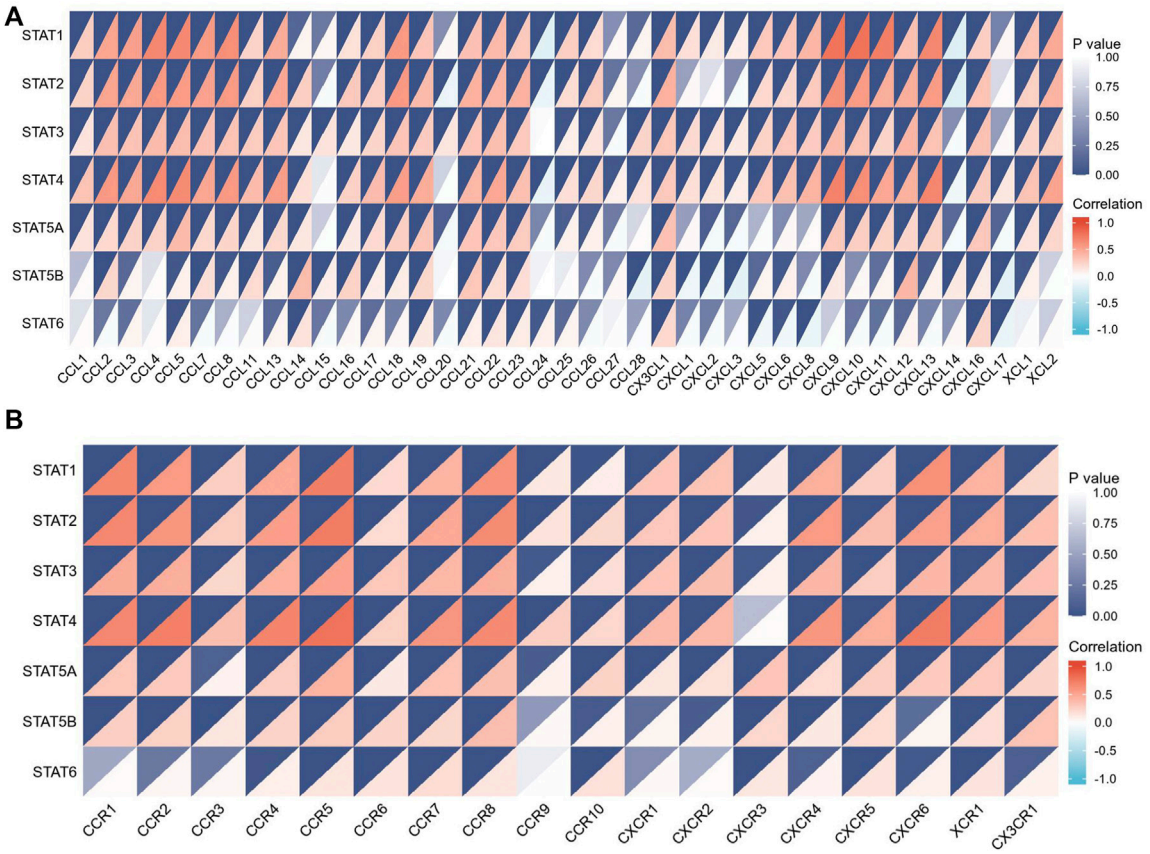


FIGURE 11
Relationship between signal transducer and activator of transcription (STAT) factor gene expression and chemokines as well as their receptors. **(A)** Correlation between STAT expression and multiple chemokines. **(B)** Correlation between STAT expression and multiple chemokine receptors.

positively correlated with MSH2, MSH6, and PMS2 ($p < 0.05$); *STAT3* and *STAT5B* were also positively correlated with all these genes, whereas *STAT2* and *STAT4* were negatively correlated with EPCAM ($p < 0.01$) (Figure 12C).

Discussion

Extensive studies have indicated that the abnormal regulation of STATs, especially for *STAT1/3/5*, is closely

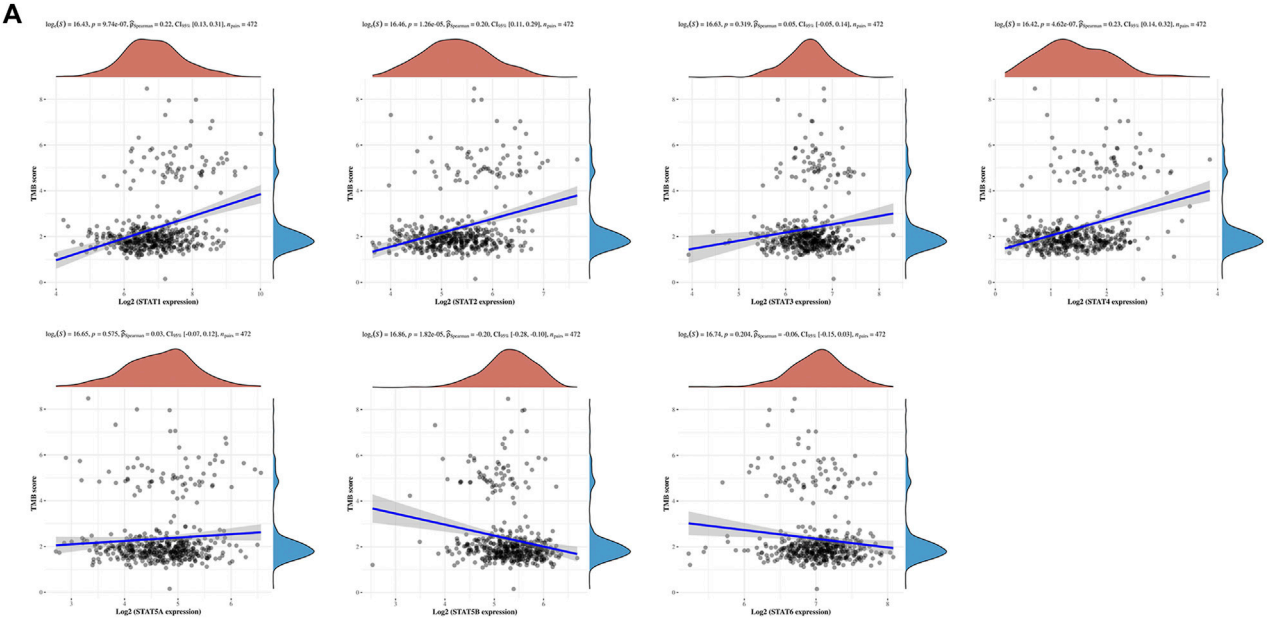


FIGURE 12
(Continued).

associated with the progression of various tumors, including solid tumors and hematologic malignancies, such as prostate cancer, breast cancer, CRC, and leukemias (Ferrajoli et al., 2006; Lassmann et al., 2007; Benekli et al., 2009; Gu et al., 2010; Koptyra et al., 2011). JAK/STAT signaling has also been identified as one of the key pathways affected by the majority of cancer gene mutations (Vogelstein and Kinzler, 2004). Although some studies have reported the role of certain STAT factors in cancer progression, there have been no studies that have comprehensively analyzed the role of different STATs in CRC. For the first time, bioinformatics analysis was used to investigate the transcription and translation levels, genetic variation, biological function, and molecular mechanism of STATs in CRC, as well as their correlation with prognosis, immune infiltration, and immunotherapy response.

Some studies demonstrate that activated *STAT1*, as a tumor suppressor, is lost in several types of malignant cells (Adámková et al., 2007), including breast cancer (Koromilas and Sexl, 2013), lung cancer (Chen et al., 2015), and esophageal squamous cell carcinoma (Zhang et al., 2014), and many reports indicate that high *STAT1* expression means better clinical prognosis (Widschwendter et al., 2002; Deng et al., 2012; Hosui et al., 2012). However, contradictory results have also been reported, showing that high *STAT1* expression levels are found in some cancers and correlate with poor prognosis compared to those with low expression levels, such as breast cancer, glioblastoma, lymphoma, and renal cell carcinoma (Khodarev et al., 2004; Duarte et al., 2012; Greenwood et al., 2012; Zhu et al., 2012; Arzt et al., 2014). In this study, database analysis showed that the transcription and translation levels of *STAT1* in CRC were higher than those in normal tissues, which was further verified by IHC staining. Furthermore, the expression of *STAT1* in patients with CRC was found to be significantly related to the tumor stage. A survival analysis revealed that a high *STAT1* transcription level had no correlation with the prognosis of patients with CRC but led to better OS in READ.

Previous reports suggest that *STAT2*, like *STAT1*, may play a dual role in cancer progression. Clifford et al. reported that sustained *STAT2* expression was required for interferon alpha-induced tumor-suppressive effects in skin squamous cell carcinoma cells, and the tumor-suppressive activity was also demonstrated in mice models (Clifford et al., 2003; Wang et al., 2003). The role of *STAT2* in the tumorigenesis of CRC and skin cancer has also been described (Gamero et al., 2010). The results of data analysis and IHC in this study indicated that *STAT2* was less expressed in CRC than in normal tissues at the transcription and translation level. However, *STAT2* expression in patients with CRC had no relationship with the tumor stage. A survival analysis showed that high *STAT2* expression was related to better OS and DFS in patients with CRC.

The bulk of evidence indicates that *STAT3* significantly correlates with cancer development and immune escape (Bromberg, 2002; Yu et al., 2009). Abnormal elevated *STAT3* activity has been found in a variety of hematological and solid malignancies such as acute myeloid leukemia (AML), multiple myeloma, and cancers of the bladder, head and neck, kidney,

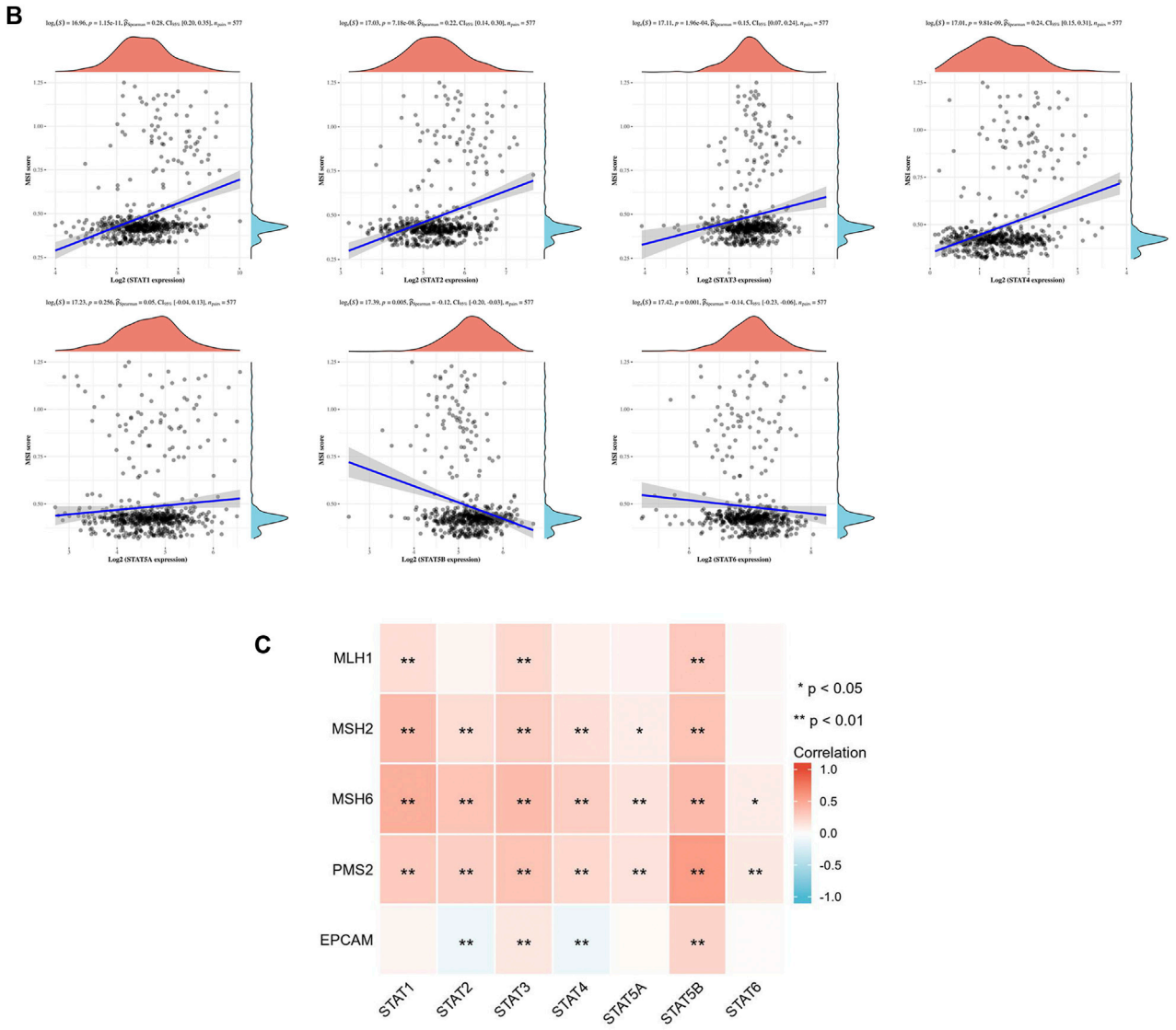


FIGURE 12
(Continued). Correlation between signal transducer and activator of transcription (STAT) factor expression and TMB score, MSI score, and MMR genes. **(A)** Correlation between STAT expression and TMB score. **(B)** Correlation between STAT expression and MSI score. **(C)** Correlation heatmap of STAT expression with MMR genes in CRC.

pancreas, uterus, ovary, esophagus, and breast (Chen et al., 2008; Sahu and Srivastava, 2009; Bar-Natan et al., 2012; Li et al., 2013; Suh et al., 2015; Geiger et al., 2016; Subramaniam et al., 2016; Zhang et al., 2016). In addition, high levels of phosphorylated *STAT3* expression often result in a poor prognosis in many types of cancer (Kusaba et al., 2006; Macha et al., 2011; Chen et al., 2013). However, controversial evidence has emerged showing that *STAT3* plays a negative role in the tumorigenesis of KRAS-induced lung cancer in mice models (Grabner et al., 2015). A high *STAT3* expression correlates with a better clinical outcome in patients with CRC and nasopharyngeal carcinoma (Hsiao et al., 2003; Gordziel et al., 2013). The analysis of the GEPIA dataset and IHC in this study showed that the expression level of *STAT3* in CRC was not different from that in adjacent normal tissues, but was associated with the tumor stage of patients with CRC. Survival analysis indicated that high *STAT3* expression was associated with better DFS, but did not affect OS in patients with CRC.

The exact role of *STAT4* in tumorigenesis remains unclear because high levels of *STAT4* expression have been shown to promote invasion and metastasis in gastric cancer and ovarian cancer (Zhou et al., 2014; Zhao et al., 2017), and potentially predict a favorable outcome in these cancers (Li et al., 2017;

Nishi et al., 2017). The results from this study demonstrated that although there was no significant difference in *STAT4* expression between CRC tissues and normal tissues, its expression was related to tumor stage, and increased *STAT4* expression level favored OS and DFS in patients with CRC.

STAT5, which consists of *STAT5A* and *STAT5B*, has been reported to be implicated in multiple malignancies. Hassel et al. (2008) reported that STAT5, a tumor promoter, aided cell proliferation and survival in melanoma by activating the antiapoptotic protein Bcl-XL. It also has been suggested that activated *STAT5* signaling promotes tumor growth, invasion, and epithelial-to-mesenchymal transition (EMT) in squamous cell carcinoma of the head and neck (SCCHN), leading to resistance to chemotherapy (Koppikar et al., 2008). Furthermore, highly activated *STAT5* results in poor prognosis in patients with prostate cancer (Li et al., 2005). However, increasing evidence suggests that *STAT5* proteins also regulate the activities of tumor suppressor genes, and activated *STAT5* is associated with a favorable prognosis in patients with breast cancer and nasopharyngeal cancer (Hsiao et al., 2003; Nevalainen et al., 2004). This study found that *STAT5A* and *STAT5B* were significantly less expressed in CRC compared to normal tissues at both the mRNA and protein levels. In patients with CRC, *STAT5B* had a significant relationship with tumor stage and exerted a favorable effect on OS and DFS but *STAT5A* did not.

Some studies have reported that activated *STAT6* signaling is important for IL-4 and IL-13-induced EMT and CRC cell aggressiveness (Cao et al., 2016; Chen et al., 2018). Furthermore, inhibition studies have indicated that targeting *STAT6* signaling can suppress tumor growth and metastasis in gastric cancer (Lu et al., 2018). The data analysis revealed that there was no significant difference in *STAT6* expression between CRC tissues and normal tissues, and *STAT6* expression level did not affect tumor stage and prognosis.

The analysis outcomes showed that there was a high mutation rate (43%) of STATs in patients with CRC, and the genes with the highest and lowest mutation rates in STATs are *STAT5B* (15%) and *STAT4* (7%), respectively. Besides, the mutation of STATs had some relationship with certain types of immune cells infiltration, but showed little effect on prognosis in patients with CRC. The co-expression relationships between different STAT members in patients with CRC indicated that these factors may play a synergistic role in the progression of CRC. To evaluate the potential interactions between STAT factors and their neighboring genes at different levels, GGI and PPI networks were constructed. GO and KEGG enrichment analyses were also performed to explore the functions of STATs and their similar genes, which are primarily related to lymphocyte activation, cytokine-mediated signaling pathway, positive regulation of immune response, and T-cell receptor signaling pathway. These pathways have a close relationship with the immune system, indicating the possible role of STATs in regulating the tumor immune microenvironment.

The study results indicated that STAT transcription levels were closely correlated with levels of immune infiltration in CRC. *STAT1/2/3/4* expressions were positively correlated with the infiltrations of B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells. The expressions of *STAT5A/5B/6* were positively associated with infiltrations of $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, and *STAT5A* expression was also positively correlated with the infiltration of B cells. Also, the expressions of *STAT1/2/3/4/5A* were positively correlated with the stromal score, immune score, and ESTIMATE score with great significance. The close relationship between the expression levels of STATs and infiltration of multiple types of immune cells further demonstrates that STATs may be the regulators of tumor immunity in CRC.

Another significant finding of this study is that STATs may be possible predictors of treatment response to immunotherapy in CRC because of the close relationship between STATs and indicators associated with immunotherapy response, such as immune checkpoint genes, TMB score, MSI score, and MMR genes. Our results showed that seven STAT members were positively co-expressed with eight immune checkpoints including CD40LG, ADORA2A, TNFSF14, ICOSLG, TNFRSF8, CD27, VSIR, and TNFRSF4. From literature searches, it was found that CD40LG and CD27, both of which are stimulatory immune checkpoints, are more likely to be used clinically. Extensive studies have been conducted to investigate the underlying mechanism by which CD40LG and CD27 modulate tumor immunity. Additionally, various types of strategies targeting them, such as agonistic/antagonistic monoclonal antibodies, cellular vaccines, and protein antagonists, have been developed and demonstrated to be safe and efficacious in early clinical trials (Starzer and Berghoff, 2020; Tang et al., 2021). Coincidentally, the factor that is most closely related to both CD40LG and CD27 is *STAT4* based on their correlation coefficient (Table 3). It was discovered that almost all the STAT factors were positively co-expressed with certain MMR genes such as MSH2, MSH6, and PMS2. A positive relationship between the TMB score and STAT1/2/4 expression and a negative relationship between the TMB score and *STAT5B* expression were observed. Moreover, *STAT1/2/3/4* expression had positive correlations with the MSI score, but *STAT5B* and *STAT6* had a negative relationship with the MSI score.

This study indicates that the expression levels of *STAT2/5A/5B* are downregulated in CRC and could inhibit the initiation and development of CRC. The close relationship

between the CRC stages and expression levels of *STAT1/3/4/5B* reveals their potential as molecular biomarkers for tumor stage classification. Moreover, the abnormal expressions of *STAT2/4/5B* have the potential to be used as prognostic predictors in patients with CRC. Besides, the strong association between the expression of STAT and infiltration of multiple types of immune cells in CRC, including B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells, demonstrates that STATs may play a role in the regulation of CRC tumor immunity. More importantly, the significant correlation between STAT expressions and immunotherapy response-associated indicators showed that they had the potential to predict response to immunotherapy in patients with CRC and could be used to assist the physician in deciding on a therapeutic regimen. These findings aid in better understanding the molecular landscape of CRC progression, providing new prognostic biomarkers, and promoting the development of more immunotherapeutic strategies for patients with CRC. However, further investigations are still needed to validate the results of this study to facilitate the clinical application of STATs as therapy targets, prognostic biomarkers, and immunotherapy predictors in CRC.

Data availability statement

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

Ethics statement

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

Author contributions

Conceptualization, XD and GD; data curation, LJ; formal analysis, SH and WZ; investigation, PC; methodology, WG and YY; project administration, XW; software, DL and WS; visualization, ZM; writing—original draft, DL; writing—review and editing, YJ and YZ. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors thank Bullet Edits Limited for the linguistic editing and proofreading 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/fgene.2022.951252/full#supplementary-material

References

Adámková, L., Soucková, K., and Kovarík, J. (2007). Transcription protein *STAT1*: biology and relation to cancer. *Folia Biol.* 53 (1), 1–6.

Arzt, L., Kothmaier, H., Halbwedl, I., Quehenberger, F., and Popper, H. H. (2014). Signal transducer and activator of transcription 1 (*STAT1*) acts like an oncogene in malignant pleural mesothelioma. *Virchows Arch.* 465 (1), 79–88. doi:10.1007/s00428-014-1584-8

Bar-Natan, M., Nelson, E. A., Xiang, M., and Frank, D. A. (2012). Stat signaling in the pathogenesis and treatment of myeloid malignancies. *Jakstat* 1 (2), 55–64. doi:10.4161/jkst.20006

Benekli, M., Baumann, H., and Wetzler, M. (2009). Targeting signal transducer and activator of transcription signaling pathway in leukemias. *J. Clin. Oncol.* 27 (26), 4422–4432. doi:10.1200/jco.2008.21.3264

Bowman, T., Garcia, R., Turkson, J., and Jove, R. (2000). Stats in oncogenesis. *Oncogene* 19 (21), 2474–2488. doi:10.1038/sj.onc.1203527

Brenner, H., Kloor, M., and Pox, C. P. (2014). Colorectal cancer. *Lancet* 383 (9927), 1490–1502. doi:10.1016/s0140-6736(13)61649-9

Bromberg, J. (2002). Stat proteins and oncogenesis. *J. Clin. Invest.* 109 (9), 1139–1142. doi:10.1172/jci15617

Cao, H., Zhang, J., Liu, H., Wan, L., Zhang, H., Huang, Q., et al. (2016). Il-13/*STAT6* signaling plays a critical role in the epithelial-mesenchymal transition of colorectal cancer cells. *Oncotarget* 7 (38), 61183–61198. doi:10.18632/oncotarget.11282

Chen, C. L., Cen, L., Kohout, J., Hutzen, B., Chan, C., Hsieh, F. C., et al. (2008). Signal transducer and activator of transcription 3 activation is associated with bladder cancer cell growth and survival. *Mol. Cancer* 7, 78. doi:10.1186/1476-4598-7-78

Chen, Y., Wang, J., Wang, X., Liu, X., Li, H., Lv, Q., et al. (2013). *STAT3*, a poor survival predicator, is associated with lymph node metastasis from breast cancer. *J. Breast Cancer* 16 (1), 40–49. doi:10.4048/jbc.2013.16.1.40

Chen, J., Zhao, J., Chen, L., Dong, N., Ying, Z., Cai, Z., et al. (2015). *STAT1* modification improves therapeutic effects of interferons on lung cancer cells. *J. Transl. Med.* 13, 293. doi:10.1186/s12967-015-0656-0

Chen, J., Gong, C., Mao, H., Li, Z., Fang, Z., Chen, Q., et al. (2018). E2f1/Sp3/*STAT6* Axis is required for il-4-induced epithelial-mesenchymal transition of colorectal cancer cells. *Int. J. Oncol.* 53 (2), 567–578. doi:10.3892/ijo.2018.4429

Clifford, J. L., Yang, X., Walch, E., Wang, M., and Lippman, S. M. (2003). Dominant negative signal transducer and activator of transcription 2 (*STAT2*) protein: Stable expression blocks interferon alpha action in skin squamous cell carcinoma cells. *Mol. Cancer Ther.* 2 (5), 453–459.

Deng, H., Zhen, H., Fu, Z., Huang, X., Zhou, H., Liu, L., et al. (2012). The antagonistic effect between *STAT1* and survivin and its clinical significance in gastric cancer. *Oncol. Lett.* 3 (1), 193–199. doi:10.3892/ol.2011.423

Duarte, C. W., Willey, C. D., Zhi, D., Cui, X., Harris, J. J., Vaughan, L. K., et al. (2012). Expression signature of ifn/*STAT1* signaling genes predicts poor survival outcome in glioblastoma multiforme in a subtype-specific manner. *PLoS One* 7 (1), e29653. doi:10.1371/journal.pone.0029653

Ferrajoli, A., Faderl, S., Ravandi, F., and Estrov, Z. (2006). The jak-stat pathway: a therapeutic target in hematological malignancies. *Curr. Cancer Drug Targets* 6 (8), 671–679. doi:10.2174/156800906779010227

Franz, M., Rodriguez, H., Lopes, C., Zuberi, K., Montojo, J., Bader, G. D., et al. (2018). Genemania update 2018. *Nucleic Acids Res.* 46 (W1), W60–W64. doi:10.1093/nar/gky311

Gamero, A. M., Young, M. R., Mentor-Marcel, R., Bobe, G., Scarzello, A. J., Wise, J., et al. (2010). *STAT2* contributes to promotion of colorectal and skin carcinogenesis. *Cancer Prev. Res.* 3 (4), 495–504. doi:10.1158/1940-6207.capr-09-0105

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

Geiger, J. L., Grandis, J. R., and Bauman, J. E. (2016). The *STAT3* pathway as a therapeutic target in head and neck cancer: barriers and innovations. *Oral Oncol.* 56, 84–92. doi:10.1016/j.oraloncology.2015.11.022

Gordziel, C., Bratsch, J., Moriggl, R., Knösel, T., and Friedrich, K. (2013). Both *STAT1* and *STAT3* are favourable prognostic determinants in colorectal carcinoma. *Br. J. Cancer* 109 (1), 138–146. doi:10.1038/bjc.2013.274

Grabner, B., Schramek, D., Mueller, K. M., Moll, H. P., Svinka, J., Hoffmann, T., et al. (2015). Disruption of *STAT3* signalling promotes kras-induced lung tumorigenesis. *Nat. Commun.* 6, 6285. doi:10.1038/ncomms7285

Greenwood, C., Metodieva, G., Al-Janabi, K., Lausen, B., Alldridge, L., Leng, L., et al. (2012). *STAT1* and Cd74 overexpression is Co-dependent and linked to increased invasion and lymph node metastasis in triple-negative breast cancer. *J. Proteomics* 75 (10), 3031–3040. doi:10.1016/j.jprot.2011.11.033

GTEx Consortium (2013). The genotype-tissue expression (gtex) project. *Nat. Genet.* 45 (6), 580–585. doi:10.1038/ng.2653

Gu, L., Vogiatzi, P., Puhr, M., Dagvadorj, A., Lutz, J., Ryder, A., et al. (2010). Stat5 promotes metastatic behavior of human prostate cancer cells *in vitro* and *in vivo*. *Endocr. Relat. Cancer* 17 (2), 481–493. doi:10.1677/erc-09-0328

Hassel, J. C., Winnemöller, D., Schartl, M., and Wellbrock, C. (2008). Stat5 contributes to antiapoptosis in melanoma. *Melanoma Res.* 18 (6), 378–385. doi:10.1097/CMR.0b013e32830ce7d7

Hosui, A., Klover, P., Tatsumi, T., Uemura, A., Nagano, H., Doki, Y., et al. (2012). Suppression of signal transducers and activators of transcription 1 in hepatocellular carcinoma is associated with tumor progression. *Int. J. Cancer* 131 (12), 2774–2784. doi:10.1002/ijc.27580

Hsiao, J. R., Jin, Y. T., Tsai, S. T., Shiau, A. L., Wu, C. L., Su, W. C., et al. (2003). Constitutive activation of *STAT3* and Stat5 is present in the majority of nasopharyngeal carcinoma and correlates with better prognosis. *Br. J. Cancer* 89 (2), 344–349. doi:10.1038/sj.bjc.6601003

Jiang, L., Zhao, X. H., Mao, Y. L., Wang, J. F., Zheng, H. J., You, Q. S., et al. (2019). Long non-coding Rna Rp11-468e2.5 curtails colorectal cancer cell proliferation and stimulates apoptosis via the Jak/Stat signaling pathway by targeting Stat5 and *STAT6*. *J. Exp. Clin. Cancer Res.* 38 (1), 465. doi:10.1186/s13046-019-1428-0

Johnson, D. E., O'Keefe, R. A., and Grandis, J. R. (2018). Targeting the il-6/jak/*STAT3* signalling Axis in cancer. *Nat. Rev. Clin. Oncol.* 15 (4), 234–248. doi:10.1038/nrclinonc.2018.8

Khodarev, N. N., Beckett, M., Labay, E., Darga, T., Roizman, B., Weichselbaum, R. R., et al. (2004). *STAT1* is overexpressed in tumors selected for radioresistance and confers protection from radiation in transduced sensitive cells. *Proc. Natl. Acad. Sci. U. S. A.* 101 (6), 1714–1719. doi:10.1073/pnas.0308102100

Koppikar, P., Lui, V. W., Man, D., Xi, S., Chai, R. L., Nelson, E., et al. (2008). Constitutive activation of signal transducer and activator of transcription 5 contributes to tumor growth, epithelial-mesenchymal transition, and resistance to epidermal growth factor receptor targeting. *Clin. Cancer Res.* 14 (23), 7682–7690. doi:10.1158/1078-0432.ccr-08-1328

Koptyra, M., Gupta, S., Talati, P., and Nevalainen, M. T. (2011). Signal transducer and activator of transcription 5a/B: Biomarker and therapeutic target in prostate and breast cancer. *Int. J. Biochem. Cell Biol.* 43 (10), 1417–1421. doi:10.1016/j.biocel.2011.06.007

Koromilas, A. E., and Sexl, V. (2013). The tumor suppressor function of *STAT1* in breast cancer. *Jakstat* 2 (2), e23353. doi:10.4161/jkst.23353

Kusaba, T., Nakayama, T., Yamazumi, K., Yakata, Y., Yoshizaki, A., Inoue, K., et al. (2006). Activation of *STAT3* is a marker of poor prognosis in human colorectal cancer. *Oncol. Rep.* 15 (6), 1445–1451.

Lassmann, S., Schuster, I., Walch, A., Göbel, H., Jütting, U., Makowiec, F., et al. (2007). *STAT3* mrna and protein expression in colorectal cancer: effects on *STAT3*-inducible targets linked to cell survival and proliferation. *J. Clin. Pathol.* 60 (2), 173–179. doi:10.1136/jcp.2005.035113

Li, H., Zhang, Y., Glass, A., Zellweger, T., Gehan, E., Bubendorf, L., et al. (2005). Activation of signal transducer and activator of transcription-5 in prostate cancer predicts early recurrence. *Clin. Cancer Res.* 11 (16), 5863–5868. doi:10.1158/1078-0432.ccr-05-0562

Li, S., Priceman, S. J., Xin, H., Zhang, W., Deng, J., Liu, Y., et al. (2013). Icaritin inhibits jak/*STAT3* signaling and growth of renal cell carcinoma. *PLoS One* 8 (12), e81657. doi:10.1371/journal.pone.0081657

Li, S., Sheng, B., Zhao, M., Shen, Q., Zhu, H., Zhu, X., et al. (2017). The prognostic values of signal transducers activators of transcription family in ovarian cancer. *Biosci. Rep.* 37 (4), BSR20170650. doi:10.1042/bsr20170650

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

Lu, G., Shi, W., and Zheng, H. (2018). Inhibition of *STAT6*/anoctamin-1 activation suppresses proliferation and invasion of gastric cancer cells. *Cancer Biother. Radiopharm.* 33 (1), 3–7. doi:10.1089/cbr.2017.2287

Macha, M. A., Matta, A., Kaur, J., Chauhan, S. S., Thakar, A., Shukla, N. K., et al. (2011). Prognostic significance of nuclear P*STAT3* in oral cancer. *Head Neck* 33 (4), 482–489. doi:10.1002/hed.21468

Mizuno, H., Kitada, K., Nakai, K., and Sarai, A. (2009). Prognoscan: a new database for meta-analysis of the prognostic value of genes. *BMC Med. Genomics* 2, 18. doi:10.1186/1755-8794-2-18

Nagy, Á., Lánczky, A., Menyhárt, O., and Győrffy, B. (2018). Validation of mirna prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci. Rep.* 8 (1), 9227. doi:10.1038/s41598-018-27521-y

Nevalainen, M. T., Xie, J., Torhorst, J., Bubendorf, L., Haas, P., Kononen, J., et al. (2004). Signal transducer and activator of transcription-5 activation and breast cancer prognosis. *J. Clin. Oncol.* 22 (11), 2053–2060. doi:10.1200/jco.2004.11.046

Nishi, M., Batsaikhan, B. E., Yoshikawa, K., Higashijima, J., Tokunaga, T., Takasu, C., et al. (2017). High *STAT4* expression indicates better disease-free survival in patients with gastric cancer. *Anticancer Res.* 37 (12), 6723–6729. doi:10.21873/anticanres.12131

Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., et al. (2019). Tisidb: an integrated repository portal for tumor-immune system interactions. *Bioinformatics* 35 (20), 4200–4202. doi:10.1093/bioinformatics/btz210

Sahu, R. P., and Srivastava, S. K. (2009). The role of stat-3 in the induction of apoptosis in pancreatic cancer cells by benzyl isothiocyanate. *J. Natl. Cancer Inst.* 101 (3), 176–193. doi:10.1093/jnci/djn470

Starzer, A. M., and Berghoff, A. S. (2020). New emerging targets in cancer immunotherapy: Cd27 (Tnfrsf7). *ESMO Open* 4, e000629. doi:10.1136/esmoopen-2019-000629

Subramaniam, K. S., Omar, I. S., Kwong, S. C., Mohamed, Z., Woo, Y. L., Mat Adenan, N. A., et al. (2016). Cancer-associated fibroblasts promote endometrial cancer growth via activation of interleukin-6/stat-3/C-myc pathway. *Am. J. Cancer Res.* 6 (2), 200–213.

Suh, Y. A., Jo, S. Y., Lee, H. Y., and Lee, C. (2015). Inhibition of il-6/*STAT3* Axis and targeting axl and Tyro3 receptor tyrosine kinases by apigenin circumvent taxol resistance in ovarian cancer cells. *Int. J. Oncol.* 46 (3), 1405–1411. doi:10.3892/ijo.2014.2808

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

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). String V11: Protein-protein association networks with increased coverage,

supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47 (D1), D607–D613. doi:10.1093/nar/gky1131

Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). Gepia2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* 47 (W1), W556–W560. doi:10.1093/nar/gkz430

Tang, T., Cheng, X., Truong, B., Sun, L., Yang, X., Wang, H., et al. (2021). Molecular basis and therapeutic implications of Cd40/Cd40l immune checkpoint. *Pharmacol. Ther.* 219, 107709. doi:10.1016/j.pharmthera.2020.107709

Verhoeven, Y., Tilborghs, S., Jacobs, J., De Waele, J., Quatannens, D., Deben, C., et al. (2020). The potential and controversy of targeting stat family members in cancer. *Semin. Cancer Biol.* 60, 41–56. doi:10.1016/j.semcancer.2019.10.002

Vogelstein, B., and Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nat. Med.* 10 (8), 789–799. doi:10.1038/nm1087

Wang, J., Pham-Mitchell, N., Schindler, C., and Campbell, I. L. (2003). Dysregulated sonic hedgehog signaling and medulloblastoma consequent to ifn-alpha-stimulated *STAT2*-independent production of ifn-gamma in the brain. *J. Clin. Invest.* 112 (4), 535–543. doi:10.1172/jci18637

Widschwendter, A., Tonko-Geymayer, S., Welte, T., Daxenbichler, G., Marth, C., Doppler, W., et al. (2002). Prognostic significance of signal transducer and activator of transcription 1 activation in breast cancer. *Clin. Cancer Res.* 8 (10), 3065–3074.

Yoshioka, R., Hasegawa, K., Mise, Y., Oba, M., Aoki, T., Sakamoto, Y., et al. (2014). Evaluation of the safety and efficacy of simultaneous resection of primary colorectal cancer and synchronous colorectal liver metastases. *Surgery* 155 (3), 478–485. doi:10.1016/j.surg.2013.10.015

Yu, H., Pardoll, D., and Jove, R. (2009). Stats in cancer inflammation and immunity: A leading role for *STAT3*. *Nat. Rev. Cancer* 9 (11), 798–809. doi:10.1038/nrc2734

Zacharakis, M., Xynos, I. D., Lazaris, A., Smaro, T., Kosmas, C., Dokou, A., et al. (2010). Predictors of survival in stage iv metastatic colorectal cancer. *Anticancer Res.* 30 (2), 653–660.

Zhang, Y., Molavi, O., Su, M., and Lai, R. (2014). The clinical and biological significance of *STAT1* in esophageal squamous cell carcinoma. *BMC Cancer* 14, 791. doi:10.1186/1471-2407-14-791

Zhang, H. F., Chen, Y., Wu, C., Wu, Z. Y., Tweardy, D. J., Alshareef, A., et al. (2016). The opposing function of *STAT3* as an oncoprotein and tumor suppressor is dictated by the expression status of *STAT3*β in esophageal squamous cell carcinoma. *Clin. Cancer Res.* 22 (3), 691–703. doi:10.1158/1078-0432.ccr-15-1253

Zhao, L., Ji, G., Le, X., Luo, Z., Wang, C., Feng, M., et al. (2017). An integrated analysis identifies *STAT4* as a key regulator of ovarian cancer metastasis. *Oncogene* 36 (24), 3384–3396. doi:10.1038/onc.2016.487

Zhou, X., Xia, Y., Su, J., and Zhang, G. (2014). Down-regulation of mir-141 induced by Helicobacter pylori promotes the invasion of gastric cancer by targeting *STAT4*. *Cell. Physiol. Biochem.* 33 (4), 1003–1012. doi:10.1159/000358671

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented Resource for the analysis of systems-level datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6

Zhu, H., Wang, Z., Xu, Q., Zhang, Y., Zhai, Y., Bai, J., et al. (2012). Inhibition of *STAT1* sensitizes renal cell carcinoma cells to radiotherapy and chemotherapy. *Cancer Biol. Ther.* 13 (6), 401–407. doi:10.4161/cbt.19291

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 September 2022
DOI 10.3389/fgene.2022.1000448

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Changle Ji,
Tongji University, China
RuiJin Xie,
Affiliated Hospital of Jiangnan University, China
Zijian Zhou,
Huashan Hospital, Fudan University, China

*CORRESPONDENCE
Yong Shao,
shaoyutong1975@163.com
Zhengxiang Han,
cnhzxyq@163.com

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 22 July 2022
ACCEPTED 24 August 2022
PUBLISHED 09 September 2022

CITATION
Liu H, Qin X, Xu Z, Wu M, Lu T, Zhou S, Yao N, Liu S, Shao Y and Han Z (2022), Comparison of effectiveness and safety of camrelizumab between HBV-related and non-B, non-C hepatocellular carcinoma: A retrospective study in China.
Front. Genet. 13:1000448.
doi: 10.3389/fgene.2022.1000448

Comparison of effectiveness and safety of camrelizumab between HBV-related and non-B, non-C hepatocellular carcinoma: A retrospective study in China

Haonan Liu[1†], Xiaobing Qin[1†], Zhiyuan Xu[2], Meng Wu[1], Tong Lu[3], Shuang Zhou[1], Nan Yao[1], Suya Liu[1], Yong Shao[4*] and Zhengxiang Han[1*]

[1]Department of Oncology, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, China, [2]Department of Emergency, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, China, [3]Department of Gastroenterology, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, China, [4]Department of General Surgery, The Affiliated Hospital of Xuzhou Medical University, Xuzhou, China

Purpose: This study aimed to compare the clinical outcomes of camrelizumab in hepatitis B virus-related hepatocellular carcinoma (HBV–HCC) patients and non-HBV, non-HCV hepatocellular carcinoma (NBNC–HCC) patients in China.

Materials and methods: A total of 54 patients with hepatocellular carcinoma who received camrelizumab were included in this retrospective study from January 2019 to December 2021. The patients were assigned to the HBV–HCC group (n = 28) and the NBNC–HCC group (n = 26). The primary endpoints were overall survival (OS) and progression-free survival (PFS), and the secondary endpoints were the objective response rate (ORR), disease control rate (DCR), and adverse events (AEs). Multivariate analysis using Cox proportional hazard regression was used to identify independent prognostic factors. A nomogram model was subsequently established based on independent prognostic factors.

Results: The mean duration of follow-up was 12.7 ± 3.6 months. The median OS was not determined. The median PFS in the HBV–HCC group was significantly longer than that in the NBNC–HCC group (9.2 vs. 6.7 months, p = 0.003). The ORR and DCR in the HBV–HCC group were significantly higher than those in the NBNC–HCC group (ORR, 28.6% vs. 7.7%, p = 0.048; DCR, 71.4% vs. 42.3%, p = 0.031). No significant differences in the total incidence of AEs were found between the HBV–HCC group and the NBNC–HCC group (75.0% vs. 69.2%, p = 0.224). Multivariate regression analysis identified etiology, AFP level, and vascular invasion as independent prognostic factors (all $p < 0.05$).

Conclusion: Our findings demonstrate that camrelizumab is more effective in HBV–HCC patients than in NBNC–HCC patients, with manageable safety.

KEYWORDS

hepatitis B virus, hepatocellular carcinoma, immunotherapy, camrelizumab, effectiveness, safety

Introduction

Hepatocellular carcinoma (HCC) is the seventh most common malignancy and the third most common cause of cancer death worldwide. Specifically, HCC cases in China comprise approximately more than half of the total number of cases globally (Zheng et al., 2018; Sung et al., 2021; Xia et al., 2022).

Chronic infection with the hepatitis B virus (HBV) and hepatitis C virus (HCV) is a major risk factor for HCC; in China, HCC caused by HBV comprises approximately 70%–80% (Wang et al., 2017). With the recent popularization of the HBV vaccine and the widespread application of antiviral drugs, the proportion of HBV-related hepatocellular carcinoma (HBV–HCC) is expected to gradually decrease, whereas that of non-HBV non-HCV hepatocellular carcinoma (NBNC–HCC) is expected to gradually increase (Thomas, 2019). The pathogenesis of NBNC–HCC is yet to be fully understood, but nonalcoholic steatohepatitis (NASH) and metabolic syndrome are regarded as important pathogenic factors (Anstee et al., 2019).

HCC is typically insidious and mostly diagnosed in the intermediate to advanced stages (Maluccio and Covey, 2012). Less than 30% of those diagnosed with the disease undergo radical resection (Ikeda et al., 2018). Immunotherapy provides an option for patients with advanced HCC (Cabibbo et al., 2010; Yoon et al., 2018). Results of the IMbrave150 study show that atezolizumab combined with bevacizumab (T + A) significantly improved overall survival (OS) and progression-free survival (PFS) (Cheng et al., 2022). In addition, the combination (T + A) was approved for the first-line treatment of advanced HCC (Finn et al., 2020; Cheng et al., 2022). It is the first treatment regimen significantly superior to sorafenib in more than 10 years since the approval of sorafenib for first-line treatment of HCC (Sangro et al., 2021). On the basis of the results of the Checkmate040 and KEYNOTE224 trials, two programmed cell death protein 1 (PD-1) inhibitors—nivolumab and pembrolizumab—were approved by the United States Food and Drug Administration (FDA) for second-line therapy of HCC (El-Khoueiry et al., 2017; Zhu et al., 2018). Qin et al. (2020) found that camrelizumab could yield an objective response rate (ORR) of 14.7% and a 6-month OS rate of 74.4% in 217 HCC patients. This finding indicates that camrelizumab could achieve clinical efficacy comparable to that of similar PD-1 inhibitors. Thus, camrelizumab became the first PD-1 inhibitor approved in China for the treatment of intermediate and advanced HCC. In addition, the most significant advantage of camrelizumab for Chinese patients is its relatively low price, hence its wide use in China (Chen et al., 2020).

Although immunotherapy has achieved satisfactory efficacy in HCC patients, international multicenter phase III studies showed that the response rate of NBNC–HCC to immunotherapy was lower than that of viral hepatitis-related HCC (Pfister et al., 2021; Ji and Nguyen, 2022; Kelley et al., 2022). The efficacy of immunotherapy may be influenced by different underlying etiologies of HCC, and different hepatic environments can significantly affect HCC cell induction and immune responses (Roderburg et al., 2020). Compared with non-viral-related HCC, the tumor immune microenvironment of HBV–HCC had a stronger immunosuppressive effect, which was reversed by PD-1 inhibitors (Liu et al., 2021). In addition, viral antigens expressed by tumor cells have been found to increase the number of antigen-specific T cells and enhance response to immunotherapy (Yarchoan et al., 2017).

However, research is lacking on immunotherapy for NBNC–HCC in China. Therefore, in order to identify clinically relevant factors potentially influencing the response to immunotherapy in HCC patients, this study aims to compare the effectiveness and safety of camrelizumab for HBV–HCC and NBNC–HCC in China and to further select HCC patients suitable for immunotherapy by etiology.

Materials and methods

Study design and patients

We retrospectively reviewed the medical records of patients with HCC who underwent camrelizumab from January 2019 to December 2021 at the Affiliated Hospital of Xuzhou Medical University.

The inclusion criteria were as follows: (Sung et al., 2021) diagnosed advanced HCC with clinical and histopathological evidence; (Zheng et al., 2018) aged 18 years or older; (Xia et al., 2022) at least one measurable lesion as defined by modified Response Evaluation Criteria in Solid Tumors (mRECIST); (Wang et al., 2017) received at least two cycles of camrelizumab; (Thomas, 2019) Child–Pugh class A or B; (Anstee et al., 2019) Barcelona Clinic Liver Cancer (BCLC) stage B or C; and (Maluccio and Covey, 2012) Eastern Cooperative Oncology Group Performance Status (ECOG PS) ≤ 2.

The exclusion criteria were as follows: (Sung et al., 2021) patients with other malignant tumors; (Zheng et al., 2018) intrahepatic cholangiocarcinoma, and combined hepatocellular–cholangiocarcinoma; (Xia et al., 2022)

coinfection with hepatitis C or other hepatitis viruses; and (Wang et al., 2017) incomplete clinical data.

Grouping and treatment protocol

All included patients were assigned to two groups—the HBV–HCC group and the NBNC–HCC group—on the basis of serum hepatitis B surface antigen, serum hepatitis C surface antigen, HBV–DNA, and HCV–DNA. The HBV–HCC group consisted of patients with positive serum hepatitis B surface antigen or positive serum HBV–DNA, and the NBNC–HCC group consisted of patients with negative serum hepatitis B surface antigen, negative serum hepatitis C surface antigen, negative serum HBV–DNA, and negative serum HCV–DNA.

Camrelizumab (200 mg/branch) was administered intravenously, one branch at a time, once every 3 weeks.

Assessment

The patients underwent contrast-enhanced chest computed tomography (CT) or contrast-enhanced magnetic resonance imaging (MRI) of the upper abdomen at baseline and every 6–8 weeks thereafter. The long-term efficacy of treatment was measured by OS and PFS. OS was defined as the time from the start of camrelizumab treatment to death from any cause, and PFS was defined as the time from the start of camrelizumab treatment to disease progression or death from any cause. The short-term efficacy was assessed in accordance with the mRECIST criteria, including complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD). All adverse events (AEs) were recorded in detail, and AEs were evaluated based on the National Cancer Institute Common Terminology Criteria for Adverse Events (version 5.0).

Statistical analysis

All data were statistically analyzed using the software SPSS 23.0 and R 4.2.0. Normally distributed continuous variables were presented as the mean ± standard deviation and analyzed using the independent *t*-test. Categorical data were expressed as counts and percentages and then analyzed using the chi-squared test or Fisher's exact probability test. The Kaplan-Meier method was used to estimate median OS and PFS. Log-rank tests were used in univariate analysis, and variables with a p value less than 0.1 were used in multivariate analysis. Multivariate Cox proportional hazard regression analysis was used to clarify the independent prognostic factors of PFS. A nomogram prediction model was constructed using the independent prognostic factors from multivariate analysis. Nomograms for 6-month PFS and 9-month PFS were also developed and then compared with the actual scenario. Internal validation and the accuracy of the nomogram were determined using the bootstrap method and the calculated concordance index (C-index). p value <0.05 was considered statistically significant.

Results

Patient characteristics

A total of 54 patients with HCC classified as BCLC stage B or C and who are receiving camrelizumab from January 2019 to December 2021 were included in this retrospective study. The HBV–HCC group consisted of 28 patients, and the NBNC–HCC group had 26 patients. The mean follow-up time was 12.7 ± 3.6 months as of March 2022. Fifteen patients were excluded for the following reasons: HCV-infected (n = 3), with incomplete clinical data (n = 3), lost to follow-up (n = 4), and accepted other PD-1 inhibitors during follow-up (n = 5) (Figure 1). The causes of NBNC–HCC were NAFLD (non-alcoholic fatty liver disease) (n = 22), alcoholic liver disease (n = 3), and autoimmune liver diseases (n = 1). Patient data, including age, gender, tumor size, cirrhosis, vascular invasion, extrahepatic metastasis, ECOG PS, Child–Pugh class, BCLC stage, previous treatment, and laboratory parameters are summarized in Table 1. No significant differences in baseline characteristics were found between the two groups (all $p > 0.05$).

Tumor response

In the HBV–HCC group, 1 patient had CR, 7 patients had PR, 12 patients had SD, and 8 patients had PD. In the NBNC–HCC group, no patients showed CR, 2 patients exhibited PR, 9 patients had SD, and 15 patients had PD (Table 2). The ORR and disease control rate (DCR) of the HBV–HCC group were greater than those of the NBNC–HCC group (ORR, 28.6% vs. 7.7%, p = 0.048; DCR, 71.4% vs. 42.3%, p = 0.031).

Progression-free survival and overall survival

Median PFS was 9.2 months (95%CI: 7.4–11.0) in the HBV–HCC group and 6.7 months (95%CI: 5.0–8.4) in the NBNC–HCC group (Figure 2). The median PFS of the HBV–HCC group was significantly longer than that of the NBNC–HCC group (p = 0.003). As of follow-up time,

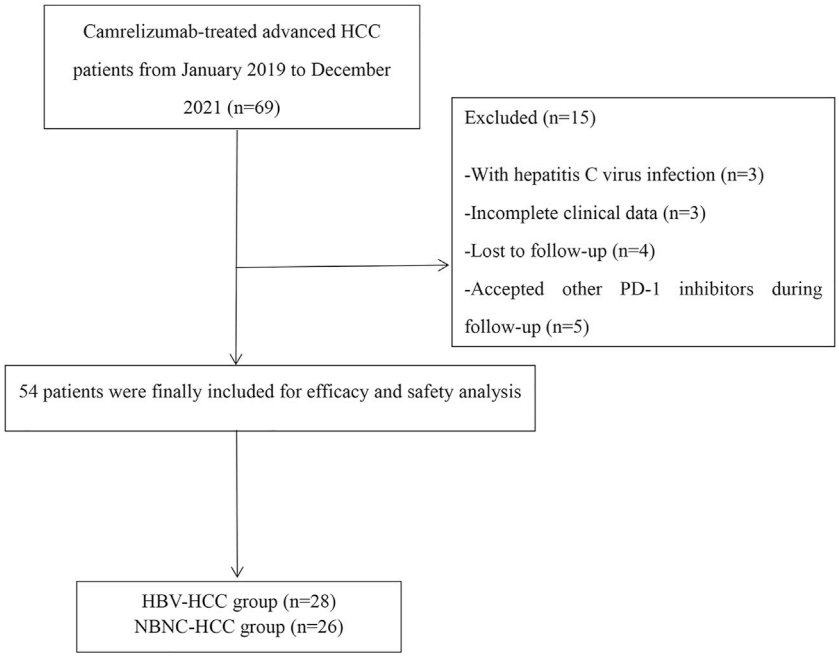


FIGURE 1
Patient flowchart.

21 patients died, including 11 from the HBV–HCC group and 10 from the NBNC–HCC group. The median OS of the HBV–HCC group was 13.3 months (95%CI: 11.4–15.2), which was significantly longer than that of the NBNC–HCC group (i.e., 11.1 months) (95%CI: 9.7–12.5) (p = 0.018) (Figure 3). The median OS still required further observation and follow-up because of the short enrollment time of patients.

Univariate and multivariate analyses of Progression-free survival

The results of the univariate and multivariate Cox proportional hazard regression analyses for PFS are summarized in Table 3. In multivariate Cox proportional hazard regression analysis, etiology (HR = 0.192, 95%CI: 0.088–0.418, p = 0.001), the AFP level (HR = 2.893, 95%CI: 1.233–6.787, p = 0.015), and vascular invasion (HR = 3.158, 95%CI: 1.436–6.942, p = 0.004) were identified as independent prognostic factors for the PFS of HCC patients. These independent prognostic factors were considered taken in the construction of a nomogram prediction model (Figure 4). The C-index of the nomogram prediction model was 0.781. The calibration curves indicated consistency between the actual and the predicted survival rates of the patients (Figure 5).

Adverse events

The two groups of patients mainly reported Grade 1–2 AEs, with reactive cutaneous capillary endothelial proliferation (RCCEP), proteinuria, hyperbilirubinemia, elevated aspartate aminotransferase (AST), elevated alanine aminotransferase (ALT), and thrombocytopenia as the most frequent AEs (Table 4). Grade ≥3 AEs occurred in 13 patients (7 in the HBV–HCC group and 6 in the NBNC–HCC group), and 1 patient in the HBV–HCC group discontinued camrelizumab treatment because of severe AEs (Grade 3 myocarditis). No significant differences in the total incidence of AEs and incidence of grade ≥3 AEs were found between the HBV–HCC group (75.0%, 25.0%) and the NBNC–HCC group (69.2%, 23.1%) (p = 0.224, p = 0.869).

Discussion

To the best of our knowledge, this research represents the first retrospective study in China that compares the tumor response, survival benefit, and tolerability of camrelizumab treatment for HBV–HCC and NBNC–HCC patients. The results demonstrated that relative to the NBNC–HCC group, the HBV–HCC group showed significant improvement in ORR and DCR. PFS was significantly

TABLE 1 Baseline characteristics of the HBV–HCC and NBNC–HCC groups.

Variable	HBV–HCC (n = 28)	NBNC–HCC (n = 26)	t/χ^2	p
Age (years)	55.9 ± 11.5	59.5 ± 10.9	−1.192	0.239
Gender			1.327	0.249
Male	24 (85.7)	19 (73.1)		
Female	4 (14.3)	7 (26.9)		
Tumor size			0.059	0.808
<5 cm	12 (42.9)	12 (30.8)		
≥5 cm	16 (57.1)	14 (53.8)		
Cirrhosis			0.017	0.897
Yes	23(82.1)	21(80.8)		
No	5(17.9)	5(19.2)		
Vascular invasion			0.247	0.619
Yes	10 (35.7)	11 (42.3)		
No	18 (64.3)	15 (57.7)		
Extrahepatic metastasis			0.051	0.821
Yes	11 (39.3)	11 (42.3)		
No	17 (60.7)	15 (57.7)		
ECGO PS			0.044	0.835
0-1	18 (64.3)	16 (61.5)		
2	10 (35.7)	10 (38.5)		
Child-Pugh class			0.027	0.869
A	21 (75.0)	20 (76.9)		
B	7 (25.0)	6 (23.1)		
BCLC stage, n (%)			1.169	0.280
B	6 (21.4)	9 (34.6)		
C	22 (78.6)	17 (65.4)		
Previous treatment			0.974	0.968
First-line	9 (32.1)	8 (30.8)		
Second-line	13 (46.4)	10 (38.5)		
Third-line or more	6 (21.4)	8 (30.8)		
Laboratory parameters				
AFP			0.350	0.554
<400 ng/ml	15 (53.6)	16 (61.5)		
≥400 ng/ml	13 (46.4)	10 (38.5)		
AST (U/L)	42.1 ± 13.9	36.9 ± 16.1	1.268	0.210
ALT (U/L)	35.1 ± 20.3	29.0 ± 14.5	1.264	0.212
ALB (g/L)	39.8 ± 5.0	40.9 ± 5.3	−7.490	0.457
TBIL (μmol/L)	18.1 ± 9.0	17.1 ± 8.2	0.460	0.648
PT (s)	12.62 ± 1.45	11.97 ± 1.01	1.907	0.062

Data in brackets represent percentages. HBV–HCC, hepatitis B virus-related hepatocellular carcinoma; NBNC–HCC, non-HBV; non-HCV, hepatocellular carcinoma; ECOG PS, Eastern Cooperative Oncology Group Performance Status; AFP, a-fetoprotein; AST, aspartate aminotransferase; ALT, alanine aminotransferase; TBIL, total bilirubin; PT, prothrombin time.

longer in the HBV–HCC group than in the NBNC–HCC group, and HBV infection was a significant independent predictor of better PFS. In addition, tolerability was similar between the two groups, and no adverse reaction-related deaths occurred.

In 2007, the molecularly targeted drug sorafenib was approved as a first-line treatment for advanced HCC; however, the median survival time was extended by only about 3 months, compared with the placebo (Boland and Wu, 2018). With the understanding of the tumor immune microenvironment and immune targets in recent years, immunotherapy has developed rapidly and achieved ideal efficacy in HCC treatment. Both National Comprehensive Cancer Network (NCCN) Guidelines for Hepatobiliary Cancers and Guidelines for Diagnosis and Treatment of Primary Liver Cancer recommend PD-1/L1 inhibitors as

TABLE 2 Tumor response between the HBV–HCC and NBNC–HCC groups.

Variable	CR	PR	SD	PD	ORR	DCR
HBV–HCC (n = 28)	1 (3.6)	7 (25.0)	12 (42.9)	8 (28.6)	8 (28.6)	20 (71.4)
NBNC–HCC (n = 26)	0 (0.0)	2 (7.7)	9 (34.6)	15 (57.7)	2 (7.7)	11 (42.3)
χ^2					3.895	4.676
P					0.048	0.031

Data in brackets represent percentages. HBV–HCC, hepatitis B virus-related hepatocellular carcinoma; NBNC–HCC, non-HBV; non-HCV, hepatocellular carcinoma; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; ORR, objective response rate; DCR, disease control rate.

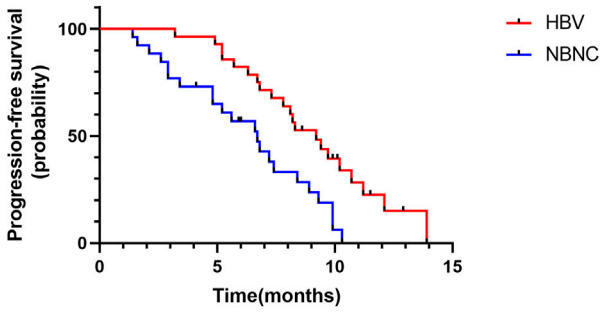


FIGURE 2
Kaplan–Meier plot of progression-free survival in the HBV–HCC and NBNC–HCC groups.

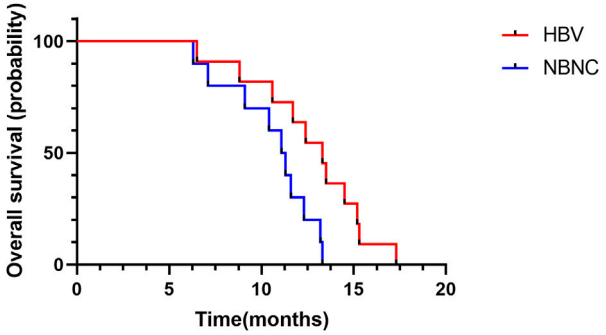


FIGURE 3
Kaplan–Meier plot of overall survival in the HBV–HCC and NBNC–HCC groups.

first-line drugs for systemic treatment of HCC (Du et al., 2022; Llovet et al., 2022). Camrelizumab, a PD-1 inhibitor, inhibits the immune escape of liver tumor cells by blocking the binding of PD-1 on CD8 T cells to programmed death ligand-1 (PD-L1) on Kupffer cells (Shen et al., 2019). However, PD-1/L1 inhibitors have exhibited only a 15%–23% response rate in HCC patients (Pinter et al., 2021). HBV infection is the most important cause of HCC; however, baseline HBV loads may not affect the clinical prognosis of PD-1 inhibitor therapy in HCC. Sun et al. (2020) found no significant correlation between clinical outcomes and HBV loads in HCC patients receiving the PD-1 inhibitor. In addition, Yuan et al. (2021) found that HBV loads did not affect the short-term efficacy of the PD-1 inhibitor combined with antiangiogenic therapy for HCC patients.

NAFLD refers to a clinical syndrome characterized by excessive deposition of fat in liver cells, apart from alcohol and other definite liver damage factors (Eslam et al., 2020). Among NAFLD, NASH is considered the most serious; up to 30% of NASH patients may develop liver inflammation and fibrosis, and some patients may further develop to HCC (Stine et al., 2018). NASH activates innate and adaptive immune cells, as well as increases metabolites and endoplasmic reticulum stress, causing hepatic necroinflammation and regenerative cycles, which may lead to HCC (Shen et al., 2017; Xu et al., 2017). Compared with HCC related to viral hepatitis, NBNC–HCC (NAFLD and/or NASH–HCC) is more insidious, and its lesions are mostly single and larger, and the prognosis of patients is worse (Li and Wang, 2022). In addition, several studies have reported that HBV–HCC and NBNC–HCC have significant differences in histopathological characteristics and prognosis (Salomao et al., 2012; Xue et al., 2020). The pathogenesis of NBNC–HCC remains unclear and is currently related to factors such as genetics, metabolism, oxidative stress, immunity, and intestinal flora imbalance (Li and Wang, 2022).

Whether viral infection affects the efficacy of immunotherapy in HCC patients remains inconclusive. In the study by Ho et al. (2020), no noticeable differences in ORR existed between virally infected and uninfected HCC patients receiving PD-1/PD-L1 inhibitors, and viral status could not be a criterion for PD-1/PD-L1 inhibitors. However, recent studies suggest that NBNC–HCC, particularly NASH–HCC, may weakly respond to immunotherapy. Kelley et al. (2022) reported that the benefit of cabozantinib plus atezolizumab treatment to PFS was mainly observed in the HBV–HCC group. Pfister et al. (2021) also found that in a mouse model of NASH, CD8/PD-1 double-positive T cells were unconventionally activated and gradually accumulated in NASH-infected livers (Dudek et al., 2021). Treatment with

TABLE 3 Univariate and multivariate analyses of the prognostic factors for progression-free survival.

Factors	Univariate			Multivariate		
	HR	95% CI	*P*	HR	95% CI	*P*
Age (≥60 vs. < 60), year	1.282	0.693–2.370	0.429	—	—	—
Gender (male vs. female)	0.606	0.275–1.335	0.214	—	—	—
Etiology (HBV vs. NBNC)	0.386	0.202–0.735	0.004	0.192	0.088–0.418	0.001
Tumor size (≥5 vs. < 5), cm	1.304	0.703–2.422	0.400	—	—	—
AFP level (≥400 vs. < 400), ng/ml	2.418	1.236–4.733	0.010	2.893	1.233–6.787	0.015
ECOG PS (0–1 vs. 2)	1.121	0.594–2.115	0.724	—	—	—
Child–Pugh class (A vs. B)	1.183	0.574–2.439	0.648			
BCLC stage (B vs. C)	1.057	0.540–2.070	0.870	—	—	—
Vascular invasion (yes vs. no)	2.889	1.528–5.461	0.001	3.158	1.436–6.942	0.004
Extrahepatic metastasis (yes vs. no)	0.703	0.372–1.326	0.276			
Cirrhosis (yes vs. no)	0.562	0.255–1.241	0.154			

HBV, hepatitis B virus; NBNC, non-HBV, non-HCV; AFP, a-fetoprotein; ECOG PS, eastern cooperative oncology group performance status.

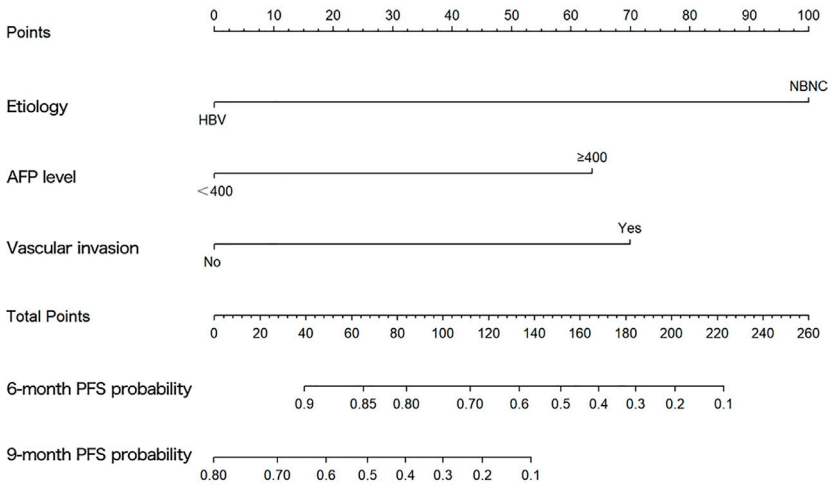


FIGURE 4
Graph showing the prognostic model for predicting 6- and 9-months progression-free survival.

PD-1 inhibitors increased the number of activated CD8/PD-1 double-positive T cells in NASH–HCC mice but failed to shrink tumors. These results indicate that activated abnormal T cells may not play a role in immune surveillance and cannot kill tumor cells. Thus, Pfister conducted a meta-analysis of three large randomized controlled phase III trials of immunotherapies in patients with advanced HCC (CheckMate-45911, IMbrave1505, and KEYNOTE-24010). The results showed that although immunotherapy improved survival in HBV–HCC patients (HR = 0.64; 95% CI = 0.49–0.83), survival in NBNC–HCC patients was not significantly improved (HR = 0.92; 95%CI = 0.77–1.11) (Pfister et al., 2021). In the current study, similar results were obtained: the median PFS of the HBV–HCC group was significantly longer than that of the NBNC–HCC group (9.2 vs. 6.7 months, $p = 0.003$).

In this study, the incidence of AEs was similar in the two groups—that is, mainly low-grade AEs. One patient in the HBV–HCC group needed to discontinue camrelizumab treatment because of Grade 3 myocarditis; the patient received symptomatic treatment such as hormones and nutritional myocardium. For camrelizumab, RCCEP was the most

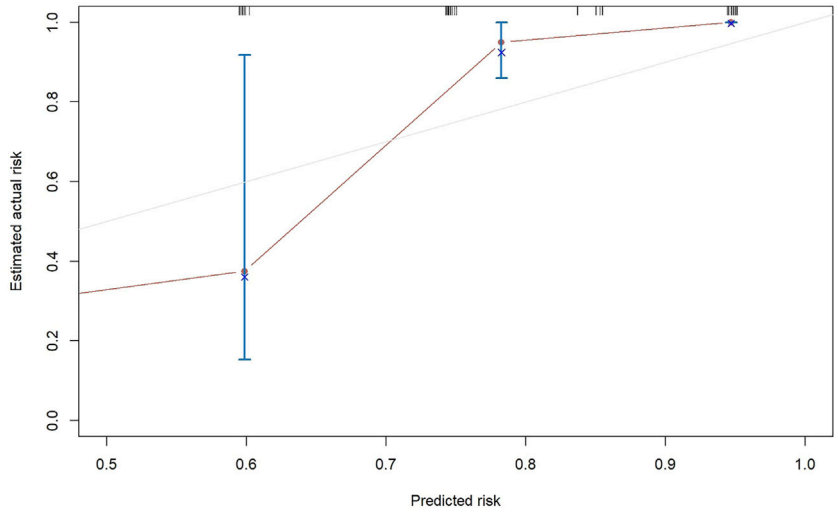


FIGURE 5
Calibration plots for 6-months progression-free survival.

TABLE 4 Treatment-related adverse events in patients with hepatocellular carcinoma.

Effect	All grades		Grade ≥3	
	HBV–HCC (*n* = 28)	NBNC–HCC (*n* = 26)	HBV–HCC (*n* = 28)	NBNC–HCC (*n* = 26)
RCCEP	13 (46.4)	11 (42.3)	0 (0.0)	0 (0.0)
Proteinuria	6 (21.4)	6 (23.1)	0 (0.0)	1 (3.8)
Thrombocytopenia	5 (17.8)	4 (15.4)	0 (0.0)	1 (3.8)
Neutropenia	3 (10.7)	2 (7.7)	0 (0.0)	0 (0.0)
Leukopenia	3 (10.7)	2 (7.7)	0 (0.0)	0 (0.0)
Hypothyroidism	3 (10.7)	4 (15.4)	0 (0.0)	1 (3.8)
Hypertension	4 (14.3)	5 (19.2)	0 (0.0)	1 (3.8)
Rash	2 (7.1)	3 (11.5)	1 (3.6)	0 (0.0)
Nausea	4 (14.3)	3 (11.5)	0 (0.0)	0 (0.0)
Diarrhea	5 (17.9)	4 (15.4)	1 (3.6)	0 (0.0)
Fatigue	4 (14.3)	3 (11.5)	0 (0.0)	0 (0.0)
Myocarditis	1 (7.1)	0 (3.8)	1 (3.6)	0 (0.0)
Hyperbilirubinemia	5 (17.9)	6 (23.1)	1 (3.6)	0 (0.0)
Elevated ALT	7 (25.0)	4 (15.4)	2 (7.1)	1 (3.8)
Elevated AST	6 (21.4)	6 (23.1)	1 (3.6)	1 (3.8)

Data in brackets represent the percentages of patients. HBV–HCC, hepatitis B virus-related hepatocellular carcinoma; NBNC–HCC, non-HBV; non-HCV, hepatocellular carcinoma; RCCEP, reactive cutaneous capillary endothelial proliferation; AST, aspartate aminotransferase; ALT, alanine aminotransferase.

common AE. The incidence rates of Grade 1–2 RCCEP in the two groups were 46.4% and 42.3%, respectively. No grade ≥3 RCCEP occurred, which was basically consistent with other studies (Qin et al., 2020). The AEs of the two groups were generally controllable, and no fatal AEs occurred.

The results of multivariable Cox proportional hazard regression analysis showed that etiology, AFP level, and vascular invasion were independent prognostic factors for PFS. We then constructed a nomogram prediction model by using these independent risk factors. The nomogram showed that etiology exerted the most effect on PFS, whereas the AFP level exerted the least effect on PFS. To validate the nomogram prediction model, we calculated the C-index and plotted the calibration curve. The C-index of the nomogram model in this

study was 0.785, and the calibration curve indicated that our model was consistent with the actual observations, proving the reliability and precision of our model. However, owing to the limited number of cases, no external validation was performed.

This study had several limitations. First, the patients in this study were only treated with camrelizumab, and our conclusions for other PD-1/PD-L1 inhibitors require verification. Second, a stratified analysis of NBNC–HCC etiologies could not be performed because of the small sample size. Third, the median OS for the majority of the population was not determined because of the relatively short follow-up period. We intend to extend the follow-up time in the future. Fourth, only 54 patients were included in the study; the sample size needs to be expanded in the future. Fifth, this study was a single-center retrospective study, and all patients included in this study were from China, where HBV is the main cause of HCC. Finally, some patients had received transarterial chemoembolization and targeted therapy, among others, which might have affected the efficacy of camrelizumab.

Conclusion

Compared with HBV–HCC patients, NBNC–HCC patients may exhibit a weaker response to PD-1 inhibitors, suggesting that clinicians can screen HCC subgroups suitable for immunotherapy on the basis of etiology to obtain improved efficacy. Large randomized controlled trials are needed in the future to prove our conclusions.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Clinical Trials, Affiliated Hospital of Xuzhou Medical University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ZH and HL designed the study. HL, ZX, MW, and TL collected the clinical data. HL, NY, and SZ analyzed the data. HL wrote the paper. ZH and SL revised the paper. 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.

References

Anstee, Q. M., Reeves, H. L., Kotsiliti, E., Govaere, O., and Heikenwalder, M. (2019). From NASH to HCC: Current concepts and future challenges. *Nat. Rev. Gastroenterol. Hepatol.* 16 (7), 411–428. doi:10.1038/s41575-019-0145-7

Boland, P., and Wu, J. J. C. C. O. (2018). Systemic therapy for hepatocellular carcinoma: Beyond sorafenib. *Chin. Clin. Oncol.* 7 (5), 50. doi:10.21037/cco.2018.10.10

Cabibbo, G., Enea, M., Attanasio, M., Bruix, J., Craxi, A., and Camma, C. (2010). A meta-analysis of survival rates of untreated patients in randomized clinical trials of hepatocellular carcinoma. *Hepatology* 51 (4), 1274–1283. doi:10.1002/hep.23485

Chen, J., Hu, X., Li, Q., Dai, W., Cheng, X., Huang, W., et al. (2020). Effectiveness and safety of toripalimab, camrelizumab, and sintilimab in a real-world cohort of Hepatitis B virus associated hepatocellular carcinoma patients. *Ann. Transl. Med.* 8 (18), 1187. doi:10.21037/atm-20-6063

Cheng, A. L., Qin, S., Ikeda, M., Galle, P. R., Ducreux, M., Kim, T. Y., et al. (2022). Updated efficacy and safety data from IMbrave150: Atezolizumab plus bevacizumab vs. sorafenib for unresectable hepatocellular carcinoma. *J. Hepatol.* 76 (4), 862–873. doi:10.1016/j.jhep.2021.11.030

Du, Y., Liu, D., and Du, Y. (2022). Recent advances in hepatocellular carcinoma therapeutic strategies and imaging-guided treatment. *J. Drug Target.* 30 (3), 287–301. doi:10.1080/1061186X.2021.1999963

Dudek, M., Pfister, D., Donakonda, S., Filpe, P., Schneider, A., Laschinger, M., et al. (2021). Auto-aggressive CXCR6(+) CD8 T cells cause liver immune pathology in NASH. *Nature* 592 (7854), 444–449. doi:10.1038/s41586-021-03233-8

El-Khoueiry, A. B., Sangro, B., Yau, T., Crocenzi, T. S., Kudo, M., Hsu, C., et al. (2017). Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): An open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* 389 (10088), 2492–2502. doi:10.1016/S0140-6736(17)31046-2

Eslam, M., Newsome, P. N., Sarin, S. K., Anstee, Q. M., Targher, G., Romero-Gomez, M., et al. (2020). A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. *J. Hepatol.* 73 (1), 202–209. doi:10.1016/j.jhep.2020.03.039

Finn, R. S., Qin, S., Ikeda, M., Galle, P. R., Ducreux, M., Kim, T. Y., et al. (2020). Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N. Engl. J. Med.* 382 (20), 1894–1905. doi:10.1056/NEJMoa1915745

Ho, W. J., Danilova, L., Lim, S. J., Verma, R., Xavier, S., Leatherman, J. M., et al. (2020). Viral status, immune microenvironment and immunological response to checkpoint inhibitors in hepatocellular carcinoma. *J. Immunother. Cancer* 8 (1), e000394. doi:10.1136/jitc-2019-000394

Ikeda, M., Morizane, C., Ueno, M., Okusaka, T., Ishii, H., and Furuse, J. (2018). Chemotherapy for hepatocellular carcinoma: Current status and future perspectives. *Jpn. J. Clin. Oncol.* 48 (2), 103–114. doi:10.1093/jjco/hyx180

Ji, F., and Nguyen, M. H. (2022). Cabozantinib plus atezolizumab in advanced hepatocellular carcinoma and the role of adjuvant antiviral therapy. *Lancet. Oncol.* 23, 962–963. doi:10.1016/S1470-2045(22)00383-7

Kelley, R. K., Rimassa, L., Cheng, A-L., Kaseb, A., Qin, S., Zhu, A. X., et al. (2022). Cabozantinib plus atezolizumab versus sorafenib for advanced hepatocellular carcinoma (COSMIC-312): A multicentre, open-label, randomised, phase 3 trial. *Lancet. Oncol.* 23, 995–1008. doi:10.1016/S1470-2045(22)00326-6

Li, J. Z. H., and Wang, S. (2022). Research progress in the pathogenesis of hepatocellular carcinoma associated with non alcoholic fatty liver disease. *China J. Dig.* 42 (03), 206–209.

Liu, T., Li, Q., Lin, Z., Wang, P., Chen, Y., Fu, Y., et al. (2021). Viral infections and the efficacy of PD-(L)1 inhibitors in virus-related cancers: Head and neck squamous cell carcinoma and hepatocellular carcinoma. *Int. Immunopharmacol.* 100, 108128. doi:10.1016/j.intimp.2021.108128

Llovet, J. M., Castet, F., Heikenwalder, M., Maini, M. K., Mazzaferro, V., Pinato, D. J., et al. (2022). Immunotherapies for hepatocellular carcinoma. *Nat. Rev. Clin. Oncol.* 19 (3), 151–172. doi:10.1038/s41571-021-00573-2

Maluccio, M., and Covey, A. (2012). Recent progress in understanding, diagnosing, and treating hepatocellular carcinoma. *Ca. Cancer J. Clin.* 62 (6), 394–399. doi:10.3322/caac.21161

Pfister, D., Nunez, N. G., Pinyol, R., Govaere, O., Pinter, M., Szydlowska, M., et al. (2021). NASH limits anti-tumour surveillance in immunotherapy-treated HCC. *Nature* 592 (7854), 450–456. doi:10.1038/s41586-021-03362-0

Pinter, M., Jain, R. K., and Duda, D. G. (2021). The current landscape of immune checkpoint blockade in hepatocellular carcinoma: A review. *JAMA Oncol.* 7 (1), 113–123. doi:10.1001/jamaoncol.2020.3381

Qin, S., Ren, Z., Meng, Z., Chen, Z., Chai, X., Xiong, J., et al. (2020). Camrelizumab in patients with previously treated advanced hepatocellular carcinoma: A multicentre, open-label, parallel-group, randomised, phase 2 trial. *Lancet. Oncol.* 21 (4), 571–580. doi:10.1016/S1470-2045(20)30011-5

Roderburg, C., Wree, A., Demir, M., Schmelzle, M., and Tacke, F. J. H. O. (2020). The role of the innate immune system in the development and treatment of hepatocellular carcinoma. *Hepat. Oncol.* 7 (1), HEP17. doi:10.2217/hep-2019-0007

Salomao, M., Remotti, H., Vaughan, R., Siegel, A. B., Lefkowitch, J. H., and Moreira, R. K. (2012). The steatohepatitic variant of hepatocellular carcinoma and its association with underlying steatohepatitis. *Hum. Pathol.* 43 (5), 737–746. doi:10.1016/j.humpath.2011.07.005

Sangro, B., Sarobe, P., Hervas-Stubbs, S., and Melero, I. (2021). Advances in immunotherapy for hepatocellular carcinoma. *Nat. Rev. Gastroenterol. Hepatol.* 18 (8), 525–543. doi:10.1038/s41575-021-00438-0

Shen, H., Yang, E. S., Conry, M., Fiveash, J., Contreras, C., Bonner, J. A., et al. (2019). Predictive biomarkers for immune checkpoint blockade and opportunities for combination therapies. *Genes. Dis.* 6 (3), 232–246. doi:10.1016/j.gendis.2019.06.006

Shen, Z. Q., Chen, Y. F., Chen, J. R., Jou, Y. S., Wu, P. C., Kao, C. H., et al. (2017). CISD2 haploinsufficiency disrupts calcium homeostasis, causes nonalcoholic fatty liver disease, and promotes hepatocellular carcinoma. *Cell. Rep.* 21 (8), 2198–2211. doi:10.1016/j.celrep.2017.10.099

Stine, J. G., Wentworth, B. J., Zimmet, A., Rinella, M. E., Loomba, R., Caldwell, S. H., et al. (2018). Systematic review with meta-analysis: Risk of hepatocellular carcinoma in non-alcoholic steatohepatitis without cirrhosis compared to other liver diseases. *Aliment. Pharmacol. Ther.* 48 (7), 696–703. doi:10.1111/apt.14937

Sun, X., Hu, D., Yang, Z., Liu, Z., Wang, J., Chen, J., et al. (2020). Baseline HBV loads do not affect the prognosis of patients with hepatocellular carcinoma receiving anti-programmed cell death-1 immunotherapy. *J. Hepatocell. Carcinoma* 7, 337–345. doi:10.2147/JHC.S278527

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

Thomas, D. L. (2019). Global elimination of chronic hepatitis. *N. Engl. J. Med.* 380 (21), 2041–2050. doi:10.1056/NEJMra1810477

Wang, M., Wang, Y., Feng, X., Wang, R., Wang, Y., Zeng, H., et al. (2017). Contribution of Hepatitis B virus and hepatitis C virus to liver cancer in China north areas: Experience of the Chinese National Cancer Center. *Int. J. Infect. Dis.* 65, 15–21. doi:10.1016/j.ijid.2017.09.003

Xia, C., Dong, X., Li, H., Cao, M., Sun, D., He, S., et al. (2022). Cancer statistics in China and United States, 2022: Profiles, trends, and determinants. *Chin. Med. J.* 135 (5), 584–590. doi:10.1097/CM9.0000000000002108

Xu, W., Zhang, X., Wu, J. L., Fu, L., Liu, K., Liu, D., et al. (2017). O-GlcNAc transferase promotes fatty liver-associated liver cancer through inducing palmitic acid and activating endoplasmic reticulum stress. *J. Hepatol.* 67 (2), 310–320. doi:10.1016/j.jhep.2017.03.017

Xue, X., Liao, W., and Xing, Y. (2020). Comparison of clinical features and outcomes between HBV-related and non-B non-C hepatocellular carcinoma. *Infect. Agent. Cancer* 15, 11. doi:10.1186/s13027-020-0273-2

Yarchoan, M., Hopkins, A., and Jaffee, E. M. (2017). Tumor mutational burden and response rate to PD-1 inhibition. *N. Engl. J. Med.* 377 (25), 2500–2501. doi:10.1056/NEJMc1713444

Yoon, S. M., Ryoo, B. Y., Lee, S. J., Kim, J. H., Shin, J. H., An, J. H., et al. (2018). Efficacy and safety of transarterial chemoembolization plus external beam radiotherapy vs sorafenib in hepatocellular carcinoma with macroscopic vascular invasion: A randomized clinical trial. *JAMA Oncol.* 4 (5), 661–669. doi:10.1001/jamaoncol.2017.5847

Yuan, G., Li, R., Li, Q., Hu, X., Ruan, J., Fan, W., et al. (2021). Interaction between Hepatitis B virus infection and the efficacy of camrelizumab in combination with apatinib therapy in patients with hepatocellular carcinoma: A multicenter retrospective cohort study. *Ann. Transl. Med.* 9 (18), 1412. doi:10.21037/atm-21-3020

Zheng, R., Qu, C., Zhang, S., Zeng, H., Sun, K., Gu, X., et al. (2018). Liver cancer incidence and mortality in China: Temporal trends and projections to 2030. *Chin. J. Cancer Res.* 30 (6), 571–579. doi:10.21147/j.issn.1000-9604.2018.06.01

Zhu, A. X., Finn, R. S., Edeline, J., Cattan, S., Ogasawara, S., Palmer, D., et al. (2018). Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): A non-randomised, open-label phase 2 trial. *Lancet. Oncol.* 19 (7), 940–952. doi:10.1016/S1470-2045(18)30351-6

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 12 September 2022
DOI 10.3389/fgene.2022.1010440

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Xinghua Victor Pan,
Southern Medical University, China
Ming Zheng,
Nanjing Medical University, China
Zijian Zhou,
Fudan University, China

*CORRESPONDENCE
Sheng Ju,
jusheng_ppk@163.com
Lijuan Tang,
lijuan_tlj@163.com
Jun Zhao,
junzhao@suda.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to *Cancer Genetics and Oncogenomics*, a section of the journal Frontiers in Genetics

RECEIVED 03 August 2022
ACCEPTED 29 August 2022
PUBLISHED 12 September 2022

CITATION
Chen Z, Yang J, Li Y, Zeng W, Bai Y, Ding C, Xu C, Li C, Chen J, Ju S, Tang L and Zhao J (2022), Integration of single-cell and bulk RNA-seq to establish a predictive signature based on the differentiation trajectory of M2 macrophages in lung adenocarcinoma. *Front. Genet.* 13:1010440. doi: 10.3389/fgene.2022.1010440

Integration of single-cell and bulk RNA-seq to establish a predictive signature based on the differentiation trajectory of M2 macrophages in lung adenocarcinoma

Zhike Chen[1,2†], Jian Yang[1,2†], Yu Li[1,2†], Weibiao Zeng[1,2], Yiling Bai[3], Cheng Ding[1,2], Chun Xu[1,2], Chang Li[1,2], Jun Chen[1,2], Sheng Ju[1,2*], Lijuan Tang[4,5*] and Jun Zhao[1,2*]

[1]Department of Thoracic Surgery, The First Affiliated Hospital of Soochow University, Suzhou, China, [2]Institute of Thoracic Surgery, The First Affiliated Hospital of Soochow University, Suzhou, China, [3]Department of Oncology, The First Affiliated Hospital of Soochow University, Suzhou, China, [4]Dalian Medical University, Dalian, Liaoning, China, [5]Department of Pathology, Affiliated Hospital of Nantong University, Nantong, Jiangsu, China

Background: Tumor-associated macrophages as important members of the tumor microenvironment, are highly plastic and heterogeneous. TAMs can be classified into two preliminary subtypes: M1 and M2 macrophages. M2 macrophages are significantly associated with the progression of lung adenocarcinoma. However, no study has investigated the heterogeneity among M2 macrophages and their differentiation-related genes at the single-cell level to guide the clinical treatment of lung adenocarcinoma.

Methods: Using the available annotation information from the Tumor Immune Single-cell Hub database, we clustered and annotated 12 lung adenocarcinoma samples using the R package 'Seurat'. Subsequently, we extracted M2 macrophages for secondary clustering analysis and performed cell trajectory analysis using the R package 'monocle2'. Based on heterogeneous genes associated with the differentiation trajectory of M2 macrophages, we established a prognostic lung adenocarcinoma model using Lasso-Cox and multivariate stepwise regression. In addition, we also performed immunotherapy and chemotherapy predictions.

Results: M2 macrophages exhibit heterogeneity among themselves. M2 macrophages in different differentiation states showed significant differences in pathway activation and immune cell communication. Prognostic signature based on heterogeneous genes can be used to classify the prognostic status and abundance of immune cell infiltration in lung adenocarcinoma patients. In addition, the calculation of the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm and the validation of the GSE126044 database indicated that lung adenocarcinoma patients with high-risk scores had poorer treatment outcomes when receiving immune checkpoint inhibitors treatment.

Conclusion: Based on scRNA-seq and Bulk-seq data, we identified M2 macrophage-associated prognostic signature with a potential clinical utility to improve precision therapy.

KEYWORDS

lung adenocarcinoma, tumor microenvironment (TME), M2 macrophages, ScRNA-seq, immunotherapy

Introduction

Lung cancer is one of the most common malignancies and the leading cause of cancer-related deaths. There are approximately two million new cases and 1.76 million deaths each year (Thai et al., 2021). It can be divided into two types: non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) (Thai et al., 2021). NSCLC is the leading type of lung cancer, accounting for about 85% of the total lung cancers (Srivastava et al., 2022). Meanwhile, Lung adenocarcinoma is the most common histological subtype of NSCLC. Lung adenocarcinoma has a strong heterogeneity and a complex tumor microenvironment (TME) (He et al., 2021). Traditional pathological stages do not fully determine the prognosis of NSCLC patients. Therefore, the development of novel and reliable prognostic models can help stratify the risk of lung adenocarcinoma patients and provide targeted immunotherapy and chemotherapy strategies (Shi et al., 2021).

TME, an ecosystem with a complex communication network, consists of tumor cells, cancer-associated stromal and immune cells, and other non-cellular components (Wu and Dai, 2017; Maacha et al., 2019). Numerous studies have shown that the development, progression, and metastasis of lung adenocarcinoma are closely related to TME (Kamata et al., 2020; Kim et al., 2020; Li et al., 2021). Macrophages are monocyte-derived immune cells with many biological functions and are also essential members of the TME (Varol et al., 2015). Tumor-associated macrophages are functionally heterogeneous and can be classified into two subtypes: M1 macrophages and M2 macrophages (Yunna et al., 2020). M1 macrophages inhibit angiogenesis and tumor progression. A growing body of literature has reported that, unlike M1 macrophages, M2 macrophages significantly promote angiogenesis, metastasis, and tumor growth (Guo et al., 2019; Zhang et al., 2019; Pan et al., 2020). Furthermore, crosstalk between M2 macrophages and immune cells (or molecules) can also promote tumor escape. Therefore, M2 macrophages are key components in the development of the tumor immunosuppressive microenvironment (Zhou et al., 2020) and would be of great scientific value to investigate the effects of M2 macrophages on lung adenocarcinoma patients (Pan et al., 2020). Traditional transcriptome sequencing techniques lose information on heterogeneity between cells as all cells in a tumor sample are treated as a whole. Thus, single-cell sequencing is a good way to characterize heterogeneity between cells (Wu F. et al., 2021). Exploring the key genes that determine cell heterogeneity in the differentiation trajectory of M2 macrophages using single-cell sequencing might help determine the prognosis of lung adenocarcinoma patients and provide valuable guidance for clinical strategies.

In this study, we first uncovered genes associated with the heterogeneity of M2 macrophages based on single-cell sequencing data. Next, we performed a univariate Cox analysis of all heterogeneous genes and extracted prognostic genes which might be relevant to lung adenocarcinoma development and progression. Based on prognostic-related genes, we performed a Lasso-Cox and multivariate stepwise regression analysis and constructed a prognostic model for lung adenocarcinoma patients (Long et al., 2021). In this model, the risk score was an independent prognostic factor for lung adenocarcinoma patients and had a higher prognostic accuracy than clinical factors. After combining the clinical factors, we constructed a nomogram for a more accurate prognostic evaluation. Together, our results showed that heterogeneous genes associated with the differentiation of M2 macrophages uncovered from single-cell sequencing data could characterize the prognostic status of lung adenocarcinoma patients. The prognostic signature we established has clinical potential to predict the efficacy of immunotherapy (ICIs) and chemotherapy.

Materials and methods

Data collection

Twelve single-cell RNA sequencing samples from five lung adenocarcinoma patients in the GSE127465 database were included in this study. Bulk sequencing data, mutation data, and clinical information for lung adenocarcinoma patients were downloaded from The *Cancer* Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) database. Microarray sequencing data and clinical information of the GSE31210 database were downloaded from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) as an external independent validation set for the prognostic signature. Furthermore, the GSE126044 database was used as the immunotherapy response validation cohort (anti-PD-1 treatment). Detailed clinical information for TCGA and GSE31210 database is listed in Supplementary Table S1.

Processing and analysis of single-cell RNA sequencing (scRNA-seq) data

The available cell clustering and cell type annotation information of GSE127465 was used in the Tumor Immune Single-cell Hub (TISCH) database (http://tisch.comp-genomics.org/) (Sun et al., 2021), and single-cell analysis was performed using the R package 'Seurat' (Hao et al., 2021). Based on the annotation results, M2 macrophages were extracted for further analysis. In this study, the number of hypervariable genes was set to 2000, and the resolution for cell clustering to 0.6. Principal component analysis (PCA) was conducted based on 2000 hypervariable genes. In addition, dimensionality reduction of single-cell data was used by the t-distributed stochastic neighbor embedding (tSNE) method (Kobak and Berens, 2019), and the 'FindAllMarkers' algorithm was performed to search for characteristic differentially expressed genes among different cell clusters. R package 'monocle2' was used for differentiation trajectory and pseudotime analysis of M2 macrophages (Zhou et al., 2022). Subsequently, the 'BEAM' (branched expression analysis modeling) statistical algorithm was used to identify heterogeneous genes that play a key role in the differentiation of M2 macrophages (Wang et al., 2022). R package 'GSVA' and 'scMetabolism' determined the enrichment of signaling pathways at the single-cell level (Hanzelmann et al., 2013; Wu et al., 2022). Finally, cell-to-cell communication analysis was performed using the R package 'iTALK' (Wang Y. et al., 2019).

Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment assays were performed using the R package 'clusterprofiler' (Wu T. et al., 2021; Kanehisa et al., 2021). GO analyses include three parts: biological process (BP), cell composition (CC), and molecular function (MF). In addition, the R package 'limma' was used to identify differentially expressed genes in the prognostic signature between high- and low-risk groups (Ritchie et al., 2015). Gene Set Variation Analysis (GSVA), an unsupervised algorithm, was performed to calculate enrichment scores of hallmark gene sets (Molecular Signatures Database (MSigDB), http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp).

Construction of prognostic signature

Univariate Cox analysis of heterogeneous genes associated with M2 macrophage differentiation was performed using the R package 'survival' to screen for prognostic genes associated with lung adenocarcinoma among them. In addition, the TCGA datasets were randomly grouped on a 3: 2 scale by the "sample_frac" function in the R package "dplyr" to obtain the training and testing datasets. Based on these prognostic genes, a Lasso-Cox regression analysis was performed using the R package 'glmnet' (Friedman et al., 2010; Wang H. et al., 2019). Next, a multivariate Cox stepwise regression approach was performed to construct a prognostic model related to the differentiation trajectory of M2 macrophages. The formula for the signature was: risk score = [Coef (gene 1) x gene Exp (1)] + [Coef (gene 2) x gene Exp (2)] +. . . . . + [Coef (gene 1) x gene Exp (i)]. R packages 'survival' and 'survminer' were used for Kaplan-Meier prognostic analysis. The R package 'timeROC' was used to assess AUC values for time-dependent ROC curves. To further improve the prediction efficiency of the risk score, the R package 'rms' was used to combine the pathological stage with the risk score to construct a more accurate nomogram (Balachandran et al., 2015).

Immune infiltration analyses of prognostic signature

Estimate, EPIC, MCPcounter, TIMER, and ssGSEA algorithms were used to calculate immune infiltration abundance in lung adenocarcinoma patients with different risk scores. Among these five algorithms, the Estimate algorithm calculated the Estimate score, tumor purity, immune score, and stromal score (Yoshihara et al., 2013). The EPIC algorithm calculated the abundance of seven immune cell types (Racle et al., 2017). The MCPcounter algorithm calculated the abundance of 10 immune cell types (Becht et al., 2016). The TIMER algorithm calculated the abundance of six immune cell types (Li et al., 2020). Subsequently, the ssGSEA algorithm was utilized to calculate the enrichment score of 24 immune cell gene sets (Bindea et al., 2013). These algorithms revealed differences in immune cell infiltration abundance between high- and low-risk groups.

Mutation analysis, and prediction of immunotherapeutic and chemotherapy responses

Based on the TCGA mutation data (maf format), mutations in the high- and low-risk groups were analyzed using the R package 'Maftools' and mapped waterfall plots (Mayakonda et al., 2018). In addition, the TIDE algorithm was utilized to analyze the sensitivity of high- and low-risk groups to immune checkpoint inhibitors (ICIs) (Jiang et al., 2018). Based on the signature formula constructed above, the risk score of lung adenocarcinoma patients was calculated in GSE126044 to assess differences in immunotherapy efficacy. Subsequently, the chemotherapeutic drug sensitivity (IC50) of patients in the high- and low-risk groups was analyzed using the R package 'pRRophetic' (Geeleher et al., 2014). These studies helped provide personalized treatment strategies.

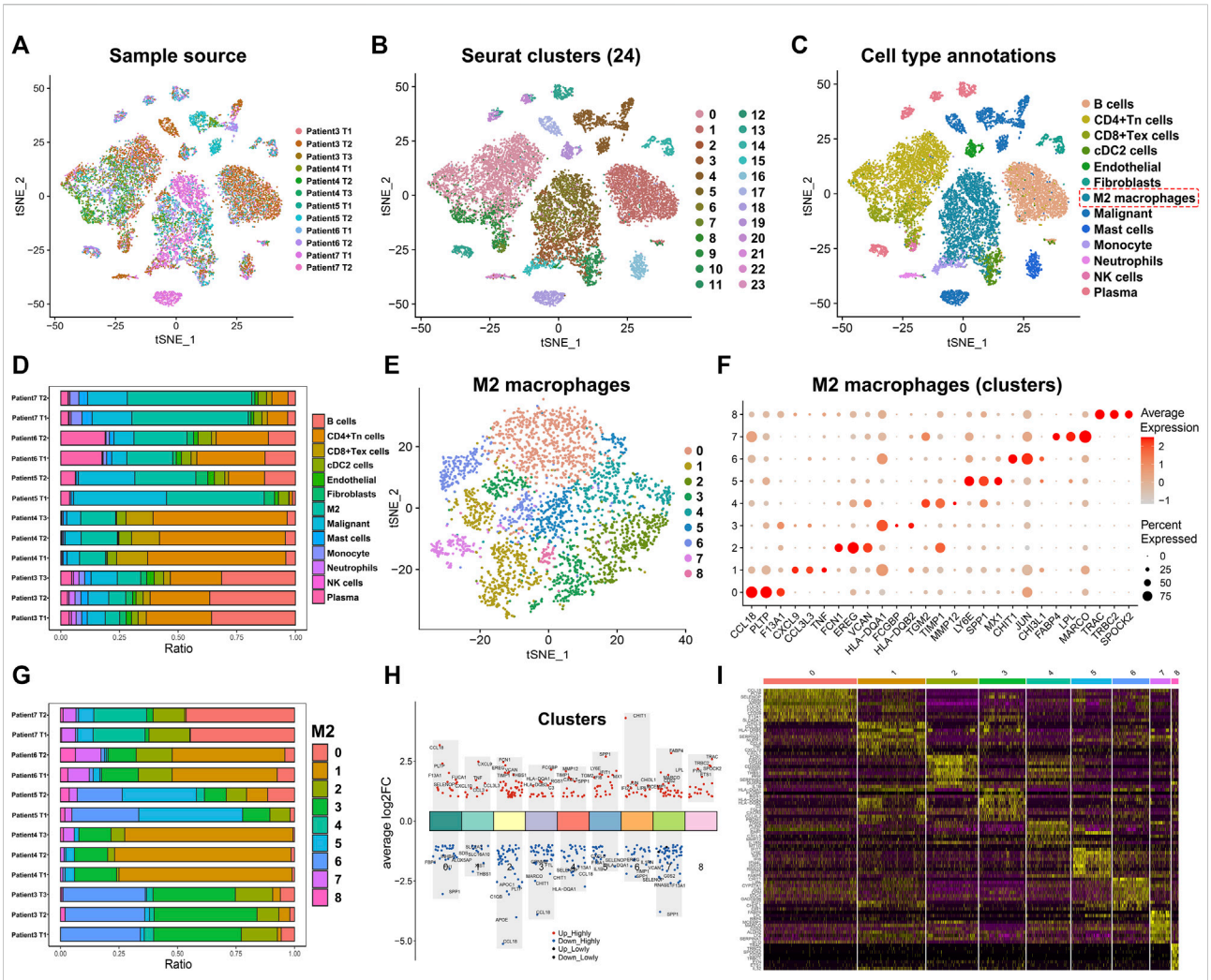


FIGURE 1
Annotation of single-cell data and extraction of M2 macrophage. **(A)** The tSNE map shows the distribution of 12 lung adenocarcinoma samples from the single-cell database GSE127465. **(B)** The tSNE plot shows that all the cells in the 12 lung adenocarcinoma samples can be classified into 24 clusters. **(C)** The tSNE map indicating that lung adenocarcinoma samples can be annotated as 13 cell types in the tumor microenvironment (different colors represent different cell types). **(D)** The histogram showing the proportion of 13 cell types in each of the 12 samples ("Patient7 T2" indicates the second tissue of patient7). **(E)** M2 macrophages classified into 9 clusters. **(F)** The bubble chart highlights the characteristic genes of different clusters **(G)** The bar chart shows the proportion of 9 clusters in each lung adenocarcinoma sample ("Patient7 T2" indicates the second tissue of patient7). **(H,I)** Biomarker genes in the 9 clusters of M2 macrophages.

Statistical analysis

Statistical analyses were performed using R software (v 4.1.3), and the results were visualized using the R packages. For non-normally distributed data, Wilcoxon rank-sum test, as a non-parametric test method, was used to examine the differences between the two groups of continuous variables, while for three and more groups we used the Kruskal–Wallis test for statistical testing. Using the Cox regression method, Kaplan-Meier prognostic analysis calculated the hazard ratio (HR). A two-sided $p < 0.05$ was considered statistically significant. Spearman method was applied for correlation analysis ($^{*}p < 0.05$, $^{**}p < 0.01$).

Results

scRNA-seq and cell annotation of lung adenocarcinoma samples

To better understand the heterogeneity of M2 macrophages in the TME of lung adenocarcinoma and its potential value for prognosis and drug treatment screening, we extracted and analyzed lung adenocarcinoma samples at the single-cell level. Based on the meta-information and cell type annotation information from the GSE127465 database on the TISCH website, we extracted 12 lung adenocarcinoma samples that

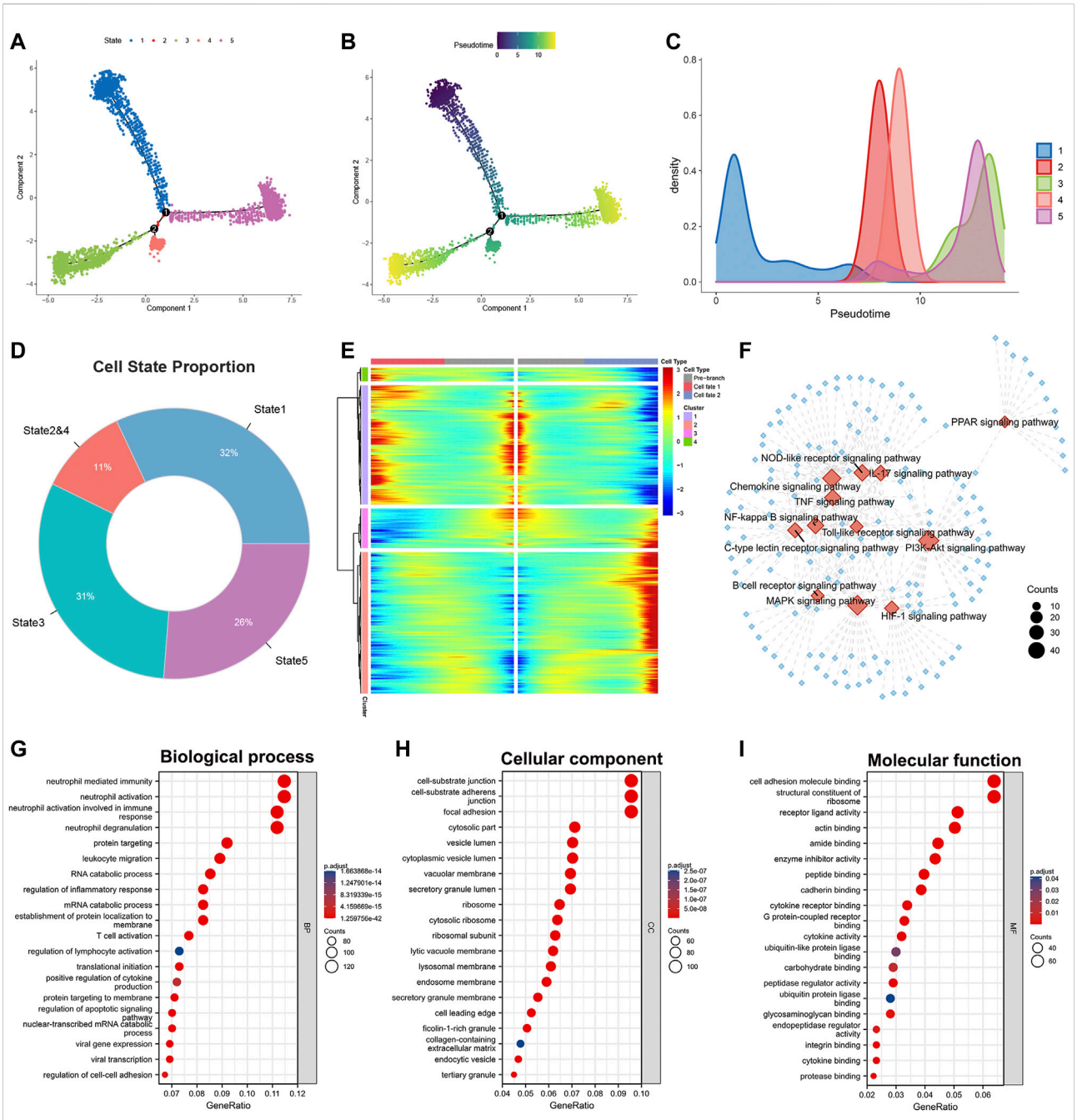


FIGURE 2
Identification of heterogeneous genes associated with differentiation trajectories of M2 macrophages. **(A)** Differentiation trajectory analysis of M2 macrophages. **(B,C)** Pseudotime analysis of M2 macrophages. **(D)** The proportion of each state in the differentiation trajectory of M2 macrophages. **(E)** The heat map revealed that M2 macrophages could exhibit four expression patterns after differentiation. **(F)** KEGG enrichment analysis of heterogeneous genes associated with differentiation of M2 macrophages. **(G–I)** GO enrichment analysis of heterogeneous genes associated with M2 macrophage differentiation.

had been quality-controlled and standardized. To overcome technical noise in scRNA-seq data, we performed Principal Component Analysis (PCA), and each principal component (PC) was considered a "meta-feature" (Supplementary Figure S1A). We identified the most suitable number of PCs (24 PCs) for downstream analysis by calculating the standard deviation of each principal component (Supplementary Figure S1B). In addition, we used tSNE, a nonlinear dimensionality reduction algorithm, to demonstrate the distribution of single-cell data from 12 lung adenocarcinoma samples (Figure 1A). We also

examined the cell distribution of lung adenocarcinoma patients of different ages and clinical stages (Supplementary Figure S1C, D). Subsequently, we used the R package 'Seurat' to classify the cells in 12 samples into 24 clusters (Figure 1B). These 24 clusters can also be categorized into 13 cell types: B cells, $CD4^+$ Tn cells, $CD8^+$ Tex cells, endothelial cells, fibroblasts, M2 macrophages, malignant cells, mast cells, monocyte cells, neutrophils cells, NK cells, and plasma cells (Figure 1C). We counted the frequency of these 13 types of immune cells and found a higher proportion of M2 macrophages in each of the 12 samples (Figure 1D). We then extracted the M2 macrophages and re-clustered them using the 'Seurat' package (Figure 1E). The result suggested that M2 macrophages can be classified into 9 clusters (0–8) based on different molecular markers (Figure 1F). M2 macrophages of lung adenocarcinoma patients showed considerable heterogeneity in the different clusters. Patient seven had a more significant proportion of cluster0 and cluster4 in the M2 macrophages (Figure 1G). Patient six and patient four had a high percentage of cluster1. Patient five had a larger ratio of cluster5 and cluster6. However, patient three had a greater portion of cluster3 and cluster6. Based on the above clustering results, we analyzed the signature genes of the 9 clusters of M2 macrophages using the 'FindAllMarkers' algorithm and visualized the results with scatter plots and heat maps (Figures 1H,I). We observed that the genes had distinct expression differences between different clusters.

Differentiation trajectory of M2 macrophages in tumor immune microenvironment

The heterogeneity among M2 macrophages was intriguing, and to further investigate the biological functions of essential genes in the differentiation of M2 macrophages, we performed differentiation trajectory analyses. We found that M2 macrophages can be divided into five differentiation states (Figure 2A). Meanwhile, we found that subpopulations of M2-type macrophages were differentially distributed on differentiation trajectory (Supplementary Figure S1E). Subsequently, we performed pseudotime analysis on M2 macrophages (Figures 2B,C). Purple indicated the initial state of cell differentiation, and yellow indicated the terminal state. Since state2 had a smaller number of cells and a high overlap with state4 in differentiation trajectory and pseudotime, we merged state2 with state4 as a whole. State1 accounted for 32%, state2&4 accounted for 11%, state3 accounted for 31%, and state5 accounted for 26% of all M2 macrophages (Figure 2D). In addition, we found that heterogeneous genes associated with the differentiation trajectory of M2 macrophages illustrated four expression patterns (Figure 2E).

We performed KEGG and GO analysis based on these statistically significant heterogeneous genes ($p < 0.0001$). KEGG enrichment analysis indicated that these heterogeneous genes were involved in activating numerous signaling pathways (Figure 2F). Examples include the Chemokine, IL-17, HIF-1, B cell receptor, and PI3K-Akt signaling pathway. The diverse activation levels of these pathways suggested that the different states of M2 macrophages might play distinct roles in the progression of lung adenocarcinoma. GO analysis phenotyped the heterogeneous genes in biological processes, cellular components, and molecular functions (Figures 2G–I, Supplementary Table S2). The results revealed that these genes activate multiple immune cells in the TME, suggesting crosstalk between M2 macrophages and immune cells.

Differential states of M2 macrophages reveal the heterogeneity of function characteristics and cellular communication levels

To further examine the functional differences between the different classes of M2 macrophages we distinguished, we performed a GSVA enrichment analysis. First, we verified the potential differences in molecular mechanisms among the 9 clusters of M2 macrophages. Although M2 macrophages exhibited pro-oncogenic activity, the different clusters of M2 macrophages showed significant differences in the activation levels of the 50 gene sets contained in Hallmark (Figure 3A). For instance, cluster7 and cluster8 have lower activation levels in numerous pathways than the other seven types of clusters. This suggested that cluster7 and cluster8 might be relatively weak in oncogenic activities. To better understand this heterogeneity within M2 macrophages, we also performed an enrichment analysis of the different states in the differentiation trajectory (Figure 3B). The results confirmed substantial heterogeneity among the different states. It was found that state1 had significantly higher enrichment levels in the epithelial-mesenchymal transition (EMT). In comparison, states2&4 had significantly higher activation levels in the reactive oxygen species (ROS) pathway, mitotic spindle, and interferon-gamma response. State3 was significantly activated on the apical junction and notch signaling and had the lowest activation on E2F targets, G2M checkpoint, and the genes upregulated by ultraviolet (UV) radiation. State5 was notably enriched in angiogenesis, hypoxia, and genes upregulated by KRAS signaling. Subsequently, we analyzed the differences in the metabolic activities of M2 macrophages in these states using the R package 'scMetabolism' (Figure 3C). The red color in the heat map corresponds to the higher activation level, and we can identify that the different states of M2 macrophages have distinct metabolic levels. This heterogeneity in metabolic levels may reveal differences in the functional levels of the M2 macrophages in different states. Finally, we analyzed the cellular communication between M2 macrophages and the other

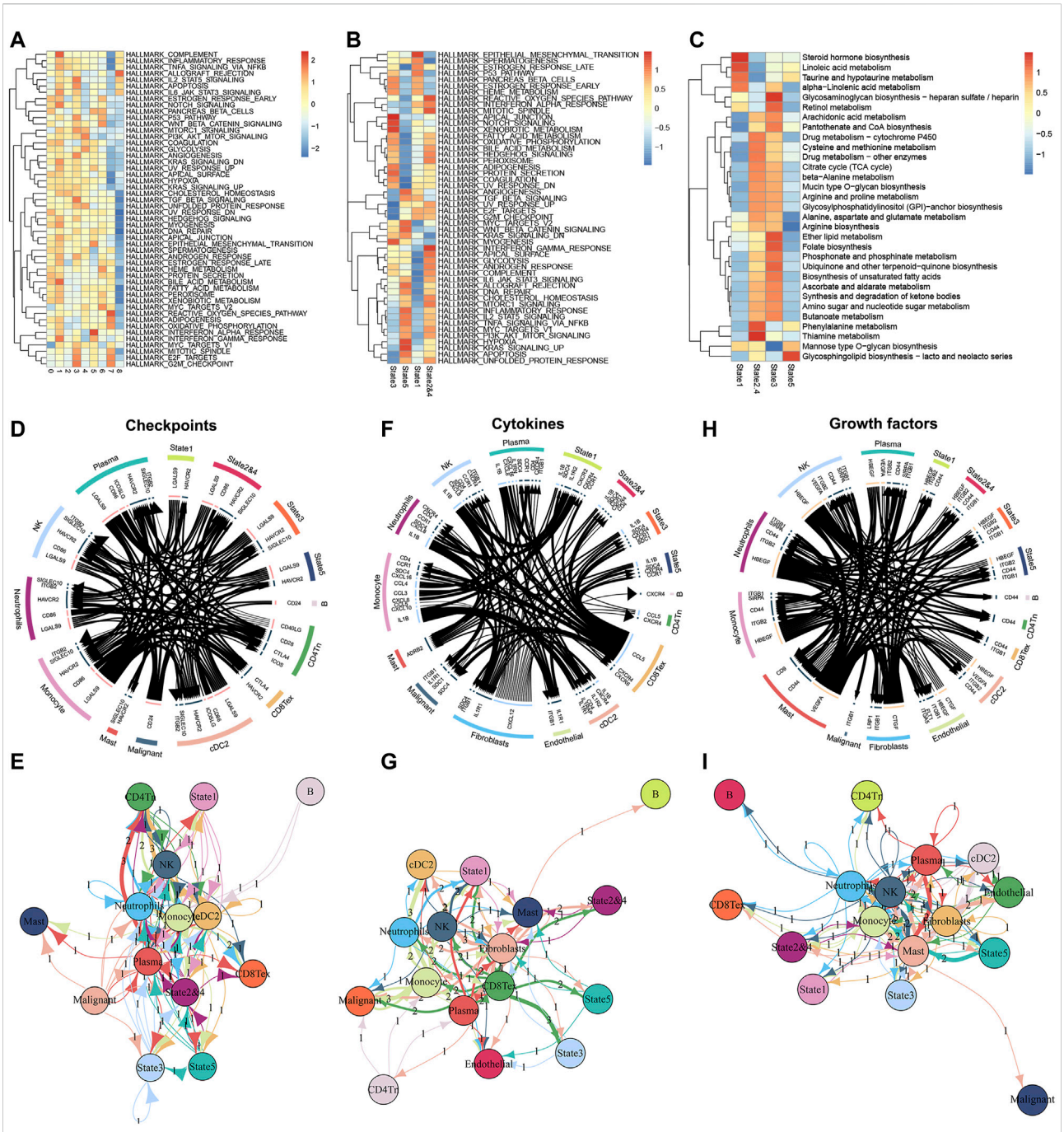


FIGURE 3
Enrichment analyses of M2 macrophages in different differentiation states and cell communication analyses in the TME. **(A)** The GSVA algorithm demonstrates different enrichment levels of 9 clusters (M2 macrophages) in the Hallmark pathway. **(B)** The GSVA analysis of differentiation states of M2 macrophages. **(C)** Enrichment scores of metabolism-related pathways in different differentiation states of M2 macrophages. **(D,E)** Communication between M2 macrophages of different differentiation states and various cell types in the tumor microenvironment at the level of checkpoints. **(F,G)** Communication between M2 macrophages of different differentiation states and various cell types in the tumor microenvironment at the level of cytokines. **(H,I)** Communication between M2 macrophages of different differentiation states and various cell types in the tumor microenvironment at the level of growth factors.

cells in the TME (immune checkpoints, cytokines, and growth factors). Regarding immune checkpoints (Figures 3D,E), we found that the M2 macrophages of state1 mainly communicated with NK cells, neutrophils, monocyte cells, and cDC2 cells. M2 macrophages in state2&4 communicated with malignant cells, $CD4^+$ Tn cells, and $CD8^+$ Tex cells in addition to

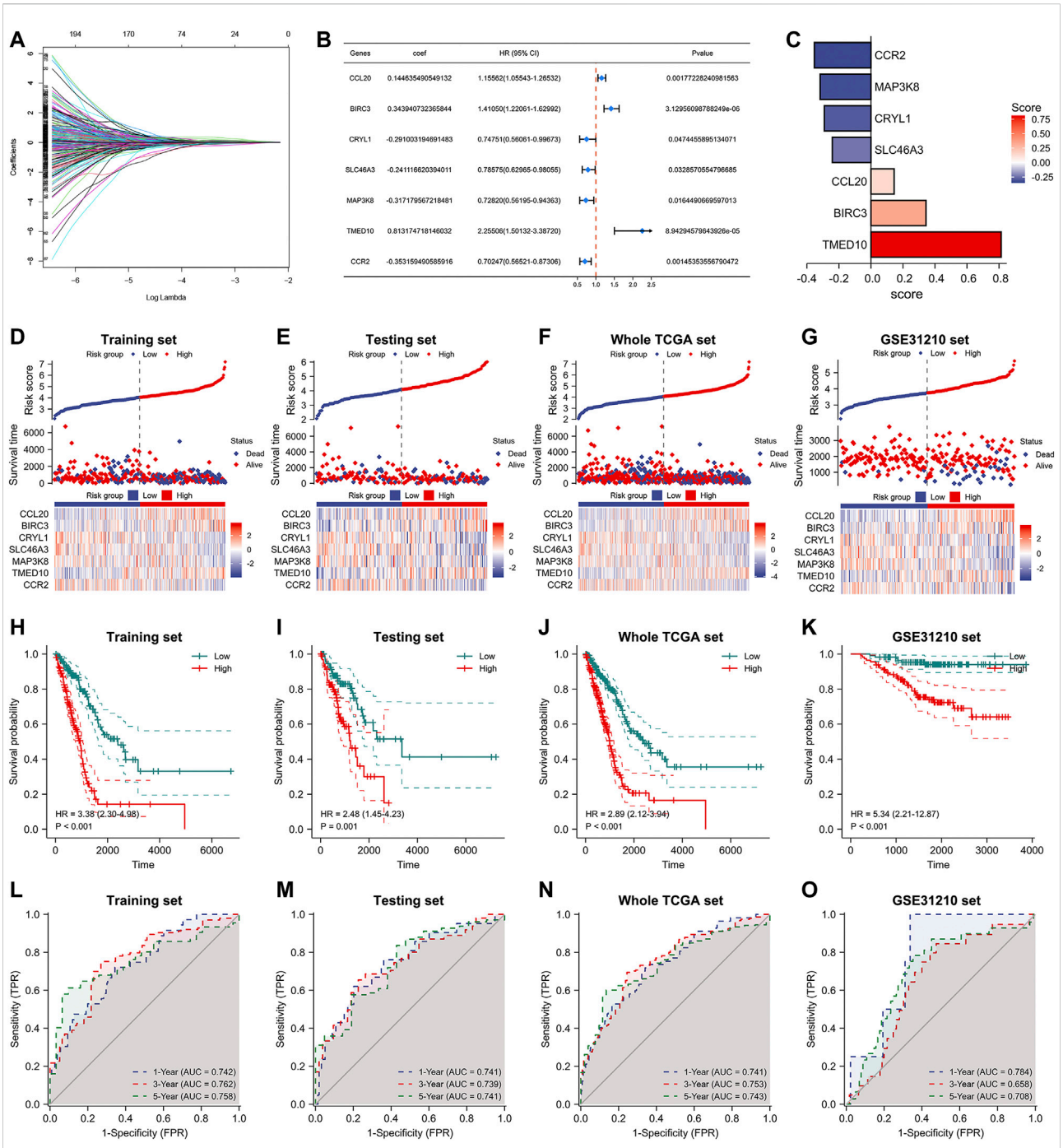


FIGURE 4
Construction of M2 macrophage differentiation-related prognostic signature. **(A)** Lasso Cox analysis of prognostic genes associated with M2 macrophage differentiation. **(B)** Multifactorial stepwise regression to construct a 7-gene prognostic model. **(C)** Coefficients of seven genes in the model formula. **(D–G)** Risk factor diagrams of signatures in the training, testing, whole TCGA, and GSE31210 dataset. **(H–K)** Kaplan-Meier prognostic analysis of signatures in the training, testing, whole TCGA, and GSE31210 dataset. **(L–O)** Time-dependent ROC curves of signatures in the training, testing, whole TCGA, and GSE31210 dataset.

the above 4 cells. M2 macrophages in state3 had extensive communication with plasma cells and malignant cells. In contrast, M2 macrophages in state5 communicated predominantly with monocyte cells. In terms of cytokines, the different states of M2 macrophages also have diverse levels of cellular communication (Figures 3F,G). However, in terms of cell

growth factors, there was no significant difference in the level of communication between these states of M2 macrophages and immune cells (Figures 3H,I). Together, the above results demonstrated heterogeneity in the level of cellular communication among different states of M2 macrophages.

Prognostic signature based on heterogeneous genes associated with differentiation of M2 macrophages can accurately predict lung adenocarcinoma patients' outcome

After extracting heterogeneous genes associated with the differentiation trajectory of M2 macrophages, we performed a univariate Cox analysis in the TCGA database and obtained 289 prognosis-related genes of lung adenocarcinoma. Utilizing Lasso-Cox with multivariate stepwise regression, we constructed a prognostic model for lung adenocarcinoma in the training set (Figures 4A,B). In addition, Figure 4C shows the coefficients of the seven genes incorporated into the formula. These seven genes were: CCL20, BIRC3, CRYL1, SLC46A3, MAP3K8, TMED10, and CCR2. The formula of the model was: risk score = [0.144635490549132 * CCL20 Exp] + [0.343940732365844 * BIRC3 Exp] + [-0.291003194691483 * CRYL1 Exp] + [-0.241116620394011 * SLC46A3 Exp] + [-0.317179567218481 * MAP3K8 Exp] + [0.813174718146032 * TMED10 Exp] + [-0.353159490585916 * CCR2 Exp]. Using this formula, we calculated the risk score values in the testing set, the entire TCGA database, and the external validation dataset GSE31210. Lung adenocarcinoma patients with high-risk scores had higher deaths in these four datasets (Figures 4D–G). The heat map results also indicated that the expression of the above seven genes had significant differences in the two risk groups. We then performed Kaplan-Meier prognostic analysis to explore the potential value of our constructed model for patients with lung adenocarcinoma. Patients with high-risk scores had a worse prognosis in the training set (HR = 3.38, $p < 0.001$), the testing set (HR = 2.48, $p = 0.001$), the entire TCGA set (HR = 2.89, $p < 0.001$), and the GSE31210 database (HR = 5.34, $p < 0.001$, Figures 4H–K). We also performed a time-dependent ROC curve analysis on these four databases to judge our model's accuracy in predicting prognosis. The AUC values of our model in the training set for 1-, 3-, and 5-years overall survivals were 0.742, 0.762, and 0.758, respectively (Figure 4L–O). In the testing set, the AUC values of our model for 1-, 3-, and 5-years survival were 0.741, 0.739, and 0.741, respectively. In the entire TCGA dataset, the AUC values of our model for 1-, 3-, and 5-years survival were 0.741, 0.753, and 0.741, respectively. While, in the GSE31210 dataset, the AUC values of our model for 1-, 3-, and 5-years survival were 0.784, 0.658, and 0.708, respectively.

Furthermore, we analyzed the relationship between numerous clinical factors and risk scores (Figures 5A–D). It was found that there was no statistical difference in the risk score between the two groups of patients aged≥65 years and those aged<65 years. The difference between the two risk groups of patients with different gender and smoking history was also not statistically different. However, the differences among patients with different pathologic stages were statistically significant. Patients with high pathological stages tended to have higher risk scores. In addition, we performed Kaplan-Meier prognostic analysis of lung adenocarcinoma patients with different clinical characteristics separately (Figure 5E–L). The results showed that patients with high-risk scores had a poor prognosis in all age groups (≥65, <65), all gender groups (Male, Female), and all pathological stage groups (Pathological stage IandII, Pathological stage IIIandIV). There was a significant prognostic difference between the high- and low-risk groups among the smoking group, with patients in the high-risk group having a poor prognosis (HR = 2.83, $p < 0.001$), however, there was no significant prognostic difference between the high- and low-risk groups in the non-smoking group. We also performed the univariate and multivariate Cox regression analyses regarding the risk scores. Our results showed that the risk score is an independent prognostic factor for lung adenocarcinoma and can be used as a clinical parameter to determine the prognosis of patients (Figure 5M, N).

Significant difference in molecular mechanisms and immune infiltration levels between high- and low-risk groups

As demonstrated in the above study, lung adenocarcinoma patients had significantly different prognoses between the high- and low-risk groups. Differential expression analysis was performed for the high-risk versus low-risk groups to investigate the mechanisms involved. Firstly, we used the R package 'limma' to identify differentially expressed genes (|FC|>1.5, FDR<0.05) and mapped the volcano (Figure 6A), Then, we performed enrichment analysis using GO and KEGG (Figure 6B, Supplementary Table S3). KEGG results showed significant enrichment in the cell cycle and IL-17 signaling pathway (Figure 6C). In addition, we performed GSVA enrichment analysis and plotted heat map and histogram for lung adenocarcinoma patients in high- and low-risk groups. Fifty gene sets from Hallmark were selected for GSVA analysis. By comparing the enrichment scores of the two groups in these 50 gene sets, 25 gene sets showed a statistically significant difference (Figures 6D,E). These results might explain the underlying mechanism for the

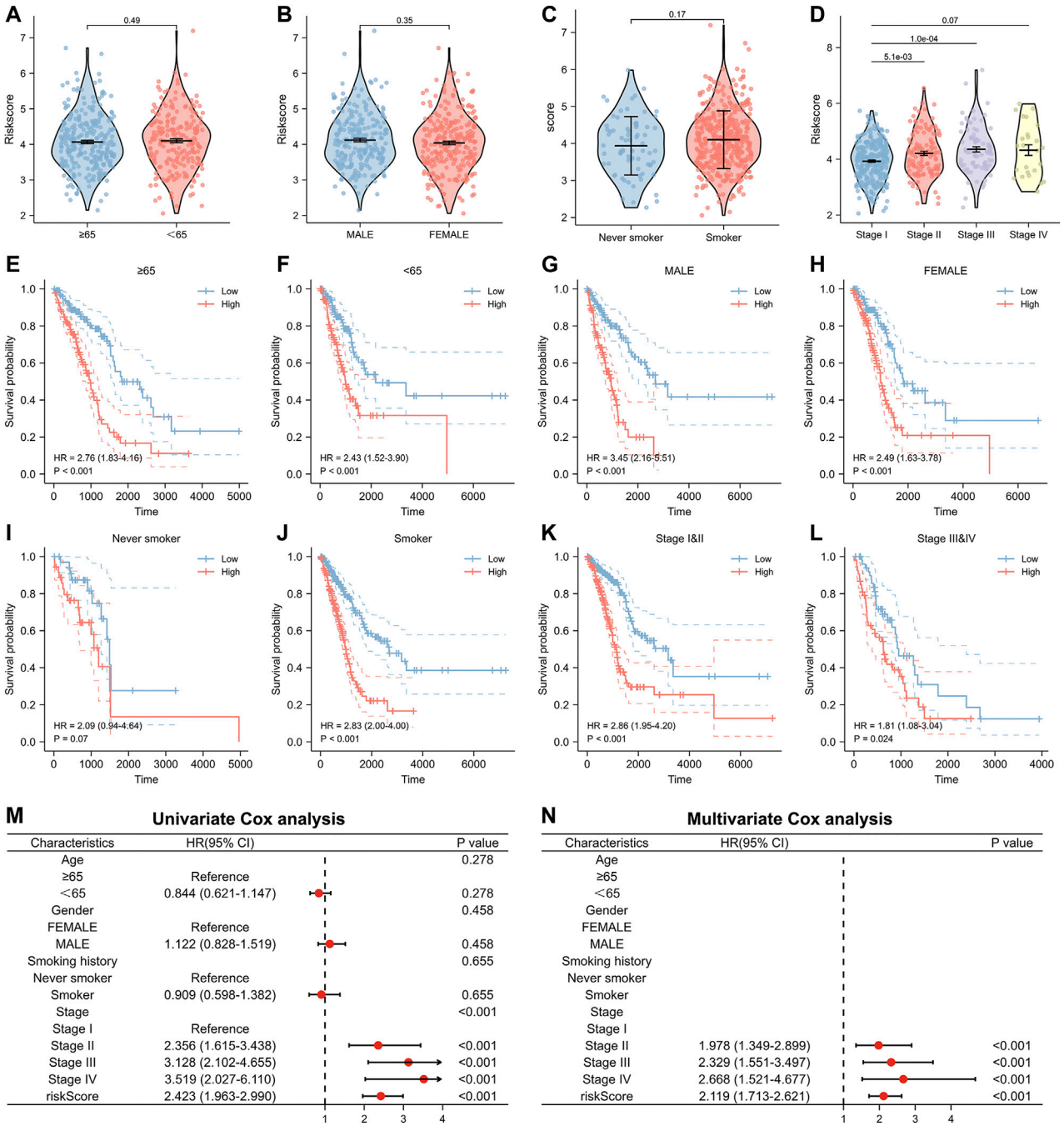


FIGURE 5
Relationship between risk score and clinical factors in the model. **(A–D)** The risk score of patients with different clinical factors. **(E,F)** Relationship between risk score and prognosis in different age groups. **(G,H)** Relationship between risk score and prognosis in different gender groups. **(I,J)** Relationship between risk score and prognosis in different smoking history groups. **(K,L)** Relationship between risk score and prognosis in different pathological stage groups. **(M)** Univariate and **(N)** multivariate Cox analysis of risk score and clinical factors.

difference in the prognosis of lung adenocarcinoma patients with different risk scores.

To understand the differences in TME among patients with different risk scores, we also analyzed the abundance of immune infiltration in high- and low-risk groups. Analysis using the Estimate algorithm showed that the high-risk group had higher tumor purity but lower immune and stromal scores (Figures 7A–D). This suggested that patients with

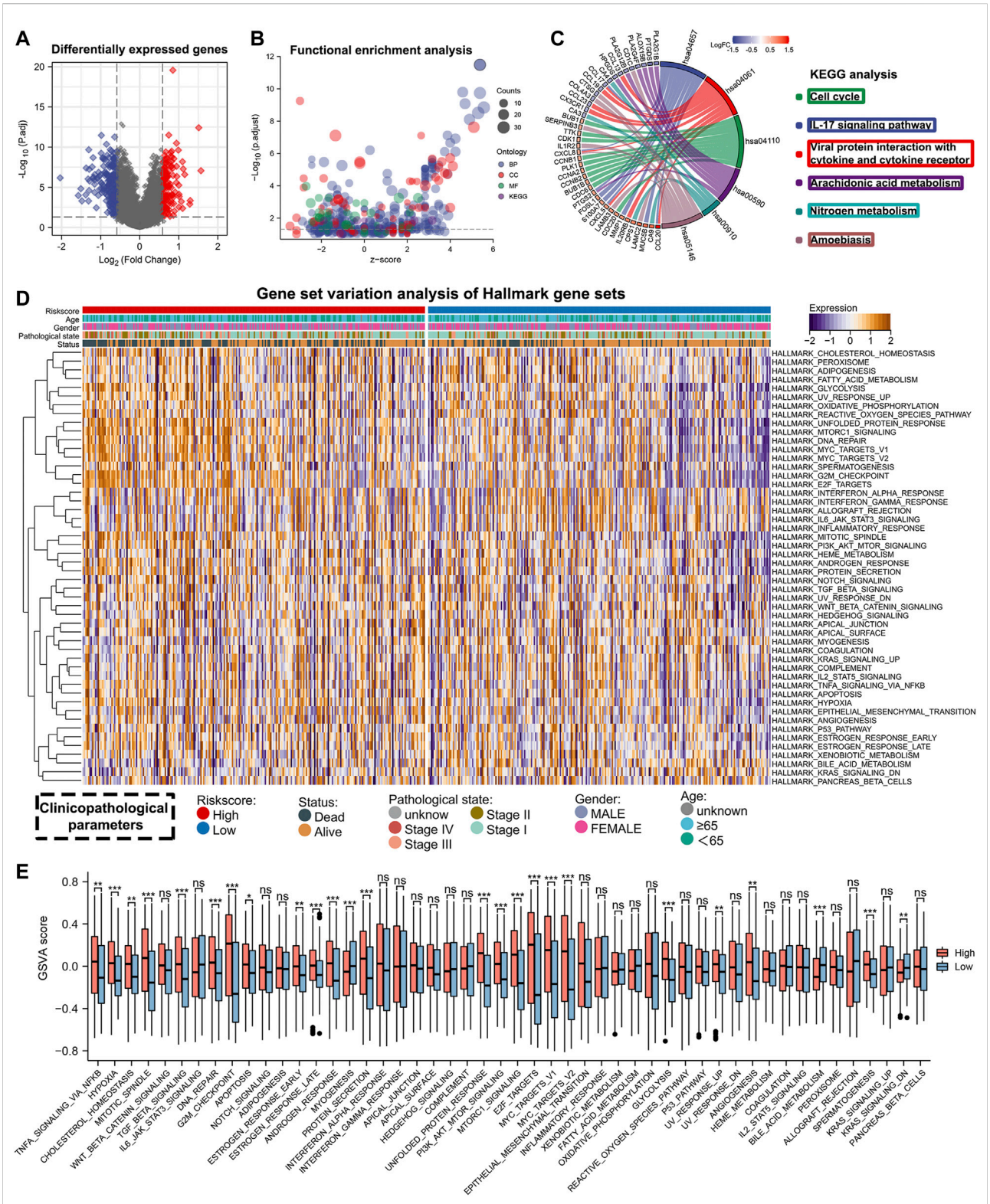


FIGURE 6
Enrichment analysis of differentially expressed genes between high- and low-risk groups. **(A)** Differentially expressed genes between high- and low-risk groups (FDR <0.05, |FC| > 1.5). **(B)** GO and KEGG enrichment analysis for differentially expressed genes. **(C)** Visualization of KEGG enrichment analysis results. **(D,E)** The heat map and bar chart shows the GSVA analysis results between the high- and low-risk groups.

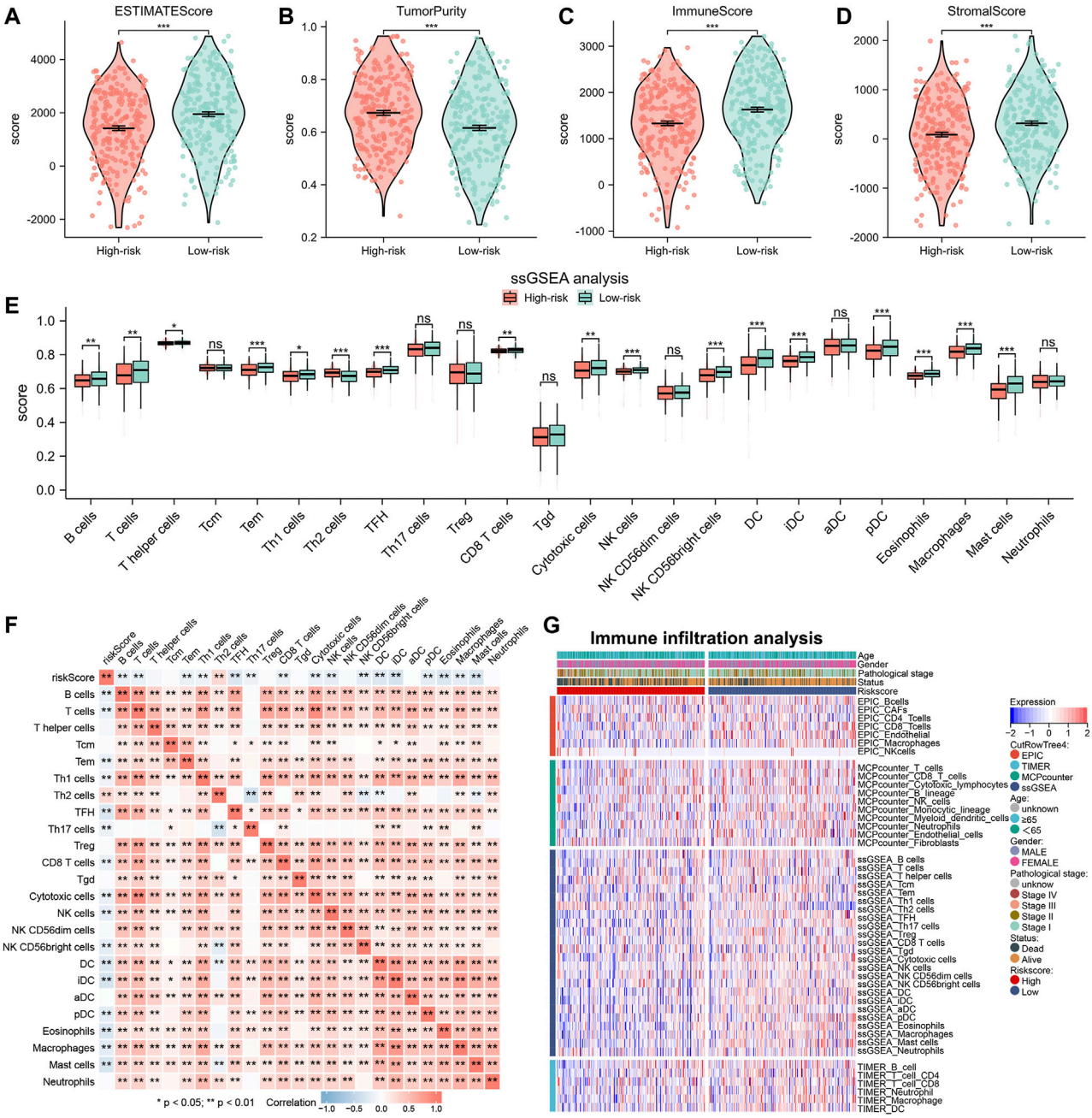


FIGURE 7
Immune infiltration analysis between high- and low-risk groups. (**A–D**) Immune infiltration abundance between high- and low-risk groups by Estimate algorithm. **(E)** The histogram displays differences in infiltrating abundance of 24 immune cell types in high- and low-risk groups analyzed by the ssGSEA algorithm. **(F)** Correlation between the risk score and 24 immune cell types. **(G)** The heat map showing the immune infiltration analysis results of EPIC, TIMER, MCPcounter, and ssGSEA.

high-risk scores exhibited a state that promoted tumor escape due to the lack of anti-tumor immune cells in the TME. Then, to investigate the immune status of the two risk groups in more detail, we used the ssGSEA algorithm to calculate the infiltration abundance of 24 immune cell types (Figure 7E). The results revealed that the high-risk group generally had a lower infiltration abundance of immune cells, including B cells, T cells, $CD8^+$ T cells, NK cells, DC cells, and mast

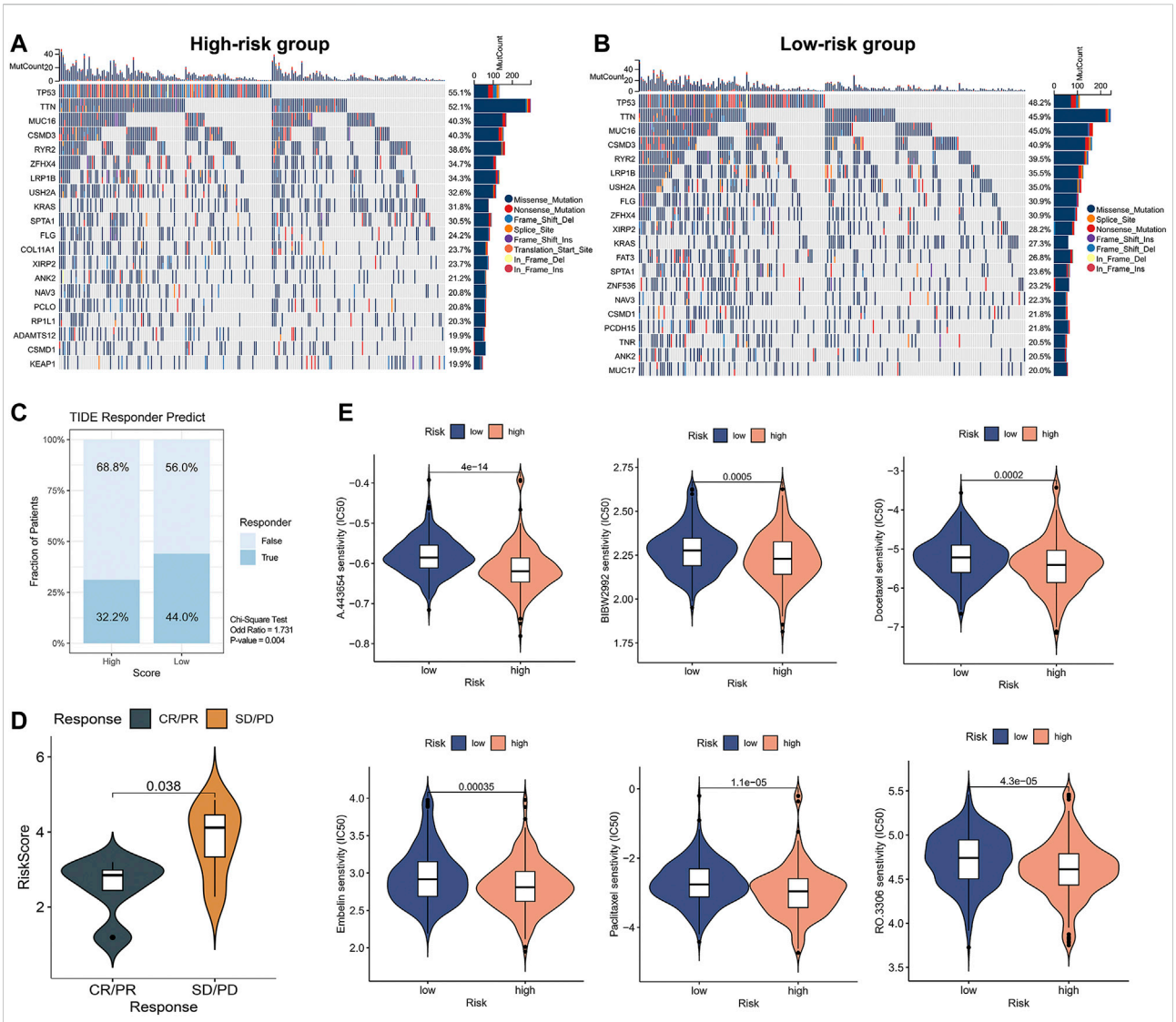


FIGURE 8
Mutation status and prediction of immunotherapy and chemotherapeutic response between high- and low-risk groups. **(A,B)** Analysis of top 20 mutation genes in high-risk versus low-risk groups. **(C)** TIDE immunotherapy response prediction in high- and low-risk groups of LUAD TCGA cohort. Chi-square test: no response to immunotherapy: high-versus low-risk, 68.8% versus 56.0%, OR = 1.731, p = 0.004. **(D)** The Independent immunotherapy dataset (GSE126044) validated immunotherapy efficacy in high- and low-risk groups. **(E)** Screening of chemotherapeutic agents sensitive to high- and low-risk groups.

cells, compared to the low-risk group. Spearman correlation analysis also showed that the abundance of immune cells was negatively correlated with the risk scores for almost all immune cell types except Th2 cells (Figure 7F). We also used the MCPcounter, TIMER, and EPIC algorithms to confirm these results. By plotting the heat map, we visualized that lung adenocarcinoma patients in the high-risk group had lower levels of immune infiltration (Figure 7G).

Risk score can suggest disparities in gene mutations and guide immunotherapy and chemotherapy

Since gene mutation status significantly affects tumor formation and progression, we performed a mutation analysis of lung adenocarcinoma patients in high- and low-risk groups. We further analyzed the differences between the top 20 genes with the highest mutation frequencies in the high

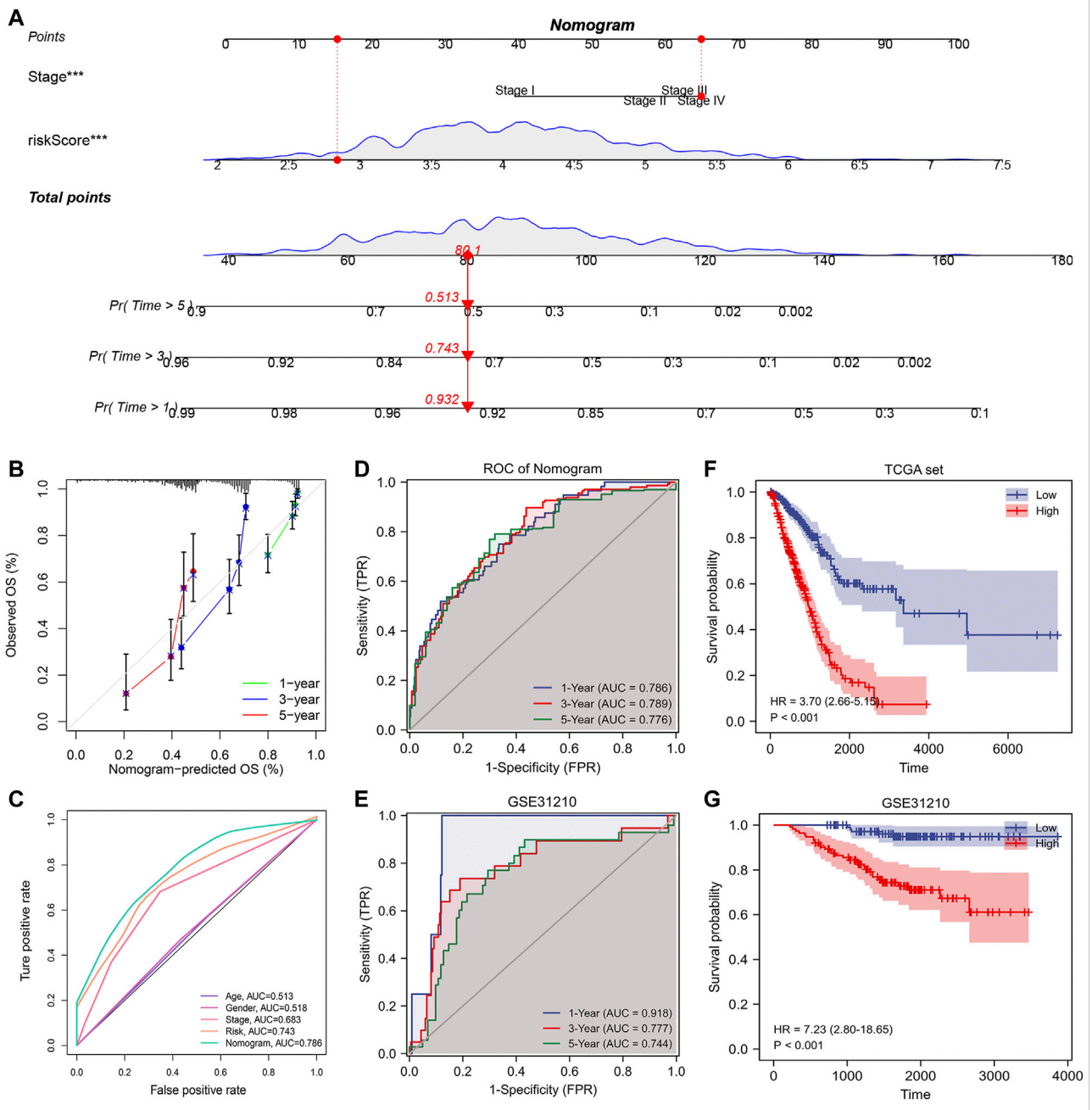


FIGURE 9
Building a more accurate nomogram. **(A)** Nomogram was constructed by combining pathological stages with risk scores. **(B)** Nomogram's 1-, 3-, and 5-years calibration curve. **(C)** The ROC curve shows AUC values for various clinical factors, risk scores, and nomogram scores. **(D)** ROC analysis of nomogram score in TCGA database. **(E)** ROC analysis of nomogram score in GSE31210 database (validation set). **(F)** Kaplan-Meier prognostic analysis of nomogram score in TCGA database. **(G)** Kaplan-Meier prognostic analysis of nomogram score in GSE31210 database.

and low-risk groups respectively (Figures 8A,B). The frequency of TP53 mutations in the high-risk group was 55.1%, while it was 48.2% in the low-risk group. The frequency of KRAS mutations was 31.8% in the high-risk group versus 27.3% in the low-risk group. The difference in gene mutation frequency might be the reason for the poorer prognosis in the high-risk group. Based on the above analyses, we explored the treatment strategies for different risk scores in depth. Since immunotherapy is commonly used in lung adenocarcinoma, we first calculated the response to immunotherapy in high- and low-risk groups using the TIDE algorithm (Figure 8C). The results suggested that the

low-risk group had a better treatment response upon immunotherapy (OR = 1.731, p = 0.004). We further validated this result using the GSE126044 database. Results showed that patients with CR/PR after immunotherapy had a significantly lower risk score than SD/PD (p = 0.038), suggesting that patients with a low-risk score are susceptible to benefit from immunotherapy (anti-PD-1 treatment) (Figure 8D). However, the high-risk group had a worse prognosis, so we performed a prediction of response to chemotherapeutic agents for patients in the high- and low-risk groups. We utilized the R package 'pRRophetic' to calculate the IC_{50} of chemotherapeutic drugs (Figure 8E). We found that the high-risk group had better sensitivity to A-443654, BIBW-2992, Docetaxel, Paclitaxel, Embelin, and RO-3306. Taken together, we provided a personalized treatment option for clinical reference by predicting the treatment effect of patients in different subgroups and compensated for the poorer effect of immunotherapy in the high-risk group.

A nomogram with a potential clinical application can be constructed based on risk score and pathological stage

The above study indicated that the risk score could act as an independent prognostic factor that can be used to determine the prognosis of patients. Therefore, to further improve our signature's predictive efficiency, we constructed a nomogram based on the TCGA database, incorporating factors such as pathological stage and risk scores (Figure 9A). We could visualize the risk assessment by calculating the score and assessing the outcome probability for each patient. In addition, we tested the predictive efficacy of the nomogram using a calibration plot (Figure 9B). The nomogram's 1-, 3-, and 5-years survival predictions were more accurate than the theoretical values. To further compare the predictive efficacy of the nomogram score with other clinical factors, we plotted the ROC curve. The nomogram score had the highest AUC value, and the predictive efficiency was further improved based on the risk score (Figure 9C). By performing the time-dependent ROC curve based on the nomogram score, we found that the AUC values of 1-, 3-, and 5-years overall survival for lung adenocarcinoma patients were 0.786, 0.789, and 0.776, respectively. To verify the accuracy of the nomogram, we constructed the nomogram again based on the external dataset GSE31210. Time-dependent ROC analysis showed that the AUC values of 1-, 3-, and 5-years overall survival for lung adenocarcinoma patients were 0.918, 0.777, and 0.744, respectively (Figures 9D,E). These results suggested that the nomogram we constructed had good accuracy. We also performed Kaplan-Meier prognostic analyses using the TCGA and external database GSE31210. The results indicated that the high nomogram score group had a significantly worse prognosis than the low score group (Figures 9F,G). In summary, the nomogram significantly improved the accuracy of determining the survival status of lung adenocarcinoma patients.

Discussion

The interconnection between tumor cells, immune cells, and stromal cells in the TME substantially influences tumorigenesis and tumor progression (Anderson and Simon, 2020). The spatial interplay of immune cells and other cells in the TME determines the immune response against tumors (Petitprez et al., 2020). As immunotherapy of tumors has been intensively studied, immune checkpoint inhibitors (ICIs) against tumor immune escape are expected to be an essential strategy to improve the prognosis of lung adenocarcinoma patients (Qiao et al., 2021; Liu et al., 2022; Reda et al., 2022). However, due to the heterogeneity and complexity of the TME, patients with the same pathological stage may also exhibit different TME characteristics, resulting in different therapeutic effects upon immunotherapy (Bagaev et al., 2021). Therefore, developing a prognostic model to determine the prognosis of lung adenocarcinoma patients early and to provide targeted immunotherapeutic strategies has significant potential for clinical application.

Tumor-associated macrophages, a vital member of the TME, have been identified in two types with different functional features, the classically activated M1 macrophages and the alternative activated M2 macrophages (Cassetta and Pollard, 2020). M1 macrophages appear in the inflammatory environment and are usually induced by cytokines from Th1, whereas M2 macrophages are primarily induced by cytokines from Th2 and counteract the inflammatory response (Sedighzadeh et al., 2021). Previous studies have shown that although all M2 macrophages exhibit anti-inflammatory and immunomodulatory effects, there is still heterogeneity among M2 macrophages. M2 macrophages can be further distinguished into four subtypes. M2a macrophages are involved in tissue fibrosis, M2b macrophages are shown to promote tumor progression, M2c macrophages are exhibited to be involved in tissue remodeling, and M2d macrophages promote angiogenesis (Wang L. X. et al., 2019). A growing number of studies have shown a significant correlation between M2 macrophages and lung adenocarcinoma progression. Lung adenocarcinoma patients with a higher density of M2 macrophages tend to have a poorer prognosis (Cao et al., 2019; Guo et al., 2019; Dai et al., 2020). M2 macrophages create an environment conducive to tumor survival by releasing growth factors, chemokines, and other inflammatory mediators (Solinas et al., 2009; Lin et al., 2019). In addition, M2 macrophages can also promote tumor metastasis and invasion by promoting angiogenesis and other pathways (Jetten et al., 2014; Xie et al., 2021). To further analyze the differences in composition and function within M2 macrophages, we used bioinformatics to perform an in-depth analysis at the single cell level.

Through annotation and clustering analysis of single-cell data, we identified differences in the composition of M2 macrophages in lung adenocarcinoma patients. These differences might indirectly contribute to the discrepancy in

biological processes and prognosis among patients. In addition, we performed differentiation trajectory analysis and pseudo-time analysis on M2 macrophages, identifying the different differentiation states of M2 macrophages. Lung adenocarcinoma patients exhibited four expression patterns based on the heterogeneous genes in the differentiation trajectory. The GSVA enrichment analysis helped us to understand the functional differences between different states of M2 macrophages. The GSVA results from our study above showed that M2 macrophages in state5 had the highest angiogenesis score, M2 macrophages in state1 had the highest epithelial-mesenchymal transition (EMT) score, and M2 macrophages in state2&4 had the highest interferon-response score. In contrast, M2 macrophages in state3 had the lowest G2M checkpoint score. This discrepancy reveals that macrophages in different states may differ in their cancer-promoting functions. Previous studies have also shown that M2 macrophages are significantly associated with angiogenesis and lymphangiogenesis, which contribute to the development of lung cancer, and also support that M2 cells are a strong indicator of poor prognosis in lung cancer (Hwang et al., 2020). Identifying particular metastasis-promoting or EMT-promoting subtypes of M2 macrophages also can help to explore the underlying molecular mechanisms further. Additionally, this heterogeneity of different states of M2 macrophages was also reflected in the metabolic and cellular communication levels. Our results provide a novel insight into the heterogeneity in M2 macrophages. Whereas the previous classification of M2 macrophages was based on different cytokine activation patterns (Colin et al., 2014), we distinguished different differentiation states of M2 macrophages based on single cell analysis. In addition, our study investigated the role of heterogeneous genes in the differentiation of M2 macrophages to guide the clinical therapy of lung adenocarcinoma.

We extracted heterogeneous genes, essential in the differentiation trajectory of M2 macrophages, and performed a univariate Cox analysis to screen for prognosis-related genes in lung adenocarcinoma. We constructed a prognostic model using Lasso-Cox and multivariate stepwise regression methods based on the prognosis-related genes in the training set. We measured the predictive efficacy of the model and explored the potential molecular mechanisms between high- and low-risk groups. Previous studies have shown that the status of the tumor microenvironment can be quantitatively assessed by risk scores (Chong et al., 2021). In our research, we found that patients in the high-risk group had an immunosuppressive microenvironment while the low-risk group had an immune-promoting microenvironment. Notably, the treatment of immune checkpoint inhibitors (ICIs) has become a hot topic in tumor therapy strategies. Immunotherapy targeting M2-type macrophages is emerging as a new direction for tumor therapy (Mills et al., 2016). The major molecules targeted by immunotherapy are programmed death receptor 1 (PD-1) and programmed death receptor ligand 1 (PD-L1). However, due to the complexity of the *in vivo* microenvironment, immunotherapy has an obvious shortcoming in that only a fraction of tumor patients respond to ICIs treatment (Wang et al., 2021). TIDE, as a novel computational architecture, has been considered as an alternative to single biomarkers for predicting the therapeutic effect of ICIs (Jiang et al., 2018). With the dual validation of the TIDE algorithm and GSE126044 set, we found that the low-risk group benefited more from immunotherapy, and this also directly indicated that the prognostic model we constructed could advance the personalization of immunotherapy.

As the high-risk group was shown to have a poor prognosis, we identified chemotherapeutic agents (A-443654, BIBW-2992, Docetaxel, Embelin, Paclitaxel, RO-3306) with better sensitivity for the high-risk group. A-443654 is an inhibitor of the AKT pathway that induces apoptosis and inhibits tumor growth (Luo et al., 2005). BIBW-2992 was reported to inhibit the kinase activity of EGFR mutants and suppress lung adenocarcinoma development (Li et al., 2008). Docetaxel and Embelin can induce the apoptosis of lung adenocarcinoma tumor cells (Avisetti et al., 2014; Jeong et al., 2021). Meanwhile, Paclitaxel, as first-line chemotherapy for patients who do not benefit from immunotherapy, together with RO-3306, can cause cell cycle G2/M phase arrest and lead to apoptosis in lung adenocarcinoma cells (Vassilev et al., 2006; Cui et al., 2020). The above chemotherapy drugs could compensate for the deficiency in immunotherapy efficacy in the high-risk group. In addition, to further improve the predictive performance of the prognostic model, we constructed a nomogram by combining the risk scores with the pathological stages. Nomogram has significantly better prognostic efficacy than the pathological stages and can be used as a complement to clinical factors by providing a more refined risk assessment.

In summary, for the first time, this research constructed a signature that can assess the prognosis of lung adenocarcinoma patients based on heterogeneous genes related to the differentiation trajectory of M2 macrophages. Our results provide a new research idea for the precision treatment of lung adenocarcinoma. However, our study still has some shortcomings. More in-depth studies are needed in the future to identify the potential molecular mechanisms of heterogeneous genes associated with the differentiation of M2 macrophages.

Conclusion

M2 macrophages, as a critical component of the lung adenocarcinoma microenvironment, promote tumor progression and metastasis. In this study, we performed differentiation trajectory and pseudotime analysis using scRNA-seq data to identify different differentiation states of M2 macrophages. By exploring the heterogeneous genes associated with M2 macrophages' differentiation, we constructed a prognostic model to predict the prognosis and

adjuvant treatment effect of lung adenocarcinoma patients, which could potentially be used as a clinical parameter for clinicians' therapy decisions in the future.

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 data analysis of the paper: ZC and JY. Drafting of the manuscript: LT and SJ. Data collection and assistants in data analysis: YL and WZ. Data visualization: YB, CD, and JC. Article Revision: CX and CL. Funding and Supervising: JZ. All authors participated in the work and agreed to take responsibility for all aspects of the work, thus ensuring that the research was accurate and that relevant issues were properly investigated and resolved.

Funding

This work was supported by the grants from National Natural Science Foundation of China (No. 81873417); Suzhou science and Technology Bureau (LCZX2019002).

Acknowledgments

We thank Bullet Edits Limited for the linguistic editing and the public databases for their data support. We are very grateful to Jun Zhao for his valuable guidance.

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/fgene.2022.1010440/full#supplementary-material

References

Anderson, N. M., and Simon, M. C. (2020). The tumor microenvironment. *Curr. Biol.* 30, R921–R925. doi:10.1016/j.cub.2020.06.081

Avisetti, D. R., Babu, K. S., and Kalivendi, S. V. (2014). Activation of p38/JNK pathway is responsible for embelin induced apoptosis in lung cancer cells: Transitional role of reactive oxygen species. *PLoS One* 9, e87050. doi:10.1371/journal.pone.0087050

Bagaev, A., Kotlov, N., Nomie, K., Svekolkin, V., Gafurov, A., Isaeva, O., et al. (2021). Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell.* 39, 845–865.e7. doi:10.1016/j.ccell.2021.04.014

Balachandran, V. P., Gonen, M., Smith, J. J., and DeMatteo, R. P. (2015). Nomograms in oncology: More than meets the eye. *Lancet. Oncol.* 16, e173–180. doi:10.1016/S1470-2045(14)71116-7

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 17, 218. doi:10.1186/s13059-016-1070-5

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

Cao, L., Che, X., Qiu, X., Li, Z., Yang, B., Wang, S., et al. (2019). M2 macrophage infiltration into tumor islets leads to poor prognosis in non-small-cell lung cancer. *Cancer Manag. Res.* 11, 6125–6138. doi:10.2147/CMAR.S199832

Cassetta, L., and Pollard, J. W. (2020). Tumor-associated macrophages. *Curr. Biol.* 30, R246–R248. doi:10.1016/j.cub.2020.01.031

Chong, W., Shang, L., Liu, J., Fang, Z., Du, F., Wu, H., et al. (2021). m(6)A regulator-based methylation modification patterns characterized by distinct tumor microenvironment immune profiles in colon cancer. *Theranostics* 11, 2201–2217. doi:10.7150/thno.52717

Colin, S., Chinetti-Gbaguidi, G., and Staels, B. (2014). Macrophage phenotypes in atherosclerosis. *Immunol. Rev.* 262, 153–166. doi:10.1111/imr.12218

Cui, H., Arnst, K., Miller, D. D., and Li, W. (2020). Recent advances in elucidating Paclitaxel resistance mechanisms in non-small cell lung cancer and strategies to overcome drug resistance. *Curr. Med. Chem.* 27, 6573–6595. doi:10.2174/0929867326666191016113631

Dai, X., Lu, L., Deng, S., Meng, J., Wan, C., Huang, J., et al. (2020). USP7 targeting modulates anti-tumor immune response by reprogramming Tumor-associated Macrophages in Lung Cancer. *Theranostics* 10, 9332–9347. doi:10.7150/thno.47137

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

Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One* 9, e107468. doi:10.1371/journal.pone.0107468

Guo, Z., Song, J., Hao, J., Zhao, H., Du, X., Li, E., et al. (2019). M2 macrophages promote NSCLC metastasis by upregulating CRYAB. *Cell. Death Dis.* 10, 377. doi:10.1038/s41419-019-1618-x

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

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., 3rd, Zheng, S., Butler, A., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell.* 184, 3573–3587.e29. doi:10.1016/j.cell.2021.04.048

He, D., Wang, D., Lu, P., Yang, N., Xue, Z., Zhu, X., et al. (2021). Single-cell RNA sequencing reveals heterogeneous tumor and immune cell populations in early-stage lung adenocarcinomas harboring EGFR mutations. *Oncogene* 40, 355–368. doi:10.1038/s41388-020-01528-0

Hwang, I., Kim, J. W., Ylaya, K., Chung, E. J., Kitano, H., Perry, C., et al. (2020). Tumor-associated macrophage, angiogenesis and lymphangiogenesis markers predict prognosis of non-small cell lung cancer patients. *J. Transl. Med.* 18, 443. doi:10.1186/s12967-020-02618-z

Jeong, M. S., Lee, K. W., Choi, Y. J., Kim, Y. G., Hwang, H. H., Lee, S. Y., et al. (2021). Synergistic antitumor activity of SH003 and Docetaxel via EGFR signaling inhibition in non-small cell lung cancer. *Int. J. Mol. Sci.* 22, 8405. doi:10.3390/ijms22168405

Jetten, N., Verbruggen, S., Gijbels, M. J., Post, M. J., De Winther, M. P., and Donners, M. M. (2014). Anti-inflammatory M2, but not pro-inflammatory M1 macrophages promote angiogenesis *in vivo*. *Angiogenesis* 17, 109–118. doi:10.1007/s10456-013-9381-6

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1

Kamata, T., So, T. Y., Ahmed, Q., Giblett, S., Patel, B., Luo, J., et al. (2020). Fibroblast-derived STC-1 modulates tumor-associated macrophages and lung adenocarcinoma development. *Cell. Rep.* 31, 107802. doi:10.1016/j.celrep.2020.107802

Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., and Tanabe, M. (2021). Kegg: Integrating viruses and cellular organisms. *Nucleic Acids Res.* 49, D545–D551. doi:10.1093/nar/gkaa970

Kim, N., Kim, H. K., Lee, K., Hong, Y., Cho, J. H., Choi, J. W., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* 11, 2285. doi:10.1038/s41467-020-16164-1

Kobak, D., and Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nat. Commun.* 10, 5416. doi:10.1038/s41467-019-13056-x

Li, D., Ambrogio, L., Shimamura, T., Kubo, S., Takahashi, M., Chirieac, L. R., et al. (2008). BIBW2992, an irreversible EGFR/HER2 inhibitor highly effective in preclinical lung cancer models. *Oncogene* 27, 4702–4711. doi:10.1038/onc.2008.109

Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 48, W509–W514. doi:10.1093/nar/gkaa407

Li, Y., Gu, J., Xu, F., Zhu, Q., Chen, Y., Ge, D., et al. (2021). Molecular characterization, biological function, tumor microenvironment association and clinical significance of m6A regulators in lung adenocarcinoma. *Brief. Bioinform.* 22, bbaa225. doi:10.1093/bib/bbaa225

Lin, Y., Xu, J., and Lan, H. (2019). Tumor-associated macrophages in tumor metastasis: Biological roles and clinical therapeutic applications. *J. Hematol. Oncol.* 12, 76. doi:10.1186/s13045-019-0760-3

Liu, B., Hu, X., Feng, K., Gao, R., Xue, Z., Zhang, S., et al. (2022). Temporal single-cell tracing reveals clonal revival and expansion of precursor exhausted T cells during anti-PD-1 therapy in lung cancer. *Nat. Cancer* 3, 108–121. doi:10.1038/s43018-021-00292-8

Long, G., Ouyang, W., Zhang, Y., Sun, G., Gan, J., Hu, Z., et al. (2021). Identification of a DNA repair gene signature and establishment of a prognostic nomogram predicting biochemical-recurrence-free survival of prostate cancer. *Front. Mol. Biosci.* 8, 608369. doi:10.3389/fmolb.2021.608369

Luo, Y., Shoemaker, A. R., Liu, X., Woods, K. W., Thomas, S. A., de Jong, R., et al. (2005). Potent and selective inhibitors of Akt kinases slow the progress of tumors *in vivo*. *Mol. Cancer Ther.* 4, 977–986. doi:10.1158/1535-7163.MCT-05-0005

Maacha, S., Bhat, A. A., Jimenez, L., Raza, A., Haris, M., Uddin, S., et al. (2019). Extracellular vesicles-mediated intercellular communication: Roles in the tumor microenvironment and anti-cancer drug resistance. *Mol. Cancer* 18, 55. doi:10.1186/s12943-019-0965-7

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

Mills, C. D., Lenz, L. L., and Harris, R. A. (2016). A breakthrough: Macrophage-directed cancer immunotherapy. *Cancer Res.* 76, 513–516. doi:10.1158/0008-5472.CAN-15-1737

Pan, Y., Yu, Y., Wang, X., and Zhang, T. (2020). Tumor-associated macrophages in tumor immunity. *Front. Immunol.* 11, 583084. doi:10.3389/fimmu.2020.583084

Petitprez, F., Meylan, M., de Reynies, A., Sautes-Fridman, C., and Fridman, W. H. (2020). The tumor microenvironment in the response to immune checkpoint blockade therapies. *Front. Immunol.* 11, 784. doi:10.3389/fimmu.2020.00784

Qiao, M., Jiang, T., Liu, X., Mao, S., Zhou, F., Li, X., et al. (2021). Immune checkpoint inhibitors in EGFR-mutated NSCLC: Dusk or dawn? *J. Thorac. Oncol.* 16, 1267–1288. doi:10.1016/j.jtho.2021.04.003

Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E., and Gfeller, D. (2017). Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife* 6, e26476. doi:10.7554/eLife.26476

Reda, M., Ngamcherdtrakul, W., Nelson, M. A., Siriwon, N., Wang, R., Zaidan, H. Y., et al. (2022). Development of a nanoparticle-based immunotherapy targeting PD-L1 and PLK1 for lung cancer treatment. *Nat. Commun.* 13, 4261. doi:10.1038/s41467-022-31926-9

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

Sedighzadeh, S. S., Khoshbin, A. P., Razi, S., Keshavarz-Fathi, M., and Rezaei, N. (2021). A narrative review of tumor-associated macrophages in lung cancer: Regulation of macrophage polarization and therapeutic implications. *Transl. Lung Cancer Res.* 10, 1889–1916. doi:10.21037/tlcr-20-1241

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

Solinas, G., Germano, G., Mantovani, A., and Allavena, P. (2009). Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation. *J. Leukoc. Biol.* 86, 1065–1073. doi:10.1189/jlb.0609385

Srivastava, S., Mohanty, A., Nam, A., Singhal, S., and Salgia, R. (2022). Chemokines and NSCLC: Emerging role in prognosis, heterogeneity, and therapeutics. *Semin. Cancer Biol.* doi:10.1016/j.semcancer.2022.06.010

Sun, D., Wang, J., Han, Y., Dong, X., Ge, J., Zheng, R., et al. (2021). Tisch: A comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. *Nucleic Acids Res.* 49, D1420–D1430. doi:10.1093/nar/gkaa1020

Thai, A. A., Solomon, B. J., Sequist, L. V., Gainor, J. F., and Heist, R. S. (2021). Lung cancer. *Lancet* 398, 535–554. doi:10.1016/S0140-6736(21)00312-3

Varol, C., Mildner, A., and Jung, S. (2015). Macrophages: Development and tissue specialization. *Annu. Rev. Immunol.* 33, 643–675. doi:10.1146/annurev-immunol-032414-112220

Vassilev, L. T., Tovar, C., Chen, S., Knezevic, D., Zhao, X., Sun, H., et al. (2006). Selective small-molecule inhibitor reveals critical mitotic functions of human CDK1. *Proc. Natl. Acad. Sci. U. S. A.* 103, 10660–10665. doi:10.1073/pnas.0600447103

Wang, H., Lengerich, B. J., Aragam, B., and Xing, E. P. (2019a). Precision Lasso: Accounting for correlations and linear dependencies in high-dimensional genomic data. *Bioinformatics* 35, 1181–1187. doi:10.1093/bioinformatics/bty750

Wang, L. X., Zhang, S. X., Wu, H. J., Rong, X. L., and Guo, J. (2019b). M2b macrophage polarization and its roles in diseases. *J. Leukoc. Biol.* 106, 345–358. doi:10.1002/JLB.3RU1018-378RR

Wang, M., Herbst, R. S., and Boshoff, C. (2021). Toward personalized treatment approaches for non-small-cell lung cancer. *Nat. Med.* 27, 1345–1356. doi:10.1038/s41591-021-01450-2

Wang, X., Miao, J., Wang, S., Shen, R., Zhang, S., Tian, Y., et al. (2022). Single-cell RNA-seq reveals the Genesis and heterogeneity of tumor microenvironment in pancreatic undifferentiated carcinoma with osteoclast-like giant-cells. *Mol. Cancer* 21, 133. doi:10.1186/s12943-022-01596-8

Wang, Y., Wang, R., Zhang, S., Song, S., Jiang, C., Han, G., et al. (2019c). iTALK: an R Package to characterize and illustrate intercellular communication. bioRxiv, 507871.

Wu, F., Fan, J., He, Y., Xiong, A., Yu, J., Li, Y., et al. (2021a). Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat. Commun.* 12, 2540. doi:10.1038/s41467-021-22801-0

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

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

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, 134–153. doi:10.1158/2159-8290.CD-21-0316

Xie, Y., Chen, Z., Zhong, Q., Zheng, Z., Chen, Y., Shangguan, W., et al. (2021). M2 macrophages secrete CXCL13 to promote renal cell carcinoma migration, invasion, and EMT. *Cancer Cell. Int.* 21, 677. doi:10.1186/s12935-021-02381-1

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

Yunna, C., Mengru, H., Lei, W., and Weidong, C. (2020). Macrophage M1/M2 polarization. *Eur. J. Pharmacol.* 877, 173090. doi:10.1016/j.ejphar.2020.173090

Zhang, J., Li, H., Wu, Q., Chen, Y., Deng, Y., Yang, Z., et al. (2019). Tumoral NOX4 recruits M2 tumor-associated macrophages via ROS/PI3K signaling-dependent various cytokine production to promote NSCLC growth. *Redox Biol.* 22, 101116. doi:10.1016/j.redox.2019.101116

Zhou, J., Tang, Z., Gao, S., Li, C., Feng, Y., and Zhou, X. (2020). Tumor-associated macrophages: Recent insights and therapies. *Front. Oncol.* 10, 188. doi:10.3389/fonc.2020.00188

Zhou, W., Bai, Y., Chen, J., Li, H., Zhang, B., and Liu, H. (2022). Revealing the critical regulators of modulated smooth muscle cells in atherosclerosis in mice. *Front. Genet.* 13, 900358. doi:10.3389/fgene.2022.900358

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 21 September 2022
DOI 10.3389/fgene.2022.1016449

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Bian Chengyu,
Nanjing Medical University, China
Kui Hu,
Guizhou Provincial People's Hospital, China

*CORRESPONDENCE
Yongxiang Song,
Songtang2004@163.com

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 11 August 2022
ACCEPTED 05 September 2022
PUBLISHED 21 September 2022

CITATION
Liu Y, Yu M, Cheng X, Zhang X, Luo Q, Liao S, Chen Z, Zheng J, Long K, Wu X, Qu W, Gong M and Song Y (2022), A novel LUAD prognosis prediction model based on immune checkpoint-related lncRNAs.
Front. Genet. 13:1016449.
doi: 10.3389/fgene.2022.1016449

A novel LUAD prognosis prediction model based on immune checkpoint-related lncRNAs

Yang Liu†, Mingyang Yu†, Xuechao Cheng, Xingshu Zhang, Qian Luo, Sijin Liao, Zhongzheng Chen, Jianhao Zheng, Kaijun Long, Xingwei Wu, Wendong Qu, Ming Gong and Yongxiang Song*

Department of Thoracic Surgery, The Affiliated Hospital of Zunyi Medical University, Zunyi, Guizhou, China

Lung adenocarcinoma (LUAD) is a malignant disease with an extremely poor prognosis, and there is currently a lack of clinical methods for early diagnosis and precise treatment and management. With the deepening of tumor research, more and more attention has been paid to the role of immune checkpoints (ICP) and long non-coding RNAs (lncRNAs) regulation in tumor development. Therefore, this study downloaded LUAD patient data from the TCGA database, and finally screened 14 key ICP-related lncRNAs based on ICP-related genes using univariate/multivariate COX regression analysis and LASSO regression analysis to construct a risk prediction model and corresponding nomogram. After multi-dimensional testing of the model, the model showed good prognostic prediction ability. In addition, to further elucidate how ICP plays a role in LUAD, we jointly analyzed the immune microenvironmental changes in LAUD patients and performed a functional enrichment analysis. Furthermore, to enhance the clinical significance of this study, we performed a sensitivity analysis of common antitumor drugs. All the above works aim to point to new directions for the treatment of LUAD.

KEYWORDS

lung adenocarcinoma, lncRNA, immune check point, tumor microenvironment, bioinformatic analyse

Abbreviations: CAFs, cancer-associated fibroblasts; CAR, chimeric antigen receptor; CNV, copy number variation; CTLA-4, cytotoxic T lymphocyte antigen-4; GDSC, genomics of drug sensitivity in cancer; ICP, immune checkpoints; ICPDEGs, immune checkpoint genes; IL-2, interleukin 2; lncRNAs, long non-coding RNAs; LUAD, lung adenocarcinoma; MDSCs, myeloid-derived macrophages; NSCLC, non-small cell lung cancer; OS, overall survival; PD-1, programmed cell death protein 1; SCLC, small cell lung cancer; TIDE, tumor immune dysfunction and exclusion; TIL, tumor-infiltrating lymphocyte; TMB, tumor mutational burden; TME, tumor microenvironment; Tregs, regulatory T cells.

Introduction

Lung cancer, as one of the most common types of cancer all over the world, has gained much attention in recent years (Cao et al., 2020; Ferlay et al., 2018). It was estimated that 2.09 million new cases were newly diagnosed, and 1.76 million patients died in 2018 (Bray et al., 2018). According to histological types, lung cancer could be classified as non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) roughly, and lung adenocarcinoma (LUAD) was the major subtype that accounted for over one million worldwide deaths annually (Zhang et al., 2020a). Smoking has become the most common risk factor for LUAD (Gould et al., 2007). Several approaches have been used in the clinical treatment of LUAD patients, mainly including radiotherapy, chemotherapy, and surgical resection according to the TNM system (Nasim et al., 2019). When progressed to advanced stages, survival decreased monthly sharply, so it is of great need for early diagnosis and intervention (Steven et al., 2016). Along with the rapid growth of large-scale genomic studies in recent decades, some somatic mutations associated with LUAD have been noticed like TP53, KRAS, EGFR, et al., which emphasized the importance of immunotherapies (Campbell et al., 2016). Meanwhile, for advanced LUAD, the effect of chemotherapy was greatly limited by its malignant nature, and immunotherapy seed to be the most effective approach to provide early diagnosis and improve survival status (Zheng et al., 2021). So, more immune therapeutic targets are needed for better and more precise clinical diagnosis and prognosis.

With the growing development in immunotherapy, several types with different mechanisms of action have been applied in clinical treatment, like vaccinations, monoclonal antibodies, and checkpoint inhibitors (Abbott and Ustoyev, 2019). Oncolytic vaccines were created in the 1920s and shelved until 1976 due to lake of understanding of the specific mechanism.

As an effective method for non-Hodgkin's lymphoma, rituximab has gradually been used in many types of cancer as an important monoclonal antibody (Ribatti, 2014). The latest immune checkpoint (ICP) proteins, like programmed cell death protein 1 (PD-1) and antibodies against cytotoxic T lymphocyte antigen-4 (CTLA-4) also have been fully investigated (Thompson, 2018). The former is a cell-surface receptor expressed on immune cell types, while the latter mainly reduces interleukin 2 (IL-2) production and T-cell proliferation (Kennedy and Salama, 2020). As for a novel T cell-target method, chimeric antigen receptor (CAR) T cell therapies have been developed and approved for clinical use mainly in hematological cancers owing to the delivery barriers faced by solid tumors (Fesnak et al., 2016). Therefore, it is of great importance to explore novel targets for solid tumors, especially LUAD.

Long non-coding RNAs (lncRNAs) represent a major class of regulatory non-coding RNAs larger than 200 nt in length (Peng et al., 2017). Altered immune infiltration is a hallmark of the tumor, and it is well recognized that lncRNAs regulate the immune response in cancer progression (Zhang et al., 2020b). Some studies demonstrated that the ectopic expression of lncRNA-cell division cycle six promoted proliferation and metastasis of breast cancer cells *via* regulation of the G1 phase checkpoint, demonstrating a critical effect in tumor development (Kong et al., 2019). Meanwhile, much emphasis has been put on the tumor microenvironment (TME) to further elucidate the immune alteration which influences tumor development apart from tumor cells. In solid tumors, TME consists of several types of immune cells and stromal cells, like cancer-associated fibroblasts (CAFs), regulatory T cells (Tregs), myeloid-derived macrophages (MDSCs), etc. (Mu and Najafi, 2021). While the correlation between lncRNAs and TME remains a mystery.

Thus, we conducted an overall immune checkpoint-related lncRNAs risk and prognostic model in patients with LUAD, trying to explore risk factors for cancer clinical care through bioinformatics technique and survival analysis, and provide potential therapeutic targets for clinical treatment.

Materials and methods

Data acquisition and processing

All relevant LUAD patients' information and data in this study were downloaded from the TCGA database (Blum et al., 2018). After excluding samples with missing prognostic information or survival time of fewer than 30 days, finally, 490 LUAD samples were included in this study. These samples are randomly divided into the training set and testing set. A total of 246 samples in the training set were used to develop a predictive risk model. The testing set included 244 samples used to validate the established risk model. The 47 ICP genes were derived from the latest research results of Liu et al. (2022) (Supplementary Table S1). ICP-related lncRNAs were obtained by Pearson's correlation test (Pearson correlation coefficient >0.4, $p < 0.001$), and 2,061 ICP-related lncRNAs were identified.

Differential RNA screening

The expression levels of lncRNAs and mRNA were extracted from the transcriptome data of LUAD and normal samples, respectively, and lncRNA expression and differential analysis were performed using the "limma" package, where genes with FDR < 0.05 and |logFC| > 1 were considered to have significant differences.

Construction of risk models

Combined with the prognostic information of patients, univariate regression analysis was used to screen the differential ICP-related lncRNAs associated with prognosis. Afterward, we used LASSO regression (R package "glmnet", version 4.1-3) to run 1,000 cycles of 10-fold cross-validation with $p < 0.05$. Finally, through multivariate regression analysis, a 14 ICP-related lncRNAs risk model was constructed.

We calculated the risk score with the following formula:

$$Risk\,score = \sum_{k=1}^{n} Coef\,(lncRNA) * expr\left(lncRNA^{k}\right)$$

where *Coef* (ln*cRNA*) represents the correlation coefficient between lncRNAs and survival, and *expr* (ln$cRNA^{K}$) represents the expression of lncRNAs. All selected LUAD samples were divided into high-risk and low-risk groups based on the mean risk scores.

Risk model testing and evaluation

Through univariate/multivariate regression analysis, ROC curves were performed ("glment," "survminer," and "survival" R packages) to test whether the risk model could be used as an independent predictor of prognosis in LUAD patients. In total, we calculated and plotted 1-, 3-, and 5-years ROC curves.

Survival analysis and principal component analysis

Kaplan–Meier (K-M) survival analysis was used to determine the overall survival (OS) of LUAD patients between two groups by the "survival" package. Principal component analysis (PCA) is used for efficient dimensionality reduction, model identification, and group visualization of high-dimensional data.

Nomogram construction

To better guide the clinical diagnosis and treatment of LUAD, we combined the risk scores, and other clinical features to construct a nomogram by "rms" package.

Tumor microenvironment and immunotherapy analysis

Using the "maftools" R package, tumor mutational burdens (TMBs) in LUAD patients were assessed. Furthermore, the CIBERSORT (Newman et al., 2015) and ssGSEA algorithm, as well as TIMER (http://timer.comp-genomics.org) {Li, 2017 #21} were performed to evaluate the immune cell infiltration status in different risk groups. To predict the efficacy of clinical immunotherapy in LUAD patients, we used Tumor Immune Dysfunction and Exclusion (TIDE) prediction.

Drug sensitivity analysis

IC50 of each LUAD patient relative to a common antineoplastic drug was determined as the patient's sensitivity to this drug using the Genomics of Drug Sensitivity in Cancer (GDSC) platform (Yang et al., 2013) and used R package pRRophetic (version 0.5) for calculation and visualization.

Functional analysis

GSEA analysis was done using gene set enrichment analyses software (https://www.gsea-msigdb.org/gsea/login.jsp) (Subramanian et al., 2005). GO and KEGG enrichment analyses based on the differential genes between high and low-risk groups were performed using DAVID online site (version 6.8), where relevant annotations with $p < 0.05$ and FDR < 0.05 were considered significantly different. Additionally, competitive endogenous RNA (ceRNA) networks were constructed and visualized using Cytoscape (version 3.6.1).

Statistical analysis

All statistical analyses were performed in R software (version 4.1.1). Differences between groups were compared using the Wilcoxon rank-sum test. K-W tests were used to compare differences between three or more groups. Statistical significance was defined as a $p < 0.05$ if the above methods were not specifically stated.

Results

Expression and copy number variation of immune checkpoints-related genes in LUAD

The workflow is shown in Figure 1. Forty-seven ICP-related genes were obtained for further analysis (Supplementary Table S1) as well as LUAD patients' clinical features can be found in Supplementary Table S2. Differences in the expression of ICP-related genes between 535 tumor tissues and 59 normal tissue samples are shown in Figure 2A. Additionally, somatic copy number variation (CNV) among 47 ICP genes was studied (Figure 2C). The ICP-related lncRNAs interaction network is shown in the form of the Sankey diagram (Figure 2B, r > 0.4, $p <$

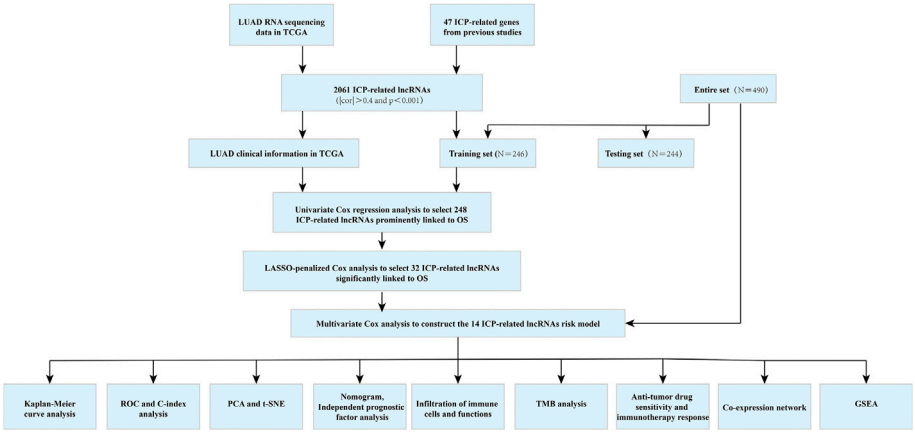


FIGURE 1
Workflow chart.

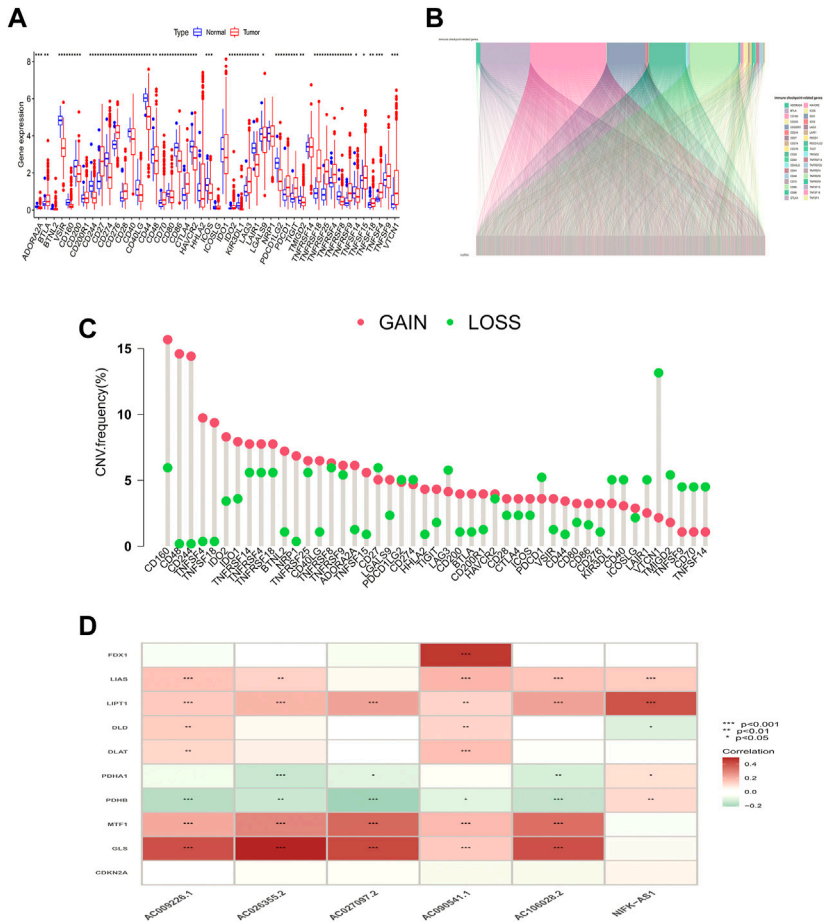


FIGURE 2
Genetic and expression variation of the MRDEGs in LUAD patients. **(A)** ICP-related gene expression profile. **(B)** Sankey relation diagram for ICP-related genes and lncRNAs. **(C)** The CNV frequency of 47 ICP genes in the LUAD cohort. **(D)** Heatmap for the correlations between randomly ICP-related genes and lncRNAs.

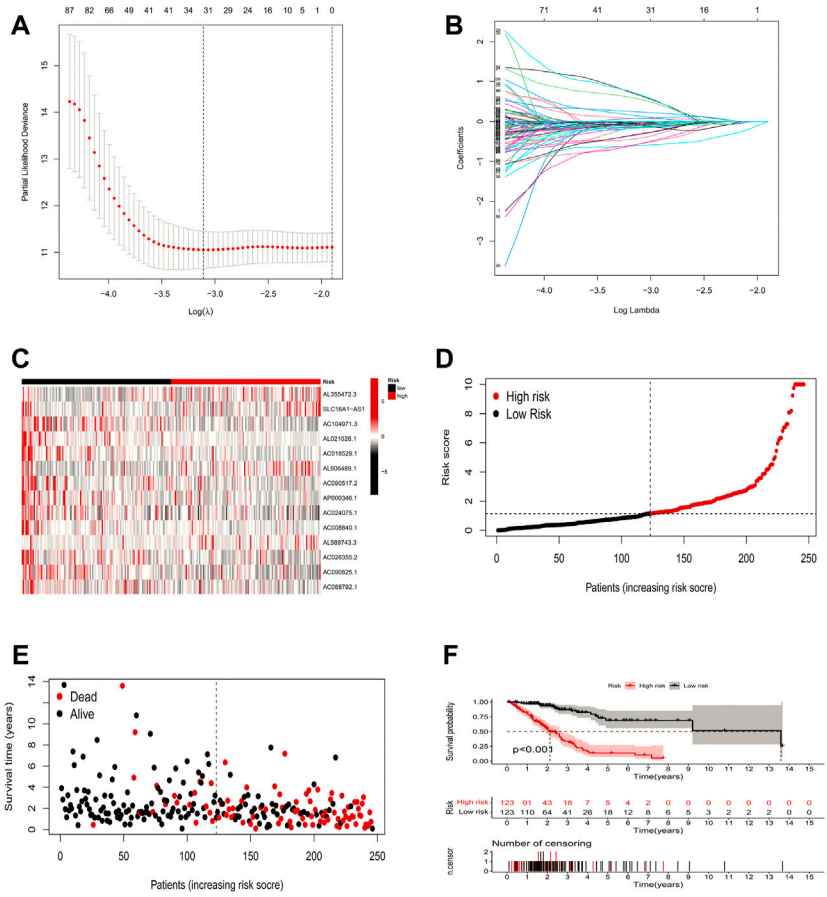


FIGURE 3
Risk model construction and validation. **(A,B)** Result of LASSO regression analysis. **(C)** Heatmap to show the expression of 14 lncRNAs between high- and low-risk groups in the training set. **(D)** Expression differences of 14 ICP-related lncRNAs in different risk groups in the training set. **(E)** Distribution of sample risk score and different patterns of survival status/time between the high-risk and low-risk groups in the training set. **(F)** Kaplan-Meier curve of high-risk and low-risk patients in the training set.

0.001), and the correlation between ICP-related genes and lncRNAs is shown in Figure 2D.

Risk model construction and validation

In this study, 248 ICP-related lncRNAs were screened by using univariate Cox regression analysis (Supplementary Table S3). To prevent overfitting prognostic features, we further performed LASSO Cox analysis and 32 lncRNAs that were highly correlated with LUAD prognosis (Figures 3A,B). Finally, 14 ICP-related lncRNAs with the strongest prognostic predictive ability were identified by multivariate COX regression analysis (Supplementary Table S4) for risk model construction.

The formula for the risk score is:

Risk scores = AL355472.3*(1.7297898132188) + SLC16A1-AS1* (2.02045359749171) + AC104971.3*(−0.861884680221223) + AL021026.1*(−1.92527800701289) +
AC018529.1*(−1.41451695188048) +
AL606489.1*(0.484465069923853) +
AC090517.2*(−0.67755983967278) +
AP000346.1*(−1.34031276829931) +
AC024075.1*(−0.456581688150894) +
AC008840.1*(−1.67186147401095) +
AL589743.3*(1.01756516292608) +
AC026355.2*(−0.383441820151563) +
AC090825.1*(−0.894102424172019) +
AC068792.1*(−0.778661427765015).

With the above signatures, the patient's prognostic risk score was calculated. For each patient, the relative expression levels of 14 ICP-related lncRNAs are presented in Figure 3C. Based on the mean risk scores, we divided all LUAD patient samples into high-risk and low-risk groups, where the patient distribution in the high-risk and low-risk groups of the training set is shown in Figure 3D. Figure 3E demonstrates the survival status and survival time of patients in the high-risk and low-risk groups in the training set. Figure 3F shows

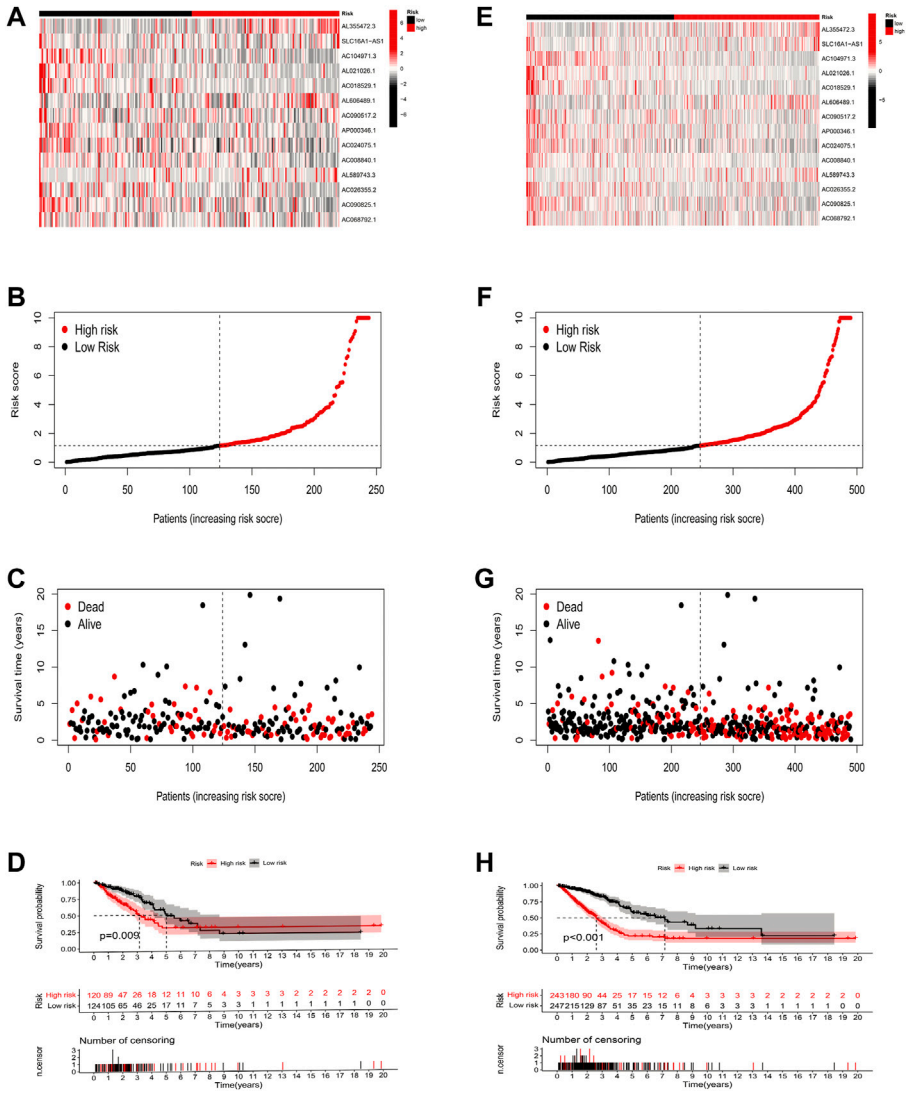


FIGURE 4
Risk model construction and validation in testing and entire sets. **(A–D)** The expression of 14 key prognostic lncRNAs in the testing set, the survival status of LUAD patients, the risk score, and the results of K-M analysis of survival analysis. **(E–H)** The expression of 14 key prognostic lncRNAs in the entire set, the survival status of LUAD patients, the risk score, and the results of K-M analysis of survival analysis.

the prognosis and survival of patients in different risk groups in the training set (based on K-M survival analysis). It can be seen that the prognosis of patients can be clearly distinguished in the training set after changing the risk model ($p < 0.001$).

To validate the predictive capability of the constructed model, we calculated the risk scores of LUAD patients by using a uniform formula. We examined the expression of ICP-related lncRNAs, survival status scores, and risk scores in LUAD patients in the testing set (Figures 4A–D) and the entire set (Figures 4E–H). In addition, the K-M analysis of the two sets also showed that patients in the low-risk group had a longer OS time than those in the high-risk group (Figures 4D,H $p = 0.009$ and $p < 0.001$).

Nomogram and independent prognostic factor analysis

To explore the independent predictive power of risk models and various clinical characteristics for patient outcomes, we performed univariate and all-factor Cox regression analyses, respectively. Univariate Cox regression analysis suggested that age, T/N grade, clinical stage, and risk score were prognostic factors for LUAD patients (Figure 5A, $p < 0.001$), and further multivariate Cox regression analysis showed that the risk score was an independent predictor of prognosis in LUAD patients, the prediction results were reliable, and the confidence level was

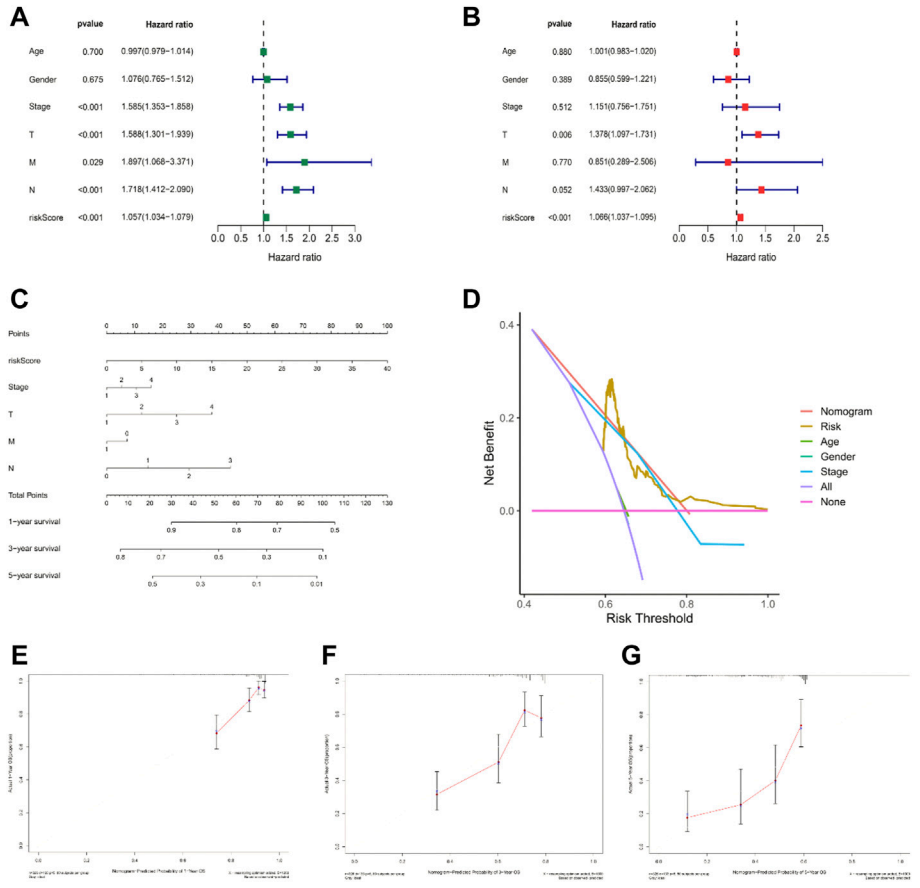


FIGURE 5
Nomogram and independent prognostic factor analysis. **(A,B)** Result of univariate/multivariate Cox regression analyses. **(C)** Nomogram predicts the probability of the 1-, 3-, and 5-years OS. **(D)** Result of DCA. **(E–G)** 1-, 3-, and 5-years predicted prognosis.

higher than that of other clinical characteristics (Figure 5B, $p < 0.001$). Therefore, it is reasonable to believe that a risk model based on 14 ICP-related lncRNAs has a significant impact on the survival and prognosis of LUAD patients and were independent prognostic factors. The nomogram (Figure 5C) was constructed with a risk score, survival rate, and other clinical features. Then the calibration curve analysis in Figures 5E–G shows the accuracy of the nomogram in predicting the 1-, 3-, and 5-years prognosis of LUAD patients. Furthermore, DCA also indicated that a nomogram has a higher prediction accuracy than a risk model alone (Figure 5D).

Otherwise, ROC curve analysis and PCA verify the efficacy of the risk model. The AUC values of the 1-, 3-, and 5-years of OS were 0.729, 0.753, and 0.735, respectively (Figure 6A), which were much higher than other clinical features (Figure 6B). This suggests that these 14 ICP-related lncRNAs are relatively reliable in the prognostic risk model of LUAD. Besides, we applied principal component analysis (PCA) to test the differences between the high-risk and low-risk groups (Figures 6C,D) to further assess the group ability of ICP-related lncRNA models. At the same time, we used PCA to verify the authenticity of the risk model constructed based on the entire gene expression profiles, ICP genes, ICP-related lncRNAs, and risk model according to the 14 hub lncRNAs (Figures 6E–H). The results confirmed that the distributional patterns of the high-risk and low-risk groups were significantly different, which elucidated that the risk model was competent to distinguish the two groups with high accuracy.

Somatic mutation landscape

We further analyzed the somatic mutation landscape of LUAD patients. As shown in Figures 7A,B, compared with the low-risk group, the high-risk group showed a higher rate of somatic mutation (92.92% vs. 83.75%), and also had a higher tumor mutational burden (TMB, Figure 7C, $p = 0.054$, with

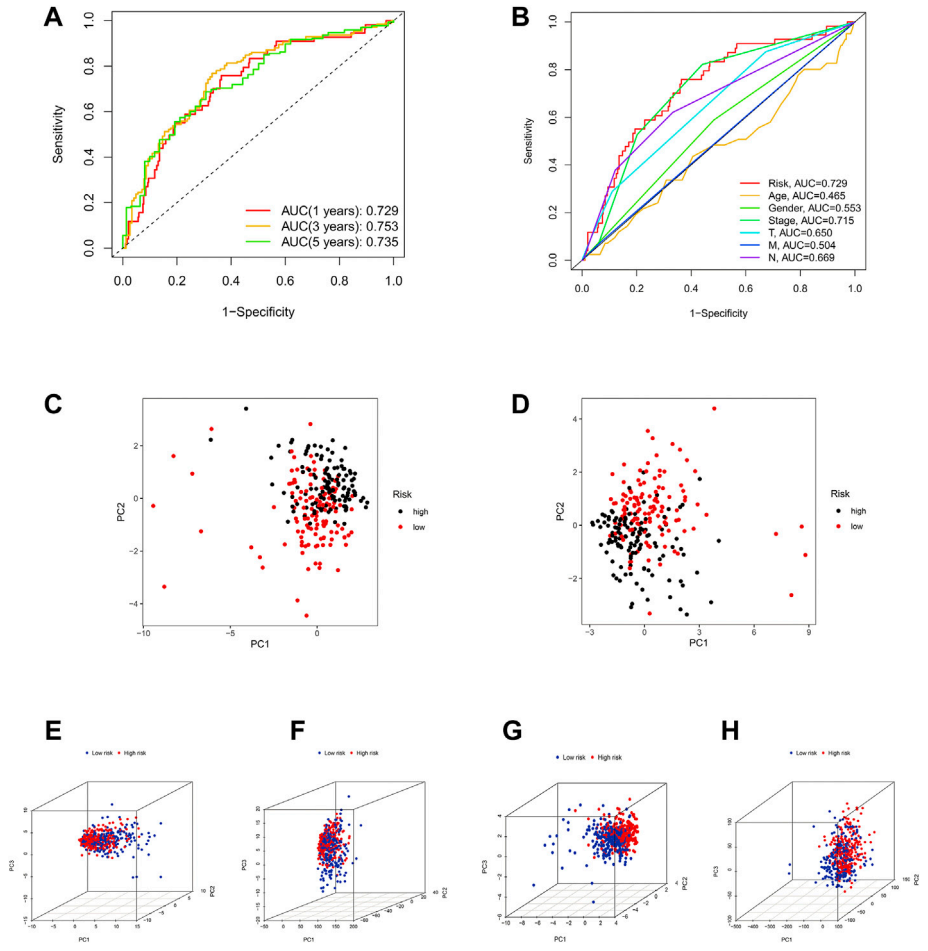


FIGURE 6
Assessment of the predictive risk model and Principal component analysis. **(A)** The entire set's 1-, 3-, and 5-years ROC curves. **(B)** ROC curves of the clinical characteristics and risk score. **(C,D)** PCA results of testing and training sets. **(E–H)** The PCA result of entire gene expression profiles, ICPDEGs, ICP-related lncRNAs, and risk model according to the 14 hub lncRNAs, respectively.

marginal statistical significance). As a classic indicator for evaluating tumor behavior, TMB has been considered reliable in evaluating the prognosis of tumor patients in the past. However, in the survival analysis, we were pleasantly surprised to find that TMB alone could not predict the prognosis of LUAD patients in the high and low TMB groups (Figure 7D, $p = 0.082$), but the combination of the TMB and risk score model can effectively distinguish the prognosis of patients with different risk levels (Figure 7E, $p < 0.0001$).

Immune signature analysis based on ICP-related lncRNAs

Figures 8A,B shows the proportion of 22 immune cells in different risk groups in the LUAD samples (Supplementary Table S5). Further ssGSEA immunoassays revealed that a variety of immune cells, including $CD8^+$ T cells, and B cells, were less infiltrated in the high-risk group samples, and more diverse in the high-risk group. The immune process activity was also lower than that of the low-risk group (Figures 8C,D). In addition, we found that all three immune scores (stromal score, immune score, and ESTIMATE score) were significantly higher in the low-risk group of LUAD patients, indicating that the TME was different from the high-risk group (Figures 8E–G). The above results suggest that patients at high risk of LUAD were in a more severe immunosuppressed state.

Interestingly, we found that monocytes and plasma cells could well predict the prognosis of patients in different risk groups (Figures 8H,I). Meanwhile, we also found that combined with risk scores, all LUAD samples could be classified into different immune subtypes (Figure 8J), which means that

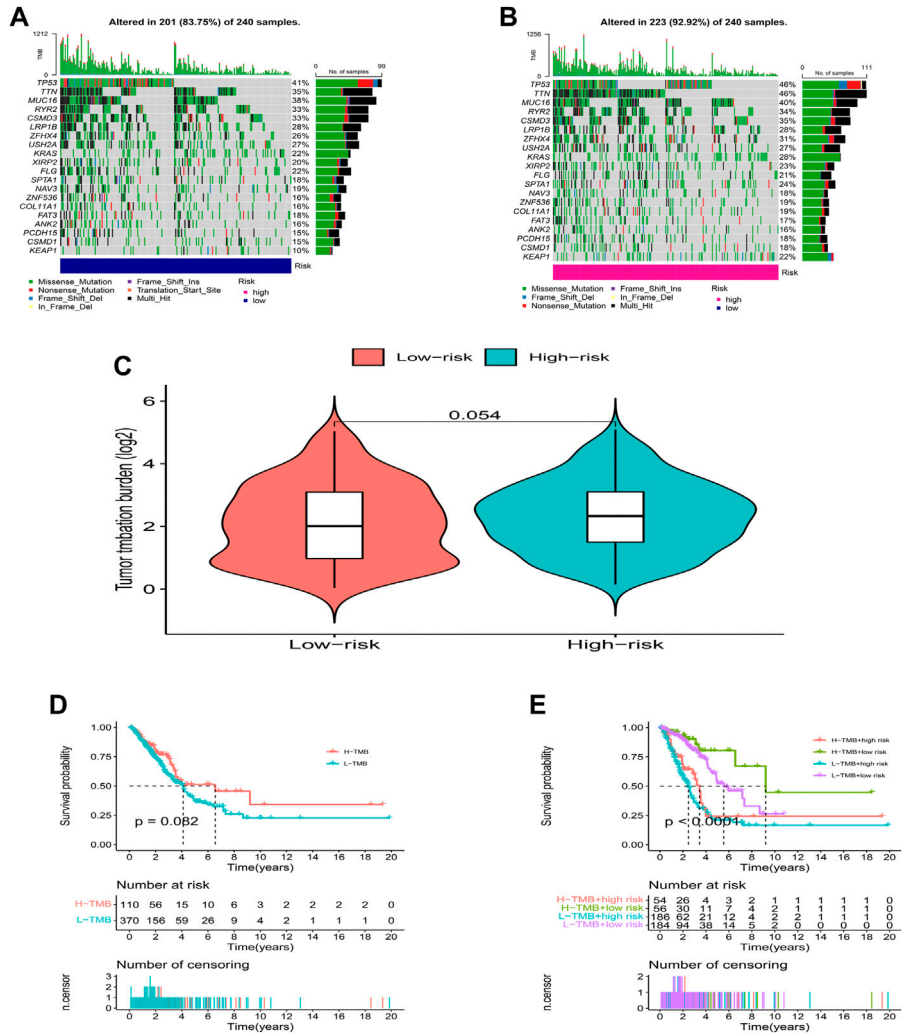


FIGURE 7
TMB analysis. **(A,B)** The waterfall plot of somatic mutation features established with high- and low-risk groups. **(C)** Tumor mutation burden in the high-risk and low-risk groups. **(D,E)** Kaplan-Meier curve of the OS among the high- and low- TMB groups.

more precise treatment strategies may be adopted for different subtypes in clinical practice.

Clinical immunotherapy analysis

For a better clinical therapeutic strategy in LUAD, a drug sensitivity analysis was conducted. The result showed that LUAD patients in the high-risk group had higher IC50s for AS601245, ATRA, ABT.888, and AP.24534, which indicated that these drugs may be clinically less effective for patients in the high-risk group. On the contrary, AG.014699, AUY922, AZD.0530, and A.443654 showed higher IC50 in the low-risk group (Figure 9A). Furthermore, we found that patients in the high-risk group had lower TIDE scores (Figure 9B), which may explain the differences in susceptibility to these drugs.

Functional enrichment analysis

To deeply explore how ICP-related lncRNAs produce biological effects, functional enrichment analysis based on differential genes between high and low-risk groups with multi-dimension was performed. The results of the GO analysis suggested that the changes of LAUD involved changes in a variety of immune processes, including humoral immunity and immune complex production (Figure 10A), and the KEGG results also suggested that the

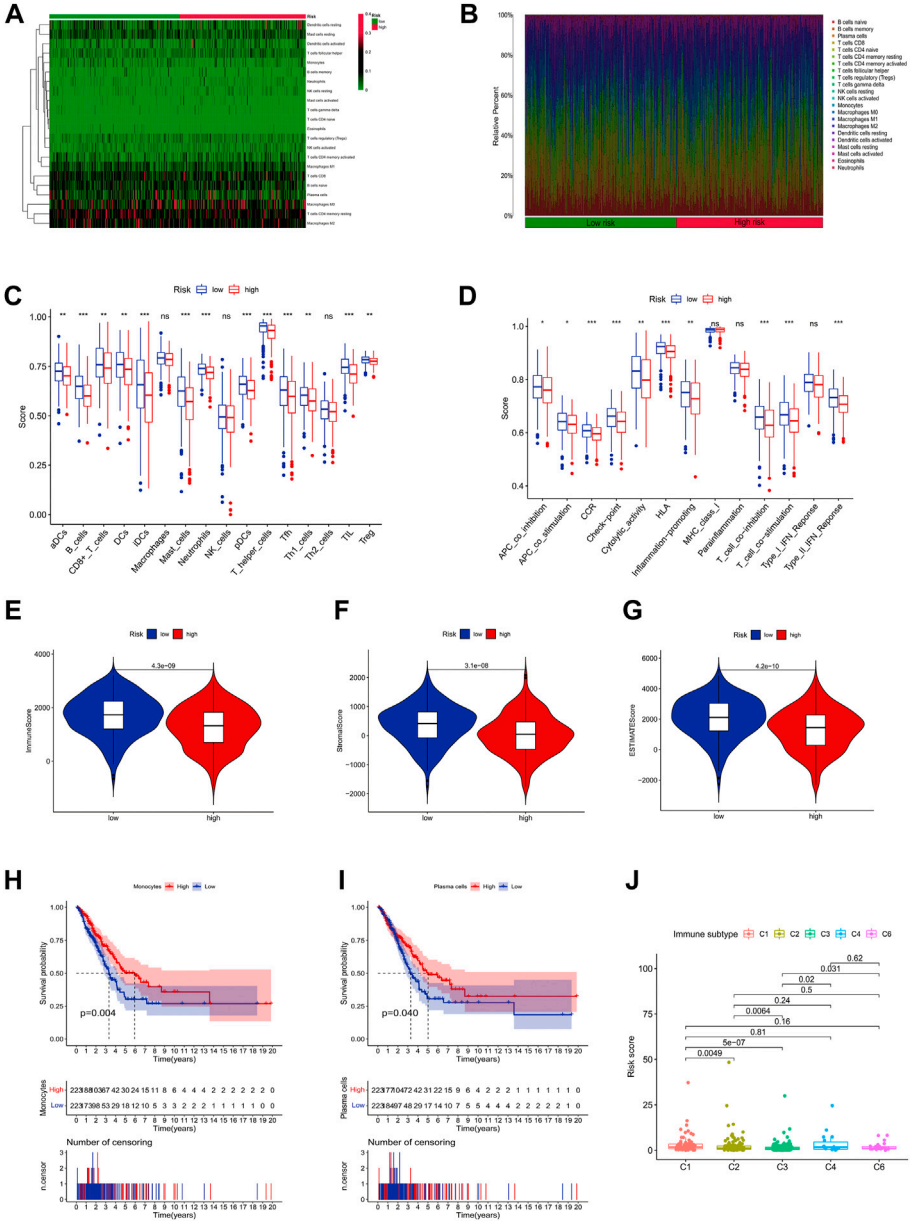


FIGURE 8
Immune infiltration discrepancy in different risk groups. **(A)** Heatmap of 22 tumor-infiltrating immune cell types in low-risk and high-risk groups. **(B)** Bar chart of the proportions for 22 immune cell types. **(C)** The ssGSEA scores of immune functions in low-risk and high-risk groups. **(D)** Immune cells in low-risk and high-risk groups. **(E–G)** The TME scores between high-risk and low-risk groups. **(H,I)** Survival analysis of combined immune cells. **(J)** Immune subtype.

disease was highly correlated with humoral immune pathways (Figure 10B). In addition, through GSEA analysis, we found that the B cell receptor pathway and cell adhesion pathway were highly enriched in the low-risk group, while those in the high-risk group were highly correlated with cell cycle and metabolic cycle (Figures 10C,D). These potential mechanisms may point to new directions for the future treatment of LUAD.

Discussion

Immunotherapy was defined as the use of materials to moderate the function of the immune system to prevent and fight disease (Lizée et al., 2013). It has been widely applied in clinical treatment for cancer like metastatic urothelial carcinoma (Sharma et al., 2016), advanced renal cell carcinoma (Motzer et al., 2015), and other types of cancer.

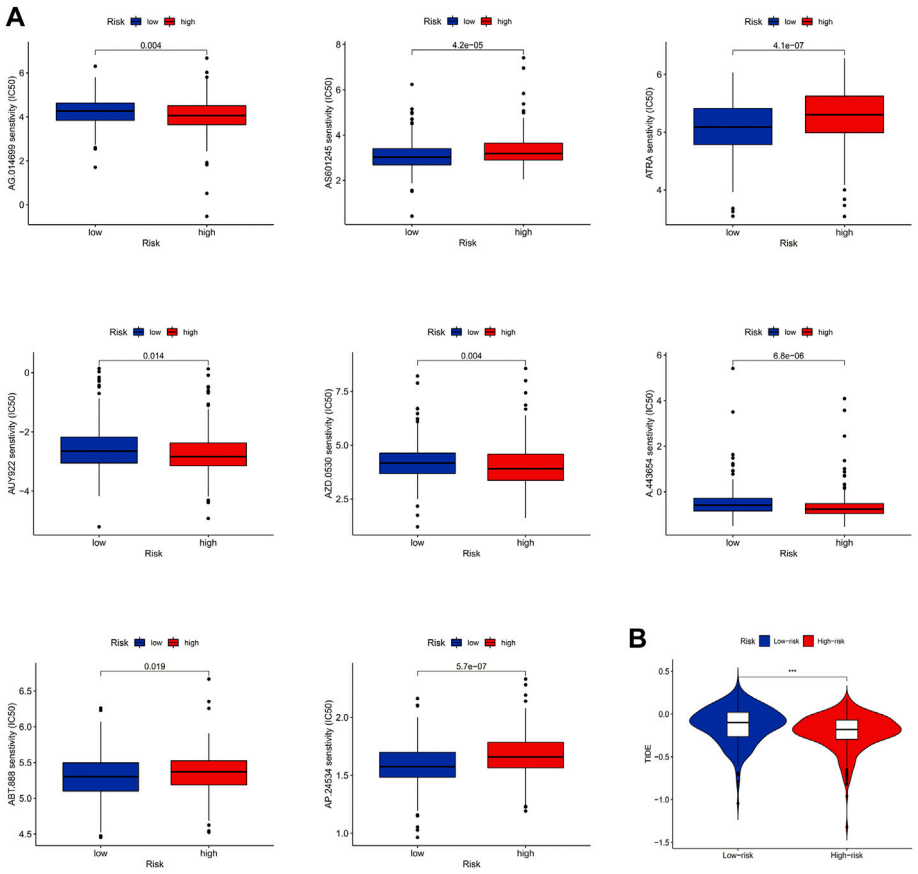


FIGURE 9
Clinical immunotherapy analysis. **(A)** Results of drug sensitivity analysis. **(B)** Result of TIDE.

With the growing investigation into immune infiltration, there was gradually reaching a consensus that long non-coding RNAs (lncRNAs) have been associated with cancer immunity regulation and the tumor microenvironment (TME) (Zhou et al., 2021a). Some researchers suggested that immune-related lncRNAs could predict immune cell infiltration and immunotherapy response in patients with liver cancer (Zhou et al., 2021b; Huang et al., 2021), and bladder cancer (Wu et al., 2020), while the association in patients with LUAD is still not understood. At the same time, TME, as a novel hotpot in cancer research, has gained much attention in recent years. Unlike tumor cells, stromal cells also a participant in the initiation, progression, and metastasis of cancer, inducing both beneficial and adverse consequences for tumorigenesis (Stepaniak et al., 1986). Current most advanced TME-directed therapies including antiangiogenic drugs and treatment directed against cancer-associated fibroblasts and the extracellular matrix were already approved or evaluated in trials (Bejarano et al., 2021). Therefore, it is of great importance to investigate the correlation between lncRNAs and immune response in patients with LUAD.

Based on the above characteristics, we conducted a prognostic model aimed at evaluating the association between ICP-related lncRNAs and TMB in patients with LUAD through bioinformatics techniques and survival analysis, providing potential treatment targets for clinical therapy and prognosis.

In this study, we found that 14 ICP-related lncRNAs were significantly associated with LUAD by Cox analysis. Of course, most of them were rarely studied and there were already several investigations about some lncRNAs. Firstly, lncRNA SLC16A1-AS1 has been identified to play a vital role in the metabolic reprogramming as targeting and co-activating of E2F1 in patients with bladder cancer (Logotheti et al., 2020). A study by Tian and Hu (2021) also demonstrated that SLC16A1-AS1 was upregulated in hepatocellular carcinoma and might downregulate miR-141 through methylation to promote cancer cell proliferation. Similarly, in patients with glioblastoma, SLC16A1-AS1 might promote cancer cell proliferation by regulating miR-149 methylation and could be considered a potential diagnostic marker in glioblastoma (Long et al., 2021). Also, there were several studies about the function of

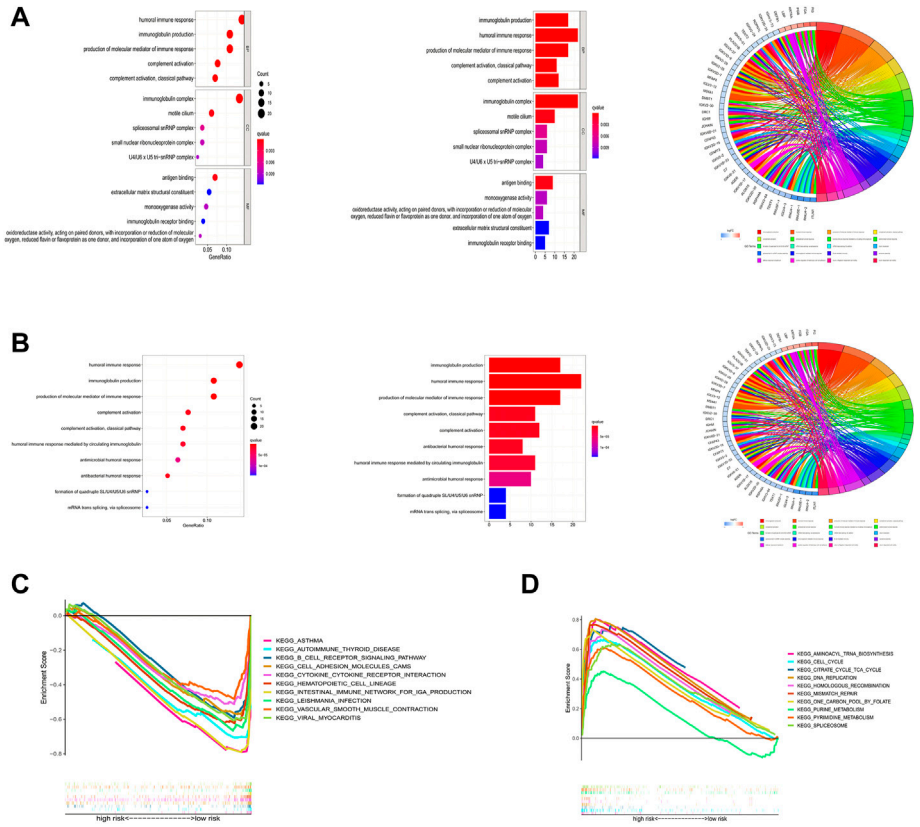


FIGURE 10
Functional enrichment analysis. **(A)** Result of GO enrichment. **(B)** Result of KEGG enrichment. **(C,D)** Result of GSEA.

SLC16A1-AS1 in oral squamous cell carcinoma (Feng et al., 2020; Li et al., 2022), and triple-negative breast cancer (Jiang et al., 2022). As for lung cancer, Liu et al. (2020) have proved that the expression of SLC16A1-AS1 was significantly lower in NSCLC tissue than that in adjacent tissue, and SLC16A1-AS1 over-expression could block the cell cycle and promote cell apoptosis *in vitro*, suggesting that it might act as a potential biomarker for patients with NSCLC. Then, when it comes to lncRNA AL606489.1, some investigations have proved that it was associated with ferroptosis in LUAD (Guo et al., 2021; Song et al., 2021; Wu et al., 2021) as well as oncosis (Chen et al., 2022), which all demonstrating a relationship between non-apoptotic cell death and LUAD and provide important predictive value for the prognosis of LUAD as well as potential clinical therapeutic targets. Similarly, AC026355.2, a vital immune-associated lncRNA, also showed its prognostic value for identifying immune and necroptosis characteristics in LUAD patients (He et al., 2021; Lu et al., 2022). As for lncRNA AC068792.1, a study by Zhou et al. (2022) proved that this TME-related lncRNA could be acted as a biomarker of clear cell renal carcinoma prognosis and immunotherapy response, while the effect in LUAD still warrants further exploration.

The GO and KEGG enrichment analysis showed that the ICP-related genes were mainly enriched in humoral immune response, immunoglobulin production, and production of the molecular immune response, emphasizing the significance of immune response in cancer development. Then the bioplot showed that the expression of immune cells in the high- and low-risk subgroups mainly focused on plasma cells, monocytes, T cells gamma delta, T cells CD4 memory resting, dendritic cells, and mast cells resting. Subsequent K-M survival analysis demonstrated that the survival probability in plasma cells high-expression subgroup was much higher than that in a low-expression subgroup, illustrating a potential protective value for patients with LUAD. Actually, in a study about the role of tumor-infiltrating B cells and intratumorally-produced antibodies in cancer-immunity interactions, Isaeva et al. (2019) found that plasma cells produced a great number of clonal IgG1, which was not much effective on prognosis, suggesting that IgG1+ tumor-infiltrating B cells might exert a beneficial effect in KRAS mutation cases. While, for the subgroup with higher expression of monocytes, the survival probability also showed the same result as that of plasma cells. As an important component in TME, monocytes were tightly connected with cancer initiation

and development. However, an investigation aiming at constructing an immune-related lncRNAs signature in patients with LUAD showed that this signature corrected negatively with B cells, $CD4^+$ T cells, and monocytes immune infiltration, and patients with low-risk scores had a higher abundance of immune cells and stromal cells around the tumor (Chen et al., 2021). This contrary result mainly could be explained that the function of tumor-associated monocyte/macrophage lineage cells (MMLCs) might be different in human tumors, especially in the early stages of the disease (Singhal et al., 2019). Classical "inflammatory" monocytes promote tumor growth and metastasis, however, nonclassical "patrolling" monocytes contribute to cancer immunosurveillance and may be targeted for cancer immunotherapy (Qian et al., 2011; Hanna et al., 2015). Thus, further studies are warranted to explore specific mechanisms in patients with LUAD.

Then we analyzed the immune score between high-risk and low-risk subgroups and found that significant differences were shown in tumor-infiltrating lymphocyte (TIL), cytolytic activity, and major histocompatibility complex. Indeed, the efficacy of clinical immunotherapy varies and depends on the amount and properties of TILs, and in general, TILs represent a favorable prognostic factor in NSCLC (Guo et al., 2018; Gueguen et al., 2021). Federico et al. (2022) demonstrated that though the number of infiltrating T cells was not associated with patient survival, the nature of the infiltrating T cells could have a prognostic value in NSCLC and became potential therapeutic approaches for clinical care. As for HLA, Datar et al. (2021) claimed that patients with cancer cell-selective HLA-B, HL-C or HLA class-II downregulation displayed decreased T cells and NK-cell infiltration, then associated with shorter overall survival, which broaden a novel insight into clinical therapeutic targets. While, for advanced NSCLC treated with immune checkpoint blockade, HLA class-I genotype was not correlated with survival, which emphasizes the correlation between immune checkpoints and HLA (Negrao et al., 2019). Further studies are needed to claim deeper relationships and provide novel insights.

In addition, there are still some limitations in our study. First, our current study was limited to the bioinformatics level and no external experiments were conducted to validate the results. Second, although we validated the model by constructing a valid prediction model, the model construction relied only on the TCGA database, which could potentially lead to less credible study results.

In conclusion, to explore the connection between lncRNAs and immune infiltration in patients with LUAD, we conduct a relatively overall and comprehensive prognostic model to evaluate the expression of various immune cells and survival probability through bioinformatics techniques, confirming that immune response played a vital role in the progression of cancer and the crosslink between immune infiltration and lncRNAs, which could provide potential therapeutic targets for clinical care.

Data availability statement

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

Author contributions

Project design: YS and YL; Database search and literature screen: YL and MY; Data extraction and analysis: XC, QL, and WQ; Quality evaluation of the included literature: YL and XZ; and manuscript writing: YL, SL, ZC, XW, JZ, KL, MG, and QL. The final draft was approved by all the authors.

Acknowledgments

We are grateful for the selfless dedication of the TCGA database.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1016449/full#supplementary-material

SUPPLEMENTARY TABLE S1
ICP-associated genes.

SUPPLEMENTARY TABLE S2
Clinical information.

SUPPLEMENTARY TABLE S3
Univariate COX regression analysis.

SUPPLEMENTARY TABLE S4
Multivariate COX regression analysis.

SUPPLEMENTARY TABLE S5
The proportion of 22 immune cells in different risk groups.

References

Abbott, M., and Ustoyev, Y. (2019). Cancer and the immune system: The history and background of immunotherapy. *Semin. Oncol. Nurs.* 35 (5), 150923. doi:10.1016/j.soncn.2019.08.002

Bejarano, L., Jordão, M. J. C., and Joyce, J. A. (2021). Therapeutic targeting of the tumor microenvironment. *Cancer Discov.* 11 (4), 933–959. doi:10.1158/2159-8290.CD-20-1808

Blum, A., Wang, P., and Zenklusen, J. C. (2018). SnapShot: TCGA-analyzed tumors. *Cell* 173 (2), 530. doi:10.1016/j.cell.2018.03.059

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

Campbell, J. D., Alexandrov, A., Kim, J., Wala, J., Berger, A. H., Pedamallu, C. S., et al. (2016). Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat. Genet.* 48 (6), 607–616. doi:10.1038/ng.3564

Cao, M., Li, H., Sun, D., and Chen, W. (2020). Cancer burden of major cancers in China: A need for sustainable actions. *Cancer Commun.* 40 (5), 205–210. doi:10.1002/cac2.12025

Chen, H., Shen, W., Ni, S., Sang, M., Wu, S., Mu, Y., et al. (2021). Construction of an immune-related lncRNA signature as a novel prognosis biomarker for LUAD. *Aging (Albany NY)* 13 (16), 20684–20697. doi:10.18632/aging.203455

Chen, H., Zhou, C., Hu, Z., Sang, M., Ni, S., Wu, J., et al. (2022). Construction of an algorithm based on oncosis-related LncRNAs comprising the molecular subtypes and a risk assessment model in lung adenocarcinoma. *J. Clin. Lab. Anal.* 36 (6), e24461. doi:10.1002/jcla.24461

Datar, I. J., Hauc, S. C., Desai, S., Gianino, N., Henick, B., Liu, Y., et al. (2021). Spatial analysis and clinical significance of HLA class-I and class-II subunit expression in non-small cell lung cancer. *Clin. Cancer Res.* 27 (10), 2837–2847. doi:10.1158/1078-0432.CCR-20-3655

Federico, L., McGrail, D. J., Bentebibel, S. E., Haymaker, C., Ravelli, A., Forget, M. A., et al. (2022). Distinct tumor-infiltrating lymphocyte landscapes are associated with clinical outcomes in localized non-small-cell lung cancer. *Ann. Oncol.* 33 (1), 42–56. doi:10.1016/j.annonc.2021.09.021

Feng, H., Zhang, X., Lai, W., and Wang, J. (2020). Long non-coding RNA slc16a1-AS1: Its multiple tumorigenesis features and regulatory role in cell cycle in oral squamous cell carcinoma. *Cell Cycle* 19 (13), 1641–1653. doi:10.1080/15384101.2020.1762048

Ferlay, J., ColoMbetM.Soerjomataram, I., Dyba, T., Randi, G., BettioM., et al. (2018). Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *Eur. J. Cancer* 103, 356–387. doi:10.1016/j.ejca.2018.07.005

Fesnak, A. D., June, C. H., and Levine, B. L. (2016). Engineered T cells: The promise and challenges of cancer immunotherapy. *Nat. Rev. Cancer* 16 (9), 566–581. doi:10.1038/nrc.2016.97

Gould, M. K., Fletcher, J., Iannettoni, M. D., Lynch, W. R., Midthun, D. E., Naidich, D. P., et al. (2007). Evaluation of patients with pulmonary nodules: When is it lung cancer?: ACCP evidence-based clinical practice guidelines. *Chest* 132 (3), 108S–130s. doi:10.1378/chest.07-1353

Gueguen, P., Metoikidou, C., Dupic, T., Lawand, M., Goudot, C., Baulande, S., et al. (2021). Contribution of resident and circulating precursors to tumor-infiltrating CD8(+) T cell populations in lung cancer. *Sci. Immunol.* 6 (55), eabd5778. doi:10.1126/sciimmunol.abd5778

Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat. Med.* 24 (7), 978–985. doi:10.1038/s41591-018-0045-3

Guo, Y., Qu, Z., Li, D., Bai, F., Xing, J., Ding, Q., et al. (2021). Identification of a prognostic ferroptosis-related lncRNA signature in the tumor microenvironment of lung adenocarcinoma. *Cell Death Discov.* 7 (1), 190. doi:10.1038/s41420-021-00576-z

Hanna, R. N., Cekic, C., Sag, D., Tacke, R., Thomas, G. D., Nowyhed, H., et al. (2015). Patrolling monocytes control tumor metastasis to the lung. *Science* 350 (6263), 985–990. doi:10.1126/science.aac9407

He, C., Yin, H., Zheng, J., Tang, J., Fu, Y., and Zhao, X. (2021). Identification of immune-associated lncRNAs as a prognostic marker for lung adenocarcinoma. *Transl. Cancer Res.* 10 (2), 998–1012. doi:10.21037/tcr-20-2827

Huang, S., Zhang, J., Lai, X., Zhuang, L., and Wu, J. (2021). Identification of novel tumor microenvironment-related long noncoding RNAs to determine the prognosis and response to immunotherapy of hepatocellular carcinoma patients. *Front. Mol. Biosci.* 8, 781307. doi:10.3389/fmolb.2021.781307

Isaeva, O. I., Sharonov, G. V., Serebrovskaya, E. O., Turchaninova, M. A., Zaretsky, A. R., ShugayM., et al. (2019). Intratumoral immunoglobulin isotypes predict survival in lung adenocarcinoma subtypes. *J. Immunother. Cancer* 7 (1), 279. doi:10.1186/s40425-019-0747-1

Jiang, B., Liu, Q., Gai, J., Guan, J., and Li, Q. (2022). LncRNA SLC16A1-AS1 regulates the miR-182/PDCD4 axis and inhibits the triple-negative breast cancer cell cycle. *Immunopharmacol. Immunotoxicol.* 44 (4), 534–540. doi:10.1080/08923973.2022.2056482

Kennedy, L. B., and Salama, A. K. S. (2020). A review of cancer immunotherapy toxicity. *Ca. Cancer J. Clin.* 70 (2), 86–104. doi:10.3322/caac.21596

Kong, X., Duan, Y., Sang, Y., Li, Y., Zhang, H., Liang, Y., et al. (2019). LncRNA-CDC6 promotes breast cancer progression and function as ceRNA to target CDC6 by sponging microRNA-215. *J. Cell. Physiol.* 234 (6), 9105–9117. doi:10.1002/jcp.27587

Li, T., Wang, D., and Yang, S. (2022). Analysis of the subcellular location of lncRNA SLC16A1-AS1 and its interaction with premature miR-5088-5p in oral squamous cell carcinoma. *Odontology*. Advance online publication. doi:10.1007/s10266-022-00712

Liu, H. Y., Lu, S. R., Guo, Z. H., Zhang, Z. S., Ye, X., Du, Q., et al. (2020). lncRNA SLC16A1-AS1 as a novel prognostic biomarker in non-small cell lung cancer. *J. Investig. Med.* 68 (1), 52–59. doi:10.1136/jim-2019-001080

Liu, J., Ling, Y., Su, N., Li, Y., Tian, S., Hou, B., et al. (2022). A novel immune checkpoint-related gene signature for predicting overall survival and immune status in triple-negative breast cancer. *Transl. Cancer Res.* 11 (1), 181–192. doi:10.21037/tcr-21-1455

Lizée, G., Overwijk, W. W., Radvanyi, L., Gao, J., Sharma, P., and Hwu, P. (2013). Harnessing the power of the immune system to target cancer. *Annu. Rev. Med.* 64, 71–90. doi:10.1146/annurev-med-112311-083918

Logotheti, S., Marquardt, S., Gupta, S. K., Richter, C., Edelhauser, B. A. H., Engelmann, D., et al. (2020). LncRNA-SLC16A1-AS1 induces metabolic reprogramming during Bladder Cancer progression as target and co-activator of E2F1. *Theranostics* 10 (21), 9620–9643. doi:10.7150/thno.44176

Long, Y., Li, H., Jin, Z., and Zhang, X. (2021). LncRNA slc16a1-AS1 is upregulated in glioblastoma and promotes cancer cell proliferation by regulating miR-149 methylation. *Cancer Manag. Res.* 13, 1215–1223. doi:10.2147/CMAR.S264613

Lu, Y., Luo, X., Wang, Q., Chen, J., Zhang, X., Li, Y., et al. (2022). A novel necroptosis-related lncRNA signature predicts the prognosis of lung adenocarcinoma. *Front. Genet.* 13, 862741. doi:10.3389/fgene.2022.862741

Motzer, R. J., Escudier, B., McDermott, D. F., George, S., Hammers, H. J., Srinivas, S., et al. (2015). Nivolumab versus everolimus in advanced renal-cell carcinoma. *N. Engl. J. Med.* 373 (19), 1803–1813. doi:10.1056/NEJMoa1510665

Mu, Q., and Najafi, M. (2021). Modulation of the tumor microenvironment (TME) by melatonin. *Eur. J. Pharmacol.* 907, 174365. doi:10.1016/j.ejphar.2021.174365

Nasim, F., Sabath, B. F., and Eapen, G. A. (2019). Lung cancer. *Med. Clin. North Am.* 103 (3), 463–473. doi:10.1016/j.mcna.2018.12.006

Negrao, M. V., Lam, V. K., Reuben, A., Rubin, M. L., Landry, L. L., Roarty, E. B., et al. (2019). PD-L1 expression, tumor mutational burden, and cancer gene mutations are stronger predictors of benefit from immune checkpoint blockade than HLA class I genotype in non-small cell lung cancer. *J. Thorac. Oncol.* 14 (6), 1021–1031. doi:10.1016/j.jtho.2019.02.008

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

Peng, W. X., Koirala, P., and Mo, Y. Y. (2017). LncRNA-mediated regulation of cell signaling in cancer. *Oncogene* 36 (41), 5661–5667. doi:10.1038/onc.2017.184

Qian, B. Z., Li, J., Zhang, H., Kitamura, T., Zhang, J., Campion, L. R., et al. (2011). CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature* 475 (7355), 222–225. doi:10.1038/nature10138

Ribatti, D. (2014). From the discovery of monoclonal antibodies to their therapeutic application: An historical reappraisal. *Immunol. Lett.* 161 (1), 96–99. doi:10.1016/j.imlet.2014.05.010

Sharma, P., Callahan, M. K., Bono, P., Kim, J., Spiliopoulou, P., Calvo, E., et al. (2016). Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): A multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. *Lancet. Oncol.* 17 (11), 1590–1598. doi:10.1016/S1470-2045(16)30496-X

Singhal, S., Stadanlick, J., Annunziata, M. J., Rao, A. S., Bhojnagarwala, P. S., O'Brien, S., et al. (2019). Human tumor-associated monocytes/macrophages and

their regulation of T cell responses in early-stage lung cancer. *Sci. Transl. Med.* 11 (479), eaat1500. doi:10.1126/scitranslmed.aat1500

Song, J., Sun, Y., Cao, H., Liu, Z., Xi, L., Dong, C., et al. (2021). A novel pyroptosis-related lncRNA signature for prognostic prediction in patients with lung adenocarcinoma. *Bioengineered* 12 (1), 5932–5949. doi:10.1080/21655979.2021.1972078

Stepaniak, P. C., Furst, J. J., and Woodard, D. (1986). Anabolic steroids as a countermeasure against bone demineralization during space flight. *Aviat. Space Environ. Med.* 57 (2), 174–178.

Steven, A., Fisher, S. A., and Robinson, B. W. (2016). Immunotherapy for lung cancer. *Respirology* 21 (5), 821–833. doi:10.1111/resp.12789

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

Thompson, J. A. (2018). New NCCN guidelines: Recognition and management of immunotherapy-related toxicity. *J. Natl. Compr. Canc. Netw.* 16 (5), 594–596. doi:10.6004/jnccn.2018.0047

Tian, J., and Hu, D. (2021). LncRNA SLC16A1-AS1 is upregulated in hepatocellular carcinoma and predicts poor survival. *Clin. Res. Hepatol. Gastroenterol.* 45 (2), 101490. doi:10.1016/j.clinre.2020.07.001

Wu, L., Wen, Z., Song, Y., and Wang, L. (2021). A novel autophagy-related lncRNA survival model for lung adenocarcinoma. *J. Cell. Mol. Med.* 25 (12), 5681–5690. doi:10.1111/jcmm.16582

Wu, Y., Zhang, L., He, S., Guan, B., He, A., Yang, K., et al. (2020). Identification of immune-related LncRNA for predicting prognosis and immunotherapeutic response in bladder cancer. *Aging (Albany NY)* 12 (22), 23306–23325. doi:10.18632/aging.104115

Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41, D955–D961. doi:10.1093/nar/gks1111

Zhang, C., Zhang, Z., Sun, N., Zhang, Z., Zhang, G., Wang, F., et al. (2020). Identification of a costimulatory molecule-based signature for predicting prognosis risk and immunotherapy response in patients with lung adenocarcinoma. *Oncoimmunology* 9 (1), 1824641. doi:10.1080/2162402X.2020.1824641

Zhang, M., Wang, N., Song, P., Fu, Y., Ren, Y., Li, Z., et al. (2020). LncRNA GATA3-AS1 facilitates tumour progression and immune escape in triple-negative breast cancer through destabilization of GATA3 but stabilization of PD-L1. *Cell Prolif.* 53 (9), e12855. doi:10.1111/cpr.12855

Zheng, Y., Tian, H., Zhou, Z., Xiao, C., Liu, H., Liu, Y., et al. (2021). A novel immune-related prognostic model for response to immunotherapy and survival in patients with lung adenocarcinoma. *Front. Cell Dev. Biol.* 9, 651406. doi:10.3389/fcell.2021.651406

Zhou, L., Fang, H., Guo, F., Yin, M., Long, H., and Weng, G. (2022). Computational construction of TME-related lncRNAs signature for predicting prognosis and immunotherapy response in clear cell renal cell carcinoma. *J. Clin. Lab. Anal.* 36, e24582. doi:10.1002/jcla.24582

Zhou, M., Zhang, Z., Bao, S., Hou, P., Yan, C., Su, J., et al. (2021). Computational recognition of lncRNA signature of tumor-infiltrating B lymphocytes with potential implications in prognosis and immunotherapy of bladder cancer. *Brief. Bioinform.* 22 (3), bbaa047. doi:10.1093/bib/bbaa047

Zhou, P., Lu, Y., Zhang, Y., and Wang, L. (2021). Construction of an immune-related six-lncRNA signature to predict the outcomes, immune cell infiltration, and immunotherapy response in patients with hepatocellular carcinoma. *Front. Oncol.* 11, 661758. doi:10.3389/fonc.2021.661758

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 23 September 2022
DOI 10.3389/fgene.2022.972212

OPEN ACCESS

EDITED BY
Qingyuan Huang,
Shanghai Cancer Center, Fudan University, China

REVIEWED BY
Huiyan Lei,
University of Texas Southwestern Medical Center, United States
Yunyun Jin,
East China Normal University, China

*CORRESPONDENCE
Shuqun Cheng,
chengshuqun@aliyun.com

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 18 June 2022
ACCEPTED 09 September 2022
PUBLISHED 23 September 2022

CITATION
Zhu H, Mao F, Wang K, Feng J and Cheng S (2022), Cuproptosis-related lncRNAs predict the clinical outcome and immune characteristics of hepatocellular carcinoma. *Front. Genet.* 13:972212. doi: 10.3389/fgene.2022.972212

Cuproptosis-related lncRNAs predict the clinical outcome and immune characteristics of hepatocellular carcinoma

Hongfei Zhu[1,2†], Feifei Mao[1†], Kang Wang[2], Jinkai Feng[2] and Shuqun Cheng[1,2*]

[1]Tongji University Cancer Center, Shanghai 10th People's Hospital, School of Medicine, Tongji University, Shanghai, China, [2]Department of Hepatic Surgery VI, Eastern Hepatobiliary Surgery Hospital, Second Military Medical University, Shanghai, China

Cuproptosis, as a novel copper-dependent and non-apoptotic form of cell death, is induced by aggregation of lipoylated mitochondrial proteins and the instability of Fe-S cluster proteins. However, the role of cuproptosis-related long noncoding RNAs (CRLncRNAs) in hepatocellular carcinoma (HCC) has not been clearly elucidated. In this study, we identified and characterized cuproptosis-related lncRNAs in HCC. 343 HCC cases from The Cancer Genome Atlas (TCGA) with gene transcriptome data and clinical data were obtained for analysis after the screening. Univariate and multivariate Cox proportional hazards analyses were performed to establish a prognostic cuproptosis-related lncRNA signature (CRlncSig). We established a prognosis-related model consisting of nine cuproptosis-related lncRNAs: GSEC, AL158166.1, AC005479.2, AL365361.1, AC026412.3, AL031985.3, LINC00426, AC009974.2, AC245060.7, which was validated in the internal cohort. High-risk group stratified by the CRlncSig was significantly related to poor prognosis ($p < 0.001$). The area under the receiver operating characteristic curve (AUC) of 1 year, 3 years, and 5 years of survival were 0.813, 0.789, and 0.752, respectively. Furthermore, a prognostic nomogram including CRlncSig with clinicopathologic factors was built with favorable predictive power. In addition, GO and KEGG enrichment analysis suggested that CRlncSig was involved in many carcinogenesis and immune-related pathways. Additionally, we found that tumor microenvironment, immune infiltration, immune function, and drug response were significantly different between the high-risk and low-risk groups based on the risk model. These results highlight the value of cuproptosis-related lncRNAs on prognosis for HCC patients and provide insight into molecular and immune features underlying cuproptosis-related lncRNAs, which might play an important role in patient management and immunotherapy.

KEYWORDS

hepatocellular carcinoma, cuproptosis, prognostic signature, tumor microenvironment, immunotherapy

Introduction

Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related death and ranks sixth among all cancers (Forner et al., 2018). Curative therapeutic approaches including liver transplantation, resection, or ablation could only be applied to patients with early-stage disease, while most patients fail to meet the criteria and have a poor prognosis (Lau et al., 2001). The mortality of HCC roughly matches its incidence because of its aggressive nature and limited treatment options (Sung et al., 2021). Thus, uncovering novel therapeutic targets and prognostic factors is an urgent need to improve treatment efficiency and prognosis.

Copper is a basic trace element for human beings, which is involved in various biological processes such as mitochondrial respiration, oxidative stress, and cytotoxicity (Ruiz et al., 2021; Ge et al., 2022). As to cancer, several studies have reported that the Cu concentration in cancer is much higher than that in normal tissues (Blockhuys et al., 2017; Ge et al., 2022). The dysregulation of copper homeostasis has been related to proliferation, angiogenesis, and metastasis, which indicates copper might play a role in tumorigenesis and tumor progression (Babak and Ahn, 2021; Shanbhag et al., 2021; Oliveri, 2022). Moreover, it also had been reported that copper might play a part in immunity and affect the expression levels of programmed death-ligand 1 (PD-L1) (Jones, 1984; Voli et al., 2020). Recently, Tsvetkov et al. found a novel form of cell death termed cuproptosis. The study revealed that increment of copper in cells could induce the aggregation of lipoylated dihydrolipoamide S-acetyltransferase (DLAT) and then affect mitochondrial tricarboxylic acid (TCA) cycle, which finally leads to proteotoxic stress and cell death (Tsvetkov et al., 2022). Metabolic reprogramming of the tricarboxylic acid (TCA) cycle usually comes with the progression of HCC, promoting tumor survival and proliferation in the context of nutrient deprivation and hypoxia (Todisco et al., 2019). So cuproptosis-related genes might be involved in tumor development and progression.

Long noncoding RNAs (lncRNAs) are a type of transcripts longer than 200 nucleotides lacking protein-coding capacity (Clark et al., 2012). And they are closely related to the development of oncogenesis, progression, metastasis, and prognosis in various tumors (Bhan et al., 2017; Wong et al., 2018). However, there are few studies on cuproptosis-related lncRNAs (CRLncRNAs) in HCC patients.

The present study identified cuproptosis-related lncRNAs and constructed a prognostic signature from these lncRNAs, which was associated with mutation landscape, the tumor microenvironment, and immunotherapy response of HCC patients. Gene enrichment analysis was also carried out to explore potential mechanisms.

Materials and methods

Data collection and processing

First, RNA-sequence data (50 normal samples and 374 tumoral samples), gene mutation data (n = 364), and clinical data (n = 377) of HCC patients were derived from the TCGA database (https://portal.gdc.cancer.gov/). The transcripts/genes expression abundance are estimated by STAR and RSEM. After eliminating the normal samples, 19895 mRNA and 16773 lncRNAs were identified in LIHC data using annotation of GENCODE project (v22) (Frankish et al., 2019). We then screened 19 cuproptosis-related genes from previous literature (Supplementary Table S1), and expression data were obtained for these genes in TCGA LIHC (Supplementary Table S2). 977 CRLncRNAs whose expression was correlated to cuproptosis-related genes were identified by Pearson correlation analysis ($|R2| > 0.4, p < 0.001$). Clinicopathological factors, including age, gender, TNM stage, pathologic grade and complete survival information were also extracted. Disease-free survival (DFS) was obtained from the previous study (Liu et al., 2018). Samples with survival time < 30 days were excluded. Finally, 343 cases with gene transcriptome data and clinical data were obtained for analysis.

Development of the cuproptosis-related lncRNAs signature

A total of 343 samples with the survival data and expression data were randomly allocated to the training sets (n = 241) and validation set (n = 102) in a 7:3 ratio. Univariate Cox regression analysis was performed to screen CRLncRNAs associated with prognosis in the training set. Then these lncRNAs were analyzed by the least absolute shrinkage and selection operator (LASSO) algorithm with 1000 cycles for the best subset of prognostic lncRNAs, and a cuproptosis-related lncRNAs signature (CRlncSig) was constructed. Risk score = $\sum$ (coef (β)*EXP_{β}), where β represents each selected lncRNA. Patients were assigned to high-risk and low-risk groups with the median risk score as the cutoff value. Kaplan-Meier survival analysis was performed to validate the clinical relevance between the two groups. The ROC curve and c-index were used to assess the predictive power of the model. Stratified analysis was conducted to further assess the additional prognostic value of CRlncSig.

Validation of the CRlncSig

Baseline characteristics were checked between training sets and validation set. The patients in the validation set were grouped

with the same method in the training set and validated using Kaplan-Meier survival analysis and risk plot.

The independently prognostic value of CRlncSig

Univariate and multivariate Cox regression analyses were used to confirm predictive power. Additionally, the correlation between CRlncSig and clinical characteristics was explored by chi-square test using TCGA.

Construction of nomogram

Risk score combined with the clinicopathological factor of age, gender, grade, and stage were used to construct a nomogram to predict the 1-, 3-, and 5-year survival of HCC patients. The calibration curve was used to test agreement between the actual overall survival (OS) and those predicted by the nomogram.

Functional enrichment analysis of risk score-associated genes

Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed to identify significant module using the "clusterprofiler" R package with adjusted p value <0.05.

Prognostic analysis of the tumor mutational burden

Somatic mutations were analyzed by "maftools" R package and illustrated in waterfall plots. TMB of each sample was calculated according to the definition of the total number of variations per million bases via Perl script (version: 5.30.2) (https://www.perl.org/). According to the median value of the TMB, patients were divided into the high-TMB group and the low-TMB group. Then we merged the mutation data with survival information and performed the Kaplan-Meier survival analysis for the two groups.

Immune-related analysis of CRlncSig

We used the single-sample gene set enrichment analysis (ssGSEA) algorithm via R packages (limma, GSVA and GSEABase) to assess immune function between high- and low-risk groups based on CRlncSig (Hänzelmann et al., 2013). ESTIMATE and CIBERSORT algorithm was performed to assess the proportions of components in the tumor microenvironment (TME) and immune cell infiltration (Yoshihara et al., 2013; Newman et al., 2015). Then we explored the relationship between the expression level of immune checkpoint genes and the two groups. Immunophenoscore (IPS) was further obtained from The Cancer Immunome Atlas (https://tcia.at/home) and used to assess the clinical response to immunotherapy between the two groups (Charoentong et al., 2017).

Significance of the CRlncSig in drug sensitivity

Fifty percent maximum inhibitory concentration (IC50) values of different groups for various antitumor drugs recommended for hepatocellular carcinoma were calculated via "pRRophetic" and "ggplot2" R package. The IC50 was then compared between low- and high-risk groups by Wilcoxon signed-rank test.

Statistical analysis

R version 4.0.2 was used to analyze the data and visualize the results. Clinicopathological parameters were compared using t-tests and chi-square tests. Spearman or Pearson correlation coefficients were performed to evaluate relationships between variables. Survival curves were created by the Kaplan-Meier method and compared by log-rank test. Univariable and multivariable analyses were performed using Cox regression models to determine prognostic factors for DFS and OS. Statistical significance was set at $p < 0.05$.

Results

Construction of the CRLncRNAs predictive signature

The flow chart of this study is shown in Figure 1A. We curated a catalog of 19 cuproptosis-related genes from previous reports (Supplementary Table S1) (Huang et al., 2015; Deigendesch et al., 2018; Tsvetkov et al., 2022). Functional annotations are shown in Supplementary Table S1. Fifteen of these genes showed significant differences ($p < 0.05$) between tumor and normal tissues in LIHC patients from TCGA (Figure 1B). The correlation between cuproptosis-related genes and prognosis of HCC patients is shown in Supplementary Figure S1.

We identified 977 CRLncRNAs (Figure 1C, Supplementary Table S2). Supplementary Table S3 showed

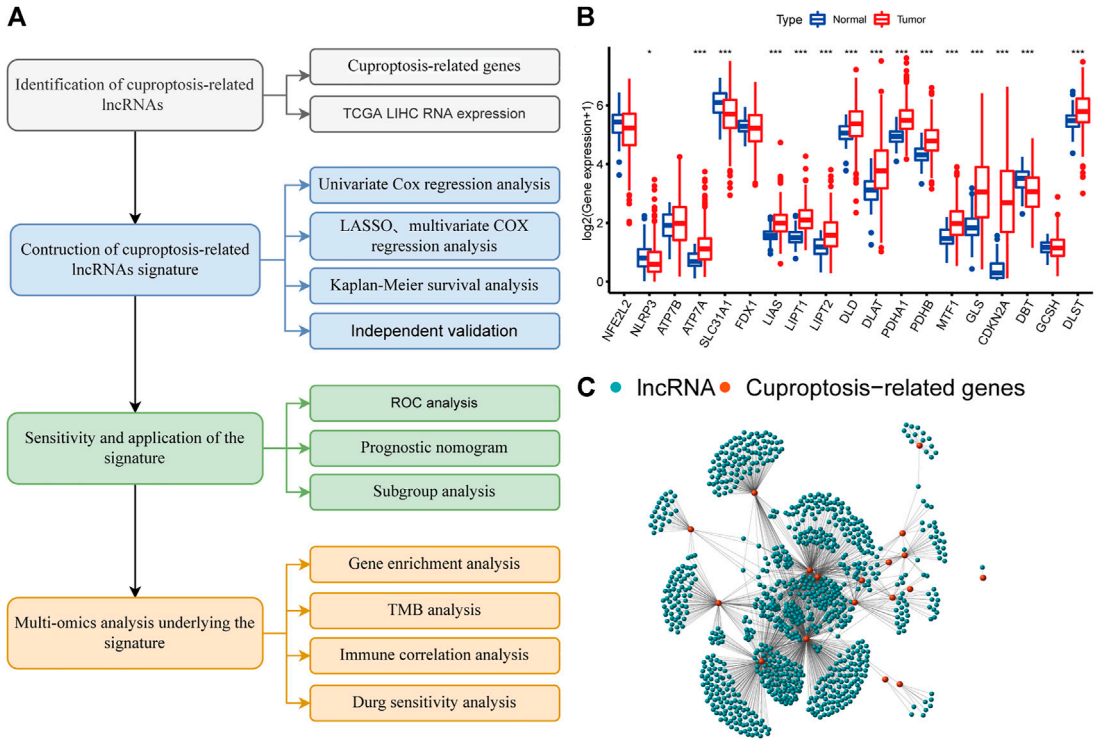


FIGURE 1
A screen of the differentially expressed cuproptosis-associated lncRNAs in hepatocellular carcinoma (HCC) **(A)** Flowchart of the present research. **(B)** Differential expression of cuproptosis-related genes in normal and HCC tissue **(C)** Network graph of cuproptosis-associated lncRNAs.

the correlation result between lncRNAS and cuproptosis genes. Then 211 CRLncRNAs were found as prognostic factors using univariate COX analysis (Supplementary Table S4). Subsequently, we performed LASSO Cox regression intending to reduce the risk of over-fitting and 13 robust genes were obtained (Figures 2A,B). Multivariate Cox regression was applied to analyze the thirteen genes and nine of them (GSEC, AL158166.1, AC005479.2, AL365361.1, AC026412.3, AL031985.3, LINC00426, AC009974.2, AC245060.7) were then used to construct a prognostic signature for HCC. Supplementary Figure S2 showed the correlation between cuproptosis-related genes and their associated lncRNAs.

Correlation between CRlncSig and prognosis of HCC patients

The coefficients of the nine CRlncSig were used to assess the scores for each patient. The risk score was calculated as follows: Risk score = (0.319,888 × expression value of GSEC) + (0.332,438 × expression value of AL158166.1) + (0.40166 × expression value of AC005479.2) + (-0.59091 × expression value of AL365361.1) + (0.764,221 × expression value of AC026412.3) + (0.457,035 × expression value of AL031985.3) + (-0.95334 × expression value of LINC00426) + (-1.61518 × expression value of AC009974.2) + (0.958,349 × expression value of AC245060.7). Then patients were assigned to low- and high-risk groups according to the median value of the risk score. Seventy percent of the 343 patients were randomly divided into the training group and the rest were in the validation group. No significant differences were found in clinical characteristics between the low- and high-groups (Supplementary Table S5). In the training set, patients in high-risk group had shorter overall survival than patients in low-risk group ($p < 0.001$, Figure 2C). This was also validated in the validation set ($p < 0.001$, Figure 2D). Next, we checked the predictive performance in disease-free survival using the dataset with DFS information (Liu et al., 2018). K-M analysis indicated significantly reduced DFS in high-risk patients ($p < 0.001$, Figure 2E). As shown in risk survival status plot, the survival of patients was inversely proportional to the risk score both in training and validation set (Figures 2F,G).

Evaluation of CRlncSig

The time-dependent ROC curve was used to assess the performance of the signature. The area under the ROC curve (AUC) of 1 year, 3 years, and 5 years of survival were 0.813, 0.789,

and 0.752, respectively (Figure 3A). The AUC of 1-year survival rate suggested that risk score (0.813) and stage (0.713) possessed a favorable prediction power (Figure 3B). The C-index of the risk score was superior to clinicopathological factors as shown in Figure 3C. The prognostic value of the risk score and other factors were evaluated with univariate and multivariate Cox regression analyses. The risk score and stage were identified as significant independent prognostic factors in both univariate Cox regression analyses (HR = 1.077, 95% CI = 1.055–1.099, $p < 0.001$ and HR = 1.804, 95% CI = 1.456-2.234, $p < 0.001$) and multivariate Cox regression analyses (HR = 1.069, 95% CI = 1.046-1.092, $p < 0.001$ and HR = 1.775, 95% CI = 1.423-2.213, $p < 0.001$) (Figures 3D,E).

Construction of nomogram

To provide a quantitative tool for clinical application, we established a nomogram with age, gender, pathological grade, stage, and risk score to predict the overall survival of patients (Figure 4A). The calibration plot showed good consistency between the actual versus predicted rates of the 1, 3, and 5-year OS (Figure 4B).

Subgroup analysis of clinicopathological variables

At last, to explore the applicability of CRlncSig, patients were assigned into groups according to age, gender, and stage. For each subgroup, patients with high-risk scores had a poor prognosis, which indicated that CRlncSig had good predictive power for all patients (Figures 4C–H).

Functional and pathway analysis

GO and KEGG analyses were performed to explore the underlying mechanisms of different prognoses between high- and low-risk groups. 1090 differentially expressed genes (DEGs) were obtained between two groups, including 947 upregulated genes and 143 downregulated genes (Supplementary Table S6). The cellular component (CC) of GO enrichment analysis indicated that DEGs were mainly enriched in “immunoglobulin complex”, and “immunoglobulin complex circulating”. Biological process (BP) showed DEGs were mainly associated with “nuclear division”, “phagocytosis, recognition”, and “humoral immune response”. While molecular function (MF) indicated DEGs were mainly concentrated in “antigen binding”, and “immunoglobulin receptor binding” (Figure 5A). According to KEGG pathway analysis, DEGs were found mainly connected with tumorigenesis and cancer progression, such as “ECM-receptor interaction”, “p53 signaling pathway”, “Central carbon metabolism in cancer”, as well as immune-related pathways, such as “HIF-1 signaling pathway” “Cytokine-cytokine receptor interaction”, (Figure 5B). These results suggested that DEGs are involved in both carcinogenesis and immune-related pathways.

Risk score-associated mutation landscape

Various basic features of somatic mutation data for low- and high-risk groups were shown in waterfall plot (Figures 5C,D). The top three mutated genes were TP53 (40%), CTNNB1 (24%) and TTN (21%) in the high-risk group, while CTNNB1 (26%), TTN (24%) and TP53 (14%) were the most common mutation genes in low-risk group. Missense mutation was the main variant classification in both groups. Then patients were divided into low- and high-TMB groups according to the median value of TMB and a significant survival difference was found between the two groups (Supplementary Figure S3A). The risk score also showed good predictive power when patients were stratified by TMB (Supplementary Figure S3B).

Immunity analysis of the risk score

To further explore the correlations between risk score and tumor immune cell infiltration, the proportions of 22 immune cell types were compared between the low- and high-risk groups with CIBERSORT algorithm. The results showed that naïve B cells, $CD8^+$ T cells (known as main immune effector cells), resting memory $CD4^+$ T cells had higher fractions in low-risk group (all $p < 0.05$) while M0 macrophages, M2 macrophages, which were known to exert immunosuppressive functions, had higher fractions in high-risk group (both $p < 0.05$) (Figure 6A). The ESTIMATE algorithms suggested a higher proportion of immune and stromal cells in the low-risk group (Figures 6B,C). Then the immune function was inferred by ssGSEA algorithm. As shown in Figure 6D, Type II IFN (IFN-γ) response, chemokine receptor (CCR), para-inflammation, T cell co-inhibition, checkpoint, T cell co-stimulation, cytolytic activity, inflammation-promoting, antigen-presenting cell (APC) coinhibition and human leukocyte antigen (HLA) were significant difference between two groups, which indicated that immune function was more active in the low-risk group. These results suggested that the signature was not only a predictive marker but also associated with immune function. Next, we explored whether levels of immune checkpoint genes were associated with risk scores. High-risk patients tended to express higher levels of 16 immune checkpoint genes, including HAVCR2, VTCN1, CD276, TNFRSF4, CD27, TNFRSF14, TNFSF4, LGALS9, CD80, TNFRSF15, CD47, HHLA2, TNFSF9, LAIR1, TNFRSF18, CD44, while low-risk

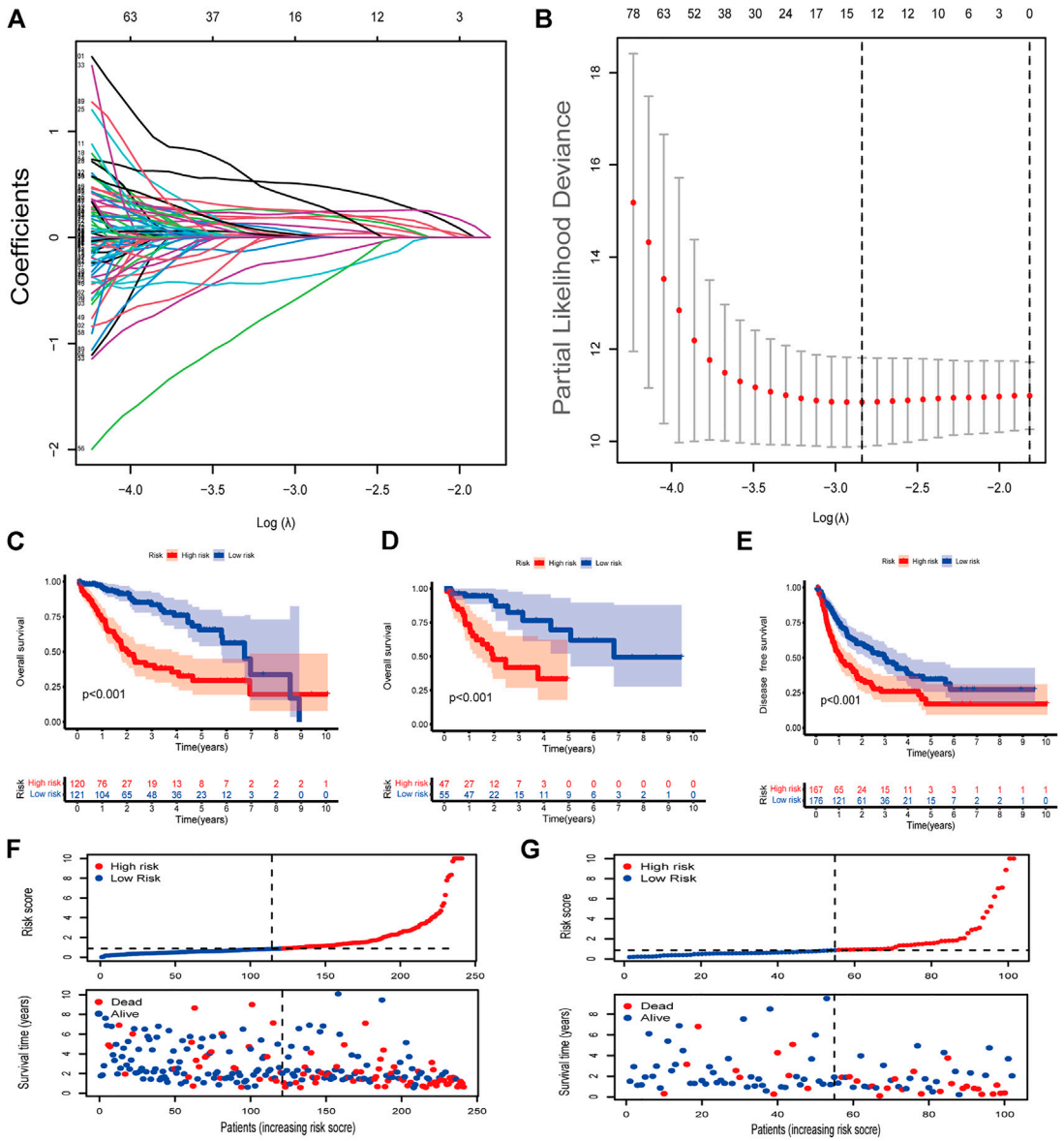


FIGURE 2
Identification of cuproptosis-associated lncRNAs with prognostic value in hepatocellular carcinoma (HCC) patients **(A,B)** LASSO Cox regression with a 10-fold cross-validation for the prognostic value of the cuproptosis-associated lncRNAs. **(C)** Kaplan-Meier analysis of the OS rate of training set patients in the high- and low-risk groups **(D)** Kaplan-Meier analysis of the OS rate of validation set patients in the high- and low-risk groups. **(E)** Kaplan-Meier analysis of the DFS rate of HCC patients in the high- and low-risk groups **(F)** Risk score distribution, survival status for patients in high- and low-risk groups from training set. **(G)** Risk score distribution, survival status for patients in high- and low-risk groups from validation set.

patients tended to express higher levels of 10 immune checkpoint genes, including LAG3, PDCD1LG2, IDO2, KIR3DL1, CD244, CD48, CD40LG, TMIGD2, CD160, CD96 (Figure 6E). To access the power of the signature for predicting the response to immunotherapy, immunophenoscore (IPS) calculated and patients in low-risk group had a higher IPS, suggesting that patients in this group might have a better response to immunotherapy (Figures 6F–I).

Drug response features underlying the CRLncRNAs

In addition to immunotherapy, we also explored the association between the risk score and the efficacy of targeted therapy and chemotherapy for HCC patients. The results suggested that the IC50 of trametinib, talazoparib was

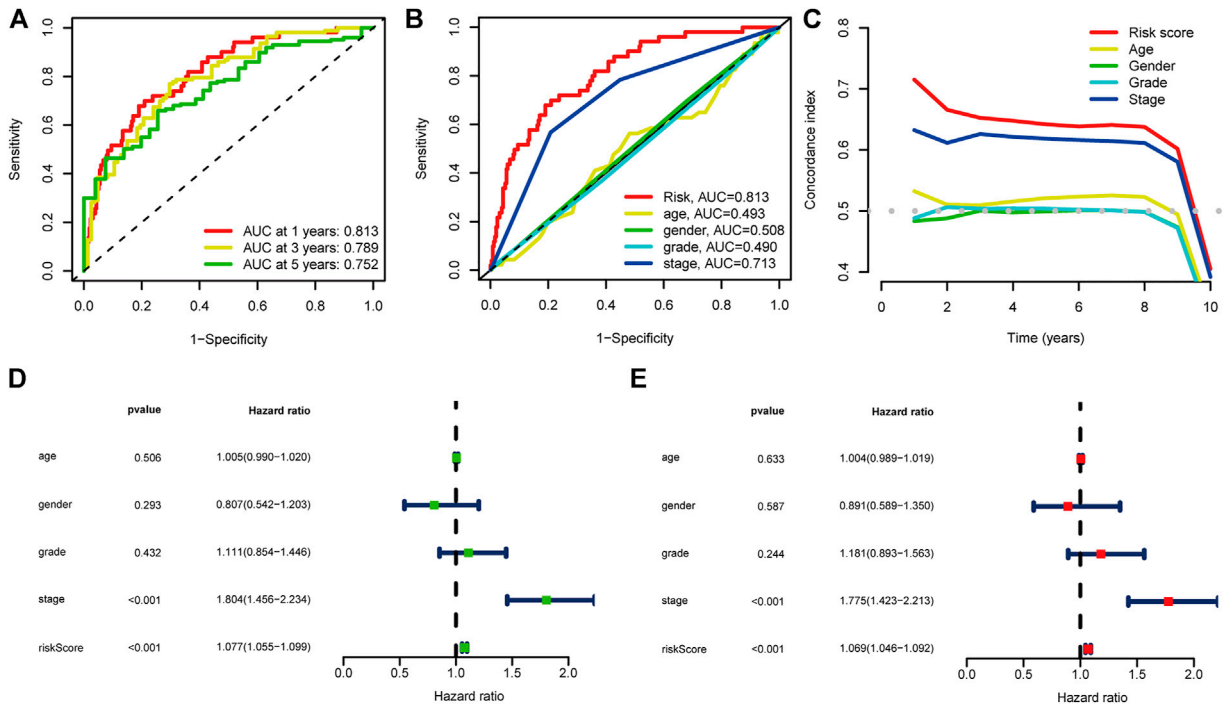


FIGURE 3
Evaluation of the prognostic cuproptosis-associated lncRNAs signature **(A)** ROC curve and AUCs at 1-year, 3-years and 5-years survival for the predictive signature. **(B)** The ROC curve of the risk score and clinicopathological variables **(C)** C index of the risk score and clinicopathological variables. **(D)** Forest plot for univariate Cox regression analysis. **(E)** Forest plot for multivariate Cox regression analysis.

positively correlated with risk score and the IC50 of 5-fluorouracil, doxorubicin, gemcitabine, mitomycin C, paclitaxel, sorafenib, sunitinib, tipifarnib, tivozanib, vinorelbine was negatively correlated with risk score (Supplementary Figure S4), which helps explore individualized treatment strategy HCC patients.

Discussion

HCC is the third leading cause of cancer-related death worldwide. The high molecular and clinical heterogeneity of HCC results in inefficient treatments and poor prognosis (Wörns and Galle, 2014). Integrating multiple biomarkers and clinical features into a single model could improve the accuracy of prediction and help formulate individualized treatment plans when compared with a single biomarker. In the present study, we identified CRLncRNAs and constructed a prognostic signature, which was associated with mutation landscape, the tumor microenvironment, and immunotherapy response of HCC patients. We also explore potential mechanisms through gene enrichment analysis.

We identified 211 CRLncRNAs associated with the overall survival of HCC patients via univariate regressions analysis. Then nine lncRNAs were conformed and developed lncRNA signature for prognostic prediction. Different kinds of predictive lncRNA signatures for HCC patients have been reported in previous studies (Huang et al., 2021; Li et al., 2021; Wang et al., 2021; Yang et al., 2021). Li et al. reported an eight m6A-related lncRNA signature with AUC of 0.633, 0.651, and 0.638 at 1-, 3–5-year (Li et al., 2021). While the highest AUC of the immune- and ferroptosis-related lncRNA signature in 5 years was 0.761 in the study by Huang (Huang et al., 2021). In our study, the lowest AUC in 5 years is 0.753, which indicates this CRLncRNAs signature has strong predictive power.

The critical contribution of this study is to demonstrate the relationship between CRLncRNAs signature and tumor microenvironment. Notably, it is worth noting that TME not only plays a vital role in the development of tumors but also has an important impact on immunotherapy response and overall survival (Hinshaw and Shevde, 2019; Fane and Weeraratna, 2020; Petitprez et al., 2020). Functional enrichment analysis showed that CRLncRNAs were mainly related to cytokine-cytokine receptor interaction, the phosphatidylinositol 3-kinases/protein kinase B (PI3K-AKT) signaling pathway and immune pathways. Cytokines are major regulators of the innate and adaptive immune systems that allow cells of the immune systems to communicate over short distances in

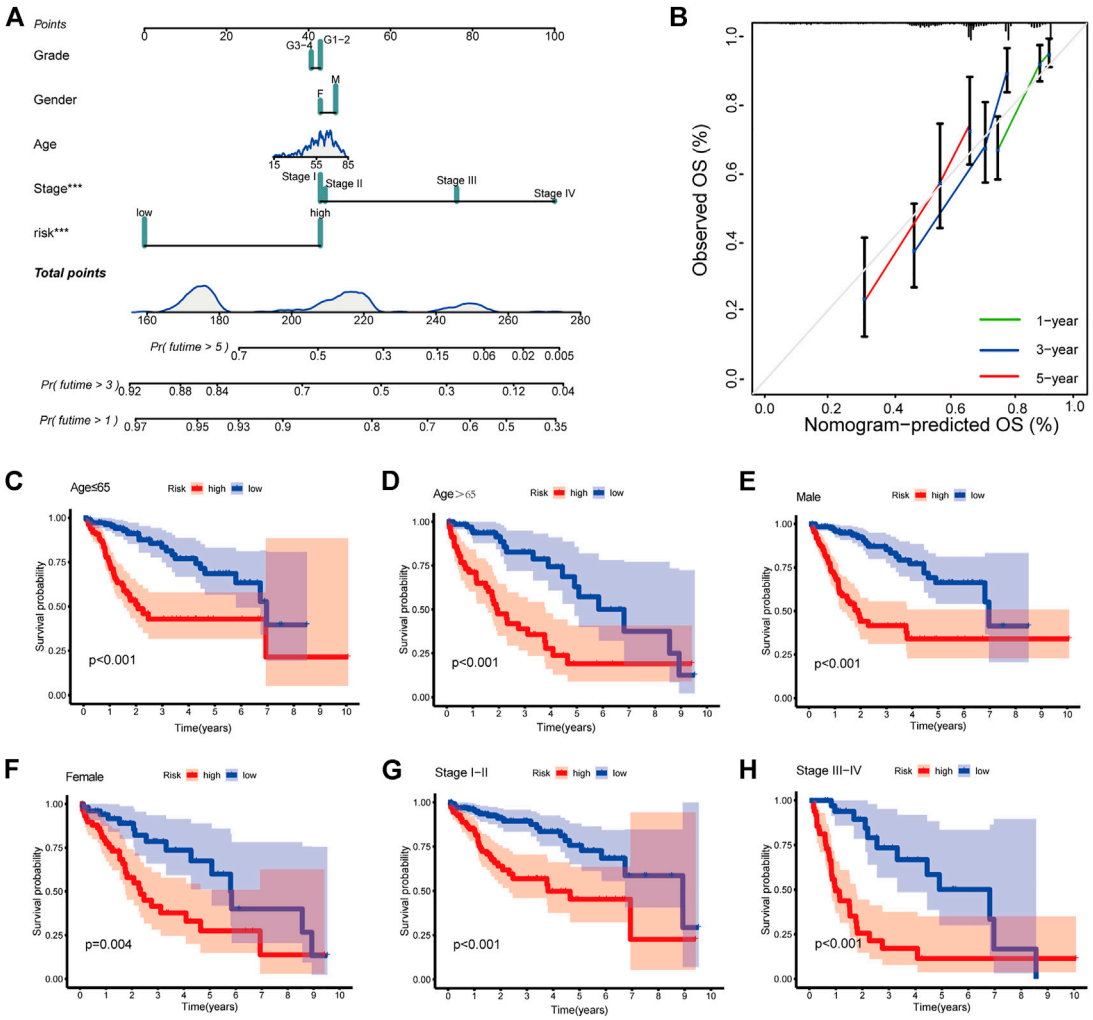


FIGURE 4
Clinical prognostic nomogram for survival prediction and subgroup analysis **(A)** A nomogram combining clinicopathological variables and risk score predicts 1, 3, and 5 years OS of HCC patients. **(B)** Calibration plots for 1-, 3-, and 5-years survival predictions **(C–H)** Subgroup survival analysis in the high- and low-risk groups, **(C)** Age ≤65 **(D)** Age > 65 **(E)** Male patients **(F)** Female patients **(G)** Stage I-II **(H)** Stage III-IV.

paracrine and autocrinefashion (Waldmann, 2018). Cytokine and cytokine receptor interaction networks were regarded as crucial effects on inflammation and oncogenesis (Dranoff, 2004). Cytokines and its receptors, such as tumor necrosis factor and interleukin 6, were important factors in the development of HCC and affected the immunotherapy effect (Kern et al., 2018; Derynck et al., 2021). PI3K-AKT signaling pathway was a classical intracellular signaling receptor to react extracellular stimulators. The PI3K/AKT pathway was dysregulated in both initiation and progression of HCC (Whittaker et al., 2010). To explore whether the signature could predict the efficiency of immunotherapy for HCC patients, we first checked the expression levels of 48 immune checkpoints genes and found that more than half of these genes were related to the risk score. Tumor immune microenvironment was also evaluated between two groups. Patients with high-risk scores had lower proportions of $CD8^+$ T cells and higher proportions of M0 macrophages and M2 macrophages, which indicated the roles of CRLncRNAs in regulating the tumor microenvironment. As we know, $CD8^+$ T cells are the main effectors in antitumor immunotherapy while M2 macrophages, working as immunosuppressive cells, promote tumor growth and invasion (Pitt et al., 2016). Patients in high-risk group are more likely to be “cold” tumors characterized by resistance to immune checkpoint therapy. IPS, derived from four major gene categories, could work as a superior predictor for immunotherapy (Charoentong et al., 2017). Then we calculated IPS to predict immunotherapy response, patients in low-risk group had a higher IPS, suggesting that patients in this group might respond better to immunotherapy. This is consistent with result of the tumor immune microenvironment analysis.

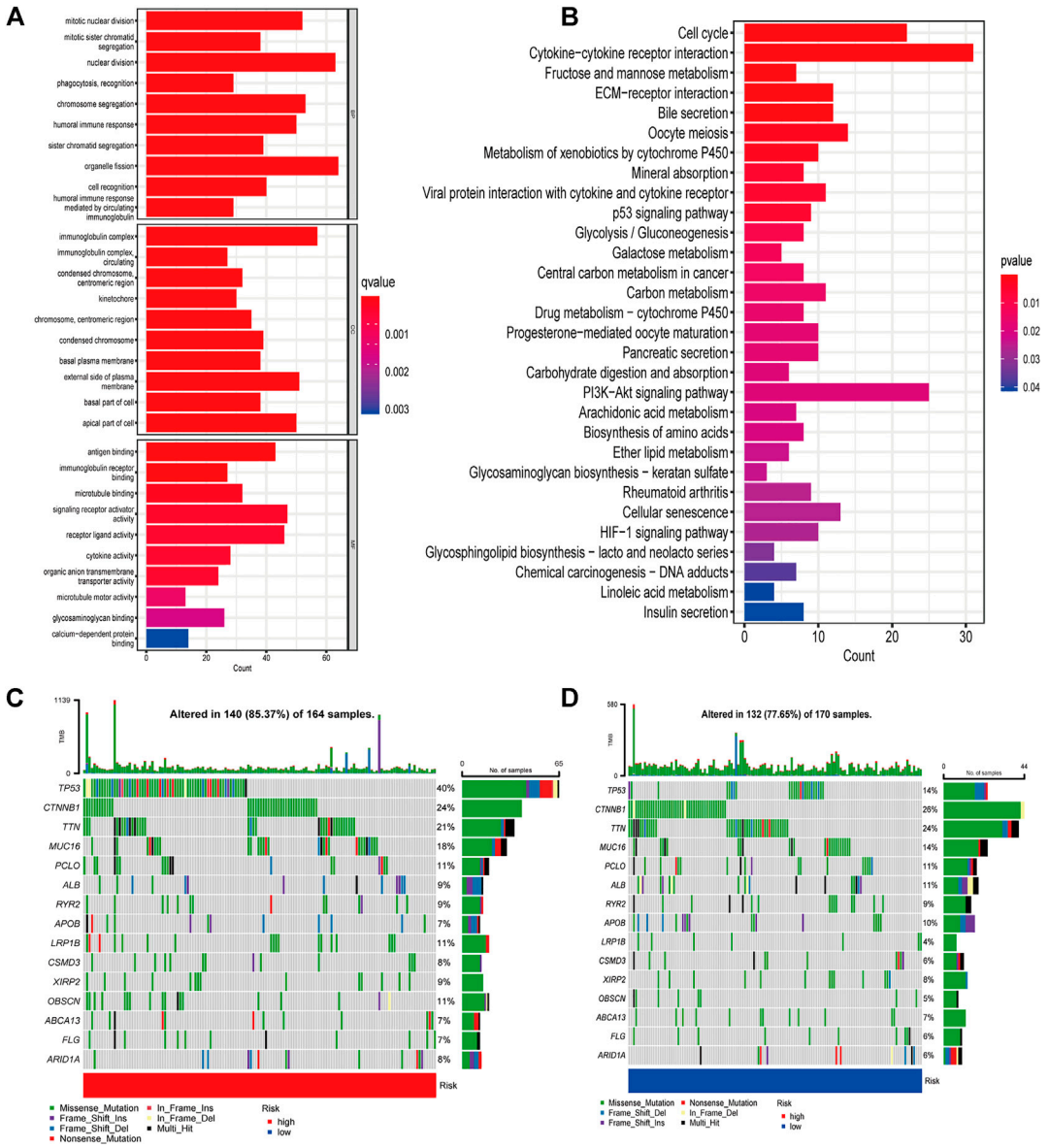


FIGURE 5
Gene enrichment and TMB in high- and low-risk groups **(A)** GO enrichment analysis **(B)** KEGG enrichment analysis **(C)** Waterfall plots displaying the mutation landscapes of the low-risk group. **(D)** Waterfall plots displaying the mutation landscapes of the high-risk group. TMB, tumor mutation burden; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, Biological process; CC, Cellular component; MF Molecular function.

Recent studies show that cuproptosis is a promising new target for cancer treatment. Copper ionophores have shown promising applications in overcoming drug resistance of cancer cells and targeting cancer stem cells. This is due to the intrinsic selectivity of copper ionophores in preferential induction of cancer cell clusters compared with normal cells (Oliveri, 2022). Another study by Voli et al. showed that copper supplementation promotes PDL1 expression and intratumor copper levels might enhance immunotherapy response (Voli et al., 2020). Our study and previous studies indicate that copper plays an important role in antitumor treatment and immunotherapy. FDX1 is the key regulators of copper ionophore–induced cell death, which encodes a reductase known to reduce Cu^{2+} to its more toxic form, Cu^{1+}, and is adirect target of elesclomol (Tsvetkov et al., 2019; Tsvetkov et al., 2022). Recent pan-cancer analysis revealed that FDX1 could be a novel biomarker in the prognosis and immunotherapy in human tumors, which could provide a basis for drug use in certain tumors (Ma et al., 2022; Zhang et al., 2022).

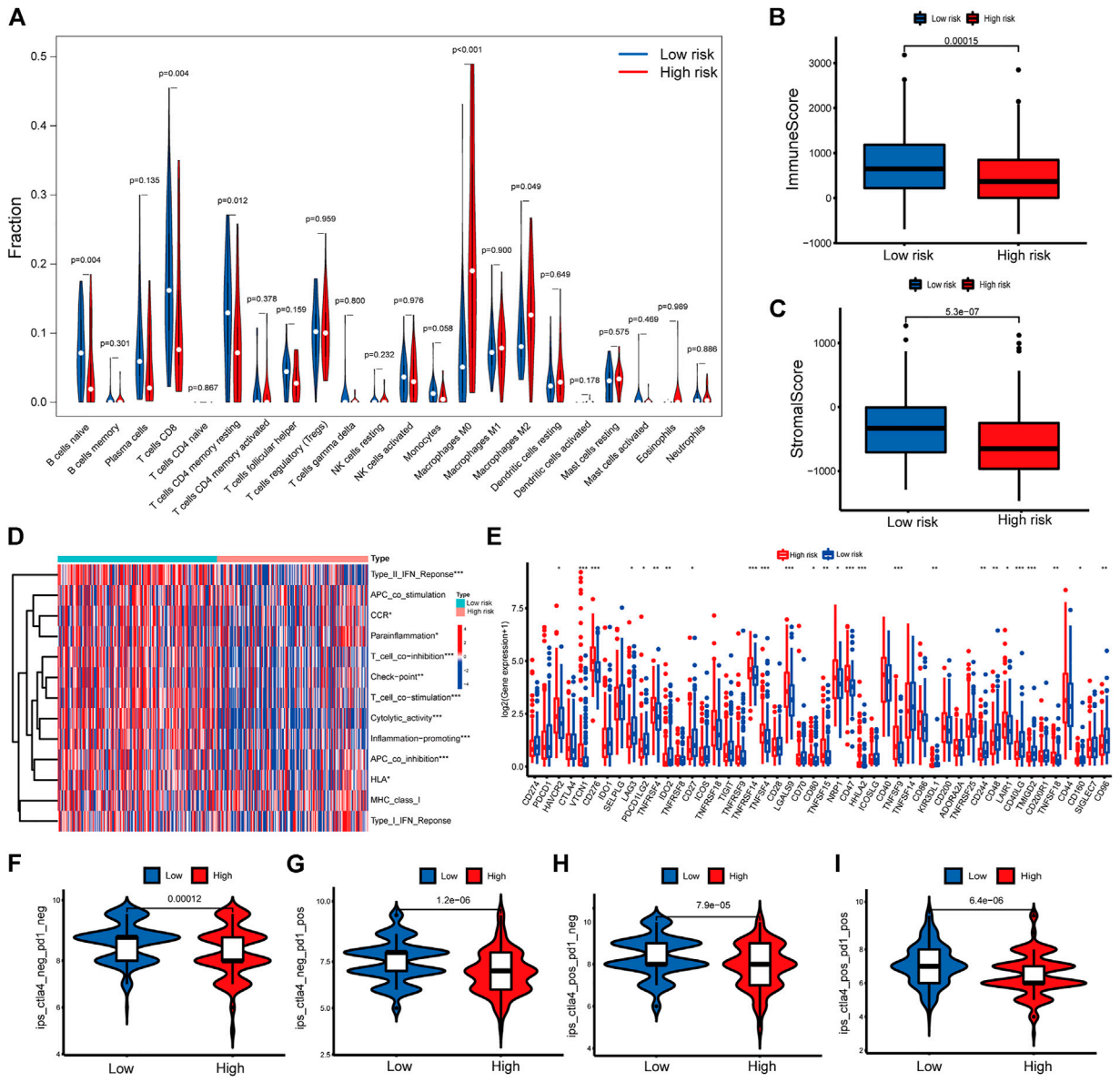


FIGURE 6
Immune related analysis in high- and low-risk groups **(A)** Differences in the infiltration of immune cells between the high- and low-risk groups. **(B–C)** Comparison of immune score **(B)**, and stromal score **(C)** between the high- and low-risk groups **(D)** The correlation between the signature and 13 immune-related functions. **(E)** Differential expression of immune checkpoint genes between the high- and low-risk groups **(F–I)** IPS values of patients categorized according to risk score of four subtypes [IPS-CTLA4-neg-PD1-neg **(F)**, IPS-CTLA4-neg-PD1-pos **(G)**, IPS-CTLA4-pos-PD1-neg **(H)**, IPS-CTLA4-pos-PD1-pos **(I)**]. IPS, Immunophenoscore. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

The current study had several limitations. First, we constructed and validate the prognostic model with a single retrospective data source. Second, some well-known prognostic factors for HCC such as AFP and microvessel invasion were not involved in the nomogram because of incomplete data for these parameters. Thus, a prospective study is needed to verify the predictive value of the signature. In addition, functional biological experiments should be carried out to further validate the results.

In summary, the cuproptosis-related lncRNA signature could effectively predict the prognosis and immunotherapy response of HCC patients. Immune analysis verified the association between the risk score and tumor microenvironment. Thus, our results offer a reasonable explanation for the distinct prognoses of patients and provide a rationale for exploring biomarkers and antitumor treatment strategies.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The Cancer Genome Atlas (https://portal.gdc.cancer.gov/).

Author contributions

HZ analyzed data and wrote the manuscript. SC, FM performed study concept and design. KW, JF analyzed data, interpreted results, and helped to write the manuscript. SC revised the manuscript and made final approval of the version. All authors have read and agreed to the published version of the manuscript.

Funding

The work was supported by National Natural Science Foundation of China (No: 82002480).

Acknowledgments

We thank the contributor of TCGA (https://cancergenome.nih.gov/) for sharing the LIHC dataset on open access. We are grateful to all the participants who have enabled this research.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.972212/full#supplementary-material

References

Babak, M. V., and Ahn, D. (2021). Modulation of intracellular copper levels as the mechanism of action of anticancer copper complexes: Clinical relevance. *Biomedicines* 9 (8), 852. doi:10.3390/biomedicines9080852

Bhan, A., Soleimani, M., and Mandal, S. S. (2017). Long noncoding RNA and cancer: A new paradigm. *Cancer Res.* 77 (15), 3965–3981. doi:10.1158/0008-5472.CAN-16-2634

Blockhuys, S., Celauro, E., Hildesjö, C., Feizi, A., Stål, O., Fierro-González, J. C., et al. (2017). Defining the human copper proteome and analysis of its expression variation in cancers. *Metallomics* 9 (2), 112–123. doi:10.1039/c6mt00202a

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

Clark, M. B., Johnston, R. L., Inostroza-Ponta, M., Fox, A. H., Fortini, E., Moscato, P., et al. (2012). Genome-wide analysis of long noncoding RNA stability. *Genome Res.* 22 (5), 885–898. doi:10.1101/gr.131037.111

Deigendesch, N., Zychlinsky, A., and Meissner, F. (2018). Copper regulates the canonical NLRP3 inflammasome. *J. Immunol.* 200 (5), 1607–1617. doi:10.4049/jimmunol.1700712

Derynck, R., Turley, S. J., and Akhurst, R. J. (2021). TGFβ biology in cancer progression and immunotherapy. *Nat. Rev. Clin. Oncol.* 18 (1), 9–34. doi:10.1038/s41571-020-0403-1

Dranoff, G. (2004). Cytokines in cancer pathogenesis and cancer therapy. *Nat. Rev. Cancer* 4 (1), 11–22. doi:10.1038/nrc1252

Fane, M., and Weeraratna, A. T. (2020). How the ageing microenvironment influences tumour progression. *Nat. Rev. Cancer* 20 (2), 89–106. doi:10.1038/s41568-019-0222-9

Forner, A., Reig, M., and Bruix, J. (2018). Hepatocellular carcinoma. *Lancet* 391 (10127), 1301–1314. doi:10.1016/S0140-6736(18)30010-2

Frankish, A., Diekhans, M., Ferreira, A. M., Johnson, R., Jungreis, I., Loveland, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 47 (D1), D766–d773. doi:10.1093/nar/gky955

Ge, E. J., Bush, A. I., Casini, A., Cobine, P. A., Cross, J. R., DeNicola, G. M., et al. (2022). Connecting copper and cancer: From transition metal signalling to metalloplasia. *Nat. Rev. Cancer* 22 (2), 102–113. doi:10.1038/s41568-021-00417-2

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

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

Huang, H. C., Hong, L., Chang, P., Zhang, J., Lu, S. Y., Zheng, B. W., et al. (2015). Chitooligosaccharides attenuate Cu2+-induced cellular oxidative damage and cell apoptosis involving Nrf2 activation. *Neurotox. Res.* 27 (4), 411–420. doi:10.1007/s12640-014-9512-x

Huang, A., Li, T., Xie, X., and Xia, J. (2021). Computational identification of immune- and ferroptosis-related LncRNA signature for prognosis of hepatocellular carcinoma. *Front. Mol. Biosci.* 8, 759173. doi:10.3389/fmolb.2021.759173

Jones, D. G. (1984). Effects of dietary copper depletion on acute and delayed inflammatory responses in mice. *Res. Vet. Sci.* 37 (2), 205–210. doi:10.1016/s0034-5288(18)31906-4

Kern, L., Mittenbühler, M. J., Vesting, A. J., Ostermann, A. L., Wunderlich, C. M., and Wunderlich, F. T. (2018). Obesity-induced TNFα and IL-6 signaling: The missing link between obesity and inflammation-driven liver and colorectal cancers. *Cancers* 11 (1), E24. doi:10.3390/cancers11010024

Lau, W. Y., Leung, T. W., Lai, B. S., Liew, C. T., Ho, S. K., Yu, S. C., et al. (2001). Preoperative systemic chemoimmunotherapy and sequential resection for unresectable hepatocellular carcinoma. *Ann. Surg.* 233 (2), 236–241. doi:10.1097/00000658-200102000-00013

Li, L., Xie, R., and Lu, G. (2021). Identification of m6A methyltransferase-related lncRNA signature for predicting immunotherapy and prognosis in patients with hepatocellular carcinoma. *Biosci. Rep.* 41 (6), BSR20210760. doi:10.1042/BSR20210760

Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D., et al. (2018). An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 173 (2), 400–416.e11. doi:10.1016/j.cell.2018.02.052

Ma, S., Hu, D., Zhu, P., Zhang, B., Zhang, Y., Zhao, X., et al. (2022). Pan-cancer analyses confirmed the cuproptosis-related gene FDX1 as an immunotherapy predictor and prognostic biomarker. *Front. Genet.* 13, 923737. doi:10.3389/fgene.2022.923737

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

Oliveri, V. (2022). Selective targeting of cancer cells by copper ionophores: An overview. *Front. Mol. Biosci.* 9, 841814. doi:10.3389/fmolb.2022.841814

Petitprez, F., Meylan, M., de Reyniès, A., Sautès-Fridman, C., and Fridman, W. H. (2020). The tumor microenvironment in the response to immune checkpoint blockade therapies. *Front. Immunol.* 11, 784. doi:10.3389/fimmu.2020.00784

Pitt, J. M., Marabelle, A., Eggermont, A., Soria, J. C., Kroemer, G., and Zitvogel, L. (2016). Targeting the tumor microenvironment: Removing obstruction to anticancer immune responses and immunotherapy. *Ann. Oncol.* 27 (8), 1482–1492. doi:10.1093/annonc/mdw168

Ruiz, L. M., Libedinsky, A., and Elorza, A. A. (2021). Role of copper on mitochondrial function and metabolism. *Front. Mol. Biosci.* 8, 711227. doi:10.3389/fmolb.2021.711227

Shanbhag, V. C., Gudekar, N., Jasmer, K., Papageorgiou, C., Singh, K., and Petris, M. J. (2021). Copper metabolism as a unique vulnerability in cancer. *Biochim. Biophys. Acta. Mol. Cell Res.* 1868 (2), 118893. doi:10.1016/j.bbamcr.2020.118893

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

Todisco, S., Convertini, P., Iacobazzi, V., and Infantino, V. (2019). TCA cycle rewiring as emerging metabolic signature of hepatocellular carcinoma. *Cancers (Basel)* 12 (1), E68. doi:10.3390/cancers12010068

Tsvetkov, P., Detappe, A., Cai, K., Keys, H. R., Brune, Z., Ying, W., et al. (2019). Mitochondrial metabolism promotes adaptation to proteotoxic stress. *Nat. Chem. Biol.* 15 (7), 681–689. doi:10.1038/s41589-019-0291-9

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science* 375 (6586), 1254–1261. doi:10.1126/science.abf0529

Voli, F., Valli, E., Lerra, L., Kimpton, K., Saletta, F., Giorgi, F. M., et al. (2020). Intratumoral copper modulates PD-L1 Expression and influences tumor immune evasion. *Cancer Res.* 80 (19), 4129–4144. doi:10.1158/0008-5472.CAN-20-0471

Waldmann, T. A. (2018). Cytokines in cancer immunotherapy. *Cold Spring Harb. Perspect. Biol.* 10 (12), a028472. doi:10.1101/cshperspect.a028472

Wang, Y., Ge, F., Sharma, A., Rudan, O., Setiawan, M. F., Gonzalez-Carmona, M. A., et al. (2021). Immunoautophagy-related long noncoding RNA (IAR-lncRNA) signature predicts survival in hepatocellular carcinoma. *Biol. (Basel)* 10 (12), 1301. doi:10.3390/biology10121301

Whittaker, S., Marais, R., and Zhu, A. X. (2010). The role of signaling pathways in the development and treatment of hepatocellular carcinoma. *Oncogene* 29 (36), 4989–5005. doi:10.1038/onc.2010.236

Wong, C. M., Tsang, F. H., and Ng, I. O. (2018). Non-coding RNAs in hepatocellular carcinoma: molecular functions and pathological implications. *Nat. Rev. Gastroenterol. Hepatol.* 15 (3), 137–151. doi:10.1038/nrgastro.2017.169

Wörns, M. A., and Galle, P. R. (2014). HCC therapies--lessons learned. *Nat. Rev. Gastroenterol. Hepatol.* 11 (7), 447–452. doi:10.1038/nrgastro.2014.10

Yang, S., Zhou, Y., Zhang, X., Wang, L., Fu, J., Zhao, X., et al. (2021). The prognostic value of an autophagy-related lncRNA signature in hepatocellular carcinoma. *BMC Bioinforma.* 22 (1), 217. doi:10.1186/s12859-021-04123-6

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

Zhang, Q., Liu, N., Wu, D., Xu, Z., Wang, Y., and Wang, P. (2022). Pan-cancer analysis reveals the value of FDX1 as a novel biomarker in the prognosis and immunotherapy in human tumors. *bioRxiv*.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 27 September 2022
DOI 10.3389/fgene.2022.990153

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Chongwu Li,
Tongji University, China
Lei Lei Liang,
Chinese Academy of Medical Sciences and Peking Union Medical College, China

*CORRESPONDENCE
Xixian Ke,
kexixian86.88@163.com
Yongxiang Song,
songtang2004@163.com

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 09 July 2022
ACCEPTED 01 September 2022
PUBLISHED 27 September 2022

CITATION
Liu Y, Zhang X, Cheng X, Luo Q, Yu M, Long K, Qu W, Tang Y, Gong M, Liang L, Ke X and Song Y (2022), Characterization of fatty acid metabolism-related lncRNAs in lung adenocarcinoma identifying potential novel prognostic targets. *Front. Genet.* 13:990153. doi: 10.3389/fgene.2022.990153

Characterization of fatty acid metabolism-related lncRNAs in lung adenocarcinoma identifying potential novel prognostic targets

Yang Liu†, Xingshu Zhang†, Xuechao Cheng, Qian Luo, Mingyang Yu, Kaijun Long, Wendong Qu, Yang Tang, Ming Gong, Lubiao Liang, Xixian Ke* and Yongxiang Song*

Department of Thoracic Surgery, The Affiliated Hospital of Zunyi Medical University, Zunyi, Guizhou, China

Lung adenocarcinoma (LUAD), a malignant respiratory tumor with an extremely poor prognosis, has troubled the medical community all over the world. According to recent studies, fatty acid metabolism (FAM) and long non-coding RNAs (lncRNAs) regulation have shown exciting results in tumor therapy. In this study, the original LUAD patient data was obtained from the TCGA database, and 12 FAM-related lncRNAs (AL390755.1, AC105020.6, TMPO-AS1, AC016737.2, AC127070.2, LINC01281, AL589986.2, GAS6-DT, AC078993.1, LINC02198, AC007032.1, and AL021026.1) that were highly related to the progression of LUAD were finally identified through bioinformatics analysis, and a risk score model for clinical reference was constructed. The window explores the immunology and molecular mechanism of LUAD, aiming to shed the hoping light on LUAD treatment.

KEYWORDS

lung adenocarcinoma, long non-coding RNA, biomarker, fatty acid metabolism, immune cells

Introduction

Lung cancer accounts for the largest share of cancer-related deaths worldwide (Thai et al., 2021). It is worth noting that lung adenocarcinoma (LUAD) accounts for up to 85% of lung cancers and is the most common subspecies (Pao and Girard, 2011) (Nicholson et al., 2021). Based on different molecular and pathological features, LUAD can be

Abbreviations: ceRNA, competitive endogenous RNA; FAM, fatty acid metabolism; GDSC, Genomics of Drug Sensitivity in Cancer; KEGG, Kyoto Encyclopedia of Genes and Genomes; lncRNAs, long non-coding RNAs; LUAD, Lung adenocarcinoma; OS, overall survival; PCA, principal component analysis; PCR, Polymerase Chain Reaction; RT-qPCR, Real-Time Quantitative Polymerase Chain Reaction; TIDE, Tumor Immune Dysfunction and Exclusion; TME, tumor microenvironment; t-SNE, t-distributed Stochastic Neighbor Embedding.

subdivided into various subtypes (Inamura, 2018). There are differences and connections between different subtypes, but the commonalities between them are high malignancy, poor prognosis, and greater difficulty in early diagnosis (Blandin Knight et al., 2017) (Rami-Porta et al., 2018). From the perspective of treatment, the current treatment methods for LUAD mainly include surgery (Zappa and Mousa, 2016), immunotherapy (Hellmann et al., 2018), targeted therapy (Arbour and Riely, 2017), etc. However, various treatment methods are limited by the histology of LUAD, mutated genes, and differences in clinical stages, and the prognosis of patients is often not exactly (Zappa and Mousa, 2016). Coupled with the low sensitivity of LUAD to radiotherapy (Zappa and Mousa, 2016) and the gradual emergence of resistance to targeted therapy drugs (Remon et al., 2021), we are forced to have a deeper understanding of LUAD.

Tumor cells are often in an abnormal metabolic environment, depending on the imbalance between the rapid proliferation of tumor cells and nutrient angiogenesis (Yi et al., 2018; Han et al., 2021). Modern thinking holds that tumor cells need to reprogram their metabolism to meet increased metabolic and synthetic demands in conjunction with their own growth needs, while simultaneously reducing the negative effects of oxidative stress during growth (Martínez-Reyes and Chandel, 2021). Collectively, tumor metabolic reprogramming is significant (Hanahan and Weinberg, 2011), and the change of aerobic glycolysis (Warburg effect), glutamine metabolism, and one-carbon *de novo* synthesis of fatty acids also confer the ability of tumors to rapidly progress in a relatively nutrient-stressed tumor microenvironment (TME) (Cluntun et al., 2017; Newman and Maddocks, 2017; Ashton et al., 2018; Zhao et al., 2020; Li et al., 2021a). Interestingly, changes in the metabolic level of tumor cells often lead to changes in the components of the TME, thereby having a significant impact on affecting the biological effects of other cellular components of the TME, and these changes will ultimately affect tumor progression (Dey et al., 2021) (Broadfield et al., 2021). It is worth noting that fatty acid metabolism (FAM), as one of the important pathways of the three major nutrients metabolism, can be coupled with a variety of metabolic pathways and participate in cell membrane formation, intracellular signal transduction, hormone secretion, and other processes, and is related with the disease and health state in human (Kimura et al., 2020) (Bogie et al., 2020). At the same time, the relationship between FAM and cancer progression has received increasing attention (Koundouros and Poulogiannis, 2020) (Bergers and Fendt, 2021). long non-coding RNAs (lncRNAs) are a class of RNAs with regulatory functions, and they have been extensively studied in the past decade (Ali and Grote, 2020). Previous studies have illustrated that lncRNAs are vital in cell cycle regulation (Jiang et al., 2021), metabolic regulation (Tan et al., 2021), and even the immune system (Fok et al., 2018), and have been recognized as playing a significant role in cancer progression (Wu et al., 2020).

Here, we constructed a 12 FAM related-lncRNAs signature risk model based on LUAD raw data in TCGA by bioinformatic methods. Further immunological and functional analysis indicated the possible mechanism of action of these lncRNAs in LUAD and their impact on the first immunotherapy of LUAD. And at the end of the study, polymerase chain reaction (PCR) technology was conducted to verify the expression of the screened lncRNAs.

Materials and methods

Data preparation and processing

All kinds of LUAD data were obtained from the TCGA database (http://portal.gdc.cancer.gov/) (Blum et al., 2018). With previous reports about FAM-related genes and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (Kanehisa et al., 2017), 1879 FAM-related lncRNAs were obtained by using the correlation test between the FAM-related genes and lncRNAs with R. The thresholds were set as $|cor|>0.4$ and $p < 0.001$. LUAD patients without overall survival (OS) values or whose OS was within 30 days were excluded. Four hundred and ninety samples were divided into training and testing sets randomly. 246 samples were contained in the training set, while contained 244 in the testing set.

Establishment and validation of the risk signature

With survival information, we screened the prognosis of FAM-related lncRNAs from 1879 differently expressed lncRNAs ($p < 0.05$). Univariate Cox regression analysis was used to screen lncRNAs related to survival. LASSO regression was performed by R package "glmnet" (version 4.1-3) with 10-fold cross-validation, 1,000 cycles. With the Multifactor Cox regression, a 12 FAM-related lncRNA risk model was finally built.

The risk score was calculated by the following formula:

$$Risk\ score = \sum_{k=1}^{n} Coef\,(lncRNA)^{*} \exp\left(lncRNA^{k}\right)$$

where Coef is the coefficient and exp is the expression level of lncRNA.

The mean score was regarded as a standard to distinguish LUAD subgroups.

Model performance estimation

The univariate and multivariate Cox (by "glment," "survminer," and "survival" R packages) regression analyses were developed to evaluate the independent predictive power

of risk models. The 1-,3-, and 5-year ROC curves were used to evaluate the effect of prognostic prediction. Principal component analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE) analysis were further used to verify the risk model.

Nomogram and calibration

A nomogram was established based on our risk score model and various clinical characteristics by the "rms" R package. The 1-, 3-, and 5-year OS and ROC curves were performed to illustrate the actual consistency of the model with the practical.

The investigation of the TME and immunotherapy

The mutation data was the sum and analyzed by R package maftools. The infiltration status of immune cells, TME scores, and immune checkpoints activation between two different subgroups were presented *via* CIBERSORT and ssGSEA algorithm and visualized by the "ggpubr" R package. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was also used to predict the likelihood of the immunotherapeutic response, Immunotherapeutic treatment data from the website (http://tide.dfci.harvard.edu/) (Wang et al., 2020). The data of the immune subtype was downloaded on TIMER (http://timer.comp-genomics.org/) (Li et al., 2017).

Exploration of the model in the clinical treatment

The R package "pRRophetic" was used to evaluate the therapy response of each LUAD patient on Genomics of Drug Sensitivity in Cancer (GDSC) (Yang et al., 2013). Drug sensitivity analyses are conducted online (https://discover.nci.nih.gov/cellminer/home.do).

Functional analysis

Differentially expressed genes (DEGs) between two groups were identified by using the package "limma" following the criteria (|Log2FC| > 1.0, p-value < 0.05). GO and KEGG enrichment analysis was applied using the package "clusterProfiler" in R. GSEA analysis was conducted to further screen functional pathways by using software GSEA 4.2.1 (http://www.gesa-msigdb.org/gsea/index.jsp) (Powers et al., 2018). Furthermore, the competitive endogenous RNA (ceRNA) network between lncRNAs and mRNAs was visualized by Cytoscape (version 3.6.1).

RNA extraction and real-time quantitative PCR

We extracted total RNA from the samples. We synthesized cDNA using a ServicebioRT First Strand cDNA Synthesis Kit (Applied Servicebio, China). Then, cDNA was subjected to a Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR) by the bio-rad CFX (Applied Bio-rad, China). We used b-actin mRNA as an internal reference to normalize the nine lncRNAs by the comparative Ct method. All three cell lines (H1299, A549, and BEAS-2B) were purchased from Procell. The ambient temperature was controlled at 37°C and the CO_2 concentration was 5%. The three cell lines were added to a 1640 medium containing 10% fetal bovine serum and incubated in a constant temperature incubator.

Statistical analysis

All statistical analyses were conducted in the R software (Version 4.1.1). Wilcoxon rank-sum test was used to compare the difference between the two groups. K-W test was performed to compare three or more groups. Kaplan-Meier analysis was used to evaluate the survival differences between the low- and high-risk score groups.

If there is no special description for the above method, statistical significance is defined as a p-value < 0.05.

Results

FAM-related lncRNAs in LUAD patients

The detailed process is shown in Figure 1. A total of 490 LUAD patients were included in this analysis, with their clinical features in Table 1. 92 FAM-related genes (Appendix D1) were obtained from previous research and the KEGG database. By using Pearson correlation analysis, 1879 FAM-related lncRNAs were discerned as FAM lncRNAs (Figure 2A). The relationship data between FAM-related genes and lncRNAs were shown in Appendix D2, and their correlation was shown in (Figure 2B) (Part of fatty acids metabolism-related genes were selected for display and Supplementary Figure S1 for all).

Construction and validation of a prognostic model

Here, 164 FAM-related lncRNAs were identified through univariate COX regression analysis (Figure 3A, results whose p < 0.01 were selected to show and all results were available in

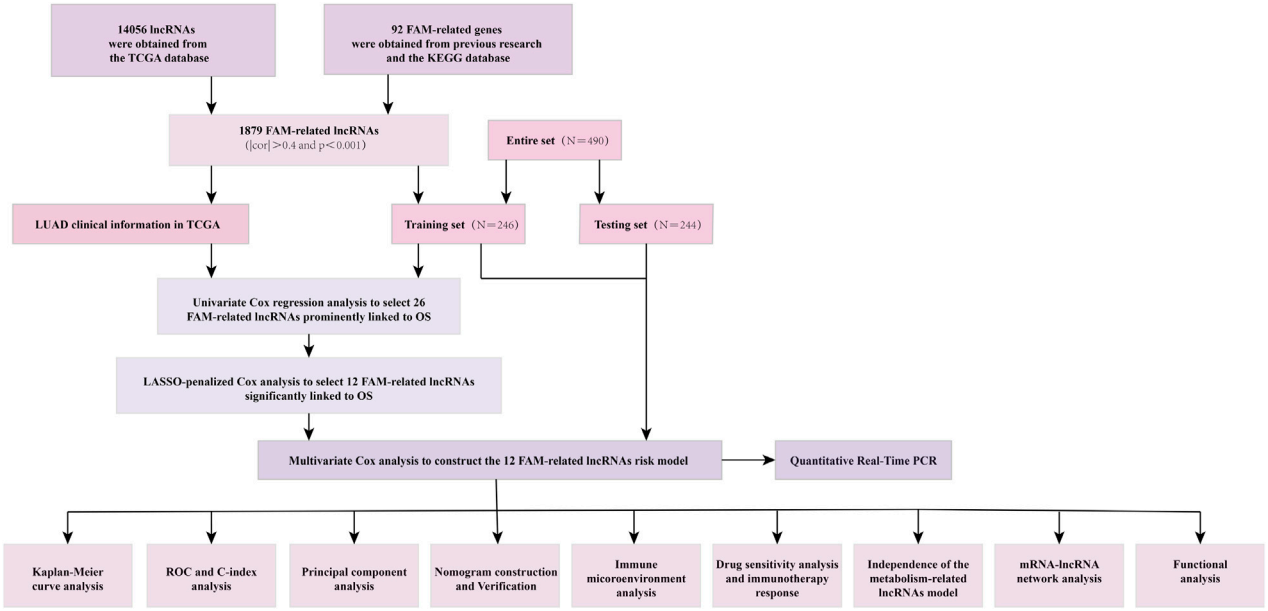


FIGURE 1
The entire analytical process of the study.

TABLE 1 The clinical characteristics of included samples.

Covariates	Type	Total	Test	Train	*p*-value
Age	≤65	231 (47.14%)	111 (45.49%)	120 (48.78%)	0.5789
	>65	249 (50.82%)	127 (52.05%)	122 (49.59%)	
	unknown	10 (2.04%)	6 (2.46%)	4 (1.63%)	
Gender	FEMALE	262 (53.47%)	136 (55.74%)	126 (51.22%)	0.3618
	MALE	228 (46.53%)	108 (44.26%)	120 (48.78%)	
Stage	Stage I-II	378 (77.14%)	184 (37.55%)	194 (39.59%)	0.8546
	Stage III-IV	69 (21.22%)	55 (11.22%)	49 (10%)	
	unknown	8 (1.63%)	5 (2.05%)	3 (1.22%)	
T	T1-2	426 (86.94%)	210 (42.86%)	216 (44.08%)	0.7023
	T3-4	61 (12.45%)	32 (6.53%)	29 (5.92%)	
	unknown	3 (0.61%)	2 (0.82%)	1 (0.41%)	
M	M0	324 (66.12%)	165 (67.62%)	159 (64.63%)	0.7872
	M1	24 (4.9%)	11 (4.51%)	13 (5.28%)	
	unknown	142 (28.98%)	68 (27.87%)	74 (30.08%)	
N	N0	317 (64.69%)	157 (64.34%)	160 (65.04%)	0.3002
	N1-3	162 (33.06%)	82 (16.73%)	80 (16.33%)	
	unknown	11 (2.24%)	5 (2.05%)	6 (2.44%)	

Supplementary Figure S2. The LASSO regression focused on 26 related lncRNAs while avoiding overfitting (Figures 3B,C). Finally, 12 FAM-related lncRNAs (Table 2) were used to construct this prognostic model (Figure 3D).

The risk score was evaluated as: risk score = AL390755.1×(−1.54813752178063)+ AC105020.6×(−2.54915590358082)+ TMPO-AS1×(0.645364340734862)+ AC016737.2×(−1.62529028103815)+

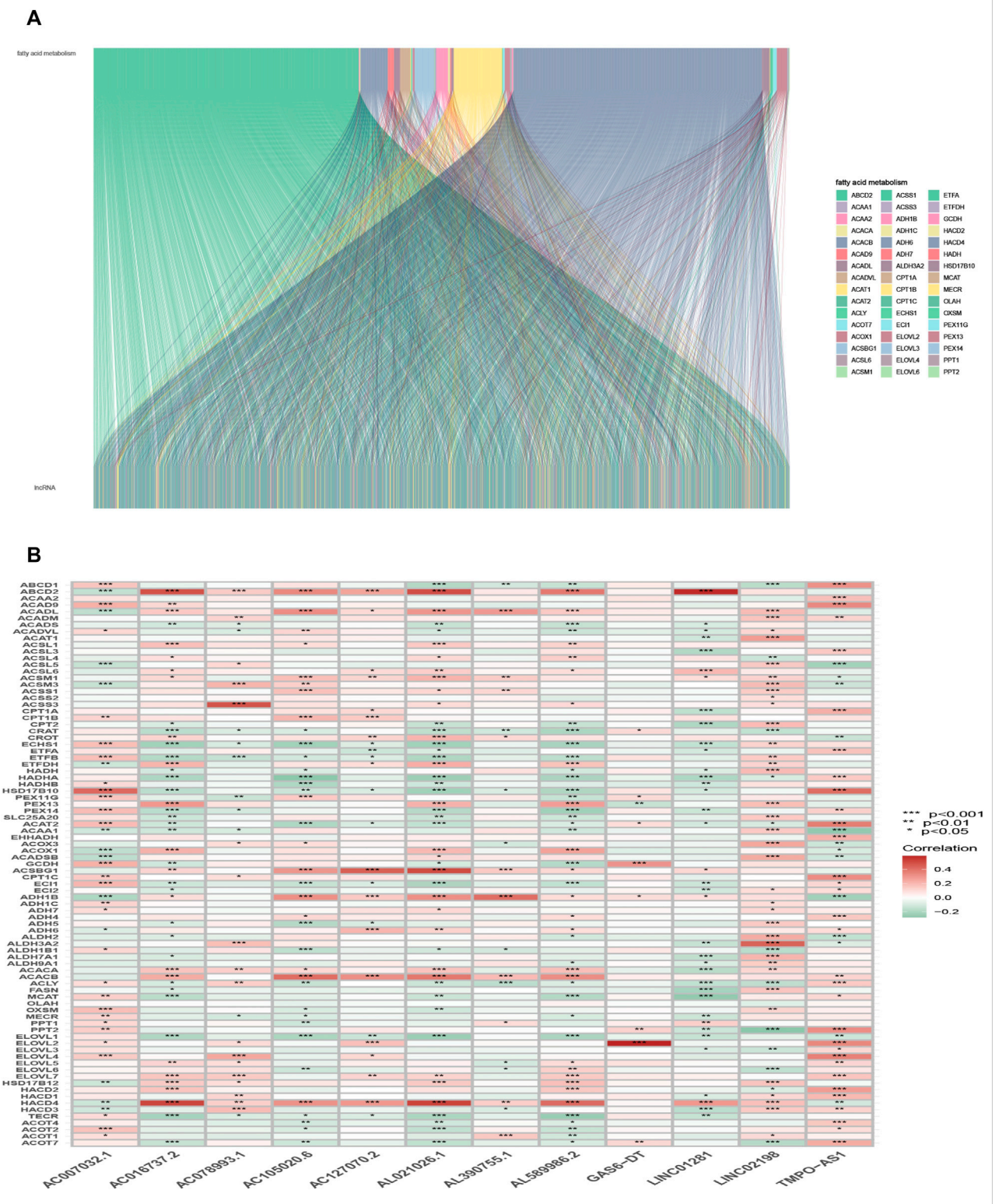


FIGURE 2
Selection of FAM-related lncRNAs in LUAD patients. **(A)** Sankey relation diagram for target lncRNAs. **(B)** Heatmap of the correlation between FAM-related genes and the 12 prognostic FAM-related lncRNAs in TCGA entire set.

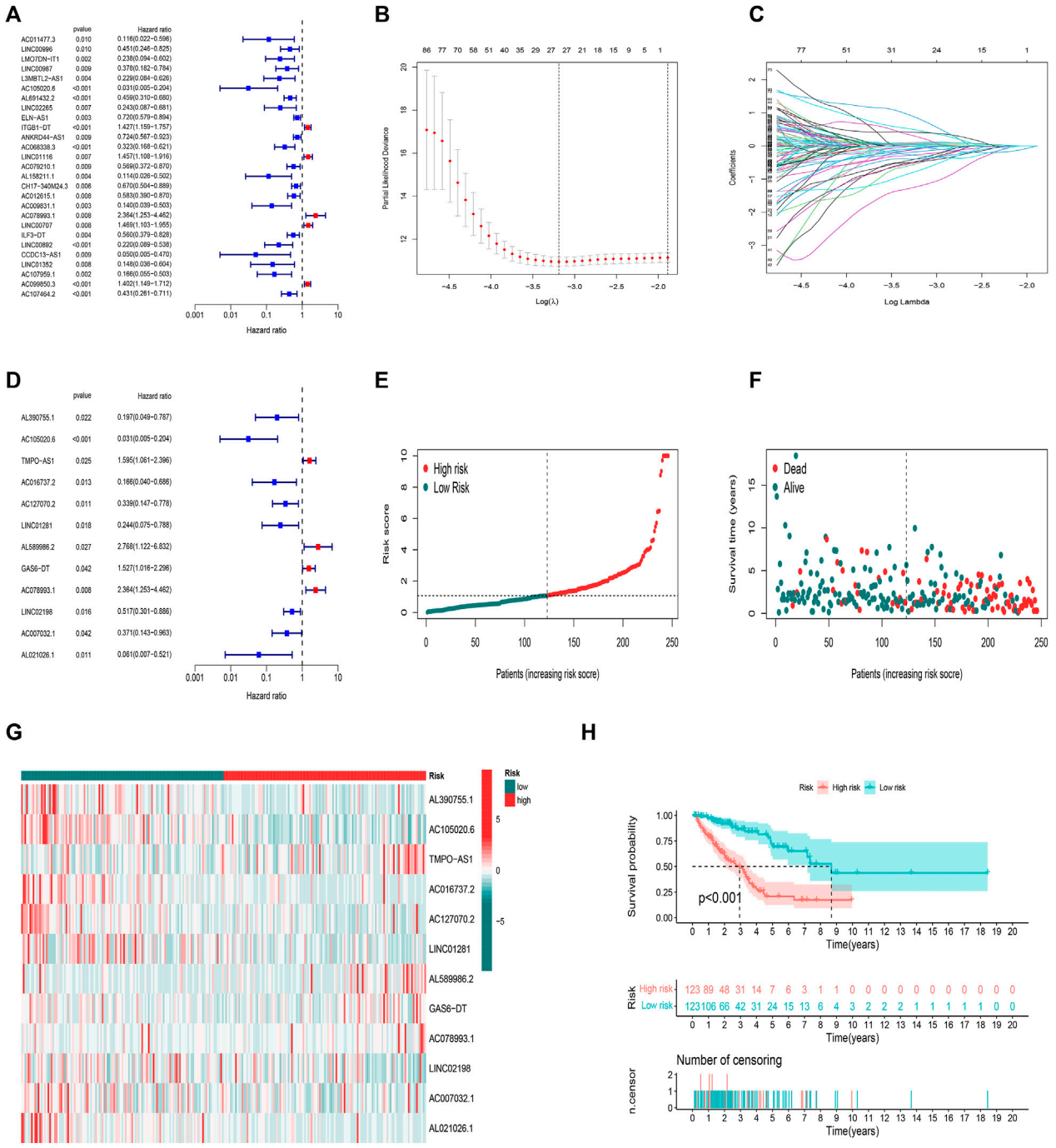


FIGURE 3
Prognostic model in training set validation. **(A)** Univariate Cox regression analysis. **(B)** The LASSO coefficient profile. **(C)** The 10-fold cross-validation for variable selection in the LASSO model. **(D)** Multivariate Cox regression analysis and 12 lncRNAs were finally selected. **(E)** Patient risk score distribution for the training set. **(F)** Survival status time between two risk groups in the training set. **(G)** 12 FAM-related lncRNAs distributed for each patient in the training set. **(H)** OS curve of the training set.

AC127070.2×(−1.1290227830944)+
LINC01281×(−1.21908027803503)+ AL589986.2×(1.97638163310763)+
GAS6-DT×(1.04880809216376)+ AC078993.1×(0.651117829305934)+
LINC02198×(−0.520620804335338)+
AC007032.1×(−1.43127474226429)+
AL021026.1×(−1.55257545730318)

TABLE 2 The 12 FAM-related prognostic lncRNAs.\

Id	Coef	HR	HR.95L	HR.95H	*p*-value
AL390755.1	−1.548137522	0.196552779	0.049061185	0.787445217	0.021594278
AC105020.6	−2.549155904	0.031057225	0.004725325	0.204123802	0.000301471
TMPO-AS1	0.645364341	1.594883853	1.061457403	2.396379259	0.024635498
AC016737.2	−1.625290281	0.165930576	0.040164005	0.685513212	0.013077694
AC127070.2	−1.129022783	0.338535109	0.147220182	0.778466772	0.010790501
LINC01281	−1.219080278	0.243707993	0.075401385	0.787698871	0.018341205
AL589986.2	1.976381633	2.768161634	1.121511954	6.832489666	0.027193217
GAS6-DT	1.048808092	1.527058692	1.015808552	2.295617855	0.041813182
AC078993.1	0.651117829	2.36441459	1.252924785	4.461924947	0.007910449
LINC02198	−0.520620804	0.516815504	0.301364474	0.886296453	0.016458722
AC007032.1	−1.431274742	0.371231383	0.143127389	0.9628677	0.041572162
AL021026.1	−1.552575457	0.060826538	0.007094814	0.521489027	0.010653806

The median value of the risk score was the standard to divide LUAD samples. All samples were divided into two low-/high-risk groups. The distribution of risk grades and survival information between the two groups is shown in (Figures 3E,F). The relative expression standards of the 12 FAM-related lncRNAs for each patient are shown in (Figure 3G). The survival analysis demonstrated that the OS of the low-risk group was longer than that of the high-risk group (Figure 3H $p < 0.001$).

We calculated risk scores for LUAD patients to validate the predictive capability of the established model by using the uniform formula. Figure 4 shows the diffusion of risk scores, survival status and time, and expression of the FAM-related lncRNAs in the testing set (Figures 4A–C) and the entire set (Figures 4E–G). The K-M survival curve based on the testing set and the entire set also showed that the patients in the low-risk group had a longer OS than those in the high-risk group (Figures 4D,H $p < 0.05$).

PCA

Heterogeneity between the two risk subgroups in the entry set and test set was examined by PCA analysis. The whole gene expression profiles, 92 FAM genes, as well as our risk model was included (Figure 5). The analysis results according to the risk model we constructed showed that the low- and high-risk groups had different distributions (Figures 5C–E). This shows that the risk model can distinguish between low- and high-risk groups.

Nomogram

The hazard ratio (HR) of the risk score and 95% confidence interval (CI) were 1.189 and 1.140–1.240 ($p < 0.001$), respectively, in univariate Cox (uni-Cox) regression while 1.176 and 1.126–1.229 ($p < 0.001$), respectively, in multivariate Cox (multi-Cox) regression (Figures 6A,B). Univariate Cox regression analysis indicated that disease stage, T stage, M stage, and risk score, were related to prognosis (Figure 6A, $p < 0.001$). Furthermore, multivariate Cox regression analysis presented that the risk score was an independent factor affecting prognosis (Figure 6B, $p < 0.001$). Therefore, we are reasonably confident that risk models based on FAM-related lncRNAs have a significant impact on the survival and prognosis of LUAD patients and are independent prognostic factors. To better predict the 1-,3-, and 5-year survival for LUAD patients, we established a nomogram combining gender, age, stage, TNM, and risk score (Figure 6C). Using calibration curve analysis, the prediction accuracy of the nomogram was assessed (Figure 6D).

Assessment of the risk model

ROC curves were utilized to evaluate the sensitivity and specificity of the model on the prognosis. The AUC (1-, 3-, and 5-year) for the train set were 0.805, 0.779, and 0.845, of the test set were 0.645, 0.576, and 0.483, and of the entire set were 0.722, 0.664, and 0.688, respectively (Figures 7A–C). The AUC value illustrated that the prognostic risk model of the 12 FAM-related lncRNAs for LUAD was comparatively dependable (Figures 7D,E).

Figures 7F–I showed the OS of patients after sub-clustering using clinical characteristics based on the risk score. Like the previous results, the OS of the low-risk group was better than that of the high-risk group.

Stratification analysis of the risk model in immune features

The infiltration status of immune cells was evaluated by the CIBERSORT algorithm. The proportions of 22 immune cells in each sample were shown in Figures 8A,B. The high-risk group

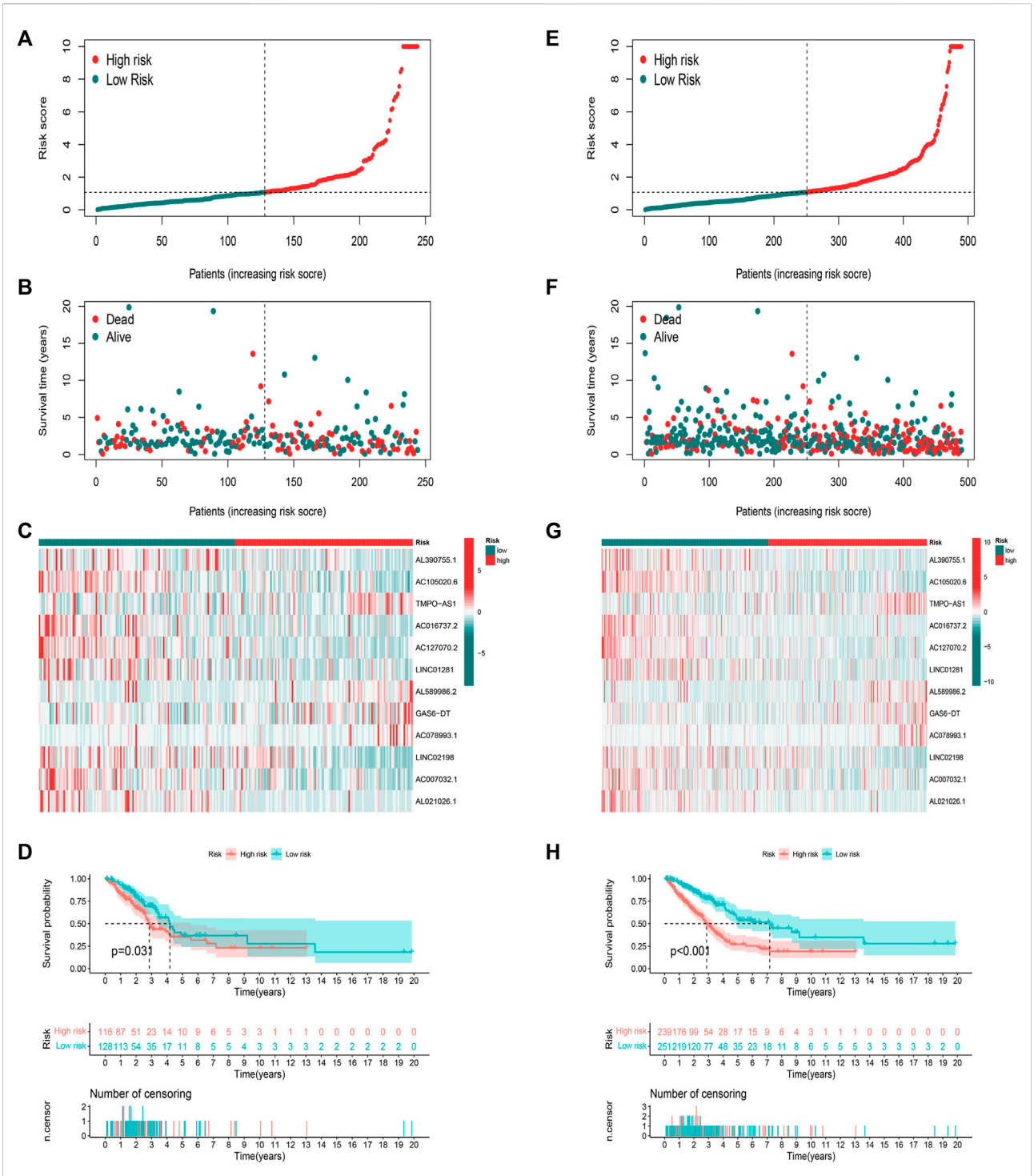


FIGURE 4
Prognostic value of risk score model in testing and entire sets. **(A–D)** Distribution of risk score, survival status, 12 hub lncRNA expression levels, and K-M survival curve (OS) in the testing set. **(E–H)** Distribution of risk score, survival status, 12 hub lncRNA expression levels, and K-M survival curve (OS) in the entire set.

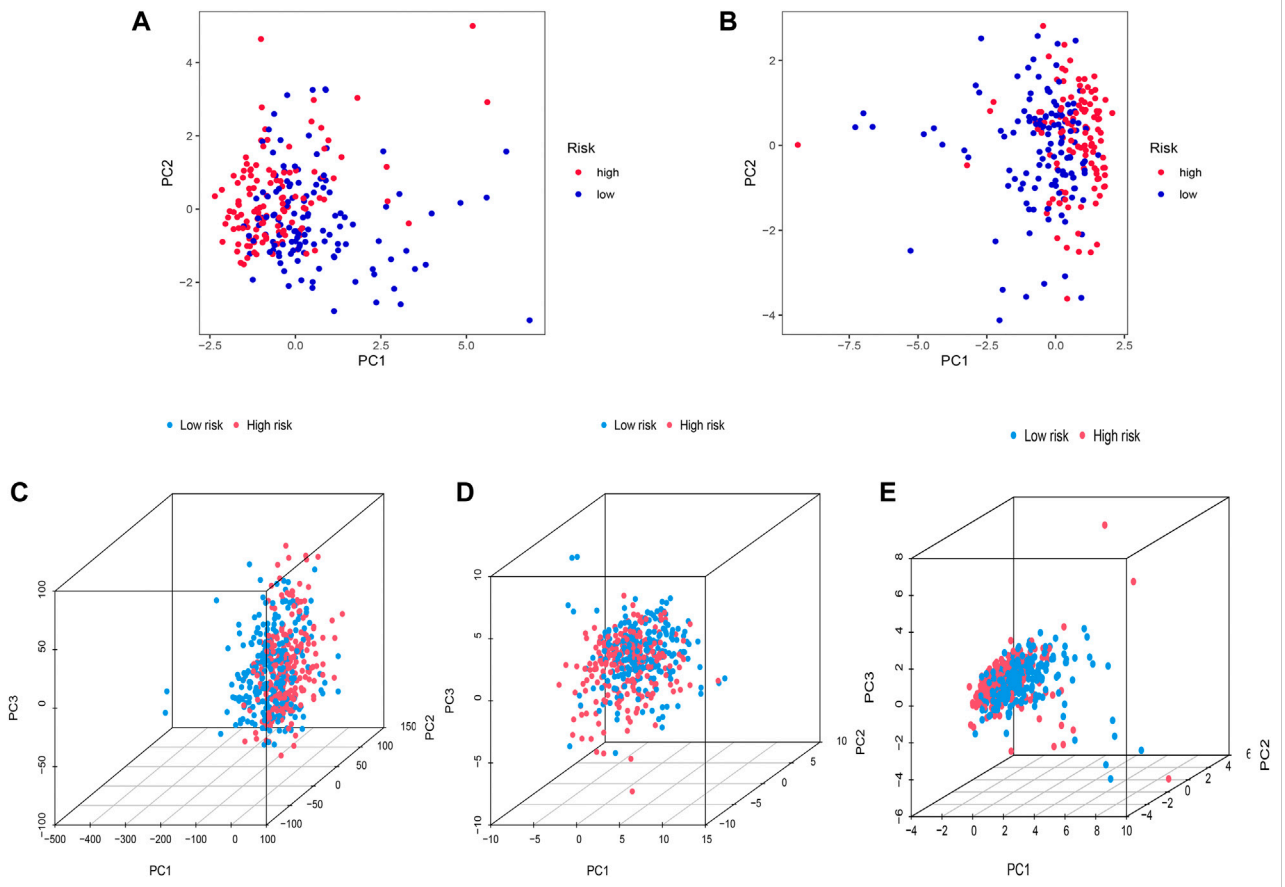


FIGURE 5
Principal component analysis. **(A,B)** 2D PCA in training and the entire set. **(C–E)** PCA between two risk groups for entire gene expression profiles, 92 FAM related-genes, and profiles of the 12 FAM-related lncRNAs as an entire set.

was associated with significantly lower levels of B cells, T cells follicular helper, and Tregs, but a higher level of eosinophils and neutrophils (Figure 8C). Subsequently, the results of the ssGSEA algorithm showed that the high-risk group had a lower mean infiltration level than the low-risk group, with T helper cells showing higher infiltration levels in both risk groups (Figures 8D,E). We obtained similar results in correlation analysis of immune responses, and overall, patients in the high-risk group had a lower immune response. In addition, LUAD patients in the high-risk group had remarkably lower stromal, immune, and ESTIMATE scores (Figures 8F–H).

We then analyzed the mutation data. Mutations were stratified according to the constructed risk model. The results of mutations analysis with those top 20 driver genes are shown in Figures 9A,B. A higher level of TP53 mutations was correlated with a worse survival state. The TMB in the high-risk group exceeded that in the low-risk group, showing that the FAM-related risk model classifier index had a high correlation with TMB (Figure 9C). Therefore, we tested the correlation between FAM-related lncRNAs and TMB based on the risk model using Spearman correlation analysis (Figure 9D r = 0.13, p = 0.0056). The results suggested a strong correlation between the FAM-based classifier index and the TMB. We further investigated the impact of TMB status on the prognosis of LUAD patients by analyzing the survival of the high and low TMB groups. However, the survival curves were similar in both groups, indicating that TMB failed to differentiate survival in LUAD (Figure 9E, $p > 0.05$). Besides, the survival outcome (OS) predictive validity of TMB was conducted, which shows a weaker predictive power than our risk model (Figure 9F, $p < 0.05$). The results show that our model may predict better than the TMB.

Furthermore, according to TIMER2.0 data (Appendix D3), we divided all samples into different immune subtypes (Figure 9G). FAM genes, 12 FAM-related lncRNAs, and risk types were included in the Sankey network (Figure 9H). The above results illustrate the high correlation of these 12 FAM-related lncRNAs with LUAD immunity from another dimension.

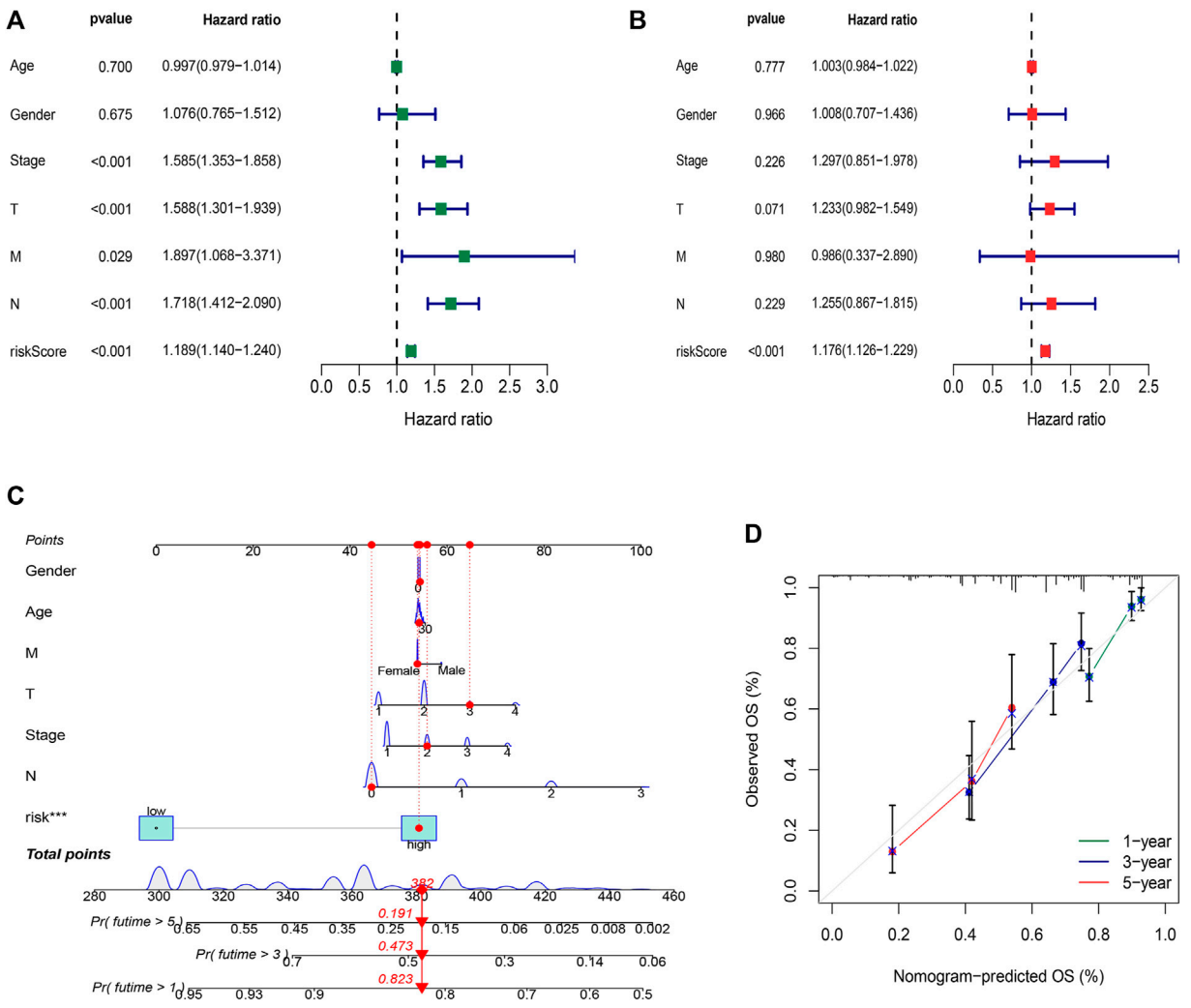


FIGURE 6
Construction and validation of the nomogram. **(A)** Univariate Cox regression analysis indicated that disease stage, T stage, M stage, and risk score, were related to prognosis ($p < 0.001$) **(B)** Multivariate Cox regression analysis presented that the risk score was an independent factor affecting prognosis ($p < 0.001$). **(C)** The nomogram predicts the probability of the 1-, 3-, and 5-year OS. **(D)** The calibration plot.

Clinical treatment and drug sensitivity analysis

Given the differences in the immune microenvironment between these two risk groups, we hypothesized that these two groups might have different responses to drugs. We then used the pRophetic algorithm to estimate treatment response against potential drugs in our model based on the IC50 of each sample in the GDSC database. The correlation between CI50 and different risk groups were shown in Figure 10A. The IC50s for AP.24534, ATRA, AS601245, and ABT.888 were significantly higher in the low-risk group (Figure 10A), suggesting that exposure to these drugs may be more appropriate for high-risk patients. Then those model-related lncRNAs and immunotherapeutic biomarkers were pooled to explore their relationship. We were pleasantly surprised to find that TMPO-AS1 was related to the sensitivity of multiple drugs (Figure 10B, the entire result was available in Appendix D4).

Unsurprisingly, the high-risk group may effect better in immunotherapy, which also means that our model might serve as a potential signature for predicting TIDE (Figure 10C).

Functional analysis

GO analysis illustrated that these risk model-related genes mainly affect the modulation of axoneme assembly, motile

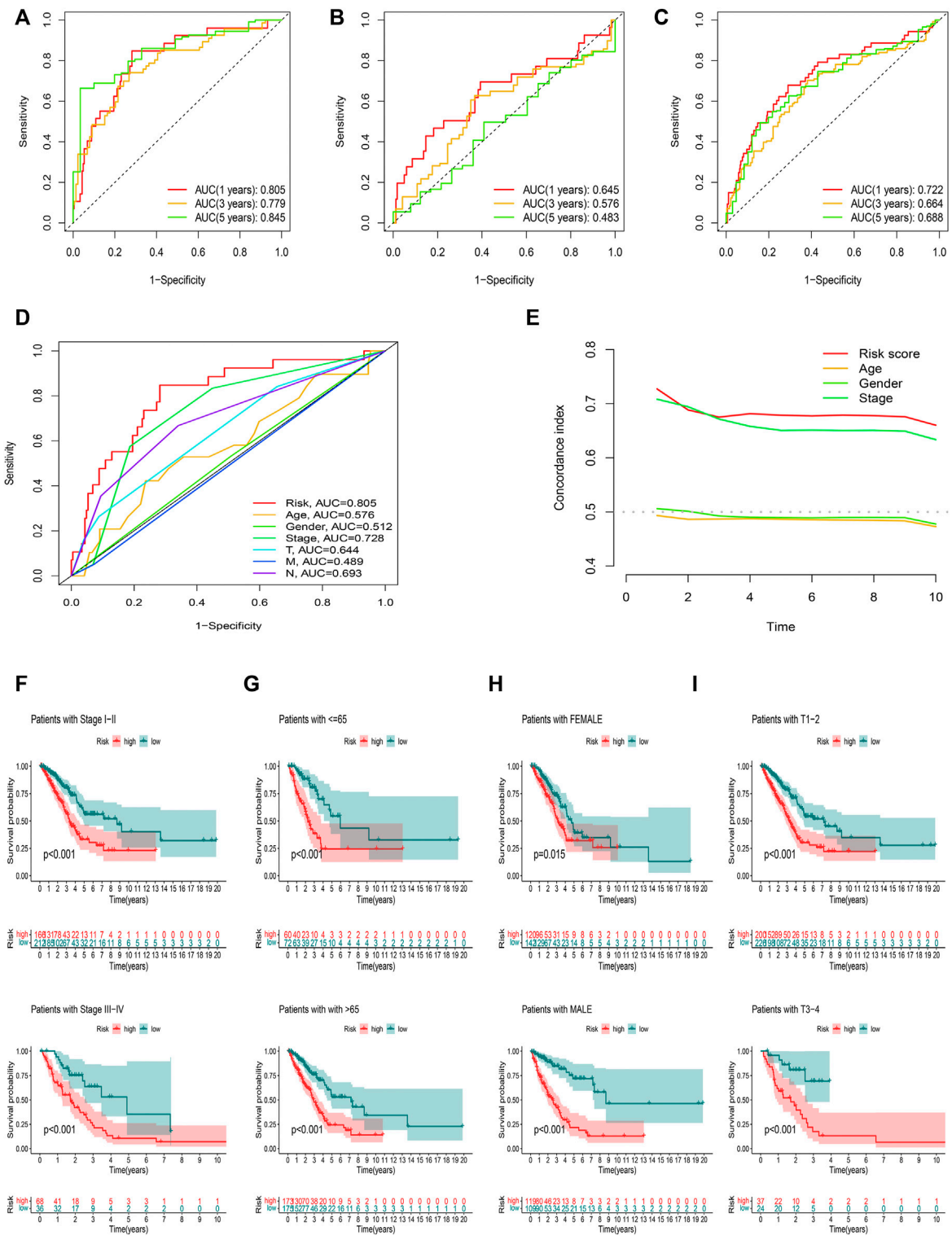


FIGURE 7
Assessment of the prognostic risk model. **(A–C)** The 1-, 3-, and 5-year ROC curves of the training, testing set, and entire set. **(D)** ROC curves of all included features. **(E)** CI of the risk score and clinical characteristics. **(F–I)** OS curve of difference clustered by LUAD clinical features between two risk groups in the entire set.

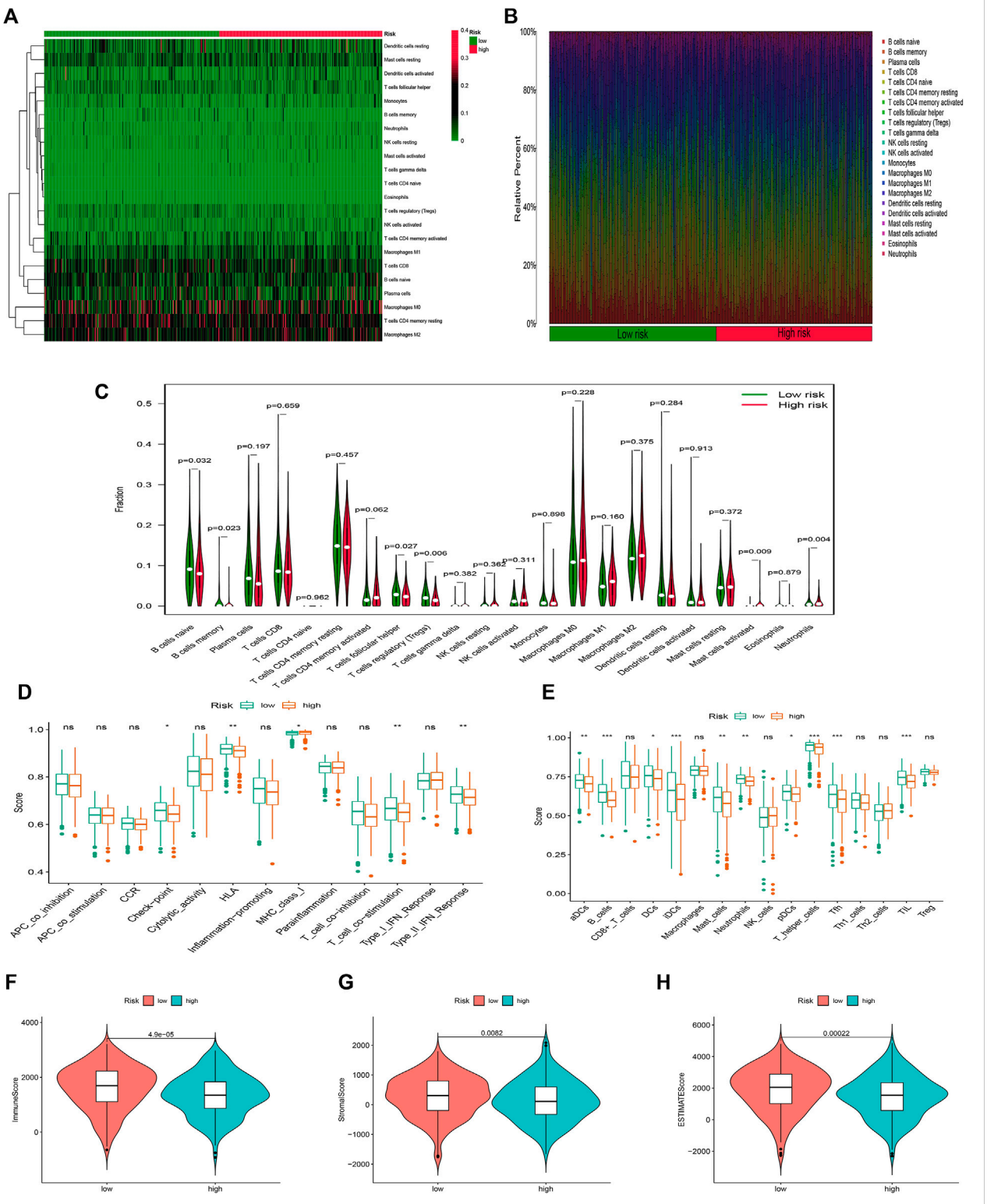


FIGURE 8
Stratification Analysis of the FAM-related lncRNA prognostic risk score in immune features. **(A–C)** Heatmap, bar chart, and relative infiltrating proportion of 22 tumor-infiltrating immune cell types in two risk groups. **(D,E)** The score of immune functions comparing two risk groups by ssGSEA or ssGSEA score. **(F–H)** The comparison of immune-related scores between high- and low-risk groups.

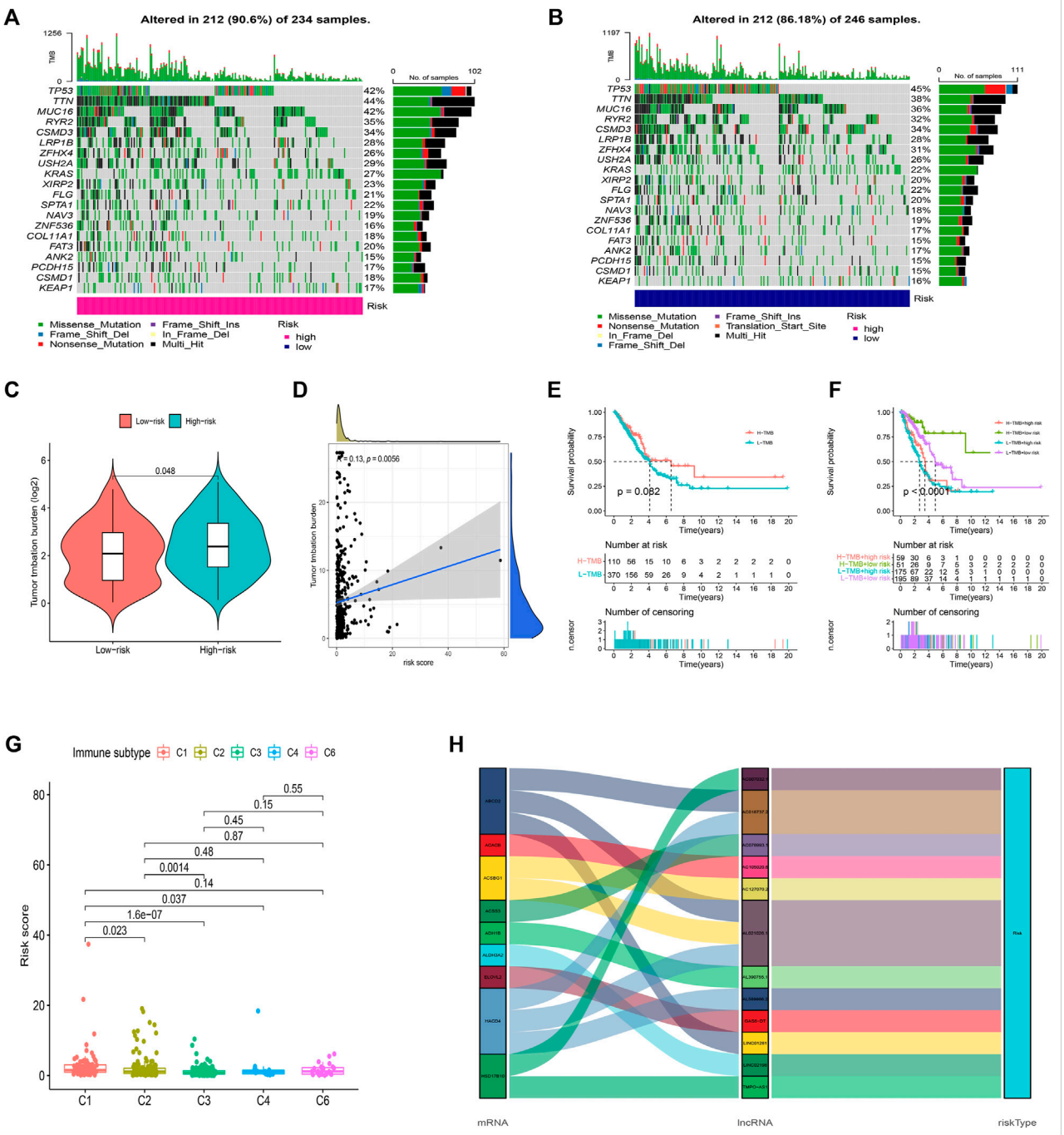


FIGURE 9
Exploration of TMB and lncRNAs networks visualization. **(A,B)** 20 genes with high mutation frequencies in different risk subgroups. **(C)** TMB difference in two risk groups. **(D)** The correlation between risk score and TMB. **(E)** K-M curves of the patient OS in the high-TMB and low-TMB groups in the entire set. **(F)** The survival outcome predictive validity of TMB. **(G)** The correlation between risk score and immune subtype. **(H)** Sankey diagram: the connection degree between the FAM-related genes, FAM-related lncRNAs, and risk types.

cilium, chemokine activity, and so on (Figure 11A). KEGG analysis illustrated that these genes were involved in multiple immune pathways such as the chemokine signaling pathway, B cell receptor signaling pathway, and so on (Figures 11B,C). KEGG analysis results were shown in Figures 11D,E. Pathways such as aminoacyl tRNA biosynthesis and biosynthesis of unsaturated fatty acids were significantly related to the high-risk group, while pathways such as allograft rejection and asthma were significantly enriched in the

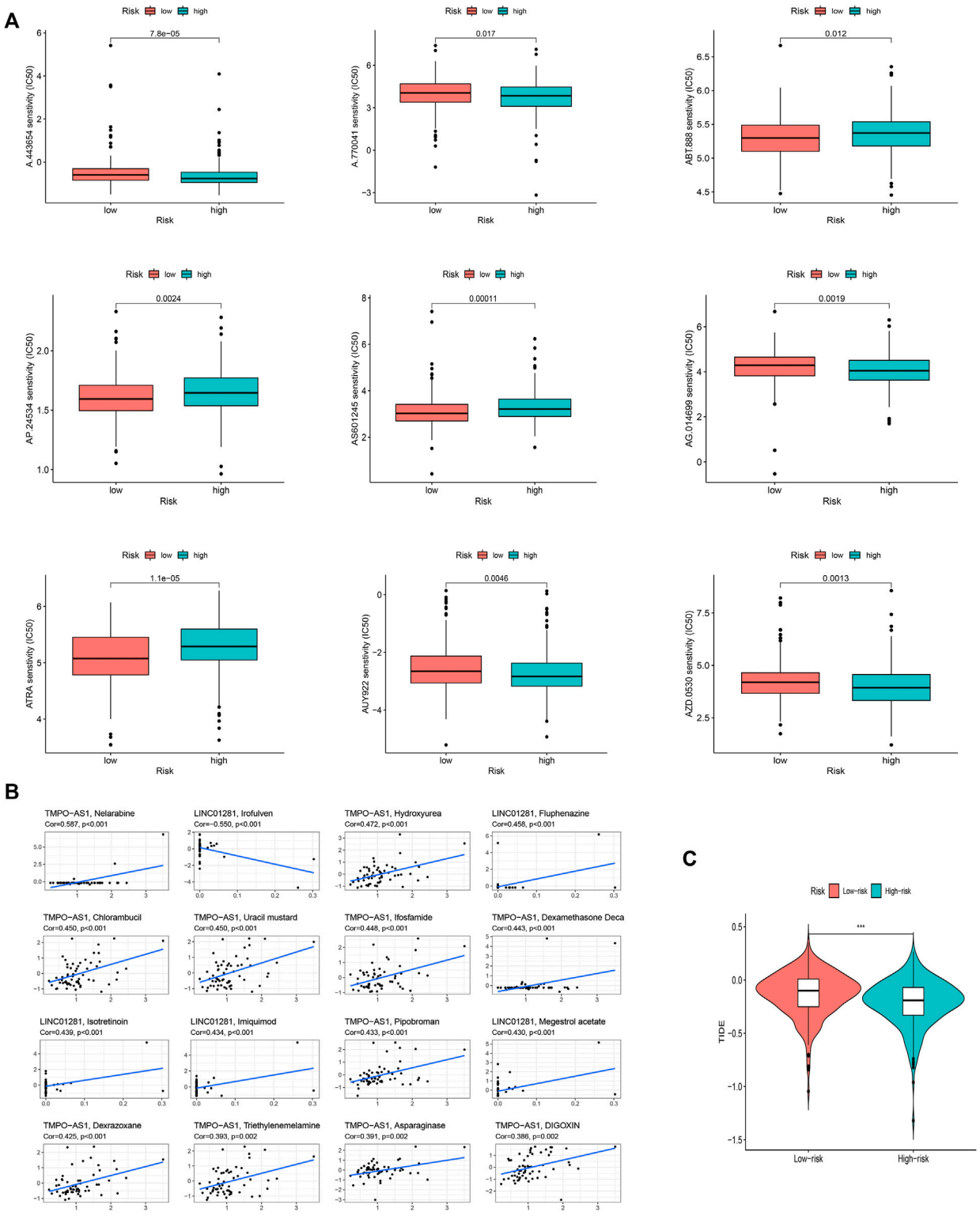


FIGURE 10
The investigation of tumor immune factors and immunotherapy. **(A)** The immunotherapy prediction of high-risk and low-risk groups. **(B)** The correlation between 12 FAM-related lncRNAs and drugs. **(C)** TIDE prediction difference in the high-risk and low-risk patients.

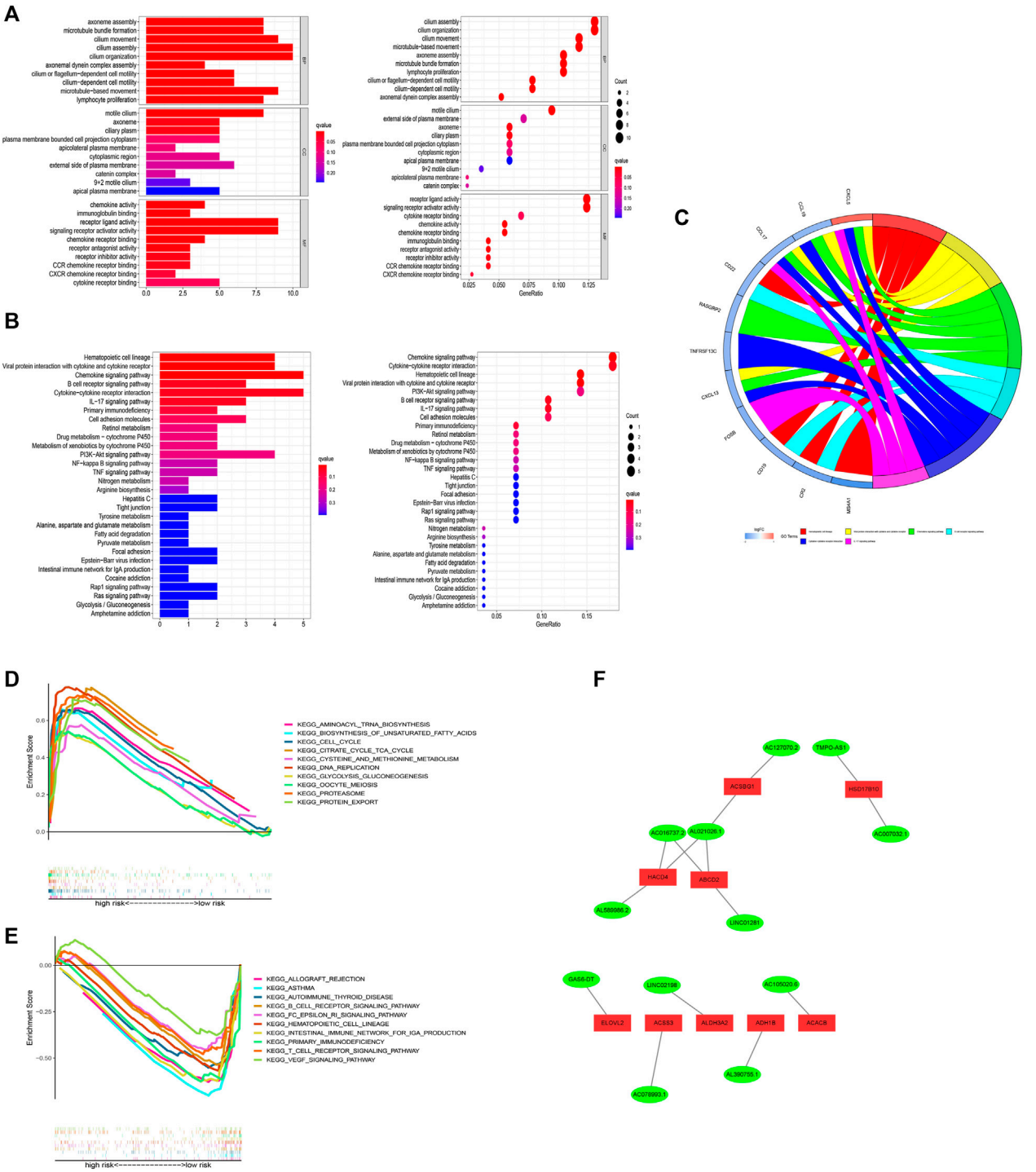


FIGURE 11
Functional analysis. **(A)** Result of GO functional enrichment (top 10). **(B)** KEGG enrichment terms (top 30). **(C)** Circle diagram in KEGG analysis. **(D)** GSEA of the top 10 pathways significantly enriched in the high-risk group. **(E)** GSEA of the top 10 pathways in the low-risk group. **(F)** 12 FAM-related lncRNAs and differential FAM genes networks.

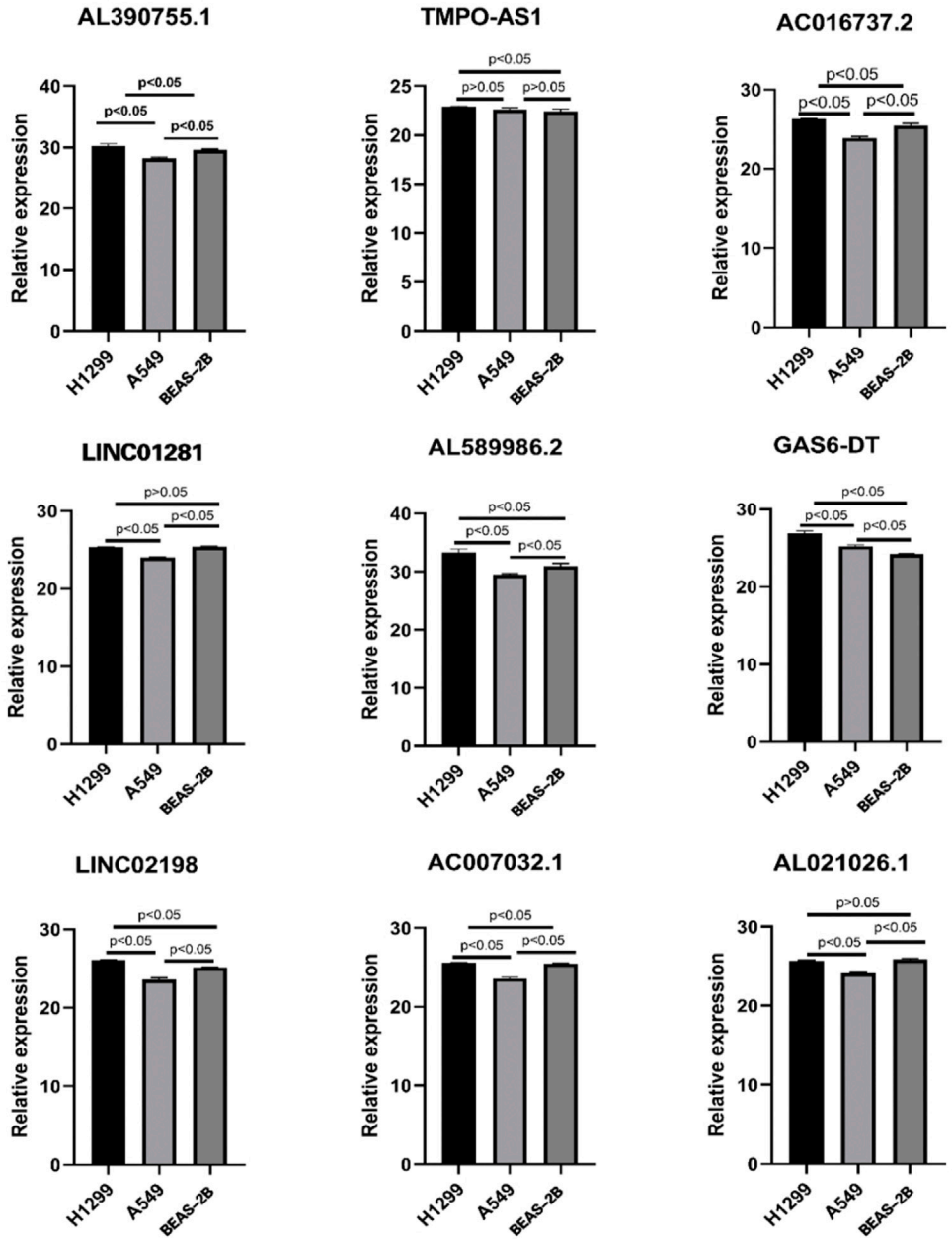


FIGURE 12
Expression of nine lncRNAs from the risk model in LUAD cell lines and bronchial epithelial cells.

low-risk group. In addition, we also established an interaction network for these key lncRNAs (Figure 11F).

Verification of expression level *in vitro* on hub lncRNAs

To verify the expression level of 12 FAM-related lncRNAs in LUAD cells, we used RT-qPCR analysis to detect BEAS-2B and LUAD cells, including A549 and H1299 (Figure 12). Unfortunately, the sequences of three lncRNAs (AC105020.6, AC127070.2, AC078993.1) did not have suitable primers, so we only verified the expression levels of the remaining nine FAM-related lncRNAs (The PCR primer sequences were available in Table 3). Among these lncRNAs, we found that the expression of GAS6-DT in the A549 and H1299 cell lines was significantly higher than that in the BEAS-2B cell line, and TMPO-AS1 was significantly higher in the H1299 cell line. Combined with the

TABLE 3 The PCR primer sequences.

Gene	F 5′-3′	R 5′-3′
TMPO-AS1	5′-CAGACCTCTACAATCGGGCACTTA-3′	5′-ATTCTTGCGGGTGGTGGGAT-3′
AC016737.2	5′-CTGGAGATGGACTTTGGCT-3′	5′-CTTGTGAGGTGGCTGTTATTATC-3′
LINC01281	5′-CAGCCCAGAGTGAAGATAAGAATAC-3′	5′-GAAGCCACCAGCAGAATGACA-3′
GAS6-DT	5′-TAGCTATTATTTCCTAAGGGTTCCAG-3′	5′-TCCATTAACTCTCTTCTCCAAAACTACA-3′
LINC02198	5′-ACTTCTGTCACCCCCTTGATTACC-3′	5′-CCAAAGACTGGTCCTCCTCTATCC-3′
AC007032.1	5′-TGATGACTTCACCCAAATACAGACC-3′	5′-ACTTTTTCCTGGCTACTTTTATCCG-3′
AL021026.1	5′-ATATCTGAGCCTGAGTTTCCCATTC-3′	5′-TTCCATAGCCGCCAATACAAGC-3′
AL390755.1	5′-GGAAAGCTATGAGGAAGAAGAAACAGA-3′	5′-CAACCTGTGCTGTGATGAATGG-3′
AL589986.2	5′-CCTGATACTGGTTTTTCTACATGCTTC-3′	5′-TCCAAGGTTGTGCTATGGTAATCTG-3′

previous research results, their high expression was associated with a poorer prognosis, with HR < 1, suggesting that GAS6-DT and TMPO-AS1 genes may be a risk factor in LUAD. The expression levels of AL021026.1 and LINC01281 in the A549 cell line were significantly lower than those in the BEAS-2B cell line, and their average expression in the H1299 cell line was lower than that in BEAS-2B, but the difference was not significant ($p > 0.05$). Model coefficients and patient outcomes are considered that they may be protective factors for LUAD. Interestingly, we also noticed that LINC02198, AC007032.1, AL589986.2, and AL390755.1 were significantly overexpressed in the H1299 cell line and significantly underexpressed in the A549 cell line, while AC016737.2 was significantly overexpressed in H1299 cell line. The high expression contradicts the result of its coefficient of less than 0 in the risk model.

Discussion

With the deepening of tumor research, the role of metabolic reprogramming in tumors cannot be underestimated any longer (Faubert et al., 2020). Simply, tumor cells are different from normal tissue cells, when tumor cells are ready to colonize other organs, they need to compete with other normal cells for the living environment and nutrients. Therefore, the metabolic demands of tumor cells are regulated to meet the needs of survival in the current environment (Schild et al., 2018). Based on the above characteristics, the metabolic reprogramming of tumor cells is also regarded as a hallmark of tumor development (Ward and Thompson, 2012). In addition, more and more evidence shows that in addition to protein-coding RNA mutations, mutations and abnormal modifications of non-coding RNAs represented by lncRNAs were also vital in tumor progression (Bhan et al., 2017). Therefore, these non-coding RNAs also play a key role in tumor progression. It is regarded as a new marker for tumor diagnosis or a new therapeutic target (Kong et al., 2019; Wang et al., 2019; Xing et al., 2021). Here, we cannot help but want to explore whether and how lncRNA can interact with the lipid metabolism reprogramming of tumor cells, and how the interaction between the two affects the process of LUAD and thus affects the prognosis and survival of patients. At the same time, it is hoped that more powerful biomarkers and therapeutic targets can be found for the clinical diagnosis and treatment of LUAD.

In this study, LUAD data were obtained from the TCGA database, while FAM-related lncRNA data were downloaded from the KEGG database. After differential gene analysis, 1879 differential lncRNAs related to FAM were found, and after survival analysis, univariate/multifactor and LASSO Cox regression. A 12-hub FAM-related lncRNA prognostic model with high reliability and validity was constructed. Further exploration was performed to figure out how those hub lncRNAs were involved in LUAD progression.

Among the 12 key FAM-related genes we finally screened for risk score modeling, most lncRNAs have not been studied, but some lncRNAs have also appeared in the construction of prognosis prediction models for different diseases. For example, AL390755.1 was used to construct a prognostic prediction model for low-grade glioblastoma, and similar LINC01281, AL589986.2, and AC007032.1 were also used for laryngeal cancer (Zhang et al., 2019), cervical cancer (Ye et al., 2021), dilated cardiomyopathy (Zhang et al., 2020) and proliferative vitreoretinopathy respectively (Ni et al., 2021), which were also considered a potential diagnostic marker. At the same time, we also noticed that GAS6-DT and TMPO-AS1 have been shown to have regulatory axes in previous studies, which can interact with another coding/non-coding RNAs, and these two lncRNAs were proved to be possible risk factors for LUAD in our PCR validation. For example, the study of Zilin Li et al. pointed out that in liver cancer cells with incomplete radiofrequency ablation, the expression of GAS6-DT is often up-regulated and can competitively inhibit the binding of microRNA-3619-5p to ARL2, thereby promoting the proliferation and migration of liver cancer cells (Li et al., 2021b). The

relevant research on TMPO-AS1 is relatively sufficient. The study of Xiaoqian Mu et al. (Mu et al., 2020) pointed out that TMPO-AS1 is highly expressed in LUAD samples and knocking down this gene negatively regulates the cell cycle of tumors and reduces the invasiveness of tumors. Targeted binding to TMPO-AS1 plays a role similar to gene knockout. A similar study by Qiu L et al. (Li et al., 2021c) also pointed out that TMPO-AS1 can also interact with miR-143-3p, ultimately affecting the expression of CDK1 and regulating the cell cycle of LUAD. Interestingly, the study by Jie Yao et al. (Yao et al., 2021) pointed out that TMPO-AS1 is also involved in the regulation of iron metabolism in LUAD.

When conducting drug sensitivity analysis, we were pleasantly surprised to find that the expression level of TMPO-AS1 is highly correlated with the sensitivity of various drugs, including ifosfamide, thiotepa, irinotecan, and other antitumor drugs that have been approved for clinical LUAD use. Among them, the expression level of this lncRNA is highly positively correlated with the CI50 of ifosfamide, which means that the higher the expression level of TMPO-AS1 (which also means a higher risk score), the worse the effect of ifosfamide for LUAD treatment. At the same time, we also noticed that trametinib showed a negative correlation. The experimental study by Toshiyuki Sumi et al. indicated that trametinib can reduce survivin expression in RB1$^+$/KRAS-mutated LUAD cells, thereby improving prognosis (Sumi et al., 2018), and there was a similar case study by Maurício Fernando Silva Almeida Ribeiro et al. (Ribeiro et al., 2021). However, the latest clinical study by Luo J, Makhnin A et al. pointed out that in the drug-resistant EGFR-mutant LUAD that had previously appeared with a tyrosine kinase inhibitor, the addition of trametinib could not reverse the sensitivity of the tumor to the drug (Luo et al., 2021), which means that TMPO-AS1 may be a potential target to solve this problem, which is worthy of further study.

In addition, in the PCR validation of these hub lncRNAs, we found that the expression levels of the four lncRNAs LINC02198, AC007032.1, AL589986.2, and AL390755.1 were not consistent in the H1299 and A549 cell lines, and they were all expressed in the H1299 cell line. Moderately high expression, but low expression in the A549 cell line. We speculate that this may be related to the differences in the genomes of the two cells themselves. H1299 is a lymph node-derived human NSCLC cell line (Giaccone et al., 1992), while A549 cells are human adenocarcinoma alveolar basal epithelial cells (Foster et al., 1998). The different sources of the two may be one of the possibilities leading to this contradiction. Secondly, the H1299 cell line is considered to be a p53 wild-type cell, while the A549 is a p53-null cell (Dorandish et al., 2021). Several studies have also pointed to heterogeneity between the two types of cells, and our findings may add some new evidence to this topic (Yang et al., 2018) (Sidorova and Petrikaitė, 2022). In addition, we also noticed that AC016737.2 was significantly highly expressed in the H1299 cell line, but the coefficient of this lncRNA in the risk model was negative. This contradiction may require more basic experiments to explain.

Furthermore, with the deepening of research, there is a lot of evidence that there is a close relationship between the metabolic reprogramming of tumor cells and the tumor immune response (Cronin et al., 2018). In a broad sense, lipid metabolites include phospholipids, fatty acids, and cholesterol, and the impact of FAM on immune cells is particularly well-studied, for example, during the transformation of monocytes like neutrophils, the demand for fatty acid synthesis is significantly increased (den Hartigh et al., 2010). In regulatory T cells (Tregs), there is high fatty acid oxidation for energy, but in effector T cells, this oxidative activity is inhibited, which also maintains the relative stability of the immune system (Amersfoort et al., 2021). When we jointly analyzed the immune characteristics of LUAD samples of different risk groups, we found that the immune characteristics of the high-risk group were all suppressed compared with the low-risk group. At the level of immune cell infiltration, T helper cells (CD4$^+$ T cells) were infiltrated to a higher degree in both high and low-risk groups, and there were significant differences between groups. Previous studies have pointed out that FAM is closely related to the phenotypic differentiation of T helper cells (Almeida et al., 2016), but our study found that Th1/Th2 subtypes did not have significant infiltration differences between the two risk groups, so we guessed that the lipid metabolism of LUAD is important. Programming may have more effect on the shift of T helper cells towards Th17. Recent work by Panagiota Mamareli indicated that *de novo* synthesis of fatty acids is necessary for the differentiation of the Th17 phenotype (Mamareli et al., 2021), In contrast, Ran You et al. observed that TIL in early-stage NSCLC was biased towards IL17A expression, whereas Th17 cells were reduced in tumor-infiltrating regional lymph nodes in advanced NSCLC (You et al., 2018). The clinical study by Chen G et al. also showed that Th17 and IL-17 increased in the peripheral blood of LUAD patients (Chen et al., 2020). Interestingly, in the functional enrichment analysis, we found that the IL-17 signaling pathway was enriched to the top position, which confirmed our conjecture to a certain extent. All the above evidence suggested that lipid metabolism reprogramming in LUAD may lead to the differentiation of T helper cells inclined towards Th17, which affects the LUAD process.

Interestingly, in the functional enrichment analysis, we found that the IL-17 signaling pathway was enriched to the top position, which confirmed our conjecture to a certain extent.

Overall, based on the lncRNA regulation of lipid metabolism in LUAD, we constructed a prognostic prediction model with good prediction results, and the model has high reliability and validity. In addition, we also

conducted a preliminary study on how these key lncRNAs participate in the LUAD process and affect LUAD immunotherapy. Everything we do aims to improve the understanding of LUAD and shed a hoping light on early clinical diagnosis and treatment of LUAD.

Limitations

Of course, almost all studies face certain challenges of limitations. Our study is no exception. First, we need to complete further validation experiments to provide reliable support for this study. Second, traditional statistical analysis methods may be of limited value in building and assessing prognostic risk models. We hope to open the door for lung adenocarcinoma research, and more questions to follow will require more investigators to join the study.

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

Project design: YS and XK; database search and literature screen: YL and XZ; data extraction and analysis: XC and WQ; quality evaluation of the included literature: YL and YT; and manuscript writing: YS, YL, XK, MG, LL, and QL. The final draft was approved by all the authors.

Funding

This study was supported by the Science and Technology Support Program of the Department of Science and Technology of Guizhou Province (grant no.: Qiankehe Support (2021) General 08), which provided financial support to conduct this study.

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/fgene.2022.990153/full#supplementary-material

References

Ali, T., and Grote, P. (2020). Beyond the RNA-dependent function of LncRNA genes. *Elife* 9, e60583. doi:10.7554/eLife.60583

Almeida, L., Lochner, M., Berod, L., and Sparwasser, T. (2016). Metabolic pathways in T cell activation and lineage differentiation. *Semin. Immunol.* 28 (5), 514–524. doi:10.1016/j.smim.2016.10.009

Amersfoort, J., Schaftenaar, F. H., Douna, H., van Santbrink, P. J., van Puijvelde, G. H. M., Slutter, B., et al. (2021). Diet-induced dyslipidemia induces metabolic and migratory adaptations in regulatory T cells. *Cardiovasc. Res.* 117 (5), 1309–1324. doi:10.1093/cvr/cvaa208

Arbour, K. C., and Riely, G. J. (2017). Diagnosis and treatment of anaplastic lymphoma kinase-positive non-small cell lung cancer. *Hematol. Oncol. Clin. North Am.* 31 (1), 101–111. doi:10.1016/j.hoc.2016.08.012

Ashton, T. M., McKenna, W. G., Kunz-Schughart, L. A., and Higgins, G. S. (2018). Oxidative phosphorylation as an emerging target in cancer therapy. *Clin. Cancer Res.* 24 (11), 2482–2490. doi:10.1158/1078-0432.CCR-17-3070

Bergers, G., and Fendt, S. M. (2021). The metabolism of cancer cells during metastasis. *Nat. Rev. Cancer* 21 (3), 162–180. doi:10.1038/s41568-020-00320-2

Bhan, A., Soleimani, M., and Mandal, S. S. (2017). Long noncoding RNA and cancer: A new paradigm. *Cancer Res.* 77 (15), 3965–3981. doi:10.1158/0008-5472.CAN-16-2634

Blandin Knight, S., Crosbie, P. A., Balata, H., Chudziak, J., Hussell, T., and Dive, C. (2017). Progress and prospects of early detection in lung cancer. *Open Biol.* 7 (9), 170070. doi:10.1098/rsob.170070

Blum, A., Wang, P., and Zenklusen, J. C. (2018). SnapShot: TCGA-analyzed tumors. *Cell* 173 (2), 530. doi:10.1016/j.cell.2018.03.059

Bogie, J. F. J., Haidar, M., Kooij, G., and Hendriks, J. J. A. (2020). Fatty acid metabolism in the progression and resolution of CNS disorders. *Adv. Drug Deliv. Rev.* 159, 198–213. doi:10.1016/j.addr.2020.01.004

Broadfield, L. A., Pane, A. A., Talebi, A., Swinnen, J. V., and Fendt, S. M. (2021). Lipid metabolism in cancer: New perspectives and emerging mechanisms. *Dev. Cell* 56 (10), 1363–1393. doi:10.1016/j.devcel.2021.04.013

Chen, G., Zhang, P. G., Li, J. S., Duan, J. J., Su, W., Guo, S. P., et al. (2020). Th17 cell frequency and IL-17A production in peripheral blood of patients with non-small-cell lung cancer. *J. Int. Med. Res.* 48 (6), 300060520925948. doi:10.1177/0300060520925948

Cluntun, A. A., Lukey, M. J., Cerione, R. A., and Locasale, J. W. (2017). Glutamine metabolism in cancer: Understanding the heterogeneity. *Trends Cancer* 3 (3), 169–180. doi:10.1016/j.trecan.2017.01.005

Cronin, S. J. F., Seehus, C., Weidinger, A., Talbot, S., Reissig, S., Seifert, M., et al. (2018). The metabolite BH4 controls T cell proliferation in autoimmunity and cancer. *Nature* 563 (7732), 564–568. doi:10.1038/s41586-018-0701-2

den Hartigh, L. J., Connolly-Rohrbach, J. E., Fore, S., Huser, T. R., and Rutledge, J. C. (2010). Fatty acids from very low-density lipoprotein lipolysis products induce lipid droplet accumulation in human monocytes. *J. Immunol.* 184 (7), 3927–3936. doi:10.4049/jimmunol.0903475

Dey, P., Kimmelman, A. C., and DePinho, R. A. (2021). Metabolic codependencies in the tumor microenvironment. *Cancer Discov.* 11 (5), 1067–1081. doi:10.1158/2159-8290.CD-20-1211

Dorandish, S., Atali, S., Ray, R., Al Khashali, H., Coleman, K. L., Guthrie, J., et al. (2021). Differences in the relative abundance of ProBDNF and mature BDNF in A549 and H1299 human lung cancer cell media. *Int. J. Mol. Sci.* 22 (13), 7059. doi:10.3390/ijms22137059

Faubert, B., Solmonson, A., and DeBerardinis, R. J. (2020). Metabolic reprogramming and cancer progression. *Science* 368 (6487), eaaw5473. doi:10.1126/science.aaw5473

Fok, E. T., Davignon, L., Fanucchi, S., and Mhlanga, M. M. (2018). The lncRNA connection between cellular metabolism and epigenetics in trained immunity. *Front. Immunol.* 9, 3184. doi:10.3389/fimmu.2018.03184

Foster, K. A., Oster, C. G., Mayer, M. M., Avery, M. L., and Audus, K. L. (1998). Characterization of the A549 cell line as a type II pulmonary epithelial cell model for drug metabolism. *Exp. Cell Res.* 243 (2), 359–366. doi:10.1006/excr.1998.4172

Giaccone, G., Battey, J., Gazdar, A. F., Oie, H., DraouiM.and Moody, T. W. (1992). Neuromedin B is present in lung cancer cell lines. *Cancer Res.* 52 (9), 2732s–2736s.

Han, J., Li, Q., Chen, Y., and Yang, Y. (2021). Recent metabolomics analysis in tumor metabolism reprogramming. *Front. Mol. Biosci.* 8, 763902. doi:10.3389/fmolb.2021.763902

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013

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 (22), 2093–2104. doi:10.1056/NEJMoa1801946

Inamura, K. (2018). Clinicopathological characteristics and mutations driving development of early lung adenocarcinoma: Tumor initiation and progression. *Int. J. Mol. Sci.* 19 (4), E1259. doi:10.3390/ijms19041259

Jiang, N., Zhang, X., Gu, X., Li, X., and Shang, L. (2021). Progress in understanding the role of lncRNA in programmed cell death. *Cell Death Discov.* 7 (1), 30. doi:10.1038/s41420-021-00407-1

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). Kegg: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45 (D1), D353–d361. doi:10.1093/nar/gkw1092

Kimura, I., Ichimura, A., Ohue-Kitano, R., and Igarashi, M. (2020). Free fatty acid receptors in health and disease. *Physiol. Rev.* 100 (1), 171–210. doi:10.1152/physrev.00041.2018

Kong, X., Duan, Y., Sang, Y., Li, Y., Zhang, H., Liang, Y., et al. (2019). LncRNA-CDC6 promotes breast cancer progression and function as ceRNA to target CDC6 by sponging microRNA-215. *J. Cell. Physiol.* 234 (6), 9105–9117. doi:10.1002/jcp.27587

Koundouros, N., and Poulogiannis, G. (2020). Reprogramming of fatty acid metabolism in cancer. *Br. J. Cancer* 122 (1), 4–22. doi:10.1038/s41416-019-0650-z

Li, Q., Bian, Y., and Li, Q. (2021). Down-regulation of TMPO-AS1 induces apoptosis in lung carcinoma cells by regulating miR-143-3p/CDK1 Axis. *Technol. Cancer Res. Treat.* 20, 1533033820948880. doi:10.1177/1533033820948880

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, Z., Jiang, M., Zhang, T., and Liu, S. (2021). GAS6-AS2 promotes hepatocellular carcinoma via miR-3619-5p/ARL2 Axis under insufficient radiofrequency ablation condition. *Cancer biother. Radiopharm.* 36 (10), 879–887. doi:10.1089/cbr.2019.3541

Li, Z., Sun, C., and Qin, Z. (2021). Metabolic reprogramming of cancer-associated fibroblasts and its effect on cancer cell reprogramming. *Theranostics* 11 (17), 8322–8336. doi:10.7150/thno.62378

Luo, J., Makhnin, A., Tobi, Y., Ahn, L., Hayes, S. A., Iqbal, A., et al. (2021). Erlotinib and trametinib in patients with EGFR-mutant lung adenocarcinoma and acquired resistance to a prior tyrosine kinase inhibitor. *JCO Precis. Oncol.* 5, 55–64. doi:10.1200/PO.20.00315

Mamareli, P., Kruse, F., Lu, C. W., Guderian, M., Floess, S., Rox, K., et al. (2021). Targeting cellular fatty acid synthesis limits T helper and innate lymphoid cell function during intestinal inflammation and infection. *Mucosal Immunol.* 14 (1), 164–176. doi:10.1038/s41385-020-0285-7

Martínez-Reyes, I., and Chandel, N. S. (2021). Cancer metabolism: Looking forward. *Nat. Rev. Cancer* 21 (10), 669–680. doi:10.1038/s41568-021-00378-6

Mu, X., Wu, H., Liu, J., Hu, X., Wu, H., Chen, L., et al. (2020). Long noncoding RNA TMPO-AS1 promotes lung adenocarcinoma progression and is negatively regulated by miR-383-5p. *Biomed. Pharmacother.* 125, 109989. doi:10.1016/j.biopha.2020.109989

Newman, A. C., and Maddocks, O. D. K. (2017). One-carbon metabolism in cancer. *Br. J. Cancer* 116 (12), 1499–1504. doi:10.1038/bjc.2017.118

Ni, Y., Liu, F., Hu, X., Qin, Y., and Zhang, Z. (2021). Coding and non-coding RNA interactions reveal immune-related pathways in peripheral blood mononuclear cells derived from patients with proliferative vitreoretinopathy. *BMC Med. Genomics* 14 (1), 30. doi:10.1186/s12920-021-00875-5

Nicholson, A. G., Tsao, M. S., Beasley, M. B., Borczuk, A. C., Brambilla, E., Cooper, W. A., et al. (2021). The 2021 WHO classification of lung tumors: Impact of advances since 2015. *J. Thorac. Oncol.* 17, 362–387. doi:10.1016/j.jtho.2021.11.003

Pao, W., and Girard, N. (2011). New driver mutations in non-small-cell lung cancer. *Lancet. Oncol.* 12 (2), 175–180. doi:10.1016/S1470-2045(10)70087-5

Powers, R. K., Goodspeed, A., Pielke-Lombardo, H., Tan, A. C., and Costello, J. C. (2018). GSEA-InContext: Identifying novel and common patterns in expression experiments. *Bioinformatics* 34 (13), i555–i564. doi:10.1093/bioinformatics/bty271

Rami-Porta, R., Call, S., Dooms, C., Obiols, C., Sanchez, M., Travis, W. D., et al. (2018). Lung cancer staging: A concise update. *Eur. Respir. J.* 51 (5), 1800190. doi:10.1183/13993003.00190-2018

Remon, J., Pignataro, D., Novello, S., and Passiglia, F. (2021). Current treatment and future challenges in ROS1- and ALK-rearranged advanced non-small cell lung cancer. *Cancer Treat. Rev.* 95, 102178. doi:10.1016/j.ctrv.2021.102178

Ribeiro, M., Knebel, F. H., Bettoni, F., Saddi, R., Sacardo, K. P., Canedo, F. S. N. A., et al. (2021). Impressive response to dabrafenib, trametinib, and osimertinib in a metastatic EGFR-mutant/BRAF V600E lung adenocarcinoma patient. *NPJ Precis. Oncol.* 5 (1), 5. doi:10.1038/s41698-021-00149-4

Schild, T., Low, V., Blenis, J., and Gomes, A. P. (2018). Unique metabolic adaptations dictate distal organ-specific metastatic colonization. *Cancer Cell* 33 (3), 347–354. doi:10.1016/j.ccell.2018.02.001

Sidorova, M., and Petrikaitė, V. (2022). The effect of beta adrenoreceptor blockers on viability and cell colony formation of non-small cell lung cancer cell lines A549 and H1299. *Molecules* 27 (6), 1938. doi:10.3390/molecules27061938

Sumi, T., Hirai, S., Yamaguchi, M., Tanaka, Y., Tada, M., Niki, T., et al. (2018). Trametinib downregulates survivin expression in RB1-positive KRAS-mutant lung adenocarcinoma cells. *Biochem. Biophys. Res. Commun.* 501 (1), 253–258. doi:10.1016/j.bbrc.2018.04.230

Tan, Y. T., Lin, J. F., Li, T., Li, J. J., Xu, R. H., and Ju, H. Q. (2021). LncRNA-mediated posttranslational modifications and reprogramming of energy metabolism in cancer. *Cancer Commun.* 41 (2), 109–120. doi:10.1002/cac2.12108

Thai, A. A., Solomon, B. J., Sequist, L. V., Gainor, J. F., and Heist, R. S. (2021). Lung cancer. *Lancet* 398 (10299), 535–554. doi:10.1016/S0140-6736(21)00312-3

Wang, L., Cho, K. B., Li, Y., Tao, G., Xie, Z., and Guo, B. (2019). Long noncoding RNA (lncRNA)-Mediated competing endogenous RNA networks provide novel potential biomarkers and therapeutic targets for colorectal cancer. *Int. J. Mol. Sci.* 20 (22), E5758. doi:10.3390/ijms20225758

Wang, Q., Li, M., Yang, M., Yang, Y., Song, F., Zhang, W., et al. (2020). Analysis of immune-related signatures of lung adenocarcinoma identified two distinct subtypes: Implications for immune checkpoint blockade therapy. *Aging (Albany NY)* 12 (4), 3312–3339. doi:10.18632/aging.102814

Ward, P. S., and Thompson, C. B. (2012). Metabolic reprogramming: A cancer hallmark even warburg did not anticipate. *Cancer Cell* 21 (3), 297–308. doi:10.1016/j.ccr.2012.02.014

Wu, P., Mo, Y., Peng, M., Tang, T., Zhong, Y., Deng, X., et al. (2020). Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. *Mol. Cancer* 19 (1), 22. doi:10.1186/s12943-020-1147-3

Xing, C., Sun, S. G., Yue, Z. Q., and Bai, F. (2021). Role of lncRNA LUCAT1 in cancer. *Biomed. Pharmacother.* 134, 111158. doi:10.1016/j.biopha.2020.111158

Yang, S. Y., Li, Y., An, G. S., Ni, J. H., Jia, H. T., and Li, S. Y. (2018). DNA damage-response pathway heterogeneity of human lung cancer A549 and H1299 cells determines sensitivity to 8-chloro-adenosine. *Int. J. Mol. Sci.* 19 (6), E1587. doi:10.3390/ijms19061587

Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41, D955–D961. doi:10.1093/nar/gks1111

Yao, J., Chen, X., Liu, X., Li, R., Zhou, X., and Qu, Y. (2021). Characterization of a ferroptosis and iron-metabolism related lncRNA signature in lung adenocarcinoma. *Cancer Cell Int.* 21 (1), 340. doi:10.1186/s12935-021-02027-2

Ye, J., Chen, X., and Lu, W. (2021). Identification and experimental validation of immune-associate lncRNAs for predicting prognosis in cervical cancer. *Onco. Targets. Ther.* 14, 4721–4734. doi:10.2147/OTT.S322998

Yi, M., Li, J., Chen, S., Cai, J., Ban, Y., Peng, Q., et al. (2018). Emerging role of lipid metabolism alterations in Cancer stem cells. *J. Exp. Clin. Cancer Res.* 37 (1), 118. doi:10.1186/s13046-018-0784-5

You, R., DeMayo, F. J., Liu, J., Cho, S. N., Burt, B. M., Creighton, C. J., et al. (2018). IL17A regulates tumor latency and metastasis in lung adeno and squamous SQ.2b and AD.1 cancer. *Cancer Immunol. Res.* 6 (6), 645–657. doi:10.1158/2326-6066.CIR-17-0554

Zappa, C., and Mousa, S. A. (2016). Non-small cell lung cancer: Current treatment and future advances. *Transl. Lung Cancer Res.* 5 (3), 288–300. doi:10.21037/tlcr.2016.06.07

Zhang, G., Fan, E., Zhong, Q., Feng, G., Shuai, Y., Wu, M., et al. (2019). Identification and potential mechanisms of a 4-lncRNA signature that predicts prognosis in patients with laryngeal cancer. *Hum. Genomics* 13 (1), 36. doi:10.1186/s40246-019-0230-6

Zhang, H., Chen, X., Zhang, D., Liu, L., Song, J., Xu, Y., et al. (2020). Identification of a novel six-long noncoding RNA signature for molecular diagnosis of dilated cardiomyopathy. *DNA Cell Biol.* doi:10.1089/dna.2020.5670

Zhao, E., Hou, J., and Cui, H. (2020). Serine-glycine-one-carbon metabolism: Vulnerabilities in MYCN-amplified neuroblastoma. *Oncogenesis* 9 (2), 14. doi:10.1038/s41389-020-0200-9

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 20 October 2022
DOI 10.3389/fgene.2022.1016085

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Qifan Li,
Soochow University, China
Xialin Yan,
Tongji University, China

*CORRESPONDENCE
Guohua Zhou,
feishiqi1@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 10 August 2022
ACCEPTED 15 September 2022
PUBLISHED 20 October 2022

CITATION
Liu S, Zhao L and Zhou G (2022), Peripheral blood markers predict immunotherapeutic efficacy in patients with advanced non-small cell lung cancer: A multicenter study. *Front. Genet.* 13:1016085. doi: 10.3389/fgene.2022.1016085

Peripheral blood markers predict immunotherapeutic efficacy in patients with advanced non-small cell lung cancer: A multicenter study

Shuai Liu[1†], Liuyuan Zhao[1,2†] and Guohua Zhou[3,4*]

[1]Department of Anesthesiology, Affiliated Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China, [2]Department of Pain Medicine, Harbin Medical University Cancer Hospital, Harbin, China, [3]Department of Anesthesiology, Ningbo First Hospital, Zhejiang, China, [4]Department of Anesthesiology, The First Affiliated Hospital of Dalian Medical University, Dalian, China

This study aims to investigate the prognostic impact of peripheral blood markers in patients with advanced non-small cell lung cancer (NSCLC) undergoing immunotherapy. In the current multicenter study, 157 advanced NSCLC cases treated by immunotherapy at three institutions were included. Biochemical parameters in baseline peripheral blood were collected. The associations between biochemical parameters and prognosis were investigated by the Kaplan–Meier survival analyses and Cox regression, and the predictive performances of biomarkers were evaluated *via* receiver operating characteristic analysis. The neutrophil-to-lymphocyte ratio (NLR) (progression-free survival [PFS]: hazard ratio [HR], 1.766; 95% confidence interval [CI], 1.311–2.380; $p < 0.001$; overall survival [OS]: HR, 1.283; 95% CI, 1.120–1.469; $p < 0.001$) and red blood cell distribution width (RDW) (PFS: HR, 1.052; 95% CI, 1.005–1.102; $p = 0.031$; OS: HR, 1.044; 95% CI, 1.001–1.091; $p = 0.042$) were revealed as independent predictors for both PFS and OS. In addition, NLR ≥3.79 (1-year PFS, 24.2% [95% CI, 15.2%–38.4%] *versus* 27.3% [95% CI, 18.2%–41.1%], $p = 0.041$; 1-year OS, 44.2% [95% CI, 32.5%–60.1%] *versus* 71.8% [95% CI, 60.6%–85.2%], $p < 0.001$) or RDW ≥44.8 g/L (1-year PFS, 19.2% [95% CI, 11.4%–32.3%] *versus* 31.7% [95% CI, 21.9%–46.0%], $p = 0.049$; 1-year OS, 54.0% [95% CI, 42.7%–68.3%] *versus* 63.1% [95% CI, 50.6%–78.6%], $p = 0.014$) was significantly correlated to poorer PFS and OS than NLR < 3.79 or RDW <44.8 g/L. Moreover, NLR and RDW achieved areas under the curve with 0.651 (95% CI, 0.559–0.743) and 0.626 (95% CI, 0.520–0.732) for predicting PFS, and 0.660 (95% CI, 0.567–0.754) and 0.645 (95% CI, 0.552–0.739), for OS. Therefore, PLR and RDW could help predict the immunotherapeutic efficacy of advanced NSCLC.

KEYWORDS

non-small cell lung cancer, biomarker, peripheral blood, neutrophil-to-lymphocyte ratio, red blood cell distribution width

Introduction

Immune checkpoint inhibitors (ICIs), which target programmed cell death 1 (PD-1) and its ligand (PD-L1), are capable of inducing sustained antitumor effects, ushering in the therapeutic era for multiple malignant neoplasms (Okazaki et al., 2013; Ribas and Wolchok, 2018). In spite of this significant breakthrough, potent immunotherapeutic responses were only observed in approximately 20% advanced non-small cell lung cancer (NSCLC) population (Borghaei et al., 2015; Brahmer et al., 2015; Reck et al., 2016). Hence, precise recognition of patients who have the potential to derive additional benefits from ICIs is essential for the personalized treatment of advanced NSCLC.

Several biomarkers for immunotherapeutic efficacy of advanced NSCLC, such as tumor mutation burden (TMB), PD-L1, and tumor-infiltrating lymphocytes, have been revealed in previous publications (Kerr et al., 2015; Meng et al., 2015; High TMB Predicts Immunotherapy Benefit, 2018). However, in the current clinical practice, recognizing these signatures primarily resorts to core biopsy, which is unable to quantify the whole heterogeneity of tumors attributable to the limited specimens and simultaneously brings about a significant morbidity risk considering its invasive manipulation (Kerr et al., 2015; McLaughlin et al., 2016). As a result, a reliable and noninvasive instrument to predict the immunotherapeutic efficacy of advanced NSCLC is urgently needed.

Previous studies indicated that tumor-related inflammation played an important role in regulating tumor progression and immune infiltration (Jomrich et al., 2021). Moreover, biochemical parameters in peripheral blood provide a convenient and cost-effective path for reflecting the inflammatory status and their predictive potentials for immunotherapeutic efficacy have been investigated in various types of cancers receiving ICIs (Fukui et al., 2019; Nenclares et al., 2021; Valero et al., 2021). However, pieces of evidence for the value of biochemical parameters in peripheral blood in advanced NSCLC are insufficient. Therefore, this study, based on a multicenter population, proposes to explore the associations between pretreatment peripheral blood markers and prognosis in advanced NSCLC populations treated with ICIs.

Materials and methods

Study population

Approval of the institutional review boards and ethics committees of Harbin Medical University Cancer Hospital, Affiliated Drum Tower Hospital, and The First Affiliated Hospital of Dalian Medical University and a waiver for informed consent were obtained. Consecutive advanced NSCLC patients who underwent ICIs treatment in the abovementioned institutions between January 2016 to December 2020 were reviewed (Figure 1). Patients were included in this study when meeting the following criteria: 1) pathologically confirmed NSCLC; 2) stage III–IV; 3) administration of treatment ICIs, regardless of pretreatment line. The exclusion criteria included incomplete baseline data and lost to follow-up. All patients completed the follow-up survey before August 2022.

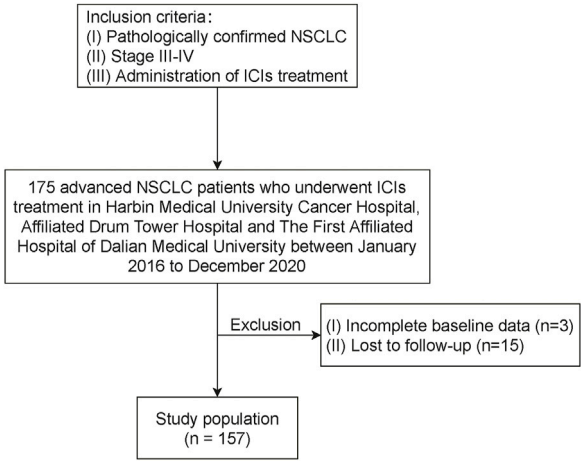


FIGURE 1
Flow chart illustrating patient inclusion. NSCLC, non-small cell lung cancer; ICI, immune checkpoint inhibitor.

Data collection

Clinicopathologic information was collected from electronic medical systems. Follow-up data were obtained through outpatient visits and telephone surveys. Overall survival (OS) was determined as the interval from initial ICI treatment to death or last follow-up. Progression-free survival (PFS) was calculated as the duration between initial ICI treatment and disease progress, death, or last follow-up.

Baseline peripheral blood samples were acquired within 7 days before immunotherapy, and routine blood biochemical parameters were collected. The inflammatory indexes were obtained based on the following formula: platelet-to-lymphocyte ratio (PLR) and absolute platelet count/absolute lymphocyte count; neutrophil-to-lymphocyte ratio (NLR) and absolute neutrophil count/absolute lymphocyte count; derived NLR (dNLR) and absolute neutrophil count/(white blood cell count-absolute neutrophil count); monocyte-to-lymphocyte ratio (MLR) and absolute monocyte count/absolute lymphocyte count; and systemic immune-inflammation index (SII), absolute neutrophil count×absolute platelet count/absolute lymphocyte count.

TABLE 1 Baseline characteristics of patients.

Characteristics	Entire cohort (n = 157)	PD + SD (n = 117)	PR + CR (n = 40)	p value
Age (years), mean ± SD	60.82 ± 10.45	60.59 ± 10.89	61.50 ± 9.16	0.636
Sex, n (%)				0.562
Male	97 (61.78)	71 (60.68)	25 (62.50)	
Female	60 (38.22)	46 (39.21)	15 (37.50)	
Smoking, n (%)				0.030
Ever	95 (60.51)	65 (55.56)	30 (75.00)	
Never	62 (39.49)	52 (44.44)	10 (25.00)	
ECOG PS, n (%)				0.551
0	12 (7.64)	7 (5.98)	5 (12.50)	
1	136 (86.62)	103 (88.03)	33 (82.50)	
2	8 (5.10)	6 (5.13)	2 (5.00)	
3	1 (0.64)	1 (0.85)	0 (0.00)	
Stage, n (%)				0.335
III	23 (14.65)	38 (32.48)	11 (27.50)	
IV	134 (85.35)	67 (57.26)	27 (67.50)	
Histology, n (%)				0.431
Squamous cell carcinoma	94 (59.87)	67 (57.26)	27 (67.50)	
Adenocarcinoma	49 (31.21)	38 (32.48)	11 (27.50)	
Others	14 (8.90)	12 (10.26)	2 (5.00)	
Treatment line, n (%)				0.856
First	84 (53.50)	62 (53.00)	22 (55.00)	
Not first	73 (46.50)	55 (47.00)	18 (45.00)	
Peripheral blood index				
PLR, mean ± SD	194.97 ± 103.36	196.66 ± 100.63	190.00 ± 112.17	0.726
NLR, mean ± SD	4.47 ± 2.86	4.68 ± 3.03	3.85 ± 2.24	0.113
dNLR, mean ± SD	2.75 ± 1.74	2.80 ± 1.68	2.61 ± 1.93	0.546
MLR, mean ± SD	0.47 ± 0.31	0.49 ± 0.33	0.40 ± 0.22	0.112
SII, mean ± SD	1,182.14 ± 943.78	1,215.66 ± 1,000.62	1,084.06 ± 756.06	0.448
HGB (g/L), mean ± SD	120.25 ± 19.10	118.56 ± 19.51	125.21 ± 17.14	0.057
RBC (10^{12}/L), mean ± SD	4.11 ± 0.61	4.06 ± 0.62	4.24 ± 0.56	0.122
WBC (10^{9}/L), mean ± SD	8.35 ± 3.33	8.39 ± 3.56	8.25 ± 2.56	0.820
NEUT%, mean ± SD	68.55 ± 12.27	68.93 ± 12.58	67.46 ± 11.40	0.514
LYM%, mean ± SD	20.28 ± 9.24	19.70 ± 9.12	21.98 ± 9.50	0.179
MONO%, mean ± SD	8.08 ± 3.28	8.09 ± 3.20	8.06 ± 3.55	0.962
EOS%, mean ± SD	2.36 ± 2.86	2.40 ± 3.15	2.24 ± 1.79	0.755
BASO%, mean ± SD	0.29 ± 0.28	0.27 ± 0.19	0.34 ± 0.44	0.202
PLT (10^{9}/L), mean ± SD	263.72 ± 96.82	259.38 ± 103.41	276.43 ± 73.91	0.338
HCT (%), mean ± SD	36.96 ± 5.34	36.47 ± 5.44	38.41 ± 4.83	0.047
MCV (fL), mean ± SD	90.26 ± 5.42	90.03 ± 5.69	90.92 ± 4.52	0.369
MCH (pg), mean ± SD	29.34 ± 1.94	29.24 ± 2.04	29.62 ± 1.62	0.287
MCHC (g/L), mean ± SD	324.82 ± 12.23	324.55 ± 13.45	325.60 ± 7.71	0.640
RDW (g/L), mean ± SD	46.10 ± 5.94	46.55 ± 5.98	44.80 ± 5.73	0.110
PCT (%), mean ± SD	26.01 ± 9.21	25.50 ± 9.69	27.53 ± 7.56	0.230
PDW (fL), mean ± SD	11.18 ± 1.77	11.16 ± 1.86	11.23 ± 1.53	0.830
P-LCR (%), mean ± SD	24.39 ± 7.47	24.30 ± 7.56	24.65 ± 7.26	0.804

PD, progressive disease; SD, stable disease; PR, partial response; CR, complete response; ECOG PS, Eastern Cooperative Oncology Group Performance Status; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; SII, systemic immune-inflammation index; HGB, hemoglobin; RBC, red blood cell count; WBC, white blood cell count; NEUT, neutrophil; LYM, lymphocyte; MONO, monocyte; EOS, eosinophils; BASO, basophil granulocytes; PLT, platelet; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red blood cell distribution width; PCT, plateletcrit; PDW, platelet distribution width; P-LCR, platelet-larger cell ratio; SD, standard deviation.

Statistical analysis

Pearson's chi-squared test and Student's *t*-test were implemented to compare the categorical and continuous parameters, respectively. Cox regressions and Kaplan–Meier survival analyses were conducted to recognize predictors for OS and PFS *via* the backward stepwise selection. The abovementioned statistical analyses were done using SPSS (version 23.0, IBM, Armonk, NY, United States) R software (version 4.1.1, http://www.R-project.org). A p value less than 0.05 was considered statistically significant.

Results

Clinicopathologic characteristics

The clinicopathologic characteristics were displayed in Table 1. The entire cohort included 97 (61.78%) men and 60 (38.22%) women, and the mean age for the whole population was 60.82 years. Smoking history was identified in 95 (60.51%) patients. ECOG PS 1 (n = 136, 86.62%) accounted for the largest proportion. Most patients were diagnosed as stage IV (n = 134, 85.35%) and squamous cell carcinoma (n = 94, 59.87%). Regarding the peripheral blood indexes, the mean level of PLR, NLR, dNLR, MLR, SII, hemoglobin (HGB), red blood cell count (RBC), white blood cell count (WBC), percentage of neutrophil (NEUT%), percentage of lymphocyte (LYM%), percentage of monocyte (MONO), percentage of eosinophils (EOS), percentage of basophil granulocytes (BASO), platelet (PLT), hematocrit (HCT), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), red blood cell distribution width (RDW), plateletcrit (PCT), platelet distribution width (PDW), and platelet-larger cell ratio (P-LCR) were 194.97, 4.47, 2.75, 0.47, 1,182.14, 120.25 g/L, 4.11 × 10^{12}/L, 8.35 × 10^{9}/L, 68.55%, 20.28%, 8.08%, 2.36%, 0.29%, 263.72 × 10^{9}/L, 36.96%, 90.26 fL, 29.34 ng, 324.82 g/L, 46.10 g/L, 26.01%, 11.18 fL and 24.39%. In addition, in subgroup analyses between 117 patients evaluated as progressive disease (PD) or stable disease (SD) and 40 patients with partial response (PR) or complete response (CR), patients classified as PR or CR were associated with a significantly higher proportion of smoking history (75% *versus* 55.56%, p = 0.030) and a higher level of HCT (38.41% *versus* 36.47%, p = 0.047).

Prognostic impact of peripheral blood markers

In the Cox survival analyses (Table 2), smoking history (hazard ratio [HR], 0.457; 95% confidence interval [CI], 0.306–0.683; p < 0.001), ECOG PS ≥ 1 (HR, 3.040; 95% CI, 1.386–6.668; p = 0.006), stage IV (HR, 0.465; 95% CI, 0.267–0.812; p = 0.007), NLR (HR, 1.766; 95% CI, 1.311–2.380; p < 0.001), dNLR (HR, 0.489; 95% CI, 0.321–0.744; p = 0.001), MLR (HR, 0.203; 95% CI, 0.044–0.929; p = 0.040), HGB (HR, 0.002; 95% CI, 0.001–0.171; p = 0.010), HCT (HR, 2.220; 95% CI, 1.183–4.166; p = 0.013), MCV (HR, 0.678; 95% CI, 0.525–0.876; p = 0.003), and RDW (HR, 1.052; 95% CI, 1.005–1.102; p = 0.031) were independent predictors for PFS. Similarly, smoking history (HR, 0.440; 95% CI, 0.250–0.775; p = 0.004), stage IV (HR, 0.445; 95% CI, 0.209–0.947; p = 0.036), NLR (HR, 1.283; 95% CI, 1.120–1.469; p < 0.001), MCH (HR, 0.852; 95% CI, 0.752–0.965; p = 0.012), and RDW (HR, 1.044; 95% CI, 1.001–1.091; p = 0.042) independently predicted OS.

As illustrated in Figure 2, ever-smoking patients achieved significantly better PFS (1-year PFS, 31.2% [95% CI, 22.2%–43.9%] *versus* 16.3% [95% CI, 8.3%–31.7%], p = 0.003) and OS (1-year OS, 64.0% [95% CI, 53.8%–76.3%] *versus* 49.1% [95% CI, 34.8%–69.1%], p = 0.042) compared with never-smoking patients. However, ECOG PS and stage failed to stratify the prognosis after immunotherapy. Moreover, as shown in Figure 3, by utilizing the median value as the cut-off, NLR ≥3.79 (1-year PFS, 24.2% [95% CI, 15.2%–38.4%] *versus* 27.3% [95% CI, 18.2%–41.1%], p = 0.041; 1-year OS, 44.2% [95% CI, 32.5%–60.1%] *versus* 71.8% [95% CI, 60.6%–85.2%], p < 0.001) or RDW ≥44.8 g/L (1-year PFS, 19.2% [95% CI, 11.4%–32.3%] *versus* 31.7% [95% CI, 21.9%–46.0%], p = 0.049; 1-year OS, 54.0% [95% CI, 42.7%–68.3%] *versus* 63.1% [95% CI, 50.6%–78.6%], p = 0.014) was significantly correlated to poorer PFS and OS than NLR< 3.79 or RDW< 44.8 g/L. In addition, patients with dNLR ≥2.41 (1-year OS, 48.5% [95% CI, 36.6%–64.4%] *versus* 67.8% [95% CI, 56.3%–81.5%], p = 0.013) or HGB <120 g/L (1-year OS, 51.1% [95% CI, 39.6%–65.9%] *versus* 67.6% [95% CI, 56.0%–81.7%], p = 0.046) showed inferiority only in OS than those with dNLR< 2.41 or HGB ≥120 g/L. However, other blood biochemical parameters did not stratify the prognosis of NSCLC receiving immunotherapy.

Predictive performance of peripheral blood markers

Considering PLR and RDW were two independent inflammatory biomarkers for both PFS and OS, the receiver operating characteristic analysis was implemented to quantify the predictive performance of PLR and RDW (Figure 4). For predicting PFS, NLR and RDW achieved areas under the curves (AUCs) with 0.651 (95% CI, 0.559–0.743) and 0.626 (95% CI, 0.520–0.732). Similarly, in the prediction for OS, the performances of NLR and RDW were shown to have AUCs of 0.660 (95% CI, 0.567–0.754) and 0.645 (95% CI, 0.552–0.739).

TABLE 2 Cox analyses for progression-free survival and overall survival.

Variables	Progression-free survival				Overall survival			
	Univariable		Multivariable		Univariable		Multivariable	
	HR (95% CI)	*p* value	HR (95% CI)	*p* value	HR (95% CI)	*p* value	HR (95% CI)	*p* value
Age	0.998 (0.980–1.018)	0.872			1.001 (0.974–1.030)	0.917		
Sex (Male)	0.649 (0.407–1.036)	0.070			0.613 (0.329–1.141)	0.123		
Smoking history (Ever)	0.562 (0.384–0.822)	0.003	0.457 (0.306–0.683)	< 0.001	0.627 (0.369–1.066)	0.085	0.440 (0.250–0.775)	0.004
ECOG PS (≥1)	1.796 (0.870–3.709)	0.113	3.040 (1.386–6.668)	0.006	1.474 (0.583–3.731)	0.413		
Stage (IV)	0.664 (0.399–1.105)	0.115	0.465 (0.267–0.812)	0.007	0.698 (0.340–1.433)	0.327	0.445 (0.209–0.947)	0.036
Histology (SCC)	0.687 (0.470–1.005)	0.053			0.689 (0.406–1.169)	0.167		
PLR	1.001 (1.000–1.003)	0.135			1.002 (1.000–1.004)	0.040		
NLR	1.077 (1.012–1.147)	0.020	1.766 (1.311–2.380)	< 0.001	1.162 (1.076–1.255)	< 0.001	1.283 (1.120–1.469)	< 0.001
dNLR	1.023 (0.935–1.121)	0.617	0.489 (0.321–0.744)	0.001	1.127 (1.008–1.260)	0.035	0.823 (0.650–1.041)	0.103
MLR	2.526 (1.433–4.452)	0.001	0.203 (0.044–0.929)	0.040	2.845 (1.392–5.816)	0.004		
SII	1.000 (1.000–1.000)	0.060			1.000 (1.000–1.001)	< 0.001		
HGB	0.990 (0.979–1.000)	0.053	0.002 (0.001–0.171)	0.010	0.980 (0.966–0.994)	0.005		
RBC	0.807 (0.593–1.099)	0.173			0.663 (0.434–1.012)	0.057		
WBC	1.029 (0.970–1.093)	0.342			1.070 (0.994–1.153)	0.073		
NEUT%	1.007 (0.991–1.022)	0.403			1.023 (1.000–1.047)	0.053		
LYM%	0.978 (0.958–0.999)	0.044			0.956 (0.926–0.987)	0.006		
MONO%	1.017 (0.966–1.071)	0.528			0.971 (0.897–1.051)	0.466		
EOS%	1.043 (0.967–1.126)	0.271			0.950 (0.837–1.077)	0.421		
BASO%	0.611 (0.297–1.258)	0.181			0.349 (0.086–1.413)	0.140		
PLT	1.000 (0.998–1.002)	0.773			1.001 (0.998–1.004)	0.512		
HCT	0.962 (0.927–0.998)	0.037	2.220 (1.183–4.166)	0.013	0.937 (0.893–0.983)	0.008		
MCV	0.968 (0.934–1.003)	0.069	0.678 (0.525–0.876)	0.003	0.961 (0.918–1.007)	0.093		
MCH	0.911 (0.824–1.008)	0.072			0.853 (0.759–0.960)	0.008	0.852 (0.752–0.965)	0.012
MCHC	0.999 (0.980–1.018)	0.894	1.020 (0.999–1.041)	0.065	0.974 (0.949–0.998)	0.036		
RDW	1.016 (0.987–1.047)	0.282	1.052 (1.005–1.102)	0.031	1.038 (0.998–1.079)	0.064	1.044 (1.001–1.091)	0.042
PCT	0.998 (0.976–1.020)	0.844			1.010 (0.980–1.041)	0.509		
PDW	0.959 (0.864–1.064)	0.431			0.977 (0.845–1.129)	0.751	0.577 (0.308–1.080)	0.085
P-LCR	0.988 (0.964–1.012)	0.333			1.002 (0.969–1.036)	0.915	1.148 (0.993–1.327)	0.062
CEA	0.796 (0.386–1.634)	0.533			0.865 (0.367–2.130)	0.755		

ECOG PS, Eastern Cooperative Oncology Group Performance Status; SCC, squamous cell carcinoma; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; SII, systemic immune-inflammation index; HGB, hemoglobin; RBC, red blood cell count; WBC, white blood cell count; NEUT, neutrophil; LYM, lymphocyte; MONO, monocyte; EOS, eosinophils; BASO, basophil granulocytes; PLT, platelet; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW, red blood cell distribution width; PCT, plateletcrit; PDW, platelet distribution width; P-LCR, platelet-larger cell ratio; CEA, carcinoembryonic antigen; HR, hazard ratio; CI, confidence interval.

Discussion

Despite immunotherapy having revolutionized the treatment paradigms of NSCLC (Okazaki et al., 2013; Ribas and Wolchok, 2018), the low response rate, therapy-related adverse effects, and high medical expense emphasize the significance of biomarkers for immunotherapeutic efficacy (Borghaei et al., 2015; Brahmer et al., 2015; Reck et al., 2016). In this study based on a multicenter population, we demonstrated that higher NLR and RDW in baseline peripheral blood were significantly correlated with poor PFS and OS in NSCLC patients undergoing ICIs treatment.

Previously, a number of studies have made investigations on this topic and revealed that TMB, PD-L1, and tumor-infiltrating lymphocytes derived from core biopsy specimens were correlated with immunotherapy prognosis of NSCLC (Kerr et al., 2015; Meng et al., 2015; High TMB Predicts Immunotherapy Benefit, 2018). However, these biomarkers suffered from biopsy-related morbidities due to their invasive nature. To overcome this limitation, further studies found that these markers in the peripheral blood also hold the potential to predict immunotherapy efficacy (Gandara et al., 2018; Wang et al., 2019; Bratman et al., 2020). Despite this breakthrough, these

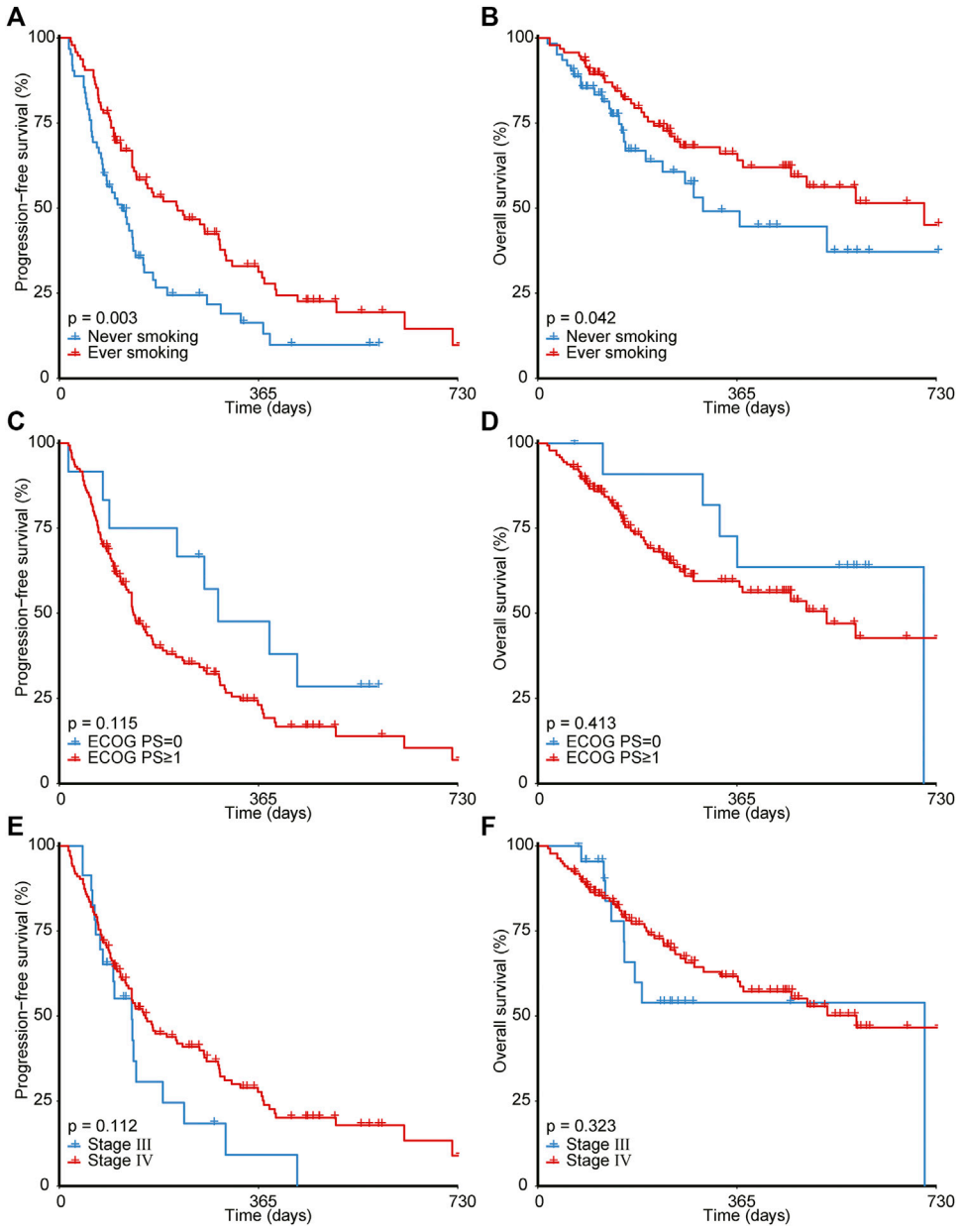


FIGURE 2
Survival analyses for patients with different smoking history **(A,B)**, ECOG PS **(C,D)**, and tumor stage **(E,F)**. ECOG PS, Eastern Cooperative Oncology Group Performance Status.

blood biomarkers were quantified based on peripheral blood mononuclear cells, which are too costly and time-consuming to acquire. In contrast, peripheral blood markers derived from routine complete blood count (CBC) are easily accessible and cost-effective, and thereby could be utilized as a convenient instrument in routine clinical practice.

Findings in our study were in line with previous publications that higher NLR was an adverse factor for the prognosis of NSCLC receiving immunotherapy (Fukui et al., 2019; Valero et al., 2021). In addition, Diem et al. (2017) concluded that PLR also played an important role in predicting immunotherapy response and prognosis and NSCLC patients with higher PLR tended to have an inferior prognosis. However, our study failed to validate the predictive efficiency of PLR: we speculated it might be attributable to that our study also included other biochemical parameters in the routine peripheral blood examination. Interestingly, we proved that increment of RDW significantly predicted poorer PFS and OS in NSCLC treated by immunotherapy, which was also observed in diffuse

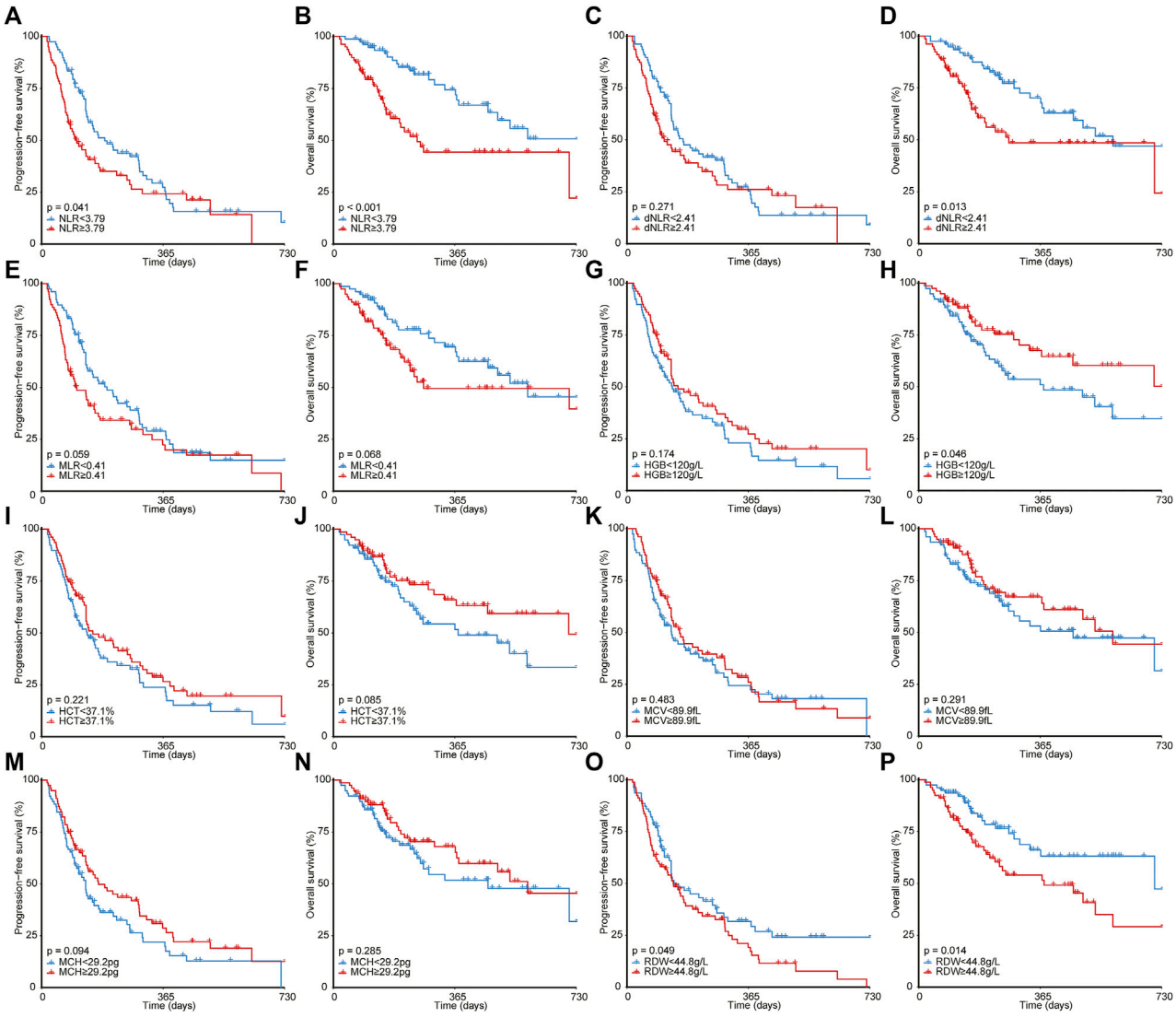


FIGURE 3
Survival analyses for patients with different NLR **(A,B)**, dNLR **(C,D)**, MLR **(E,F)**, HGB **(G,H)**, HCT **(I,J)**, MCV **(K,L)**, MCH **(M,N)**, and RDW **(O,P)**. NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; HGB, hemoglobin; HCT, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; RDW, red blood cell distribution width.

large B-cell lymphoma receiving immunotherapy (Beltran et al., 2019), but limited previous studies demonstrated its value in the NSCLC population. As such, we first indicated the capability of RDW as the biomarker for immunotherapeutic efficacy, and this finding might imply further insight into the prediction of immunotherapy response.

In addition to clinical implications, it is important to understand the biological basis underlying the prediction of NLR and RDW. The predictive mechanism of NLR might be rooted in its contributions to an immunosuppressive tumor microenvironment. On the one hand, as neutrophils were capable of releasing components mediating immunosuppression and tumor angiogenesis, neutrophil infiltration, thereby, established a microenvironment promoting cancer initiation, proliferation, and metastasis (Gonzalez et al., 2018; Shaul and Fridlender, 2019). On the other hand, reduced densities of lymphocyte infiltration contributed to the decreased response of antitumor T-cell, and the high level of neutrophils might further restrain T-cell response (Restifo et al., 2012; Zito Marino et al., 2017).

RDW, as an indicator representing the variations in the shape and size of red blood cells, is easily accessible in a routine CBC examination. The increased level of RDW implies a sign of impairments in erythropoiesis and red blood cell metabolism. The mechanism underlying the correlation of RDW with

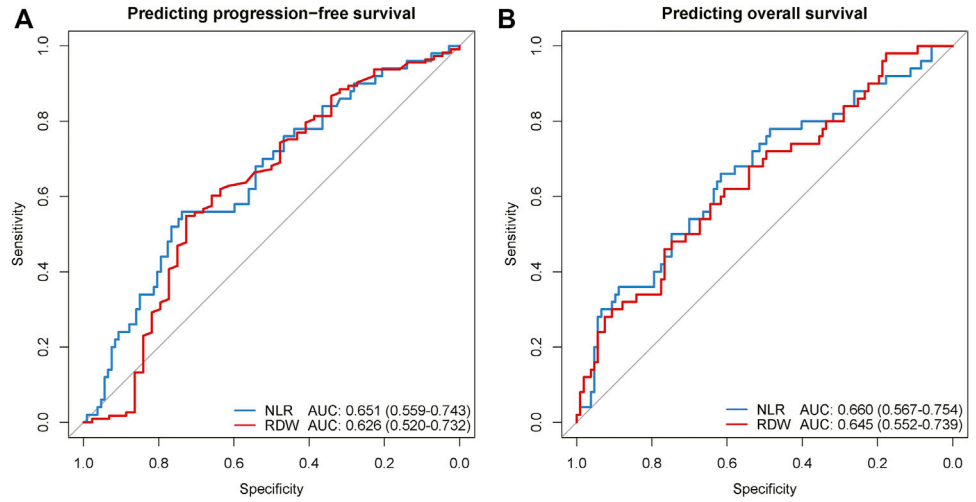


FIGURE 4
The ROC curves of NLR and RDW for predicting regression-free survival **(A)** and overall survival **(B)**. ROC, receiver operating characteristic; AUC, area under the curve of the receiver operating characteristic; NLR, neutrophil-to-lymphocyte ratio; RDW, red blood cell distribution width.

immunotherapeutic efficacy has not been clarified. However, several publications revealed that increasing RDW might result from oxidative stress, inflammation, and poor nutritional status *via* variation of erythropoiesis (Salvagno et al., 2015), and emerging pieces of evidence indicate that RDW was an adverse predictor for the prognosis of multiple malignancies (Koma et al., 2013; Albayrak et al., 2014; Ay et al., 2015).

Still, several limitations existed in the current study. First, despite the inclusion of a multicenter population, this study was limited by its retrospective nature, which suffered from selection bias and potential confounders. We utilized the multivariable regression to adjust prognostic predictors, but the impact of some known biomarkers, such as TMB, could not be evaluated. Thus, future prospective studies are required to validate our conclusions. Second, the small sample size reduces the power of the current study, and to be confirmed, further follow-up studies enrolling a larger sample size need to be performed. Finally, the underlying mechanism of the biomarkers has not been elucidated, and future studies focusing on the biological basis of NLR and RDW are warranted.

Conclusion

Our study demonstrated that NLR and RDW in baseline peripheral blood could help stratify the prognosis of advanced NSCLC patients receiving immunotherapy. Thus, NLR and RDW harbor the potential to serve as effective biomarkers for immunotherapeutic efficacy in NSCLC.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by the institutional review boards of Harbin Medical University Cancer Hospital, Affiliated Drum Tower Hospital, and The First Affiliated Hospital of Dalian Medical University. The ethics committee waived the requirement of written informed consent for participation.

Author contributions

GZ conceived the original idea. LZ collected the clinical data. SL analyzed the clinical data and wrote the manuscript. LZ helped modify the manuscript. All authors read and approved the final 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

References

Albayrak, S., Zengin, K., Tanik, S., Bakirtas, H., Imamoglu, A., and Gurdal, M. (2014). Red cell distribution width as a predictor of prostate cancer progression. *Asian pac. J. Cancer Prev.* 15 (18), 7781–7784. doi:10.7314/apjcp.2014.15.18.7781

Ay, S., Eryilmaz, M. A., Aksoy, N., Okus, A., Unlu, Y., and Sevinc, B. (2015). Is early detection of colon cancer possible with red blood cell distribution width? *Asian pac. J. Cancer Prev.* 16 (2), 753–756. doi:10.7314/apjcp.2015.16.2.753

Beltran, B. E., Paredes, S., Castro, D., Cotrina, E., Sotomayor, E. M., and Castillo, J. J. (2019). High red cell distribution width is an adverse predictive and prognostic factor in patients with diffuse large B-cell lymphoma treated with chemoimmunotherapy. *Clin. Lymphoma Myeloma Leuk.* 19 (9), e551–e557. doi:10.1016/j.clml.2019.06.005

Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., Ready, N. E., et al. (2015). Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* 373 (17), 1627–1639. doi:10.1056/NEJMoa1507643

Brahmer, J., Reckamp, K. L., Baas, P., Crino, L., Eberhardt, W. E. E., Poddubskaya, E., et al. (2015). Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N. Engl. J. Med.* 373 (2), 123–135. doi:10.1056/NEJMoa1504627

Bratman, S. V., Yang, S. Y. C., Iafolla, M. A. J., Liu, Z., Hansen, A. R., Bedard, P. L., et al. (2020). Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat. Cancer* 1 (9), 873–881. doi:10.1038/s43018-020-0096-5

Diem, S., Schmid, S., Krapf, M., Flatz, L., Born, D., Jochum, W., et al. (2017). Neutrophil-to-Lymphocyte ratio (NLR) and Platelet-to-Lymphocyte ratio (PLR) as prognostic markers in patients with non-small cell lung cancer (NSCLC) treated with nivolumab. *Lung Cancer* 111, 176–181. doi:10.1016/j.lungcan.2017.07.024

Fukui, T., Okuma, Y., Nakahara, Y., Otani, S., Igawa, S., Katagiri, M., et al. (2019). Activity of nivolumab and utility of neutrophil-to-lymphocyte ratio as a predictive biomarker for advanced non-small-cell lung cancer: A prospective observational study. *Clin. Lung Cancer* 20 (3), 208–214. doi:10.1016/j.cllc.2018.04.021

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

Gonzalez, H., Hagerling, C., and Werb, Z. (2018). Roles of the immune system in cancer: From tumor initiation to metastatic progression. *Genes Dev.* 32 (19-20), 1267–1284. doi:10.1101/gad.314617.118

High TMB predicts immunotherapy benefit. *Cancer Discov.*, 2018. 8(6): p. 668, doi:10.1158/2159-8290.CD-NB2018-048

Jomrich, G., Paireder, M., Kristo, I., Baierl, A., Ilhan-Mutlu, A., Preusser, M., et al. (2021). High systemic immune-inflammation index is an adverse prognostic factor for patients with gastroesophageal adenocarcinoma. *Ann. Surg.* 273 (3), 532–541. doi:10.1097/SLA.0000000000003370

Kerr, K. M., Tsao, M. S., Nicholson, A. G., Yatabe, Y., Wistuba, I. I., Hirsch, F. R., et al. (2015). Programmed death-ligand 1 immunohistochemistry in lung cancer: In what state is this art? *J. Thorac. Oncol.* 10 (7), 985–989. doi:10.1097/JTO.0000000000000526

Koma, Y., Onishi, A., Matsuoka, H., Oda, N., Yokota, N., Matsumoto, Y., et al. (2013). Increased red blood cell distribution width associates with cancer stage and prognosis in patients with lung cancer. *PLoS One* 8 (11), e80240. doi:10.1371/journal.pone.0080240

McLaughlin, J., Han, G., Schalper, K. A., Carvajal-Hausdorf, D., Pelekanou, V., Rehman, J., et al. (2016). Quantitative assessment of the heterogeneity of PD-L1 expression in non-small-cell lung cancer. *JAMA Oncol.* 2 (1), 46–54. doi:10.1001/jamaoncol.2015.3638

Meng, X., Huang, Z., Teng, F., Xing, L., and Yu, J. (2015). Predictive biomarkers in PD-1/PD-L1 checkpoint blockade immunotherapy. *Cancer Treat. Rev.* 41 (10), 868–876. doi:10.1016/j.ctrv.2015.11.001

Nenclares, P., Gunn, L., Soliman, H., Bover, M., Trinh, A., Leslie, I., et al. (2021). On-treatment immune prognostic score for patients with relapsed and/or metastatic head and neck squamous cell carcinoma treated with immunotherapy. *J. Immunother. Cancer* 9 (6), e002718. doi:10.1136/jitc-2021-002718

Okazaki, T., Chikuma, S., Iwai, Y., Fagarasan, S., and Honjo, T. (2013). A rheostat for immune responses: The unique properties of PD-1 and their advantages for clinical application. *Nat. Immunol.* 14 (12), 1212–1218. doi:10.1038/ni.2762

Reck, M., Rodriguez-Abreu, D., Robinson, A. G., Hui, R., Csoszi, T., Fulop, A., et al. (2016). Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N. Engl. J. Med.* 375 (19), 1823–1833. doi:10.1056/NEJMoa1606774

Restifo, N. P., Dudley, M. E., and Rosenberg, S. A. (2012). Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat. Rev. Immunol.* 12 (4), 269–281. doi:10.1038/nri3191

Ribas, A., and Wolchok, J. D. (2018). Cancer immunotherapy using checkpoint blockade. *Science* 359 (6382), 1350–1355. doi:10.1126/science.aar4060

Salvagno, G. L., Sanchis-Gomar, F., Picanza, A., and Lippi, G. (2015). Red blood cell distribution width: A simple parameter with multiple clinical applications. *Crit. Rev. Clin. Lab. Sci.* 52 (2), 86–105. doi:10.3109/10408363.2014.992064

Shaul, M. E., and Fridlender, Z. G. (2019). Tumour-associated neutrophils in patients with cancer. *Nat. Rev. Clin. Oncol.* 16 (10), 601–620. doi:10.1038/s41571-019-0222-4

Valero, C., Lee, M., Hoen, D., Weiss, K., Kelly, D. W., Adusumilli, P. S., et al. (2021). Pretreatment neutrophil-to-lymphocyte ratio and mutational burden as biomarkers of tumor response to immune checkpoint inhibitors. *Nat. Commun.* 12 (1), 729. doi:10.1038/s41467-021-20935-9

Wang, Z., Duan, J., Cai, S., Han, M., Dong, H., Zhao, J., et al. (2019). Assessment of blood tumor mutational burden as a potential biomarker for immunotherapy in patients with non-small cell lung cancer with use of a next-generation sequencing cancer gene panel. *JAMA Oncol.* 5 (5), 696–702. doi:10.1001/jamaoncol.2018.7098

Zito Marino, F., Ascierto, P. A., Rossi, G., Staibano, S., Montella, M., Russo, D., et al. (2017). Are tumor-infiltrating lymphocytes protagonists or background actors in patient selection for cancer immunotherapy? *Expert Opin. Biol. Ther.* 17 (6), 735–746. doi:10.1080/14712598.2017.1309387

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 November 2022
DOI 10.3389/fgene.2022.966365

OPEN ACCESS

EDITED BY
Amitabha Chaudhuri,
SciGenom Labs, India

REVIEWED BY
Laura Sterian Ward,
State University of Campinas, Brazil
Wei Zhou,
Cancer Hospital, Chongqing University, China

*CORRESPONDENCE
Libo Chen,
lbchen@sjtu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 10 June 2022
ACCEPTED 28 October 2022
PUBLISHED 09 November 2022

CITATION
Jin Y, Qiu X, He Z, Wang J, Sa R and Chen L (2022), ERBB2 as a prognostic biomarker correlates with immune infiltrates in papillary thyroid cancer. *Front. Genet.* 13:966365. doi: 10.3389/fgene.2022.966365

ERBB2 as a prognostic biomarker correlates with immune infiltrates in papillary thyroid cancer

Yuchen Jin†, Xian Qiu†, Ziyan He†, JunYao Wang, Ri Sa and Libo Chen*

Department of Nuclear Medicine, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China

Epidermal growth factor receptor 2 (*ERBB2*) is commonly over-expressed in advanced or metastatic tissues of papillary thyroid cancer (PTC) with poor prognosis, while it remains unknown whether *ERBB2* plays a role in the progression of PTC. Thus, we analyzed the data derived from online repositories, including TCGA, KEGG, GO, GeneMANIA, and STRING, to explore the relationship between *ERBB2* expression and prognosis, tumor phenotypes of interest, and immune infiltrates in PTC. Compared to normal thyroid tissue, *ERBB2* was up-regulated in PTC samples ($p < 0.001$); In comparison with the group with low expression of *ERBB2*, the group with high expression of *ERBB2* had poorer progression-free interval in stage III/IV patients ($p = 0.008$) and patients aged >45 years ($p = 0.019$). The up-regulated *ERBB2* was associated with iodine metabolism dysfunction, proliferation, metastasis, angiogenesis, and drug resistance. The expression of *ERBB2* negatively correlated with enrichment scores of B cells ($r = -0.176$, $p < 0.001$), $CD8^+$ T cells ($r = -0.160$, $p < 0.001$), cytotoxic cells ($r = -0.219$, $p < 0.001$), NK CD56dim cells ($r = -0.218$, $p < 0.001$), plasmacytoid dendritic cells ($r = -0.267$, $p < 0.001$), T cells ($r = -0.164$, $p < 0.001$), T follicular helper cells ($r = -0.111$, $p = 0.012$), gamma delta T cells ($r = -0.105$, $p = 0.017$), and regulatory T cells ($r = -0.125$, $p = 0.005$). In conclusion, *ERBB2* may serve as a prognostic biomarker and an immunotherapeutic target in PTC, deserving further exploration.

KEYWORDS

ERBB2, prognosis, biomarker, immune microenviroment, papillary thyroid cancer

Introduction

Thyroid cancer is one of the most common endocrine neoplasms, and the incidence rate is overgrowing, with 53,815 new cases expected in the US (Siegel et al., 2021) and 224,023 in China (Xia et al., 2022) in 2022. Papillary thyroid cancer (PTC) is the most widely recognized thyroid tumor, accounting for around 84% of all thyroid cancer patients (Fagin and Wells, 2016). Alarmingly, the incidence of invasive and metastatic

PTC has risen 1.5–5-fold over the past 30 years (Lim et al., 2017). Within 10 years after initial surgery and radioiodine therapy, approximately 1/10 to 1/5 of PTCs developed local recurrence and distant metastasis; Once distant metastases of PTC occurred, the 5-year survival rate of patients would decrease from 98% to 63–75% (Cancer Facts 2022). The high rates of recurrence, metastasis, and radioiodine refractoriness have become the key bottleneck, stymying the cure of PTC (Jin et al., 2018).

Epidermal growth factor receptor 2 (*ERBB2*), a well-known oncogene of multiple cancers, such as breast cancer and ovarian cancer (Baxevanis et al., 2004), was found to be firmly connected with cancer occurrence, proliferation, metastasis, drug resistance, immune escape and poor prognosis (Dawood et al., 2009; Bates et al., 2018). Owe to the research on *ERBB2*, multiple diagnostic and therapeutic methods were established for tumor management (Mitri et al., 2012). In thyroid cancer, however, the deep-going research on *ERBB2* is limited. Although *ERBB2* is over-expressed in progressive and metastatic PTC cases (Kremser et al., 2003), and high expression of *ERBB2* was found to be responsible for resistance to mitogen-activated extracellular signal-regulated kinase inhibitors (MEKi) (Montero-Conde et al., 2013), the *ERBB2* upregulation-related unfavorable phenotypes, such as tumor occurrence, proliferation, metastasis, and poor progression-free interval (PFI), have not been comprehensively explored in PTC, and the relations between *ERBB2* expression and the immune microenvironment in PTC remains unknown.

Given that the application of *ERBB2*-related management approaches, such as imaging probes and immunotherapeutic drugs, is limited by lacking knowledge of the association between *ERBB2* expression and oncological features of PTC; we, therefore, conducted the bioinformatic study with integrated data from online repositories, including TCGA, KEGG, GO, GeneMANIA, and STRING to explore the potency and translational value of *ERBB2* as a prognostic biomarker and immunotherapeutic target for PTC management.

Materials and methods

Data collection

The mRNA data and clinical characteristics of PTC patients were downloaded from the TCGA database (https://www.cancer.gov/tcga/), an open data portal that compiled clinical information and RNA-Seq data of 33 cancers. The thyroid cancer (THCA) project of the TCGA database includes 568 samples (510 PTC and 58 paracancerous normal samples). Clinical data collection for the THCA project was mainly completed in 2014; thus, PTC staging was adopted in AJCC 7th edition (Agrawal et al., 2014). The mRNA sequencing data of all samples were converted to the format of transcripts per million (TPM) for subsequent analysis.

OncoPrint analysis online

OncoPrint is a method for visualizing samples by integrating gene expression heat map and gene variant distribution map. The gene expression and gene mutation data in OncoPrint analysis were derived from the TCGA-THCA dataset and analyzed online (https://www.cbioportal.org). The OncoPrint analysis was utilized to show the association between *ERBB2* expression and *BRAF* and *RAS* mutant PTC sample distribution.

Progression-free interval analysis

The Kaplan-Meier Progression-free interval (PFI) curve compared prognosis differences between patients with high and low expression of *ERBB2*. Since the death reports of PTC in the TCGA database were few (n = 16), the clinical prognosis data included in this study were PFI. Due to the limited number of patients with distant metastases in the database (n = 9), we could not analyze PFI in the subgroup of patients with distant metastases.

Co-expression heat map

Co-expression heat map assesses the correlation between the expression of *ERBB2* and gene sets, including the iodine metabolism-related gene set of *TSHR*, *SLC5A8*, *SLC26A4*, and *TPO* (Portulano et al., 2013), the tumor angiogenesis gene set of *VEGFA*, *FLT1*, *KDR*, *FLT4*, *PECAM1*, *VWF*, *TIE1*, *TEK*, *ANGPT1*, *ANGPT2*, *CDH5*, and *CLDN5* (Smith et al., 2010; Haibe et al., 2020), Lymph node metastasis gene set of *EVA1A*, *TIMP1*, *SERPINA1*, *FAM20A*, *FN1*, *TNC*, and *MXRA8* (Wu et al., 2021), distant metastatic gene sets of *MMP2*, *PTTG1*, *VEGFC*, *CXCR4*, and *FGF2*, tumor cell proliferation set of *MKI67*, *PCNA*, and *MCM2*) (Liang et al., 2011), and MEKi resistance marker gene set of *SPRY2*, *SPRY4*, *ETV4*, *ETV5*, *DUSP4*, *DUSP6*, *CCND1*, *EPHA2*, and *EPHA4* (Mazzoni et al., 2019; Degirmenci et al., 2021).

Analysis of immune infiltrates

The enrichment scores of 24 immune cells were based on the reported literature (Bindea et al., 2013). The correlation analysis between *ERBB2* expression and enrichment scores of immune infiltration was tested with a single sample gene set enrichment analysis (ssGSEA) with R packages of clusterProfiler (version 3.8.0).

Analysis of ERBB2-related differentially expressed genes (DEGs)

To further understand the role of *ERBB2* in the progression of PTC, we screened the DEGs between samples with high and

low expression of *ERBB2*. We applied Log_2(fold change) > 1.5 and adjusted $p < 0.001$ to select DEGs. The analysis is performed using the R package of DESeq2 (version 3.8). All the DEGs were graphed in a volcano plot.

Enrichment analysis of KEGG and GO terms

GO analysis is a widely applied bioinformatics tool for determining *ERBB2*-related biological processes, cellular components, and molecular functions. *KEGG* analysis is used to analyze *ERBB2*- related signal path changes. We applied GO and *KEGG* to analyze the biological function of *ERBB2* in PTC. To understand how *ERBB2* is involved in tumorigenesis, we used *KEGG* and GO online tools to analyze the signaling pathways and gene functions in which *ERBB2*-related DEGs participated. Gene set enrichment analysis (GSEA) with R package clusterProfiler (version 3.8.0) evaluated the *ERBB2*- related DEGs contributing to annotated gene functions, cell phenotypes, signal pathways, and diseases.

GeneMANIA and string analysis online

The network of gene-gene relations was constructed with the online tool GeneMANIA (https://genemania.org/). The website extensively integrated data on gene-gene interactions, gene co-expression networks, and gene function enrichment. To understand the *ERBB2*-involved mechanisms, the STRING tool (version 11.5, https://string-db.org/) was utilized to analyze the protein-protein interaction (PPI) network of corresponding proteins of *ERBB2*-related DEGs. The PPI network was displayed with Cytoscape software (version 3.9.1).

Statistical analysis

All statistical analyses and graphs were performed using R software (version 3.6.3). The *ERBB2* expression in unpaired and paired samples was analyzed with Wilcoxon rank-sum test and Wilcoxon signed-rank test, respectively. Kruskal-Wallis test was tested to assess the relationship between clinical/cytogenetic features and *ERBB2* expression. The differences in enrichment scores of immune infiltrate were analyzed using the Wilcoxon rank sum tests. The Spearman statistical method was adopted for the correlation analysis between ERBB2 expression and the gene expression of interest or enrichment scores of 24 immune cells. Hazard ratio, log-rank tests, and the Kaplan-Meier curve were applied to assess the role of *ERBB2* as a prognostic biomarker. $P < 0.05$ was considered statistically significant.

Results

Aberrantly upregulated *ERBB2* expression in PTC samples

By analyzing the RNA-Seq data downloaded from TCGA data, we found that the *ERBB2* expression in PTC samples was generally higher than that in normal thyroid tissue (unpaired test: $p < 0.001$, Figure 1A; paired test: $p < 0.001$, Supplementary Figure S1). OncopPrint plot displayed that PTC tissues with high expression of *ERBB2* were mainly distributed in *BRAF* mutation samples rather than *RAS* mutation samples (Figure 1B). Subgroup analysis showed that the *ERBB2* expression in *BRAF* mutant PTC, classic PTC, and PTC with bilateral foci were significantly higher than those in corresponding subgroups of *RAS* mutant PTC ($p < 0.001$), follicular variant PTC (FVPTC) ($p < 0.001$), and PTC with unilateral foci ($p = 0.02$), respectively (Figure 2). To determine the *ERBB2*-related clinical value and the underlying mechanism, PTC samples with high and low expression of *ERBB2* were divided based on the median TPM value. The characteristics of patients with the high and low expression of *ERBB2* were summarized in Supplementary Table S1.

Elevated *ERBB2* expression was associated with the poor prognosis

The relations between *ERBB2* expression and PFI were analyzed using the Kaplan-Meier curve. As is shown in Figure 3, the group with high expression of *ERBB2* had a significantly poorer prognosis than that with low expression of *ERBB2* in stage III/IV patients (HR: 2.7, CI: 1.29–5.66, $p = 0.008$) and patients aged >45 years (HR: 2.22, CI: 1.09–4.54, $p = 0.019$) (Figure 3); the detailed subgroup analyses of PFI Kaplan-Meier curves were plotted in Supplementary Figures S2–4.

Highly expressed *ERBB2* was correlated to unfavorable tumor phenotypes

The association between *ERBB2* expression and clinical phenotypes was performed with correlation analysis. The co-expression heat map showed that the expression of *ERBB2* positively correlated with *TSHR* but negatively correlated with *SLC5A8*, *SLC26A4*, and *TPO* (Figure 4A); In addition, the expression of *ERBB2* positively correlated with gene sets of tumor angiogenesis (*VEGFA*, *FLT*1, *KDR*, *FLT4*, *PECAM1*, *VWF*, *TIE1*, *TEK*, *ANGPT1*, *ANGPT2*, *CDH5*, and *CLDN5*) (Figure 4B), lymph node metastasis (Gene sets: *EVA1A*, *TIMP1*, *SERPINA1*, *FAM20A*, *FN1*, *TNC*, and *MXRA8*)

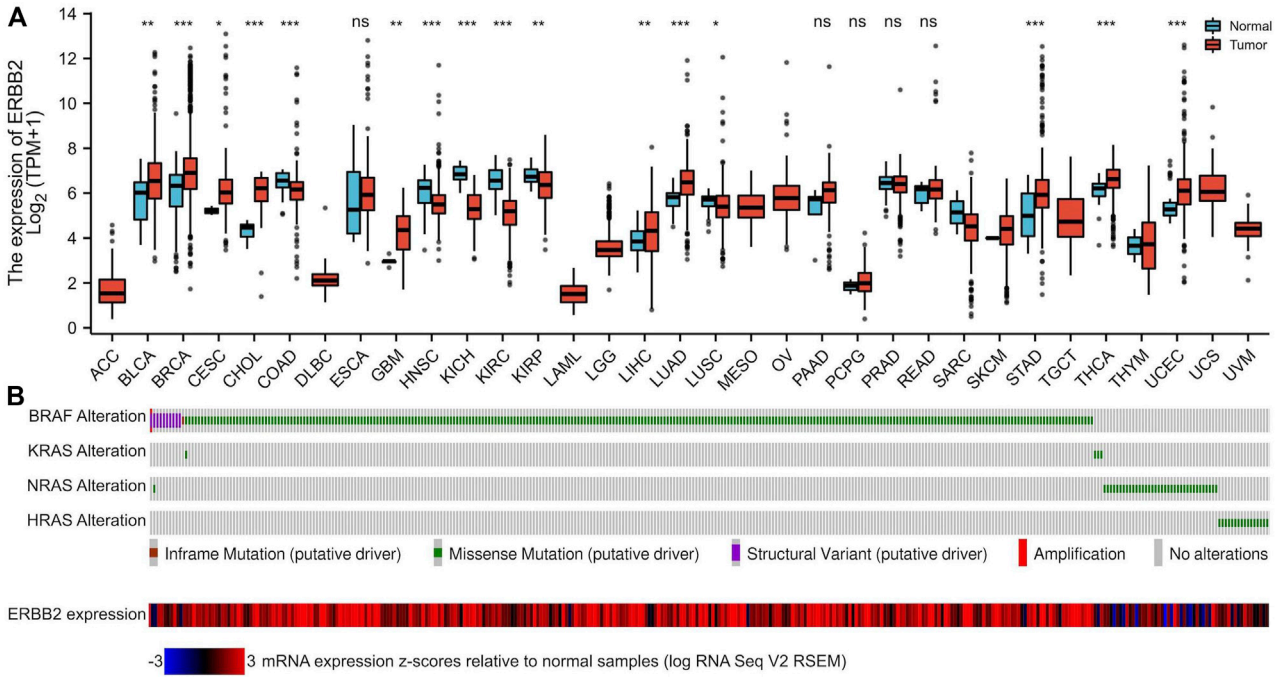


FIGURE 1
The mRNA expression of *ERBB2* in pan-cancers and tumor-adjacent normal samples. **(A)** Compared with paired normal tissues, mRNA expression of ERBB2 was upregulated in 16 cancer datasets: BLCA ($p = 0.007$), BRCA ($p < 0.001$), CESC ($p = 0.034$), CHOL ($p < 0.001$), GBM ($p = 0.002$), LIHC ($p = 0.008$), LUAD ($p < 0.001$), STAD ($p < 0.001$), and THCA ($p < 0.001$) and UCEC ($p < 0.001$). **(B)** The OncoPrint plot displays an overview of gene alterations of the *BRAF*, *KRAS*, *NRAS*, and *HRAS*, and the expression of *ERBB2* in papillary thyroid cancer samples from TCGA. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

(Figure 4C), distant metastases (Gene sets: *MMP2*, *VEGFC*, *CXCR4*, and *FGF2*) (Figure 4D), tumor cell proliferation (Gene sets: *MKI67*, *PCNA*, and *MCM2*) (Figure 4E), and MEKi resistance (Gene sets: *SPRY2*, *SPRY4*, *ETV4*, *ETV5*, *DUSP4*, *DUSP6*, *CCND1*, *EPHA2*, and *EPHA4*) (Figure 4F).

Up-regulated *ERBB2* was associated with suppressed tumor-infiltration of immune cells

The tumor-infiltrating immune cells were quantified by ssGSEA. Compared to the samples with low expression of *ERBB2*, the enrichment scores of T cells, $CD8^+$ T cells, cytotoxic cells, NK CD56dim cells, plasmacytoid dendritic (pDC) cells, gamma delta T cells (Tgd), and regulatory T cells (TReg) were lower in samples with high expression of *ERBB2* (Figure 5A). The correlation analysis showed that *ERBB2* expression was negatively correlated with the enrichment scores of B cells, $CD8^+$ T cells, cytotoxic cells, NK CD56dim cells, pDC, T cells, follicular helper T cells (TFH), Tgd cells, and TReg cells (Figure 5B, Supplementary Table S2).

Identification of DEGs between samples with high and low expression of *ERBB2*

A total of 146 DEGs between samples with high and low expression of *ERBB2* were yielded (Supplementary Table S3). The five most up-regulated genes and the five most downregulated genes were shown in the volcano plot (Figure 6). Spearman correlation analysis found that the expression of *ERBB2* negatively correlated with *KLK15*, *KLK1*, *ARSF*, and *FGF21*; and positively correlated with *TAGLN3*, *GLRA1*, *RSPO1*, *SPAG11B*, and *SPAG11A* (Supplementary Figure S5).

ERBB2-related DEGs participated in the primary immunodeficiency signaling pathway

Annotations from KEGG and GO were enriched with 146 *ERBB2*-related DEGs, showing that the *ERBB2*-related DEGs were involved in the KEGG term of primary immunodeficiency and GO term of the humoral immune response, regulation of execution phase of apoptosis, and

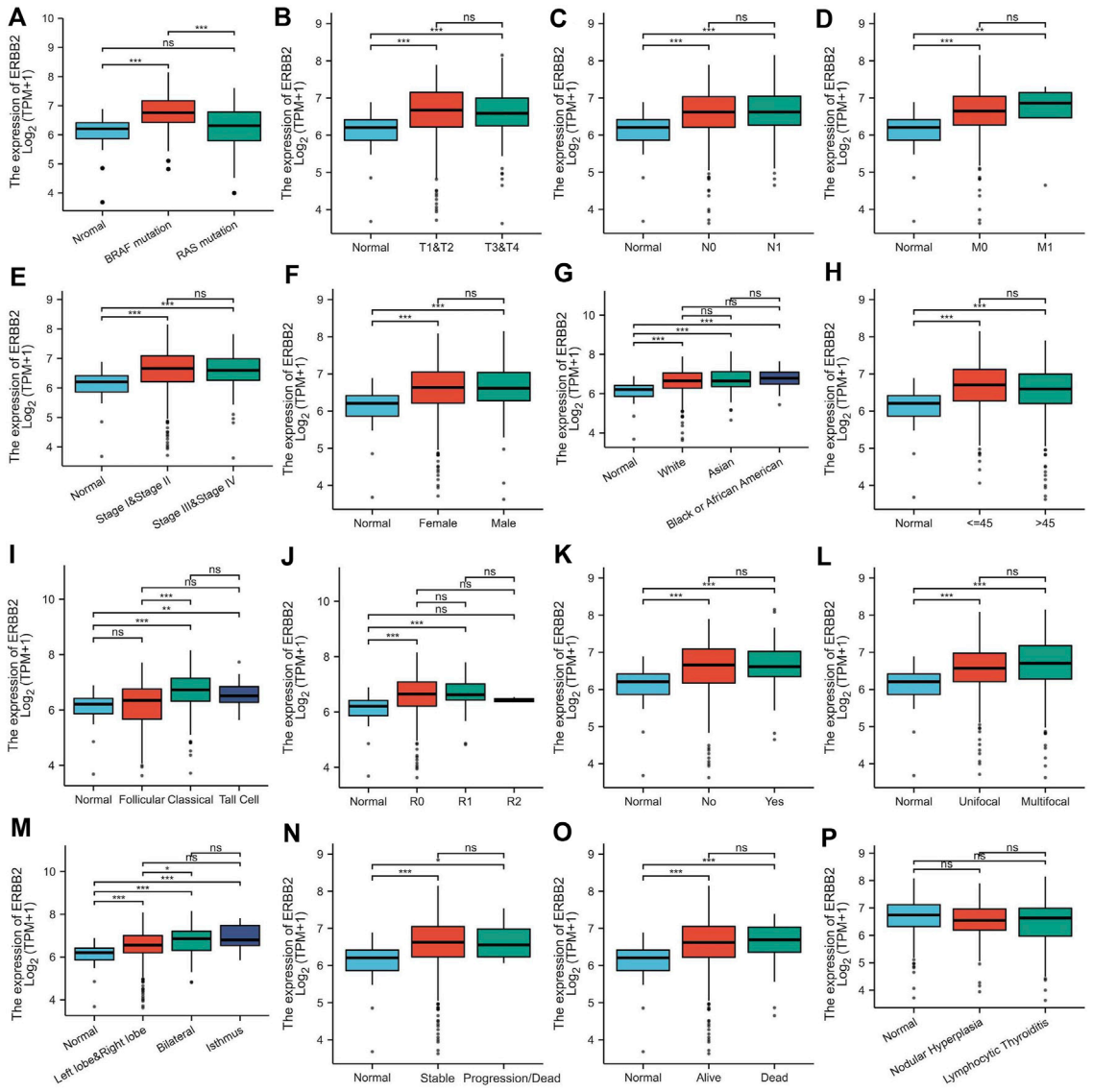


FIGURE 2
ERBB2 expression in papillary thyroid cancer classified by characteristics: **(A)** mutation, **(B)** T stage, **(C)** N stage, **(D)** M stage, **(E)** pathological stage, **(F)** gender, **(G)** race, **(H)** age, **(I)** histological type, **(J)** residual tumor, **(K)** extrathyroidal extension, **(L)** primary tumor focus type, **(M)** tumor location, **(N)** overall survival event, **(O)** progression-free interval event, and **(P)** thyroid gland disorder history. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

negative regulation of execution phase of apoptosis (Figure 7A). Thereinto, the majority of *ERBB2*-related DEGs (n = 23) were enriched into the primary immunodeficiency signaling pathway, attracting more attention. GSEA analysis adds more insight into the 23 *ERBB2*-related DEGs and primary immunodeficiency signaling pathway, showing that *ERBB2*-related DEGs associated with primary immunodeficiency were mainly down-regulated *ERBB2*-related DEGs, such as *CD79A*, *CD19*, *CD3D*, *CD3E*, *CD8A*, and *CD4* (Figure 7B).

ERBB2-related DEGs constructed the core PPI networks

Network analysis showed that the 23 *ERBB2*-related DEGs had gene interaction and co-expression relationships with each other; the functional enrichment found that the 23 *ERBB2*-related DEGs involved in the function of a variety of immune cells, including the functions of B cells, monocytes, and T cells (Supplementary Figure S6). The PPI network of corresponding proteins of the *ERBB2*-related DEGs was screened with the

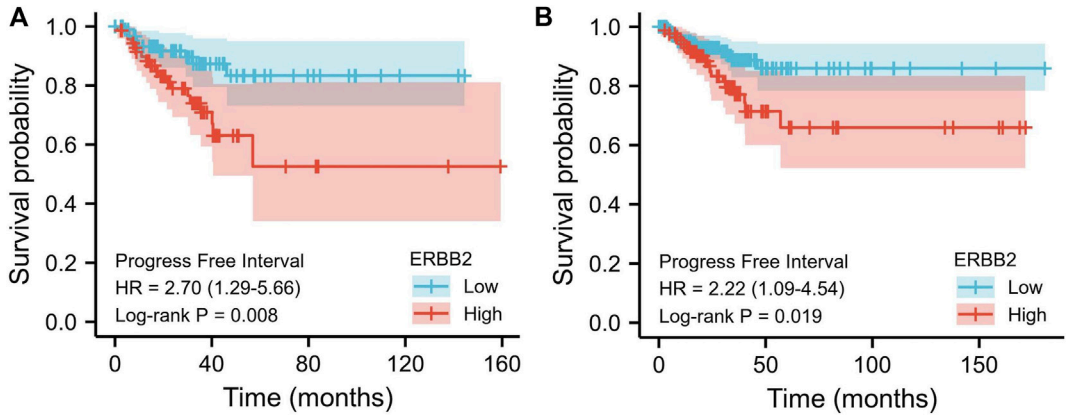


FIGURE 3
Kaplan-Meier curves of progression-free interval comparing the high and low expression of *ERBB2* in thyroid cancer patients. **(A)** In stage III/IV patients, the progression-free interval of the high *ERBB2* expression group was poorer than that of the low expression group (p = 0.008). **(B)** In patients aged >45 years, the progression-free survival of the high *ERBB2* expression group was poorer than that of the low expression group (p = 0.019).

STRING tool with a confidence threshold of 0.4. In total, 84 nodes and 127 edges were connected, and the networks with nodes ≥3 were displayed with Cytoscape (Supplementary Figure S7A). The core network was analyzed with the Cytoscape-MCODE function, showing the networks with an MCODE score of 4.815 (Supplementary Figure S7B) and 4.612 (Supplementary Figure S7C); the proteins of *NPY* and *ESR1* were located in the center of the PPI network (Supplementary Figure S7).

Discussion

Despite *ERBB2* being a well-known oncogene and regarded as a crucial diagnostic/prognostic biomarker and therapeutic target in multiple tumors, such as breast cancer (Wynn and Tang, 2022), the exact function of *ERBB2* in tumors remains largely unknown. Only a few data are available in PTC, showing an 18%–100% rate of positive expression of *ERBB2* protein *via* immunohistochemistry in 43–45 cases (Kremser et al., 2003; Ruggeri et al., 2016). Thus, expression profiles of the *ERBB2* gene in the PTC need to be comprehensively investigated in the context of a large-scale study. In the present study, we took advantage of a large-scale analysis of gene expression profiles in 510 PTC patients in the TCGA database. The bioinformatic analysis adds the knowledge of detailed relations between *ERBB2* expression and the clinical characteristics, prognosis, and immunoenvironment of PTC. Moreover, it is a novel finding that the expression levels of *ERBB2* were significantly correlated to the markers of the immune cells in the PTC tissues, suggesting that ERB2 may play a crucial role in regulating the tumor-immunoenvironment. The present study uncovered the biological and clinical roles of *ERBB2* in PTC, which might help imply novel ERBB2-based management strategies to improve the prognosis of PTC patients.

ERBB2 is up-regulated in multiple tumors located in breast, bladder, pancreas, ovary, and esophagus, especially in tumors with poor prognostic characteristics; in breast cancer, high expression of *ERBB2* has become one of the hallmarks of poor prognosis (Dawood et al., 2009). Similarly, our study showed that the expression of *ERBB2* in PTC was aberrantly up-regulated. A step further, we found that *ERBB2* expression in patients with bilateral PTCs was higher than that in those with unilateral foci, which may be related to the fact that bilateral PTC is more aggressive and metastatic than the PTC with unilateral foci (Qu et al., 2016). Meanwhile, we found that the expression of *ERBB2* in classic PTC was higher than that in FVPTC, which may be related to the poorer prognosis of PTC than that of FVPTC (Liu et al., 2018). The prognosis analysis in our study uncovered that the high expression of *ERBB2* is a risk factor for poor prognosis in PTC.

The highly expressed *ERBB2* was deemed to participate in cancer progression, promoting tumor proliferation, invasion, and metastasis (Mitri et al., 2012). PTC patients with high expression of *ERBB2* were prone to suffer from distant metastasis (Kremser et al., 2003). In our study, the expression of *ERBB2* was found to be positively correlated with VEGF-associated genes of *VEGFA*, *FLT*, *KDR*, and *FLT4* (Shibuya, 2011), vascular endothelial cell markers of *PECAM1* and *VWF* (Bauer et al., 2015), and vascular support-related genes of *TEK*, *ANGPT1*, *ANGPT2*, *CDH5*, *CLDN5*, and *JAM2* (Oshi et al., 2021). Furthermore, the expression of *ERBB2* was positively correlated with lymph node metastatic signature genes such as *EVA1A*, *TIMP1*, *SERPINA1*, *FAM20A*, *FN1*, *TNC*, and *MXRA8* (Wu et al., 2021), and distant metastatic signature genes of

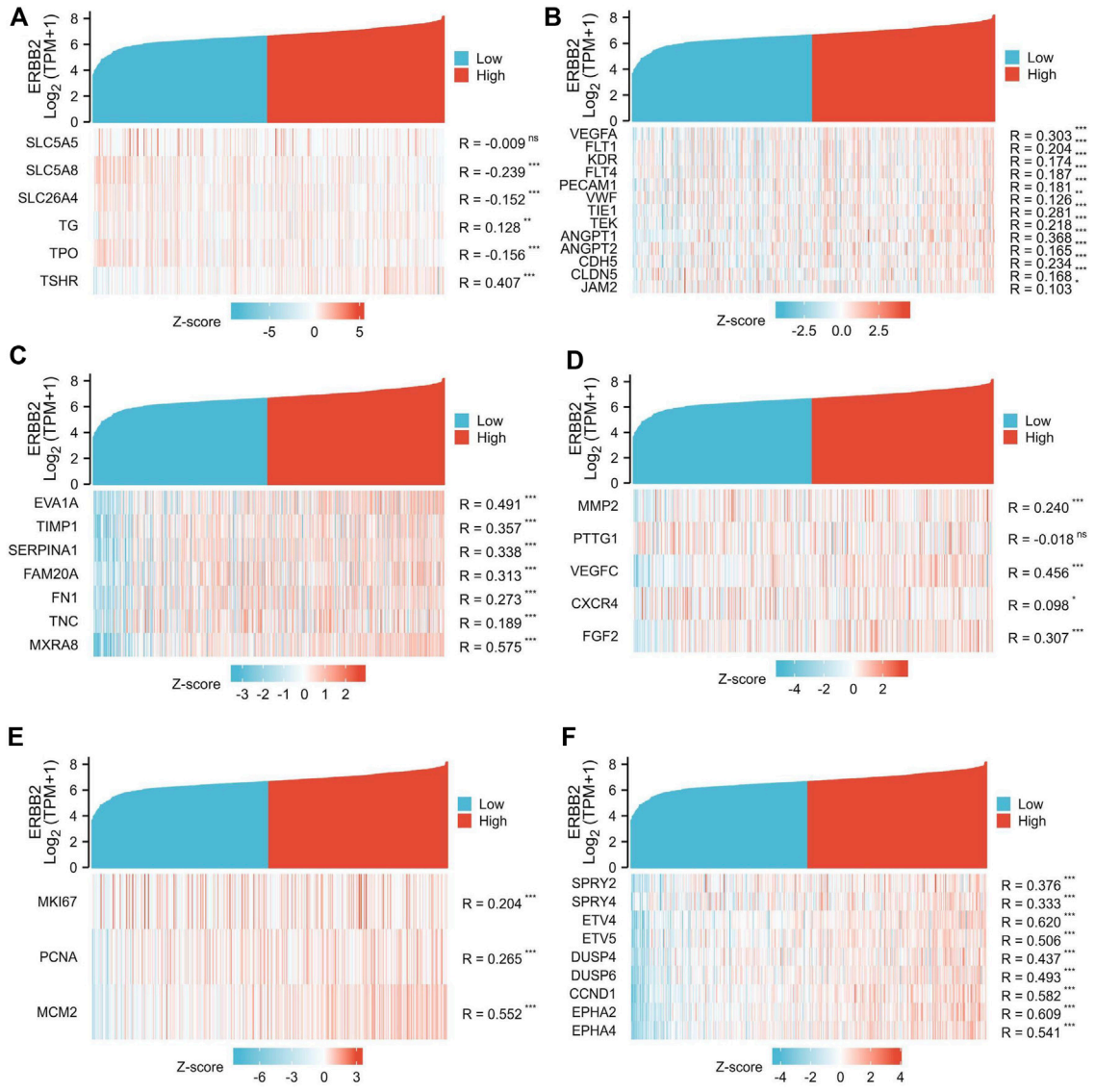


FIGURE 4
Co-expression analysis of *ERBB2* and genes related to unfavorable thyroid cancer phenotypes. **(A)** the expression of *ERBB2* is negatively correlated with iodine metabolism genes of *SLC5A8*, *SLC26A4*, and *TPO*. **(B)** the expression of *ERBB2* was positively correlated with the tumor angiogenesis gene set. **(C)** the expression of *ERBB2* is positively correlated with the gene set of lymph node metastasis. **(D)** the expression of *ERBB2* was positively correlated with the distant metastasis gene set. **(E)** the expression of *ERBB2* was positively correlated with the gene set of cell proliferation. **(F)** the expression of *ERBB2* was positively correlated with the MEK inhibitor resistance gene set. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

MMP2, *PTTG1*, *VEGFC*, *CXCR4*, and *FGF2* (Liang et al., 2011). Besides, *ERBB2* expression was also positively correlated with proliferation-related genes, such as *MKI67*, *PCNA*, and *MCM2* (Juríková et al., 2016). Collectively, highly expressed *ERBB2* is involved in progressive oncological behaviors of PTC, which might be responsible for the poor prognosis.

Interestingly, high expression of *ERBB2* was found to be associated with dysfunction of iodine metabolism of PTC in our study. The expression of *ERBB2* was negatively correlated with *SLC5A8*, *SCL26A4*, and *TPO*, which relate to iodine absorption, transport, and organization. The downregulated *SLC5A8* was associated with impaired resorption of organic iodine (Porra et al., 2005), making it challenging to transport iodine across the basal membrane to cell cytoplasm in thyroid follicular cells (Anekpuritanang et al., 2021). Meanwhile, *TPO* is the crucial protein to organize iodine with the participation of H_2O_2, and *TPO* deficiency would prevent iodine from taking part in the organization process (Kessler et al., 2008). Although the expression of *ERBB2* was positively correlated with that of *TSHR*, the role of *TSHR* in radioiodine refractoriness remains

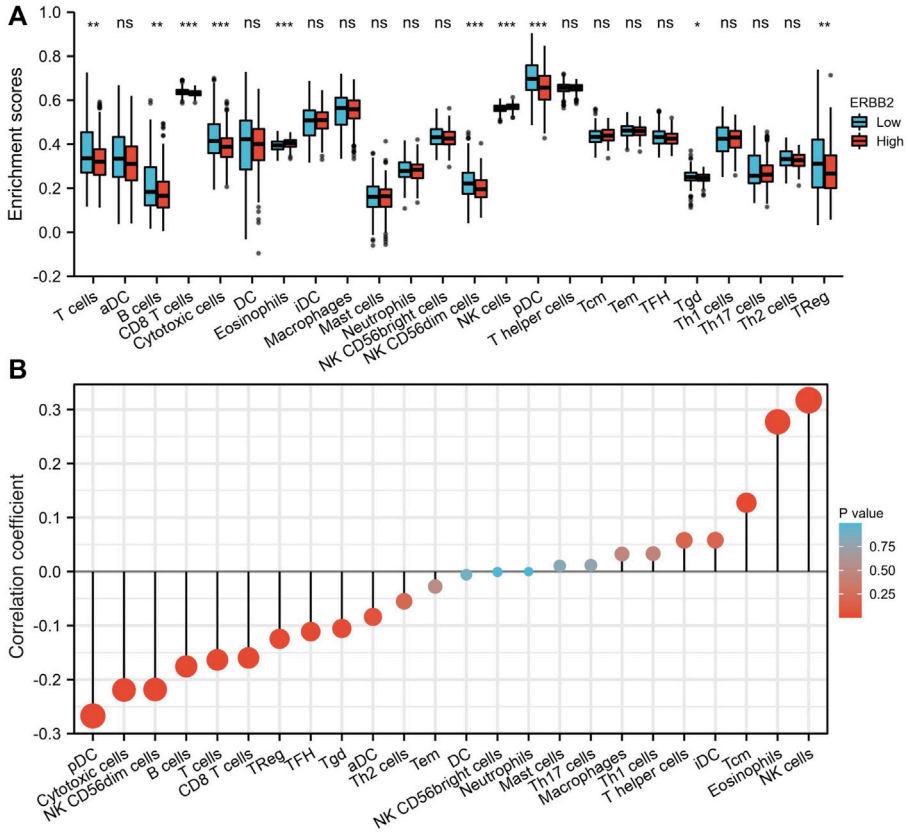


FIGURE 5
Immune cell enrichment scores in thyroid cancer tissues with different *ERBB2* expression. **(A)** The enrichment scores of T cells, B cells, $CD8^+$ T cells, cytotoxic cells, NK CD56dim cells, pDCs, Tgd, and TReg in the *ERBB2* high expression group were lower than those in the *ERBB2* low expression group. **(B)** Correlation analysis between the expression of *ERBB2* and the enrichment scores of immune infiltrating cells in the tumor. *ERBB2* expression negatively correlated with the enrichment scores of $CD8^+$ T cell, cytotoxic cell, NK CD56dim cell, pDC, T cell, TFH, Tgd, and TReg. DC, dendritic cell; aDC, activated dendritic cell; pDC, plasmacytoid dendritic cell; iDC, interdigitating dendritic cell; Tcm, central memory T cell; Tem, effector memory T cell; TFH, follicular helper T cell; Tgd, T gamma delta, γδ; Th1, T helper type 1; Th17, T helper type 17; Th2, T helper type 2; Treg, regulatory T cells; NK, natural killer. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

debated because the highly expressed *TSHR* is not only associated with high radioiodine uptake (Hou et al., 2010) but also causes fast thyroid tumor growth (Lu et al., 2010). Therefore, the high expression of *ERBB2* is likely to associate with the poor efficacy of radioactive iodine to some extent, and this mechanism needs to be further studied.

In recent years, targeted drugs, including commonly applied receptor tyrosine kinase inhibitors (RTKi) and novel MEKi, have become valuable antitumor drugs for advanced or progressive thyroid cancer (Jin et al., 2018). Notwithstanding the broad application of RTKi, the occurrence of RTKi-related severe adverse events is frequent, and drug resistance eventually develops in almost all PTC patients (Cheng et al., 2020). Compared to RTKi, the MEKi has the additional capability of inducing tumor differentiation, activating immune recognition, and possesses the characteristic of a lower incidence of severe adverse events (Neuzillet et al., 2014), attracting more attentions and expectations in the field of anti-PTC. Still, MEKi resistance is the typical scenario, often calling for a combination of MEKi and other therapeutics; nevertheless, the improvement of efficacy by the combination is generally limited (Shah et al., 2017), requiring a crucial key to solving the most pressing challenge. MEKi resistance has been proven to be closely related to the high expression of *ERBB2* (Sa et al., 2022). For instance, high *ERBB2* expression is one characteristic phenotype of thyroid cancer resistance to MEKi selumetinib (Montero-Conde et al., 2013). Our study further explored the correlation between the expression of *ERBB2* and MEKi resistance genes of *SPRY2*, *SPRY4*, *ETV4*, *ETV5*, *DUSP4*, *DUSP6*, *CCND1*, *EPHA2*, and *EPHA4* (Wagle et al., 2018). In *BRAF* or *RAS* mutant tumor cells, *DUSP4*, *DUSP6*, *SPRY2*, and *SPRY4* tend to be highly expressed, allowing tumors to evade regular MAPK signaling pathway feedback (Pratilas et al., 2009). Other genes, such as *ETV4*, *ETV5*, *CCND1*, *EPHA2*, and *EPHA4,* also play a role in

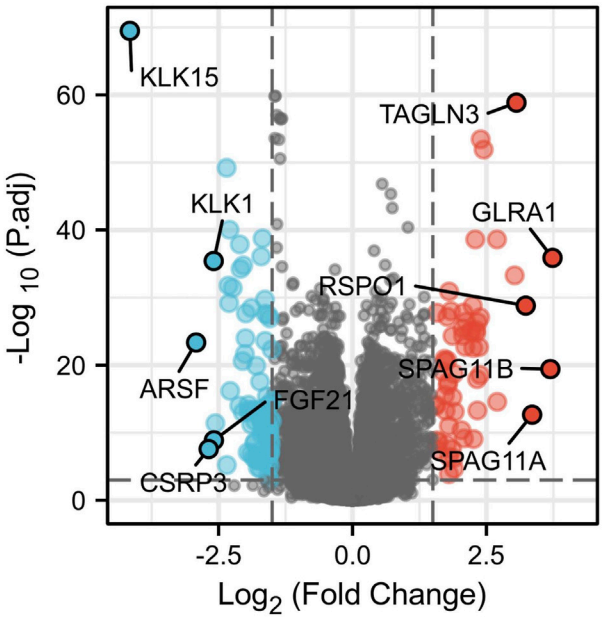


FIGURE 6
Volcano plot of differentially expressed genes (DEGs) between high and low *ERBB2* expression groups. The significantly up-regulated genes are red dots; the significantly down-regulated genes are blue dots. The dots marked with gene names are the most up-regulated five DEGs and the most down-regulated five DEGs.

activating the MAPK pathway, promoting tumor resistance to MEKi (Chesnokov et al., 2022). Together with the previous study (Montero-Conde et al., 2013), our study confirmed the relationship between highly expressed *ERBB2* and MEKi resistance in PTC.

Over the past decade, immune-related diagnostic techniques have successfully become new tools in tumor management. For instance, Michael D Farwell et al. conducted a phase I trial of PET/CT scan with ^{89}Zr-labeled IAB22M2C minibody targeting CD8$^+$ cells, which visualized the biodistribution of tumor-infiltrating CD8$^+$ T cells and predicted early response to immunotherapy (Farwell et al., 2021). In the present research, *ERBB2* expression was inversely correlated with the enrichment scores of CD8$^+$ T cells. Therefore, radioactive or non-radioactive probes, such as ^{68}Ga-labeled anti-ERBB2 Nanobody (Keyaerts et al., 2016) or polyethylene glycol-conjugated anti-ERBB2 peptides targeting ERBB2 protein (Guan et al., 2018) might be of value for predicting the response to immunotherapy in PTCs. In addition, ERBB2-targeted imaging would overcome the false positive uptake of ^{89}Zr-labeled IAB22M2C in CD8$^+$ normal tissues, such as the bone marrow and lymphnodes.

The finding, *ERBB2* overexpression related to suppressed T cell infiltration in tumors, is not unique to PTC. Liu et al. (2021) also found that *ERBB2* expression was negatively correlated with the infiltration of B cells and CD8$^+$ T cells in cutaneous melanoma, similar to our findings. In breast cancer, *ERBB2* is related to low expression of MHC class I surface antigen, causing impaired recognition between tumor cells and CD8$^+$ T cells (Herrmann et al., 2004; Seliger and Kiessling, 2013). Besides, the PD-1/PD-L1 antibodies have been used in clinical trials to treat advanced, progressive, or metastatic thyroid cancer. However, the results were unsatisfactory; most patients still have an unfavorable response to PD-1 antibody therapy (Mehnert et al., 2019), related to the limited T-cell infiltration in PTC (Bastman et al., 2016). DEG analysis showed that in PTC samples with high expression of *ERBB2*, *KLK15* and *KLK1* were downregulated. The down-regulated *KLK*s were previously reported to be responsible for impaired immune microenvironment reprogramming during an antitumor immune response (Srinivasan et al., 2022). In our research, KEGG and GO annotation analysis found that the DEGs were mainly associated with the primary immunodeficiency pathway, which is closely related to tumor occurrence and the failure of antitumor immunotherapy (Tangye et al., 2020). Therefore, increasing tumor-infiltrating immune cells by applying bispecific antibodies or adding cytokines, e.g., PD-L1×CD3 bispecific antibody (Yang et al., 2021) or interleukin-17 (Nagaoka et al., 2020), in *ERBB2* highly expressed tumors may be beneficial to improving the response to immunotherapy in PTC patients.

In addition, we found that *NPY* and *ESR1* were located at the center of the *ERBB2*-related PPI network. *NPY* is a pleiotropic gene initially thought to be an endogenous anxiolytic peptide whose expression can be regulated by stress; however, *NPY* has been found to promote the growth and migration of breast cancer cells in recent years (Lin et al., 2021). *ESR1* encodes an estrogen receptor, plays a crucial role in the occurrence and progression of breast and endometrial cancer, and is the main reason for resistance to estrogen suppression therapy (Piscuoglio et al., 2018). These results suggest that *ERBB2* may be involved in the occurrence and progression of PTC together with *NPY* and *ESR1*; the mechanism needs further investigation.

Despite the fact that our data analysis enhanced the understanding of the roles of *ERBB2* associated with PTC progression, there remain some limitations. First, the role of *ERBB2* in follicular thyroid cancer, poorly differentiated thyroid cancer, and anaplastic thyroid cancer could not be investigated because those pathological types or endpoint information were lacking in online databases. Second, despite large sample studies in public databases prone to possess the superiority of strong evidence, public databases remain to lack some clinical information on PTC tissue around the prognostic tumor status of radioiodine uptake, angiogenesis, proliferation, drug resistance, and serum thyroglobulin levels; Therefore, we cannot perform a clinically based analysis comparing those prognostic factors above, and directly investigating their relations with *ERBB2* expression.

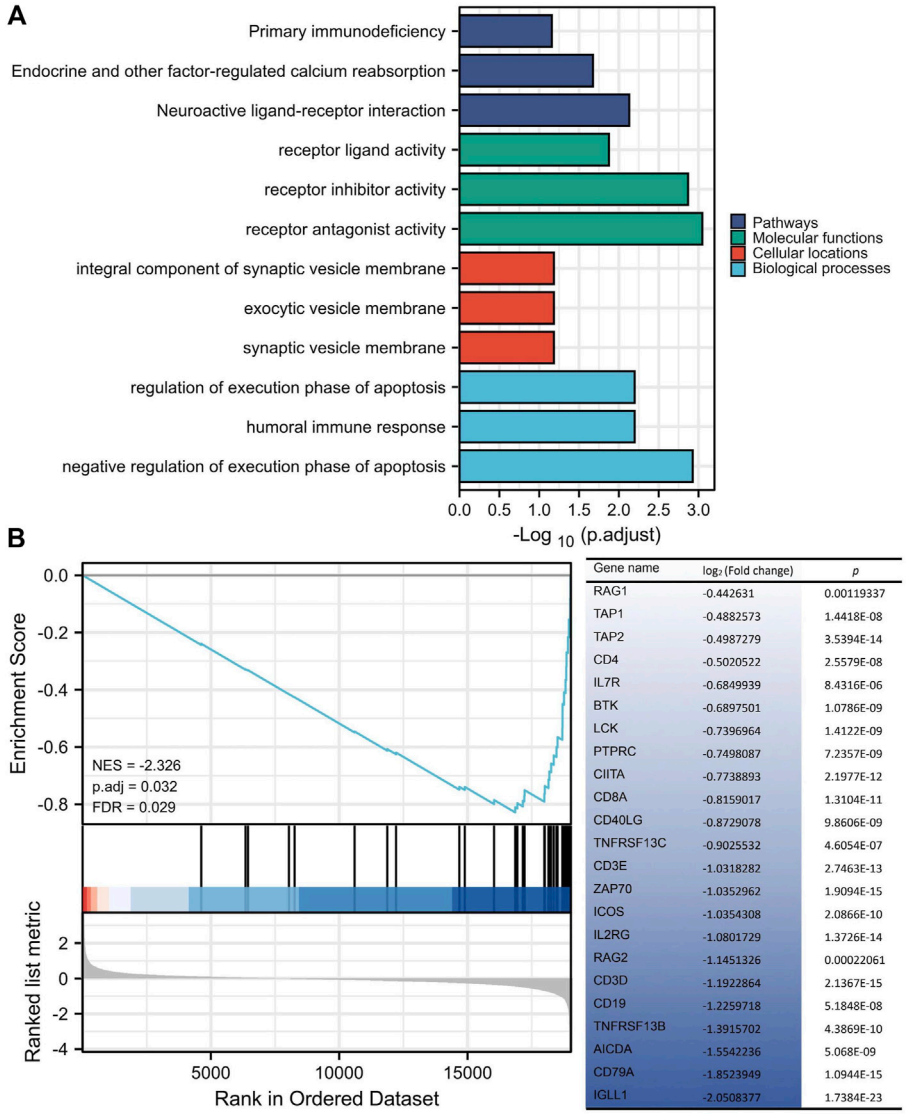


FIGURE 7
KEGG and GO enrichment analysis of differentially expressed genes between high and low expression of *ERBB2* groups. **(A)** 146 ERBB2-related differentially expressed genes were enriched in KEGG signaling pathways and GO annotations, such as the primary immunodeficiency. **(B)** GSEA analysis of 19433 protein-coding genes and 146 ERBB2-related differentially expressed genes showed that the genes involved in the primary immunodeficiency pathway were mainly 23 down-regulated genes in the *ERBB2* high expression group; Gene expression was sorted by $\log_2$ (fold change).

Thus, a more comprehensive prospective study is of value in the future to detail the role of *ERBB2* in thyroid cancer. Lastly, this research mainly relied on mRNA data from the TCGA database; therefore, it is necessary to further experimentally investigate the direct evidence of *ERBB2*-related mechanisms in PTC tumorigenesis and progression *in vitro* or *in vivo*.

Because of the complex interaction between *ERBB2* and mutant genes, we didn't separate poor prognosis contributed by *ERBB2* overexpression and gene mutations. To step further, we have planned a more profound analysis in the subsequent study to separate their contributions to poor prognosis. As is known, *BRAF* mutation is a driver gene for tumorigenesis, which causes *ERBB2* overexpression and poor prognosis of PTC (Kebebew et al., 2007; Caria et al., 2016). It remains unknown whether *ERBB2* overexpression independently contributes to the poor prognosis of PTC, which need further validation.

The present work, the first to document, provides a comprehensive study of the relationship between the *ERBB2* expression and clinical characteristics of a large-scale PTC cohort. We found that *ERBB2* was highly expressed in PTC, and

elevated *ERBB2* expression was associated with iodine metabolism dysfunction, tumor proliferation, metastasis, angiogenesis, MEKi resistance, and poor prognosis. Meanwhile, *ERBB2* expression was inversely correlated with the infiltration of immune cells in PTC tissues. It seems that *ERBB2* might be a prognostic biomarker and an immunotherapeutic target in PTC, warranting further clinical validation. The study might also be a new start to expect future investigations on the latent mechanisms that bridge *ERBB2* expression, clinical characteristics, and immunosuppressive environment in PTC.

Data availability statement

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

Author contributions

Conceptualization, YJ, XQ, and ZH; methodology, YJ; software, YJ; validation, XQ, and ZH; formal analysis, YJ; investigation, YJ and ZH; resources, LC; data curation, ZH, and RS; writing—original draft preparation, YJ; writing—review and editing, XQ, ZH, JW, and RS; visualization, YJ; supervision, LC; project administration, LC; funding acquisition, LC. All authors have read and agreed to the published version of the manuscript.

Funding

The study was funded by the National Natural Science Foundation of China, grant numbers 81671711 and 82171981.

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/fgene.2022.966365/full#supplementary-material

References

Agrawal, N., Akbani, R., Aksoy, B. A., Ally, A., Arachchi, H., Asa, S. L., et al. (2014). Integrated genomic characterization of papillary thyroid carcinoma. *Cell.* 159, 676–690. doi:10.1016/j.cell.2014.09.050

Anekpuritanang, T., Uataya, M., Claimon, A., Laokulrath, N., Pongsapich, W., and Pithuksurachai, P. (2021). The association between radioiodine refractory in papillary thyroid carcinoma, sodium/iodide symporter expression, and BRAFV600E mutation. *Onco. Targets. Ther.* 14, 3959–3969. doi:10.2147/ott.s308910

Bastman, J. J., Serracino, H. S., Zhu, Y., Koenig, M. R., Mateescu, V., Sams, S. B., et al. (2016). Tumor-infiltrating T cells and the PD-1 checkpoint pathway in advanced differentiated and anaplastic thyroid cancer. *J. Clin. Endocrinol. Metab.* 101, 2863–2873. doi:10.1210/jc.2015-4227

Bates, J. P., Derakhshandeh, R., Jones, L., and Webb, T. J. (2018). Mechanisms of immune evasion in breast cancer. *Bmc Cancer* 18, 556. doi:10.1186/s12885-018-4441-3

Bauer, A. T., Suckau, J., Frank, K., Desch, A., Goertz, L., Wagner, A. H., et al. (2015). von Willebrand factor fibers promote cancer-associated platelet aggregation in malignant melanoma of mice and humans. *Blood* 125, 3153–3163. doi:10.1182/blood-2014-08-595686

Baxevanis, C. N., Sotiropoulou, P. A., Sotiriadou, N. N., and Papamichail, M. (2004). Immunobiology of HER-2/neu oncoprotein and its potential application in cancer immunotherapy. *Cancer Immunol. Immunother.* 53, 166–175. doi:10.1007/s00262-003-0475-7

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

Cancer Facts (2022). Cancer-facts-statistics. Available at: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html (Accessed May 3, 2022).

Caria, P., Cantara, S., Frau, D. V., Pacini, F., Vanni, R., and Dettori, T. (2016). Genetic heterogeneity of HER2 amplification and telomere shortening in papillary thyroid carcinoma. *Int. J. Mol. Sci.* 17, 1759. doi:10.3390/ijms17101759

Cheng, L., Fu, H., Jin, Y., Sa, R., and Chen, L. (2020). Clinicopathological features predict outcomes in patients with radioiodine-refractory differentiated thyroid cancer treated with sorafenib: A real-world study. *Oncologist* 25, e668–e678. doi:10.1634/theoncologist.2019-0633

Chesnokov, M. S., Yadav, A., and Chefetz, I. (2022). Optimized transcriptional signature for evaluation of MEK/ERK pathway baseline activity and long-term modulations in ovarian cancer. Biorxiv. doi:10.1101/2022.03.21.485160

Dawood, S., Broglio, K., Buzdar, A. U., Hortobagyi, G. N., and Giordano, S. H. (2009). Prognosis of women with metastatic breast cancer by HER2 status and trastuzumab treatment: An institutional-based review. *J. Clin. Oncol.* 28, 92–98. doi:10.1200/jco.2008.19.9844

Degirmenci, U., Yap, J., Sim, Y. R. M., Qin, S., and Hu, J. (2021). Drug resistance in targeted cancer therapies with RAF inhibitors. *Cancer Drug resist.* 4, 665–683. doi:10.20517/cdr.2021.36

Fagin, J. A., and Wells, S. A. (2016). Biologic and clinical perspectives on thyroid cancer. *N. Engl. J. Med.* 375, 1054–1067. doi:10.1056/nejmra1501993

Farwell, M. D., Gamache, R. F., Babazada, H., Hellmann, M. D., Harding, J. J., Korn, R., et al. (2021). CD8-Targeted PET imaging of tumor-infiltrating T cells in patients with cancer: A phase I first-in-humans study of 89Zr-Df-IAB22M2C, a radiolabeled anti-CD8 minibody. *J. Nucl. Med.* 63, 720–726. doi:10.2967/jnumed.121.262485

Guan, S. S., Wu, C. T., Chiu, C. Y., Luo, T. Y., Wu, J. Y., Liao, T. Z., et al. (2018). Polyethylene glycol-conjugated HER2-targeted peptides as a nuclear imaging probe for HER2-overexpressed gastric cancer detection *in vivo*. *J. Transl. Med.* 16, 168. doi:10.1186/s12967-018-1550-3

Haibe, Y., Kreidieh, M., Hajj, H. E., Khalifeh, I., Mukherji, D., Temraz, S., et al. (2020). Resistance mechanisms to anti-angiogenic therapies in cancer. *Front. Oncol.* 10, 221. doi:10.3389/fonc.2020.00221

Herrmann, F., Lehr, H.-A., Drexler, I., Sutter, G., Hengstler, J., Wollscheid, U., et al. (2004). HER-2/neu-Mediated regulation of components of the MHC class I antigen-processing pathway. *Cancer Res.* 64, 215–220. doi:10.1158/0008-5472.can-2522-2

Hou, P., Bojdani, E., and Xing, M. (2010). Induction of thyroid gene expression and radioiodine uptake in thyroid cancer cells by targeting major signaling pathways. *J. Clin. Endocrinol. Metab.* 95, 820–828. doi:10.1210/jc.2009-1888

Jin, Y., Nostrand, D. V., Cheng, L., Liu, M., and Chen, L. (2018). Radioiodine refractory differentiated thyroid cancer. *Crit. Rev. Oncol. Hematol.* 125, 111–120. doi:10.1016/j.critrevonc.2018.03.012

Juríková, M., Danihel, Ľ., Polák, Š., and Varga, I. (2016). Ki67, PCNA, and MCM proteins: Markers of proliferation in the diagnosis of breast cancer. *Acta Histochem.* 118, 544–552. doi:10.1016/j.acthis.2016.05.002

Kebebew, E., Weng, J., Bauer, J., Ranvier, G., Clark, O. H., Duh, Q.-Y., et al. (2007). The prevalence and prognostic value of BRAF mutation in thyroid cancer. *Ann. Surg.* 125, 466–470. doi:10.1097/sla.0b013e318148563d

Kessler, J., Obinger, C., and Eales, G. (2008). Factors influencing the study of peroxidase-generated iodine species and implications for thyroglobulin synthesis. *Thyroid* 18, 769–774. doi:10.1089/thy.2007.0310

Keyaerts, M., Xavier, C., Heemskerk, J., Devoogdt, N., Everaert, H., Ackaert, C., et al. (2016). Phase I study of 68Ga-HER2-Nanobody for PET/CT assessment of HER2 expression in breast carcinoma. *J. Nucl. Med.* 57, 27–33. doi:10.2967/jnumed.115.162024

Kremser, R., Obrist, P., Spizzo, G., Erler, H., Kendler, D., Kemmler, G., et al. (2003). Her2/neu overexpression in differentiated thyroid carcinomas predicts metastatic disease. *Virchows Arch.* 442, 322–328. doi:10.1007/s00428-003-0769-3

Liang, H., Zhong, Y., Luo, Z., Huang, Y., Lin, H., Zhan, S., et al. (2011). Diagnostic value of 16 cellular tumor markers for metastatic thyroid cancer: An immunohistochemical study. *Anticancer Res.* 31, 3433–3440.

Lim, H., Devesa, S. S., Sosa, J. A., Check, D., and Kitahara, C. M. (2017). Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. *Jama* 317, 1338–1348. doi:10.1001/jama.2017.2719

Lin, S., Li, Y., Sun, X., Chen, Q., Huang, S., Lin, S., et al. (2021). Update on the role of neuropeptide Y and other related factors in breast cancer and osteoporosis. *Front. Endocrinol.* 12, 705499. doi:10.3389/fendo.2021.705499

Liu, S., Geng, R., Lin, E., Zhao, P., and Chen, Y. (2021). ERBB1/2/3 expression, prognosis, and immune infiltration in cutaneous melanoma. *Front. Genet.* 12, 602160. doi:10.3389/fgene.2021.602160

Liu, Z., Zeng, W., Huang, L., Wang, Z., Wang, M., Zhou, L., et al. (2018). Prognosis of FTC compared to PTC and FVPTC: Findings based on SEER database using propensity score matching analysis. *Am. J. Cancer Res.* 8, 1440–1448.

Lu, C., Zhao, L., Ying, H., Willingham, M. C., and Cheng, S.-Y. (2010). Growth activation alone is not sufficient to cause metastatic thyroid cancer in a mouse model of follicular thyroid carcinoma. *Endocrinology* 151, 1929–1939. doi:10.1210/en.2009-1017

Mazzoni, M., Mauro, G., Erreni, M., Romeo, P., Minna, E., Vizioli, M. G., et al. (2019). Senescent thyrocytes and thyroid tumor cells induce M2-like macrophage polarization of human monocytes via a PGE2-dependent mechanism. *J. Exp. Clin. Cancer Res.* 38, 208. doi:10.1186/s13046-019-1198-8

Mehnert, J. M., Varga, A., Brose, M. S., Aggarwal, R. R., Lin, C.-C., Prawira, A., et al. (2019). Safety and antitumor activity of the anti–PD-1 antibody pembrolizumab in patients with advanced, PD-L1–positive papillary or follicular thyroid cancer. *Bmc Cancer* 19, 196. doi:10.1186/s12885-019-5380-3

Mitri, Z., Constantine, T., and O'Regan, R. (2012). The HER2 receptor in breast cancer: Pathophysiology, clinical use, and new advances in therapy. *Chemother. Res. Pract.* 2012, 743193. doi:10.1155/2012/743193

Montero-Conde, C., Ruiz-Llorente, S., Dominguez, J. M., Knauf, J. A., Viale, A., Sherman, E. J., et al. (2013). Relief of feedback inhibition of HER3 transcription by RAF and MEK inhibitors attenuates their antitumor effects in BRAF-mutant thyroid carcinomas. *Cancer Discov.* 3, 520–533. doi:10.1158/2159-8290.cd-12-0531

Nagaoka, K., Shirai, M., Taniguchi, K., Hosoi, A., Sun, C., Kobayashi, Y., et al. (2020). Deep immunophenotyping at the single-cell level identifies a combination of anti-IL-17 and checkpoint blockade as an effective treatment in a preclinical model of data-guided personalized immunotherapy. *J. Immunother. Cancer* 8, e001358. doi:10.1136/jitc-2020-001358

Neuzillet, C., Tijeras-Raballand, A., Mestier, L., de Cros, J., Faivre, S., and Raymond, E. (2014). MEK in cancer and cancer therapy. *Pharmacol. Ther.* 141, 160–171. doi:10.1016/j.pharmthera.2013.10.001

Oshi, M., Satyananda, V., Angarita, F. A., Kim, T. H., Tokumaru, Y., Yan, L., et al. (2021). Angiogenesis is associated with an attenuated tumor microenvironment, aggressive biology, and worse survival in gastric cancer patients. *Am. J. Cancer Res.* 11, 1659–1671.

Piscuoglio, S., Ng, C. K. Y., Weigelt, B., Chandarlapaty, S., and Reis-Filho, J. S. (2018). ESR1 and endocrine therapy resistance: More than just mutations. *Ann. Oncol.* 29, 787–789. doi:10.1093/annonc/mdy081

Porra, V., Ferraro-Peyret, C., Durand, C., Selmi-Ruby, S., Giroud, H., Berger-Dutrieux, N., et al. (2005). Silencing of the tumor suppressor gene SLC5A8 is associated with BRAF mutations in classical papillary thyroid carcinomas. *J. Clin. Endocrinol. Metab.* 90, 3028–3035. doi:10.1210/jc.2004-1394

Portulano, C., Paroder-Belenitsky, M., and Carrasco, N. (2013). The Na+/I– symporter (NIS): Mechanism and medical impact. *Endocr. Rev.* 35, 106–149. doi:10.1210/er.2012-1036

Pratilas, C. A., Taylor, B. S., Ye, Q., Viale, A., Sander, C., Solit, D. B., et al. (2009). V600EBRAF is associated with disabled feedback inhibition of RAF–MEK signaling and elevated transcriptional output of the pathway. *Proc. Natl. Acad. Sci. U. S. A.* 106, 4519–4524. doi:10.1073/pnas.0900780106

Qu, N., Zhang, L., Wu, W. L., Ji, Q. H., Lu, Z. W., Zhu, Y. X., et al. (2016). Bilaterality weighs more than unilateral multifocality in predicting prognosis in papillary thyroid cancer. *Tumour Biol.* 37, 8783–8789. doi:10.1007/s13277-015-4533-5

Ruggeri, R. M., Campennì, A., Giuffrè, G., Giovanella, L., Siracusa, M., Simone, A., et al. (2016). HER2 analysis in sporadic thyroid cancer of follicular cell origin. *Int. J. Mol. Sci.* 17, 2040. doi:10.3390/ijms17122040

Sa, R., Liang, R., Qiu, X., He, Z., Liu, Z., and Chen, L. (2022). IGF2BP2-dependent activation of ERBB2 signaling contributes to acquired resistance to tyrosine kinase inhibitor in differentiation therapy of radioiodine-refractory papillary thyroid cancer. *Cancer Lett.* 527, 10–23. doi:10.1016/j.canlet.2021.12.005

Seliger, B., and Kiessling, R. (2013). The two sides of HER2/neu: Immune escape versus surveillance. *Trends Mol. Med.* 19, 677–684. doi:10.1016/j.molmed.2013.08.003

Shah, M. H., Wei, L., Wirth, L. J., Daniels, G. A., Souza, J. A. D., Timmers, C. D., et al. (2017). Results of randomized phase II trial of dabrafenib versus dabrafenib plus trametinib in BRAF-mutated papillary thyroid carcinoma. *J. Clin. Oncol.* 35, 6022. doi:10.1200/jco.2017.35.15_suppl.6022

Shibuya, M. (2011). Vascular endothelial growth factor (vegf) and its receptor (vegfr) signaling in angiogenesis: A crucial target for anti- and pro-angiogenic therapies. *Genes. Cancer* 2, 1097–1105. doi:10.1177/1947601911423031

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

Smith, N. R., Baker, D., James, N. H., Ratcliffe, K., Jenkins, M., Ashton, S. E., et al. (2010). Vascular endothelial growth factor receptors VEGFR-2 and VEGFR-3 are localized primarily to the vasculature in human primary solid cancers. *Clin. Cancer Res.* 16, 3548–3561. doi:10.1158/1078-0432.ccr-09-2797

Srinivasan, S., Kryza, T., Batra, J., and Clements, J. (2022). Remodelling of the tumour microenvironment by the kallikrein-related peptidases. *Nat. Rev. Cancer* 22, 223–238. doi:10.1038/s41568-021-00436-z

Tangye, S. G., Al-Herz, W., Bousfiha, A., Chatila, T., Cunningham-Rundles, C., Etzioni, A., et al. (2020). Human inborn errors of immunity: 2019 update on the classification from the international union of immunological societies expert committee. *J. Clin. Immunol.* 40, 24–64. doi:10.1007/s10875-019-00737-x

Wagle, M.-C., Kirouac, D., Klijn, C., Liu, B., Mahajan, S., Junttila, M., et al. (2018). A transcriptional MAPK Pathway Activity Score (MPAS) is a clinically relevant biomarker in multiple cancer types. *NPJ Precis. Oncol.* 2, 7. doi:10.1038/s41698-018-0051-4

Wu, L., Zhou, Y., Guan, Y., Xiao, R., Cai, J., Chen, W., et al. (2021). Seven genes associated with lymphatic metastasis in thyroid cancer that is linked to tumor immune cell infiltration. *Front. Oncol.* 11, 756246. doi:10.3389/fonc.2021.756246

Wynn, C. S., and Tang, S.-C. (2022). Anti-HER2 therapy in metastatic breast cancer: Many choices and future directions. *Cancer Metastasis Rev.* 41, 193–209. doi:10.1007/s10555-022-10021-x

Xia, C., Dong, X., Li, H., Cao, M., Sun, D., He, S., et al. (2022). Cancer statistics in China and United States, 2022: Profiles, trends, and determinants. *Chin. Med. J.* 135, 584–590. doi:10.1097/cm9.0000000000002108

Yang, R., Shen, S., Gong, C., Wang, X., Luo, F., Luo, F., et al. (2021). Bispecific antibody PD-L1 x CD3 boosts the anti-tumor potency of the expanded Vγ2Vδ2 T cells. *Front. Immunol.* 12, 654080. doi:10.3389/fimmu.2021.654080

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fgene.2022.1084206

OPEN ACCESS

EDITED BY
Amr Ahmed El-Arabey,
Al-Azhar University, Egypt

REVIEWED BY
Nan Wang,
First Affiliated Hospital of Zhengzhou University, China
Xinxing Wang,
Stony Brook University, United States

*CORRESPONDENCE
Xianhai Zeng,
zxhklwx@163.com
Yongyan Wu,
wuyongyan@sxent.org
Wei Gao,
gaoweisxent@sxent.org

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 30 October 2022
ACCEPTED 15 December 2022
PUBLISHED 04 January 2023

CITATION
Zheng X, Zhang C, Zheng D, Guo Q, Maierhaba M, Xue L, Zeng X, Wu Y and Gao W (2023), An original cuproptosis-related genes signature effectively influences the prognosis and immune status of head and neck squamous cell carcinoma.
Front. Genet. 13:1084206.
doi: 10.3389/fgene.2022.1084206

An original cuproptosis-related genes signature effectively influences the prognosis and immune status of head and neck squamous cell carcinoma

Xiwang Zheng[1,2‡], Chunming Zhang[1,2,3‡], Defei Zheng[4], Qingbo Guo[1,2], Mijiti Maierhaba[1,2], Lingbin Xue[5,6], Xianhai Zeng[5,6*†], Yongyan Wu[5,6*†] and Wei Gao[5,6*†]

[1]Shanxi Key Laboratory of Otorhinolaryngology Head and Neck Cancer, First Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, [2]Shanxi Province Clinical Medical Research Center for Precision Medicine of Head and Neck Cancer, First Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, [3]Department of Otolaryngology Head and Neck Surgery, First Hospital of Shanxi Medical University, Taiyuan, Shanxi, China, [4]Department of Hematology/Oncology, Children's Hospital of Soochow University, Suzhou, Jiangsu, China, [5]Department of Otolaryngology Head and Neck Surgery, Longgang Otolaryngology Hospital, Shenzhen, Guangdong, China, [6]Shenzhen Institute of Otolaryngology and Key Laboratory of Otolaryngology, Longgang Otolaryngology Hospital, Shenzhen, Guangdong, China

Background: Recently, a non-apoptotic cell death pathway that is dependent on the presence of copper ions was proposed, named as cuproptosis. Cuproptosis have been found to have a strong association with the clinical progression and prognosis of several cancers. Head and neck squamous cell carcinoma (HNSC) are among the most common malignant tumors, with a 5-year relative survival rate ranging between 40% and 50%. The underlying mechanisms and clinical significance of cuproptosis-related genes (CRGs) in HNSC progression have not been clarified.

Methods: In this study, expression pattern, biological functions, Immunohistochemistry (IHC), gene variants and immune status were analyzed to investigate the effects of CRGs on HNSC progression. Moreover, a 12-CRGs signature and nomogram were also constructed for prognosis prediction of HNSC.

Results: The results revealed that some CRGs were dysregulated, had somatic mutations, and CNV in HNSC tissues. Among them, ISCA2 was found to be upregulated in HNSC and was strongly correlated with the overall survival (OS)

Abbreviations: ANM, Adjacent normal mucosa; AUC, Area under time dependent ROC curve; CRGs, Cuproptosis-related genes; CNV, Copy-number variation; DSS, Disease-specific survival; GEO, Gene expression omnibus; GO, Gene ontology; HNSC, Head and neck squamous cell carcinoma; IHC, Immunohistochemistry; KEGG, Kyoto encyclopedia of genes and genomes; LASSO, Least absolute shrinkage and selection operator; LSCC, Laryngeal squamous cell carcinoma; MAF, Mutation annotation format; OS, Overall survival; PPIs, Protein-protein interactions; PFI, Progression-free interval; PCA, Principal component analysis; ROC, Receiver operating characteristic; ssGSEA, Single sample gene set enrichment analysis; TCGA, The cancer genome atlas; TISIDB, Tumor and immune system interaction database.

of HNSC patients (HR = 1.13 [1.01–1.26], *p*-value = 0.0331). Functionally, CRGs was mainly associated with the TCA cycle, cell cycle, iron-sulfur cluster assembly, p53 signaling pathway, chemical carcinogenesis, and carbon metabolism in cancer. A 12-CRGs signature for predicting the OS was constructed which included, CAT, MTFR1L, OXA1L, POLE, NTHL1, DNA2, ATP7B, ISCA2, GLRX5, NDUFA1, and NDUFB2. This signature showed good prediction performance on the OS (HR = 5.3 [3.4–8.2], *p*-value = 3.4e-13) and disease-specific survival (HR = 6.4 [3.6–11], *p*-value = 2.4e-10). Furthermore, 12-CRGs signature significantly suppressed the activation of $CD4^+$ T cells and antigen processing and presentation. Finally, a nomogram based on a 12-CRGs signature and clinical features was constructed which showed a significantly adverse effect on OS (HR = 1.061 [1.042–1.081], *p*-value = 1.6e-10) of HNSC patients.

Conclusion: This study reveals the association of CRGs with the progression of HNSC based on multi-omics analysis. The study of CRGs is expected to improve clinical diagnosis, immunotherapeutic responsiveness and prognosis prediction of HNSC.

KEYWORDS

head and neck squamous cell carcinoma, cuproptosis, ISCA2, immune infiltration, prognosis

Introduction

Copper is a transition metal that is required for essential enzymes and has a key role in cellular metabolism and bioenergy conversion (Kim et al., 2021). Aberrant levels of copper ions have been associated with anemia, cell proliferation and death, metabolic disease and cancer (Ren et al., 2019; Ge et al., 2022). Various forms of cell death, such as apoptosis, necroptosis, pyroptosis and ferroptosis, have been explored in recent years. However, there is limited knowledge on the mechanism of copper overdose-induced cell death (Li et al., 2022). A recent study by Tsvetkov et al. proposed a novel concept of copper-dependent cell death and termed it as cuproptosis (Tsvetkov et al., 2022). The study revealed that copper toxicity was highly associated with mitochondrial activity since key components in mitochondrial metabolism participated in copper ions-induced cell death. The genes coding for key components in copper-dependent cell death have been identified, and may provide a new strategy for the diagnosis, therapy and outcomes prediction of cancers.

Head and neck cancer is one of the most common malignant tumor, accounting for 5.7% of the global cancer mortality (Qin et al., 2022). Head and neck squamous cell carcinoma (HNSC) is the major histological subtype of head and neck cancer, with a 5-year relative survival rate of only 40%–50% (Dai et al., 2020). Analysis of clinical cases revealed that copper levels were significantly correlated with the initiation and progression of HNSC (Ressnerova et al., 2016; Kudva et al., 2021). Although cuproptosis is just a recent concept, there are already studies exploring the ability of cuproptosis-related genes (CRGs) to predict cancer prognosis and their effects on cancer progression (Bian et al., 2022). However, the role CRGs in HNSC is yet to be determined.

In the current study, we comprehensively investigated the effects of CRGs on HNSC progression based on multi-omics analysis, and a multi-genes CRGs signature and nomogram for HNSC prognosis prediction were also constructed. Briefly, we first analyzed the expression profiles of the CRGs and carried out differential expression analysis of CRGs between HNSC and normal tissues. We also analyzed the somatic mutations and copy-number variation (CNV), as wells as protein-protein interactions (PPIs) and biological functions of CRGs. ISCA2 was chosen for further analysis and was found to be up-regulated at the RNA and protein levels, and significantly associated with the prognosis of HNSC. A 12-CRGs gene signature, including CAT, MTFR1L, OXA1L, POLE, NTHL1, DNA2, ATP7B, ISCA2, GLRX5, NDUFA1, NDUFB2, and DLAT, was developed, and exhibited significant ability to predict overall survival (OS) and disease-specific survival (DSS). The 12-CRGs gene signature was also associated with the immune status of HNSC, particularly, the suppression of $CD4^+$ T cells activation and antigen processing and presentation. Finally, a nomogram consisting of the 12-CRGs gene signature and clinical features was constructed for clinic utility. In conclusion, the significant association between the expression of CRGs and HNSC progression indicate that CRGs have potential roles as diagnostic, therapeutic and prognostic biomarkers for HNSC.

Methods and materials

Data collection

Forty-three cuproptosis-related genes (CRGs) involved in lipoic acid pathway, mitochondria complex I and Fe-S cluster

regulation were manually identified from published literature (Tsvetkov et al., 2022). RNA expression and somatic mutation data of 499 head and neck squamous cell carcinoma (HNSC) samples and 45 normal tissues samples were obtained from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). The corresponding clinical features of the TCGA samples were downloaded from UCSC Xena database (https://xenabrowser.net/). Firstly, the TCGA samples were first sorted based on the length of overall survival (OS) time, and then pseudo-randomized into two groups in an alternating fashion, designated TCGA-first cohort and TCGA-second cohort. Data for another HNSC cohort consisting of 270 tumor samples and corresponding clinical features (GSE65858) were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/gds/). The TCGA-first cohort was used as the training cohort, whereas TCGA-second and GSE65858 were used as the validation cohorts. The clinical features of the three cohorts are summarized in the Supplementary Table S1.

Expression pattern analysis

The expression profile of 43 CRGs was visualized using a heat map that was generated using the pheatmap package (v1.0.12) in R (v4.1.0) based on unsupervised clustering. Principal Component Analysis (PCA) was used to determine the ability of the 43 CRGs to discriminate between HNSC and normal tissues, and was carried out using the stats (v4.1.0) and pca3d (v0.10.2) packages in R. The correlation amongst the 43 CRGs was analyzed using the stats and corrplot (v0.92) packages in R. The boxplot showing the expression of the 43 CRGs was generated using the ggpubr package in R (v0.4.0), while Student's t-test was used for statistical analysis between HNSC and normal tissues.

Somatic mutation analysis

The maftools (v2.8.05) package in R was used to visualize and analyze the somatic mutation of the 43 CRGs based on Mutation Annotation Format (MAF) from TCGA database. In addition, GISTIC2 (v2.0.23) for Linux (Ubuntu; v20.04.3) was utilized to analyze the somatic copy-number variations (CNV) of HNSC tissues.

Protein-protein intersections and biological function analysis

Protein-protein intersections (PPIs) among the 43 CRGs were built in the STRING database (v11.5; https://string-db.org/). In addition, PPI networks for 479 additional proteins together with the 43 CRGs were also generated from the STRING database, and visualized using Cytoscape (v3.7.1). Functional enrichment analysis including Gene Ontology (GO) and KEGG analysis of CRGs was implemented using R package clusterProfiler (v4.0.5) and the results visualized using the ggplot2 package (v3.3.6) in R.

Cox regression and survival analysis

First, univariate Cox regression analysis was used to determine the effects of the 43 CRGs and clinical features on survival of the patients. Next, the risk factors that showed significant effect on survival in the univariate analysis including CRGs ISCA2, gender, age, pathologic stage and alcohol history, were used for multivariate Cox regression analysis. Cox regression analysis was carried out using the survival (v3.3–1) package in R, while the results were visualized using the forestplot (v2.0.1) package. Kaplan-Meier estimate curves were used for estimation of survival probability over 5 years and were generated and visualized using survival and survminer (v0.4.9) packages, respectively. Wald test in Cox regression and log-rank test in Kaplan-Meier estimate were applied to assess the statistical differences among different groups. OS probability of ISCA2 was also verified in all TCGA HNSC samples using GEPIA2 webserver (http://gepia2.cancer-pku.cn/).

Immunohistochemistry analysis

Immunohistochemistry (IHC) was used to validate the expression of ISCA2 at proteome level. IHC images of ISCA2 in HNSC and normal tissues were obtained from the Human Protein Atlas (https://www.proteinatlas.org/). The ISCA2 antibody used for immunohistochemistry was HPA030492 (Atlas antibodies, Bromma, Sweden).

Least absolute shrinkage and selection operator analysis and construction of CRGs risk score model

Least absolute shrinkage and selection operator (LASSO) was used to shrink the data dimensionality, extract representative features and build the CRGs risk-score model. LASSO analysis was carried out using the glmnt (v4.1–4) package in R. We then calculated the risk score for each individual using the following formula: risk score = $\sum_{i=1}^{n}(co_{CRGsi} \times ex_{CRGsi})$, where ex_{CRGsi} is expression level of the risk factor *CRGsi*; co_{CRGsi} is the coefficient of the risk factor *CRGsi*; *n* is the number of the risk factors in the CRGs risk score model. A Kaplan-Meier curve was used to validate the predictive ability of upper risk score model. The time-independent Receiver Operating Characteristic (ROC) curve and area under time dependent ROC curve (AUC) were

utilized to evaluate the sensitivity and specificity of the risk score model. The calculation and visualization of the time-independent ROC was done using the timeROC (v0.4) and ggplot2 packages, respectively.

Validation of CRGs expression in tumor tissues

To validate the expression levels of genes in CRGs risk score model, expression of selected genes in 57 paired laryngeal squamous cell carcinoma (LSCC, one of the most common HNSC) and matched adjacent normal mucosa (ANM) tissues were analyzed using RNA sequencing technology and bioinformatic analysis, of which detailed workflows and methods were described in our previous study (Wu et al., 2020).

The effects of CRGs risk score model on immune infiltration

The 28 immune cell types and the 21 immune-related pathways analyzed in this study were identified from the tumor and immune system interaction database (TISIDB; http://cis.hku.hk/TISIDB) and the Kyoto Encyclopedia of Genes and Genomes (KEGG; https://www.kegg.jp/), respectively. The single sample gene set enrichment analysis (ssGSEA) was employed to estimate the immune infiltration levels based on HNSC expression profile. Enrichment scores of ssGSEA were calculated using the GSVA (v1.40.1) package in R, and Student's t-test was used to compare immune infiltration level and CRGs risk score model.

Nomogram construction

A nomogram that combined the CRGs risk score model and clinical features was built using the rms (v6.3–0) package in R. The nomogram was used to predict the 1-, 2-, three- and 5-year OS probability. The calibration curve used to evaluate the prediction ability of nomogram was generated using the rms package in R. Furthermore, decision curve analysis was employed to assess the clinical utility of the nomogram, and was visualized using the rmda (v1.6) package in R. The nomogram score of each patient was calculated, and the patients were divided into high risk and low risk groups based on the median nomogram score. Cox regression analysis and Kaplan-Meier estimate curves were utilized to estimate the effects of nomogram risk score on 5-year OS.

Statistical analysis

All statistical analyses were performed in R program in this study. Student's t-test was used for statistical analysis of CRGs expression between HNSC and normal tissues. Wald test was used for assessing the statistical significance in Cox regression analysis. Significance was tested using the Log-rank test in Kaplan-Meier estimate. Over-representation test that is the default testing method in clusterProfiler package in R language was used to significant analysis in enrichment analysis. Unless the significant cutoff was specifically indicated, it was assumed to be $p\text{-}value < 0.05$.

Results

The expression patterns of CRGs was significantly correlated with HNSC

A total of 43 CRGs were manually identified and used to investigate the effects of cuproptosis in HNSC in current study (Supplementary Table S2). The workflow of the bioinformatics analysis is shown in Supplementary Figure S1. The heat map revealed that there was significant difference in the expression patterns of CRGs between HNSC and normal tissues (Figures 1A, B). Moreover, the 3D plot of PCA showed that the expression pattern of CRGs in HNSC was distinct from that in normal tissues (Figures 1C,D). To investigate the internal correlation of CRGs, the expression correlation analysis was performed. Figures 1E, F indicate that internal correlation of CRGs was similar in both TCGA-first and TCGA-second cohort. Especially, genes in Fe-S cluster complex I were highly correlated with each other. Results of statistical analysis showed that eleven CRGs (POLD1, NTHL1, MTFR1, CDK5RAP1, TIMMDC1, NDUFA8, PPAT, DNA2, ISCA2, POLE, RTEL1) were upregulated in HNSC compared with normal tissues, while eight genes (NDUFC1, MTFR1L, LIAS, FDX1, DLAT, PDHB, ETFDH, IDH2) were downregulated expression in HNSC tissues in both TCGA-first and TCGA-second cohorts (Figures 1G, H). Moreover, statistical analysis based on paired samples extracted from all TCGA HNSC samples also validated the expression difference of the above nineteen CRGs between HNSC and normal tissues (Supplementary Figure S2). In addition, GO enrichment analysis revealed that the CRGs were mainly enriched in the tricarboxylic acid cycle, iron-sulfur cluster assembly, copper ion response/transport/homeostasis process and other pathways related to respiratory metabolism (Figure 1I), which was consistent with the results of the previous study (Tsvetkov et al., 2022).

CRGs had somatic mutations in HNSC

To investigate the mutational patterns of CRGs in HNSC, MAF files of HNSC samples were used for somatic mutation analysis. Oncoplots showed that thirteen genes (MTFR1, DNA2, ACO2, PDHA1, CDK5RAP1, RTEL1, CAT, POLD1, DLAT, POLE, ATP7B, DLST, ATP7A) had at least one somatic mutation in both TCGA-first and TCGA-second cohorts

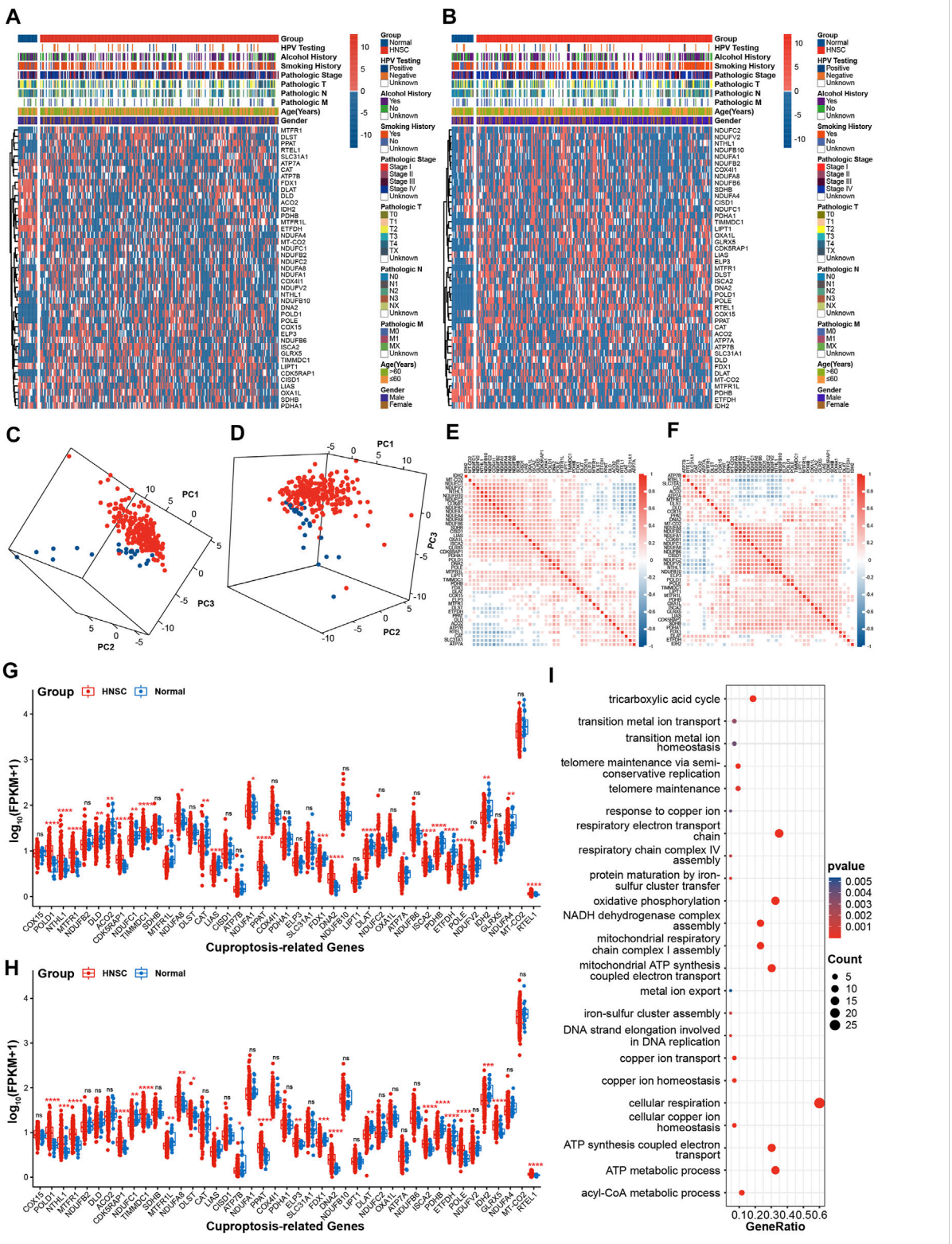


FIGURE 1
Expression patterns of CRGs in HNSC **(A,B)** The heatmap of CRGs in TCGA-first **(A)** and TCGA-second **(B)** cohort **(C, D)** 3D PCA plots of CRGs in TCGA-first **(C)** and TCGA-second **(D)** cohorts, red: HNSC tissues; blue: normal tissues **(E,F)** The correlations analysis of CRGs in TCGA-first **(E)** and TCGA-second **(F)** cohorts **(G–H)** Boxplot showing of CRGs expression in TCGA-first **(G)** and TCGA-second **(H)** cohorts **(I)** Enrichment analysis of GO biological process terms. **p-value* < 0.05; ***p-value* < 0.01; ****p-value* < 0.001; *****p-value* < 0.0001.

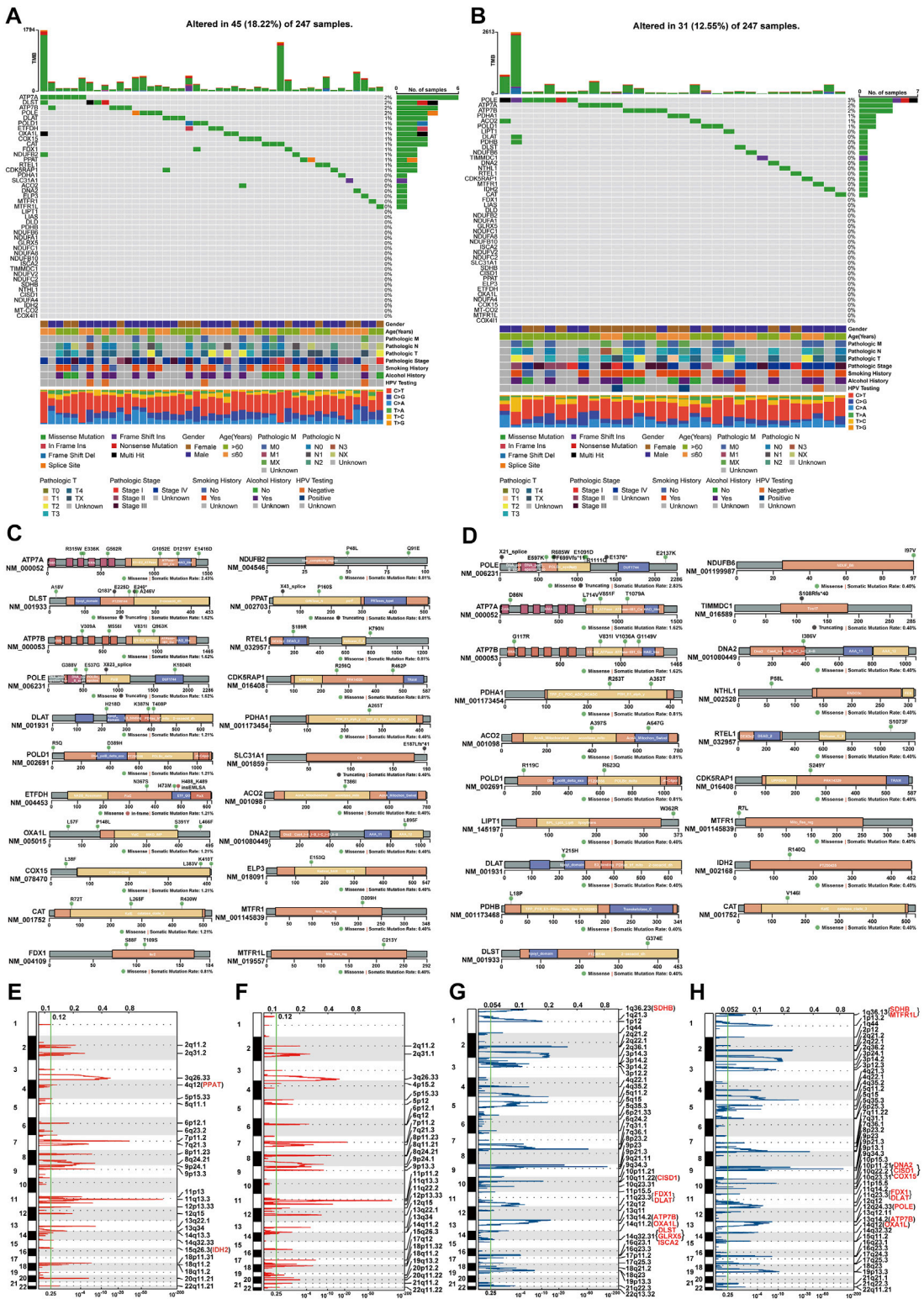


FIGURE 2
Analysis of gene variants of CRGs in HNSC **(A,B)** Mutational oncoplot of CRGs in TCGA-first **(A)** and TCGA-second **(B)** cohorts **(C,D)** Changes in proteins structure of CRGs due to somatic mutations in TCGA-first **(C)** and TCGA-second **(D)** cohorts **(E,F)** Amplification of chromosomal sub-bands of CRGs in TCGA-first **(E)** and TCGA-second **(F)** cohort **(G,H)** Chromosomal deletion of sub-bonds of CRGs in TCGA-first **(G)** and TCGA-second **(H)** cohorts.

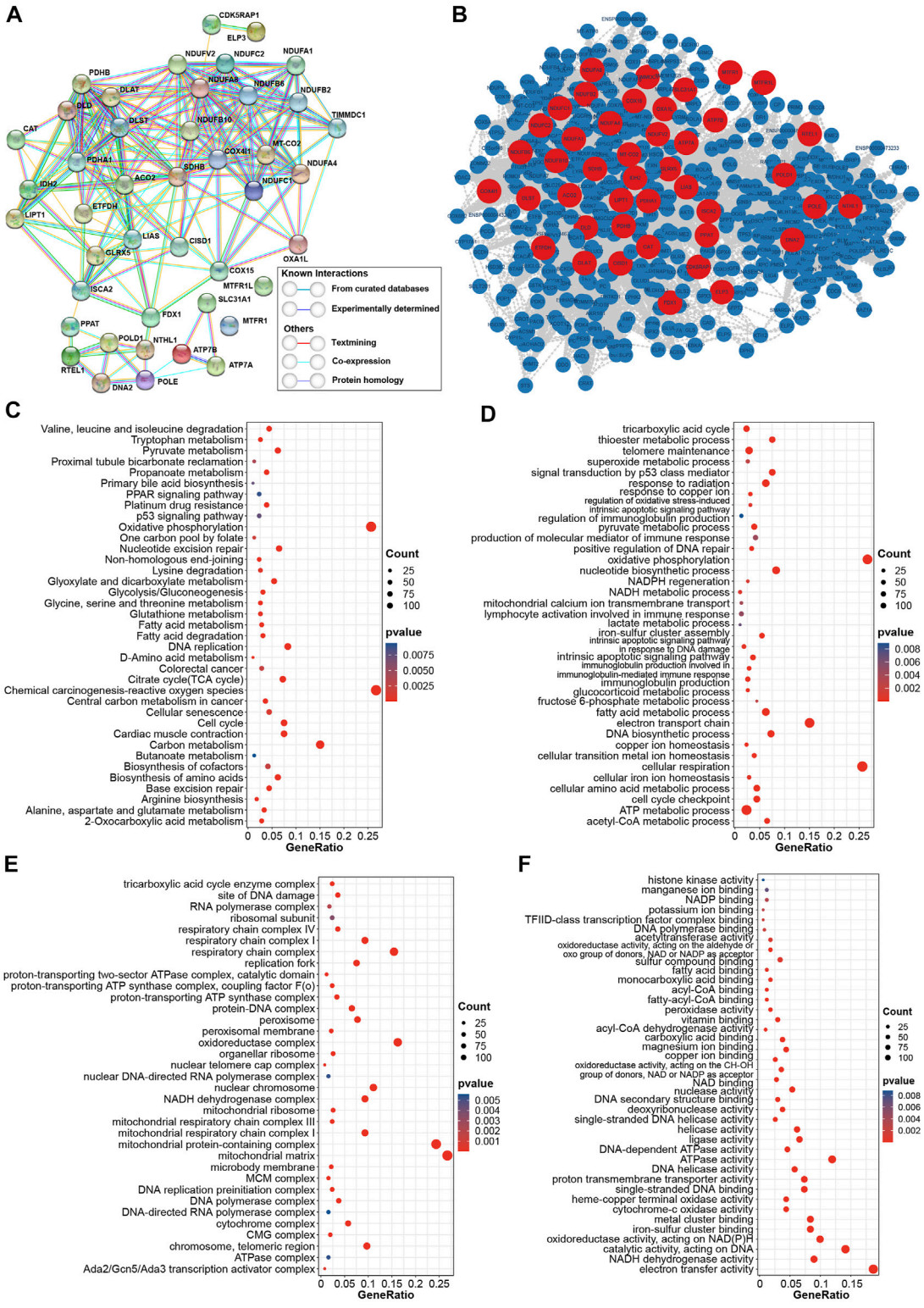


FIGURE 3
Construction of PPIs network and functional enrichment analysis **(A)** Internal PPIs network of CRGs obtained from STRING database **(B)** PPIs network of 522 CRPs from the STRING database, red: proteins coded by CRGs; blue: additional proteins obtained from the STRING database **(C)** KEGG enrichment analysis of CRPs **(D)** Enrichment analysis for CRPs based on GO biological process terms **(E)** Enrichment analysis for CRPs based on GO cellular component terms **(F)** Enrichment analysis for CRPs based on GO biological function terms.

(Figures 2A, B). The altered ratios in TCGA-first and TCGA-second cohorts were 18.22% (45 of 247 samples) and 12.55% (31 of 247 samples), respectively. Further, the mutational lollipop plots indicated that the gene mutations affected the spatial structure of proteins (Figures 2C, D). Since CNV affects gene expression (Fang et al., 2016), we also investigated the amplifications and deletions of genes in CRGs. The heat map of raw CNV number indicated that TCGA-first and TCGA-second cohorts had a similar CNV landscape (Supplementary Figures S3A, B). Figures 2E, F show the somatic chromosomal amplifications of genes in CRGs. Deletion analysis revealed that there were clear copy number deletions in the chromosomal sub-bands where SDHB, CISD1, FDX1, DLAT ATP7B, and OXA1L are located (Figures 2G, H). It should be noted that FDX1 and DLAT had chromosomal deletions as well as downregulated expression, indicating that chromosomal deletions of FDX1 and DLAT might influence their expression levels.

Construction of PPIs networks and functional enrichment analysis

To study the intrinsic interactions amongst CRGs, the PPIs network was built in STRING database. The PPIs network in Figure 3A revealed the potential connection among CRGs based on diverse interactions. Proteins in lipoic acid pathway and Fe-S cluster complex I were highly respective correlation. Further, 479 additional proteins that interacted with CRGs were also obtained from the STRING database, and their PPIs network are shown in Supplementary Table S3 and Figure 3B. The up- and downstream collaboration of various proteins in a signaling pathway is one of the important PPIs ways. All cuproptosis-related proteins (CRPs) included proteins coded by 43 CRGs and 479 additional proteins were subsequently utilized for functional enrichment analysis. KEGG analysis revealed that CRPs were mainly enriched in p53 signaling pathway, TCA cycle, cell cycle, glycolysis/gluconeogenesis, chemical carcinogenesis, carbon metabolism in cancer and other diseases and metabolism pathways (Figure 3C). Similarly, GO of biological process analysis showed that CRPs were enriched in p53-related pathway, TCA cycle and cell cycle pathway. In addition, CRPs were also enriched in iron ion regulation, iron-sulfur cluster assembly, immune-related regulation, synthetic and repair of DNA and energy metabolism signaling pathway in biological process analysis (Figure 3D). Figures 3E,F show the enrichment results of CRPs in the cellular component terms and molecular function terms of GO analysis. The similar biological functions, such as TCA cycle, mitochondrial components, DNA polymerase complex, ATPase activity, DNA helicase activity and iron-sulfur cluster binding, have validated the analysis process of GO biological process terms.

ISCA2 was upregulated and strongly associated with prognosis of HNSC

To investigate the prognostic effects of CRGs in HNSC, univariate Cox regression was used to analyze the correlation between the single gene expression of CRGs and the OS of HNSC. Figure 4A shows that ISCA2 and POLE were significantly correlated with OS of HNSC, with ISCA2 being a poor prognostic factor for HNSC OS (Hazard ratio (HR) = 1.5 [1.0–2.3], p-value = 0.038). Expression of ISCA2 was upregulated in HNSC at the transcriptome level (Figures 1G, H, 4B,C). Moreover, IHC analysis was utilized to validate the expression of ISCA2 in proteome. IHC results in the Human Protein Atlas database showed that ISCA2 was lowly expressed in normal tissues, while high level of ISCA2 was observed in HNSC tissues (Figure 4D). We also analyzed the effects of clinical features on the prognosis. Results of Cox regression analysis indicated that age, gender, pathologic stage and alcohol history were risk factors (Figure 4E). We further investigated the synergistic effects of ISCA2 and clinical features using multivariate Cox regression analysis (Figure 4F). In the multivariate Cox model, ISCA2 (HR = 1.13 [1.01–1.26], p-value = 0.0331), age (HR = 0.44 [0.24–0.82], p-value = 0.0098), pathologic stage (HR = 2.23 [1.06–4.66], p-value = 0.0337) and alcohol history (HR = 0.54 [0.3–0.99], p-value = 0.0447) demonstrated significant effects on the OS of HNSC. The Kaplan-Meier method was employed to validate the effects of ISCA2 in different prognosis. Results in Figures 4G–I indicate that ISCA2 had significant effects on the probability of OS, progression-free interval (PFI; HR = 1.9 [1.2–2.9], p-value = 0.0042) and diseases-specific survival (DSS; HR = 2.1 [1.3–3.6], p-value = 0.0046) in the training cohort. The effect trends of ISCA2 on prognosis in TCGA-first cohort was validated in the TCGA-second and GSE65858 cohorts (Supplementary Figure S4). In addition, effects of ISCA2 on OS probability of all TCGA HNSC samples was also verified in GEPIA2 webserver (Figure 4J).

Construction of a 12-CRGs signature for OS prediction of HNSC

Multi-gene-combination signatures have a higher sensitivity than the single index in prediction of clinical outcomes (Ekanem et al., 2019). Forty-three CRGs were used to construct a multi-genes signature based on LASSO algorithm. A 12-CRGs signature, including CAT, MTFR1L, OXA1L, POLE, NTHL1, DNA2, ATP7B, ISCA2, GLRX5, NDUFA1, NDUFB2, and DLAT, was finally developed (Figure 5A). Risk scores based on the expression levels of the 12-CRGs signature for each patient were calculated utilizing follow formula: risk scores = $0.1407 \times ex_{CAT} + -0.0002 \times ex_{MTFR1L} + 0.0179 \times ex_{OXA1L} + -0.2615 \times ex_{POLE} + -0.0043 \times ex_{NTHL1} + -0.0667 \times ex_{DNA2} + 0.0628 \times$

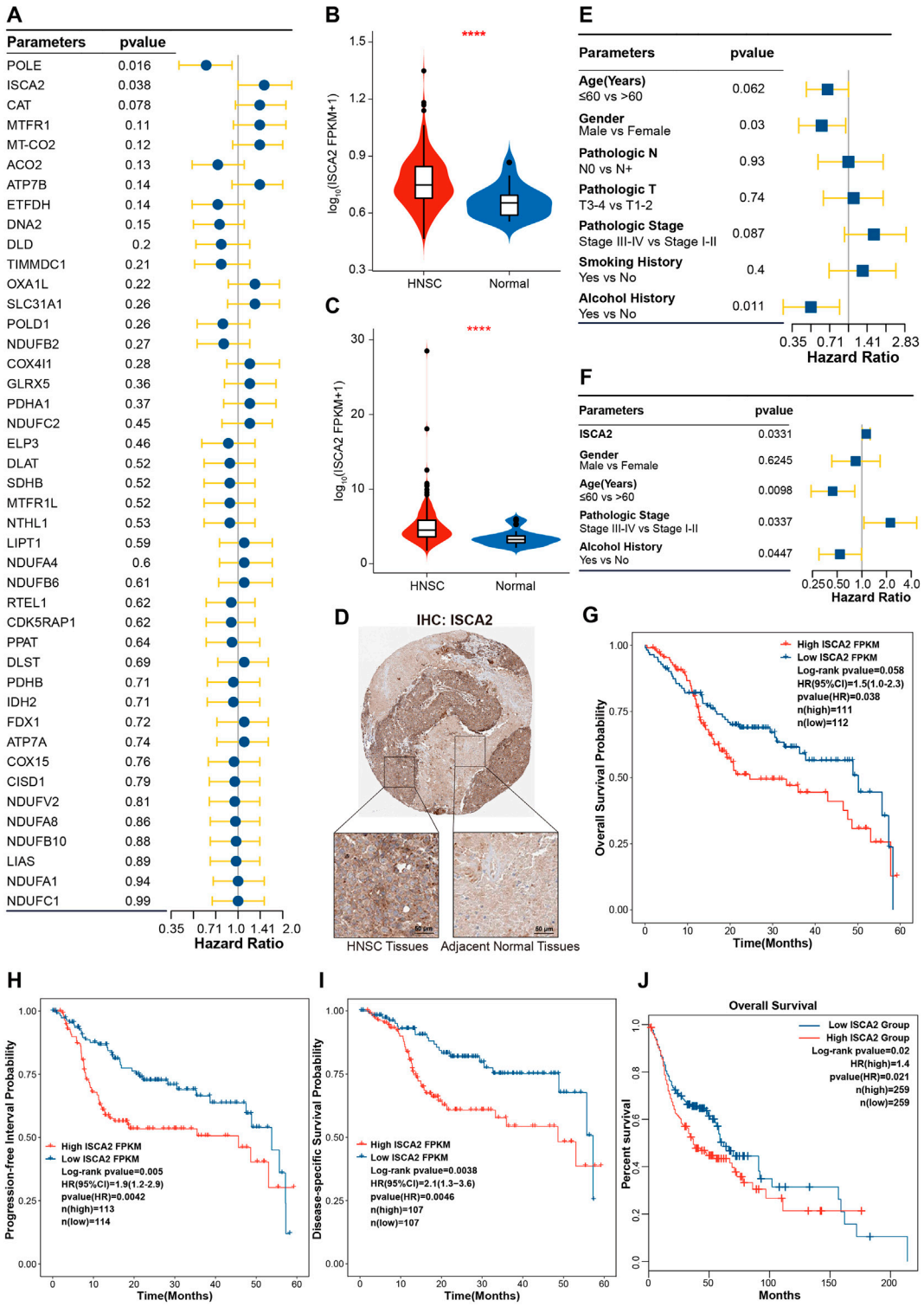


FIGURE 4
ISCA2 was upregulated and significantly correlated with HNSC progression **(A)** Univariate Cox regression analysis of CRGs **(B,C)** Boxplot showing ISCA2 expression in TCGA-first **(B)** and TCGA-second **(C)** cohorts **(D)** Validation of ISCA2 expression by IHC in HNSC and normal tissues **(E)** Univariate Cox regression analysis of clinical features **(F)** Multivariate Cox regression analysis of risk factors consisting of ISCA2, gender, age pathologic stage, and alcohol history **(G–I)** Kaplan-Meier curve showing the association of ISCA2 expression with the 5-year OS **(G)**, PFI **(H)**, and DSS **(I,J)** Association of ISCA2 with the OS in all TCGA HNSC samples as verified GEPIA2 webserver. ****p-value* < 0.0001.

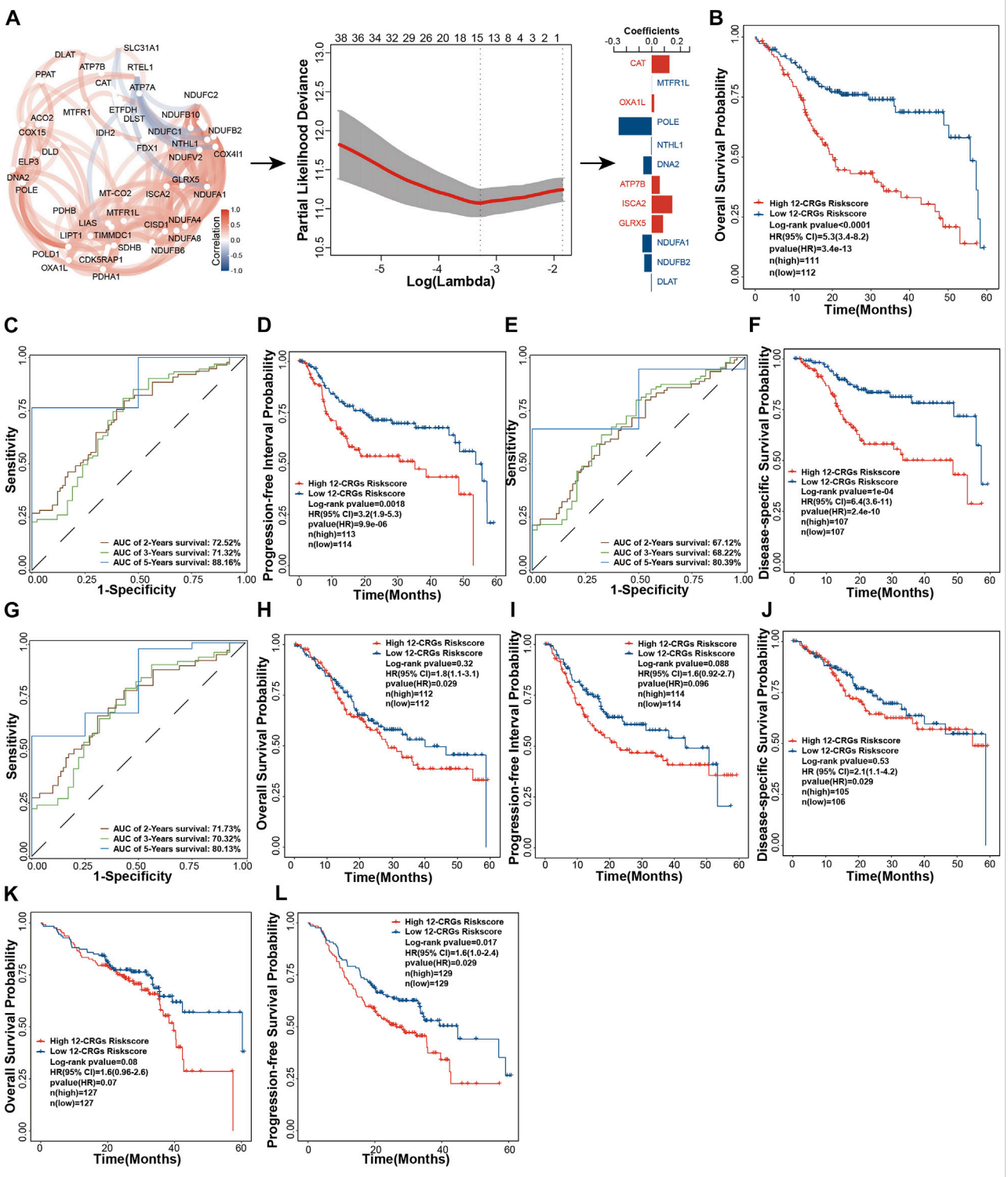


FIGURE 5
Construction and validation of 12-CRGs signature for OS prediction of HNSC **(A)** Schematic diagram of 12-CRGs signature construction **(B)** Kaplan-Meier curve showing the OS predicted by the 12-CRGs signature over 5 years **(C)** AUC values of the time-dependent ROC curves showing the predicted ability of the 12-CRGs signature for OS in 2-, 3- and 5-year **(D)** Kaplan-Meier curve showing the PPI predicted by the 12-CRGs signature over 5 years **(E)** AUC values of the time-dependent ROC curves showing the predicted ability of the 12-CRGs signature for PPI in 2-, 3- and 5-year **(F)** Kaplan-Meier curve showing the DSS predicted by the 12-CRGs signature over 5 years **(G)** AUC values of the time-dependent ROC curves showing the predicted ability of the 12-CRGs signature for DSS in 2-, 3- and 5-year **(H–J)** Kaplan-Meier curve of OS **(H)**, PPI **(I)** and DSS **(J)** predicted by 12-CRGs signature over 5 years in TCGA-second cohort **(K,L)** Kaplan-Meier curve of OS **(K)** and progression-free survival **(L)** predicted by 12-CRGs signature over 5 years in GSE65858 cohort.

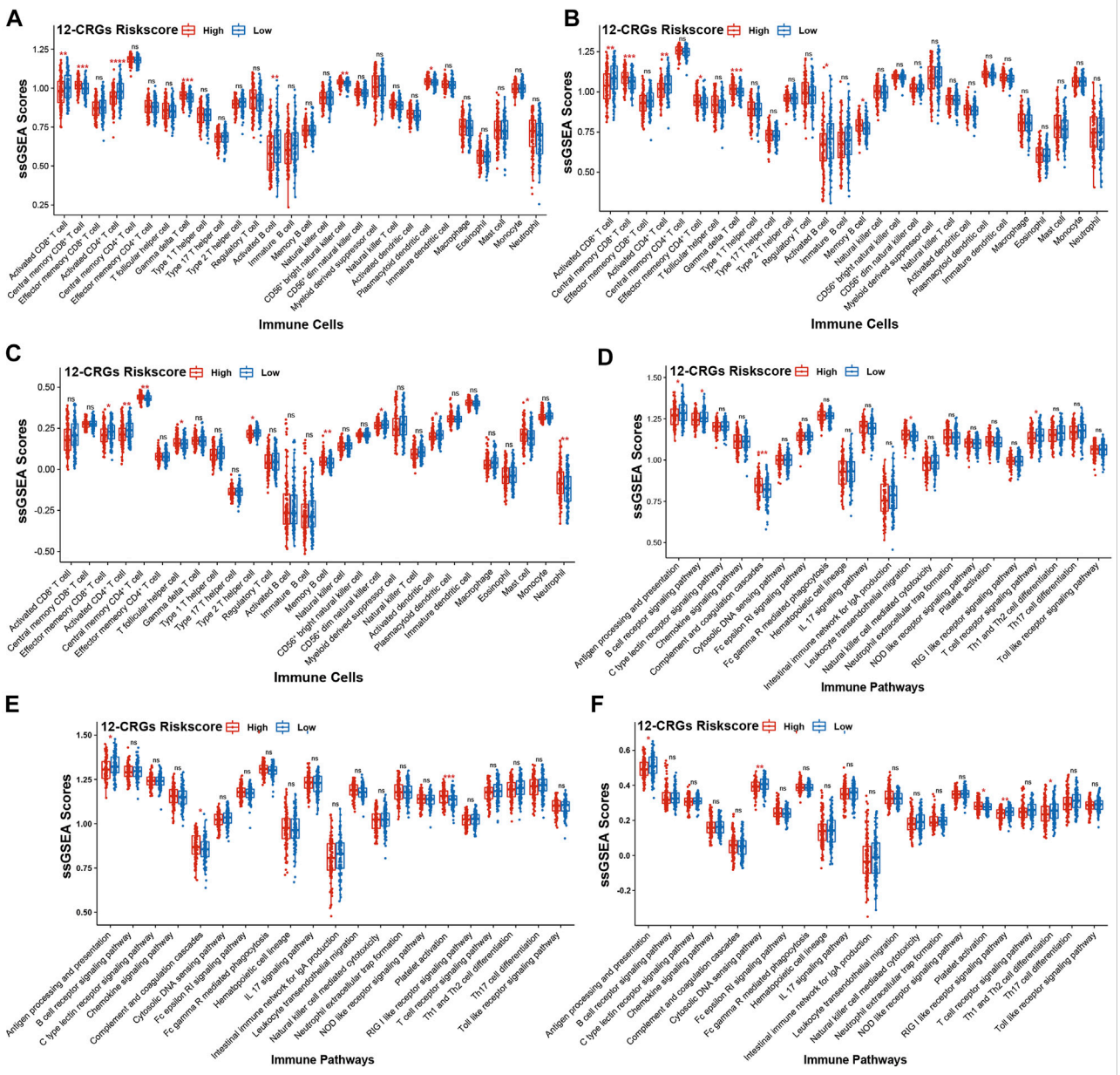


FIGURE 6
Association of the 12-CRGs signature with immune status **(A–C)** Association of the 12-CRGs signature with immune-related cells in TCGA-first **(A)**, TCGA-second **(B)** and GSE65858 **(C)** cohorts **(D–F)** Association of the 12-CRGs signature with the immune-related pathways in TCGA-first **(D)**, TCGA-second **(E)** and GSE65858 **(F)** cohorts. **p-value* < 0.05; ***p-value* < 0.01; ****p-value* < 0.001; *****p-value* < 0.0001.

ex_{ATP7B} + 0.1622 × ex_{ISCA2} + 0.0890 × ex_{GLRX5} + −0.0733 × ex_{NDUFA1} + −0.0589 × ex_{NDUFB2} + −0.0045 × ex_{DLAT}. In TCGA-first cohort, patients were divided into high and low risk groups based on the median risk score of the 12-CRGs signature. Figures 5B, D,F show that patients in the high-risk group had the more adverse outcomes of OS (HR = 5.3 [3.4–8.2], *p*-value = 3.4e-13), PFI (HR = 3.2 [1.9–5.3], *p*-value = 9.9e-06), and DSS (HR = 6.4 [3.6–11], *p*-value = 2.4e-10). The AUC of time dependent ROC was used to evaluate the sensitivity and specificity of 12-CRGs risk scores in prognostic prediction. The AUC for OS, PFI and DSS was 72.52%, 71.32% and 88.16%, 67.12%, 68.22% and 80.39, 71.73, 70.32% and 80.13% in the time period of 2-, three- and 5-year, respectively (Figures 5C,E,G). Unsurprisingly, 12-CRGs signature also had the significant effects on the probability of OS (HR = 1.8 [1.1–3.1], *p*-value = 0.029), and DSS (HR = 2.1 [1.1–4.2], *p*-value = 0.029) in TCGA-second cohort (Figures 5H,J). In addition, although its effect on PFI in TCGA-second cohort was not significantly different (HR = 1.6 [0.92–2.7],

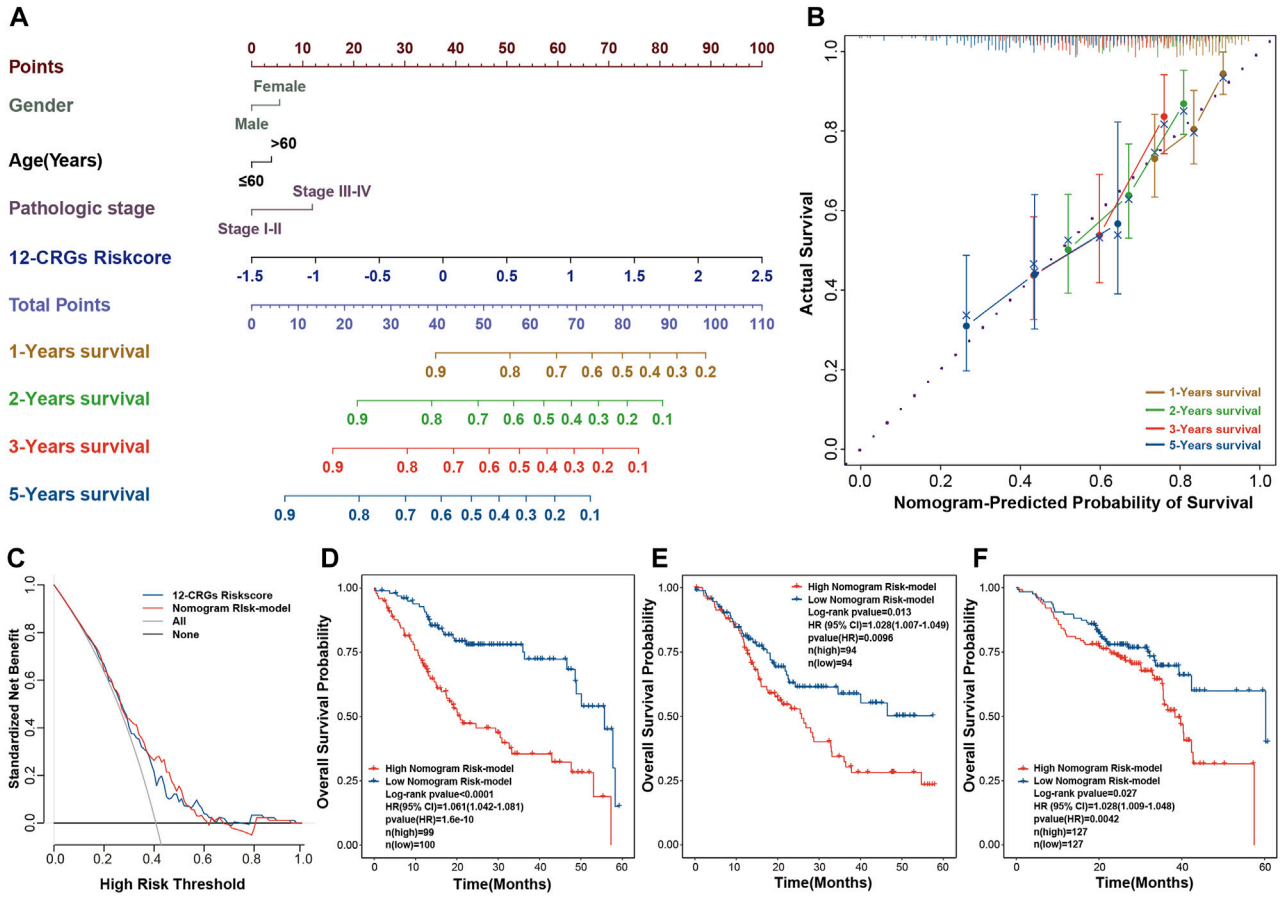


FIGURE 7
Construction of a nomogram for clinic utilization **(A)** A visual nomogram for predicting the OS at 1-, 2-, three- and 5-year after diagnosis **(B)** Calibration curve for assessing the predictive ability of the nomogram **(C)** Decision curve for evaluating the clinic use of the nomogram **(D–F)** Association of the nomogram scores with OS in TCGA-first **(D)**, TCGA-second **(E)** and GSE65858 **(F)** cohorts.

p-value = 0.096), the 12-CRGs risk scores have showed a tendency toward being a risk factor based on the clear separation of probability curves between high and low risk group (Figure 5I). Further, the trend of 12-CRGs as the risk factor was also validated in the prognosis analysis of GSE65858 cohort that included OS (HR = 1.6 [0.96–32.6], *p*-value = 0.07) and progression-free survival (PFS; HR = 1.6 [1.0–2.4], *p*-value = 0.029) analysis (Figures 5K,L).

The interactions between cancer cells and immune cells plays an important role in regulation of cancer progression (Hara et al., 2021). To investigate the correlation between 12-CRGs risk scores and immune status in patient with HNSC, ssGSEA was used to calculate the immune scores for each patient based on 28 immune cell types and 21 immune-related pathways. In the TCGA-first and TCGA-second cohort, $CD8^+$ T cell, $CD4^+$ T cell and B cell immune scores were negatively correlated with high 12-CRGs risk group (Figures 6A,B). Moreover, the activation of $CD4^+$ T cell was significantly suppressed in all three cohort (Figures 6A–C). Analysis of the effects of the 12-CRGs signature on the immune-related pathways revealed that antigen processing and presentation were impaired in high 12-CRGs risk group (Figures 6D–F).

Construction and validation of nomogram for HNSC clinical utilization

The combination of multi-genes signature and clinical features can improve the prognosis prediction of cancer patients (Yang et al., 2021). A visual nomogram that combined the 12-CRGs risk scores and clinical characters was built to predict the OS of HNSC patients (Figure 7A). Calibration curves were used to assess the predictive ability of nomogram, with the 45° line in the calibration plot represents the best prediction (Figure 7B). The nomogram scores of each HNSC patient were calculated using the nomogram, and then a nomogram risk model was developed based on

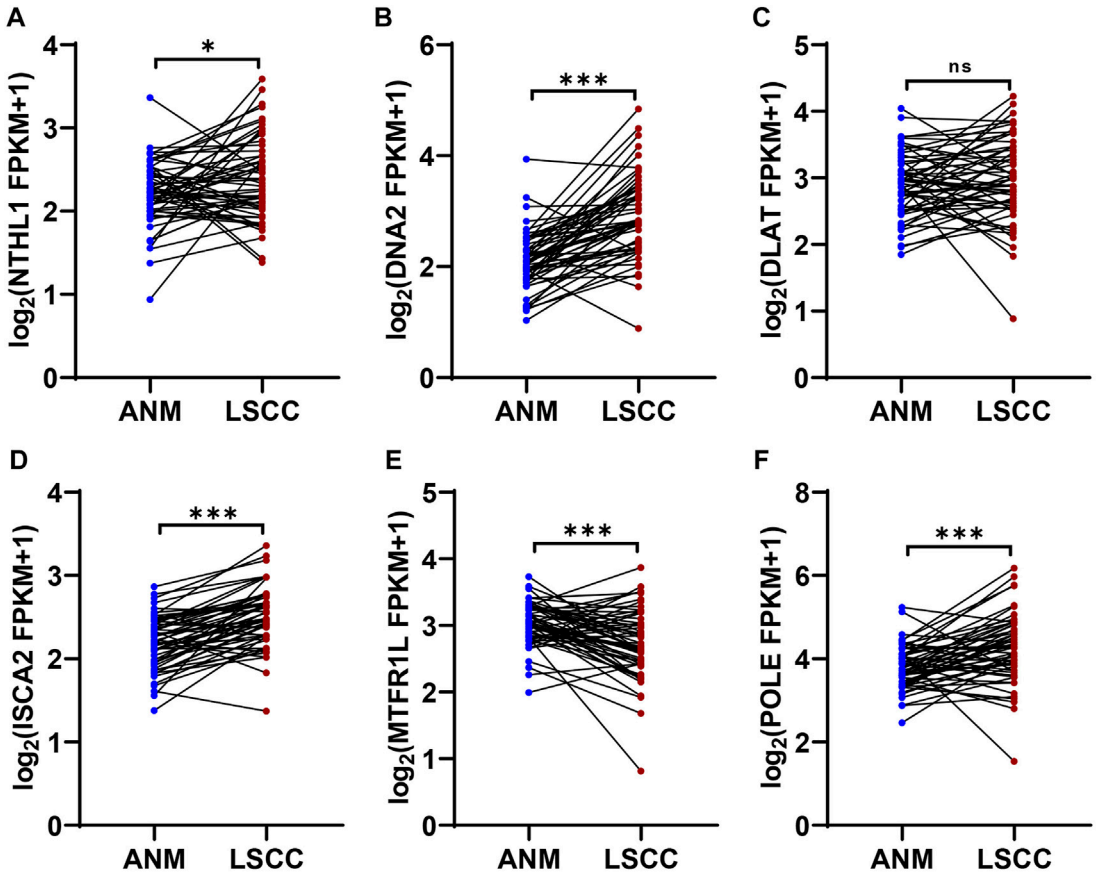


FIGURE 8
Expression validation of genes in 12-CRGs signature using RNA-Seq data of LSCC and ANM tissues **(A–F)** Expression levels of NTHL1 **(A)**, DNA2 **(B)**, DLAT **(C)**, ISCA2 **(D)**, MTFR1L **(E)** and POLE **(F)**.

nomogram scores. The decision curve verified that the nomogram risk model and 12-CRGs risk scores were clinically useful (Figure 7C). Finally, effects of nomogram scores on OS probability were also validated, and results showed that the OS probability of patients in high nomogram score was significantly decreased in all cohorts: TCGA-first (HR = 1.061 [1.042–1.081], *p*-value = 1.6e-10), TCGA-second (HR = 1.028 [1.007–1.049], *p*-value = 0.0096) and GSE65858 (HR = 1.028 [1.009–1.048], *p*-value = 0.0042) cohort (Figures 7D–F).

Validation of genes expression in 12-CRGs signature

In 12-CRGs signature, POLE, NTHL1, DNA2 and ISCA2 was upregulated expression in HNSC tissues while MTFR1L and DLAT was downregulated expression in HNSC tissues in both TCGA-first and TCGA-second cohort (Figures 1G, H). In order to validate the aberrant expression of POLE, NTHL1, DNA2, ISCA2, MTFR1L and DLAT in HNSC tissues, we analyzed the expression levels of these genes using RNA sequencing data of 57 LSCC and paired ANM tissues. The results showed that POLE, NTHL1, DNA2 and ISCA2 was upregulation and MTFR1L was downregulation in LSCC tissues compared to ANM tissues, however, DLAT was not statistically significant difference between LSCC and ANM tissues (Figure 8A–F).

Discussion

To the best of our knowledge, there were no reports about the analysis of the potential correlation between CRGs and progression of HNSC so far. In current study, we investigated the effects of CRGs on HNSC development, examined the genes variation of CRGs in HNSC, analyzed the biological functions of CRGs, explored the immune influence of CRGs on HNSC, and finally developed a 12-CRGs signature and nomogram for HNSC prognosis prediction and clinic use, respectively.

Previous studies have indicated that CRGs have potential as diagnostic, and predictive biomarkers as well as therapeutic targets of cancers in clinic (Bian et al., 2022; Tsvetkov et al., 2022). In our study, we found that eleven CRGs (POLD1, NTHL1, MTFR1, CDK5RAP1, TIMMDC1, NDUFA8, PPAT, DNA2, ISCA2, POLE, RTEL1) were upregulated in HNSC compared with normal tissues, while eight genes (NDUFC1, MTFR1L, LIAS, FDX1, DLAT, PDHB, ETFDH, IDH2) were downregulated in HNSC tissues compared with normal tissues. Particularly, overexpression of ISCA2 in HNSC was observed at the RNA and protein levels, and ISCA2 was strongly correlated with OS (HR = 1.5 [1.0–2.3], *p*-value = 0.038), PFI (HR = 1.9 [1.2–2.9], *p*-value = 0.0042) and DSS (HR = 2.1 [1.3–3.6], *p*-value = 0.0046) of HNSC. ISCA2 is required for Iron-Sulfur cluster assembly and plays an essential role in the pathogenesis of multiple mitochondrial dysfunction syndromes (Weiler et al., 2020; Lebigot et al., 2021). The inhibition of ISCA2 can significantly reduce the xenograft growth of clear cell renal cell carcinoma. Mechanistically, the suppression of ISCA2 both decreases HIF-α levels and induces ferroptosis through triggering singal pathways that does not rely on ISCA2's role in mitochondrial [4Fe-4S] assembly (Green et al., 2022), indicating that ISCA2 may be one of the potential regulators of HNSC progression.

Given that genetic mutations are one of the direct causes of cancers (Cheek and Naxerova, 2022). We investigated the mutational status of CRGs in HNSC. Thirteen genes, including MTFR1, DNA2, ACO2, PDHA1, CDK5RAP1, RTEL1, CAT, POLD1, DLAT, POLE, ATP7B, DLST and ATP7A, had somatic mutations in both training and validation cohort. Previous study reported that POLE *p*. L424V mutation was more frequent in patients with polyposis, colorectal cancer and oligodendroglioma, and four POLD1 variants (*p*.D316H, *p*. D316G, *p*. R409W, and *p*. L474P) were identified in non-polyposis colorectal cancer families. Moreover, POLE/POLD1 variants carriers direct to associated phenotype characterized by attenuated or oligo-adenomatous colorectal polyposis, colorectal cancer, and brain tumors (Bellido et al., 2016). Furthermore, POLE/POLD1 mutations were also found to be a promising predictive biomarkers for positive Immune-checkpoint inhibitor outcomes (Wang et al., 2019). ATP7B tag SNPs rs9535828 and rs9535826 were found to be correlated with platinum resistance in Chinese Han lung cancer patients (Li et al., 2014). The recurrent variant c.1121G>A (*p*.Gly374Glu; dbSNP: rs1270341616) in TCA-cycle-related gene DLST was more frequent in pheochromocytomas and paragangliomas (Remacha et al., 2019). It has been reported that the mutations in DNA2 (Zheng et al., 2020), ACO2 (Sajnani et al., 2017), and ATP7A (Li et al., 2017) were correlated with the initiation or poor clinical outcomes of cancers. CNV is another type of genetic variation in cancer (Dixon et al., 2018). We found that the chromosomal sub-bands where FDX1 and DLAT are located had deletions, and the expression levels of FDX1 and DLAT were downregulated. The downregulation of DLAT has been associated with its chromosomal deletions in liposarcoma (Crago et al., 2012), thus suggesting that the chromosomal deletions of FDX1 and DLAT may affect their own expression levels. Gene variants of CRGs may have the potential correlations with the initiation and poor clinical outcomes of HNSC.

PPIs govern a majority of cellular pathways and processes in organism (Lenz et al., 2021). We analyzed the internal interactions amongst CRGs and constructed a PPIs network consisting of 522 CRPs that were obtained from STRING database. PPIs have a non-substitutable role in all relevant biological processes, and functional enrichment analysis is an important method to investigate the collaboration of various proteins in a signaling pathway (Zhang et al., 2021). All CRPs were utilized for functional enrichment analysis based on GO and KEGG method. We found that CRPs were mainly enriched in p53 signaling pathway, TCA cycle, cell cycle, chemical carcinogenesis, carbon metabolism in cancer, iron-sulfur cluster assembly, immune-related regulation, synthetic and repair of DNA and energy metabolism signaling pathway. The enrichment of the CRGs in the TCA cycle and iron-sulfur cluster assembly was consistent with reports from a previous study by Tsvetkov et al. (Tsvetkov et al., 2022). Components required by TCA cycle and iron-sulfur cluster assembly are mainly regulatory targets for key proteins in cuproptosis process (Li et al., 2022; Tsvetkov et al., 2022). The enrichment of CRPs in p53 signaling, chemical carcinogenesis and carbon metabolism in cancer pathways might bring the potential research perspective in the relation between cuproptosis and cancers.

Multi-gene signatures have been employed for predicting prognosis and have exhibited a significantly effective ability in classifying individuals with multiple clinical-pathological risk factors, such as total mortality, chemotherapy response, and metastasis risk (Ahluwalia et al., 2021). In our study, we constructed a 12-CRGs signature consisting of CAT, MTFR1L, OXA1L, POLE, NTHL1, DNA2, ATP7B, ISCA2, GLRX5, NDUFA1, NDUFB2, and DLAT for OS prediction of HNSC. The 12-CRGs signature had the significantly adverse effects on OS (HR = 5.3 [3.4–8.2], *p*-value = 3.4e-13) of HNSC patients. Moreover, the 12-CRGs signature also exhibited the trends of poor effect on PFI (HR = 1.6 [0.92–2.7], *p*-value = 0.096) and DSS (HR = 6.4 [3.6–11], *p*-value = 2.4e-10). Some components in 12-CRGs signature have been found to have high correlation with development and progression of cancers. DNA2 have functions as both a tumor promoter and suppressor in cancers. On the one hand, DNA2 can suppress the initiation of tumors by maintaining the genomic integrity, on the other hand, it can promote the cancer cells survival through counteracting replication stress (Zheng et al., 2020). The suppression of GLRX5 activates iron-starvation response, increase intracellular free iron and in turn results in Fenton reaction and ferroptosis. In addition, GLRX5 inhibition predisposes therapy-resistant HNSC cells to ferroptosis (Lee J et al., 2020).

DLAT can catalyze the conversion of pyruvate into Acetyl CoA, promote oxidative phosphorylation, ATP generation and catabolic reactions, which are important in the development of cancer (Goh et al., 2015). In addation, CAT (Galasso et al., 2021), POLE (Bellido et al., 2016), NTHL1 (Magrin et al., 2021), ATP7B (Li et al., 2017), ISCA2 (Weiler et al., 2020; Green et al., 2022), NDUFA1 (Mamelak et al., 2005) and NDUFB2(Shan et al., 2020) have also been reported to play a key role in the development of cancers. The combination of multi-genes signature and clinical characters can improve the predictive ability of prognosis for cancer patients (Yang et al., 2021). A nomogram for clinic utilization was constructed based on the 12-CRGs signature and clinical features including gender, age and pathologic stage. The nomogram scores were significantly associated with OS of HNSC in TCGA-first (HR = 1.061 [1.042–1.081], *p*-value = 1.6e-10), TCGA-second (HR = 1.028 [1.007–1.049], *p*-value = 0.0096) and GSE65858 (HR = 1.028 [1.009–1.048], *p*-value = 0.0042) cohorts. In 12-CRGs signature, POLE, NTHL1, DNA2 and ISCA2 was upregulated while MTFR1L and DLAT was downregulated in HNSC tissues in both TCGA-first and TCGA-second cohort. Furthermore, the aberant expressions of POLE, NTHL1, DNA2, ISCA2 and MTFR1L were validated by our RNA-Seq data. These results indicated that POLE, NTHL1, DNA2, ISCA2 and MTFR1L may have greater potential for clinical applications.

Immune cell response to cancer cells plays a crucial role in the regulation of cancer progression (Hara et al., 2021). In this study, correlation analysis between the 12-CRGs signature and immune status of HNSC revealed that $CD4^+$ T cell activation and antigen processing and presentation were suppressed in high risk group of 12-CRGs signature. $CD4^+$ T cells are a class of T helper cells, which are strongly correlated with the biological process of antitumor, and can improve the activity of other antitumor immune cells, such as $CD8^+$ T cells and macrophages (Miggelbrink et al., 2021). Antigen processing and presentation are most important events in the recognition of antigens by T cells. Moreover, specific T cell tumor antigens generated through antigen processing and presentation are potential agents in the field of cancer immunotherapy (Lee MY et al., 2020). The above results demonstrated that 12-CRGs signature might influence the progression of HNSC by regulating the immune response.

Besides, some limitations should be noted in this study. Firstly, the numbers of CRGs under investigated were still limited in this study. When more CRGs are identified in the literature, a more representative number of CRGs will be included in the study. Secondly, the CRGs risk score model was constructed and validated in public database in this study. We will govern prospective multi-center clinical data to vertify the ability of CRGs risk score model in the future. Further, the effects of 12-CRGs signature on immune status of HNSC will be further investigated at the level of molecular mechanism.

Conclusion

We comprehensively investigated the effects of CRGs on progression of HNSC at multi-omics levels, and constructed a 12-CRGs signature and nomogram for the prediction of HNSC prognosis and clinical use, respectively. Our study demonstrated that there was significant difference in expression and genes variants of CRGs between HNSC and normal tissues. ISCA2 is a CRG chosen for further analysis and was found to be upregulated in HNSC and was closely related to the prognosis of HNSC patients. Importantly, the 12-CRGs signature and nomogram had significant effects on prognosis of HNSC patients. The 12-CRGs signature was significantly associated with suppression of $CD4^+$ T cell activation and antigen processing and presentation. The significant association between the expression of CRGs and HNSC progression indicated that CRGs may potential roles as diagnostic, therapeutic and prognostics biomarkers for HNSC.

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

Human participants involved in this study were reviewed by TCGA, GEO or Human Protein Atlas belonging to public databases, and have obtained ethical approval from respective ethics committee. Users can download relevant data from above public databases for free for research and publish relevant articles.

Author contributions

XZ and CZ: Investigation, Methodology, Software, Formal analysis, Data Curation, Writing—Original Draft, Visualization; DZ and QG: Investigation, Formal analysis, Data Curation, Visualization; MM: Methodology, Formal analysis, Data Curation, Visualization; LX: Data Curation and Formal analysis; WG, YW, and XZ: Conceptualization, Resources, Writing—review and editing, Supervision, Project administration. All authors read and approved the manuscript.

Funding

This work was supported by Fundamental Research Program of Shanxi Province (grant numbers 20210302124088 and 201805D211007), the Research Project of Shanxi Province Health and Family Planning Commission (grant number 2019033); the

National Natural Science Foundation of China (grant number 82073101); Shenzhen Science and Technology Program (grant number RCJC20210706091950028); Research Project Supported by Shanxi Scholarship Council of China (grant number 2020165); Fund for the Scientific Activities of Selected Return Overseas Professionals in Shanxi Province (grant number 20200034); Research Funds for China Central Government-guided Development of Local Science and Technology (grant number 2020-165-19).

Acknowledgments

All authors thank TCGA, GEO, UCSC Xena, Human Protein Atlas, STRING, GEPIA2 databases and corresponding data contributors for sharing the data about HNSC.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1084206/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Graphical presentation of the workflow of this study.

SUPPLEMENTARY FIGURE S2
Boxplot showing of CRGs expression in paired samples extracted from all TCGA HNSC samples.

SUPPLEMENTARY FIGURE S3
Heat map of CNVs in HNSC. **(A,B)** Heat map of CNV in TCGA-1st **(A)** and TCGA-2nd **(B)** cohorts.

SUPPLEMENTARY FIGURE S4
Effects of ISCA2 on HNSC progression in the validation cohort. **(A–C)** Kaplan-Meier curve showing the association of ISCA2 with OS **(A)**, PFI **(B)**, and DSS **(C)** over 5 years in TCGA-2nd cohort. **(D,E)** Kaplan-Meier curve showing the association of ISCA2 with OS **(D)** and progression-free survival **(E)** over 5 years in the GSE65858 cohort.

SUPPLEMENTARY TABLE S1
Clinical features of HNSC patients in current study.

SUPPLEMENTARY TABLE S2
Cuproptosis-related gene list was analyzed in current study.

SUPPLEMENTARY TABLE S3
CRPs list obtained from the STRING database.

References

Ahluwalia, P., Kolhe, R., and Gahlay, G. K. (2021). The clinical relevance of gene expression based prognostic signatures in colorectal cancer. *Biochim. Biophys. Acta Rev. Cancer* 1875 (2), 188513. doi:10.1016/j.bbcan.2021.188513

Bellido, F., Pineda, M., Aiza, G., Valdés-Mas, R., Navarro, M., Puente, D. A., et al. (2016). POLE and POLD1 mutations in 529 kindred with familial colorectal cancer and/or polyposis: Review of reported cases and recommendations for genetic testing and surveillance. *Genet. Med.* 18 (4), 325–332. doi:10.1038/gim.2015.75

Bian, Z., Fan, R., and Xie, L. (2022). A novel cuproptosis-related prognostic gene signature and validation of differential expression in clear cell renal cell carcinoma. *Genes (Basel)* 13 (5), 851. doi:10.3390/genes13050851

Cheek, D. M., and Naxerova, K. (2022). Mapping the long road to cancer. *Cell* 185 (6), 939–940. doi:10.1016/j.cell.2022.02.020

Crago, A. M., Socci, N. D., DeCarolis, P., O'Connor, R., Taylor, B. S., Qin, L., et al. (2012). Copy number losses define subgroups of dedifferentiated liposarcoma with poor prognosis and genomic instability. *Clin. Cancer Res.* 18 (5), 1334–1340. doi:10.1158/1078-0432.CCR-11-2820

Dai, F., Dai, L., Zheng, X., Guo, Y., Zhang, Y., Niu, M., et al. (2020). Non-coding RNAs in drug resistance of head and neck cancers: A review. *Biomed. Pharmacother.* 127110231, 110231. doi:10.1016/j.biopha.2020.110231

Dixon, J. R., Xu, J., Dileep, V., Zhan, Y., Song, F., Le, V. T., et al. (2018). Integrative detection and analysis of structural variation in cancer genomes. *Nat. Genet.* 50 (10), 1388–1398. doi:10.1038/s41588-018-0195-8

Ekanem, T. I., Huang, C., Wu, M., Lin, D., Lai, W. T., and Lee, K. (2019). Glycidamide promotes the growth and migratory ability of prostate cancer cells by changing the protein expression of cell cycle regulators and epithelial-to-mesenchymal transition (EMT)-Associated proteins with prognostic relevance. *Int. J. Mol. Sci.* 20 (9), 2199. doi:10.3390/ijms20092199

Fang, Q., George, A. S., Brinkmeier, M. L., Mortensen, A. H., Gergics, P., Cheung, L. Y. M., et al. (2016). Genetics of combined pituitary hormone deficiency: Roadmap into the genome era. *Endocr. Rev.* 37 (6), 636–675. doi:10.1210/er.2016-1101

Galasso, M., Gambino, S., Romanelli, M. G., Donadelli, M., and Scupoli, M. T. (2021). Browsing the oldest antioxidant enzyme: Catalase and its multiple regulation in cancer. *Free Radic. Biol. Med.* 172, 172264–172272. doi:10.1016/j.freeradbiomed.2021.06.010

Ge, E. J., Bush, A. I., Casini, A., Cobine, P. A., Cross, J. R., DeNicola, G. M., et al. (2022). Connecting copper and cancer: From transition metal signalling to metalloplasia. *Nat. Rev. Cancer* 22 (2), 102–113. doi:10.1038/s41568-021-00417-2

Goh, W. Q. J., Ow, G. S., Kuznetsov, V. A., Chong, S., and Lim, Y. P. (2015). DLAT subunit of the pyruvate dehydrogenase complex is upregulated in gastric cancer-implications in cancer therapy. *Am. J. Transl. Res.* 7 (6), 1140–1151.

Green, Y. S., Santos, M. C. F. d., Fuja, D., Reichert, E., Campos, A. R., Cowman, S. J., et al. (2022). ISCA2 inhibition decreases HIF and induces ferroptosis in clear cell renal carcinoma. *bioRxiv*. doi:10.1101/2022.06.01.494206

Hara, T., Chanoch-Myers, R., Mathewson, N. D., Myskiw, C., Atta, L., Bussema, L., et al. (2021). Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. *Cancer Cell* 39 (6), 779–792.e11. e11. doi:10.1016/j.ccell.2021.05.002

Kim, J., Matsubara, T., Lee, J., Fenollar-Ferrer, C., Han, K., Kim, D., et al. (2021). Lysosomal SLC46A3 modulates hepatic cytosolic copper homeostasis. *Nat. Commun.* 12 (1), 290. doi:10.1038/s41467-020-20461-0

Kudva, A. K., Raghu, S. V., Achar, P. K., Rao, S., Suresh, S., and Baliga, M. S. (2021). Study of serum zinc and copper levels and tumor pathology: A pilot study in people affected with head and neck cancers. *Indian J. Otolaryngology Head Neck Surg.* doi:10.1007/s12070-021-02589-7

Lebigot, E., Schiff, M., and Golinelli-Cohen, M. (2021). A review of multiple mitochondrial dysfunction syndromes, syndromes associated with defective Fe-S protein maturation. *Biomedicines* 9 (8), 989. doi:10.3390/biomedicines9080989

Lee J, J., You, J. H., Shin, D., and Roh, J. (2020). Inhibition of glutaredoxin 5 predisposes cisplatin-resistant head and neck cancer cells to ferroptosis. *Theranostics* 10 (17), 7775–7786. doi:10.7150/thno.46903

Lee My, M. Y., Jeon, J. W., Sievers, C., and Allen, C. T. (2020). Antigen processing and presentation in cancer immunotherapy. *J. Immunother. Cancer* 8 (2), e001111. doi:10.1136/jitc-2020-001111

Lenz, S., Sinn, L. R., O'Reilly, F. J., Fischer, L., Wegner, F., and Rappsilber, u. (2021). Reliable identification of protein-protein interactions by crosslinking mass spectrometry. *Nat. Commun.* 12 (1), 3564. doi:10.1038/s41467-021-23666-z

Li, S., Bu, L., and Cai, L. (2022). Cuproptosis: Lipoylated TCA cycle proteins-mediated novel cell death pathway. *Signal Transduct. Target Ther.* 7 (1), 158. doi:10.1038/s41392-022-01014-x

Li, T., Peng, J., Zeng, F., Zhang, K., Liu, J., Li, X., et al. (2017). Association between polymorphisms in CTR1, CTR2, ATP7A, and ATP7B and platinum resistance in epithelial ovarian cancer. *Int. J. Clin. Pharmacol. Ther.* 55 (10), 774–780. doi:10.5414/CP202907

Li, X., Yin, J., Wang, Y., He, H., Li, X., Gong, W., et al. (2014). The ATP7B genetic polymorphisms predict clinical outcome to platinum-based chemotherapy in lung cancer patients. *Tumour Biol.* 35 (8), 8259–8265. doi:10.1007/s13277-014-2072-0

Magrin, L., Fanale, D., Brando, C., Fiorino, A., Corsini, L. R., Sciacchitano, R., et al. (2021). POLE, POLD1, and NTHL1: The last but not the least hereditary cancer-predisposing genes. *Oncogene* 40 (40), 5893–5901. doi:10.1038/s41388-021-01984-2

Mamelak, A. J., Kowalski, J., Murphy, K., Yadava, N., Zahurak, M., Kouba, D. J., et al. (2005). Downregulation of NDUFA1 and other oxidative phosphorylation-related genes is a consistent feature of basal cell carcinoma. *Exp. Dermatol.* 14 (5), 336–348. doi:10.1111/j.0906-6705.2005.00278.x

Miggelbrink, A. M., Jackson, J. D., Lorrey, S. J., Srinivasa, E. S., Waibl-Polania, J., Wilkinson, D. S., et al. (2021). CD4 T-cell exhaustion: Does it exist and what are its roles in cancer. *Clin. Cancer Res.* 27 (21), 5742–5752. doi:10.1158/1078-0432.CCR-21-0206

Qin, C., Lu, Y., Zhang, H., Zhang, Z., Xu, W., Wen, S., et al. (2022). Biological roles and clinical significance of estrogen and androgen receptors in head and neck cancers. *J. Cancer* 13 (7), 2189–2199. doi:10.7150/jca.66707

Remacha, L., Pirman, D., Mahoney, C. E., Coloma, J., Calsina, B., Currás-Freixes, M., et al. (2019). Recurrent germline DLST mutations in individuals with multiple pheochromocytomas and paragangliomas. *Am. J. Hum. Genet.* 104 (4), 1008–1010. doi:10.1016/j.ajhg.2019.04.010

Ren, F., Logeman, B. L., Zhang, X., Liu, Y., Thiele, D. J., and Yuan, P. (2019). X-ray structures of the high-affinity copper transporter Ctr1. *Nat. Commun.* 10 (1), 1386. doi:10.1038/s41467-019-09376-7

Ressnerova, A., Raudenska, M., Holubova, M., Svobodova, M., Polanska, H., Babula, P., et al. (2016). Zinc and copper homeostasis in head and neck cancer: Review and meta-analysis. *Curr. Med. Chem.* 23 (13), 1304–1330. doi:10.2174/0929867323666160405111543

Sajnani, K., Islam, F., Smith, R. A., Gopalan, V., and Lam, A. K. (2017). Genetic alterations in Krebs cycle and its impact on cancer pathogenesis. *Biochimie*, 135164–135172. doi:10.1016/j.biochi.2017.02.008

Shan, N., Li, N., Dai, Q., Hou, L., Yan, X., Amei, A., et al. (2020). Interplay of tRNA-derived fragments and T cell activation in Breast cancer patient survival. *Cancers (Basel)* 12 (8), 2230. doi:10.3390/cancers12082230

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science* 375 (6586), 1254–1261. doi:10.1126/science.abf0529

Wang, F., Zhao, Q., Wang, Y., Jin, Y., He, M., Liu, Z., et al. (2019). Evaluation of POLE and POLD1 mutations as biomarkers for immunotherapy outcomes across multiple cancer types. *JAMA Oncol.* 5 (10), 1504–1506. doi:10.1001/jamaoncol.2019.2963

Weiler, B. D., Brück, M., Kothe, I., Bill, E., Lill, R., and Mühlenhoff, U. (2020). Mitochondrial [4Fe-4S] protein assembly involves reductive [2Fe-2S] cluster fusion on ISCA1-ISCA2 by electron flow from ferredoxin FDX2. *Proc. Natl. Acad. Sci. U.S.A.* 117 (34), 20555–20565. doi:10.1073/pnas.2003982117

Wu, Y., Zhang, Y., Zheng, X., Dai, F., Lu, Y., Dai, L., et al. (2020). Circular RNA circCORO1C promotes laryngeal squamous cell carcinoma progression by modulating the let-7c-5p/PBX3 axis. *Mol. Cancer* 19 (1), 99. doi:10.1186/s12943-020-01215-4

Yang, P. 1., Yu, H., Jin, Y., Qu, F., Ren, H., Tang, Z., et al. (2021). Construction and validation of an immune infiltration-related gene signature for the prediction of prognosis and therapeutic response in Breast cancer. *Front. Immunol.* 12666137, 666137. doi:10.3389/fimmu.2021.666137

Zhang, Y., Zeng, T., Chen, L., Huang, T., and Cai, Y. (2021). Determining protein-protein functional associations by functional rules based on gene ontology and KEGG pathway. *Biochim. Biophys. Acta Proteins Proteom* 1869 (6), 140621. doi:10.1016/j.bbapap.2021.140621

Zheng, L., Meng, Y., Campbell, J. L., and Shen, B. (2020). Multiple roles of DNA2 nuclease/helicase in DNA metabolism, genome stability and human diseases. *Nucleic Acids Res.* 48 (1), 16–35. doi:10.1093/nar/gkz1101

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fgene.2022.1070630

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Maria Caffo,
University of Messina, Italy
Xulei Huo,
Beijing Tiantan Hospital, China

*CORRESPONDENCE
Roujun Peng,
pengrj@sysucc.org.cn

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 15 October 2022
ACCEPTED 05 December 2022
PUBLISHED 04 January 2023

CITATION
Luo W, Quan Q, Jiang J and Peng R (2023), An immune and epithelial–mesenchymal transition-related risk model and immunotherapy strategy for grade II and III gliomas. *Front. Genet.* 13:1070630. doi: 10.3389/fgene.2022.1070630

An immune and epithelial–mesenchymal transition-related risk model and immunotherapy strategy for grade II and III gliomas

Wei Luo†, Qi Quan†, Jiaxin Jiang and Roujun Peng*

Department of VIP Section, State Key Laboratory of Oncology in South China, Sun Yat-sen University Cancer Center, Guangzhou, Guangzhou, China

Grade II and III gliomas are heterogeneous and aggressive diseases. More efficient prognosis models and treatment methods are needed. This study aims to construct a new risk model and propose a new strategy for grade II and III gliomas. The data were downloaded from The Cancer Genome Atlas (TCGA), the Gene Expression Omnibus (GEO), gene set enrichment analysis (GSEA), and the EMTome website for analysis. The Human Cell Landscape website and the Genomics of Drug Sensitivity in Cancer website were used for single-cell analysis and drug susceptibility analysis. Gene set enrichment analysis, gene function enrichment analysis, univariate and multivariate Cox regression analyses, Pearson's correlation analysis, log-rank test, Kaplan–Meier survival analysis, and ROC analysis were performed. We constructed an immune-related prognostic model associated with the isocitrate dehydrogenase 1 (IDH1) mutation status. By analyzing the immune microenvironment of patients with different risk scores, we found that high-risk patients were more likely to have an inflammatory immune microenvironment and a higher programmed death ligand-1 (PD-L1) expression level. Epithelial–mesenchymal transition (EMT)-related gene sets were significantly enriched in the high-risk group, and the epithelial–mesenchymal transition phenotype was associated with a decrease in $CD8^+$ T cells and an increase in M2 macrophages. Transforming growth factor-β (TGF-β) signaling was the most important signaling in inducing epithelial–mesenchymal transition, and TGFB1/TGFBR1 was correlated with an increase in $CD8^+$ T cytopenia and M2 macrophages. Survival analysis showed that simultaneous low expression of TGFBR1 and PD-L1 had better survival results. Through single-cell analysis, we found that TGFB1 is closely related to microglia and macrophages, especially M2 macrophages. Finally, we discussed the sensitivity of TGFB1 inhibitors in gliomas using cell line susceptibility data. These results demonstrated a potential immunotherapy strategy in combination with the TGFB1/TGFBR1 inhibitor and PD-1/PD-L1 inhibitor for grade II and III gliomas.

KEYWORDS

glioma, IDH1, immune, EMT, TGF-β, PD-L1

1 Introduction

Grade II and III gliomas are the most common primary brain tumors and proved to be with substantial heterogeneity in terms of pathological features and clinical outcomes (Louis et al., 2021). In order to distinguish the different pathological features of patients and develop individualized treatment strategies, glioma-related biomarkers have been identified. The IDH1 was reported to mutate frequently in gliomas (Yan et al., 2009). It had been proved that gliomas with the IDH1 mutation were more sensitive to chemotherapy and radiotherapy, resulting in a better prognosis (Sanson et al., 2009). Other biomarkers, for example, codeletion of 1p and 19q (1p/19q) (Smith et al., 2000), Capicua (CIC) transcriptional repressor mutation (Gleize et al., 2015), loss of chromosome 9p, mutation of phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (PIK3CA) and phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1) (Draaisma et al., 2015), and deletion of cyclin-dependent kinase inhibitor 2 A (CDKN2A) (Reis et al., 2015), were confirmed to have prognostic value and important for rational selection of surgery, radiotherapy, and chemotherapy treatment. However, over 50% of grade II and III gliomas eventually develop into highly aggressive gliomas, indicating the need for a more efficient prognosis model and treatment methods (Zhu et al., 2021).

In the past decades, there has been little progress in the treatment of gliomas. Although immunotherapy successfully promoted the treatment results of other cancer types and was the major research direction for gliomas, limited progression had been made in immunotherapy treatment of gliomas (Yang et al., 2022). In CheckMate-498, a randomized clinical trial, nivolumab combined with bevacizumab and nivolumab combined with chemoradiotherapy in newly diagnosed glioma patients with O^6-methylguanine DNA methyltransferase (MGMT) promoter unmethylation were both ineffective (Reardon et al., 2020). Comparing with IDH-wild gliomas, the IDH-mutant gliomas have significantly low tumor-infiltrating lymphocytes and PD-L1 expression (Berghoff et al., 2017). Numerous studies had demonstrated that gliomas were infiltrated by immune cells that made up to 30% of a tumor's mass (Kaminska et al., 2021). Such extensive accumulation of innate immune cells in gliomas might be misleading as these events did not reflect the effective anti-tumor immunity. This phenomenon enlightened us that tumor immune infiltration in gliomas might be accompanied by other pathological processes that promote tumor progression.

As mentioned previously, the construction of a new risk model and improving the efficacy of immunotherapy are essential for the treatment of gliomas. In this study, we attempted to construct an immune-related risk model and proposed a feasible strategy for improving the immunotherapy efficacy for gliomas.

2 Materials and methods

2.1 RNA sequencing data

The IDH1 somatic mutation status for 500 samples, gene expression data for 525 samples, and the corresponding clinical datasheets for 515 samples were obtained from TCGA website (https://portal.gdc.cancer.gov/). Among these grade II and III glioma samples, 493 samples with RNA sequencing data and IDH1 mutation information were subjected to subsequent analyses. Log2 (x+1) normalization was performed for all gene expression data. Rows and columns with more than 50% missing values were removed. The study report fully met TCGA publication requirements.

2.2 Microarray data

The gene expression profile matrix files from GSE107850 (including 195 samples) and GSE43388 (including 43 samples, 15 from GSE43388-GPL570 and 28 from GSE43388-GPL8542) were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). We used the R package inSilicoMerging [DOD: 10.1186/1471–2105–13–335] to merge the datasets. Also, we used empirical Bayes methods (Johnson et al., 2007) to remove batch effects.

2.3 Construction and validation of the risk model

GSEA was performed to determine how the immunological pathways and corresponding immune genes differ between IDH1-wild (IDH1-wt) (n = 117) and IDH1 mutation (IDH1-mut) (n = 376) samples in TCGA cohort. An annotated gene set file (c7. immunesigdb.v7.4. symbols.gmt; downloaded from the Molecular Signatures Database) was selected for use as the reference gene set. The expression profiles of the top 50 genes expressed in the IDH1-wt and IDH1-mut groups were analyzed *via* univariate Cox regression analysis. In this analysis, genes were regarded as significant at $p < 0.05$. For the significant genes, least absolute shrinkage and selection operator (LASSO) Cox analysis was performed by using the glmnet R package. Then, the nine candidate genes were analyzed by multivariate Cox regression analysis based on progression-free survival (PFS). Finally, two independent prognostic factors for PFS were analyzed by multivariate Cox regression analysis based on overall survival (OS) to construct the risk model ($\text{Risk Score} = \sum_i Coefficient\ of\ (i) * Expression of\ gene\ (i)$). The log-rank test and Kaplan–Meier survival analysis were used to assess the predictive ability of the prognostic model. The maxstat R package was applied to determine the best cutoff value,

and then patients were classified into low-risk and high-risk groups. We performed ROC analysis using the R package pROC.

2.4 Correlation of immunophenotype with the risk model

Using the deconvo_CIBERSOR method of the IOBR R package, the immune cell infiltration score was calculated. The R package ESTIMATE was used to calculate the immune scores for each patient based on gene expression. Pearson's correlation coefficients for risk scores and immune infiltration were calculated. The rank sum test was used to compare differences in immune cells and immune-related molecules between high-risk and low-risk groups. $p < 0.05$ was considered significant.

2.5 Functional enrichment analysis

The STRING website (https://cn.string-db.org/) was used to construct a 32-gene functional protein association network. We used the Kyoto Encyclopedia of Genes and Genomes Application Programming Interface (KEGG API) to obtain genetic annotations for the latest KEGG pathway. Enrichment analysis was performed using the R package clusterProfiler.

2.6 Epithelial–mesenchymal transition (EMT) gene set-related analysis

The 61 EMT gene sets were downloaded from the EMTome website (http://emtome.org/) (Vasaikar et al., 2021). GSEA was performed between the high-risk and low-risk groups. The enrichment scores of the EMT gene sets were analyzed by univariate Cox analysis through the EMTome website. The gene set with the highest prognostic significance and the highest GSEA enrichment score was selected for gene set variation analysis (GSVA) to obtain the enrichment scores of samples. Pearson's correlation coefficients for enrichment scores, risk scores, immune cells, and immune-related molecules were calculated. The log-rank test and Kaplan–Meier survival analysis were applied to assess the predictive ability of the enrichment score. By reviewing the literature, we identified five EMT phenotype-related signaling pathways and downloaded corresponding gene sets from the GSEA website (http://www.gsea-msigdb.org/), namely, REACTOME_SIGNALING_BY_TGFB_FAMILY_MEMBERS, GOBP_CANONICAL_WNT_SIGNALING_PATHWAY, BIOCARTA_RAS_PATHWAY, GOBP_NOTCH_SIGNALING_PATHWAY, and GOBP_PHOSPHATIDYLINOSITOL_3_KINASE_SIGNALING (Dongre and Weinberg, 2019). We divided the samples into two groups according to the enrichment score, with 263 samples in the high-enrichment score group and 262 samples in the low-enrichment score group, and performed GSEA to obtain the most significant enrichment signaling pathway.

2.7 Effects of transforming growth factor-β (TGF-β)-associated molecules, PD-L1 (CD274), and CTLA-4 on immune cell infiltration and prognosis

Pearson's correlation coefficients for enrichment scores, TGFB1, TGFBR1, TGFB2, and TGFB3 were calculated, and the correlation matrix was plotted. Multivariate Cox regression analysis was performed on TGFB1, TGFBR1, TGFB2, TGFB3, PD-L1, and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4). For TGFB1, TGFBR1, CD274, CTLA-4, and immune cells, we calculated the Pearson's correlation coefficient and plotted the correlation scatterplot. The log-rank test and Kaplan–Meier survival analysis were used to assess the predictive ability of TGFBR1 and PD-L1. The maxstat R package was applied to determine the best cutoff value, and then patients were classified into the low-expression group and the high-expression group. The samples were then divided into four groups (TGFBR1-H+CD274-L, TGFBR1-H+CD274-H, TGFBR1-L+CD274-L, and TGFBR1-l+CD274-H) for survival analysis, and the differences in immune cell infiltration between groups were compared by the rank sum test, and violin charts were plotted.

2.8 Single-cell analysis was performed to determine cell localization of TGF-β signaling-related molecules

This analysis was conducted through the Human Cell Landscape website (https://db.cngb.org/HCL/index.html) (Han et al., 2020). Platform creators analyzed >700,000 single cells from >50 human tissues (2–4 replicates per tissue in general) and cultures. Through the brain section of the Gallery module, we can acquire the single-cell data matrix and analysis results related to brain tissue.

2.9 Anti-TGFB1 drug susceptibility analysis

We adopted the Genomics of Drug Sensitivity in Cancer website (https://www.cancerrxgene.org/) and the EMTome website (http://emtome.org/) for drug sensitivity analysis. The two websites provided drug susceptibility data on the TGFB1 inhibitor LY2109761 in different cell lines, as well as online analysis tools. Using online tools from both websites, we performed drug susceptibility analysis.

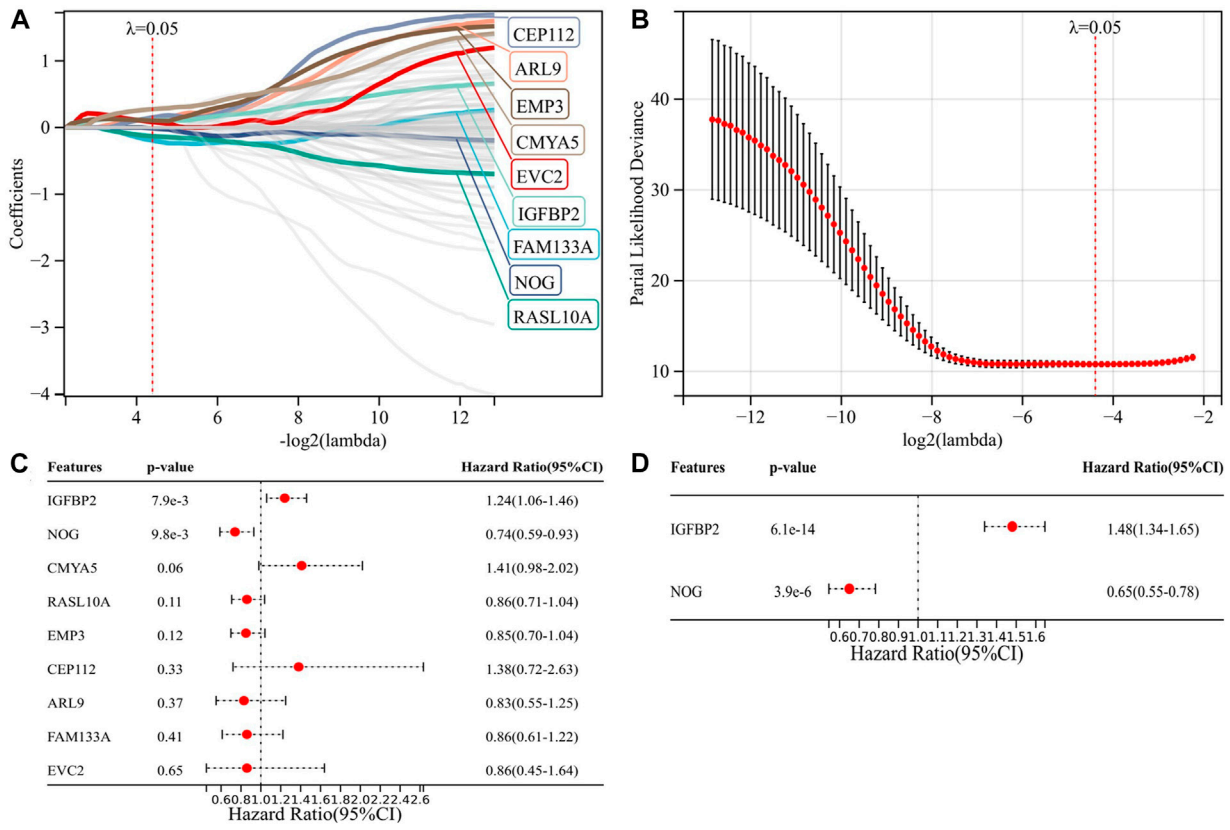


FIGURE 1
Construction of a risk model associated with IDH1 and immune status. **(A–B)** Nine candidate genes for risk modeling by LASSO Cox regression analysis; **(C)** multivariate analysis of nine genes on PFS; and **(D)** multivariate analysis of two genes on OS.

2.10 Data processing platform

All the data processing was performed on Sangerbox (http://vip.sangerbox.com/home.html), a powerful platform based on R (Shen et al., 2022), including RNA sequencing data normalization, merging of datasets, removing batch effects, gene set enrichment analysis, gene set variation analysis, gene function enrichment analysis, univariate and multivariate Cox regression analyses, LASSO-Cox regression analysis, Pearson's correlation analysis, log-rank test, Kaplan–Meier survival analysis, ROC analysis, the immune cell infiltration score and immune infiltration score calculation, and KEGG enrichment analysis.

3 Results

3.1 Construction and validation of a risk model associated with IDH1 and immune status

We applied GSEA between IDH1-wt (n = 117) and IDH1-mut (n = 376) groups using c7. immunesigdb.v7.4. symbols.gmt as the reference gene set. All 3919 immune-related gene sets with FDR < 0.05 were enriched in the IDH1-wt group, suggesting that the IDH1-wt group was more correlated with immune response than the IDH1-mut group. The top 50 genes enriched in the two groups were used for univariate Cox regression analysis. The results showed that all 100 genes were significantly associated with prognosis (Figure S1). These 100 genes were then analyzed by LASSO Cox, and nine candidate genes were obtained for risk modeling (Figures 1A–B). In order to better predict PFS, we performed multivariate analysis on PFS for nine genes and obtained two genes with independent prognostic value for PFS, NOG, and IGFBP2 (Figure 1C). Through multivariate analysis of NOG and IGFBP2 in OS, we obtained the expression coefficients of these two genes and constructed a risk model (Figure 1D). Risk score = -0.4253 * expression of NOG + 0.3954 *expression of IGFBP2.

We calculated the risk scores of all 493 TCGA samples and performed survival analysis, plotting K-M curves and ROC curves. The results showed that the risk model was a good predictor of patients' OS (p = 1.3e-24, HR = 5.52) and PFS (p = 3.2e-13, HR = 3.20) (Figures 2A–D). Clinically, the

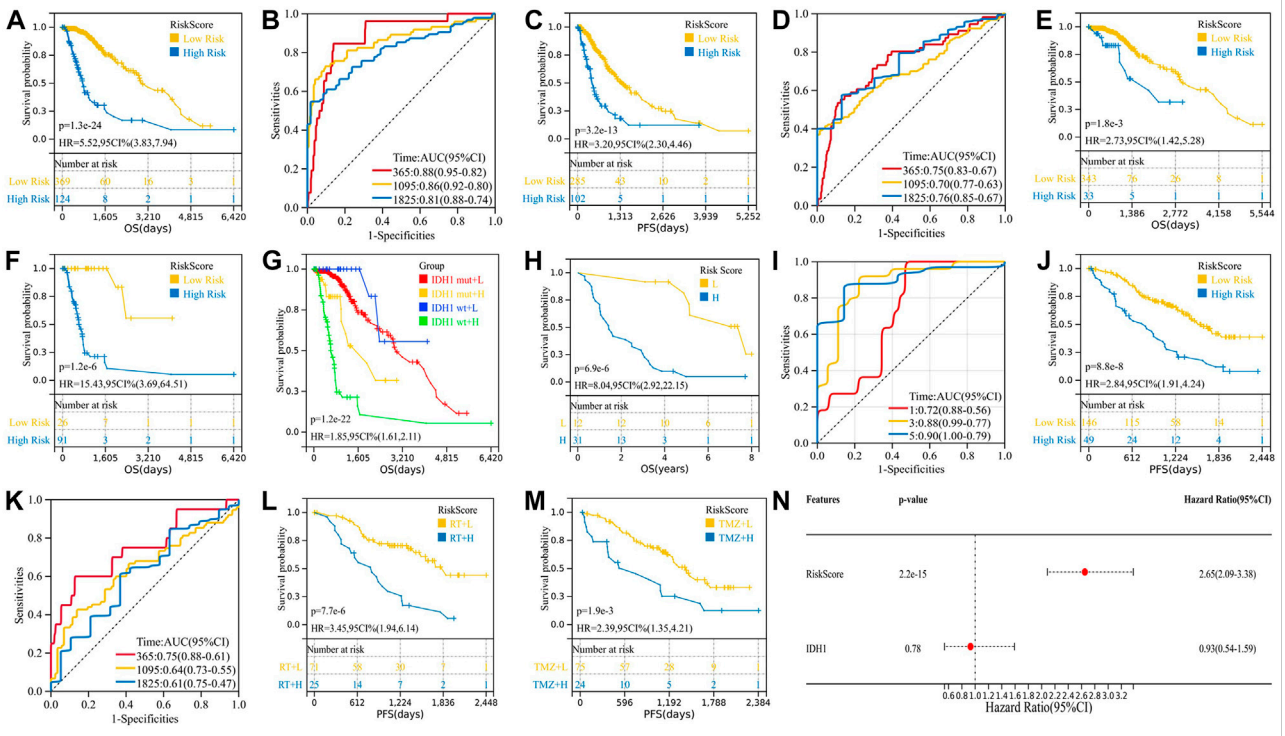


FIGURE 2
Validation of the two-gene risk model. **(A)** K-M survival curve for OS of TCGA cohort; **(B)** ROC curve for OS of TCGA cohort; **(C)** K-M survival curve for PFS of TCGA cohort; **(D)** ROC curve for PFS of TCGA cohort; **(E)** high-risk patients in the IDH1-mut group have shorter OS; **(F)** high-risk patients in the IDH1-wt group have shorter OS; **(G)** K-M curve combining the IDH1 mutation status and risk score; **(H–I)** validation of the two-gene risk model in GSE43388; **(J–M)** validation of the two-gene risk model in GSE107850; and **(N)** multivariate analysis of IDH1 and risk score.

IDH1 mutation status was used to determine patient prognosis, and we explored whether our risk model could further stratify the IDH1 mutation status. In both the IDH1-mut and IDH1-wt groups, a high-risk score indicates a worse prognosis (Figures 2E–F), whereas the IDH1 status in the low-risk group does not affect prognosis (Figure 2G). Multivariate analysis of IDH1 and risk score suggested that risk score is an independent prognostic factor relative to IDH1 (Figure 2N).

GSE107850 (including 195 samples) and GSE43388 (including 43 samples) were used to validate the risk prediction capabilities of risk models. Among them, the study endpoint of the GSE107850 sample was PFS after radiotherapy (RT) and temozolomide (TMZ) treatment, and the study endpoint of the GSE43388 sample was OS, so we used these two sets of sample data to verify the predictive efficacy of risk models on PFS and OS, respectively. Furthermore, we used the GSE107850 dataset to analyze the predictive power of risk models for response to radiotherapy and chemotherapy treatment. As shown in Figures 2H–M, the risk model can predict not only the patient's OS (p = 3.2e-6, HR = 7.89) and PFS (p = 8.8e-8, HR = 2.84) but also the patient's response to radiation (p = 7.7e-6, HR = 3.45) and chemotherapy (p = 1.9e-3, HR = 2.39).

3.2 High-risk patients are more likely to have a "hot" immune microenvironment

We analyze the correlation of the risk model with immune infiltration scores, immune cell infiltration, and expression of immune molecules. There is a significant positive correlation between immune risk scores and immune infiltration scores, including stromal score (p = 6.2e-36, r = 0.52), immune score (p = 1.8e-18, r = 0.38), and ESTIMATE score (p = 1.0e-25, r = 0.45) (Figures 3A–C). The high-risk group had higher expression of immunosuppressive molecules (CD274, CTLA-4, IDO1, and IL10) and also had higher expression of immunostimulant molecules (CD27, CD28, CD40, CD40LG, and ICOS) (Figures 3D–E). Regulatory T cells (Tregs) were elevated in the high-risk group (p = 1.1e-6). However, B naïve cells (p = 4.0e-4), mast resting cells (p = 4.3e-3), CD4+ T memory resting cells (p = 6.1e-10), $CD8^+$ T cells (p = 6.6e-5), neutrophils (p = 4.4e-5), and M0 (p = 3.7e-5) and M1 (p = 2.2e-6) macrophages were also elevated in the high-risk group. Risk stratification did not appear to affect M2 macrophage infiltration results (p = 0.62) (Figure 3F). From these results, a higher level of inflammation coexisted with a higher level of immunosuppression, but unfortunately, although this immunosuppression was accompanied by increased PD-L1

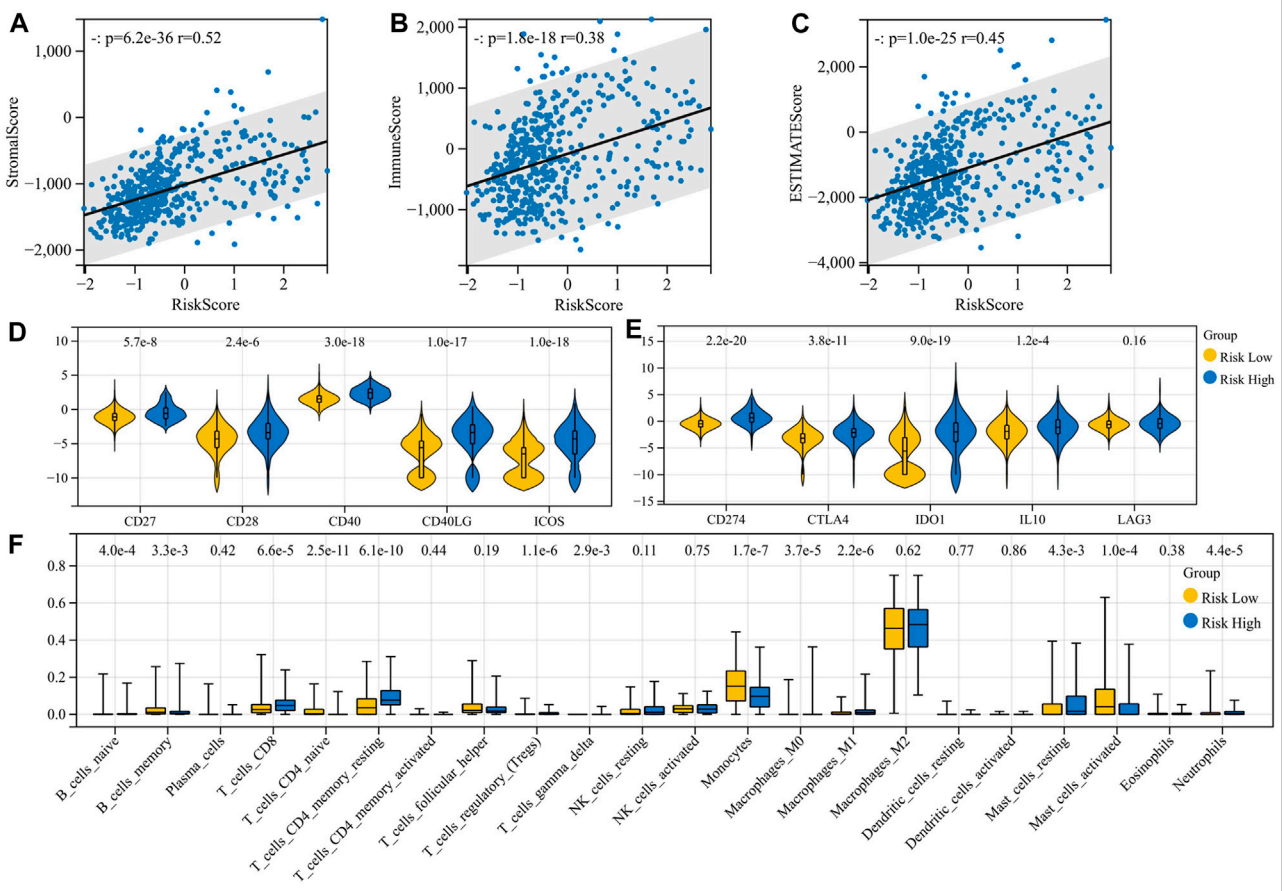


FIGURE 3
Immunocorrelation analysis of the risk model. **(A–C)** Correlation of the risk model with immune infiltration scores; **(D–E)** correlation of the risk model with the expression of immune molecules; and **(F)** correlation of the risk model with immune cell infiltration.

expression, it could not be reversed by PD-L1 inhibitors and transformed into clinical benefit. We hypothesized that along with this immunosuppressive process, there were other malignant phenotypes that promoted tumor progression.

3.3 EMT-related signaling pathways are closely related to the risk model

In order to explore the signaling pathways related to the risk model, we first used the STRING website to construct a gene regulatory network of 32 genes for the risk model (Figure 4A), and then we performed KEGG enrichment analysis on this regulatory network (Figure 4B). Some signaling pathways that are closely related to EMT had been enriched, including TGF-β signaling pathway (Hao et al., 2019), signaling pathways regulating pluripotency of stem cells (Mani et al., 2008; Shibue and Weinberg, 2017), Hippo signaling pathway (Cordenonsi et al., 2011), PI3K-Akt signaling pathway, and mTOR signaling pathway (Song et al., 2014). This suggested that the EMT phenotype may be also formed at the same time as the tumor formed an inflammatory immune microenvironment.

3.4 EMT phenotypes are significantly associated with M2 macrophages and are regulated by the TGF-β signaling pathway

The EMTome website provided 61 gene sets that are associated with the EMT phenotype, and we downloaded all of them for GSEA between the two risk groups. At the same time, we used the online tool of the EMTome website to perform univariate Cox regression analysis of each gene set enrichment score. The top gene sets are listed in Table 1; Table 2. PMID29212455: Wang_et_al. 2017 was the gene set with the most significant prognostic value and highest normalized enrichment score (NES), and we plotted the GSEA enrichment curve for this gene set (Figure 5A).

We then performed GSVA with the PMID29212455: Wang_et_al. 2017 gene set and obtained an enrichment score for each

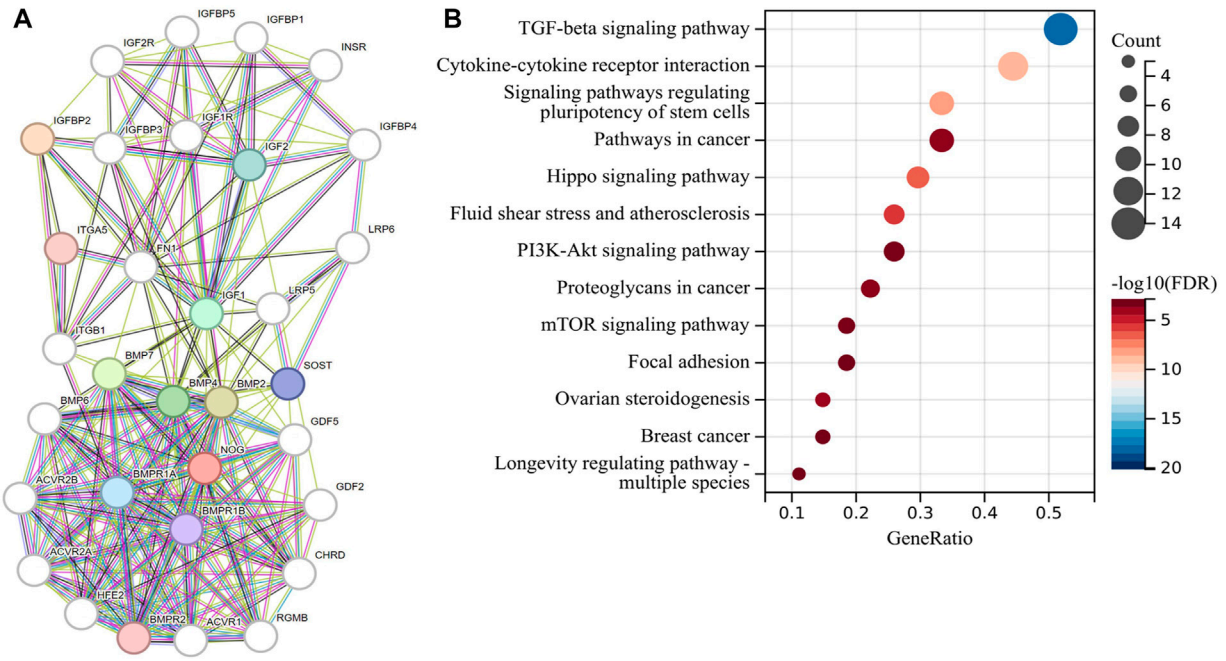


FIGURE 4
Risk model-related gene function enrichment analysis. **(A)** Gene regulatory network of 32 genes related to the risk model and **(B)** KEGG enrichment analysis of this regulatory network.

TABLE 1 Univariate Cox regression analysis for EMT gene sets by the EMTome website.

Signature	Cox coefficient	Hazard ratio	Log-rank *p*-value
PMID29212455: Wang_et_al. 2017	1.38	4 (2.6–6.1)	8.00E-12
PMID26061747: Huang_et_al. 2015	1.22	3.4 (2.2–5.1)	5.40E-10
PMID29700419: Liang_et_al. 2018	1.16	3.2 (2.1–4.8)	3.90E-09
PMID29440769: Chae_et_al. 2018	1.15	3.1 (2.1–4.7)	6.00E-09
PMID20215510: Choi_et_al. 2010	1.13	3.1 (2.1–4.6)	4.80E-09
PMID24004852: Cieslik_et_al. 2013	1.13	3.1 (2.1–4.7)	1.40E-08
PMID19666588: Creighton_et_al. 2009	1.11	3 (2–4.5)	1.50E-08
PMID26088755: Kim_et_al. 2015	1.09	3 (2–4.4)	1.20E-08
PMID26771021: MsigDB_v7.0	1.1	3 (2–4.4)	1.30E-08

Note: By the online tool of the EMTome website, univariate Cox regression analysis was applied to determine the effects of different EMT phenotypes and gene set enrichment scores on the survival of grade II and III gliomas. $p < 0.05$ was considered statistically significant.

sample. Enrichment scores had a significant positive correlation with risk scores (p = 1.1e-29, r = 0.48) (Figure 5B), and high enrichment scores indicated shorter survival (p = 1.0e-9, HR = 2.92) (Figure 5C). The enrichment score was positively correlated with immunosuppressive cells and immunosuppressive molecules, for example, Tregs, M2 macrophages, PD-L1, and CTLA-4, but negatively correlated with B memory cells and CD4+ T cells, suggesting the immunosuppressive properties of the EMT phenotype (Figures 5D–J). In particular, unlike risk scores, EMT enrichment scores were positively correlated with M2 macrophage infiltration (p = 2.1e-8, r = 0.25) but not with $CD8^+$ T-cell infiltration (p = 0.45, r = -0.03).

In order to determine the main signaling pathways that induce EMT phenotypes, we reviewed the relevant literature, identified five candidate signaling pathways, downloaded the gene sets of each pathway through the GSEA website, and

TABLE 2 GSEA rank list for EMT gene sets.

Term	ES	NES	p-value	FDR	FWER
PMID29212455: Wang_et_al. 2017	0.5657	2.3413	0.0000	0.001	0.001
PMID25744723: Schliekelman_et_al. 2015	0.4556	2.2449	0.0000	0.0005	0.001
PMID30728376: Soo Min_et_al. 2019	0.5228	2.2231	0.0000	0.0003	0.001
PMID24510113: Reka_et_al. 2014	0.7297	2.2166	0.0000	0.0003	0.001
PMID29346386: Hollern_et_al. 2018	0.5984	2.2083	0.0000	0.0002	0.001
PMID20713713: Taube_et_al. 2010	0.6191	2.1953	0.0000	0.0001	0.001
PMID23734191: Zarkoob_et_al. 2013	0.6186	2.1953	0.0000	0.0002	0.001
PMID19340593: Joyce_et_al. 2009	0.6535	2.184	0.0000	0.0001	0.001
PMID25214461: Tuan_et_al. 2014	0.594	2.175	0.0000	0.0001	0.001
PMID29700419: Liang_et_al. 2018	0.6864	2.1655	0.0000	0.0003	0.002

Note: EMT gene sets were acquired from the EMTome website, and GSEA was performed between the high-risk and low-risk groups. FDR <0.05 was considered significantly different in gene set enrichment between the two groups.

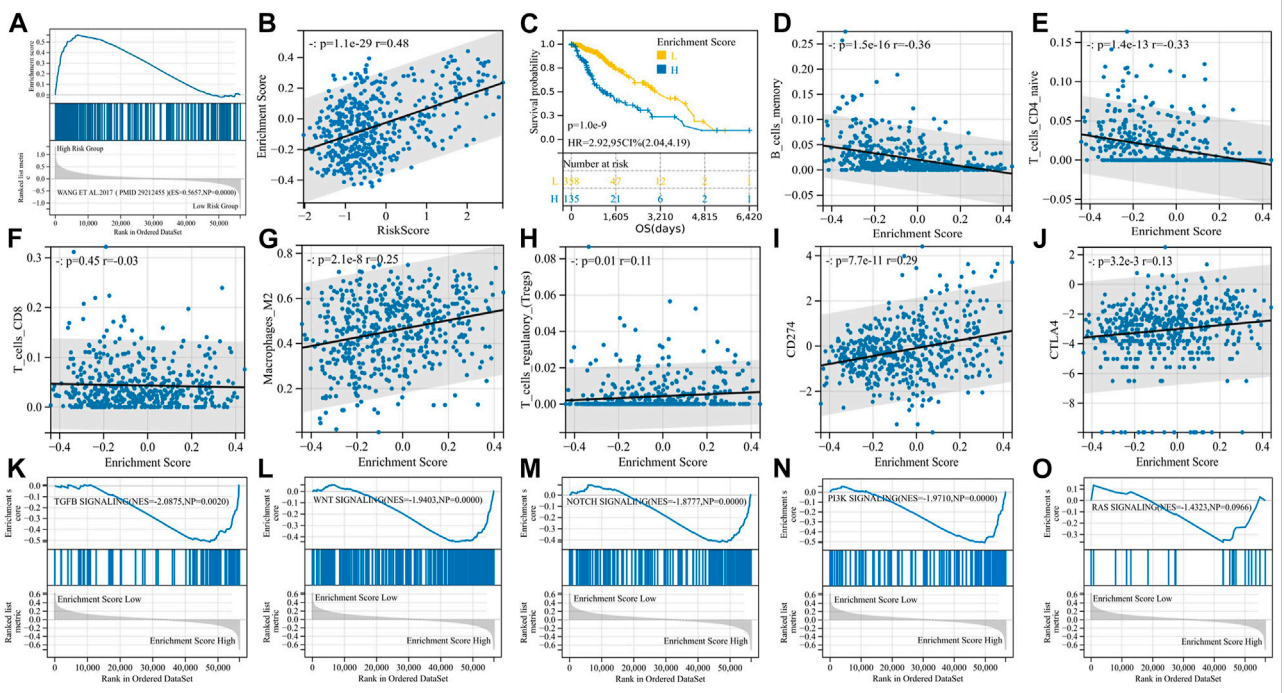


FIGURE 5
EMT phenotypes are significantly associated with M2 macrophages and are regulated by the TGF-β signaling pathway. **(A)** GSEA enrichment curve for the PMID29212455: Wang_et_al. 2017 gene set; **(B)** enrichment score and risk score are significantly positively correlated; **(C)** K-M curve for the enrichment score; **(D–J)** correlation analysis of enrichment scores and immunity; and **(K–O)** GSEA enrichment curve for EMT-related signaling pathways.

then divided the samples into two groups according to the EMT enrichment score for GSEA. The TGF-β signaling pathway, WNT signaling pathway, Notch signaling pathway, and PI3K signaling pathway were all significantly enriched in the high-enrichment score group, with the TGF-β signaling pathway having the largest NES (|NES| = 2.0875), and we believed that

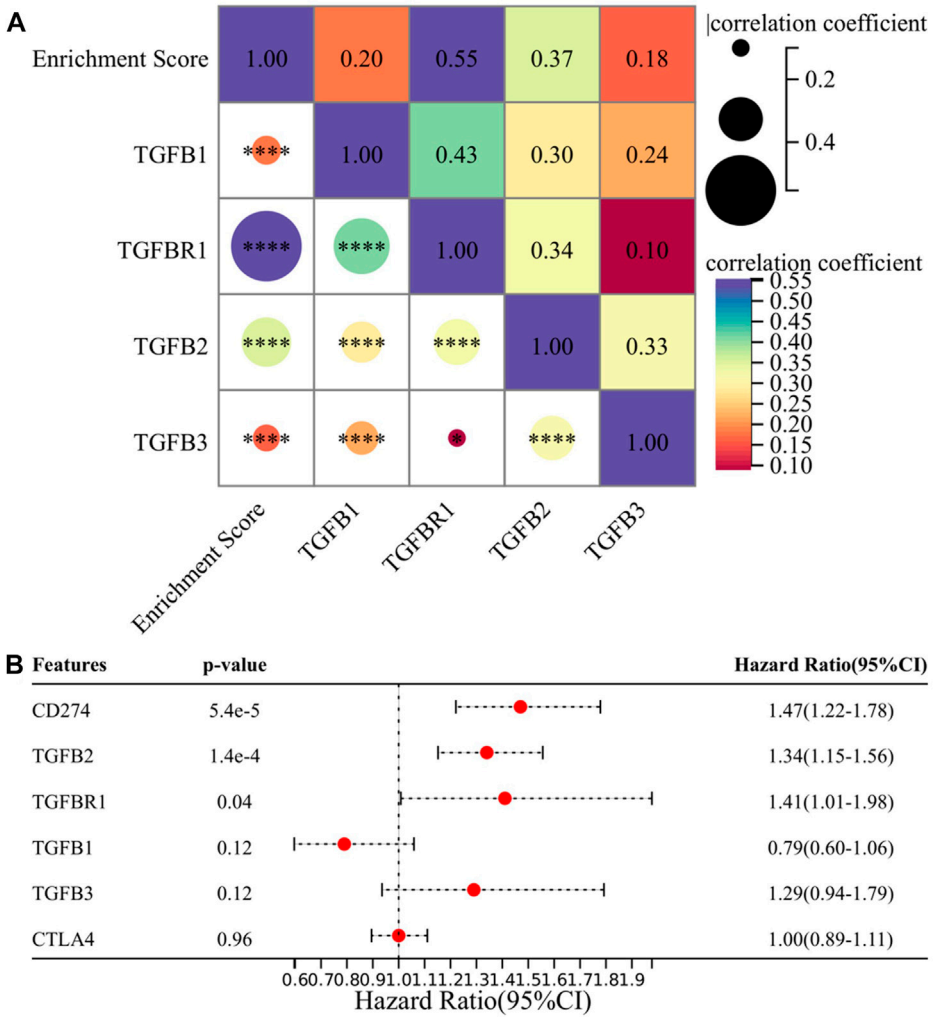


FIGURE 6
TGFB1/TGFBR1 may be a suitable target in the treatment of gliomas. **(A)** Correlation between TGFB1/TGFBR1, TGFB2, TGFB3, and enrichment score and **(B)** multivariate survival analysis of TGFB1/TGFBR1, TGFB2, TGFB3, PD-L1, and CTLA-4.

TGF-β signaling played an important role in inducing EMT phenotypes in gliomas (Figures 5K–O).

3.5 Compared with PD-L1 and CTLA-4, TGFB1/TGFBR1 is much associated with the immunosuppressive microenvironment

In this section, we explored specific molecules of the TGF-β signaling pathway, and we selected TGFB1, TGFBR1, TGFB2, and TGFB3 as research subjects because the drugs targeting these molecules are currently in the clinical research stage (Tauriello et al., 2022), and exploring them will make our research conclusions more likely to guide clinical practice. First, we analyzed the correlation between four candidate molecules and EMT enrichment scores, and the results showed that all four molecules were significantly correlated with EMT enrichment scores, among which TGFBR1 had the highest correlation ($p < 0.001$, r = 0.55) (Figure 6A). We then performed multivariate Cox regression analysis on four candidate molecules together with PD-L1 and CTLA-4, and the results suggested that PD-L1, TGFB2, and TGFBR1 were independent prognostic factors (Figure 6B). Based on the previous conclusions, we selected TGFB1/TGFBR1 for subsequent analysis.

Next, we focused on the correlation between TGFB1/TGFBR1, PD-L1, CTLA-4, and immune cell infiltration. With the increase in TGFB1/TGFBR1 expression, the infiltration of $CD4^+$ T naïve cells and $CD8^+$ T cells (p = 8.0e-5, r = -0.18; p =

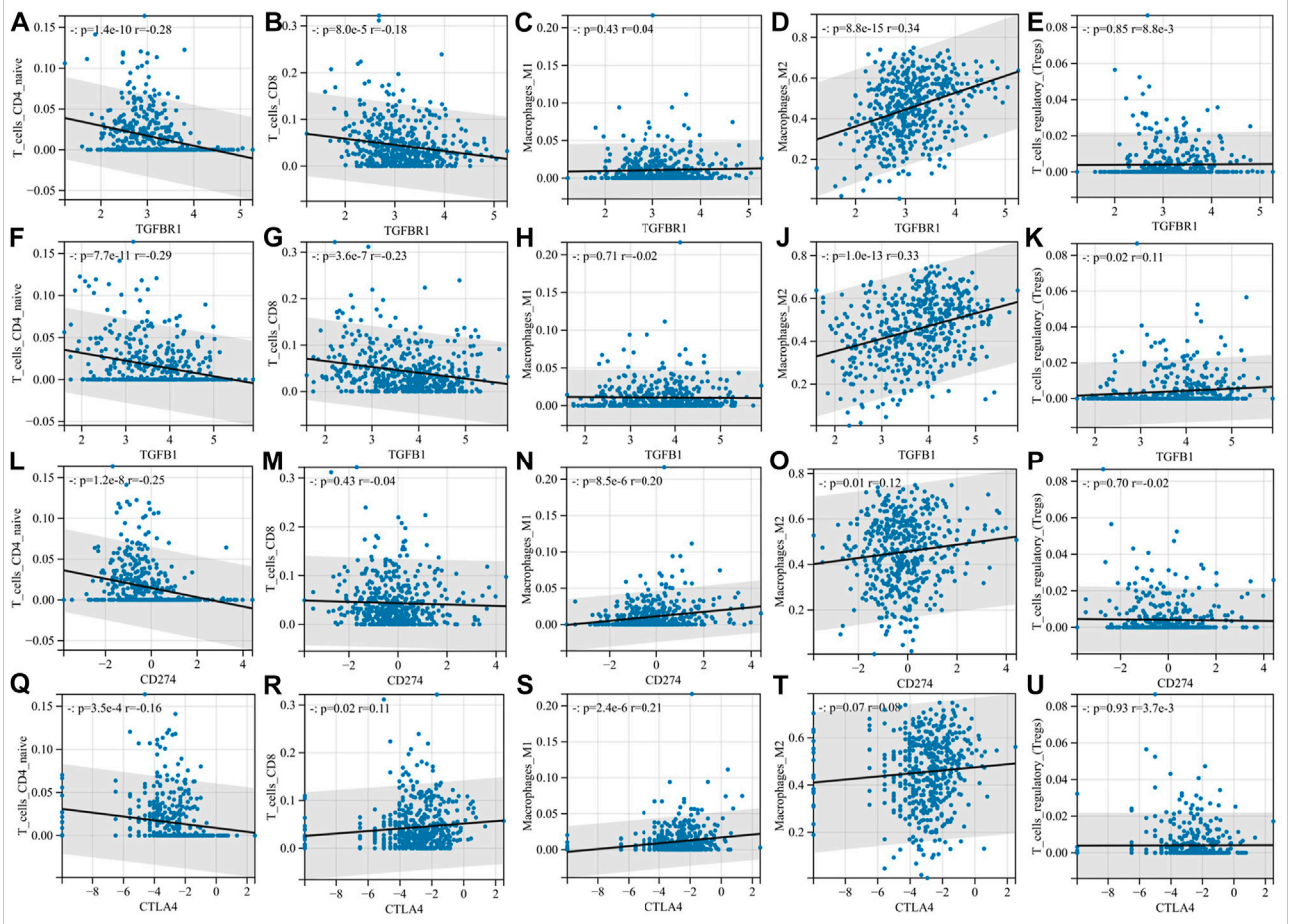


FIGURE 7
Compared with PD-L1 and CTLA-4, TGFB1/TGFBR1 is more associated with the immunosuppressive microenvironment. **(A–E)** Correlation of TGFBR1 with immune cell infiltration; **(F–K)** correlation of TGFB1 with immune cell infiltration; **(L–P)** correlation of PD-L1 with immune cell infiltration; and **(Q–U)** correlation of CTLA-4 with immune cell infiltration.

3.6e-7, r = -0.23) decreased significantly, while the infiltration of M2 macrophages (p = 8.8e-15, r = -0.34; p = 1.0e-13, r = 0.33) and Tregs increased significantly (Figures 7A–K). Although $CD4^+$ T naïve cell infiltration was decreased with PD-L1 expression, there was no significant change in $CD8^+$ T cells and Tregs. In addition, M1 macrophage infiltration was also increased (p = 8.5e-6, r = 0.20), while M2 macrophage infiltration was relatively low (p = 0.01, r = 0.12), suggesting that M0 macrophages were more likely polarized toward M1 macrophages (Figures 7L–P). As for CTLA-4, $CD4^+$ T naïve cell infiltration also decreased, while $CD8^+$ T cell (p = 0.02, r = 0.11) and M1 macrophage (p = 2.4e-6, r = 0.21) infiltration increased, and M2 macrophages and Tregs showed no significant changes (Figures 7Q–U). Compared with PD-L1 and CTLA-4, TGFB1/TGFBR1 was tightly associated with a decrease in $CD8^+$ T cells and an increase in M2 macrophages and Tregs infiltration, suggesting that TGFB1/TGFBR1 may have more powerful immunosuppressive properties.

3.6 Simultaneous blocking of TGFB1/TGFBR1 and PD-L1 might significantly improve survival

We calculated the optimal cutoff value for TGFBR1 and PD-L1 using the R package maxstat (maximally selected rank statistics with several p-value approximations, version: 0.7-25). We then grouped the samples according to cutoff values and performed survival analysis. Patients with higher PD-L1 and TGFBR1 expression had shorter survival (p = 1.4e-9, HR = 2.92; p = 1.0e-7, HR = 2.58) (Figures 8A–B). Patients with both high expression of TGFBR1 and PD-L1 had the worst prognosis, and those with low expression of both TGFBR1 and PD-L1 had the best prognosis, while those with high expression of one of the two molecules had a moderate prognosis (Figure 8C; Supplementary Table S1). The group with the best prognosis tended to have more $CD4^+$ T naïve cells and $CD8^+$ T-cell infiltration (although no statistically significant difference was achieved) and less

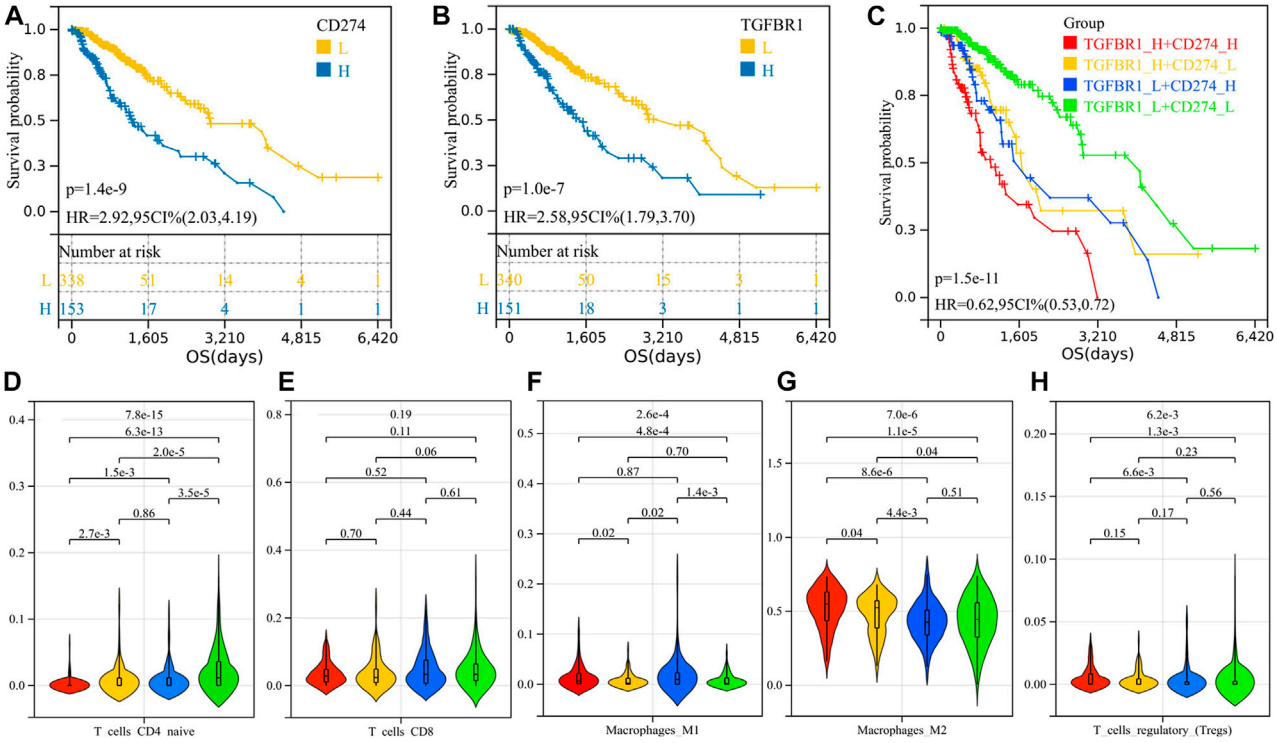


FIGURE 8
Survival analysis of PD-L1 and TGFBR1 and comparison of immune cell infiltration between groups. **(A)** K-M curve for PD-L1; **(B)** K-M curve for TGFBR1; **(C)** K-M curve for patients with different TGFBR1 and PD-L1 expression levels; and **(D–H)** immune cell infiltration between groups.

M2 macrophage and Tregs infiltration (Figures 8D–H). These results suggested that simultaneous blocking of TGFB1/TGFBR1 and PD-L1 is more likely to confer a survival benefit than blocking PD-L1 alone.

3.7 TGFB1 is closely related to microglia and macrophages in brain tissues

We analyzed the cellular localization of TGF-β-related molecules in brain tissue through Human Cell Landscape, a single-cell analysis website. In four single-cell samples of brain tissue, we found that TGFB1 and TGFBI (transforming growth factor-beta induced, a protein induced by TGFB1) were mainly expressed in microglia and macrophages, especially M2 macrophages. The results of the clustering analysis with marker genes are shown in Figures 9A–D and Table 3.

3.8 Gliomas may be one of the tumors most sensitive to the TGFB1 inhibitor

LY2109761 is an inhibitor of TGFB1 and is widely used in *in vitro* studies of various tumors. Through the Genomics of Drug Sensitivity in Cancer website and the EMTome website, we analyzed the susceptibility results of LY2109761 in various tumors. Figure 10A shows the IC_{50} values of LY2109761 in different LGG cell lines, and the IC_{50} values of LY2109761 in all available tumor cell lines are shown in Figure 10B. After standardizing the susceptibility data of all tumor cell lines, we found that LGG ranked sixth in drug sensitivity among 29 tumor species (Figure 10C). This suggested that TGFB1/TGFBR1 inhibitors may be sensitive in gliomas.

4 Discussion

Previous studies had demonstrated that the IDH1 mutation status has a significant effect on the immune microenvironment of diffuse glioma, and IDH1 wild-type patients generally have higher lymphocyte infiltration and PD-L1 expression (Yan et al., 2009; Berghoff et al., 2017). Also, the IDH1 mutation status has a significant correlation with prognosis (Sanson et al., 2009). Therefore, we can utilize these features of IDH1 to construct a risk model related to the inflammatory immune microenvironment and survival. Through GSEA, we found that the 3,919 immune-related gene sets with significant differences between the two groups were enriched in the

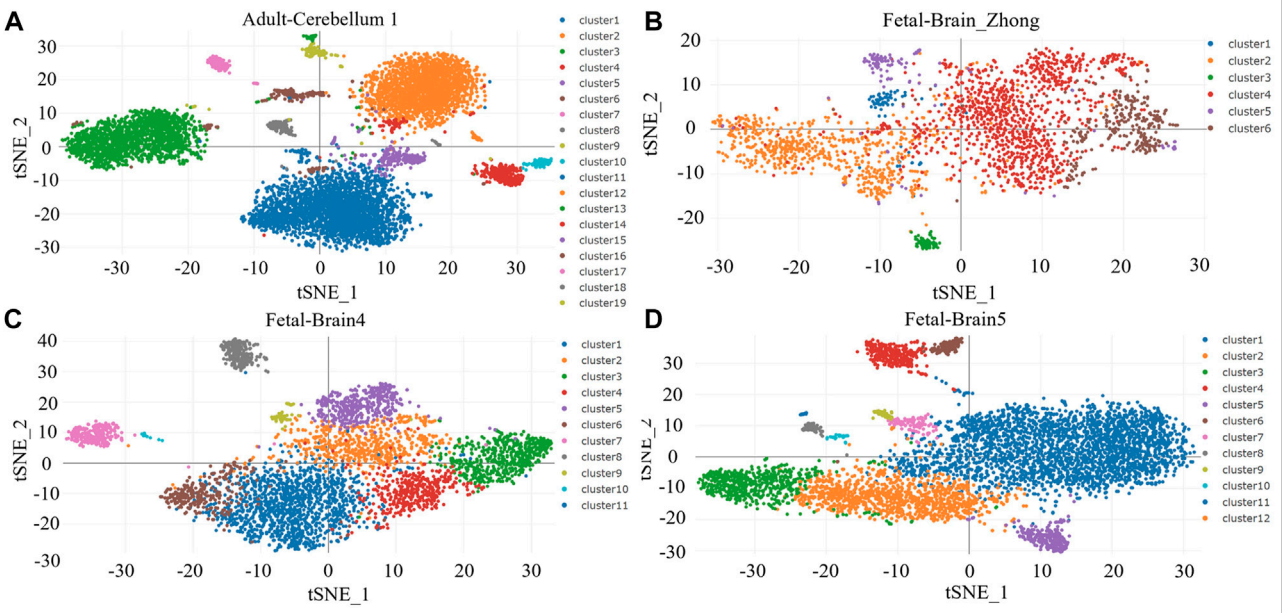


FIGURE 9
(A–D) Clustering analysis for brain tissues.

TABLE 3 TGFB1 and TGFBI are the top markers for microglia/macrophage in the brain.

Cluster	Annotation	Gene	*p*_value	avg_diff	pct.1	pct.2
Adult-Cerebellum1_Cluster16	Macrophage	TGFBI	5.82E-63	1.8518	0.367	0.011
Fetal-Brain_Zhong_Cluster3	Microglia	TGFB1	1.46E-48	1.1991	0.284	0.013
Fetal-Brain4_Cluster10	Macrophage	TGFBI	1.33E-15	0.8792	0.158	0.005
Fetal-Brain5_Cluster11	M2 macrophage	TGFBI	4.94E-30	1.6913	0.348	0.016
Fetal-Brain5_Cluster6	Fibroblast	TGFBI	6.48E-91	1.5493	0.254	0.011

Note: Single-cell differential gene expression analysis for four brain tissues was performed on the Human Cell Landscape website; $p < 0.05$ was considered statistically significant.

IDH1-wild group, suggesting that the IDH1-wild type is more related to immune response, which is consistent with the conclusions mentioned previously. According to the gene rank list of GSEA, top 50 genes in each group were selected for univariate analysis, and the results showed that these 100 genes had significant prognostic value. By performing LASSO Cox and multivariate Cox regression analyses on 100 genes, we constructed a two-gene risk model consisting of IGFBP2 and NOG.

IGFBP2 is a pleiotropic oncogene and plays a role in the occurrence and development of a variety of tumors (Brouwer-Visser and Huang, 2015). It has been confirmed that IGFBP2 can induce tumor epithelial–mesenchymal transformation and metastasis through the NF-κB signaling pathway (Gao et al., 2016). NOG is a natural inhibitor of bone morphogenetic protein (BMP), especially BMP2 and BMP4, which are the members of the TGF-β family. BMPs suppressed the tumorigenic function of human glioma-initiating cells by inducing cell differentiation, cell cycle arrest, and apoptosis (Bao et al., 2013). Several reports have shown that BMP4 is expressed in low-grade gliomas and that it serves as a favorable prognostic marker in gliomas (Bao et al., 2013; Nayak et al., 2020; Zhou et al., 2020). In addition, BMP4 was able to abolish cancer stem cell populations in human cancers, including malignant gliomas (Piccirillo et al., 2006; Piccirillo and Vescovi, 2006). Interestingly, as an inhibitor of BMPs, the expression of NOG and BMP2/4 was positively correlated. The high-risk group had lower levels of BMPs, consistent with the studies mentioned previously. We verified that the risk model has good prognostic value in TCGA cohort and two GEO cohorts and can further stratify the IDH1-mut and IDH1-wt groups. The multivariate Cox regression analyses further confirmed the independent prognostic value of the

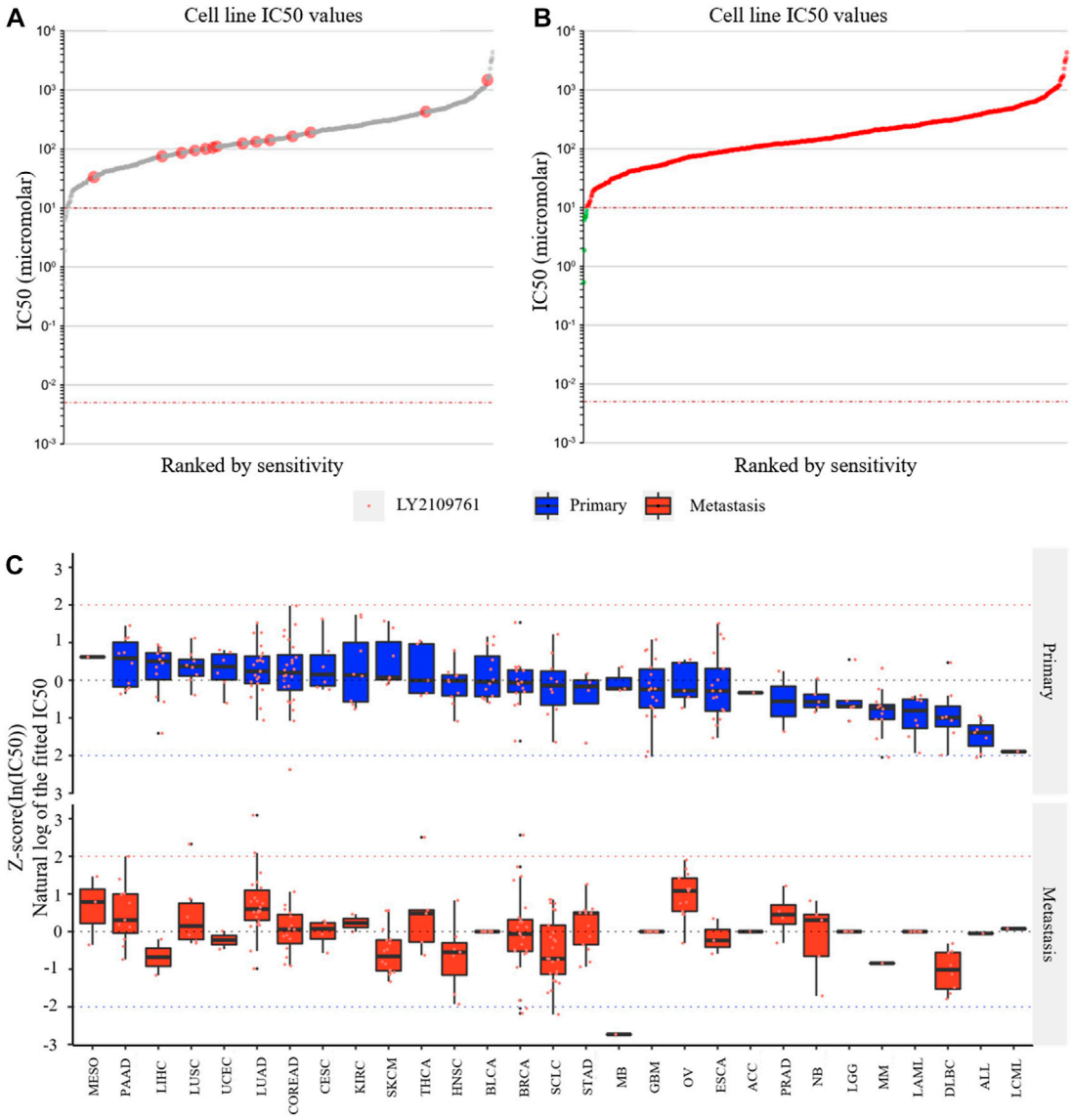


FIGURE 10
Drug susceptibility analysis of the TGFB1 inhibitor. **(A)** IC_{50} values of LY2109761 in different LGG cell lines; **(B)** IC_{50} values of LY2109761 in all available tumor cell lines; and **(C)** LGG ranked sixth in drug sensitivity among 29 tumor species.

risk model. Thus, this model might be used as a risk indicator in clinical practice.

To further investigate the relationship between the risk model and the immune microenvironment, we analyzed the correlation between immune infiltration scores, immune-related molecules, and immune cell infiltration with immune-related prognostic models. The stromal score, immune score, and ESTIMATE score were positively correlated with the risk score. The high-risk group had higher expression of immunosuppressive molecules and also had higher expression of immunostimulant molecules. Tregs were elevated in the high-risk group. However, B naïve cells, mast resting cells, CD4+ T memory resting cells, CD8+ T cells, neutrophils, and M0 and M1 macrophages were also elevated in the high-risk group. Similar conclusions had been reached in other people's studies. Berghoff et al. (2017) found significantly higher levels of PD-1-positive tumor-infiltrating lymphocytes and PD-L1 expression in IDH-wild-type gliomas than IDH-mutant gliomas. Liu et al. (2020) demonstrated higher CTLA-4 expression in higher-grade IDH-wild-type tumors than lower-grade IDH-mutant tumors. We would summarize these phenotypes as "hot" tumor microenvironments. "Hot" tumors and high expression of PD-L1 were known as hallmarks of sensitivity to immunotherapy (Xiong et al., 2018; Gao et al.,

2022), while inexplicably, the efficiency of immunotherapy was limited in gliomas, regardless of the IDH1 status and tumor grades (Blumenthal et al., 2016; Bouffet et al., 2016; Johanns et al., 2016; Lukas et al., 2018; Reardon et al., 2020). We hypothesized that the patient's inflammatory immune microenvironment is accompanied by other malignant phenotypes.

We then established a 32-gene functional protein association network for NOG and IGFBP2 on the STRING website. These 32 genes were analyzed for gene function enrichment using the KEGG database. The TGF-β signaling pathway, Hippo signaling pathway, PI3K-AKT–mTOR pathway, and signaling pathways regulating pluripotency of stem cells were found to be closely linked to the risk model. It had been reported that stem cell properties could be acquired by tumor cells through EMT. Induction of EMT in immortalized human mammary epithelial cells was sufficient to induce the expression of stem cell markers, enhance self-renewal, and increase the number of tumor-initiating cells (Mani et al., 2008; Shibue and Weinberg, 2017). TGF-β had been thought to be the most important factor inducing EMT *via* the classic Smad and non-Smad pathways (Lamouille et al., 2014; Hao et al., 2019). The Hippo pathway and PI3K-AKT–mTOR pathway had also been proved to be related to EMT (Cordenonsi et al., 2011; Song et al., 2014). It had been reported that there exists a strong correlation between EMT and immune activation. Further analysis demonstrated high expression of immune checkpoints and other druggable immune targets such as PD-1, PD-L1, CTLA-4, OX40L, and PD-L2 in patients with the EMT phenotype (Mak et al., 2016). Therefore, we speculated that the inflammatory immune microenvironment of gliomas is accompanied by the EMT phenotype.

In subsequent analyses, we found that EMT enrichment scores were significantly positively correlated with risk scores, M2 macrophage infiltration, Tregs, PD-L1, and CTLA-4 expression and negatively correlated with $CD8^+$ T-cell infiltration. Although Mak et al. (2016) discovered that the EMT phenotype is always accompanied with immune activation and higher expression of immune checkpoint molecules and declared that immunotherapy might have potential, the reality was far from that (10). The inflammatory immune microenvironment is accompanied by EMT, which in turn induces immunosuppression against the inflammatory immune microenvironment. Reversing the EMT phenotype might be necessary for immunotherapy treatment.

Dongre and Weinberg (2019) reviewed the main signaling pathways that induce EMT, including the TGF-β signaling pathway, the WNT signaling pathway, the Notch signaling pathway, the PI3K signaling pathway, and the RAS signaling pathway. Through the enrichment analysis of the aforementioned pathway, we found that the TGF-β signaling pathway is the most important. TGFB1/TGFBR1, TGFB2, and TGFB3 were elected for further investigation. Among these four molecules, TGFBR1 not only had the highest correlation with EMT enrichment scores but also was an independent prognostic factor relative to PD-L1. Immunocorrelation analysis showed that TGFB1/TGFBR1 had more powerful immunosuppressive properties than PD-L1 and CTLA-4, especially in inducing M2 macrophage infiltration and $CD8^+$ T-cell depletion. According to the different expressions of TGFBR1 and PD-L1, we performed survival analysis in groups, and the results showed that patients with both TGFBR1 and PD-L1 expression had obvious survival advantages, and the high expression of either molecule led to poor prognosis. This indicates that the combined inhibition of TGFB1/TGFBR1 and PD-1/PD-L1 has a good clinical application prospect.

Through single-cell analysis, we further determined that TGFB1 and TGFBI are mainly derived from microglia and M2 macrophages. As resident macrophages of the central nervous system (CNS), microglia are associated with diverse functions essential to the developing and adult brain during homeostasis and disease (Borst et al., 2021). Microglia-derived TAM (tumor-associated macrophages) increased angiogenesis and suppressed T-cell proliferation. Depletion of TAM provides survival advantages and delays recurrence when combined with standard-of-care treatment such as irradiation (Akkari et al., 2020). Numerous studies have demonstrated that gliomas are infiltrated by immune cells that make up to 30% of a tumor's mass (Nduom et al., 2015). The predominant population consists of glioma-associated microglia and macrophages, and their numbers inversely correlate with patients' survival (Gieryng et al., 2017). We speculated that by synthesizing and secreting TGFB1, microglia and M2 macrophages simultaneously induced EMT phenotype and immunosuppression.

Finally, we explored the relative drug sensitivity of the TGFB1 inhibitor in glioma cell lines through the Genomics of Drug Sensitivity in Cancer website and the EMTome website. Although these studies are *in vitro* experiments, the relative sensitivity between different tumor species can still give us some hints that gliomas have relatively good sensitivity relative to most tumors.

In this study, we constructed an immune-related prognostic model associated with the IDH1 mutation status. This model enables further risk stratification of patients with different IDH1 mutation states. By analyzing the immune microenvironment of patients with different risk scores, we found that high-risk patients were more likely to have an inflammatory immune microenvironment and a higher PD-L1 expression level, although clinical studies showed that patients with different IDH1 mutation states did not benefit from PD-1/PD-L1 inhibitors. We speculated that there were other malignant phenotypes that accompanied the inflammatory immune microenvironment, so we performed KEGG analysis on the risk model gene and found that it may be closely related to the EMT phenotype. This hypothesis was confirmed because

EMT-related gene sets were significantly enriched in the high-risk group. Subsequently, we found that the EMT phenotype was associated with a decrease in $CD8^+$ T cells and an increase in M2 macrophages, which is different from the risk model. By analyzing the main signaling pathways that induce the EMT phenotype, we found that TGF-β was the most important one in gliomas, and TGFB1/TGFBR1 showed stronger immunosuppressive properties than PD-L1 and CTLA-4, especially in inducing an increase in $CD8^+$ T cytopenia and M2 macrophages. It is clinically instructive that simultaneous low expression of TGFBR1 and PD-L1 has obvious survival advantages over other expression modes. Through single-cell analysis, we also found that TGFB1 is closely related to microglia and macrophages, especially M2 macrophages, which can explain why the increase in TGFB1/TGFBR1 expression is accompanied by a significant increase in M2 macrophages. Finally, we discussed the sensitivity of TGFB1 inhibitors in gliomas using cell line susceptibility data. From these analyses, we demonstrated a viable clinical strategy in combination with TGFB1/TGFBR1 inhibitors and PD-1/PD-L1 inhibitors for the treatment of high-risk gliomas.

Data availability statement

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

Author contributions

WL: conceptualization, methodology, software, investigation, formal analysis, and writing—original draft; QQ and RP: supervision and writing—review and editing; JJ and RP: data analysis.

Funding

This work was supported by the Basic and Applied Basic Research Fund of Guangdong Province (2022A1515012387).

Acknowledgments

The authors would like to thank the staff members of TCGA Research Network and the Gene Expression Omnibus (GEO) database and all the authors for making their valuable research data public.

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/fgene.2022.1070630/full#supplementary-material

References

Akkari, L., Bowman, R. L., Tessier, J., Klemm, F., Handgraaf, S. M., de Groot, M., et al. (2020). Dynamic changes in glioma macrophage populations after radiotherapy reveal CSF-1R inhibition as a strategy to overcome resistance. *Sci. Transl. Med.* 12 (552), eaaw7843. doi:10.1126/scitranslmed.aaw7843

Bao, Z., Zhang, C., Yan, W., Liu, Y., Li, M., Zhang, W., et al. (2013). BMP4, a strong better prognosis predictor, has a subtype preference and cell development association in gliomas. *J. Transl. Med.* 11, 100. doi:10.1186/1479-5876-11-100

Berghoff, A. S., Kiesel, B., Widhalm, G., Wilhelm, D., Rajky, O., Kurscheid, S., et al. (2017). Correlation of immune phenotype with IDH mutation in diffuse glioma. *Neuro. Oncol.* 19 (11), 1460–1468. doi:10.1093/neuonc/nox054

Blumenthal, D. T., Yalon, M., Vainer, G. W., Lossos, A., Yust, S., Tzach, L., et al. (2016). Pembrolizumab: First experience with recurrent primary central nervous system (CNS) tumors. *J. Neurooncol.* 129 (3), 453–460. doi:10.1007/s11060-016-2190-1

Borst, K., Dumas, A. A., and Prinz, M. (2021). Microglia: Immune and non-immune functions. *Immunity* 54 (10), 2194–2208. doi:10.1016/j.immuni.2021.09.014

Bouffet, E., Larouche, V., Campbell, B. B., Merico, D., de Borja, R., Aronson, M., et al. (2016). Immune checkpoint inhibition for hypermutant glioblastoma multiforme resulting from germline biallelic mismatch repair deficiency. *J. Clin. Oncol.* 34 (19), 2206–2211. doi:10.1200/JCO.2016.66.6552

Brouwer-Visser, J., and Huang, G. S. (2015). IGF2 signaling and regulation in cancer. *Cytokine Growth Factor Rev.* 26 (3), 371–377. doi:10.1016/j.cytogfr.2015.01.002

Cordenonsi, M., Zanconato, F., Azzolin, L., Forcato, M., Rosato, A., Frasson, C., et al. (2011). The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell* 147 (4), 759–772. doi:10.1016/j.cell.2011.09.048

Dongre, A., and Weinberg, R. A. (2019). New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* 20 (2), 69–84. doi:10.1038/s41580-018-0080-4

Draaisma, K., Wijnenga, M. M., Weenink, B., Gao, Y., Smid, M., Robe, P., et al. (2015). PI3 kinase mutations and mutational load as poor prognostic markers in diffuse glioma patients. *Acta Neuropathol. Commun.* 3, 88. doi:10.1186/s40478-015-0265-4

Gao, S., Sun, Y., Zhang, X., Hu, L., Liu, Y., Chua, C. Y., et al. (2016). IGFBP2 activates the NF-κB pathway to drive epithelial-mesenchymal transition and invasive character in pancreatic ductal adenocarcinoma. *Cancer Res.* 76 (22), 6543–6554. doi:10.1158/0008-5472.CAN-16-0438

Gao, W., Wang, X., Zhou, Y., Wang, X., and Yu, Y. (2022). Autophagy, ferroptosis, pyroptosis, and necroptosis in tumor immunotherapy. *Signal Transduct. Target. Ther.* 7 (1), 196. doi:10.1038/s41392-022-01046-3

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

Gleize, V., Alentorn, A., Connen de Kerillis, L., Labussiere, M., Nadaradjane, A. A., Mundwiller, E., et al. (2015). CIC inactivating mutations identify aggressive subset of 1p19q codeleted gliomas. *Ann. Neurol.* 78 (3), 355–374. doi:10.1002/ana.24443

Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., et al. (2020). Construction of a human cell landscape at single-cell level. *Nature* 581 (7808), 303–309. doi:10.1038/s41586-020-2157-4

Hao, Y., Baker, D., and Ten Dijke, P. (2019). TGF-beta-Mediated epithelial-mesenchymal transition and cancer metastasis. *Int. J. Mol. Sci.* 20 (11), 2767. doi:10.3390/ijms20112767

Johanns, T. M., Miller, C. A., Dorward, I. G., Tsien, C., Chang, E., Perry, A., et al. (2016). Immunogenomics of hypermutated glioblastoma: A patient with germline POLE deficiency treated with checkpoint blockade immunotherapy. *Cancer Discov.* 6 (11), 1230–1236. doi:10.1158/2159-8290.CD-16-0575

Johnson, W. E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8 (1), 118–127. doi:10.1093/biostatistics/kxj037

Kaminska, B., Ochocka, N., and Segit, P. (2021). Single-cell omics in dissecting immune microenvironment of malignant gliomas-challenges and perspectives. *Cells* 10 (9), 2264. doi:10.3390/cells10092264

Lamouille, S., Xu, J., and Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* 15 (3), 178–196. doi:10.1038/nrm3758

Liu, F., Huang, J., Liu, X., Cheng, Q., Luo, C., and Liu, Z. (2020). CTLA-4 correlates with immune and clinical characteristics of glioma. *Cancer Cell Int.* 20, 7. doi:10.1186/s12935-019-1085-6

Louis, D. N., Perry, A., Wesseling, P., Brat, D. J., Cree, I. A., Figarella-Branger, D., et al. (2021). The 2021 WHO classification of tumors of the central nervous system: A summary. *Neuro. Oncol.* 23 (8), 1231–1251. doi:10.1093/neuonc/noab106

Lukas, R. V., Rodon, J., Becker, K., Wong, E. T., Shih, K., Touat, M., et al. (2018). Clinical activity and safety of atezolizumab in patients with recurrent glioblastoma. *J. Neurooncol.* 140 (2), 317–328. doi:10.1007/s11060-018-2955-9

Mak, M. P., Tong, P., Diao, L., Cardnell, R. J., Gibbons, D. L., William, W. N., et al. (2016). A patient-derived, pan-cancer EMT signature identifies global molecular alterations and immune target enrichment following epithelial-to-mesenchymal transition. *Clin. Cancer Res.* 22 (3), 609–620. doi:10.1158/1078-0432.CCR-15-0876

Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133 (4), 704–715. doi:10.1016/j.cell.2008.03.027

Nayak, S., Mahenthiran, A., Yang, Y., McClendon, M., Mania-Farnell, B., James, C. D., et al. (2020). Bone morphogenetic protein 4 targeting glioma stem-like cells for malignant glioma treatment: Latest advances and implications for clinical application. *Cancers (Basel)* 12 (2), 516. doi:10.3390/cancers12020516

Nduom, E. K., Weller, M., and Heimberger, A. B. (2015). Immunosuppressive mechanisms in glioblastoma. *Neuro. Oncol.* 17 (7), vii9–vii14. doi:10.1093/neuonc/nov151

Piccirillo, S. G., Reynolds, B. A., Zanetti, N., Lamorte, G., Binda, E., Broggi, G., et al. (2006). Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature* 444 (7120), 761–765. doi:10.1038/nature05349

Piccirillo, S. G., and Vescovi, A. L. (2006). Bone morphogenetic proteins regulate tumorigenicity in human glioblastoma stem cells. *Ernst Scher. Found. Symp. Proc.* 5, 59–81. doi:10.1007/2789_2007_044

Reardon, D. A., Brandes, A. A., Omuro, A., Mulholland, P., Lim, M., Wick, A., et al. (2020). Effect of Nivolumab vs Bevacizumab in patients with recurrent glioblastoma: The CheckMate 143 phase 3 randomized clinical trial. *JAMA Oncol.* 6 (7), 1003–1010. doi:10.1001/jamaoncol.2020.1024

Reis, G. F., Pekmezci, M., Hansen, H. M., Rice, T., Marshall, R. E., Molinaro, A. M., et al. (2015). CDKN2A loss is associated with shortened overall survival in lower-grade (World Health Organization Grades II-III) astrocytomas. *J. Neuropathol. Exp. Neurol.* 74 (5), 442–452. doi:10.1097/NEN.0000000000000188

Sanson, M., Marie, Y., Paris, S., Idbaih, A., Laffaire, J., Ducray, F., et al. (2009). Isocitrate dehydrogenase 1 codon 132 mutation is an important prognostic biomarker in gliomas. *J. Clin. Oncol.* 27 (25), 4150–4154. doi:10.1200/JCO.2009.21.9832

Shen, W., Song, Z., Zhong, X., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 1 (3), e36. doi:10.1002/imt2.36

Shibue, T., and Weinberg, R. A. (2017). EMT, CSCs, and drug resistance: The mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* 14 (10), 611–629. doi:10.1038/nrclinonc.2017.44

Smith, J. S., Perry, A., Borell, T. J., Lee, H. K., O'Fallon, J., Hosek, S. M., et al. (2000). Alterations of chromosome arms 1p and 19q as predictors of survival in oligodendrogliomas, astrocytomas, and mixed oligoastrocytomas. *J. Clin. Oncol.* 18 (3), 636–645. doi:10.1200/JCO.2000.18.3.636

Song, Y., Luo, Q., Long, H., Hu, Z., Que, T., Zhang, X., et al. (2014). Alpha-enolase as a potential cancer prognostic marker promotes cell growth, migration, and invasion in glioma. *Mol. Cancer* 13, 65. doi:10.1186/1476-4598-13-65

Tauriello, D. V. F., Sancho, E., and Batlle, E. (2022). Overcoming TGFβ-mediated immune evasion in cancer. *Nat. Rev. Cancer* 22 (1), 25–44. doi:10.1038/s41568-021-00413-6

Vasaikar, S. V., Deshmukh, A. P., den Hollander, P., Addanki, S., Kuburich, N. A., Kudaravalli, S., et al. (2021). EMTome: A resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. *Br. J. Cancer* 124 (1), 259–269. doi:10.1038/s41416-020-01178-9

Xiong, Y., Wang, K., Zhou, H., Peng, L., You, W., and Fu, Z. (2018). Profiles of immune infiltration in colorectal cancer and their clinical significant: A gene expression-based study. *Cancer Med.* 7 (9), 4496–4508. doi:10.1002/cam4.1745

Yan, H., Parsons, D. W., Jin, G., McLendon, R., Rasheed, B. A., Yuan, W., et al. (2009). IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.* 360 (8), 765–773. doi:10.1056/NEJMoa0808710

Yang, K., Wu, Z., Zhang, H., Zhang, N., Wu, W., Wang, Z., et al. (2022). Glioma targeted therapy: Insight into future of molecular approaches. *Mol. Cancer* 21 (1), 39. doi:10.1186/s12943-022-01513-z

Zhou, Y., Liu, Y., Zhang, J., Yu, D., Li, A., Song, H., et al. (2020). Autocrine BMP4 signaling enhances tumor aggressiveness via promoting wnt/β-catenin signaling in IDH1-mutant gliomas. *Transl. Oncol.* 13 (2), 125–134. doi:10.1016/j.tranon.2019.10.019

Zhu, H., Hu, X., Gu, L., Jian, Z., Li, L., Hu, S., et al. (2021). TUBA1C is a prognostic marker in low-grade glioma and correlates with immune cell infiltration in the tumor microenvironment. *Front. Genet.* 12, 759953. doi:10.3389/fgene.2021.759953

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fgene.2022.1090640

OPEN ACCESS

EDITED BY
Chang Gu,
Tongji University, China

REVIEWED BY
Brian Piening,
Earle A. Chiles Research Institute, United States
Wenji Ma,
Columbia University, United States

*CORRESPONDENCE
Hongluan Mao,
hongluanmao@sdu.edu.cn

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 05 November 2022
ACCEPTED 27 December 2022
PUBLISHED 10 January 2023

CITATION
Hou Y, Qiao S, Li M, Han X, Wei X, Pang Y and Mao H (2023), The gene signature of tertiary lymphoid structures within ovarian cancer predicts the prognosis and immunotherapy benefit.
Front. Genet. 13:1090640.
doi: 10.3389/fgene.2022.1090640

The gene signature of tertiary lymphoid structures within ovarian cancer predicts the prognosis and immunotherapy benefit

Yue Hou[1,2,3,4], Sijing Qiao[1,2,3,4], Miao Li[1,2,3,4], Xue Han[1,2,3,4], Xuan Wei[1,2,3,4], Yingxin Pang[1,2,3,4] and Hongluan Mao[1,2,3,4]*

[1]Department of Obstetrics and Gynecology, Qilu Hospital of Shandong University, Jinan, Shandong, China, [2]Division of Gynecology Oncology, Qilu Hospital of Shandong University, Jinan, Shandong, China, [3]Key Laboratory of Gynecology Oncology of Shandong Province, Qilu Hospital of Shandong University, Jinan, Shandong, China, [4]Shandong Engineering Laboratory for Urogynecology, Qilu Hospital of Shandong University, Jinan, Shandong, China

Ovarian cancer (OC) has the lowest survival rate among gynecologic malignancies. Ectopic lymphocyte aggregates, namely tertiary lymphoid structures (TLSs), have been reported as positive biomarkers for tumor prognosis. However, the related gene signature of tertiary lymphoid structure in ovarian cancer was less understood. Therefore, this study first exhibited the organizational patterns of tertiary lymphoid structure by H&E staining and immunohistochemistry (IHC), and confirmed the improved survival values of tertiary lymphoid structure and quantified tumor-infiltrating lymphocytes ($CD20^+$ B cells and $CD8^+$ T cells) in ovarian cancer patients. Secondly, we collected the genes involved in tertiary lymphoid structure from databases. By the univariate regression analysis, the tertiary lymphoid structure gene signature (CETP, CCR7, SELL, LAMP3, CCL19, CXCL9, CXCL10, CXCL11, and CXCL13) with prognostic value, characteristically of ovarian cancer, was constructed in the TCGA dataset and validated in the GSE140082 dataset. Thirdly, by performing CIBERSORT and Tumor Immune Dysfunction and Exclusion (TIDE) analysis, we found that the high expression of this gene signature was positively correlated with developed immune infiltration and reduced immune escape. The improved IPS score and application in the IMvigor210 dataset received PD-L1 proved the predictive value of immunotherapy for this gene signature. Furthermore, this signature showed a better correlation between tumor mutation burden and classical checkpoint genes. In conclusion, Tertiary lymphoid structure plays important role in tumor immunity and the gene signature can be evaluated as a biomarker for predicting prognosis and guiding immunotherapy in ovarian cancer.

KEYWORDS

tertiary lymphoid structures, gene signature, prognosis, immunotherapy, ovarian cancer

1 Introduction

Ovarian cancer is just like a "secluded killer," menacing the health of the female reproductive system. The classical treatment regimens for ovarian cancer focus on tumor-reducing surgery and platinum-based chemotherapy (Marchetti et al., 2021). However, accounting for tardy diagnosis, extensive metastasis, recurrence, and resistance to chemotherapy drugs, the 5-year survival rate is less than 50% (Torre et al., 2018; Yang et al., 2018; An and Yang 2021). The poly-ADP-ribose polymerase (PARP) inhibitors (PARPi) and bevacizumab have been approved as first-line maintenance therapy. Despite

the prolonged progression-free survival, patients did not show a significant long-term survival benefit (Mirza et al., 2020; Shoji et al., 2022). Therefore, increasing treatment trials attempt to extend overall survival.

In recent years, immunotherapy has presented new opportunities. Studies of various cancers have confirmed that immunotherapy significantly improves the outcomes of patients. Emerging Immune checkpoint inhibitors (ICI) rejuvenate $CD8^+$ T cells by targeting PD-1, PD-L1, or CTLA4, which can deliver rapid and durable effects for patients with advanced tumors (Mahoney et al., 2015; Van Allen et al., 2015; Sharpe and Pauken 2018). As the treatment of ovarian cancer turns to immunotherapy, new challenges also emerge. The therapeutic of experimental drugs, such as CD274 antibody and CAR-T, are very limited. Ovarian cancer is defined as a "cold" tumor without marked lymphocytic infiltration, which indicates that failure to effectively stimulate the immunity to benefit from the treatment of immunosuppressive checkpoint inhibitors. Immunosuppression networks composed of myeloid-derived suppressor cells (MDSCs), Tregs, TAMs, cancer-associated fibroblasts (CAFs), and adipocytes may negatively affect the immunotherapy (Yang et al., 2020). The tumor immune microenvironment plays a key role in balancing immune escape and immune invasion. Herein, identifying reliable markers based on the relationship between tumor biology and TME characteristics is a crucial development for OC.

Tertiary lymphoid structures (TLS) are ectopic lymphoid organs stimulated by inflammatory signals (Dieu-Nosjean et al., 2014). The character of TLS diverges in different inflammatory atmospheres (Sautes-Fridman et al., 2019; Lutge et al., 2021). When the body is exposed to an infection or tissue damage, TLS invokes powerful immune responses by mobilizing lymphocytes to antigen sediment, while SLO grapples ineffectively (Aloisi and Pujol-Borrell 2006; Moyron-Quiroz et al., 2006). Conversely, TLS would locally generate auto-reactive T and B cells, accelerating the progression of autoimmune disease or incapacitating functions of grafts (Thaunat et al., 2010; Lucchesi and Bombardieri 2013). Increasing evidence substantiated that TLS favorably impacts the prognosis of cancer patients. The superior TLS density positively correlated with overall survival and disease-free survival has been observed in studies of oral squamous cell carcinoma, lung cancer, invasive breast cancer, and colorectal cancer (Gu-Trantien et al., 2013; Di Caro et al., 2014; Germain et al., 2014; Li et al., 2020). As for ovarian cancer, previous studies proposed that TLS coordinates the infiltration of T cells, such as the cytotoxic T cell and CXCL13-producing $CD4^+$ T cells, and antibody-producing PCs to advance antitumor responses (Kroeger et al., 2016; Yang et al., 2021; Ukita et al., 2022). However, the impact of TLS on gynecologic malignancies, especially ovarian cancer, demands a better understanding.

Similar to SLO, typical TLSs are organized by T cell zones and B cell zones containing germinal centers (GCs). Multitudinous techniques can be implemented to describe the landscape of TLS. HE staining is the earliest approach to distinguish TLS from morphological mirrors (Barmpoutis et al., 2021). Afterward, various studies clarified the highlights of TLS as a niche for T and B cells responding antagonistically to tumors by performing H&E staining, multi-immunohistochemistry (mIHC), gene expression analysis, and flow cytometry on considerable series of cancer samples (Silina et al., 2018). Nevertheless, there is no unified expert consensus on the presence and quantification of TLS. Subjective factors such as cell markers selected by investigators and working experience in morphological probes increased the experiment bias. Recent research on TLS progressed to the genetic sequencing level. In a study of malignant melanoma, the TLS signature containing nine genes was determined through differential analysis of transcriptome data of B cells and T cells (Cabrita et al., 2020). The gene signature reflected the predicted value of TLS in superior clinical outcomes and better response to ICI therapy. One review constructed TLS hallmark genes by summarizing the TLS features (Dieu-Nosjean et al., 2014). Another study reported a 12-chemokine signature associated with TLS (Lin et al., 2020). Considering the formation of TLS is a combination of lymphocytes and stromal cells under the action of a series of cytokines in the TME, the three gene sets could be generalized as the original collection of TLSs for subsequent analysis.

Therefore, our study was designed to define the appearance and prognostic significance of TLS. Then we established and examined the TLS gene signature of OC strongly correlated to prognosis. Our study discusssed the prognostic impact of the TLS gene signature in OC patients and further examined its predictive value for immunotherapy baed on the immune landscape. The relationship between the gene signature and representative biomarkers was also explored.

2 Materials and methods

2.1 Data collection and pre-processing

A retrospective study was performed on 60 ovarian cancer patients treated at Qilu Hospital of Shandong University (between 2014 and 2017). The Ethics Committee of Qilu Hospital of Shandong University has approved the collection of experimental specimens. Inclusion criteria were as follows: initial diagnosis of ovarian cancer, unaccepting of neoadjuvant chemotherapy, complete clinical data, and denial of other tumors. We interviewed the patient cohort until 1 June 2022. All patients' information was anonymized.

RNA-seq and clinical data of ovarian cancer (OC) patients were downloaded from The Cancer Genome Atlas (https://portal.gdc.cancer.gov/). After eliminating the expression profile data with a high degree of variation and incomplete clinical information, finally, 362 patients were selected. Another part of ovarian cancer patients (n = 380, GSE140082, GPL14951) from Gene Expression Omnibuswere (https://www.ncbi.nlm.nih.gov/geo) also were recruited for verification.

2.2 HE and immunohistochemistry protocol

The HE-stained sections were directly derived from the Pathology Department of Qilu Hospital of Shandong University. After being heated at a stationary temperature of 65° for 1.5 h, paraffin-embedded slides were dewaxed by xylene and immersed in decreasing ethanol concentrations for hydration. We applied heat-induced epitope retrieval for 15 min to expose the antigen. These slices were separately covered in hydrogen peroxide solution and goat serum for 30 min. Primary antibodies (CD20, 1:200, Rabbit mAb, Cell Signaling Technology, United States; CD3, 1:100, Rabbit mAb, Abcam, Cambridge, United Kingdom; CD8, 1:200, ab3516, Abcam, United Kingdom) were used to specifically bind to antigens under 4°C lasting for 12 h. And we applied immunopotentiator and secondary antibody modified by horseradish peroxidase to combine the primary

antibody. Samples stained by 3,3′-diaminobenzidine (DAB) and dehydrated under alcohol were continued to be counterstained by hematoxylin. Lastly, these slices were faxed by neutral gum. All images of sections were analyzed by three mature pathological professors. The organized dense lymphocyte aggregates in HE-stained samples were identified as tertiary lymphoid structures. Histology of $CD20^+$ B cells and $CD3^+$ T cells were evaluated to judge the maturity of TLS. The evaluation criteria are Semi-quantitative of lymphocytes: a. Proportion of lymphocytes in stroma: <5% = 0; 5%–25% = 1; 26%–50% = 2; 51%–75% = 3; 76%–100% = 4; b. The staining intensity: without any staining = 0; faint yellow = 1; claybank or brown = 2; The total scores = a+b. The cohort was divided into high and low groups according to average scores.

2.3 The establishment of TLS gene signature of ovarian cancer

There are 25 genes in the three published TLS-associated datasets, but only 21 genes were found in the TCGA-OC expression profile data. These 21 genes were used as a background set (Dieu-Nosjean et al., 2014; Cabrita et al., 2020; Lin et al., 2020). We implemented univariate cox regression analysis to screen the genes signature related to the prognosis of ovarian cancer patients in the TCGA database ($p < .05$). Simple sample Gene Set Enrichment Analysis (ssGSEA) is an algorithm to calculate the enrichment fraction of the pairing of each sample and gene set, whose most common application is to calculate immune infiltration (Barbie et al., 2009; Chen et al., 2022). Mimicking its computational process in immune infiltration, we take the established TLS gene set as the target gene set. The enrichment score of each sample was calculated through the R package "GSEA," and used as the score of the TLS gene signature. Then the study cohort (the TCGA dataset as train set, the GSE140082 dataset as test set) were divided into high and low expression groups based on median scores.

2.4 Analysis of CIBERSORT and ESTIMATE

CIBERSORT and ESTIMATE are the two most common algorithms for assessing immune infiltration. CIBERSORT is a tool for deconvolution of the expression matrix of human immune cell subtypes based on the principle of linear support vector regression (Chen et al., 2018). ESTIMATE is a method developed to evaluate tumor purity by using gene expression characteristics to infer the ratio of stromal cells to immune cells in tumor samples (Yoshihara et al., 2013). We discussed the functional lymphocytes among 22 kinds of immune cells by the R package "CIBERSORT" based on the TLS signature respectively (Newman et al., 2015). And the R package "ESTIMATE" were used to compare the differences in immune scores, stromal scores and ESTIMATE scores between high and low TLS signature groups (Yoshihara et al., 2013).

2.5 The analysis of TIDE scores and IPS scores

The developers of TIDE proposed two mechanisms of immune escape: 1) Dysfunction: T cells are inhibited by high levels of CTL; 2) Exclusion: suppression of T cell at low levels of CTL infiltration (Jiang et al., 2018). Higher TIDE score means a greater likelihood of immune escape and thus a poorer response to ICI. Data of OC patients were also downloaded from The Cancer Immunome Atlas (TCIA) (https://tcia.at/home). The IPS score (0–10 score) of each sample was standardized by the gene expression in four antibodies as immunotherapy (Gui et al., 2021). If the IPS score is higher, ICI is more effective.

2.6 The prediction of PD-L1 efficacy of the TLS signature

Most solid tumors can be classified into three immunological phenotypes: immune-inflamed, immune-excluded, or immune desert (Hegde et al., 2016). Cohered with the RECIST (1.1) criteria, the patients were assessed as responders with complete remission rate (CR), or partial remission rate (PR), and non-responders with stable disease (SD) or progressive disease (PD). We extracted the expression profile and clinical data from the IMvigor210 dataset, in which patients with locally advanced or metastatic urothelial carcinoma. The nine genes established before were applicated again to excogitate a risk score equation by the same route to compare the differences in risk scores and survival among immune subtypes and response subgroups (Mariathasan et al., 2018).

2.7 Mutation and check point genes analysis

The mutation frequency of TCGA-OC patients was calculated by the "maftools" package. These samples were divided into high and low TMB groups based on median of per/MB of mutation. And the oncoplot waterfall plot was analyzed to summarize the details of the single-nucleotide variant (SNV), including the top 10 genes, the base of the mutation site, and types of mutation. Copy number variations (CNV) of samples was also calculated. And we used PD-L1 and CTLA-4 as representatives of immune checkpoint genes to analyze the relationship with TLS gene characteristics.

2.8 Statistical analysis

Statistical analyses were conducted using R software (version4.0.3) and SPSS (23.0, IBM). The samples were grouped according to median. The relationship between TLS and clinical characteristics was assessed by chi-square test. Wilcoxon test was employed to analyze differences between two TLS signature groups. Differences in Overall Survival (OS) and Progression-Free Survival (PFS) between different groups were evaluated by Kaplan—Meier survival curve and verified by the log-rank test. The prognostic capability of the emergence of TLS and gene signature were evaluated by singular and multiple Cox regression analysis. A two-sided p-value of <.05 was considered statistically significant.

3 Results

3.1 TLSs are organized structures in ovarian cancer

Our study defined tertiary lymphoid structures (TLSs) as organized clusters of lymphocytes present in HE-stained tumor

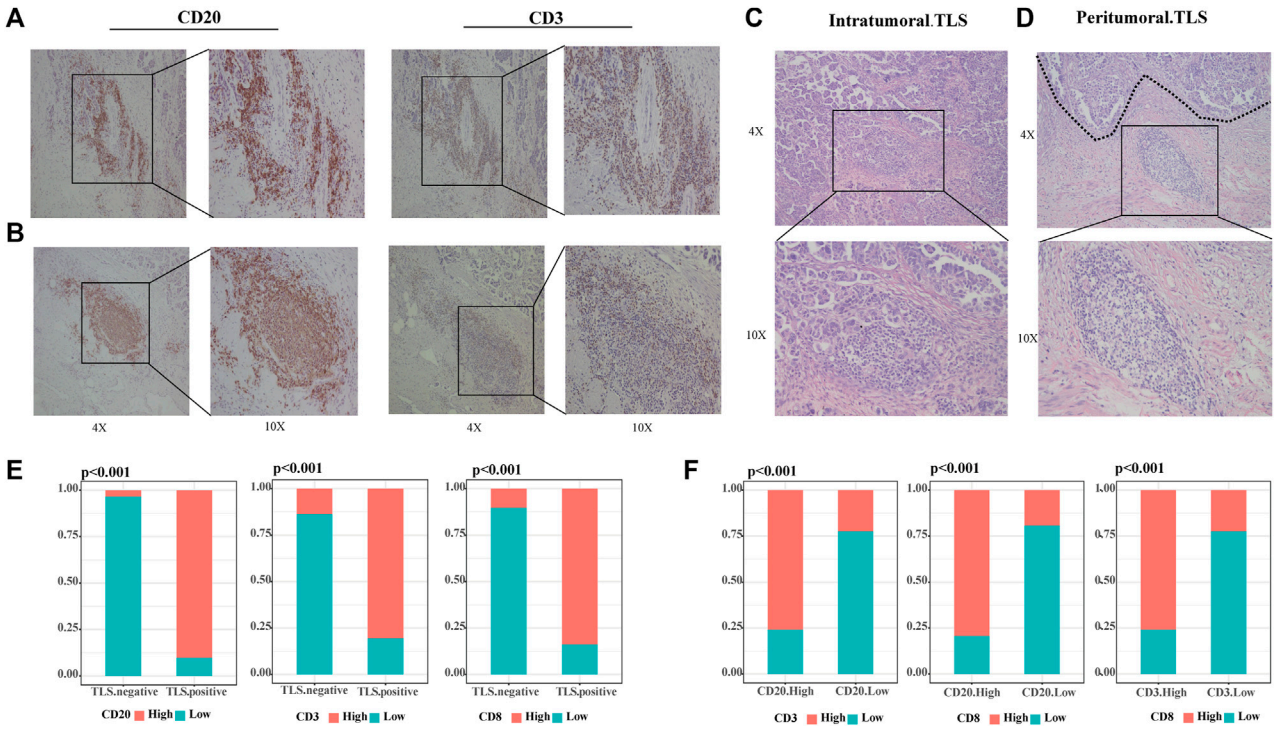


FIGURE 1
TLSs are dense clusters of lymphocytes in ovarian cancer. **(A)** Representative IHC images of immature TLS, scattered CD20$^+$ B cells without any follicular structure and CD3$^+$T cells. **(B)** Representative IHC images of mature TLS, CD20$^+$ B zone form a follicular structure surrounded by the CD3$^+$ T-cell zone. **(C)** Representative HE images of intratumoral TLS, immature TLS. **(D)** Representative HE images of peritumoral TLS, mature TLS. **(E)** Comparison of lymphocyte distribution between TLS-positive and negative groups, *p*-value was determined by the chi-square test. **(F)** Correlation between lymphocyte distribution, *p*-value was determined by the chi-square test. (Scale bar, 200 μm in "4X" pictures, 100 μm in "10X" pictures; $p < .05$).

specimens from OC patients (31/60 36.67%). We categorized TLS into two classes according to the number of CD20$^+$ B cells and CD3$^+$ T cells, and the formation of lymphoid follicles through IHC images. Scattered lymphocytes were found in TLS-negative specimens. The samples with small concentrations of B and T cells in the stroma are judged of immature TLS (12/31 38.71%) (Figure 1A). When CD20$^+$ B cells assemble and evolve into lymphoid follicles, surrounded by CD3$^+$ T cells marginal zone, mature TLS (19/31 61.29%) were settled down (Figure 1B). This phenomenon suggested that the maturation of TLS is a dynamic process. TLSs recruit lymphocytes to form dense ectopic tissue at the tumor boundaries. We observed that in TLS-positive samples, most mature TLS were located in the tumor margin (Figure 1C), while immature TLS were nearly shown to be intratumoral TLS (Figure 1D). Compared to TLS-negative cases, the high CD20$^+$ B cells, CD3$^+$ T cells, and CD8$^+$ T cells groups accounted for significantly more in those with TLS (Figure 1E, all $p < .001$). Further, the infiltration of CD20$^+$ B cells, CD3$^+$ T cells, and CD8$^+$ T cells were strongly correlated with each other (Figure 1F, all $p < .001$).

3.2 The existence of TLSs represents a superior survival outcome

The correlation between TLS and clinical factors is shown in Supplementary Table S1. These variables were included in the univariate regression analysis through the log-rank test. Stage I/II and TLS-positive exhibited a favorable trend towards advanced 5-year OS. Furthermore, TLS-positive plays a role of independence in multivariate analysis (Supplementary Table S2). Correspondingly, TLS-positive also is an independent factor for 5-year PFS sequentially by univariate and multivariate analyses (Supplementary Table S3). As portrayed in Figures 2A, B, the TLS-positive group performed the superior 5-year overall survival and progression-free survival (5-year-OS: $p = .0078$; 5-year-PFS: $p = .041$). However, the survival curves showed no significant difference in 5-year OS and 5-year PFS between immature and mature TLS groups (Figures 2C, D, 5-year-OS: $p = .54$; 5-year-PFS: $p = .6$). And the better prognosis of high expression of CD20$^+$ B cells and CD8$^+$ T cells groups than low expression groups interpreted that TLS achieves an anti-tumor immune response through the synergistic performance of critical immune cells (Figures 2E, F).

3.3 The correlation of TLS signature with OS in OV patients

Our study established the gene signature (CETP, CCR7, SELL, LAMP3, CCL19, CXCL9, CXCL10, CXCL11, and CXCL13) of TLS associated with the prognosis of ovarian cancer patients after including 21 genes in univariate regression analysis (Supplementary Figure S1, Figure 3A). Based on the matched scores of the genes set, OC patients were separated into TLS signature high (above 50%) and low queues (below 50%). As exhibited in Figure 3B, the high

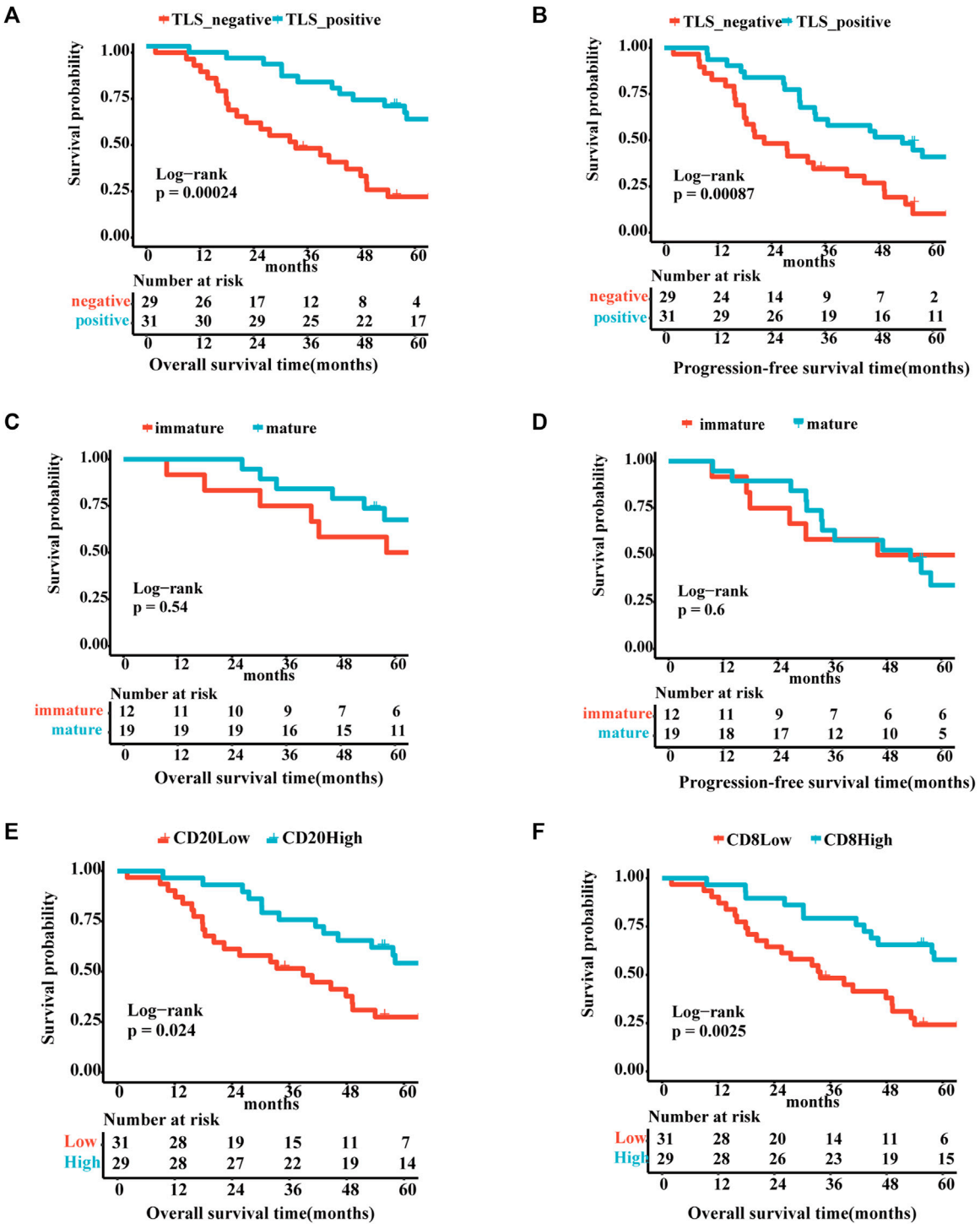


FIGURE 2
The existence of TLSs represents superior survival outcome Survival analysis of OV patients between the positive and negative TLS subgroups **(A,B)**: **(A)** The 5-year overall survival. **(B)** The 5-year progression-free survival. Survival analysis of OV patients between the immature and mature TLS subgroups: **(C)** The 5-year overall survival. **(D)** The 5-year progression-free survival. Survival analysis of OV patients between the high and low TIL subgroups **(E,F)**: **(E)** The 5-year overall survival of CD20. **(F)** The 5-year overall survival of CD8. ($p < .05$).

expression group presented a superior trend towards improved OS ($p = .00044$). Univariate and multivariate regression analyses were evaluated by comparing the effect of the gene feature with other parameters on the overall survival of patients. Age, the recurrence and progression of the tumor, and TLS gene signature were screened as a statistically prognostic factor in univariate analysis (Figure 3D). Furthermore, the TLS signature was associated with a significantly better OS in multivariate analyses (Figure 3F). Considering the lack of

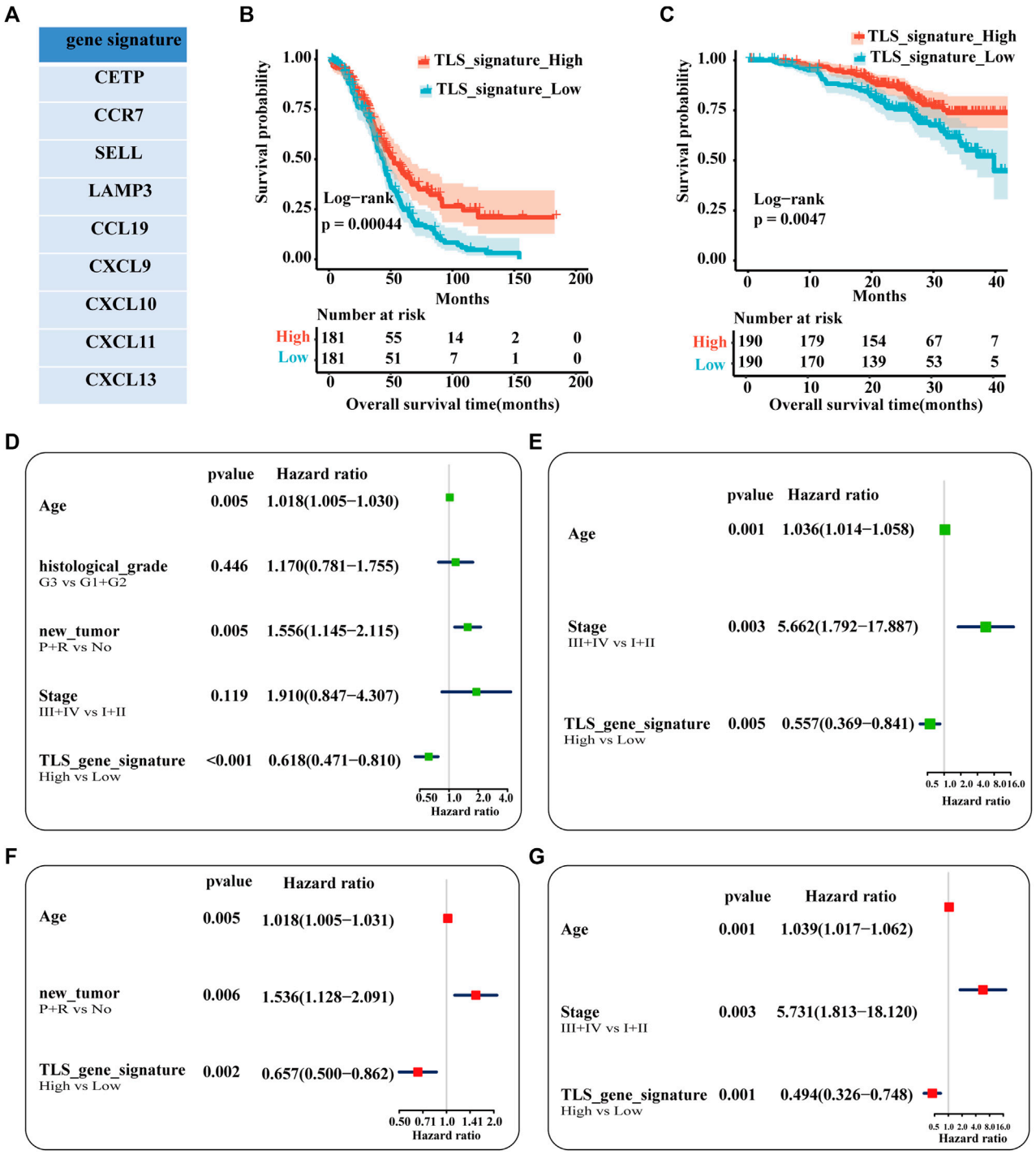


FIGURE 3
Analysis of the independence of TLS gene signature on prognosis. **(A)** The nine genes associated with better prognosis in TCGA-OV dataset. Survival analysis of OV patients between the high and low TLS signature subgroups **(B,C)**: **(B)** The TCGA cohort as training set. **(B)** The GSE queue as validation set. Univariable analysis **(D)** and multivariable analysis **(F)** of the overall survival in TCGA-OV cohort. Univariable analysis **(E)** and multivariable analysis **(G)** of the overall survival in GSE140082 cohort. (tumor status: No = without-tumor; P = progression; R = recurrence, $p < .05$).

significance of tumor stage in the prognosis of OC patients in the TCGA-OC database, we used the GSE140082 database as the validation cohort. Likewise, the TLS gene signature was also an independent predictor of OS (Figures 3C, E, G). Besides, the prognostic significance of this TLS gene signature in PFS is illustrated in Supplementary Figure S2.

3.4 The relationship between TLS signature and immune cells infiltration

As displayed in Figure 4A, the distribution of 22 kinds of immune cells differs with the TLS signature subgroups. The degree of $CD8^+$ T cells, activated and resting memory $CD4^+$ T cells, regulatory T cells

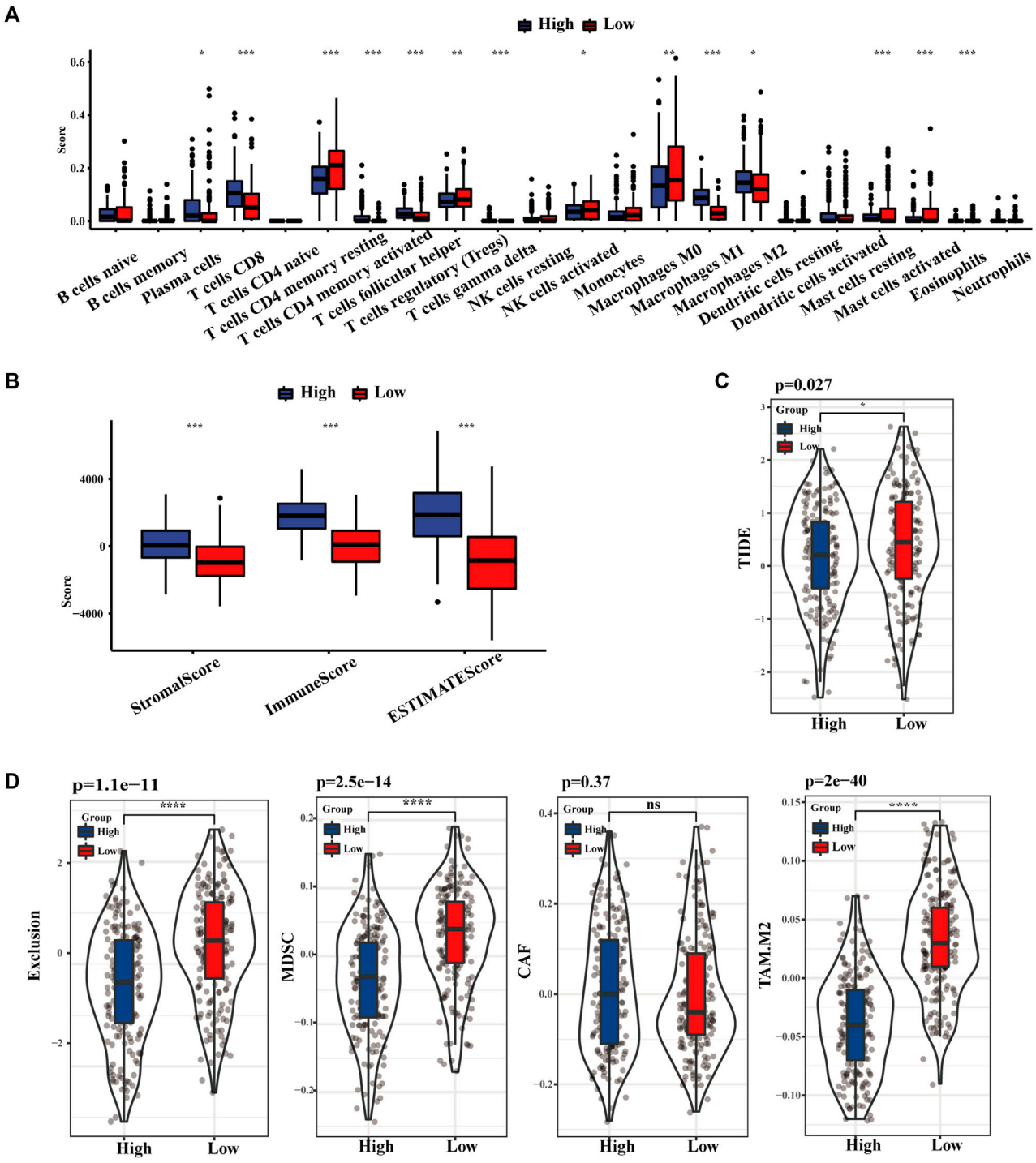


FIGURE 4
The landscape of Immune infiltration based on TLS signature. **(A)** The distribution of 28 immune cells between high and low TLS signature groups. **(B)** Box plots of TLS signature score in stromal, immune, and estimate score. **(C)** Violin plot of the differences in calculated TIDE scores between high and low TLS signature groups. **(D)** Differences in immune escape mechanisms and immunosuppressive cells between the two groups of TIDE. (NS: no significance, $*p < .05$, $**p < .01$, and $***p < .001$).

(Tregs), and M1 phenotype macrophages in the high TLS expression group was statistically higher than in the low group. Similarly, the TLS signature high group obtained distinctly higher immune score, stromal score, and estimate score, which indicated that gene expression levels associated with TLS prominently affect the infiltration situation of immune cells and stromal cells (Figure 4B). Correspondingly, the TIDE score in the low-risk group was notably higher than in the lowgroups, indicating the existence of tumor immune evasion in the low-risk group (Figure 4C). We further discovered that the immunosuppressive cell expression models, namely T cell exclusion, dominated the immune escape of the low TLS signature group. Meanwhile, the degree of IFNG, CD274, and

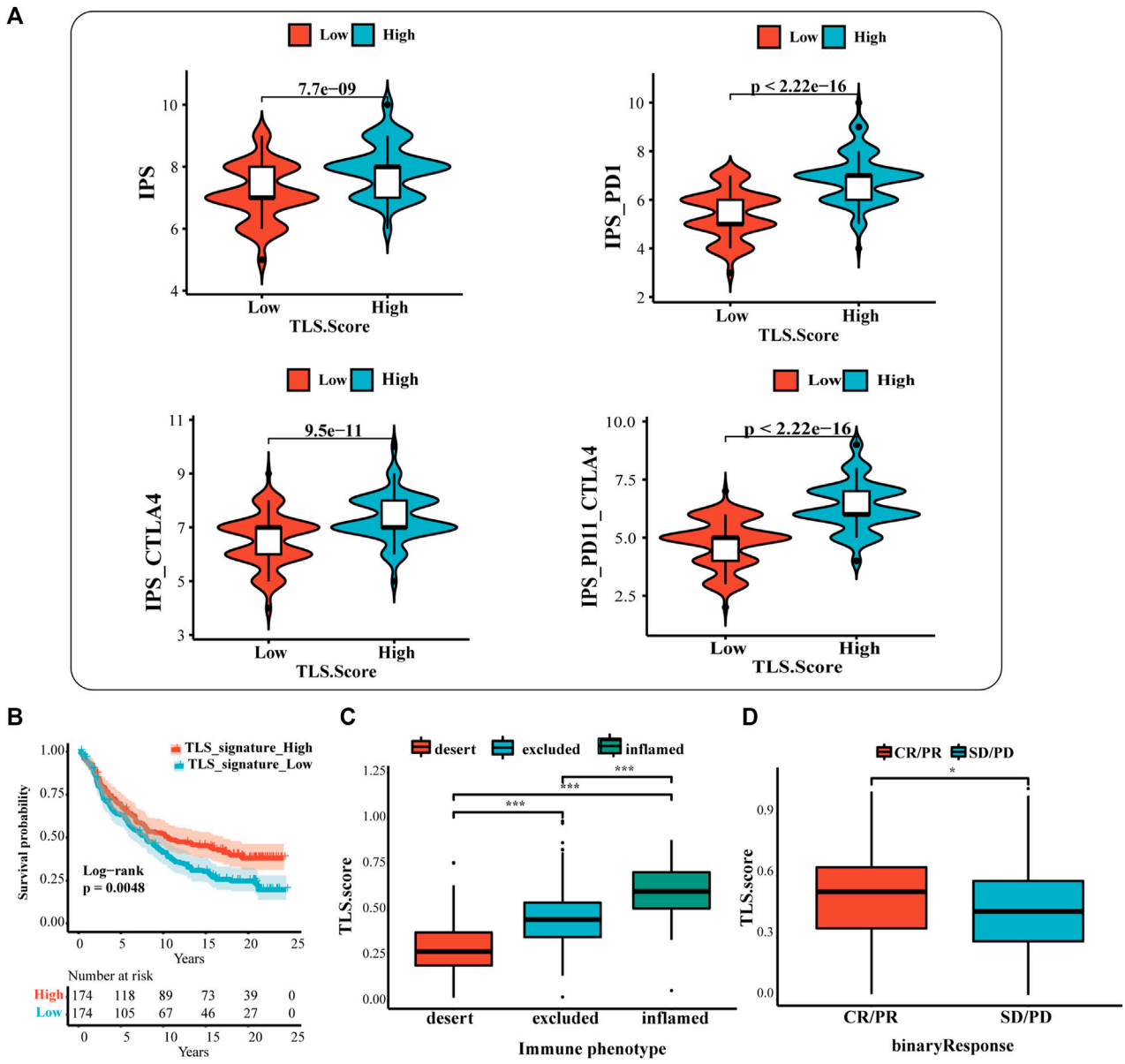


FIGURE 5
The prognostic value of TLS features in immunotherapy response. **(A)** The association between IPS analysis and TLS signature score. The analysis of TLS signature in the IMvigor210 cohort **(B,D)**: **(B)** The survival analysis between the high and low TLS signature subgroups; **(C)** Box plots of TLS signature score in the desert, excluded and inflamed immune phenotypes; **(D)** Distribution of ISS in groups with different anti-PD-L1 clinical response statuses. (NS: no significance, $*p < .05$, $**p < .01$, and $***p < .001$).

CD8 in the high TLS signature group was significantly higher than in the low TLS signature group (Supplementary Figure S3). Conclusively, our study indicated the high TLS signature group as the cohort most potentially benefiting from immunotherapy (Figure 4D).

3.5 The predictive value of TLS features in immunotherapy response

The results showed that the IPS, IPS-CTLA4, IPS-PD1, and IPS-PD1-CTLA4 scores were higher in the high TLS signature group (Figure 5A). Additionally, the difference reinforces the better response of the high-expression group to immunotherapy. Based on the above results, we applied the TLS signature to the urothelial carcinoma patients who received the treatment of PD-L1 inhibitors (atezolizumab) to inspect the estimation of TLS signature in ICI. The results displayed that the high TLS signature exhibited the prognostic value (Figure 5B). Meanwhile, the prediction of the TLS signature in immunological phenotype was investigated. In this cohort, the TLS signature score of immune-inflamed phenotype was more distinguished than that of the immune-desert or the immune-excluded phenotype (Figure 5C). Then, we further discussed the differences in PDL1 immune checkpoint inhibitor benefits between high- and low TLS signature groups. It turned out that patients in the high group performed a higher complete response (CR)and partial response (PR) rate (Figure 5D).

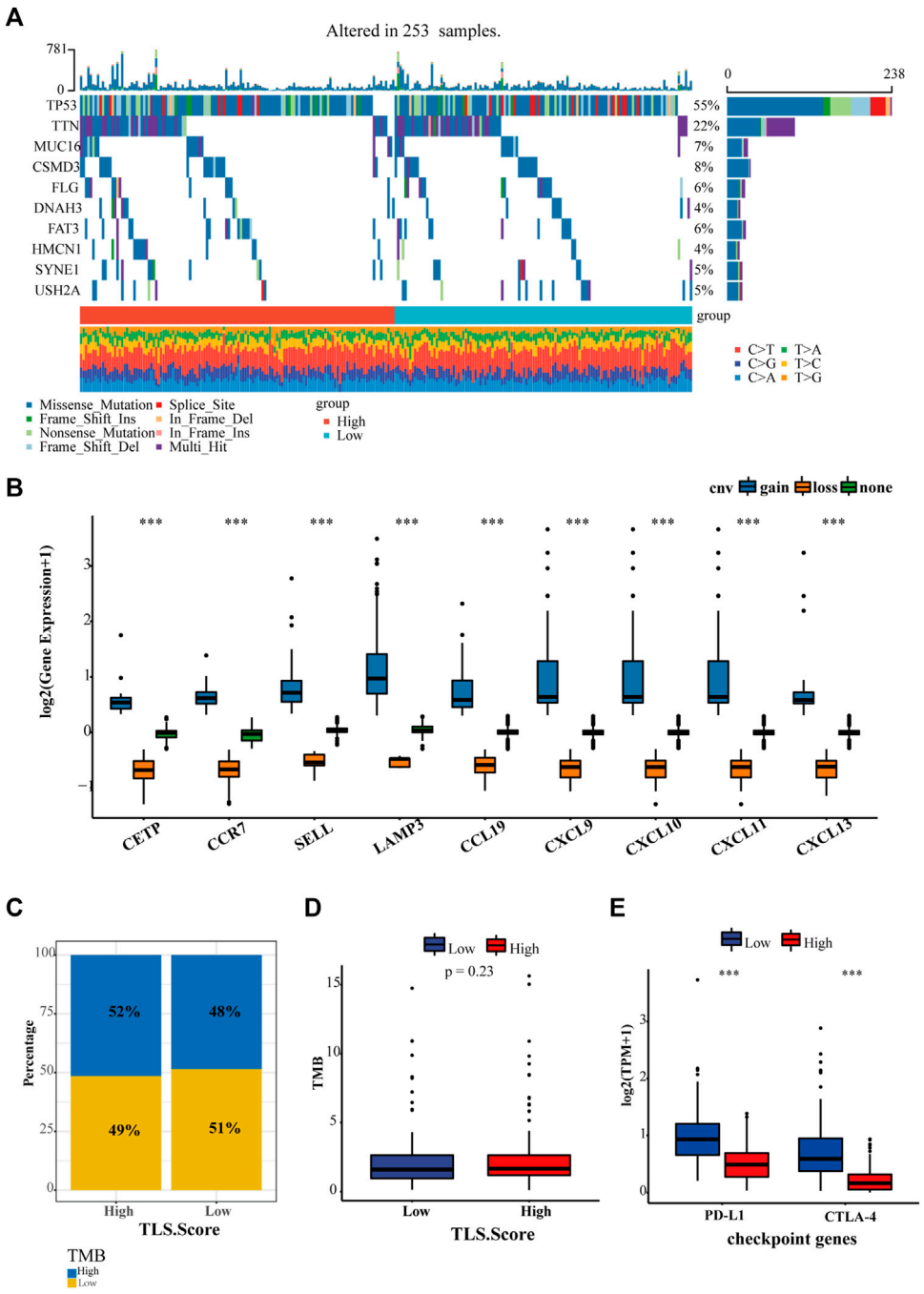


FIGURE 6
The relationship between TLS gene signature and existing markers of immunotherapy. **(A)** The landscape of genetic variation of the TLS gene signature in the OV cohort. **(B)** The CNV variation frequency of the TLS gene signature in the OV cohort. The proportion **(C)** and frequency **(D)** of tumor mutation load between high and low TLS gene signature groups. **(E)** The distinction of PD-L1 and CTLA-4 between high and low TLS gene signature groups. (*$p < .05$, **$p < .01$, and ***$p < .001$).

3.6 The relationship between TLS gene signature and existing markers of immunotherapy

Except for tumor-infiltrating lymphocytes, tumor mutational burden (TMB) and immune checkpoints, such as PD-1, PD-L1, and CTLA4, have been developed to predict immunotherapy. Hence our study proceeded to estimate the diversity between the landscape of tumor mutation and typical checkpoint genes between high and low TLS signature groups. In total, 235 samples were genetically altered. The top 10 genes with mutation frequency between TLS high group and the low group were generalized in Figure 6A. The most common variant classification is a missense mutation. SNP ranks first in variant type, and C>A is the most frequent SNV class (Figure 6A, Supplementary Figure S4). In addition, we found that the gene signature was characterized by a

precise copy number amplification (Figure 6B). The high TLS signature group occupied an increased proportion of high-TMB than the low group, despite the statistically significant difference in TMB between the two groups being not shown (Figures 6C, D). Moreover, the expression of two representative checkpoint genes, PD-L1 and CTLA4, significantly increased in the high TLS signature group (Figure 6E).

4 Discussion

Considering that the effect of immunotherapy on ovarian cancer is restricted, exploring biomarkers or targets to filtrate ideal beneficiaries is a future development. The three objectives of this study are as follows. First and foremost was to assess the presence of TLS and its prognostic impact on ovarian cancer. Identification of an independent signature of TLS associated with prognosis is the second goal. The third aspect is to explore the immune landscape and predictiveness of immunotherapy of this gene signature.

Recently the remarkable manifestation of tertiary lymphatic structures in the tumor microenvironment has begun to be noticed. Various studies on neoplasms have confirmed the prognostic value of TLS (Cabrita et al., 2020; Vella et al., 2021). Consistent with their results, our study revealed that TLS is strongly associated with prolonged OS and PFS. Beyond that, the higher density of $CD20^+$ B cells, and $CD8^+$ T cells play an excellent role in advanced 5 years overall survival. These data suggest that TLSs exert efficient antitumor response by recruiting activated B cells and T cells in TME. Tumor-infiltrating lymphocytes accumulate orderly in ectopic organs to form TLS, but the specific antitumor immune reactivity remains obscure. As the most abundant cells in TLS, B cells progressively enrich into GC as TLS develops. Moreover, B cells within GC tend to mature into the IgG and IgA-producing plasma cells (PCs), which can migrate along the fibroblast track to the nest of tumor (Meylan et al., 2022). Driven by memory B cells and a range of cytokines (including TNF, IL-2, IL-6, and IFNγ), T cells home to the tumor bed and release perforin and granzyme or Fas/FasL pathways to destroy tumor cells. However, the single presence of TIL was suboptimal. Researches confirmed that tumor-infiltrating CD8+T cells showed their cytotoxic capacity function only when B cells were enhanced within the tumor (Kroeger et al., 2016). The simultaneous occurrence of intratumoral $CD8^+$ T cells and $CD20^+$ B cells was also found to be independently associated with improved survival (Wouters and Nelson 2018; Cabrita et al., 2020). Noticeably, the maturation stage of TLS could significantly affect the prognosis in theory, but this study found no significant imparity between immature and mature TLS groups. Finite-sized and single-center samples of this study may contribute to the dispute. However, these results showed that the tumor immunology advantage of TLS is that it increases the opportunity of binding tumor antigen by the recruitment of TIL, while serving as a collection point, strengthening the interaction between TIL and cytokines. For immune-deficient ovarian cancer, TLS brings positive forces for increasing sensitivity to ICI. TLS-related drugs or vaccines can improve the effect of ICI under development at OC.

To date, there is no unified expert consensus on the identification and quantification of TLSs. Consistent with previous studies, this study also conducted H&E staining and immunohistochemistry (IHC) with selected markers to detect TLS, while these measures are inconvenient and subjective. In the context that TLS is a combination of lymphocytes and cytokines, this study innovatively selected and constructed the TLS signature related to prognosis in ovarian cancer from background genes. Satisfactorily, ovarian cancer patients in public datasets with high TLS signature performed a preferable survival outcome. The established genetic traits include CETP, CCR7, SELL, LAMP3, CCL19, CXCL9, CXCL10, CXCL11, and CXCL13. CXCL13 acts as a B lymphocyte chemoattractant in TLS, mediating the recruitment of B cells and promoting T/B separation and the formation of a germinal center (GC) within TLS (Rodriguez et al., 2021). The consequence of CXCL13 in ovarian cancer is gradually understood. A large study of high-grade serous ovarian cancer found CXCL13 enhanced immune efficacy in combination with PDL1 by promoting the expansion and activation of CXCR5 + CD8 + T cells. We also observed that the CXCL13-producing CD4 + T cells could facilitate the formation of TLS and strengthen the cooperative antitumor activity of cellular and humoral immunity in ovarian cancer (Ukita et al., 2022). CXCL9, CXCL10. CXCL11/CXCR3 axis contains the four genes of this gene signature. The excellent performance of the axis in TME reflects in the migration, differentiation, and activation of immune cells. In *in vitro* experiments of mice, as the downregulation or inhibition of the expression of the CXCR3 axis, the migration of Th1, NK, and CTL cells were all significantly decreased, and the survival outcome also worsened (Tokunaga et al., 2018; House et al., 2020). In high density of LAMP3-DC-mature-type TLS, CXCL9, 10 and 11 are involved in the migration of Th1 cells by binding to the CXCR3 receptor on tumor-infiltrating T cells (Sautes-Fridman et al., 2019). A sequencing analysis for 1,310 breast cancer patients demonstrated the relevance of CXCL10 and HRD, identifying CXCL10 as biomarker for anti-PD-1/PD-L1 therapy (Shi et al., 2021). Paradoxically, this axis induces the onset of immunosuppression by attracting Treg migration to the focal site. In the context of the contribution of HRD in ovarian cancer, CXCL10 provides a neoteric perspective on markers of immunotherapy. Improved survival, a preferable anti-PD-L1 therapy caused by the overexpression of CXCL9 in ovarian cancer, determined it as a stable predictive target (Seitz et al., 2022). The role of CETP, CCR7, and SELL in ovarian cancer is less studied. The chemokine ligand CCR7 and the adhesion molecule CD62L (expressed by the SELL, L-selectin) act as lymph node homing receptors that regulate T cells' activation and migration patterns. The loss of KLF2 resulted from the activation of PI3K and mTOR and subsequently prevented the target genes, such as CD62L, and the reduction of CCR7 (Sinclair et al., 2008). L-selectin shedding and transcriptional shutdown allow terminal differentiation of naive T cells to effector memory T cells (TEM), while TEM flow from secondary lymphoid organs to peripheral tissues for repositioning (Ivetic et al., 2019). And the downregulation of KLF 2 and its target gene S1PR1 antagonistically prompted the upregulation of the type-C lectin CD69. Meanwhile, the locally produced TGF-βis all involved in the expression of CD103, leading to the retention of T cells in the tissue, namely the generation of tissue-resident T cells (TRM). Workel et al. (2019) found that TGF-βprompted $CD8CD103^+$ T cells to secrete CXCL13, which recruits B cells in TME and is essential for the formation of TLS. Therefore, we hypothesize that there is a strong association between the expression of SELL and CCR7, TRM, and TLS, and exploring their interactions may provide new ideas for tumor treatment. For example, a new immunotherapy concern is restoring control of selectins to modulate tumor immune expansion. In a mouse model of adoptive T cell cancer immunotherapy, overexpression of L-selectin improved the control of tumor cells by promoting infiltration and proliferation of T cells (Watson et al., 2019). Retrospecting the existing achievement, this gene signature is feasible for predicting prognosis and immunotherapy efficacy. However, the specific mechanisms and how to

adequately activate its potency in immunotherapy still have a long way to go.

As noted above, tumor-infiltrating lymphocytes are extensively involved in the antitumor response of TME and are considered a positive indicator of the prognosis and efficacy of immune checkpoint inhibitors (Germain et al., 2014). Our data discovered that the percentage and distribution of immune cells is positively correlated with the expression level of the TLS gene signature. And interestingly, this study also found that immunosuppressive subsets also occupied a high proportion in the TLS high expression group. This result revealed that immunosuppressive subsets are also essential parts of TLSs. TLSs undertake the regulator in balancing immune infiltration and immune tolerance of TME. One must be vigilant that TLS may also produce auto-reactive lymphocytes with self-mediated immune toxicity, which cripple immune checkpoint blockade.

PD-L1 has been reported to be negatively correlated with prognosis in malignancy (Nakanishi et al., 2007). Atezolizumab, as the PD-L1 inhibitor, is currently the only FDA approved for BCa treatment. Combined with TIDE, IPS analysis, we found that TLS signature is correlated with a better response to ICB. The high TLS signature group are determined as a potentially beneficial population for TLS immunotherapy after verification in BCa data. Considering the ICB only provides remission in a small percentage of patients (Gong et al., 2018), the TLS signature developed by our study would help clinicians stratify patients to select those with potential benefit from ICB. Due to the lack of public data on ovarian cancer immunotherapy, the accuracy of this feature prediction is subject to further debate. Now molecular profiling that guides ovarian cancer immunotherapy includes tumor mutational burden (TMB), homologous repair deficient and proficient (HRD, HRP) phenotypes, neoantigen intratumoral heterogeneity (ITH), and tumor-infiltrating lymphocytes (TILs) (Yang et al., 2020; Morand et al., 2021). Considering the ovarian cancer is a “cold” tumor without high TMB, combining TMB and other potential markers is redefined as necessary to improve the predictive ability. Interestingly, this study discovered that the mutation burden, PDL1 and CTLA4 in high TLS gene signature group are higher than in the low group. So, these could be applied to the combined prediction of OC. As for HRD, it is mainly used to predict the response to poly-(ADP ribose) polymerase (PARP) inhibitors and platinum chemotherapy in OC. Several groups have generalized the HRD gene set, including MMR, BRCA1/2, POLD1 or POLE, MUTYH, and ERCC1 (Zhang et al., 2020). Research on the relationship between mutational details of these genes and TLS may lead to new directions for OC immunotherapy.

5 Conclusion

This study proved the existence of TLS s and its application as a positive prognostic marker in ovarian cancer. And we constructed an ovarian cancer-associated TLS gene signature, which has a high predictive value in the prognosis and response to immunotherapy.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Ethics Committee of Qilu Hospital of Shandong University (KLYY-202107-119-1). The patients/participants provided their written informed consent to participate in this study.

Author contributions

HY, QSJ, and WX designed the research. HY conducted the analysis. HY, QSJ, and LM collected data. HY and QSJ performed experiments. HY, QSJ, and HX analyzed data. HY wrote the paper. MHL and PYX revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Funding

This work was supported by the China Postdoctoral Science Fund (21510077311145 and 21300076311047), Natural Science Foundation of Shandong Province (ZR2020QH252 and ZR2020QH044) and Science Foundation of Qilu Hospital of Shandong Province.

Acknowledgments

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors.

Conflict of interest

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

Publisher’s note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1090640/full#supplementary-material

References

Aloisi, F., and Pujol-Borrell, R. (2006). 'Lymphoid neogenesis in chronic inflammatory diseases. *Nat. Rev. Immunol.* 6, 205–217. doi:10.1038/nri1786

An, Y., and Yang, Q. (2021). 'Tumor-associated macrophage-targeted therapeutics in ovarian cancer. *Int. J. Cancer* 149, 21–30. doi:10.1002/ijc.33408

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). 'Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462, 108–112. doi:10.1038/nature08460

Barmpoutis, P., Di Capite, M., Kayhanian, H., Waddingham, W., Alexander, D. C., Jansen, M., et al. (2021). 'Tertiary lymphoid structures (TLS) identification and density assessment on H&E-stained digital slides of lung cancer. *PLoS One* 16, e0256907. doi:10.1371/journal.pone.0256907

Cabrita, R., Lauss, M., Sanna, A., Donia, M., Skaarup Larsen, M., Mitra, S., et al. (2020). 'Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* 577, 561–565. doi:10.1038/s41586-019-1914-8

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). 'Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Chen, H., Luo, H., Wang, J., Li, J., and Jiang, Y. (2022). 'Identification of a pyroptosis-related prognostic signature in breast cancer. *BMC Cancer* 22, 429. doi:10.1186/s12885-022-09526-z

Di Caro, G., Bergomas, F., Grizzi, F., Doni, A., Bianchi, P., Malesci, A., et al. (2014). 'Occurrence of tertiary lymphoid tissue is associated with T-cell infiltration and predicts better prognosis in early-stage colorectal cancers. *Clin. Cancer Res.* 20, 2147–2158. doi:10.1158/1078-0432.CCR-13-2590

Dieu-Nosjean, M. C., Goc, J., Giraldo, N. A., Sautes-Fridman, C., and Fridman, W. H. (2014). 'Tertiary lymphoid structures in cancer and beyond. *Trends Immunol.* 35, 571–580. doi:10.1016/j.it.2014.09.006

Germain, C., Gnjatic, S., Tamzalit, F., Knockaert, S., Remark, R., Goc, J., et al. (2014). 'Presence of B cells in tertiary lymphoid structures is associated with a protective immunity in patients with lung cancer. *Am. J. Respir. Crit. Care Med.* 189, 832–844. doi:10.1164/rccm.201309-1611OC

Gong, J., Chehrazi-Raffle, A., Reddi, S., and Salgia, R. (2018). 'Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: A comprehensive review of registration trials and future considerations. *J. Immunother. Cancer* 6, 8. doi:10.1186/s40425-018-0316-z

Gu-Trantien, C., Loi, S., Garaud, S., Equeter, C., Libin, M., de Wind, A., et al. (2013). CD4(+) follicular helper T cell infiltration predicts breast cancer survival. *J. Clin. Invest.* 123, 2873–2892. doi:10.1172/JCI67428

Gui, C. P., Wei, J. H., Chen, Y. H., Fu, L. M., Tang, Y. M., Cao, J. Z., et al. (2021). *'A new thinking: extended application of genomic selection to screen multiomics data for development of novel hypoxia-immune biomarkers and target therapy of clear cell renal cell carcinoma*, 22.*Brief. Bioinform*

Hegde, P. S., Karanikas, V., and Evers, S. (2016). 'The where, the when, and the how of immune monitoring for cancer immunotherapies in the era of checkpoint inhibition. *Clin. Cancer Res.* 22, 1865–1874. doi:10.1158/1078-0432.CCR-15-1507

House, I. G., Savas, P., Lai, J., Chen, A. X. Y., Oliver, A. J., Teo, Z. L., et al. (2020). Macrophage-derived CXCL9 and CXCL10 are required for antitumor immune responses following immune checkpoint blockade. *Clin. Cancer Res.* 26, 487–504. doi:10.1158/1078-0432.CCR-19-1868

Ivetic, A., Hoskins Green, H. L., and Hart, S. J. (2019). 'L-selectin: A major regulator of leukocyte adhesion, migration and signaling. *Front. Immunol.* 10, 1068. doi:10.3389/fimmu.2019.01068

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). 'Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1

Kroeger, D. R., Milne, K., and Nelson, B. H. (2016). 'Tumor-Infiltrating plasma cells are associated with tertiary lymphoid structures, cytolytic T-cell responses, and superior prognosis in ovarian cancer. *Clin. Cancer Res.* 22, 3005–3015. doi:10.1158/1078-0432.CCR-15-2762

Li, Q., Liu, X., Wang, D., Wang, Y., Lu, H., Wen, S., et al. (2020). 'Prognostic value of tertiary lymphoid structure and tumour infiltrating lymphocytes in oral squamous cell carcinoma. *Int. J. Oral Sci.* 12, 24. doi:10.1038/s41368-020-00092-3

Lin, Z., Huang, L., Li, S., Gu, J., Cui, X., and Zhou, Y. (2020). Pan-cancer analysis of genomic properties and clinical outcome associated with tumor tertiary lymphoid structure. *Sci. Rep.* 10, 21530. doi:10.1038/s41598-020-78560-3

Lucchesi, D., and Bombardieri, M. (2013). 'The role of viruses in autoreactive B cell activation within tertiary lymphoid structures in autoimmune diseases. *J. Leukoc. Biol.* 94, 1191–1199. doi:10.1189/jlb.0413240

Lutge, M., Pikor, N. B., and Ludewig, B. (2021). 'Differentiation and activation of fibroblastic reticular cells. *Immunol. Rev.* 302, 32–46. doi:10.1111/imr.12981

Mahoney, K. M., Rennert, P. D., and Freeman, G. J. (2015). 'Combination cancer immunotherapy and new immunomodulatory targets. *Nat. Rev. Drug Discov.* 14, 561–584. doi:10.1038/nrd4591

Marchetti, C., De Felice, F., Romito, A., Iacobelli, V., Sassu, C. M., Corrado, G., et al. (2021). 'Chemotherapy resistance in epithelial ovarian cancer: Mechanisms and emerging treatments. *Semin. Cancer Biol.* 77, 144–166. doi:10.1016/j.semcancer.2021.08.011

Mariathasan, S., Turley, S. J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., et al. (2018). TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 554, 544–548. doi:10.1038/nature25501

Meylan, M., Petitprez, F., Becht, E., Bougouin, A., Pupier, G., Calvez, A., et al. (2022). 'Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity* 55, 527–541.e5. doi:10.1016/j.immuni.2022.02.001

Mirza, M. R., Coleman, R. L., Gonzalez-Martin, A., Moore, K. N., Colombo, N., Ray-Coquard, I., et al. (2020). The forefront of ovarian cancer therapy: Update on PARP inhibitors. *Ann. Oncol.* 31, 1148–1159. doi:10.1016/j.annonc.2020.06.004

Morand, S., Devanaboyina, M., Staats, H., Stanbery, L., and Nemunaitis, J. (2021). 'Ovarian cancer immunotherapy and personalized medicine. *Int. J. Mol. Sci.* 22, 6532. doi:10.3390/ijms22126532

Moyron-Quiroz, J. E., Rangel-Moreno, J., Hartson, L., Kusser, K., Tighe, M. P., Klonowski, K. D., et al. (2006). 'Persistence and responsiveness of immunologic memory in the absence of secondary lymphoid organs. *Immunity* 25, 643–654. doi:10.1016/j.immuni.2006.08.022

Nakanishi, J., Wada, Y., Matsumoto, K., Azuma, M., Kikuchi, K., and Ueda, S. (2007). 'Overexpression of B7-H1 (PD-L1) significantly associates with tumor grade and postoperative prognosis in human urothelial cancers. *Cancer Immunol. Immunother.* 56, 1173–1182. doi:10.1007/s00262-006-0266-z

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

Rodriguez, A. B., Peske, J. D., Woods, A. N., Leick, K. M., Mauldin, I. S., Meneveau, M. O., et al. (2021). 'Immune mechanisms orchestrate tertiary lymphoid structures in tumors via cancer-associated fibroblasts. *Cell Rep.* 36, 109422. doi:10.1016/j.celrep.2021.109422

Sautes-Fridman, C., Petitprez, F., Calderaro, J., and Fridman, W. H. (2019). 'Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* 19, 307–325. doi:10.1038/s41568-019-0144-6

Seitz, S., Dreyer, T. F., Stange, C., Steiger, K., Brauer, R., Scheutz, L., et al. (2022). 'CXCL9 inhibits tumour growth and drives anti-PD-L1 therapy in ovarian cancer. *Br. J. Cancer* 126, 1470–1480. doi:10.1038/s41416-022-01763-0

Sharpe, A. H., and Pauken, K. E. (2018). 'The diverse functions of the PD1 inhibitory pathway. *Nat. Rev. Immunol.* 18, 153–167. doi:10.1038/nri.2017.108

Shi, Z., Shen, J., Qiu, J., Zhao, Q., Hua, K., and Wang, H. (2021). 'CXCL10 potentiates immune checkpoint blockade therapy in homologous recombination-deficient tumors. *Theranostics* 11, 7175–7187. doi:10.7150/thno.59056

Shoji, T., Enomoto, T., Abe, M., Okamoto, A., Nagasawa, T., Oishi, T., et al. (2022). 'Efficacy and safety of standard of care with/without bevacizumab for platinum-resistant ovarian/fallopian tube/peritoneal cancer previously treated with bevacizumab: The Japanese Gynecologic Oncology Group study JGOG3023. *Cancer Sci.* 113, 240–250. doi:10.1111/cas.15185

Silina, K., Soltermann, A., Attar, F. M., Casanova, R., Uckeley, Z. M., Thut, H., et al. (2018). 'Germinal centers determine the prognostic relevance of tertiary lymphoid structures and are impaired by corticosteroids in lung squamous cell carcinoma. *Cancer Res.* 78, 1308–1320. doi:10.1158/0008-5472.CAN-17-1987

Sinclair, L. V., Finlay, D., Feijoo, C., Cornish, G. H., Gray, A., Ager, A., et al. (2008). 'Phosphatidylinositol-3-OH kinase and nutrient-sensing mTOR pathways control T lymphocyte trafficking. *Nat. Immunol.* 9, 513–521. doi:10.1038/ni.1603

Thaunat, O., Patey, N., Caligiuri, G., Gautreau, C., Mamani-Matsuda, M., Mekki, Y., et al. (2010). 'Chronic rejection triggers the development of an aggressive intragraft immune response through recapitulation of lymphoid organogenesis. *J. Immunol.* 185, 717–728. doi:10.4049/jimmunol.0903589

Tokunaga, R., Zhang, W., Naseem, M., Puccini, A., Berger, M. D., Soni, S., et al. (2018). 'CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - a target for novel cancer therapy. *Cancer Treat. Rev.* 63, 40–47. doi:10.1016/j.ctrv.2017.11.007

Torre, L. A., Trabert, B., DeSantis, C. E., Miller, K. D., Samimi, G., Runowicz, C. D., et al. (2018). 'Ovarian cancer statistics, 2018. *CA Cancer J. Clin.* 68, 284–296. doi:10.3322/caac.21456

Ukita, M., Hamanishi, J., Yoshitomi, H., Yamanoi, K., Takamatsu, S., Ueda, A., et al. (2022). 'CXCL13-producing CD4+ T cells accumulate in the early phase of tertiary lymphoid structures in ovarian cancer. *JCI Insight* 7, e157215. doi:10.1172/jci.insight.157215

Van Allen, E. M., Miao, D., Schilling, B., Shukla, S. A., Blank, C., Zimmer, L., et al. (2015). Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350, 207–211. doi:10.1126/science.aad0095

Vella, G., Guelfi, S., and Bergers, G. (2021). 'High endothelial venules: A vascular perspective on tertiary lymphoid structures in cancer. *Front. Immunol.* 12, 736670. doi:10.3389/fimmu.2021.736670

Watson, H. A., Durairaj, R. R. P., Ohme, J., Alatsatianos, M., Almutairi, H., Mohammed, R. N., et al. (2019). 'L-Selectin enhanced T cells improve the efficacy of cancer immunotherapy. *Front. Immunol.* 10, 1321. doi:10.3389/fimmu.2019.01321

Workel, H. H., Lubbers, J. M., Arnold, R., Prins, T. M., van der Vlies, P., de Lange, K., et al. (2019). 'A transcriptionally distinct CXCL13(+)CD103(+)CD8(+) T-cell population is associated with B-cell recruitment and neoantigen load in human cancer. *Cancer Immunol. Res.* 7, 784–796. doi:10.1158/2326-6066.CIR-18-0517

Wouters, M. C. A., and Nelson, B. H. (2018). 'Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer. *Clin. Cancer Res.* 24, 6125–6135. doi:10.1158/1078-0432.CCR-18-1481

Yang, C., Xia, B. R., Zhang, Z. C., Zhang, Y. J., Lou, G., and Jin, W. L. (2020). 'Immunotherapy for ovarian cancer: Adjuvant, combination, and neoadjuvant. *Front. Immunol.* 11, 577869. doi:10.3389/fimmu.2020.577869

Yang, M., Lu, J., Zhang, G., Wang, Y., He, M., Xu, Q., et al. (2021). 'CXCL13 shapes immunoactive tumor microenvironment and enhances the efficacy of PD-1 checkpoint blockade in high-grade serous ovarian cancer. *J. Immunother. Cancer* 9, e001136. doi:10.1136/jitc-2020-001136

Yang, T., Cheng, J., You, J., Yan, B., Liu, H., and Li, F. (2018). S100B promotes chemoresistance in ovarian cancer stem cells by regulating p53. *Oncol. Rep.* 40, 1574–1582. doi:10.3892/or.2018.6527

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

Zhang, J., Shih, D. J. H., and Lin, S. Y. (2020). 'Role of DNA repair defects in predicting immunotherapy response. *Biomark. Res.* 8, 23. doi:10.1186/s40364-020-00202-7

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 07 April 2023
DOI 10.3389/fgene.2023.1085584

OPEN ACCESS

EDITED BY
Anton A. Buzdin,
European Organisation for Research and Treatment of Cancer, Belgium

REVIEWED BY
Yuhao Zhang,
Zhejiang Provincial People's Hospital, China
Weiren Luo,
The Second Affiliated Hospital of Southern University of Science and Technology, China

*CORRESPONDENCE
Shiyin Fu,
fu_shiyin@126.com

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

SPECIALTY SECTION
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 31 October 2022
ACCEPTED 27 March 2023
PUBLISHED 07 April 2023

CITATION
Hu Y, Yang Q, Cai S, Wang W and Fu S (2023), The integrative analysis based on super-enhancer related genes for predicting different subtypes and prognosis of patient with lower-grade glioma.
Front. Genet. 14:1085584.
doi: 10.3389/fgene.2023.1085584

The integrative analysis based on super-enhancer related genes for predicting different subtypes and prognosis of patient with lower-grade glioma

Yungang Hu[1†], Qingqing Yang[2†], Shuzhou Cai[1], Wei Wang[1] and Shiyin Fu[3*]

[1]Department of Neurosurgery, Wuhan University of Science and Technology Affiliated Xiaogan Central Hospital, Xiaogan, Hubei, China, [2]Department of Thyroid and Breast Surgery, Wuhan University of Science and Technology Affiliated Xiaogan Central Hospital, Xiaogan, Hubei, China, [3]Department of Pediatric, Jinchu University of Technology Affiliated Central Hospital, Jingmen, Hubei, China

Objective: Emerging evidence revealed that super-enhancer plays a crucial role in the transcriptional reprogramming for many cancers. The purpose aimed to explored how the super-enhancer related genes affects the prognosis and tumor immune microenvironment (TIME) of patients with low-grade glioma (LGG).

Methods: In this study, the differentially expressed genes (DEGs) between LGG cohorts and normal brain tissue cohort were identified by the comprehensive analysis of the super-enhancer (SE) related genes. Then non-negative matrix factorization was performed to seek the optimal classification based on the DEGs, while investigating prognostic and clinical differences between different subtypes. Subsequently, a prognostic related signature (SERS) was constructed for the comprehensive evaluation in term of individualized prognosis, clinical characteristics, cancer markers, genomic alterations, and immune microenvironment of patients with LGG.

Results: Based on the expression profiles of 170 DEGs, we identified three SE subtypes, and the three subtypes showed significant differences in prognostic, clinicopathological features. Then, nine optimal SE-related genes were selected to construct the SERS through the least absolute shrinkage and selection operator Cox regression analysis. Survival analysis showed that SERS had strong and stable predictive ability for the prognosis of LGG patients in the The Cancer Genome Atlas, China Glioma Genome Atlas, and Remdrandt cohorts, respectively. We also found that SERS was highly correlated with clinicopathological features, tumor immune microenvironment, cancer hallmarks, and genomic alterations in LGG patients. In addition, the predictive power of SERS for immune checkpoint inhibitor treatment is also superior. The qRT-PCR results and immunohistochemical results also confirmed the difference in the expression of four key genes in normal cells and tumors, as well as in normal tissues and tumor tissues.

Conclusion: The SERS could be suitable to utilize individualized prognosis prediction and immunotherapy options for LGG patients in clinical application.

KEYWORDS

super-enhancer, lower-grade glioma, prognostic signature, tumor immune microenvironment, immunotherapy

1 Introduction

Gliomas are the most common intracranial malignant tumor, accounting for more than 80% of primary malignant tumor in central nervous system (Ostrom et al., 2022). Low-grade gliomas (LGG), also known as WHO grade II and III tumors defined by the World Health Organization, are composed of diffuse low-grade and intermediate-grade gliomas (Brat et al., 2015). Compared with patients with glioblastoma (GBM), LGG patients have a relatively low degree of malignancy (Chen et al., 2022). However, even with comprehensive treatment including surgical resection, radiotherapy and chemotherapy, some LGG patients still have the characteristics of high recurrence and progression rates (Liu et al., 2018; Jiang et al., 2021). Significant heterogeneity in patient outcomes and treatment response remains a major clinical challenge for neurosurgeons. Traditionally, WHO grade II gliomas were considered to have a better prognosis than WHO grade III gliomas, but since the WHO reclassification of gliomas in 2016, molecular alterations have been considered more objective and precise than grading (Gittleman et al., 2020). Although there have been some progress in the onlooker research on LGG in recent years (Xu et al., 2021a), few drugs are currently approved for the treatment of LGG patients, and the prognosis has not been significantly improved (Ye et al., 2021). Therefore, there is an urgent need to explore new biomarkers to predict the prognosis of LGG patients and find potential therapeutic targets.

Gene regulation plays a major role in tumor pathogenesis, and the regulation of long non-coding RNAs (lncRNAs) on tumors is the hotspot of current research (Lou et al., 2020). Aberrant gene expression promotes tumorigenesis, progression and metastasis (Mansour et al., 2014). Enhancers in gene regulatory elements can bidirectionally transcribe enhancer RNA, a non-coding RNA transcribed by enhancers, that not only drives tumorigenesis, but also regulates genes and immune checkpoints (Lee et al., 2020). Super-enhancers (SE) are clusters of enhancers formed by contiguously arranged enhancers in tandem. SE usually appear near most of the key genes that determine cell identity and function, and play a more effective role than typical enhancers (Hnisz et al., 2013; Whyte et al., 2013). They have the ability to flexibly regulate, by combining unneeded regions to form highly concentrated regional transcriptional machinery, thereby affecting epigenetics and regulating tumorigenesis and progression (Chen et al., 2018). The researchers also found that SE operate covertly in a particularly latent manner, but control across multiple cancer lineages, with cancer cells assembling their own super-enhancers, thereby overproducing malignant oncogenes, exhibiting cancer hallmarks of hyperplasia, invasion and metastasis (Whyte et al., 2013). Yang believed that identifying, mapping out, and disrupting SE has the potential to transform how clinical cancer is managed (Whyte et al., 2013). Hence, as we concentrated on personalized therapy for patients with cancers, SE can serve as the potential biomarkers to track and understand the evolution of individual cancers, and ultimately may become important targets in therapeutic interventions.

The lncRNA HCCL5 in human tissue cells was identified as a SE-driven oncogenic factor that promotes the malignant development of hepatocellular carcinoma by promoting HCC cell viability, migration, and classical epithelial-mesenchymal transition (Peng et al., 2019). TCOF1 depletion in triple-negative breast cancer patients significantly inhibited the growth and invasiveness of triple-negative breast cancer cells (Hu et al., 2022a). Heparanase (HPSE) is a cancer metastasis protein that is regulated by the hnRNPU/p300/EGR1/HPSE axis, promotes high expression of HPSE enhancer RNA, is an independent prognostic factor for poor prognosis in cancer patients (Jiao et al., 2018). Regarding whether SE affect progression and overall survival in patients with LGG, the jury is still out.

Therefore, to solve the above problems, this study investigated the effect of different types of SE-related genes on the survival of patients with LGG by collecting data from The Cancer Genome Atlas (TCGA), China Glioma Genome Atlas (CGGA) and Rembrandt Database for LGG. At the same time, we constructed and evaluated prognostic score (SERS) based on 8 SE-related genes for patients with LGG. On top of that, the relationship was also explored between SERS and prognosis, clinicopathological features, tumor immune microenvironment, cancer hallmarks, genomic alterations and immunotherapy efficacy in patients with LGG. We provided a new strategy for predicting the prognosis of and assessing treatment effects for patients with LGG, and thus the findings of this study will help individualized prognosis prediction and immunotherapy decisions in patients with LGG.

2 Materials and methods

2.1 Data collection and study population

The RNA sequencing data and clinical information of LGG patients were extracted from TCGA (https://portal.gdc.cancer.gov/), CGGA (http://www.cgga.org.cn/) and Rembrandt (http://gliovis.bioinfo.cnio.es/) databases. A total of 5 LGG cohorts were gather in this study, namely, the TCGA, CGG693, CGGA325, CGGA301 and Rembrandt cohorts, respectively. Patients with no survival data or overall survival (OS) < 30 days were excluded from further analysis. Zakharova et al. (Zakharova et al., 2022) had re-classified the TCGA sampling according to the updated WHO CNS Tumor Classification in 2021, we therefore used the updated glioma diagnoses for analysis in this study. The transcriptome data with normal brain tissue were also obtained from Genotype-Tissue Expression (GTEx; https://gtexportal.org/home/). Furthermore, SE-related gene can be downloaded from the SEA v. 3.0 database (http://sea.edbc.org). The clinicopathological characteristics of LGG patients in five cohorts are generalized in Table 1.

The differential expression analysis firstly performed based on GTEx dataset and TCGA dataset, and finally 1,672 differentially expressed genes (DEGs) were extracted with the cutoff values of log2 fold-change |logFC|>2 and p-value < 0.05. Then, in the same way, 285 DEGs were extracted between GTEx dataset and TCGA

TABLE 1 Characteristics of glioma patients in training and validation cohorts.

Clinicopathological characteristics	Training cohort	Validation cohorts			
	TCGA	CGGA693	CGGA325	CGGA325	Rembrandt
Number of patients	331	420	170	158	119
Age (mean ± SD; years)	41.3 ± 13.2	40.3 ± 10.4	40.4 ± 10.9	39.6 ± 10.6	NA
Gender					
Female	146	185	65	68	37
Male	185	235	105	90	59
NA	0	0	0	0	0
Survival status					
Alive	272	223	82	85	34
Dead	59	197	88	73	85
Preoperative KPS					
<80	50	NA	NA	NA	NA
≥80	93	NA	NA	NA	NA
NA	188	NA	NA	NA	NA
Histology					
Astrocytoma	193	254	110	102	80
Oligoastrocytoma	0	29	0	18	0
Oligodendroglioma	138	137	60	38	34
NA	0	0	0	0	0
WHO grade					
II	179	172	97	105	63
III	152	248	73	53	56
NA	0	0	0	0	0
IDH status					
Mutant	331	288	125	104	NA
Wild type	0	94	44	1	NA
NA	0	38	1	53	NA
1p19q codeletion					
Codeletion	138	125	55	16	8
Non-codeletion	193	257	113	33	13
NA	0	38	2	109	98
MGMT promoter status					NA
Methylated	271	200	84	43	NA
Unmethylated	60	129	70	106	NA
NA	0	38	16	9	NA
TERT status					NA
Mutant	116	NA	NA	NA	NA
Wild type	137	NA	NA	NA	NA
NA	78	NA	NA	NA	NA

dataset. Eventually, the differentially expressed SE-related genes were shared by two cohorts were considered eligible.

2.2 Identification of SE subtypes of LGG patients

Based on the above DEGs, non-negative matrix factorization (NMF) consensus clustering analysis was performed to obtain the optimal SE subtypes of LGG patients (Hillman et al., 2018). The commonality, dispersion and contour indicators are used to judge the optimal number of subtypes. The t-distributed stochastic neighbor embedding (tSNE) algorithm we applied to confirm the reliability of clustering results by naked eyes. The Kaplan-Meier survival curves were then used to identify differences in survival difference among different SE subtypes. In addition, we compared differences in clinicopathological features among different SE subtypes.

2.3 Construction and validation of a prognostic SERS

The univariate Cox regression was conducted to select the prognostic SE-related DEGs. Then the least absolute shrinkage

and selection operator (LASSO) Cox regression analysis was performed to identify the SE-related prognostic signature (SERS) in the TCGA cohort (Friedman et al., 2010). The prognostic risk score of each LGG patient was calculated with the regression coefficient and the expression of the corresponding gene. The calculation formula of SERS was shown below:

$$Risk\ score = \sum_{i=1}^{n} (Coef_i * X_i)$$

where n represents the number of all the selected gene; i represents the serial number of each gene; X_i and $Coef_i$ refer to the expression level of each selected gene and corresponding coefficient, respectively. The cut-off value, defined as the median risk score was divide the patients into high- or low-risk group. The Kaplan-Meier survival curve analysis were conducted to evaluate the accuracy of prognosis of LGG patients between the high- and low-risk groups. The receiver operating characteristic (ROC) curves and the area under the ROC curves (AUC) were plotted and calculated to describe the accuracy of predicting OS. The above analyses were performed simultaneously in the TCGA cohort and four independent validation cohorts. What is more, we finally conducted meta-analysis to calculate the pooled hazard ratio of SERS.

2.4 Development of a nomogram

Initially, the univariate Cox regression analysis were performed based on SERS and clinicopathological features, and then multivariate Cox regression analysis was used to identify independent predictors in the TCGA cohort. The nomogram was developed in the TCGA cohort to individually predict 1, 3, and 5-year survival probabilities in LGG patients. And the predicted outcomes for LGG patients were presented in the form of ROC curves. To evaluate the stability of this nomogram, a 10-fold cross-validation algorithm was performed in the TCGA cohort for the internal validation, and the external validation was conducted in the other four independent cohorts. In addition, calibration curves and C-index were performed in the TCGA and validation cohorts to evaluate the usability of this nomogram.

2.5 Evaluation of genomic alterations

Tumor mutational burden (TMB) was calculated as the total number of somatic, coding, base substitution, and indel mutations examined per megabase of genome (Mayakonda et al., 2018). The somatic mutation profile ordered in Mutation Annotation Format (MAF) was obtained from the TCGA database. The mutation spectrum and frequency differences were analyzed between high and low risk genes (Bi et al., 2020). In addition, copy number alteration (CNA) data in LGG patients were obtained from the TCGA database. We used GISTIC2.0 to identify significant amplifications or deletions genome-wide. CNA burden was defined as the total number of genes with copy number changes at the focal and arm levels (Shen et al., 2019).

2.6 Assessment of TIME and immunotherapeutic responses

For the purpose of better understanding the underlying biological functions of DEGs between high-risk and low-risk groups, the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed to identify annotated functions and gene enrichment pathways (Hu et al., 2022b). DEGs between high- and low-risk groups were set the cutoff values of $|\log2FC|>2$ and the BH method adjusted $p < 0.05$.

There has been an increasing recognition that the interaction of cancer cells and tumor microenvironment may best be conceptualized as an ecological process (Kenny et al., 2006). Hence, the ESTIMATE algorithm was used to calculate the immune score, stromal score, ESTIMATE score and tumor purity in LGG patients (Yoshihara et al., 2013) for assessing the difference of stromal and immune cells in LGG. Simultaneously, CIBERSORT was performed to calculate the proportions of 22 immune cells from LGG based on gene expression (Newman et al., 2015). In addition, Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was also applied to assess potential response to immune checkpoint inhibitions (ICI) therapy for LGG patients (Jiang et al., 2018).

2.7 Quantitative real-time polymerase chain reaction (qRT -PCR) and immunohistochemistry (IHC)

The normal human astrocyte line HA1800 and glioma cell lines U87, U251, A172 and LN229, were purchased from the Cell Bank of the Chinese Academy of Sciences. The clinical specimens of 10 LGG patients were collected in the Department of Neurosurgery of Wuhan Union Hospital from June 2021 to December 2021. Ten non-tumor brain tissues were obtained from patients with brain tissue resection due to craniocerebral injury from June 2021 to December 2021. The study was approved by the Medical Ethics Committee of our hospital, and the informed consent was obtained from each patient. Total RNA was extracted from cell lines and sample tissues using RNAiso Plus (Takara 9109). According to the instruction, cDNA was synthesized by reverse transcription through using HiScript® III RT SuperMix for qPCR (+gDNA wiper) (Vazyme R323-01). The qRT-PCR analyses were performed using the AceQ® qPCR SYBR Green Master Mix (Vazyme Q111-02) with PCR LightCycler480 (Roche Diagnostics, Basel, Switzerland). All expression data was normalized to GAPDH as an internal control using the $2^{-\Delta\Delta CT}$ method. All primers used were synthesized by GeneCreate Biological Engineering Co., Ltd. (Wuhan, China). The protein levels of the selected genes were then verified by IHC experiments. In addition, the relations between the selected gene and tumor immune features also analyzed in LGG patients.

2.8 Statistical analysis

The PERL language (version, 5.30.2, http://www.perl.org) was used to preprocess RNA-seq transcriptome information. The R software (version 4.0.1, http://www.R-project.org) were conducted

for statistical analyses and graph visualization. Continuous variables are described as mean ± standard deviation, and categorical variables are described as frequency (n) and proportion (%). Chi-square test or Fisher's exact test was performed to compare categorical variables between two groups. Student's *t*-test or one-way ANOVA was used to compare continuous variables with normal distribution between two or among more groups. The Mann-Whitney *U* test was used to compare non-normally distributed continuous variables between two groups, while the Kruskal Wallis test was used to compare non-normally distributed continuous variables among three or more groups. Survival differences between groups were assessed using Kaplan-Meier curves. Univariate and multivariate cox proportional hazards models were applied to estimate hazard ratios for variables and to identify independent prognostic factors. The cutoff value with statistical significance was set at two-tailed $p < 0.05$.

3 Results

3.1 Overall structure of this study

First of all, the GTEs between LGG and normal brain tissues were screen out. Based on the expression profiles of theses selected GTEs, NMF consensus clustering was performed to construct SE subtypes of LGG patients. Then, we explored the heterogeneities of prognosis and clinicopathological features for SE subtypes. Subsequently, the Univariate Cox regression analysis LASSO Cox algorithm were combined to screen for robust SERS and presented as a nomogram. The effectiveness of SERS was assessed in multiple dimensions. The overall flow diagram of this study was presented in Figure 1.

3.2 Identification of SE subtypes in TCGA cohort based on the DEGs

The differential expression analysis based on GTEx dataset and TCGA dataset was shown as the volcano in Figure 2A, and differential expression analysis based on GTEx dataset and CGGA693 dataset was also shown as the volcano in Figure 2B. Then, a total of 170 DEGs (Figure 2C) shared by two cohorts were used for subsequent analysis, they can be found in Supplementary Table S1.

Based on the expression profiles of 170 SE associated DEGs, the NMF was performed in the TCGA cohort to identify SE subtypes. As shown in Figure 2D, we chose 3 as the optimal number of clusters based on common, scatter, and contour metrics. Then, a total of 469 LGG patients were divided into three subtypes (Figure 2E), named SE1 ($n = 125$), SE2 ($n = 75$), and SE3 ($n = 125$). The heatmap of the consensus matrix exhibits clear boundaries, indicating the accuracy and robustness of the clustering results. t-SNE plot showed clear differences in the distribution between the three SE subtypes (Figure 2F). Significant differences in the expression of 170 prognostic SE-related DEGs can also be observed in the heatmap in Figure 2G. Kaplan-Meier survival curves showed obvious survival differences among the three SE subtypes (Figure 2H). The LGG patients in SE1 subtype had the best survival outcome, while SE2 had the worst survival outcome. At the same time, the heterogeneity of clinicopathological characteristics of these three subtypes were analyzed and found interestingly no significant differences among these clinicopathological characteristics (Supplementary Figure S1).

3.3 Development and validation of the SERS

The Univariate Cox regression analyses were conducted based on DEGs to identify prognostic SERG. The results of the analysis indicated 33 genes were obviously related to the OS of LGG, and detailed information for these prognostic SERG was shown in Supplementary Table S2. Then, the LASSO analysis was performed on above 33 prognostic SERG in the TCGA cohort to explore simplest and most accurate model. Finally, a total of 9 optimal prognostic SERG (AQP7, MYOD1, CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2) were stood out and incorporated in the SERS (Figures 3A, B). Figure 3C exhibited the LASSO coefficients of each selected gene in this signature. Among them, the coefficients of seven genes (CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2) are positive number, which means that they are related to bad prognosis for LGG patients, whereas the coefficients AQP7 and MYOD1 are negative number, indicating a good prognosis. The Kaplan-Meier survival curves of these nine optimal genes were shown in Supplementary Figure S2. The risk score of each patient was calculated as follows: SERS score = (−0.408 × AQP7) + (−0.107 × MYOD1) + 0.186 × CDCA2 + 0.625 × FAM92B + 0.163 × HOXA11 − AS + 0.454 × E2F7 + 0.344 × KIF18A + 0.130 × MC5R + 0.201 × SKOR2. Subsequently, the median SERS score in was set as the cut-off value to stratified the 325 LGG patients into the high- and low-risk groups. Heatmap analysis of nine genes showed markedly different distributions between high- and low-risk groups, the risky genes were upregulated in the high-risk group and the protective genes were upregulated in the low-risk group (Figure 3D).

The SERS was calculated with LASSO coefficients obtained from the TCGA cohort to stratified into with the median score high- and low-risk groups in other 4 cohorts. The Kaplan-Meier survival curves demonstrated that patients with high-SERS showed worse OS than low-SERS in the TCGA cohort (log-rank test $p < 0.001$; Figure 3E). Consistent results were also observed in four other independent validation cohorts (log-rank test $p < 0.001$; Figures 3F–I). The distribution plot of the risk score and survival status showed that the SERS had the positively correlation with the deaths of LGG patients (Supplementary Figures S3A–E). Furthermore, the ROC curves confirmed the satisfactory predictive performance of the of SERS in predicting 1-, 3-, and 5-year OS (Supplementary Figures S3F–I). Thus, SERS were sufficiently discriminative on both the validation cohorts. In addition, a meta-analysis was performed to assess the overall predicting accuracy, and the results indicated that the overall pooled HR for SERS was 3.2 (95% CI = 1.69–6.08; Figure 3J).

3.4 Relationship between SERS and clinicopathological characteristics, genomic alterations

The clinical relationships of SERS were attempted to explore in the TCGA cohort. As shown in Figure 4A, SERS were arranged

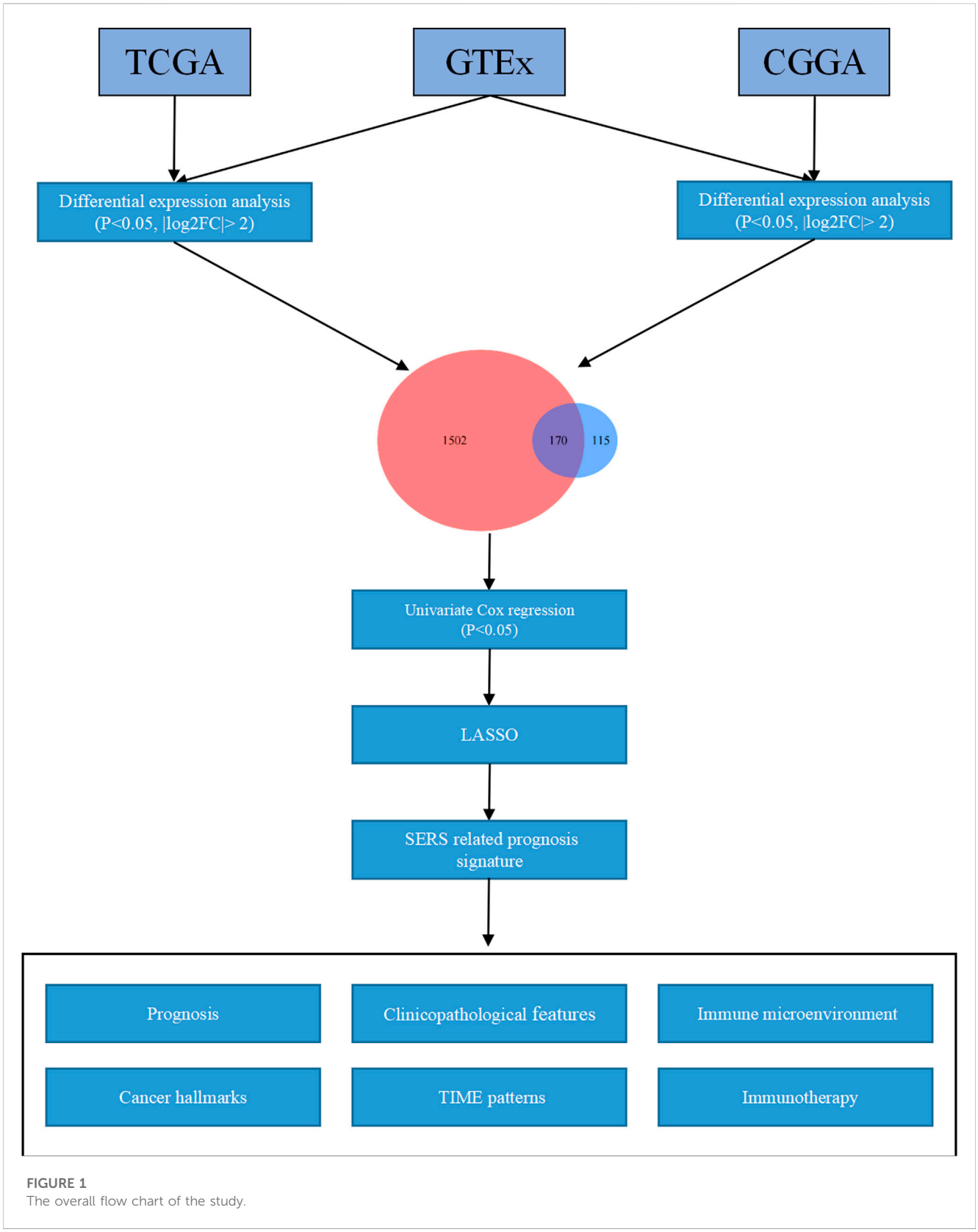


FIGURE 1
The overall flow chart of the study.

from low to high to show the correlation between SERS and clinicopathological characteristics. There were significant differences in, survival status, Histology, 1p19q status, and SE subtype between high and low SERS groups, but no significant differences in -age, gender, WHO grade, MGMT status, TERT status and Transcriptome subtype. Furthermore, SERS levels

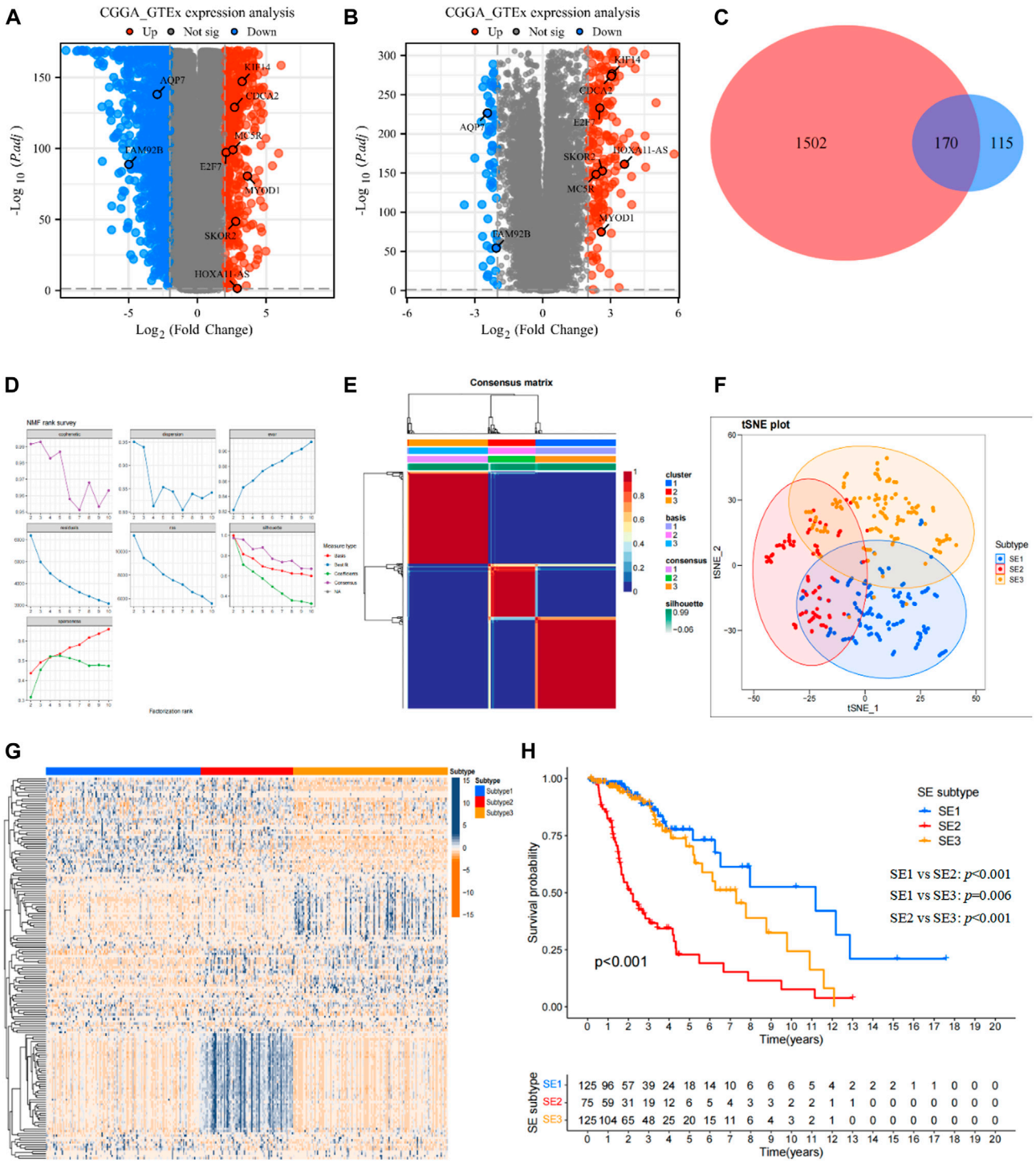


FIGURE 2
Identification of SE subtypes of LGG by using NMF algorithm. **(A)** Volcano plot showed DEGs ($p < 0.05$ and |log2FC|>2) between LGG tissues in TCGA cohort and normal brain tissues in GTEx database. **(B)** Volcano plot showed DEGs ($p < 0.05$ and |log2FC|>2) between LGG tissues in CGGA cohort and normal brain tissues in GTEx database. **(C)** Venn diagram identified prognostic super-enhancer related DEGs. **(D)** The NMF rank survey of TCGA cohort using theSE-related DEGs. **(E)** Consensus map of NMF clustering. **(F)** tSNE plot of 170 SE-related DEGs to distinguish SE subtypes. **(G)** Heatmap showed the expression levels of 170 SE-related DEGs among SE subtypes. **(H)** Kaplan–Meier survival analysis exhibited significantly different OS among three SE subtypes.

between LGG patients stratified by various clinicopathological features were compared. The results of the analyses showed that LGG patients with, death status, Oligodendroglioma and SE2 subtype showed significant higher SERS, while no significant differences of SERS were observed in other subgroups (Figures 4B–K). Likewise, the relationship between

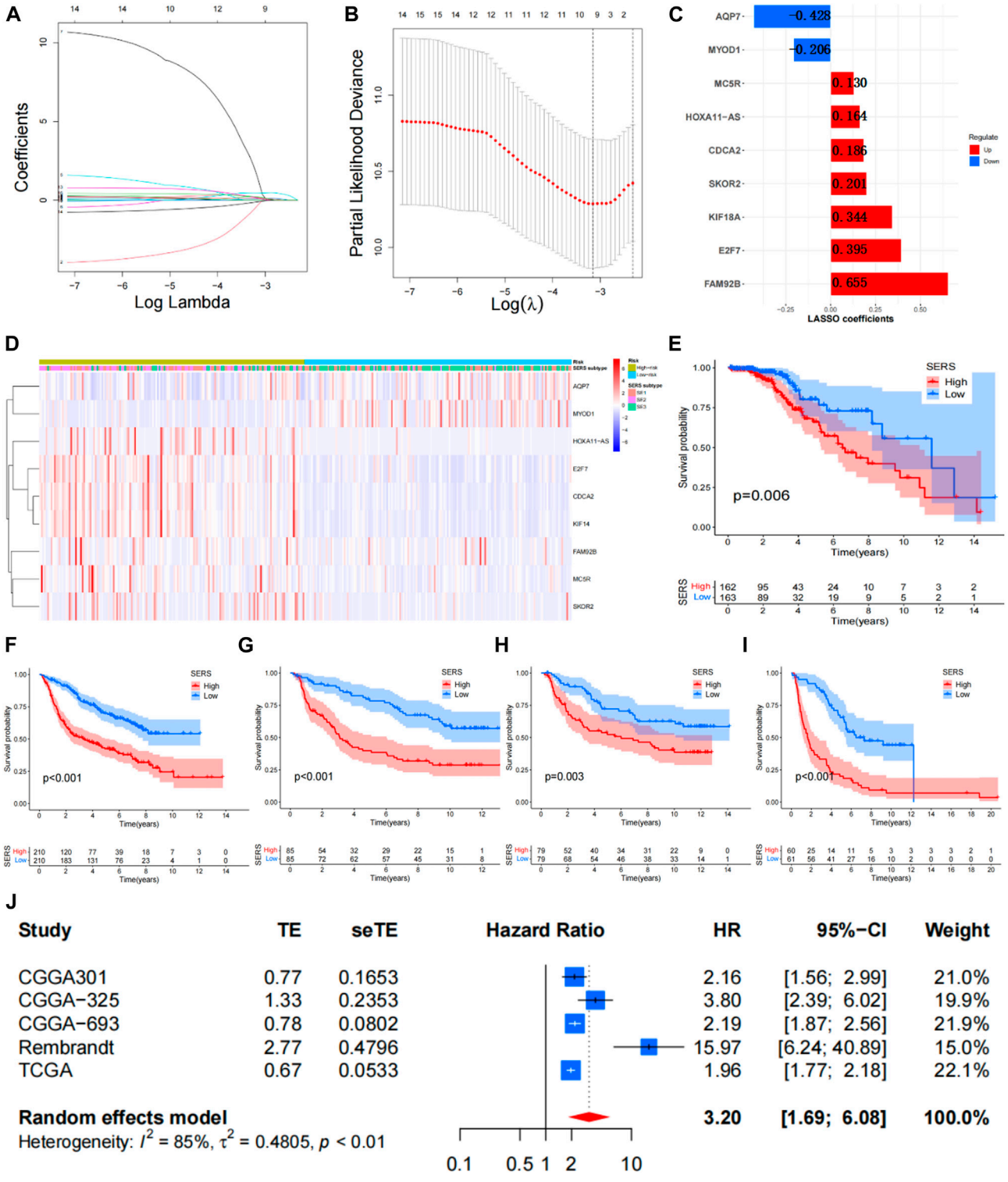


FIGURE 3
Construction of the SERS for LGG patients. **(A,B)** The LASSO regression was performed to minimize the risk of overfitting with the minimum criteria. **(C)** LASSO coefficients of selected the SERS. **(D)** Heatmap was represented expression levels of 8 SE-related genes in the high- and low-risk groups, respectively. **(E–I)** The Kaplan–Meier survival curves of SERS in TCGA, CGGA693, CGGA325, CGGA301, and Rembrandt cohorts, respectively. **(J)** Meta-analysis with random_effects showed a pooled hazard ratio (HR) of SERS.

SERS and clinicopathological characteristics of LGG patients in the CGGA693, CGGA325, CGGA301 and Rembrandt cohorts was also identified the similar results to the TCGA cohort (Supplementary Figures S4–S7).

To better address the prognostic features associated with SERS, the correlation between common cancer markers and SERS were also explored. The correlation heatmap showed that SERS was significantly positively correlated with many well-known cancer hallmarks including DNA repair, cell cycle, hypoxia, and metabolism (Figure 5A). The correlation between SERS and 29 immune signatures was illustrated by a correlation heatmap in TCGA cohort (Figure 5B). Subsequently, further analysis showed that SERS was significantly positively associated with TMB, mutation count, copy number gain and loss burden at the focal level, and copy number gain burden at the arm level (Figures 5C–H). The distribution of TMB, mutation counts, copy number burdens at focal and arm levels between high and low-risk groups were also compared in TCGA cohort (Supplementary Figures S10A–F). Based on the above data, it is indicated that high SERS may represent a higher frequency of genomic alterations to some extent.

3.5 Establishment and evaluation of a nomogram

The univariate Cox regression and multivariate Cox regression analyses were performed to identify independent prognostic factors in the TCGA (Figures 6A, B), CGGA693, CGGA325, CGGA301, and Rembrandt cohorts (Supplementary Figures S8A–H). As we expected, SERS including nine selected genes was confirmed as an independent prognostic factor in all cohorts. The nomogram was established to predict 1-, 3-, and 5-year survival time in LGG patients based on the independent prognostic factors (age, WHO grade, and SERS) identified in the TCGA cohort (Figure 6C). The nomogram was firstly internally assessed, and the C-index was 0.862 (95% CI: 0.811–0.896), 0.833 (95% CI: 0.786–0.896), 0.812 (95% CI: 0.761–0.856) at 1, 3, and 5 years, respectively. The 1-year, 3-year and 5-year ROC curves showed that compared with SERS or age, the nomogram had the highest AUC values with the 1-year, 3-year and 5-year AUC values were 0.911, 0.913, and 0.812, respectively, which indicating that the nomogram had the optimal prediction effect (Figures 6D–F). The calibration curves showed a good fit between the actual and nomogram-predicted results for 1-, 3-, and 5-year OS (Figure 6G). In the same way, external validation of this nomogram was performed in the CGGA693, CGGA325, CGGA301 and Rembrandt cohorts. The accuracy in predicting 1-, 3-, and 5-year survival was good, and calibration curve analysis showed that the predicted and actual outcomes were basically conformity in all 4 cohorts (Supplementary Figures S9A–P). Therefore, this nomogram has potential as a quantitative predictor of prognosis in LGG patients.

3.6 Correlation of SERS with the LGG immune microenvironment and immunotherapy

Based on the differential expression analysis of high and low risk groups in the TCGA cohort, there were 462 DEGs (|log2FC|>2 and the BH method adjusted $p < 0.05$.) We then further performed functional enrichment analysis to characterize the biological functions of DEGs between the two risk subgroups. The results of GO analysis revealed that DEGs are enriched in several immune-related biological processes, such as regulation of T cell activation, positive regulation of T cell activation, negative regulation of immune system process, and positive regulation of lymphocyte activation (Figure 7A). Following, KEGG pathway analysis also showed significant enrichment of immune-related pathways, including cytokine-cytokine receptor interactions and chemokine signaling pathways (Figure 7B).

Given the findings that DEGs are enriched in immune-related functions, we further investigated the correlation of SERS with the immune microenvironment of LGG in the TCGA cohort. It turned out that SERS was significantly positively correlated with immune, stromal, and ESTIMATE scores, but negatively correlated with tumor purity, suggesting that the infiltration levels of immune cells and stromal cells increased with SERS (Figures 7C–F). The distribution of ESTIMATE scores, immune scores, stromal scores and tumor purity were no significant between high and low-risk groups in TCGA cohort (Supplementary Figures S10G–J). Further correlation analysis was performed between SERS and the infiltration levels of 22 immune cells quantified by the CIBERSORT algorithm. The results showed that the expressions of cells CD4 memory resting, T cells CD4 memory activated, NK cells activated, Monocytes, Macrophages M1, Mast cells activated and Neutrophils were significantly different in high and low risk groups. Among of them, the abundance of T cells CD4 memory resting, Macrophages M1, and Neutrophils was lower in the high-risk group, but the abundance of T cells CD4 memory activated, NK cells activated, Mast cells activated and Monocytes was higher in the high-risk group (Figure 7G).

In addition, we evaluated the correlation of SERS with immune checkpoints (PD-1, PD-L1, LAG-3, and B7-H3) and macrophage-associated molecules (CCL2, CCR2, CXCR4, and CSF1). The results showed that all immune checkpoints and macrophage-associated molecules were upregulated in the high-risk group except for LAG-3 (Figure 8A). We next determined whether there is a correlation between immune checkpoints and prognostic SERGs. The heat map showed that immune checkpoint proteins were significantly positively correlated with CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2, and significantly negatively correlated with AQP7 and MYOD1 (Figure 8B). In the TCGA cohort, SERS was positively correlated with TIDE and T-cell exclusion score, and negatively correlated with MSI score and T-cell dysfunction score (Figure 8C). The distribution difference can also be clearly observed in the high and low risk groups (Supplementary Figures S10K–N). In view of the TIDE algorithm, the distribution of SERS for the non-responder and responder groups to ICI indicated that the non-responder group had a significantly higher SERS, which just happened to explain the poorer prognosis of LGG patients who did not respond to ICI (Figure 8D). The high SERS subgroup had a lower proportion of responders to ICI treatment compared with the low SERS subgroup ($p < 0.05$, Figure 8E). Figure 8F showed that SERS had a satisfactory prediction in immunotherapy effect, which can provide a reference for whether patients should undergo immunotherapy.

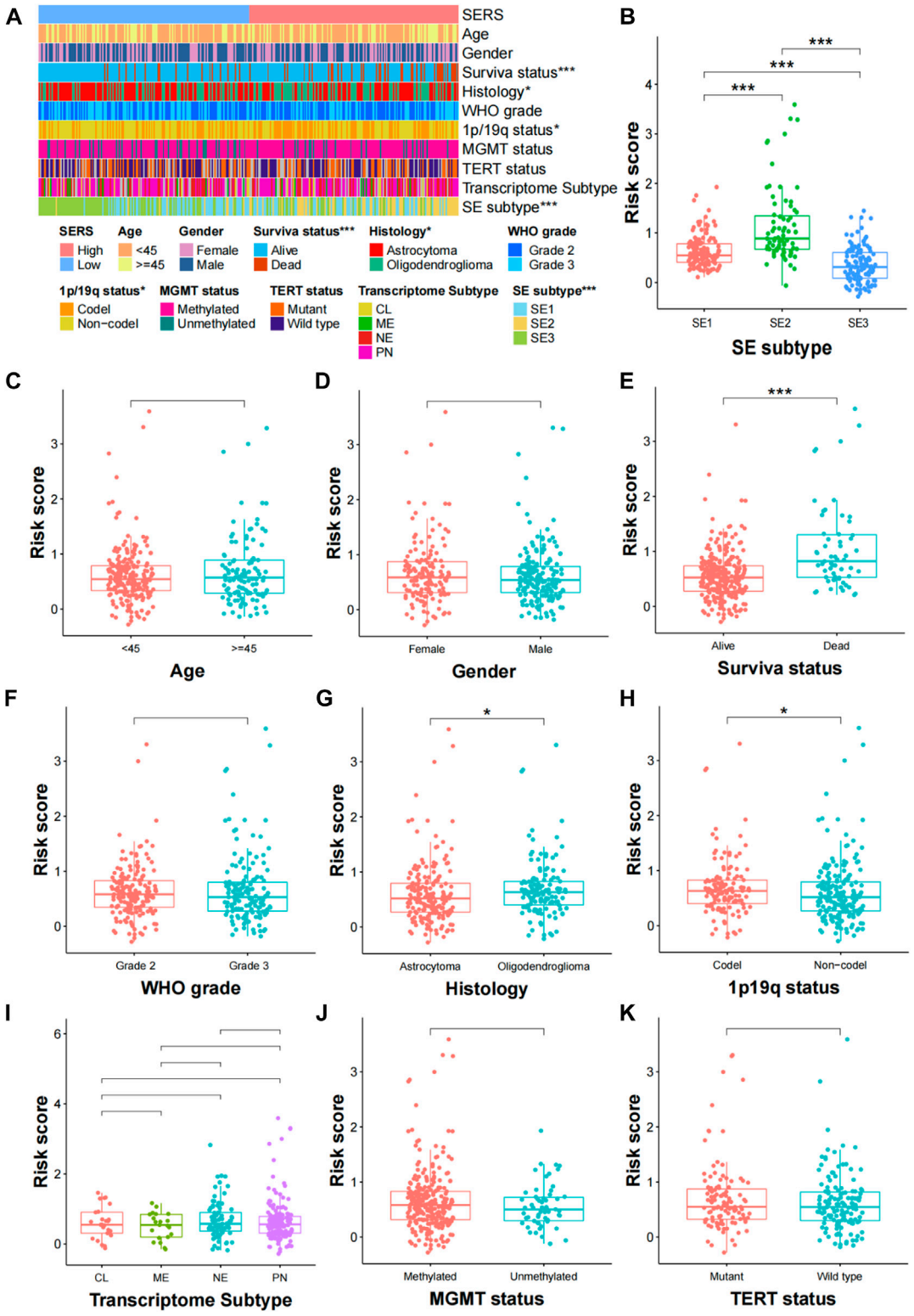


FIGURE 4
Correlation analysis between the prognostic SERS and clinicopathological characteristics in the TCGA cohort. **(A)** A heatmap was represented expression levels of eight selected SERS and the distribution of clinicopathological characteristics in the high- and low-risk groups, respectively. **(B–K)** Different levels of risk scores in glioma patients stratified by age, gender, Survival status, WHO grade, Histology, 1p19q codeletion, MGMT status, SERS subtype, TERT status and Transcriptome subtype. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

10.3389/fgene.2023.1085584

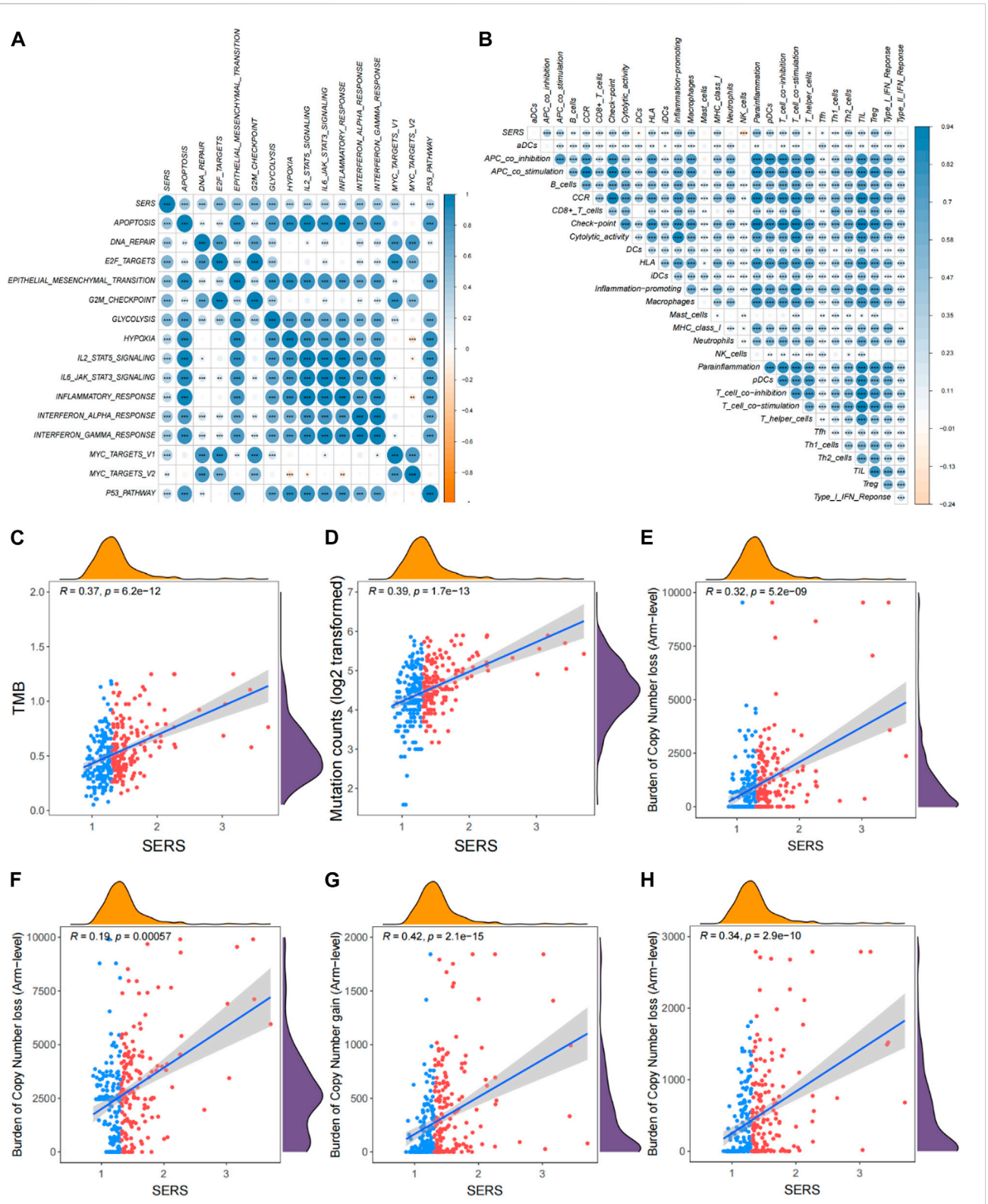


FIGURE 5
Correlation of SERS with clinicopathological features, genomic alterations and TIME patterns in TCGA cohort. **(A)** Correlation between SERS and the known cancer hallmarks of LGG patients. **(B)** Correlation between SERS and the 29 immune signatures. **(C–H)** Correlation of SERS with TMB, mutation counts, and copy number burdens at focal and arm levels. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

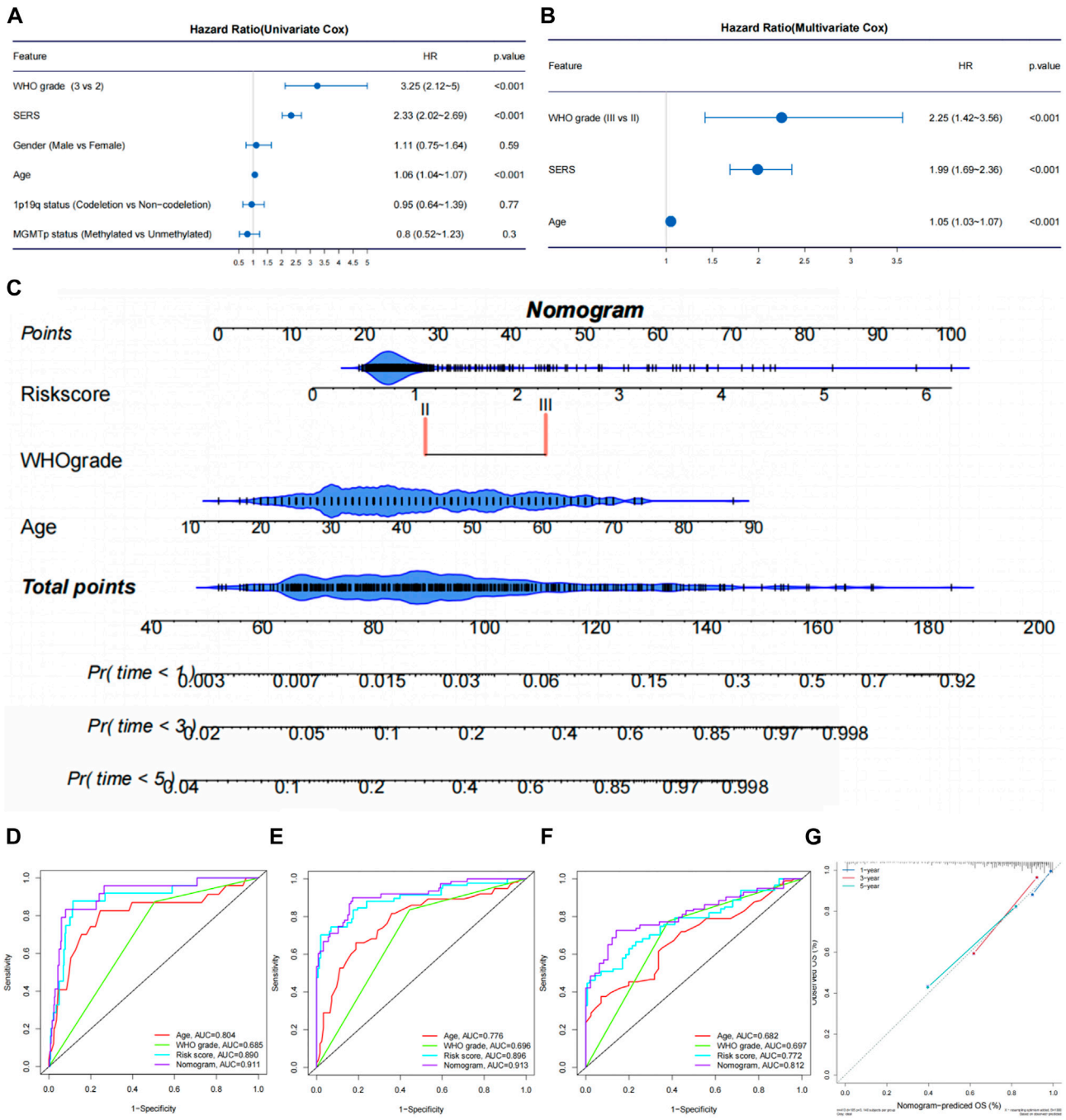


FIGURE 6
Establishment and evaluation of a nomogram in the TCGA cohort. **(A, B)** The univariate Cox regression and multivariate Cox regression were performed on SERS and other clinicopathological features TCGA cohort. **(C)** Nomogram based on SERS, WHO grade and age. **(D–F)** The receiver operating characteristic (ROC) curves of the nomogram predicted 1-, 3-, and 5-year OS in the TCGA cohort, respectively. **(G)** Calibration curves showed the good consistency between predicted and observed 1-, 3-, and 5-year overall survival (OS) in the TCGA cohort.

3.7 The expression levels of selected SE-related genes

Two SE-related genes (AQP7, and E2F7) were selected to detect their transcriptional levels in cell lines, LGG tissues and normal brain tissues. The qRT-PCR results showed that compared with HA1800, the mRNA expression levels of AQP7 in human glioma cell lines were generally decreased, while the mRNA expression levels of E2F7 were generally increased (Figure 9A). Subsequently, we also detected their expression levels in 10 normal brain tissues and 10 glioma tissues. The qRT-PCR results of the tissue samples were consistent with those of the cell lines (Figure 9B). The representative IHC staining images of AQP7 and E2F7were shown in Figure 9C.

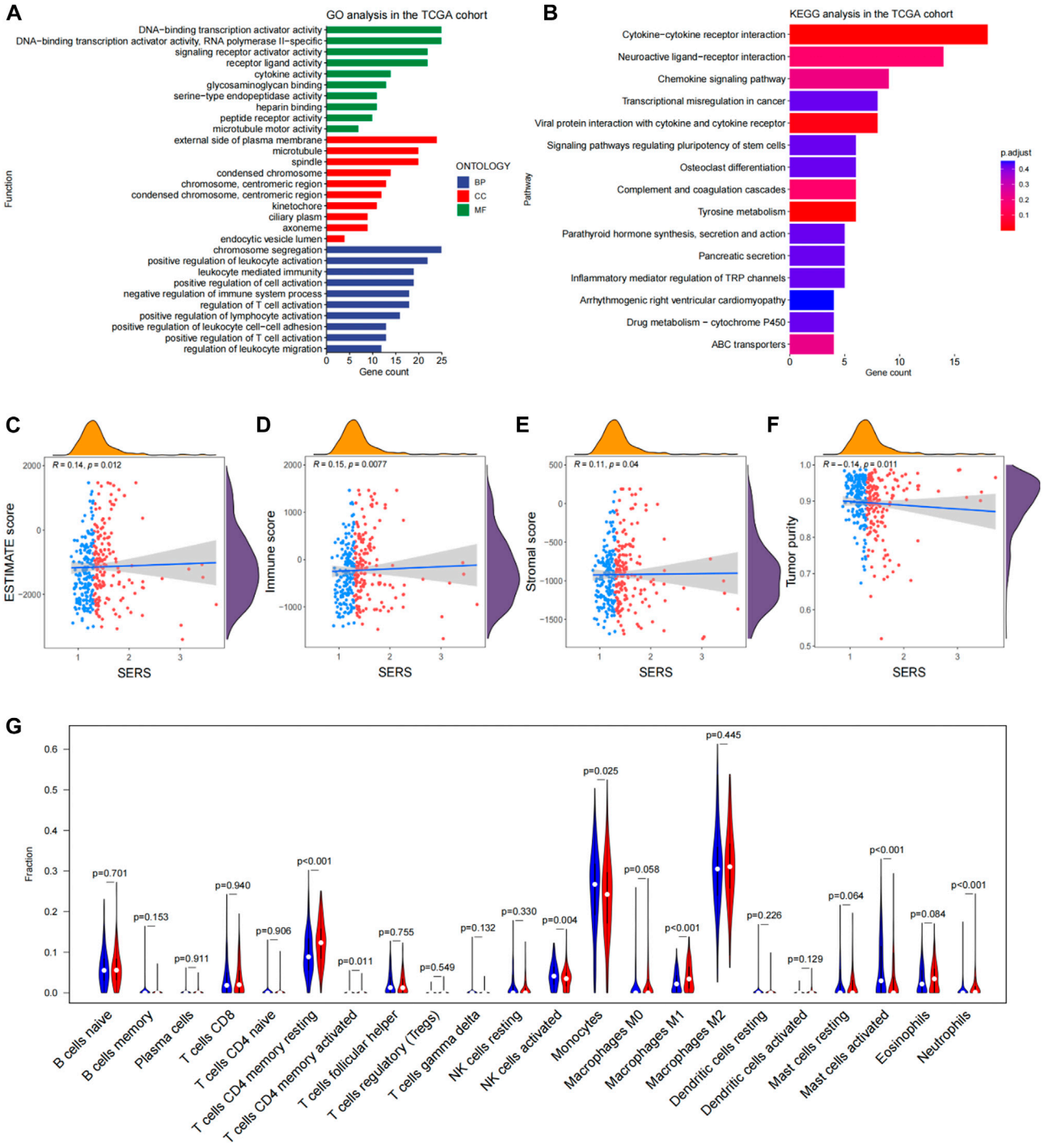


FIGURE 7
Functional enrichment analysis and immune landscape of glioma microenvironmental in the TCGA cohort. **(A, B)** Go analysis and KEGG analysis in the TCGA cohort. **(C–F)** Correlation of SERS with immune scores, stromal scores, ESTIMATE scores, tumor purity and SERS. **(G)** The abundance of 22 immune cells in the high- and low-risk groups. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

In addition, the relations between the selected gene (AQP7 and E2F7) and tumor immune features also analyzed. It turned out that AQP7 was significantly positively correlated with immune, stromal, and ESTIMATE scores, but negatively correlated with tumor purity (Figure 10A). While E2F7 was not significantly related with immune, stromal, and ESTIMATE scores, and tumor purity (Figure 10C). Further correlation analysis was also performed between the selected gene and the infiltration levels of 22 immune cells. The results showed that the expressions of 22 immune cells were significantly different whatever in high and low expression of AQP7 or E2F7 groups (Figures 10B, D).

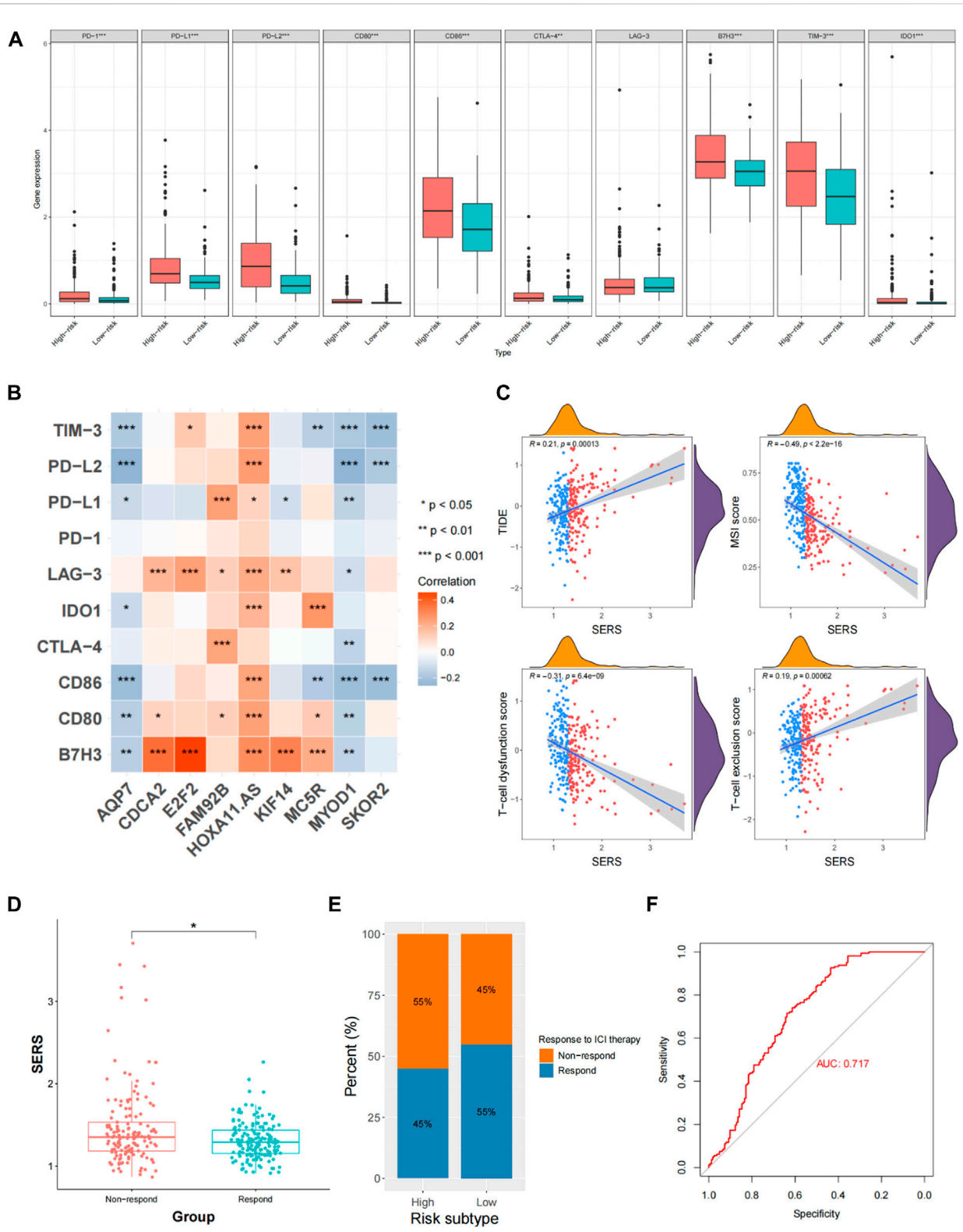


FIGURE 8
Evaluation of immune checkpoints and immunotherapy responsiveness in the TCGA cohort. **(A)** The expression levels of immune checkpoints and macrophage associated molecules in the high- and low-risk groups. **(B)** Correlation analysis between the prognostic SERS and immune checkpoints. **(C)** Correlation of SERS with T-cell dysfunction score, TIDE, MSI score and T-cell exclusion score. **(D)** The distributions of risk scores between non-respond and respond groups. **(E)** Comparative analysis of the response rates to ICI treatment in the high- and low-risk groups. **(F)** The ROC curve of predicting immunotherapeutic benefit.

4 Discussion

LGG patients, with the better prognosis than GBM patients, account for about half of all glioma patients. But their survival time varies widely, ranging from 1 year to more than 10 years (Chen et al., 2022), notwithstanding LGG patients with the same WHO grade had same standardized sequential therapy including surgery, radiotherapy, and chemotherapy. The high heterogeneity of LGG, which results in inconsistent treatment effects and prognosis, is a clinical conundrum faced by most neurosurgeons. And so, there is an urgent need to develop accurate and robust prognostic prediction models for data-assisted clinical decision-making. With the rapid development of bioinformatics and sequencing technologies, some studies have reported gene markers as prognostic indicators to predict the prognosis of LGG, such as hypoxia-related genes (Dao et al., 2018), ferroptosis-related genes (Wan et al., 2021), immune-related genes (Zhou et al., 2018) and the corresponding lncRNAs. Compared with other biomarkers, SE, as important distal regulatory DNA elements, are direct drivers of carcinogenesis and are highly tissue-specific. Therefore, SE are good candidates for predicting prognosis in various cancers. Ropri et al. (2021) and Huang et al. (2022) performed mechanistic exploration and prognostic prediction in breast cancer and hepatocellular carcinoma, respectively. However, whether super-enhancer related genes can serve as prognostic markers for LGG needs further discussion.

In this study, the NMF algorithm was used to identify three LGG subtypes in 325 LGG patients based on the expression profiles of DEGs between LGG and NBT. Then, significant differences in prognostic, clinicopathological features of the three LGG subtypes were observed with the naked eyes. A prognostic signature, called SERS, was constructed by univariate Cox regression and LASSO Cox regression for an individualized comprehensive assessment. The results showed that SERS was significantly associated with the prognosis, clinicopathological features, genomic alterations and TIME pattern of LGG patients, and the predictive ability of SERS for ICI treatment was also outstanding. In addition, a clinically accessible nomogram was constructed based on SERS, age, and WHO classification, which maintained excellent predictive accuracy in both the internal cohort and 4 external cohorts (CGGA693, CGGA325, CGGA301, and Rembrandt). So, it can provide a good net clinical benefit for screening LGG patients at high risk of death.

The SERS was constructed on 9 SE-related genes in our study, incorporating AQP7, MYOD1, CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2. Among these genes, CDCA2, FAM92B, HOXA11-AS, E2F7, KIF18A, MC5R, and SKOR2 were risky genes, which are associated with poor prognosis for LGG patients. Whereas the remaining two genes with good prognosis. Conversely, AQP7 and MYOD1 are related to good prognosis. AQP7, named Aquaporin 7, is a water and glycerol channel. Chen et al. demonstrated that low expression of AQP7 correlates with tumor grade and aggressive features of hepatocellular carcinoma (Chen et al., 2016). In a mouse model of breast cancer, lower AQP7 expression resulted in a reduction in primary tumor burden and lung metastases, thus suggesting that AQP7 is a prognostic indicator of overall survival in breast cancer patients (Dai et al., 2020). Myogenic differentiation 1 (MYOD1), as a transcription factor, promoted expression of muscle-specific genes. Wu et al. (2020) found that the expression of MYOD1D is positively correlated with the migration and invasion of gastric cancer cells. The cell division cycle associated 2 (CDCA2) has been proved to play an important role in the tumorigenesis of some cancers. The study showed that the high expression of CDCA2 was significantly correlated with the expression of related components of cell cycle phase transition and G2/M phase transition pathway, and suggested that CDCA2 could be a potential target for regulating tumor growth and radiation resistance in patients with oesophageal square cell carcinoma (Xu et al., 2021b). FAM92B, HOXA11-AS, E2F7, MC5R, and SKOR2 are important epigenetic regulators that can be targeted for cancer therapy. Specifically, E2F7 is an atypical E2F transcription factor family member with two independent DNA-binding domains. Some studies have found that E2F7 is upregulated in endometrial cancer, skin squamous cell carcinoma and other malignant tumors, promote tumor progression and metastasis in these cancers (Endo-Munoz et al., 2009; Li et al., 2015). KIF18A, a member of the kinesin-8 subfamily, has low expression in most human normal tissues and abnormally high expression in a variety of malignant tumor tissues (Marquis et al., 2021), which is associated with malignant pathological features and poor prognosis of cancer patients, and it promotes the proliferation, invasion and metastasis of tumor cells (Sepaniac et al., 2021). KIF18A may be a novel molecular targeted therapy for cancers. PTCRA (pre-T cell antigen receptor) is a protein-coding gene, together with the TCRB and CD3 complexes, encodes a protein that forms the T-cell pre-receptor complex, which regulates early T cell development (Liu et al., 2010).

The SERS is very effective in predicting prognosis of LGG patients in this study. To further explore the specific mechanism, we identified DEGs in high- and low-risk groups without hesitation. Then, the GO and KEGG analysis were performed to explore the detailed biological processes and pathways of these genes affecting the prognosis of LGG patients. Functional enrichment analysis revealed that DEGs between different risk subgroups were enriched in many immune-related biological processes and pathways. Therefore, we subsequently analyzed immune scores and immune cell infiltration between the two risk subgroups. Further analysis found that high risk was positively correlated with immune score, the abundance of T cells CD4 memory resting, and T cells CD4 memory activated. On the contrary, activated NK cells (tumor killer cells) showed higher abundance in the low-risk group. These results suggest that super-enhancer-related genes are related to the LGG immune microenvironment to a certain extent. From the above results, it can be concluded that the anti-tumor immunity of LGG patients in the high-risk group is significantly weakened, so we speculated that this may be one of the important reasons for their poor prognosis. The research of cancer immunotherapy has been very hot in recent years, especially immune checkpoint inhibitors have been quite mature as the first generation of immunotherapy, they play a therapeutic role in various cancers through mainly blocking PD-1/PD-L1 pathway and molecular receptors and/or ligands such as CTLA-4 (Topalian et al., 2015). Several previous studies have described therapeutic effect for immune checkpoints in some cancers, with findings consistent with favorable clinical outcomes in patients with many cancers, such as glioma (Puigdelloses et al., 2021), hepatocellular carcinoma (Sangro et al., 2020), lung cancer (Kartolo et al., 2021), and more. Therefore, we also evaluated the relationship between SERS and the expression levels of immune

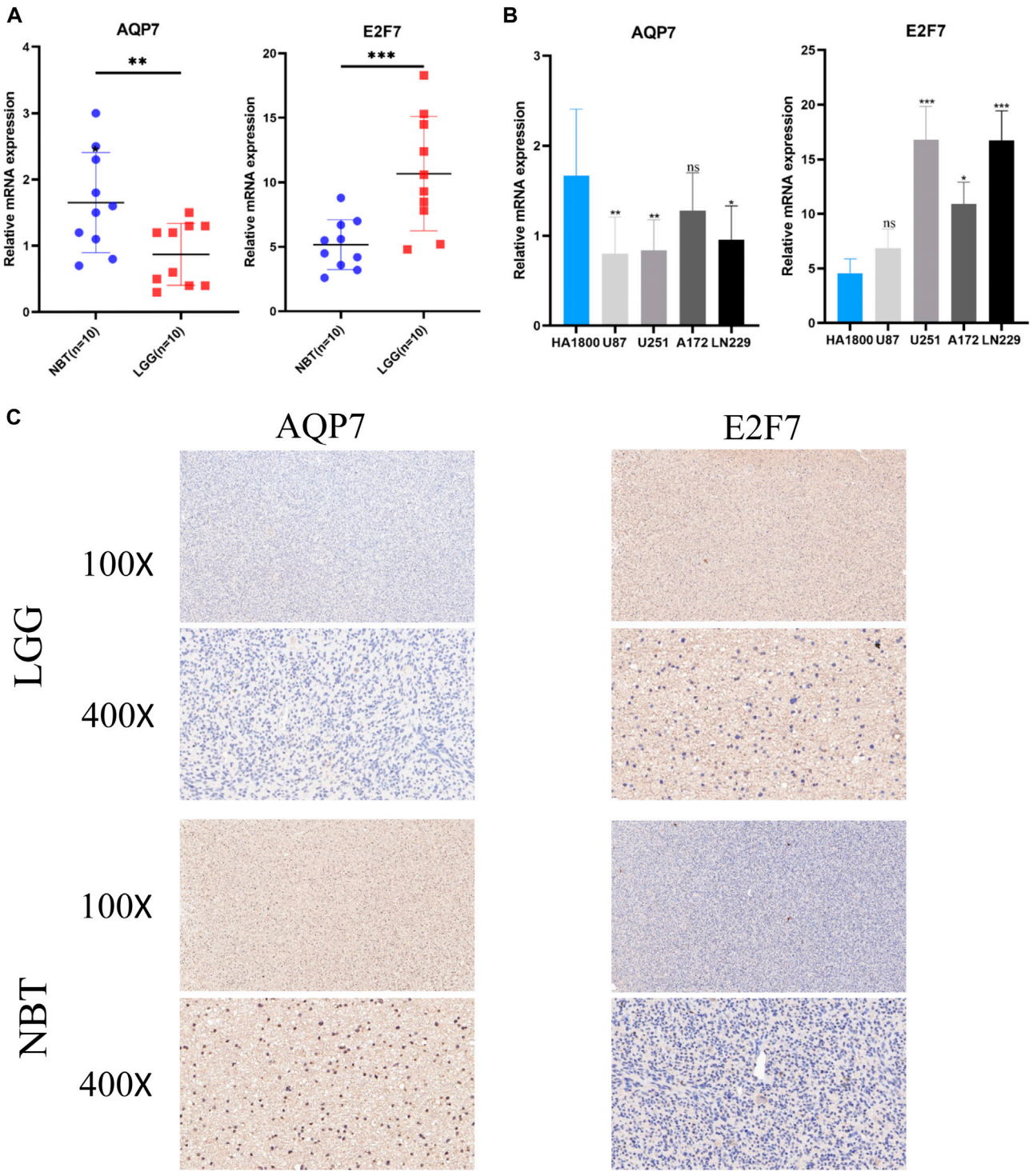


FIGURE 9
Validation of the expression levels of selected super-enhancer related genes. **(A)** Scatter plots of differential transcript levels between AQP7 and E2F7 in glioma cell lines and normal human astrocytes cell lines (HA1800). **(B)** Scatter plots of differential transcript levels between AQP7 and E2F7 in LGG and NBT. **(C)** The representative IHC staining images of AQP7 and E2F7. LGG low-grade glioma, NBT non-tumor tissues. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

checkpoints, macrophage-related molecules, and immunotherapy response. It found that SERS was positively correlated with the expression levels of immune checkpoints and macrophage-related molecules. The response rate to ICI was significantly lower than that of the low-risk group. Therefore, we surmised that this may be another reason for the poor prognosis of LGG patients in the high-

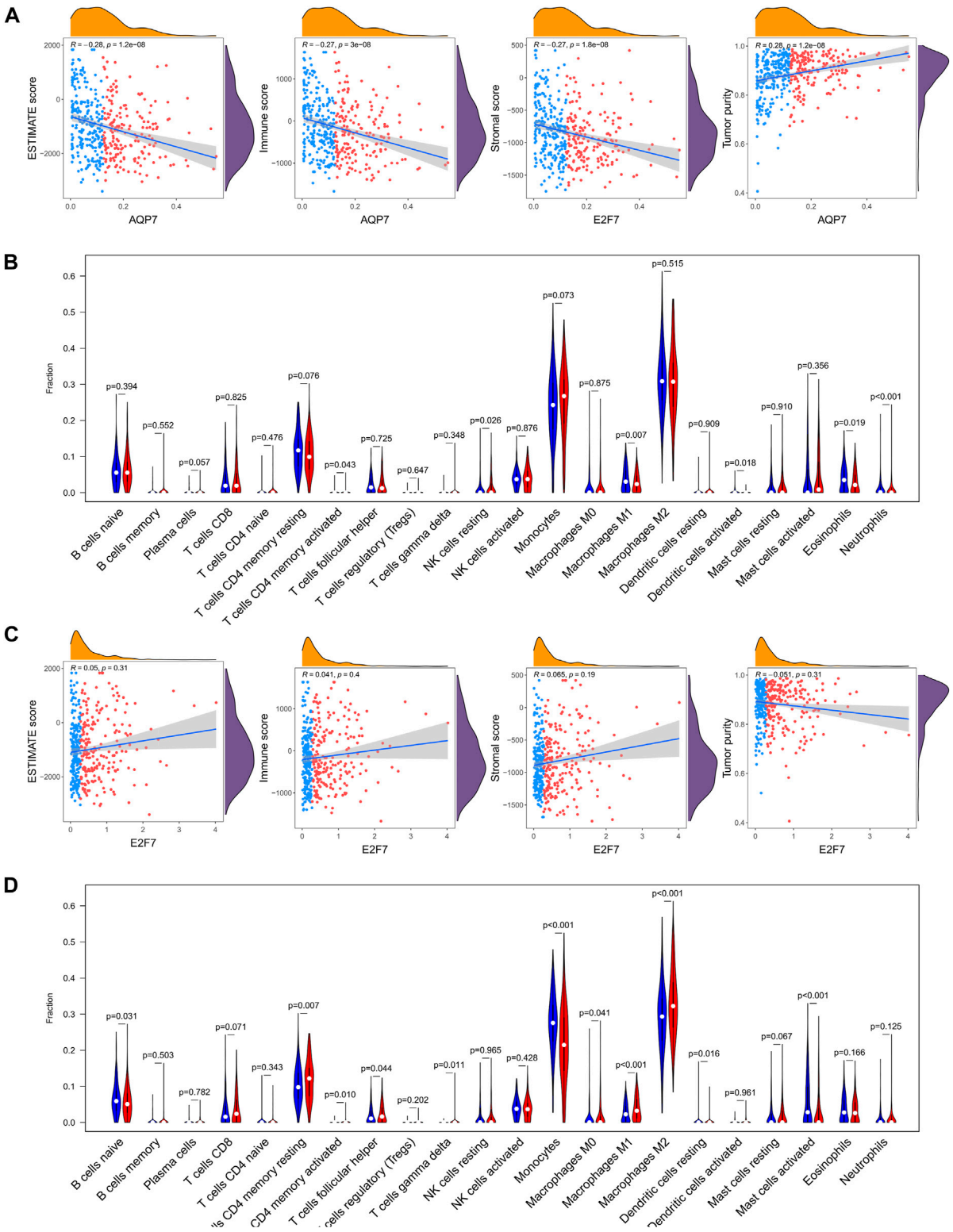


FIGURE 10
The relations between the selected gene and tumor immune features. **(A)** The abundance of 22 immune cells in the high-expression and low-expression of AQP7. **(B)** Correlation of the expression of AQP7 with immune scores, stromal scores, ESTIMATE scores, and tumor purity. **(C)** The abundance of 22 immune cells in the high-expression and low-expression of E2F7. **(D)** Correlation of the expression of E2F7 with immune scores, stromal scores, ESTIMATE scores, and tumor purity. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

risk group. Taken together, the SERS proposed in our study can be used to screen clinically high-risk LGG patients, and then to prescribe professionally-informed treatment.

Without doubt, there are some inevitable shortcomings in this study. Firstly, this is a retrospective study based on public databases, thus yielding more reliable results in a prospective study. Secondly, the five cohorts of LGG patients have varying degrees of lack of clinical information, which may lead to varying degrees of selection bias. Thirdly, GO and KEGG enrichment analysis and subsequent immune microenvironment and immune checkpoint analysis were not validated in the other 4 cohorts. Fourthly, we only analyzed transcriptome information and did not perform multi-omics analysis including methylation and gene copy number. Finally, further experiments are needed to explore the specific molecular mechanism of super-enhancer related genes for further elucidation.

5 Conclusion

In conclusion, three novel LGG subtypes were established based on SE-related genes. Subsequently, an accurate and independently validated model were proposed for predicting overall survival in LGG. In addition, we also found that SERS was associated with prognosis, clinicopathological features, tumor immune microenvironment, cancer hallmarks, and genomic alterations and the effect of immunotherapy in patients with LGG. The findings can be as the novel biomarkers for predicting prognosis and potential therapeutic targets for LGG, which will help physicians and patients to evaluate prognosis, determine follow-up period, and make immunotherapy decisions.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Wuhan University of Science and Technology Affiliated Xiaogan Central Hospital. The ethics committee waived the requirement of written informed consent for participation.

Author contributions

Conception and design: SF. Data collection and analysis and interpretation: YH, QY, and WW. Language editing and grammar correction: SC. Manuscript writing: YH and QY. Final approval of the manuscript: All authors.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1085584/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Comparisons of age, gender, KPS, survival status, histology, WHO grade, MGMT status, Transcriptome subtype and TERT status among SE subtypes in TCGA cohord.

SUPPLEMENTARY FIGURE S2
The Kaplan–Meier curves of TCGA cohort showed that there is different overall survival for patients with different expression levels of nine selected super-enhancer related prognostic genes.

SUPPLEMENTARY FIGURE S3
Risk scores, survival status in high and low-risk group and Kaplan–Meier survival curves of SERS in TCGA, CGGA693, CGGA325, CGGA301 and Rembrandt cohorts, respectively.

SUPPLEMENTARY FIGURE S4
Correlation analysis between the prognostic SERS and clinicopathological characteristics in the CGGA693 cohort. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

SUPPLEMENTARY FIGURE S5
Correlation analysis between the prognostic SERS and clinicopathological characteristics in the CGGA325 cohort. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

SUPPLEMENTARY FIGURE S6
Correlation analysis between the prognostic SERS and clinicopathological characteristics in the CGGA301 cohort. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

SUPPLEMENTARY FIGURE S7
Correlation analysis between the prognostic SERS and clinicopathological characteristics in the Rembrandt cohort. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns No significance.

SUPPLEMENTARY FIGURE S8
The univariate Cox regression and multivariate Cox regression were performed on SERS and other clinicopathological features CGGA693, CGGA325, CGGA301, Rembrandt cohorts, respectively.

SUPPLEMENTARY FIGURE S9
The ROC curves of the nomogram predicted 1-, 3-, and 5-year OS and corresponding calibration curves in CGGA693, CGGA325, CGGA301, Rembrandt cohorts, respectively.

SUPPLEMENTARY FIGURE S10
The distribution of TMB, mutation counts, copy number burdens at focal and arm levels, immune scores, stromal scores, ESTIMATE scores and tumor purity, TIDE, T-cell exclusion score, MSI score and T-cell dysfunction score between high and low-risk groups in TCGA cohort. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$.

References

Bi, F., Chen, Y., and Yang, Q. (2020). Significance of tumor mutation burden combined with immune infiltrates in the progression and prognosis of ovarian cancer. *Cancer Cell. Int.* 20, 373. doi:10.1186/s12935-020-01472-9

Brat, D. J., Verhaak, R. G., Aldape, K. D., Yung, W. K., Salama, S. R., Cooper, L. A., et al. (2015). Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med.* 372, 2481–2498. doi:10.1056/NEJMoa1402121

Chen, X. F., Li, C. F., Lu, L., and Mei, Z. C. (2016). Expression and clinical significance of aquaglyceroporins in human hepatocellular carcinoma. *Mol. Med. Rep.* 13, 5283–5289. doi:10.3892/mmr.2016.5184

Chen, S., Pu, J., Bai, J., Yin, Y., Wu, K., Wang, J., et al. (2018). Ezh2 promotes hepatocellular carcinoma progression through modulating mir-22/galectin-9 axis. *J. Exp. Clin. Cancer Res.* 37, 3. doi:10.1186/s13046-017-0670-6

Chen, J., Shen, S., Li, Y., Fan, J., Xiong, S., Xu, J., et al. (2022). Apollo: An accurate and independently validated prediction model of lower-grade gliomas overall survival and a comparative study of model performance. *Ebiomedicine* 79, 104007. doi:10.1016/j.ebiom.2022.104007

Dai, C., Charlestin, V., Wang, M., Walker, Z. T., Miranda-Vergara, M. C., Facchine, B. A., et al. (2020). Aquaporin-7 regulates the response to cellular stress in breast cancer. *Cancer Res.* 80, 4071–4086. doi:10.1158/0008-5472.CAN-19-2269

Dao, T. P., Rosch, S., Mairbaurl, H., Pusch, S., Unterberg, A., Herold-Mende, C., et al. (2018). Identification of a prognostic hypoxia-associated gene set in idh-mutant glioma. *Int. J. Mol. Sci.* 19, 2903. doi:10.3390/ijms19102903

Endo-Munoz, L., Dahler, A., Teakle, N., Rickwood, D., Hazar-Rethinam, M., Abdul-Jabbar, I., et al. (2009). E2f7 can regulate proliferation, differentiation, and apoptotic responses in human keratinocytes: Implications for cutaneous squamous cell carcinoma formation. *Cancer Res.* 69, 1800–1808. doi:10.1158/0008-5472.CAN-08-2725

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

Gittleman, H., Sloan, A. E., and Barnholtz-Sloan, J. S. (2020). An independently validated survival nomogram for lower-grade glioma. *Neuro Oncol.* 22, 665–674. doi:10.1093/neuonc/noz191

Hillman, R. T., Chisholm, G. B., Lu, K. H., and Futreal, P. A. (2018). Genomic rearrangement signatures and clinical outcomes in high-grade serous ovarian cancer. *J. Natl. Cancer Inst.* 110, 265–272. doi:10.1093/jnci/djx176

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-Andre, V., Sigova, A. A., et al. (2013). Super-enhancers in the control of cell identity and disease. *Cell.* 155, 934–947. doi:10.1016/j.cell.2013.09.053

Hu, J., Lai, Y., Huang, H., Ramakrishnan, S., Pan, Y., Ma, V., et al. (2022). Tcof1 upregulation in triple-negative breast cancer promotes stemness and tumour growth and correlates with poor prognosis. *Br. J. Cancer* 126, 57–71. doi:10.1038/s41416-021-01596-3

Hu, X., Ni, S., Zhao, K., Qian, J., and Duan, Y. (2022). Bioinformatics-led discovery of osteoarthritis biomarkers and inflammatory infiltrates. *Front. Immunol.* 13, 871008. doi:10.3389/fimmu.2022.871008

Huang, P., Zhang, B., Zhao, J., and Li, M. D. (2022). Integrating the epigenome and transcriptome of hepatocellular carcinoma to identify systematic enhancer aberrations and establish an aberrant enhancer-related prognostic signature. *Front. Cell. Dev. Biol.* 10, 827657. doi:10.3389/fcell.2022.827657

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of t cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1

Jiang, T., Nam, D. H., Ram, Z., Poon, W. S., Wang, J., Boldbaatar, D., et al. (2021). Clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett.* 499, 60–72. doi:10.1016/j.canlet.2020.10.050

Jiao, W., Chen, Y., Song, H., Li, D., Mei, H., Yang, F., et al. (2018). Hpse enhancer rna promotes cancer progression through driving chromatin looping and regulating hnrnpu/p300/egr1/hpse axis. *Oncogene* 37, 2728–2745. doi:10.1038/s41388-018-0128-0

Kartolo, A., Feilotter, H., Hopman, W., Fung, A. S., and Robinson, A. (2021). A single institution study evaluating outcomes of pd-l1 high kras-mutant advanced non-small cell lung cancer (nsclc) patients treated with first line immune checkpoint inhibitors. *Cancer Treat. Res. Commun.* 27, 100330. doi:10.1016/j.ctarc.2021.100330

Kenny, P. A., Nelson, C. M., and Bissell, M. J. (2006). The ecology of tumors: By perturbing the microenvironment, wounds and infection may be key to tumor development. *Scientist* 20, 30.

Lee, J. H., Xiong, F., and Li, W. (2020). Enhancer rnas in cancer: Regulation, mechanisms and therapeutic potential. *Rna Biol.* 17, 1550–1559. doi:10.1080/15476286.2020.1712895

Li, Q., Qiu, X. M., Li, Q. H., Wang, X. Y., Li, L., Xu, M., et al. (2015). Microrna-424 may function as a tumor suppressor in endometrial carcinoma cells by targeting e2f7. *Oncol. Rep.* 33, 2354–2360. doi:10.3892/or.2015.3812

Liu, H., Chi, A. W., Arnett, K. L., Chiang, M. Y., Xu, L., Shestova, O., et al. (2010). Notch dimerization is required for leukemogenesis and t-cell development. *Genes. Dev.* 24, 2395–2407. doi:10.1101/gad.1975210

Liu, X., Li, Y., Qian, Z., Sun, Z., Xu, K., Wang, K., et al. (2018). A radiomic signature as a non-invasive predictor of progression-free survival in patients with lower-grade gliomas. *Neuroimage Clin.* 20, 1070–1077. doi:10.1016/j.nicl.2018.10.014

Lou, J., Hao, Y., Lin, K., Lyu, Y., Chen, M., Wang, H., et al. (2020). Circular rna cdr1as disrupts the p53/mdm2 complex to inhibit gliomagenesis. *Mol. Cancer* 19, 138. doi:10.1186/s12943-020-01253-y

Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., et al. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* 346, 1373–1377. doi:10.1126/science.1259037

Marquis, C., Fonseca, C. L., Queen, K. A., Wood, L., Vandal, S. E., Malaby, H., et al. (2021). Chromosomally unstable tumor cells specifically require kif18a for proliferation. *Nat. Commun.* 12, 1213. doi:10.1038/s41467-021-21447-2

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

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

Ostrom, Q. T., Price, M., Ryan, K., Edelson, J., Neff, C., Cioffi, G., et al. (2022). Cbtrus statistical report: Pediatric brain tumor foundation childhood and adolescent primary brain and other central nervous system tumors diagnosed in the United States in 2014-2018. *Neuro Oncol.* 24, i1–i38. doi:10.1093/neuonc/noac161

Peng, L., Jiang, B., Yuan, X., Qiu, Y., Peng, J., Huang, Y., et al. (2019). Super-enhancer-associated long noncoding rna hccl5 is activated by zeb1 and promotes the malignancy of hepatocellular carcinoma. *Cancer Res.* 79, 572–584. doi:10.1158/0008-5472.CAN-18-0367

Puigdelloses, M., Garcia-Moure, M., Labiano, S., Laspidea, V., Gonzalez-Huarriz, M., Zalacain, M., et al. (2021). Cd137 and pd-l1 targeting with immunovirotherapy induces a potent and durable antitumor immune response in glioblastoma models. *J. Immunother. Cancer* 9, e002644. doi:10.1136/jitc-2021-002644

Ropri, A. S., DeVaux, R. S., Eng, J., Chittur, S. V., and Herschkowitz, J. I. (2021). Cis-acting super-enhancer lncrnas as biomarkers to early-stage breast cancer. *Breast Cancer Res.* 23, 101. doi:10.1186/s13058-021-01479-8

Sangro, B., Chan, S. L., Meyer, T., Reig, M., El-Khoueiry, A., and Galle, P. R. (2020). Diagnosis and management of toxicities of immune checkpoint inhibitors in hepatocellular carcinoma. *J. Hepatol.* 72, 320–341. doi:10.1016/j.jhep.2019.10.021

Sepaniac, L. A., Martin, W., Dionne, L. A., Stearns, T. M., Reinholdt, L. G., and Stumpff, J. (2021). Micronuclei in kif18a mutant mice form stable micronuclear envelopes and do not promote tumorigenesis. *J. Cell. Biol.* 220, e202101165. doi:10.1083/jcb.202101165

Shen, R., Li, P., Li, B., Zhang, B., Feng, L., and Cheng, S. (2019). Identification of distinct immune subtypes in colorectal cancer based on the stromal compartment. *Front. Oncol.* 9, 1497. doi:10.3389/fonc.2019.01497

Topalian, S. L., Drake, C. G., and Pardoll, D. M. (2015). Immune checkpoint blockade: A common denominator approach to cancer therapy. *Cancer Cell.* 27, 450–461. doi:10.1016/j.ccell.2015.03.001

Wan, R. J., Peng, W., Xia, Q. X., Zhou, H. H., and Mao, X. Y. (2021). Ferroptosis-related gene signature predicts prognosis and immunotherapy in glioma. *Cns Neurosci. Ther.* 27, 973–986. doi:10.1111/cns.13654

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., et al. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell.* 153, 307–319. doi:10.1016/j.cell.2013.03.035

Wu, F., Qin, Y., Jiang, Q., Zhang, J., Li, F., Li, Q., et al. (2020). Myod1 suppresses cell migration and invasion by inhibiting fut4 transcription in human gastric cancer cells. *Cancer Gene Ther.* 27, 773–784. doi:10.1038/s41417-019-0153-3

Xu, S., Tang, L., Liu, Z., Luo, C., and Cheng, Q. (2021). Hypoxia-related lncrna correlates with prognosis and immune microenvironment in lower-grade glioma. *Front. Immunol.* 12, 731048. doi:10.3389/fimmu.2021.731048

Xu, B., Chen, H., Xu, Z., Yao, X., Sun, X., and Cheng, H. (2021). Cdca2 promotes tumorigenesis and induces radioresistance in oesophageal squamous cell carcinoma cells. *Mol. Med. Rep.* 24, 530. doi:10.3892/mmr.2021.12169

Ye, L., Wang, L., Yang, J., Hu, P., Zhang, C., Tong, S., et al. (2021). Identification of tumor antigens and immune subtypes in lower grade gliomas for mrna vaccine development. *J. Transl. Med.* 19, 352. doi:10.1186/s12967-021-03014-x

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

Zakharova, G., Efimov, V., Raevskiy, M., Rumiantsev, P., Gudkov, A., Belogurova-Ovchinnikova, O., et al. (2022). Reclassification of tcga diffuse glioma profiles linked to transcriptomic, epigenetic, genomic and clinical data, according to the 2021 who cns tumor classification. *Int. J. Mol. Sci.* 24, 157. doi:10.3390/ijms24010157

Zhou, M., Zhang, Z., Zhao, H., Bao, S., Cheng, L., and Sun, J. (2018). An immune-related six-lncrna signature to improve prognosis prediction of glioblastoma multiforme. *Mol. Neurobiol.* 55, 3684–3697. doi:10.1007/s12035-017-0572-9

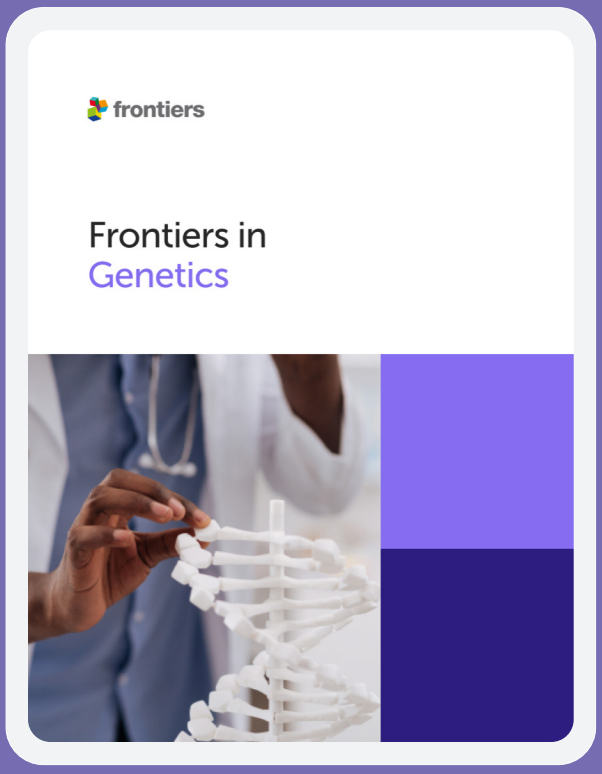
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact