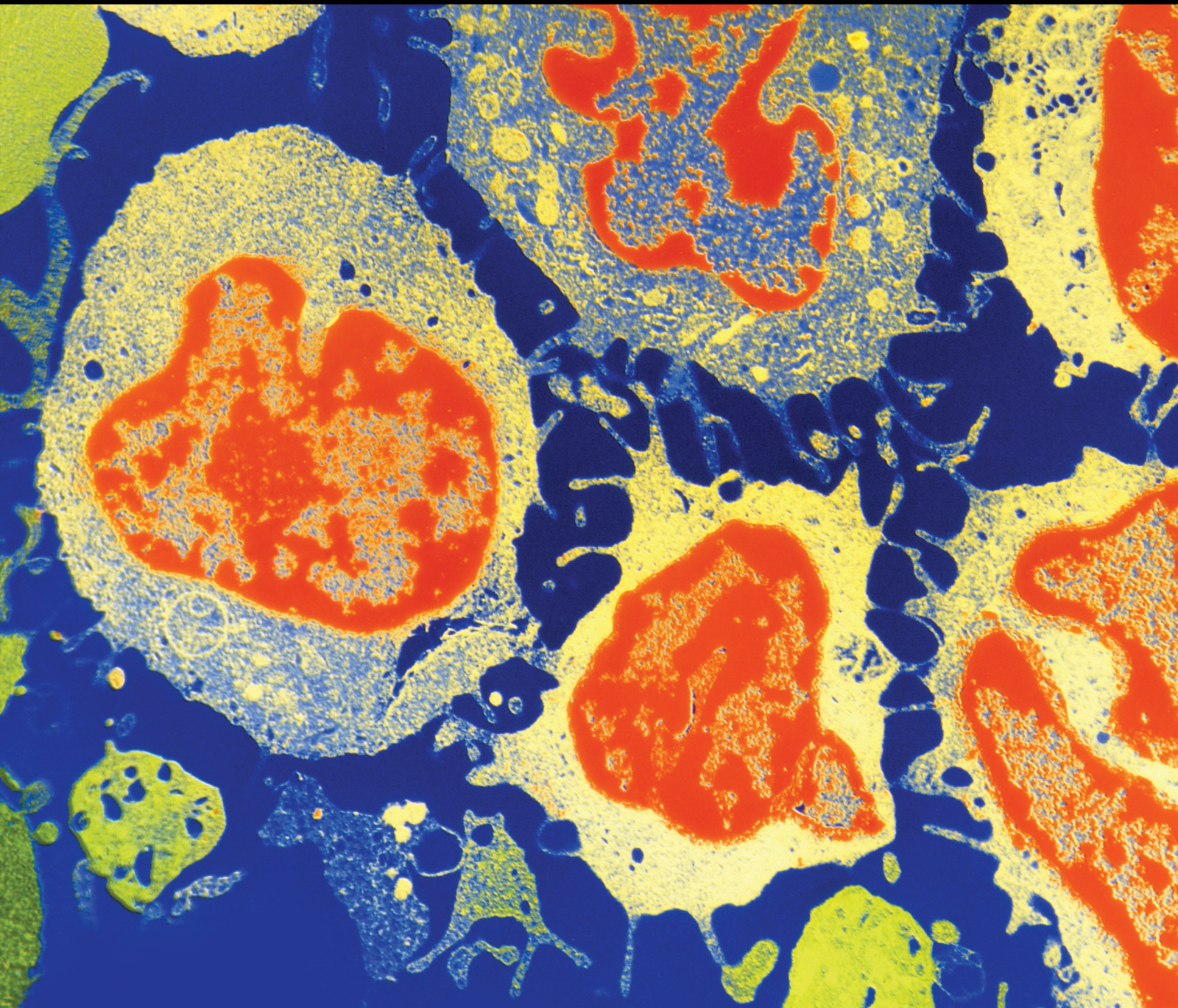


Bioinformatics in Cancer and Immune Microenvironment 2022

Lead Guest Editor: Feng Jiang

Guest Editors: Song Cao and Jian Song





**Bioinformatics in Cancer and Immune
Microenvironment 2022**

Journal of Oncology

**Bioinformatics in Cancer and Immune
Microenvironment 2022**

Lead Guest Editor: Feng Jiang

Guest Editors: Song Cao and Jian Song



Copyright © 2023 Hindawi Limited. All rights reserved.

This is a special issue published in "Journal of Oncology" All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chief Editor

Bruno Vincenzi, Italy

Academic Editors

Thomas E. Adrian, United Arab Emirates

Ruhai Bai , China

Jiaolin Bao, China


Rossana Berardi, Italy

Benedetta Bussolati, Italy

Sumanta Chatterjee, USA


Thomas R. Chauncey, USA

Gagan Chhabra, USA

Francesca De Felice , Italy

Giuseppe Di Lorenzo, Italy


Xiangya Ding , China

Peixin Dong , Japan

Xingrong Du, China

Elizabeth R. Dudnik , Israel

Pierfrancesco Franco , Italy


Ferdinand Frauscher , Austria

Rohit Gundamaraju, USA


Han Han , USA

Jitti Hanprasertpong , Thailand


Yongzhong Hou , China

Wan-Ming Hu , China


Jialiang Hui, China

Akira Iyoda , Japan

Reza Izadpanah , USA

Kaiser Jamil , India

Shuang-zheng Jia , China

Ozkan Kanat , Turkey

Zhijia Kang , USA

Pashtoon M. Kasi , USA

Jorg Kleeff, United Kingdom

Jayaprakash Kolla, Czech Republic

Goo Lee , USA


Peter F. Lenehan, USA


Da Li , China

Rui Liao , China

Rengyun Liu , China

Alexander V. Louie, Canada

Weiren Luo , China

Cristina Magi-Galluzzi , USA

Kanjoormana A. Manu, Singapore


Riccardo Masetti , Italy

Ian E. McCutcheon , USA

Zubing Mei, China

Giuseppe Maria Milano , Italy

Nabiha Missaoui , Tunisia

Shinji Miwa , Japan

Sakthivel Muniyan , USA

Magesh Muthu , USA

Nandakumar Natarajan , USA


P. Neven, Belgium


Patrick Neven, Belgium

Marco Noventa, Italy

Liren Qian , China

Shuanglin Qin , China

Dongfeng Qu , USA

Amir Radfar , USA

Antonio Raffone , Italy


Achuthan Chathrattil Raghavamenon, India

Faisal Raza, China

Giandomenico Roviello , Italy


Subhadeep Roy , India


Prasannakumar Santhekadur , India

Chandra K. Singh , USA

Yingming Sun , China


Mohammad Tarique , USA

Federica Tomao , Italy

Vincenzo Tombolini , Italy

Maria S. Tretiakova, USA

Abhishek Tyagi , USA

Satoshi Wada , Japan


Chen Wang, China

Xiaosheng Wang , China

Guangzhen Wu , China

Haigang Wu , China

Yuan Seng Wu , Malaysia


Yingkun Xu , China

WU Xue-liang , China

ZENG JIE YE , China

Guan-Jun Yang , China





Junmin Zhang , China

Dan Zhao , USA

Dali Zheng , China










Contents

Upregulation of APOC1 Promotes Colorectal Cancer Progression and Serves as a Potential Therapeutic Target Based on Bioinformatics Analysis

Weiwei Tang , Hanyuan Liu, Xiao Li , Theng Choon Ooi, Nor Fadilah Rajab, Hongyong Cao , and Razinah Sharif 

Research Article (13 pages), Article ID 2611105, Volume 2023 (2023)

TREM2 as a Prognostic Biomarker for Osteosarcoma Microenvironment Remodeling

Zhi-Long Shen , Zhao-Yu Chen , Yong Ji , Hao Jiang , Zhi-Peng Zhu , Hao Yuan , Bo Li , Wei Xu , and Jianru Xiao 





Research Article (14 pages), Article ID 3677789, Volume 2023 (2023)

Association of LAMA1 Single-Nucleotide Polymorphisms with Risk of Esophageal Squamous Cell Carcinoma among the Eastern Chinese Population

Shaoyuan Zhang , Yong Fang, Feng Su, Tian Jiang , Jinjie Yu , Siyun Lin, Lu Lv, Tao Long, Huiwen Pan, Junqing Qi, Qiang Zhou, Weifeng Tang , Guowen Ding, Liming Wang, Lijie Tan , and Jun Yin 



Research Article (9 pages), Article ID 6922909, Volume 2023 (2023)

A Prognostic Signature for Colon Adenocarcinoma Patients Based on m6A-Related lncRNAs

Su-Zhe Zhou , Ying-Lian Pan, Qing-Chun Deng, Chang-Jun Yin, De-Jiang Zhou, Ming-Liang Liu , Jun Zhou , and Xiao-Jing Wu 



Research Article (13 pages), Article ID 7797710, Volume 2023 (2023)

Chromatin Regulator-Related Gene Signature for Predicting Prognosis and Immunotherapy Efficacy in Breast Cancer

Dongxu Feng, Wenbing Li, Wei Wu, Ulf Dietrich Kahlert, Pingfa Gao, Gangfeng Hu, Xia Huang, Wenjie Shi , and Huichao Li 

Research Article (12 pages), Article ID 2736932, Volume 2023 (2023)

Cancer-Associated Fibroblasts Affect Tumor Metabolism and Immune Microenvironment in Gastric Cancer and Identification of Its Characteristic Genes

Chanchan Gao , Fei Liu, Qiuju Ye, and Aiping Guo 







Research Article (14 pages), Article ID 1424589, Volume 2023 (2023)

HPV-Related Prognostic Signature Predicts Survival in Head and Neck Squamous Cell Carcinoma

Hongyu Zhao , Fengxu Wang, Xuehai Wang, Xinyuan Zhao , and Jinfeng Ji 

Research Article (10 pages), Article ID 7357566, Volume 2022 (2022)

Identification of Immune and Hypoxia Risk Classifier to Estimate Immune Microenvironment and Prognosis in Cervical Cancer

Yujing Shi , Qing Gao , Zeyuan Liu , Gefenqiang Shen , Xinchun Sun , and Xiaoke Di 



Research Article (20 pages), Article ID 6906380, Volume 2022 (2022)

A Connexin-Based Biomarker Model Applicable for Prognosis and Immune Landscape Assessment in Lung Adenocarcinoma

Junqing Qi , Jun Yin , and Guowen Ding 

Research Article (12 pages), Article ID 9261339, Volume 2022 (2022)

Identification of a Metabolic Reprogramming-Associated Risk Model Related to Prognosis, Immune Microenvironment, and Immunotherapy of Stomach Adenocarcinoma

Yan Zhao , Dongsheng Zhang , and Yueming Sun 

Research Article (15 pages), Article ID 7248572, Volume 2022 (2022)

Research Article

Upregulation of APOC1 Promotes Colorectal Cancer Progression and Serves as a Potential Therapeutic Target Based on Bioinformatics Analysis

Weiwei Tang ^{1,2}, Hanyuan Liu,³ Xiao Li ³, Theng Choon Ooi,¹ Nor Fadilah Rajab,¹ Hongyong Cao ³ and Razinah Sharif ^{1,4}

¹Center for Healthy Ageing and Wellness, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia

²Hepatobiliary, Liver Transplantation Center, The First Affiliated Hospital of Nanjing Medical University, Key Laboratory of Living Donor Transplantation, Chinese Academy of Medical Sciences, Nanjing, Jiangsu, China

³General Surgery, Nanjing First Hospital, Nanjing Medical University, Nanjing, Jiangsu, China

⁴Biocompatibility Laboratory, Centre for Research and Instrumentation, University Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia

Correspondence should be addressed to Razinah Sharif; razinah@ukm.edu.my

Received 16 September 2022; Accepted 11 October 2022; Published 1 March 2023

Academic Editor: Song Cao

Copyright © 2023 Weiwei Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Approximately 10% of cancer patients worldwide have colorectal cancer (CRC), a prevalent gastrointestinal malignancy with substantial mortality and morbidity. The purpose of this work was to investigate the APOC1 gene's expression patterns in the CRC tumor microenvironment and, using the findings from bioinformatics, to assess the biological function of APOC1 in the development of CRC. **Methods.** The TCGA portal was employed in this investigation to find APOC1 expression in CRC. Its correlation with other genes and clinicopathological data was examined using the UALCAN database. To validate APOC1's cellular location, the Human Protein was employed. In order to forecast the relationship between APOC1 expression and prognosis in CRC patients, the Kaplan–Meier plotter database was used. TISIDB was also employed to evaluate the connection between immune responses and APOC1 expression in CRC. The interactions of APOC1 with other proteins were predicted using STRING. In order to understand the factors that contribute to liver metastasis from CRC, single-cell RNA sequencing (scRNA-seq) was done on one patient who had the disease. This procedure included sampling preoperative blood and the main colorectal cancer tissues, surrounding colorectal cancer normal tissues, liver metastatic cancer tissues, and normal liver tissues. Finally, an in vitro knockdown method was used to assess how APOC1 expression in tumor-associated macrophages (TAMs) affected CRC cancer cell growth and migration. **Results.** When compared to paracancerous tissues, APOC1 expression was considerably higher in CRC tissues. The clinicopathological stage and the prognosis of CRC patients had a positive correlation with APOC1 upregulation and a negative correlation, respectively. APOC1 proteins are mostly found in cell cytosols where they may interact with APOE, RAB42, and TREM2. APOC1 was also discovered to have a substantial relationship with immunoinhibitors (CD274, IDO1, and IL10) and immunostimulators (PVR, CD86, and ICOS). According to the results of scRNA-seq, we found that TAMs of CRC tissues had considerably more APOC1 than other cell groups. The proliferation and migration of CRC cells were impeded in vitro by APOC1 knockdown in TAMs. **Conclusion.** Based on scRNA-seq research, the current study shows that APOC1 was overexpressed in TAMs from CRC tissues. By inhibiting APOC1 in TAMs, CRC progression was reduced in vitro, offering a new tactic and giving CRC patients fresh hope.

1. Introduction

Having a high rate of morbidity and death, colorectal cancer (CRC) is a malignant tumor of the digestive system [1]. Around 10% of all cancer patients globally have CRC, affecting 1.36 million persons [2]. The liver is the most common site of hepatic metastasis in CRC because of its close anatomical proximity. At the time of initial diagnosis, 20–25% of patients had CRC liver metastases at some point, later on, 50–60% will [3]. Currently, hepatectomy, which has a 60% 5-year survival rate, is the best course of treatment for CRC patients with liver metastases. Unfortunately, approximately 20–25% of patients with CRC liver metastases are eligible for resection at the time of diagnosis, leading to the majority of patients eventually passing away from advanced metastases [4].

The influence of immunology on the development, prognosis, and therapeutic response of CRC has been shown in recent investigations. A better prognosis, for instance, is linked to T and NK cell enrichment in original CRC or metastases [5, 6]. Programmed cell death 1 (PD1), T cell control of CD28 superfamily members, and PD ligand 1 (PD-L1) are examples of immune checkpoint molecules that have recently been identified as potential targets for CRC immunotherapy [7]. Pembrolizumab (Keytruda®) was the first PD1 inhibitor authorized by the FDA for the treatment of metastatic malignant melanoma [8]. The historic clinical trial of pembrolizumab for CRC, NCT01876511, is noteworthy. The clinical trial included 11 patients with deficient mismatch repair gene expression (dMMR) colorectal cancer, 21 patients with proficient mismatch repair gene expression (pMMR) colorectal cancer, and 9 patients with dMMR other malignancies. In dMMR CRC patients and pMMR CRC patients, respectively, the immune-related objective response rate and the immune-related progression-free survival (PFS) rate were 40% and 78% and 0% and 11%. While the pMMR CRC group's median PFS and overall survival (OS) were 2.2 months and 5.0 months, respectively, the dMMR group did not accomplish these milestones [9]. Patients with CRC now have a lot of optimism because of PD1 monoclonal antibody therapy, but medication resistance is still a concern that needs to be fixed. The study of the molecular mechanisms underlying the CRC tumor microenvironment is so crucial.

In the current work, we employed bioinformatics technology to evaluate the TCGA database and discovered that apolipoprotein C1 (APOC1) expression was considerably higher expressed in CRC tissues than in nearby tissues and that it was associated with clinical stage and a bad prognosis. The selection of APOC1 to study its role in the CRC tumor microenvironment was inspired by our research group [10]. Although APOC1 has been observed to be crucial for the growth and metastasis of a number of malignancies [11], the underlying mechanisms have not been fully understood, particularly with regard to its function and part in tumor immunity [12]. Liwen Ren et al. established that APOC1 is an immunological biomarker that controls macrophage polarization and encourages tumor dissemination through extensive pan-cancer studies, which revealed

that APOC1 is intimately connected to the infiltration of different immune cells in a range of malignancies [13]. We conducted single-cell RNA sequencing (scRNA-seq) in one patient with CRC liver metastasis to further examine why APOC1 is highly expressed in CRC tissues. We covered primary colorectal cancer tissues (CT), neighboring colorectal cancer tissues (CP), liver metastatic cancer tissues (LT), normal liver tissues (LP), and preoperative blood (PB) in order to determine which type of cell population APOC1 is significantly enriched for and address causes of liver metastasis from CRC. Furthermore, we conducted a preliminary evaluation of APOC1's role in encouraging CRC migration and proliferation in culture.

2. Materials and Methods

2.1. APOC1 Expression Level Analysis and Clinicopathological Analysis in CRC. The expression of APOC1 was examined in 24 different tumor tissue types, including CRC and related para-carcinoma tissues, using the TCGA portal. Here, UALCAN was utilized to compare the expression of APOC1 in CRC patients with various stages and lymph node metastases [14, 15].

2.2. Tools for APOC1 Location in Cells. A large number of tissue, cellular, and pathological results as well as gene data in cells and tissues are compiled in the Human Protein Atlas, through which we obtained the location of APOC1 in cells. The subcellular portion of the Human Protein provides high-resolution insight into the expression and spatial and temporal distribution of proteins encoded by 13,041 genes, representing 65% of human protein-coding genes. For each gene, the subcellular distribution of proteins was studied by immunofluorescence (ICC-IF) and confocal microscopy in up to three different cell lines selected from the 36 cell lines found in cell line sections. We showed the colocalization of APOC1 in three cell lines including A-431, U-2 OS, and U-251 MG cell lines.

2.3. Interaction Analysis of APOC1. STRING is a public database for searching for interactions and connections between proteins, both direct and indirect [16]. We perform a thorough analysis and forecast of the outcomes, as well as a summary of information exchange and contact with other (primary) databases. STRING was used to build a network of connections between APOC1, APOE, RAB42, and TREM2 among other significant proteins. The relationship between APOC1 and other genes in CRC was examined using the TCGA portal.

2.4. scRNA-Seq Analysis. In accordance with the principles outlined in the Declaration of Helsinki, all participants were given information about the study, and patients signed informed consent was obtained before it could begin. After surgical resection at the First Affiliated Hospital of Nanjing Medical University, primary colorectal cancer tissues and corresponding intestine tissues, liver metastatic cancer

tissues, and corresponding liver tissues were collected. Additionally, preoperative blood was taken. An Illumina HiSeq4000 sequencer evaluated the samples (Singleron Biotechnologies, China). Using the 10X Genomics Cell Ranger workflow, the raw counts were compared to the human reference provided by 10X Genomics (GRCh38 version) (version 2.1.0). The filter expression matrix Cell Ranger generates for each sample are read and processed using the R program Seurat. Additionally, Seurat was used to checking the quality of single-cell expression matrices (version 3.2.0). First, cells that met the following standards for quality were eliminated: mitochondrial transcripts with less than 15% uniqueness, less than 100 unique genes mapped, and more than 500 unique molecular identification counts (UMI). Using the default settings of the R program DoubletFinder, double peaks in cells were found. The left-over cells in all samples were kept and subsequently merged with Seurat for additional analysis, assuming that the duplex was eliminated.

2.5. Cell Cluster Analysis and Cell Type Identification. Using FindVariable features from Seurat, 2000 highly variable genes (HVG) were generated, and these genes were employed in principal component analysis (PCA) with parameter $NPCS = 30$. The Harmony program (version 0.1.0) was used to eliminate probable batches from samples with the parameter $NPCA = 12$ based on the PCA results. FindClusters from Seurat are then used to identify cell clusters using a shared nearest-neighbor graph. Seurat's RunTSNE and RunUMAP are used to reduce the harmony dimension of visualization by T-distributed random neighborhood embedding (tSNE) and unified manifold approximation and projection (UMAP).

2.6. Tool for Immune-Related Analysis of APOC1. The spearman correlations between APOC1 and immune-modulator expression were investigated using TISIDB, a digital portal for tumor and immune system interactions that integrates a variety of heterogeneous data [16].

2.7. Primary Culture of THP-1 and Cell Transfection. 10% fetal bovine serum (FBS) was added to RPMI 1640 medium (BI, USA) when cultivating THP-1 cells (Gibco, USA), following the transfection of THP-1 cells with lentiviral vectors, including sh-NC and sh-APOC1 (GeneChem, China). sh-NC:5'-TTCTCCGAACGTGTCACGT-3'; sh-APOC1:5'-GCATCAAACAGAGTGAAC TTT-3'. After two days of PMA-mediated macrophage differentiation, THP-1 cells were collected. THP-1 cells were exposed to HCT116/LOVO culture supernatant in RPMI 1640 media for 2 days in tumor-associated macrophages (TAMs) stimulation tests, which resulted in the production of TAMs.

2.8. Cell Proliferation Experiments. For CCK8 experiments, CRC cells were cocultured with TAMs supernatant containing sh-NC or sh-APOC1. Ten microliters of CCK8 solution (RiboBio, China) were applied at 0 hours, 24 hours,

48 hours, and 72 hours after the cocultured tumor cells were implanted in 96 wells. 4 hours after adding the CCK8 solution, analyses were carried out using a microplate reading element at 450 nm in accordance with the manufacturer's instructions (Synergy4, USA).

2.9. Scratch Wound Experiment. At 48 h after transfection, after cells were adherent into monolayer cells, the scratched cells were evenly crossed with a sterile gun tip, gently washed with PBS, and then replaced with 1% FBS medium and cultured in a 37°C and 5% CO₂ incubator. At 0 h and 48 h, 5 fields were randomly selected under an inverted microscope to observe the wound healing and take photos. Cell migration distance was measured and calculated.

2.10. Statistics-Related Analyzing Process. Continuous information was compared between the two groups by one individual *t*-test procedure. GraphPad Prism 8.0 was used to carry out the statistically significant analytical method and present the results graphically. It was deemed statistically significant with a *P* value of 0.05.

3. Results

3.1. Expression and of Clinical Role of APOC1 in CRC Based on TCGA Data. The TCGA portal revealed that the expression of APOC1 was higher in tumor tissues, including CRC, than in normal tissues (Figure 1(a)). Based on subgroup analysis of CRC individual cancer stages and lymph node metastasis, it was discovered that APOC1 expression increased with increasing cancer stage and lymph node metastasis (Figures 1(b) and 1(c)). Using the Kaplan–Meier plotter, the prognostic value of APOC1 in CRC was further investigated. The findings revealed that, albeit not statistically significant ($p > 0.05$), CRC patients with high APOC1 expression had considerably worse prognoses than those with low expression (Figure 1(d)). Data from the Human Protein Atlas analysis showed that CRC patients, including those with rectum and colon cancer, had high or low expression of the APOC1 protein (Figure 1(e)).

3.2. Genes and Proteins Cointeracted with APOC1 in CRC. The human protein Atlas database revealed that APOC1 was found in the cytoplasm of A-431, U-2 OS, and U-251 MG cells (Figure 2(a)). It was possible to find proteins that interact with APOC1 using a STRING interactive network (Figure 2(b)). Further investigation revealed a high correlation between the expression of APOC1 and proteins that may interact with it, including APOE, RAB42, and TREM2 (Figure 2(c)).

3.3. Acquisition of scRNA-Seq Profiles of Samples and Data Generation in CRC Liver Metastasis. In this study, we carried out scRNA-seq [17] in one CRC liver metastasis patient covering primary colorectal cancer tissues (CT), adjacent tissues of colorectal cancer (CP), liver metastatic cancer tissues (LT), normal liver tissues (LP), and preoperative

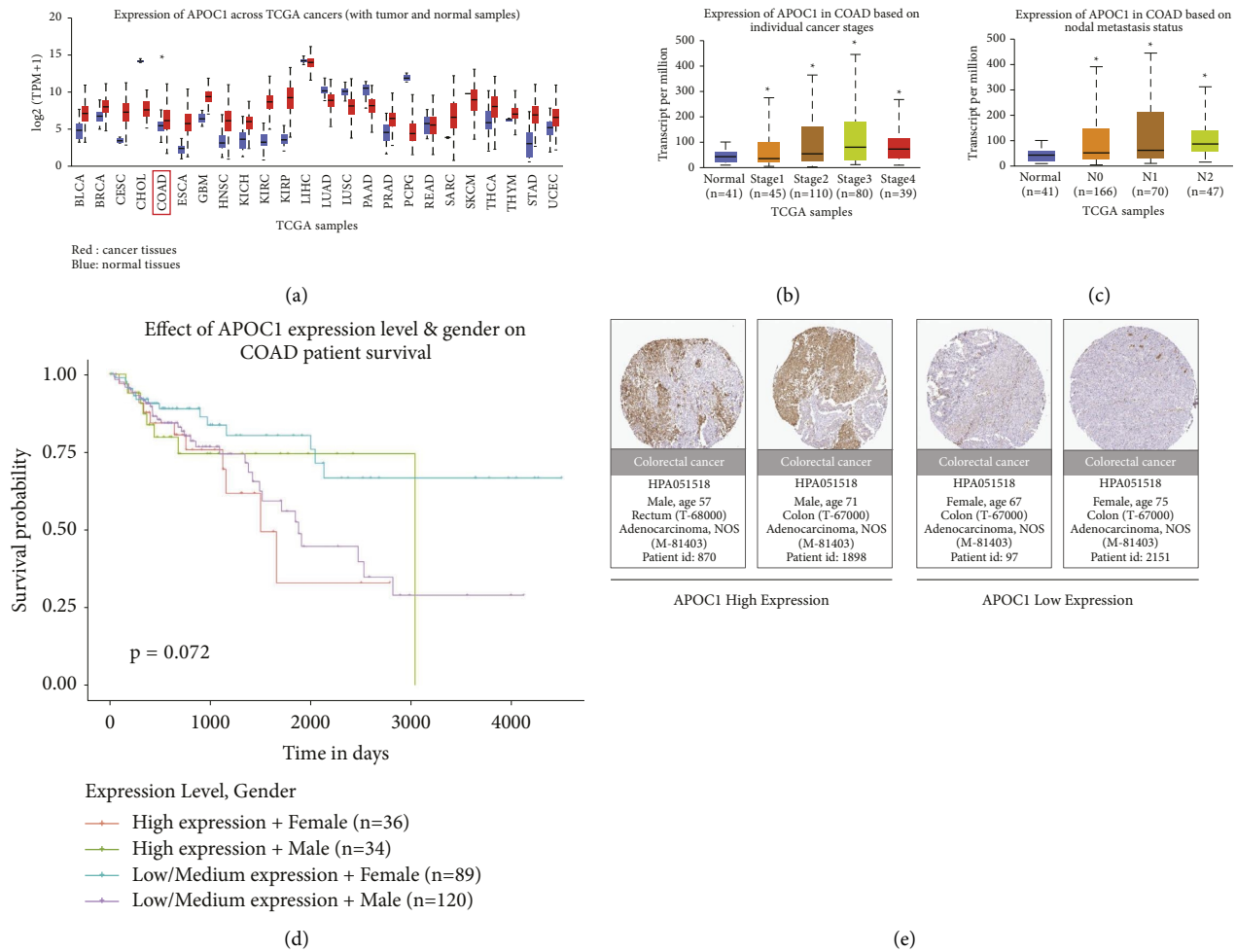


FIGURE 1: Expression of APOC1 in CRC tissues. (a) The expression level of APOC1 mRNA in different types of cancer tissues compared to normal tissue. (b) The correlation between APOC1 mRNA expression and tumor stage. (c) The correlation between APOC1 mRNA expression and lymph node metastatic status. (d) The relationship between APOC1 expression and CRC patient's prognosis. (e) Immunohistochemical of APOC1 expression in CRC tissues from different patients. * $P < 0.05$.

blood (PB) and aimed to address the causes of liver metastasis from CRC. Through the definition of classification, we finally identified 16 cell clusters in immune cells using a UMAP plot (Figure 3(a)). Each cell type has unique marker genes (Figure 3(b)). For example, the B cell cluster specifically expresses MS4A1, CD79A, and CD79B. CD4-IL7R expresses IL7R, IL32, and MALAT1; while CD8-GZMB expresses CD8A, CD8B, and KLRD1. TAM-APOC1 expresses CD14, C1QC, APOC1, and SPP1. In addition, we used the violin chart to show the expression of some marker genes (CD79A, FCN1, and APOC1) in various cell populations (Figure 3(c)). A cluster map and histogram were applied to show the expression of these cell clusters in different tissues and results revealed that there was less TAM but more B and plasma cell clusters in CP compared with CT. CT showed more TAM and fewer CD8 T cells. Moreover, there were more CD8 T cells and NK enrichment in LP compared with LT (Figures 4(a) and 4(b)). These results demonstrate that TAMs might play an important role in both the metastasis of primary tissues and the colonization of metastatic foci and that the lethality of NK cells in cancer

tissues is insufficient in the colonization process after metastasis of CRC.

3.4. APOC1 Was Highly Expressed in TAMs of CRC Tissues.

We discovered that APOC1 may be crucial to the TAMs of CRC based on the results of the scRNA-seq analysis. As a result, we carefully examined the expression of APOC1 in each sample and each cluster of cells. The exact distribution of various subgroups in various samples is also displayed in the UMAP graphic (Figure 5(a)). The enrichment of APOC1 in various cell clusters in various samples was more clearly displayed by the UMAP map (Figure 5(b)). TAMs had much higher levels of APOC1 than other cell clusters such as CD8 T, CD4 T, and NK cells, which were both less expressed (Figure 5(b)). Furthermore, we looked at the data from Hae-Ock Lee et al. [18] and discovered that APOC1 was primarily expressed in myeloid CRC tissues (Figures 6(a) and 6(b)), which is similar to our findings. Our intense curiosity about the function of APOC1 in TAMs from CRC was piqued by all of these analyses.

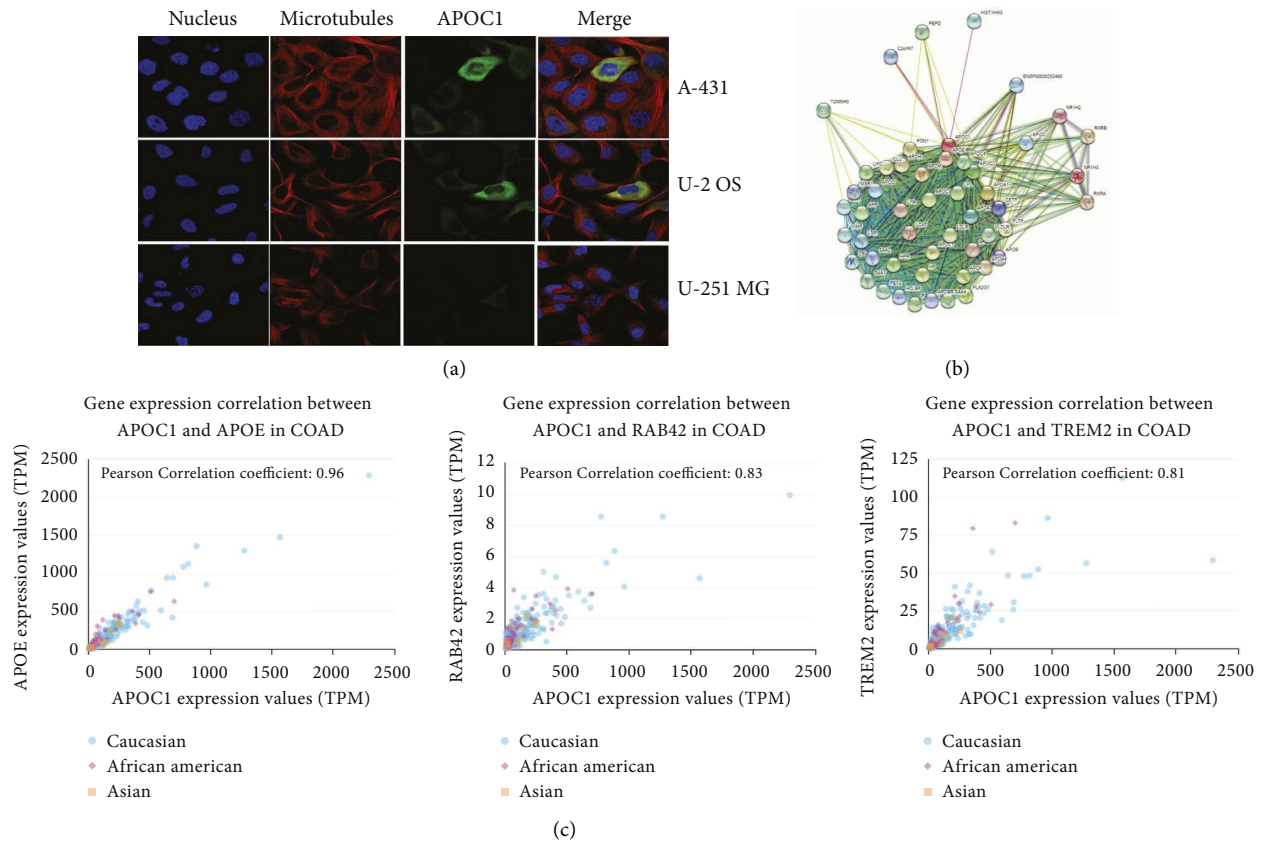


FIGURE 2: Genes and proteins cointeracted with APOC1. (a) APOC1 located in the cytosol. (b) Interactions between APOC1 and other proteins. (c) Relationship analysis between APOC1 and APOE, RAB42, and TREM2 in CRC.

3.5. APOC1 Expression Was Correlated with Immune Factors.

We investigated the connection between the expression of APOC1 and immunological components in CRC because the aforementioned findings showed that APOC1 is linked to immunity, particularly TAMs, in CRC. As shown in Figures 7 and 8, there was a significant link between the expression of immunostimulators (PVR, CD86, and ICOS) and immunoinhibitors (CD274, IDO1, and IL10) and APOC1 itself.

3.6. Inhibition of APOC1 of TAM Reduced CRC Progression In Vitro.

We stimulated CRC cells with TAM supernatant in order to further confirm the function of APOC1 in TAMs from CRC in vitro. By using the CCK8 and scratch assays, we discovered that TAMs in the sh-APOC1 group significantly decreased the proliferation and migration of CRC cells in comparison to the control group (Figures 9(a) and 9(b)).

4. Discussion

An earlier study found that the mitogenic impact of high-density lipoprotein cholesterol (HDL) on bovine vascular endothelial cells in vitro was caused by APOC1 purified from HDL [19], and APOC1 has recently been identified as a molecule involved in the advancement of cancer. According to research, APOC1 functions as an oncogene in cervical cancer, and its knockdown both in vitro and in vivo

reduces the proliferation of cervical cancer cells. The clinical outcome of cervical cancer patients is highly correlated with the relative expression of APOC1 [20]. Li Yangling et al. discovered that APOC1 activated STAT3 to increase renal clear cell carcinoma metastasis [11]. According to Huaying Xiao et al., clear cell renal cell carcinoma (ccRCC) tissues had a greater expression level of APOC1 than the normal group did. Poor prognosis was linked to high APOC1 expression in female patients but not in male patients. In ccRCC patients older than 60 years, high APOC1 expression also decreased survival [21]. Through the MAPK signaling pathway, APOC1 increases the growth of CRC tumors, according to research by Ren Hui et al. [12]. Based on the TCGA portal, we discovered in the current study that the expression of APOC1 in tumor tissues, including CRC, was obviously higher than that in normal tissues. Subgroup analysis also revealed that the higher the cancer stage and lymph node metastasis, the higher the expression of APOC1. Results from the Kaplan–Meier plotter demonstrated that, despite being not statistically significant, the prognosis of CRC patients with high APOC1 expression was significantly worse than that of those with low expression. Our conclusion is generally in line with the previous conclusion.

A STRING interactive network revealed a favorable correlation between the expression of APOC1 and proteins like APOE, RAB42, and TREM2 that may interact with APOC1. Both APOC1 and APOE are apolipoproteins, which

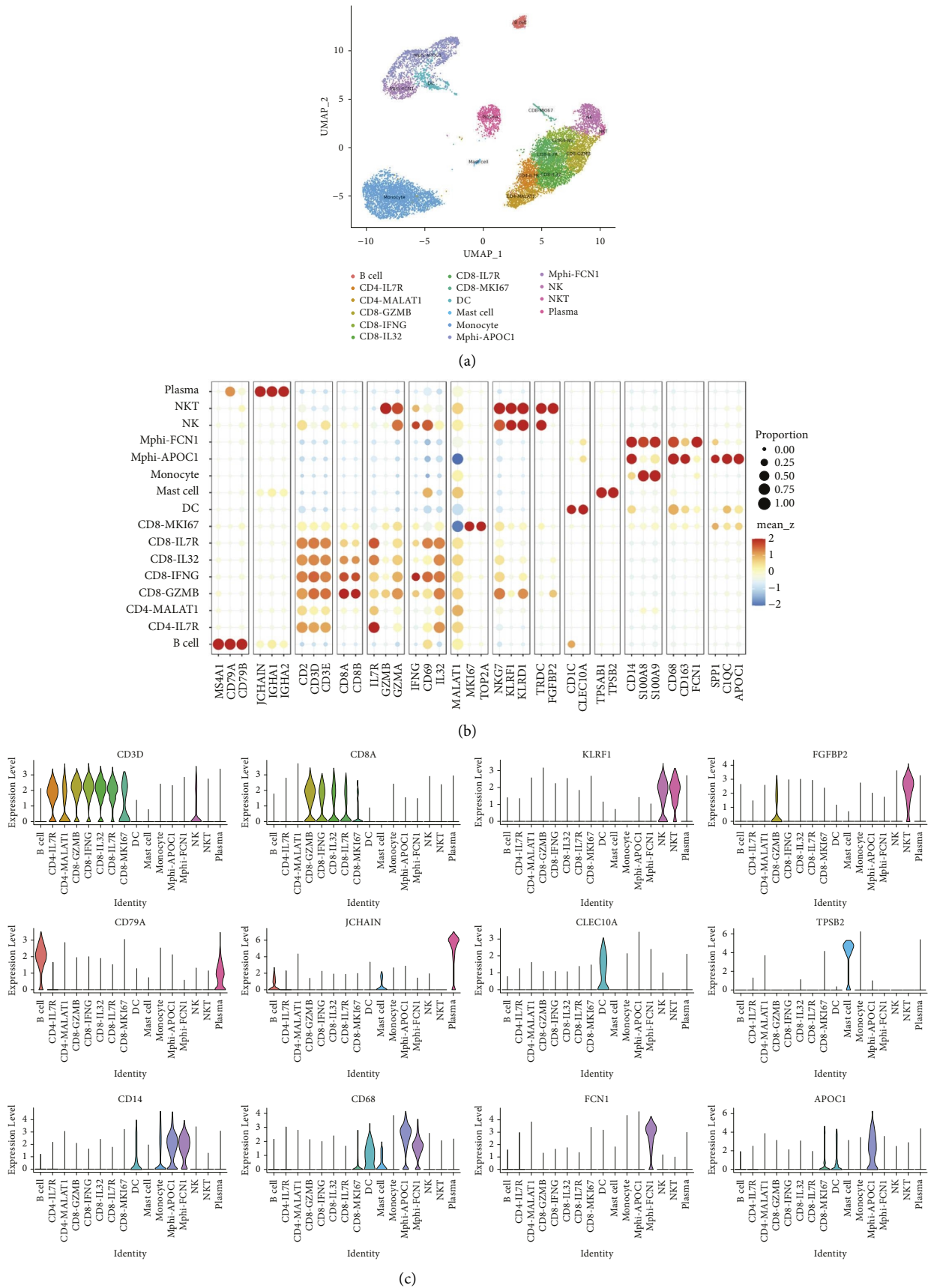
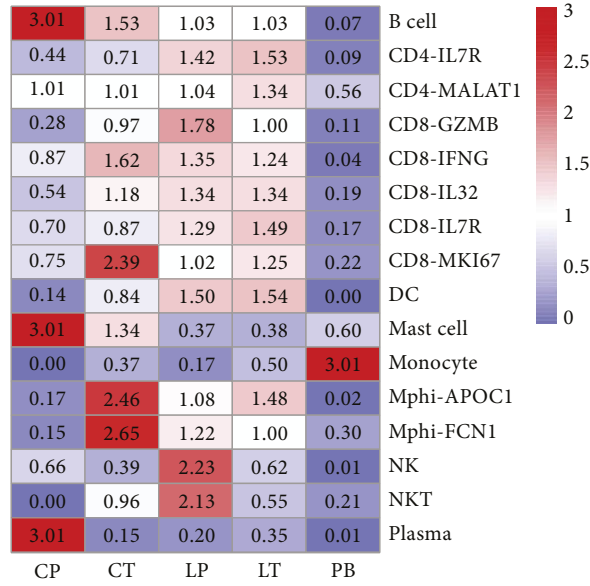


FIGURE 3: Diverse cell types in CRC delineated by single-cell transcriptomic analysis. (a) UMAP plot showing 16 clusters of immune cells. (b) Dot plot showing the clustering of immune cell types in each sample. (c) The violin diagram showing expression levels of specific markers in each cell type.



(a)



(b)

FIGURE 4: Expression of individual cell populations in individual samples. (a) The distribution number of each immune cell subgroup in each sample. (b) The histogram showing the distribution number of each immune cell subgroup in each sample. CT, primary colorectal cancer tissues; CP, adjacent tissues of colorectal cancer; LT, liver metastatic cancer tissues; LP, normal liver tissues; PB, preoperative blood.

function as physiological carriers of hydrophobic lipids in aqueous fluids throughout the body [22]. Apolipoproteins and different malignancies may be related, according to

some research studies. In lung cancer cells and B16F10 cells, APOE expression was knocked down, which reduced tumor development and metastasis [23]. To describe

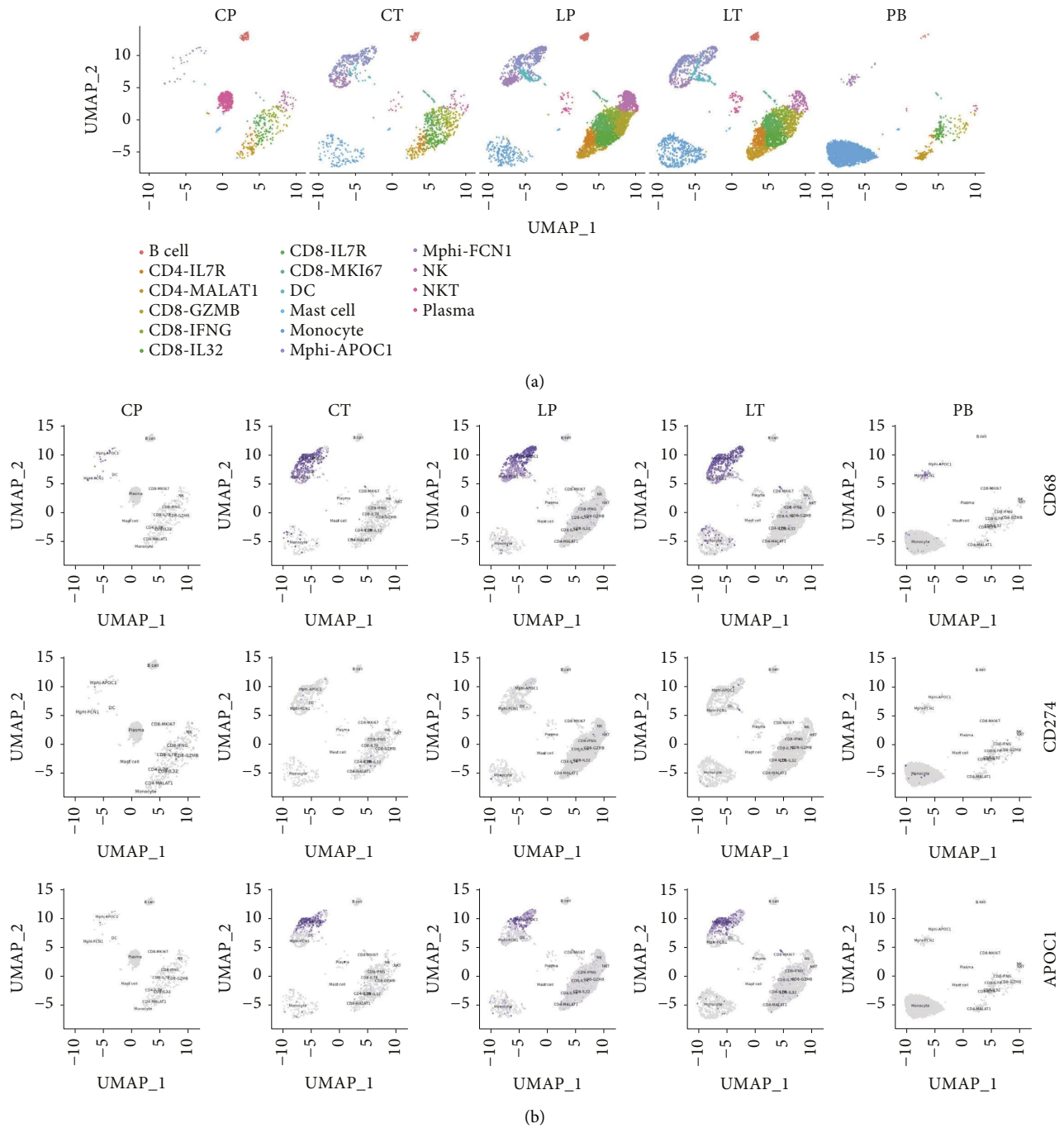


FIGURE 5: APOC1 expression in individual cell populations and individual samples. (a) UMAP plot showing the distribution of each cell subgroup in each sample. (b) The expression of APC1 in immune cells in each sample. CT, primary colorectal cancer tissues; CP, adjacent tissues of colorectal cancer; LT, liver metastatic cancer tissues; LP, normal liver tissues; PB, preoperative blood (PB).

APOE-TREM2 interactions, molecular docking and molecular dynamics (MD) investigations were carried out. Additionally, it was examined how a significant TREM2-disease-related mutation (R47H) affected TREM2 affinity for APOE. The outcomes demonstrated that the binding energy occurred between APOE and TREM2 in an isomer-dependent manner, with the potency order being APOE4 > APOE3 > APOE2. Furthermore, the R47H mutation decreased the connection between the APOE and

TREM2 proteins, which may be a result of hydrogen bond interactions, hydrophobic interactions, or a weaker electrostatic interaction between APOE and TREM2 [24]. RAB42 is linked to cancer prognosis and progression, according to earlier studies. RAB42 expression levels in hepatocellular carcinoma (HCC) tissues were higher than in normal tissues, according to a prior investigation. Significant correlations were found between highly expressed RAB42 and a number of clinical indicators in HCC patients.

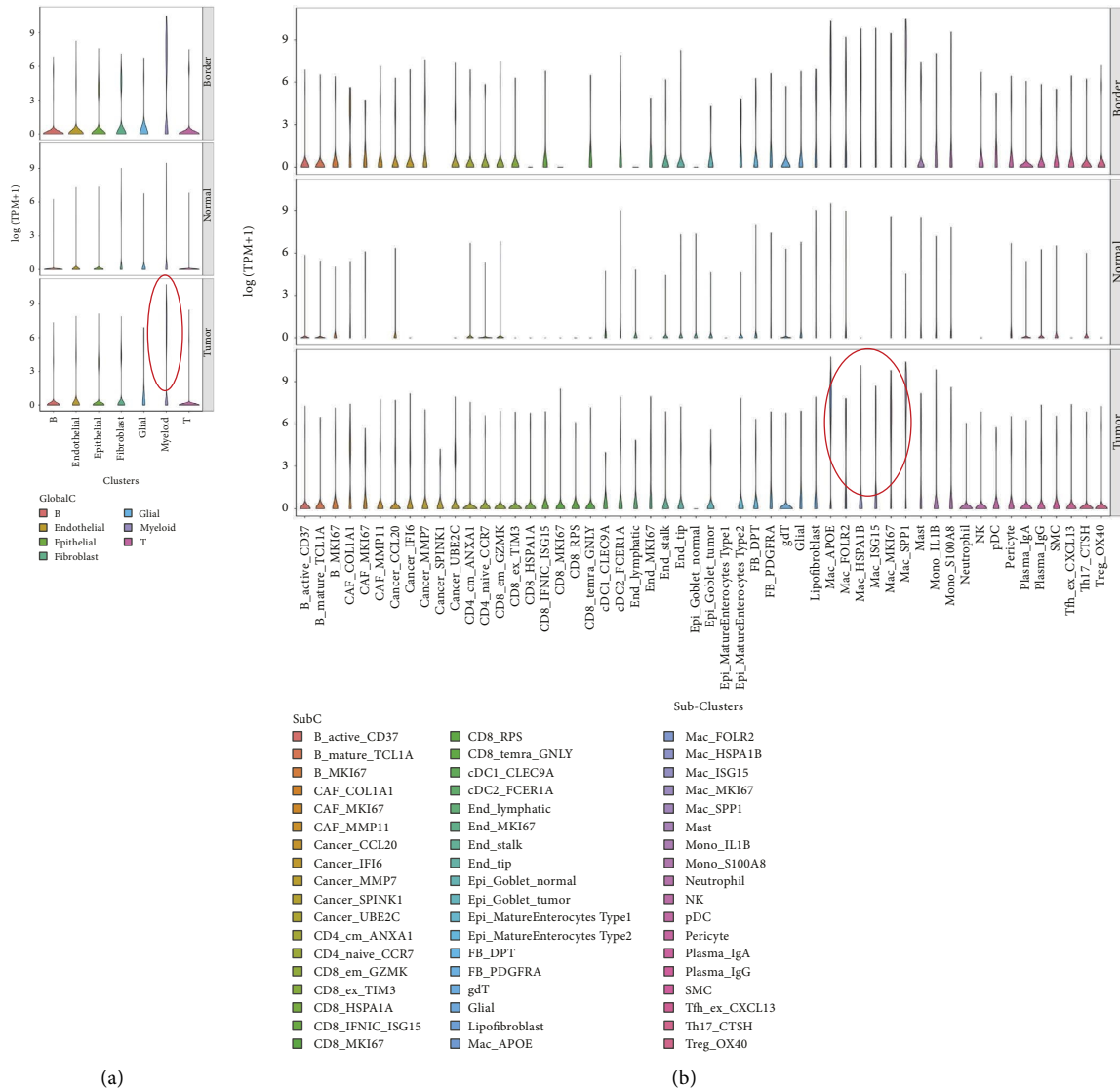


FIGURE 6: Research results of APOC1 at single-cell level from Hae-Ock Lee's study. (a) The violin diagram displaying the distribution of APOC1 expression in different cells from CRC tissues in total analysis. (b) The violin diagram displaying the distribution of APOC1 expression in different cells from CRC tissues in subanalysis.

Additionally, elevated RAB42 expression unmistakably indicated a bad prognosis for HCC [25]. In comparison to normal samples, glioblastoma (GBM) samples showed higher expression of RAB42. Patients with high RAB42 expression have a worse prognosis than those with low RAB42 expression in GBM. A total of 35 pathways, including the P53 pathway, were significantly activated in GBM samples with elevated RAB42 expression [26]. A greater understanding of the direct interactions between APOC1, APOE, RAB42, and TREM2 in cancer is, however, required due to the paucity of studies in this area.

The role and mechanism of APOC1 in the tumor microenvironment have only been partially studied. With the quick advancement of scRNA-seq technology, diverse cell populations can be characterized and identified, and new cell markers and regulatory pathways can be found. It is interesting to note that APOC1 has been linked to a number of

immune cell infiltrations in different malignancies. scRNA-seq research revealed that TAMs expressed the bulk of APOC1 in expression. TAMs with the M2 phenotype are produced when renal cell cancer cells are cocultured; this is prevented by silencing APOC1. By interacting with CD163 and CD206, APOC1 boosted macrophage polarization toward M2 by increasing its expression in M2 or TAM. Additionally, through secreting CCL5, macrophages overexpressing APOC1 aided in the spread of renal cell cancer cells [13]. According to Chan et al.'s research, TAMs have high levels of APOE, APOC1, and SPP1 expression, which results in an anti-inflammatory macrophage phenotype [27]. Based on the findings of this study's scRNA-seq, we show that basic CRC and liver metastatic tissues exhibit APOC1 overexpression in TAMs. We also looked at the relationship between the expression of APOC1 and immune factors in CRC and discovered that there was a significant positive

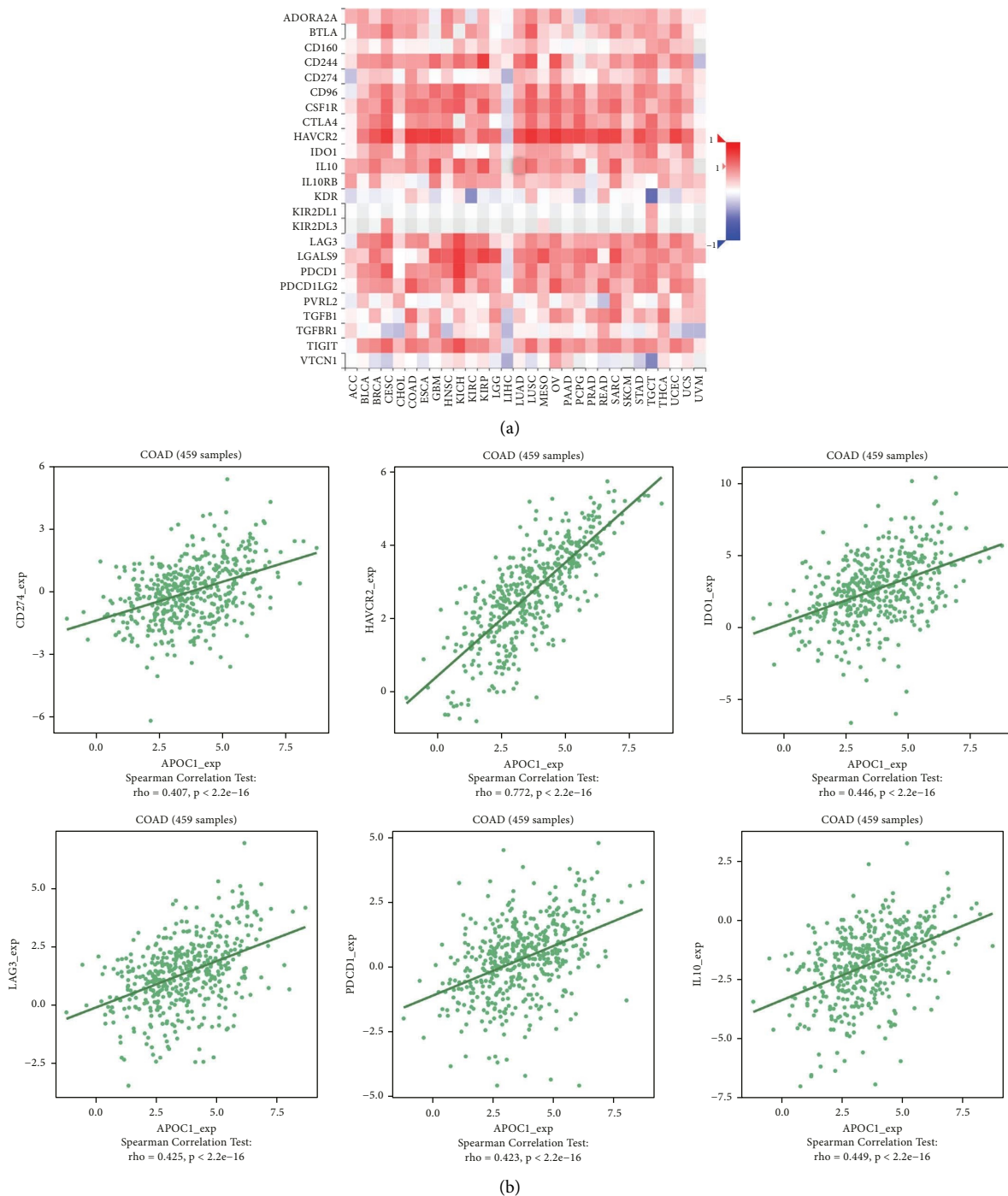


FIGURE 7: Correlation between APOC1 expression and immunoinhibitors in CRC. (a) The heat map showing the correlation between APOC1 and immunoinhibitor factors in different cancers. (b) The line graph showing the correlation of APOC1 with specific immune indicators in CRC.

correlation between the expression of immunoinhibitors (CD274, IDO1, and IL10) and the expression of APOC1. This result suggested that APOC1 is important in the development of the immunosuppressive tumor microenvironment. We stimulated CRC cancer cells with TAM supernatant in order to further confirm the function of APOC1 in CRC TAMs in vitro. TAMs in the sh-APOC1 group drastically decreased CRC cell proliferation and

migration by CCK8 and scratch assays. The significance of APOC1 in the tumor immune microenvironment is substantially expanded by our findings. Massimo Pancione et al. proposed that many different functions of TAMs during tumor progression may depend on their intrinsic adaptation to positional schemes that are acquired through factors that control the balance between a tumor suppressor and tumor-promoting activities. In the primary tumor, oncogenic

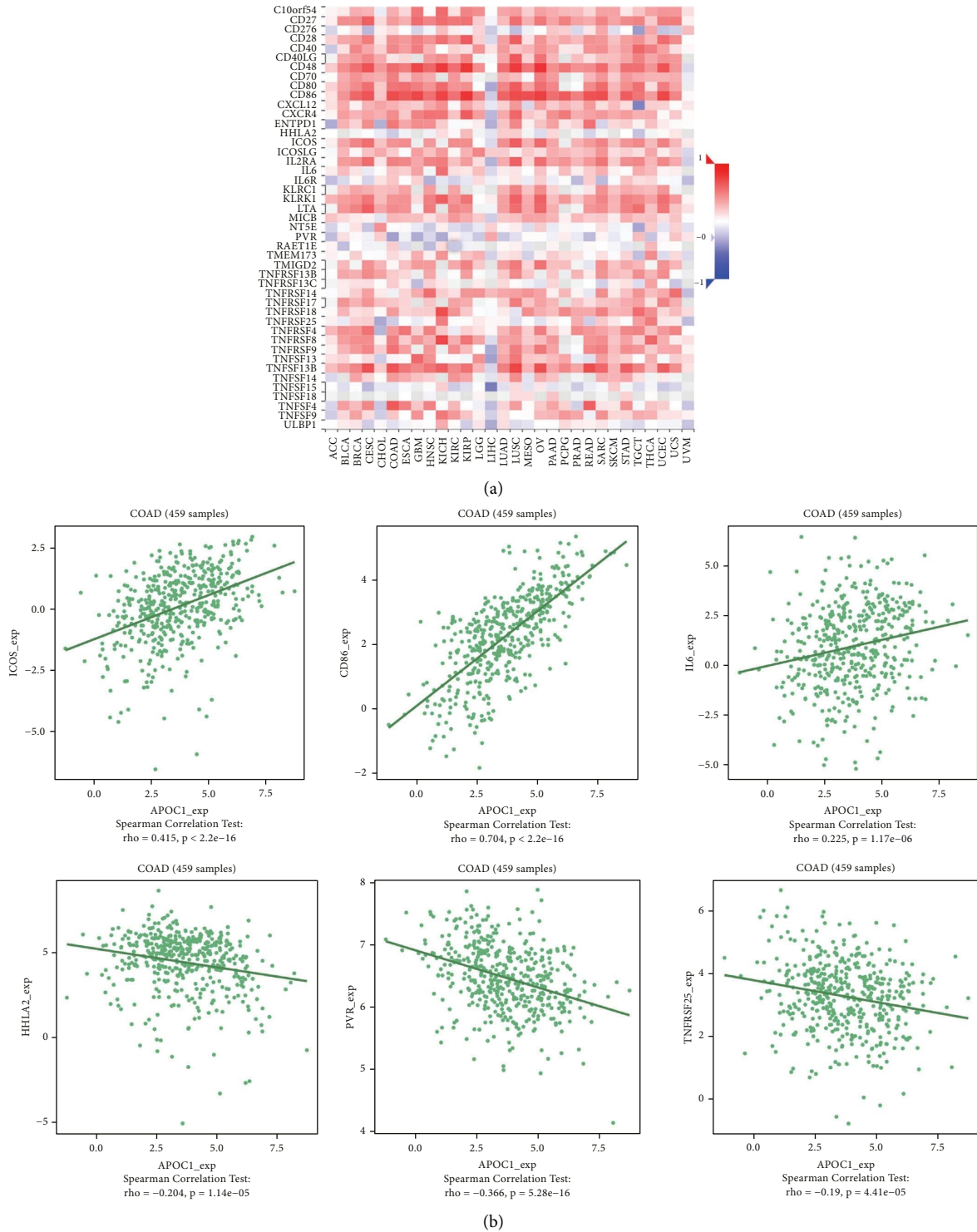


FIGURE 8: Correlation between APOC1 expression and immunostimulators in CRC. (a) The heat map showing the correlation between APOC1 and immunostimulator factors in different cancers. (b) The line graph showing the correlation of APOC1 with specific immune indicators in CRC.

alterations or changes in the tumor microenvironment establish a new equilibrium that can be further modified during metastasis. There are at least two mechanisms supporting the prometastatic function of TAMs:

- (1) M2-macrophages can form a dense barrier around invasive cancer cells, leading to heterotypic interactions between tumor cells and the surrounding matrix, disrupting host tissue integrity;
- (2) Invasive cancer cells can acquire

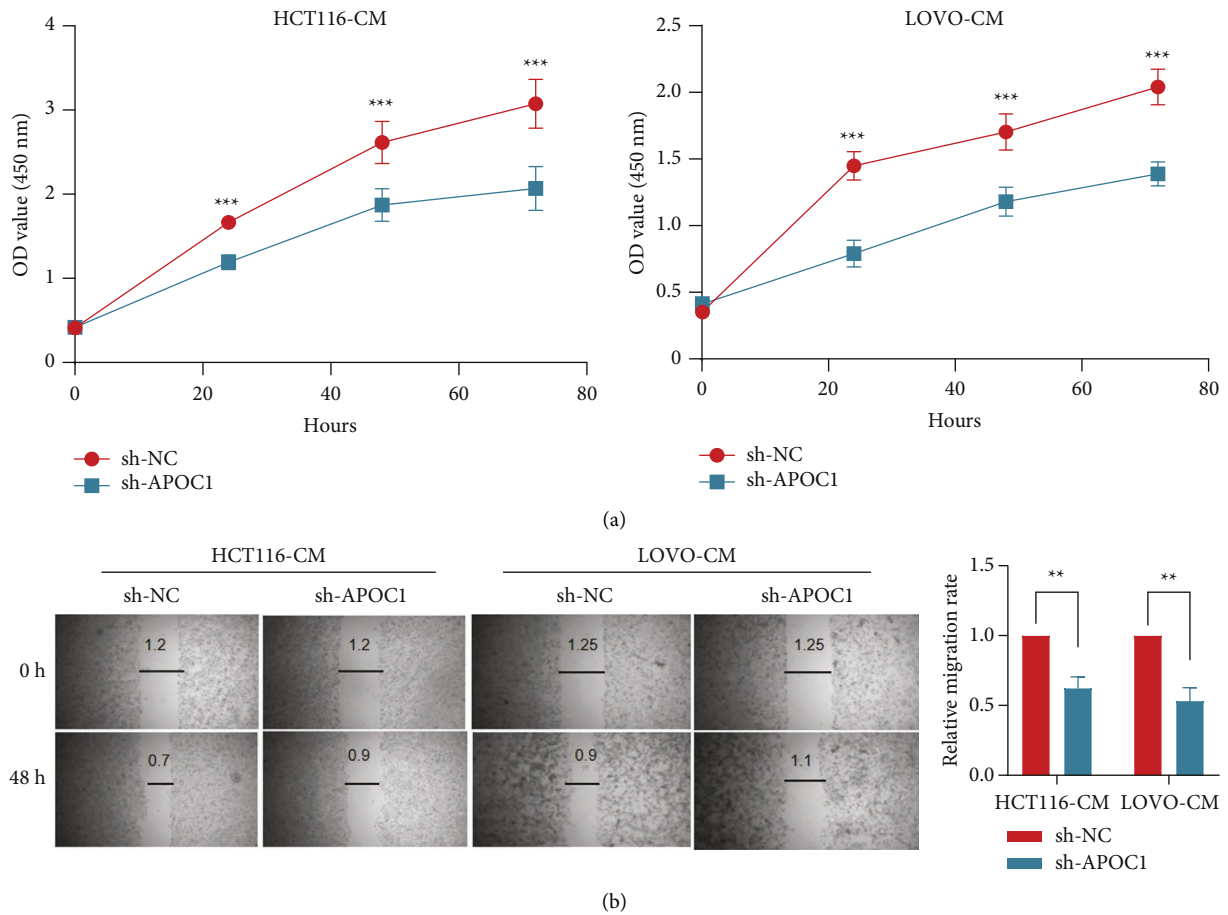


FIGURE 9: Inhibition of APOC1 of TAMs reduced CRC progression in vitro. (a) CCK8 assay of CRC cancer cells with TAM supernatant in different groups. (b) Scratch assay of cancer cells with TAM supernatant in different groups. ** $P < 0.01$, *** $p < 0.001$.

immunophenotypic features, such as macrophage fusion with cancer cells, which promotes homotypic interactions between host matrix and TAMs [28–30]. Therefore, TAMs play an indispensable role in the progression of liver metastasis of CRC.

The association between APOC1 and CRC was extensively investigated in this research using bioinformatics analysis and a few trials, although there are still numerous gaps in our understanding. First off, no mechanistic investigation was done; only APOC1's expression and function in TAMs were confirmed. Second, there are not many research studies on in vivo efficacy in animals. Third, the impact of APOC1 knockdown on other cells was not investigated in TAMs. We anticipate publishing more information about the connection between APOC1 function and cancer in many cell types.

5. Conclusion

In conclusion, the current study shows that APOC1 was highly expressed in TAMs of CRC tissues based on scRNA-seq and bioinformatics analysis and that inhibiting APOC1 of TAMs slowed CRC progression in vitro, offering a novel approach and giving CRC patients fresh hope.

Data Availability

All data relevant to the study are included in the article.

Consent

Not applicable.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] V. Aran, A. P. Victorino, L. C. Thuler, and C. G. Ferreira, "Colorectal cancer: epidemiology, disease mechanisms and interventions to reduce onset and mortality," *Clinical Colorectal Cancer*, vol. 15, no. 3, pp. 195–203, 2016.
- [2] N. Keum and E. Giovannucci, "Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies," *Nature Reviews Gastroenterology & Hepatology*, vol. 16, no. 12, pp. 713–732, 2019.
- [3] F. C.-L. Chow and K. S.-H. Chok, "Colorectal liver metastases: an update on multidisciplinary approach," *World Journal of Hepatology*, vol. 11, no. 2, pp. 150–172, 2019.

- [4] S. Pathak, R. Jones, J. M. F. Tang et al., "Ablative therapies for colorectal liver metastases: a systematic review," *Colorectal Disease*, vol. 13, no. 9, pp. e252–e265, 2011.
- [5] M. Donadon, K. Hudspeth, M. Cimino et al., "Increased infiltration of natural killer and T cells in colorectal liver metastases improves patient overall survival," *Journal of Gastrointestinal Surgery*, vol. 21, no. 8, pp. 1226–1236, 2017.
- [6] P. S. Yoon, N. Del Piccolo, V. S. Shirure et al., "Advances in modeling the immune microenvironment of colorectal cancer," *Frontiers in Immunology*, vol. 11, Article ID 614300, 2020.
- [7] Y. Xiao and G. J. Freeman, "The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy," *Cancer Discovery*, vol. 5, no. 1, pp. 16–18, 2015.
- [8] A. Barone, M. Hazarika, M. R. Theoret et al., "FDA approval summary: pembrolizumab for the treatment of patients with unresectable or metastatic melanoma," *Clinical Cancer Research*, vol. 23, no. 19, pp. 5661–5665, 2017.
- [9] D. T. Le, J. N. Uram, H. Wang et al., "PD-1 blockade in tumors with mismatch-repair deficiency," *New England Journal of Medicine*, vol. 372, no. 26, pp. 2509–2520, 2015.
- [10] X. Hao, Z. Zheng, H. Liu et al., "Inhibition of APOC1 promotes the transformation of M2 into M1 macrophages via the ferroptosis pathway and enhances anti-PD1 immunotherapy in hepatocellular carcinoma based on single-cell RNA sequencing," *Redox Biology*, vol. 56, Article ID 102463, 2022.
- [11] Y.-l. Li, L.-w. Wu, L.-h. Zeng et al., "ApoC1 promotes the metastasis of clear cell renal cell carcinoma via activation of STAT3," *Oncogene*, vol. 39, no. 39, pp. 6203–6217, 2020.
- [12] H. Ren, Z. Chen, L. Yang et al., "Apolipoprotein C1 (APOC1) promotes tumor progression via MAPK signaling pathways in colorectal cancer," *Cancer Management and Research*, vol. 11, pp. 4917–4930, 2019.
- [13] L. Ren, J. Yi, Y. Yang et al., "Systematic pan-cancer analysis identifies APOC1 as an immunological biomarker which regulates macrophage polarization and promotes tumor metastasis," *Pharmacological Research*, vol. 183, Article ID 106376, 2022.
- [14] D. S. Chandrashekar, S. K. Karthikeyan, P. K. Korla et al., "UALCAN: an update to the integrated cancer data analysis platform," *Neoplasia*, vol. 25, pp. 18–27, 2022.
- [15] D. S. Chandrashekar, B. Bashel, S. A. H. Balasubramanya et al., "UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses," *Neoplasia*, vol. 19, no. 8, pp. 649–658, 2017.
- [16] B. Ru, C. N. Wong, Y. Tong et al., "TISIDB: an integrated repository portal for tumor-immune system interactions," *Bioinformatics*, vol. 35, no. 20, pp. 4200–4202, 2019.
- [17] D. Szklarczyk, A. L. Gable, K. C. Nastou et al., "The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets," *Nucleic Acids Research*, vol. 49, no. D1, pp. D605–D612, 2021.
- [18] D. Rong, G. Sun, Z. Zheng et al., "MGP promotes CD8+ T cell exhaustion by activating the NF- κ B pathway leading to liver metastasis of colorectal cancer," *International Journal of Biological Sciences*, vol. 18, no. 6, pp. 2345–2361, 2022.
- [19] J.-F. Tournier, F. Bayard, and J.-P. Tauber, "Rapid purification and activity of apolipoprotein CI on the proliferation of bovine vascular endothelial cells in vitro," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 804, no. 2, pp. 216–220, 1984.
- [20] C. Gao, C. Zhou, J. Zhuang et al., "MicroRNA expression in cervical cancer: novel diagnostic and prognostic biomarkers," *Journal of Cellular Biochemistry*, vol. 119, no. 8, pp. 7080–7090, 2018.
- [21] H. Xiao and Y. Xu, "Overexpression of apolipoprotein C1 (APOC1) in clear cell renal cell carcinoma and its prognostic significance," *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*, vol. 27, Article ID e929347, 2021.
- [22] E. V. Fuior and A. V. Gafencu, "Apolipoprotein C1: its pleiotropic effects in lipid metabolism and beyond," *International Journal of Molecular Sciences*, vol. 20, no. 23, p. 5939, 2019.
- [23] Y. S. Lee, I. J. Yeo, K. C. Kim, S.-B. Han, and J. T. Hong, "Corrigendum: inhibition of lung tumor development in ApoE knockout mice via enhancement of TREM-1 dependent NK cell cytotoxicity," *Frontiers in Immunology*, vol. 12, p. 1379, 2022.
- [24] Z. Mai, W. Wei, H. Yu, Y. Chen, Y. Wang, and Y. Ding, "Molecular recognition of the interaction between ApoE and the TREM2 protein," *Translational Neuroscience*, vol. 13, no. 1, pp. 93–103, 2022.
- [25] H. Peng, X. Du, and Y. Zhang, "RAB42 is a potential biomarker that correlates with immune infiltration in hepatocellular carcinoma," *Frontiers in Molecular Biosciences*, vol. 9, Article ID 898567, 2022.
- [26] L. Sun, T. Yan, and B. Yang, "The progression related gene RAB42 affects the prognosis of glioblastoma patients," *Brain Sciences*, vol. 12, no. 6, p. 767, 2022.
- [27] Z. Chen, Y. Huang, Z. Hu et al., "Dissecting the single-cell transcriptome network in patients with esophageal squamous cell carcinoma receiving operative paclitaxel plus platinum chemotherapy," *Oncogenesis*, vol. 10, no. 10, pp. 71–13, 2021.
- [28] M. Pancione, G. Giordano, A. Remo et al., "Immune escape mechanisms in colorectal cancer pathogenesis and liver metastasis," *Journal of immunology research*, vol. 2014, pp. 1–11, 2014.
- [29] I. Kareva and P. Hahnfeldt, "The emerging "hallmarks" of metabolic reprogramming and immune evasion: distinct or linked? Reprogramming and immune evasion: distinct or linked?" *Cancer Research*, vol. 73, no. 9, pp. 2737–2742, 2013.
- [30] A. Mantovani, A. Sica, P. Allavena, C. Garlanda, and M. Locati, "Tumor-associated macrophages and the related myeloid-derived suppressor cells as a paradigm of the diversity of macrophage activation," *Human Immunology*, vol. 70, no. 5, pp. 325–330, 2009.

Research Article

TREM2 as a Prognostic Biomarker for Osteosarcoma Microenvironment Remodeling

Zhi-Long Shen ¹, Zhao-Yu Chen ¹, Yong Ji ², Hao Jiang ¹, Zhi-Peng Zhu ¹,
Hao Yuan ¹, Bo Li ¹, Wei Xu ¹, and Jianru Xiao ¹

¹Department of Orthopedic Oncology, Changzheng Hospital, Second Military Medical University, 415 Fengyang Road, Shanghai 200003, China

²The Department of Spine Surgery II, Strategic Support Force Medical Center, Beijing, China

Correspondence should be addressed to Bo Li; 1014505443@qq.com, Wei Xu; xuwei_spine@163.com, and Jianru Xiao; jianruxiao83@163.com

Received 23 August 2022; Revised 23 October 2022; Accepted 24 November 2022; Published 17 February 2023

Academic Editor: Feng Jiang

Copyright © 2023 Zhi-Long Shen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The tumor microenvironment (TME) acts as a crucial role in the occurrence and development of osteosarcoma (OS). Despite this, the mechanism controlling the components of immunity and stroma in the tumor microenvironment remains a mystery. To conduct this study, we download and collate transcriptome data from the TARGET database, whose full name is Therapeutically Applicable Research to Generate Effective Treatments, as well as available clinical information of OS. The CIBERSORT and ESTIMATE methodology are used to acquire the proportions of components of immunity and stroma and tumor-infiltrating immune cells (TICs). Protein-protein interaction (PPI) networks and Cox regression analysis are used to select differentially expressed genes (DEGs). A prognostic biomarker is determined by intersecting univariate COX and PPI results, which lead to the finding of Triggering receptor expressed on myeloid cells-2 (TREM2). Based on the next analysis, TREM2 expression is positively correlated with OS survival time. Immune function-related genes have enrichment in the group with high expression of TREM2, according to gene set enrichment analysis (GSEA). The percentage of TICs by CIBERSORT methodology revealed that the expression of TREM2 is positively associated with follicular helper T cells, CD8-positive T cells, and M2 macrophages and negatively correlated with plasma cells, M0 macrophages, and naive CD4-positive T cells. All results suggest a possible integral role of TREM2 in the immune-related events of TME. Therefore, TREM2 may be a potential indicator of remodeling of TME in osteosarcoma, which is useful and helpful in predicting the clinical prognostic outcome of OS patients and provide a unique perspective for immunotherapy for OS.

1. Introduction

Chiefly affecting children and young adults and occupying about nine percent of cancer-related deaths in youngsters whose age range is between 10 and 24 years old [1, 2], the exact cause of osteosarcoma is still unclear as a primary malignant bone tumor. In addition, the local invasiveness and metastasis of osteosarcoma remain an enormous challenge of therapy and poor prognosis [3]. With the advent of adjuvant and neoadjuvant chemotherapy, the five-year survival rate of OS had a substantial increase which is upto about 70% since the 1970s, but after lung metastasis, the

five-year survival rate still maintains as low as 20–30% [4]. In addition to this, osteosarcoma is highly heterogeneous which makes the prediction of treatment outcomes complicated [5]. The OS includes distinct histological subtypes: osteoblastic, chondroblastic, fibroblastic, giant-cell rich, epithelioid, small-cell, and telangiectatic types [4]. Therefore, exploring new diagnostic and predictive biomarkers and validating more therapeutic targets are continuously essential and critical.

Acting decisive roles in tumor occurrence, progression, metastasis, and sensitivity to therapy, TME has aroused tremendous interest in basic and clinical research as

a therapeutic target in cancer [6]. Resident stromal cells and recruited immune cells are the primary component of TME in OS. There is convincing evidence to prove that the stromal cell acts a prominent role in angiogenesis and the remodeling of extracellular matrix in tumors [7]. The occurrence, growth, and progression of tumors are critically affected by the mutual effects between host tumor cells and stromal cells. However, the stromal components of different tumors vary widely. The understanding of the mechanism of crosstalk among tumors is still at a low level [7]. In the meantime, several studies keep close tabs on how immune-related cells impact tumor occurrence, growth, and progression. An increasing number of research studies reveal that TICs acted as an up-and-coming indicator for the understanding and therapeutic effects of TME [8]. Studies have shown that osteosarcoma's immune environment is primarily composed of T-lymphocytes and macrophages. Osteosarcoma cells can control the recruitment, differentiation, and development of immune-infiltrating cells, which results in a local environment of immune tolerance. This kind of environment is favorable to the development of tumors, the resistance of drugs, and even metastases [9, 10]. Therefore, to properly demonstrate the mechanism of TME immune and stromal components regulation, precise genetic analysis is a research hotspot as well as a challenge.

In our study, CIBERSORT and ESTIMATE methodology is used to count on the proportions and composition of the components of immunity and stroma of OS patients from the TARGET database and selected interesting biomarker TREM2. Several researchers identified the TREM2 receptor as a dominating signaling hub of pathology-induced immunity, which can sense tissue damage and activate robust remodeling immunity as responding to it [11]. By playing a part in tumor-associated macrophages (TAMs) and myeloid-derived suppressor cells (MDSCs), TREM2 participated in facilitating an immune-suppressive TME in numerous cancers, including lung cancer, gastric cancer, and glioma [11–13]. In our study, embarking on a comparison between components of immunity and stroma in TME, differentially expressed genes (DEGs) are generated, which revealed that TREM2 may be a potential biomarker of TME remodeling in osteosarcoma.

2. Data Collection and Processing

2.1. Data Source. All data of transcriptome RNA-seq of 101 OS samples and clinical data (including age and sex) of 253 clinical cases are downloaded and collated from the TARGET database (<https://ocg.cancer.gov/programs/target>) on May 24, 2022. In the genetic screening phase, we used all transcriptomic data, but in the prognostic analysis, we used only those data that had both transcriptomic data and survival status (survival status and survival time). After integration, a total of 95 samples had both transcriptome and survival data. There were 55 men and 40 women. Eleven patients were younger than 10 years, 62 were between 10 and 18 years, and 22 were older than 18 years.

2.2. Calculation of Three Kinds of Score. To estimate the components of immunity and stroma in TME for every sample, the ESTIMATE algorithm is loaded with estimate package [14] in R software (version 4.2.0). The three kinds of scores (ImmuneScore, StromalScore, and ESTIMATEScore) increase with the increase of each of the three levels (immunity-related, stroma-related, and the summation of both), respectively. The larger the scores are, the higher the respective composition of the corresponding TME components is.

2.3. Survival Analysis. We combined the three kinds of scores in TME with survival information of OS patients, using the Limma package in R. On account of the median value immune score and stromal score, ninety-five OS patients are split into two different groups, low- and high-score groups, respectively. Using the survival and survminer packages in R software, survival and survminer analyses are calculated. Survival curves are plotted using the Kaplan–Meier methodology, and statistical significance is ascertained by log-rank test; $P < 0.05$ is accepted as significant statistically.

2.4. Identification of Differently Expression Genes between the Low and High Groups. The median value allows the sample to be divided into two equal parts, so we use it as the split line. One hundred and one patients are distinguished as low or high scores, respectively, in comparison with the median ImmuneScore and StromalScore values. Differences between high- and low-scoring samples are achieved by using the R and limma package, also low and high subgroups are compared to obtain the corresponding differentially expressed genes. Genes with $FDR < 0.05$ and \log_2 -transformed fold change > 1 (high subgroup/low subgroup) are regarded as significantly differentially expressed genes.

2.5. Enrichment Analyses of GO and KEGG. The enrichplot, clusterProfiler, and ggplot2 packages of R software are used to classify 118 DEGs according to genomic annotation information, i.e., gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG). Genomic annotation information with both p and q values < 0.05 are regarded as an important and statistically significant role in the development and progression of osteosarcoma.

2.6. Heatmaps. R with the heatmap package is applied to establish the heatmap of DEGs.

2.7. The Difference Analysis of Scores with Clinical Characteristics. Data on clinical information of OS patients are also of interest. R software is used to perform statistical analysis, and Wilcoxon or Kruskal–Wallis rank sum tests are used to determine whether there are statistical differences between clinical indicators between two groups.

2.8. Establishment of a PPI Network. PPI networks reveal the interactions between proteins, and we chose to use the STRING database to construct the corresponding network graphs. What is worth mentioning is that the nodes used to set up the network contain only those nodes whose confidence level of interaction is greater than 0.9.

2.9. Analysis of COX Regression. Univariate Cox regression analysis allows initial screening out of potentially non-significant variables, which is achieved through R software and survival package. As shown, those ascertained and significant genes met $p < 0.05$ in both analyses of univariate Cox and Kaplan–Meier tests.

2.10. Gene Set Enrichment Analysis. Briefly, GSEA can determine the contribution of a predefined gene set to the phenotype, our gene set is all transcriptomic data as described previously, and the analysis is based on the C7 and HALLMARK target sets (v6.2). Just gene sets with corrected $p < 0.05$ and FDR $q < 0.05$ are regarded as significant sets. All GSEA analyses are performed on GSEA-4.2.3 software.

2.11. TIC Profile. The TIC abundance profiles of tumor samples can reflect the immune cell composition in osteosarcoma to some extent and can be calculated by CIBERSORT. The calculated results are screened, and only samples of $p < 0.05$ are retained for subsequent analysis.

2.12. Statistical Analysis. All statistical analyses were conducted by R software (version 4.1.3). The Wilcoxon test was used to compare the differences between the two groups. p value < 0.05 was considered statistically significant.

3. Result

3.1. Analysis Process of This Study. This study can be divided into two major steps: the discovery of TREM2 and the follow-up study of TREM2 (Figure 1). First, osteosarcoma tissue consists of tumor cells and stromal cells, which correspond to ImmuneScore and StromalScore. Each score is used to divide samples into two groups, respectively, using the median value as the cut-off value, and the intersection of DEGs between the high and low groups of each score is used for subsequent PPI and regression analysis, while the intersection of PPI and regression analysis results in turn, eight key genes (ITGAM, HLA-DMA, LY96, C1QA, C1QB, C1QC, TREM2, and C3AR1) are identified. TREM2 is used as our gene of interest for subsequent studies including survival analysis, GSEA, and analysis of immune-related functions.

3.2. Scores Are Associated with OS Patient Survival and Clinical Characteristics. An important indication of whether the immune and stromal ratios are significant in patients with osteosarcoma is the relationship with survival, so we performed a Kaplan–Meier analysis of three kinds of scores, and not surprisingly, the scores correlated positively with

survival (Figures 2(a) and 2(b)). To assess the combined composition of two components in TME, we add ImmuneScore and StromalScore to get ESTIMATEScore (Supplement Table 1). Despite the result showing there is no significant correlation between ESTIMATEScore and the overall survival rate (Figure 2(c)), its p value is still less than 0.1. These entire results implied that the composition of TME is clinically important and the compositional aspects of TME can forecast patients' prognosis of OS, especially immune and stromal components.

In addition to the survival rate, it is worth discussing whether these three kinds of scores are correlated with other clinical indicators such as age and gender. The results indicated that gender is significantly correlated with ImmuneScore and ESTIMATEScore in patients ($P < 0.05$, Figures 2(d) and 2(f)), except StromalScore ($p = 0.1$, Figures 2(e)), while age is not significantly correlated with any score ($p > 0.05$, Figures 2(g)–2(i)). We found that scores in female patients are higher than in male patients.

3.3. Immune-Related Genes Are Mainly Shared DEGs between the ImmuneScore and StromalScore. Analysis of comparing patients between low and high scores is executed to ascertain if there are definitive genetic profile alterations of components of immunity and stroma in TME. Eight hundred and ninety DEGs (Five hundred and twenty-nine downregulated and three hundred and sixty-one upregulated genes) are received by comparing two groups (low- and high-ImmuneScore patients), with the median value as the cut-off (Figures 3(a), 3(c), 3(d), Supplement Table 2). Correspondingly, five hundred and thirty-one DEGs (Two hundred and twenty-four downregulated and three hundred and seven upregulated) are received from the StromalScore (Figures 3(b)–3(d), Supplement Table 3). Furthermore, twenty-nine low-score downregulated genes and eighty-nine high-score upregulated genes are cross-linked between the ImmuneScore and StromalScore by an analysis of Venn diagrams (Figures 3(c) and 3(d), Supplement Table 4). The entire DEGs (118 genes in all) are deemed as determinants of status in TME. GO enrichment analysis results give evidence that the DEGs ordinarily have a corresponding in terms linked to immunity, including innate and acquired immunity (Figure 3(e), Supplement Figure 1A and 1B). The KEGG enrichment analysis similarly gives evidence of that DEG enrichment in the disease spectrum is related to the immune system, including infection and autoimmune disease (Figure 3(f), Supplement Figure 1C and 1D). Hence, the overall function of differentially expressed genes appears to have a corresponding immune-related event, hinting the participation of immunity-related elements is a principal signature in the TME of OS.

3.4. Cross-Tabulation Analysis between Univariate Cox Regression and PPI Network. To move forward a single step in exploring the latent mechanism, we worked with Cytoscape software to set up the PPI network in the STRING database. Figure 4(a) show the mutual interplay among the 118 genes, and ranked in the top thirty genes are listed in the picture as

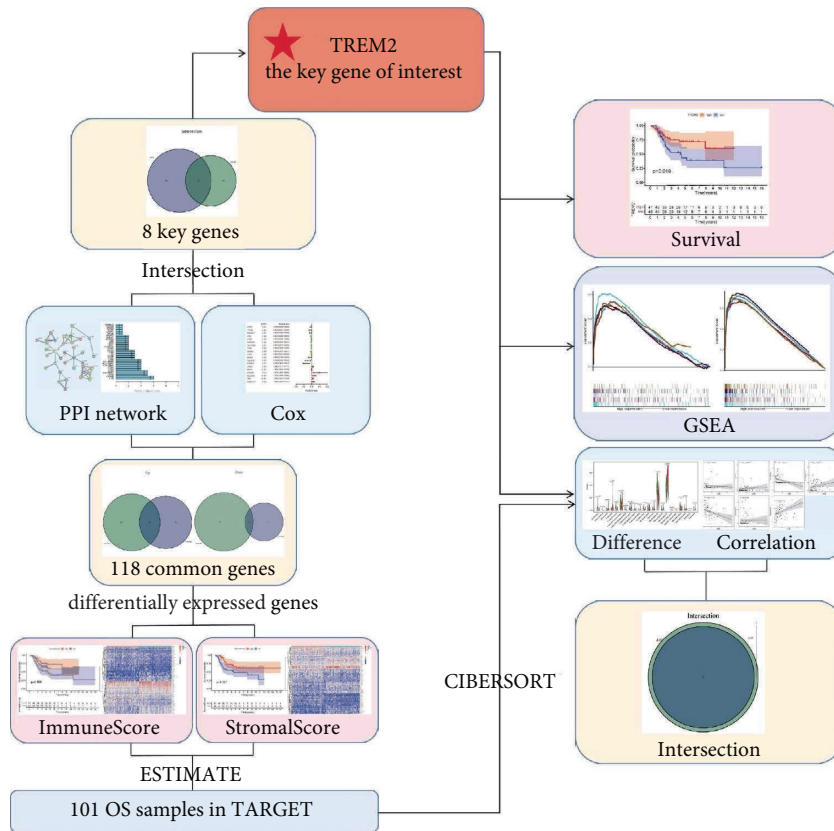


FIGURE 1: Schematic diagram of the study design.

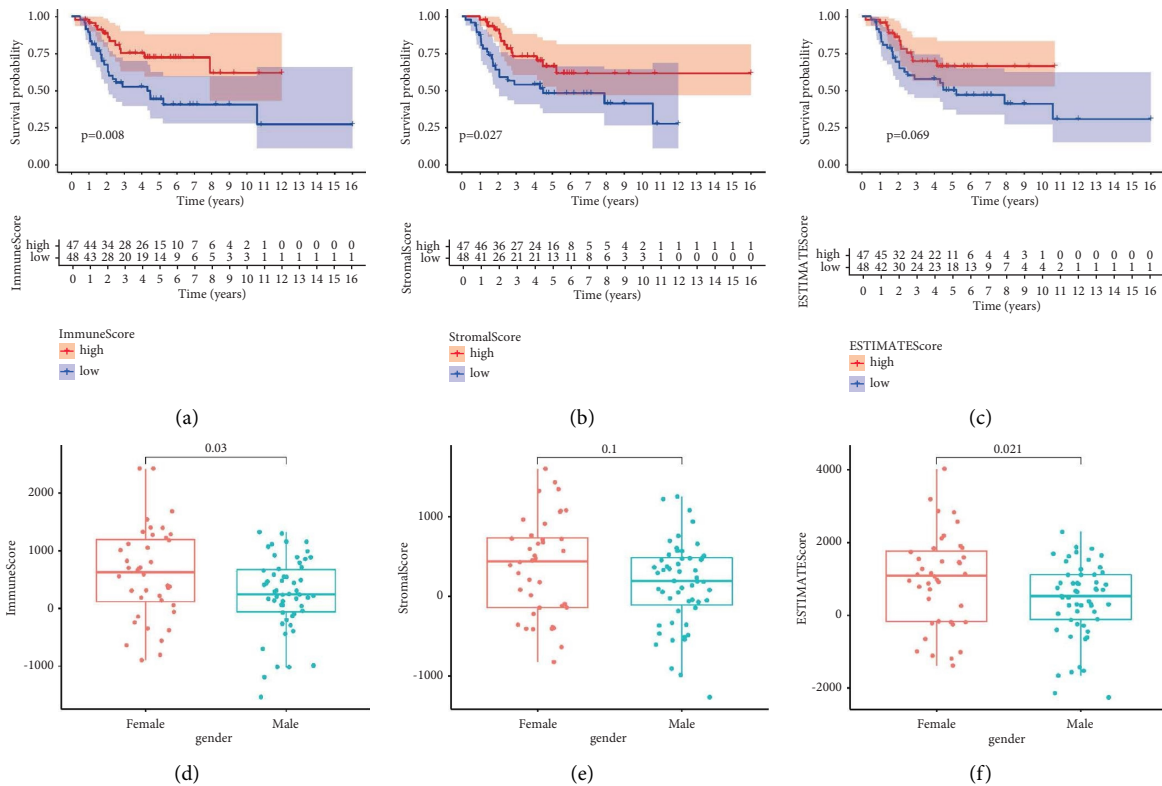


FIGURE 2: Continued.

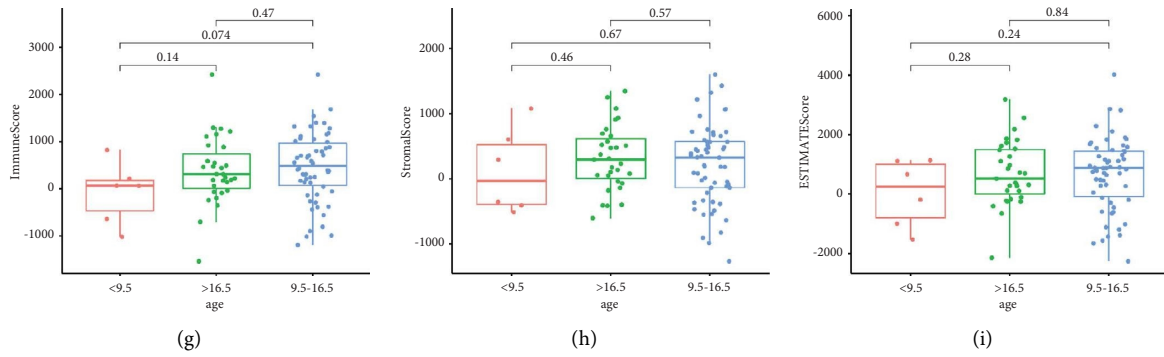
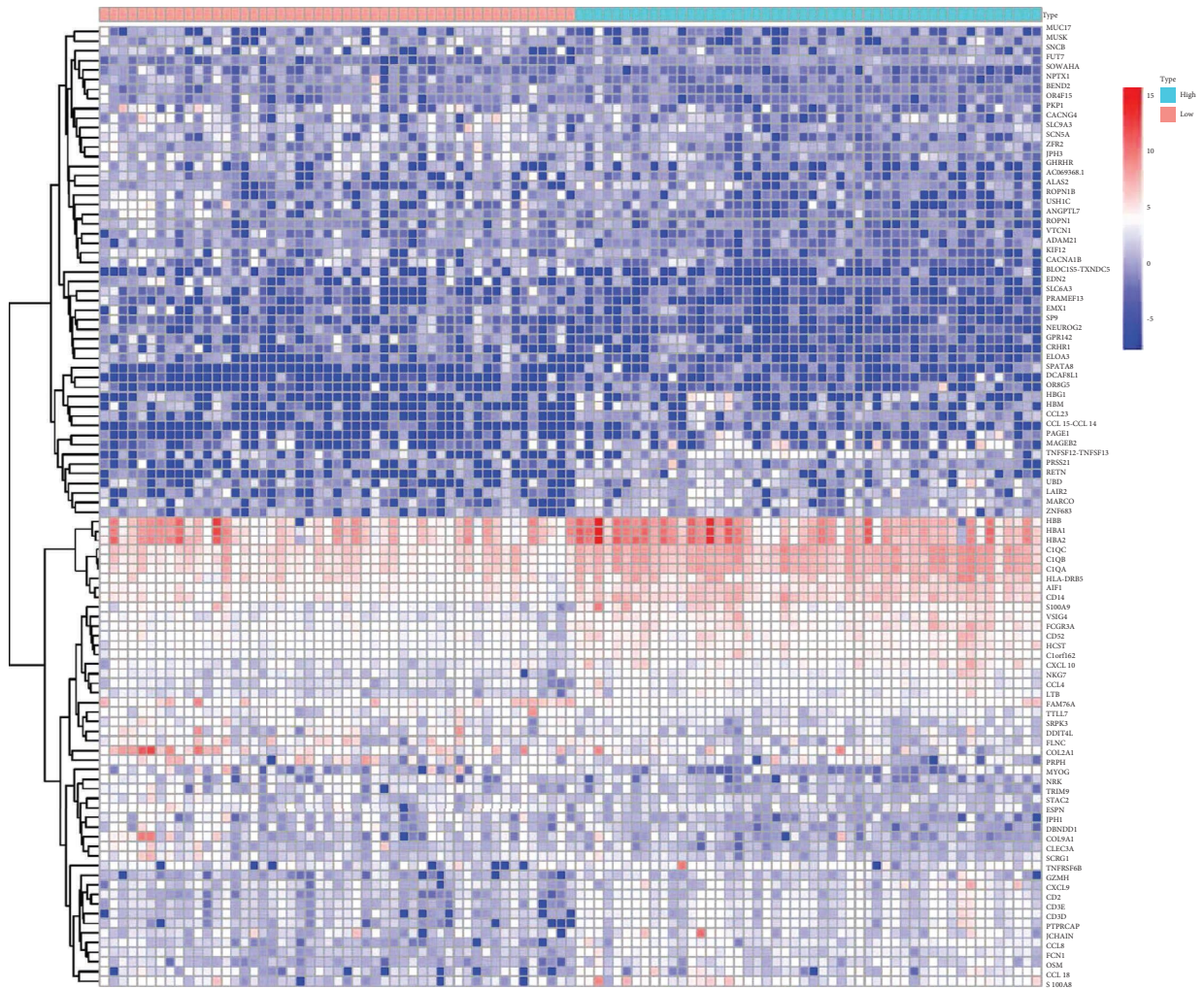
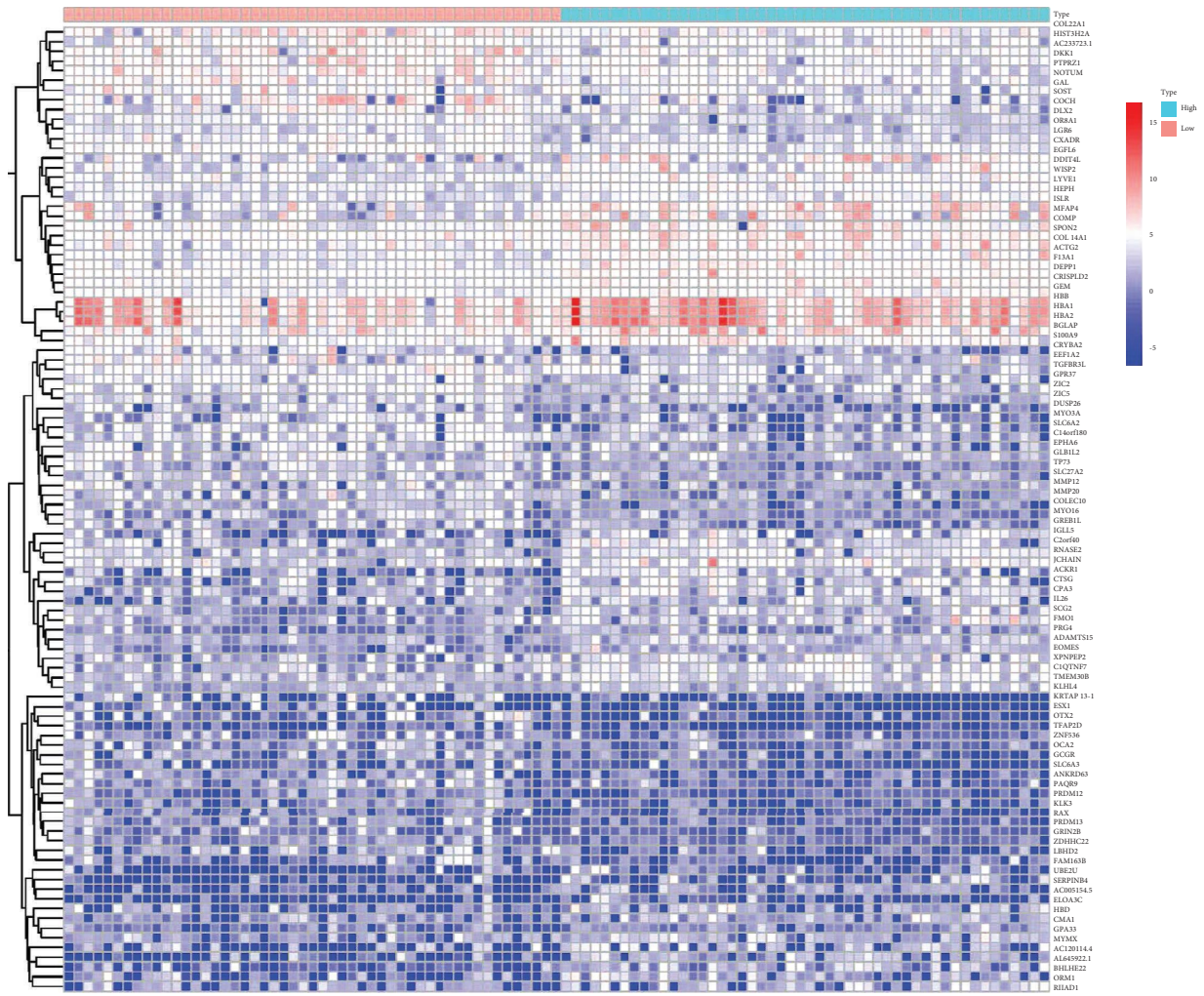


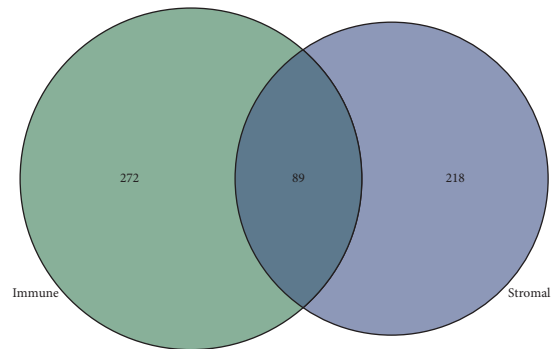
FIGURE 2: Correlation of scores with the survival and clinical characteristics of patients with osteosarcoma (OS). (a–c) Kaplan–Meier survival analysis for OS patients grouped into high or low score in ImmuneScore, StromalScore, and ESTIMATEScore determined by the comparison with the median, respectively. $p = 0.008, 0.027, 0.069$ by log-rank test, respectively. (d–f) Distribution of ImmuneScore, StromalScore, and ESTIMATEScore in gender. The $p = 0.03, 0.1,$ and $0.021,$ respectively, by Kruskal–Wallis rank sum test. (g–i) Distribution of ImmuneScore, StromalScore, and ESTIMATEScore in age. The p value was calculated by Kruskal–Wallis rank sum test as shown in this figure.



(a)
FIGURE 3: Continued.

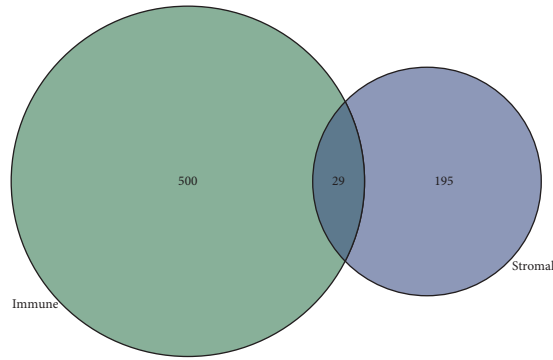


(b)

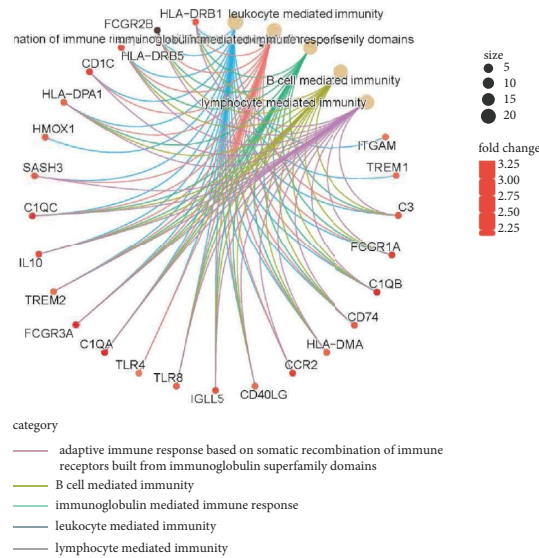


(c)

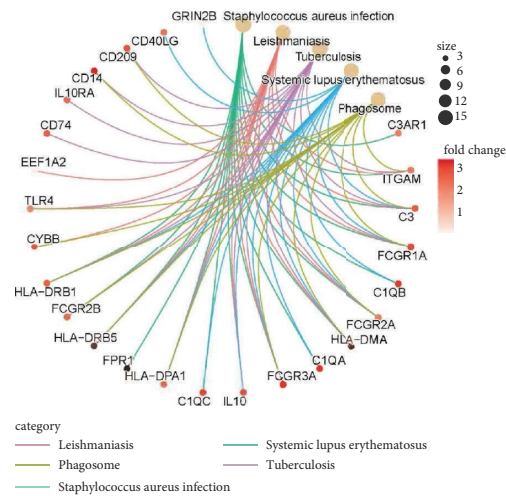
FIGURE 3: Continued.



(d)



(e)



(f)

FIGURE 3: Heatmaps, Venn plots, and enrichment analysis of GO and KEGG for DEGs. (a, b) Heatmap for DEGs generated by comparison of the high-score group vs. the low-score group in ImmuneScore and StromalScore, respectively. Row name of heatmap is the gene name, and column name is the ID of samples which not shown in plot. Differentially expressed genes were determined by Wilcoxon rank sum test with $q = 0.05$ and fold change > 1 . (c, d) Venn plots showing common upregulated and downregulated DEGs shared by ImmuneScore and StromalScore, and $q < 0.05$ and fold change > 1 as the DEGs significance filtering threshold. (e, f) GO and KEGG enrichment analysis for 118 DEGs, terms with p and $q < 0.05$ were considered to be enriched significantly.

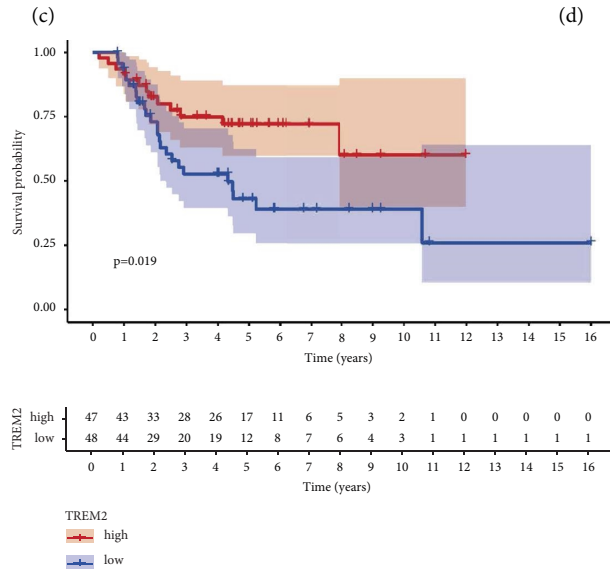
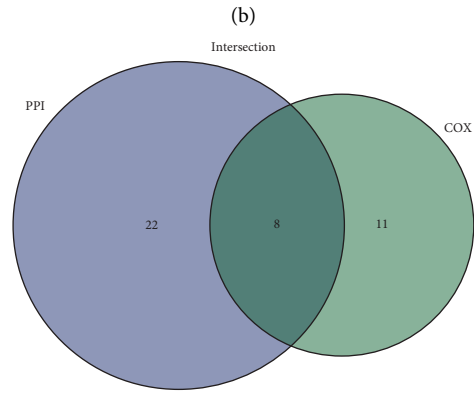
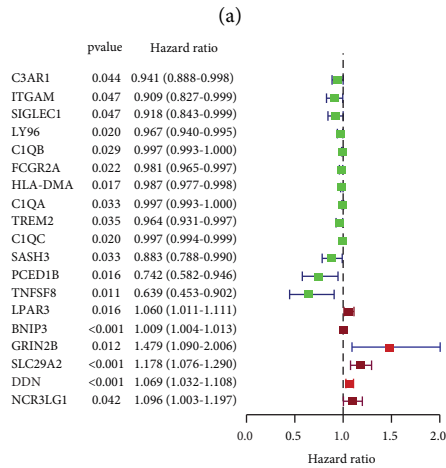
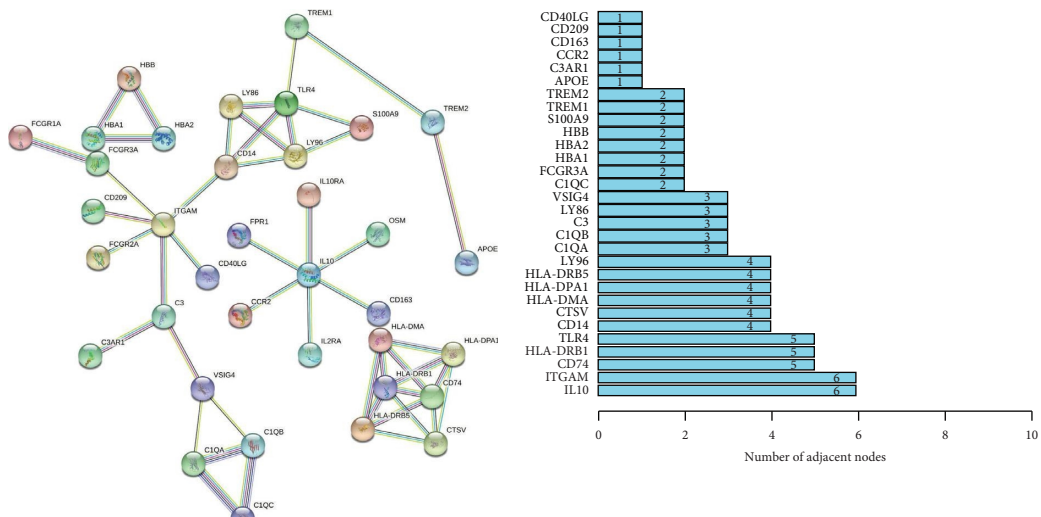
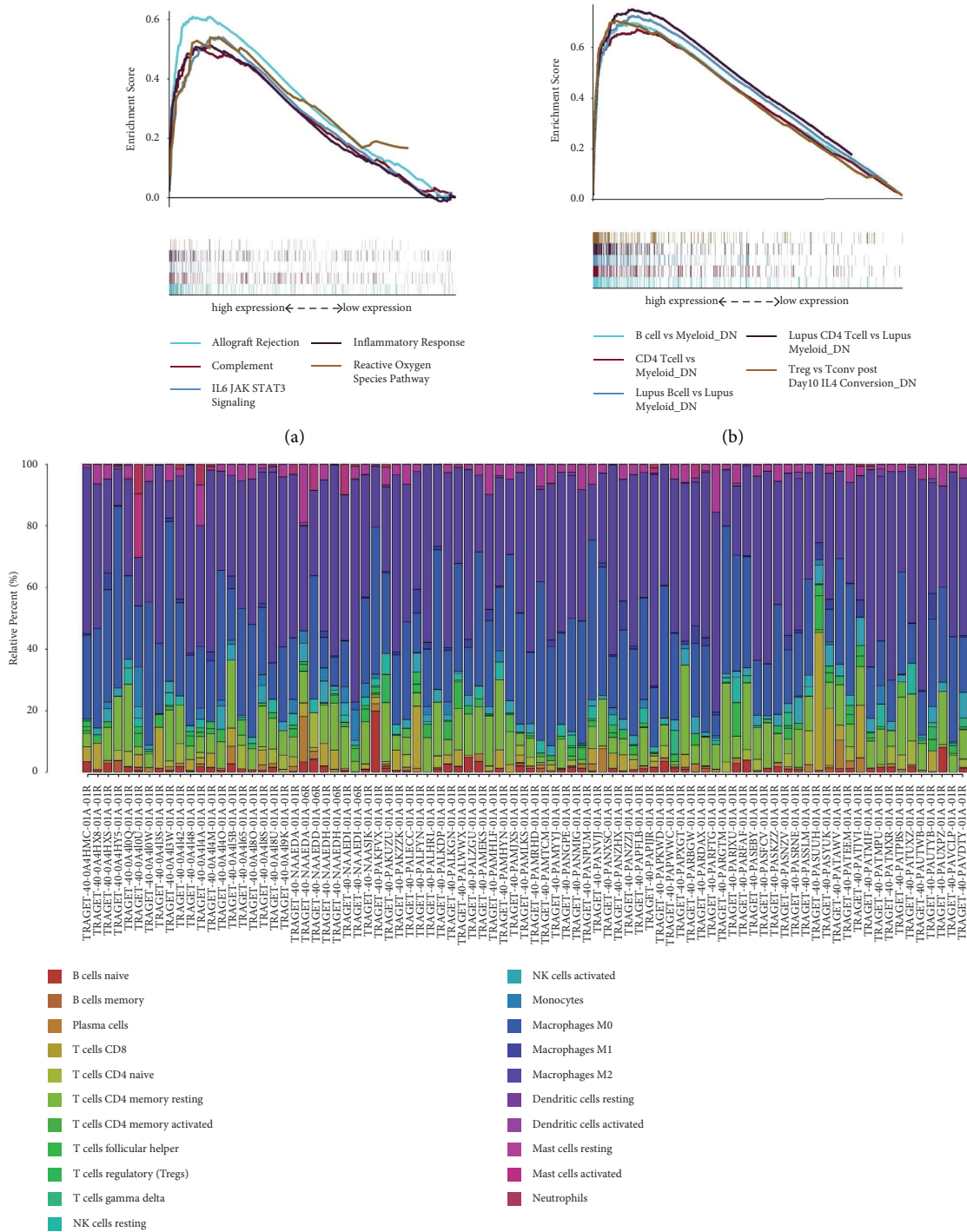


FIGURE 4: Protein-protein interaction network and univariate Cox analysis. (a) Interaction network constructed with the nodes with interaction confidence value >0.90. (b) The top 30 genes ordered by the number of nodes. (c) Univariate Cox regression analysis with 118 DEGs, listing the top significant factors with $p < 0.05$. (d) Venn plot showing the common factors shared by leading 30 nodes in PPI and top significant factors in univariate Cox. (e) Survival analysis for OS patients with different TREM2 expression. $p = 0.019$ by log-rank test.



(c)
FIGURE 5: Continued.

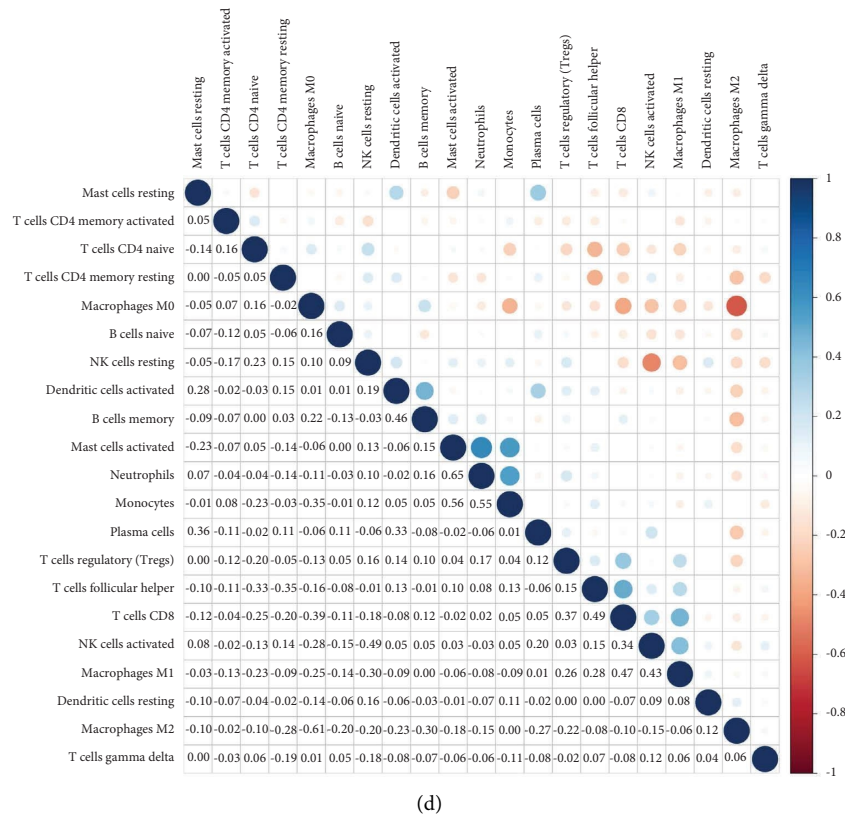


FIGURE 5: GSEA for samples with high TREM2 expression and low expression and TIC profile and correlation analysis in tumor samples. (a, b) The enriched gene sets in Hallmark and C7 sets by the high TREM2 expression sample. (c) Bar plot showing the proportion of 21 kinds of TICs in OS tumor samples. (d) Heatmap showing the correlation between 21 kinds of TICs and numeric in each tiny box indicating the p value of correlation between two kinds of cells. The shade of each tiny color box represented corresponding correlation value between two cells, and Pearson coefficient was used for significance test.

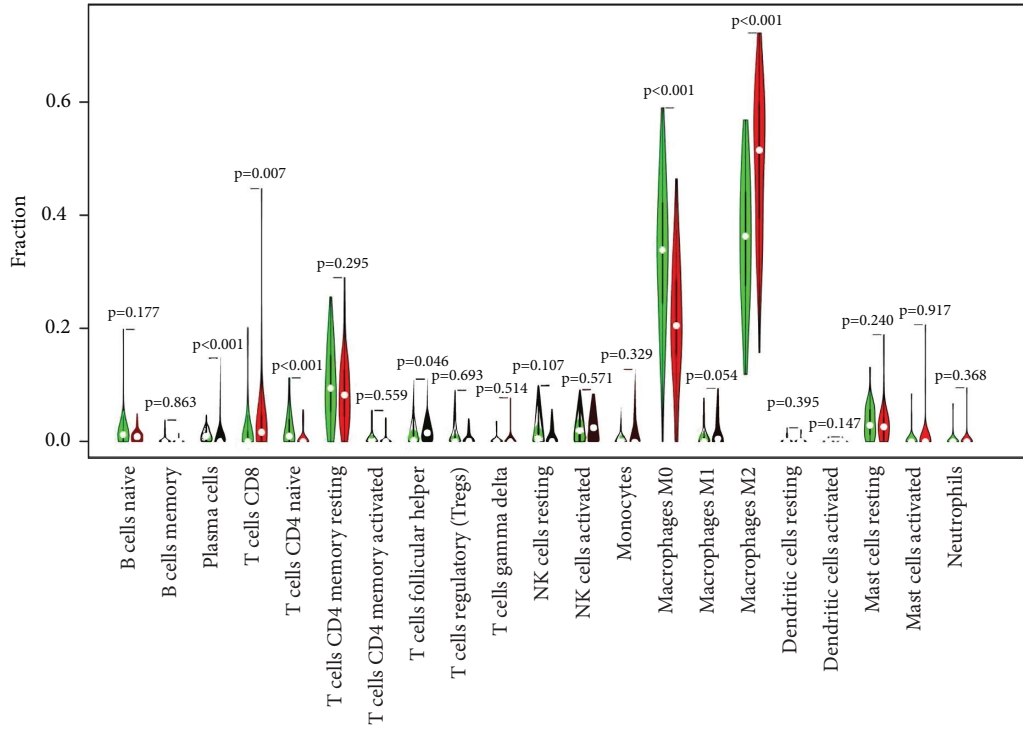
rank order (Figure 4(b)). The vital factors impacting the survival of OS patients among 118 DEGs are selected by applying univariate COX regression analysis (Figure 4(c)). Then, these intersecting sets between the core nodes of PPI and the top nineteen Cox regressors is carried out, and eight superimposed factors are in place, which are identified (ITGAM, HLA-DMA, LY96, C1QA, C1QB, C1QC, TREM2, and C3AR1, Figure 4(d)).

3.5. Relationships between TREM2 and Survival Time and Clinical Characteristics in OS Patients. Based on previous report, we chose TREM2 for further study [15]. According to the median expression of TREM2 gene, we separated the OS patients into two groups, low- and high-expression TREM2 expression groups. There is a significant difference of survival rate statistically between two groups by the high TREM2 expression group has a higher survival rate than patients with corresponding low expression (Figure 4(e)). What's more, there is no statistical difference between TREM2 expression and clinical characteristics (Supplement Figure 2).

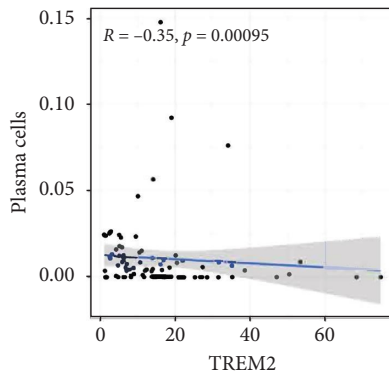
3.6. TREM2 as a Potential Indicator of TME Remodeling. Taking the fact that the levels of TREM2 expression have positive correlation with OS patient survival into consideration, these two groups are in comparison in GSEA.

Hallmark and C7 sets of both demonstrated that the groups with high expression of TREM2 have observably more enrichment in immunity-related gene sets, suggesting immunity-related functions, such as the complement response, allograft rejection, IL6/JAK/STAT3 signaling, and acquired immunity are substantially more vibrant (Figures 5(a) and 5(b)). Therefore, it is implicit that the status of the TME can be mirrored by the TREM2 expression.

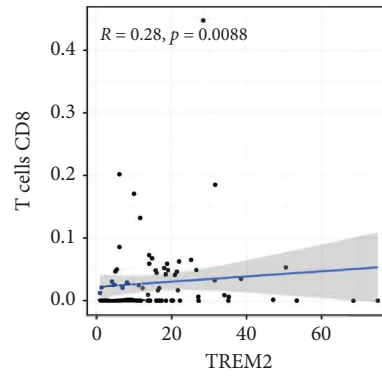
3.7. Correlation Analysis of the Levels of TREM2 Expression and TICs. To move forward a single step in confirming the relevance between TME and expression levels of TREM2. CIBERSORT methodology is utilized to acquire the immune subpopulation composition of tumor-infiltrating. The establishment of twenty-two kinds of immunity-related cell profiles is executed as follows (Figure 5(c)), and the relevance among TICs is figured up (Figure 5(d)). The discrepancy and connection between the expression of TREM2 expression and the proportions of TICs are analyzed. Six kinds of TREM2-related TICs are obtained (Figures 6(a)–6(h)). Of these, three types of TICs are associated positively with the expression of TREM2, including CD8-positive T cells, follicular helper T cells, and M2 macrophages. Three types of TICs, including plasma cells, naive CD4-positive T cells, and M0 macrophages, are associated



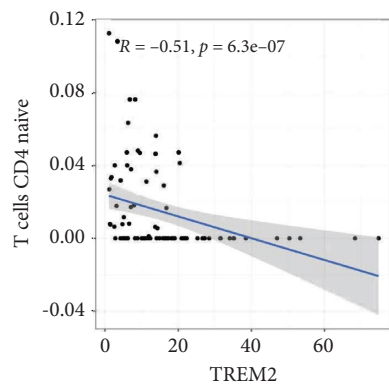
(a)



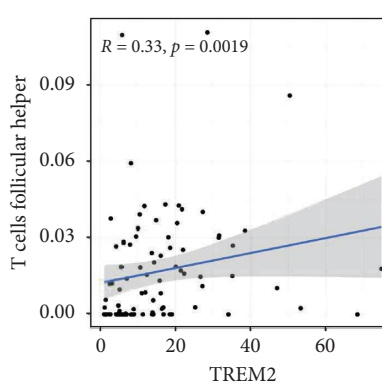
(b)



(c)



(d)



(e)

FIGURE 6: Continued.

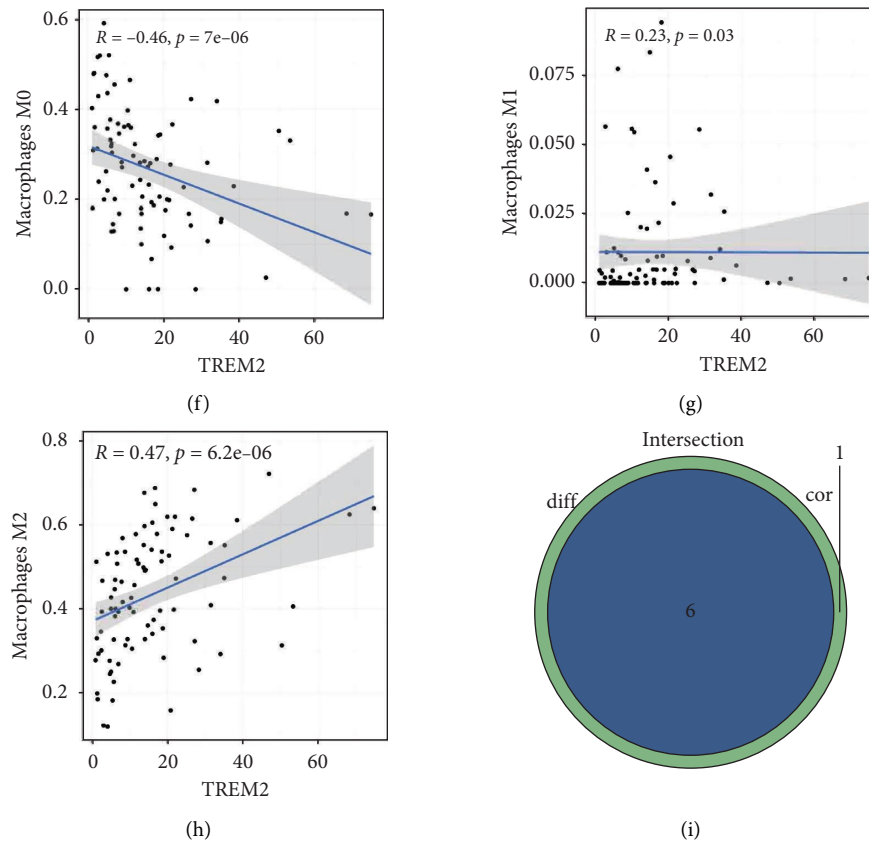


FIGURE 6: Correlation of TICs proportion with TREM2 expression. (a) Violin plot showed the ratio differentiation of 21 kinds of immune cells between OS tumor samples with low or high TREM2 expression relative to the median of TREM2 expression level, and Wilcoxon rank sum was used for the significance test. (b–h) Scatter plot showed the correlation of 7 kinds of TICs proportion with the TREM2 expression ($p < 0.05$). The red line in each plot was fitted linear model indicating the proportion tropism of the immune cell along with TREM2 expression, and Pearson coefficient was used for the correlation test. (i) Venn plot displayed eight kinds of TICs correlated with TREM2 expression codetermined by difference and correlation tests displayed in violin and scatter plots, respectively.

negatively with the expression of TREM2. What's more, there are 6 kinds of TICs of TREM2 expression (Figure 6(i)). These findings are a further indication of the effect of TREM2 expression levels on TME immunoactivities.

4. Discussion

In the current study, genes of the tumor microenvironment that related to the survival of OS patients from the TARGET database are what we attempted to appraise. TREM2 is appraised to be engaged in immunity-related activities. More significantly, a battery of research on bioinformatics revealed that TREM2 is a prognostic biomarker for osteosarcoma microenvironment remodeling.

The tumor microenvironment played a pivotal part in tumorigenesis and its progression. It is of strategic meaning to detect the underlying therapeutic targets which can contribute to the remodeling and facilitating the transition of the tumor microenvironment from a developmental state to an inhibitory state.

Numerous research studies had elucidated the significance of tumor microenvironment in tumorigenesis [16]. In osteosarcoma tumor microenvironment-related literature,

we take notice that the connection between the score of immunity and survival state has been investigated, and C3AR1, PPARG, PDK1, IGHG3, and C1Q are recognized as prognostic biomarkers [17–19]. The immune components in TME sever the purpose of the prognosis of patients by analyzing the OS data in the TARGET database. In particular, the composition of immunity and stroma in TME has a strong correlation with the overall survival in OS patients. These results demonstrated and emphasized the importance of pursuing the connection between stromal cells and tumor cells, which will give a novel perception for discovering and developing more efficient therapy. What's more, this paper also substantiated that TICs have relevance with the clinical prognostic outcome of OS [20]. The relevance offers a brand new theoretical footstone for the evolution of more efficient immunotherapeutic methods.

For the last few years, tremendous progress was acquired in immunotherapy, and the inhibition of immune checkpoint inhibitors (ICIs) in OS made significant progress [21, 22]. However, the inhibition of immune checkpoint inhibitor (ICI) immunotherapy for osteosarcoma (OS) is severely restricted by the lacking of immunogenicity and poor T cell infiltration [23, 24]. Therefore, the immunotherapy of OS

is in urgent need of some novel candidate exploitation. Here, the decreased expression of TREM2 has a significant association with poor prognosis by analyzing the transcriptomic of OS in the TARGET database. Consequently, we will center on the relevance between the expression of TREM2 and TME to supply a novel treatment idea for OS immunotherapeutic methods.

As a dominating signaling hub of pathology-induced immunity, TREM2 catch the attention of the leading role of myeloid cells in various pathological processes which can mediate immunosuppression [25]. Many markers of tissue injury are ligands for the TREM2 receptor, and binding of the TREM2 receptor and ligand contributes to cell survival and resistance to inflammation, affecting cell phenotype by regulating phagocytosis and metabolism [15]. In cancer research, TREM2 is observed in macrophages beyond 200 cancer cases of humans in fostering an immune-suppressive TME [15]. There, TREM2 is perhaps a biomarker to alter tumor bone marrow infiltration and reinforce immunotherapy of ICIs [25].

CIBERSORT methodology is applied to accomplish the analysis of the proportion of TICs and completed the composition of twenty-two profiles of immune cells. The results exhibited that macrophages accounted for the highest proportion in the TME of OS, especially M2 macrophages. The fraction of M2 macrophages in high expression groups of TREM2 is higher, which may have a relation to the immune-suppressive TME. In addition, CD8-positive T cells and follicular helper T cells have a positive correlation with the differential expression of TREM2. Plasma cells, naive CD4-positive T cells, and M0 macrophages have a negative correlation with the expression of TREM2. All results suggest that the differential expression of TREM2 is linked to the levels of immune cell infiltration and is a critical target for ameliorating the prognosis of OS. As an attractive biomarker for modulation of individual immunotherapy who are intractable to therapy of ICIs and have a TME rich in TAM, TREM2 is tightly associated with TAMs [26].

Applying the ESTIMATE algorithm, functional enrichment analysis is applied to acquire the gene of the tumor microenvironment of OS in the TARGET database. TREM2 catches our eye as a potential prognostic biomarker for OS patients. What is of interest is that, although TREM2 may mediate the immunosuppressive tumor microenvironment through macrophage M2 polarization, its expression level has positive relevance with the overall survival time of osteosarcoma patients. Further research is indispensable to disclose the mechanism of regulating and exploit novel immunotherapeutic strategies.

Data Availability

The datasets analyzed for this study can be found in the TARGET database (<https://ocg.cancer.gov/programs/target>).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

SZL conceived and designed this study. SZL, CZY, and JY performed data processing and analysis and wrote the manuscript. JH, ZZP, and YH performed a literature search. LB, XW, and XJR performed a review and specialty consultation. All authors read and approved the final manuscript. Zhi-Long Shen, Zhao-Yu Chen, and Yong Ji contributed equally to this work and share the first authorship.

Supplementary Materials

The Supplementary Figures and Tables for this article can be found online at: <https://doi.org/10.6084/m9.figshare.21972614.v1>. Supplementary Figure 1: the GO and KEGG enrichment analysis. (A, B) The GO enrichment analysis. (C, D) The KEGG enrichment analysis. Supplementary Figure 2: relationship between TREM2 expression and clinical features. (A) Relationship between TREM2 expression and age. (B) Relationship between TREM2 expression and gender. Supplementary Table 1: immune infiltration score. The row is the ID of the sample and the column is the score, including StromalScore, ImmuneScore, and ESTIMATEScore. Supplementary Table 2: differentially expressed genes between high and low ImmuneScore groups. The row is the gene name, and the column is the fold change after taking the logarithm with the base of 2 and the adjusted *P* value in turn. Supplementary Table 3: differentially expressed genes between high and low StromalScore groups. The row is the gene name, and the column is the fold change after taking the logarithm with the base of 2 and the adjusted *P* value in turn. Supplementary Table 4: common differentially expressed genes shared by ImmuneScore groups and StromalScore groups. The row is the gene name and the column is the fold change after taking the logarithm with the base of 2. (*Supplementary Materials*)







References

- [1] L. Kager, G. Tamamyan, and S. Bielack, "Novel insights and therapeutic interventions for pediatric osteosarcoma," *Future Oncology*, vol. 13, no. 4, pp. 357–368, 2017.
- [2] A. Biazzo and M. De Paolis, "Multidisciplinary approach to osteosarcoma," *Acta Orthopaedica Belgica*, vol. 82, 2016.
- [3] C. Yang, Y. Tian, F. Zhao et al., "Bone microenvironment and osteosarcoma metastasis," *International Journal of Molecular Sciences*, vol. 21, no. 19, p. 6985, 2020.
- [4] L. Mirabello, R. J. Troisi, and S. A. Savage, "Osteosarcoma incidence and survival rates from 1973 to 2004: data from the surveillance, epidemiology, and end results program," *Cancer*, vol. 115, no. 7, pp. 1531–1543, 2009.
- [5] Y. Zhou, D. Yang, Q. Yang et al., "Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma," *Nature Communications*, vol. 11, no. 1, p. 6322, Dec 10 2020.
- [6] Y. Xiao and D. Yu, "Tumor microenvironment as a therapeutic target in cancer," *Pharmacology & Therapeutics*, vol. 221, Article ID 107753, 2021.
- [7] K. M. Bussard, L. Mutkus, K. Stumpf, C. Gomez-Manzano, and F. C. Marini, "Tumor-associated stromal cells as key

- contributors to the tumor microenvironment,” *Breast Cancer Research*, vol. 18, no. 1, p. 84, 2016.
- [8] G. S. T, “Innate and adaptive immune cells in Tumor microenvironment,” *Gulf J Oncolog*, vol. 1, no. 35, pp. 77–81, 2021.
- [9] F. Cersosimo, S. Lonardi, G. Bernardini et al., “Tumor-associated macrophages in osteosarcoma: from mechanisms to therapy,” *International Journal of Molecular Sciences*, vol. 21, no. 15, p. 5207, 2020.
- [10] M. F. Heymann, F. Lezot, and D. Heymann, “The contribution of immune infiltrates and the local microenvironment in the pathogenesis of osteosarcoma,” *Cellular Immunology*, vol. 343, Article ID 103711, Sep. 2019.
- [11] D. L. Kober and T. J. Brett, “TREM2-Ligand interactions in health and disease,” *Journal of Molecular Biology*, vol. 429, no. 11, pp. 1607–1629, Jun 2 2017.
- [12] Y. Yao, H. Li, J. Chen et al., “TREM-2 serves as a negative immune regulator through Syk pathway in an IL-10 dependent manner in lung cancer,” *Oncotarget*, vol. 7, no. 20, Article ID 29620, 2016.
- [13] X. Zhang, W. Wang, P. Li, X. Wang, and K. Ni, “High TREM2 expression correlates with poor prognosis in gastric cancer,” *Human Pathology*, vol. 72, pp. 91–99, Feb. 2018.
- [14] R. C. Team, “A language and environment for statistical Computing. Vienna: R foundation for statistical computing,” 2018, <https://www.R-project.org/>.
- [15] A. Deczkowska, A. Weiner, and I. Amit, “The physiology, pathology, and potential therapeutic applications of the TREM2 signaling pathway,” *Cell*, vol. 181, no. 6, pp. 1207–1217, 2020.
- [16] N. M. Anderson and M. C. Simon, “The tumor microenvironment,” *Current Biology*, vol. 30, no. 16, pp. R921–R925, 2020.
- [17] L. H. Chen, J. F. Liu, Y. Lu, X. Y. He, C. Zhang, and H. H. Zhou, “Complement C1q (C1qA, C1qB, and C1qC) may Be a potential prognostic factor and an index of tumor microenvironment remodeling in osteosarcoma,” *Frontiers in Oncology*, vol. 11, Article ID 642144, 2021.
- [18] T. Zou, W. Liu, Z. Wang et al., “C3AR1 mRNA as a potential therapeutic target associates with clinical outcomes and tumor microenvironment in osteosarcoma,” *Frontiers of Medicine*, vol. 8, Article ID 642615, 2021.
- [19] C. Zhang, J. H. Zheng, Z. H. Lin et al., “Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma,” *Aging*, vol. 12, no. 4, pp. 3486–3501, 2020.
- [20] X. Li, Y. Chen, X. Liu et al., “Tim3/Gal9 interactions between T cells and monocytes result in an immunosuppressive feedback loop that inhibits Th1 responses in osteosarcoma patients,” *International Immunopharmacology*, vol. 44, pp. 153–159, 2017.
- [21] L. Xie, J. Xu, X. Sun et al., “Apatinib plus camrelizumab (anti-PD1 therapy, SHR-1210) for advanced osteosarcoma (APFAO) progressing after chemotherapy: a single-arm, open-label, phase 2 trial,” *Journal Immunother Cancer*, vol. 8, no. 1, Article ID e000798, 2020.
- [22] K. Boye, A. Longhi, T. Guren et al., “Pembrolizumab in advanced osteosarcoma: results of a single-arm, open-label, phase 2 trial,” *Cancer Immunology Immunotherapy*, vol. 70, no. 9, pp. 2617–2624, 2021.
- [23] W. Yu, Y. Wang, J. Zhu et al., “Autophagy inhibitor enhance ZnPc/BSA nanoparticle induced photodynamic therapy by suppressing PD-L1 expression in osteosarcoma immunotherapy,” *Biomaterials*, vol. 192, pp. 128–139, 2019.
- [24] Y. X. Ge, T. W. Zhang, L. Zhou et al., “Enhancement of anti-PD-1/PD-L1 immunotherapy for osteosarcoma using an intelligent autophagy-controlling metal organic framework,” *Biomaterials*, vol. 282, Article ID 121407, Mar. 2022.
- [25] M. Molgora, E. Esaulova, W. Vermi et al., “TREM2 modulation remodels the tumor myeloid landscape enhancing anti-PD-1 immunotherapy,” *Cell*, vol. 182, no. 4, pp. 886–900, 2020.
- [26] M. Binnewies, J. L. Pollack, J. Rudolph et al., “Targeting TREM2 on tumor-associated macrophages enhances immunotherapy,” *Cell Reports*, vol. 37, no. 3, Article ID 109844, 2021.

Research Article

Association of LAMA1 Single-Nucleotide Polymorphisms with Risk of Esophageal Squamous Cell Carcinoma among the Eastern Chinese Population

Shaoyuan Zhang ^{1,2}, Yong Fang^{1,2}, Feng Su^{1,2}, Tian Jiang ^{1,2}, Jinjie Yu ^{1,2}, Siyun Lin,^{1,2} Lu Lv,³ Tao Long,³ Huiwen Pan,³ Junqing Qi,³ Qiang Zhou,⁴ Weifeng Tang ⁵, Guowen Ding,³ Liming Wang,⁶ Lijie Tan ^{1,2} and Jun Yin ^{1,2}

¹Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, Shanghai 200032, China

²Cancer Center, Zhongshan Hospital, Fudan University, Shanghai 200032, China

³Department of Cardiothoracic Surgery, Affiliated People's Hospital of Jiangsu University, Jiangsu 212002, China

⁴Department of Thoracic Surgery, Sichuan Cancer Hospital and Institute, Sichuan 610042, China

⁵Department of Cardiothoracic Surgery, Nanjing Drum Tower Hospital, Jiangsu 210008, China

⁶Department of Respiratory, Shanghai Xuhui Central Hospital, Shanghai 200031, China

Correspondence should be addressed to Lijie Tan; tan.lijie@zs-hospital.sh.cn and Jun Yin; yin.jun2@zs-hospital.sh.cn

Received 4 September 2022; Revised 3 October 2022; Accepted 10 October 2022; Published 14 February 2023

Academic Editor: Song Cao

Copyright © 2023 Shaoyuan Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction. LAMA1, also known as laminin subunit $\alpha 1$, is a member of the laminin family, which is widely reported to be a key basement membrane molecule that affects various biological activities and is associated with many kinds of diseases. We aimed to investigate the association between LAMA1 single-nucleotide polymorphisms and the occurrence and progression of esophageal squamous cell carcinoma in the Chinese population. **Method.** 2,186 participants were collected retrospectively between October 2008 and January 2017, including 1,043 ESCC patients and 1,143 noncancer patients. A 2 mL blood sample was obtained intravenously for the LDR for SNP analysis. The 6 SNP loci of LAMA1 were selected and examined. We analyzed the association of several genetic models of 6 LAMA1 SNP loci, sex, age, smoking and drinking status, and the occurrence of esophageal squamous cell carcinoma. **Results.** In the rs62081531 G > A locus, genotype GA was a protective factor for ESCC compared with GG (OR: 0.830, $P = 0.046$), especially among the younger and nondrinkers. At rs607230 T > C, genotype TC was linked with a lower risk of ESCC compared with TT. (OR: 0.613, $P = 0.034$). Haplotype Frequencies revealed that $A_{rs62081531}G_{rs621993}A_{rs539713}T_{rs566655}A_{rs73938538}C_{rs607230}$ (OR: 0.803, $P = 0.028$) and $G_{rs62081531}G_{rs621993}A_{rs539713}T_{rs566655}C_{rs73938538}C_{rs607230}$ (OR: 0.679, $P = 0.010$) were strongly associated with lower susceptibility of ESCC. **Conclusion.** The LAMA1 rs62081531, rs539713, rs566655, and rs607230 polymorphisms were demonstrated to be related to susceptibility to ESCC in the Chinese population. LAMA1 SNPs may have a significant impact on the occurrence of esophageal cancer and may serve as potential diagnostic biomarkers.

1. Introduction

Esophageal cancer is a prevalent malignant tumor that has a high rate of morbidity and mortality worldwide. According to the Global Cancer Statistics 2020, esophageal cancer is more prevalent in East Asia, particularly in China, as well as West Asia and Africa [1]. There was a clear correlation

between the pathological type of esophageal cancer and its geographic distribution. Squamous cell carcinoma is the most common type of esophageal cancer in developing countries. China has a high prevalence of esophageal squamous cell carcinoma (ESCC) of up to 90%. A variety of pathogenic factors may lead to esophageal squamous cell carcinoma, including smoking, drinking, eating habits, and

viral infections [2, 3]. However, not everyone exposed to these risk factors develops esophageal cancer, suggesting that genetic susceptibility, particularly single nucleotide polymorphisms (SNPs), plays a significant role in the development of esophageal squamous cell carcinoma.

LAMA1 is also known as laminin subunit $\alpha 1$. Laminins are a family of glycoproteins found in the extracellular matrix that comprises the basement membrane [4, 5]. Laminins have a heterotrimeric structure composed of an α , β , and γ chain [6]. Numerous biological processes are known to be directed by them, including cell adhesion, mitogenesis, differentiation, and metastasis, all of which contribute to carcinogenesis [7–9]. Tissue distribution of LAMA1 occurred mainly in early epithelial development and some adult epithelia. Recent reports have shown that laminin-1 acts as an efficient attachment protein for a large variety of cultured cell types *in vitro* [10]. Mutations in the *LAMA1* gene result in a deficiency of the laminin $\alpha 1$ chain, which may lead to tumorigenesis and progression [11]. Recent research indicates that *LAMA1* mutations or overexpression are linked with the occurrence and development of various malignant tumors, including colon cancer, pancreatic cancer, and ovarian cancer [12–15].

However, the relationship between the *LAMA1* single-nucleotide polymorphism and ESCC remains unclear. Through multicenter large-sample case-control research, we aim to thoroughly investigate the association between *LAMA1* single-nucleotide polymorphisms and the incidence and progression of esophageal cancer.

2. Method

2.1. Patients and Study Design. Between October 2008 and January 2017, 2,186 participants were collected from the Affiliated People's Hospital and the Affiliated Hospital of Jiangsu University (Zhenjiang, China). Totally, 1,043 cases of esophageal cancer were diagnosed and histologically confirmed as squamous cell carcinoma by two pathologists independently. Patients with a history of any other types of cancer or with metastasis or those who had received neoadjuvant therapy were excluded. Around the same time, 1,143 noncancer patients from both hospitals were enrolled, with a frequency matching by age (± 5 years) and gender, and the majority of them were admitted for trauma.

For 1,043 patients and 1,143 negative controls, baseline information such as age, gender, and other ESCC-related risk factors, such as smoking and drinking, were gathered through a questionnaire. A total of 1,143 control individuals and all case subjects provided feedback. Each participant had a 2 mL blood sample obtained intravenously for analysis.

The protocol adhered to the *Declaration of Helsinki* on the ethical conduct of research involving human/animal subjects and was approved by the Ethics Committee of Jiangsu University (Zhenjiang, China). All participants signed an informed consent form before recruitment.

2.2. Genomic DNA Extraction and Single-Nucleotide Polymorphism Analysis. The QIAamp DNA Blood Mini Kit was

applied to amplify genomic DNA isolated from peripheral blood using PCR (Qiagen, Berlin, Germany). Samples were genotyped further using the ligation detection reaction (LDR) approach (supported by Genesky Biotechnology Inc., Shanghai, China). Six *LAMA1* SNP loci (rs62081531, rs621993, rs539713, rs566655, rs73938538, and rs607230) were selected and analyzed. As a methodology of quality control, the analysis is repeated on 10% of randomly selected samples. In a preliminary study, we performed a linkage disequilibrium analysis on the 1000Genomes database, identified SNP loci with correlations, and further explored tag SNPs.

2.3. Surgical and Histological Evaluation. All patients underwent esophagectomy by qualified surgeons. Following the procedure, surgical specimens will be fixed to a cork and immersed in 10% formalin. All these patients' specimens were systematically reevaluated by experienced pathologists specialized in thoracic oncology and restaged under the 8th edition AJCC/UICC staging of cancers of the esophagus and esophagogastric junction. The histopathological examination includes tumor size, grade of differentiation, the margin of resection, and lymph node status.

2.4. Statistical Analysis. Statistical analyses were performed using SPSS version 26 (IBM, Chicago, IL) and R software (Version 4.0.3, R Foundation for Statistical Computing, Vienna, Austria). The baseline characteristics were summarized using the R software package "tableone". Clinical characteristics were compared using Fisher's exact test or the χ^2 test for categorical variables and the Student's *t*-test for comparing continuous variables. For the genetic model (A as a major allele and B as a minor allele), (a) *dominant model*: allele B increases risk; (b) *Recessive model*: two copies of minor allele B are required for increased risk; (c) *additive model*: r -fold increased risk for AB and $2r$ increased risk for BB; (d) *multiplicative model*: r -fold increased risk for AB and r^2 increased risk for BB [16]. Two-tailed p value < 0.05 is considered as statistical significance, whereas p values between 0.05 and 0.10 are considered borderline statistically significant. The crude odd ratio (OR) and its corresponding 95% confidence interval (CI) are calculated according to genotypes between the two groups. In most cases, the risk is compared using the parametric test. When the sample size in the group is small, we use the nonparametric test. For stratified analyses, we included age, gender, smoking and drinking status, and the SNP model in the analysis, resulting in an adjusted OR for the SNP model. The adjusted OR and its corresponding CI are calculated by logistic regression analysis, and the hierarchical analysis is carried out. Demographic data, including age, sex, alcohol, and smoking, were covariates. The genotype was a dummy variable, and the group was a dependent variable. The SHEsis online platform [17] was utilized to conduct linkage disequilibrium studies and visualize the results by using R 4.0.3 and the R software packages "LDheatmap" and "genetics," and SHEsis was also used to conduct haplotype frequency analyses [18].

TABLE 1: Distribution of clinicopathological characteristics in the ESCC case and control groups.

| | Case (<i>n</i> = 1043) | Control (<i>n</i> = 1143) | <i>p</i> value |
|----------------------------|----------------------------|-------------------------------|----------------|
| Age (median (IQR)) | 63.00 (59.00, 68.00) | 63.00 (54.00, 70.00) | 0.257 |
| Gender (%) | | | |
| Female | 285 (27.3) | 315 (27.6) | 0.941 |
| Male | 758 (72.7) | 828 (72.4) | |
| Smoke status (%) | | | |
| No | 589 (56.5) | 803 (70.3) | <0.001 |
| Former/current | 454 (43.5) | 340 (29.7) | |
| Alcohol consumption (%) | | | |
| No | 714 (68.5) | 957 (83.7) | <0.001 |
| Former/current | 329 (31.5) | 186 (16.3) | |
| BMI (median (IQR)) | 22.27 (20.20, 24.35) | 23.88 (21.89, 25.88) | <0.001 |
| Chronic disease (%) | | | |
| No | 797 (76.4) | 604 (52.8) | <0.001 |
| Yes | 246 (23.6) | 539 (47.2) | |
| Hypertension (%) | 234 (22.4) | 416 (36.4) | |
| Diabetes (%) | 31 (3.0) | 216 (18.9) | |
| Cardiovascular disease (%) | 0 (0.0) | 59 (5.2) | |
| pT stage (%) | | | |
| Tis | 1 (0.1) | | |
| T1a | 33 (3.2) | | |
| T1b | 112 (10.7) | | |
| T2 | 382 (36.6) | | |
| T3 | 513 (49.2) | | |
| T4 | 2 (0.2) | | |
| pN stage (%) | | | |
| N0 | 775 (74.3) | | |
| N1 | 201 (19.3) | | |
| N2 | 45 (4.3) | | |
| N3 | 22 (2.1) | | |
| Differentiation (%) | | | |
| High | 368 (35.3) | | |
| Moderate | 537 (51.5) | | |
| Low | 138 (13.2) | | |

Data are no. (%) or mean (SD) or median (IQR); BMI: body mass index; pT stage: pathological T stage; pN stage: pathological N stage.

3. Results

Between October 2008 and January 2017, we included 1,043 patients and 1,143 negative controls. There was no statistically significant difference in median age between the case and control groups [63.00 (59.00, 68.00) versus 63.00 (54.00, 70.00), $p = 0.257$]. Women accounted for 27.3% of the case group and 27.6% of the control group ($P = 0.941$). The proportion of smoking and drinking in ESCC patients was higher than that in the control group (43.5% vs. 29.7%, $P < 0.001$, 31.5% vs. 16.3%, $P < 0.001$). Detailed clinicopathological information was described in Table 1.

The brief information on the six genotyped SNPs of *LAMA1* is shown in Table 2. All SNP genotyping experiments were successful at a rate greater than 95%. The control group's minor allele frequencies (MAF) were similar to those found in East Asian groups in the 1000 Genomes and gnomAD-Genomes databases. The Hardy-Weinberg equilibrium (HWE) test revealed that all six SNPs in the control group had p values greater than 0.05, indicating that the control group was genetically equilibrium. Linkage disequilibrium of 6 SNP loci of *LAMA1* is shown in Figure 1

and the coefficient of linkage disequilibrium and correlation coefficient test are described in Tables S7a and S7b.

The association of 6 SNPs in *LAMA1* with esophageal squamous cell carcinoma is shown in Table 3. In the rs62081531 locus, G/A was a protective factor for ESCC compared with G/G (OR: 0.820, 95% CI: 0.677–0.993, $P = 0.042$). We can also find in the dominant model that the minor allele A is a protective factor for ESCC (OR: 0.830, 95% CI: 0.691–0.997, $P = 0.046$). At rs607230, T/C was a protective factor for ESCC compared with T/T (OR: 0.613, 95% CI: 0.389–0.965, $P = 0.034$).

Simultaneously, we performed stratified analyses on 6 loci of *LAMA1*, and detailed information is present in Tables S1–S6. In rs62081531 G > A, in the younger population, the dominant model (OR: 0.680, 95% CI: 0.527–0.877, $P = 0.003$), the recessive model (OR: 0.741, 95% CI: 0.594–0.923, $P = 0.008$), or the multiplicative model (OR: 0.754, $P = 0.010$) all imply a negative link with ESCC incidence. In those without alcohol consumption, we found that either the dominant model (OR: 0.777, 95% CI: 0.627–0.963, $P = 0.022$), the recessive model (OR: 0.796, 95% CI: 0.660–0.961, $P = 0.018$), or the multiplicative model

TABLE 2: Primary information of LAMA1.

| Gene | LAMA1 | | | | | |
|---|--------------------|--------------------|--------------|------------------|--------------------|--------------|
| | rs62081531 | rs621993 | rs539713 | rs566655 | rs73938538 | rs607230 |
| Genotyped SNP | G > A | G > A | A > G | T > G | A > C | T > C |
| Allele | | | | | | |
| Consequence | Synonymous variant | Synonymous variant | Stop gained | Missense variant | Synonymous variant | Stop gained |
| RegulomeDB rank [†] | 7 | 4 | 5 | 5 | 5 | 4 |
| Chromosome | | | 18 | | | |
| Chromosome position [‡] | 6986260 | 7033038 | 7017323 | 7034509 | 7008584 | 6980524 |
| MAF in control | A = 0.1810 | A = 0.1731 | G = 0.2768 | G = 0.1432 | C = 0.1142 | T = 0.1913 |
| MAF in 1000 genomes | | | | | | |
| Global | A = 0.1378 | A = 0.4065 | G = 0.5475 | G = 0.1502 | C = 0.1717 | T = 0.3063 |
| East Asian | A = 0.1915 | A = 0.1319 | G = 0.2262 | G = 0.1042 | C = 0.1091 | T = 0.1756 |
| MAF in gnomAD-genomes | | | | | | |
| Global | A = 0.178903 | A = 0.468719 | G = 0.612173 | G = 0.203225 | C = 0.157585 | T = 0.319456 |
| East Asian | A = 0.1986 | A = 0.1629 | G = 0.2540 | G = 0.1313 | C = 0.1097 | T = 0.1621 |
| <i>p</i> value for the HWE test in the controls | 0.9432 | 0.9743 | 0.9996 | 0.7212 | 0.8633 | 0.6770 |
| Genotyping method | | | LDR | | | |
| % genotyping value | 98.90% | 98.95% | 98.95% | 98.95% | 98.86% | 98.03% |

[†]<https://www.regulomedb.org/>. [‡]Based on Genome Reference Consortium Human Build 38(GRCh38).

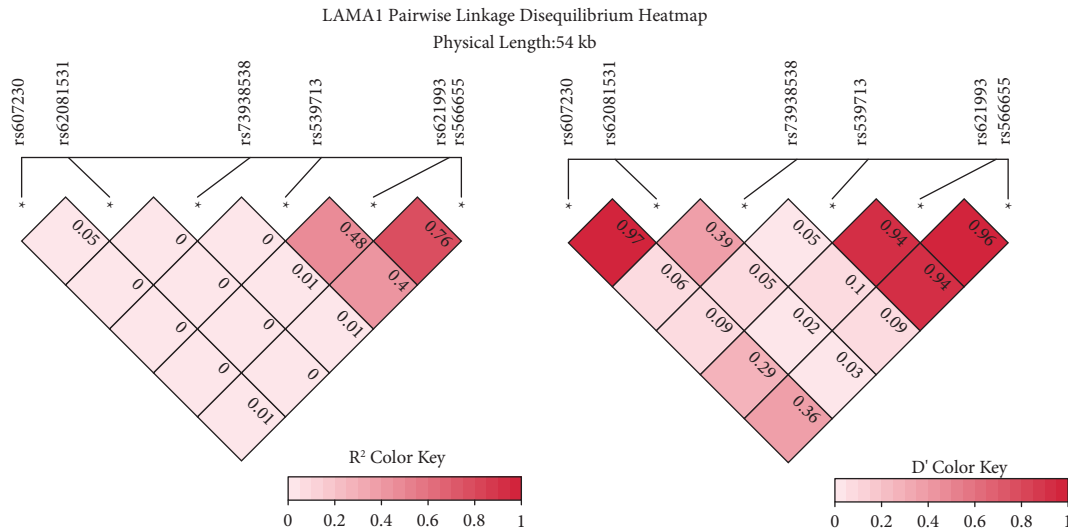


FIGURE 1: D' (normalized D) or r² (correlation coefficient) of linkage disequilibrium on 6 SNP loci of *LAMA1*.

(OR: 0.800, 95% CI: 0.664–0.964, $P = 0.019$) reveals that it is a protective factor for esophageal cancer (Table S1). At the rs539713 A > G locus, in the former or current drinking population, we found that in the recessive or multiplicative model, it was associated with a reduced risk of ESCC (OR: 0.443, 95% CI: 0.224–0.876, $P = 0.019$ and OR: 0.744, 95% CI: 0.556–0.996, $P = 0.047$) (Table S3). At the rs566655 T > G locus, we discovered that both the additive model (OR: 1.270, 95% CI: 1.006–1.603, $P = 0.044$) and the multiplicative model (OR: 1.274, 95% CI: 1.007–1.611, $P = 0.044$) were linked with an elevated risk of ESCC in individuals younger than 65 years old (Table S4). It was found that the rs607230 T > C variant was associated with a decreased risk of esophageal cancer in adults aged 65 years (OR: 0.427, 95% CI: 0.196–0.930, $P = 0.032$) (Table S6). After stratified analysis, rs621993 G > A and rs73938538 A > C did not seem to be linked with esophageal cancer susceptibility (Table S2 and S5). We also analyzed the association of *LAMA1* 6 SNPs with tumor differentiation and lymph node metastasis, and the detailed information is shown in Table S8. We found that neither the lymph node positivity rate nor the degree of differentiation was significantly associated with the mutation status of the 6 loci. Haplotype frequencies analysis revealed that $A_{rs62081531}G_{rs621993}A_{rs539713}T_{rs566655}A_{rs73938538}C_{rs607230}$ (OR: 0.803, 95% CI: 0.660~0.977, $P = 0.028$) and $G_{rs62081531}G_{rs621993}A_{rs539713}T_{rs566655}C_{rs73938538}C_{rs607230}$ (OR: 0.679, 95% CI: 0.504~0.913, $P = 0.010$) were associated with less susceptibility of ESCC (Table 4).

4. Discussion

Through our multicenter large-sample case-control study, we found that rs62081531 locus G > A and rs607230 locus T > C of *LAMA1* were independent protective factors for esophageal squamous cell carcinoma, especially in those younger than 65 and nondrinkers.

Haplotype frequency analysis revealed that $A_{rs62081531}G_{rs621993}A_{rs539713}T_{rs566655}A_{rs73938538}C_{rs607230}$ and

$G_{rs62081531}G_{rs621993}A_{rs539713}T_{rs566655}C_{rs73938538}C_{rs607230}$ were associated with less susceptibility to ESCC among the Chinese population. To our knowledge, this is the first report on *LAMA1* single-nucleotide polymorphisms and susceptibility to esophageal squamous cell carcinoma based on large-scale multicenter clinical research.

LAMA1 (laminin subunit $\alpha 1$) is a part of laminin, a glycoprotein found in the extracellular matrix that constitutes the basement membrane, and has been shown to be involved in the occurrence and development of various diseases [11]. Tissue distribution of laminin subunit $\alpha 1$ is mainly in early epithelial development and some adult epithelia [8]. In terms of carcinogenesis, laminin plays an essential role in cell adhesion, mitosis, differentiation, and even metastasis [19]. Laminin is a fundamental functional component of the basement membrane of several tissues, including the endothelium of the vessel wall, and different isoforms may contribute to vascular homeostasis [20]. The $\alpha 1$ subunit of laminin is typically confined to capillary walls and is expressed in the basal layer of capillaries in the central nervous system [21]. A recent study demonstrates that laminin-1 functions as a chemoattractant for both stromal and vascular cells, as well as in epithelial/stromal cell interactions for the structure of the basement membrane and segregation of integrins, hence signaling the proliferation of epithelial cells [13]. Similarly, in colorectal cancer, Wu et al. reported novel driver mutations occurring during adenoma and cancer evolution by single-cell whole-exome sequencing (scWES), with *LAMA1* (PI3K-Akt signaling pathway) being one of the most critical pathways for CRC evolution [22]. Likewise, Gudjonsson and coworkers revealed that laminin-1 plays a vital role in the replacement of myoepithelial cells in polarity reversal in breast cancer [23]. *LAMA1* (laminin $\alpha 1$) mutations are highly related to retinal avascularity and neovascularization in nontumor fields, such as the Porcetti-Boltshausen syndrome [24]. Regardless of tumor or other nontumor diseases, *LAMA1* is essential for vascular homeostasis and the basal layer of blood vessels.

TABLE 3: Association between LAMA1 single-nucleotide polymorphism and esophageal squamous cell carcinoma while controlling age and gender

| Locus | Genotype | ESCC case | Control | Codominant model | | Dominant model | | Recessive model | | Additive model | | Multiplicative model | |
|------------|----------|------------|------------|------------------------|--------------|------------------------|---------|------------------------|---------|------------------------|---------|------------------------|---------|
| | | | | Or (95%CI) | P value | Or (95%CI) | P value | Or (95%CI) | P value | Or (95%CI) | P value | Or (95%CI) | P value |
| rs62081531 | G/G | 729 (69.9) | 765 (66.9) | Ref | | | | | | | | | |
| | G/A | 261 (25.0) | 334 (29.2) | 0.820 (0.677-0.993) | 0.042 | 0.830 (0.691-0.997) | 0.839 | 0.968 (0.606-1.545) | 0.891 | 0.870 (0.743-1.017) | 0.081 | 0.866 (0.739-1.015) | 0.076 |
| | A/A | 34 (3.3) | 39 (3.4) | 0.915 (0.571-1.465) | 0.711 | | | | | | | | |
| rs621993 | G/G | 704 (67.5) | 777 (68.0) | Ref | | | | | | | | | |
| | G/A | 288 (27.6) | 328 (28.7) | 0.969 (0.803-1.17) | 0.744 | 0.981 (0.818-1.177) | 0.839 | 1.114 (0.682-1.818) | 0.666 | 0.997 (0.852-1.167) | 0.970 | 0.997 (0.851-1.167) | 0.970 |
| | A/A | 33 (3.2) | 33 (2.9) | 1.104 (0.674-1.807) | 0.695 | | | | | | | | |
| rs539713 | A/A | 525 (50.3) | 595 (52.1) | Ref | | | | | | | | | |
| | G/A | 436 (41.8) | 456 (39.9) | 1.084 (0.909-1.292) | 0.371 | 1.044 (0.881-1.236) | 0.621 | 0.805 (0.576-1.124) | 0.202 | 0.991 (0.866-1.135) | 0.900 | 0.992 (0.868-1.133) | 0.902 |
| | G/G | 64 (6.1) | 87 (7.6) | 0.834 (0.591-1.175) | 0.299 | | | | | | | | |
| rs566655 | T/T | 738 (70.8) | 832 (72.8) | Ref | | | | | | | | | |
| | T/G | 263 (25.2) | 286 (25.0) | 1.037 (0.853-1.259) | 0.717 | 1.057 (0.875-1.277) | 0.563 | 1.340 (0.736-2.441) | 0.338 | 1.071 (0.904-1.268) | 0.430 | 1.070 (0.904-1.266) | 0.432 |
| | G/G | 24 (2.3) | 20 (1.7) | 1.353 (0.741-2.469) | 0.325 | | | | | | | | |
| rs73938538 | A/A | 829 (79.5) | 891 (78.0) | Ref | | | | | | | | | |
| | C/A | 184 (17.6) | 234 (20.5) | 0.845 (0.682-1.048) | 0.125 | 0.844 (0.684-1.042) | 0.115 | 0.854 (0.373-1.957) | 0.710 | 0.857 (0.706-1.042) | 0.122 | 0.859 (0.707-1.042) | 0.124 |
| | C/C | 10 (1.0) | 13 (1.1) | 0.827 (0.361-1.896) | 0.653 | | | | | | | | |
| rs607230 | T/T | 49 (4.7) | 37 (3.2) | Ref | | | | | | | | | |
| | T/C | 293 (28.1) | 361 (31.6) | 0.613 (0.389-0.965) | 0.034 | 0.657 (0.425-1.016) | 0.059 | 1.046 (0.875-1.25) | 0.624 | 0.981 (0.844-1.140) | 0.803 | 0.981 (0.842-1.142) | 0.801 |
| | C/C | 664 (63.7) | 739 (64.7) | 0.678 (0.437-1.053) | 0.084 | | | | | | | | |

OR: odds ratio; CI: confidence interval.

TABLE 4: Haplotype frequencies in the case and control groups and risk of ESCC.

| Haplotype (LAMA1) | Case (freq) | Control (freq) | Chi2 | Or [95% CI] | p value |
|---|----------------|----------------|-------|---------------------|--------------|
| A _{rs62081531} A _{rs621993} G _{rs539713} G _{rs566655} A _{rs73938538} C _{rs607230} | 38.73 (1.9%) | 63.18 (2.8%) | — | — | |
| A _{rs62081531} A _{rs621993} G _{rs539713} G _{rs566655} C _{rs73938538} C _{rs607230} | 4.99 (0.2%) | 2.60 (0.1%) | — | — | |
| A _{rs62081531} A _{rs621993} G _{rs539713} T _{rs566655} A _{rs73938538} C _{rs607230} | 2.33 (0.1%) | 4.12 (0.2%) | — | — | |
| A _{rs62081531} G _{rs621993} A _{rs539713} T _{rs566655} A _{rs73938538} C _{rs607230} | 200.31 (10.0%) | 269.91 (11.9%) | 4.823 | 0.803 [0.660~0.977] | 0.028 |
| A _{rs62081531} G _{rs621993} A _{rs539713} T _{rs566655} C _{rs73938538} C _{rs607230} | 13.40 (0.7%) | 15.93 (0.7%) | — | — | |
| A _{rs62081531} G _{rs621993} G _{rs539713} T _{rs566655} A _{rs73938538} C _{rs607230} | 38.33 (1.9%) | 56.03 (2.5%) | — | — | |
| G _{rs62081531} A _{rs621993} A _{rs539713} G _{rs566655} A _{rs73938538} C _{rs607230} | 7.10 (0.4%) | 4.11 (0.2%) | — | — | |
| G _{rs62081531} A _{rs621993} A _{rs539713} T _{rs566655} A _{rs73938538} C _{rs607230} | 1.29 (0.1%) | 3.52 (0.2%) | — | — | |
| G _{rs62081531} A _{rs621993} G _{rs539713} G _{rs566655} A _{rs73938538} C _{rs607230} | 168.31 (8.4%) | 171.67 (7.6%) | 0.731 | 1.102 [0.881~1.379] | 0.393 |
| G _{rs62081531} A _{rs621993} G _{rs539713} G _{rs566655} A _{rs73938538} T _{rs607230} | 18.19 (0.9%) | 30.69 (1.4%) | — | — | |
| G _{rs62081531} A _{rs621993} G _{rs539713} G _{rs566655} C _{rs73938538} C _{rs607230} | 29.83 (1.5%) | 39.59 (1.7%) | — | — | |
| G _{rs62081531} A _{rs621993} G _{rs539713} G _{rs566655} C _{rs73938538} T _{rs607230} | 3.90 (0.2%) | 7.69 (0.3%) | — | — | |
| G _{rs62081531} A _{rs621993} G _{rs539713} T _{rs566655} A _{rs73938538} C _{rs607230} | 29.88 (1.5%) | 35.63 (1.6%) | — | — | |
| G _{rs62081531} A _{rs621993} T _{rs566655} A _{rs73938538} T _{rs607230} | 12.55 (0.6%) | 18.16 (0.8%) | — | — | |
| G _{rs62081531} A _{rs621993} G _{rs539713} T _{rs566655} C _{rs73938538} C _{rs607230} | 6.78 (0.3%) | 3.95 (0.2%) | — | — | |
| G _{rs62081531} G _{rs621993} A _{rs539713} T _{rs566655} A _{rs73938538} C _{rs607230} | 845.42 (42.2%) | 909.38 (40.1%) | 1.066 | 1.072 [0.939~1.224] | 0.302 |
| G _{rs62081531} G _{rs621993} A _{rs539713} T _{rs566655} A _{rs73938538} T _{rs607230} | 259.49 (12.9%) | 275 (11.9%) | 0.720 | 1.083 [0.900~1.303] | 0.396 |
| G _{rs62081531} G _{rs621993} A _{rs539713} T _{rs566655} C _{rs73938538} C _{rs607230} | 74.63 (3.7%) | 119.97 (5.3%) | 6.626 | 0.679 [0.504~0.913] | 0.010 |
| G _{rs62081531} G _{rs621993} A _{rs539713} T _{rs566655} C _{rs73938538} T _{rs607230} | 40.47 (2.0%) | 43.28 (1.9%) | — | — | |
| G _{rs62081531} G _{rs621993} G _{rs539713} T _{rs566655} A _{rs73938538} C _{rs607230} | 117.31 (5.9%) | 114.20 (5.0%) | 1.127 | 1.155 [0.885~1.508] | 0.288 |
| G _{rs62081531} G _{rs621993} G _{rs539713} T _{rs566655} A _{rs73938538} T _{rs607230} | 34.02 (1.7%) | 59.00 (2.6%) | — | — | |
| G _{rs62081531} G _{rs621993} G _{rs539713} T _{rs566655} C _{rs73938538} C _{rs607230} | 8.56 (0.4%) | 15.83 (0.7%) | — | — | |

Freq: frequency; OR: odds ratio; CI: confidence interval; p < 0.05 is regarded as statistically significant value.

Interestingly, Velling et al. found that none of the colon cancer cell lines synthesized the laminin $\alpha 1$ protein, and they suggested that mutations in the *LAMA1* gene may underlie the lack of laminin $\alpha 1$ chains observed in some colon cancers [25]. In our study, *LAMA1* mutation showed protective factors in both rs62081531 G > A and rs607230 T > C, and basic research also found that *LAMA1* deficiency could inhibit the proliferation and invasion of esophageal cancer [26, 27]. Clearly, not all *LAMA1* SNPs are protective factors. People younger than 65 years of age indicate that the mutation at rs566655 T > G increases the risk of esophageal cancer, which may be correlated to the function of a certain SNP.

However, there are few reports on *LAMA1* and esophageal cancer. Most of the research is limited to the genetic function of *LAMA1*. Meng and colleagues found that laminin $\alpha 1$ (*LAMA1*) is highly expressed in ESCC tissue and mediates the FAK-PI3K-Akt signaling pathway [27]. Zhou et al. revealed that *LAMA1* was significantly upregulated in ESCC tissues and positively correlated with an aggressive oncogenic phenotype [26]. Nevertheless, the relationship between *LAMA1* SNP and disease in malignancies has not been demonstrated. Previous research has focused chiefly on nontumor studies such as chronic disease or degenerative disease. Zhao et al. showed that rs2089760 T > G, which is located in the *LAMA1* promoter region, may be associated with myopia in Chinese populations [28]. Similarly, the *LAMA1* rs2089760 G > A mutation was reported to reduce transcription factor binding ability and transcription initiation activity and negatively control the gene transcription of *LAMA1*, playing a crucial role in pathological myopia [29]. In a study on degenerative diseases, D'Aoust and

colleagues were the first to discover that *LAMA1* rs73938538 A > C was positively related to Alzheimer's disease in the Amish community [30]. Due to the single-nucleotide polymorphism of *LAMA1*, the mutated site seems unable to effectively translate *LAMA1* into laminin subunit $\alpha 1$ so as to exert its specific biological function, hence preventing esophageal cancer susceptibility. We observed that the consequences of two SNP loci that were related to esophageal susceptibility were stop-gain mutations, while the rest were synonymous and missense variants. This is largely in accordance with our expectations, especially when terminal gain mutations and missense variants can dramatically alter protein function, even with single nucleotide changes. However, we also reveal that in the ESCC population, *LAMA1* polymorphisms did not show a statistically significant association with the degree of differentiation, lymph node positivity, or T stage.

The main limitation of our study is that we only included populations from a specific region in eastern China, which may result in a certain geographical specificity and may not be generalized to the entire ESCC population. In addition, our study lacks replication in independent cohorts. Furthermore, we only discovered SNP sites in peripheral blood in our investigation. How these SNPs of *LAMA1* translate into biological function in the evolution of esophageal carcinoma is definitely a primary subject of our future investigation, which is currently technically difficult due to the lack of biological tools in our lab.

In conclusion, we found a strong association of *LAMA1* rs62081531, rs539713, rs566655, and rs607230 polymorphisms with esophageal cancer susceptibility in the Chinese population. *LAMA1* SNPs may significantly impact

the occurrence of esophageal cancer and may serve as effective diagnostic biomarkers.

Abbreviations

EC: Esophageal cancer
 ESCC: Esophageal squamous cell carcinoma
 SNP: Single-nucleotide polymorphism.

Data Availability

Statistical results of this study are available from corresponding authors upon reasonable request. Specific patient clinical information and genetic data to support the study results have not yet been made available due to the National BioSafety Law of the People's Republic of China.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Shaoyuan Zhang, Feng Su, and Jun Yin were involved in the conceptualization of the study. Lu Lv, Tao Long, Huiwen Pan, Junqing Qi, Qiang Zhou, Weifeng Tang, Guowen Ding, Liming Wang, and Jun Yin were involved in methodology and managed the resources. Data analysis was performed by Shaoyuan Zhang, Yong Fang, Feng Su, Jinjie Yu, and Tian Jiang. Shaoyuan Zhang and Siyun Lin visualized the study. Shaoyuan Zhang and Yong Fang wrote the manuscript. Lijie Tan, Liming Wang, and Jun Yin administered the project. Lijie Tan and Jun Yin supervised the study.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (81370001, 81300037, 81000028, 81570031, 81101889, 81472332, and 81341006); the Key Research and Development Program of Jiangsu Province (BE2016714); the Natural Science Foundation of Jiangsu Province (BK2010333 and BK2011481); the "333" Elitist Training Program, Jiangsu, China (BRA2013135 and BRA2017129); the "Six Talent Peaks" Training Program, Jiangsu, China (2015-WSN-117 and 2014-WSN-078); the "Distinguished Medical Specialist" Program, Jiangsu, China; the "Innovative and Entrepreneurial Elite Team" Program (2016), Jiangsu, China; the research funding of the Shanghai Hospital Development Center (SHDC12018X12 and SHDC2020CR4039); the Natural Science Foundation of Shanghai (19ZR1449900 and 20ZR1411600); 2021 Clinical Research Navigation Program of the Shanghai Medical College of Fudan University (grant number: DGF501035/002); and the Shanghai Xuhui District Medical Research Project (SHXH201805).

Supplementary Materials

Table S1: Stratified analyses between rs62081531 G > A polymorphism and ESCC risk by age, gender, smoking

status, and alcohol consumption. Table S2: Stratified analyses between rs621993 G > A polymorphism and ESCC risk by age, gender, smoking status, and alcohol consumption. Table S3: Stratified analyses between rs539713 A > G polymorphism and ESCC risk by age, gender, smoking status, and alcohol consumption. Table S4: Stratified analyses between rs566655 T > G polymorphism and ESCC risk by age, gender, smoking status, and alcohol consumption. Table S5: Stratified analyses between rs73938538 A > C polymorphism and ESCC risk by age, gender, smoking status, and alcohol consumption. Table S6: Stratified analyses between rs607230 T > C polymorphism and ESCC risk by age, gender, smoking status, and alcohol consumption. Table S7a: The linkage disequilibrium test of LAMA1 in the case and control groups. Table S7b: The linkage disequilibrium test of LAMA1 in the case and control groups. Table S8: The pathological baseline of 1043 ESCC patients on LAMA1 single-nucleotide polymorphism. pT stage: pathological T stage; pN stage: pathological N stage. (*Supplementary Materials*)

References

- [1] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] S. Ohashi, Si Miyamoto, O. Kikuchi, T. Goto, Y. Amanuma, and M. Muto, "Recent advances from basic and clinical studies of esophageal squamous cell carcinoma," *Gastroenterology*, vol. 149, no. 7, pp. 1700–1715, 2015.
- [3] Y. Song, L. Li, Y. Ou et al., "Identification of genomic alterations in oesophageal squamous cell cancer," *Nature*, vol. 509, no. 7498, pp. 91–95, 2014.
- [4] M. Aumailley and N. Smyth, "The role of laminins in basement membrane function," *Journal of Anatomy*, vol. 193, no. 1, pp. 1–21, 1998.
- [5] T. R. Patel, D. Nikodemus, T. M. D. Besong et al., "Biophysical analysis of a lethal laminin alpha-1 mutation reveals altered self-interaction," *Matrix Biology*, vol. 49, pp. 93–105, 2016.
- [6] R. Timpl, H. Rohde, P. G. Robey, S. I. Rennard, J.-M. Foidart, and G. R. Martin, "Laminin-A glycoprotein from basement membranes," *Journal of Biological Chemistry*, vol. 254, no. 19, pp. 9933–9937, 1979.
- [7] H. Colognato and P. D. Yurchenco, "Form and function: the laminin family of heterotrimeric," *Developmental Dynamics*, vol. 218, no. 2, pp. 213–234, 2000.
- [8] M. Patarroyo, K. Tryggvason, and I. Virtanen, "Laminin isoforms in tumor invasion, angiogenesis and metastasis," *Seminars in Cancer Biology*, vol. 12, no. 3, pp. 197–207, 2002.
- [9] A. De Arcangelis, P. Neuville, R. Boukamel, O. Lefebvre, M. Kedinger, and P. Simon-Assmann, "Inhibition of laminin alpha 1-chain expression leads to alteration of basement membrane assembly and cell differentiation," *Journal of Cell Biology*, vol. 133, no. 2, pp. 417–430, 1996.
- [10] P. Ekblom, P. Lonai, and J. F. Talts, "Expression and biological role of laminin-1," *Matrix Biology*, vol. 22, no. 1, pp. 35–47, 2003.
- [11] R. S. Bresalier, B. Schwartz, Y. S. Kim, Q. Y. Duh, H. K. Kleinman, and P. M. Sullam, "The laminin alpha 1 chain Ile-Lys-Val-Ala-Val (IKVAV)-containing peptide promotes

- liver colonization by human colon cancer cells," *Cancer Research*, vol. 55, no. 11, pp. 2476–2480, 1995.
- [12] J. J. Grzesiak, K. C. Smith, D. W. Burton, L. J. Deftos, and M. Bouvet, "Integrin-mediated laminin-1 adhesion upregulates CXCR4 and IL-8 expression in pancreatic cancer cells," *Surgery*, vol. 141, no. 6, pp. 804–814, 2007.
- [13] A. De Arcangelis, O. Lefebvre, A. Méchine-Neuville et al., "Overexpression of laminin alpha 1 chain in colonic cancer cells induces an increase in tumor growth," *International Journal of Cancer*, vol. 94, no. 1, pp. 44–53, 2001.
- [14] B. Diao and P. Yang, "Comprehensive analysis of the expression and prognosis for laminin genes in ovarian cancer," *Pathology and Oncology Research*, vol. 27, Article ID 1609855, 2021.
- [15] O. Fotouhi, M. Adel Fahmideh, M. Kjellman et al., "Global hypomethylation and promoter methylation in small intestinal neuroendocrine tumors: an in vivo and in vitro study," *Epigenetics*, vol. 9, no. 7, pp. 987–997, 2014.
- [16] C. M. Lewis, "Genetic association studies: design, analysis and interpretation," *Briefings in Bioinformatics*, vol. 3, no. 2, pp. 146–153, 2002.
- [17] Y. Yong and L. He, "SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci," *Cell Research*, vol. 15, no. 2, pp. 97–98, 2005.
- [18] Z. Li, Z. Zhang, Z. He et al., "A partition-ligation-combination-subdivision EM algorithm for haplotype inference with multiallelic markers: update of the SHEsis (<http://analysis.bio-x.cn>)," *Cell Research*, vol. 19, no. 4, pp. 519–523, 2009.
- [19] Y. Kuratomi, M. Nomizu, P. K. Nielsen et al., "Identification of metastasis-promoting sequences in the mouse laminin alpha 1 chain," *Experimental Cell Research*, vol. 249, no. 2, pp. 386–395, 1999.
- [20] L. F. Yousif, J. Di Russo, and L. Sorokin, "Laminin isoforms in endothelial and perivascular basement membranes," *Cell Adhesion & Migration*, vol. 7, no. 1, pp. 101–110, 2013.
- [21] I. Virtanen, D. Gullberg, J. Rissanen et al., "Laminin α 1-chain shows a restricted distribution in epithelial basement membranes of fetal and adult human tissues," *Experimental Cell Research*, vol. 257, no. 2, pp. 298–309, 2000.
- [22] H. Wu, X. Y. Zhang, Z. Hu et al., "Evolution and heterogeneity of non-hereditary colorectal cancer revealed by single-cell exome sequencing," *Oncogene*, vol. 36, no. 20, pp. 2857–2867, 2017.
- [23] T. Gudjonsson, L. Rønnev-Jessen, R. Villadsen, F. Rank, M. J. Bissell, and O. W. Petersen, "Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition," *Journal of Cell Science*, vol. 115, no. 1, pp. 39–50, 2002.
- [24] E. Marlow, R. V. P. Chan, E. Oltra, I. Rusu, and M. P. Gupta, "Retinal avascularity and neovascularization associated with LAMA1 (laminin1) mutation in porette-boltshausen syndrome," *JAMA Ophthalmol*, vol. 136, no. 1, pp. 96–97, 2018.
- [25] T. Velling, C. F. Tiger, P. Ekblom, and D. Gullberg, "Laminin α chains in colon carcinoma cell lines: detection of a truncated laminin α 1 mRNA in c-2 cells," *Experimental Cell Research*, vol. 248, no. 2, pp. 627–633, 1999.
- [26] P. L. Zhou, Z. Wu, W. Zhang et al., "Circular RNA hsa_circ_0000277 sequesters miR-4766-5p to upregulate LAMA1 and promote esophageal carcinoma progression," *Cell Death & Disease*, vol. 12, no. 7, p. 676, 2021.
- [27] X. Meng, X. Chen, P. Lu et al., "MicroRNA-202 inhibits tumor progression by targeting LAMA1 in esophageal squamous cell carcinoma," *Biochemical and Biophysical Research Communications*, vol. 473, no. 4, pp. 821–827, 2016.
- [28] Y. Y. Zhao, F. J. Zhang, S. Q. Zhu et al., "The association of a single nucleotide polymorphism in the promoter region of the LAMA1 gene with susceptibility to Chinese high myopia," *Molecular Vision*, vol. 17, pp. 1003–1010, 2011.
- [29] Y. Liang, Y. Song, F. Zhang, M. Sun, and N. Wang, "Effect of a single nucleotide polymorphism in the LAMA1 promoter region on transcriptional activity: implication for pathological myopia," *Current Eye Research*, vol. 41, no. 10, pp. 1379–1386, 2016.
- [30] L. N. D'Aoust, A. C. Cummings, R. Laux et al., "Examination of candidate exonic variants for association to Alzheimer disease in the Amish," *PLoS One*, vol. 10, no. 2, Article ID e0118043, 2015.

Research Article

A Prognostic Signature for Colon Adenocarcinoma Patients Based on m6A-Related lncRNAs

Su-Zhe Zhou ¹, Ying-Lian Pan,² Qing-Chun Deng,³ Chang-Jun Yin,⁴ De-Jiang Zhou,⁴ Ming-Liang Liu ⁵, Jun Zhou ^{4,6} and Xiao-Jing Wu ⁷

¹Department of General Practice, Hefei BOE Hospital, Hefei 230013, China

²Department of Medical Oncology, The First Affiliated Hospital of Hainan Medical University, Haikou 570102, China

³Department of Gynecology, The Second Affiliated Hospital of Hainan Medical University, Haikou 570102, China

⁴Department of Gastrointestinal Surgery, Sun Yat-Sen Memorial Hospital of Sun Yat-Sen University, Guangzhou 510120, China

⁵Department of Gastrointestinal Surgery, Affiliated Hospital of Jining Medical College, Jining 272007, China

⁶Department of Gastrointestinal Surgery, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou 510150, China

⁷Department of Surgery Critical Care Medicine, Beijing Shijitan Hospital, Capital Medical University, Beijing 100038, China

Correspondence should be addressed to Ming-Liang Liu; lml0481@163.com, Jun Zhou; zhoujun2@mail.sysu.edu.cn, and Xiao-Jing Wu; wuxiaojing@tmu.edu.cn

Received 17 August 2022; Revised 15 November 2022; Accepted 28 November 2022; Published 13 February 2023

Academic Editor: Feng Jiang

Copyright © 2023 Su-Zhe Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

N6-methyladenosine (m6A) modification is a common epigenetic modification. It is reported that lncRNA can be regulated by m6A modification. Previous studies have shown that lncRNAs associated with m6A regulation (m6A-lncRNAs) serve as ideal prognostic biomarkers. However, whether lncRNAs are involved in m6A modification in colon adenocarcinoma (COAD) needs further exploration. The objective of this study was to construct an m6A-lncRNAs-based signature for patients with COAD. We obtained the RNA sequencing data and clinical information from The Cancer Genome Atlas (TCGA). Pearson correlation analysis was employed to recognize lncRNAs associated with m6A regulation (m6A-lncRNAs). 24 prognostic m6A-lncRNAs was identified by univariate Cox regression analysis. Gene set enrichment analysis (GSAE) was used to investigate the potential cellular pathways and biological processes. We have also explored the relationship between immune infiltrate levels and m6A-lncRNAs. Then, a predictive signature based on the expression of 13 m6A-lncRNAs was constructed by the Lasso regression algorithm, including UBA6-AS1, AC139149.1, U91328.1, AC138207.5, AC025171.4, AC008760.1, ITGB1-DT, AP001619.1, AL391422.4, AC104532.2, ZEB1-AS1, AC156455.1, and AC104819.3. ROC curves and K M survival curves have shown that the risk score has a well-predictive ability. We also set up a quantitative nomogram on the basis of risk score and prognosis-related clinical characteristics. In summary, we have identified some m6A-lncRNAs that correlated with prognosis and tumor immune microenvironment in COAD. In addition, a potential alternative signature based on the expression of m6A-lncRNAs was provided for the management of COAD patients.

1. Introduction

Colon adenocarcinoma is a common pathological type of colon cancer, and its prognosis is poor [1, 2]. Therapies for colon adenocarcinoma (COAD) include surgery, chemotherapy, and radiation therapy [3]. Surgery can cure about half of the COAD patients, but the recurrence rate stays high after surgery. Chemotherapy and radiation are not

effective due to their side effects and drug resistance. Besides, the significant heterogeneity of COAD limits the utilization of traditional methods [4]. With the approach of the era of personalized therapy managements, traditional diagnosing failed to satisfy advanced diagnoses and therapies. It is expected to build more useful prognostic signatures to help improve personalized treatment management.

N6-methyladenosine (m6A) modification is one of the principal internal modifications of RNA and participates in many biological processes [5, 6]. m6A modification has been proven to be a reversible process, which is regulated by methyltransferase (writer), demethylase (eraser), and signal sensor (reader) [7]. Studies have reported that m6A modification plays critical roles in the progression of different malignant tumors, including COAD [8–14]. For instance, METTL3 has been identified to promote COAD occurrence and progression by relying on IGF2BP1/IGF2BP2 [15, 16]. It has been found that METTL14 suppresses COAD occurrence and progression by relying on YTHDF2 [17, 18].

lncRNA is a kind of RNA molecule that does not encode a protein, with a length of more than 200 bp, and plays important roles in the carcinogenesis and progression of cancers, including COAD [19, 20]. It has been proven that m6A modification can regulate the physiological functions of lncRNAs [21, 22]. For example, the structure of lncRNA can be regulated by binding to m6A readers, allowing m6A residues to be accessed by specific RNA-binding proteins [23, 24]. m6A modification modulates the structure of lncRNA MALAT1 to play the function of the structural switch, which is related with cancer malignancies [25]. METTL16 (writer) was identified as an RNA-binding protein of lncRNA MALAT1 [26]. m6A modification can stabilize lncRNA FAM225A that served as a sponge for miR-1275 and miR-590-3p in nasopharyngeal carcinoma [27]. METTL3 (writer) can stabilize and upregulate LINC00958, which is involved with the malignancy of liver cancer progression by sponging miR-3619-5p [28]. Previous studies have shown that m6A-lncRNAs serve as ideal prognostic biomarkers. Nevertheless, whether lncRNAs are involved in the regulation of m6A modification in COAD still need to be elucidated. The objective was to identify prognostic m6A-lncRNAs and construct an m6A-lncRNAs-based prognostic signature for patients with COAD.

In this study, we obtained the RNA sequencing data in TCGA and identified 24 prognostic m6A-lncRNAs and 13 m6A-lncRNAs were selected by the Lasso regression algorithm to construct prognostic signature. We have also verified the reliability of the prognostic signature. In addition, a quantitative prognostic nomogram was constructed based on signature and clinical features.

2. Methods

2.1. Data Source and Preparation. RNA sequencing data of 398 COAD patients with clinical information were obtained from TCGA. 19 COAD patients were excluded due to a lack of necessary clinical information. Subsequently, RNA sequencing data were divided into mRNAs and lncRNAs using the Ensembl Genome Browser [29]. The corresponding clinical data included age, gender, tumor-node-metastasis (TNM) stage, pathological stage, and survival time. We randomly divided COAD patients at a ratio of 7 : 3 into the training cohort (267 patients) and validation cohort (112 patients).

2.2. Identifying Prognosis-Related m6A-lncRNAs. 23 m6A regulators, based on published articles, including 8 writers, 13 readers, and 2 erasers, were collected in this study (Table 1) [30, 31]. Pearson correlation analysis was applied to the expression of lncRNAs and the 23 m6A regulators to recognize m6A-associated lncRNAs. Univariate Cox regression analysis was applied to recognize prognostic m6A-lncRNAs. We used the “limma” R software package to analyse the differential expression of prognosis-related m6A lncRNA.

2.3. Consensus Clustering Analysis. For a better understanding of the role of m6A-lncRNAs, consensus clustering was performed by the “ConsensusClusterPlus” R package to divide all samples into different clusters based on the expression of prognosis-related m6A-lncRNAs [32]. Subsequently, Kaplan Meier analysis was performed in different clusters. We also applied the CIBERSORT algorithm and the Wilcoxon test to analyse different immune cell infiltration between clusters. Besides, immune, stromal, and ESTIMATE scores were calculated by the “ESTIMATE” R package. GSEA software was applied to investigate the potential cellular pathways and biological processes in different clusters.

2.4. Development and Evaluation of Prognostic Signature. After prognosis-related m6A-lncRNAs were identified, LASSO regression analysis was applied to setup a risk model by “glmnet” R package, which could avoid overfitting by disposing of highly correlated lncRNAs [33, 34]. The signature was calculated in the following format:

$$\text{risk score} = \sum_{i=1}^n \text{coef}(i) * \text{lncRNA}(i) \text{ expression.} \quad (1)$$

Then, COAD patients were classified into high- and low-risk subgroups. K M survival analysis was employed to compare whether there were differences in survival between the two subgroups using “survival” R packages. To testify the prediction efficacy of the risk model, we employed ROC curves and measured the AUC values by R package “timeROC” [35].

2.5. Establishment of Prognostic Nomogram. To further evaluate the reliability of the signature, a comprehensive analysis of risk score and clinical features was performed. Subsequently, a quantitative nomogram was developed on the basis of risk score and clinical features using the “rms” R package [36]. We applied the calibration curves to outline the accuracy of the nomogram.

2.6. Statistical Analysis. All statistical analysis in this study was performed using R software (V 4.0.4). The Wilcoxon’s test was employed to compare the difference. Kaplan Meier (K M) survival analysis was performed by using the log-rank

TABLE 1: N6-methyladenosine (m6A) modification regulators.

| | |
|--------|---|
| Eraser | FTO; ALKBH5 |
| Writer | ZC3H13; METTL16; METTL14; METTL3; VIRMA; RBM15B; RBM15; WTAP |
| Reader | FMR1; YTHDF3; YTHDF2; YTHDF1; YTHDC2; YTHDC1; RBMX; HNRNPC; HNRNPA2B1; LRPPRC; IGFBP3; IGFBP2; IGFBP1 |

test. Unless otherwise stated, $P < 0.05$ was considered a statistically significant difference.

3. Results

3.1. Identification of Prognosis-Related m6A-lncRNAs. By Pearson correlation analysis, a total of 1505 lncRNAs were identified as m6A lncRNA, with an absolute correlation coefficient > 0.4 ($P < 0.001$). Then, 24 prognostic m6A-lncRNAs were identified by univariate Cox regression analysis. Among them, UBA6-AS1, AC139149.1, NIFK-AS1, AC245041.1, U91328.1, SFTA1P, AC138207.5, SNHG26, AC025171.4, AC008760.1, AC026367.1, ITGB1-DT, AP001619.1, LINC01138, LINC01545, AL391422.4, AC104532.2, AC005229.4, ZKSCAN2-DT, ZEB1-AS1, AC107308.1, AC156455.1, and ATP2B1-AS1 were recognized as risky lncRNAs for $HR > 1$ ($P < 0.05$) and AC104819.3 was recognized as protective lncRNA with $HR < 1$ ($P < 0.05$) (Figure 1(a)). The correlation between lncRNAs and m6A regulators is shown in Figure 1(b). The expression of 24 prognostic m6A-lncRNAs in normal tissues and tumor tissues was displayed by the box plot and heatmap. (Figures 1(c) and 1(d)).

3.2. Consensus Clustering Analysis. For further understanding the roles of prognostic m6A-lncRNAs, patients were clustered according to the expression of m6A-lncRNAs by consensus clustering analysis. As displayed in the consensus matrix map and cumulative distribution function (CDF) plot for $k=2$, the interference was the smallest (Figures 2(a) and 2(b)). The overall survival results of patients in cluster 1 are better than those in cluster 2 (Figure 2(c)). Figure 2(d) displays the correlation between clinicopathological features and clusters.

3.3. Gene Set Enrichment Analysis (GSAE). To explore the potential cellular pathways and biological processes of prognostic m6A-lncRNAs, the GSEA was employed between two clusters. As displayed in Figure 3, genes in cluster 2 were enriched in the p53-signaling-pathway, proteasome, cell cycle, and peroxisome. Besides, genes in cluster 1 were enriched in TGF-beta signaling-pathway, ERBB-signaling-pathway, ECM-receptor interaction, MAPK-signaling-pathway, and JAK-STAT-signaling-pathway.

3.4. Immune Cell Infiltration and Distribution of Immunity. To explore the roles of m6A-lncRNAs in the tumor immune microenvironment (TIME), we compared the scores of 22 different immune cell types in two clusters. As shown in Figure 4(a), activated memory CD4 T cells are rich in cluster 2 ($P < 0.05$). Figure 4(b) shows the positive correlation

between the m6A-lncRNAs and PD-L1. Furthermore, we analysed the distribution of immunity in two clusters. Cluster 1 have a higher stromal score, immune score, and ESTIMATE score ($P < 0.05$, Figures 4(c)–4(e)).

3.5. Construction and Validation of m6A-lncRNAs Signature. The LASSO regression algorithm was used to avoid overfitting and for constructing risk scores. As displayed in Figures 5(a) and 5(b), at penalty factor (λ) 13, the coefficients of some variables are near to 0 [37]. Ultimately, 13 m6A-lncRNAs were identified as independent prognostic factors. The m6A-lncRNAs risk model for predicting prognosis in COAD was established on the basis of coefficients of 13 m6A-lncRNAs in the following format: risk score = UBA6-AS1 * 0.7684476 + AC139149.1 * 0.7685387 + AC104819.3 * (-1.96469) + U91328.1 * 0.3846686 + AC138207.5 * 0.1605522 + AC025171.4 * 0.1220795 + AC008760.1 * 0.1164466 + ITGB1-DT * 0.5387446 + AP001619.1 * 0.0992952 + AL391422.4 * 0.4690258 + AC104532.2 * 0.1403134 + ZEB1-AS1 * 0.3567363 + AC156455.1 * 0.1279327 (Figure 5(c)).

The K M curves showed that the survival outcome of the low-risk subgroup was better ($P < 0.05$) (Figures 5(d) and 5(g)). The time-dependent ROC curves indicated that, in the training cohort, the AUC value at 1 year was 0.845, 0.797 at 3 years, and 0.813 at 5 years (Figure 5(e)). In the validation cohort, the AUC value at 1 year was 0.750, 0.821 at 3 years, and 0.935 at 5 years (Figure 5(h)). Furthermore, the risk score distribution plot and scatter plot also showed that the survival outcome of the high-risk subgroup was worse. The heatmap illustrated that the expressions of risky lncRNAs were upregulated in the high-risk group, while protective lncRNA was downregulated (Figures 5(f) and 5(i)).

To further validate the risk model, the correlation between risk score and clinical characteristics was analysed. As shown in Figure 6(a), the risk score was significantly correlated with pathological stage and immune score ($P < 0.05$). Besides, the scatter plot showed that the risk score was also significantly correlated with PD-L1 expression and immune score (Figures 6(b) and 6(d)).

3.6. Development of Survival Prognostic Nomogram. We comprehensively analysed the risk score and clinical characteristics. Risk scores, age, and pathological stage were identified as significant independent prognostic variables ($P < 0.05$, Figures 7(a) and 7(b)), which revealed that the risk model served as a reliable tool for COAD patients. Subsequently, we constructed a quantitative nomogram on the basis of risk score and clinical characteristics (Figure 7(c)). The calibration curves displayed the concordance between observed and predicted overall survival (Figures 7(d)–7(f)).

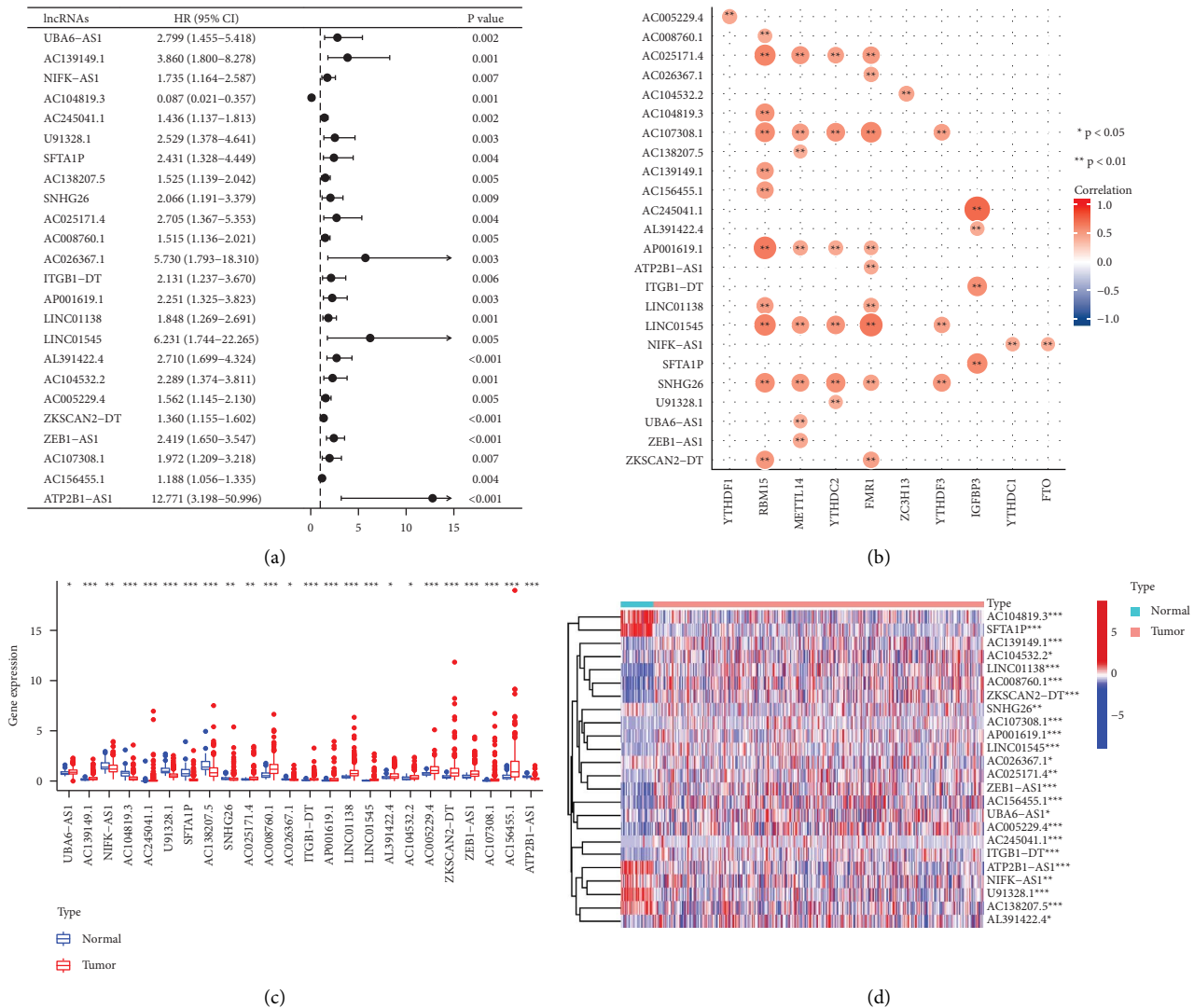


FIGURE 1: Identification of prognostic m6A-lncRNAs. (a) Forest plots of 24 prognosis-related m6A-lncRNAs. (b) The correlation between lncRNAs and m6A regulators. (c) Boxplots of 24 prognosis-related m6A-lncRNAs. (d) Heatmap of 24 prognosis-related m6A-lncRNAs.

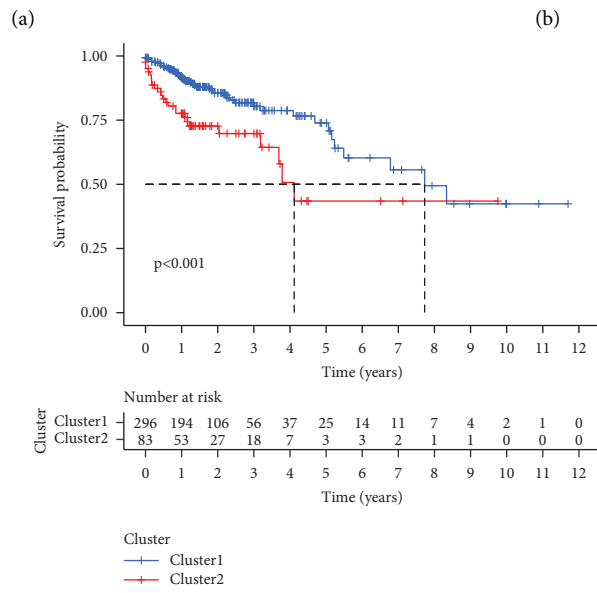
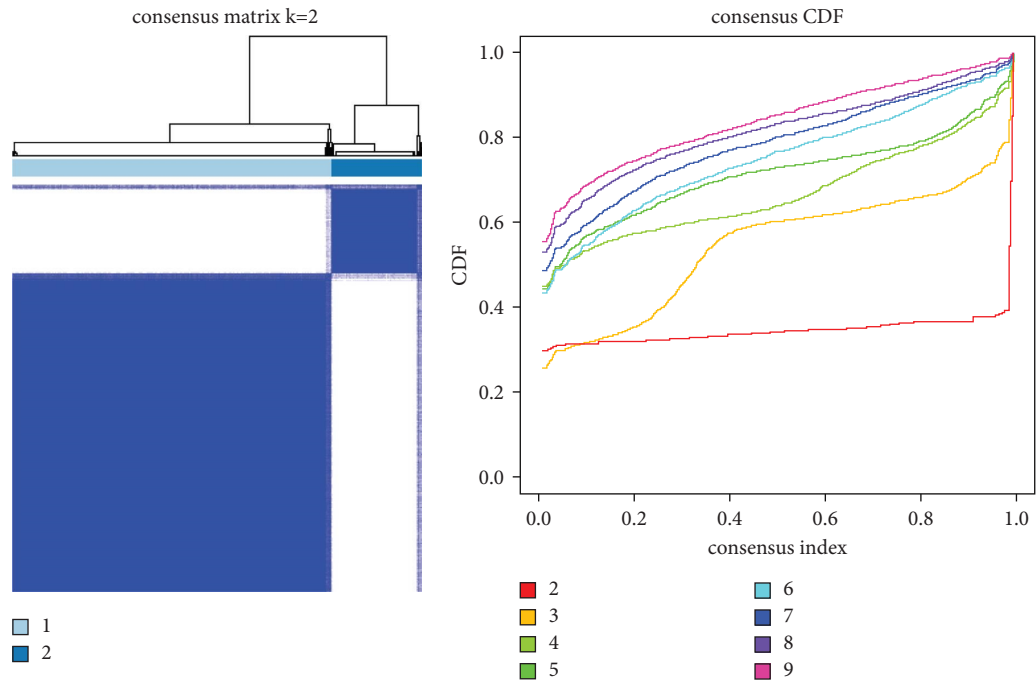
4. Discussions

N6-methyladenosine (m6A) modification is a common epigenetic modification in eukaryotes, involving many biological processes, such as RNA splicing, translation, and expression. M6A modification and lncRNAs play critical roles in the biological processes of COAD. For example, m6A regulators, such as METTL3, METTL14, and YTHDF2, have been proven to be involved in regulating the pathological process of COAD [16, 17]. It is reported that lncRNA can be regulated by m6A modification. However, whether lncRNAs are involved in the regulation of m6A modification in the progression of COAD needs further exploration.

With the approach of the era of personalized therapy managements, traditional diagnosing failed to satisfy advanced diagnoses and therapies. It is expected to build more useful prognostic signatures to help improve personalized treatment management. In this study, we focused on identifying prognosis-related lncRNAs associated with m6A

modification (m6A-lncRNAs) and used bioinformatics methods to establish a reliable risk model for patients with COAD.

We downloaded RNA sequencing data in TCGA. M6A-lncRNAs were recognized by Pearson correlation analysis according to the expression of lncRNAs and m6A regulators. There are 24 prognostic m6A-lncRNAs identified by univariate Cox regression analysis. For exploring the biological features of these 24 prognostic m6A-lncRNAs, we divided the COAD patients into two clusters. GSAE has been applied to investigate the potential cellular pathways and biological processes. Accumulated studies have illustrated that tumor immune microenvironment (TIME) is correlated with tumorigenesis and the development of COAD [38-41]. In this study, we have explored the relationship between immune infiltrate levels and m6A-lncRNAs. There are also significant differences between the two clusters in ESTIMATE score, stromal score, and immune score.



(c)
FIGURE 2: Continued.

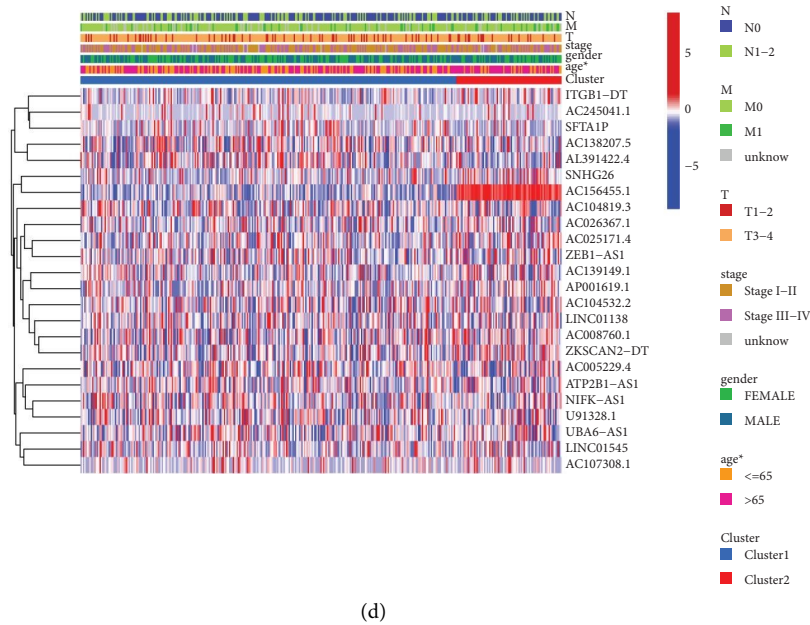


FIGURE 2: Consensus cluster classification. (a) Consensus clustering matrix when $k=2$. (b) The cumulative distribution functions (CDFs) plot for $k=2-9$. (c) The K M survival curves for COAD patients in 2 clusters. (d) Heatmap of 24 prognosis-related m6A-lncRNAs expression along with clinicopathological features in 2 subgroups. Note. * $P < 0.05$.

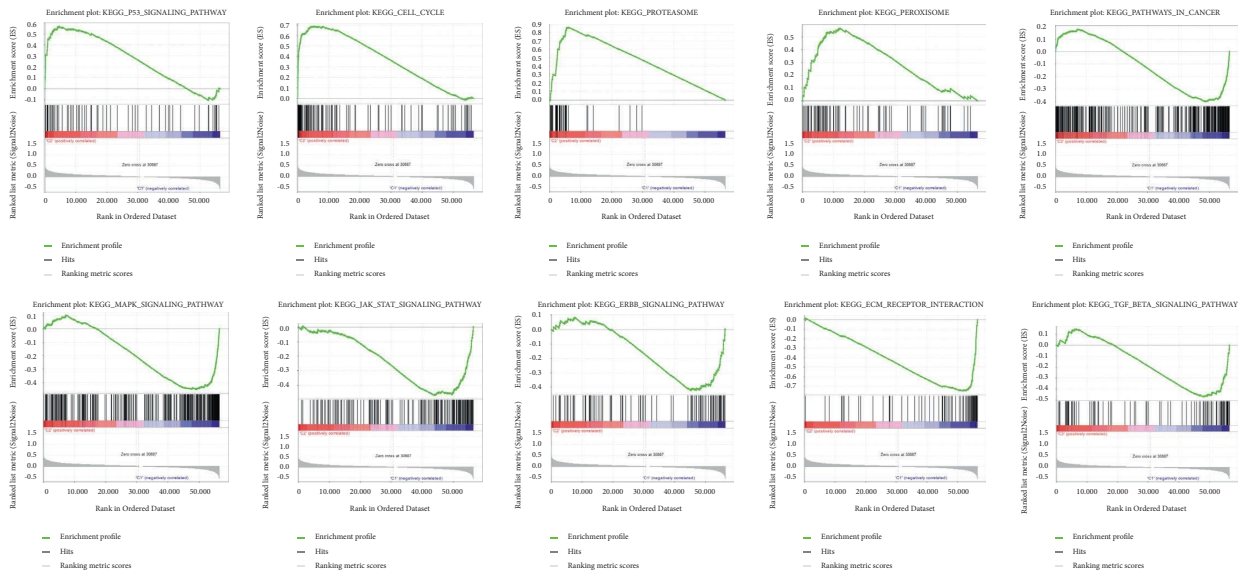


FIGURE 3: Gene set enrichment analysis.

The Lasso regression algorithm is widely used to construct risk models. The most important difference between lasso regression analysis and traditional stepwise Cox regression analysis is that it can process all variables simultaneously, instead of step by step [34], which greatly improves the stability of the model. In this study, LASSO regression analysis was conducted to avoid overfitting and to construct risk scores. Ultimately, 13 m6A-lncRNAs, including UBA6-AS1, AC139149.1, U91328.1, AC138207.5, AC025171.4, AC008760.1, ITGB1-DT, AP001619.1,

AL391422.4, AC104532.2, ZEB1-AS1, AC156455.1, and AC104819.3 were selected to construct the signature. To testify the reliability of the signature, COAD patients were divided, at a ratio of 7:3, into a training set and validation set. K M survival analysis and the ROC curves showed well discrimination of the signatures. In addition, risk score and clinicopathological characteristics were integrated into the analysis. Age, risk score, and pathological stage were recognized as independent prognostic factors, which also demonstrated the reliability of the risk model. Then,

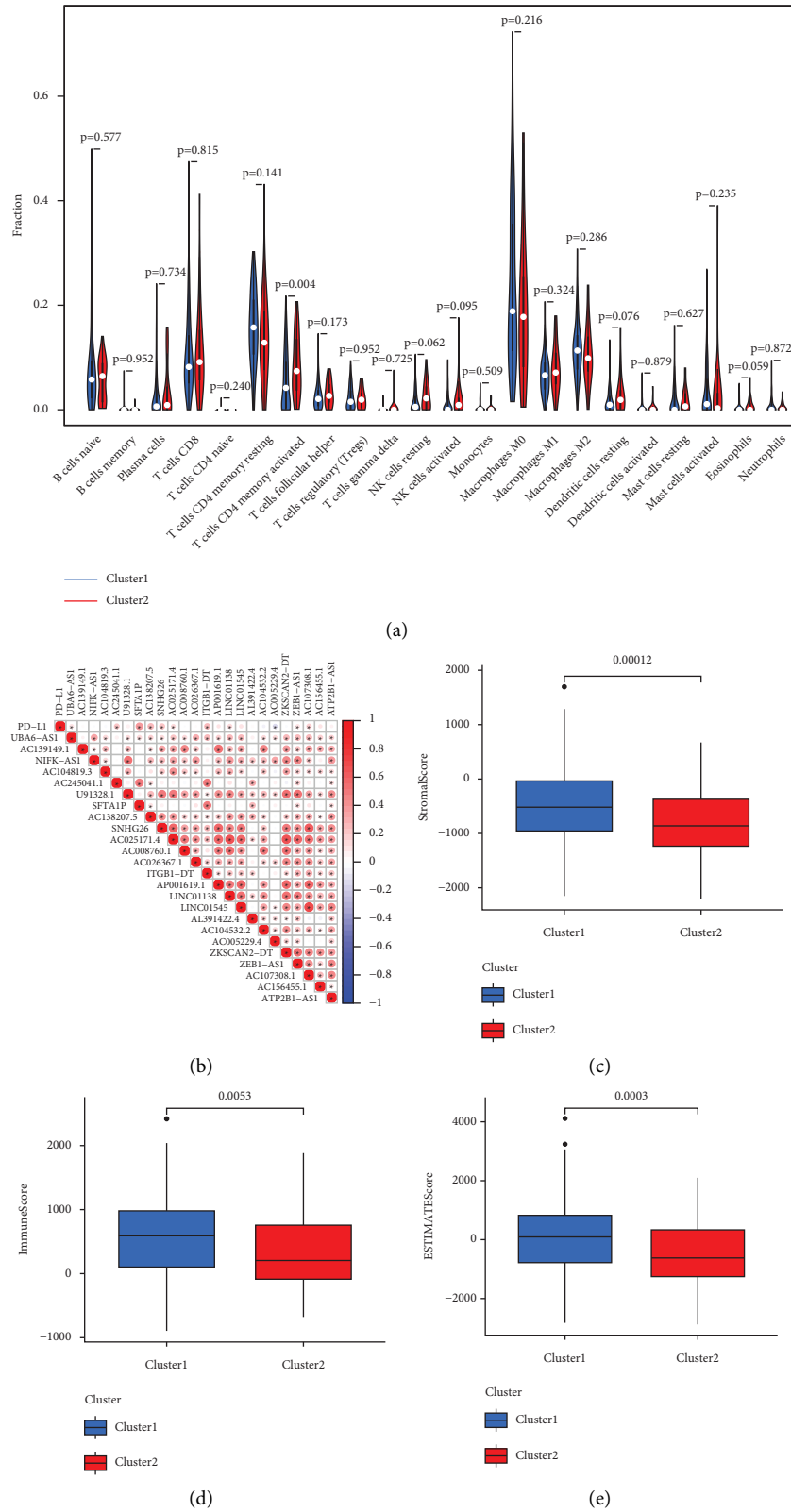


FIGURE 4: Immune cell infiltration and the distribution of immunity. (a) Infiltrating levels of 22 types of immune cells. (b) Correlation between the m6A-lncRNAs and PD-L1. (c-e) Stromal score, immune score, and ESTIMATE score in two clusters.

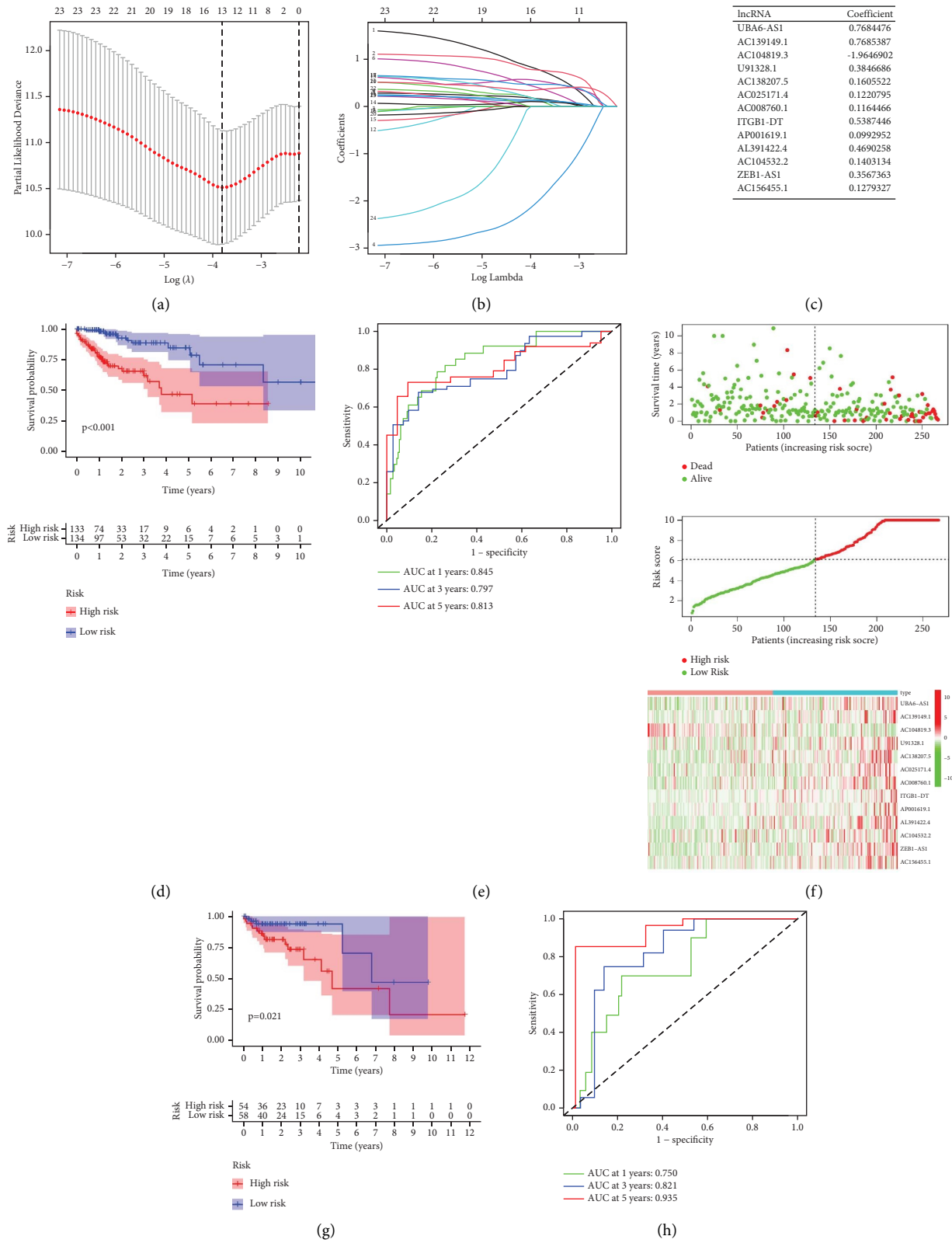


FIGURE 5: Continued.

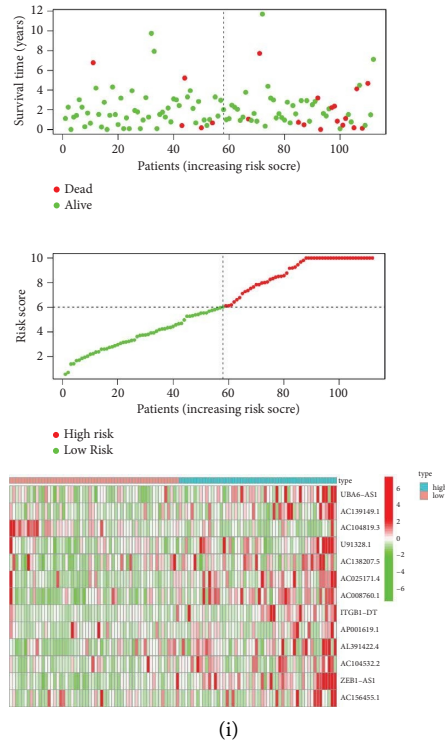


FIGURE 5: Construction and validation of prognostic signature. (a, b) Calculation of the optimal λ in Lasso regression analysis. (c) 13 prognosis lncRNAs and each of its coefficients. (d, e) The K M survival curves and ROC curves in training cohort. (g, h) The K M survival curves and ROC curves in validation cohort. (f, i) Survival time scatter plots and risk score distribution plots in training set and internal validation set.

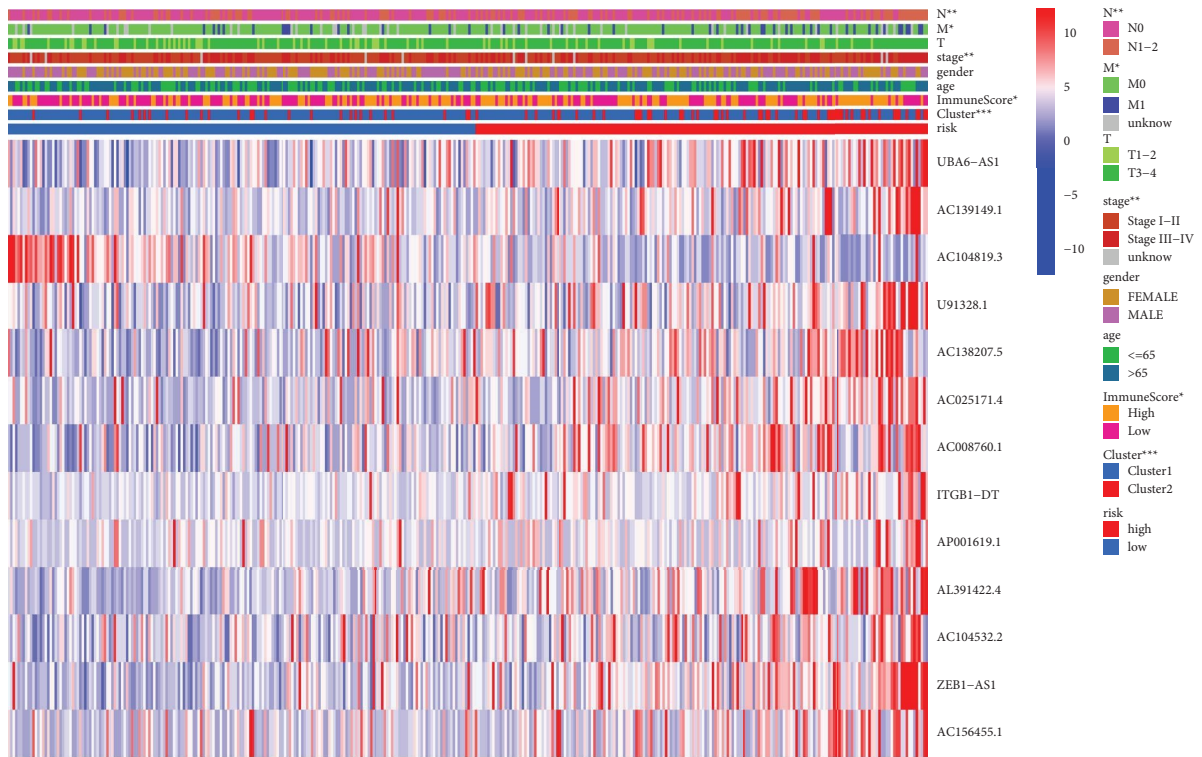


FIGURE 6: Continued.

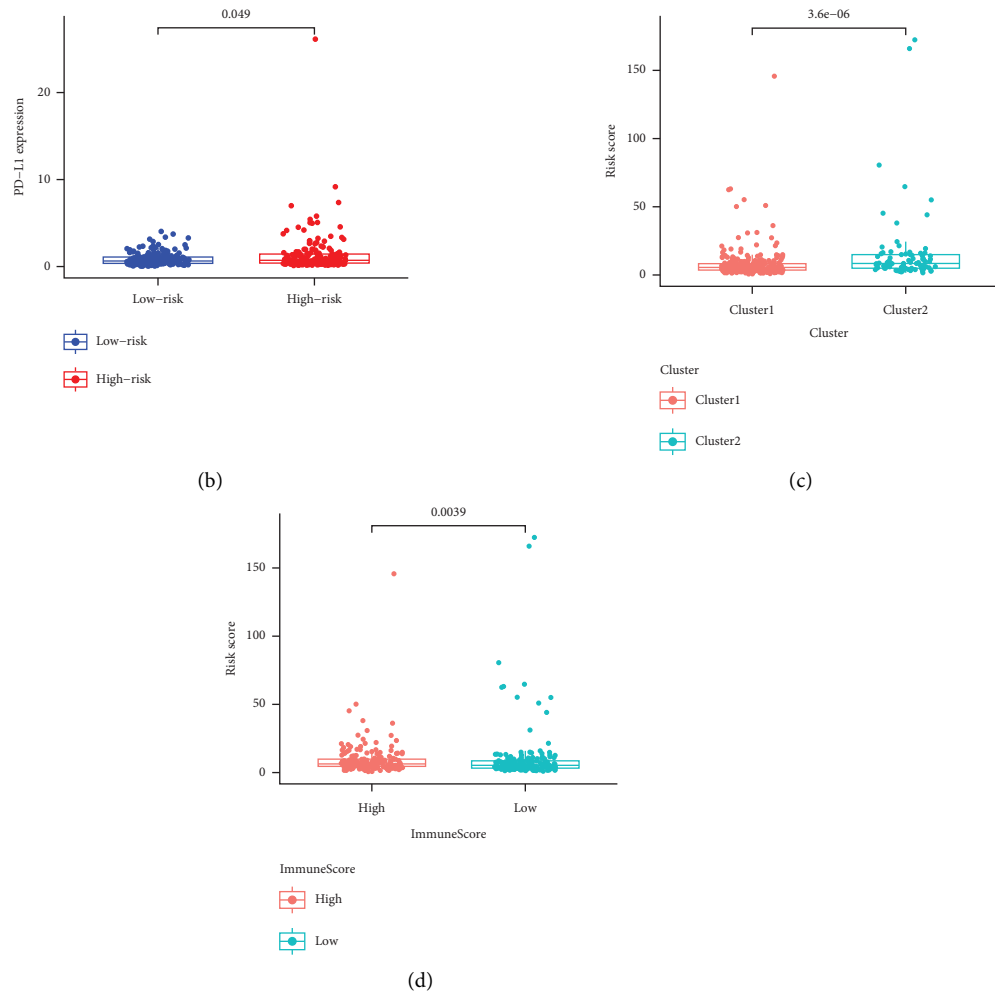


FIGURE 6: The correlation of risk score and clinical characteristics. (a) Heatmap of m6A-lncRNAs expression, immune score, and clinical features in low- and high-risk score subgroup. (b) Distribution of PD-L1 expression stratified by risk scores. (c, d) Distribution of risk scores stratified by clusters and immune score.

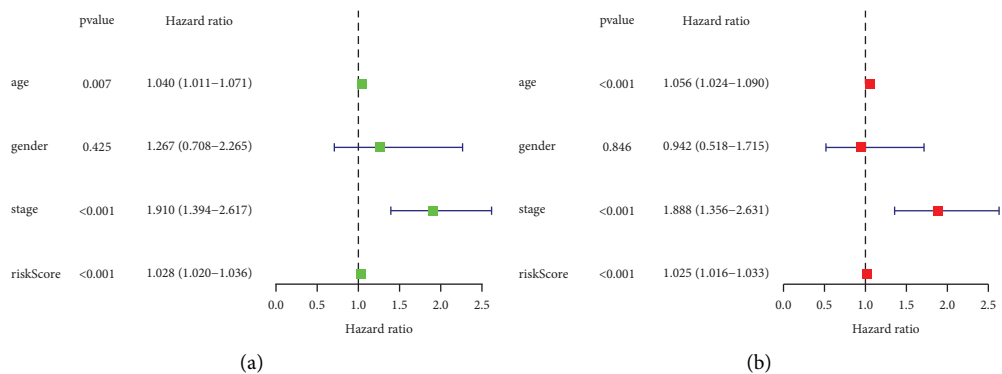


FIGURE 7: Continued.

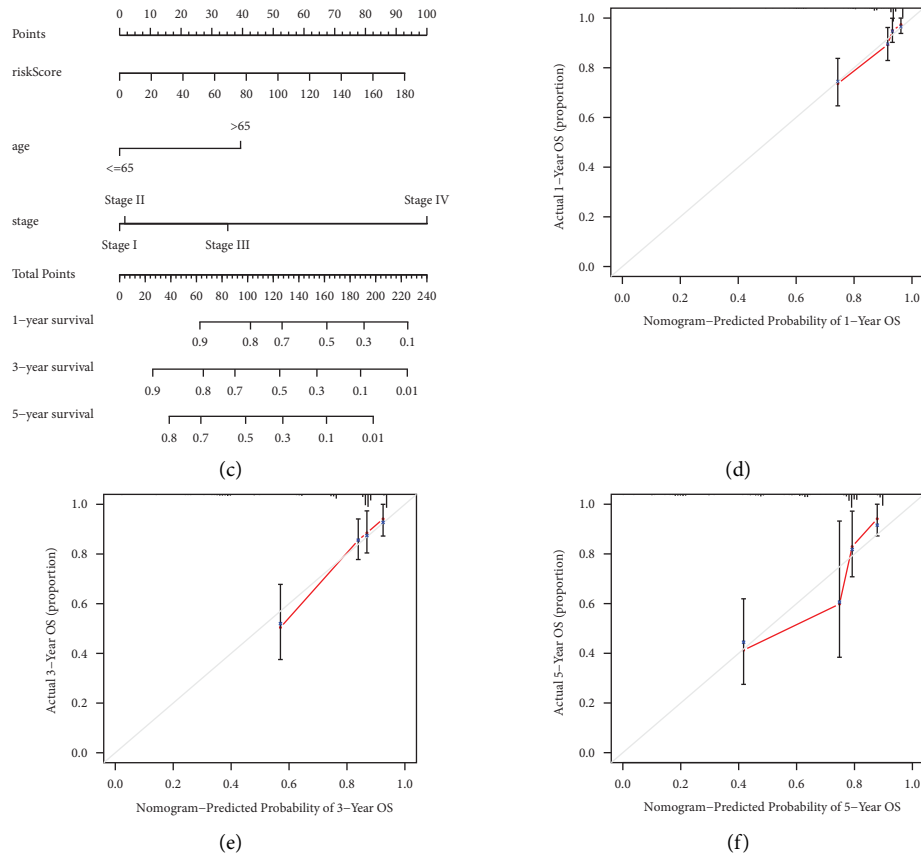


FIGURE 7: Construction of survival prediction nomogram. (a, b) The univariate and multivariate Cox regression analysis. (c) Nomogram based on risk score and prognostic clinical characteristics. (d–f) The calibration curves of the prognostic nomogram.

a quantitative prognostic nomogram was constructed. The calibration curves demonstrated the accuracy of the nomogram.

There were also limitations in our study. First, the m6A-lncRNAs risk model was constructed and validated based on the TCGA database. We did not verify the prognostic signatures in external independent cohorts. Second, the interaction between prognosis-related m6A-lncRNA and m6A regulator lacks experimental confirmation.

5. Conclusions

To sum up, we have identified some m6A-related lncRNAs which were correlated with prognosis and tumor immune microenvironment. A reliable alternative prognostic signature was provided for the management of COAD patients. We also combined risk scores with clinical characteristics to establish a quantitative prognostic nomogram.

Data Availability

The data analysed in this study are free available in The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov>).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The final version of the manuscript was approved by all the authors. SZZ designed the study and drafted the manuscript. YLP, QCD, CJY, and DJZ were in charge of data obtainment and statistical analysis. XJW, JZ, and MLL supervised this study and revised the manuscript. Su-Zhe Zhou, Ying-Lian Pan, Qing-Chun Deng, and Chang-Jun Yin contributed equally to this research and should be regarded as co-first authors.

Acknowledgments

This study was supported by the Hainan Natural Science Foundation High Level Talent Program (no. 821RC712) and Hainan Key R&D Program (no. ZDYF2022SHFZ068).

References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

- [3] K. D. Miller, L. Nogueira, T. Devasia et al., "Cancer treatment and survivorship statistics, 2022," *CA: A Cancer Journal for Clinicians*, vol. 72, no. 5, pp. 409–436, 2022.
- [4] Y. Okugawa, W. M. Grady, and A. Goel, "Epigenetic alterations in colorectal cancer: emerging biomarkers," *Gastroenterology*, vol. 149, no. 5, pp. 1204–1225.e12, 2015.
- [5] N. Liu and T. Pan, "N6-methyladenosine–encoded epitranscriptomics," *Nature Structural & Molecular Biology*, vol. 23, no. 2, pp. 98–102, 2016.
- [6] B. S. Zhao, I. A. Roundtree, and C. He, "Post-transcriptional gene regulation by mRNA modifications," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 1, pp. 31–42, 2017.
- [7] S. Zaccara, R. J. Ries, and S. R. Jaffrey, "Reading, writing and erasing mRNA methylation," *Nature Reviews Molecular Cell Biology*, vol. 20, no. 10, pp. 608–624, 2019.
- [8] T. Wang, S. Kong, M. Tao, and S. Ju, "The potential role of RNA N6-methyladenosine in Cancer progression," *Molecular Cancer*, vol. 19, no. 1, p. 88, 2020.
- [9] Z. Fang, Y. Hu, J. Hu, Y. Huang, S. Zheng, and C. Guo, "The crucial roles of N(6)-methyladenosine (m(6)A) modification in the carcinogenesis and progression of colorectal cancer," *Cell & Bioscience*, vol. 11, no. 1, p. 72, 2021.
- [10] D. Dai, H. Wang, L. Zhu, H. Jin, and X. Wang, "N6-methyladenosine links RNA metabolism to cancer progression," *Cell Death & Disease*, vol. 9, no. 2, p. 124, 2018.
- [11] M. Chen, L. Wei, C. T. Law et al., "RNA N6-methyladenosine methyltransferase-like 3 promotes liver cancer progression through YTHDF2-dependent posttranscriptional silencing of SOCS2," *Hepatology*, vol. 67, no. 6, pp. 2254–2270, 2018.
- [12] C. Zhao, Y. Liu, S. Ju, and X. Wang, "Pan-cancer analysis of the N6-methyladenosine eraser FTO as a potential prognostic and immunological biomarker," *International Journal of General Medicine*, vol. 14, pp. 7411–7422, 2021.
- [13] L. P. Vu, B. F. Pickering, Y. Cheng et al., "The N(6)-methyladenosine (m(6)A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells," *Nature Medicine*, vol. 23, no. 11, pp. 1369–1376, 2017.
- [14] G. Yan, Y. An, B. Xu, N. Wang, X. Sun, and M. Sun, "Potential impact of ALKBH5 and YTHDF1 on tumor immunity in colon adenocarcinoma," *Frontiers Oncology*, vol. 11, Article ID 670490, 2021.
- [15] S. Xiang, X. Liang, S. Yin, J. Liu, and Z. Xiang, "N6-methyladenosine methyltransferase METTL3 promotes colorectal cancer cell proliferation through enhancing MYC expression," *Am J Transl Res*, vol. 12, no. 5, pp. 1789–1806, 2020.
- [16] T. Li, P. S. Hu, Z. Zuo et al., "METTL3 facilitates tumor progression via an m(6)A-IGF2BP2-dependent mechanism in colorectal carcinoma," *Molecular Cancer*, vol. 18, no. 1, p. 112, 2019.
- [17] X. Chen, M. Xu, X. Xu et al., "METTL14-mediated N6-methyladenosine modification of SOX4 mRNA inhibits tumor metastasis in colorectal cancer," *Molecular Cancer*, vol. 19, no. 1, p. 106, 2020.
- [18] X. Yang, S. Zhang, C. He et al., "METTL14 suppresses proliferation and metastasis of colorectal cancer by down-regulating oncogenic long non-coding RNA XIIST," *Molecular Cancer*, vol. 19, no. 1, p. 46, 2020.
- [19] S. Chen and X. Shen, "Long noncoding RNAs: functions and mechanisms in colon cancer," *Molecular Cancer*, vol. 19, no. 1, p. 167, 2020.
- [20] S. J. Liu, H. X. Dang, D. A. Lim, F. Y. Feng, and C. A. Maher, "Long noncoding RNAs in cancer metastasis," *Nature Reviews Cancer*, vol. 21, no. 7, pp. 446–460, 2021.
- [21] L. Zuo, H. Su, Q. Zhang et al., "Comprehensive analysis of lncRNAs N(6)-methyladenosine modification in colorectal cancer," *Aging (Albany NY)*, vol. 13, no. 3, pp. 4182–4198, 2021.
- [22] X. Tan, Q. Li, Q. Zhang, G. Fan, Z. Liu, and K. Zhou, "Integrative analysis reveals potentially functional N6-methyladenosine-related long noncoding RNAs in colon adenocarcinoma," *Frontiers in Genetics*, vol. 12, Article ID 739344, 2021.
- [23] F. Fazi and A. Fatica, "Interplay between N (6)-methyladenosine (m(6)A) and non-coding RNAs in cell development and cancer," *Frontiers in Cell and Developmental Biology*, vol. 7, p. 116, 2019.
- [24] N. Liu, Q. Dai, G. Zheng, C. He, M. Parisien, and T. Pan, "N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions," *Nature*, vol. 518, no. 7540, pp. 560–564, 2015.
- [25] X. Zhang, M. H. Hamblin, and K. J. Yin, "The long noncoding RNA Malat1: its physiological and pathophysiological functions," *RNA Biology*, vol. 14, no. 12, pp. 1705–1714, 2017.
- [26] J. A. Brown, C. G. Kinzig, S. J. DeGregorio, and J. A. Steitz, "Methyltransferase-like protein 16 binds the 3'-terminal triple helix of MALAT1 long noncoding RNA," *Proceedings of the National Academy of Sciences of the U S A*, vol. 113, no. 49, pp. 14013–14018, 2016.
- [27] Z. Q. Zheng, Z. X. Li, G. Q. Zhou et al., "Long noncoding RNA FAM225A promotes nasopharyngeal carcinoma tumorigenesis and metastasis by acting as ceRNA to sponge miR-590-3p/miR-1275 and upregulate ITGB3," *Cancer Research*, vol. 79, no. 18, pp. 4612–4626, 2019.
- [28] X. Zuo, Z. Chen, W. Gao et al., "M6A-mediated upregulation of LINC00958 increases lipogenesis and acts as a nano-therapeutic target in hepatocellular carcinoma," *Journal of Hematology & Oncology*, vol. 13, no. 1, p. 5, 2020.
- [29] A. D. Yates, P. Achuthan, W. Akanni et al., "Ensembl 2020," *Nucleic Acids Research*, vol. 48, no. 1, pp. D682–d688, 2020.
- [30] J. Zheng, J. Guo, B. Cao, Y. Zhou, and J. Tong, "Identification and validation of lncRNAs involved in m6A regulation for patients with ovarian cancer," *Cancer Cell International*, vol. 21, no. 1, p. 363, 2021.
- [31] J. Wang, C. Shen, D. Dong, X. Zhong, Y. Wang, and X. Yang, "Identification and verification of an immune-related lncRNA signature for predicting the prognosis of patients with bladder cancer," *International Immunopharmacology*, vol. 90, Article ID 107146, 2021.
- [32] M. D. Wilkerson and D. N. Hayes, "ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, pp. 1572–1573, 2010.
- [33] R. Alhamzawi and H. T. M. Ali, "The Bayesian adaptive lasso regression," *Mathematical Biosciences*, vol. 303, pp. 75–82, 2018.
- [34] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.
- [35] P. Blanche, J. F. Dartigues, and H. Jacqmin-Gadda, "Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks," *Statistics in Medicine*, vol. 32, no. 30, pp. 5381–5397, 2013.
- [36] A. Iasonos, D. Schrag, G. V. Raj, and K. S. Panageas, "How to build and interpret a nomogram for cancer prognosis,"

- Journal of Clinical Oncology*, vol. 26, no. 8, pp. 1364–1370, 2008.
- [37] R. Tibshirani, “The lasso method for variable selection in the Cox model,” *Statistics in Medicine*, vol. 16, no. 4, pp. 385–395, 1997.
- [38] W. Q. Liu, W. L. Li, S. M. Ma, L. Liang, Z. Y. Kou, and J. Yang, “Discovery of core gene families associated with liver metastasis in colorectal cancer and regulatory roles in tumor cell immune infiltration,” *Translational Oncology*, vol. 14, no. 3, Article ID 101011, 2021.
- [39] R. Ouyang, Z. Li, P. Peng et al., “Exploration of the relationship between tumor mutation burden and immune infiltrates in colon adenocarcinoma,” *International Journal of Medical Sciences*, vol. 18, no. 3, pp. 685–694, 2021.
- [40] P. Zhang, G. Liu, and L. Lu, “N6-Methyladenosine-Related lncRNA signature is a novel biomarkers of prognosis and immune response in colon adenocarcinoma patients,” *Frontiers in Cell and Developmental Biology*, vol. 9, Article ID 703629, 2021.
- [41] H. Zhang, L. Zhao, S. Li et al., “N6-Methyladenosine-Related lncRNAs in tumor microenvironment are potential prognostic biomarkers in colon cancer,” *Frontiers Oncology*, vol. 11, Article ID 697949, 2021.

Research Article

Chromatin Regulator-Related Gene Signature for Predicting Prognosis and Immunotherapy Efficacy in Breast Cancer

Dongxu Feng,¹ Wenbing Li,¹ Wei Wu,¹ Ulf Dietrich Kahlert,² Pingfa Gao,¹ Gangfeng Hu,¹ Xia Huang,¹ Wenjie Shi ^{2,3} and Huichao Li ⁴

¹Department of General Surgery, Chongming Hospital Affiliated to Shanghai University of Medicine and Health Sciences, Chongming District, Shanghai 202150, China

²University Clinic for General, Visceral, Vascular-and Transplantation Surgery, Faculty of Medicine, Otto-von-Guericke-University, Magdeburg 39120, Germany

³University Hospital for Gynaecology, Pius-Hospital, University Medicine Oldenburg, Oldenburg 26121, Germany

⁴Department of Thyroid Surgery, The Affiliated Hospital of Qingdao University, Qingdao 266555, Shandong, China

Correspondence should be addressed to Wenjie Shi; wenjie.shi@uni-oldenburg.de and Huichao Li; lihuichao@qdu.edu.cn

Received 3 September 2022; Revised 22 October 2022; Accepted 24 November 2022; Published 30 January 2023

Academic Editor: Feng Jiang

Copyright © 2023 Dongxu Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Many studies have found that chromatin regulators (CRs) are correlated with tumorigenesis and disease prognosis. Here, we attempted to build a new CR-related gene model to predict breast cancer (BC) survival status. **Methods.** First, the CR-related differentially expressed genes (DEGs) were screened in normal and tumor breast tissues, and the potential mechanism of CR-related DEGs was determined by function analysis. Based on the prognostic DEGs, the Cox regression model was applied to build a signature for BC. Then, survival and receiver operating characteristic (ROC) curves were performed to validate the signature's efficacy and identify its independent prognostic value. The CIBERSORT and tumor immune dysfunction and exclusion (TIDE) algorithms were used to assess the immune cells infiltration and immunotherapy efficacy for this signature, respectively. Additionally, a novel nomogram was also built for clinical decisions. **Results.** We identified 98 CR-related DEGs in breast tissues and constructed a novel 6 CR-related gene signature (ARID5A, ASCL1, IKZF3, KDM4B, PRDM11, and TFF1) to predict the outcome of BC patients. The prognostic value of this CR-related gene signature was validated with outstanding predictive performance. The TIDE analysis revealed that the high-risk group patients had a better response to immune checkpoint blockade (ICB) therapy. **Conclusion.** A new CR-related gene signature was built, and this signature could provide the independent predictive capability of prognosis and immunotherapy efficacy for BC patients.

1. Introduction

Breast cancer (BC) is a common cancer in the world [1]. Although the widespread use of adjuvant chemotherapy and hormonal drugs has reduced mortality in BC patients, there are still individual differences in treatment outcomes and different clinical benefits for BC patients [2]. Fortunately, with the continuous updating of new therapies, the use of cancer biomarkers has become an aid in BC diagnosis, prognosis, treatment response prediction, and disease monitoring during and after treatment [3]. Nowadays, many researchers have tried to use various bioinformatics techniques to identify the

biomarkers or build the risk model in BC and achieved good research results, such as the machine learning model in predicting the immune subtype [4] and the eight-lncRNA prognostic model [5]. Therefore, in order to provide a more effective prediction of survival in tumor patients, continuous efforts are needed to identify new prognostic key molecules and explore their prognostic values.

Chromatin regulators (CRs) are essential upstream regulatory factors in epigenetics that can cooperate to connect the organizational scales of chromatin from nucleosome assembly to the establishment of functional chromatin domains [6]. CRs can highly regulate chromatin structure by

four broad classes of nuclear factors, including histone variants, histone chaperones, chromatin remodelers, and histone modifiers [7]. In human cancer pathogenesis, it has been found that the function of CRs is frequently disrupted by genetic mutations and epigenetic alterations, resulting in perturbed gene expression profiles. The role of CRs in cancer is complex and highly specific [8]. To begin with, the carcinogenic effects of chromatin regulators are well established. Meanwhile, recent new studies have demonstrated their tumor suppressive properties in the regulation of multiple cellular processes [9]. For example, MLL3/4, one of the chromatin regulators, may exert broad tumor suppressor effects in various cancers [10]. In addition, growing evidence has shown that CRs and tumor prognosis are closely related. In cervical cancer, 57 overexpressed chromatin regulators were identified to have prognostic significance [11]. Similarly, CR-related genes were found to be significantly associated with postoperative outcomes in astrocytomas [12]. However, CRs affecting prognosis of breast cancer is still lacking more understanding and research.

The tumor microenvironment (TME), consisting of extracellular matrix, stromal cells, and immune cells [13], is important in cancer initiation, progression, and therapy [14]. With the emergence of new technologies, immunotherapy has become a widely concerned research direction for cancer treatment, which aims to use the immune system as a tool for treating oncological diseases [15]. The effect of immunotherapy is associated with the TME, especially the tumor-infiltrating immune cells (TICs) [16]. Currently, there are 3 types of immunotherapies that target tumor-specific T cells, including immune checkpoint blockade (ICB), adoptive cellular therapies, and cancer vaccines [17]. In immune checkpoint blockade therapy, multiple cancers have been treated with two different antiprogrammed cell death protein 1 (PD1) drugs (nivolumab and pembrolizumab) [18] and anticytotoxic T lymphocyte-associated antigen 4 (CTLA4) (ipilimumab) [19]. The use of therapeutic antibodies does provide clinical benefit to a small patient population, but adverse effects associated with immune checkpoint blockade complicate immunotherapy and limits its use in cancer patients [20]. Therefore, in order to promote the efficacy of ICB therapy, it is necessary to identify and explore the predictive biomarkers for immune checkpoint-blocking therapies [21, 22].

In this work, we investigated the potentiality of CR-related DEGs as prognostic markers in BC through several bioinformatic analyses. Furthermore, CR-related DEGs were successfully employed to construct a six-gene prognostic model and a new CR-related gene signature for predicting patients' outcome and immunotherapy efficacy.

2. Materials and Methods

2.1. Data Collection. The training set was downloaded from the The Cancer Genome Atlas (TCGA) database ($n = 1167$; 113 normal samples vs. 1054 tumor samples). The validate set was obtained from GSE20685 ($n = 327$) [23]. Considering that a patient's follow-up for less than 30 days is too short to evaluate the prognosis, we excluded those patients for

a more accurate evaluation. The complete clinical information of patients is shown in Supplementary Table 1. Lastly, there were 870 chromatin regulator (CR)-related genes obtained from previous research [6].

2.2. Identification of CR-Related DEGs. First, the data of the training set were normalized with the log₂ transformation. Next, the expression profiles of 853 CRs were extracted from the normalized TCGA-BRCA matrix based on the obtained CR names. In the training set, DEGs related to CRs were performed by the "limma" package. Differential expression was set as the adjusted p value less than 0.05, and the absolute value of log₂FC (fold change) was greater than 1. The result of the analysis was visualized in a volcano map utilizing the "ggplot2" R package.

2.3. Biology Function Enrichment. ClusterProfiler was conducted to analyze gene oncology and the KEGG pathway, and this procedure aims at discovering the potential mechanism of the CR-related DEGs. In addition, the enrichment analysis results were visualized via circos diagrams.

2.4. Signature Model Construction and Validation. Batch Cox regression was applied to screen the prognosis-related DEGs of CRs in the training set. And the important candidate genes were analyzed by the multivariate Cox analysis for identifying independent prognosis risk factors. Based on the multivariate regression results, we then developed an optimal signature to predict prognosis based on the coefficients (Coef). The model was constructed as the risk formula: risk score = A gene \times Coef + B gene \times Coef + . . . + X gene \times Coef. The forest map was plotted by the "forestplot" R package. The area under curve of ROC will be used to evaluate the model in the training and validation set, respectively.

2.5. Signature Prognostic Value Evaluation. Other clinical variables also affect patient clinical prognosis, and when compared with risk score, if risk score is also associated with patients outcomes, here, we use Cox regression model to demonstrate it, and we also conduct subgroup analyses of risk score and survival status.

2.6. Immune Cell Infiltration and Immunotherapy Efficacy Estimation. To evaluate the association between this signature and tumor-infiltrating immune cells, CIBERSORT was used to analyze the proportion of various TICs between high-score and low-score patients in the TCGA training set. We also use the tumor immune dysfunction and exclusion (TIDE) indicator to estimate ICB therapy response in BC patients. Among them, lower TIDE scores mean that tumor cells have less chance with immune escape, indicating a higher response rate to ICBs therapy. Finally, we assessed the correlation between this signature and immunotherapeutic markers and ICB-related genes via Wilcoxon test.

2.7. Predictive Nomogram Establishment and Evaluation. A nomogram, a visual scoring system, was established in the training set by integrating the risk score of CR-related gene signature and clinicopathological variables to be used to evaluate the OS of the BC patients. At the same time, the bootstrap method was used to calculate the concordance index (C-index), ROC, and calibration curves. Besides, we further utilized the Kaplan–Meier analysis for OS, DSS, and PFS to prove the clinical prognostic value of the nomogram. All the analyses and plots were done by the R software and suitable R packages.

3. Results

3.1. CR-Related DEGs and Function Enrichment. In the training dataset, a total of 98 CR-related DEGs (Figure 1(a)) (Supplementary Table 2) were screened from 1154 BC tissue samples and 113 normal tissue samples. We predicted the biological mechanisms of 98 CR-related DEGs. GO enrichment analysis showed that the most highly enriched molecular functions (MF) of these DEGs were histone binding, transcription corepressor activity, and histone kinase activity. Cellular components (CC) of these DEGs were covalent chromatin modification, histone modification, and DNA replication. The biological process (BP) of these DEGs was remarkably involved in covalent chromatin modification, histone modification, and DNA replication. Moreover, they are also enriched in the cell cycle, homologous recombination, and lysine degradation by KEGG enrichment analysis (Figure 1(b)).

3.2. Prognostic Signature Construction Based on the CR-Related Gene. We used 98 CR-related DEGs to build the prognostic signature in the TCGA training set. First, ten important CR-related genes associated with OS in BC patients were selected (Figure 1(c)). Then, six hub CR-related genes (ARID5A, ASCL1, IKZF3, KDM4B, PRDM11, and TFF1) were identified as associated with the OS and contributed to the risk signature (Figure 1(d)) (Table 1). The risk score = $(-0.201073277) * ARID5A + 0.1231206 * ASCL1 + (-0.268713644) * IKZF3 + (-0.197882736) * KDM4B + (-0.538035757) * PRDM11 + (-0.050431061) * TFF1$.

3.3. Evaluation and Validation of Signature Efficacy. The high-risk group was defined as risk score more than the median value, others were defined as low-risk groups, respectively, in the training dataset. Then, the risk score distribution between the two groups was compared (Figure 2(a)). Patients with high risk have a poor outcome while the low-risk patients always suggest a better prognosis ($p = 1.179e - 04$) (Figure 2(b)). Moreover, the area under the ROC curve (AUC) was 0.713, which proved the predictive efficacy of our signature for predicting the OS in the training set (Figure 2(c)), which was also validated in the GEO. The risk score distribution in the two groups is shown in Figure 2(d). In addition, patients with low risk also have a better prognosis ($p = 7.955e - 04$) (Figure 2(e)), and the

results further supported the above conclusion: AUC is equal to 0.821 (Figure 2(f)).

3.4. Applicability of the CR-Related Gene Signature. We integrated the signature with other clinical risk factors in the total data set to further assess the independent prognostic value of the risk model for BC. As shown in the TCGA training dataset, age, TNM stage, ER status, and risk score were associated with the OS ($p < 0.05$), and risk score was an independent prognostic factor ($p < 0.05$) (Figures 3(a) and 3(b)). Additionally, in the GEO validation dataset, analysis results also support the abovementioned conclusions (Figures 3(c) and 3(d)).

3.5. Subgroup Analysis to Evaluate Gene Signature. We performed survival subgroup analysis in the TCGA training set to demonstrate that the signature is associated with clinical features. First of all, the clinical patients were classified into two groups, such as age ≤ 60 vs. >60 groups, the T1/2 stage vs. T3/4 stage, the N (-) stage vs. N (+) stage (N0 and N1–N3, respectively), and the M0 stage vs. M1 stage. Subgroup analysis suggests that subgroup patients also have survival rate differences in age, T stage, N1–N3 stage, and M0 stage between high- and low-risk groups. However, patients with the N0 or M1 stage has no significance ($p = 0.080$ and $p = 0.486$, respectively) (Figures 4(a)–4(h)).

3.6. Immune Cell Infiltration and Immunotherapy Efficacy Estimation for the Signature. As for the relationship between the signature and TME, 19 of the 22 TICs were significantly related to the risk score (Figure 5(a)). Here, the CD8⁺ T cells, resting memory CD4⁺ T cells, regulatory T cells (Tregs), gamma delta T cells, follicular helper T cells, M1 macrophages, memory B cells, naive B cells, activated NK cells, monocytes, resting myeloid dendritic cells, and activated mast cells were negatively correlated with the risk score, while the M0 macrophages, M2 macrophages, resting NK cells, activated myeloid dendritic cells, resting mast cells, eosinophils, and neutrophils were positively correlated with the risk score. Meanwhile, the proportion of 12 TICs did differ significantly between the two groups ($p < 0.05$, Figure 5(b)). The proportion of M0 and M2 macrophages was significantly higher in the high-risk group, while the proportion of plasma cells, naive B cells, CD8⁺ T cells, gamma delta T cells, resting memory CD4⁺ T cells, resting NK cells, and resting mast cells was significantly higher in the low-risk group.

Regarding the CR-related signature's potential for predicting response to immunotherapy in BC patients, the result showed that the low-risk group had a higher TIDE score ($p < 0.001$, Figure 5(c)), meaning a poor response to immunotherapy. Likewise, 13 immune checkpoint molecules (CTLA4, PDL1, PD1, PDL2, LAG3, TIM3, CD86, BTLA, ICOS, CD96, CD160, TIGIT, and IDO1) are also positive with low-risk group gene expression ($p < 0.05$, Figure 5(d)). Taken together, these results illustrated the importance of the risk score in breast cancer immunotherapy prediction.

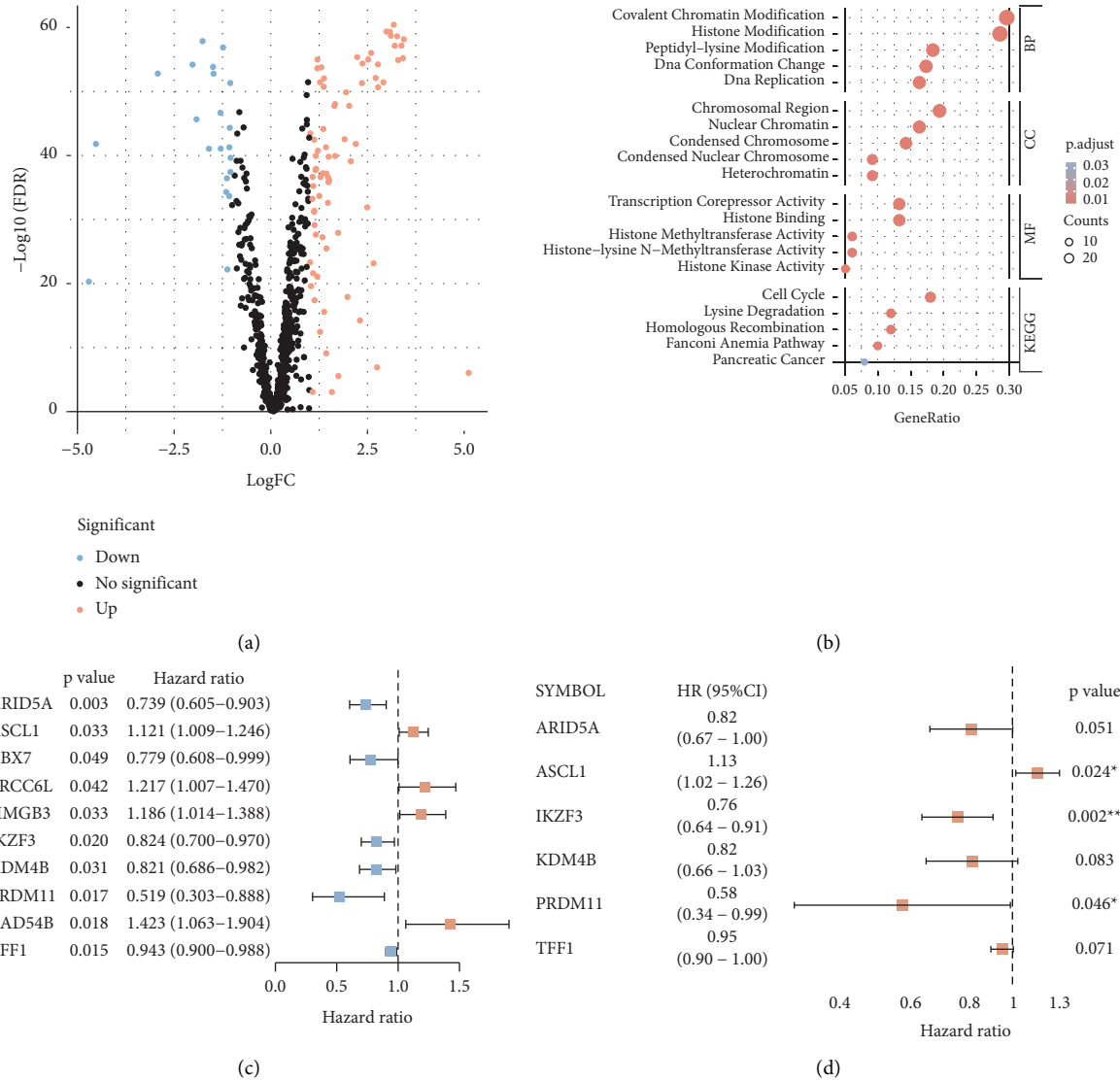


FIGURE 1: Differential expression analysis and the construction of CR-related prognosis signature in the TCGA training set. Volcano map of the CR-related DEGs. Red represents the up-regulated genes; blue represents the down-regulated genes ($|\log_{2}FC| > 1$, FDR q value < 0.05) (a). GO enrichment analysis (consisting of BP, CC and MF) and KEGG pathway enrichment analysis for the CR-related DEGs ($p < 0.05$, FDR q value < 0.05) (b). Univariate Cox regression analysis selected 10 CR-related genes correlated with OS (c). Multivariate Cox regression analysis identified a 6-gene (ARID5A, ASCL1, IKZF3, KDM4B, PRDM11, and TFF1) prognostic signature (d). * $p < 0.05$ and ** $p < 0.01$.

TABLE 1: Six CR-related prognostic genes significantly associated with OS in breast cancer patients.

| Gene | Multivariate cox regression analysis | | | | |
|--------|--------------------------------------|-------------|-------------|-------------|-------------|
| | Coef | HR | HR 0.95L | HR 0.95 H | p value |
| ARID5A | -0.201073277 | 0.8178525 | 0.668022015 | 1.001288425 | 0.051475659 |
| ASCL1 | 0.1231206 | 1.131020814 | 1.016273871 | 1.258723773 | 0.02408809 |
| IKZF3 | -0.268713644 | 0.764362104 | 0.642458524 | 0.909396333 | 0.002434433 |
| KDM4B | -0.197882736 | 0.820466059 | 0.655829072 | 1.026432927 | 0.083334851 |
| PRDM11 | -0.538035757 | 0.583894036 | 0.344413023 | 0.989893597 | 0.045751465 |
| TFF1 | -0.050431061 | 0.950819475 | 0.900181977 | 1.004305459 | 0.070902358 |

CRs, chromatin regulators; OS, overall survival; Coef, β coefficient; HR, hazard ratio.

3.7. Nomogram for Clinical Decision. A visualized nomogram was built and was used to predict OS probability at three periods (1, 3, and 5 years) (Figure 6(a)). The

concordance C-index was 0.734, which illustrated a good ability in predicting OS for BC patients. Furthermore, we found that the calibration curves indicated that the predicted

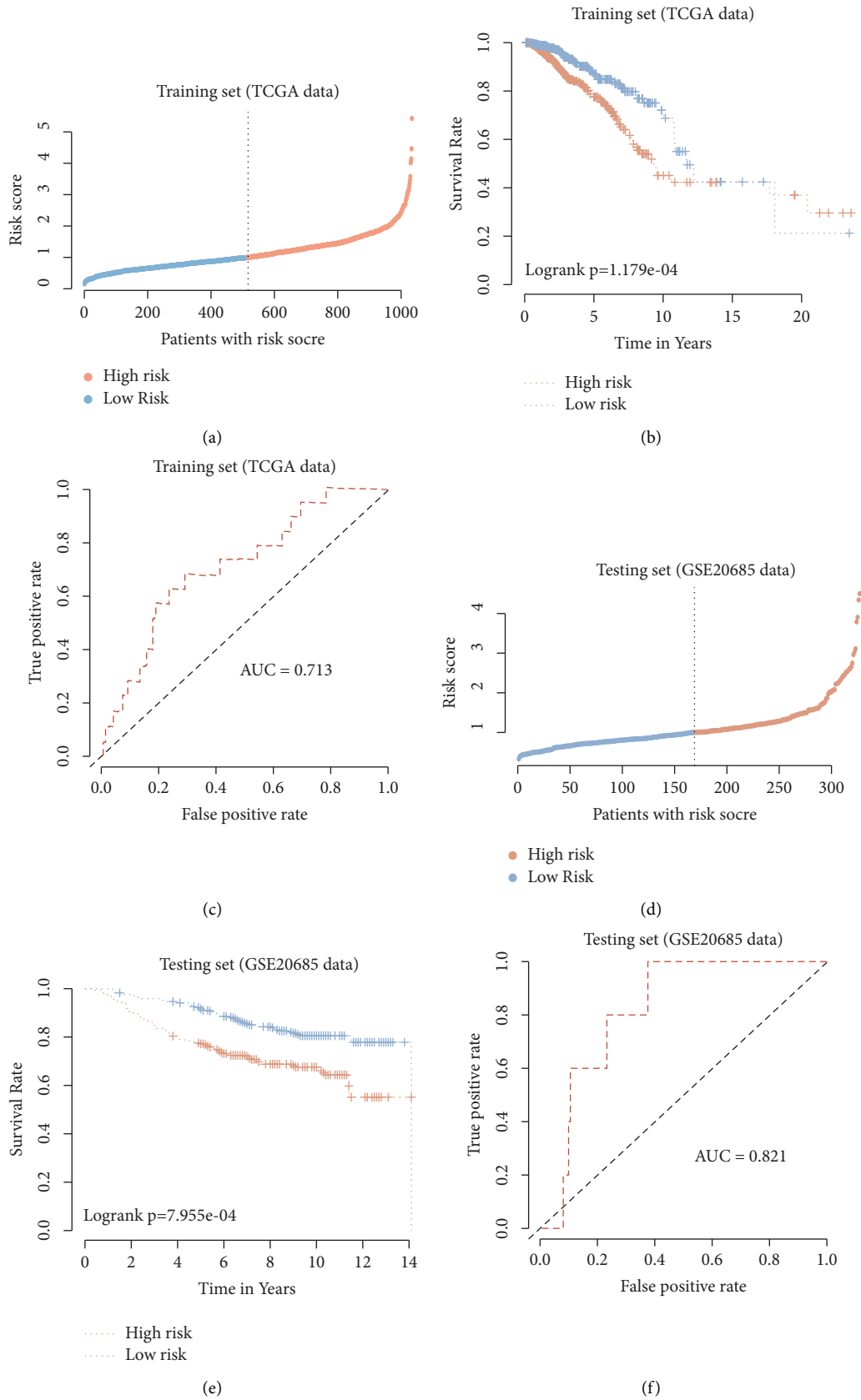


FIGURE 2: Validation of the CR-related prognostic signature in BC patients. Risk score distribution for patients in low- and high-risk groups from the TCGA training set (a). Kaplan–Meier survival analysis of OS between the low- and high-risk groups from the TCGA training set (b). AUC in ROC analysis for risk scores predicting the OS from the TCGA training set (c). Risk score distribution for BC patients in low- and high-risk groups from the GEO validation set (d). Kaplan–Meier survival analysis of OS between the low- and high-risk groups from the GEO validation set (e). AUC in ROC analysis for risk scores predicting the OS from the GEO validation set (f).

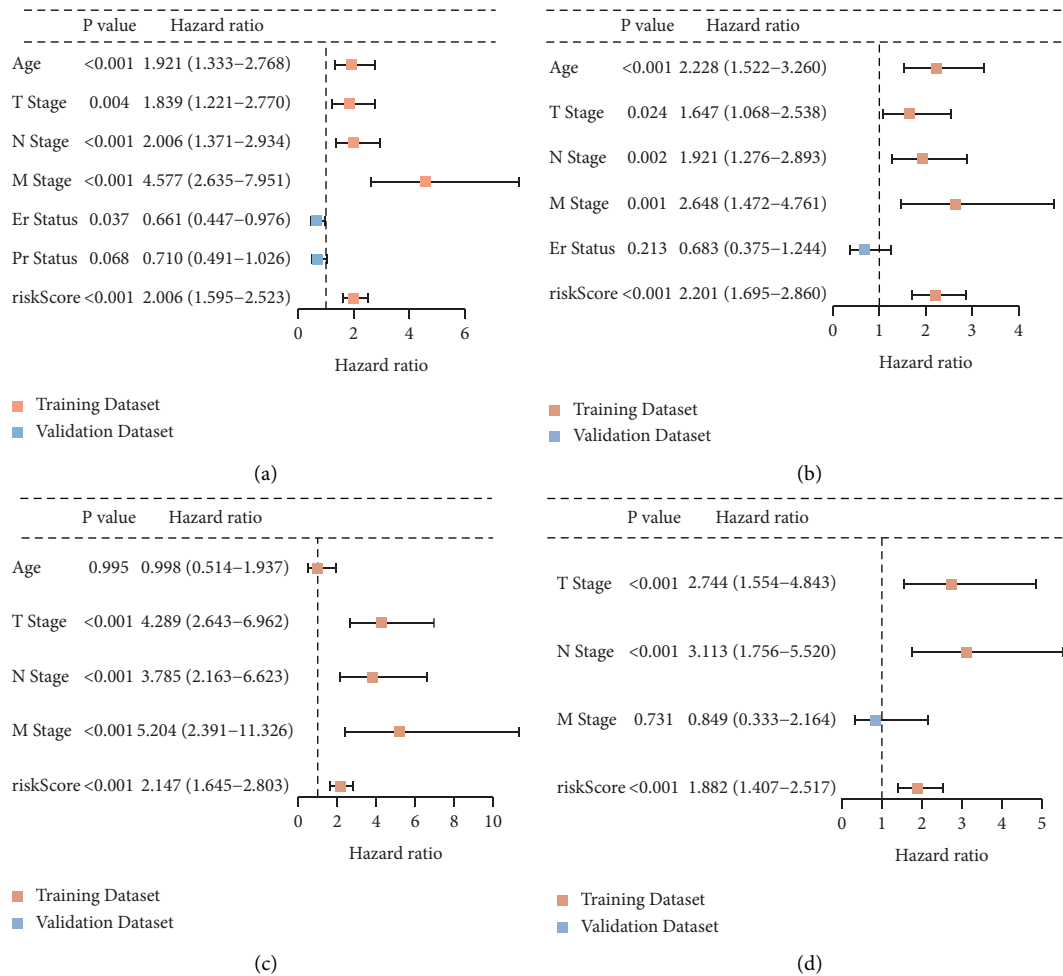


FIGURE 3: Independent prognostic analysis of the clinical traits and risk score in BC patients. Univariate and multivariate Cox regression analyses of the OS in the TCGA training set (a) and (b). Univariate and multivariate cox regression analyses of the OS in the GEO validation set (c) and (d).

curve is close to the ideal curve (Figure 6(b)). At the same time, the ROC curves of this nomogram also showed a good accuracy to individually predict OS for BC patients (AUC = 0.788, 0.731, and 0.713, respectively) (Figures 6(c)–6(e)). Lastly, we also assessed the prognostic value of the nomogram, finding that it was remarkably associated with OS, DSS, and PFS ($p < 0.05$) (Figure 6(f)).

4. Discussion

A significantly poor prognosis for the stage IV female BC patients who were diagnosed between 2007 and 2013 can be observed [24]. To improve the survival rate of BC patients, an increasing number of genetic signatures have been established. For example, Shi et al. [25] developed a five-mRNA signature model based on the ceRNA network for predicting the survival of BC. However, few studies have focused on prognostic signatures based on key genes in CRs that have been shown to have potential as prognostic markers. What's more, studies have shown that epigenetic factors and mechanisms can be involved in regulating the TME and the ICBs response. In lung adenocarcinomas,

ASF1A deficiency could sensitize lung adenocarcinomas to anti-PD-1 therapy by inducing immunogenic M1-like macrophage differentiation and enhancing T cell activation of the TME [26]. Then, in checkpoint-blocked refractory mouse melanoma, histone demethylase LSD1 depletion enhanced tumor immunogenicity and T-cell infiltration in poorly immunogenic tumors and elicited a significant response to anti-PD-1 treatment [27]. Thus, exploring and evaluating the CR-related gene expression in BC patients is important. In our study, we constructed a CR-related gene signature, a useful tool to predict patients' outcomes and immunotherapy sensitivity.

The CR-related gene signature has 6 hub genes; ASCL1, as one of the 6 genes, was positively associated with outcome, while the levels of ARID5A, IKZF3, KDM4B, PRDM11, and TFF1 were negatively associated with survival. As the results demonstrated, the model could predict the patient's outcome.

The ROC curves also confirmed the favorable predictive performance of this signature. Besides, the independent prognostic analysis determined that this signature, age, and TNM stage were independent predictors for BC prognosis.

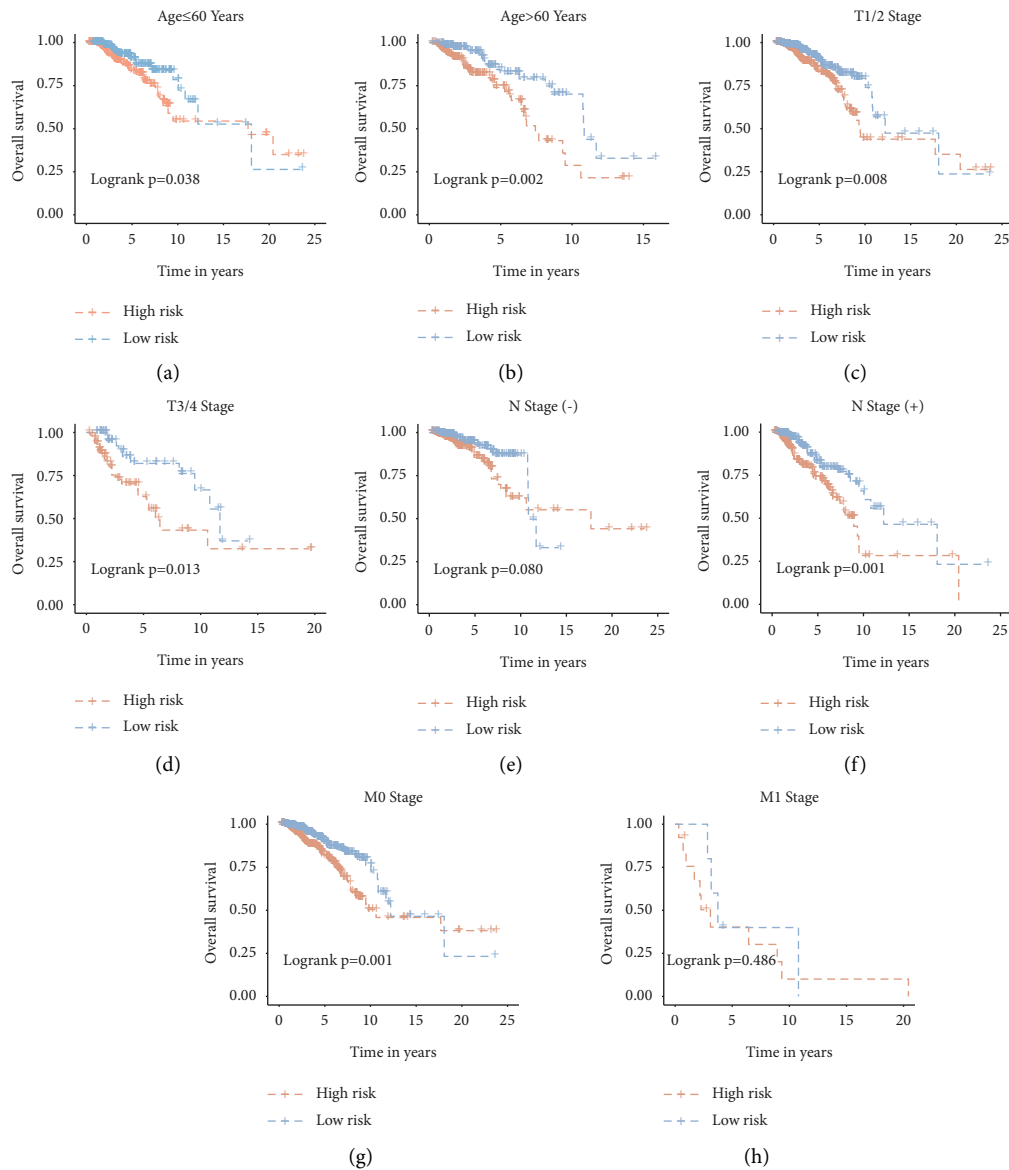


FIGURE 4: Kaplan–Meier survival subgroup analysis in BC patients from the TCGA training set based on the 6-gene signature stratified by clinical characteristics. The OS differences between low- and high-risk group stratified by age (≤ 60 years and > 60 years, respectively) (a) and (b). The OS differences between low- and high-risk group stratified by T stage (T1-2 and T3-4, respectively) (c) and (d). The OS differences between low- and high-risk group stratified by N ($-$) and N ($+$) stage. (N0 and N1–3, respectively) (e) and (f). The OS differences between low- and high-risk group stratified by M stage (M0 and M1, respectively) (g) and (h).

As expected, survival subgroup analysis also suggests the effective prediction ability of the signature in subgroups. Next, we found that patients with high-risk conditions will obtain a better response to immunotherapy and were more suitable for ICB therapy. Moreover, a nomogram model consisting of clinical factors and signature was established, which showed good power and accuracy with a high AUC value and C-index in BC patients. Therefore, the CR-related gene signature was a reliable model to predict prognosis and immunotherapy efficacy, which might have potential implications in clinical practice for BC.

All six genes involved in the signature model are associated with chromatin regulation or tumorigenesis.

Achaete-scute complex homolog 1 (ASCL1) is a key regulator of neuroendocrine differentiation [28]. ASCL1 was highly expressed in classic small cell lung cancer (SCLC); additionally, it was a key driver of tumorigenesis in classic SCLC and correlated with the survival and development of lung cancers with neuroendocrine (NE) features [29]. Jiang et al. [30] confirmed that when ASCL1 was successfully overexpressed with SV40 large T-antigen, it could synergistically inhibit retinoblastoma protein and p53 to promote the development of aggressive adenocarcinoma with NE characteristics; however, in the developing mouse lung, knockout of ASCL1 resulted in specific ablation of lung NE cells. Moreover, in recent studies, ASCL1 demonstrated that

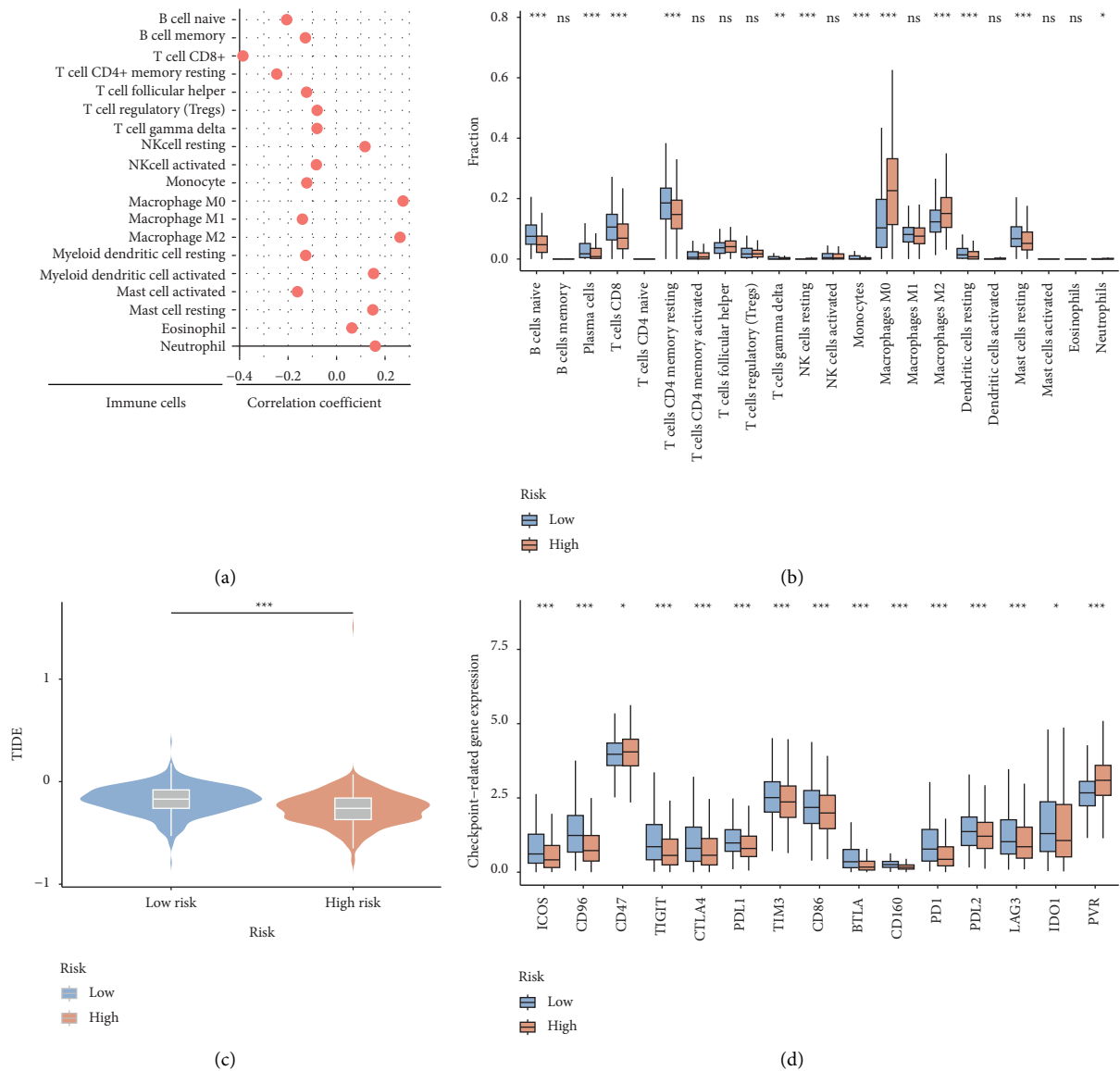


FIGURE 5: The estimation of immune cells infiltration and immunotherapy prediction for the prognostic model in the TCGA training set. Ns, not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. The correlation analysis between the risk score and 19 tumor-infiltrating immune cells evaluated by CIBERSORT algorithm ($p < 0.05$) (a). The differences of 22 tumor-infiltrating immune cells between low- and high-risk groups evaluated by CIBERSORT algorithm (b). Violin plot showing the differential TIDE scores between the high-risk and low-risk groups (c). Box plot showing the expression of immune checkpoint-related markers in low- and high-risk groups (d).

it is involved in the differentiation, cell proliferation, and E-cadherin expression of NE cells in SCLC cell lines by regulating the Wnt signaling pathway [31]. Our work also found that ASCL1 was a risk factor for BC prognosis patients, implying that it might promote breast cancer tumorigenesis. ARID5A (AT-rich interactive domain-containing protein 5a) is one of the Arid family of proteins and possesses the ability to bind nucleic acids, which exist in the nucleus under normal conditions [32]. Meanwhile, ARID5A has been shown to mainly regulate inflammatory and autoimmune disease development by regulating the expression of Interleukin-6 (IL-6) mRNA [33]. Subsequent research revealed that ARID5A could regulate IL-6 mRNA stability through NF- κ B and MAPK signaling pathways [34]. As for

IKZF3 (Aiolos), it belongs to the family of Cys2-His2 zinc finger proteins, which is a lymphocyte lineage transcription factor necessary for the survival of the malignant [35]. For multiple myeloma, immunomodulatory drugs, such as thalidomide could activate E3-ubiquitin ligases and induce degradation of key transcription factors to exert direct antimyeloma effects and promote the survival of myeloma [36]. Then, IKZF3 was a frequently mutated tumor suppressor gene in acute lymphoblastic leukemia (ALL), and its deletion could block lymphocytic lineage differentiation and increase the susceptibility to developing ALL [37]. Consistently, we also found that IKZF3 was a protective factor for BC and facilitated the prognosis of BC patients. Lysine-specific histone demethylase 4B (KDM4B) is a histone

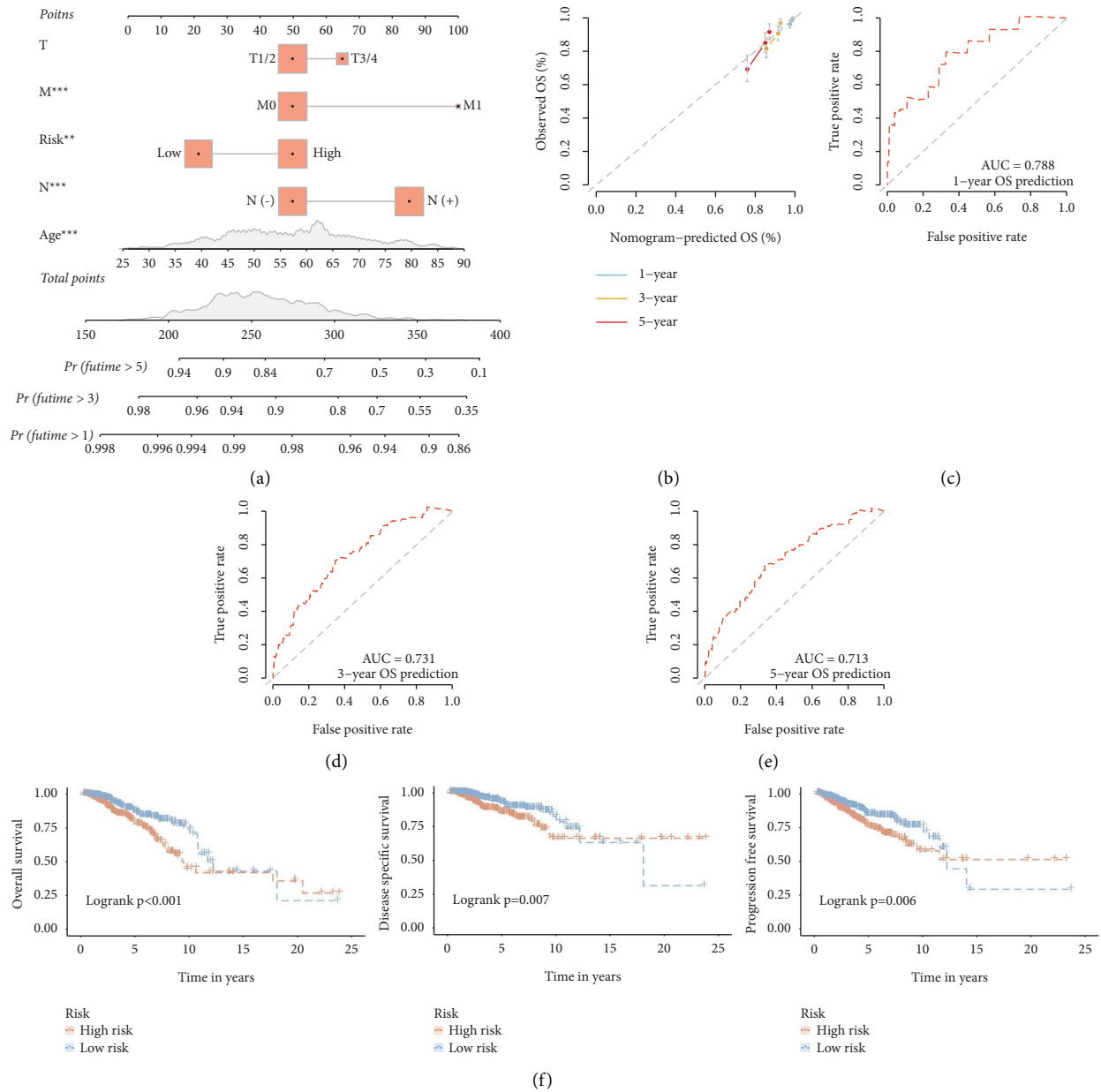


FIGURE 6: Construction and evaluation of a nomogram in the TCGA training set. The nomogram consists of age, T stage, N stage, M stage and risk score to predict the probability of 1-, 3-, and 5-year OS in BC patients (a). Calibration curves of 1-, 3- and 5- year OS in BC patients predicted by the Nomogram (b). ROC curves of 1-, 3-, and 5-year OS predicted by the nomogram (c)–(e). Kaplan–Meier survival curves for OS, DSS and PFS in BC patients based on the nomogram (f).

demethylase for H3K9me3 [38]. According to the genome-wide analysis, KDM4B might be a cancer-specific regulator of alternative splicing by regulating additional alternative splicing-related genes involved in tumorigenesis [39]. In breast cancer, studies have revealed that KDM4B not only antagonizes H3K9 tri-methylation in peripheral heterochromatin and affects H3K4/H3K9 methylation but also plays a role in estrogen receptor α -regulated breast cancer development and mammary epithelial cells proliferation [40, 41]. At the same time, KDM4B was the first identified androgen receptor (AR)-regulated demethylase with effects on AR signaling and turnover and might be a therapeutic

target for prostate cancer [42]. The PR-domain (PRDM) family of genes and the putative transcriptional regulators belong to the SET domain family of histone methyltransferases, which can directly catalyze histone lysine methylation or work by recruiting transcriptional cofactors [43]. Some of the PRDMs are deregulated in cancer and act as tumor suppressors or oncogenes, especially in hematologic malignancies and solid cancers [44]. In diffuse large B-cell lymphomas (DLBCLs), the study showed that the overexpression of PRDM11 (PR-domain family member Prdm11) could induce apoptosis in the E μ -Myc mouse model, then, the DLBCLs patients with low levels of

PRDM11 correlate with shorter overall survival [45]. Additionally, PRDM11 was also identified as a novel locus associated with forced vital capacity, which could be a new target for lung diseases in a genome-wide association analysis [46]. TFF1 (trefoil factor 1), one of the trefoil factor family (TFF), is a small molecule peptide and prevalent in the mucosal environment [47]. In human gastric cancer, it is widely accepted that TFF1 is markedly low-expressed and functions as a gastric tumor suppressor [48]. In BC, although the serum and tissue levels of TFF1 are typically overexpressed [49], many clinical studies have also reported that TFF1 deficiency increases tumorigenicity of human breast cancer cells, and TFF1 expression in BC has an effect on good clinical outcomes for patients [50, 51]. Taken together, all 6 CR-related genes have been reported as taking part in the development of tumors, playing a role in cancer suppression or carcinogenesis. However, the prognostic role in tumors is still less elucidated, and more studies and investigations are needed to understand their prognostic value and mechanisms.

To our knowledge, this is the first time to establish and validate the chromatin regulator-related gene prognostic signature using a large sample size and a high AUC value for breast cancer. Regardless, several limitations can be further improved in the study. In the first place, although our results showed the predictive potentiality and clinical value of our signature, the potential mechanisms of these 6 CR-related genes in BC still require more in-depth experimental investigation. Secondly, the data and information from a total of 1,362 BC patients in the public database used to build the prognostic signature and validate the predictive efficiency of this model are inadequate; therefore, prospective clinical studies are supposed to further confirm our findings.

In conclusion, we identified CR-related DEGs and their predictive ability of prognosis in breast cancer. After that, a novel 6-gene signature model using CR-related DEGs was developed and validated to predict the OS and immunotherapeutic sensitivity for BC patients. Furthermore, a nomogram integrating this novel gene signature and clinical factors was constructed to accurately predict the prognosis for breast cancer, which might provide individualized treatment and aid clinical decision-making for BC patients through prospective validation experiments in the future.

Data Availability

The data used in this study can be obtained from the public databases, namely, The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO).

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

DF, WL, WW, and UDK analyzed the data and wrote the original draft. PG, GH, and XH reviewed and edited the article. WS and HL acquired fund, developed methodology,

administered the project, collected resources, and supervised the study. DF, WL, WS, and HL contributed equally to this study.

Supplementary Materials

Supplementary Table 1: clinical information of BC patients in this study. Supplementary Table 2: chromatin regulator-related differentially expressed genes in breast cancer. (*Supplementary Materials*)

References



- [1] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022," *CA: A Cancer Journal for Clinicians*, vol. 72, no. 1, pp. 7–33, 2022.
- [2] E. Schwich and V. Rebmann, "The inner and outer qualities of extracellular vesicles for translational purposes in breast cancer," *Frontiers in Immunology*, vol. 9, p. 584, 2018.
- [3] A. Braden, R. Stankowski, J. Engel, and A. Onitilo, "Breast cancer biomarkers: risk assessment, diagnosis, prognosis, prediction of treatment efficacy and toxicity, and recurrence," *Current Pharmaceutical Design*, vol. 20, no. 30, pp. 4879–4898, 2014.
- [4] Z. Chen, M. Wang, R. L. De Wilde et al., "A machine learning model to predict the triple negative breast cancer immune subtype," *Frontiers in Immunology*, vol. 12, Article ID 749459, 2021.
- [5] Z. Liu, M. Li, Q. Hua, Y. Li, and G. Wang, "Identification of an eight-lncRNA prognostic model for breast cancer using WGCNA network analysis and a Cox-proportional hazards model based on L1-penalized estimation," *International Journal of Molecular Medicine*, vol. 44, no. 4, pp. 1333–1343, 2019.
- [6] J. Lu, J. Xu, J. Li et al., "FACER: comprehensive molecular and functional characterization of epigenetic chromatin regulators," *Nucleic Acids Research*, vol. 46, no. 19, pp. 10019–10033, 2018.
- [7] A. L. Steunou, M. Cramet, D. Rossetto et al., "Combined action of histone reader modules regulates NuA4 local acetyltransferase function but not its recruitment on the genome," *Molecular and Cellular Biology*, vol. 36, no. 22, pp. 2768–2781, 2016.
- [8] M. A. Morgan and A. Shilatifard, "Chromatin signatures of cancer," *Genes & Development*, vol. 29, no. 3, pp. 238–249, 2015.
- [9] X. S. Shu, L. Li, and Q. Tao, "Chromatin regulators with tumor suppressor properties and their alterations in human cancers," *Epigenomics*, vol. 4, no. 5, pp. 537–549, 2012.
- [10] C. H. Yeh, X. T. Bai, R. Moles et al., "Mutation of epigenetic regulators TET2 and MLL3 in patients with HTLV-I-induced acute adult T-cell leukemia," *Molecular Cancer*, vol. 15, no. 1, p. 15, 2016.
- [11] A. M. Paul, M. R. Pillai, and R. Kumar, "Prognostic significance of dysregulated epigenomic and chromatin modifiers in cervical cancer," *Cells*, vol. 10, p. 2665, 2021.
- [12] K. Zhang, H. Zhao, K. Zhang, C. Hua, X. Qin, and S. Xu, "Chromatin-regulating genes are associated with post-operative prognosis and isocitrate dehydrogenase mutation in astrocytoma," *Annals of Translational Medicine*, vol. 8, no. 23, p. 1594, 2020.

- [13] F. Chen, X. Zhuang, L. Lin et al., "New horizons in tumor microenvironment biology: challenges and opportunities," *BMC Medicine*, vol. 13, no. 1, p. 45, 2015.
- [14] D. F. Quail and J. A. Joyce, "Microenvironmental regulation of tumor progression and metastasis," *Nature Medicine*, vol. 19, no. 11, pp. 1423–1437, 2013.
- [15] S. A. Rosenberg, "Cancer immunotherapy comes of age," *Nature Clinical Practice Oncology*, vol. 2, no. 3, p. 115, 2005.
- [16] Y. Zhang and Z. Zhang, "The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications," *Cellular and Molecular Immunology*, vol. 17, no. 8, pp. 807–821, 2020.
- [17] A. D. Waldman, J. M. Fritz, and M. J. Lenardo, "A guide to cancer immunotherapy: from T cell basic science to clinical practice," *Nature Reviews Immunology*, vol. 20, no. 11, pp. 651–668, 2020.
- [18] E. J. Lipson, P. M. Forde, H. J. Hammers, L. A. Emens, J. M. Taube, and S. L. Topalian, "Antagonists of PD-1 and PD-L1 in cancer treatment," *Seminars in Oncology*, vol. 42, no. 4, pp. 587–600, 2015.
- [19] J. Larkin, V. Chiarion-Sileni, R. Gonzalez et al., "Combined nivolumab and ipilimumab or monotherapy in untreated melanoma," *New England Journal of Medicine*, vol. 373, pp. 23–34, 2015.
- [20] M. A. Postow, R. Sidlow, and M. D. Hellmann, "Immune-related adverse events associated with immune checkpoint blockade," *New England Journal of Medicine*, vol. 378, no. 2, pp. 158–168, 2018.
- [21] S. L. Topalian, J. M. Taube, R. A. Anders, and D. M. Pardoll, "Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy," *Nature Reviews Cancer*, vol. 16, no. 5, pp. 275–287, 2016.
- [22] G. T. Gibney, L. M. Weiner, and M. B. Atkins, "Predictive biomarkers for checkpoint inhibitor-based immunotherapy," *The Lancet Oncology*, vol. 17, no. 12, pp. e542–e551, 2016.
- [23] K. J. Kao, K. M. Chang, H. C. Hsu, and A. T. Huang, "Correlation of microarray-based breast cancer molecular subtypes and clinical outcomes: implications for treatment optimization," *BMC Cancer*, vol. 11, no. 1, p. 143, 2011.
- [24] J. Ferlay, M. Colombet, I. Soerjomataram et al., "Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018," *European journal of cancer*, vol. 103, pp. 356–387, 2018.
- [25] W. Shi, D. Hu, S. Lin, and R. Zhuo, "Five-mRNA signature for the prognosis of breast cancer based on the ceRNA network," *BioMed Research International*, vol. 2020, Article ID 9081852, 17 pages, 2020.
- [26] F. Li, Q. Huang, T. A. Luster et al., "In vivo epigenetic CRISPR screen identifies *Asf1a* as an immunotherapeutic target in *k*-mutant lung adenocarcinoma," *Cancer Discovery*, vol. 10, no. 2, pp. 270–287, 2020.
- [27] W. Sheng, M. W. LaFleur, T. H. Nguyen et al., "LSD1 ablation stimulates anti-tumor immunity and enables checkpoint blockade," *Cell*, vol. 174, no. 3, pp. 549–563.e19, 2018.
- [28] F. F. Vasconcelos and D. S. Castro, "Transcriptional control of vertebrate neurogenesis by the proneural factor *Ascl1*," *Frontiers in Cellular Neuroscience*, vol. 8, p. 412, 2014.
- [29] B. A. Westerman, S. Neijenhuis, A. Poutsma et al., "Quantitative reverse transcription-polymerase chain reaction measurement of HASH1 (ASCL1), a marker for small cell lung carcinomas with neuroendocrine features," *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, vol. 8, no. 4, pp. 1082–1086, 2002.
- [30] T. Jiang, B. J. Collins, N. Jin et al., "Achaete-scute complex homologue 1 regulates tumor-initiating capacity in human small cell lung cancer," *Cancer Research*, vol. 69, no. 3, pp. 845–854, 2009.
- [31] Y. Tenjin, S. Kudoh, S. Kubota et al., "Ascl1-induced Wnt11 regulates neuroendocrine differentiation, cell proliferation, and E-cadherin expression in small-cell lung cancer and Wnt11 regulates small-cell lung cancer biology," *Laboratory Investigation*, vol. 99, no. 11, pp. 1622–1635, 2019.
- [32] J. Iwahara and R. T. Clubb, "Solution structure of the DNA binding domain from Dead ringer, a sequence-specific AT-rich interaction domain (ARID)," *The EMBO Journal*, vol. 18, no. 21, pp. 6084–6094, 1999.
- [33] K. Masuda, B. Ripley, R. Nishimura et al., "Arid5a controls IL-6 mRNA stability, which contributes to elevation of IL-6 level in vivo," *Proceedings of the National Academy of Sciences*, vol. 110, no. 23, pp. 9409–9414, 2013.
- [34] K. K. Nyati, K. Masuda, M. M. U. Zaman et al., "TLR4-induced NF- κ B and MAPK signaling regulate the IL-6 mRNA stabilizing protein Arid5a," *Nucleic Acids Research*, vol. 45, no. 5, pp. 2687–2703, 2017.
- [35] Q. L. Sievers, G. Petzold, R. D. Bunker et al., "Defining the human c2h2 zinc finger degrome targeted by thalidomide analogs through crbn," *Science*, vol. 362, 2018.
- [36] G. Lu, R. E. Middleton, H. Sun et al., "The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins," *Science*, vol. 343, pp. 305–309, 2014.
- [37] S. Winandy, P. Wu, and K. Georgopoulos, "A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma," *Cell*, vol. 83, no. 2, pp. 289–299, 1995.
- [38] B. D. Fodor, S. Kubicek, M. Yonezawa et al., "Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells," *Genes & Development*, vol. 20, no. 12, pp. 1557–1562, 2006.
- [39] L. Duan, Z. Chen, J. Lu et al., "Histone lysine demethylase KDM4B regulates the alternative splicing of the androgen receptor in response to androgen deprivation," *Nucleic Acids Research*, vol. 47, no. 22, pp. 11623–11636, 2019.
- [40] L. Shi, L. Sun, Q. Li et al., "Histone demethylase JMJD2B coordinates H3K4/H3K9 methylation and promotes hormonally responsive breast carcinogenesis," *Proceedings of the National Academy of Sciences*, vol. 108, no. 18, pp. 7541–7546, 2011.
- [41] L. Gaughan, J. Stockley, K. Coffey et al., "KDM4B is a master regulator of the estrogen receptor signalling cascade," *Nucleic Acids Research*, vol. 41, no. 14, pp. 6892–6904, 2013.
- [42] Y. W. Chin and S. Y. Han, "KDM4 histone demethylase inhibitors for anti-cancer agents: a patent review," *Expert Opinion on Therapeutic Patents*, vol. 25, no. 2, pp. 135–144, 2015.
- [43] F. Di Tullio, M. Schwarz, H. Zorgati, S. Mzoughi, and E. Guccione, "The duality of PRDM proteins: epigenetic and structural perspectives," *FEBS Journal*, vol. 289, no. 5, pp. 1256–1275, 2022.
- [44] C. K. Fog, G. G. Galli, and A. H. Lund, "PRDM proteins: important players in differentiation and disease," *BioEssays*, vol. 34, no. 1, pp. 50–60, 2012.
- [45] C. K. Fog, F. Asmar, C. Côme et al., "Loss of PRDM11 promotes MYC-driven lymphomagenesis," *Blood*, vol. 125, no. 8, pp. 1272–1281, 2015.
- [46] D. W. Loth, M. S. Artigas, S. A. Gharib et al., "Genome-wide association analysis identifies six new loci associated with

- forced vital capacity,” *Nature Genetics*, vol. 46, no. 7, pp. 669–677, 2014.
- [47] L. Thim, “A new family of growth factor-like peptides. ”Trefoil’ disulphide loop structures as a common feature in breast cancer associated peptide (pS2), pancreatic spasmolytic polypeptide (PSP), and frog skin peptides (spasmolysins),” *FEBS Letters*, vol. 250, no. 1, pp. 85–90, 1989.
- [48] M. Soutto, Z. Chen, A. M. Katsha et al., “Trefoil factor 1 expression suppresses Helicobacter pylori-induced inflammation in gastric carcinogenesis,” *Cancer*, vol. 121, no. 24, pp. 4348–4358, 2015.
- [49] M. H. Elnagdy, O. Farouk, A. K. Seleem, and H. A. Nada, “TFF1 and TFF3 mRNAs are higher in blood from breast cancer patients with metastatic disease than those without,” *Journal of oncology*, vol. 2018, Article ID 4793498, pp. 1–8, 2018.
- [50] E. Buache, N. Etique, F. Alpy et al., “Deficiency in trefoil factor 1 (TFF1) increases tumorigenicity of human breast cancer cells and mammary tumor development in TFF1-knockout mice,” *Oncogene*, vol. 30, no. 29, pp. 3261–3273, 2011.
- [51] L. K. Dunnwald, M. A. Rossing, and C. I. Li, “Hormone receptor status, tumor characteristics, and prognosis: a prospective cohort of breast cancer patients,” *Breast Cancer Research*, vol. 9, no. 1, p. R6, 2007.

Research Article

Cancer-Associated Fibroblasts Affect Tumor Metabolism and Immune Microenvironment in Gastric Cancer and Identification of Its Characteristic Genes

Chanchan Gao ¹, Fei Liu,² Qiuju Ye,² and Aiping Guo ²

¹Department of Oncology, Zhongda Hospital, School of Medicine, Southeast University, Nanjing 210009, China

²Department of Medical Oncology, Luhe People's Hospital of Nanjing, Nanjing 21009, China

Correspondence should be addressed to Aiping Guo; 929813607@qq.com

Received 11 October 2022; Revised 30 October 2022; Accepted 25 November 2022; Published 30 January 2023

Academic Editor: Feng Jiang

Copyright © 2023 Chanchan Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Cancer-associated fibroblasts (CAFs) have reported widely involved in cancer progression. However, its underlying mechanism in gastric cancer is still not clarified. **Methods.** The data used in this study were all downloaded from the Cancer Genome Atlas database. R software and the R packages were used for all the analyses. **Results.** In our study, we first quantified the CAFs infiltration using the ssGSEA algorithm. The clinical correlation result showed that CAFs were associated with a worse prognosis and clinical features. Pathway enrichment also indicated several oncogenic pathways in GC patients with high CAFs infiltration, including epithelial-mesenchymal transition (EMT), myogenesis, allograft rejection, the inflammatory response, and IL2/STAT5 signaling. Furthermore, FNDC1 and RSPO3 were identified as the characteristic genes of CAFs through two machine learning algorithms, LASSO logistic regression and SVM-RFE. The following analysis showed that FNDC1 and RSPO3 were associated with more progressive clinical features and had a good prediction efficiency of the CAFs infiltration status in GC patients. Pathway enrichment and genomic instability were performed to explore the underlying mechanisms of FNDC1 and RSPO3. Immune infiltration analysis showed that CAFs were positively correlated with M2 macrophages. Moreover, we found that the GC patients with low CAFs infiltration were more sensitive to immunotherapy. Also, the CAFs, FNDC1, and RSPO3 could generate a certain effect on the sensitivity of doxorubicin, mitomycin, and paclitaxel. **Conclusions.** In summary, our study comprehensively investigated the role of CAFs in GC, which might be associated with immunotherapy sensitivity. Meanwhile, FNDC1 and RSPO3 were identified as the underlying targets of GC.

1. Introduction

Gastric cancer (GC) is the fifth most common cancer around the world, with over one million new cases diagnosed annually [1]. There has been a noticeable increase in the incidence of GC worldwide, along with its high mortality and metastasis rate [1]. At present, surgery is still the first-line therapy option for early-staged GC and can lead to persistent prognosis benefits [2]. Meanwhile, combined therapies, including chemotherapy and targeted therapy, have also prolonged the overall survival (OS) of advanced GC patients [3]. Despite this, however, the five years survival rate of advanced GC patients is still less than 20% [3]. Therefore,

early diagnosis and precise therapy of GC patients remain the focus of research.

Tumor cells are continuously affected by the tumor microenvironment (TME) they exist in, the components of which mainly consist of immune and stromal cells [4]. Cancer-associated fibroblasts (CAFs) are one of the most prominent cell types in TME that can influence tumor progression in multiple manners [5]. CAFs can secrete specific biological factors such as EGF, TGF- β , and IL6 to facilitate tumor malignant phenotype, including tumor neovascularization and immune escape, leading to tumor deterioration [6]. Meanwhile, CAFs can regulate tumor metabolism. CAFs can enhance glycolysis and excrete plenty

of lactic acid and hydrogen ions, forming an acidic microenvironment to inhibit the activity of immune cells. Also, the metabolites of lactic acid and pyruvate produced by CAFs can be used as nutrients for tumor cells to stimulate their growth [7]. Recently, increasing attention has been paid to the role of CAFs in cancers for its diverse biological functions. For instance, Liubomirski et al. found that in breast cancer, the interactions between cancer cells and CAFs can significantly enhance the prometastatic phenotypes of the TME, further resulting in the higher angiogenesis, migratory, and invasive potential of cancer cells [8]. In esophageal squamous cell carcinoma, Jolly et al. revealed that CAFs can secrete IL-6 and exosomal miR-21 to induce the generation of monocytic myeloid-derived suppressor cells, which not only suppressed immune function but also enhanced drug resistance [9]. However, few studies have focused on the role of CAFs in GC, and therefore, it is meaningful to explore the underlying effect of CAFs to guide the treatment of GC.

Advancements in bioinformatic analysis provide a great convenience for researchers in investigating the underlying biological mechanisms of diseases [10]. In our study, we quantified the CAFs infiltration using the ssGSEA algorithm and comprehensively explored its role in GC. CFNDC1 and RSPO3 were identified as the characteristic genes of CAFs through two machine learning algorithms, LASSO logistic regression and SVM-RFE. Further following analysis showed that FNDC1 and RSPO3 were associated with more progressive clinical features and had a good prediction efficiency of the CAFs infiltration status in GC patients. Pathway enrichment and genomic instability were performed to explore the underlying mechanisms of FNDC1 and RSPO3. Immune infiltration analysis showed that CAFs were positively correlated with M2 macrophages. Moreover, we found that the GC patients with low CAFs infiltration were more sensitive to immunotherapy. Also, the CAFs, FNDC1, and RSPO3 could generate a certain effect on the sensitivity of doxorubicin, mitomycin, and paclitaxel.

2. Methods

2.1. Available Data Acquisition. The public transcription profiles and clinical information of GC patients were downloaded from The Cancer Genome Atlas database-TCGA-STAD project. The expression profile was in TPM form and was annotated based on the Homo sapiens.GRCh38.107.gtf file. Clinical information was in a “bcr-xml” file and extracted using the Perl code. Differentially expressed genes (DEGs) analysis was performed using the limma package with the threshold of $|\log_{2}FC| > 1$ and $adj.P < 0.05$. The basic information of enrolled patients is shown in Table 1.

2.2. Single Sample Gene Set Enrichment Analysis. Single sample gene set enrichment analysis (ssGSEA) was used to quantify the relative enrichment score of CAFs [11]. The genes used for quantification were ACTA2, FAP, PDGFRB, CAV1, PDPN, PDGFRA, ZEB1, FOXF1, SPARC, MMP2,

TABLE 1: Basic information of enrolled patients.

| Features | Numbers (<i>n</i>) | Percentage (%) | |
|----------|----------------------|----------------|------|
| Age | ≤65 | 197 | 44.5 |
| | >65 | 241 | 54.4 |
| | Unknown | 5 | 1.1 |
| Gender | Female | 158 | 35.7 |
| | Male | 285 | 64.3 |
| Grade | G1 | 12 | 2.7 |
| | G2 | 159 | 35.9 |
| | G3 | 263 | 59.4 |
| | Unknown | 9 | 2.0 |
| Stage | Stage I | 59 | 13.3 |
| | Stage II | 130 | 29.3 |
| | Stage III | 183 | 41.3 |
| | Stage IV | 44 | 9.9 |
| | Unknown | 27 | 6.1 |
| T stage | T1 | 23 | 5.2 |
| | T2 | 93 | 20.9 |
| | T3 | 198 | 44.7 |
| | T4 | 119 | 26.9 |
| | Unknown | 10 | 2.3 |
| M stage | M0 | 391 | 88.3 |
| | M1 | 30 | 6.8 |
| | Unknown | 22 | 4.9 |
| N stage | N0 | 132 | 29.8 |
| | N1 | 119 | 26.9 |
| | N2 | 85 | 19.2 |
| | N3 | 88 | 19.9 |
| | Unknown | 19 | 4.3 |

and FN1 from the CellMarker website (<https://bio-bigdata.hrbmu.edu.cn/CellMarker/>). The metabolism and immune-related pathways were also quantified using ssGSEA analysis.

2.3. Pathway Enrichment Analysis. Pathway enrichment analysis was performed using the gene set enrichment analysis (GSEA) algorithm, and the analyzed gene set was the Hallmark signature. The terms with $|\text{normalized enrichment score (NES)}| > 1$ and $adj.P < 0.05$ were considered statistically significant.

2.4. Characteristic Gene Identification. Two machine learning algorithms, LASSO logistic regression and support vector machine recursive feature elimination (SVM-RFE), were utilized to identify the characteristic genes of specific features [12]. Receiver operating characteristic (ROC) curves were used to evaluate the prediction efficiency of characteristic genes. Principal component analysis (PCA) was performed using the ade4 package in R environments.

2.5. Immune Infiltration and Genomic Analyses. The quantification of the immune microenvironment of GC was conducted using the CIBERSORT algorithm, and 22 types of infiltrating immune cells were extracted [13]. The scores of TMB and MSI were downloaded from the TCGA database. The tumor stemness index mRNasi and EREG-mRNasi were calculated according to the one-class logistic regression

(OCLR) machine learning algorithm of the previous study [14].

2.6. Immunotherapy and Drug Sensitivity Analyses. Tumor Immune Dysfunction and Exclusion (TIDE) analysis (<https://tide.dfci.harvard.edu/>) and submap algorithm were utilized to evaluate the immunotherapy response rate of GC patients. Drug sensitivity analysis was conducted based on the data from the Genomics of Drug Sensitivity in Cancer (GDSC) database (<https://www.cancerrxgene.org>).

2.7. Statistical Analysis. R software was responsible for all the analysis. Here, the comparison with P value less than 0.05 was considered statistically significant. The ggplot2 package was utilized for most plots [15]. The correlation of continuous variables was compared using the Spearman method. The comparison of variables with a normal distribution was performed using the Student's T -test. Kaplan–Meier (KM) survival curves were used to evaluate the prognosis effect of specific index.

3. Results

3.1. Quantification of CAFs in TCGA Data. The flowchart of whole study is shown in Figure S1. First, based on the marker genes mentioned above, the relative infiltration of CAFs in GC tissue was quantified using the ssGSEA algorithm (Figure 1(a)). KM survival curve showed that the patients with higher CAFs infiltration might have a worse overall survival (OS) (Figure 1(b), HR = 1.41, $P = 0.041$). Furthermore, we explored the CAFs differences in patients with different clinical features. The result showed that CAFs might be associated with a more progressive grade and T stage (Figures 1(c) and 1(d)). However, no significant difference was observed in M and N stages (Figures 1(e) and 1(f)). Pathway enrichment analysis showed that in the patients with higher CAFs infiltration, the pathway of epithelial-mesenchymal transition (EMT), myogenesis, allograft rejection, inflammatory response, and IL2/STAT5 signaling were remarkably enriched in (Figure 1(g)).

3.2. Identification of the Characteristic Genes of CAFs. Then, we performed the DEGs analysis with the threshold of $|\log_{2}FC| > 1$ and $\text{adj.}P < 0.05$. A total of 268 downregulated and 1697 upregulated DEGs were identified (Figure 1(h)). LASSO logistic regression and the SVM-RFE algorithm were used to identify the characteristic genes of CAFs (Figures 2(a)–2(c)). LASSO logistic regression identified four genes, including FNDC1, SGCD, FGF7, and RSPO3. Further, among these four genes, the SVM-RFE algorithm screened two genes FNDC1 and RSPO3, as the characteristic genes of CAFs (Figure 2(d)). ROC curves showed that FNDC1 and RSPO3 had great prediction in the CAFs infiltration status of GC patients (Figures 2(e) and 2(f), FNDC1, AUC = 0.890; RSPO3, AUC = 0.885). Then, logistic regression was performed based on the FNDC1 and RSPO3. The formula was “score = $-6.691 + 0.9797 *$

FNDC1 + 1.2415 * RSPO3.” The ROC curve showed that the logistic score had an excellent prediction ability of the CAFs infiltration of GC patients (Figure 2(g)). PCA analysis indicated that the genes FNDC1 and RSPO3 could effectively distinguish the GC patients with high and low CAFs infiltration (Figure 2(h)).

3.3. Prognosis Effect and Clinical Correlation of FNDC1 and RSPO3. KM survival curves showed that the patients with high FNDC1 and RSPO3 expression might have a worse OS, DSS and PFI (Figures 3(a)–3(f)). Also, we found that the patients with higher CAFs infiltration might have a higher FNDC1 and RSPO3 expression (Figure 3(g)). Meanwhile, the young patients (≤ 65 years old) tend to have a higher RSPO3 expression (Figure 3(h)); the G3 GC patients might have a higher FNDC1 and RSPO3 expression than G1-2 patients (Figure 3(i)); the stage III-IV patients might have a higher RSPO3 expression (Figure 3(j)); the T3-4 GC patients might have a higher FNDC1 and RSPO3 expression than T1-2 patients (Figure 3(k)); the N1-3 GC patients might have a higher RSPO3 expression than N0 patients (Figure 3(l)).

3.4. Biological Explorations of FNDC1 and RSPO3. CAFs have been reported to affect tumor metabolism. Pathway correlation analysis indicated that CAFs was negatively correlated with TRN- α metabolism, KREBS cycle metabolism, amino acid metabolism, vitamin metabolism, abnormal metabolism, and vitamin metabolism, yet positively correlated with folate metabolism (Figure 4(a)). We next explored the underlying pathways of FNDC1 and RSPO3. Pathway enrichment analysis of RSPO3 showed that the pathway of the apical junction, inflammatory response, KRAS signaling, and EMT were significantly enriched in the patients with high RSPO3 expression (Figure 4(b)). For FNDC1, the pathway of NOTCH signaling, angiogenesis, hedgehog signaling, TGF- β signaling, and IL6/JAK/STAT3 signaling were significantly enriched in (Figure 4(c)). Pan-cancer analysis revealed the expression patterns of FNDC1 and RSPO3 in solid cancers. The result showed that FNDC1 was upregulated, while RSPO3 was downregulated in GC tissue (Figures 5(a) and 5(b)). Genomic instability analysis showed that FNDC1 had no significant effect on TMB, MSI, and tumor stemness index (Figures 5(c)–5(e)). However, RSPO3 might be associated with a lower TMB, MSI and stemness index (Figures 5(f)–5(h)). Immune analysis showed that FNDC1 was positively correlated with NK cells, macrophages, and iDC, while negatively correlated with Th17 cells (Figure S2A); RSPO3 was positively correlated with NK cells, mast cells, and pDC yet negatively correlated with Th17 cells and Th2 cells (Figure S2B).

3.5. CAFs Is Positively Correlated with M2 Macrophages. The crosstalk between different cells can significantly affect the TME of GC. The CIBERSORT algorithm was used for immune cell infiltration. The correlation of CAFs and the quantified immune cells are shown in Figure 6(a). The result showed that CAFs was positively correlated with naive

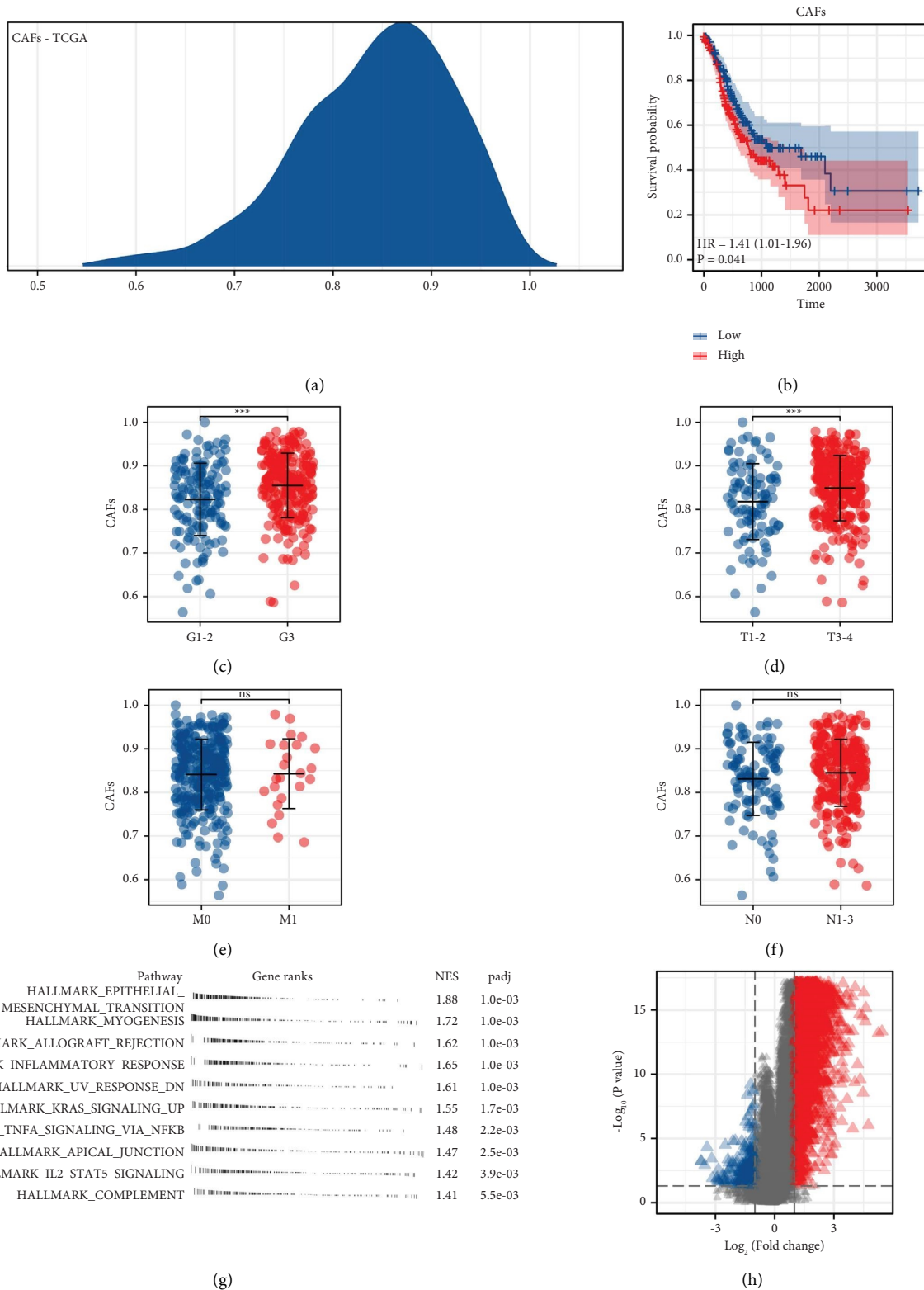


FIGURE 1: Exploration of CAFs in GC. (a) ssGSEA was performed to quantify the relative content of CAFs in TCGA database. (b) KM survival curves showed that CAFs were associated with a worse prognosis. (c–f) The difference of CAFs infiltration in patients with different clinical features. (g) Pathway enrichment analysis of CAFs. (h) DEGs analysis between high and low CAFs infiltration with the threshold of $|\log_2(\text{FC})| > 1$ and $\text{adj.}P < 0.05$.

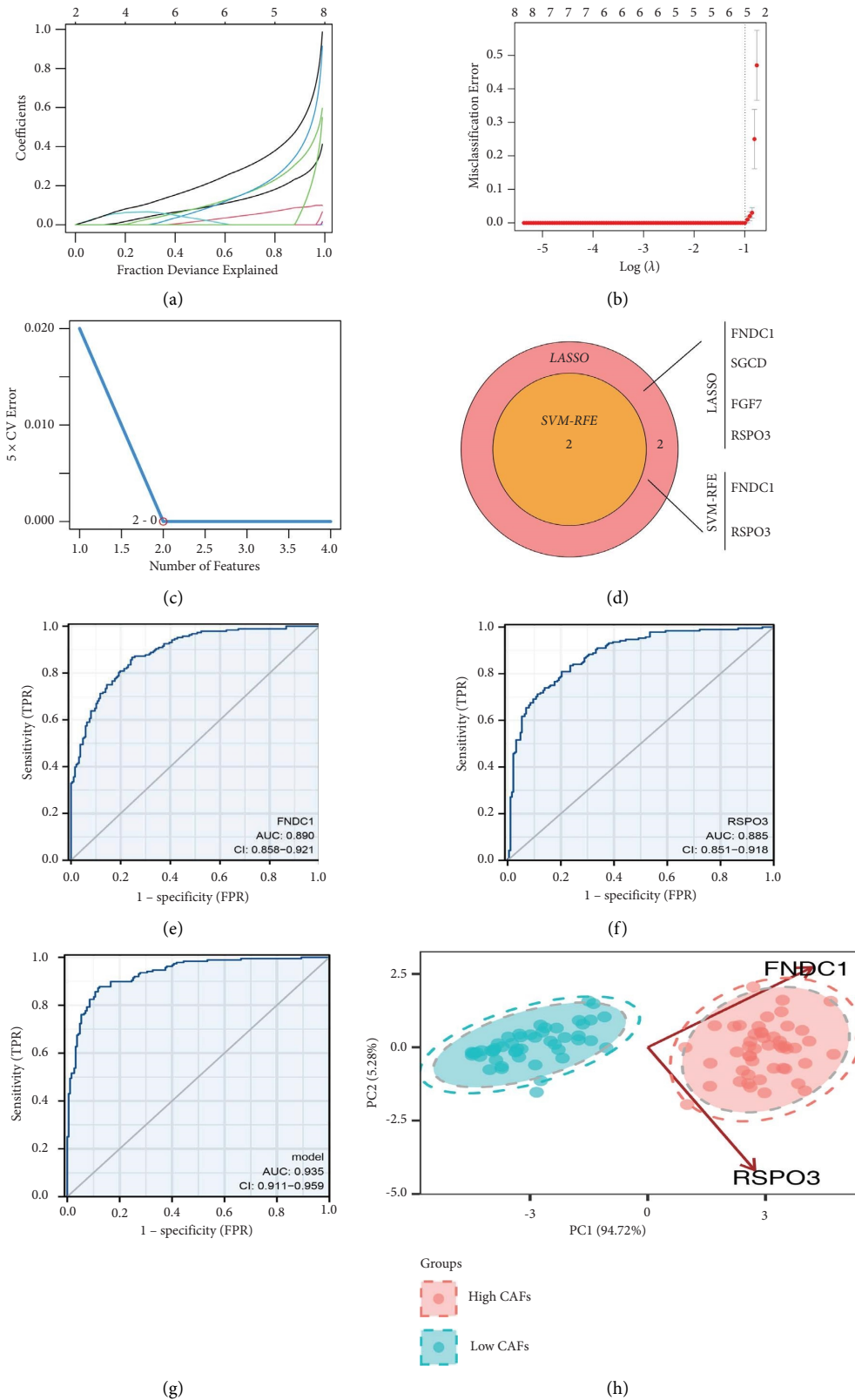


FIGURE 2: Identification of the characteristic genes of CAFs. (a, b) LASSO logistic regression was used to identify the characteristic genes of CAFs. (c) SVM-RFE was used to identify the characteristic genes of CAFs. (d) FNDC1 and RSPO3 were identified as the characteristic genes of CAFs. (e, f) ROC curves to evaluate the prediction efficiency of FNDC1 and RSPO3 on CAFs infiltration status. (g) A logistic regression was performed based on the FNDC1 and RSPO3. The formula was “score = -95.2708 + 15.9714 * FNDC1 + 13.5927 * RSPO3.” (h) PCA analysis of FNDC1 and RSPO3 in different CAFs infiltration patients.

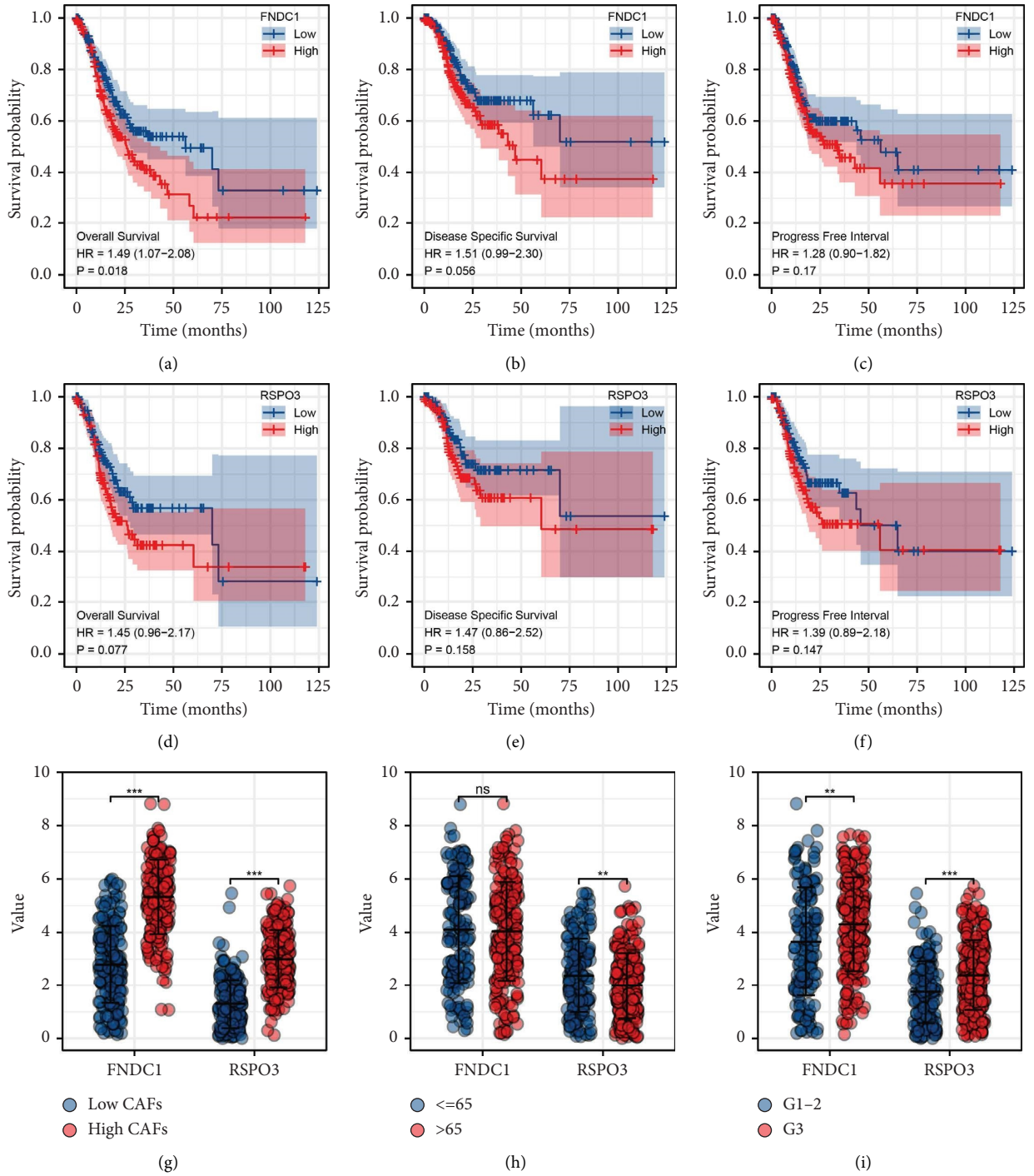


FIGURE 3: Continued.

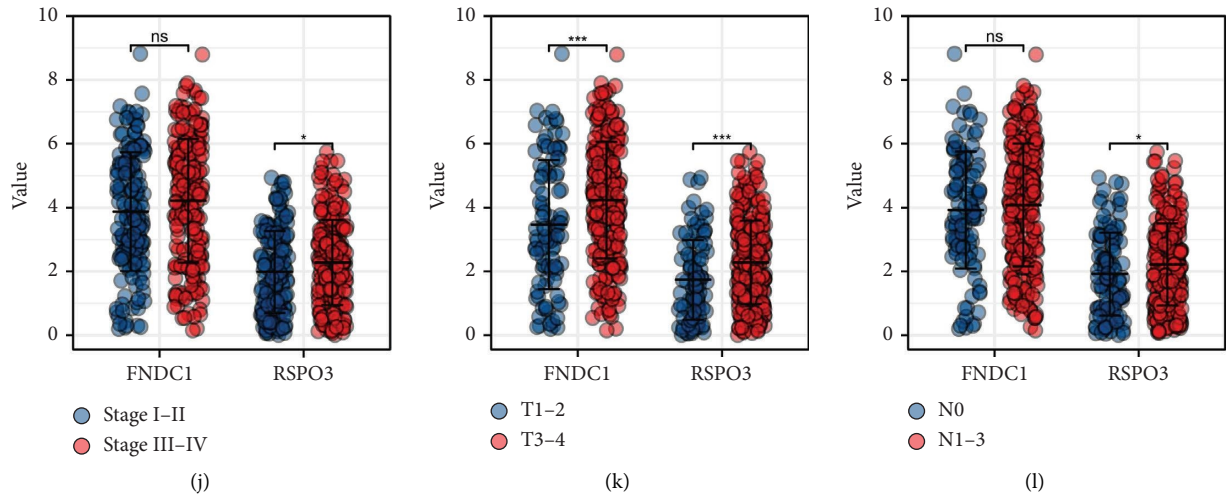


FIGURE 3: Prognosis effect and clinical correlation of FNDC1 and RSPO3. (a–c) The OS, DSS, and PFI of FNDC1 in TCGA database. (d–f) The OS, DSS, and PFI of RSPO3 in the TCGA database. (g–l) The expression level of FNDC1 and RSPO3 in different GC patients.

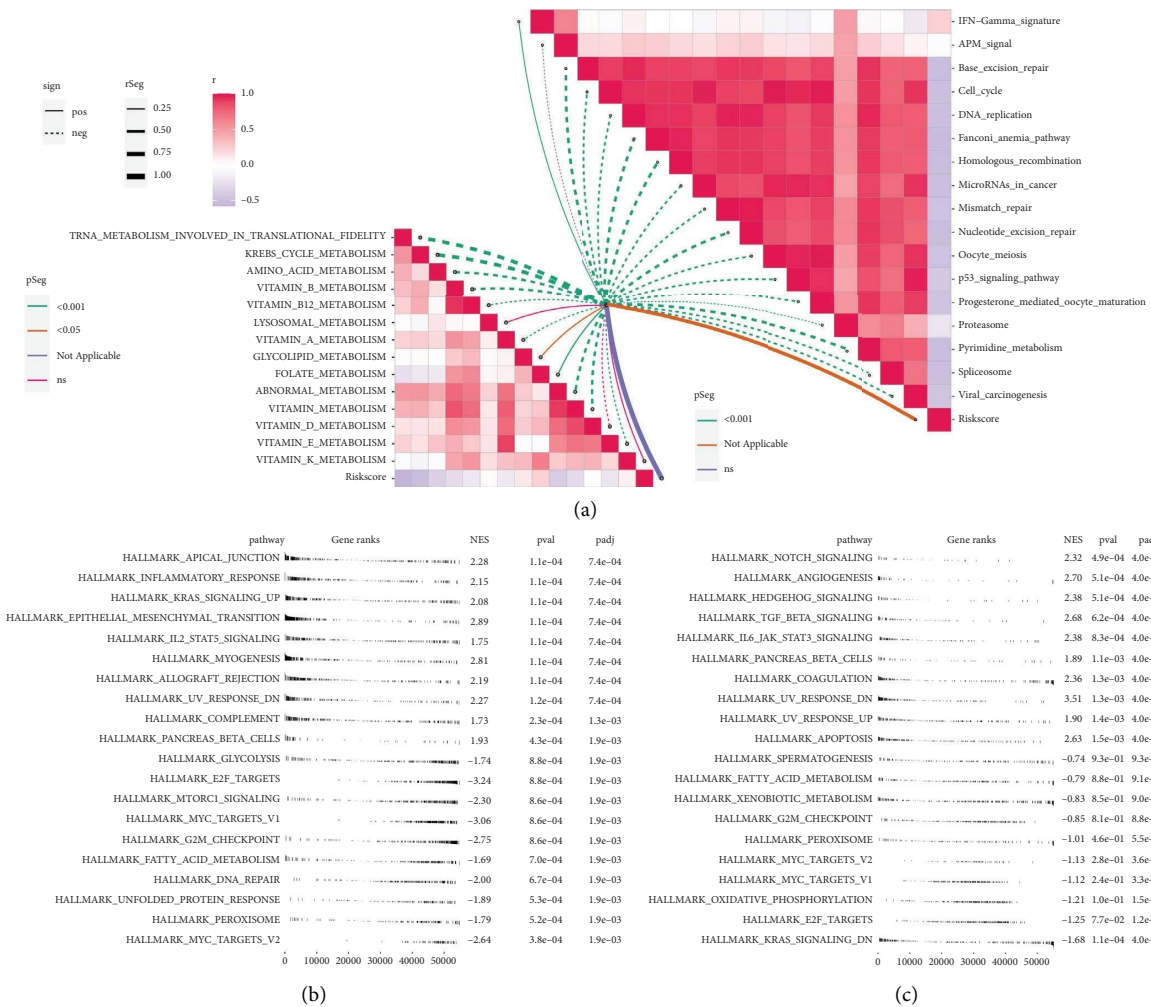
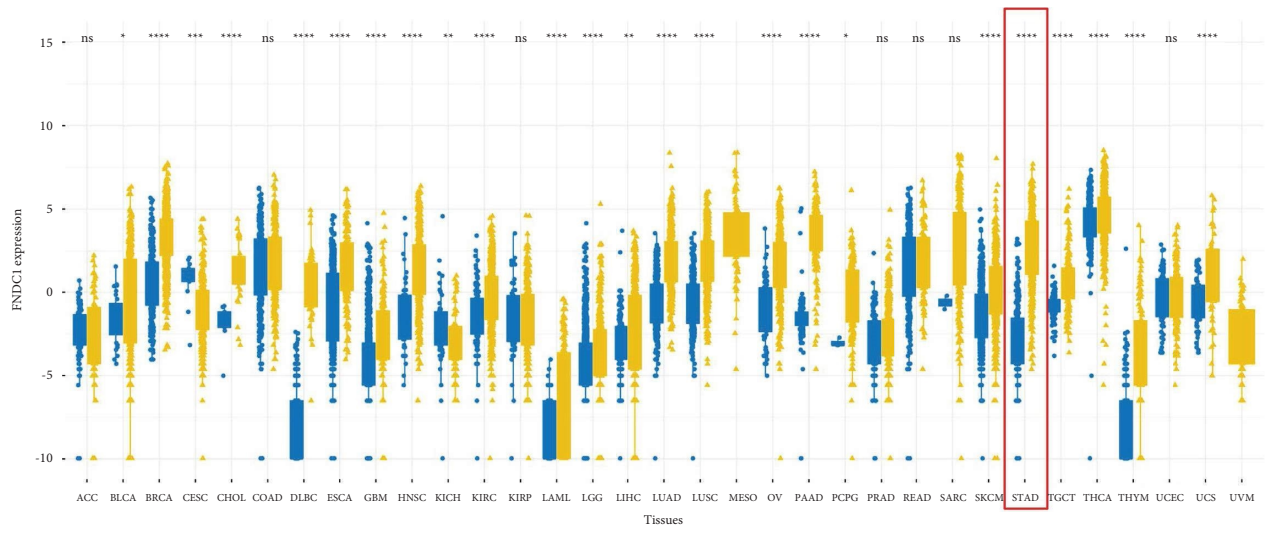
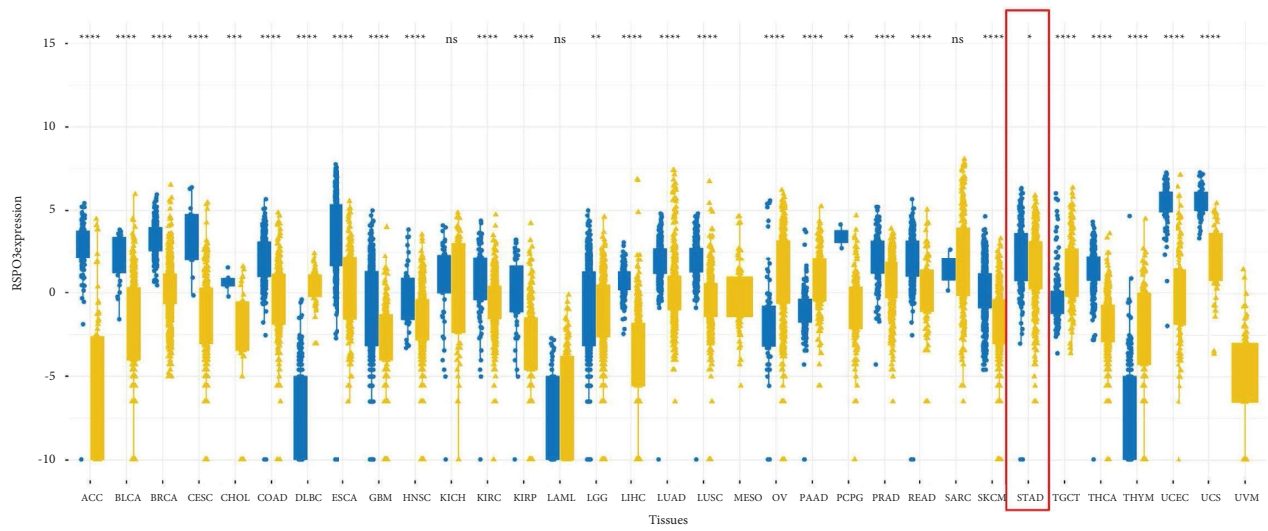


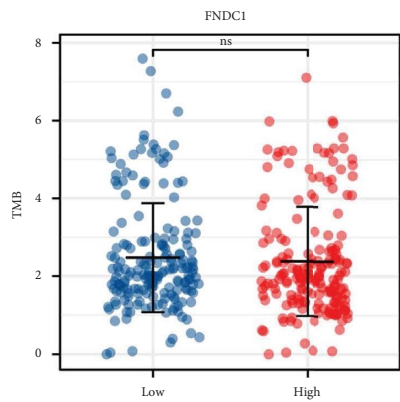
FIGURE 4: Pathway enrichment analyses of FNDC1 and RSPO3. (a) Pathway enrichment analysis of FNDC1 using the GSEA analysis. (b) Pathway enrichment analysis of RSPO3 using the GSEA analysis. (c) Pathway enrichment analysis of FNDC1 using the GSEA analysis.



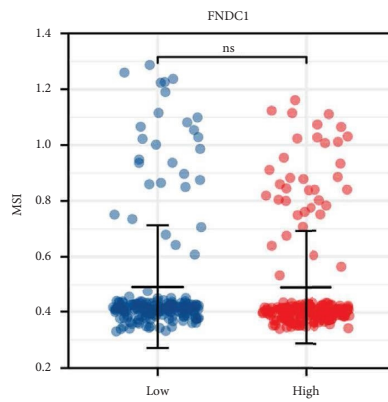
(a)



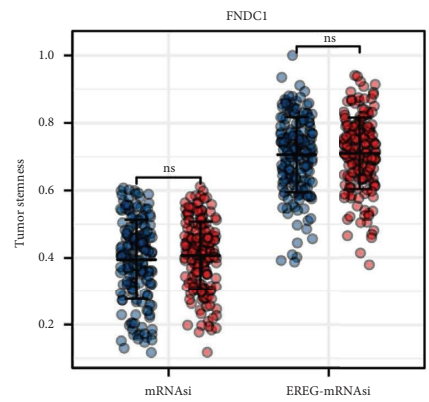
(b)



(c)



(d)



(e)

FIGURE 5: Continued.

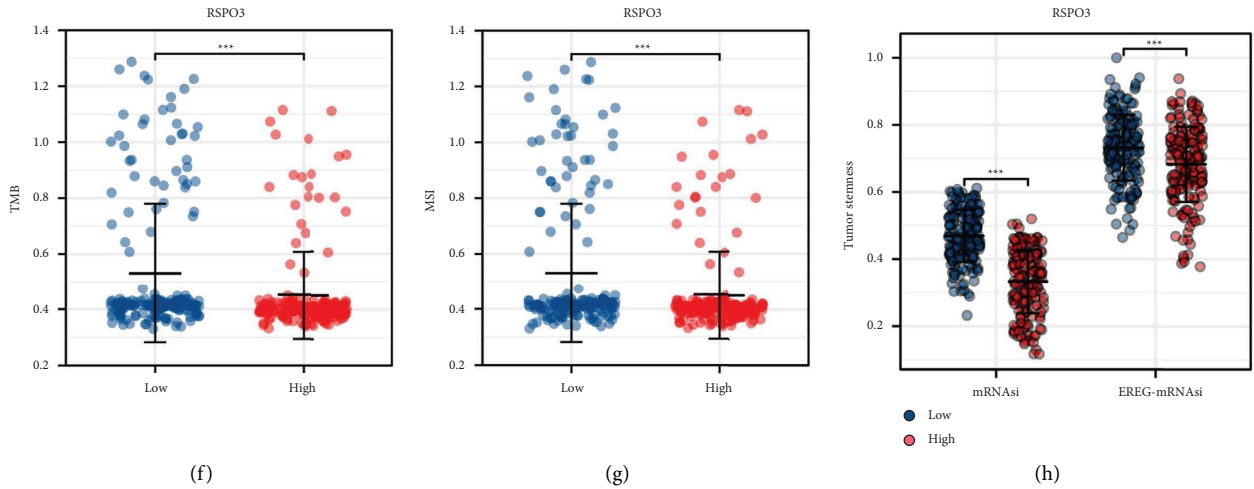


FIGURE 5: Further exploration of FNDC1 and RSPOS. (a, b) Pan-cancer analysis illustrates the expression pattern of FNDC1 and RSPOS. (c-e) The correlation between FNDC1 and TMB, MSI, and tumor stemness index. (f-h) The correlation between RSPO3 and TMB, MSI, and tumor stemness index.

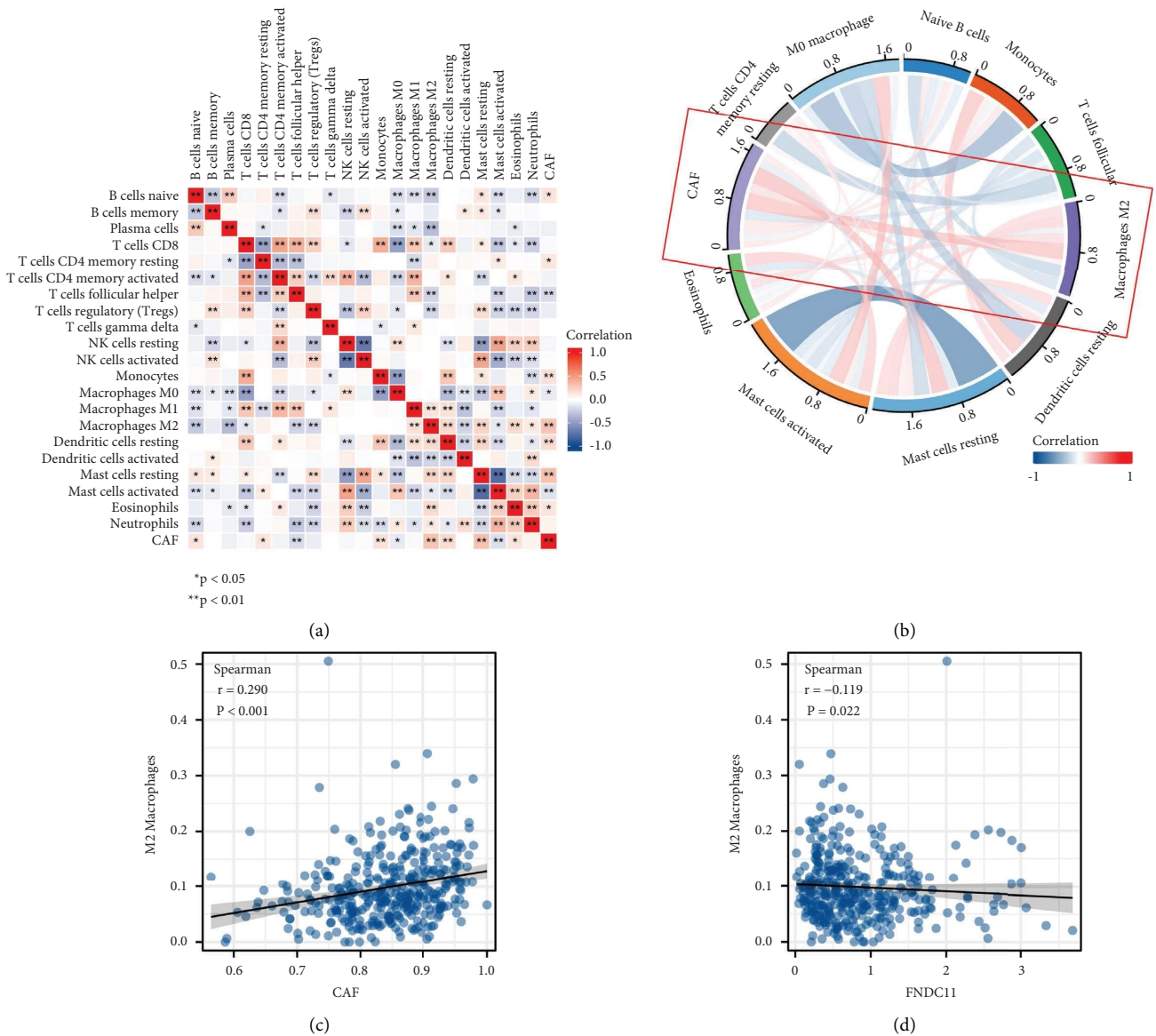


FIGURE 6: Continued.

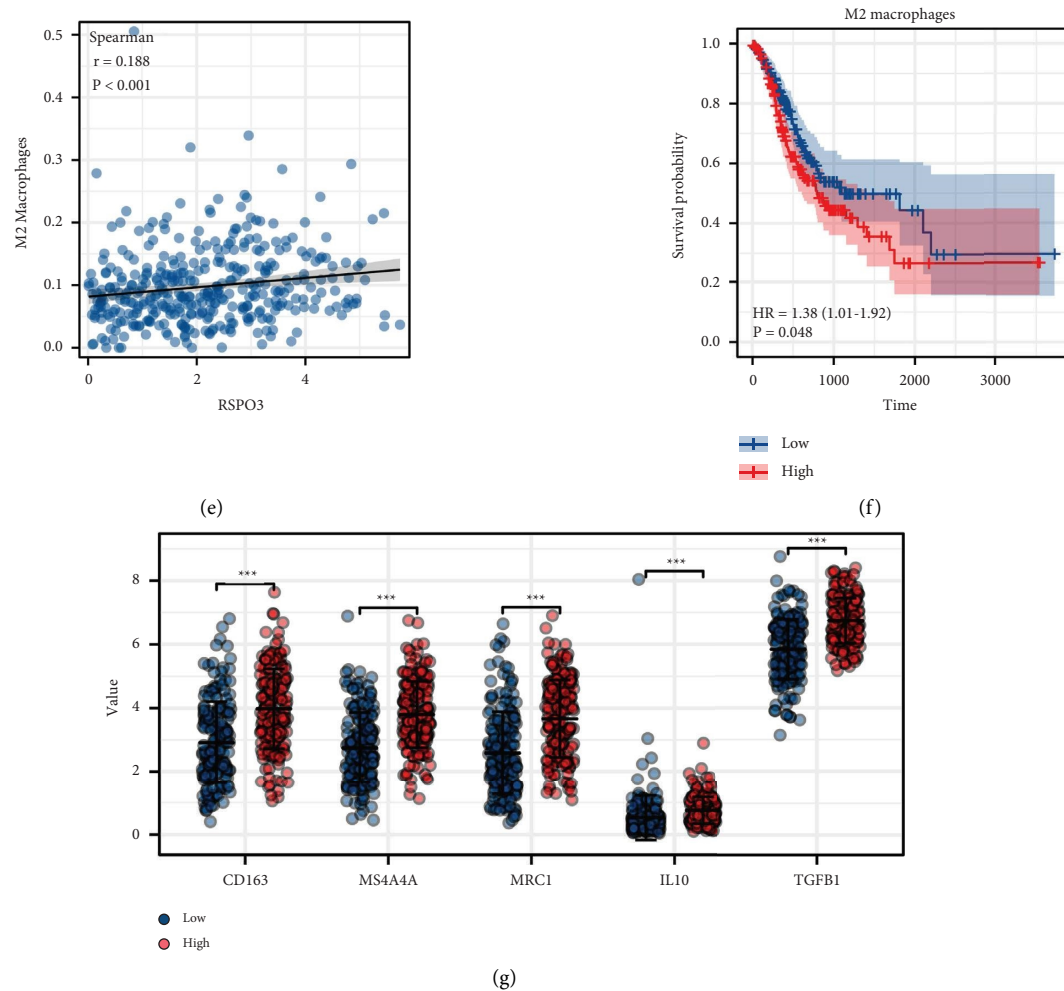


FIGURE 6: Immune infiltration analysis. (a) The correlation of CAFs and the quantified immune cells. (b) CAFs were positively correlated with M2 macrophages. (c) The correlation between M2 macrophages and CAFs, FNDC1, and RSPO3. (d) KM survival curves of the M2 macrophages. (e) The characteristic makers and factors were all highly expressed in the samples with high CAFs infiltration. (f) The KM survival curve showed that M2 macrophages might be associated with a poor prognosis. (g) The characteristic makers and factors were all highly expressed in the samples with high CAFs infiltration.

B cells, resting CD4⁺ memory T cells, monocytes, M2 macrophages, resting dendritic cells, resting mast cells, and eosinophils, yet negatively correlated with follicular helper T cells, M0 macrophages, and activated mast cells (Figures 6(b) and 6(c)). Moreover, we found that FNDC1 was negatively, while RSPO3 was positively correlated with M2 macrophages (Figures 6(d) and 6(e)). Also, the KM survival curve showed that M2 macrophages might be associated with a poor prognosis (Figure 6(f)). Meanwhile, the characteristic makers and factors were all highly expressed in the samples with high CAFs infiltration (Figure 6(g)).

3.6. CAFs and Its Characteristic Genes Were Associated with the Sensitivity of Immunotherapy and Chemotherapy. Immunotherapy is a novel therapeutic option for advanced GC. Thus, we explored the underlying difference in immunotherapy sensibility between high and low CAFs infiltration patients. Immune checkpoint correlation analysis

showed that CTLA4, HAVCR2, PDCD1LG2, PDCD1, and TIGIT were differentially expressed in high and low CAFs infiltration patients (Figure 7(a)). The TIDE analysis was then performed, in which the patients with TIDE a score >0 were defined as nonresponders and <0 were defined as responders. The result showed in low CAFs infiltration patients, the proportion of immunotherapy responders was 53.2%. However, in high CAFs infiltration patients, the proportion of immunotherapy responders was only 20.9%, indicating that low CAFs infiltration GC patients might be more sensitive to immunotherapy (Figure 7(b)). Submap analysis indicated that the patients with low CAFs infiltration might be more sensitive to both PD-1 and CTLA4 therapies (Figure S3). Considering the significant correlation between CAFs and M2 macrophages, we further explored the effect of M2 macrophages on immunotherapy. Results showed a positive correlation between the TIDE score and M2 macrophages (Figure S4A). Moreover, we found that the patients with high M2 macrophages infiltration tend to have

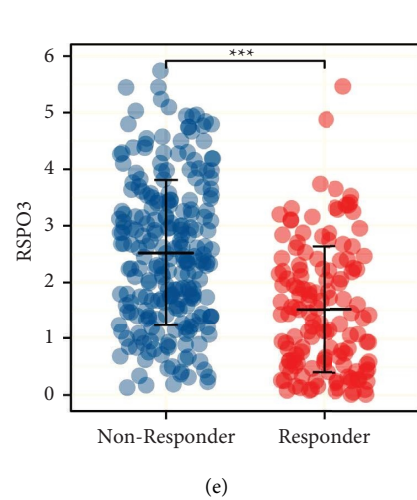
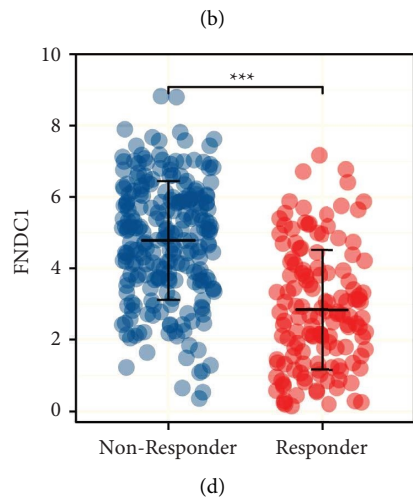
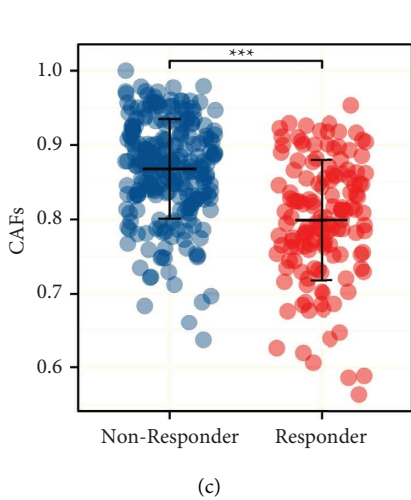
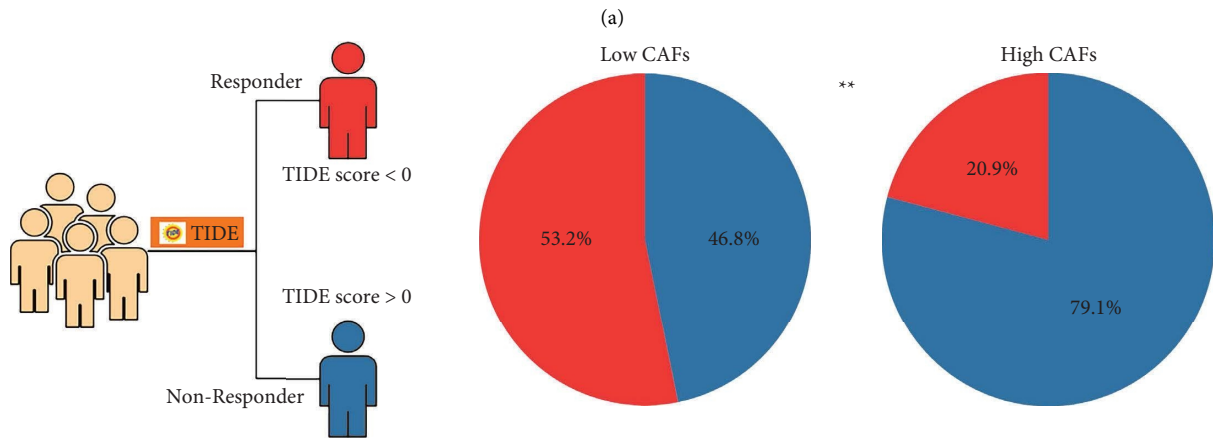
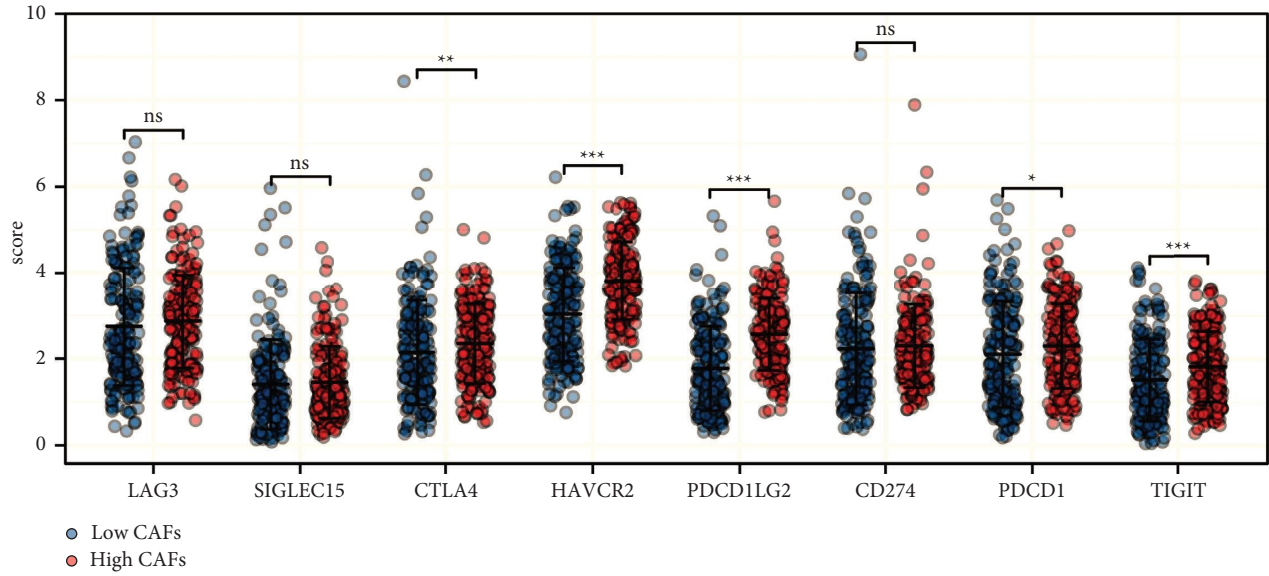


FIGURE 7: CAFs can affect the immunotherapy response of GC. (a) Multiple immune checkpoints level in GC patients with high and low CAFs infiltration. (b) TIDE analysis was performed to evaluate the immunotherapy response difference between high and low CAFs infiltration patients. (c-e) The difference of CAFs, FNDC1, and RSP03 in immunotherapy responders and nonresponders patients.

a higher TIDE score, as well as a lower percentage of immunotherapy responders (Figures S4B and S4C). Moreover, the immunotherapy responder had a low CAFs level (Figure 7(c)), as well as a lower FNDC1 and RSPO3 expression (Figures 7(d) and 7(e)). Drug sensitivity analysis showed that CAFs were negatively correlated with the IC₅₀ of doxorubicin, while positively correlated with the IC₅₀ of mitomycin and paclitaxel (Figures 8(a)–8(c)); FNDC1 was negatively correlated with the IC₅₀ of doxorubicin, while positively correlated with the IC₅₀ of paclitaxel (Figures 8(d)–8(f)); RSPO3 was negatively correlated with the IC₅₀ of doxorubicin, while positively correlated with the IC₅₀ of mitomycin and paclitaxel (Figures 8(g)–8(i)).

4. Discussion

A common cancer, GC poses one of the most serious public health problems [1]. CAFs are an important part of the TME in GC that can significantly affect cancer progression. Therefore, a deep investigation of CAFs and their related molecule targets would contribute to understanding the intrinsic biological mechanism of GC. In medical research, the investigation and analysis of the classification or prediction of response variables in biomedical research are often challenging due to the data sparsity generated by limited sample sizes and a moderate or very large number of predictors. Bioinformatic analysis can effectively solve this contradiction and is a powerful tool for screening clinical predictors [16].

In our study, we first quantified the CAFs infiltration using the ssGSEA algorithm. The clinical correlation result showed that CAFs were associated with a worse prognosis and clinical features. Pathway enrichment also indicated several oncogenic pathways in GC patients with high CAFs infiltration. Further, FNDC1 and RSPO3 were identified as the characteristic genes of CAFs through two machine learning algorithms, LASSO logistic regression and SVM-RFE. The following analysis showed that FNDC1 and RSPO3 were associated with more progressive clinical features and had a good prediction efficiency of the CAFs infiltration status in GC patients. Pathway enrichment and genomic instability were performed to explore the underlying mechanisms of FNDC1 and RSPO3. Immune infiltration analysis showed that CAFs were positively correlated with M2 macrophages. Moreover, we found that the GC patients with low CAFs infiltration were more sensitive to immunotherapy. Also, the CAFs, FNDC1, and RSPO3 could generate a certain effect on the sensitivity of doxorubicin, mitomycin, and paclitaxel.

Generally, in TME, the content of CAF is the most abundant, and it can affect the occurrence and development of cancer through intercellular contact, the release of various regulatory factors, and the remodeling of the extracellular matrix [17]. In colon cancer, Hu et al. indicated that CAFs could secrete the exosome miR-92a-3p that was engulfed by colon cancer cells, further activating Wnt/ β -catenin pathway and inhibiting mitochondrial apoptosis, leading to

metastasis and chemotherapy resistance [18]. Su et al. revealed that CD10⁺ GPR77⁺ CAFs could induce cancer formation and chemoresistance through sustaining tumor stemness [19]. Wen et al. indicated that CAFs-derived IL32 could promote breast cancer cell invasion and metastasis through integrin β 3-p38 MAPK signaling [20]. Pathway enrichment analysis showed that CAFs could activate the EMT, KRAS, and IL2/STAT5 signaling. In GC, Li et al. found that cancer-associated neutrophils could induce EMT through IL-17a to facilitate the invasion and migration of cancer cells [21]. Also, Wang et al. indicated that the downregulation of miRNA-214 in CAFs could enhance the migration and invasion of GC cells by targeting FGF9 and inducing EMT [22]. Our results were consistent with previous studies, which reflect the validity of the analysis.

Through machine learning algorithms, FNDC1 and RSPO3 were identified as the characteristic genes of CAFs. FNDC1, whose full name is “fibronectin type III domain containing 1”, has been reported to promote GC development. Jiang et al. demonstrated that FNDC1 could facilitate the invasion of GC by regulating the Wnt/ β -catenin signaling and is correlated with peritoneal metastasis [23]. RSPO3 has been reported as being widely involved in cancer progression. For example, Chen et al. revealed that RSPO3 could enhance the aggressiveness of bladder cancer through Wnt/ β -catenin and Hedgehog signaling pathways [24]. Fischer et al. found that in colon cancer with Wnt mutations, RSPO3 antagonism could hamper the malignant biological behavior of cancer cells [25]. However, virtually no study explored the RSPO3 in GC. Our study comprehensively investigated the underlying role of RSPO3 in GC, which can provide direction for future studies. In clinical practice, detecting the relative expression levels of FNDC1 and RSPO3 could indicate the CAFs infiltration level of patients, as well as their response on GC immunotherapy.

Interestingly, immune infiltration analysis showed that CAFs were associated with M2 macrophages. The interaction between different cells can significantly affect the remodeling effects of TME [26]. Previous studies have shown the underlying crosstalk between CAFs and M2 macrophages. Based on a coculture system, Cho et al. found that cancer-stimulated CAFs could promote M2 macrophage activation through secreting IL6 and GM-CSF [27]. Meanwhile, from a review summarized by Gunaydin, the interaction between CAFs and tumor-associated macrophages in TME can enhance tumorigenesis and immune escape [28]. Notably, our results also showed that in patients with low CAFs infiltration, the response rate to immunotherapy is higher (53.2% vs. 23.9%). Immunotherapy has shown a promising effect for specific advanced GC patients.

Although our research is based on high-quality bioinformatics analysis, some limitations should be noticed. First, the potential race bias is hard to ignore. Most patients enrolled in our study were from Western populations, which might decrease the credibility of our conclusions. Second, detailed laboratory examinations are hard to obtain. If all the

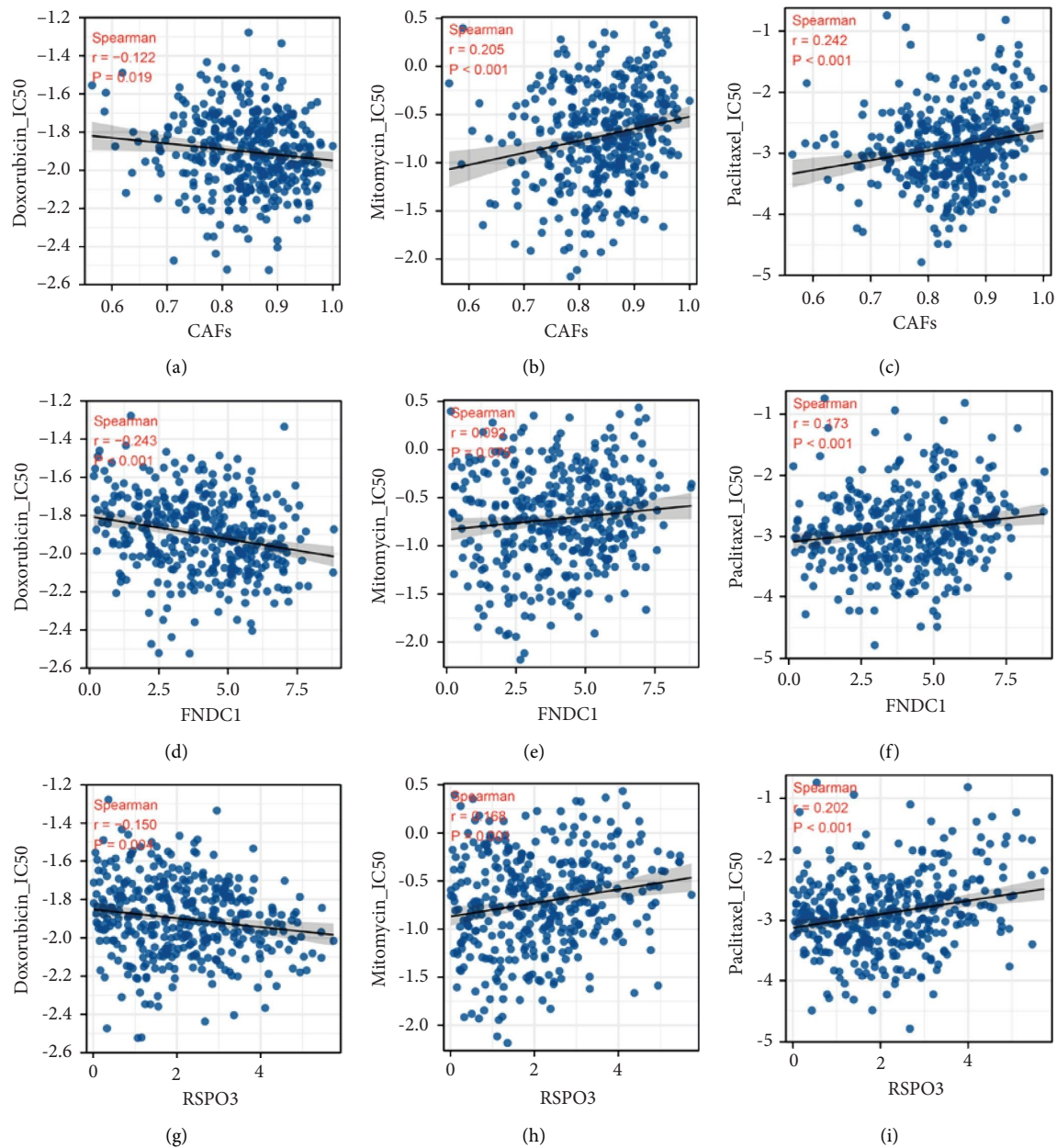


FIGURE 8: Drug sensitivity analysis. (a–c) The correlation of CAFs and the IC₅₀ of doxorubicin, mitomycin, and paclitaxel. (d–f) The correlation of FNDC1 and the IC₅₀ of doxorubicin, mitomycin, and paclitaxel. (g–i) The correlation of RSPO3 and the IC₅₀ of doxorubicin, mitomycin, and paclitaxel.

data from all examinations can be obtained, our conclusion will be more abundant.

Data Availability

The raw data mentioned in this study can be downloaded from online databases. The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

All the authors declare that there are no conflicts of interest.

Authors' Contributions

G-CC analyzed the data and wrote the manuscript. L-F helped data discussion. Y-QJ provided specialized expertise and collaboration in data analysis. G-AP conceived and designed the whole project and drafted the manuscript. All authors have read and approved the final version of the manuscript.

Acknowledgments

The authors greatly appreciate the analytical data provided by the TCGA and GEO databases. This work was supported

by the Jiangsu Provincial Health Commission (Grant no. H2019107).

Supplementary Materials

Figure S1: the flowchart of whole study. Figure S2: immune correlation analysis of FNDC1 and RSPO3. Notes: A: immune correlation analysis of FNDC1; B: immune correlation analysis of RSPO3. Figure S3: submap analysis was used to indicate patients' sensitivity to PD-1 and CTLA4 therapy. Figure S4: effect of M2 macrophages on GC immunotherapy. Notes: A: correlation of M2 macrophages and TIDE score; B: TIDE score in patients with high and low M2 macrophages infiltration; C: the percentage of immunotherapy responders and nonresponders in patients with high and low M2 macrophages infiltration. (*Supplementary Materials*)

References

- [1] E. C. Smyth, M. Nilsson, H. I. Grabsch, N. C. Van Grieken, and F. Lordick, "Gastric cancer," *The Lancet*, vol. 396, no. 10251, pp. 635–648, 2020.
- [2] F. Liu, C. Huang, Z. Xu et al., "Morbidity and mortality of laparoscopic vs. open total gastrectomy for clinical stage I gastric cancer: the Class02 multicenter randomized clinical trial," *JAMA Oncology*, vol. 6, no. 10, pp. 1590–1597, 2020.
- [3] Z. Song, Y. Wu, J. Yang, D. Yang, and X. Fang, "Progress in the treatment of advanced gastric cancer," *Tumor Biology*, vol. 39, no. 7, Article ID 101042831771462, 2017.
- [4] T. Wu and Y. Dai, "Tumor microenvironment and therapeutic response," *Cancer Letters*, vol. 387, pp. 61–68, 2017.
- [5] G. Biffi and D. A. Tuveson, "Diversity and biology of cancer-associated fibroblasts," *Physiological Reviews*, vol. 101, no. 1, pp. 147–176, 2021.
- [6] X. Mao, J. Xu, W. Wang et al., "Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives," *Molecular Cancer*, vol. 20, no. 1, 131 pages, 2021.
- [7] G. Kharashvili, D. Simkova, K. Bouchalova, M. Gachechiladze, N. Narsia, and J. Bouchal, "The role of cancer-associated fibroblasts, solid stress and other microenvironmental factors in tumor progression and therapy resistance," *Cancer Cell International*, vol. 14, no. 1, 41 pages, 2014.
- [8] Y. Liubomirski, S. Lerrer, T. Meshel et al., "Tumor-stroma-inflammation networks promote pro-metastatic chemokines and aggressiveness characteristics in triple-negative breast cancer," *Frontiers in Immunology*, vol. 10, 10757 pages, 2019.
- [9] L. A. Jolly, S. Novitskiy, P. Owens et al., "Fibroblast-Mediated Collagen Remodeling within the Tumor Microenvironment Facilitates Progression of Thyroid Cancers Driven by Brafv600e and Pten Loss," *Cancer Research*, vol. 76, no. 7, pp. 1804–1813, 2016.
- [10] Z. Dai, X. Y. Gu, S. Y. Xiang, D. D. Gong, C. F. Man, and Y. Fan, "Research and application of single-cell sequencing in tumor heterogeneity and drug resistance of circulating tumor cells," *Biomarker research*, vol. 8, no. 1, 60 pages, 2020.
- [11] S. Hänzelmann, R. Castelo, and J. Guinney, "Gsva: gene set variation analysis for microarray and rna-seq data," *BMC Bioinformatics*, vol. 14, no. 1, 7 pages, 2013.
- [12] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Campbell, "Introduction to machine learning, neural networks, and deep learning," *Translational vision science & technology*, vol. 9, no. 2, 14 pages, 2020.
- [13] B. Chen, M. S. Khodadoust, C. L. Liu, A. M. Newman, and A. A. Alizadeh, "Profiling tumor infiltrating immune cells with cibersort," *Methods in Molecular Biology*, vol. 1711, pp. 243–259, 2018.
- [14] T. M. Malta, A. Sokolov, A. J. Gentles et al., "Machine learning identifies stemness features associated with oncogenic dedifferentiation," *Cell*, vol. 173, no. 2, pp. 338–354, 2018.
- [15] K. Ito and D. Murphy, "Application of Ggplot2 to pharmacometric graphics," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 2, no. 10, e79 pages, 2013.
- [16] Z. Yin, D. Wu, J. Shi et al., "Identification of Aldh3a2 as a novel prognostic biomarker in gastric adenocarcinoma using integrated bioinformatics analysis," *BMC Cancer*, vol. 20, no. 1, 1062 pages, 2020.
- [17] Y. Chen, K. M. McAndrews, and R. Kalluri, "Clinical and therapeutic relevance of cancer-associated fibroblasts," *Nature Reviews Clinical Oncology*, vol. 18, no. 12, pp. 792–804, 2021.
- [18] J. L. Hu, W. Wang, X. L. Lan et al., "Cafs secreted exosomes promote metastasis and chemotherapy resistance by enhancing cell stemness and epithelial-mesenchymal transition in colorectal cancer," *Molecular Cancer*, vol. 18, no. 1, 91 pages, 2019.
- [19] S. Su, J. Chen, H. Yao et al., "Cd10(+)/Gpr77(+) cancer-associated fibroblasts promote cancer formation and chemoresistance by sustaining cancer stemness," *Cell*, vol. 172, no. 4, pp. 841–856, 2018.
- [20] S. Wen, Y. Hou, L. Fu et al., "Cancer-associated fibroblast (Caf)-Derived Il32 promotes breast cancer cell invasion and metastasis via integrin B3-P38 mapk signalling," *Cancer Letters*, vol. 442, pp. 320–332, 2019.
- [21] S. Li, X. Cong, H. Gao et al., "Tumor-associated neutrophils induce emt by il-17a to promote migration and invasion in gastric cancer cells," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, 6 pages, 2019.
- [22] R. Wang, Y. Sun, W. Yu et al., "Downregulation of mirna-214 in cancer-associated fibroblasts contributes to migration and invasion of gastric cancer cells through targeting Fgf9 and inducing emt," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, 20 pages, 2019.
- [23] T. Jiang, W. Gao, S. Lin et al., "Fndc1 promotes the invasiveness of gastric cancer via wnt/B-catenin signaling pathway and correlates with peritoneal metastasis and prognosis," *Frontiers in Oncology*, vol. 10, Article ID 590492, 2020.
- [24] Z. Chen, L. Zhou, L. Chen et al., "Rspo3 promotes the aggressiveness of bladder cancer via wnt/B-catenin and hedgehog signaling pathways," *Carcinogenesis*, vol. 40, no. 2, pp. 360–369, 2019.
- [25] M. M. Fischer, V. P. Yeung, F. Cattaruzza et al., "Rspo3 antagonism inhibits growth and tumorigenicity in colorectal tumors harboring common Wnt pathway mutations," *Scientific Reports*, vol. 7, no. 1, Article ID 15270, 2017.
- [26] J. Chen, S. Chen, X. Dai et al., "Exploration of the underlying biological differences and targets in ovarian cancer patients with diverse immunotherapy response," *Frontiers in Immunology*, vol. 13, Article ID 1007326, 2022.
- [27] H. Cho, Y. Seo, K. M. Loke et al., "Cancer-stimulated cafs enhance monocyte differentiation and protumoral tam activation via Il6 and gm-csf secretion," *Clinical Cancer Research*, vol. 24, no. 21, pp. 5407–5421, 2018.
- [28] G. Gunaydin, "Cafs interacting with tams in tumor microenvironment to enhance tumorigenesis and immune evasion," *Frontiers in Oncology*, vol. 11, Article ID 668349, 2021.

Research Article

HPV-Related Prognostic Signature Predicts Survival in Head and Neck Squamous Cell Carcinoma

Hongyu Zhao ¹, Fengxu Wang,^{2,3} Xuehai Wang,³ Xinyuan Zhao ³ and Jinfeng Ji ²

¹Department of Radiotherapy Oncology, Affiliated Hospital of Nantong University, Nantong 226000, China

²Department of Integrated Traditional Chinese and Western Internal Medicine, Affiliated Tumor Hospital of Nantong University, Nantong Tumor Hospital, Nantong 226631, China

³Department of Occupational Medicine and Environmental Toxicology, Nantong Key Laboratory of Environmental Toxicology, School of Public Health, Nantong University, Nantong 226019, China

Correspondence should be addressed to Xinyuan Zhao; zhaoxinyuan@ntu.edu.cn and Jinfeng Ji; jijfeng@163.com

Received 16 September 2022; Revised 3 October 2022; Accepted 6 October 2022; Published 15 November 2022

Academic Editor: Feng Jiang

Copyright © 2022 Hongyu Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Head and neck squamous cell carcinoma (HNSCC) is one of the most common cancers, worldwide. Considering the role of human papilloma virus (HPV) in tumor development and sensitivity to treatment of HNSCC, we aimed to explore the prognostic classification ability of HPV-related signatures in head and neck cancer. **Methods.** HPV-related signatures were screened out based on Gene Expression Omnibus (GEO) and the Cancer Genome Atlas (TCGA) databases. HPV-related signatures with prognostic value were identified through univariate Cox regression analysis and a risk signature was established by least absolute shrinkage and selection operator (LASSO). Further, we developed a nomogram by integrating independent prognostic factors. **Results.** A total of 55 HPV-associated signatures were differentially expressed and ten of them were associated with prognosis of HNSCC patients. The prognostic signature based on CDKN2A, CELSR3, DMRTA2, SERPINE1, TJP3, FADD, and IGF2BP2 expression was constructed. Univariate and multivariate regression analyses demonstrated that the novel prognostic signature was an independent prognostic factor of HNSCC. The nomogram integrating the prognostic signature and other independent prognostic factors was developed. **Conclusion.** In summary, the prognostic signature of the HPV-related signatures might serve as an important prognostic biomarker for patients with HNSCC.

1. Introduction

Nearly 95% of head and neck cancer cases are caused by head and neck squamous cell carcinoma (HNSCC), considered the sixth most common form of cancer globally [1]. It has been reported that, despite advances in treatment modalities, outcomes of HNSCC patients have not been improved significantly [2]. Local invasion and distant metastasis at first diagnosis are the main reasons for their worse prognosis [1]. Therefore, it is urgent to construct new prognostic models to aid early clinical diagnosis.

As far as etiology is concerned, HNSCC is primarily caused by tobacco and alcohol use [3]. Throughout the world, research studies on HPV have proliferated in recent years [4]. Globally, approximately 25% of the number of

head and neck cancer cases are caused by HPV infection [4]. Also, HPV causes almost all cases of cervical cancer and a subset of other anogenital cancers [5].

In this study, we tried to identify HPV-related signatures and construct a novel prognostic model and nomogram by using sequencing data and clinically relevant information of patients with HNSCC. We expect to improve the diagnosis and prognosis of patients with head and neck cancer from the perspective of HPV.

2. Materials and Methods

2.1. Data Collection. To identify HPV-related signatures, we selected and downloaded GSE65858 cohort from Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/>)

geo/) [6], which included 73 HPV-positive and 196 HPV-negative patients with HNSCC, and clinical information is shown in Supplementary Table 2. The package “limma” [7] was used for differential expression analysis with the screening criteria ($\text{adj.}p < 0.05$). Transcript expression profiling of 501 HNSCC tumor samples and 44 adjacent normal tissues was obtained from TCGA [8], and clinical information is shown in Supplementary Table 3.

2.2. Identification of Differentially Expressed HPV-Related Signatures. HPV-related signature expression in HNSCC patients was analyzed by using the “limma” package [7], with selection standard of $|\log(\text{FC})| \geq 2$ and $\text{adj.}p < 0.01$.

2.3. Enrichment Analysis. To reveal the biological features of hub HPV-related signatures, Gene Ontology (GO) [9] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [10] pathway analyses were conducted in this part. Tightly linked interaction network may facilitate the exploration of specific mechanisms in HNSCC, and closely linked genes may lead to an accurate prognostic model. In parallel, we constructed a protein interaction network by using the STRING database (<https://string-db.org/>) [11] in order to visualize the hub signatures.

2.4. Construction and Verification of the Prognostic Signature. Prognosis-associated signatures were picked out through the univariate Cox regression analysis with $p < 0.01$ in TCGA cohort. To avoid overfitting, least absolute shrinkage and selection operator (LASSO) analysis was applied to further screen HPV-related signatures. We calculated the risk score for each HNSCC patient by using the following formula: $\text{Risk score} = \text{gene (A) expression} \times \text{coef (A)} + \text{gene (B) expression} \times \text{coef (B)} + \text{gene (i) expression} \times \text{coef (i)}$ [12]. Patients with HNSCC were divided into the high- and low-risk groups according to the median risk score. Kaplan–Meier survival analysis was used to compare the overall survival rate between the high- and low-risk groups. Principal component analysis (PCA) was used to evaluate the clustering ability of the prognostic signature.

2.5. Immune Infiltration Analysis. In this study, we attempted to perform immune infiltration analysis by using the CIBERSORT [13] and single-sample Gene Set Enrichment Analysis (ssGSEA) algorithm [14]. Moreover, we attempted to mine drugs targeting the high-risk group with poor prognosis by using the Genomics of Drug Sensitivity in Cancer (GDSC) database [15].

2.6. Cox Regression Analyses and Construction of a Nomogram. Univariate and multivariate Cox regression analyses were used to evaluate independent prognostic factors in HNSCC. A nomogram integrating the independent prognostic factors was established by “rms”

package [16]. Calibration curves and area under the curve (AUC) were used to verify the validity of the nomogram [17, 18] we constructed.

3. Results

3.1. Expression and Pathway Enrichment of HPV-Related Signatures in HNSCC Patients. We derived 1136 HPV-associated signatures with significant differences from the GSE65858 cohort (Figure 1(a), Supplemental Table 1). A total of 55 HPV-associated signatures were differentially expressed in TCGA cohort according to screening criteria (Figures 1(b)–1(c)). In the light of GO and KEGG pathway analyses, HPV-associated signatures were significantly accumulated in focal adhesion, ECM-receptor interaction, AGE-RAGE signaling pathway in diabetic complications, protein digestion and absorption, and relaxin signaling pathway (Figures 1(d)–1(e)). Also, a protein interaction network was constructed with tightly linked HPV-associated signatures (Figure 1(f)).

3.2. Construction and Validation of a Prognostic Model Based on HPV-Related Signatures in HNSCC Patients. Ten prognosis-related signatures were identified by univariate Cox regression analysis (Figure 2(a)). To avoid overfitting, LASSO regression was performed and selected seven HPV-related signatures for constructing the prognostic signature (Figures 2(b)–2(c)). $\text{Risk score} = (-0.0413836017 \times \text{Exp CDKN2A}) + (-0.365345241419471 \times \text{Exp CELSR3}) + (-0.0633132503419892 \times \text{Exp DMRTA2}) + (0.0162329920245241 \times \text{Exp SERPINE1}) + (-0.0468942153601447 \times \text{Exp TJP3}) + (0.12874495442016 \times \text{Exp FADD}) + (0.139702586686539 \times \text{Exp IGF2BP2})$.

Kaplan–Meier survival curves showed a worse overall survival and progression-free survival in the high-risk group when compared with the low-risk group (Figures 3(a)–3(b)). Meanwhile, the prognostic signature showed a great clustering ability compared with the cohort without genotyping (Figures 3(c)–3(d)). As shown in Figure 3(e), seven HPV-related signatures are ranked in different groups. The distribution of the risk scores and survival time are shown in Figures 3(f)–3(g). Meanwhile, the time-dependent ROC curves of the prognostic signature were performed and shown in Supplemental Figure 1A. The relationship between prognostic signature and clinicopathological factors is shown in Figures 4(a)–4(f). Interestingly, we found that the risk score was highly expressed in smokers compared to nonsmokers ($p = 0.00046$, Figure 4(e)). In addition, HPV-negative patients had a significantly higher risk value than HPV-positive patients with HNSCC (Supplemental Figure 1B).

3.3. Immune Infiltration Analysis. Through CIBERSORT algorithm, we found that naive B cells, plasma cells, CD8 T cells, activated memory CD4 T cells, follicular helper T cells, regulatory T cells (Tregs), gamma delta T cells, and

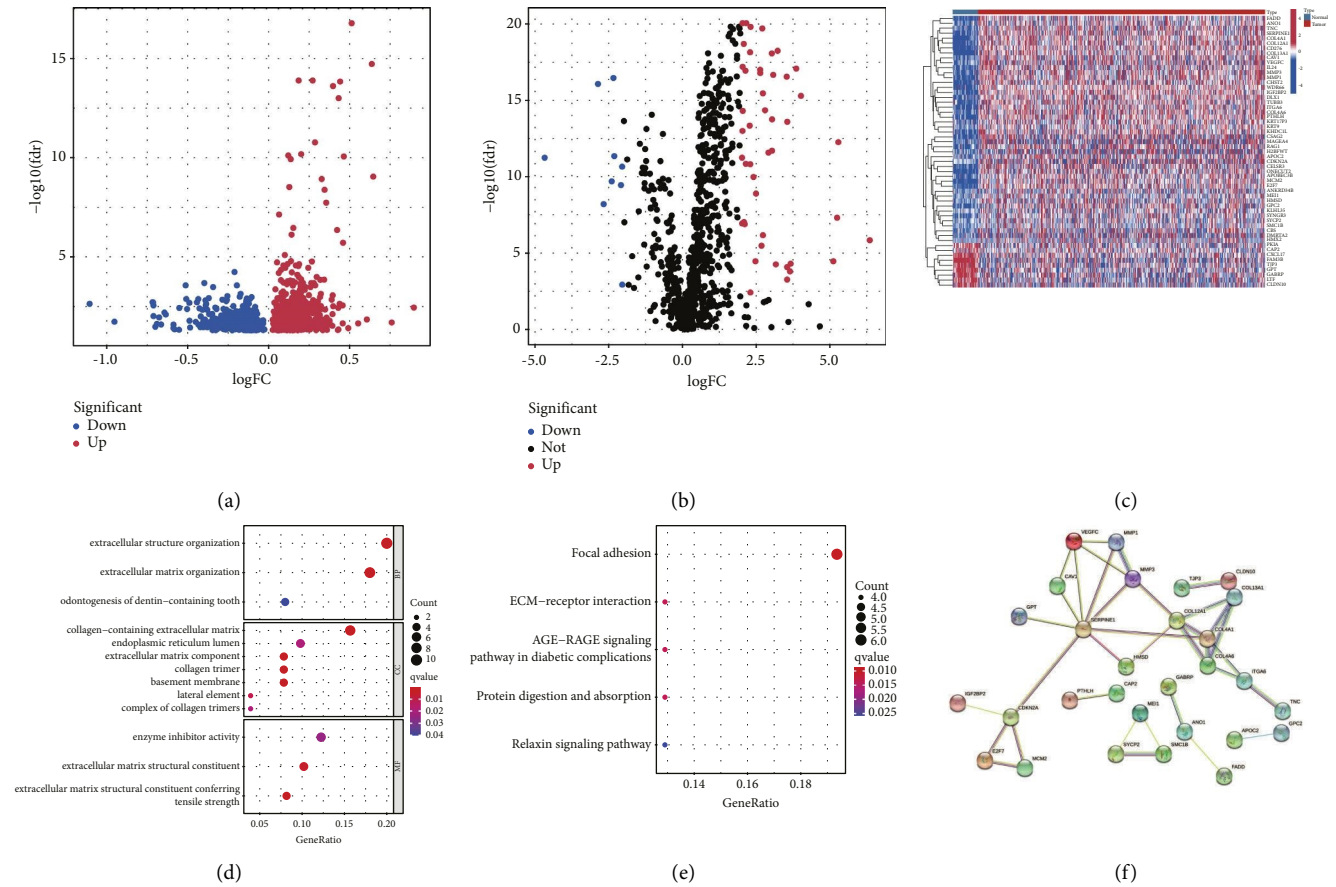


FIGURE 1: Expression and pathway enrichment of HPV-related signatures in HNSCC patients. (a) Volcano plot of HPV-associated signatures, red dots represent upregulated and blue dots represent downregulated genes; (b) volcano plot of TCGA cohort, red dots represent upregulated and blue dots represent downregulated genes; (c) heat map of TCGA cohort, the horizontal axis represents HNSCC samples and the vertical axis represents hub genes; (d) GO terms of HPV-associated signatures; (e) KEGG pathways of HPV-associated signatures; and (f) PPI network of HPV-associated signatures, the thicker the lines, the tighter the relationships.

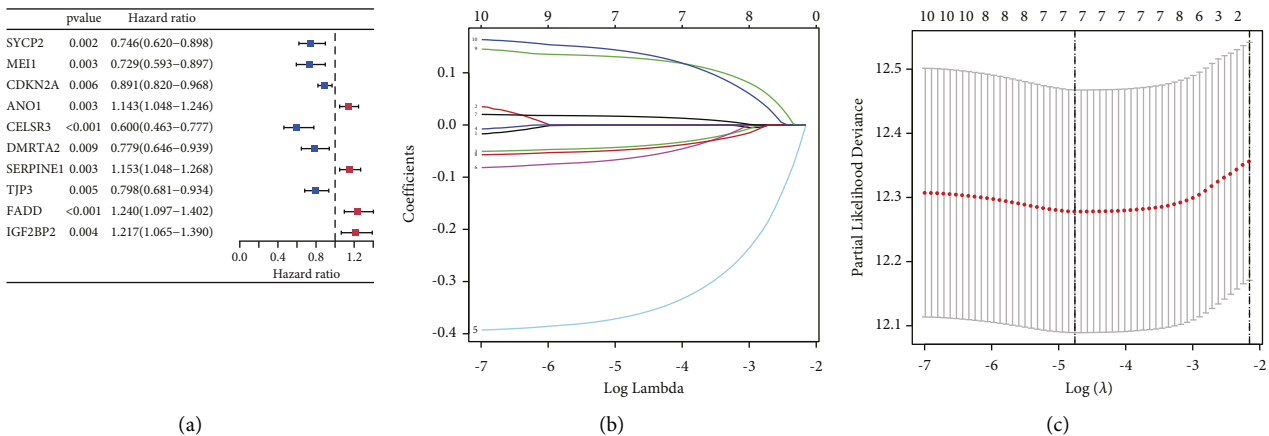


FIGURE 2: Construction of a prognostic signature based on HPV-related signatures. (a) Univariate Cox regression analysis, blue dots represent protective factors and red dots represent risk factors; (b) LASSO coefficient profiles, curves represent prognostic HPV-related signatures screened from univariate Cox regression analysis; and (c) LASSO deviance profiles.

resting mast cells were highly enriched in the low-risk group (Figure 5(a)) while resting memory CD4 T cells, M0 macrophages, activated dendritic cells, and eosinophils were

highly enriched in the high-risk group (Figure 5(a)). Based on the ssGSEA algorithm, moreover, check-point, cytolytic activity, HLA, inflammation-promoting, T cell co-

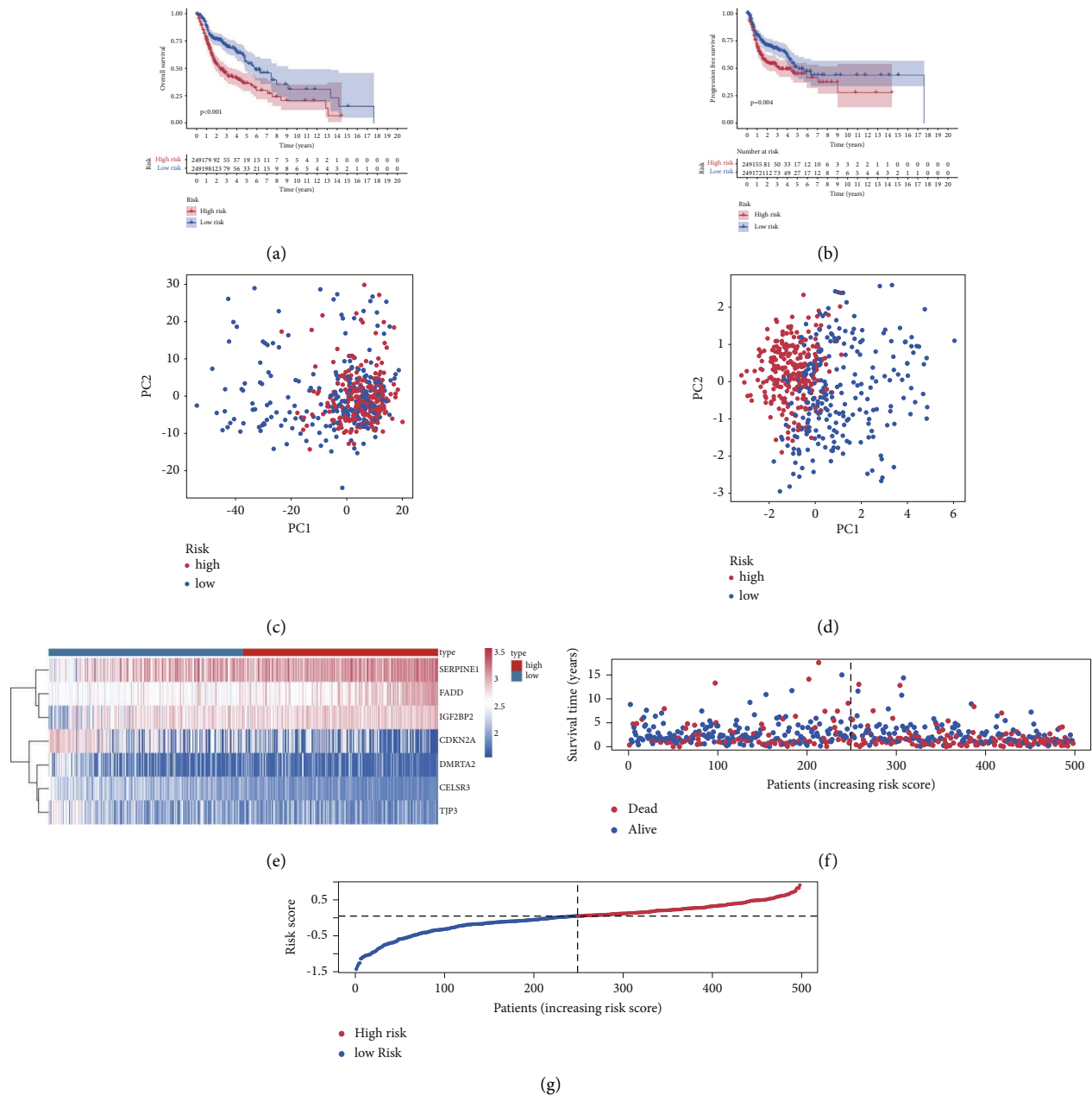


FIGURE 3: Validation of the prognostic signature. (a) Kaplan–Meier curve for overall survival, the blue curve represents the low-risk group and red curve represents the high-risk group; (b) Kaplan–Meier curve for progression-free survival, the blue curve represents the low-risk group and red curve represents the high-risk group; (c) principal component analysis without prognostic signature, the blue dots represent patients in the low-risk group and red dots represent patients in the high-risk group; (d) principal component analysis with prognostic signature, the blue dots represent patients in the low-risk group and red dots represent patients in the high-risk group; (e) heat map of HPV-related signatures expression; (f) survival status plot based on prognostic signature; and (g) risk score plot classified by prognostic signature.

inhibition, T cell co-stimulation, and type II IFN response were highly enriched in the low-risk group (Figure 5(b)). Moreover, WH-4-023, an inhibitor of LCK/SRC, was found to be useful in the high-risk group through pRRophetic algorithm (Figures 5(c)-5(d)).

3.4. The Prognostic Signature Was an Independent Prognostic Factor for HNSCC Patients. Using both univariate and multivariate Cox regression analyses, age, stage, and the

prognostic signature were found to be reliable independent predictors of HNSCC (Figures 6(a)-6(b)). Subsequently, a nomogram integrating independent prognostic factors with significant differences was developed to predict the one-, three-, and five- year OS of HNSCC patients (Figure 6(c)). In calibration curves, the predictive curves originated from the nomogram showed a high agreement with the ideal curve (Figure 6(d)). AUC values of the nomogram were 0.665, 0.679, and 0.645 for one-, three-, and five-year OS,

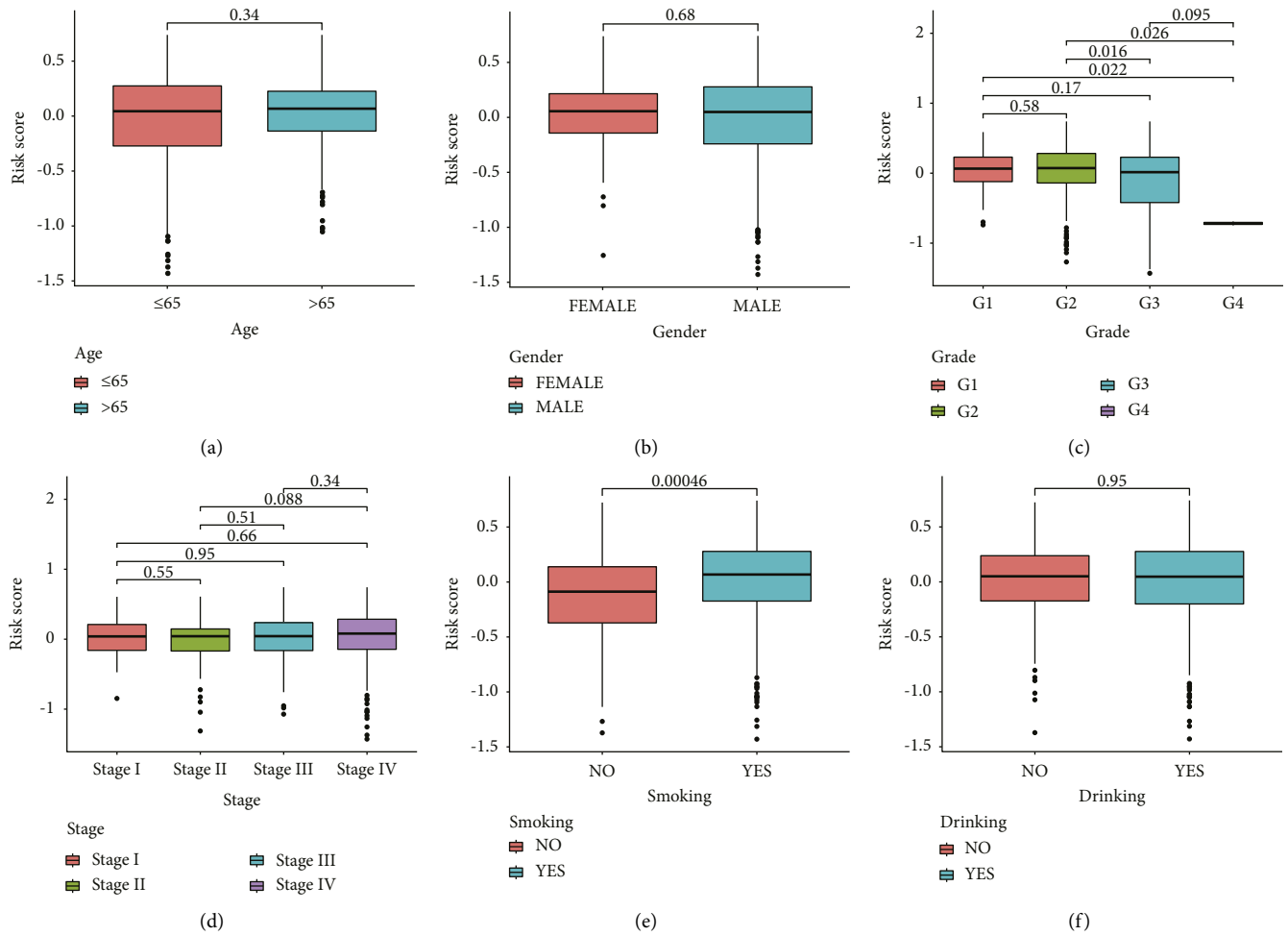


FIGURE 4: Relationship between prognostic signature expression and age (a), gender (b), grade (c), stage (d), smoking status (e), and drinking (f).

respectively, demonstrating a better predictive accuracy of the nomogram compared with that of a single clinical factor (Figures 6(e)–6(g)).

4. Discussion

Head and neck squamous cell carcinoma is a collection of epithelial tumors originating mainly from the mucosa of the oral cavity, oropharynx, larynx, or hypopharynx [19]. Despite continuous advances in treatment modalities, the prognosis of patients with head and neck cancer has not improved significantly over the past decades [20]. The high recurrence rate of patients and the low response to intervention treatments such as chemoradiotherapy are the main reasons for the poor prognosis of patients [21, 22].

In recent years, the popularization and rapid development of next-generation sequencing (NGS) technology have enabled us to have a new understanding and insight into the molecular landscape of different tumors [23]. For head and neck cancer, these advances provided new insights into the different molecular mechanisms of HNSCC, the tumor microenvironment (TME) and heterogeneity of HNSCC,

and the diversity of clinical responses among HNSCC subtypes [23–25].

In this study, we attempted to construct a new prognostic model from the perspective of HPV by microarray and high-throughput sequencing data. The flowchart of this study is shown in Figure 7. On the basis of the important impact of HPV on HNSCC, we identified HPV-associated signatures by comparing the transcriptome data of HPV-positive and HPV-negative HNSCC patients from the GEO dataset. To achieve a larger sample size to justify the scientific quality of this research, TCGA database was used for subsequent analyses. Ultimately, 55 HPV-associated signatures were differentially expressed and were significantly accumulated in focal adhesion, ECM-receptor interaction, AGE-RAGE signaling pathway in diabetic complications, protein digestion and absorption, and relaxin signaling pathway. Subsequently, seven HPV-related signatures (CDKN2A, CELSR3, DMRTA2, SERPINE1, TJP3, FADD, and IGF2BP2) that enabled the classification of high- and low-risk HNSCC patients were screened out.

Among these prognostic HPV-associated signatures, we noted one of the interesting. Cyclin-dependent kinase

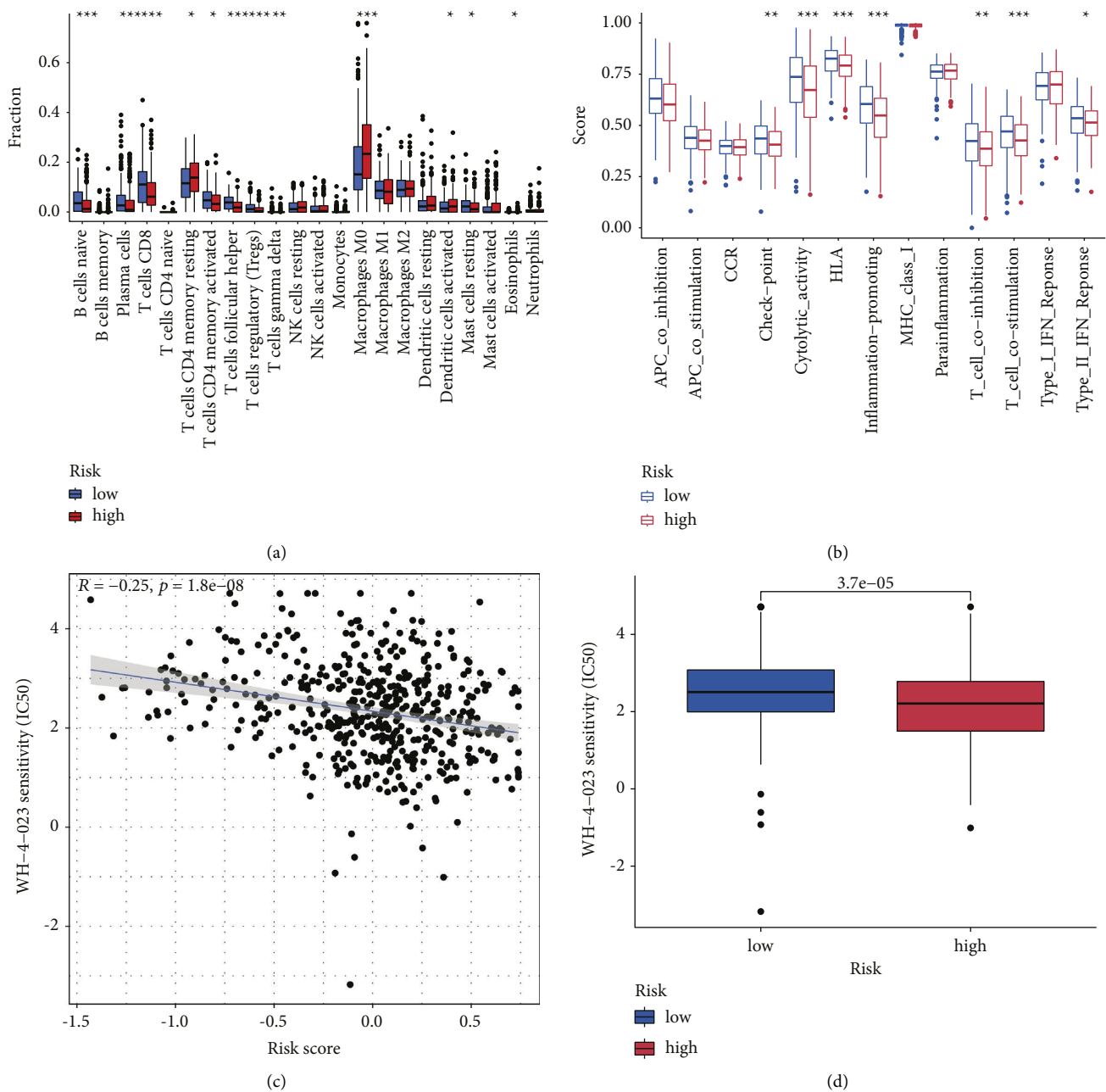


FIGURE 5: Relationship between prognostic signature and immune infiltration. (a) The infiltrating levels of 22 immune cells in low- and high-risk groups via CIBERSORT algorithm; (b) the levels of immune activity in low- and high-risk groups via the ssGSEA algorithm; and (c)-(d) relationship between WH-4-023 sensitivity and risk scores.

inhibitor 2A, also known as p16, plays a critical role in cell-cycle regulation [26]. In addition, infection with HPV leads to overexpression of p16, and thus p16 is often used as a common marker of HPV positivity [26, 27]. Recently, multiple studies have surfaced that p16 expression in head and neck cancers was independent of HPV infection and should not serve as a reliable marker for HPV infection [28, 29]. Consistent with previous studies, p16 functioned as a tumor suppressor in HNSCC patients of this study and as a risk score reducer in the prognostic model with a negative coef value.

Interestingly, in the present study, we found that smoking HNSCC patients had a higher risk value. These results indicated that smoking would affect HPV status as well as the expression of HPV-related signatures and may give clues to the mechanism of head and neck cancers. Through prognostic analysis, we validated the prognostic model's validity and accuracy. In addition, immune infiltration analysis showed a higher level of immune cell infiltration in the low-risk group than that of the high-risk group, which may explain the worse prognosis in the high-risk group. To our knowledge, this study is the first attempt

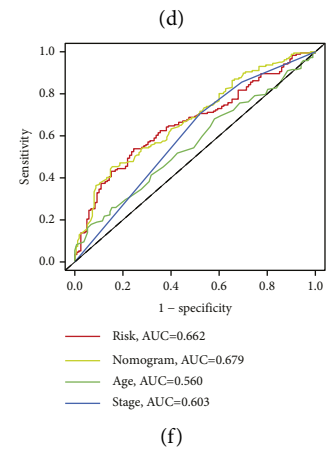
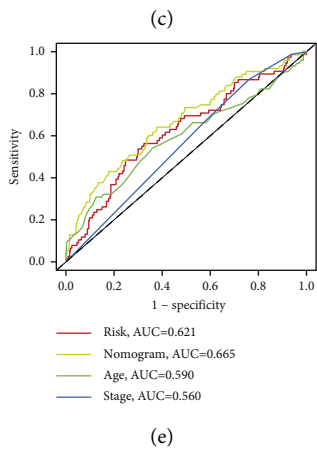
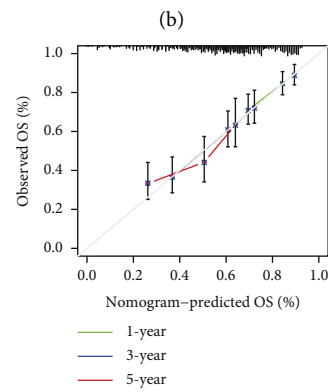
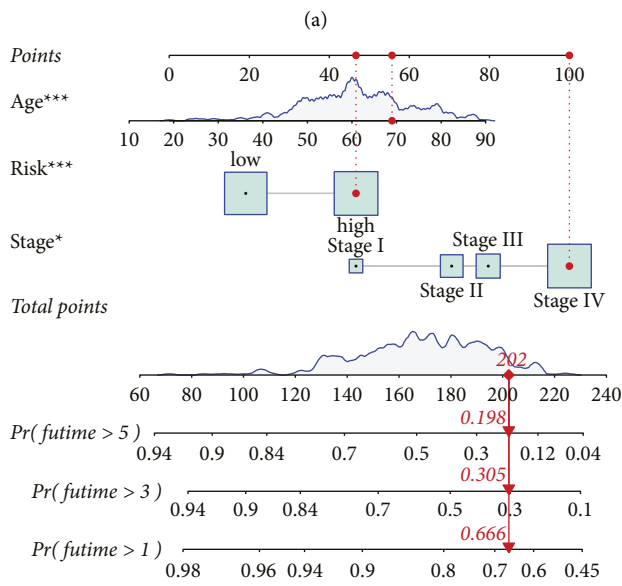
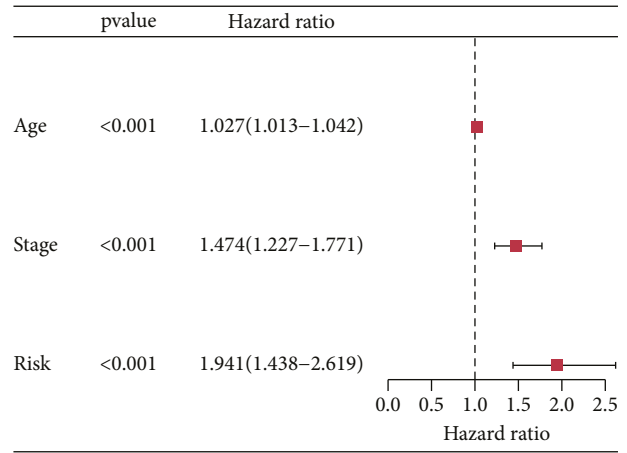
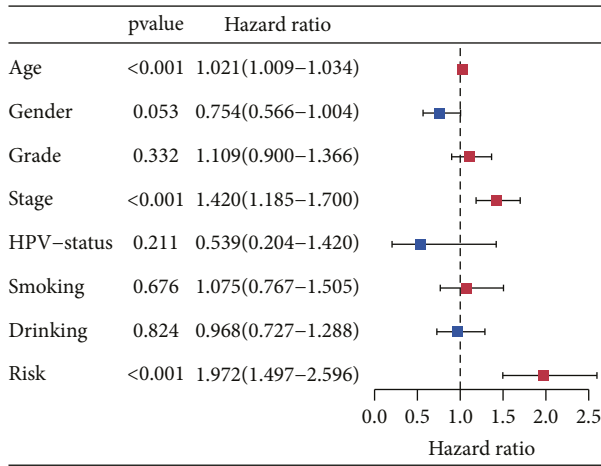


FIGURE 6: Continued.

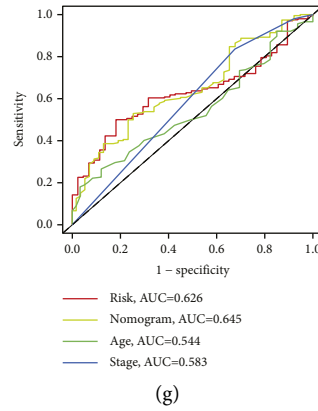


FIGURE 6: Independent prognostic analysis and construction of a novel nomogram. (a) Univariate Cox regression analysis for the overall survival of HNSCC patients; (b) multivariate Cox regression analysis for the overall survival of HNSCC patients; (c) the nomogram on the basis of independent prognostic factors; (d) time-dependent calibration curves of the nomogram; (e) ROC curves for one-year OS; (f) ROC curves for three-year OS; and (g) ROC curves for five-year OS.

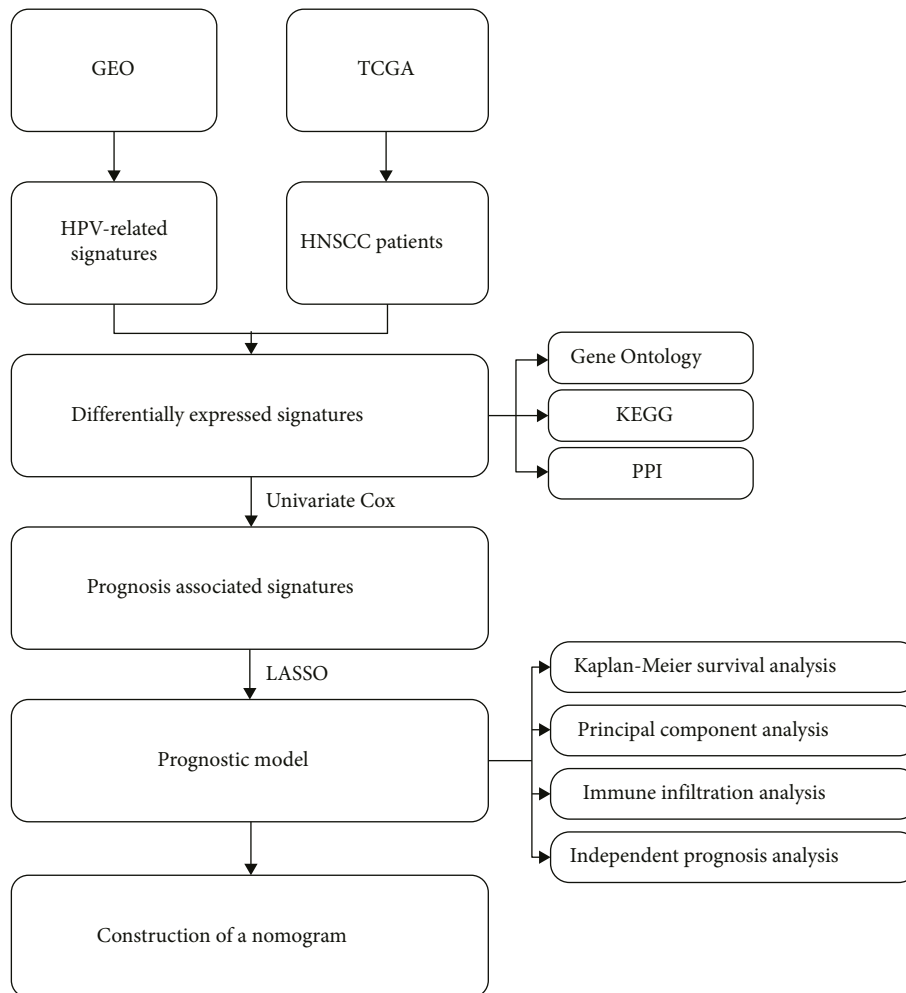


FIGURE 7: Flowchart of this study.

to construct a prognostic signature for head and neck cancer based on HPV-related genes. For the convenience of clinical application, a novel nomogram based on independent prognostic factors of head and neck cancer was constructed.

Meanwhile, in our study, we encountered a few limitations. The sample size and cohort size need to be expanded to guarantee their accuracy for this study. More patients with HNSCC and more prospective clinical trials need to be

included in the calculation of the prognostic model for reducing statistical bias. Moreover, further biochemical experiments in vivo and in vitro on the seven HPV-related signatures should be conducted further.

5. Conclusion

We screened out differentially expressed signatures between HPV⁺ and HPV⁻ HNSCC patients and developed a novel prognostic signature based on large sample datasets. Meanwhile, a novel nomogram was constructed by integrating independent prognostic factors.

Data Availability

The datasets analyzed during the current study are available in TCGA repository (<https://portal.gdc.cancer.gov>) and GEO (<https://www.ncbi.nlm.nih.gov/geo/>).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Hongyu Zhao and Fengxu Wang performed the data curation and analysis. Xuehai Wang and Xinyuan Zhao analyzed and interpreted the results. Jinfeng Ji drafted and reviewed the manuscript. All authors read and approved the final manuscript. Hongyu Zhao and Fengxu Wang contributed equally to this work.

Acknowledgments

This study was supported by Natural Science Foundation of Jiangsu Province (BK20201444); Qing Lan Project for Excellent Young Key Teachers of Colleges and Universities of Jiangsu Province (2020).

Supplementary Materials

Supplemental Table 1: HPV-associated signatures with significant differences from the GSE65858 cohort. Supplemental Table 2: clinical information of the GEO cohort. Supplemental Table 3: clinical information of TCGA cohort. Supplemental Figure 1: ROC curves of the prognostic signature (A); the relationship between the prognostic signature and HPV status (B). (*Supplementary Materials*)







References

- [1] A. Argiris, M. V. Karamouzis, D. Raben, and R. L. Ferris, "Head and neck cancer," *The Lancet*, vol. 371, no. 9625, pp. 1695–1709, 2008.
- [2] S. H. Huang and B. O'Sullivan, "Overview of the 8th edition TNM classification for head and neck cancer," *Current Treatment Options in Oncology*, vol. 18, no. 7, p. 40, 2017.
- [3] E. M. Rettig and G. D'Souza, "Epidemiology of head and neck cancer," *Surgical Oncology Clinics of North America*, vol. 24, no. 3, pp. 379–396, 2015.
- [4] E. Tumban, "A current update on human papillomavirus-associated head and neck cancers," *Viruses*, vol. 11, no. 10, p. E922, 2019.
- [5] L. Zhai and E. Tumban, "Gardasil-9: a global survey of projected efficacy," *Antiviral Research*, vol. 130, pp. 101–109, 2016.
- [6] T. Barrett, S. E. Wilhite, P. Ledoux et al., "NCBI GEO: archive for functional genomics data sets--update," *Nucleic Acids Research*, vol. 41, no. D1, pp. D991–D995, 2012.
- [7] M. E. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, p. e47, 2015.
- [8] A. Blum, P. Wang, and J. C. Zenklusen, "SnapShot: TCGA-analyzed tumors," *Cell*, vol. 173, no. 2, p. 530, 2018.
- [9] The Gene Ontology Consortium, "Gene Ontology consortium: going forward," *Nucleic Acids Research*, vol. 43, no. D1, pp. D1049–D1056, 2015.
- [10] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.
- [11] C. v Mering, M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, and B. Snel, "STRING: a database of predicted functional associations between proteins," *Nucleic Acids Research*, vol. 31, no. 1, pp. 258–261, 2003.
- [12] Y. Shen, H. Zhou, S. Dong, W. Dong, and L. Zhang, "Smoking patients with laryngeal cancer screened with a novel immunogenomics-based prognostic signature," *Frontiers in Genetics*, vol. 13, Article ID 961764, 2022.
- [13] A. M. Newman, C. L. Liu, M. R. Green et al., "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.
- [14] S. Shen, G. Wang, R. Zhang et al., "Development and validation of an immune gene-set based Prognostic signature in ovarian cancer," *EBioMedicine*, vol. 40, pp. 318–326, 2019.
- [15] W. Yang, J. Soares, P. Greninger et al., "Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells," *Nucleic Acids Research*, vol. 41, no. D1, pp. D955–D961, 2012.
- [16] Y. Shen, Q. Huang, Y. Zhang, C.-Y. Hsueh, and L. Zhou, "A novel signature derived from metabolism-related genes GPT and SMS to predict prognosis of laryngeal squamous cell carcinoma," *Cancer Cell International*, vol. 22, no. 1, p. 226, 2022.
- [17] A. C. Alba, T. Agoritsas, M. Walsh et al., "Discrimination and calibration of clinical prediction models: users' guides to the medical literature," *JAMA*, vol. 318, no. 14, pp. 1377–1384, 2017.
- [18] D. L. Streiner and J. Cairney, "What's under the ROC? An introduction to receiver operating characteristics curves," *Canadian Journal of Psychiatry*, vol. 52, no. 2, pp. 121–128, 2007.
- [19] D. Hashim, E. Genden, M. Posner, M. Hashibe, and P. Boffetta, "Head and neck cancer prevention: from primary prevention to impact of clinicians on reducing burden," *Annals of Oncology*, vol. 30, no. 5, pp. 744–756, 2019.
- [20] X. Li, A. Lee, M. A. Cohen, E. J. Sherman, and N. Y. Lee, "Past, present and future of proton therapy for head and neck cancer," *Oral Oncology*, vol. 110, Article ID 104879, 2020.
- [21] G. Tolstonog and C. Simon, "Trends in surgical research in head and neck cancer," *Current Treatment Options in Oncology*, vol. 18, no. 6, p. 38, 2017.
- [22] W. M. Lydiatt, S. G. Patel, B. O'Sullivan et al., "Head and Neck cancers-major changes in the American Joint Committee on

- cancer eighth edition cancer staging manual,” *CA: A Cancer Journal for Clinicians*, vol. 67, no. 2, pp. 122–137, 2017.
- [23] P. S. Hammerman, D. N. Hayes, and J. R. Grandis, “Therapeutic insights from genomic studies of head and neck squamous cell carcinomas,” *Cancer Discovery*, vol. 5, no. 3, pp. 239–244, 2015.
- [24] L. Tonella, M. Giannoccaro, S. Alfieri, S. Canevari, and L. De Cecco, “Gene expression signatures for head and neck cancer patient stratification: are results ready for clinical application?” *Current Treatment Options in Oncology*, vol. 18, no. 5, p. 32, 2017.
- [25] C. R. Leemans, P. J. F. Snijders, and R. H. Brakenhoff, “The molecular landscape of head and neck cancer,” *Nature Reviews Cancer*, vol. 18, no. 5, pp. 269–282, 2018.
- [26] C. Ndiaye, M. Mena, L. Alemany et al., “HPV DNA, E6/E7 mRNA, and p16INK4a detection in head and neck cancers: a systematic review and meta-analysis,” *The Lancet Oncology*, vol. 15, no. 12, pp. 1319–1331, 2014.
- [27] M. Lechner, A. R. Chakravarthy, V. Walter et al., “Frequent HPV-independent p16/INK4A overexpression in head and neck cancer,” *Oral Oncology*, vol. 83, pp. 32–37, 2018.
- [28] C. Doll, C. Steffen, B. Beck-Broichsitter et al., “The prognostic significance of p16 and its role as a surrogate marker for human papilloma virus in oral squamous cell carcinoma: an analysis of 281 cases,” *Anticancer Research*, vol. 42, no. 5, pp. 2405–2413, 2022.
- [29] V. Shyamsundar, S. V. Thangaraj, A. Krishnamurthy et al., “Exome sequencing with validations and expression of p16/cdkn2a shows no association with HPV in oral cancers,” *Asian Pacific Journal of Cancer Prevention*, vol. 23, no. 1, pp. 191–200, 2022.

Research Article

Identification of Immune and Hypoxia Risk Classifier to Estimate Immune Microenvironment and Prognosis in Cervical Cancer

Yujing Shi ¹, Qing Gao ², Zeyuan Liu ³, Gefenqiang Shen ², Xinchen Sun ^{1,2}
and Xiaoke Di ²

¹Department of Oncology, Jurong People's Hospital, Huayang Town, Jurong City, China

²Department of Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

³Department of Radiation Oncology, Nanjing Jiangning Hospital and the Affiliated Jiangning Hospital of Nanjing Medical University, Nanjing, China

Correspondence should be addressed to Xinchen Sun; sunxinchen@njmu.edu.cn and Xiaoke Di; dixiaoke2017@163.com

Received 13 August 2022; Revised 4 September 2022; Accepted 1 October 2022; Published 17 October 2022

Academic Editor: Song Cao

Copyright © 2022 Yujing Shi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Purpose. Cervical cancer (CC) is one of the most common gynecologic neoplasms. Hypoxia is an essential trigger for activating immunosuppressive activity and initiating malignant tumors. However, the determination of the role of immunity and hypoxia on the clinical outcome of CC patients remains unclear. **Methods.** The CC independent cohort were collected from TCGA database. Consensus cluster analysis was employed to determine a molecular subtype based on immune and hypoxia gene sets. Cox relevant analyses were utilized to set up a risk classifier for prognosis assessment. The underlying pathways of classifier genes were detected by GSEA. Moreover, we conducted CIBERSORT algorithm to mirror the immune status of CC samples. **Results.** We observed two cluster related to immune and hypoxia status and found the significant difference in outcome of patients between the two clusters. A total of 251 candidate genes were extracted from the two clusters and enrolled into Cox relevant analyses. Then, seven hub genes (CCL20, CXCL2, ITGA5, PLOD2, PTGS2, TGFBI, and VEGFA) were selected to create an immune and hypoxia-based risk classifier (IHBRC). The IHBRC can precisely distinguish patient risk and estimate clinical outcomes. In addition, IHBRC was closely bound up with tumor associated pathways such as hypoxia, P53 signaling and TGF β signaling. IHBRC was also tightly associated with numerous types of immunocytes. **Conclusion.** This academic research revealed that IHBRC can be served as predictor for prognosis assessment and cancer treatment estimation in CC.

1. Introduction

Cervical cancer (CC) is the fourth most frequently diagnosed cancer and the second mortal cancer in female population, which poses a serious health threat to women globally [1]. According to the GLOBOCAN 2020 database, there were 604127 new cases and 341831 new deaths from CC, and the death rate is 12.4 versus 5.2 per 100,000 people in transitioning and in transited countries, respectively [2]. Etiologically, accumulating evidence has implied that infection with high-risk human papillomavirus (HPV) is the primary factor for CC [3]. Up to 90% of cases are driven by high-risk HPV strains including 16, 18, 31, 33, and 35, with other low-risk HPV types generally produce benign cervical lesions [4, 5]. Despite the

promotion of HPV vaccine immunoprevention, many patients are diagnosed with advanced stage at their first diagnosis, making the exploration of early diagnosis biomarkers and effective prognostic model urgently needed [6, 7].

Recently, tumor microenvironment (TME) is causing general interest in various cancer settings. TME is composed of multiple cells residing in cancers, including immune cells, fibroblasts, endothelial cells, and mesenchymal cells [8]. These cells closely interact with each other and organize into distinct cellular communities [9]. Distinct immune cell response categories tumors into 3 types named “hot”, “altered”, and “cold” tumors [10]. Accumulating evidence has identified the immunotherapy as a promising intervention for cancer patients [11]. By reprogramming the immunosuppressive state in the

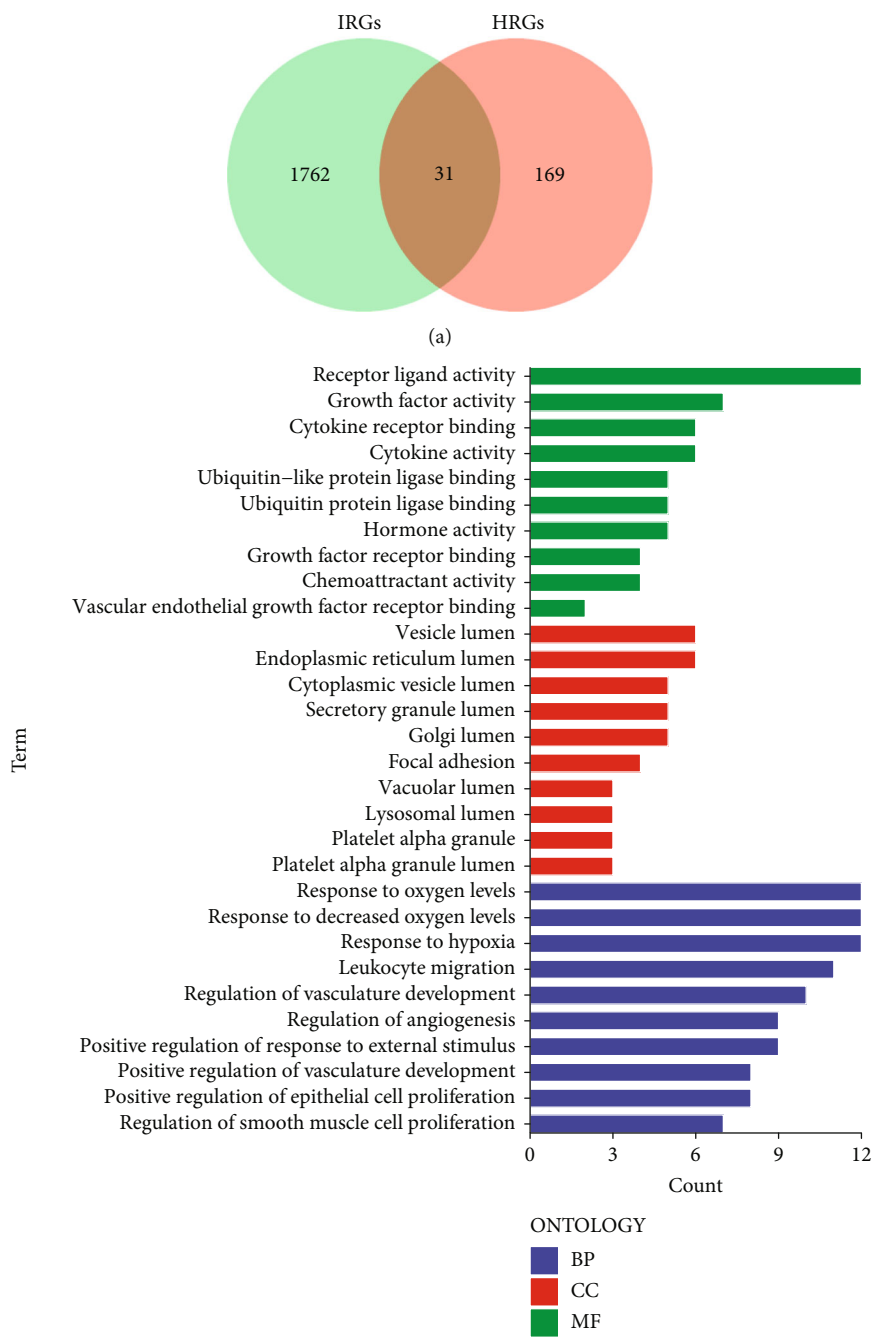


FIGURE 1: Continued.

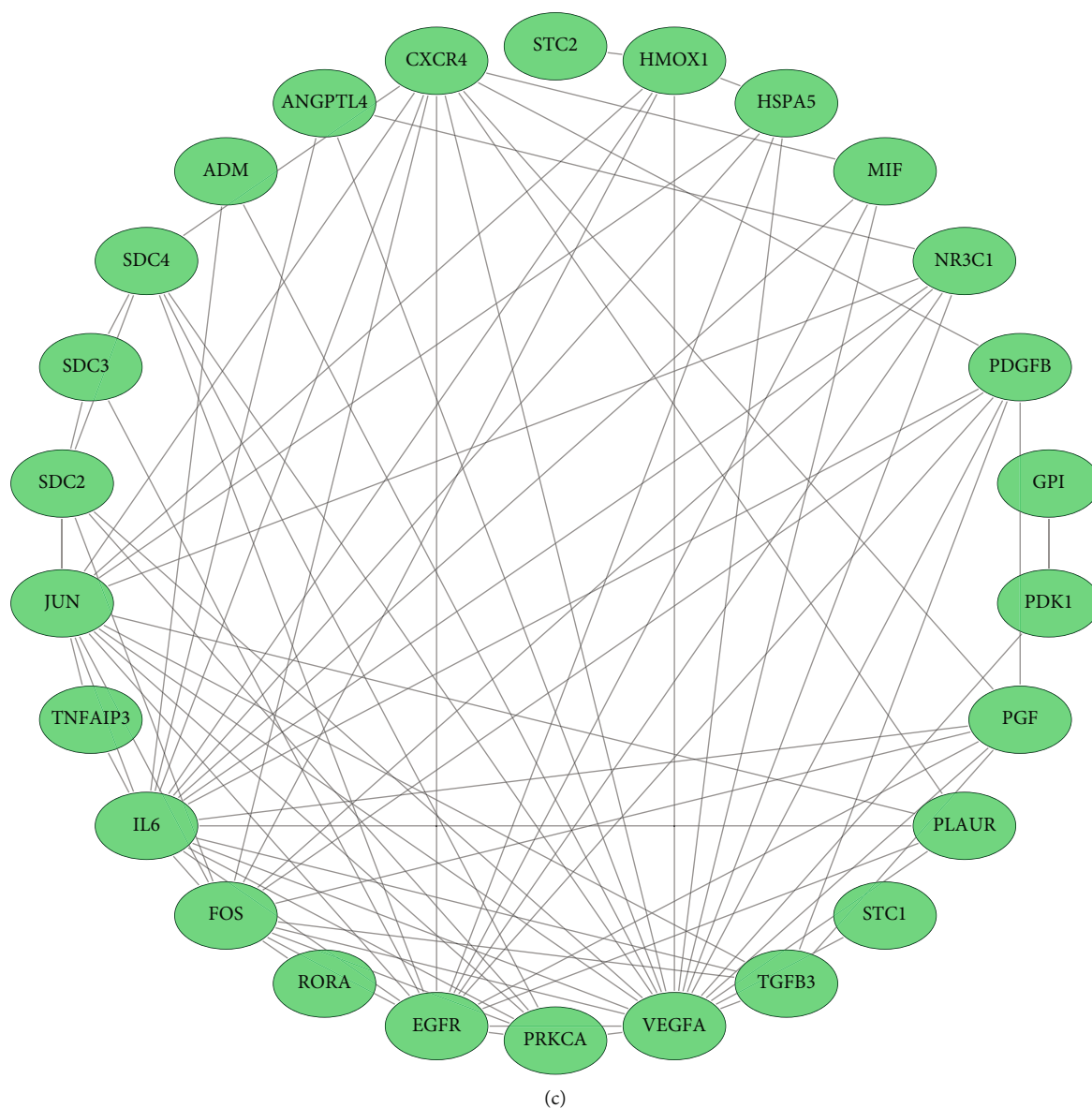


FIGURE 1: Characterization of immune and hypoxia genes. (a) The Venn plot of overlapped genes; (b) GO function enrichment analysis; (c) the PPI network of the overlapped genes.

“cold tumor” into an activated one, the usage of immune checkpoint inhibitors as well as some cell-specific compounds has achieved exciting clinic outcome in multiple cancers [12, 13]. However, immunotherapy in CC remains largely unexplored.

Hypoxia is one of common characteristics of tumors and is closely related to tumor progression and poor prognosis [14, 15]. Cells respond to hypoxia environment by regulating various metabolism pathways, which subsequently causes deficient hypervascularization, enhanced tumor cell proliferation, and distant metastasis tendency [16–18]. Emerging evidence has validated the crosstalk between hypoxia and immunophenotype in tumors. For instance, HIF2 α has been reported to exert its protective role in pancreatic ductal adenocarcinoma by improving immune responses [19]. Moreover, Zhang et al. once reported that hypoxia condition

elevated the tumor cell resistance to cytotoxic T lymphocytes mediated lysis, which is dependent on the upregulation of HIF1 α and PD-1 expression [20]. Taken together, it is reasonable to speculate that novel approaches targeting alleviating hypoxia condition could augment the current outcome for CC patients.

Most of the indicators proposed in previous studies to predict clinical outcomes of CC patients are limited to single genes, such as HPV, PTEN, and FHIT [21]. However, using only a single biomarker to assess prognosis is greatly partial, as the mechanisms affecting the development of CC are extremely complex. Currently, prognostic signature consisting of multiple genes has been proven to present independent prognostic ability by several reports, which has also attracted the attention of scholars in the field of oncology [22, 23]. Compared to the traditional TNM system, the

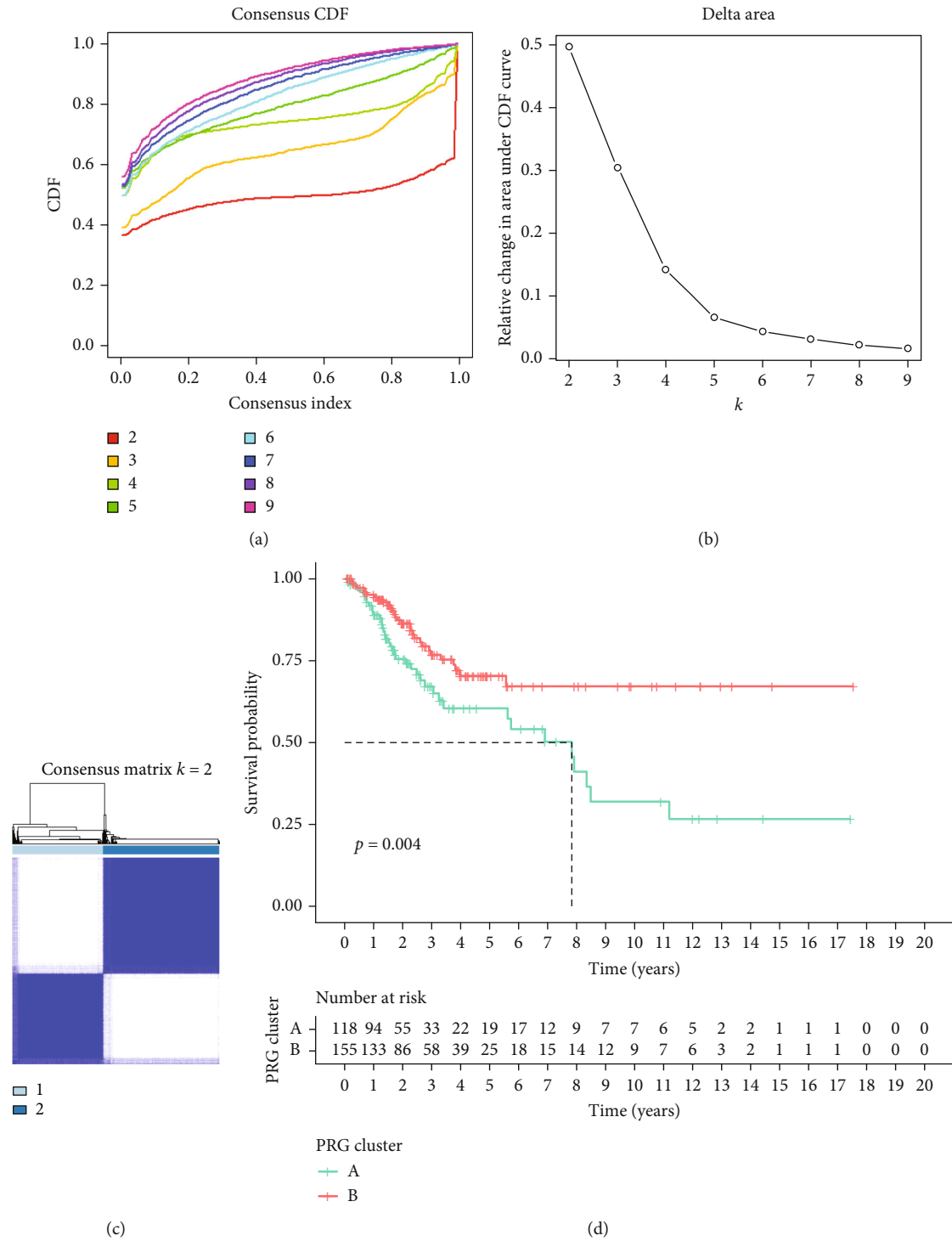


FIGURE 2: Continued.

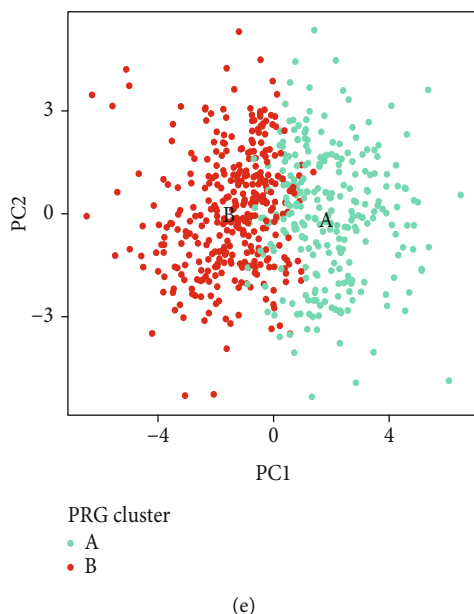


FIGURE 2: Consensus clustering determined a molecular subtype related to immune and hypoxia. (a) The CDF score of consensus index; (b) relative change of CDF curve; (c) consensus matrix for $k = 2$; (d) the Kaplan–Meier survival analysis; (e) principal component analysis of the two clusters.

prognostic model is capable of accurately predicting not only clinical outcomes but also the patient’s immune status and treatment benefits.

The alteration of metabolic state and immunophenotype in tumors largely restrain the therapy response for CC patients, while the relevant study is still in very early stages. In our current research, we combined the immune-related genes (IRGs) and hypoxia-related genes (HRGs) to establish a prognostic signature with high accuracy for CC. In addition, immune cell infiltration analysis was performed in two risk groups of CC samples. Altogether, our exploration will help clarify the specific immune environment in different populations and provide new ideas and insights for the prevention and treatment of CC.

2. Methods

2.1. Data Acquisition. The TCGA-CSCC dataset containing gene expression and simple nucleotide variation was collected from TCGA website (<https://cancergenome.nih.gov/>). And the clinical data of TCGA-CSCC dataset was obtained from cBioPortal website (<http://www.cbioportal.org/>). Next, we combined the clinical traits of the two databases by patient ID. The exclusion criteria were set as follows: (1) histologic diagnosis is not CC; (2) samples without completed data for analysis; and (3) survival time of less than 30 days. Moreover, we extracted IRGs from ImmPort database (<https://www.immport.org/shared/genelists/>) and collected HRGs from MSigDB website (<https://www.gsea-msigdb.org/gsea/msigdb/>, Supplementary Table 1).

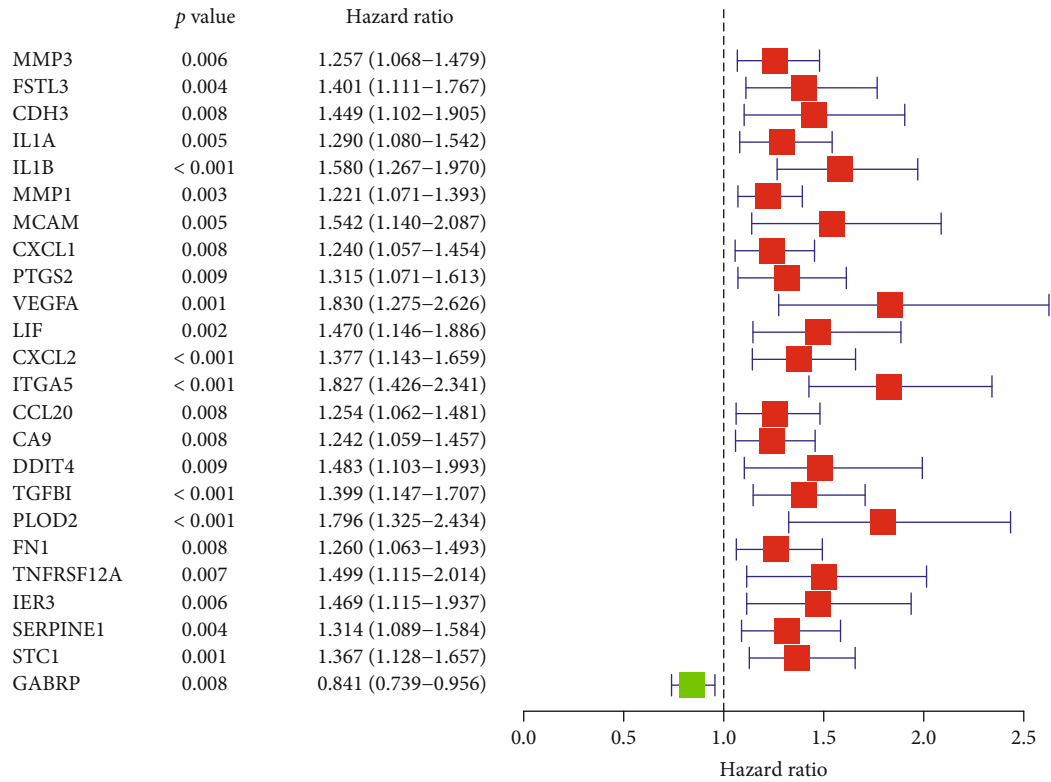
2.2. Gene Cluster Analysis. The consensus cluster algorithm was performed using the “ConsensusClusterPlus” package [24]. To determine the optimal cluster score, we assessed the

Delta area and cumulative distribution function (CDF). Next, we compared clinical outcome discrepancies between different subtypes by survival analysis. We also utilized differential analysis to screen differentially expressed genes (DEGs) between different subtypes for subsequent analysis [25].

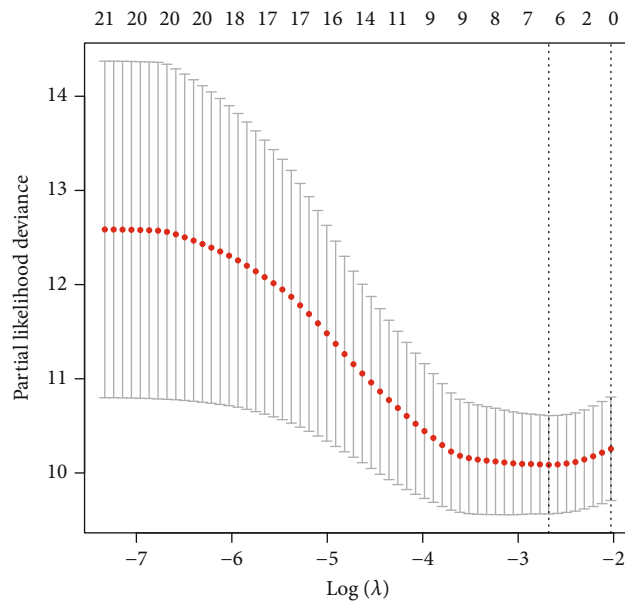
2.3. Development of a Risk Classifier. All CC samples were randomly divided into training set and validation set. The DEGs from cluster analysis were first subject to univariate analysis. Then, we enrolled the potential genes with prognostic value in multivariate analysis. Finally, we created immune- and hypoxia-based risk classifier (IHBRC) according to regression coefficients of each model factors. The risk equation is as follows: risk factor = $\sum_{i=1}^n (Coef_i \times Exp_i)$; $Coef_i$ is the coefficient of the classifier generated by Cox analyses, and Exp_i is the expression level of each model genes. The patients were divided into high- and low-risk groups according to the median risk score.

2.4. Survival Analysis. The differences in clinical outcome were detected between two risk groups by Kaplan–Meier analysis. ROC curves were plotted to test the reliability of IHBRC in assessing patients’ outcomes. Univariate and multivariate analyses were applied to confirm the independent value of IHBRC in CC.

2.5. Gene Set Enrichment Analysis (GSEA). The transcriptome data and risk groups information were enrolled into GSEA [26]. Next, we selected the hallmark, all v7. 5. symbols. Gmt in the MSigDB database as the reference gene set. The default weighted enrichment method was applied for 1000 enrichment analysis. The gene sets with $p < 0.05$ and $FDR < 0.25$ were considered as significantly enriched gene sets.



(a)



(b)

FIGURE 3: Continued.

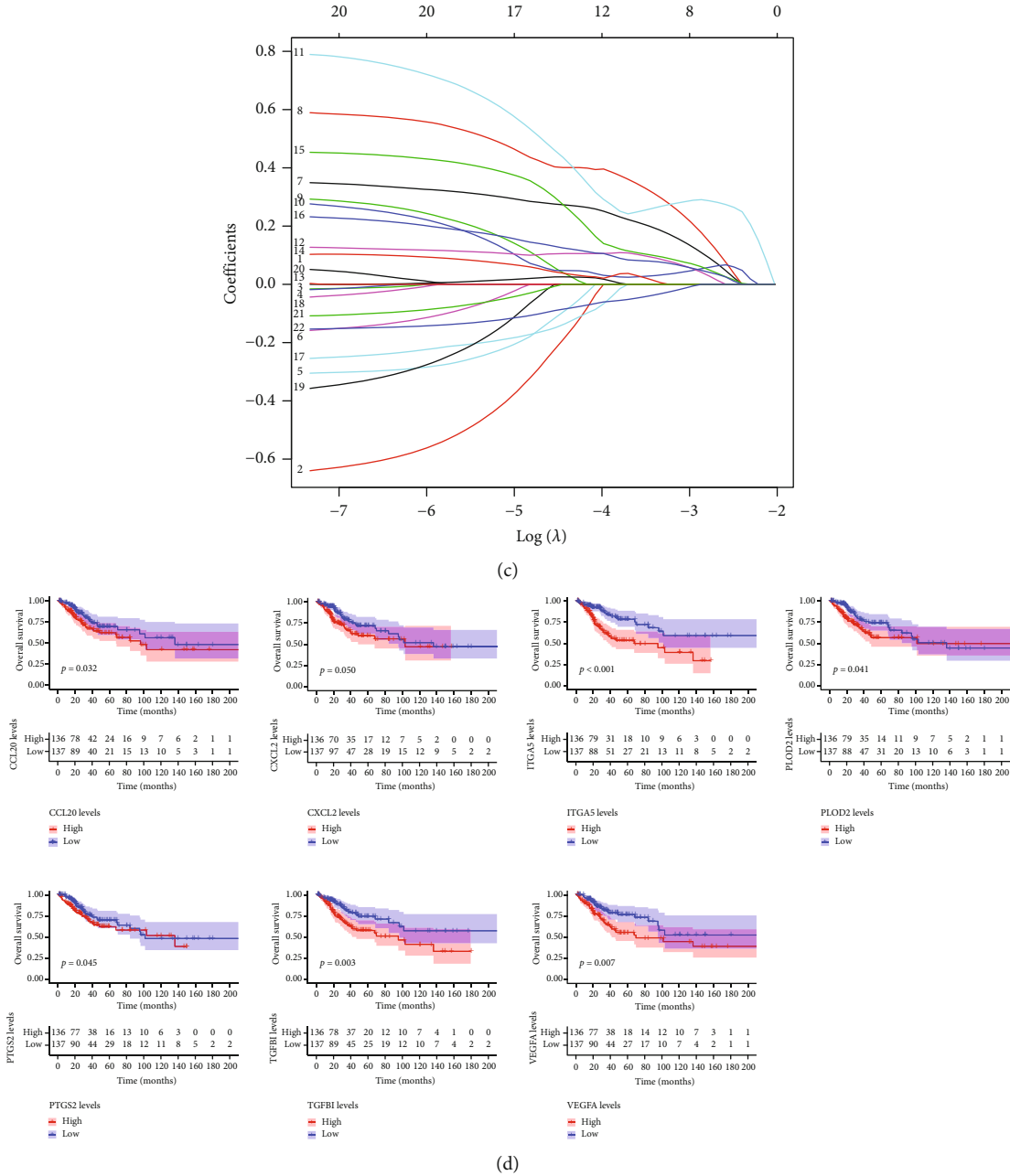


FIGURE 3: Construction of a risk classifier. (a) Univariate Cox regression analysis; (b–c) LASSO coefficients for risk classifier; (d) the survival analysis of classifier genes.

TABLE 1: Multivariate analysis of the seven model genes in CC.

| Gene | Coefficient | P value |
|-------|-------------|---------|
| CCL20 | 0.0131 | 0.007 |
| CXCL2 | 0.0638 | 0.001 |
| ITGA5 | 0.2812 | 0.001 |
| PLOD2 | 0.0340 | 0.001 |
| PTGS2 | 0.0697 | 0.008 |
| TGFBI | 0.0374 | 0.001 |
| VEGFA | 0.1113 | 0.001 |

2.6. Immune Infiltration Analysis. CIBERSORT is a powerful algorithm proposed by Newman et al. to mirror the infiltration status of immunocytes [27]. Performing an immunocytes gene set including 547 genes, CIBERSORT was applied to determine 22 immunocyte types containing B cells (naive B cells and memory B cells), T cells (CD8 T cells, naive CD4 T cells, resting memory CD4 T cells, activated memory CD4 T cells, follicular helper T cells), immunosuppressive cells (T cells regulatory (Tregs), M2 macrophages and eosinophils) as well as other cells (resting NK cells, activated NK cells, monocytes, macrophages, dendritic cells,

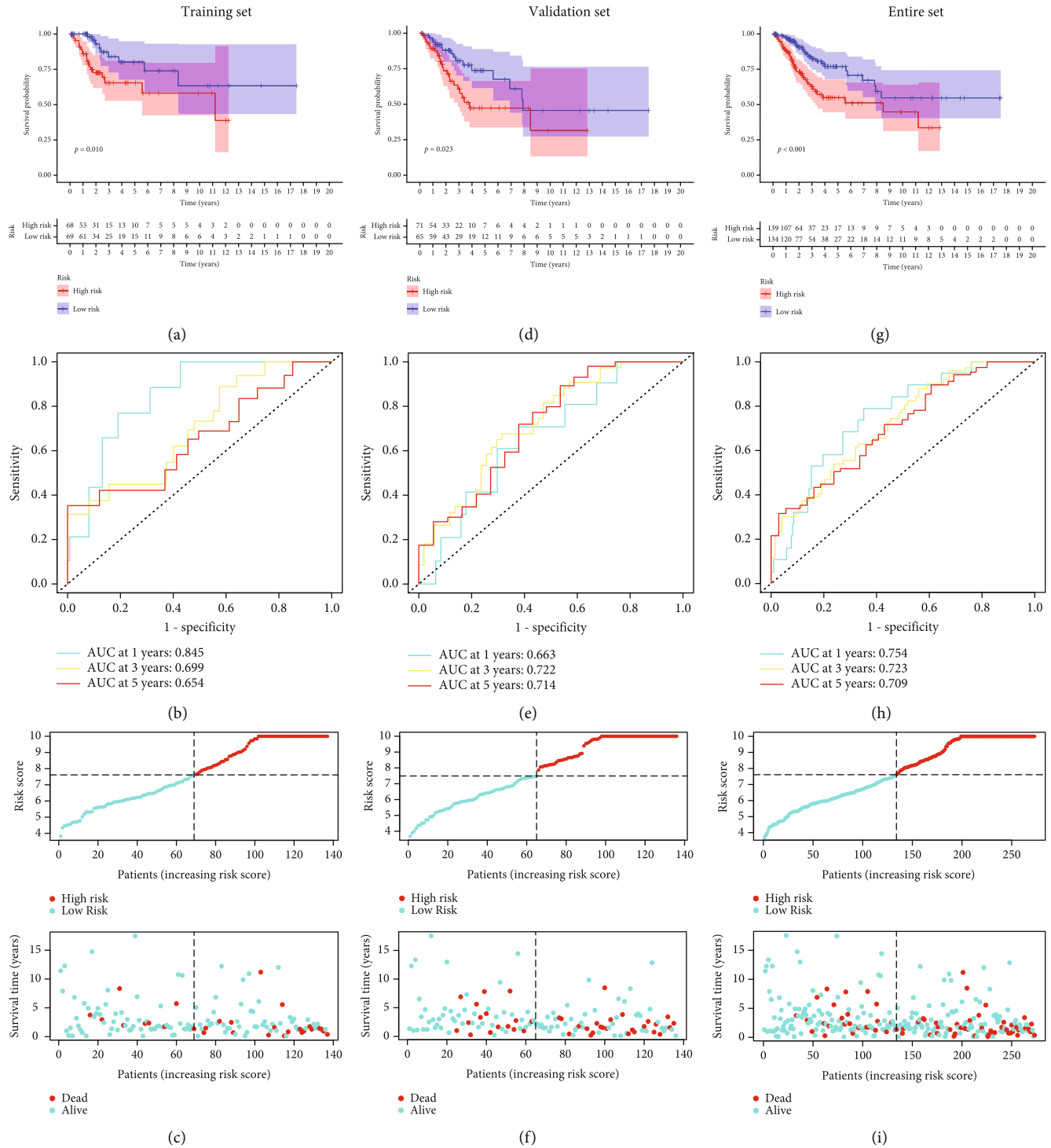


FIGURE 4: Predictive value of the classifier. (a) Survival curves of prognostic difference between two risk groups in the training set; (b) ROC curve of the assessment reliability of the classifier in the training set; (c) the distribution of risk score and survival status in the training set. (d–f) and (g–i) the testing set and the entire set were used to confirm the predictive value of the classifier.

most cells, eosinophils, and plasma cells). To detect the TME of CC cases, we conducted correlation analysis to analyze the relationship between risk score and 22 immunocytes types.

2.7. Tumor Mutation Burden Analysis. We employed the mutation data of CC cases to compare the tumor mutation

burden (TMB) in two subgroups. The TMB value was generated using following equation: $TMB = (\text{total mutation}/\text{total coverbased}) \times 10^6$.

2.8. Chemotherapy Drug Sensitivity Analysis. To estimate the predictive power of the IHBC for chemotherapeutic drug

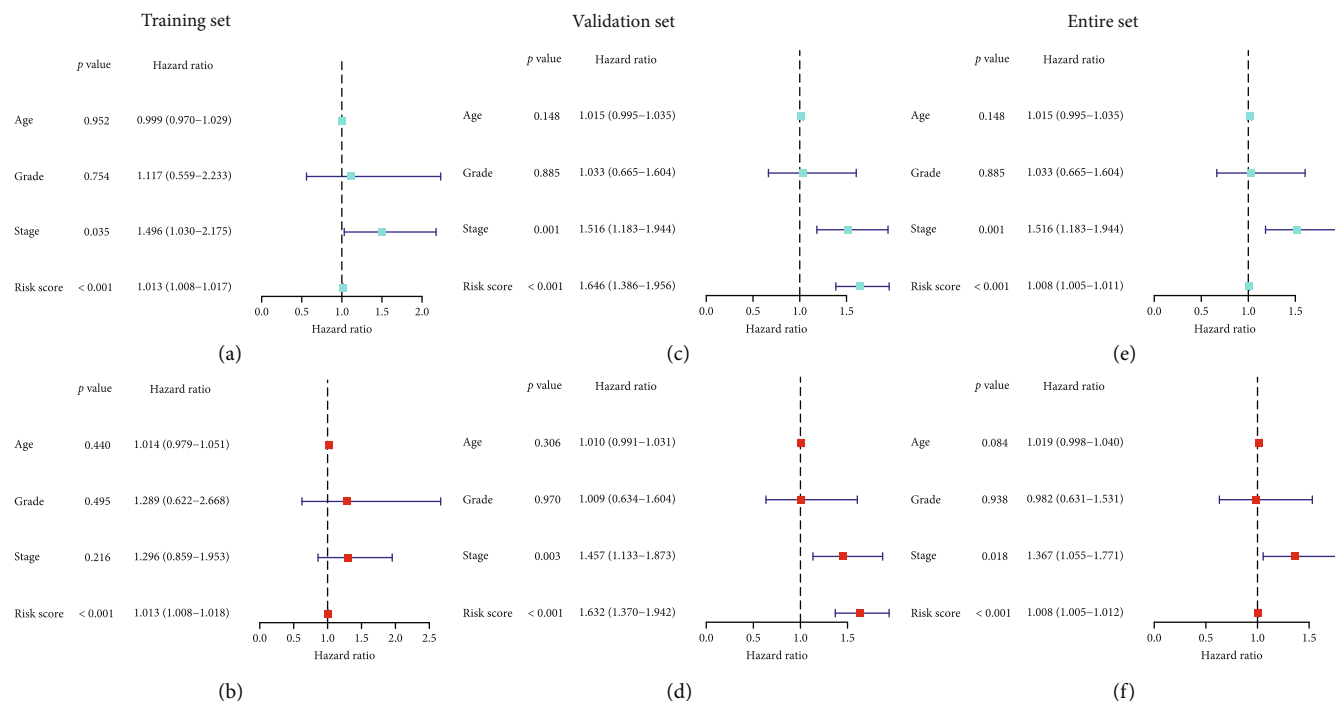


FIGURE 5: Independent prognosis analysis of the classifier. (a–c) Univariate Cox regression analysis; (d–f) multivariate Cox regression analysis.

efficacy, the half-maximal inhibitory concentration (IC₅₀) was taken as an index to measure the drug sensitivity. The difference in the IC₅₀ between two risk groups was compared by pRRophetic of R.

2.9. Identification of the Target miRNAs. To explore the target miRNAs of model genes, a prediction approach with starBase (<http://starbase.sysu.edu.cn/>) was conducted. The criteria for determination was set by five prediction programs.

3. Results

3.1. Characterization of Immune and Hypoxia Genes. To discover the hub genes which could regulate both immunity and hypoxia process, we screened 31 overlapped genes by intersection of IRGs and HRGs lists (Figure 1(a)). Then, we performed function analysis on these 31 genes and found that they were enriched in response to hypoxia, leukocyte migration, and regulation of angiogenesis (Figure 1(b)). Meanwhile, we created a PPI network to better clarify the interaction of 31 genes at protein level (Figure 1(c)).

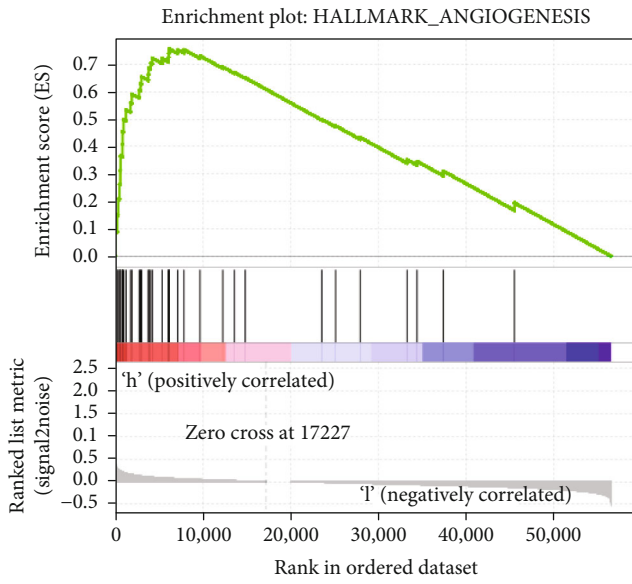
3.2. Consensus Cluster Analysis. A total of 31 hub genes were incorporated into cluster analysis. The results indicated that CDF value growth was flat when $k=2$ and Delta area increased insignificantly at $k>3$ (Figures 2(a) and 2(b)). The fractal matrix showed the favorable intergroup difference and intragroup association, suggesting these pivot genes could categorize all CC samples into two subtypes (Cluster 1 ($n=130$) and Cluster 2 ($n=174$)). Therefore, the clustering stability was best for $k=2$ (Figure 2(c)). Sur-

vival analysis illustrated the significant difference in patient outcome between two clusters (Figure 2(d)). PCA analysis uncovered the favorable distinction between the two clusters (Figure 2(e)). Furthermore, 251 DEGs were collected from differential analysis between two clusters.

3.3. Development of a Risk Classifier. In the training set, we first determined 24 survival-associated indicators based on above 251 DEGs via univariate analysis (Figure 3(a)). Then, the candidate genes were enrolled into LASSO regression to remove the over fitting genes (Figures 3(b) and 3(c)). Finally, multivariate analysis was employed, and seven hub genes were selected to develop an IHBRC (Table 1): risk score = $(0.0131 \times \text{CCL20}) + (0.0638 \times \text{CXCL2}) + (0.2812 \times \text{ITGA5}) + (0.0340 \times \text{PLOD2}) + (0.0697 \times \text{PTGS2}) + (0.0374 \times \text{TGFBI}) + (0.1113 \times \text{VEGFA})$. In addition, Figure 3(d) demonstrated the prognostic power of seven hub predictors.

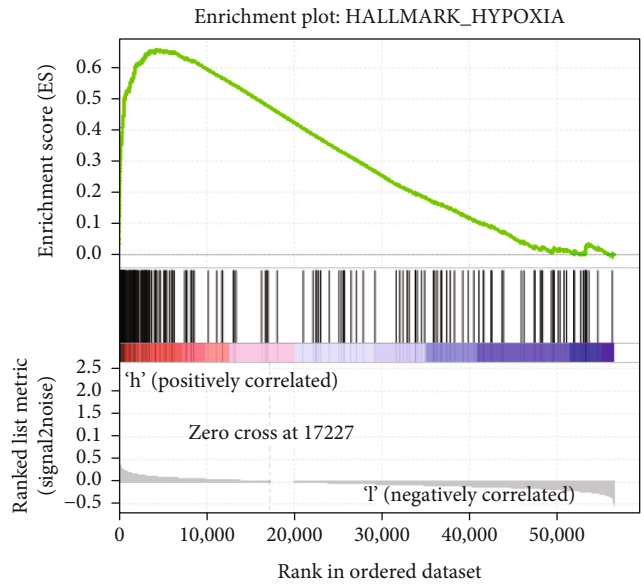
As suggested by Figure 4(a), high-risk group presented a dismal prognosis benefit in the training set. The AUC values of 1-, 3-, and 5-year survival were 0.845, 0.699, and 0.654, respectively (Figure 4(b)). We measured the survival outcome of patients in both groups and found that patients' outcomes were dismal as the risk score elevated (Figure 4(c)). Meanwhile, we confirmed the performance of IHBRC in the validation and the entire cohorts using the same analysis described above and obtained the same results for the trend (Figures 4(d)–4(i)).

3.4. Independent Prognostic Analysis. To examine the independent value of IHBRC in terms of survival of CC cases, univariate and multivariate analyses were employed. In the training set, univariate analysis demonstrated that low risk



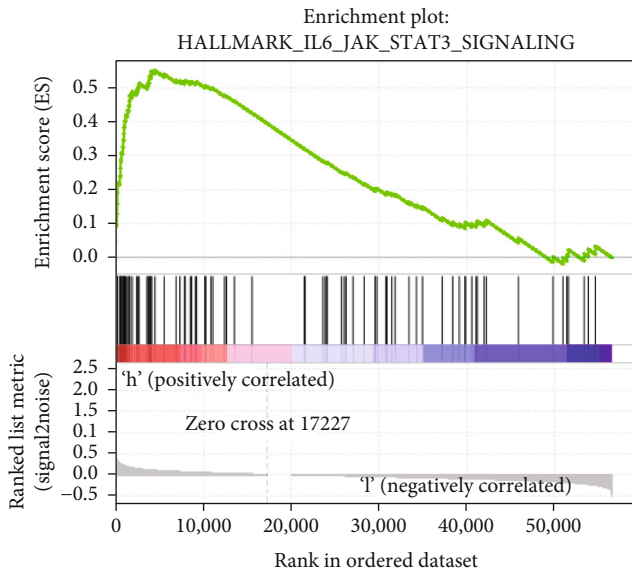
— Enrichment profile
— Hits
— Ranking metric scores

(a)



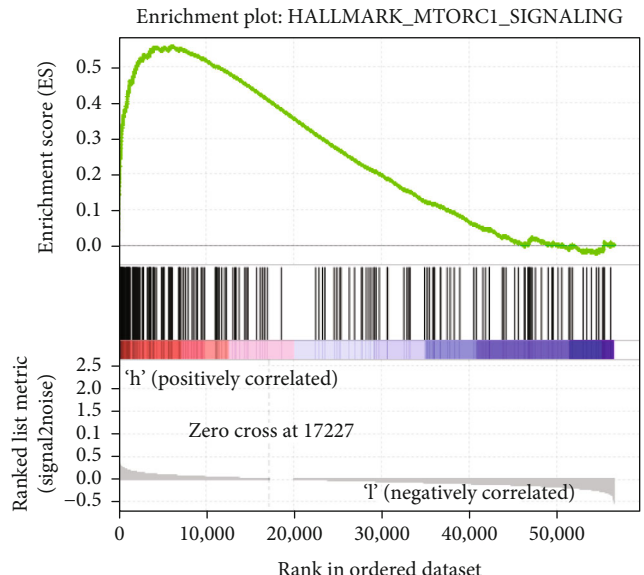
— Enrichment profile
— Hits
— Ranking metric scores

(b)



— Enrichment profile
— Hits
— Ranking metric scores

(c)



— Enrichment profile
— Hits
— Ranking metric scores

(d)

FIGURE 6: Continued.

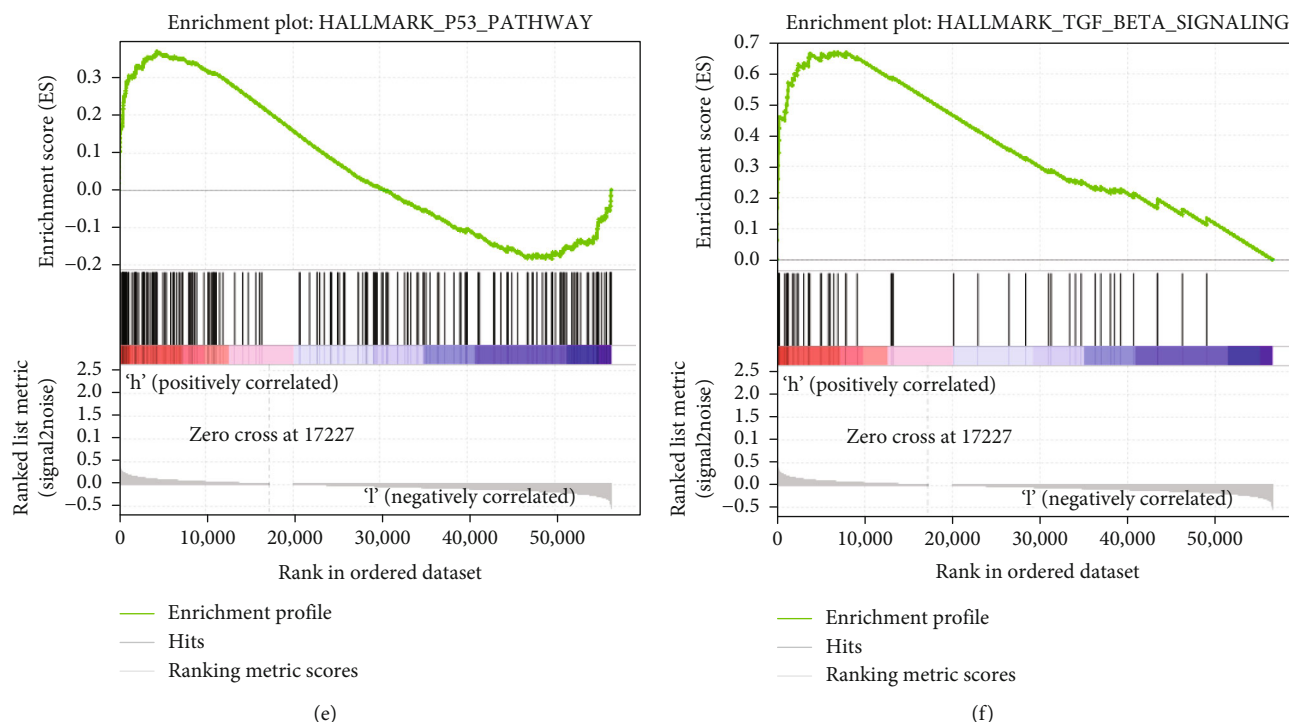


FIGURE 6: Gene set enrichment analysis. (a) Angiogenesis; (b) hypoxia; (c) IL6/JAK/STAT3 signaling; (d) MTORC1 signaling; (e) P53 signaling; (f) TGF β signaling.

score was remarkably correlated with favorable prognosis (Figure 5(a)). Furthermore, multivariate analysis still revealed that low risk score was independently associated with favorable outcome of CC patients (Figure 5(b)), which could serve as an independent prognostic factor for glioma. These were confirmed by the test and the entire sets (Figures 5(c)–5(f)).

3.5. GSEA Enrichment Analysis. To explore the distinction in molecular pathways between the two groups, we applied GSEA based on hallmarks gene sets. The results disclosed that hallmarks including angiogenesis, hypoxia, IL6/JAK/STAT3 signaling, MTORC1 signaling, P53 signaling, and TGF β signaling were markedly enriched in high-risk group (Figure 6).

3.6. Immune Infiltration Analysis. In order to mirror the immune status of two groups, we estimated enrichment value of different immunocytes. Figure 7(a) illustrated the relationship between seven model biomarkers and immunocytes. As shown in Figure 7(b), risk score was negatively correlated with the infiltration level of memory B cells, naïve B cells, resting dendritic cells, and macrophages M1 and CD8 T cells, while neutrophils were activated in IHBRC-high cohort.

3.7. Immune Checkpoints Analysis for Risk Classifier. Subsequently, we detected the relationship between signature and the expression of immune checkpoints. Figure 8(a) revealed six immune checkpoints that were greatly differentially expressed in the two risk groups. As suggested by

Figure 8(b), BTLA was significantly downregulated in the high-risk group, while PDL2, ICAM1, CCL2, IL10, and TGFBI were markedly enriched in the high-risk group, indicating that patients with high risk are likely to be immunosuppressive status.

3.8. Analysis of Immunotherapy and Chemotherapy Response. Waterfall diagrams indicated the mutational differences in the 20 genes between the two groups. We observed that the IHBRC-high cohort had a higher PIK3CA mutation rate than the IHBRC-low group (31 vs. 20%), (Figures 9(a) and 9(b)). Given the importance of TMB in evaluating immunotherapy response for patients with CC, we observed IHBRC-high group had lower TMB value (Figure 9(c)). In addition, high risk score was correlated with a lower IC50 of docetaxel, doxorubicin, and gemcitabine ($p < 0.05$), suggesting that the IHBRC served as a favorable indicator for chemosensitivity (Figures 9(d)–9(f)).

3.9. Construction of IHBRC-Related Regulatory Network. The reciprocal regulation of mRNA and miRNA is closely bound up with tumor development. Based on the starbase online tool, we identified the target miRNAs of seven model genes with high relevance scores (Figure 10). Moreover, miRNA set enrichment analysis was performed to explore the function of the target miRNAs by TAM 2.0 tool. The results showed these miRNAs were mainly involved in cell aging, apoptosis, immune response, inflammation, and regulation of Stem Cell (Supplementary Table 2).

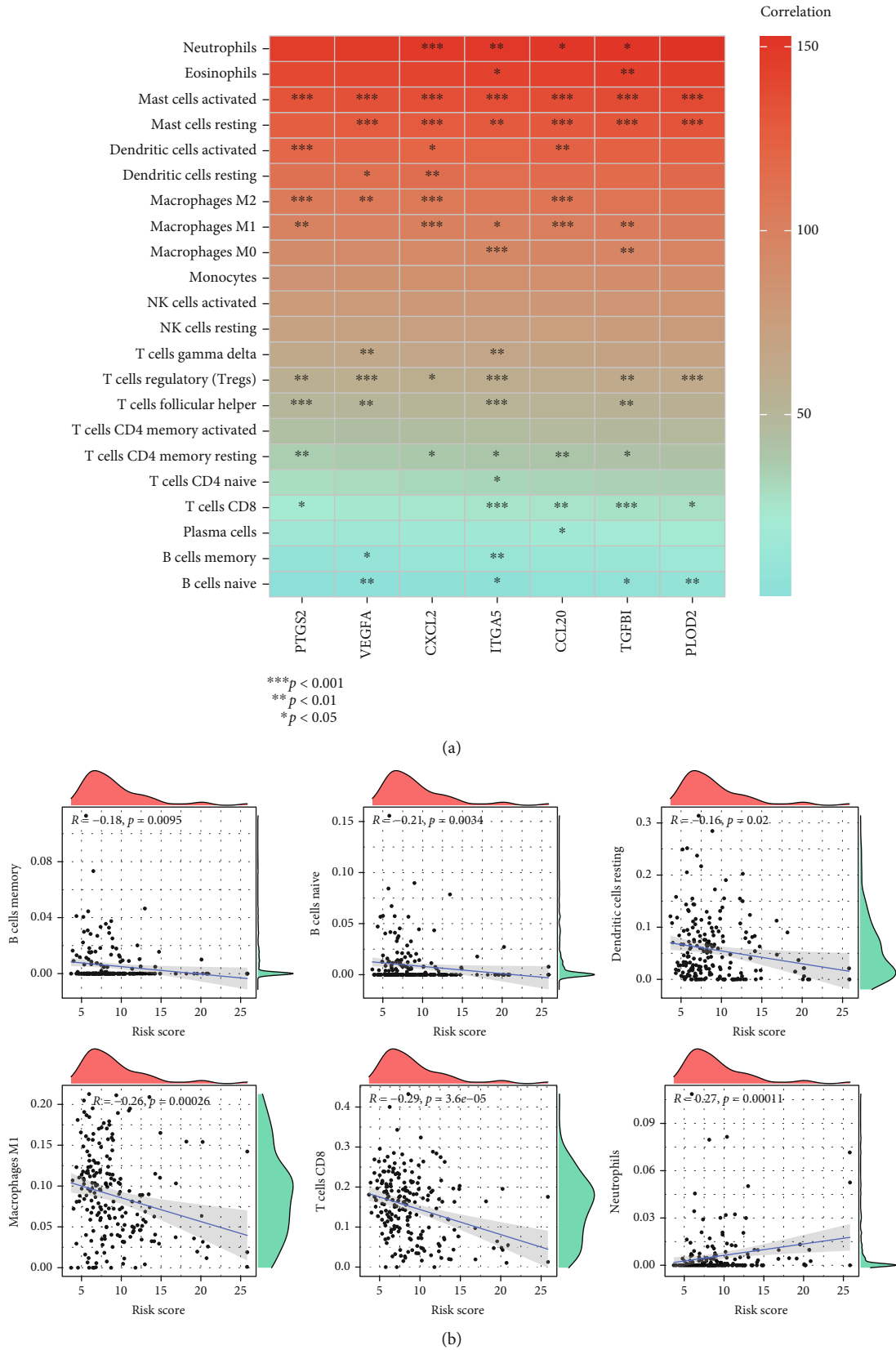


FIGURE 7: Immune infiltration analysis. (a) The relationship between seven model biomarkers and immunocytes; (b) correlation analysis of risk score and immunocytes (memory B cells, naïve B cells, resting dendritic cells, macrophages M1, CD8 T cells, and neutrophils).

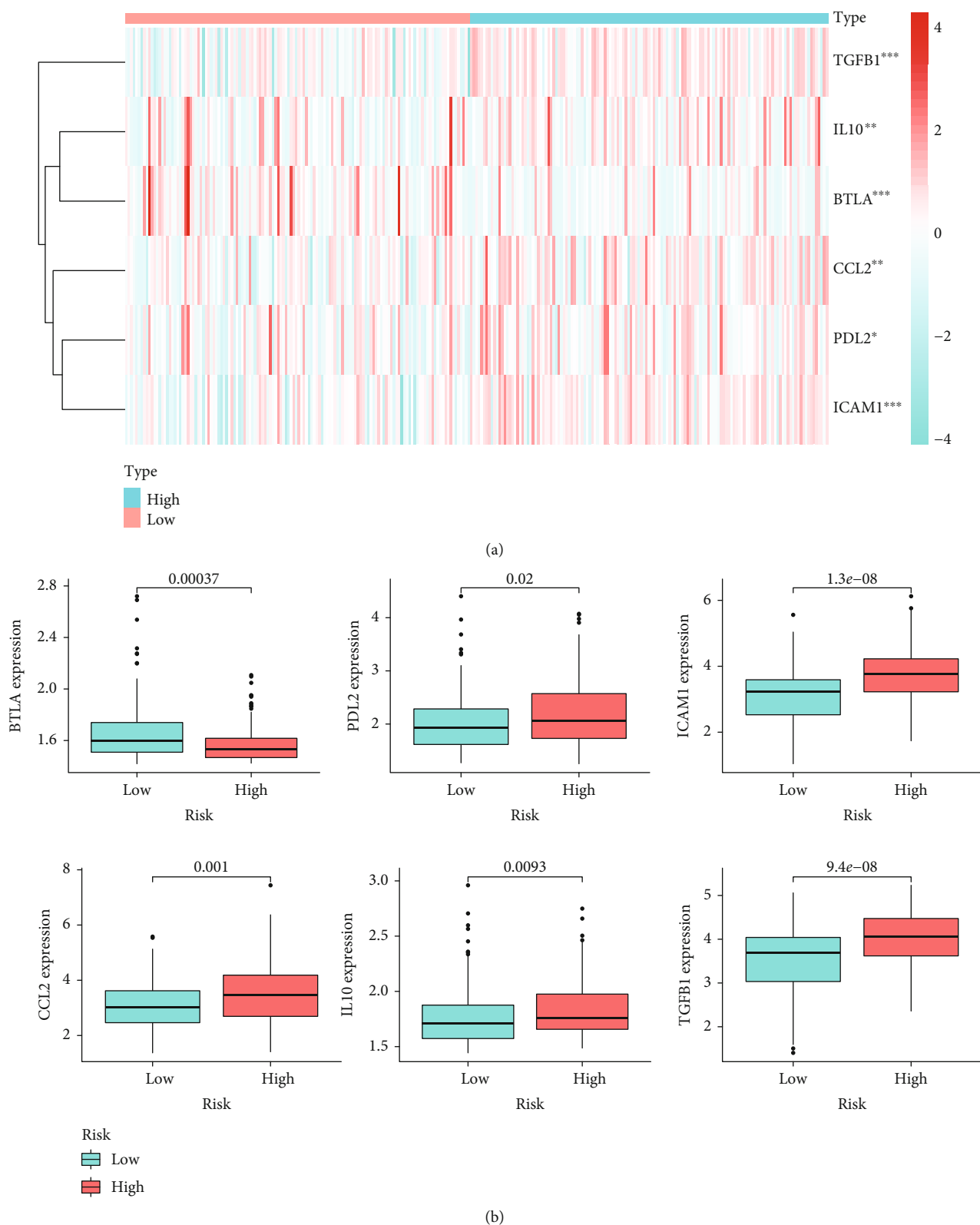
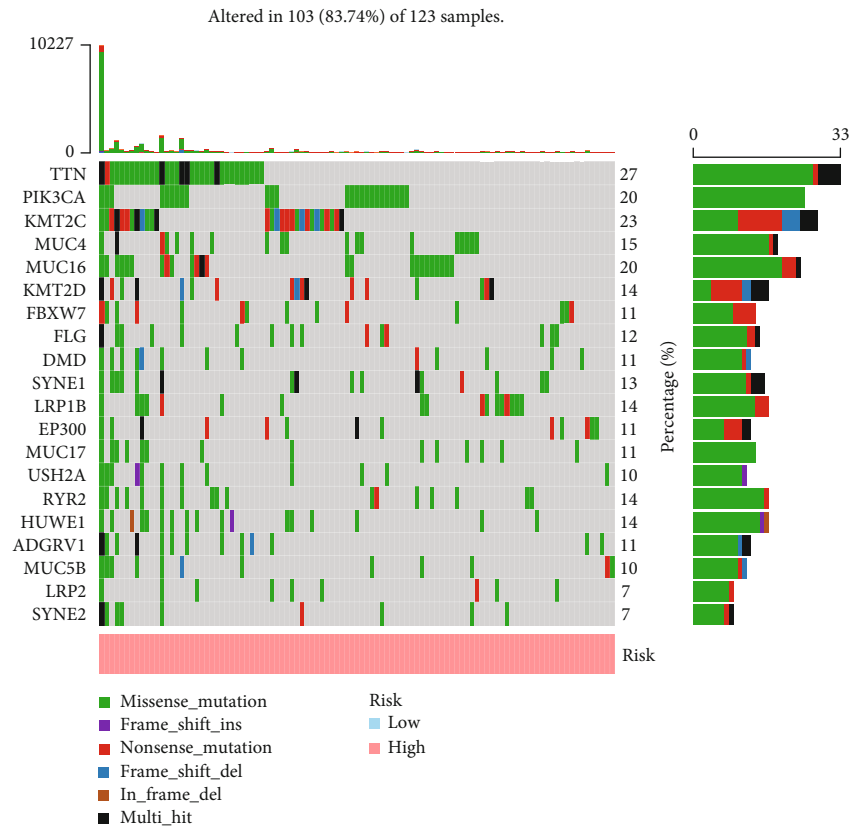
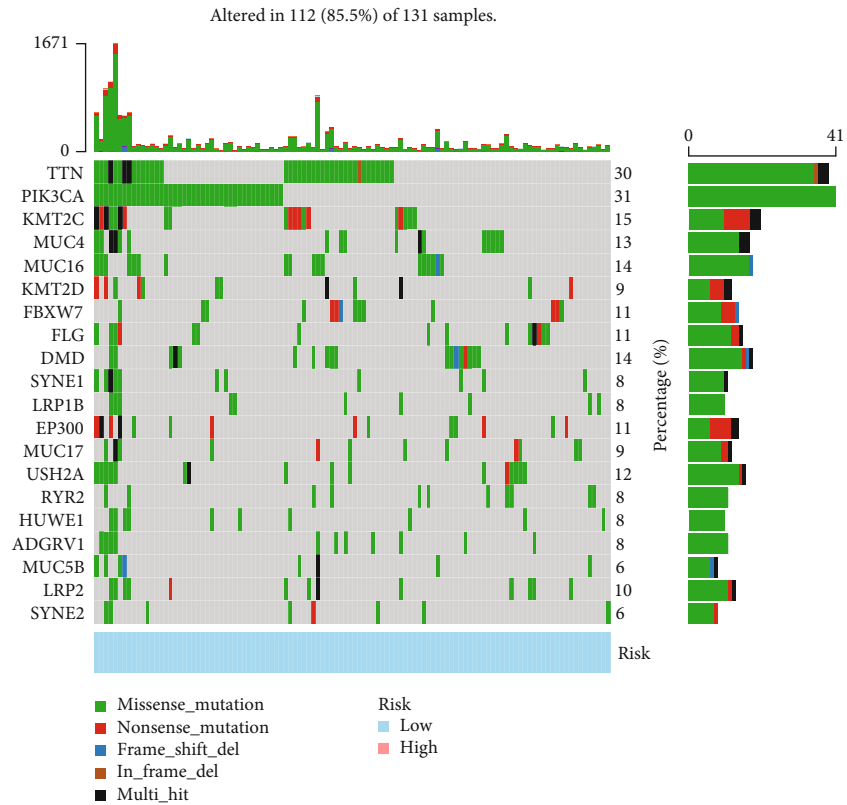


FIGURE 8: Correlation analysis of immune checkpoints and risk score. (a) Heatmap displaying expression patterns of immune checkpoints between two groups; (b) box line diagram showing the expression differences of six immune checkpoints between two groups.



(a)



(b)

FIGURE 9: Continued.

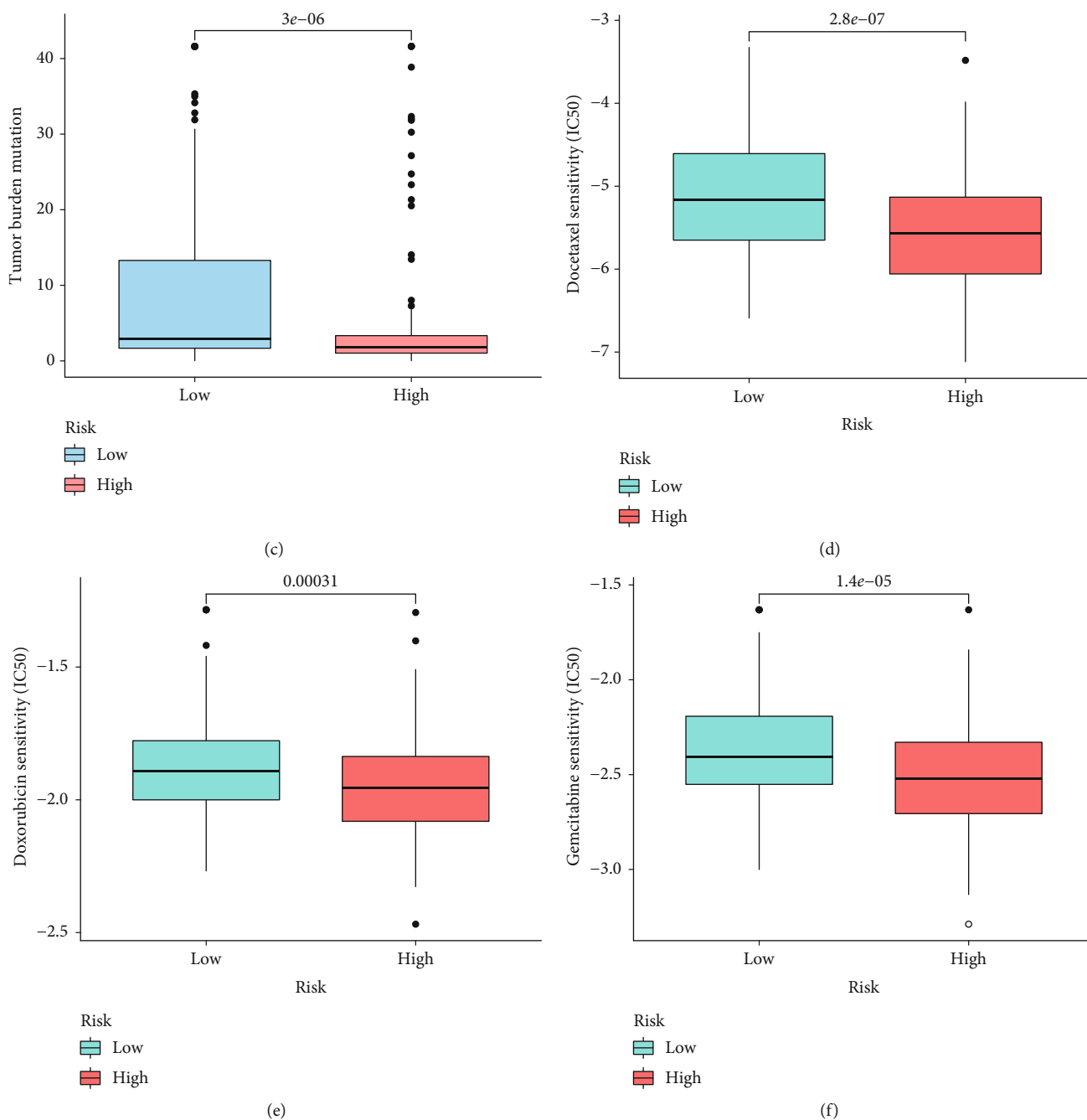


FIGURE 9: Analysis of immunotherapy and chemotherapy response. (a-b) The top 20 mutated genes in the two groups; (c) the TMB in the two groups; (d-f) chemotherapeutic response in the two groups.

4. Discussion

Antitumor effects of immune cells could be largely influenced by TME, including intercellular crosstalk between different cell types, chemokines concentrations, and metabolism environment, thus it is crucial to establish a comprehensive understanding on the genetic and population characteristics of TME. In our study, we categorize CC patients into two distinct clusters, in which they have totally differed prognosis, based on the expression level of immune- and hypoxia-related genes. Our proposed classifier is a favorable biomarker to assess the prognosis of CC cases.

Meanwhile, the classifier can serve as an indicator for predicting immune infiltration levels, TMB value and chemotherapy response, providing a novel insight for future research and clinical practice.

A total of hub seven genes (CCL20, CXCL2, ITGA5, PLOD2, PTGS2, TGFBI, and VEGFA) were identified as risky indicators in our prognostic model, and the involvement of some genes in CC has been reported before. PTGS2, also named COX-2, is a crucial target to prevent progression in various cancer types [28–30]. Early in 2004, Kulkarni et al. reported that the COX-2 expression was elevated in CC samples compared to normal cervical tissue. A number of

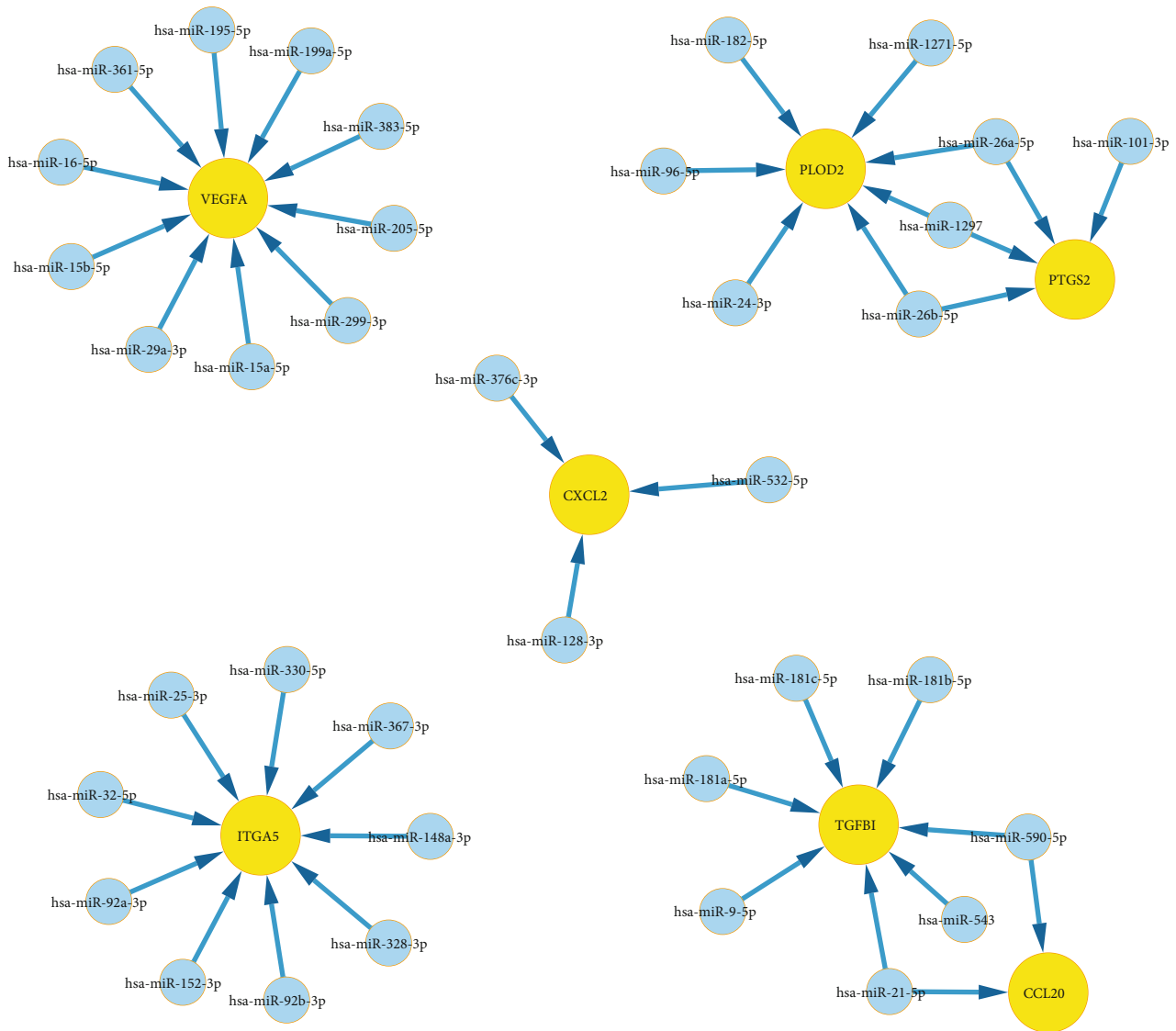


FIGURE 10: Construction of IHBRC-related regulatory network.

signaling including EGF and nuclear factor κ B (NF- κ B) pathway has been validated to mediate COX-2 expression in CC [31, 32]. Moreover, the usage of COX-2 selective inhibitors selectively enhances radio responsiveness in CC cell line under both normoxic and hypoxic conditions [33]. VEGFA is considered to play a crucial role in physiological and pathological angiogenesis [34]. In stimulation of VEGFA, endothelial cells proliferate and migrate to form new vessels [35]. The cross talk between VEGF signaling and immune response has been recently demonstrated. Briefly, VEGFA contributes to the polarization of macrophages into an M2 immunosuppressive phenotype [36–38]. In turn, these immunosuppressive cells can further produce proangiogenic factors including VEGFA and MMP9 [39]. The role of CXCL2 in CC has been intensively reported before. Zhang et al. once revealed that CXCL2 may promote tumor proliferation and metastasis induced by the overexpression of A-kinase-interacting protein 1 (AKIP1) in CC [40]. In agreement with our result, Yang et al. recently indi-

cated that the expression level CXCL2 is strongly associated with lymph node metastasis and prognosis in CC patients [41]. Four other genes including ITGA5, CCL20, TGFB1, and PLOD2 were previously studied in various malignancies, while their involvement in CC remains largely unexplored, and more basic researches are needed to reveal their biological function in CC [42–44].

As an endogenous noncoding RNA, miRNA could regulate 30% of protein-coding genes in human cells. Numerous studies have reported that miRNA is an upstream regulator of tumor-associated genes and engages in regulating biological processes such as proliferation and migration of cancer cells [45]. Our results revealed that hsa-miR-26a-5p, hsa-miR-26b-5p, hsa-miR-1297, hsa-miR-590-5p, and hsa-miR-21-5p were shared modulators of model genes. In cervical cancer, miR-590-5p was proven to facilitate tumor viability by inhibiting CHL1 [46]. Also, miR-590-5p could boost the malignant behaviors of liver cancer by interacting with FOXO1 [47]. Gu et al. disclosed that DUXAP8 could boost

cells growth and angiogenesis by targeting miR-1297 in CC [48]. Moreover, miR-21-5p also serves as an important factor regulating the effect of HAND2-AS1 on CC [49].

Molecular signaling was further analyzed in our research to unveil the mechanism underlying CC progression. In general, defective vasculatures and overweighing demands of oxygen contribute to the hypoxia environment in solid tumors [50]. HIFs induced by the hypoxic microenvironment play a central part in several aspects of tumor formation, especially in the regulation of tumor angiogenesis. HIF has a bidirectional regulatory effect on tumor angiogenesis. *In vitro* studies revealed that when HIF-1 α activity was inhibited, it had different effects on the expression of proangiogenic factors. VEGF, angiogenin, and TGF β -1 expressions were diminished, while IL-6 and MCP-1 were significantly increased. *In vivo* tests showed that RNA inhibition of HIF-1 α also showed a decrease in VEGF expression and an increase in IL-8 expression. Consequently, when HIF- α is inhibited, one proangiogenic factor may be increased when another proangiogenic factor is inhibited, and as a result, there may still be an actual increase in tumor vascularization [51, 52].

As a result, ATP production shifts from oxidative phosphorylation to glycolysis, and the acidic microenvironments subsequently confer the alterations of gene expression and activation of multiple molecular pathways, accelerating the cancer progression [53, 54]. The genetic alternations of mTOR protein have a significant role in tumorigenesis [55, 56]. A number of molecules are involved in the modulation of mTOR signaling, and specific inhibitors show a good performance in prevention and treatment of various tumors including oral cancer, ovarian cancer, and lung carcinoma [57–59]. TP53, which encodes a sequence-specific DNA-binding transcription factor, is one of the most frequently mutated genes in cancers [60]. Studies show that depletion of TP53 can remarkably increase the incidence of carcinogen-induced carcinogenesis and accelerate the tumor growth and invasiveness [61]. TGFB is a critical regulator of numerous biological processes in both normal and cancer cells [62]. Timmins and Ringshausen recently reviewed that in B-cell malignancies, targeting the TGFB axis, should be considered a promising approach in the context of immunotherapy [63]. The IL-6/STAT3 pathway is a classic signaling that can induce enhanced EMT process in cancers [64]. You et al. revealed the function role of IL-6/STAT3 pathway in promoting the malignant progression in oral squamous cell carcinoma patients, and further research is urgently needed to establish a more applicable therapeutic strategy targeting STAT3 pathway [65].

Of note, the immune landscape results validated that the infiltration level of M1-like macrophage and antitumor CD8+ T cells is significantly low in high-risk group, which is associated with poor clinical outcome. It has been indicated that M1-like macrophage serve as a protective factor in tumor microenvironment by promoting antitumor response [66, 67]. For instance, a recent study pointed out that irradiation in CC can bring a subtype shift from M2-like to the M1-like phenotype and eventually lead to an enhanced antitumor immune status [68]. It is well established that CD8+ T cells play key roles in the elimination of HPV in CC [69]. Previous studies have uncovered the

higher ratios of CD8+ to CD4+ T cells being closely related to improved survival [70]. On the contrary, the infiltration of neutrophils is proven to be positively correlated with survival of CC patients in our model, which is consistent with the common view that neutrophils are regarded as the most important leukocytes involving in first line defense to tissue damage [71–73]. Compared to the classic discipline to divide tumor immunophenotype into three subtypes (hot, altered, and cold), our finding compared the immune cell infiltration in high- and low-risk populations, may provide a more accurate model to guide the cellular based immunotherapy in CC.

Considerable research has suggested that docetaxel, doxorubicin, and gemcitabine can be the major chemotherapy drugs to control CC [74–76]. Exploring the relationship between risk and chemotherapy sensitivity by our model, we observed that high-risk patients had a higher sensitivity to the above drugs, which provides a favorable reference for the chemotherapy strategy of cases with CC.

Although our model was confirmed to possess promising potential for clinical application in CC, our project has some shortcomings. The clinical cohort in our study was drawn from the TCGA database of samples. We still need external datasets to validate our model. In addition, our research was mainly based on bioinformatics analyses, the expression pattern and underlying mechanisms of the model be detected with *in vivo* and *in vitro* experiments.

In summary, we developed a favorable risk classifier according to immune and hypoxia molecular subtypes. Our proposed risk classifier can be served as predictor for prognosis assessment and cancer treatment estimation in CC.

Data Availability

The public datasets to support the results of this subject can be gained from TCGA (<https://portal.gdc.cancer.gov/>).

Disclosure

A preprint of this article has previously been published in Research Square (<https://www.researchsquare.com/article/rs-1813951/v1>) [77].

Conflicts of Interest

All the authors declare that they have no conflict of interest.

Authors' Contributions

Yujing Shi, Qing Gao, and Zeyuan Liu have contributed equally to this work. Xiaoke Di and Xinchun Sun visualized the study and took part in the study design and performance. Yujing Shi, Qing Gao, Zeyuan Liu, and Gefenqiang Shen conducted the manuscript writing and bioinformatics analysis. All authors read and approved the final manuscript.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (82003228 and 82102831); Natural

Science Foundation of Jiangsu Province (BK20201080); Research Project of Clinical Medical Science and Technology Development Fund of Jiangsu University (JLY2021097).

Supplementary Materials

Supplementary 1. Supplementary Table 1: the list of hypoxia-related genes.

Supplementary 2. Supplementary Table 2: potential function of the target miRNAs.

References

- [1] V. Bouvard, N. Wentzensen, A. Mackie et al., "The IARC perspective on cervical cancer screening," *The New England Journal of Medicine*, vol. 385, no. 20, pp. 1908–1918, 2021.
- [2] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [3] L. Ferrall, K. Y. Lin, R. B. S. Roden, C. F. Hung, and T. C. Wu, "Cervical cancer immunotherapy: facts and hopes," *Clinical Cancer Research*, vol. 27, no. 18, pp. 4953–4973, 2021.
- [4] K. O'Leary, "HPV vaccines beat cervical cancer," *Nature Medicine*, vol. 25, 2021.
- [5] D. M. Harper and M. Jimbo, "Elimination of cervical cancer depends on HPV vaccination and primary HPV screening," *The Lancet Infectious Diseases*, vol. 21, no. 10, pp. 1342–1344, 2021.
- [6] D. T. Liss, T. Uchida, C. L. Wilkes, A. Radakrishnan, and J. A. Linder, "General health checks in adult primary care: a review," *JAMA*, vol. 325, no. 22, pp. 2294–2306, 2021.
- [7] K. Alfaro, M. Maza, M. Cremer, R. Masch, and M. Soler, "Removing global barriers to cervical cancer prevention and moving towards elimination," *Nature Reviews. Cancer*, vol. 21, no. 10, pp. 607–608, 2021.
- [8] D. F. Quail and J. A. Joyce, "Microenvironmental regulation of tumor progression and metastasis," *Nature Medicine*, vol. 19, no. 11, pp. 1423–1437, 2013.
- [9] B. A. Luca, C. B. Steen, M. Matusiak et al., "Atlas of clinically distinct cell states and ecosystems across human solid tumors," *Cell*, vol. 184, no. 21, pp. 5482–5496.e28, 2021.
- [10] J. Galon and D. Bruni, "Approaches to treat immune hot, altered and cold tumours with combination immunotherapies," *Nature Reviews. Drug Discovery*, vol. 18, no. 3, pp. 197–218, 2019.
- [11] P. S. Hegde and D. S. Chen, "Top 10 challenges in cancer immunotherapy," *Immunity*, vol. 52, no. 1, pp. 17–35, 2020.
- [12] C. Ohe, T. Yoshida, J. Ikeda et al., "Histologic-based tumor-associated immune cells status in clear cell renal cell carcinoma correlates with gene signatures related to cancer immunity and clinical outcomes," *Biomedicine*, vol. 10, no. 2, p. 323, 2022.
- [13] E. I. Buchbinder and A. Desai, "CTLA-4 and PD-1 pathways: similarities, differences, and implications of their inhibition," *American Journal of Clinical Oncology*, vol. 39, no. 1, pp. 98–106, 2016.
- [14] L. Davis, M. Recktenwald, E. Hutt et al., "Targeting HIF-2 α in the tumor microenvironment: redefining the role of HIF-2 α for solid cancer therapy," *Cancers*, vol. 14, no. 5, p. 1259, 2022.
- [15] I. B. Barsoum, C. A. Smallwood, D. R. Siemens, and C. H. Graham, "A mechanism of hypoxia-mediated escape from adaptive immunity in cancer cells," *Cancer Research*, vol. 74, no. 3, pp. 665–674, 2014.
- [16] A. C. Koong, V. K. Mehta, Q. T. Le et al., "Pancreatic tumors show high levels of hypoxia," *International Journal of Radiation Oncology • Biology • Physics*, vol. 48, no. 4, pp. 919–922, 2000.
- [17] C. J. Garcia Garcia, Y. Huang, N. R. Fuentes et al., "Stromal HIF2 regulates immune suppression in the pancreatic cancer microenvironment," *Gastroenterology*, vol. 162, no. 7, pp. 2018–2031, 2022.
- [18] A. Eustace, N. Mani, P. N. Span et al., "A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer," *Clinical Cancer Research*, vol. 19, no. 17, pp. 4879–4888, 2013.
- [19] B. Muz, P. de la Puente, F. Azab, and A. K. Azab, "The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy," *Hypoxia*, vol. 3, pp. 83–92, 2015.
- [20] C. Zhang, A. Quinones, and A. Le, "Metabolic reservoir cycles in cancer," *Seminars in Cancer Biology*, vol. 118, 2022.
- [21] G. Curty, P. S. de Carvalho, and M. A. Soares, "The role of the cervicovaginal microbiome on the genesis and as a biomarker of premalignant cervical intraepithelial neoplasia and invasive cervical cancer," *International Journal of Molecular Sciences*, vol. 21, no. 1, p. 222, 2020.
- [22] P. Song, W. Li, X. Wu et al., "Integrated analysis of single-cell and bulk RNA-sequencing identifies a signature based on B cell marker genes to predict prognosis and immunotherapy response in lung adenocarcinoma," *Cancer Immunology, Immunotherapy*, vol. 71, no. 10, pp. 2341–2354, 2022.
- [23] M. K. Halle, A. C. Munk, B. Engesæter et al., "A gene signature identifying CIN3 regression and cervical cancer survival," *Cancers*, vol. 13, no. 22, p. 5737, 2021.
- [24] M. D. Wilkerson and D. N. Hayes, "ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, pp. 1572–1573, 2010.
- [25] M. E. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, article e47, 2015.
- [26] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, pp. 15545–15550, 2005.
- [27] A. M. Newman, C. L. Liu, M. R. Green et al., "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.
- [28] B. Singh, J. A. Berry, A. Shohar, G. D. Ayers, C. Wei, and A. Lucci, "COX-2 involvement in breast cancer metastasis to bone," *Oncogene*, vol. 26, no. 26, pp. 3789–3796, 2007.
- [29] S. Kulkarni, J. S. Rader, F. Zhang et al., "Cyclooxygenase-2 is overexpressed in human cervical cancer," *Clinical Cancer Research*, vol. 7, no. 2, pp. 429–434, 2001.
- [30] H. Choy and L. Milas, "Enhancing radiotherapy with cyclooxygenase-2 enzyme inhibitors: a rational advance?," *Journal of the National Cancer Institute*, vol. 95, no. 19, pp. 1440–1452, 2003.
- [31] J. L. Souza, K. Martins-Cardoso, I. S. Guimaraes et al., "Interplay between EGFR and the platelet-activating factor/PAF receptor signaling axis mediates aggressive behavior of cervical cancer," *Frontiers in Oncology*, vol. 10, article 557280, 2020.

- [32] M. J. Kim, H. S. Kim, S. H. Lee, Y. Yang, M. S. Lee, and J. S. Lim, "NDRG2 controls COX-2/PGE2-mediated breast cancer cell migration and invasion," *Molecules and Cells*, vol. 37, no. 10, pp. 759–765, 2014.
- [33] S. Anoopkumar-Dukie, T. Conere, A. Houston et al., "The COX-2 inhibitor NS398 selectively sensitizes hypoxic HeLa cells to ionising radiation by mechanisms both dependent and independent of COX-2," *Prostaglandins & Other Lipid Mediators*, vol. 148, article 106422, 2020.
- [34] T. K. Choueiri and W. G. Kaelin Jr., "Targeting the HIF2-VEGF axis in renal cell carcinoma," *Nature Medicine*, vol. 26, no. 10, pp. 1519–1530, 2020.
- [35] S. Wang, X. Li, M. Parra, E. Verdin, R. Bassel-Duby, and E. N. Olson, "Control of endothelial cell proliferation and migration by VEGF signaling to histone deacetylase 7," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 22, pp. 7738–7743, 2008.
- [36] A. J. Petty and Y. Yang, "Tumor-associated macrophages: implications in cancer immunotherapy," *Immunotherapy*, vol. 9, no. 3, pp. 289–302, 2017.
- [37] C. Montemagno and G. Pages, "Resistance to anti-angiogenic therapies: a mechanism depending on the time of exposure to the drugs," *Frontiers in Cell and Development Biology*, vol. 8, p. 584, 2020.
- [38] L. Hirsch, R. Flippot, B. Escudier, and L. Albiges, "Immunomodulatory roles of VEGF pathway inhibitors in renal cell carcinoma," *Drugs*, vol. 80, no. 12, pp. 1169–1181, 2020.
- [39] L. Yang, L. M. DeBusk, K. Fukuda et al., "Expansion of myeloid immune suppressor gr+CD11b+ cells in tumor-bearing host directly promotes tumor angiogenesis," *Cancer Cell*, vol. 6, no. 4, pp. 409–421, 2004.
- [40] W. Zhang, Q. Wu, C. Wang, L. Yang, P. Liu, and C. Ma, "AKIP1 promotes angiogenesis and tumor growth by upregulating CXC-chemokines in cervical cancer cells," *Molecular and Cellular Biochemistry*, vol. 448, no. 1–2, pp. 311–320, 2018.
- [41] P. Yang, Y. Ruan, Z. Yan, Y. Gao, H. Yang, and S. Wang, "Comprehensive analysis of lymph nodes metastasis associated genes in cervical cancer and its significance in treatment and prognosis," *BMC Cancer*, vol. 21, no. 1, p. 1230, 2021.
- [42] X. Zhang, F. Chen, P. Huang et al., "Exosome-depleted miR-148a-3p derived from hepatic stellate cells promotes tumor progression via ITGA5/PI3K/Akt axis in hepatocellular carcinoma," *International Journal of Biological Sciences*, vol. 18, no. 6, pp. 2249–2260, 2022.
- [43] L. Xia, E. Tian, M. Yu et al., "ROR γ t agonist enhances anti-PD-1 therapy by promoting monocyte-derived dendritic cells through CXCL10 in cancers," *Journal of Experimental & Clinical Cancer Research*, vol. 41, no. 1, p. 155, 2022.
- [44] K. H. G. Mills, "Innate lymphoid cells recruit T cells to turn up the heat on tumors," *Cancer Cell*, vol. 40, no. 4, pp. 362–364, 2022.
- [45] Y. S. Lee and A. Dutta, "MicroRNAs in cancer," *Annual Review of Pathology*, vol. 4, no. 1, pp. 199–227, 2009.
- [46] Y. Chu, Y. Ouyang, F. Wang et al., "MicroRNA-590 promotes cervical cancer cell growth and invasion by targeting CHL1," *Journal of Cellular Biochemistry*, vol. 115, no. 5, pp. 847–853, 2014.
- [47] G. Jia, Y. Tang, G. Deng et al., "miR-590-5p promotes liver cancer growth and chemotherapy resistance through directly targeting FOXO1," *American Journal of Translational Research*, vol. 11, no. 4, pp. 2181–2193, 2019.
- [48] J. Gu, Y. Liu, T. Qi, W. Qian, D. Hu, and W. Feng, "Long non-coding RNA DUXAP8 elevates RCN2 expression and facilitates cell malignant behaviors and angiogenesis in cervical cancer via sponging miR-1297," *Diagnostic Pathology*, vol. 16, no. 1, p. 105, 2021.
- [49] Y. Gao, T. Zou, W. Liang, Z. Zhang, and M. Qie, "Long non-coding RNA HAND2-AS1 delays cervical cancer progression via its regulation on the microRNA-21-5p/TIMP3/VEGFA axis," *Cancer Gene Therapy*, vol. 28, no. 6, pp. 619–633, 2021.
- [50] W. Li and X. Sun, "Recent advances in developing novel anti-cancer drugs targeting tumor hypoxic and acidic microenvironments," *Recent Patents on Anti-Cancer Drug Discovery*, vol. 13, no. 4, pp. 455–468, 2018.
- [51] N. Yatabe, S. Kyo, Y. Maida et al., "HIF-1-mediated activation of telomerase in cervical cancer cells," *Oncogene*, vol. 23, no. 20, pp. 3708–3715, 2004.
- [52] B. J. Monk, L. J. Willmott, and D. A. Sumner, "Anti-angiogenesis agents in metastatic or recurrent cervical cancer," *Gynecologic Oncology*, vol. 116, no. 2, pp. 181–186, 2010.
- [53] J. Peixoto and J. Lima, "Metabolic traits of cancer stem cells," *Disease Models & Mechanisms*, vol. 11, no. 8, 2018.
- [54] A. Emami Nejad, S. Najafgholian, A. Rostami et al., "The role of hypoxia in the tumor microenvironment and development of cancer stem cell: a novel approach to developing treatment," *Cancer Cell International*, vol. 21, no. 1, p. 62, 2021.
- [55] Y. Peng, Y. Wang, C. Zhou, W. Mei, and C. Zeng, "PI3K/Akt/mTOR pathway and its role in cancer therapeutics: are we making headway?," *Frontiers in Oncology*, vol. 12, article 819128, 2022.
- [56] C. Harsha, K. Banik, H. L. Ang et al., "Targeting AKT/mTOR in Oral cancer: mechanisms and advances in clinical trials," *International Journal of Molecular Sciences*, vol. 21, no. 9, p. 3285, 2020.
- [57] Q. Wang, X. Zhang, X. Song, and L. Zhang, "Overexpression of T-cadherin inhibits the proliferation of oral squamous cell carcinoma through the PI3K/AKT/mTOR intracellular signalling pathway," *Archives of Oral Biology*, vol. 96, pp. 74–79, 2018.
- [58] M. K. Ediriweera, K. H. Tennekoon, and S. R. Samarakoon, "Role of the PI3K/AKT/mTOR signaling pathway in ovarian cancer: biological and therapeutic significance," *Seminars in Cancer Biology*, vol. 59, pp. 147–160, 2019.
- [59] J. T. Beck, A. Ismail, and C. Tolomeo, "Targeting the phosphatidylinositol 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) pathway: an emerging treatment strategy for squamous cell lung carcinoma," *Cancer Treatment Reviews*, vol. 40, no. 8, pp. 980–989, 2014.
- [60] S. Schwitalla, P. K. Ziegler, D. Horst et al., "Loss of p53 in enterocytes generates an inflammatory microenvironment enabling invasion and lymph node metastasis of carcinogen-induced colorectal tumors," *Cancer Cell*, vol. 23, no. 1, pp. 93–106, 2013.
- [61] S. Zhang, L. Carlsen, L. Hernandez Borrero, A. A. Seyhan, X. Tian, and W. S. El-Deiry, "Advanced strategies for therapeutic targeting of wild-type and mutant p53 in cancer," *Bio-molecules*, vol. 12, no. 4, p. 548, 2022.
- [62] S. Mariathasan, S. J. Turley, D. Nickles et al., "TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells," *Nature*, vol. 554, no. 7693, pp. 544–548, 2018.

- [63] M. A. Timmins and I. Ringshausen, "Transforming growth factor-beta orchestrates tumour and bystander cells in B-cell non-Hodgkin lymphoma," *Cancers*, vol. 14, no. 7, p. 1772, 2022.
- [64] R. B. Bednarczyk, N. Y. Tuli, E. K. Hanly et al., "Macrophage inflammatory factors promote epithelial-mesenchymal transition in breast cancer," *Oncotarget*, vol. 9, no. 36, pp. 24272–24282, 2018.
- [65] Y. You, Z. Tian, Z. Du et al., "M1-like tumor-associated macrophages cascade a mesenchymal/stem-like phenotype of oral squamous cell carcinoma via the IL6/Stat3/THBS1 feedback loop," *Journal of Experimental & Clinical Cancer Research*, vol. 41, no. 1, p. 10, 2022.
- [66] K. Ley, C. Laudanna, M. I. Cybulsky, and S. Nourshargh, "Getting to the site of inflammation: the leukocyte adhesion cascade updated," *Nature Reviews. Immunology*, vol. 7, no. 9, pp. 678–689, 2007.
- [67] K. Kashfi, J. Kannikal, and N. Nath, "Macrophage reprogramming and cancer therapeutics: role of iNOS-derived NO," *Cell*, vol. 10, no. 11, p. 3194, 2021.
- [68] P. X. Liew and P. Kubes, "The neutrophil's role during health and disease," *Physiological Reviews*, vol. 99, no. 2, pp. 1223–1248, 2019.
- [69] S. J. Otter, J. Chatterjee, A. J. Stewart, and A. Michael, "The role of biomarkers for the prediction of response to checkpoint immunotherapy and the rationale for the use of checkpoint immunotherapy in cervical cancer," *Clinical Oncology (Royal College of Radiologists)*, vol. 31, no. 12, pp. 834–843, 2019.
- [70] R. S. Jayshree, "The immune microenvironment in human papilloma virus-induced cervical lesions-evidence for estrogen as an immunomodulator," *Frontiers in Cellular and Infection Microbiology*, vol. 11, article 649815, 2021.
- [71] Y. Zhang, S. He, C. Xu, Y. Jiang, Q. Miao, and K. Pu, "An activatable polymeric nanoprobe for fluorescence and photoacoustic imaging of tumor-associated neutrophils in cancer immunotherapy," *Angewandte Chemie (International Ed. in English)*, vol. 134, no. 27, 2022.
- [72] H. Wang, M. M. H. Yung, H. Y. S. Ngan, K. K. L. Chan, and D. W. Chan, "The impact of the tumor microenvironment on macrophage polarization in cancer metastatic progression," *International Journal of Molecular Sciences*, vol. 22, no. 12, p. 6560, 2021.
- [73] J. Ren, L. Li, B. Yu et al., "Extracellular vesicles mediated pro-inflammatory macrophage phenotype induced by radiotherapy in cervical cancer," *BMC Cancer*, vol. 22, no. 1, p. 88, 2022.
- [74] Y. Xia, M. Xiao, M. Zhao et al., "Doxorubicin-loaded functionalized selenium nanoparticles for enhanced antitumor efficacy in cervical carcinoma therapy," *Materials Science & Engineering. C, Materials for Biological Applications*, vol. 106, article 110100, 2020.
- [75] Q. Wang, Y. T. Yen, C. Xie et al., "Combined delivery of salinomycin and docetaxel by dual-targeting gelatinase nanoparticles effectively inhibits cervical cancer cells and cancer stem cells," *Drug Delivery*, vol. 28, no. 1, pp. 510–519, 2021.
- [76] D. G. Mutch and J. D. Bloss, "Gemcitabine in cervical cancer," *Gynecologic Oncology*, vol. 90, no. 2, pp. S8–15, 2003.
- [77] Y. Shi, H. Chen, Z. Liu, G. Shen, X. Sun, and X. Di, "Identification of Immune and Hypoxia Risk Classifier to Estimate Immune Microenvironment and Prognosis in Cervical Cancer," vol. 2022, 2022.

Research Article

A Connexin-Based Biomarker Model Applicable for Prognosis and Immune Landscape Assessment in Lung Adenocarcinoma

Junqing Qi ¹, Jun Yin ^{2,3} and Guowen Ding ¹

¹Department of Cardiothoracic Surgery, Affiliated People's Hospital of Jiangsu University, Zhenjiang, China

²Department of Cardiothoracic Surgery, Zhongshan Hospital, Fudan University, Shanghai, China

³School of Medicine, Jiangsu University, Zhenjiang, China

Correspondence should be addressed to Guowen Ding; dgw2006@hotmail.com

Received 12 August 2022; Revised 14 September 2022; Accepted 26 September 2022; Published 12 October 2022

Academic Editor: Song Cao

Copyright © 2022 Junqing Qi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Purpose. Gap junction protein (Connexin) family is the basic unit of cellular connection, whose multiple members were recently demonstrated to be associated with tumor progression. However, the expression pattern and prognostic value of connexin in lung adenocarcinoma (LUAD) have not yet been elucidated. **Methods.** Consensus cluster algorithm was first applied to determine a novel molecular subtype in LUAD based on connexin genes. The differentially expressed genes (DEGs) between two clusters were obtained to include in Cox regression analyses for the model construction. To examine the predictive capacity of the signature, survival curves and ROC plots were conducted. We implemented GSEA method to uncover the function effects enriched in the risk model. Moreover, the tumor immune microenvironment in LUAD was depicted by CIBERSORT and ssGSEA methods. **Results.** The integrated LUAD cohort (TCGA-LUAD and GSE68465) were clustered into two subtypes (C1 = 217 and C2 = 296) based on 21 connexins and the clinical outcomes of LUAD cases in the two clusters showed remarkable discrepancy. Next, we collected 222 DEGs among two subclusters to build a prognostic model using stepwise Cox analyses. Our proposed model consisted of six genes that accurately forecast patient outcomes and differentiate patient risk. GSEA indicated that high-risk group was involved in tumor relevant pathways were activated in high-risk group, such as PI3K/AKT signaling, TGF- β pathway, and p53 pathway. Furthermore, LUAD cases with high-risk presented higher infiltration level of M2 macrophage and neutrophil, suggesting high-risk group were more likely to generate an immunosuppressive status. **Conclusion.** Our data identified a novel connexin-based subcluster in LUAD and further created a risk signature which plays a central part in prognosis assessment and clinical potency.

1. Introduction

Lung cancer is one of the most common malignancies worldwide, and the prevention of lung cancer is a worldwide public health issue. According to the latest statistics published by the International Agency for Research on Cancer (IARC), the global incidence and mortality rates of lung cancer in 2020 are among the highest in the world [1]. The risk of lung cancer will continue to intensify and become prominent within the future given the huge population base, aging, and high levels of tobacco consumption [2]. The incidence of lung adenocarcinoma (LUAD) is increasing every year and accounts for more than half of nonsmall cell lung cancer [3]. Front-line clinicians have been pushing for the

promotion of new technologies for comprehensive treatment (such as radiotherapy, immunotherapy, and targeted therapy), which have greatly reduced intraoperative injuries and postoperative complications for LUAD patients, but the diagnosis and treatment of LUAD is still encountering critical challenges [4]. For example, most patients have obvious symptoms at the time of consultation. In addition, the high incidence of resistance to radiotherapy and immunotherapy has contributed to unfavorable clinical outcomes for patients. Lung cancer is a highly heterogeneous tumor, and its occurrence is the result of coregulation of multiple genes [5]. In-depth investigation of the molecular mechanism of LUAD will provide valuable guidance for early diagnosis and individualized treatment of LUAD.

Tumor-infiltrating immune cells (TIICs) are an integral part of the tumor microenvironment (TME), including tumor-associated macrophages (TAMs), lymphocytes, and natural killer (NK) cells [6]. These immunocytes play a central part in killing management of tumors (e.g. CD8+ T cells and NK cells) on the one hand and in fostering tumor development on the other. In view of its vital role in tumor progression, the TME has emerged as an essential therapeutic target [7]. Immunosuppression of CD8+ T cells within the TME can be relieved by the use of PD1 inhibitors. It has achieved remarkable effect on the treatment of melanoma, lymphoma, and other tumors, suggesting that immunotherapy holds favorable prospects [8]. However, most patients are still experiencing poor outcomes after immunotherapy. Therefore, the immune landscape of TME in LUAD needs to be further elucidated.

The gap junction (GJ) is a special membrane structure consisting of an arrangement of connecting channels between two adjacent cells. Gap junction proteins (Connexins) are the basic units of GJ formed mainly in the cell membrane and cytoplasm [9]. Connexin participates in the exchange of messages and substances between cells and serves as an important regulator of physiological processes such as cell metabolism, internal environment stability, proliferation, and differentiation. Posttranslational modifications of connexin are often precisely regulated by cellular signaling networks [10]. Studies have demonstrated that connexin is closely bound up with a variety of classical cellular signaling pathways including MAPK, TGF- β , and Wnt pathways [11, 12]. Previous findings indicated that tumor cells present defective gap junction communication and abnormal expression of gap junction protein (connexin, Cx) [13]. As the most widely expressed gene in the Cx family, Cx43 shows the closest relationship with tumors. Poyet et al. revealed that downregulation of Cx43 expression correlates with gastric cancer tissue type, tumor differentiation degree, and clinical stage [14]. In bladder cancer, overexpression of Cx43 boosts tumor cell survival and progression by reinforcing the activity of intercellular gap junctions [15]. Moreover, GJA1 was proved to be a target gene of miR-30b-5p which could contribute to pancreatic cancer angiogenesis [16]. Nevertheless, up to now, it remains very little research on the role of connexins in LUAD.

With the advent of histological technologies and big data analysis, researchers can obtain more detailed information from tumor cells and effectively identify complex molecular features of tumors from massive amounts of data, enabling a deeper understanding of tumor biological features and clinical phenotypes [17]. Advances in bioinformatic analysis technologies have permitted researchers to observe a panoramic view of the biological process of tumor progression directly through clinical samples, which has furthered our insights into the identification of novel multiple biomarker-based signatures for clinical prediction [18, 19]. Consequently, exploring important clinically relevant variables and validating their reliable correlation with patient prognosis is a pivotal factor in facilitating the evolution of precision tumor therapy.

In this academic research, the genetic characteristics of connexins in LUAD were detected according to the data from public databases. Furthermore, we determine a novel molecular subtype based on connexins and uncover the clinical potency of the connexin-related model in LUAD cases.

2. Methods

2.1. Data Collection. The gene expression profile and the corresponding clinical information were obtained from the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) and TCGA (<https://portal.gdc.cancer.gov/>) databases, respectively. The LUAD cohort from the TCGA database containing the gene expression and the clinical information of 535 LUAD patients was utilized as the training set to establish the prognostic model, and the GSE68465 dataset containing RNA sequencing of 442 LUAD samples was selected as the validation set. The exclusion benchmarks were set as follows: (1) histologic diagnosis is not LUAD, (2) cases without completed data, and (3) overall survival time of less than 30 days. A total of 21 connexins were retrieved from previous research [20]. The gene information of all connexins is summarized in Supplementary Table 1.

2.2. Connexins Gene Cluster Analysis. A total of 21 connexins were subjected to determine the connexin-based molecular subtype using the R package “ConsensusClusterPlus” [21]. The difference between different subclusters was evaluated using the Kaplan–Meier survival analysis. The differentially expressed genes (DEGs) were screened by the “limma” package [22], before being processed for subsequent analysis.

2.3. Construction of Connexin-Related Signature. All the samples in training cohort were randomly divided into training and internal validation cohorts at a 1:1 ratio. Univariate Cox analysis was employed to discover prognostic genes in the training cohort. Next, the corresponding coefficients of these model genes were calculated to establish a prognostic model by multivariate analysis. The formula was established as follows: the risk score = $\sum_{i=1}^n (\text{coef} \times \text{Exp}_i)$. The Exp_i was the expression level of each gene and the coef was the risk coefficient of each gene. All the patients were divided into high- and low-risk group base on median risk value. To verify the predictive performance of the connexin-related gene signature, an external dataset, GSE68465, was enrolled into subsequent validation.

2.4. Immune Activity Analysis. The CIBERSORT algorithm (<https://cibersort.stanford.edu/>) was used to quantify the relative infiltration levels of 21 types of immune cells, as described before. The immune activity between the two risk subgroups, as demonstrated by normalized enrichment score (NES), were compared by the single sample gene set enrichment analysis (ssGSEA) [23].

2.5. Functional Enrichment Analysis. GSEA analysis was performed to reveal the potential molecular mechanisms of prognosis related genes and adjusted $p < 0.05$ was set as the

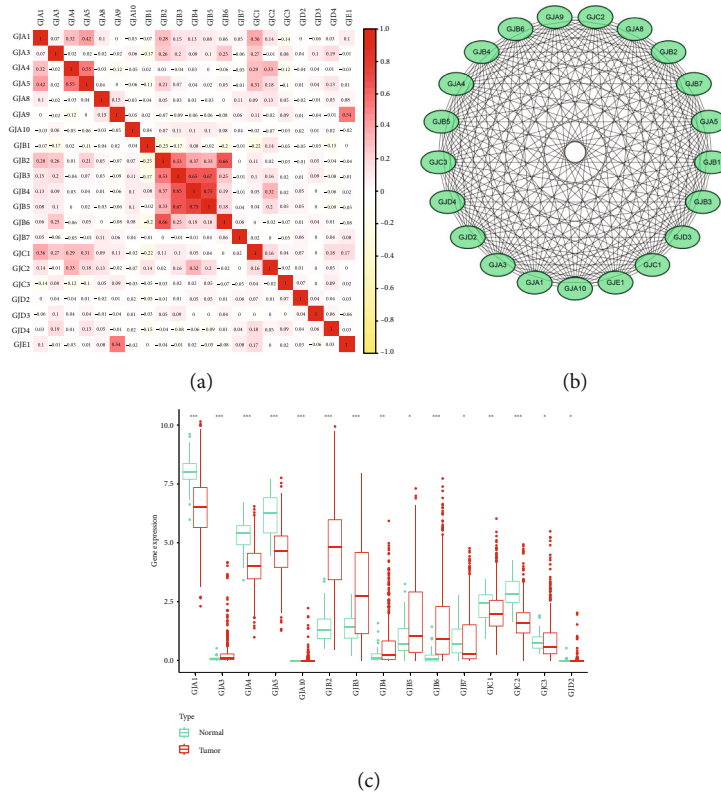


FIGURE 1: The genetic characteristics of Connexins in LUAD. (a) Correlation between expression levels of 21 connexins. (b) PPI network of 21 connexins. (c) Expression patterns of 21 connexins.

cutoff value [24]. To obtain the signaling pathways for LUAD patients, the Kyoto Encyclopedia of Genes and Genomes (KEGG) was performed and visualized by the use of “clusterProfiler” and “ggplot2” R package, respectively.

2.6. Statistical Analysis. All statistical data in this research was analyzed by R version 4.0.5. In order to further assess the predictive capacity of the established signature, the Kaplan–Meier survival analysis was performed using the “survival” R package, and the time-dependent receiver operational feature curves (ROC) were drawn based on the “survival ROC” R packages. The area under the ROC (AUC) values for 1-, 3-, and 5-year survival rate were calculated. Univariate and multivariate Cox analyses were implemented to confirm the independence of the model.

3. Results

3.1. The Genetic Characteristics of Connexins. First, we detected the correlation between 21 connexins in TCGA-LUAD dataset. The results suggested that there was a significant coexpression relationship between GJA4 and GJA5, GJA9 and GJE1, and GJB2 and GJB6 (Figure 1(a)). To explore the interaction relationship of 21 connexins at protein level, a PPI network was set up by STRING tool (Figure 1(b)). As suggested by Figure 1(c), the GJA3, GJA10, GJB2, GJB3, GJB4, GJB5, and GJB6 were remarkably

enriched in LUAD tissues, while GJA1, GJA4, GJA5, GJB7, GJC1, GJC2, GJC3, and GJD2 were greatly downregulated.

3.2. Determination of a Connexin-Based Molecular Subtype. With the 21 connexins included in consensus cluster analysis, we found that all LUAD cases were clustered into two subgroups (Figure 2(a)). The intergroup relationship between two subtypes was lowest when $k = 2$ (Figures 2(b) and 2(c)). Survival curves illustrated that there were notable discrepancies in patient outcomes between two subgroups (Figure 2(d)). PCA analysis revealed two groups of significant cluster characteristics (Figure 2(e)). In Figure 2(f), there was a tight correlation between cluster and different clinical traits. Then, a total 222 DEGs were obtained between two clusters for next Cox analysis.

3.3. Construct of a Prognostic Signature. In the training set, we first employed univariate Cox regression to discover 20 genes with prognostic values in LUAD (Figure 3(a)). Next, these 20 genes were enrolled into multivariate Cox regression and six model genes (LOXL2, PTPRH, DKK1, PKP2, NKX2-1, and KRT6A) were determined to create a prognostic model (Figure 3(b)). The risk score = $(-0.2665 \times \text{LOXL2}) + (0.1905 \times \text{PTPRH}) + (0.1281 \times \text{DKK1}) + (0.5798 \times \text{PKP2}) + (0.4434 \times \text{NKX2-1}) + (0.0103 \times \text{KRT6A})$. Then, we performed GEPIA database to explore the expression patterns of six model genes. As shown in Figure 3(c), LOXL2, PTPRH, PKP2, and KRT6A were greatly upregulated in LUAD tissues.

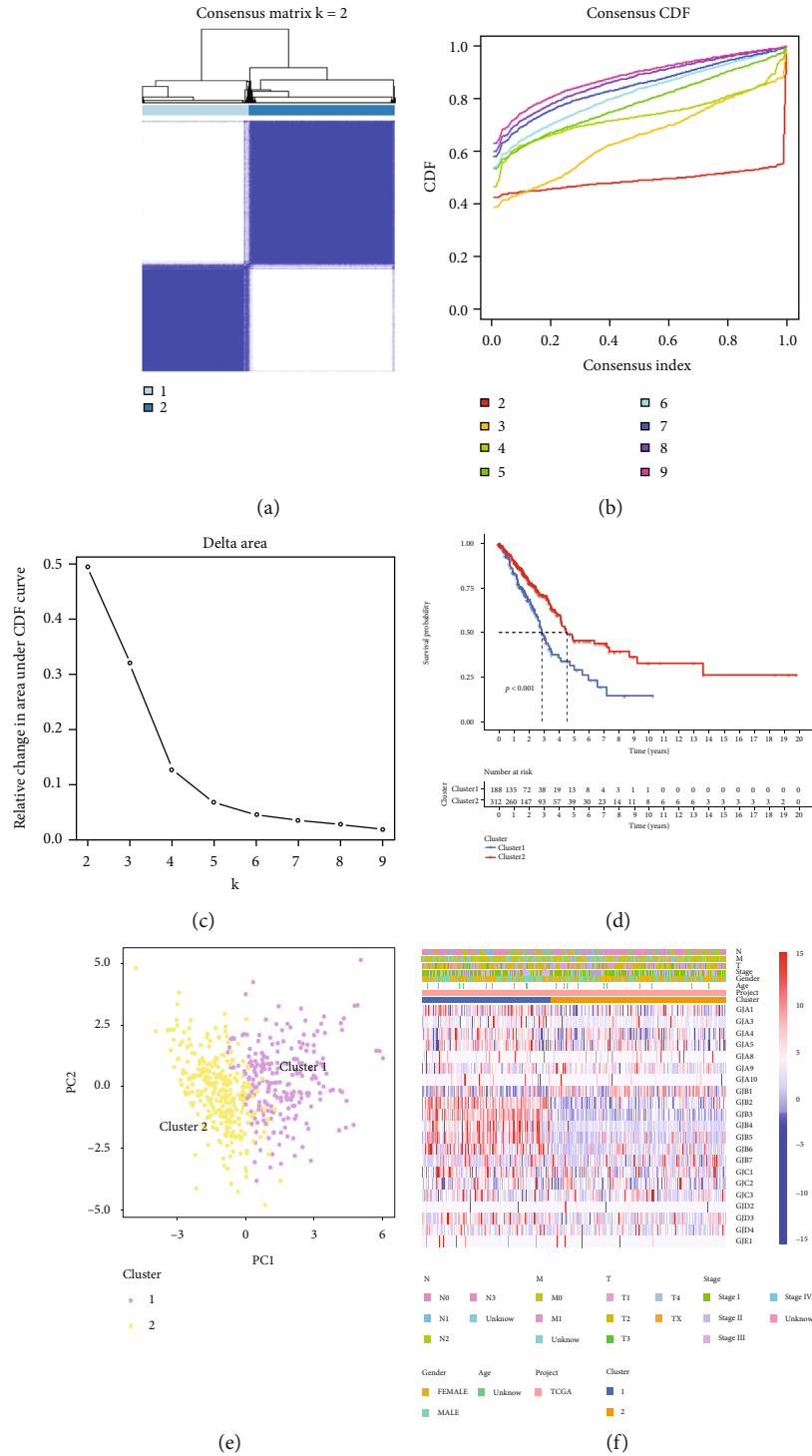


FIGURE 2: Connexin-based consensus clustering analysis. (a) Consensus cluster analysis. (b)-(c) Relative change of CDF curve. (d) The Kaplan–Meier survival analysis. (e) Principal component analysis of the two clusters. (f) Heatmap of connexin-related cluster.

Survival analysis indicated that patients with high-risk displayed a dismal clinical outcome (Figure 4(a)). The AUC (area under the curve) values of 1-, 3-, and 5-year survival rate generated by the model were 0.717, 0.702, and 0.627, respectively (Figure 4(b)). The risk plot of six genes signature is shown in Figure 4(c). Moreover, the same methods were conducted in GSE68465 cohort to confirm

the performance of the model, and the similar results were observed (Figures 4(d)–4(f)).

3.4. Independent Prognostic Analysis and Subgroup Analysis. To examine the independence of the risk model, univariate and multivariate methods were applied. In the TCGA cohort, univariate analysis showed that stage and the risk

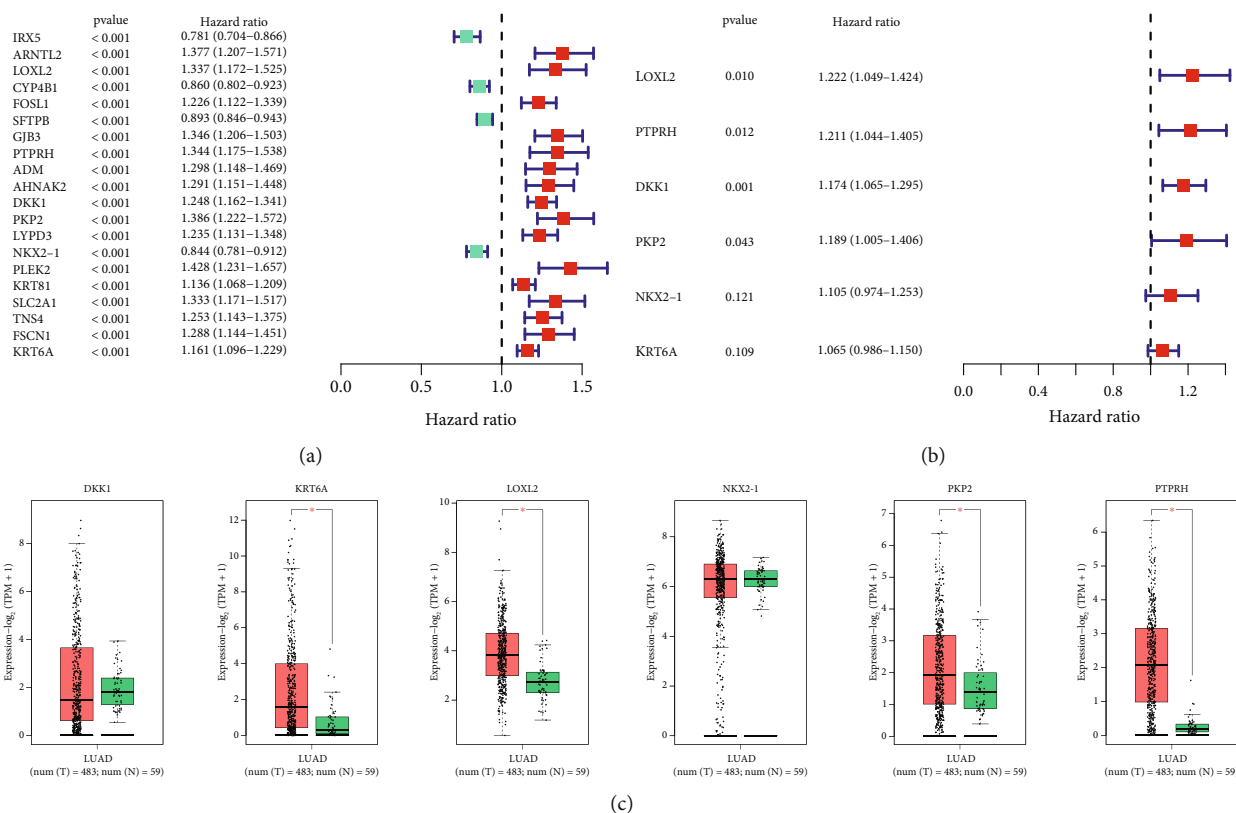


FIGURE 3: Development of a prognostic model. (a) Univariate Cox regression analysis. (b) Multivariate regression analysis for model construction. (c) Expression level of six model genes (LOXL2, PTPRH, DKK1, PKP2, NKX2-1, and KRT6A) from the GEPIA database.

score were hazard factors for evaluating patient outcome (Figure 5(a)). Multivariate Cox analysis showed that risk score ($p < 0.001$) was independent factor for assessing prognosis of LUAD (Figure 5(b)). Meanwhile, the independence of our established signature was validated in the GSE68465 cohort (Figure 5(c) and 5(d)). Next, we further detected whether the risk model is a prognostic factor for the survival assessment in different subgroups with various clinical traits. In Figure 5(e), the survival rates of the high-risk patients based on age, gender, stage, T stage, and N stage were lower than those of the low-risk patients.

3.5. GSEA of the Risk Model. GSEA showed that top five Hallmarks were greatly enriched in high-risk group, including epithelial-mesenchymal transition, glycolysis, hypoxia, PI3K/AKT/MTOR signaling, and TGF- β signaling (Figures 6(a) and 6(b)). KEGG analysis revealed that high-risk group was positively correlated with pathway in cancer, cell cycle, and p53 pathway (Figures 6(c) and 6(d)).

3.6. The Immune Landscape of LUAD. In order to characterize the immune microenvironment of LUAD cases, we first calculate the proportion of 21 different immunocytes by CIBERSORT algorithms. The results revealed that macrophages M0, macrophages M2, activated CD4 memory T cells, and neutrophils were enriched in high-risk cohort, whereas memory B cells and resting CD4 memory T cells were upregulated in low-risk cohort (Figure 7). Further-

more, we compared the difference in immune activity between the two groups by ssGSEA. As revealed by Figure 8, APC-related function, immune checkpoints, inflammation-promoting, and IFN type II were activated greatly in high-risk groups.

4. Discussion

LUAD is one of the most frequently diagnosed malignancies globally and is currently the leading cause of cancer death [25, 26]. Jemal et al. once reported that nearly 70% of LUAD patients were discovered at terminal stages at the first time of diagnosis, with 60% of them already developed distant metastasis by then [27]. Although great efforts have been made in the exploration of gene mutation targeted therapy, the five-year survival rate of LUAD patients remain dismal, which is mainly due to the lack of specific and reliable biomarkers [28]. Taken together, it is urgently needed to explore effective and less invasive surrogate molecular biomarkers that can help determine the clinical outcome of LUAD patients and further develop more promising therapeutic targets for cancer treatment.

Connexin hemichannels have long been recognized as structural precursors to form gap junctions [29]. Thus, connexins play important role in the maintaining tissue homeostasis, and their mutation is implied to induce the onset of multiple disorders. Following that, accumulating evidence has unveiled the involvement of connexins in

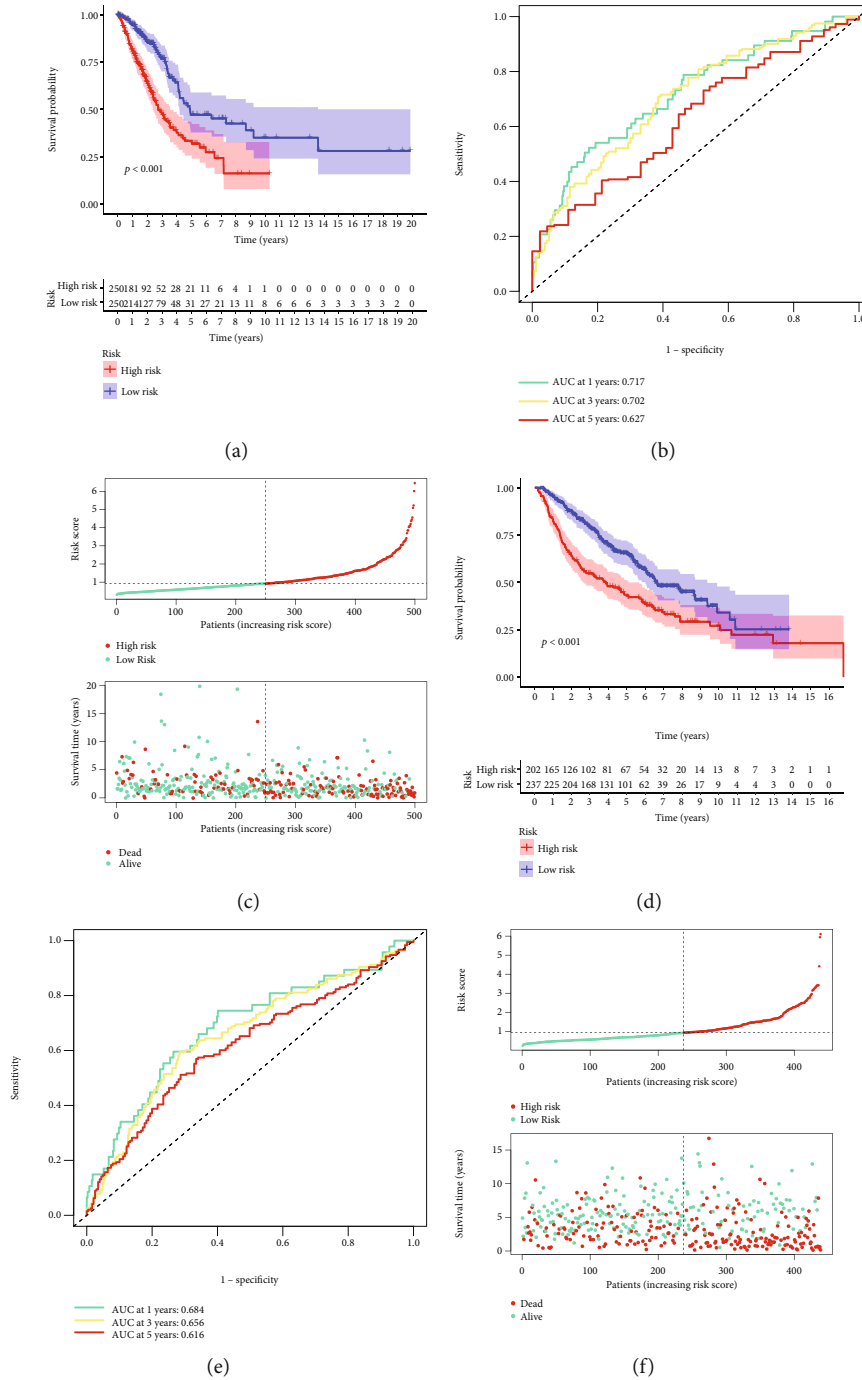


FIGURE 4: Predictive performance of the signature. (a) and (d) Survival analysis in the TCGA and the GEO datasets. (b) and (e) ROC curves of the signature. (c) and (f) The risk distribution plots in two independent cohorts.

carcinogenesis, including prostate cancer, renal cancer, and glioma cancer [30–33]. Intriguingly, connexins are reported to have distinct expression patterns at different stages of tumor progression. More specifically, connexins showed declined expression in the primary stage, while can be an overexpression when tumor cells developed a more invasive phenotype [34]. Until now, the understanding of connexin channels in LUAD is rather restrained [35, 36]. In the current study, we divided LUAD into two distinct subtypes based on expression profiles of 21 types of connexins. In

principle, LUAD patients in cluster 1 showed much poorer outcome compared to their counterparts in cluster 2. Subsequently, we identified 222 DEGs between the two populations for the establishment of prognostic model.

Advances in “Next-generation” sequencing technology have laid the foundation for the development of gene signatures in which clinical outcome of patients can be assessed on the basis of their transcriptomic data as well as the pathological grading. A total of six-gene based were verified to play a critical role in predicting clinical outcome patients

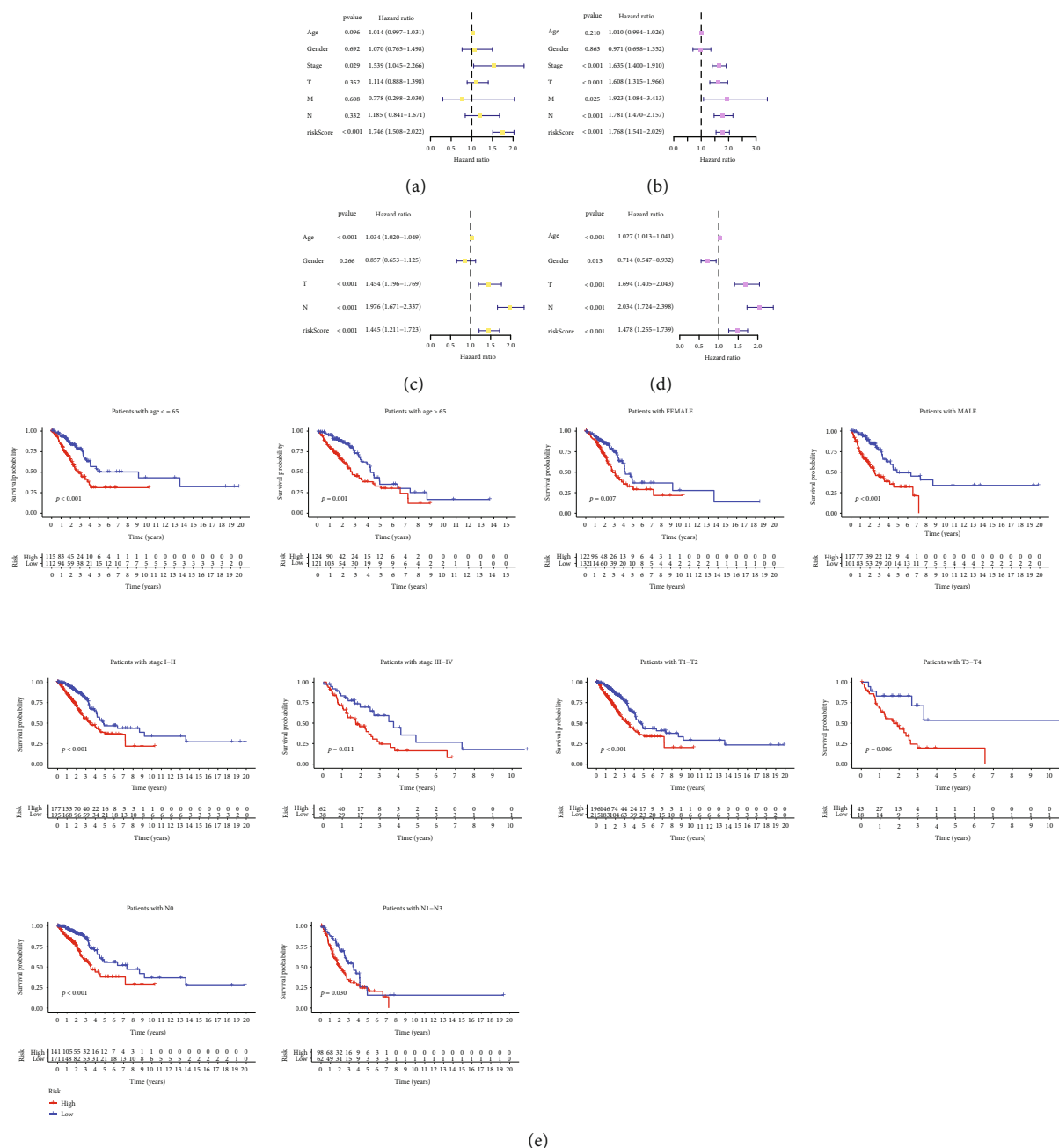


FIGURE 5: Independent prognosis analysis. (a) and (c) Univariate regression analysis the TCGA and the GEO datasets. (b) and (d) Multivariate regression analysis in two cohorts. (e) Subgroup analysis of the risk model based on age, gender, stage, T stage, and N stage.

with LUAD. LOXL2, which is strongly induced by hypoxia condition, has been identified to exert its protumor effects by promoting tumor progression in various cancers, including breast cancer, colorectal cancer, cervical cancer, and LUAD [37, 38]. In LUAD, LOXL2 was demonstrated to contribute to cell surface matrix remodeling and subsequently bring dissemination of tumor cell aggregates [39]. Protein tyrosine phosphatases (PTP) family is well known for its role in regulating tumor cell proliferation, migration, and invasion in pathology of cancers. Chen et al. once validated the prognostic value of PTPRH in LUAD tissues. The transcription as well as the protein level of PTPRH was found to be

noticeably upregulated in LUAD tissues, as demonstrated by qRT-PCR and immunohistochemistry, respectively [40]. The role of DKK1 in cancer development remains unelucidated. Although DKK1 has been reported to act as a tumor suppressor in various malignant tumors, opposing results regarding DKK1 expression and its role in cancer have been achieved recently [41, 42]. For instance, Zeybek et al. reported that the expression levels of the DKK1 in early-stage LUAD tissue were significantly downregulated compared to their counterparts in normal tissues and were closely related to the tumor progression [43]. Aberrantly expressed PKP2 has been found in a number of tumors,

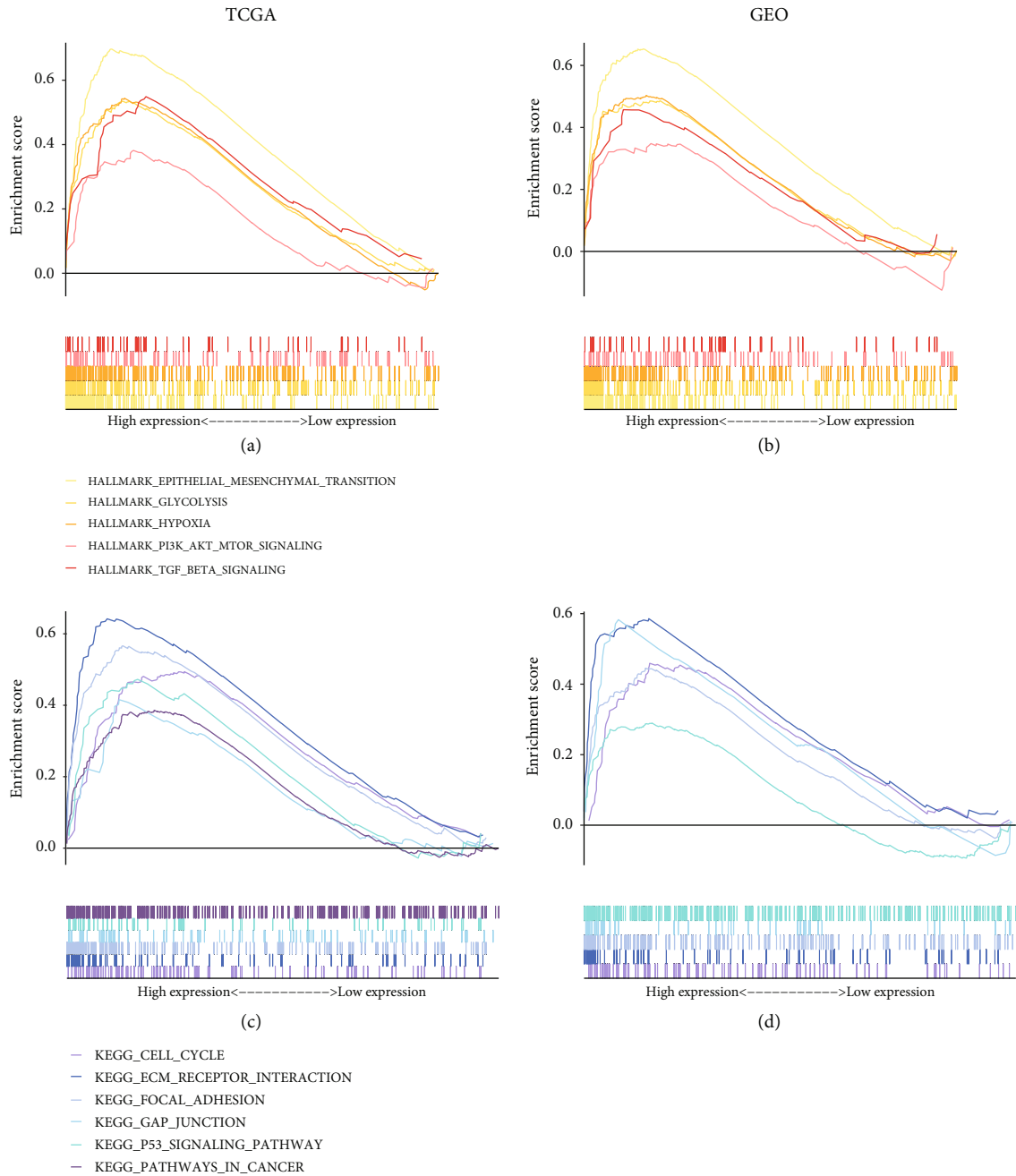


FIGURE 6: Gene set enrichment analysis. (a) and (b) Hallmark analysis of the two risk groups. (c) and (d) KEGG analysis of the two risk groups.

including bladder, osteosarcoma, and ovarian cancers [44, 45]. GSEA analysis revealed that PKP2 expression is positively associated with EGFR signaling in LUAD. It is worth noting that studies regarding the precise functions of these genes in regulating the development of LUAD remain rare until now, further research should focus on elucidating their biological functions on the basis of our work.

Molecular mechanisms participating in the regulation of LUAD were validated using the GSEA analysis. Top five Hallmarks including “EMT”, “hypoxia”, “glycolysis” “PI3K/AKT”, and “TGF- β ” were observed to be associated with the prognosis of LUAD patients in our gene signature. Acti-

vation of EMT, characterized by the loss of cell polarity and the breakdown of basement membrane, can bring mesenchymal characteristics to epithelial cells and finally promote tumor metastasis [46, 47]. EMT can also interact with “hypoxia” and “glycolysis” signaling to induce metabolic reprogramming in cancer cells [48]. The tumor progression renders the nutrients limited supply. As a result, tumors attempt to adapt to the hypoxia TME by switching to glycolysis from mitochondrial oxidative phosphorylation for their energy production, which is now known as the Warburg effect [49]. The involvement of genetic alterations of PI3K/AKT signaling in promoting the onset and development of

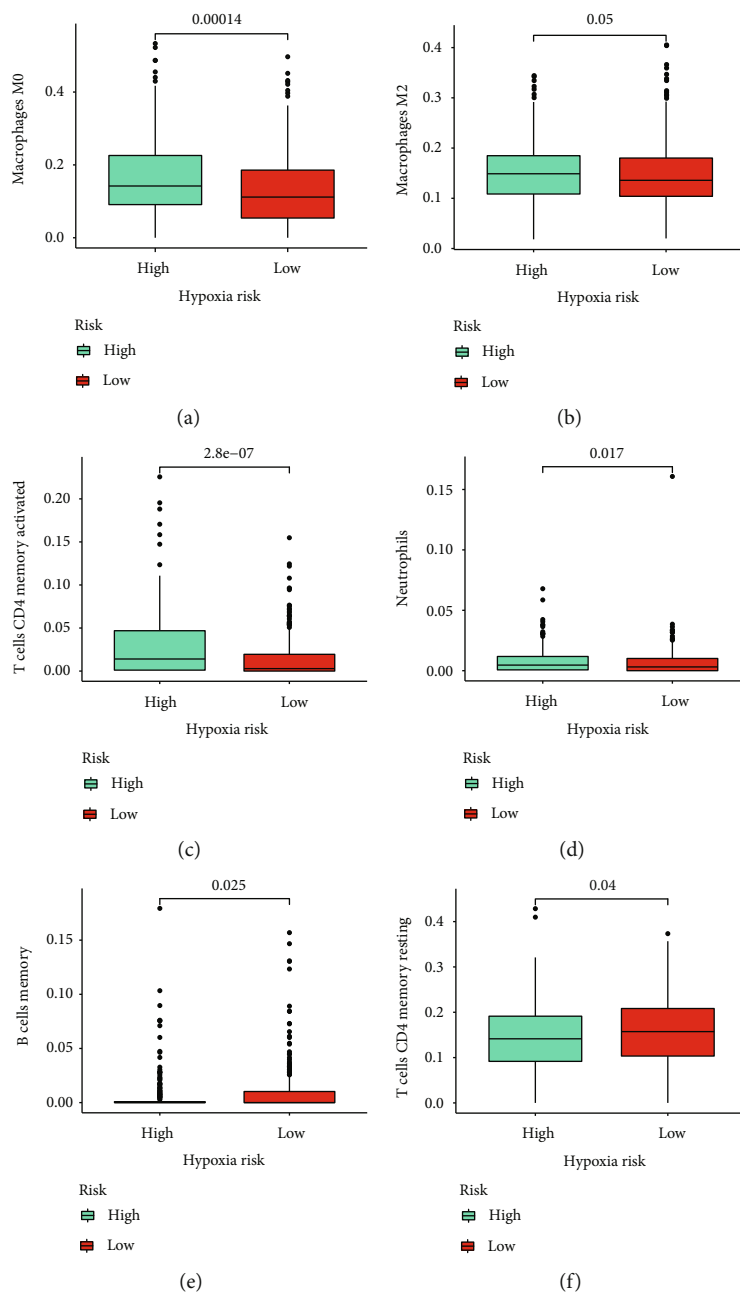


FIGURE 7: Immune infiltration analysis. (a) Macrophages M0. (b) Macrophages M2. (c) Activated CD4 memory T cells. (d) Neutrophils. (e) Memory B cells. (f) Resting CD4 memory T cells.

LUAD has been largely reported [50]. In line with previous studies, the PI3K/AKT pathway was found to be aberrantly activated in high-risk LUAD patients. Altogether, these hallmarks represent attractive therapeutic targets for the detection of novel anticancer therapies.

Additionally, we determined the distinct immunocyte infiltration status in high- and low- risk LUAD patients. M0 macrophages, M2 macrophages, activated CD4 memory T cells, and neutrophils were enriched in high-risk patients, whereas memory B cells and resting CD4 memory T cells were relatively abundant in low-risk patients. Our results revealed that there may be some existing interactions between the expression pattern of connexins and infiltration

situation in LUAD patients, which sheds lights on the detection of novel tumor immunotherapy.

Numerous reports have demonstrated a tight relationship between inflammation and cancer. The inflammatory component of tumor development involves a various population of leukocytes. These immune cells could be served as a crucial inflammatory contributor to cancer progression by releasing cytokines, chemokines, and cytotoxic mediators. Cancer-associated inflammation has impact on malignancies in many ways, including cell growth, cancer metastasis, and therapeutic resistance [51]. Although short-term IFN- γ stimulation can enhance the expression of MHC class I and antigen presentation in tumor cells, prolonged IFN- γ

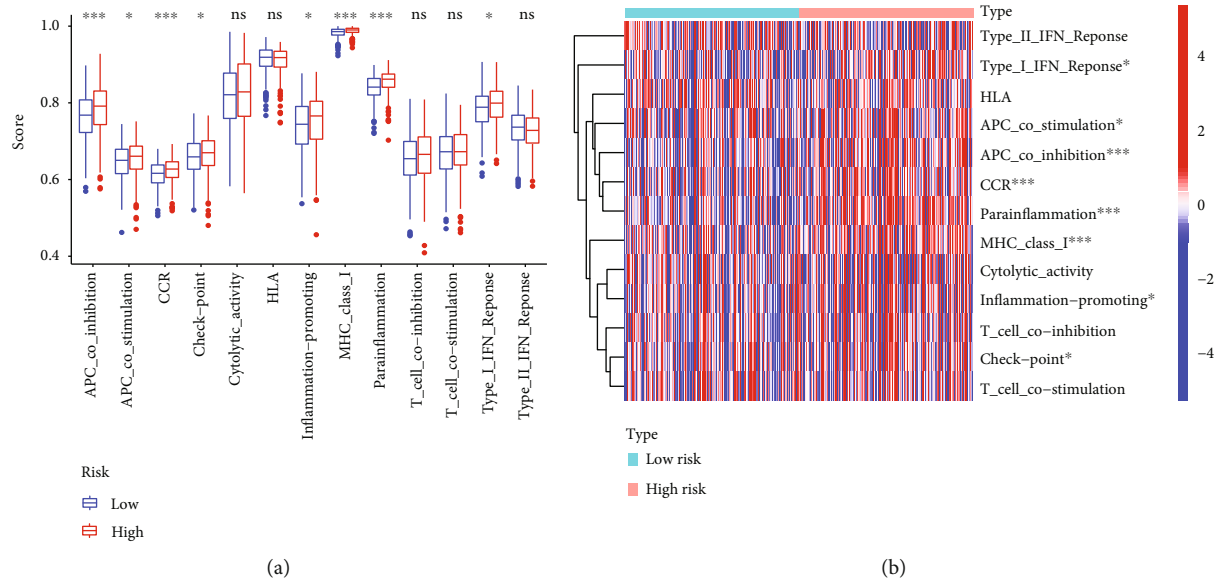


FIGURE 8: Immune function analysis. (a) The boxplot demonstrated the differences in immune function between two groups. (b) Heatmap of immune function analysis (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

exposure may lead to immune escape. On the one hand, tumor cells can reduce the IFN- γ -dependent immunosurveillance by affecting the expression and activity of IFN- γ , leading to the occurrence of immune escape. Also, IFN- γ can activate crucial immune escape genes such as PD-L1 and CTLA-4 [52]. In our data, we found that promoting inflammation and IFN response were activated in high-risk group, suggesting patients are prone to be immunosuppressive status. In addition, LUAD cases with high-risk may benefit from Immuno-Checkpoint Inhibitor (ICI) since these patients presented higher expression of immune checkpoints.

However, there are some limitations in our analysis. The data for building the model were mainly from public databases. Although the model has been confirmed in two independent datasets, its reliability still needs further validation in more real-world cohorts. The expression patterns of six model genes should be detected based on clinical LUAD specimens. Moreover, various experiments need to be conducted to explore the underlying molecular functions and mechanisms of connexin-related biomarkers. In the present study, we observed that the risk model displayed robust predictive power for assessing patient outcomes and could be stably applied to patients with LUAD. Furthermore, our constructed model could be served as a predictor for mirroring immune status of LUAD cases and provide valuable reference for therapeutic strategies.

In conclusion, we established an effective prognostic model consist of six genes on the basis of connexins molecular subtypes. Molecular signaling, immune phenotypes, and immune activities in two risk cohorts were further assessed. Taken together, our gene signature can help provide potential therapeutic targets for the different subclusters of LUAD patients and may aid in helping them choose personalized immunotherapy.

Data Availability

The public datasets to support the results of this research can be collected from TCGA (<https://portal.gdc.cancer.gov/>) and GEO (<https://www.ncbi.nlm.nih.gov/geo/>).

Conflicts of Interest

All the authors declare that they have no conflict of interest.

Authors' Contributions

Junqing Qi and Guowen Ding visualized the study and took part in the study design. Junqing Qi, Jun Yin, and Guowen Ding performed the manuscript writing and bioinformatics analysis. All authors read and approved the final manuscript.

Supplementary Materials

Supplementary Table 1: the list of 21 connexin genes. (*Supplementary Materials*)

References

- [1] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] J. Song, S. Zhang, Y. Sun et al., "A radioresponse-related lncRNA biomarker signature for risk classification and prognosis prediction in non-small-cell lung cancer," *Journal of Oncology*, vol. 2021, Article ID 4338838, 16 pages, 2021.
- [3] J. R. Molina, P. Yang, S. D. Cassivi, S. E. Schild, and A. A. Adjei, "Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship," *Mayo Clinic Proceedings*, vol. 83, no. 5, pp. 584–594, 2008.

- [4] A. Chi, X. He, L. Hou et al., "Classification of non-small cell lung cancer's tumor immune micro-environment and strategies to augment its response to immune checkpoint blockade," *Cancers (Basel)*, vol. 13, no. 12, 2021.
- [5] L. Seguin, M. Durandy, and C. C. Feral, "Lung adenocarcinoma tumor origin: a guide for personalized medicine," *Cancers (Basel)*, vol. 14, no. 7, 2022.
- [6] L. Horvath, B. Thienpont, L. Zhao, D. Wolf, and A. Pircher, "Overcoming immunotherapy resistance in non-small cell lung cancer (NSCLC) - novel approaches and future outlook," *Molecular Cancer*, vol. 19, no. 1, p. 141, 2020.
- [7] D. C. Hinshaw and L. A. Shevde, "The tumor microenvironment innately modulates cancer progression," *Cancer Research*, vol. 79, no. 18, pp. 4557–4566, 2019.
- [8] C. Genova, C. Dellepiane, P. Carrega et al., "Therapeutic implications of tumor microenvironment in lung cancer: focus on immune checkpoint blockade," *Frontiers in Immunology*, vol. 12, article 799455, 2021.
- [9] M. Delmar, D. W. Laird, C. C. Naus, M. S. Nielsen, V. K. Verselis, and T. W. White, "Connexins and disease," *Cold Spring Harbor Perspectives in biology*, vol. 10, no. 9, 2018.
- [10] J. I. Wu and L. H. Wang, "Emerging roles of gap junction proteins connexins in cancer metastasis, chemoresistance and clinical application," *Journal of Biomedical Science*, vol. 26, no. 1, p. 8, 2019.
- [11] P. J. Marie, "Signaling pathways affecting skeletal health," *Current Osteoporosis Reports*, vol. 10, no. 3, pp. 190–198, 2012.
- [12] M. Ishikawa and Y. Yamada, "The role of pannexin 3 in bone biology," *Journal of Dental Research*, vol. 96, no. 4, pp. 372–379, 2017.
- [13] C. X. Chen, K. J. Luo, J. P. Yang et al., "Connexins and cAMP cross-talk in cancer progression and metastasis," *Cancers (Basel)*, vol. 13, no. 1, 2020.
- [14] C. Poyet, L. Buser, F. Roudnicky et al., "Connexin 43 expression predicts poor progression-free survival in patients with non-muscle invasive urothelial bladder cancer," *Journal of Clinical Pathology*, vol. 68, no. 10, pp. 819–824, 2015.
- [15] M. Choudhary, C. Naczki, W. Chen, K. D. Barlow, L. D. Case, and L. J. Metheny-Barlow, "Tumor-induced loss of mural connexin 43 gap junction activity promotes endothelial proliferation," *BMC Cancer*, vol. 15, p. 427, 2015.
- [16] K. Chen, Q. Wang, X. Liu, F. Wang, Y. Yang, and X. Tian, "Hypoxic pancreatic cancer derived exosomal miR-30b-5p promotes tumor angiogenesis by inhibiting GJA1 expression," *International journal of biological sciences*, vol. 18, no. 3, pp. 1220–1237, 2022.
- [17] M. Kunz, B. Wolf, H. Schulze et al., "Non-coding RNAs in lung cancer: contribution of bioinformatics analysis to the development of non-invasive diagnostic tools," *Genes (Basel)*, vol. 8, no. 1, 2016.
- [18] D. Wu and X. Wang, "Application of clinical bioinformatics in lung cancer-specific biomarkers," *Cancer Metastasis Reviews*, vol. 34, no. 2, pp. 209–216, 2015.
- [19] J. Song, Y. Sun, H. Cao et al., "A novel pyroptosis-related lncRNA signature for prognostic prediction in patients with lung adenocarcinoma," *Bioengineered*, vol. 12, no. 1, pp. 5932–5949, 2021.
- [20] Y. J. Liu, M. Han, J. P. Li et al., "An analysis regarding the association between connexins and colorectal cancer (CRC) tumor microenvironment," *Journal of Inflammation Research*, vol. 15, pp. 2461–2476, 2022.
- [21] M. D. Wilkerson and D. N. Hayes, "ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, pp. 1572–1573, 2010.
- [22] M. E. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, p. e47, 2015.
- [23] M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, and N. Hacohen, "Molecular and genetic properties of tumors associated with local immune cytolytic activity," *Cell*, vol. 160, no. 1–2, pp. 48–61, 2015.
- [24] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences*, vol. 102, no. 43, pp. 15545–15550, 2005.
- [25] V. Rufini, M. Lorusso, F. Inzani et al., "Correction to: correlation of somatostatin receptor PET/CT imaging features and immunohistochemistry in neuroendocrine tumors of the lung: a retrospective observational study," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 68, 2022.
- [26] A. Bhopal, M. D. Peake, D. Gilligan, and P. Cosford, "Lung cancer in never-smokers: a hidden disease," *Journal of the Royal Society of Medicine*, vol. 112, no. 7, pp. 269–271, 2019.
- [27] A. Jemal, E. M. Ward, C. J. Johnson et al., "Annual report to the nation on the status of cancer, 1975–2014, featuring survival," *JNCI: Journal of the National Cancer Institute*, vol. 109, no. 9, 2017.
- [28] C. L. Chaffer and R. A. Weinberg, "A perspective on cancer cell metastasis," *Science*, vol. 331, no. 6024, pp. 1559–1564, 2011.
- [29] W. H. Evans, E. De Vuyst, and L. Leybaert, "The gap junction cellular internet: connexin hemichannels enter the signalling limelight," *Biochemical Journal*, vol. 397, no. 1, pp. 1–14, 2006.
- [30] H. Sato, K. Fukumoto, S. Hada et al., "Enhancing effect of connexin 32 gene on vinorelbine-induced cytotoxicity in A549 lung adenocarcinoma cells," *Cancer Chemotherapy and Pharmacology*, vol. 60, no. 3, pp. 449–457, 2007.
- [31] C. Lamiche, J. Clarhaut, P. O. Strale et al., "The gap junction protein Cx43 is involved in the bone-targeted metastatic behaviour of human prostate cancer cells," *Clinical & Experimental Metastasis*, vol. 29, no. 2, pp. 111–122, 2012.
- [32] E. Fujimoto, H. Sato, S. Shirai et al., "Inhibition of Src activity enhances the tumor-suppressive effect of the connexin 32 gene in Caki-1 renal cancer cells," *Oncology Reports*, vol. 15, no. 5, pp. 1359–1365, 2006.
- [33] S. Crespin, J. Bechberger, M. Mesnil, C. C. Naus, and W. C. Sin, "The carboxy-terminal tail of connexin43 gap junction protein is sufficient to mediate cytoskeleton changes in human glioma cells," *Journal of Cellular Biochemistry*, vol. 110, no. 3, pp. 589–597, 2010.
- [34] M. Mesnil, S. Crespin, J. L. Avanzo, and M. L. Zaidan-Dagli, "Defective gap junctional intercellular communication in the carcinogenic process," *Biochimica et Biophysica Acta*, vol. 1719, no. 1–2, pp. 125–145, 2005.
- [35] J. Z. Zhou and J. X. Jiang, "Gap junction and hemichannel-independent actions of connexins on cell and tissue functions—an update," *FEBS Letters*, vol. 588, no. 8, pp. 1186–1192, 2014.
- [36] S. Ravera, M. Bartolucci, E. Adriano et al., "Support of nerve conduction by respiring myelin sheath: role of connexons," *Molecular Neurobiology*, vol. 53, no. 4, pp. 2468–2479, 2016.

- [37] X. Yang, S. Li, W. Li et al., "Inactivation of lysyl oxidase by β -aminopropionitrile inhibits hypoxia-induced invasion and migration of cervical cancer cells," *Oncology Reports*, vol. 29, no. 2, pp. 541–548, 2013.
- [38] R. Kittler, L. Pelletier, A. K. Heninger et al., "Genome-scale RNAi profiling of cell division in human tissue culture cells," *Nature Cell Biology*, vol. 9, no. 12, pp. 1401–1412, 2007.
- [39] K. Mizuno, N. Seki, H. Mataka et al., "Tumor-suppressive microRNA-29 family inhibits cancer cell migration and invasion directly targeting LOXL2 in lung squamous cell carcinoma," *International Journal of Oncology*, vol. 48, no. 2, pp. 450–460, 2016.
- [40] A. Chen, S. Ding, X. Shen, and X. Lin, "The high expression of PTPRH is associated with poor prognosis of human lung adenocarcinoma," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 9932088, 9 pages, 2021.
- [41] S. Yaccoby, W. Ling, F. Zhan, R. Walker, B. Barlogie, and J. D. Shaughnessy Jr., "Antibody-based inhibition of DKK1 suppresses tumor-induced bone resorption and multiple myeloma growth in vivo," *Blood*, vol. 109, no. 5, pp. 2106–2111, 2007.
- [42] H. Hirata, Y. Hinoda, K. Nakajima et al., "Wnt antagonist DKK1 acts as a tumor suppressor gene that induces apoptosis and inhibits proliferation in human renal cell carcinoma," *International Journal of Cancer*, vol. 128, no. 8, pp. 1793–1803, 2011.
- [43] A. Zeybek, N. Oz, S. Kalemci et al., "The role of Wnt pathway antagonists in early-stage lung adenocarcinoma," *Molecular Biology Reports*, vol. 49, no. 1, pp. 9–17, 2022.
- [44] H. Yan, Y. Zhou, Z. Chen, X. Yan, and L. Zhu, "Long non-coding RNA HCG11 enhances osteosarcoma phenotypes by sponging miR-1245b-5p that directly inhibits plakophilin 2," *Bioengineered*, vol. 13, no. 1, pp. 140–154, 2022.
- [45] H. Takahashi, H. Nakatsuji, M. Takahashi et al., "Up-regulation of plakophilin-2 and down-regulation of plakophilin-3 are correlated with invasiveness in bladder cancer," *Urology*, vol. 79, no. 1, pp. 240e1–240e8, 2012.
- [46] M. W. Nam, C. W. Kim, and K. C. Choi, "Epithelial-mesenchymal transition-inducing factors involved in the progression of lung cancers," *iomolecules & Therapeutics*, vol. 30, no. 3, pp. 213–220, 2022.
- [47] Y. Liu, X. Ao, W. Yu, Y. Zhang, and J. Wang, "Biogenesis, functions, and clinical implications of circular RNAs in non-small cell lung cancer," *Molecular Therapy-Nucleic Acids*, vol. 27, pp. 50–72, 2022.
- [48] N. C. Denko, "Hypoxia, HIF1 and glucose metabolism in the solid tumour," *Nature Reviews. Cancer*, vol. 8, no. 9, pp. 705–713, 2008.
- [49] M. G. Vander Heiden, L. C. Cantley, and C. B. Thompson, "Understanding the Warburg effect: the metabolic requirements of cell proliferation," *Science*, vol. 324, no. 5930, pp. 1029–1033, 2009.
- [50] A. C. Tan, "Targeting the PI3K/Akt/mTOR pathway in non-small cell lung cancer (NSCLC)," *Thoracic Cancer*, vol. 11, no. 3, pp. 511–518, 2020.
- [51] M. Gomes, A. L. Teixeira, A. Coelho, A. Araujo, and R. Medeiros, "The role of inflammation in lung cancer," *Advances in Experimental Medicine and Biology*, vol. 816, pp. 1–23, 2014.
- [52] C. F. Lin, C. M. Lin, K. Y. Lee et al., "Escape from IFN-gamma-dependent immunosurveillance in tumorigenesis," *Journal of Biomedical Science*, vol. 24, no. 1, p. 10, 2017.

Research Article

Identification of a Metabolic Reprogramming-Associated Risk Model Related to Prognosis, Immune Microenvironment, and Immunotherapy of Stomach Adenocarcinoma

Yan Zhao ¹, Dongsheng Zhang ², and Yueming Sun ²

¹Department of Gastrointestinal Surgery, Affiliated Hospital of Jiangnan University, Wuxi, China

²Department of General Surgery, First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Correspondence should be addressed to Yueming Sun; sunyueming@njmu.edu.cn

Received 5 August 2022; Revised 22 August 2022; Accepted 27 August 2022; Published 21 September 2022

Academic Editor: Song Cao

Copyright © 2022 Yan Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Stomach adenocarcinoma (STAD) is one of the most common malignant digestive tumors. Metabolic reprogramming is an essential feature of tumorigenesis. The roles of metabolic reprogramming in STAD patients were investigated to explore the tumor immune microenvironment (TME) and potential therapeutic strategies. STAD samples' transcriptomic and clinical data were collected from The *Cancer* Genome Atlas (TCGA) set and the GSE84437 set. The signature based on the metabolism-related genes (MRGs) was built using the Cox regression model to predict prognosis in STAD. Notably, this MRG-based signature (MRGS) accurately predicted STAD patients' clinical survival in multiple datasets and could serve as an indicator independently. STAD patients with high scores on the MRGS were eligible for generating a type I/II interferon (IFN) response, according to a complete examination of the link between the MRGS and TME. Tumor Immune Dysfunction and Exclusion (TIDE) and immunophenoscore (IPS) analyses revealed that STAD patients with different MRGS scores had different reactions to immunotherapy. Consequently, assessing the pattern of these MRGs increases the understanding of TME features in STAD, hence directing the development of successful immunotherapy regimens.

1. Introduction

Stomach adenocarcinoma (STAD) is among the most common digestive malignant tumors. In 2018, approximately one million new cases were reported worldwide, the bulk of which was identified at an advanced stage locally [1, 2]. The prevalence and development of STAD continue to be poorly understood. Existing treatments for STAD mainly include surgery and chemotherapy. After surgery, the rate of local recurrence or distant metastasis varies from 40 to 70 percent, and the adverse effects of radiation and chemotherapy are quickly visible [3]. Consequently, the prevention of STAD has become a pressing public health concern. It is vital to explore the underlying mechanism of STAD to discover novel therapeutic and diagnostic targets that might help to raise the patient survival rate.

Cancers are characterized by metabolic reprogramming, which may contribute to carcinogenesis [4–6]. A large number of studies have pointed out that metabolic phenotypes can be used to image tumors and offer prognostic information, as well as treat malignancies [7]. Targeting certain metabolic pathways as a therapy technique may be beneficial in cancers. For instance, 5-fluorouracil (5-FU) possesses anticancer properties [8]. Previous studies have revealed that the progression of STAD is strongly associated with many different metabolic pathways [9, 10]. In addition, energy metabolism could be a therapeutic focus for STAD patients in the clinic. Nonetheless, the expression patterns of metabolism-related genes (MRGs) involved in metabolic reprogramming remain unclear, as well as their clinical values in STAD. Consequently, systematically evaluating the expression

features and clinical importance of those MRGs may be essential for the treatment of patients with STAD.

In this investigation, an MRG-based signature (MRGS) was generated and adequately confirmed by evaluating the transcriptome and clinical data of STAD samples in depth. This research next investigated the connection between the MRGS and other clinicopathologic variables and developed a predictive nomogram. Intriguingly, subsequent investigation revealed that MRGS was strongly linked to immune-related pathways. Consequently, we investigated the associations between the MRGS and tumor immune microenvironment (TME), checkpoint genes, as well as response to immunotherapy and sensitivity to chemotherapeutic treatment.

2. Methods

2.1. Data Collection. The STAD cohort from The Cancer Genome Atlas (TCGA) data portal containing 350 samples and the GSE84437 cohort containing 433 samples were selected for collecting information on STAD samples. Thereafter, the whole TCGA-STAD set was subdivided into a training set and an internal testing set in random order. Besides, we used the whole TCGA-STAD set as another internal validation set, and the GSE84437 set as an external validation set. A total of 1916 specific MRGs that are involved in all the metabolism-associated pathways were downloaded from the `c2.cp.kegg.v7.2.symbols.gmt` at the GSEA website [11], as shown in Table S1. Besides, data from immunotherapeutic cohorts were obtained from the IMvigor210 (<http://research-pub.Gene.com/IMvigor210CoreBiologies>) [12].

2.2. Identification of Candidate MRGs and Construction of the MRGS. Differentially expressed MRGs were identified between STAD and noncarcinoma samples from the entire TCGA set by using the “limma” package [13]. Then, the candidate MRGs were subsequently extracted from all the differentially expressed ones. The associations of candidate MRGs with the overall survival (OS) of STAD patients from the training set were analyzed using univariate Cox regression. The most optimal genes were selected via using the LASSO regression through a package named “glmnet” [14]. Thereafter, the multivariate Cox regression based on the optimal genes was used for confirming hub genes to construct the MRGS. Based on the median one of all MRGS scores in the training set, STAD patients in all sets were separately subdivided into the low- or high-risk group.

2.3. Evaluation of the Constructed Model’s and Nomogram’s Prognostic Value. The “survival” R package plotted the Kaplan–Meier analysis of all STAD groups [15]. In addition, the plotted ROC curves were to determine the signature’s specificity, as well as its sensitivity [16]. The entire TCGA cohort was utilized for analyzing the independence of the MRGS along with several common clinical variables. Combining these clinical variables with the constructed

MRGS, we built a prognostic nomogram to help to assess the survival probability of STAD patients quantitatively [17].

2.4. Analysis of Immune Cell Infiltration Level and Enriched Pathways. To analyze the correlation between the built MRGS and immune cell infiltration, we estimated the 22 immune cell subtype infiltration levels by CIBERSORT [18]. GSEA analysis was carried out on the gene expression through the package named “GSEA” to explore the biological process distinction.

2.5. Immunotherapy Efficacy Based on the MRGS. The tumor mutation burden (TMB) was calculated for each sample from the entire TCGA set. The checkpoint gene level was analyzed for confirming their relationship with the clinical OS of patients with STAD. Tumor Immune Dysfunction and Exclusion (TIDE, <http://tide.dfci.harvard.edu/>) [19] is designed to examine immune evasion mechanisms. It serves as an additional reliable biomarker that is usually used to predict immunotherapy efficacy. Greater TIDE scores suggest that tumor cells are more likely to elude immunosurveillance, hence implying a lower rate of immunotherapy response. The immunophenogram (IPS) of The Cancer Immunome Atlas (TCIA, <https://home.at/>) database was also used to assess the response of STAD patients to immune checkpoint inhibitors (ICIs) [20]. A higher IPS score frequently implies a more favorable immunotherapy response.

2.6. Chemotherapy Sensitivity Analysis. CellMiner [21] (<https://discover.nci.nih.gov/cellminer>) was used to access the NCI-60 database, which comprises a total of 60 cancer cell lines derived from various kinds of malignancies. We carried out Pearson correlation analyses to determine the relationship between the MRGS values and sensitivity to chemotherapeutic drugs.

2.7. Statistical Analysis. In this study, statistical analysis was carried out using the SPSS and R software (version 3.5.1). The “survival” package was used for the Kaplan–Meier analyses, as well as the univariate and multivariate Cox regression analyses. And, the risk ratios and accompanying confidence intervals of 95 percent were gathered. For the in vitro experiments, independent sample *t*-tests were performed for the comparisons between groups. *P* values less than 0.05 were considered statistically significant.

3. Results

3.1. Identification of Hub MRGs and Construction of the MRGS. By comparing the genes’ expression levels in the STAD and normal samples, we obtained 676 differentially expressed MRGs (Figure 1(a)). Meanwhile, we acquired 140 MRGs associated with the OS of STAD samples via the univariate analysis. Then, a total of 47 MRGs were extracted (Figure 1(b)). To explore the MRGs that were closely related to the STAD prognosis in the clinic, we carried out a univariate

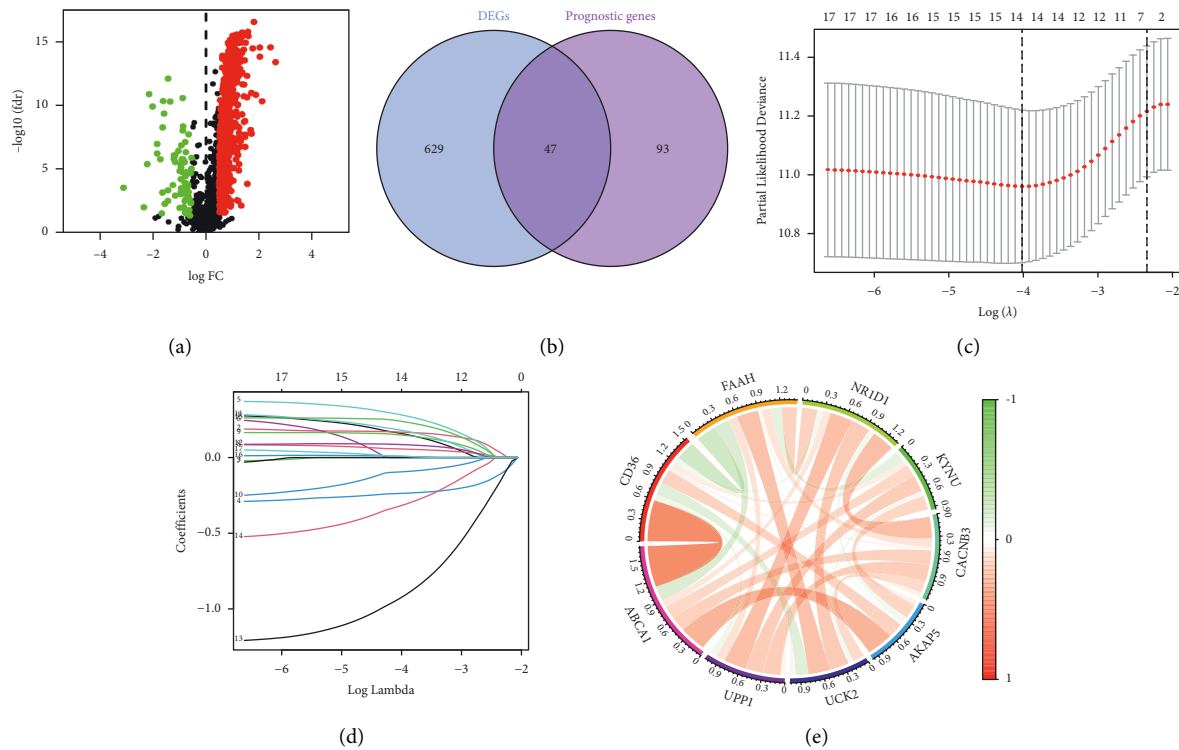


FIGURE 1: Establishment of MRGS. (a) Volcano plot regarding MRGS that differentially expressed between STAD samples and normal samples. (b) The intersections of the differentially expressed MRGS and the MRGS with prognostic value for STAD. (c) LASSO Cox regression analysis for STAD samples based on the MRGS in the intersections. (d) Coefficient profiles from the LASSO Cox analysis. (e) Correlation network of the nine candidates MRGS.

analysis based on the 47 chosen genes and found 18 related genes. The most appropriate tuning parameter from the LASSO Cox analysis was chosen later for preventing the overfitting based on the 18 ones (Figures 2(c) and 2(d)). Finally, a total of 9 MRGs were selected as hub ones, including ABCA1, CD36, FAAH, NR1D1, KYNU, CACNB3, AKAP5, UCK2, and UPP1. Afterward, the MRGS was established according to the expression of these 9 hub MRGs along with their multivariate Cox regression coefficients. The formula is as followed: score = $(0.2684 \times \text{ABCA1 level}) + (0.2350 \times \text{CD36 level}) + (-0.3043 \times \text{FAAH level}) + (0.3981 \times \text{NR1D1 level}) + (0.1651 \times \text{KYNU level}) + (0.2927 \times \text{CACNB3 level}) + (-1.2282 \times \text{AKAP5 level}) + (-0.5845 \times \text{UCK2 level}) + (0.2684 \times \text{UPP1 level})$. Moreover, the 9 prognostic MRGs were correlated with each other (Figure 1(e)).

3.2. Valuation of the Predicting Ability of MRGS. Based on the MRGS, we calculated each STAD sample's score and divided all STAD samples into the low-risk group or the high-risk group (Figure 2). Figures 2(a)–2(d) show that STAD samples in the low-risk group owned favorable survival when they were compared with those in the high-risk group in multiple sets. The risk scores and survival status of STAD samples from the multiple sets are shown in Figures 2(e)–2(h). The expression of hub MRGs in the proposed signature was similar in multiple sets (Figures 2(i)–2(l)). Furthermore, ROC analyses were carried out to evaluate the risk model's prediction (Figures 2(m)–2(p)). The values of the area under the ROC curves performed in the TCGA training set were

0.659, 0.758, and 0.783, separately for 1-, 3-, and 5- year survival, suggesting that MRGS had a good performance in monitoring survival. Meanwhile, MRGS had highly accurate predictions for the survival of STAD samples in the TCGA testing set, the whole TCGA set, and the GSE84437 set.

3.3. Association of the MRGS and Clinical Variables in STAD.

We carried out the univariate and multivariate Cox regression analyses on the MRGS and common clinical features in the entire TCGA-STAD cohort. Figures 3(a) and 3(b) demonstrate that the MRGS was significantly associated with the OS of STAD patients, suggesting that the generated MRGS may serve as a factor independently predicting the clinical prognosis. Noteworthy, age and tumor stage were also significantly correlated with the OS (Figures 3(a) and 3(b)). In addition, we developed a clinical nomogram based on the MRGS and multiple chosen clinical factors to objectively estimate the survival likelihood of individuals (Figure 3(c)). In general, the calibration curves of our generated prognostic nomogram were very congruent with the anticipated and observed survival rate in the whole TCGA cohort (Figure 3(d)). In addition, the AUC of the nomogram was greater than that of other clinical variables in the ROC curve (Figure 3(e)), demonstrating the nomogram's superior performance.

Subgroup analysis was carried out to see whether the built signature still had independent predictive value for the most important clinical characteristics. Figure 4 demonstrate that the MRGS retained its predictive power in subgroups defined

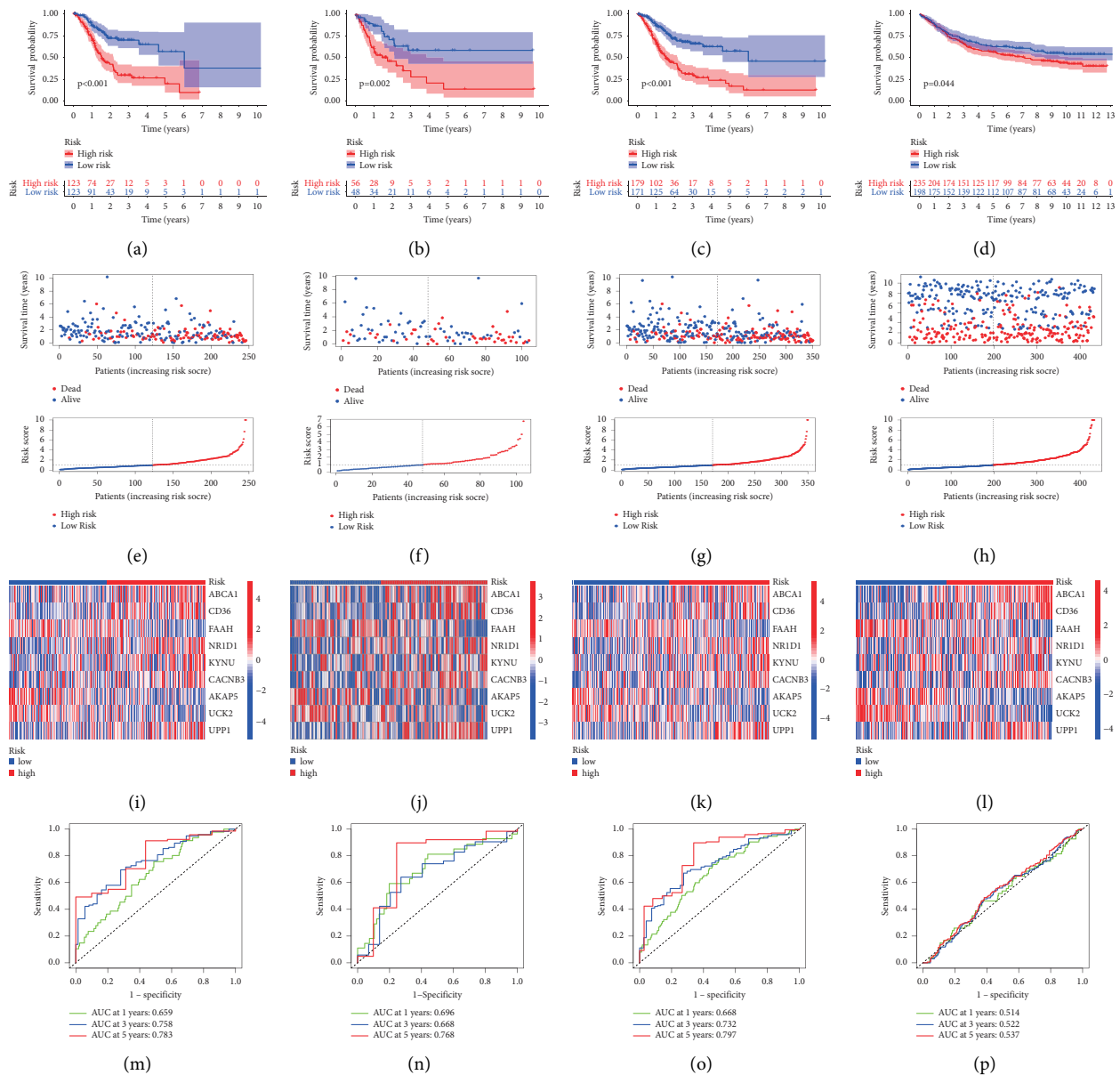
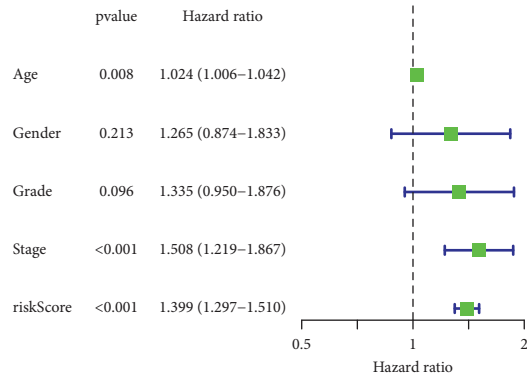


FIGURE 2: Evaluation of MRGS in predicting the survival of STAD samples from different cohorts. Distribution of KM survival (a-d), risk scores and survival status (e-h), hub MRGS' expression levels in different STAD groups (i-l), and time-dependent ROC analyses (m-p) on the TCGA training set (a, e, I, and m), TCGA testing set (b, f, j, and n), entire TCGA cohort (c, g, k, and o), and GSE84437 cohort (d, h, l, and p).

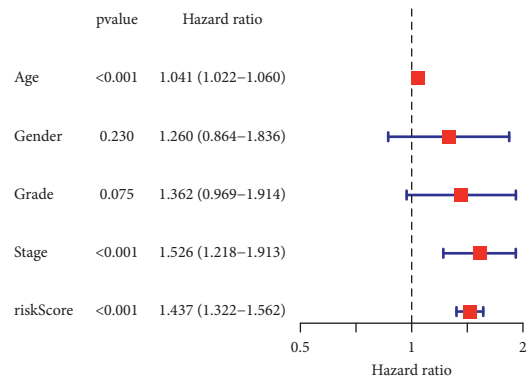
by age (age > 65 or ≤ 65, Figures 4(a) and 4(b)), gender (female or male, Figures 4(c) and 4(d)), tumor grade (G1-2 + G3, Figures 4(e) and 4(f)), tumor stage (SI-II + SIII-IV, Figures 4(g) and 4(h)), *T* stage (T1-2 + T3-4, Figures 4(i) and 4(j)), *M* stage (*M*0 + *M*1, Figures 4(k) and 4(l)), and *N* stage (*N*0-1 + *N*2-3, Figures 4(m) and 4(n)). In subgroups with distinct clinical features, the OS duration of low-risk samples was manifestly longer than that of high-risk samples.

3.4. Interrelation of the MRGS, Immune Cell Infiltration, TME, and TMB. To completely characterize the immunological aspects of STAD, CIBERSORT was carried out to examine the infiltration of immune cell subtypes in the whole TCGA-STAD set. The relative abundance of activated

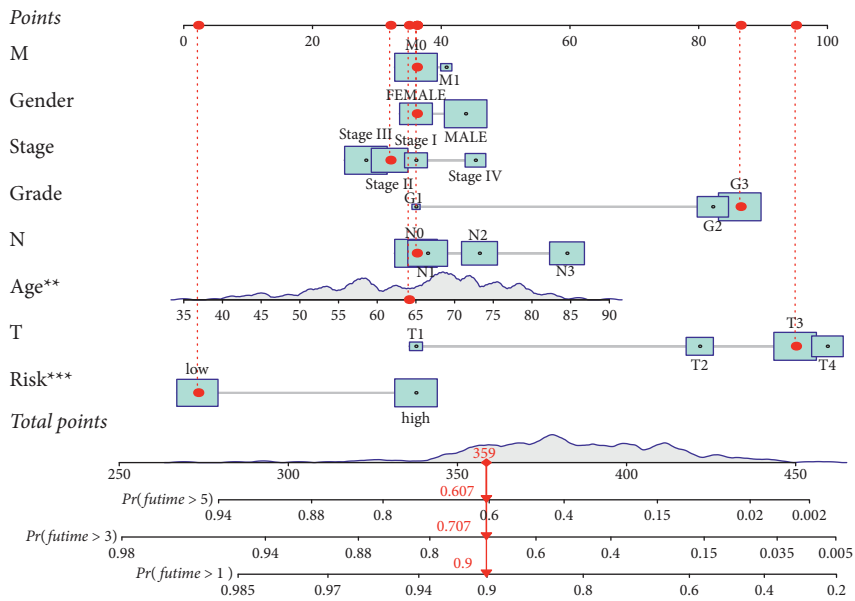
memory CD4 T cells was significantly negatively related to the score, and so did that of the follicular helper T cells (Figure 5(a)), whereas the relative abundance of M2 macrophages and resting mast cells were significantly positively correlated with the score (Figure 5(a)). In addition, type I and type II interferon (IFN) responses were both activated in the group with high risks, suggesting that immunosuppressed STAD patients might react to immunotherapy (Figure 5(b)). To further investigate the biological behaviors, a GSVA enrichment analysis was undertaken. Interestingly, many metabolism pathways, including selenoamino acid metabolism, glyoxylate and dicarboxylate metabolism, and cysteine and methionine metabolism, were significantly enriched in the low score group (Figure S1).



(a)



(b)



(c)

FIGURE 3: Continued.

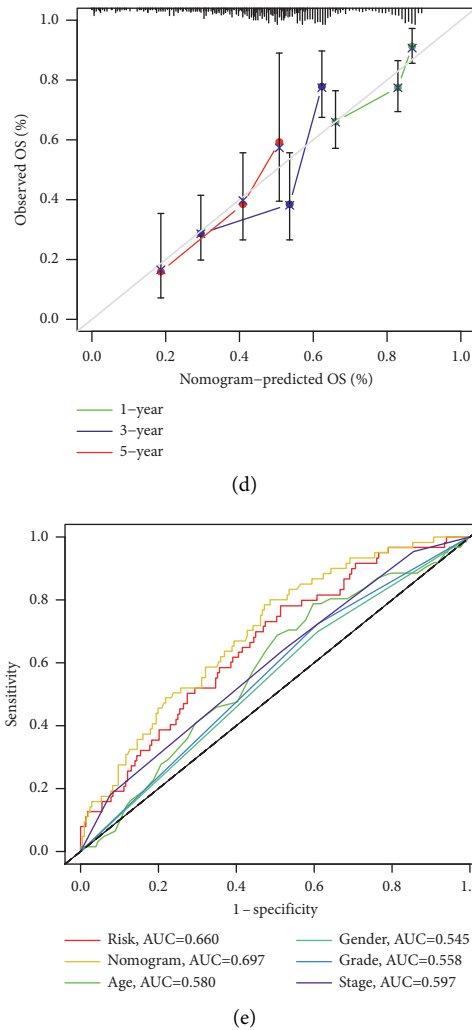


FIGURE 3: The dependence of the MRGS for prognostic prediction in STAD. (a) Results of the univariate Cox analyses of the MRGS and multiple clinical features in patients from the entire TCGA-STAD set. (b) Results of the further multivariate Cox analyses. (c) Nomogram predicting the OS in the entire TCGA cohort. (d) Calibration curves of nomogram on the consistency. (e) ROC analysis of the constructed clinical nomogram by comparing it with other chosen clinical variables.

On the basis of MRGS scores and the hierarchical clustering algorithm, all STAD samples from the entire TCGA set were neatly divided into two groups (Figure 5(c)). The features of the TME between the two STAD groups were discovered based on the findings of ESTIMATE. We discovered that the groups with the higher MRGS scores had higher estimate score and stromal score levels than the other group, which had lower values (Figure 5(d)). The mutation data were examined using the maftool package, and the mutations were stratified according to the variant effect predictor. Figures 6(a) and 6(b) depict the top 20 driver genes with the greatest frequency of modification between the high- and low-risk STAD groups. The difference in TMB between groups was also shown to be statistically significant (Figure 6(c)). Clearly, a high TMB was connected with a healthy clinical OS (Figure 6(d)). We investigated if the combination of the MRGS and TMB may be a more accurate prognostic biomarker. Therefore, we used MRGS and TMB to stratify all STAD samples from the entire TCGA set into four distinct groups. As seen in Figure 6(e), there were

substantial disparities between all four groups. Moreover, the individuals with the highest TMB and lowest MRGS scores had the greatest OS. These findings indicated conclusively that MRGS was positively associated with tumor malignancy.

3.5. Correlation of Checkpoint Genes and the MRGS and Their Impact on Clinical Outcome in the Entire TCGA-STAD Cohort. Previous research has shown the significance of immune checkpoint genes in regulating immune infiltration [22–24]. To further study the complicated interplay between immune checkpoints and the established MRGS, we evaluated their expression patterns across MRGS-based groups. As shown in Figures 7(a)–7(c), STAD patients with higher MRGS scores expressed lower levels of three chosen immune checkpoint genes (PD-1, CTLA4, and LAG3) in the entire TCGA set. Meanwhile, the expression levels of three chosen checkpoint genes all showed negative correlations to the MRGS scores (Figures 7(d)–7(f)). Then, we

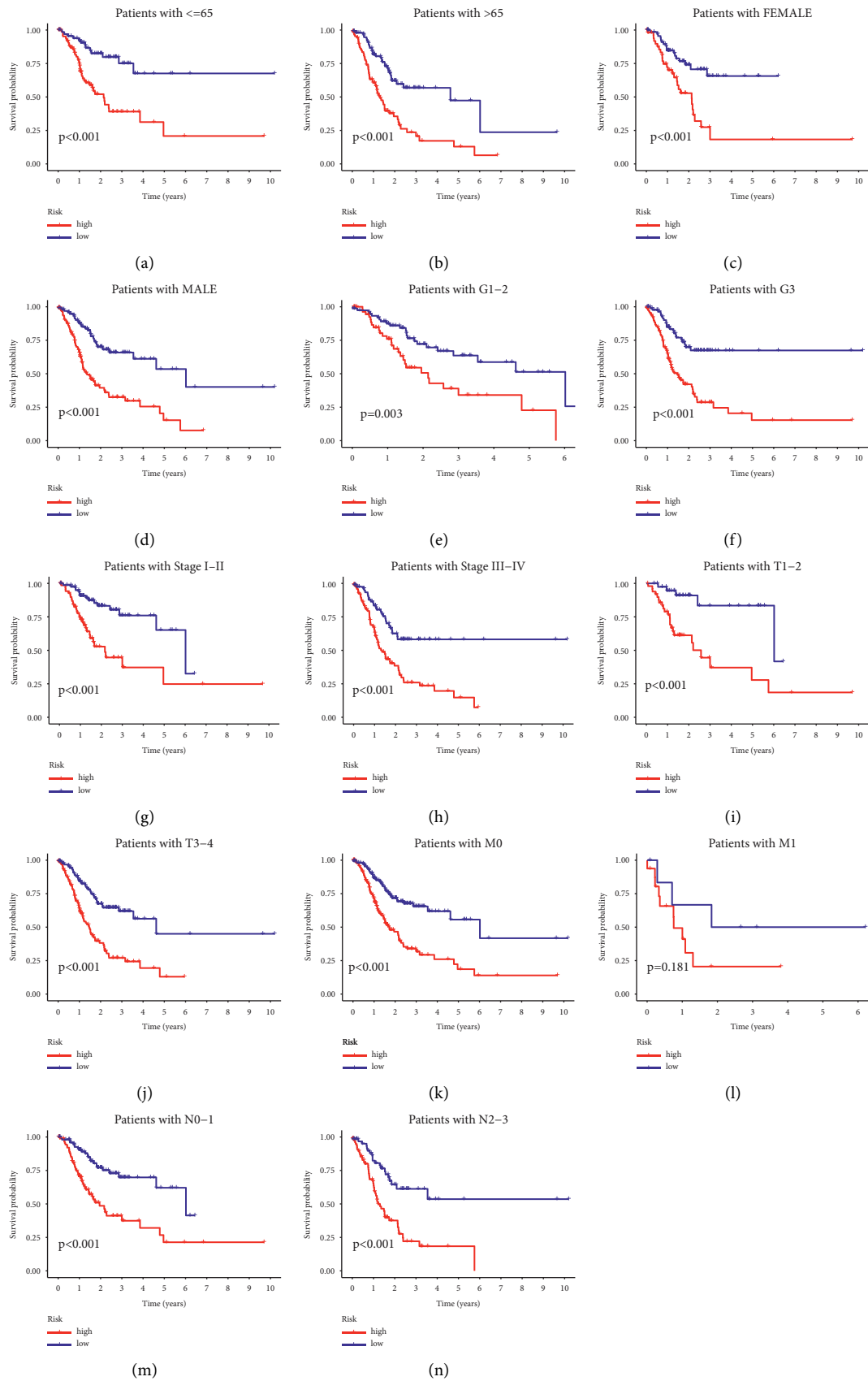


FIGURE 4: Stratified analysis based on the built model and clinical stratifications. (a-m) Longer survival time was obviously observed in STAD patients with low scores in most clinical stratifications, including patients' age (a and b), patients' gender (c and d), tumor grade (e and f), tumor stage (g and h), tumor T stage (I and j), tumor M stage (k and l) and tumor N stage (m and n).

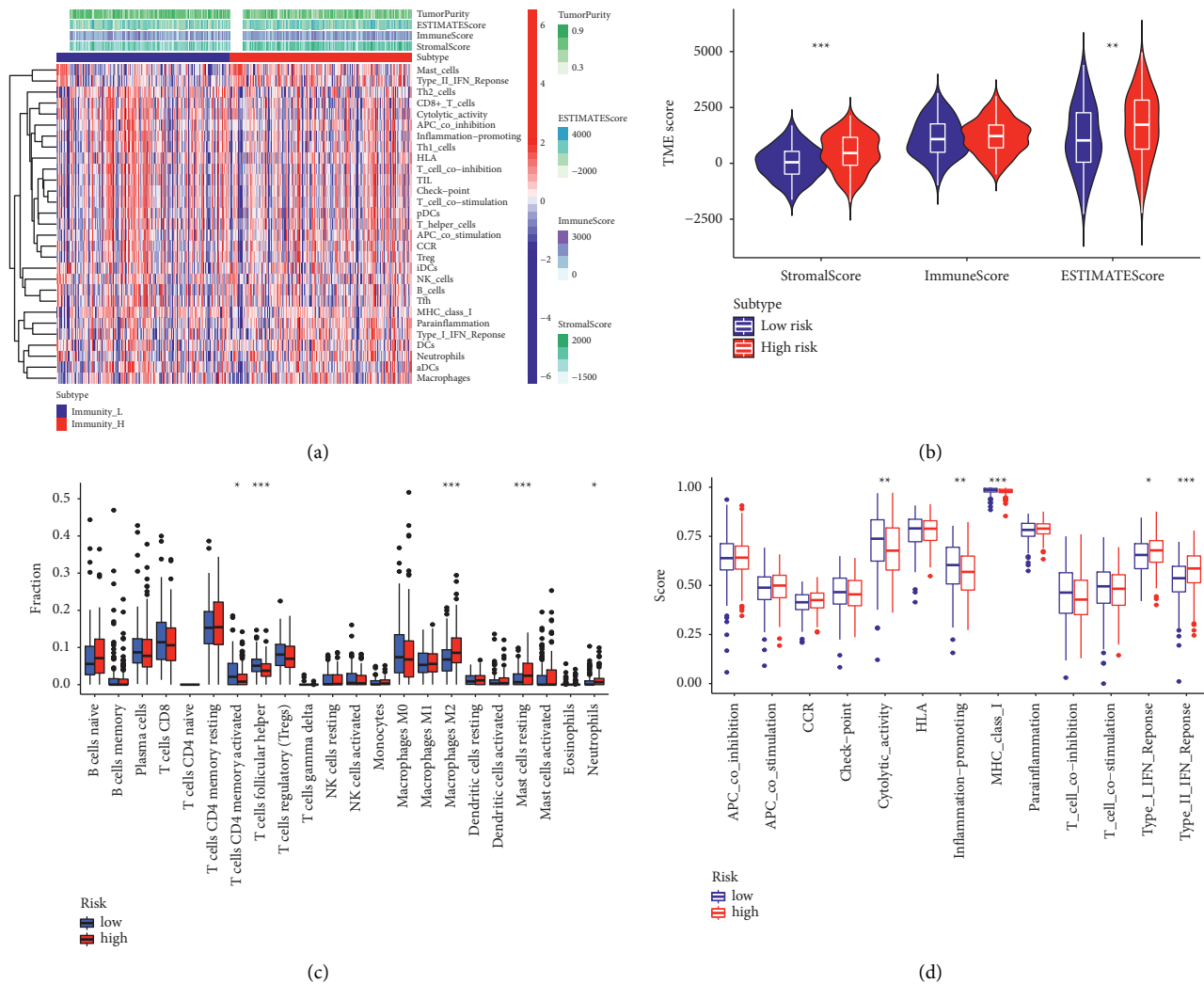


FIGURE 5: Correlation between MRGS and immune cell infiltration, TME. (a) Correlation between the MRGS and immune cell infiltration. (b) Results of ssGSEA analysis on the immune-related functions between two risk groups. (c) The landscape of the immune characteristics and TME. (d) Correlations between MRGS score and TME score.

analyzed MRGS in conjunction with immune checkpoint expression to determine if MRGS affects the OS of STAD patients with comparable checkpoint genes' expression. Survival analysis was carried out on four groups that were stratified by MRGS and immune checkpoint gene expression. Figure 7(g) illustrate that those individuals with higher PD-1 expression levels and lower MRGS scores had a longer OS than those with higher PD-1 expression levels and higher MRGS scores. In individuals with low PD-1 expression levels, a lower risk score indicated a survival rate that was significantly improved. In the entire TCGA-STAD cohort, similar survival trends were identified across the four STAD patient groups stratified by the MRGS scores and CTLA4 (Figure 7(h)) or LAG3 (Figure 7(i)) expression.

3.6. Predictive Potential of the MRGS in Immunotherapy Response and Drug Sensitivity. There is mounting evidence that ICIs increase STAD survival, although responses vary.

Therefore, precise prognostic biomarkers are urgently required. In light of the link between the MRGS and immune infiltration, as well as the checkpoint gene levels, we investigated the predictive ability of MRGS by analyzing its correlation with known immunotherapy predictors, such as TIDE [25, 26] and IPS [27]. High-risk STAD patients tended to attain greater TIDE scores in the TCGA cohort, suggesting that those in the group with low scores may benefit from ICIs in the clinic (Figure 8(a)). IPS serves as a superior predictor for the response to anti-CTLA-4 antibodies and anti-PD-1 antibodies. Although our results showed that there was no difference in IPS between the two groups shown in Figures 8(b) and 8(d), the IPS scores in the low-risk STAD group in Figures 8(c) and 8(e) were significantly elevated, suggesting that these patients may have better responses to ICIs. In addition, given the immunotherapy response prediction capacity of the MRGS, we ran Kaplan-Meier analyses on the immunotherapy cohort (IMvigor210) to evaluate the predictive significance of the immunotherapeutic OS. The anti-PD-L1 clinical response

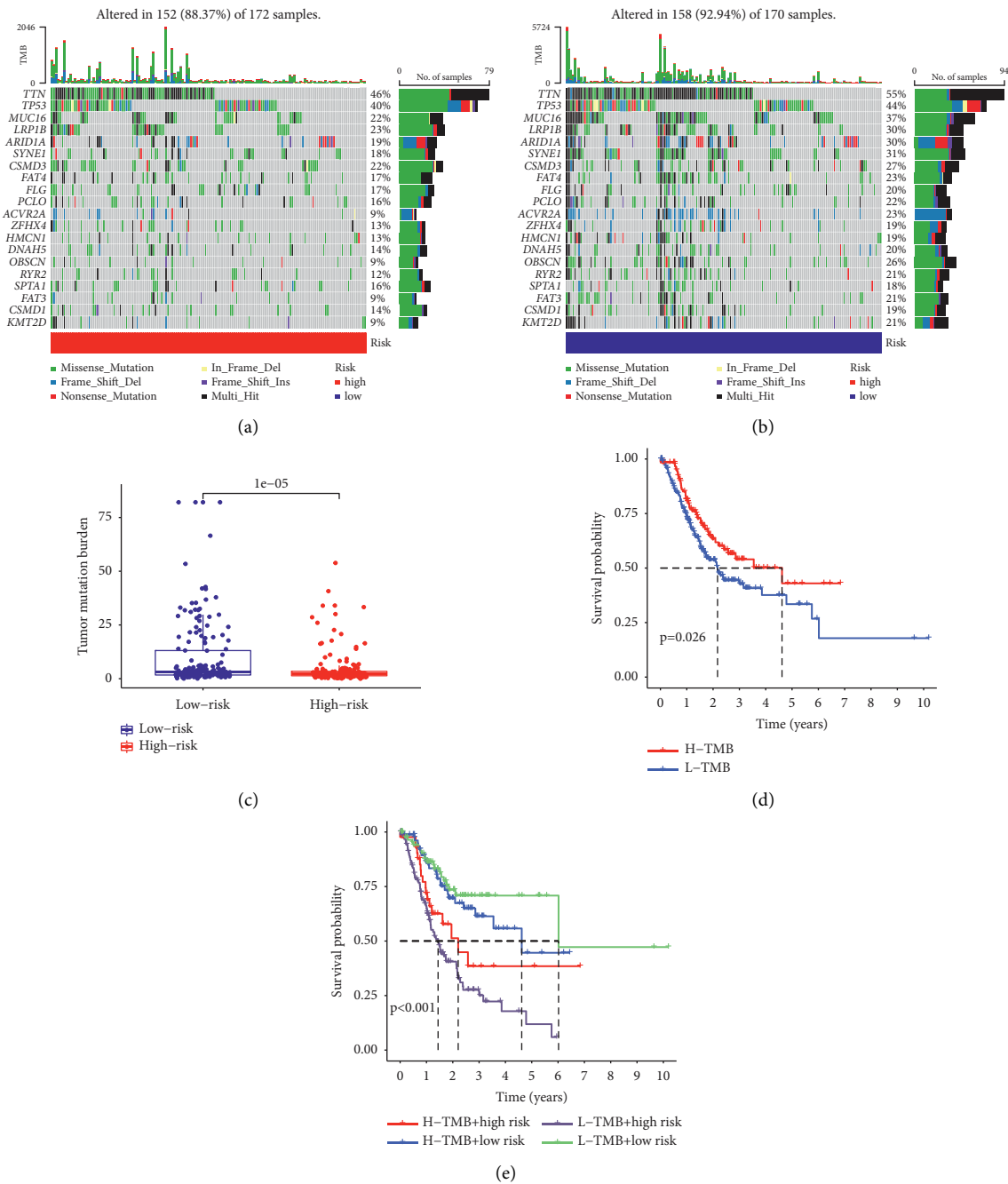


FIGURE 6: Estimation of TME and TMB in the entire TCGA-STAD set. (a) and (b) Genes with high mutation frequencies in different risk STAD groups. (c) Difference of TMB between two STAD groups. (d) Kaplan–Meier analysis based on the TMB. (e) Kaplan–Meier analysis for four groups that were stratified by combining the constructed MRGS and the TMB.

in the IMvigor210 cohort was classified as partial response (PR), complete response (CR), progressing disease (PD), and stable disease (SD). As anticipated, low scores were found to have a favorable trend in the immunotherapeutic OS (Figure 8(f)). The MRGS also had meaningful differences between the CR/PR and SD/PD groups from the IMvigor210 cohort (Figure 8(g)). All the results above showed that the constructed MRGS performs well in predicting the response to immunotherapy for STAD patients.

Moreover, we investigated the expression of prognostic MRGs in the proposed signature in cancer cells from the NCI-60 database. The results in Figure 9 showed that the majority of these prognostic MRGs were closely associated with sensitivity to some chemotherapy drugs. For example, increased expression of AKAP5, ABCA1, UCK2, and CACNB3 was obviously related to the increased drug resistance to lapatinib, afatinib, osimertinib, dacomitinib, neratinib, ibrutinib, etc. On the contrary, elevated expression levels of UPP1, CACNB3, and AKAP5 were

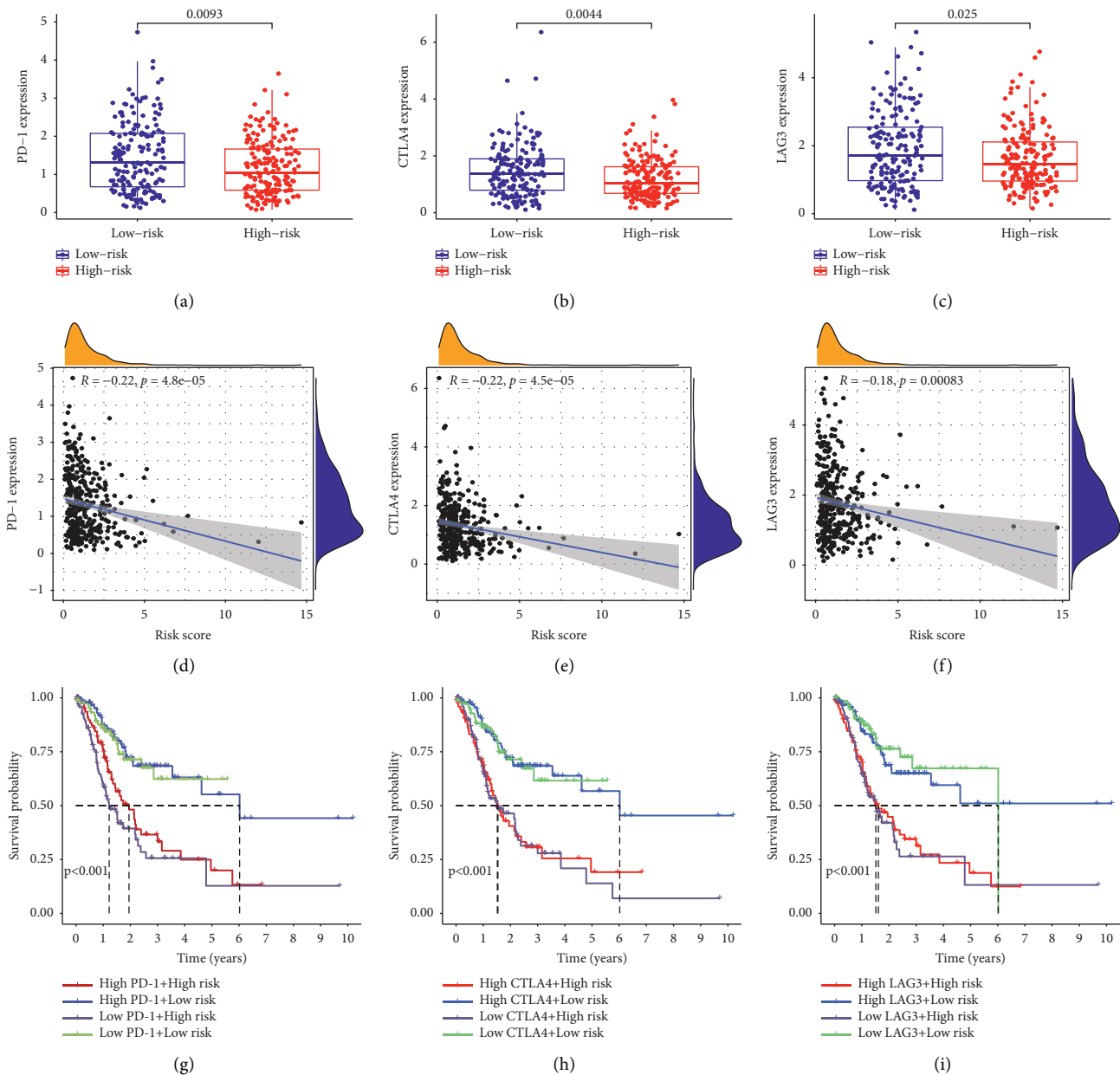


FIGURE 7: Correlation of checkpoint genes and the MRGS and their impact on clinical outcome. (a-c) Comparison of the PD-1, CTLA-4, or LAG3 expression levels between groups with different risks. (d-f) Correlation between the MRGS score and checkpoint gene (PD-1, CTLA-4, or LAG3) expression level. (g-i) Kaplan-Meier analyses the clinical OS in the four groups grouped by the MRGS and the level of PD-1, CTLA-4, or LAG3.

significantly related to the increased drug sensitivity to fulvestrant, dexrazoxane, actinomycin *D*, selumetinib, pipamperone, etc.

4. Discussion

Extensive research shows that cells in cancers usually exhibit abnormal metabolic characteristics [28]. Metabolic reprogramming is a crucial characteristic of cancer genesis. Changes in cellular metabolic activity are a characteristic of cancer [29]. One of the physiological hallmarks of a human malignant tumor is an elevated glycolytic metabolism, for

instance [30]. Several studies have shown that metabolic markers like cysteine metabolism, nucleotide metabolism, and 2-hydroxyglutarate may be used to categorize and treat gliomas [31, 32]. Given the above, metabolic therapy is a viable therapeutic option for STAD.

Here, we analyzed the MRGS' levels, together with their prognostic value based on the transcriptomic data of patients with STAD. As a result, 9 identified MRGs were adopted for building the MRGS. The mRNA level of NR1D1 was up-regulated in STAD tissues when compared with normal tissues [33]. The prognosis of STAD patients likely benefited from lower expression levels of KYNU [34]. Low expression of

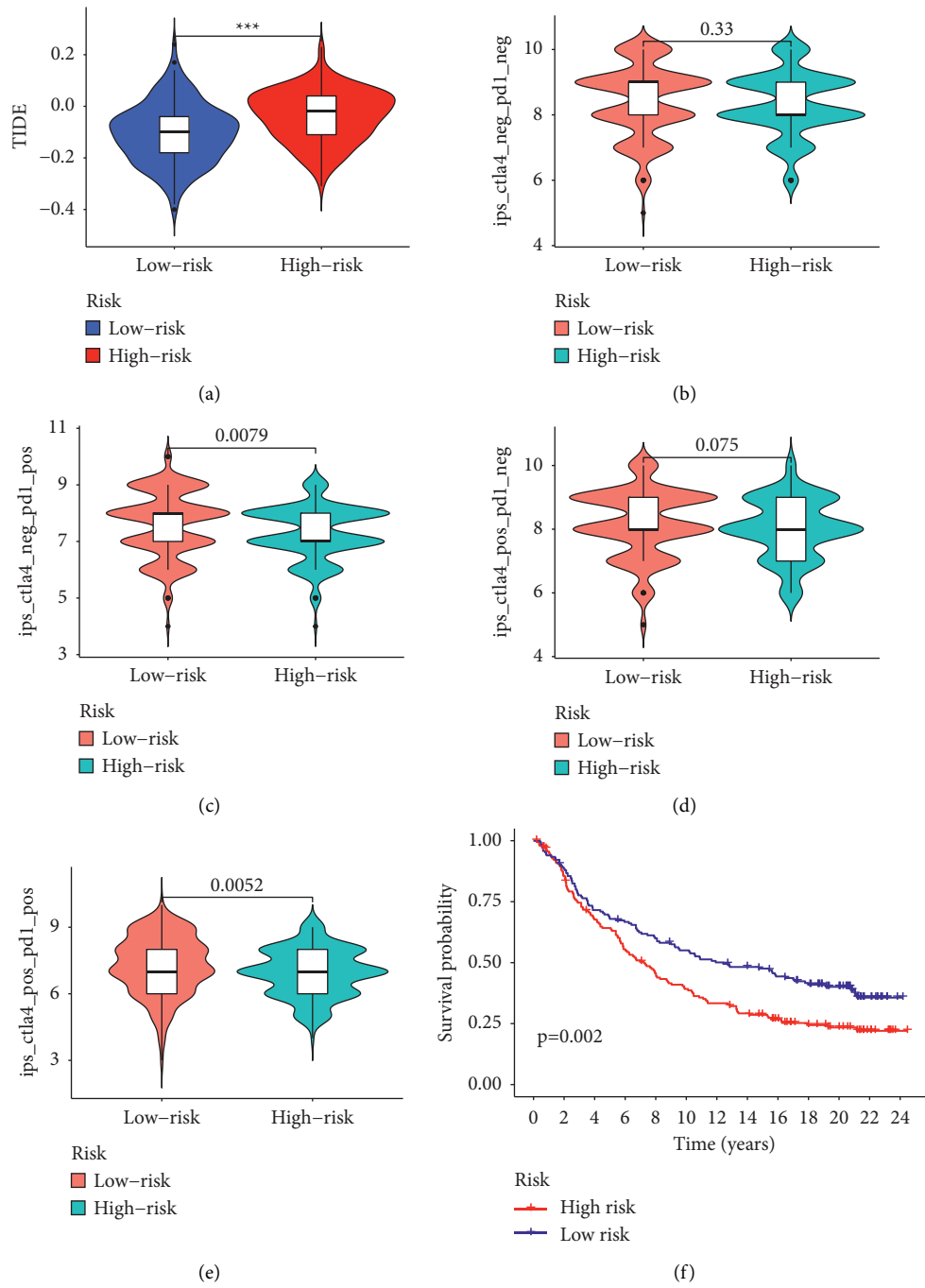
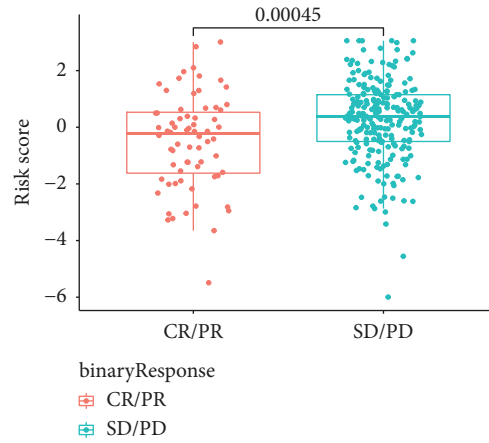


FIGURE 8: Continued.



(g)

FIGURE 8: Predictive potential of the MRGS in immunotherapeutic benefits. (a) The TIDE scores between STAD patients with different MRGS scores. (b-e) The association between the MRGS score and IPS. (f) Kaplan–Meier analysis based on the IMvigort210 cohort. (g) The distribution of MRGS in the binary response.

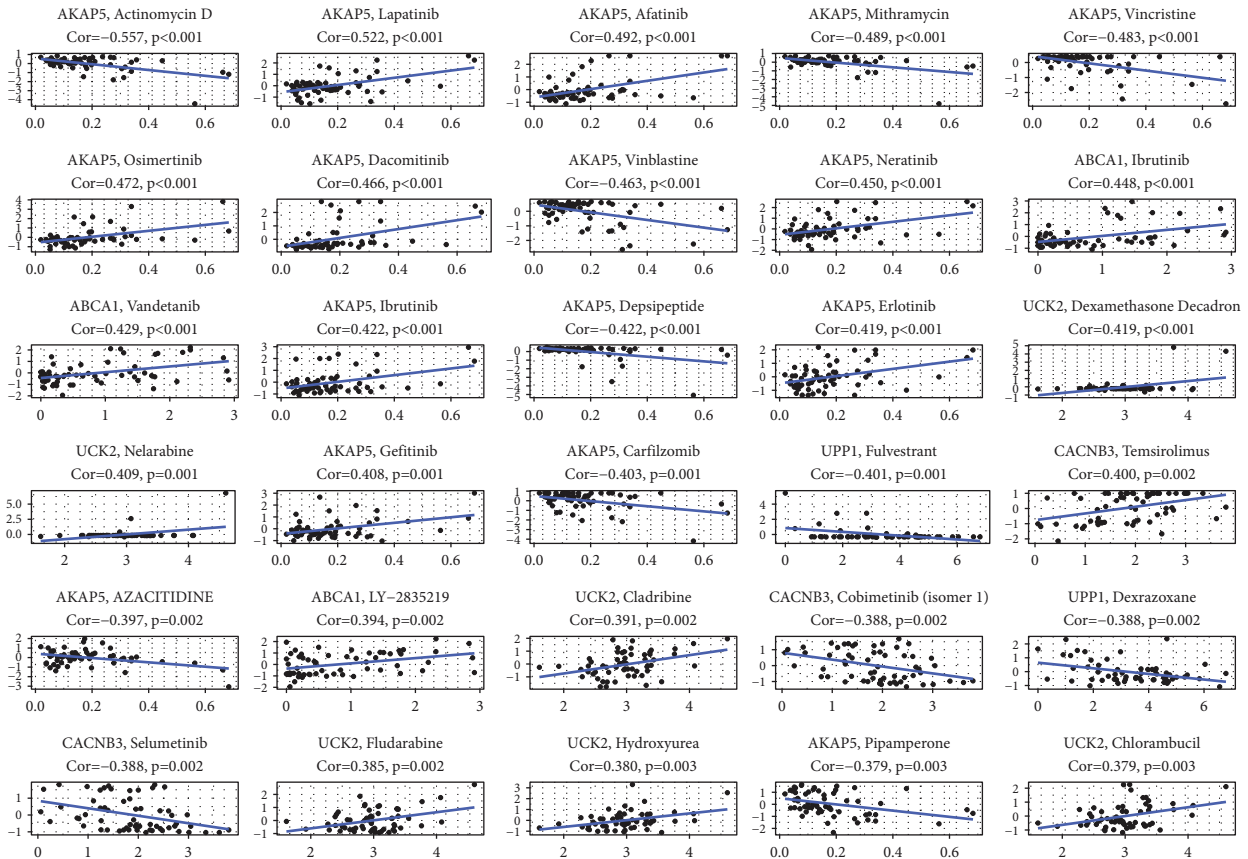


FIGURE 9: The relationship between MRGS’ expression and drug sensitivity.

AKAP5 may be a potential molecular marker for predicting poor prognosis of non-mucin-producing stomach adenocarcinoma (NMSA) via regulating cholesterol homeostasis, estrogen response, glycolysis, notch signaling, and adipogenesis pathways [35]. The malignant cell marker, UPP1, was selected to generate a signature for STAD patients [36]. Although other MRGs in the proposed signature have not been

investigated in STAD previously, they have been confirmed to influence the progression of cancers [37–41]. According to the median MRGS value of all scores, we grouped the STAD samples from different sets into low-risk or high-risk groups. The findings demonstrated that STAD patients with lower model scores had superior survival to those with higher MRGS scores. The ROC curve analysis confirmed that the

MRGS had a high potential for predicting survival. Moreover, the univariate or multivariate analysis based on MRGS and other clinical variables confirmed the independence of the MRGS's prognostic value, while the OS time of STAD samples varied significantly in different clinical feature subgroups. The nomogram we built based on the MRGS and common clinical variables could better predict the survival of STAD patients since its AUC value was a little higher.

Numerous landmark studies have pointed out that metabolic changes exert vital roles in immune regulation [42, 43]. The aberrant metabolism in cancers may have a substantial effect on TME, which is often acidic, hypoxic, and lacks nutrients necessary by immunological cells [44]. In addition, aerobic glycolysis inside cancer cells shapes the immune system by upregulating cytokine production and inhibiting the development of monocytes into dendritic cells [45]. Meanwhile, immune cells' metabolism is a significant factor in determining their survivability and roles [46]. Consequently, metabolism is strongly tied to immunity. The creation of a prognostic signature related to metabolism may aid a lot in forecasting the status of immune responses. To better understand the relationship between the built MRGS and immunity, we compared TME between groups and found several immune cell infiltration levels were obviously elevated in the STAD group with low risk. These findings indicate that metabolically active tumor cells in STAD formed a microenvironment that was harmful to immune cells, consistent with previous reports [47]. The immune response against cancer is an important component of the complicated tumor immunophenotype that underlies the TME [48]. By modifying the immune cell state inside TME, the aberrant metabolism may contribute to a different outcome. This is consistent with the finding that tailored metabolism may facilitate the regulation of antitumor immune response.

Interestingly, our data showed that the main indicators of exhausted T cells were abnormally elevated in STAD samples with higher MRGS scores, suggesting T cells may have become more hypofunctional and hyporesponsive as metabolism activates in cancer cells. This result may explain why elderly persons have a lower immunotherapy response rate. In order to confirm our findings, we also collected immunotherapy data from TICA. STAD patients with lower MRGS scores may have a stronger immunotherapy response, as evidenced by the findings. Moreover, we collected immunotherapy information from the IMvigor210 set and further confirmed the ability of the MRGS in predicting the STAD patients' immunotherapy responses.

The current work has a number of strengths. This is the first signature that is created on the basis of metabolic reprogramming that may represent the prognosis and TME, immunotherapy of STAD patients. Second, different data sets were used in order to validate and assess the predictive efficacy of our created signature. Thirdly, we determined the possible molecule that matched our signature. However, more prospective cohort studies are required to evaluate the therapeutic utility of this predictive signature. In the meanwhile, the possible chemical requires additional investigation.

5. Conclusion

In conclusion, we performed a complete examination of the expression of MRGs in STAD patients and then developed an MRGS with the ability to predict clinical outcomes and immunological microenvironment. As a useful tool, this built signature may aid in searching for possible combination immunotherapy drugs and offer a therapeutic approach for the treatment of STAD patients.

Data Availability

Publicly available datasets were analyzed in this study. This data can be found in TCGA-STAD and GSE84437.

Conflicts of Interest

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

Authors' Contributions

The authors thank all the patients for providing organization and related scholars for sharing test results in public. YZ and YS designed the experiment; YZ, DZ, and YS collected data; YZ and YS were involved in the methodology; YZ and YS wrote the original draft; DZ and YS wrote the review and edited. All authors agree to publish the article.

Supplementary Materials

Figure S1. The GSEA enrichment heatmap between two risk score groups from the entire TCGA-STAD set. Table S1. The details of MRGs.() (*Supplementary Materials*)

References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] H. Katai, T. Ishikawa, K. Akazawa et al., "Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001–2007)," *Gastric Cancer*, vol. 21, no. 1, pp. 144–154, 2018.
- [3] A. D. Wagner, N. L. Syn, M. Moehler et al., "Chemotherapy for advanced gastric cancer," *Cochrane Database of Systematic Reviews*, vol. 8, Article ID CD004064, 2017.
- [4] M. Yamamoto, H. Inohara, and T. Nakagawa, "Targeting metabolic pathways for head and neck cancers therapeutics," *Cancer and Metastasis Reviews*, vol. 36, no. 3, pp. 503–514, 2017.
- [5] J. Li, C. Bolyard, G. Xin, and Z. Li, "Targeting metabolic pathways of myeloid cells improves cancer immunotherapy," *Frontiers in Cell and Developmental Biology*, vol. 9, Article ID 747863, 2021.

- [6] Y. A. Shen, C. C. Chen, B. J. Chen et al., "Potential therapies targeting metabolic pathways in cancer stem cells," *Cells*, vol. 10, no. 7, p. 1772, 2021.
- [7] M. G. Vander Heiden and R. J. DeBerardinis, "Understanding the intersections between metabolism and cancer biology," *Cell*, vol. 168, no. 4, pp. 657–669, 2017.
- [8] S. Hemaiswarya and M. Doble, "Combination of phenylpropanoids with 5-fluorouracil as anti-cancer agents against human cervical cancer (HeLa) cell line," *Phytomedicine*, vol. 20, no. 2, pp. 151–158, 2013.
- [9] Z. Ye, M. Zheng, Y. Zeng et al., "A 13-gene metabolic prognostic signature is associated with clinical and immune features in stomach adenocarcinoma," *Frontiers Oncology*, vol. 11, Article ID 612952, 2021.
- [10] K. Yao, L. Wei, J. Zhang et al., "Prognostic values of GPNMB identified by mining TCGA database and STAD microenvironment," *Aging (Albany NY)*, vol. 12, no. 16, Article ID 16254, 2020.
- [11] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, Article ID 155, 2005.
- [12] S. Mariathasan, S. J. Turley, D. Nickles et al., "TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells," *Nature*, vol. 554, no. 7693, pp. 544–548, 2018.
- [13] M. E. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, p. e47, 2015.
- [14] S. Engebretsen and J. Bohlin, "Statistical predictions with glmnet," *Clinical Epigenetics*, vol. 11, no. 1, p. 123, 2019.
- [15] G. D'Arrigo, D. Leonardis, S. Abd ElHafeez, M. Fusaro, G. Tripepi, and S. Roumeliotis, "Methods to analyse time-to-event data: the kaplan-meier survival curve," *Oxidative Medicine and Cellular Longevity*, vol. 2021, Article ID 2290120, 7 pages, 2021.
- [16] P. J. Heagerty, T. Lumley, and M. S. Pepe, "Time-dependent ROC curves for censored survival data and a diagnostic marker," *Biometrics*, vol. 56, no. 2, pp. 337–344, 2000.
- [17] J. K. Holodinsky, A. Y. X. Yu, M. K. Kapral, and P. C. Austin, "Comparing regression modeling strategies for predicting hometime," *BMC Medical Research Methodology*, vol. 21, no. 1, p. 138, 2021.
- [18] A. M. Newman, C. L. Liu, M. R. Green et al., "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.
- [19] Z. Wang, Y. Wang, T. Yang et al., "Machine learning revealed stemness features and a novel stemness-based classification with appealing implications in discriminating the prognosis, immunotherapy and temozolomide responses of 906 glioblastoma patients," *Briefings in Bioinformatics*, vol. 22, no. 5, Article ID bbab032, 2021.
- [20] P. Charoentong, F. Finotello, M. Angelova et al., "Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade," *Cell Reports*, vol. 18, no. 1, pp. 248–262, 2017.
- [21] S. Wang, M. Gribskov, T. R. Hazbun, and P. E. Pascuzzi, "CellMiner Companion: an interactive web application to explore CellMiner NCI-60 data," *Bioinformatics*, vol. 32, no. 15, pp. 2399–2401, 2016.
- [22] S. Kumagai, Y. Togashi, T. Kamada et al., "The PD-1 expression balance between effector and regulatory T cells predicts the clinical efficacy of PD-1 blockade therapies," *Nature Immunology*, vol. 21, no. 11, pp. 1346–1358, 2020.
- [23] V. R. Juneja, K. A. McGuire, R. T. Manguso et al., "PD-L1 on tumor cells is sufficient for immune evasion in immunogenic tumors and inhibits CD8 T cell cytotoxicity," *Journal of Experimental Medicine*, vol. 214, no. 4, pp. 895–904, 2017.
- [24] F. Jiang, X. Y. Wang, M. Y. Wang et al., "An immune checkpoint-related gene signature for predicting survival of pediatric acute myeloid leukemia," *Journal of Oncology*, vol. 2021, Article ID 5550116, 14 pages, 2021.
- [25] J. Fu, K. Li, W. Zhang et al., "Large-scale public data reuse to model immunotherapy response and resistance," *Genome Medicine*, vol. 12, no. 1, p. 21, 2020.
- [26] P. Jiang, S. Gu, D. Pan et al., "Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response," *Nature Medicine*, vol. 24, no. 10, pp. 1550–1558, 2018.
- [27] L. Franco-Luzon, S. Garcia-Mulero, R. Sanz-Pamplona et al., "Genetic and immune changes associated with disease progression under the pressure of oncolytic therapy in A neuroblastoma outlier patient," *Cancers*, vol. 12, no. 5, p. 1104, 2020.
- [28] M. G. Vander Heiden, L. C. Cantley, and C. B. Thompson, "Understanding the Warburg effect: the metabolic requirements of cell proliferation," *Science*, vol. 324, no. 5930, pp. 1029–1033, 2009.
- [29] R. J. DeBerardinis and N. S. Chandel, "Fundamentals of cancer metabolism," *Science Advances*, vol. 2, no. 5, Article ID e1600200, 2016.
- [30] X. Wei, T. Mao, S. Li et al., "DT-13 inhibited the proliferation of colorectal cancer via glycolytic metabolism and AMPK/mTOR signaling pathway," *Phytomedicine*, vol. 54, pp. 120–131, 2019.
- [31] R. Pandey, L. Cafilisch, A. Lodi, A. J. Brenner, and S. Tiziani, "Metabolomic signature of brain cancer," *Molecular Carcinogenesis*, vol. 56, no. 11, pp. 2355–2371, 2017.
- [32] L. Moren, A. Bergenheim, S. Ghasimi, T. Brannstrom, M. Johansson, and H. Antti, "Metabolomic screening of tumor tissue and serum in glioma patients reveals diagnostic and prognostic information," *Metabolites*, vol. 5, no. 3, pp. 502–520, 2015.
- [33] Z. Huang, A. He, J. Wang et al., "The circadian clock is associated with prognosis and immune infiltration in stomach adenocarcinoma," *Aging (Albany NY)*, vol. 13, no. 12, Article ID 16655, 2021.
- [34] Y. Gong, S. Wu, S. Dong et al., "Development of a prognostic metabolic signature in stomach adenocarcinoma," *Clinical and Translational Oncology*, vol. 24, no. 8, pp. 1615–1630, 2022.
- [35] Z. Zhong, Z. Ye, G. He, W. Zhang, J. Wang, and S. Huang, "Low expression of A-kinase anchor protein 5 predicts poor prognosis in non-mucin producing stomach adenocarcinoma based on TCGA data," *Annals of Translational Medicine*, vol. 8, no. 4, p. 115, 2020.
- [36] Q. Zou, Y. Lv, Z. Gan, S. Liao, and Z. Liang, "Identification and validation of a malignant cell subset marker-based polygenic risk score in stomach adenocarcinoma through integrated analysis of bulk and single-cell RNA sequencing data," *Frontiers in Cell and Developmental Biology*, vol. 9, Article ID 720649, 2021.
- [37] A. Herzine, G. Sekkat, S. Kaminski et al., "Lipolysis-stimulated lipoprotein receptor acts as sensor to regulate ApoE release in astrocytes," *International Journal of Molecular Sciences*, vol. 23, no. 15, p. 8630, 2022.

- [38] A. Altuna-Coy, X. Ruiz-Plazas, S. Sanchez-Martin et al., "The lipidomic profile of the tumoral periprostatic adipose tissue reveals alterations in tumor cell's metabolic crosstalk," *BMC Medicine*, vol. 20, no. 1, p. 255, 2022.
- [39] L. Brunetti, F. Loiodice, L. Piemontese, P. Tortorella, and A. Laghezza, "New approaches to cancer therapy: combining fatty acid amide hydrolase (FAAH) inhibition with peroxisome proliferator-activated receptors (PPARs) activation," *Journal of Medicinal Chemistry*, vol. 62, no. 24, Article ID 11003, 2019.
- [40] R. Mitra, J. Lee, J. Jo et al., "Prediction of postoperative recurrence-free survival in non-small cell lung cancer by using an internationally validated gene expression model," *Clinical Cancer Research*, vol. 17, no. 9, pp. 2934–2946, 2011.
- [41] H. Wu, H. Xu, D. Jia, T. Li, and L. Xia, "METTL3-induced UCK2 m6A hypermethylation promotes melanoma cancer cell metastasis via the WNT/ β -catenin pathway," *Annals of Translational Medicine*, vol. 9, no. 14, p. 1155, 2021.
- [42] R. J. Kishton, M. Sukumar, and N. P. Restifo, "Metabolic regulation of T cell longevity and function in tumor immunotherapy," *Cell Metabolism*, vol. 26, no. 1, pp. 94–109, 2017.
- [43] R. M. Loftus and D. K. Finlay, "Immunometabolism: cellular metabolism turns immune regulator," *Journal of Biological Chemistry*, vol. 291, no. 1, pp. 1–10, 2016.
- [44] R. D. Leone and J. D. Powell, "Metabolism of immune cells in cancer," *Nature Reviews Cancer*, vol. 20, no. 9, pp. 516–531, 2020.
- [45] B. Ghesquiere, B. W. Wong, A. Kuchnio, and P. Carmeliet, "Metabolism of stromal and immune cells in health and disease," *Nature*, vol. 511, no. 7508, pp. 167–176, 2014.
- [46] X. Chen, L. Yan, F. Jiang et al., "Identification of a ferroptosis-related signature associated with prognosis and immune infiltration in adrenocortical carcinoma," *International Journal of Endocrinology*, vol. 2021, Article ID 4654302, 12 pages, 2021.
- [47] M. Cerezo and S. Rocchi, "Cancer cell metabolic reprogramming: a keystone for the response to immunotherapy," *Cell Death & Disease*, vol. 11, no. 11, p. 964, 2020.
- [48] F. Jiang, Y. Hu, X. Liu, M. Wang, and C. Wu, "Methylation pattern mediated by m(6)A regulator and tumor microenvironment invasion in lung adenocarcinoma," *Oxidative Medicine and Cellular Longevity*, vol. 2022, Article ID 2930310, 15 pages, 2022.